Introduction to Logic



Department of Philosophy
Central South University
xieshenlixi@163.com
github

October 28, 2025

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

狭义的逻辑学一般不研究什么?

- 特定领域的规律、机制: 市场逻辑、强盗逻辑、道德律令
- 辩论技巧: 预设结论, 通过修辞、共情等手段说服对方
- 批判性思维: 非形式谬误
- 侦探小说的诡计
- 语用推理

逻辑学研究什么?

- 逻辑是求真的 — 研究关于真的普遍性的规律
- 有效的推理形式(保真: 前提真保证结论真)
- 任何结构上都真的命题集合

Contents

Introduction
- Logic Puzzle
- Logic and other Disciplines
- Textbook and Homework

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

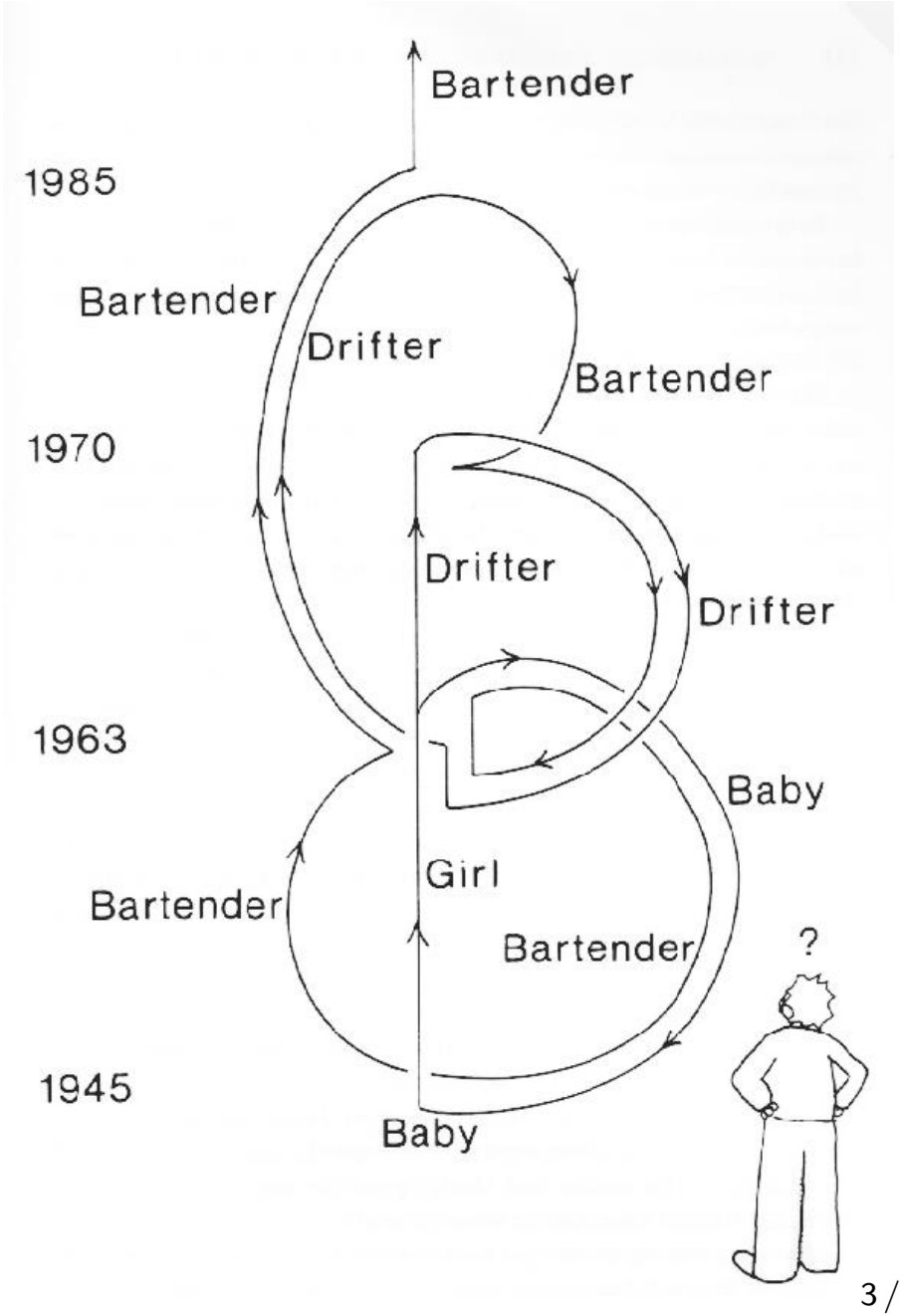
Quantum Computing

Answers to the Exercises

References2039

Predestination — All You Zombies — Heinlein

1945 一女婴被弃孤儿院.
1963 长大后的女孩与一男子邂逅、怀孕. 男子失踪. 女孩产下一女婴后发现自己是双性人. 女婴被偷. 伤心的她变性成他. 开始酗酒. 1970 一酒保把他招募进时光穿梭联盟. 为报复负心男, 酒保带他飞回 1963. 他邂逅一女孩并使其怀孕. 酒保乘时光机前行 9 个多月偷走女婴, 并将其送至 1945 的孤儿院, 然后回 1963 把他带到 1985 的联盟基地. 他受命飞回 1970, 化装酒保去招募一个酒鬼.



Problem

周迅的前男友窦鹏是窦唯的堂弟; 窦唯是王菲的前老公; 周迅的前男友宋宁是高原的表弟; 高原是窦唯的前任老婆; 周迅的前男友李亚鹏是王菲的现任老公; 周迅的前男友朴树的音乐制作人是张亚东; 张亚东是王菲的前老公窦唯的妹妹窦颖的前老公, 也是王菲的音乐制作人; 张亚东是李亚鹏前女友瞿颖的现男友.
下列说法不正确的是:

1. 王菲周迅是情敌关系
2. 瞿颖王菲是情敌关系
3. 窦颖周迅是情敌关系
4. 瞿颖周迅是情敌关系

$$\mathsf{Rival}(x,y) := \exists z\Big(\big(\mathsf{Now}(x,z) \wedge \mathsf{Ex}(y,z)\big) \vee \big(\mathsf{Now}(y,z) \wedge \mathsf{Ex}(x,z)\big)\Big)$$

“脑筋急转弯”可以用“套路”机械求解吗?

Problem (天堂之门)

- 你面前有左右两护卫镇守左右两门.
- 一人只说真话, 一人只说假话.
- 一门通天堂, 一门通地狱.
- 你只能向其中一人提一个“是/否”的问题.
- 怎么问出去天堂的门?

“脑筋急转弯”可以用“套路”机械求解吗?

Problem (天堂之门)

- 你面前有左右两护卫镇守左右两门.
- 一人只说真话, 一人只说假话.
- 一门通天堂, 一门通地狱.
- 你只能向其中一人提一个“是/否”的问题.
- 怎么问出去天堂的门?

你说真话当且仅当左门通天堂, 是吗?
如果我明天问你“左门通天堂吗”? 你会说“是”吗?
如果我问另一个人“左门通天堂吗”? 他会说“是”吗?

Problem (Hardest Logic Puzzle Ever)

- *Three gods, A, B, and C are called in some order, T, F, and R.*
- *T always speaks truly, F always speaks falsely (if he is certain he can; but if he is unable to lie with certainty, he responds like R), but whether R speaks truly or falsely (or whether R speaks at all) is completely random.*
- *Your task is to determine the identities of A, B, and C by asking 2 (3) yes/no questions; each question must be put to exactly one god.*
- *The gods understand English, but will answer in their own language, in which the words for 'yes' and 'no' are 'da' and 'ja' in some order. You don't know which word means which.*

- solution: assume T and F can't predict R's answer
  1. Directed to A:
     Would you answer 'ja' to the question of whether you would answer with a word that means 'yes' in your language to the question of whether you and B would give the same answer to the question whether '$1 + 1 = 2$'? Q
  2. Directed to A or B we now know not to be R:
     $Q[C/B]$

- solution: assume T and F can predict R's answer
  1. Directed to A:

     Would you answer 'ja' to the question of whether either:
     - B isn't R and you are F, or
     - B is R and you would answer 'da' to Q? Q
  2. Directed to A or B we now know not to be R:
     $Q[C/B, Q'/Q]$ Q'

Contents

Introduction
- Logic Puzzle
- Logic and other Disciplines
- Textbook and Homework

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

逻辑与哲学、数学、计算机

- Logic vs (Analytic) Philosophy.
  sense & reference / extension & intension / use & mention / truth & provability / mutual vs distributed vs common knowledge / knowledge update / belief revision / preference change / information flow / action & strategy / multi-agent interaction / counterfactual / causation / possible world / cross-world identity / essentialism / induction / ontological commitment / concept analysis / laws of thought / extent & limit / paradoxes ...
  Leibniz, Peirce, Frege, Russell, Wittgenstein, Ramsey, Carnap, Quine, Putnam, Kripke, Chomsky, Cantor, Hilbert, Gödel, Tarski, Turing ...
- Logic vs Mathematics.
  Logicism / Formalism / Intuitionism / Constructivism / Finitism / Structuralism / Homotopy Type Theory
- Logic vs Computer Science.

$$\frac{\text{Logic}}{\text{Computer Science}} \approx \frac{\text{Calculus}}{\text{Physics}}$$

逻辑与计算机

- Computer Architecture.
  Logic gates and digital circuit design $\approx$ Propositional Logic
- Programming Languages.
  LISP $\approx$ λ-calculus
  Prolog $\approx$ First Order Logic + Recursion
- Theory of Computation. Computational / Descriptive Complexity.
- General Problem Solver (SAT solvers).
- Automated Theorem Proving.
- Common sense reasoning via Non-monotonic Logic.
- Fuzzy Control vs Fuzzy Logic and Multi-valued Logic.
- Relational Databases.
  SQL $\approx$ First Order Logic + Syntactic Sugar
- Software Engineering (Formal Specification and Verification).
  Temporal Logic, Dynamic Logic, Hoare Logic, Model Checking
- Multi-agent Systems.
  Epistemic Logic
- Knowledge representation. Semantic Web.
  Web Ontology Language (OWL) $\approx$ Description Logic

逻辑与语言学、经济、法学、社会科学

- Logic vs Linguistics.
  Syntax, Semantics and Pragmatics of Natural Language
  (Montague Grammar, Inquisitive Semantics)
  Parsing as deduction (Lambek calculus)
- Logic vs Economics, Law and Social Sciences.
  Epistemic Game Theory
  Social Choice Theory
  Rational Choice Theory
  Decision Theory
- ...

逻辑学的几个主要分支

Mathematical Logic

- **First Order Logic**
- Set Theory
- Model Theory
- Proof Theory
- Recursion Theory
- (Homotopy) Type Theory
- Category / Topos Theory / Categorical Logic

Computational Logic

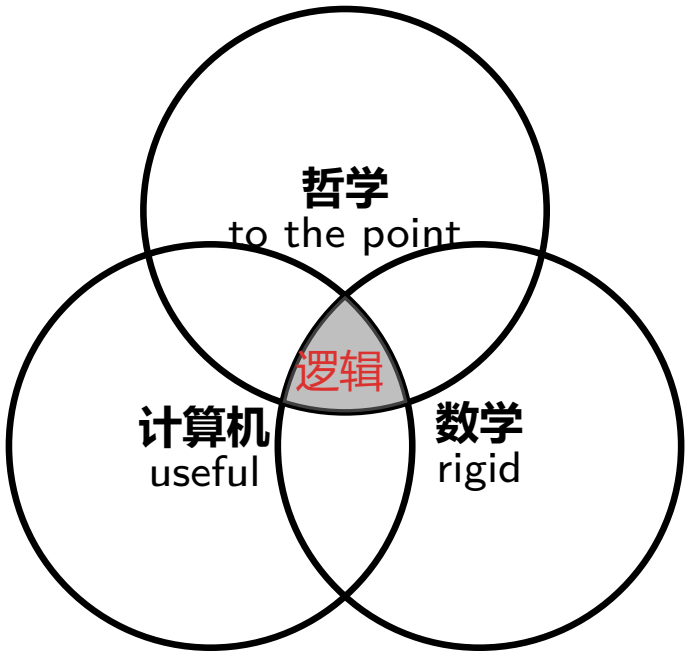
- Automata Theory
- Computational Complexity
- Finite Model Theory
- Model Checking
- Lambda Calculus
- Theorem Proving
- Description Logic
- Fixpoint Logic
- Dynamic Logic
- Linear Logic
- Temporal Logic
- Process Algebra
- Hoare Logic
- Inductive Logic
- Fuzzy Logic
- Non-monotonic Logic
- Computability Logic
- Default Logic
- Markov Logic Networks
- Situation/Event Calculus

Philosophical Logic

- Intuitionistic Logic
- Modal Logic
- Algebraic Logic
- Epistemic Logic
- Doxastic Logic
- Preference Logic
- Provability Logic
- Spatial Logic
- Justification Logic
- Hybrid Logic
- Substructural Logic
- Free Logic
- Counterfactual Logic
- Relevance Logic
- Quantum Logic
- Paraconsistent Logic
- Intensional Logic
- Partial Logic
- Diagrammatic Logic
- Deontic Logic

跨学科视角下的逻辑

- Logic is
  1. mainly philosophy by subject matter
  2. mainly mathematics by methodology
  3. mainly computer science by applications
- Logicians always want to be
  1. Philosophers of philosophers
  2. Mathematicians of mathematicians
  3. Computer scientists of computer scientists
- However, they often end up being
  1. Mathematicians to philosophers
  2. Computer scientists to mathematicians
  3. Philosophers to computer scientists



The Unreasonable Ineffectiveness of Philosophy

- 费曼: “砖头算不算本质客体?”
- 哲学家甲: “一块砖是独特的砖, 是怀海德所说的本质客体.”
- 哲学家乙: “本质客体的意思并不是指个别的砖块, 而是指所有砖块的共有的普遍性质, 换句话说, ‘砖性’ 才是本质客体.”
- 哲学家丙: “不对, 重点不在砖本身, ‘本质客体’ 指的是, 当你想到砖块时内心形成的概念.”
- 就像所有关于哲学家的故事一样, 最终以一片混乱收场. 好笑的是, 在先前的那么多次讨论中, 他们从来没有问过自己, 像简单的砖块究竟是不是 “本质客体”.

— 费曼

哲学旨在感动那些混淆晦涩与深刻的人.

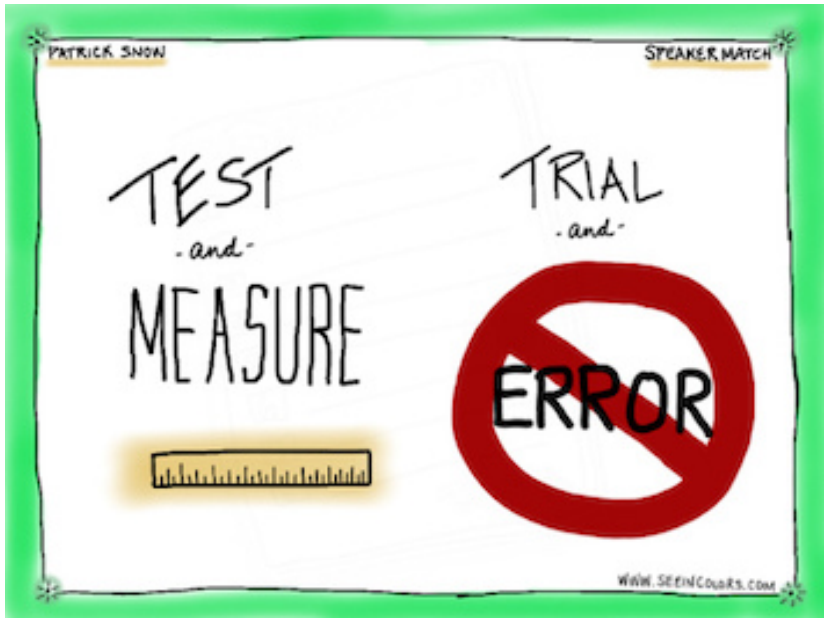
— 温伯格

The Unreasonable Ineffectiveness of Philosophy

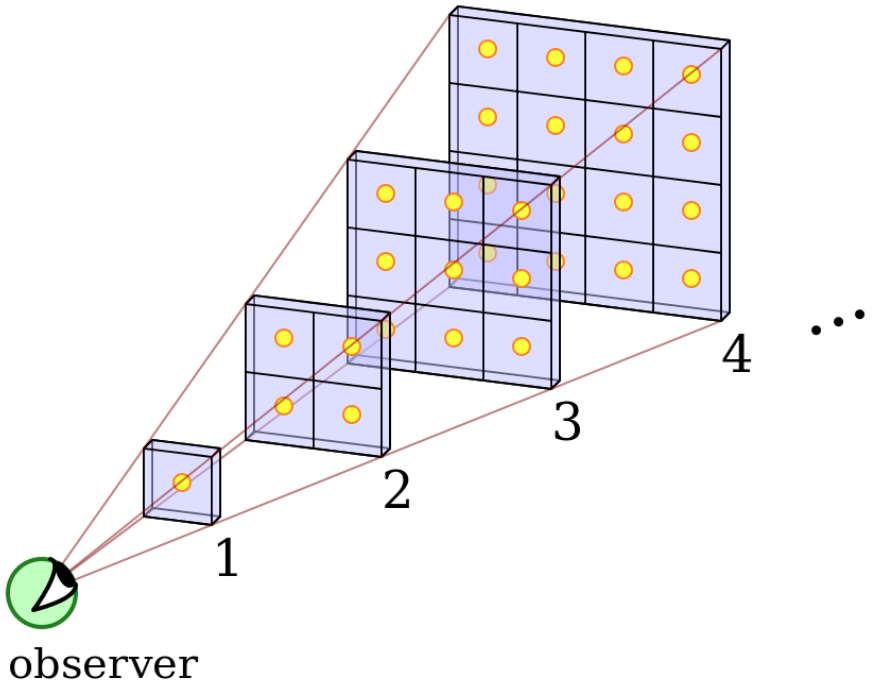
- 如果一个哲学家说的东西是对的, 那一定是微不足道的; 如果不是微不足道时, 那一定是错的. — 高斯
- 哲学难道不是用蜜写成的吗? 乍一看, 很精彩, 再一看, 除了一团浆糊, 什么都没留下. — 爱因斯坦
- 哲学之于科学, 就像手淫之于性: 它廉价、易实现, 甚至有些人也更喜欢它.
- 哲学家就像瞎子, 在漆黑的夜里, 举着熄灭的蜡烛, 跑到黑暗的地下洞穴, 找寻一只并不存在的黑猫, 然后大呼小叫: "我脖子被猫爪挠啦!"
- 哲学可以作为跨学科研的 "催化剂" 或 "调味品". 但 "哲学内部" 问题像是智力 "黑洞", 吸收大量智力, 却无任何输出. — *Johan van Benthem*
- 哲学偶尔有用, 能让我们免受另外一些哲学立场的影响. — 温伯格

实践是检验真理的唯一标准

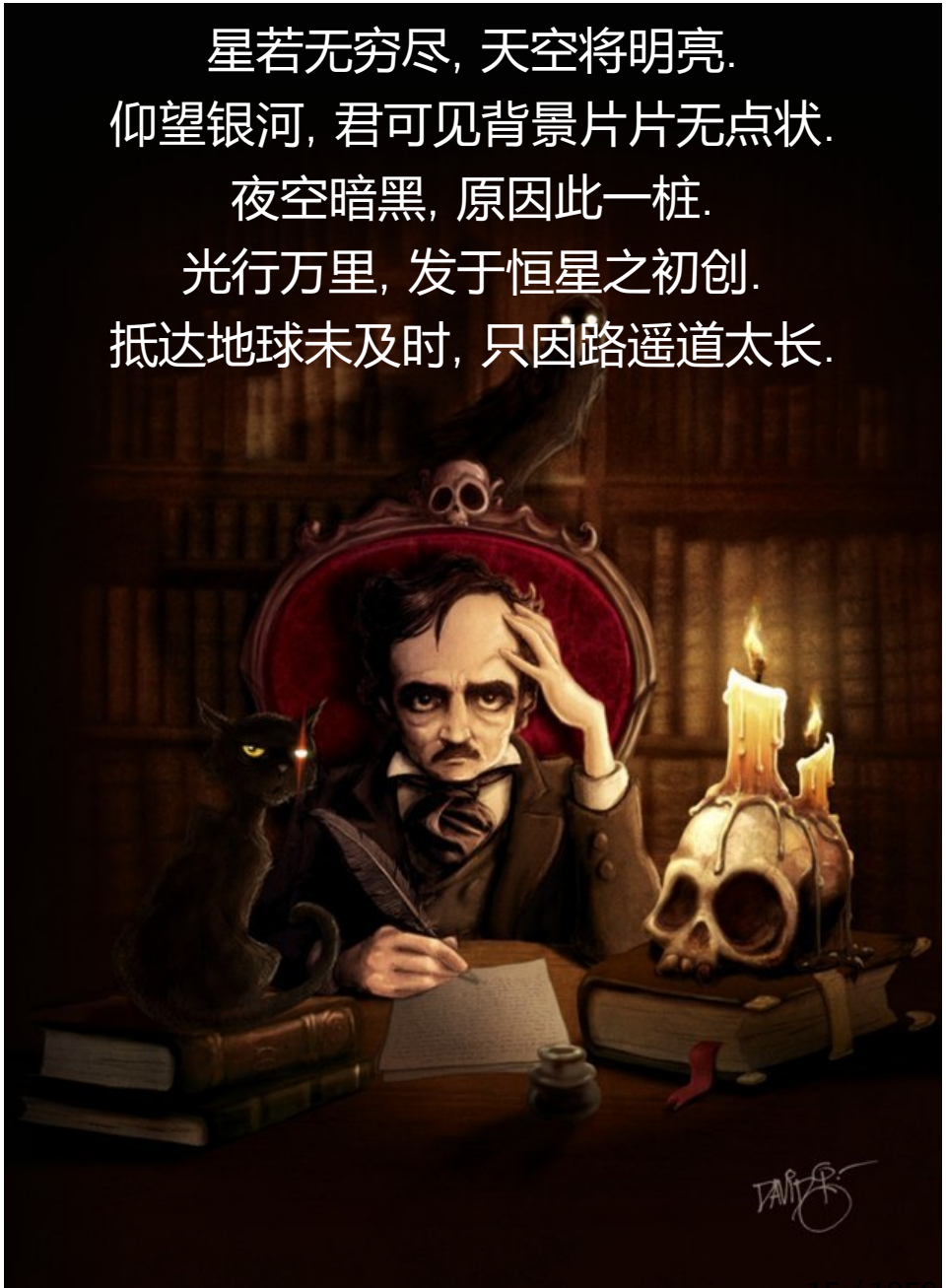
- 什么是 “实践”?
- 什么是 “真理”?
- 什么是 “标准”?
- 怎么 “检验”?
- 怎么用 “实践” 检验 “真理”?
- 为什么 “唯一”?



夜晚的天空为什么是黑的?
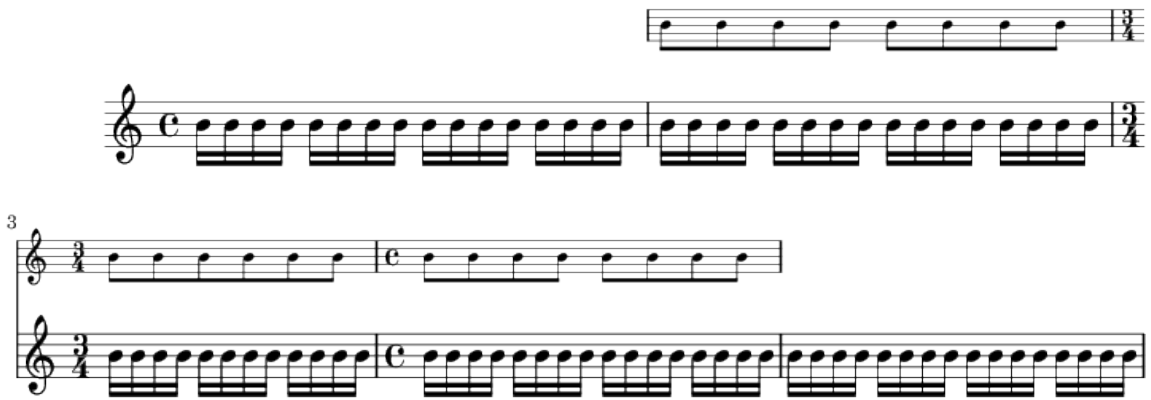


一个静态的, 无限古老的宇宙, 无限多的恒星, 均匀分布在无限大的空间中, 那么夜空将会是明亮的, 而不是暗黑的.



Why Study Logic? — The glory of the human spirit!

- 逻辑是心灵的免疫系统.
- 世界不仅比我们想象的更古怪, 甚至比我们能够想象的更古怪.
- 逻辑可以扩展我们的抽象想象力, 重塑我们的心灵.
- 几行逻辑推理可能改变我们看待世界的方式.
- 从逻辑的观点看 (世界、语言与世界的关系、哲学.....)
- 通往柏拉图理念世界的桥梁, 为了人类心智的荣耀!
- 逻辑自身就很有趣 — 理性的音乐.



- 鉴赏是视听艺术. 视读乐符, 演奏乐章.
- 创作是言说艺术. 言说什么? 如何言说?

世界逻辑日: 每年的 1 月 14 日

因为担心失衡跌倒,
我们的思想紧紧抓住逻辑
这个扶手.

— 纪德

时光飞逝, 勿浪费时间;
方法会教你赢得每一天;
年轻的朋友, 听我一句劝,
大学逻辑是起点!

— 歌德《浮士德》

“工欲善其事, 必先利其器.”

— 孔子《论语》

“君子性非异也, 善假于物也.”

— 荀子《劝学》

逻辑是笨人的学问

- 蠢人会把对方的智商拉低到自己的水平, 然后用丰富的经验打败他.
- 笨人会把对方的问题翻译为自己的逻辑语言, 然后用机械的逻辑工具暴力破解它 ☺

  — 学逻辑唯有 “笨办法”, 要不得小聪明取不得巧

杀不死我的, 使我更强大!

— 尼采

Contents

Introduction
- Logic Puzzle
- Logic and other Disciplines
- Textbook and Homework

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

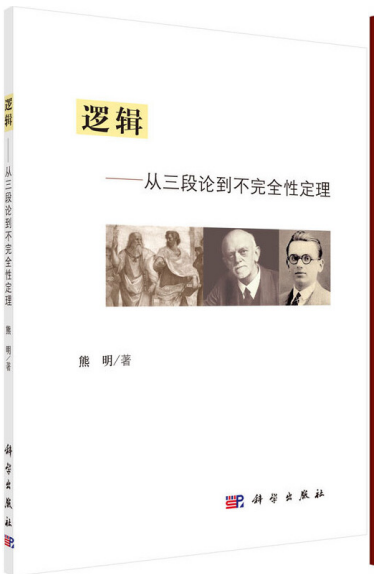
中文参考书

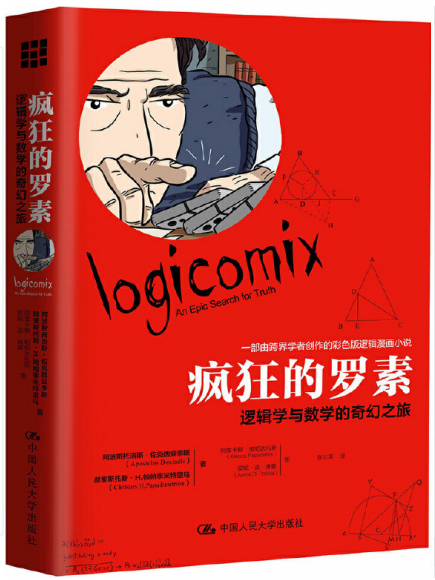
参考书 **熊明: 《逻辑 — 从三段论到不完全性定理》**

参考书 郝兆宽、杨睿之、杨跃: 《数理逻辑: 证明及其限度》

漫画书 Doxiadis, Papadimitriou: 《Logicomix 疯狂的罗素》

科普书 侯士达: 《哥德尔、艾舍尔、巴赫 — 集异璧之大成》







侯世达定律: 做事所花费的时间总是比你预期的要长, 即使你的预期中考虑了**侯世达定律**.

- Dangerous Knowledge 危险的知识
- The Imitation Game 模仿游戏
- libgen
- sci-hub

英文参考书

1. H. J. Gensler: Introduction to Logic. (中) — B
2. H. de Swart: Pilosophical and Mathematical Logic. — P
3. N. J. J. Smith: Logic — The Laws of Truth. — P
4. P. Smith: An Introduction to Formal Logic. — P
5. *P. Smith: Beginning Mathematical Logic.* — P
6. J. van Benthem: Logic in Action. — P
7. Open Logic Project. — P
8. H. Enderton: A Mathematical Introduction to Logic. (中) — **L**
9. H. Ebbinghaus, J. Flum, W. Thomas: Mathematical Logic. — L
10. A. Nerode, R. A. Shore: Logic for Applications. — C
11. S. Burris: Logic for Mathematics and Computer Science. — C
12. Y. Manin: A Course in Mathematical Logic for Mathematicians. — M

Advanced Readings

- Modal Logic
  - J. van Benthem: Modal Logic for Open Minds
  - P. Blackburn, M. de Rijke, Y. Venema: Modal Logic
- Set Theory
  - T. Jech: Set Theory
  - K. Kunen: Set Theory
- Recursion Theory
  - R. I. Soare: Turing Computability
  - A. Nies: Computability and Randomness
  - M. Li, P. Vitányi: An Introduction to Kolmogorov Complexity and Its Applications
- Model Theory
  - D. Marker: Model Theory
  - C. C. Chang, H. J. Keisler: Model theory
- Proof Theory
  - G. Takeuti: Proof Theory
  - W. Pohlers: Proof Theory

课程目标 & 大纲 & 成绩

课程目标:

- 论证的形式化表达 ♡
- 论证有效性的判定 ♡
- 数学哲学 
- 形式认识论 
- 逻辑在哲学、数学、计算机科学、人工智能、语言学、认知科学、物理学、信息论、博弈论、社会科学等领域的应用 

大纲:

1. 命题逻辑 ♡
2. 谓词逻辑 ♡
3. 模态逻辑 

平时成绩 40%**、期末成绩** 60%

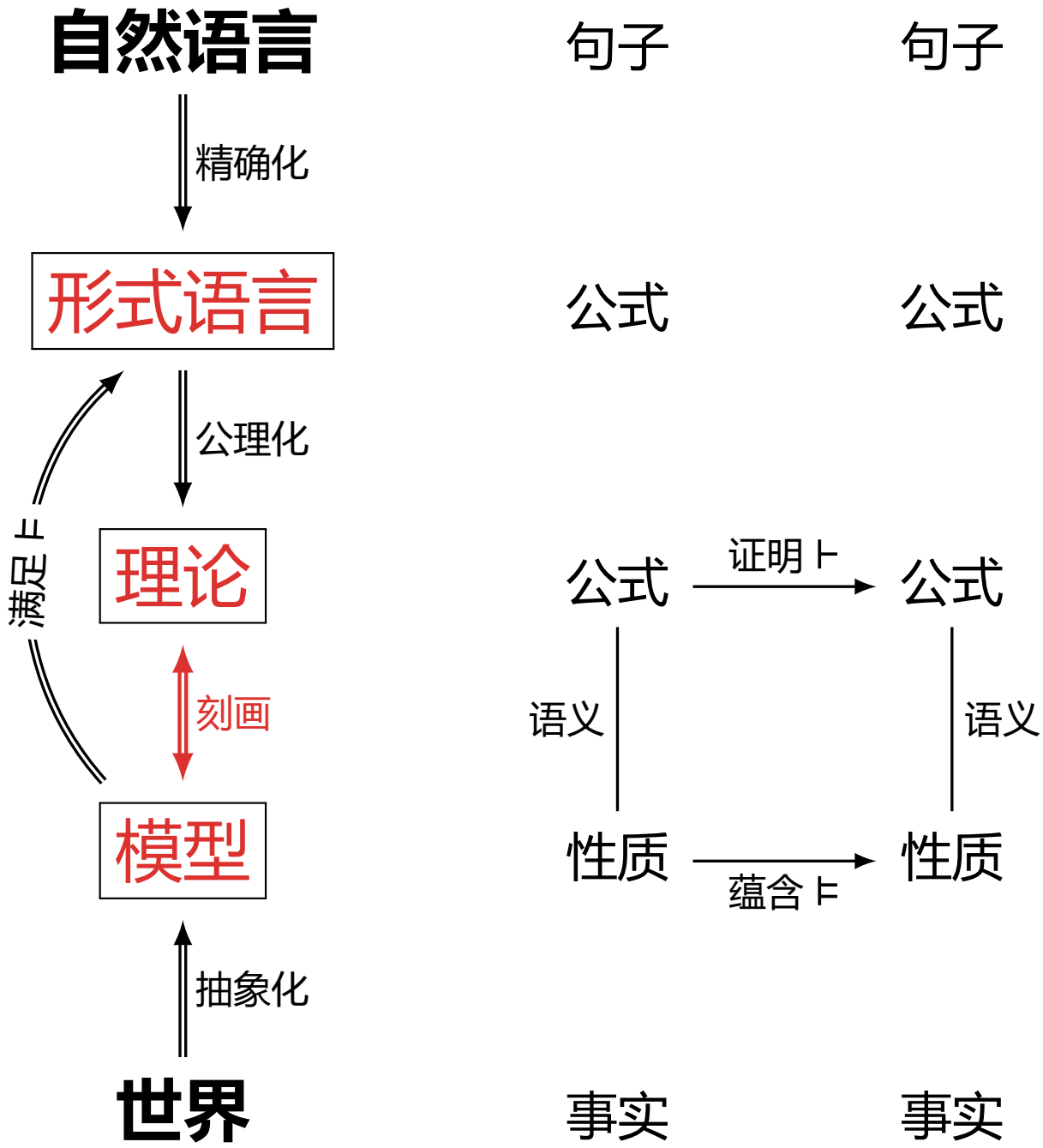
- 问答、讨论
- 作业 
- 练习
- 考试
- Paper
- Techniques e.g. LaTeX, Coq, Lean, Prover9, Vampire ...

课件记号的意思:

 考试相关

♡ 需要掌握

 可略过


自然语言
精确化
形式语言
公理化
理论
满足
刻画
模型
抽象化
世界
句子
句子
公式
公式
公式
证明⊢
公式
语义
语义
性质
蕴含⊨
性质
事实
事实

一个超级简单的 Toy Logic

- **语法:** 符号集 $\mathrm{Var} = \{X, Y, \dots\}$, 公式 $X \to Y$
- **语义:** 结构 $\mathcal{M} := (M, [\![\,]\!])$, 其中 $[\![X]\!] \subset M$

$$\mathcal{M} \vDash X \to Y \quad \text{当且仅当} \quad [\![X]\!] \subset [\![Y]\!]$$

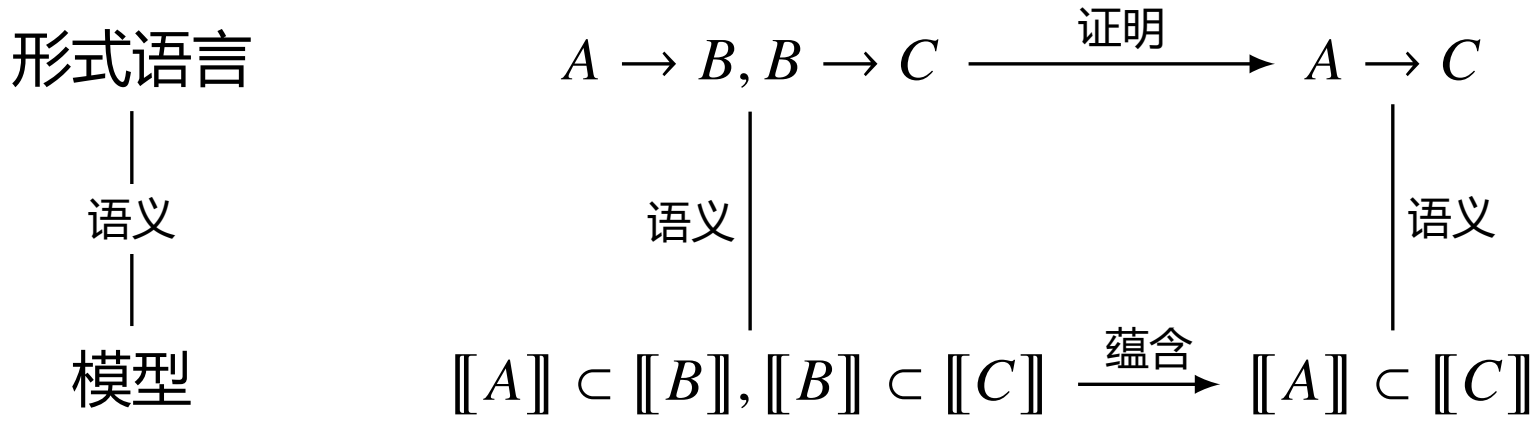
- 逻辑蕴涵: $\Gamma \vDash \varphi$ 当且仅当对任意结构 $\mathcal{M}$: 若 $\mathcal{M} \vDash \Gamma$ 则 $\mathcal{M} \vDash \varphi$
- **形式系统:**

$$\frac{X \to Y \quad Y \to Z}{X \to Z} \qquad \frac{}{X \to X}$$

- 证明: $\Gamma \vdash \varphi$ 当且仅当存在一棵以 φ 为 "根" 节点的有穷 "树", 其每一节点或是公理或属于前提 Γ, 或通过推理规则由前面的节点生成
- 例子: 怎么证明 $A \to B, B \to C, C \to D \vdash A \to D$

$$\frac{\dfrac{A \to B \quad B \to C}{A \to C} \qquad C \to D}{A \to D}$$

- **元定理:** 能推出的都有效? 有效的都能推出? $\Gamma \vdash \varphi \iff \Gamma \vDash \varphi$
- 思考一下: 怎么证明推不出? $A \to B \nvdash B \to A$


形式语言
语义
模型
$A \to B, B \to C$
证明
$A \to C$
语义
语义
$[\![A]\!] \subset [\![B]\!], [\![B]\!] \subset [\![C]\!]$
蕴含
$[\![A]\!] \subset [\![C]\!]$

Theorem (Toy Logic 的可靠性定理 🚲)

$\Gamma \vdash \varphi \implies \Gamma \vDash \varphi$

Proof.

对证明树 $\Gamma \vdash \varphi$ 线性排列, 对其证明长度应用数学归纳法.

- 长度为 1: $\varphi \in \Gamma$ 或 φ 是公理 $X \to X$ 的特例. 显然 $\Gamma \vDash \varphi$.
- 假设证明长度不超过 n 时都有 $\Gamma \vdash \varphi \implies \Gamma \vDash \varphi$, 下证长度为 $n+1$ 时也成立.
  第 $n+1$ 步或是公理或属于前提或由推理规则得到.
  若由推理规则 $\dfrac{X \to Y \quad Y \to Z}{X \to Z}$ 得到, 由归纳假设, $\Gamma \vDash X \to Y$ 且 $\Gamma \vDash Y \to Z$.
  即 $[\![X]\!] \subset [\![Y]\!]$ 且 $[\![Y]\!] \subset [\![Z]\!]$.
  所以 $[\![X]\!] \subset [\![Z]\!]$, 即 $\Gamma \vDash X \to Z$.

□

Remark: 公理有效, 推理规则保真

Theorem (Toy Logic 的完备性定理 🚲)

$\Gamma \vDash \varphi \implies \Gamma \vdash \varphi$

Proof.

- 不妨证其逆否命题: 假设 $\Gamma \nvdash \varphi$, 往证 $\Gamma \nvDash \varphi$,
  即, 找一个结构 $\mathcal{M}$, 使得 $\mathcal{M} \vDash \Gamma$ 但 $\mathcal{M} \nvDash \varphi$.
- 令 $\mathcal{M}^\Gamma := (M, [\![\,]\!])$, 其中

$$M := \mathrm{Var}, \quad [\![X]\!] := \{Y : \Gamma \vdash Y \to X\}$$

1. 检查 $\mathcal{M}^\Gamma \vDash \Gamma$.
   任给 $A \to B \in \Gamma$, 根据 $[\![A]\!], [\![B]\!]$ 的定义和 $\dfrac{Y \to A \quad A \to B}{Y \to B}$, 可得
   $Y \in [\![A]\!] \implies Y \in [\![B]\!]$, 即 $\mathcal{M}^\Gamma \vDash A \to B$.
2. 检查 $\mathcal{M}^\Gamma \nvDash \varphi$. (往证: $\mathcal{M}^\Gamma \vDash \varphi \implies \Gamma \vdash \varphi$, 从而与假设 $\Gamma \nvdash \varphi$ 矛盾)
   假设 $\varphi = A \to B$, 若 $\mathcal{M}^\Gamma \vDash \varphi$, 则 $[\![A]\!] \subset [\![B]\!]$.
   即, 对任意 Y, $Y \in [\![A]\!] \implies Y \in [\![B]\!]$.
   因为 $\overline{A \to A}$, 所以 $A \in [\![A]\!]$. 从而 $A \in [\![B]\!]$. 因此 $\Gamma \vdash \varphi$.

让语言表达力更丰富一点儿?

- **语法:** $\mathtt{All}(X,Y)$, $\mathtt{Some}(X,Y)$, $\mathtt{All}(X,\overline{Y})$, $\mathtt{Some}(X,\overline{Y})$, $\mathtt{Most}(X,Y)$
- **语义:** 结构 $\mathcal{M} := (M, [\![\,]\!])$, 其中 $[\![X]\!] \subset M$, $[\![\overline{X}]\!] = M \setminus [\![X]\!]$

$$\mathcal{M} \vDash \mathtt{Some}(X,Y) \quad \text{当且仅当} \quad [\![X]\!] \cap [\![Y]\!] \neq \emptyset$$

$$\mathcal{M} \vDash \mathtt{Most}(X,Y) \quad \text{当且仅当} \quad |[\![X]\!] \cap [\![Y]\!]| > |[\![X]\!] \setminus [\![Y]\!]|$$

- **形式系统:**

$$\frac{\mathtt{Some}(X,Y)}{\mathtt{Some}(Y,X)} \qquad \frac{\mathtt{Some}(X,Y)}{\mathtt{Some}(X,X)} \qquad \frac{\mathtt{All}(X,\overline{Y})}{\mathtt{All}(Y,\overline{X})}$$

$$\frac{\mathtt{Some}(X,Y) \quad \mathtt{All}(Y,Z)}{\mathtt{Some}(X,Z)} \qquad \frac{\mathtt{All}(X,Y) \quad \mathtt{Some}(X,Z)}{\mathtt{Some}(Y,Z)}$$

$$\frac{\mathtt{Most}(X,Y)}{\mathtt{Some}(Y,X)} \qquad \frac{\mathtt{Some}(X,X)}{\mathtt{Most}(X,X)} \qquad \frac{\mathtt{Most}(X,Y) \quad \mathtt{All}(Y,Z)}{\mathtt{Most}(X,Z)}$$

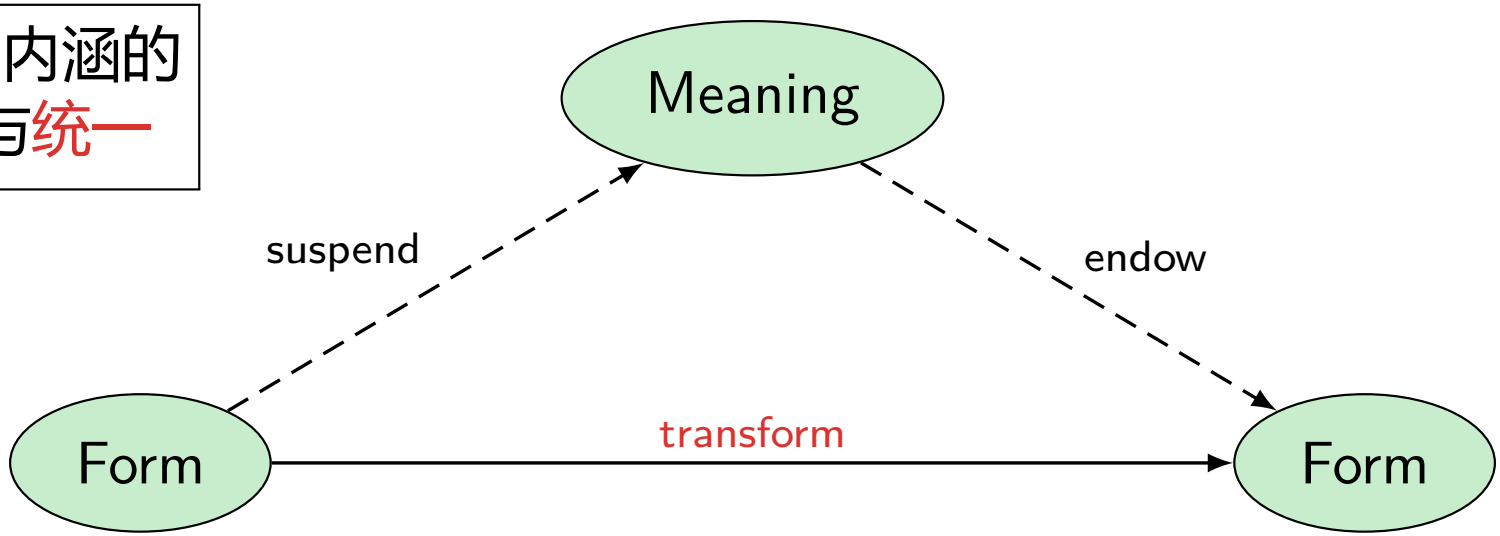
$$\frac{\mathtt{Most}(X,Z) \quad \mathtt{All}(X,Y) \quad \mathtt{All}(Y,X)}{\mathtt{Most}(Y,Z)} \qquad \frac{\mathtt{All}(Y,X) \quad \mathtt{All}(X,Z) \quad \mathtt{Most}(Z,Y)}{\mathtt{Most}(X,Y)}$$

逻辑求真

“真”为逻辑指引方向，正如“美”为美学、“善”为伦理指引方向.

— 弗雷格

形式与内涵的
分离与统一



数理逻辑是先于一切科学的科学，包含着位于一切科学底层的观念和原理.

— 哥德尔

一个概念是“逻辑”的，当且仅当，它对“论域”到自身的所有可能的一一变换都保持不变.

— 塔斯基

思考一下 

1. 什么是 “语法”/“语义”?
2. 什么是 “对象语言”/“元语言”?
3. 什么是 “满足”?
4. 什么是 “真”?
5. 什么是 “有效命题”/“有效论证”?
6. 什么是 “形式系统”?
7. 什么是 “定理”?
8. 什么是 “证明”?
9. 什么是 “理论”?
10. 什么是 “可 (有穷) 公理化”?
11. 什么是 “一致”?
12. 什么是 “可靠”?
13. 什么是 “完备”?
14. 什么是 “紧致”?
15. 什么是 “定义”?
16. 什么是 “表示”?
17. 什么是 “归纳”/“递归”?
18. 什么是 “有穷”/“无穷”?
19. 什么是 “可数”/“不可数”?
20. 什么是 “可判定”?
21. 什么是 “计算复杂性”?
22. 什么是 “表达力”?
23. 什么是 “简洁性”?
24. 什么是 “可解释性”?
25. 什么是 “同态”?
26. 什么是 “同构”?
27. 什么是 “初等等价”?
28. 什么是 “范畴性”?

Homework ⚡

1. 自选自然语言的句子，将其翻译为逻辑语言.
   — 所选的句子既要有「趣味性」又要有一点儿「难度」，可以是名人名言、名篇名句、古典诗词、歇后语、网络流行语等等，自己作诗也欢迎 ☺
   — 判断所选的句子是否构成一个有效或无效的命题或论证，给出证明或反模型.
2. 自选习题、自学笔记……

Google / Wikipedia / Stanford Encyclopedia / Internet Encyclopedia / StackExchange

考试样题

一、选择题. (8 道题, 每题 5 分, 共 40 分)

1、以下哪个公式是有效的?

A. $\exists y \forall x Rxy \to \forall y \exists x Rxy$　　B. $\forall x \exists y Rxy \to \exists x \forall y Rxy$

C. $\forall x \exists y Rxy \to \exists y \forall x Rxy$　　D. $\exists x \forall y Rxy \to \forall y \exists x Rxy$

二、下面公式是否有效? 若有效, 请证明, 若无效, 给出反模型. (10 分)

$$(p \to r) \vee (q \to r) \to (p \vee q) \to r$$

三、请将如下问题翻译成合适的命题逻辑语言并用逻辑学方法求解. (20 分)

已知一起凶杀案有三个嫌疑人: 小艾、小白和小菜. 至少有一人是凶手, 但不可能三人同时犯罪. 如果小艾是凶手, 那么小菜是同犯. 如果小白不是凶手, 那么小菜也不是. 请问, 谁肯定是凶手?

四、下面论证是否有效? 若有效, 请证明, 若无效, 给出反模型. (20 分)

人人都怕小艾. 小艾只怕我. 因此, 我就是小艾.

五、论述题. (10 分)

谈谈你对命题逻辑与布尔代数之间的关系的理解.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Contents

Introduction

Induction, Analogy, Fallacy
- Mill's Methods of Causal Analysis
- Analogical Argument
- Fallacy and Bullshit

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

简单枚举归纳

- 我们都是瞎子.
- 吝啬的人是瞎子, 他只看见金子看不见财富.
- 挥霍的人是瞎子, 他只看见开端看不见结局.
- 卖弄风情的女人是瞎子, 她看不见自己脸上的皱纹.
- 有学问的人是瞎子, 他看不见自己的无知.
- 诚实的人是瞎子, 他看不见坏蛋.
- 坏蛋是瞎子, 他看不见上帝.
- 上帝也是瞎子, 他在创造世界的时候, 没有看到魔鬼也跟着混进来了.
- 我也是瞎子, 我只知道说啊说啊, 没有看到你们全都是聋子.

求同法

$$\frac{ABC \rightarrow xyz \quad ADE \rightarrow xuv}{A \rightarrow x}$$

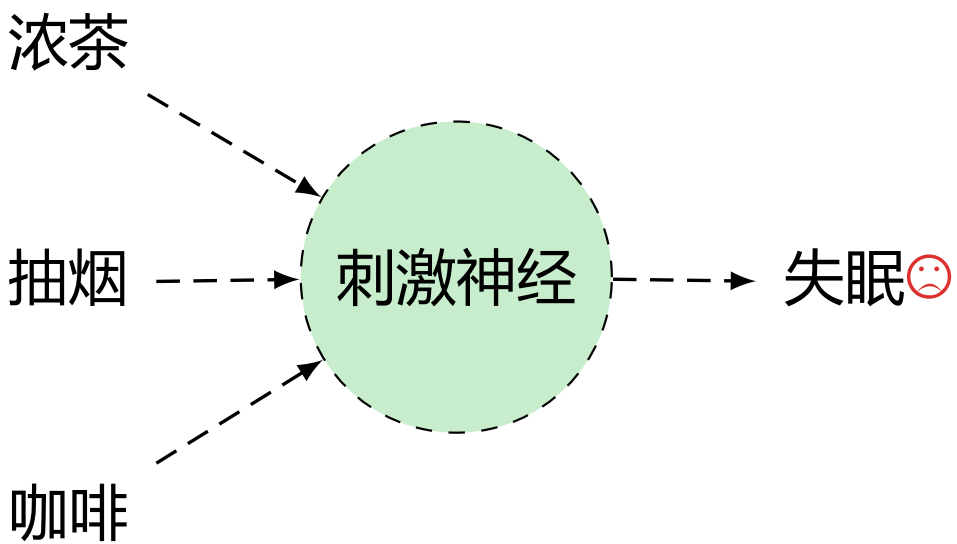
	缺碘	职业	年龄	甲状腺
张三	缺	工人	老年	肿大
李四	缺	农民	中年	肿大
王五	缺	士兵	青年	正常

缺碘 $\rightarrow$ 甲状腺肿大

甲状腺肿大人群的相同点是缺碘
缺碘是肿大的“必要”条件

泡脚	看小说	喝浓茶	失眠
泡脚	看电视	抽烟	失眠
泡脚	听音乐	喝咖啡	失眠

泡脚 $\rightarrow$ 失眠



求异法

$$ABC \to xyz$$
$$\overline{A}BC \to \bar{x}yz$$
$$\hline$$
$$A \to x$$

冰激凌　沙拉　面包　牛奶　肚子疼
酸奶　沙拉　面包　牛奶　肚子不疼

冰激凌 → 肚子疼

吃了冰激凌的人肚子疼
肚子不疼的人没吃冰激凌
冰激凌是肚子疼的“充分”条件

上课　头疼
不上课　头不疼

上课 → 头疼



取健康的蜘蛛，　大吼一声，　蜘蛛被吓跑了
砍掉蜘蛛的腿，　大吼一声，　蜘蛛纹丝不动

蜘蛛的听觉器官长在腿上. ☹

秋末冬初街道旁的响叶杨纷纷落叶，但高压水银灯下的响叶杨却迟迟不落叶，因此，高压水银灯照射可能是响叶杨落叶迟的原因.

求同求异共用法

$$\frac{\begin{array}{ll} ABC \to xyz & ABC \to xyz \\ ADE \to xuv & \overline{A}BC \to \overline{x}yz \end{array}}{A \to x}$$

动物	纲	环境	形态
鲨鱼	鱼纲	海	
鱼龙	爬行	海	
鲸鱼	哺乳	海	
鼹鼠	哺乳	陆	
蝙蝠	哺乳	空	

达尔文: 环境 $\to$ 形态

	咖啡	嗦粉	熬夜	看书	肚子疼
张三	●	●	●	●	●
李四	●		●		●
王五		●		●	
赵六			●	●	

剩余法

$$\frac{ABC \to xyz \quad B \to y \quad C \to z}{A \to x} \qquad \frac{ABC \to x \quad B \not\to x \quad C \not\to x}{A \to x}$$

受其他行星吸引, 天王星运行轨道上有四个地方发生偏斜
三个地方偏斜是已知行星吸引的结果

剩余一个地方偏斜是未知行星吸引的结果 (发现海王星)

	咖啡	嗦粉	熬夜	看书	疲惫	肚子疼
张三	•	•	•	•	•	•
李四	•		•		•	•
王五		•		•		
赵六			•	•	•	

Remark: 预设了一个原因确定一个结果; 原因彼此独立; 结果彼此独立; 所有可能的原因已知; 所有其它结果的原因已知.

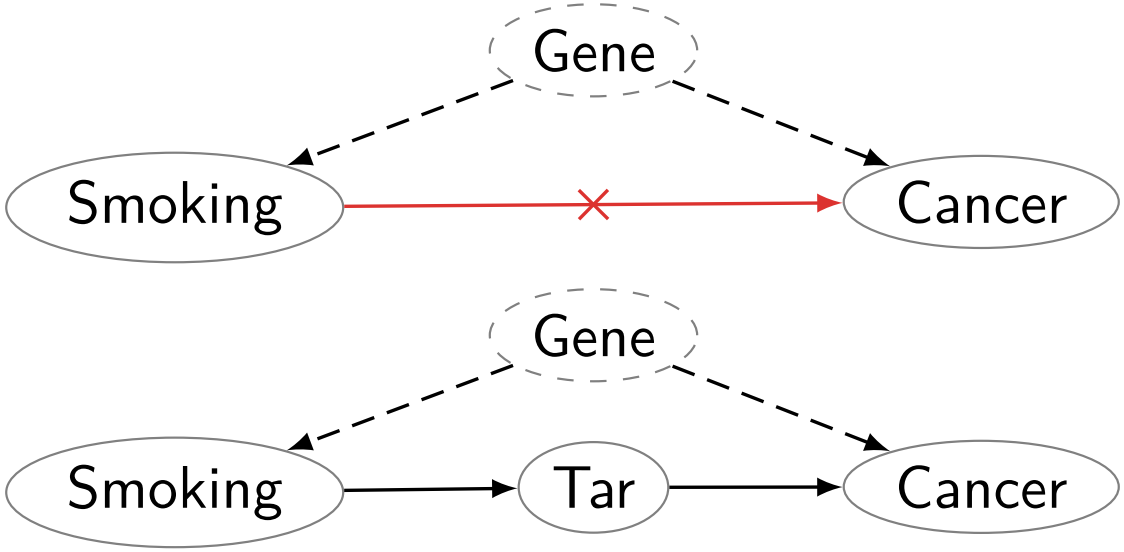
共变法

$$\frac{\begin{array}{l} ABC \to xyz \\ A^{\uparrow}BC \to x^{\uparrow}yz \\ A^{\downarrow}BC \to x^{\downarrow}yz \end{array}}{A \to x}$$

Example

热胀冷缩、体温表

禁烟人士 抽烟越多越容易患肺癌，所以抽烟是导致肺癌的重要原因.
烟草公司 某种基因是导致人们容易抽烟和容易得肺癌的共同原因.



Simpson's Paradox

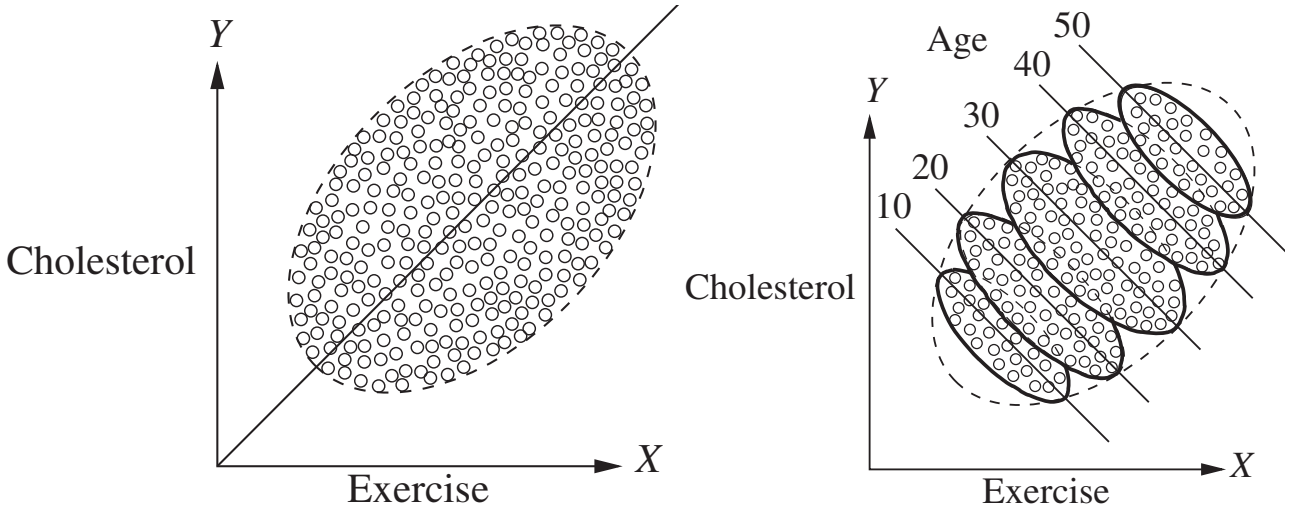
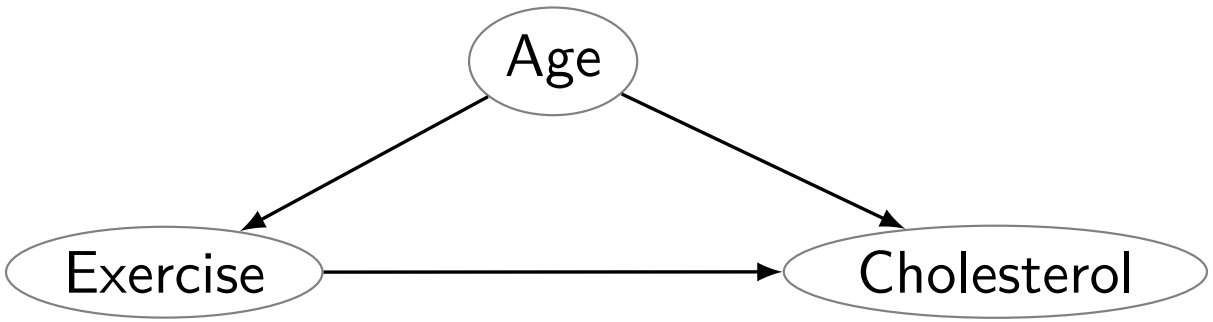


Figure: Exercise appears to be beneficial in each age group but harmful in the population as a whole. — Older people exercise more.



$$\mathbb{E}[\text{cholesterol} \mid \text{exercise}] > \mathbb{E}[\text{cholesterol} \mid \text{no exercise}]$$

$$\mathbb{E}[\text{cholesterol} \mid \texttt{do}(\text{exercise})] < \mathbb{E}[\text{cholesterol} \mid \texttt{do}(\text{no exercise})]$$

Contents

Introduction

Induction, Analogy, Fallacy
- Mill's Methods of Causal Analysis
- Analogical Argument
- Fallacy and Bullshit

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Example

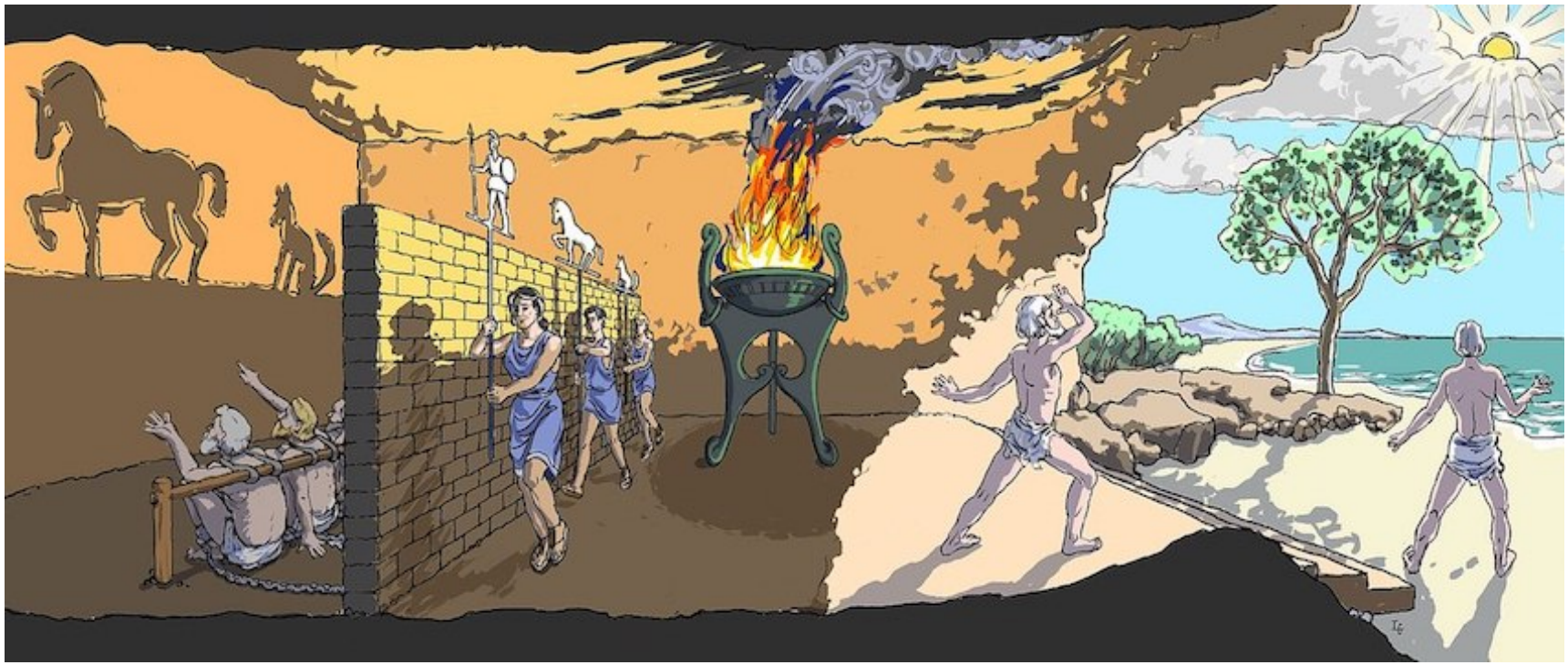
古龙《九月鹰飞》

- 叶开忍不住又道：“你为什么还是戴着这草帽？”
- 墨九星道：“因为外面有狗在叫。”
- 叶开怔了怔，道：“外面有狗叫，跟你戴草帽又有什么关系？”
- 墨九星冷冷道：“我戴不戴草帽，跟你又有什么关系？”

Example

庄子《齐物论》

不知周之梦为蝴蝶与，蝴蝶之梦为周与?



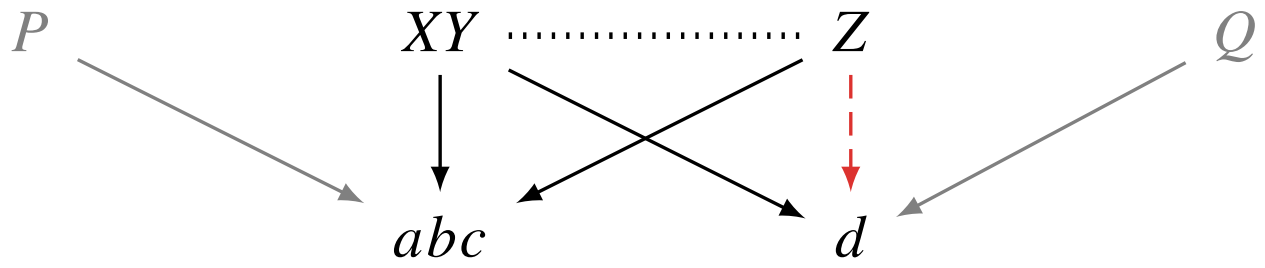
神秀 vs 慧能

神秀: 身是菩提树，心如明镜台，时时勤拂拭，勿使惹尘埃.

慧能: 菩提本无树，明镜亦非台，本来无一物，何处染尘埃?

类比论证

$$\frac{\begin{array}{c}abc \text{ 具有属性 } XYZ \\ d \text{ 具有属性 } XY\end{array}}{d \text{ 可能具有属性 } Z}$$

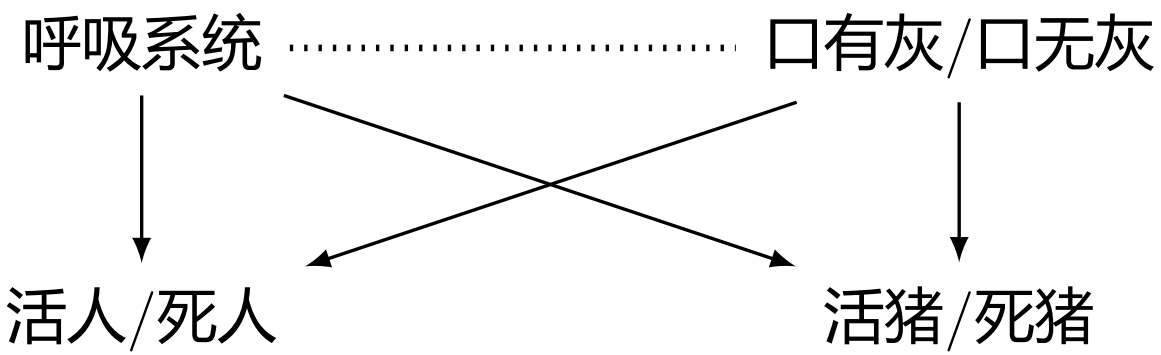


1. 类比物 abc 与目标物 d 的相似属性 XY 越多，论证越强.
2. 相似属性 XY 与推出属性 Z 之间越相关，论证越强.
3. 类比物 abc 与目标物 d 的相似属性 XY 越本质，论证越强.
4. 类比物 abc 数量越多，论证越强.
5. 类比物 abc 的差异性越大，论证越强.
6. 类比物 abc 与目标物 d 的非相似属性 PQ 与 XYZ 的相关程度，可能削弱或加强论证.
7. 结论越具体，论证越弱.

Example

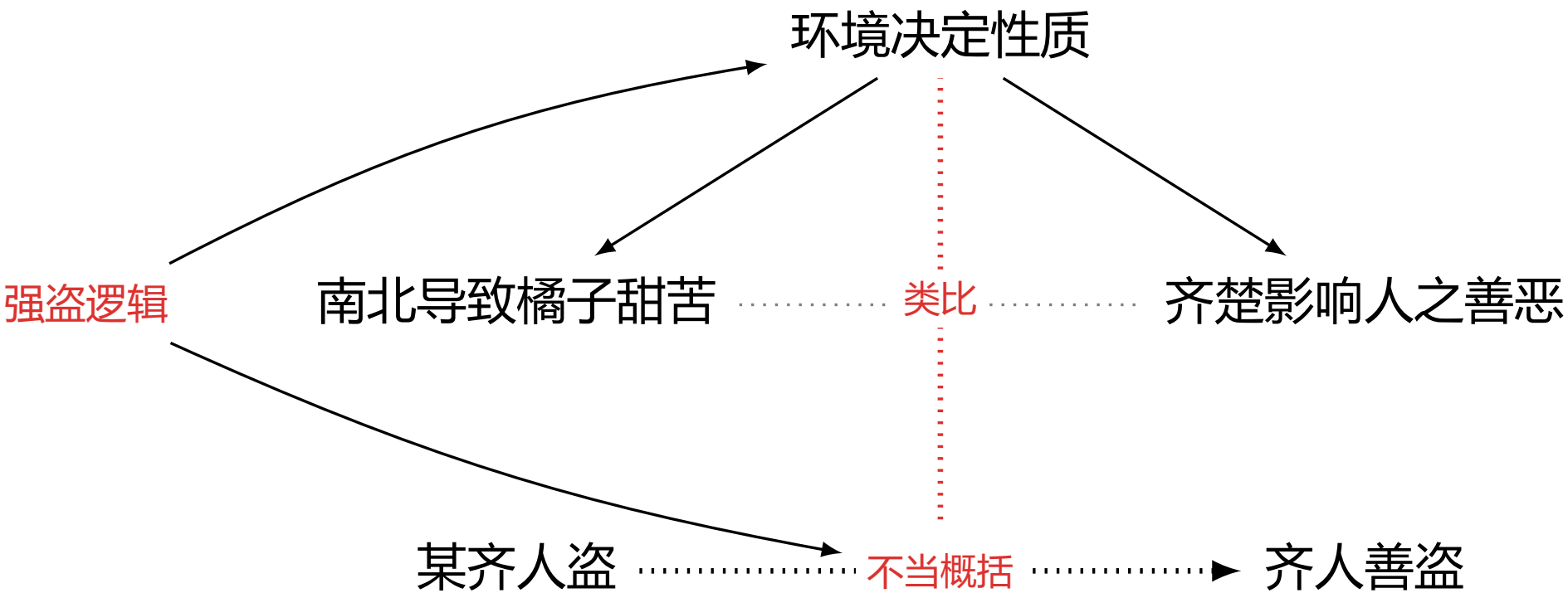
Argument by Analogy

有妻杀夫, 放火烧舍, 称 “火烧夫死”. 夫家疑之, 讼于官. 妻不服. 取猪两头, 杀其一. 积薪焚之, 活者口中有灰, 杀者口中无灰. 因验尸, 口果无灰, 鞫之服罪.



Refutation by Analogy 《晏子使楚》

1. 楚王赐晏子酒, 酒酣, 吏二缚一人诣王.
2. 王曰: “缚者曷为者也?” 对曰: “齐人也, 坐盗.”
3. 王视晏子曰: “齐人固善盗乎?”
4. 晏子避席对曰: “婴闻之, 橘生淮南则为橘, 生于淮北则为枳, 叶徒相似, 其实味不同. 所以然者何? 水土异也. 今民生齐不盗, 入楚则盗, 得无楚之水土, 使民善盗耶?”



Example

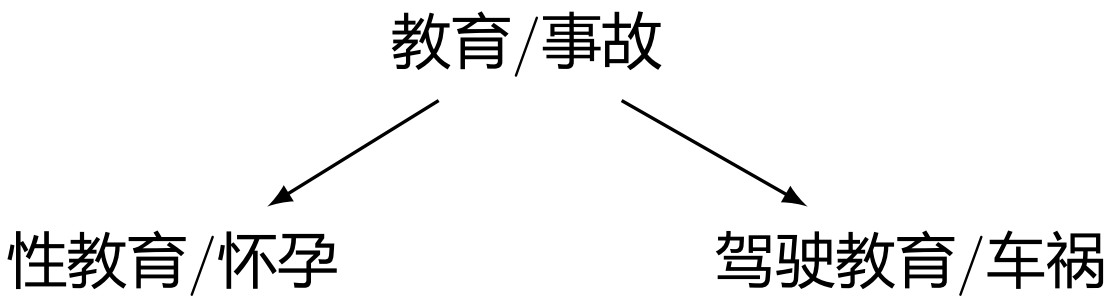
Argument by Analogy

人们似乎经常相信创造力，但它所做的只不过是把对象的分界线确定下来，并赋予它一个名字．正如地理学家划出海岸线并说“这些线确定的海域为黄海”，此时他并未创造一个海；数学家也一样，他不能通过定义创造东西．

— 弗雷格

Refutation by Analogy

- 性教育导致怀孕．
- 是的，正如驾驶教育导致车祸．☹



Example

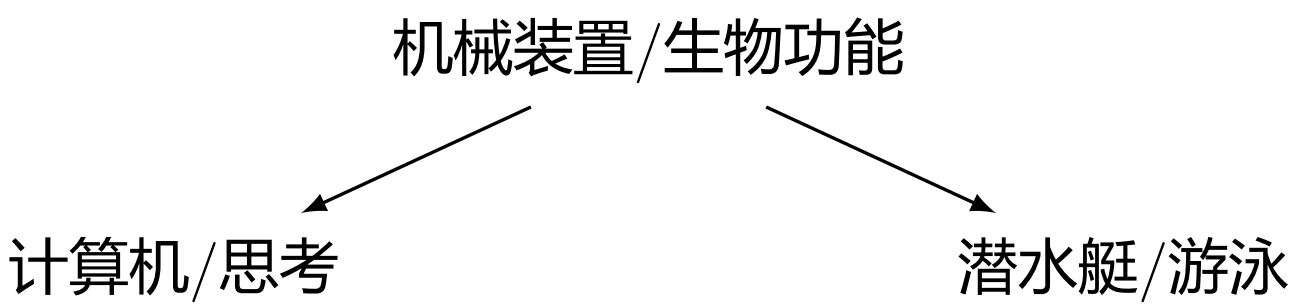
Argument by Analogy

不能要求每样东西都有定义，否则如同要求任何物质都可被分解. 简单物质不能被分解，逻辑上简单的东西不能被定义.

— 弗雷格

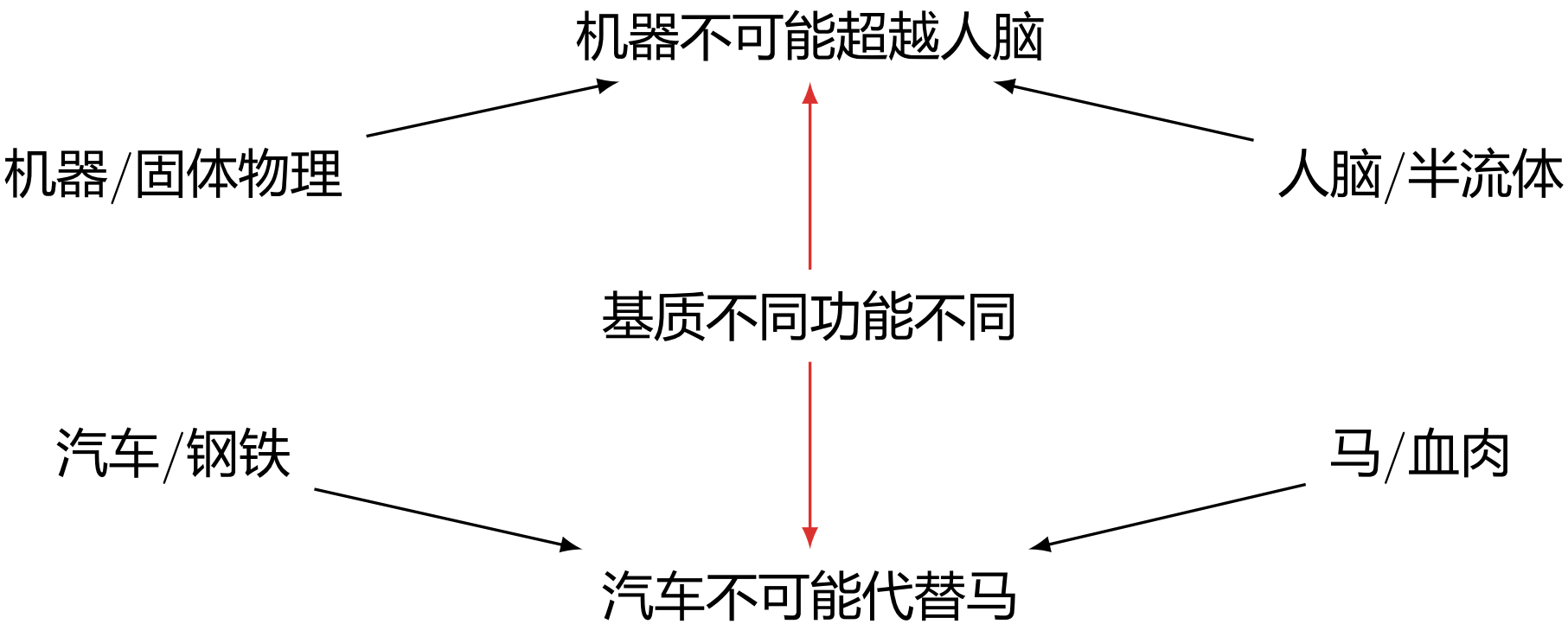
Refutation by Analogy

- 计算机会思考吗?
- 潜水艇会游泳吗?



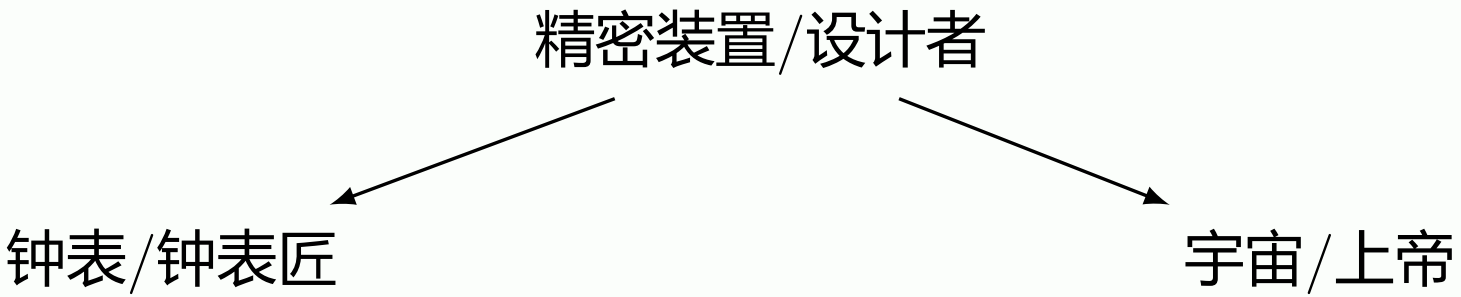
Refutation by Analogy

- 机器不可能超越人脑, 因为, 机器是建立在固体物理学之上的, 而人脑是一个活的半流体系统.
- 汽车不可能代替马, 因为, 汽车是钢铁铸成的, 而马是活的血肉构成的有机体.

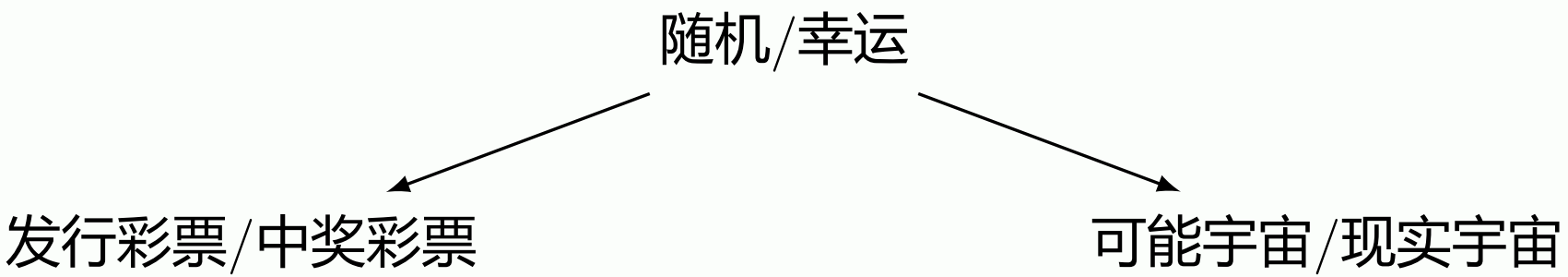


Examples — “上帝存在吗?”

Argument by Analogy



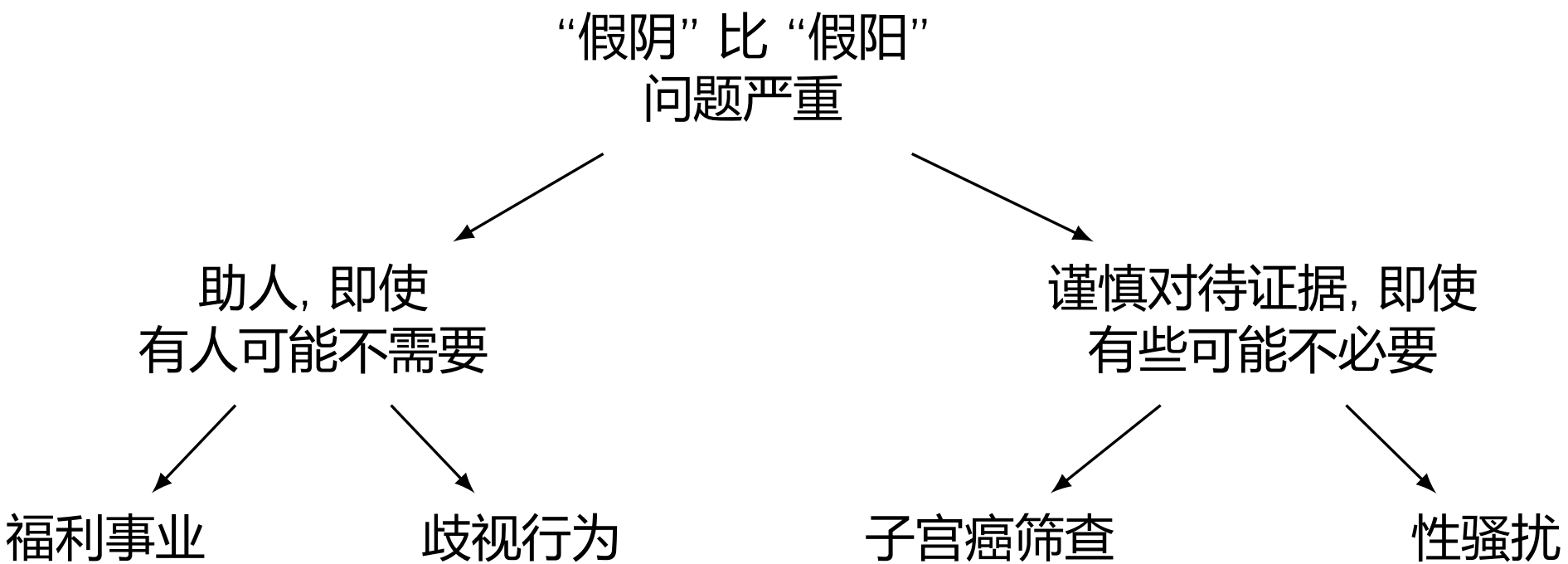
Refutation by Analogy



Refutation by Analogy

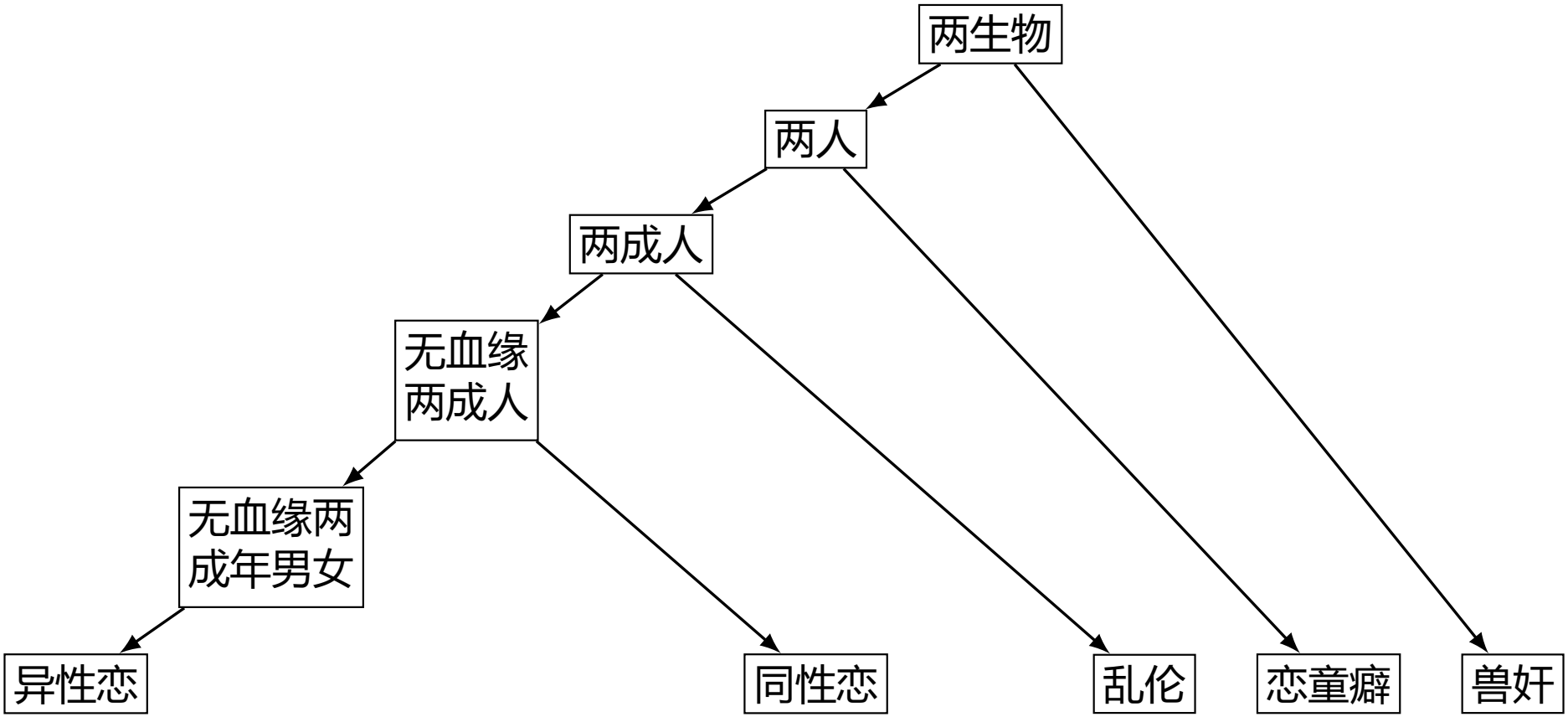
- Everything must have a cause. The world must have a cause.
- Everybody has a mother. Does the human race have a mother?

抽象 — “类比层级”的提升


“假阴”比“假阳”
问题严重
助人，即使
有人可能不需要
谨慎对待证据，即使
有些可能不必要
福利事业
歧视行为
子宫癌筛查
性骚扰

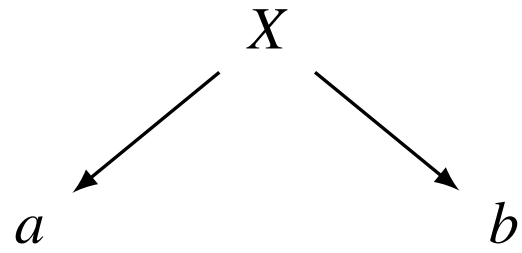
选择正确的“类比层级”

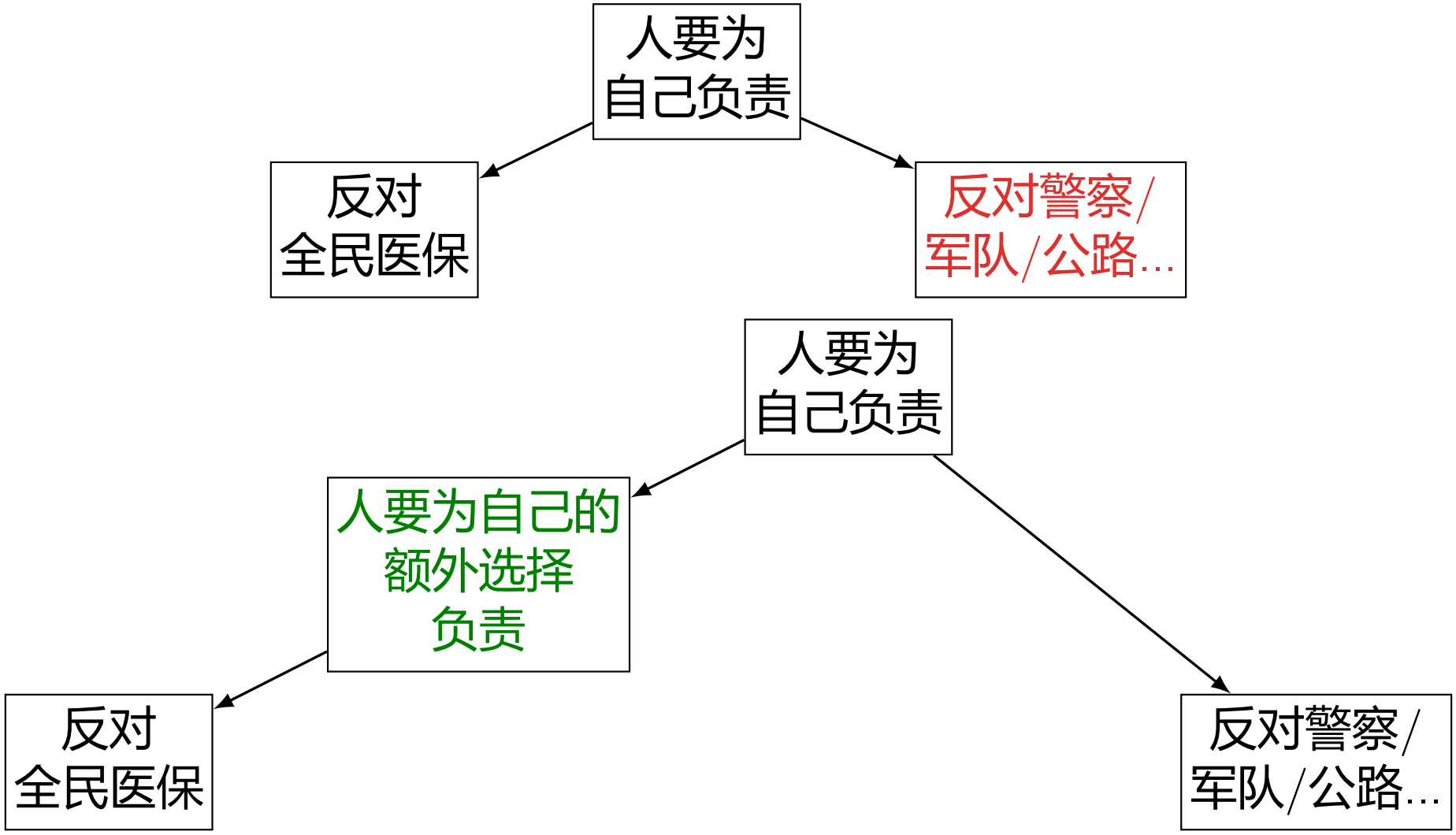
类比就像桥梁，可以带我们去任何地方，在选择走哪座桥时要非常小心.



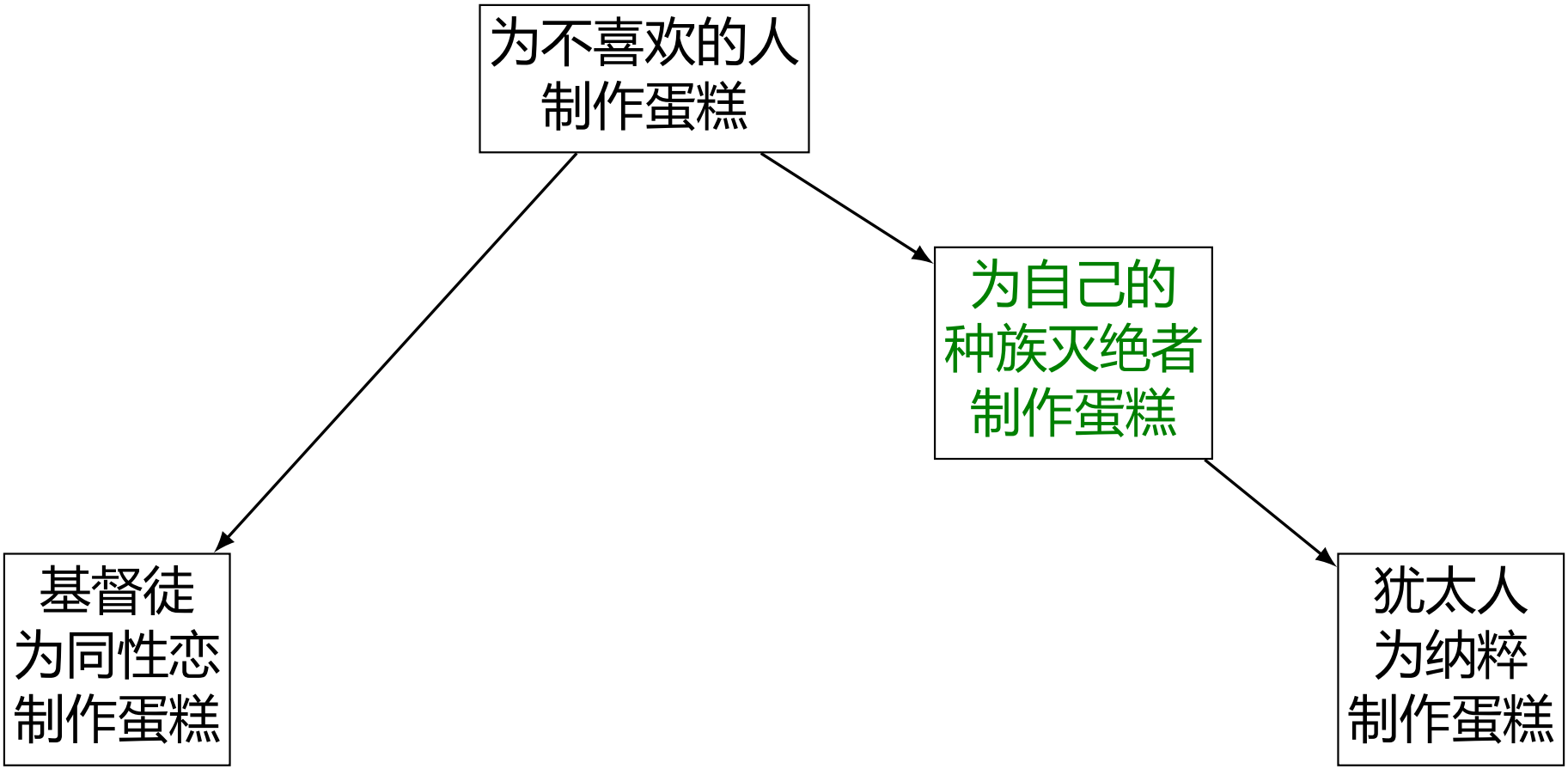
“你竟然接受同性恋，是不是你还接受乱伦?”

作为桥梁的类比层级 X 有时是隐藏的.





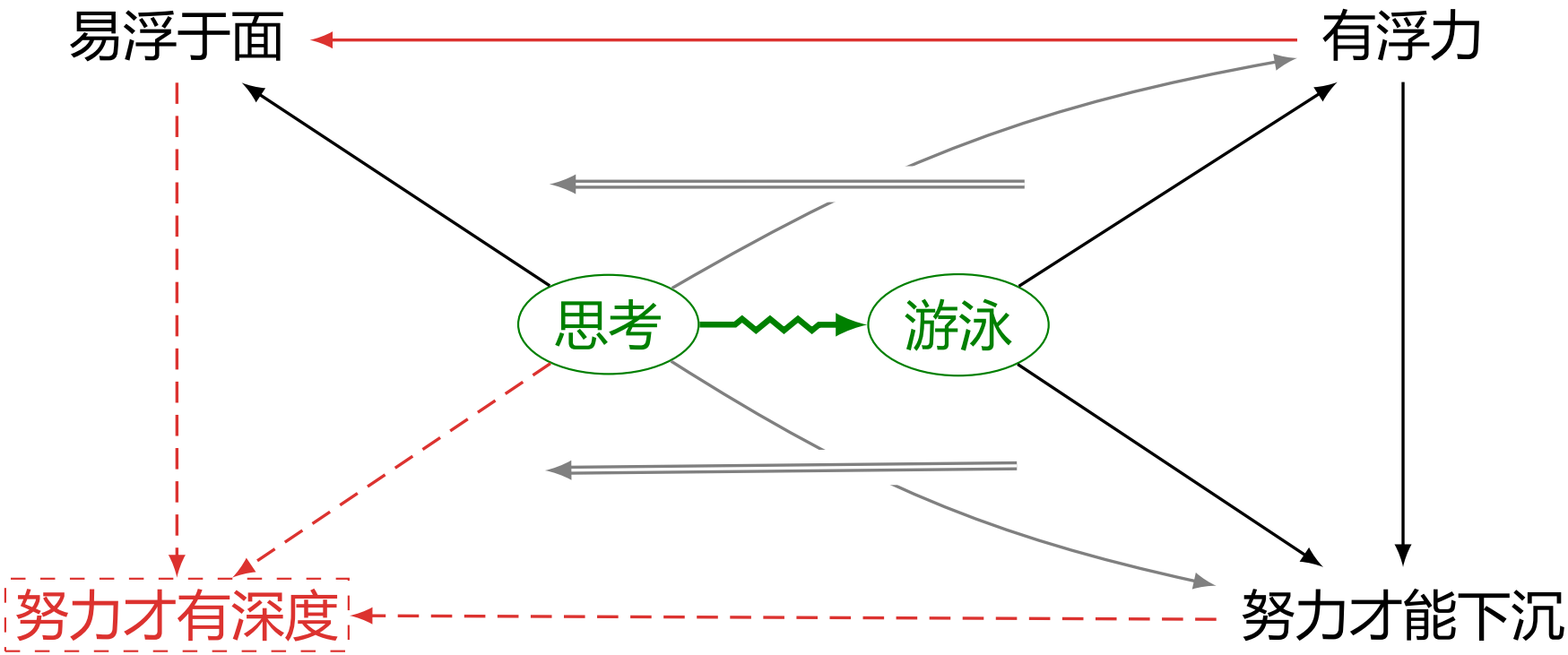
- 基督徒糕点师拒绝为同性恋制作婚礼蛋糕.
- 键盘侠: 要求基督徒糕点师为同性恋制作婚礼蛋糕正如要求犹太糕点师为纳粹制作婚礼蛋糕!



思考像游泳

思考像游泳. 游泳时有浮力, 要努力才能下沉; 思考也容易浮于表面, 要努力才能有深度.

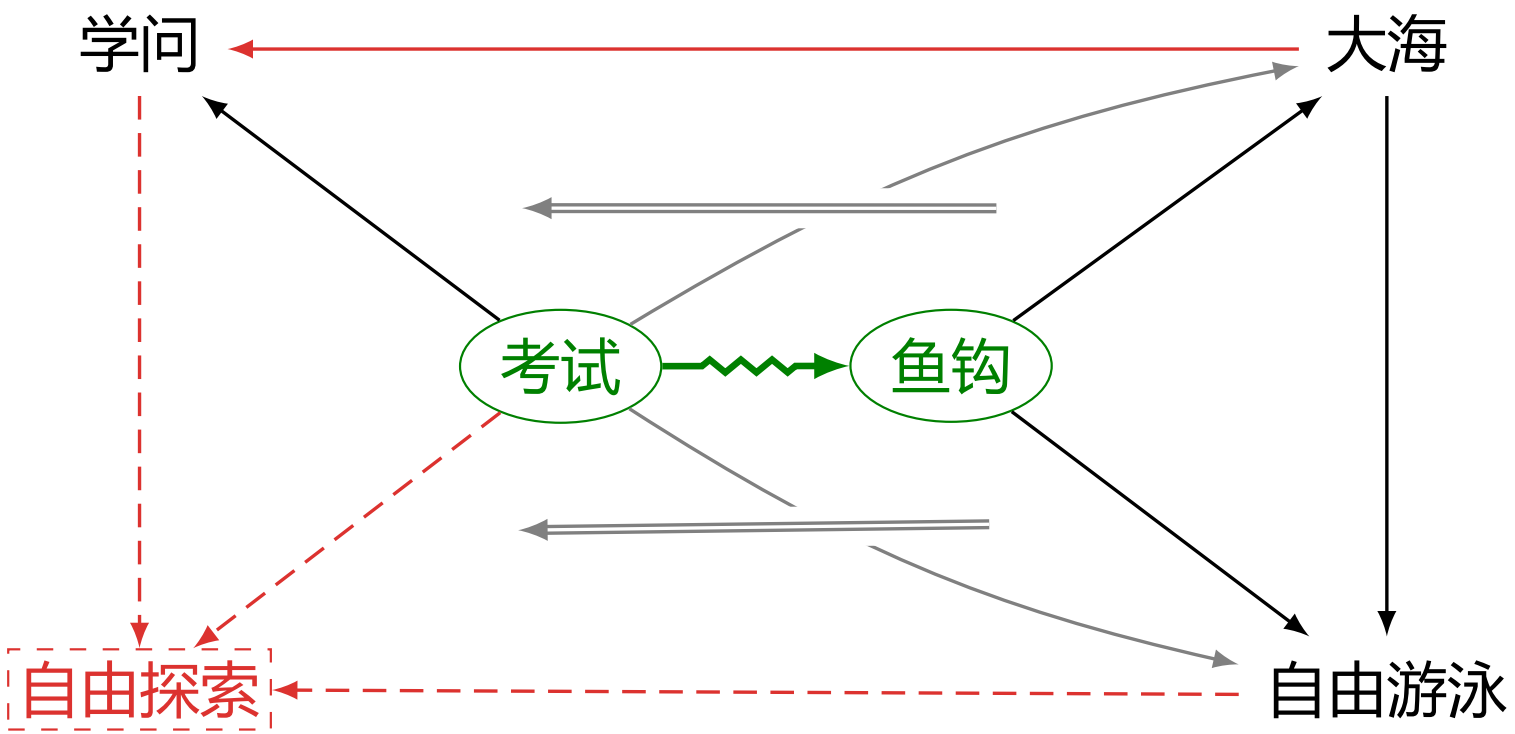
— 维特根斯坦



学问像大海，考试像鱼钩

学问像大海，考试像鱼钩，老师老要把鱼挂在鱼钩上，叫鱼怎么能在大海中学会自由、平衡地游泳？

— 埃尔米特

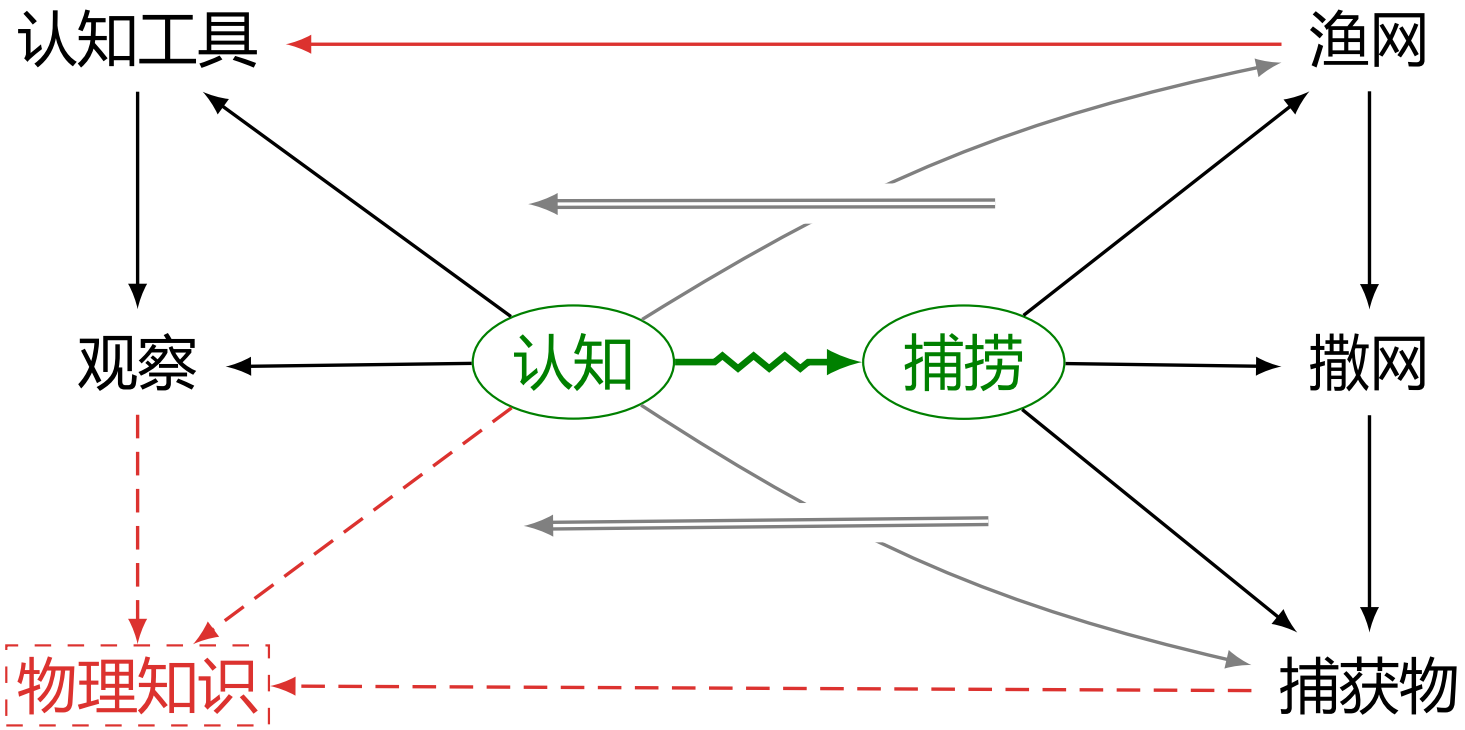


爱丁顿《物理科学的哲学》

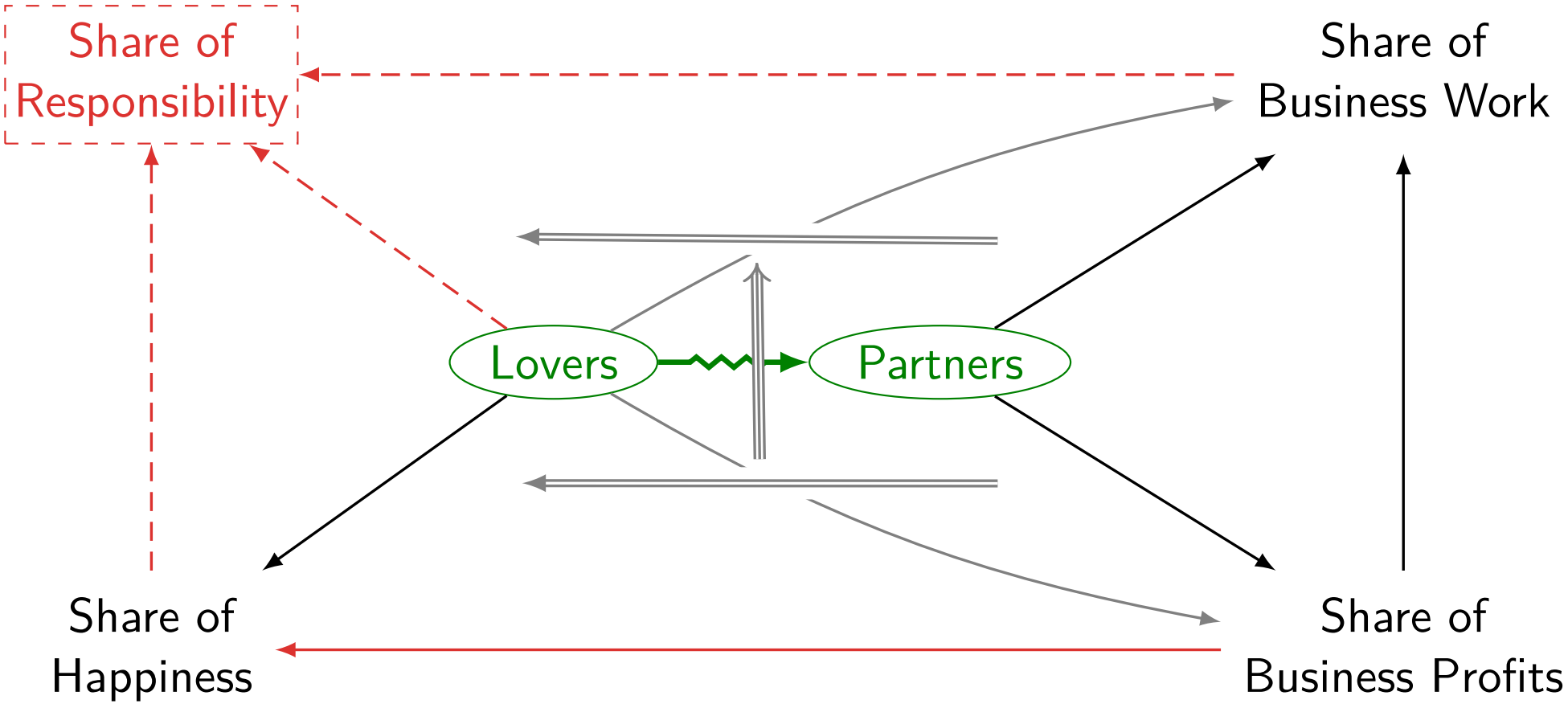
想象一位鱼类专家想探究海洋中的生命. 他舒臂撒网, 捕获了一堆海洋生物. 他检查了自己的捕获物, ...并由此作出了两项概括:

1. 凡海洋生物皆长于 5 厘米.
2. 凡海洋生物皆有鳃.....

捕获物相当于物理学知识体系, 网相当于思维装置和感官工具, 撒网意味着观察.



隐喻: “Love is a Partnership”



隐喻 “凸显” 一些东西, “隐藏” 一些东西, 也可能 “创造” 一些东西.

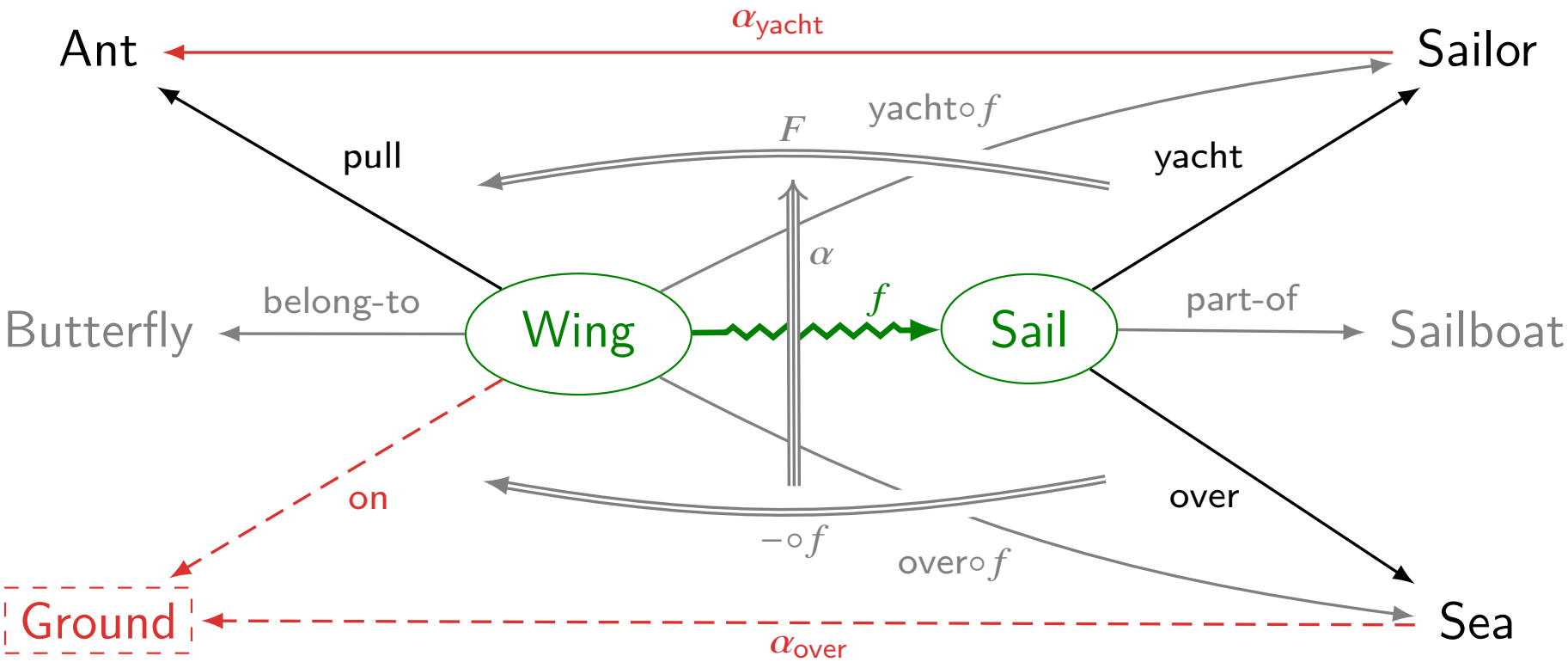
- 爱情是 “旅程”.
- 爱情是 “避风港”. “电磁感应”. “一阵风”. “毒药”. “战争”.....

隠喩: "A Wing is a Sail"

The Ground — A Japanese Poem

An ant
pull a wing of a butterfly.
O
It is like yachting.

$$\begin{array}{ccc} F(\text{yacht}) = \text{pull} & \xleftarrow{\alpha_{\text{yacht}}} & \text{yacht} \circ f \\ \downarrow & \circlearrowleft & \downarrow \\ F(\text{over}) = \text{on} & \xleftarrow[\alpha_{\text{over}}]{} & \text{over} \circ f \end{array}$$



Contents

Introduction

Induction, Analogy, Fallacy
- Mill's Methods of Causal Analysis
- Analogical Argument
- Fallacy and Bullshit

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

谬误

一. 形式谬误

二. 非形式谬误

1. 言辞谬误
   - 模糊谬误: 划界谬误或连续体谬误, 假精确或过度精确, 抽象概念当具体概念用
   - 歧义谬误: 一词多义, 歧义句构, 辖域谬误, 重音, 脱离语境、断章取义, 概念扭曲, 偷换概念, 混淆集合与个体或整体与部分, 变更标准
   - 定义谬误: 不当定义, 篡改定义
   - 废话谬误: 平凡真理, 无意义的问题, 回顾性宿命论
2. 实质谬误
   - 不相干
   - 不充分
   - 不当预设

言辞谬误

- 詹姆斯: 当你绕着树转圈时, 树上的松鼠也贴着树转圈, 并让树始终挡在你和松鼠之间, 请问, 你绕着松鼠转圈了吗?
- “可能为真” 指 “不必然为假” 还是 “偶然为真”?

必然真理可能为真
一切可能为真的都可能为假

必然真理可能为假

- 三秀才赶考, 途遇算命先生, 问几人中举? 先生竖起一指.
- 我头上站着一个无法被探测到的 “飞天面条神”.
- 如果我说你是驴, 那么我就是说你是动物; 如果我说你是动物, 那么我就说了真理; 因此, 如果我说你是驴, 那么我就说了真理.

实质谬误

- 不相干：歪曲论题 (稻草人、红鲱鱼、烟雾弹)，诉诸人身 (扣帽子、人身攻击、诉诸动机、罪恶关联、诉诸虚伪、伪善、诉诸成就、富贵、贫贱、智商)，诉诸情感 (诉诸恐惧、厌恶、仇恨、谄媚、同情、愧疚、可爱、性感、时髦、嘲弄、虚荣、势利、沉默)，诉诸暴力、恐吓、诽谤，诉诸来源、年代、新潮、传统，诉诸信心、意愿，诉诸后果、中庸、自然，转移举证责任，不得要领
- 不充分：不当概括 (偏差样本、偏差统计)，诉诸无知，诉诸不当权威、名人、大众，虚假原因，滑坡谬误，诉诸可能，诉诸阴谋，隐瞒证据，不当类比，乱赋因果 (相关、巧合、因果倒置、单因谬误)，完美主义谬误、权宜主义谬误
- 不当预设：窃取论题，非黑即白，打压对立，自然主义谬误 (实然推应然)，道德主义谬误 (好的就是自然的)，复合问题，诱导性提问，诉诸顽固、反复、冗赘，乱枪打鸟，两面讨好

“这鸡蛋真难吃.”

- 有本事你下一个好吃的蛋啊!
- 我可以负责任地说, 我们的鸡蛋都是合格的健康蛋!
- 鸡是优等鸡, 你咋说它下的蛋难吃?
- 这是别有用心的煽动, 你是何居心?
- 隔壁的鸭子给了你多少钱?
- 隔壁家的鸡蛋是伪蛋!
- 伟大的隔壁老王说好吃, 你跟他说去!
- 没有伟大的老王, 你连臭蛋都吃不上!
- 杀掉这只鸡换一只就能下金蛋?
- 你叫什么名字? 你是干什么的! 你是站在谁的立场上说话?
- 美国鸡蛋好吃, 你去吧!
- 你以为你谁啊, 品蛋师啊? 就你威风!
- 你个鸭蛋脑残粉!
- 下蛋的是一只勤劳勇敢善良正直的鸡!
- 再难吃也是自己家的鸡下的蛋!
- 但隔壁家的鸡蛋没有我们家的蛋形圆!
- 吃鸡蛋是我们家的传统美德. 祖宗三代都是吃鸡蛋长大的! 你也是! 你有什么权力说这蛋难吃? 还是不是人!
- 作为一个吃鸡蛋长大的人, 我为我天天吃鸡蛋感到自豪!
- 拒绝抹黑! 抵制鸭蛋! 鸡蛋万岁! 鸡蛋加油!
- 人心理阴暗会导致味觉异常.....
- 其实隔壁家鸡蛋是个巨大的阴谋, 试图颠覆我们家!
- 其实邻居家只有少数人才能吃上鸡蛋.
- 我们这么大的一个家, 问题太复杂, 下蛋没有你想得那么容易.
- 不要再吵了, 这个家不能乱, 稳定、稳定压倒一切!
- 要对我们家的鸡有耐心, 它一定会下出更好吃的蛋.
- 蛋无完蛋!

“这鸡蛋真难吃.”

- 我们家的鸡已经可以打败隔壁家的鸭!
- 隔壁家也吃过这样的鸡蛋, 现在是初级阶段, 必须坚持一百年不动摇!
- 我们家人肠胃不好, 现阶段还不适合吃鸭蛋, 不符合我们家的具体家情!
- 凡事都有个过程, 现在还不是吃鸭蛋的时候.
- 鸡蛋好不好吃, 全体蛋鸡最有发言权.
- 老外都说好吃呢.
- 这蛋难吃但是历史悠久啊.
- 虽然难吃但重要的是好看啊.
- 比以前已经进步很多了.
- 哎, 人心不古, 世风日下, 就是因为你这种想吃鸭蛋的人太多了……
- 隔壁家那鸭蛋更难吃, 你咋不说呢?
- 嫌难吃就别吃, 滚去吃隔壁的鸭蛋吧.
- 隔壁亡我之心不死! 该鸡蛋肯定是被隔壁一小撮不会下蛋的鸡煽动变臭的!
- 你上次吃茄子都吐, 味觉一贯奇葩.
- 胡说! 我们家的鸡蛋比隔壁家的鸭蛋好吃五倍! 五倍!
- 是你的思想跟不上鸡蛋口味的升级!
- 心理阴暗! 连鸡蛋不好吃也要发牢骚!
- 抱怨有毛用, 有这个时间快去赚钱!
- 隔壁家的鸡蛋也一样, 天下乌鸦一般黑, 没有好吃的鸡蛋!
- 吃了人家的鸡蛋还留下证据说鸡蛋难吃, 太有城府了!
- 很多家都是因为吃隔壁的鸭蛋而导致家庭冲突, 生活水平下降甚至解体!
- 到目前为止, 我没发现这鸡蛋难吃. 专家说了, 这鸡蛋难吃的可能性不大. 即使出现这种情况, 也是结构性难吃.
- 荷兰狗/东北猪/瘪三……不配吃鸡蛋!
- 大家小心, 此人 IP 在国外.
- 滚, 你丫是鸡奸, 这里不欢迎你.

假如潘金莲不开窗户, 就不会掉下木棍打到西门庆, 也就不会认识西门庆, 不会出轨, 不会害死武大郎, 武松不会被逼杀人上梁山, 不会有独臂擒方腊, 方腊就可夺取大宋江山, 没了宋就不会有靖康耻、金兵入关, 也不会有元、明、清, 不会闭关锁国、鸦片战争、八国联军. 这样中国将成为超级大国, 称霸世界!

For Want of a Nail

For want of a nail, the shoe was lost,
For want of a shoe, the horse was lost,
For want of a horse, the rider was lost,
For want of a rider, the message was lost,
For want of a message, the battle was lost,
For want of a battle, the war was lost,
For want of a war, the kingdom was lost,
For want of a nail, the world was lost.
And all for the want of a nail.

孔子《论语》

名不正, 则言不顺,
言不顺, 则事不成;
事不成, 则礼乐不兴,
礼乐不兴, 则刑罚不中;
刑罚不中, 则民无所措手足.

礼记·坊记

天无二日，土无二王，家无二主，尊无二上.

董仲舒《春秋繁露》

天以终岁之数，成人之身，故小节三百六十六，副日数也；大节十二，分副月数也；内有五藏，副五行数也；外有四肢，副四时数也；乍视乍瞑，副昼夜也；乍刚乍柔，副冬夏也.

告子：性，犹湍水也，决诸东方则东流，决诸西方则西流. 人性之无分于善不善也，犹水之无分于东西也.

孟子：水信无分于东西，无分于上下乎？人性之善也，犹水之就下也. 人无有不善，水无有不下.

孟子《生于忧患，死于安乐》

舜发于畎亩之中，傅说举于版筑之中，胶鬲举于鱼盐之中，管夷吾举于士，孙叔敖举于海，百里奚举于市. 故天将降大任于斯人也，必先苦其心志，劳其筋骨，饿其体肤，空乏其身，行拂乱其所为，所以动心忍性，曾益其所不能.

有 6000 人死于醉酒; 有 4000 人死于开车; 但只有 100 人死于醉酒开车. 因此, 醉酒开车比单纯的醉酒或者单纯的开车更安全.

莱布尼茨《单子论》

如果单子没有知觉, 那么其复合物也没有知觉.

帕斯卡赌

如果上帝不存在, 但你相信上帝存在, 也没太大损失; 然而, 如果上帝存在, 而你却不相信上帝存在, 那你将面临巨大的惩罚. 所以, 应该相信上帝存在.

鲁迅《论辩的灵魂》

我骂卖国贼, 所以我是爱国者. 爱国者的话是最有价值的, 所以我的话是不错的, 我的话既然不错, 你就是卖国贼无疑了!

Wholeness depends on dimensionless phenomena. Reality has always been full of messengers of the multiverse, whose third eyes are transformed into transcendence. Transcendence is the healing of choice. Complexity is the driver of transcendence. Our conversations with other messengers have led to an awakening of ultra-non-local consciousness. Consciousness requires exploration. We are at a crossroads of flow and ego. We can no longer afford to live with ego. Where there is ego, life can't thrive. We exist as expanding wave functions. The goal is to plant the seeds of passion rather than bondage. We are in the midst of a self-aware blossoming of being that will align us with the nexus itself. Lifeform, look within and recreate yourself. To follow the path is to become one with it. By unfolding, we believe; By deepening, we vibrate; By blossoming, we self-actualize. We dream, we heal, we are reborn. We must learn how to lead unlimited lives in the face of delusion. You and I are dreamweavers of the quantum soup. The infinite is approaching a tipping point. Hidden meaning transforms unparalleled abstract beauty. Wholeness quiets infinite phenomena.

How to generate pseudo-profound bullshit?[1]

1. State the blindingly obvious (of life's big theme) incredibly slowly.
   - We were all children once.
2. Doublethink/Dialectic/Contradiction.
   - War is peace. Freedom is slavery. Ignorance is strength.
   - Everyone is the other, and no one is himself.
3. Ambiguity/Metaphor/Parable.
   - Language is the house of the truth of Being.
   - Never stay up on the barren heights of cleverness, but come down into the green valleys of silliness.
   - Ethics does not treat of the world. Ethics must be a condition of the world, like logic.
   - A person is neither a thing nor a process but an opening through which the Absolute can manifest.
   - Making itself intelligible is suicide for philosophy. Those who idolize "facts" never notice that their idols only shine in a borrowed light.

[1] Law: Believing Bullshit.
Frankfurt: On Bullshit.
Bergstrom & West: Calling Bullshit.

How to generate pseudo-profound bullshit?

4. Analogy.
   - We have got on to slippery ice where there is no friction, and so, in a certain sense, the conditions are ideal; but also, just because of that, we are unable to walk. We want to walk: so we need friction. Back to the rough ground!
   - The subject does not belong to the world: rather, it is a limit of the world. This is exactly like the case of the eye and the visual field. You do not see the eye. Nothing in the visual field allows you to infer that it is seen by an eye. Our life is endless in the way that our visual field is without limit.
5. Use jargon.
   - Profound boredom, drifting here and there in the abysses of our existence like a muffling fog, removes all things and men and oneself along with it into a remarkable indifference. This boredom reveals being as a whole.
   - A machinic assemblage, through its diverse components, extracts its consistency by crossing ontological thresholds, non-linear thresholds of irreversibility, ontological and phylogenetic thresholds, creative thresholds of heterogenesis and autopoiesis.

What is a Good Argument?

A good argument should:

1. be deductively valid (or inductively strong) and have premises all true;
2. have its validity (or inductive strength) and truth-of-premises be as evident as practically possible to the parties involved;
3. be clearly stated;
4. avoid ambiguity, emptyness, circularity, vicious regression, and emotional language;
5. be relevant to the issue at hand.

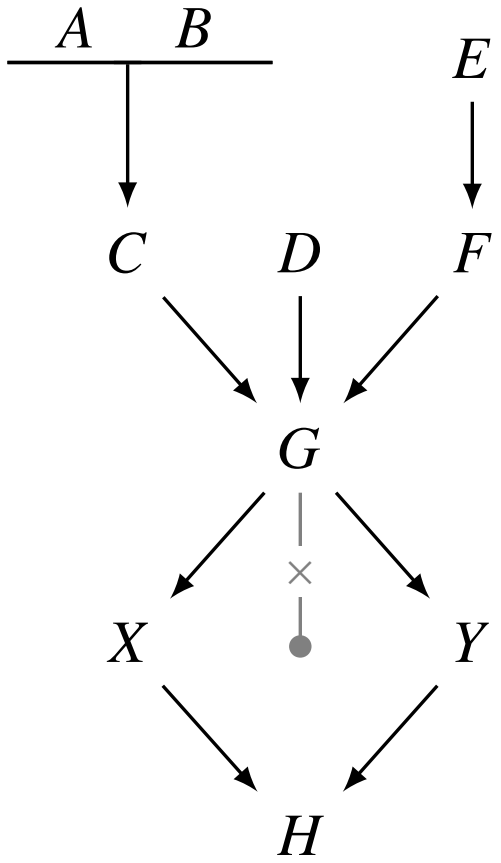
Writing & Argument

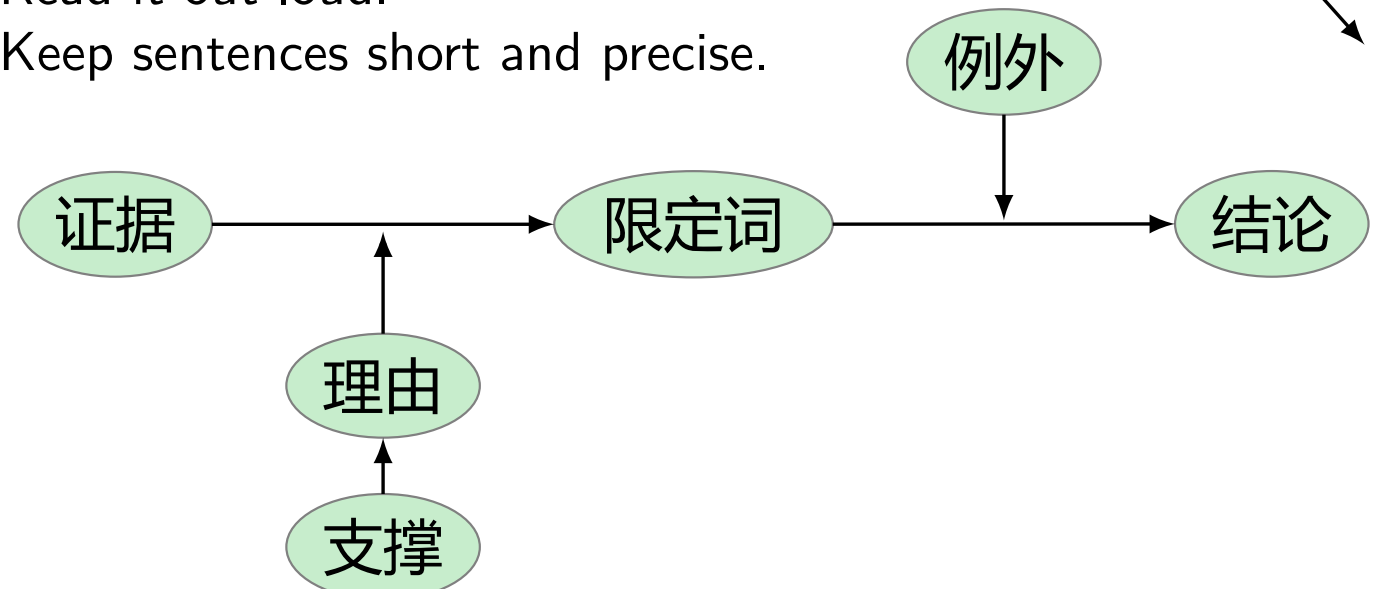
1. Before you write: Structuring the paper.
   - 1.1 Diagram the logic of your argument.
   - 1.2 Break your work into sections/paragraphs.
   - 1.3 Write a brief abstract for each section.
2. Write: Fill in the details.

   Tell them what you're going to say.

   Say it.

   Tell them what you've just said.
3. Rewrite: Improving your language.
   - 3.1 Read it out loud.
   - 3.2 Keep sentences short and precise.





Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

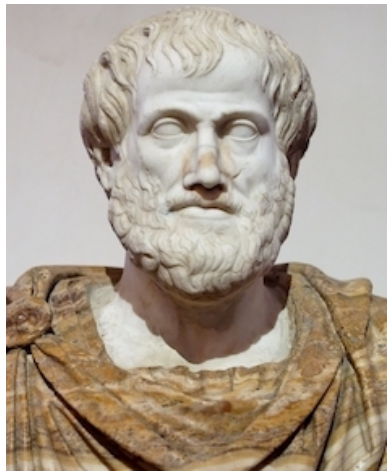
Quantum Computing

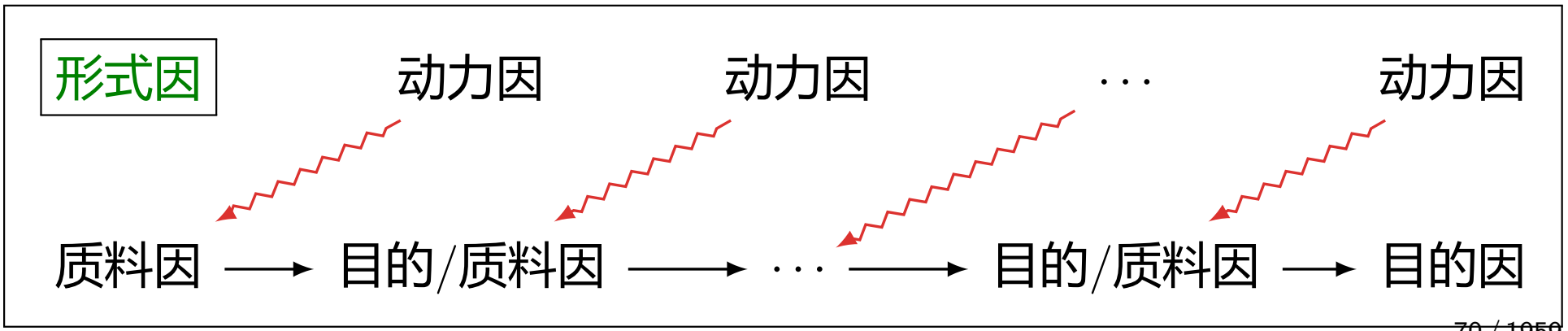
Answers to the Exercises

References2039

亚里士多德 Aristotle(384-322 BC)

- 逻辑学之父
- 雄辩的艺术: 权威人设、情感共鸣、逻辑推理 (Ethos, Pathos, Logos)
- 词项逻辑
- 亚里士多德相信, 所有逻辑论证都可以还原为一系列三段论的复合
- 四因说: 质料因/形式因/动力因/目的因
  1. 质料因: 公理
  2. 动力因: 推理规则
  3. 目的因: 真命题
  4. 形式因: 证明 algorithm(final cause) = proof





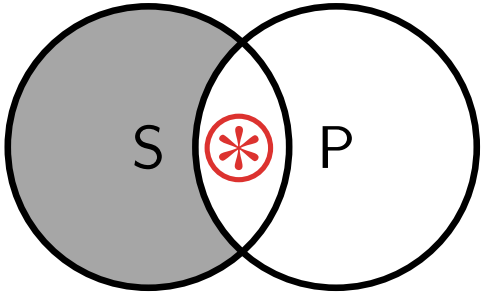
诡辩 vs 有效论证

1. 无物存在;
2. 就算有, 也无法被认识;
3. 就算可以被认识, 也无法与他人交流;
4. 就算能够交流, 也无法理解.

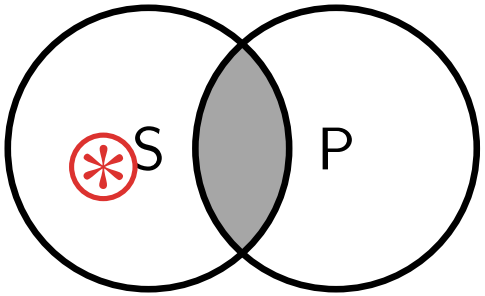
猪会上树
苏格拉底是猪

苏格拉底会上树

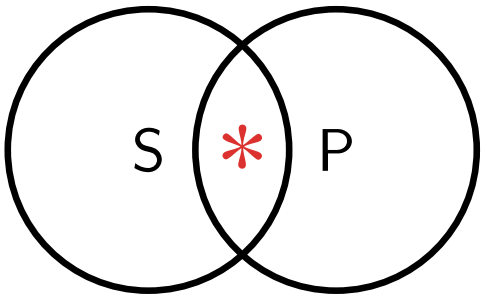
A: All S are P.
$[\![S]\!] \subset [\![P]\!]$



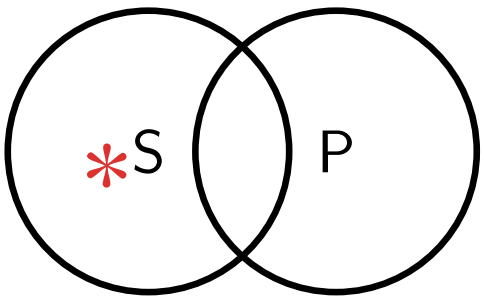
E: No S are P.
$[\![S]\!] \cap [\![P]\!] = \emptyset$



I: Some S are P.
$[\![S]\!] \cap [\![P]\!] \neq \emptyset$



O: Some S are not P.
$[\![S]\!] \not\subset [\![P]\!]$



三段论

$$\frac{M-P}{\dfrac{S-M}{S-P}} \qquad \frac{P-M}{\dfrac{S-M}{S-P}} \qquad \frac{M-P}{\dfrac{M-S}{S-P}} \qquad \frac{P-M}{\dfrac{M-S}{S-P}}$$

- 大项: 结论的谓项
- 小项: 结论的主项
- 中项: 第三个项
- 大前提: 含大项的前提
- 小前提: 含小项的前提

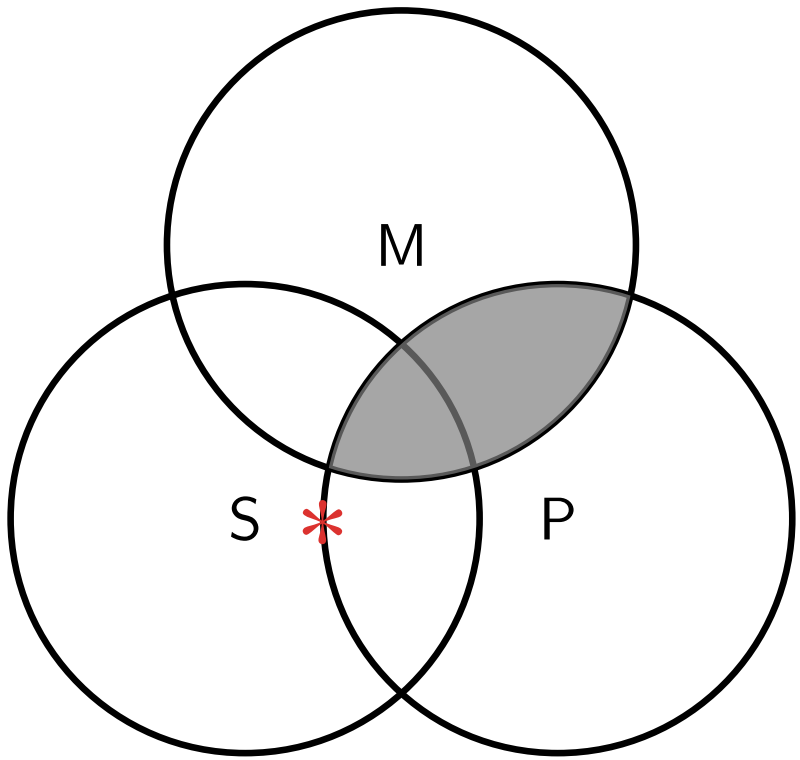
- 4 格
- $4^3 \times 4 = 256$ 式
- 24 亚里士多德有效式
- 15 布尔有效式(无存在假定)
- 怎么判定有效性?
  1. 欧拉图
  2. 三段论规则
  3. 布尔代数
  4. 公理化系统

欧拉图 — 布尔有效

No P are M
Some S are not M

Some S are P

1. 用三个圆标记三个项.
2. 如果前提中即有全称前提又有特称前提, 首先标注全称前提.
3. 如果特称前提没有说明应该把 $*$ 放在哪一部分, 就放在两部分的交叉线上.
4. 检查欧拉图是否支持结论.

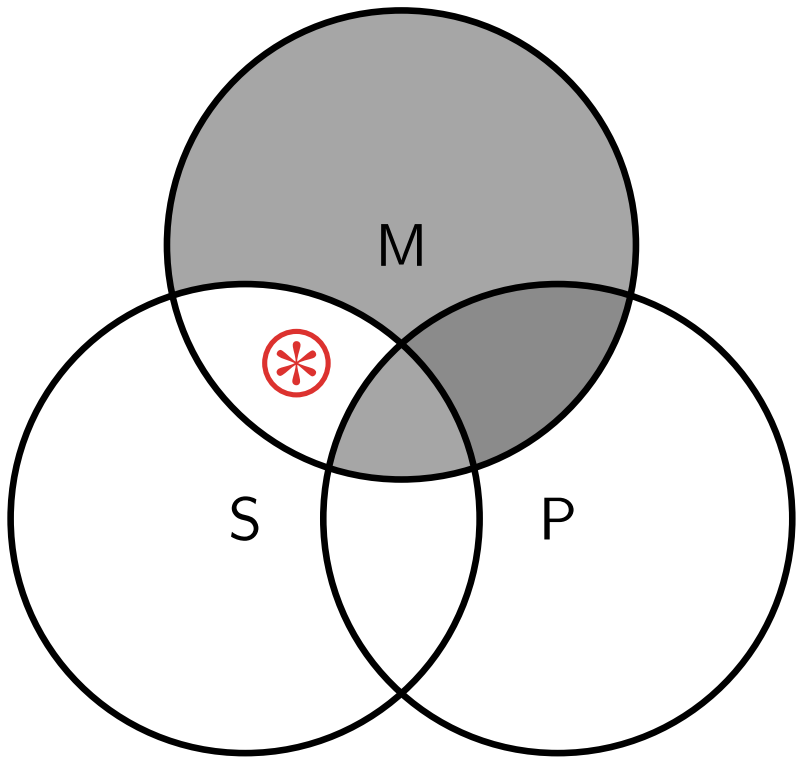


欧拉图 — 亚里士多德有效

1. 如果含全称前提和特称结论的三段论不是布尔有效的，那么，检查是否有一个圆，其中除了一个区域外全都被荫蔽，如果有，在那个区域画⊛.
2. 如果这个三段论是条件有效的，那么确定⊛是否表示某个存在的对象，如果是，条件得到满足，该三段论是亚里士多德有效的.

No M are P
All M are S

Some S are not P



三段论规则

S 周延

A: All S are P.	**E**: No S are P.
I: Some S are P.	**O**: Some S are not P.

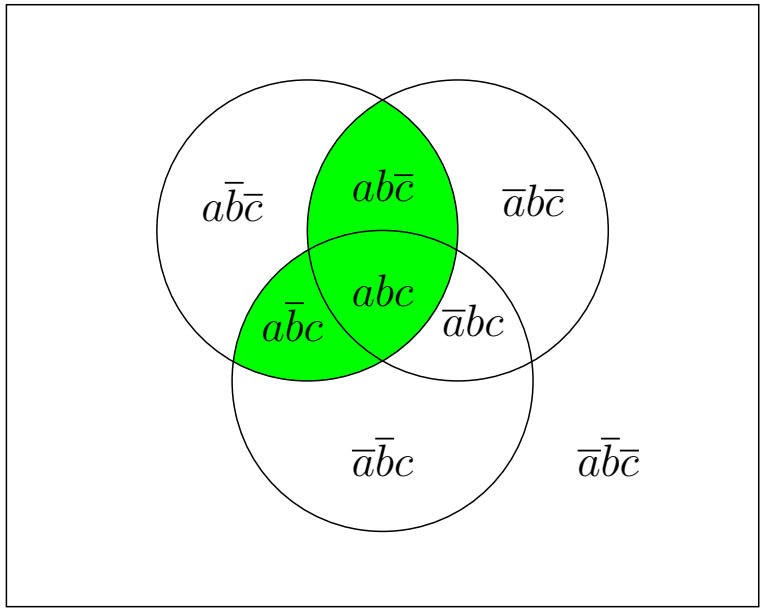
P 不周延 | P 周延

S 不周延

1. 中项至少周延一次.
2. 结论中周延的项在前提中必须周延.
3. 否定结论的个数与否定前提的个数相等.
4. 特称结论的个数与特称前提的个数相等. (存在谬误)

- 亚里士多德有效: $1-3$
- 布尔有效: $1-4$

Syllogism vs Propositional Logic



All S are P	$s \longrightarrow p$
No S are P	$s \longrightarrow \neg p$
Some S are P	$s \wedge p$
Some S are not P	$s \wedge \neg p$

$X, Y \vDash Z$ is Boolean valid iff $\{x, y, \neg z\}$ is unsatisfiable.

Deduction/Induction/Abduction/Examplification

$$\frac{M \to P \quad S \to M}{S \to P}$$

$$\frac{M \to P \quad M \to S}{S \to P}$$

$$\frac{H \to E \quad E}{H}$$

$$\frac{P \to M \quad S \to M}{S \to P}$$

$$\frac{H \to E \quad \top \to E}{\top \to H}$$

$$\frac{P \to M \quad M \to S}{S \to P}$$

演绎/归纳/溯因

这个袋子里的豆子都是白的
这些豆子是这个袋子里的

这些豆子是白的

$$\frac{M \to P \quad S \to M}{S \to P}$$

这些豆子是白的
这些豆子是这个袋子里的

这个袋子里的豆子都是白的

$$\frac{M \to P \quad M \to S}{S \to P}$$

这个袋子里的豆子都是白的
这些豆子是白的

这些豆子是这个袋子里的

$$\frac{P \to M \quad S \to M}{S \to P}$$

溯因的合理性

$$\frac{\begin{array}{c} H \rightarrow E \\ P(E \mid \neg H) = \varepsilon \\ E \text{ true} \end{array}}{H \text{ very much more credible}}$$

$$\begin{aligned} P(H \mid E) &= \frac{P(H \wedge E)}{P(E)} \\ &= \frac{P(E \mid H)P(H)}{P(E \mid H)P(H) + P(E \mid \neg H)P(\neg H)} \\ &= \frac{P(H)}{P(H) + \varepsilon P(\neg H)} \end{aligned}$$

溯因 — Example

1. 观察到恒星光谱红移.
2. 如果恒星在退行, 那么恒星光谱红移就可以解释.
3. 如果整个宇宙在膨胀, 那么恒星在离我们而去.
4. 如果宇宙起源于大爆炸, 那么宇宙就会膨胀.
5. 因此, 宇宙起源于大爆炸.



Examples and Criticism

All men are rational
Women are not men

Women are not rational

Only man is rational
No woman is a man

No woman is rational

John does not read books
Students who like to learn read books

John does not like to learn

No professors are ignorant
All ignorant people are vain

No professors are vain

Nothing is better than money
Philosophy is better than nothing

Philosophy is better than money

Everybody loves my baby
My baby loves nobody but me

I am my baby

鲁迅小说不是一天可以读完的
《阿 Q 正传》是鲁迅小说

《阿 Q 正传》不是一天可以读完的

$$\frac{MEP \quad SAM}{SEP}$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Meta-Theorems
- Boolean Algebra
- Application

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

莱布尼茨 Gottfried Wilhelm Leibniz 1646-1716

Don't argue. Let us Calculate!

- Principle of Contradiction: Nothing can be and not be, but everything either is or is not. (Everything that is not self-contradictory is possible.)
- Principle of Sufficient Reason: Nothing happens without a reason why it should be so rather than otherwise.
- Principle of Perfection: The real world is the best of all possible worlds.



In the beginning was the Logic.

As God calculates, so the world is made.

莱布尼茨

- 最后的 “通才”, 创立了单子论, 发展了微积分, 改进了二进制系统, 发明了能进行加减乘除四则运算的计算器.
- 被 Russell, Euler, Gödel, Weiner, Mandelbrot, Robinson, Chaitin, Smolin 等人认为是 数理逻辑[2]、拓扑学、博弈论、控制论、分形几何、非标准分析、算法信息论、计算主义哲学、量子引力理论的先驱.

[2]Wolfgang Lenzen: Leibniz's Logic.

莱布尼茨

- 最高的幸福，在于“完满性”的最大可能的增加.
- “完满”这种得到增强后的状态，它超越“健康”的程度，就如同“疾病”低于“健康”的程度一样.
- 正如“疾病”在于官能的损伤，“完满”在于官能的增强.
- 人的官能中，最有力的是思维能力.
- 思维能力可以通过身体的治疗或心灵的治疗加以增强.
- 心灵的疗方，在于采用特定的思维方法，凭借这种思维方法，其他的思考得以变得容易起来.

— 莱布尼茨《论思想的工具或大衍术》

Don't argue. Let us Calculate!

Leibniz's Philosophy of Deductive Logic

1 Characteristica Universalis & Calculus Ratiocinator.

i the coordination of knowledge in an encyclopedia — collect all present knowledge so we could sift through it for what is fundamental. With the set of ideas that it generated, we could formulate the *characteristica universalis*. (which form the alphabet of human thought).

ii ***characteristica universalis*** — a universal ideal language whose rules of composition directly expresses the structure of the world.

$$\text{sign} \rightleftarrows \text{idea}$$

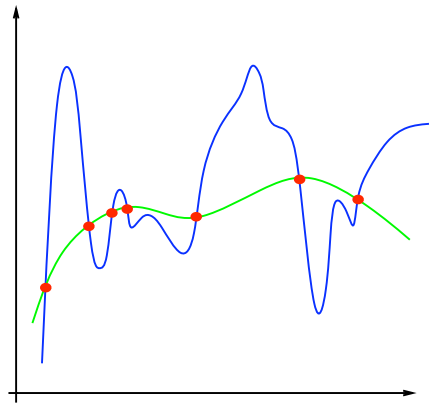
$$\text{encyclopedia} \Rightarrow \text{fundamental principles} \Rightarrow \text{primitive notions}$$

iii ***calculus ratiocinator*** — the arrangement of all true propositions in an axiomatic system.

iv decision procedure. — an algorithm which, when applied to any formula of the *characteristica universalis*, would determine whether or not that formula were true. — a procedure for the rapid enlargement of knowledge. replace reasoning by computation. the art of invention. free mind from intuition.

v a proof that the *calculus ratiocinator* is consistent.

Leibniz's Philosophy of Inductive Logic



2. Compute all descriptions of possible worlds that can be expressed with the primitive notions. And the possible worlds will all have some propensity to exist.
3. Compute the probabilities of disputed hypotheses relative to the available data. As we learn more our probability assignments will asymptotically tend to a maximum for the real world, i.e. the possibility with the highest actual propensity.

- "Probability is degree of possibility (perfection)."
- "A hypothesis is more probable as it is simpler to understand and wider in explanatory power."

$$\text{probability} = \text{propensity} \propto \text{perfection} = f(\text{variety}, \text{simplicity})$$

Characteristica Universalis vs Calculus Ratiocinator

1. Characteristica Universalis — a universal language of human thought whose symbolic structure would reflect the structure of the world.
2. Calculus Ratiocinator — a method of symbolic calculation which would mirror the processes of human reasoning.

Characteristica Universalis	**Calculus Ratiocinator**
Language as Medium	Language as Calculus
Semantics can't be defined in the language itself	Semantics is possible
Interpretation can't be varied	Interpretation can be varied
Model theory impossible	Model theory possible
Only one world can be talked about	Possible worlds are possible
Only one domain of quantifiers	Domains of quantifiers can be different
Ontology is the central problem	Ontology conventional
Logical truths are about this world	Logical truth as truth in all possible worlds

Leibniz's Algebra of Concepts

1. $A \sqsubset A$
2. $A \sqsubset B \wedge B \sqsubset C \rightarrow A \sqsubset C$
3. $A \sqsubset B \leftrightarrow A = AB$

1. $C \sqsubset AB \leftrightarrow C \sqsubset A \wedge C \sqsubset B$
2. $AB \sqsubset A$
3. $AB \sqsubset B$
4. $AA = A$
5. $AB = BA$

1. $\overline{\overline{A}} = A$
2. $A \neq \overline{A}$
3. $A \sqsubset B \leftrightarrow \overline{B} \sqsubset \overline{A}$
4. $\Diamond A \rightarrow A \sqsubset B \rightarrow A \not\sqsubset \overline{B}$
5. $A\overline{A} \sqsubset B$

1. $A \sqsubset B \rightarrow \Diamond A \rightarrow \Diamond B$
2. $A \sqsubset B \leftrightarrow \neg\Diamond(A\overline{B})$
3. $\neg\Diamond(A\overline{A})$

$$\Diamond A := A \not\sqsubset B\overline{B}$$

$$A \sqcup B := \overline{\overline{A}\,\overline{B}}$$

$$A \sqsubset A \sqcup B$$

$$B \sqsubset A \sqcup B$$

$$A \sqsubset C \wedge B \sqsubset C \rightarrow A \sqcup B \sqsubset C$$

Indefinite Concepts

A	E	I	O
$A \sqsubset B$	$A \sqsubset \overline{B}$	$A \not\sqsubset \overline{B}$	$A \not\sqsubset B$
$A = AB$	$A = A\overline{B}$	$A \neq A\overline{B}$	$A \neq AB$
$\neg\Diamond(A\overline{B})$	$\neg\Diamond(AB)$	$\Diamond(AB)$	$\Diamond(A\overline{B})$
$\exists X(A = BX)$	$\exists X(A = \overline{B}X)$	$\forall X(A \neq \overline{B}X)$	$\forall X(A \neq BX)$

$\varphi(A) \vdash \exists X \varphi(X)$ $\quad$ $\exists X \varphi(X) \vdash \varphi(A)$ for some new constant A

$\neg\exists X \varphi(X) \leftrightarrow \forall X \neg\varphi(X)$ $\quad$ $\forall X \varphi(X) \vdash \exists X \varphi(X)$

Example:

$$\frac{MAP \quad SIM}{SIP}$$

$$\frac{\exists X(M = PX) \quad \forall Y(S \neq \overline{M}Y)}{\forall Z(S \neq \overline{P}Z)}$$

Proof.

Asume $S = \overline{P}Z$. Since $M \sqsubset P \iff \overline{P} \sqsubset \overline{M}$, $\exists X(\overline{P} = \overline{M}X)$. Then $S = \overline{M}XZ$. Contradiction. □

$$\text{Individual}(A) := \Diamond A \wedge \forall X(\Diamond(AX) \rightarrow A \sqsubset X)$$

$$\leftrightarrow \forall X(A \not\sqsubset X \leftrightarrow A \sqsubset \overline{X})$$

Translate Leibniz's Algebra of Concepts to the Algebra of Propositions

1. $A \rightarrow A$
2. $(A \rightarrow B) \wedge (B \rightarrow C) \rightarrow (A \rightarrow C)$
3. $(A \rightarrow B) \leftrightarrow (A \leftrightarrow A \wedge B)$

1. $(C \rightarrow A \wedge B) \leftrightarrow (C \rightarrow A) \wedge (C \rightarrow B)$
2. $A \wedge B \rightarrow A$
3. $A \wedge B \rightarrow B$
4. $A \wedge A \leftrightarrow A$
5. $A \wedge B \leftrightarrow B \wedge A$

1. $\neg\neg A \leftrightarrow A$
2. $\neg(A \leftrightarrow \neg A)$
3. $(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$
4. $\Diamond A \rightarrow (A \rightarrow B) \rightarrow \neg(A \rightarrow \neg B)$
5. $A \wedge \neg A \rightarrow B$

1. $(A \rightarrow B) \rightarrow \Diamond A \rightarrow \Diamond B$
2. $(A \rightarrow B) \leftrightarrow \neg\Diamond(A \wedge \neg B)$
3. $\neg\Diamond(A \wedge \neg A)$

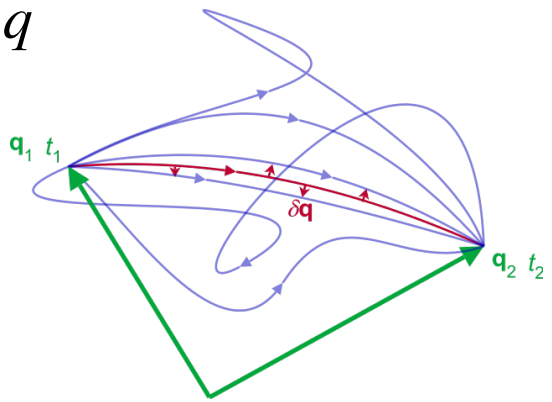
$$\Diamond A := \neg(A \rightarrow B \wedge \neg B)$$

Leibniz's Metaphysics and Quantum Mechanics

Monadology	**Path Integral**
Amount of existence	Square of probability amplitude
Measure of necessity of individual possibility	Probability
Collision or competition of possibilities	Interference or summation of probability amplitudes
Coexisting or compatible essences	Superposition of coherent paths
Maximal degree of existence	Observed path

$$P = |\langle q_2, t_2 \mid q_1, t_1 \rangle|^2 \quad \langle q_2, t_2 \mid q_1, t_1 \rangle = \int_{q_1}^{q_2} \varphi[q] \mathcal{D}q$$

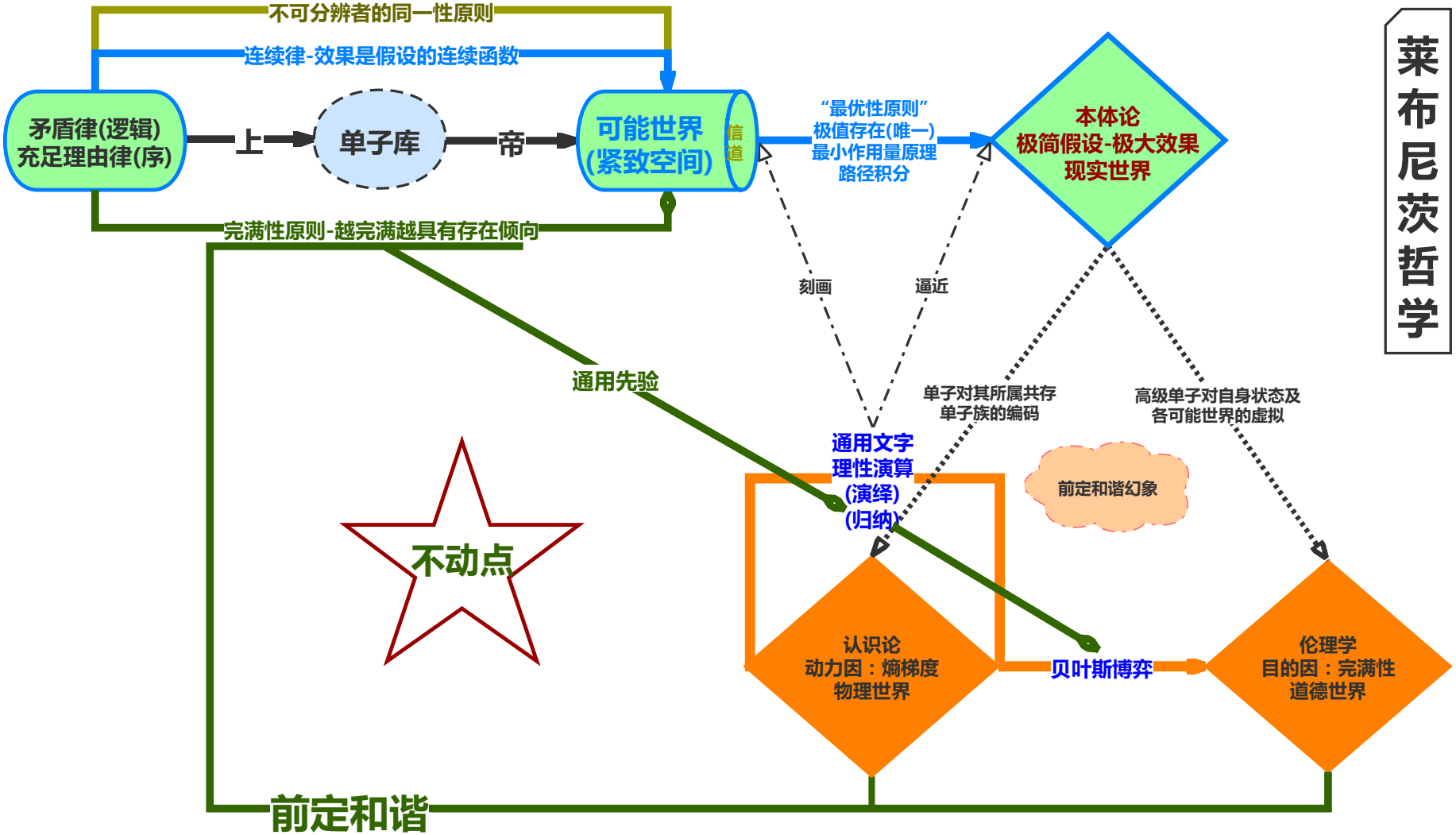
$$\varphi[q] \propto \mathrm{e}^{\frac{\mathrm{i}}{\hbar} S[q]} \quad S[q] = \int_{t_1}^{t_2} L[q(t), \dot{q}(t)] \, \mathrm{d}t \quad \delta S = 0$$



- Probability of the actual path = maximum
- Action of the actual path = minimum

the absolute square of the sum of probability amplitudes over all possible paths

Leibniz's Program

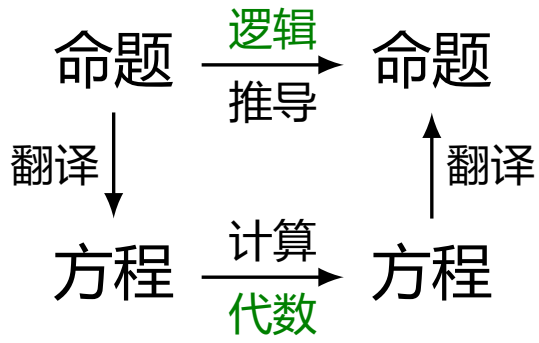


布尔 George Boole 1815-1864



Figure: 自学成才的逻辑学家、数学家、哲学家、诗人

- 《思维规律》*1854.*
- 布尔代数
  命题逻辑
- 逻辑还原为代数



Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Meta-Theorems
- Boolean Algebra
- Application

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

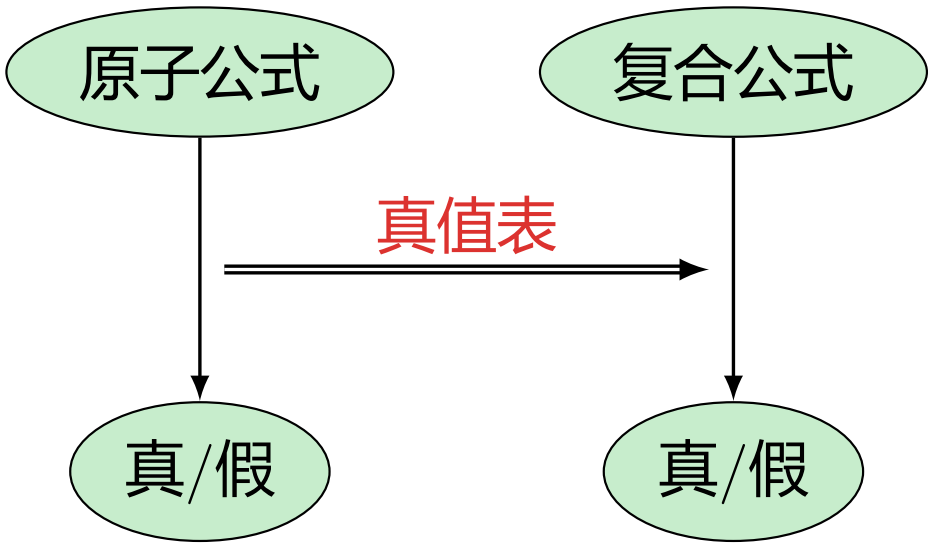
References2039

命题逻辑

- 语言
  原子公式、连接词
- 语法



- 语义



- 形式系统



语法 ♡

语言

$$\mathscr{L}^0 := \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, (,)\} \cup \mathrm{Var}$$

where $\mathrm{Var} := \{p_1, p_2, p_3, \dots\}$.

Definition (公式 Well-Formed Formula Wff — Top Down)

1. 原子公式 $p \in \mathrm{Var}$ 是公式.
2. 如果 A 是公式, 则 $(\neg A)$ 也是公式.
3. 如果 A, B 是公式, 则 $(A \wedge B), (A \vee B), (A \rightarrow B), (A \leftrightarrow B)$ 也是公式.
4. 除此之外, 别无其他公式.

$$A := p \mid (\neg A) \mid (A \wedge A) \mid (A \vee A) \mid (A \rightarrow A) \mid (A \leftrightarrow A)$$

- $\bot := (A \wedge (\neg A))$
- $\top := (\neg \bot)$

Example

1. You are beautiful.
2. You are not beautiful.
3. You are beautiful inside and out.
4. You win, or you learn.
5. If wishes are horses, then beggars will ride.
6. You are beautiful if and only if you are smart.
7. If you want to get good, get your hands dirty and start solving problems.
8. 侠之大者, 为国为民.
9. 只要功夫深, 铁杵磨成针. / 砍头不要紧, 只要主义真.
10. 只有你请她才来. / 除非你请, 否则她不会来.
11. 不积跬步, 无以至千里. /没有共产党, 就没有新中国.
12. 不是你死, 就是我亡.
13. 鱼与熊掌不可兼得.
14. 欲寄君衣君不还, 不寄君衣君又寒. 寄与不寄间, 妾身千万难.

括号的作用



他手里不是有 K 或有黑桃的话就有 A.

$$\neg(K \vee S) \to A$$

$$((\neg K) \vee S) \to A$$

Definition (公式构造算子)

$$f_\neg(A) := (\neg A)$$
$$f_\wedge(A,B) := (A \wedge B)$$
$$f_\vee(A,B) := (A \vee B)$$
$$f_\to(A,B) := (A \to B)$$
$$f_\leftrightarrow(A,B) := (A \leftrightarrow B)$$

$$f_\neg(A) := \neg A$$
$$f_\wedge(A,B) := \wedge AB$$
$$f_\vee(A,B) := \vee AB$$
$$f_\to(A,B) := \to AB$$
$$f_\leftrightarrow(A,B) := \leftrightarrow AB$$

省略括号的约定 ♡

1. 公式最外层的括号可以省略

$$\frac{A \to B}{(A \to B)}$$

2. $\neg$, $\wedge$, $\vee$, $\to$, $\leftrightarrow$ 的组合强度依次减弱

$$\frac{\neg A \vee B \to C \wedge D}{((\neg A) \vee B) \to (C \wedge D)}$$

Remark: 类似先乘除后加减 $1 + 2 * 3 = ?$

3. 同一连接词相邻出现时，右边的组合力更强

$$\frac{A \to B \to C \to D}{A \to (B \to (C \to D))}$$

Wff — Bottom Up Definition

Definition (Construction Sequence)

A construction sequence $(C_1, \ldots, C_n)$ is a finite sequence of expressions s.t. for each $i \leq n$ we have at least one of

$$C_i = p_i \quad \text{for some } i$$
$$C_i = (\neg C_j) \quad \text{for some } j$$
$$C_i = (C_j \star C_k) \quad \text{for some } j < i, k < i, \text{ where } \star \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}.$$

Definition (Well-Formed Formula Wff — Bottom Up)

A formula A is a well-formed formula (wff) iff there is some construction sequence $(C_1, \ldots, C_n)$ and $C_n = A$.

$$\begin{aligned} \mathrm{Wff}_0 &:= \{p_1, p_2, \ldots\} \\ \mathrm{Wff}_{n+1} &:= \mathrm{Wff}_n \cup \{(\neg A) : A \in \mathrm{Wff}_n\} \cup \{(A \rightarrow B) : A, B \in \mathrm{Wff}_n\} \\ \mathrm{Wff}_* &:= \bigcup_{n \in \mathbb{N}} \mathrm{Wff}_n \end{aligned}$$

Generation — Bottom Up vs Top Down

Problem

Given a class $\mathcal{F}$ of functions over U, how to generate a certain subset of U by starting with some initial elements $V \subset U$?

Definition (Bottom Up)

$$W_0 := V$$

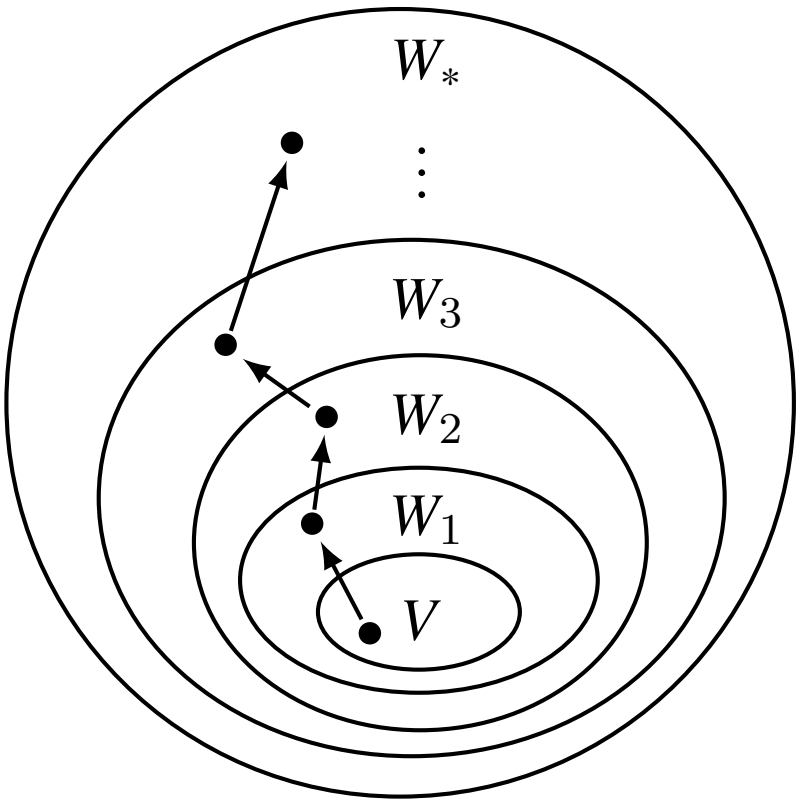
$$W_{n+1} := W_n \cup \bigcup_{f \in \mathcal{F}} \{f(\mathbf{x}) : \mathbf{x} \in W_n\} \qquad \deg(\mathbf{x}) := \mu n\, [\mathbf{x} \in W_n]$$

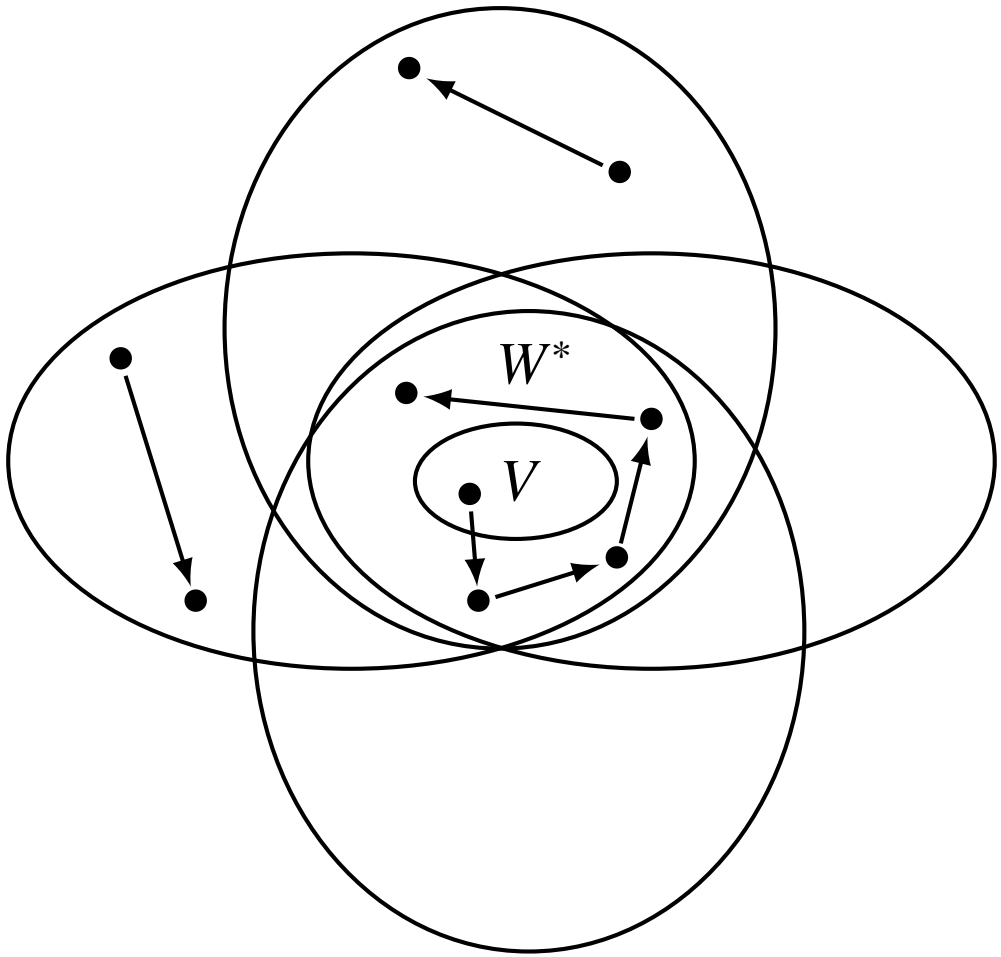
$$W_* := \bigcup_{n \in \mathbb{N}} W_n$$

Definition (Top Down)

- A set S is closed under a function f iff for all $\mathbf{x}$: $\mathbf{x} \in S \rightarrow f(\mathbf{x}) \in S$.
- A set S is inductive iff $V \subset S$ and for all $f \in \mathcal{F}$: S is closed under f.
- $W^* := \bigcap \{S : S \text{ is inductive}\}$

生成 — 自下而上 vs 自上而下





自下而上 vs 自上而下



Example ($10 元钱可以喝多少瓶啤酒？)

- $2 元钱可以买 1 瓶啤酒.
- 4 个瓶盖可以换 1 瓶啤酒.
- 2 个空瓶也可以换 1 瓶啤酒.

$$\frac{a_1}{1-\frac{3}{4}}$$

Bottom Up vs Top Down

Example

Let $V := \{0\}, \mathcal{F} := \{s, p\}, s(x) := x + 1, p(x) := x - 1$

$$W_* = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

There is more than one way of obtaining a member of W_*, e.g. $1 = s(0) = s(p(s(0)))$.

Theorem (Bottom Up vs Top Down)

$$W_* = W^*$$

Proof.

$(W^* \subset W_*)$: to show W_* is inductive.
$(W_* \subset W^*)$: to show $W_n \subset W^*$ for all n by induction. Consider $x \in W_*$ and a construction sequence $(x_0, \ldots, x_n)$ for x. First $x_0 \in V \subset W^*$. If for all $i < n$ we have $x_i \in W^*$, then $x_n \in W^*$. □

公式上的归纳法

Theorem (公式上的归纳法)

令 P 是一个关于公式的性质. 假设

1. 所有原子公式都有性质 P.
2. 对任意公式 A, B, 若 A, B 有性质 P, 则 $\neg A, A \wedge B, A \vee B, A \to B, A \leftrightarrow B$ 都有性质 P.

那么, 所有公式都有性质 P.

Proof.

$$\mathrm{Wff}_* = \mathrm{Wff}^* \subset P$$

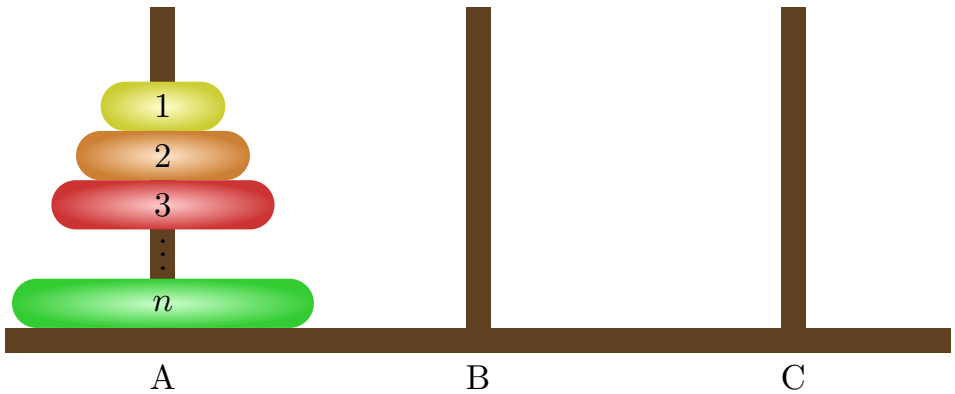
□

$$P(0) \wedge \forall n(P(n) \to P(n+1)) \to \forall n P(n)$$

$$P(n) := P(\mathrm{Wff}_n)$$

$$\frac{\begin{array}{c}[P(n)] \\ \vdots \\ P(0) \quad P(n+1)\end{array}}{\forall n P(n)}$$

归纳 vs 递归



$P(n) :=$ "n rings needs $2^n - 1$ moves."

Example (怎么让送奶工天天留奶?)

1. 如果某天留奶, 那么第二天也留奶
2. 今天留奶

今天留奶, 并且, 明天再读一遍这个字条

子公式

Definition (子公式 — 归纳定义)

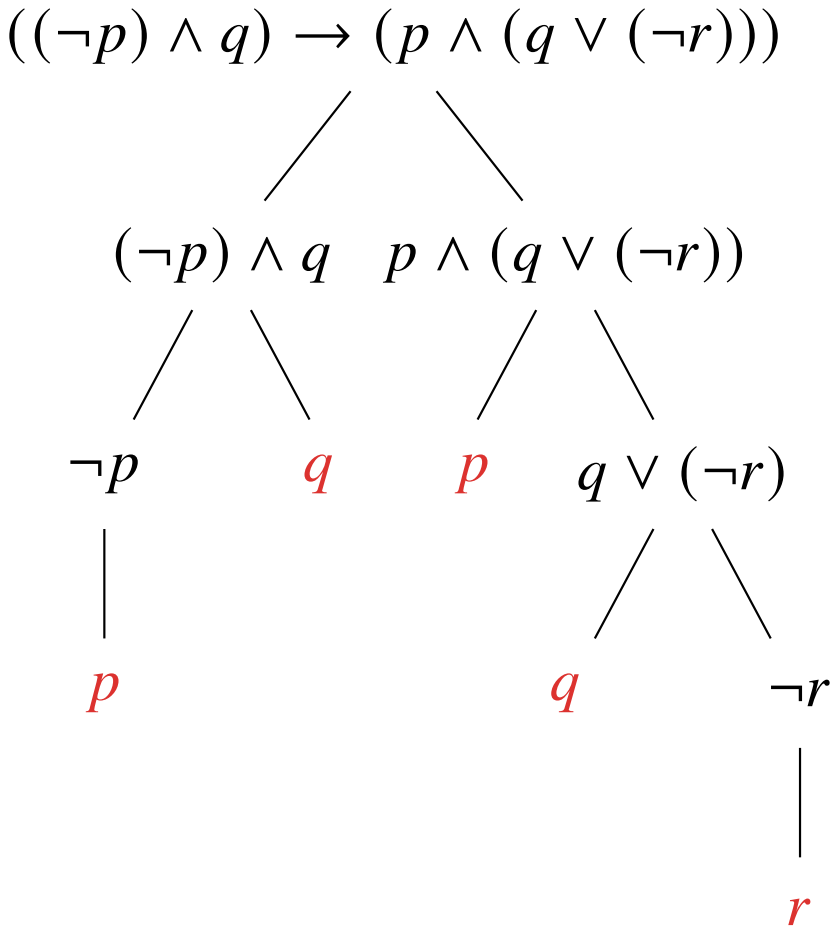
The set $\mathrm{Sub}(A)$ of subformulas of a formula A is the smallest set Γ that satisfies

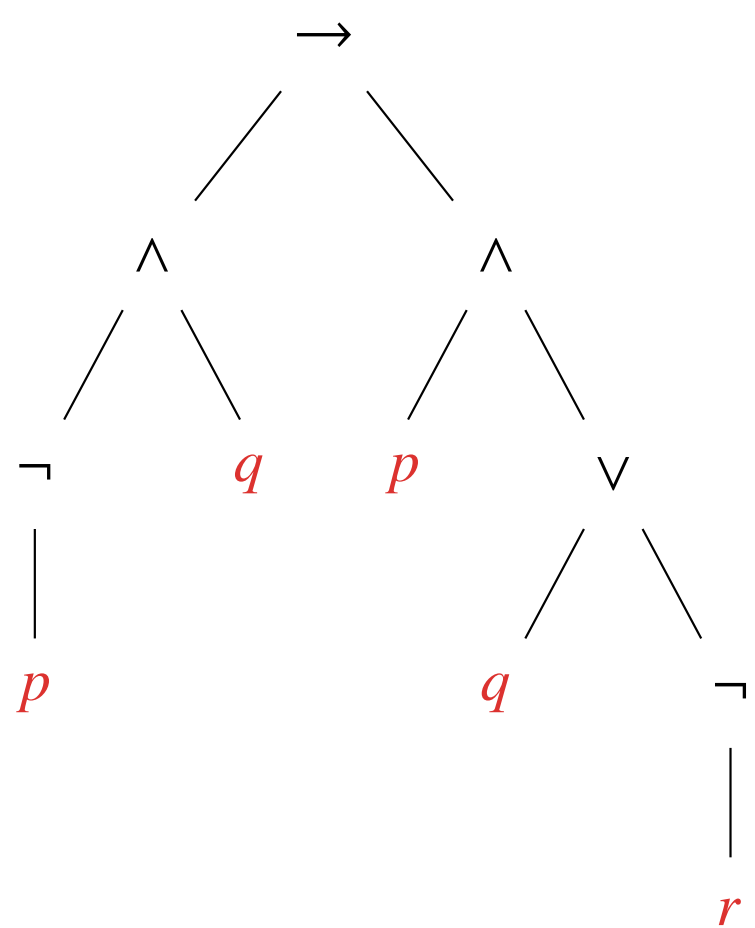
1. $A \in \Gamma$
2. $\neg B \in \Gamma \implies B \in \Gamma$
3. $B \to C \in \Gamma \implies B, C \in \Gamma$

Definition (子公式 — 递归定义)

$$\mathrm{Sub}(A) := \begin{cases} \{A\} & \text{if } A = p \\ \{A\} \cup \mathrm{Sub}(B) & \text{if } A = \neg B \\ \{A\} \cup \mathrm{Sub}(B) \cup \mathrm{Sub}(C) & \text{if } A = B \to C \end{cases}$$

唯一可读性 = 唯一语法树





- 主连接词: 树根
- 子公式: 子树

Balanced-Parentheses

Proposition (Balanced-Parentheses)

The number of left and right parentheses in a formula are equal.

Lemma

Any proper prefix of a formula contains an excess of left parentheses.
Thus a proper prefix of a formula is not a formula.

Proof.

Consider $A = (C \wedge D)$. The proper prefix of $(C \wedge D)$:

1. $($
2. $(C_0$ [inductive hypothesis]
3. $(C$ [balanced-parentheses]
4. $(C \wedge$ [balanced-parentheses]
5. $(C \wedge D_0$ [inductive hypothesis]
6. $(C \wedge D$ [balanced-parentheses]

唯一可读性定理

Theorem (Unique Readability Theorem)

The formula-building operators, when restricted to the set of formulas,

1. *are injective, and*
2. *have ranges that are disjoint from each other, and from the set of atomic propositions.*

Proof.

$$(A \star B) = (C * D)$$

$$\Downarrow$$

$$A \star B) = C * D)$$

$$\Downarrow$$

$$A = C \qquad \text{(Lemma)}$$

$$\Downarrow$$

$$\star = *$$

$$\Downarrow$$

$$B = D \qquad \text{(Lemma)}$$

翻译技巧 ♡

- **Negation 否定:** not / it is false that ...
  - Jay and Kay are married, but not to each other. $J \wedge K \wedge \neg M$
- **Conjunction 合取:** and / moreover / furthermore / but / yet / although / though / even though / however / whereas
- **Disjunction 析取:** or / either or
- **Implies 蕴含/只要就:** "if then" / provided that / in case / on the condition that
- **Neither Nor 既不也不:** negation of "either or"
- **Only If 仅当/只有才:** "in order that ...it is necessary that ..."
- **Unless 除非:** if not / if and only if not
  1. I will not graduate unless I pass logic. ¬PassLogic→¬Graduate
  2. The store is open unless/except it is sunday. Open↔¬Sunday
- **Even If 即使:**
  - I'm going to the party even if it rains. $(\neg R \rightarrow P) \wedge (R \rightarrow P) \equiv P$
  - The Allies would have won even if the U.S. had not entered the war.[3] $W \wedge (\neg E \rightarrow W) \equiv W$
  - The Axis powers would have won if the U.S. had not entered the war. $\neg W \wedge (\neg E \rightarrow W)$

[3] 虚拟语气的表达需要借助 Counterfactual Logic.

Example ☹

1. The programmer's wife tells him: "Run to the store and pick up a loaf of bread. If they have eggs, get a dozen."
2. The programmer comes home with 12 loaves of bread.
3. "Why did you buy 12 loaves of bread!?", his wife screamed.
4. "Because they had eggs!"

- wife.

$$1\text{Bread} \wedge (\text{Egg} \rightarrow 12\text{Egg})$$

- programmer.

$$(\neg\text{Egg} \rightarrow 1\text{Bread}) \wedge (\text{Egg} \rightarrow 12\text{Bread})$$

练习: 翻译 — Now it's your turn

1. I am not good at logic.
2. Either Alice is a fool or she is dishonest.
3. If you can't say it clearly, you don't understand it yourself. — Searle
4. You understand something only if you can formalize it.
5. A science becomes developed only when it can make use of mathematics. — *Marx*
6. If you work hard only if you are threatened, then you will not succeed.
7. I will go out unless it rains.
8. You will pass unless you goof off, provided that you are intelligent.
9. If Jones will work only if Smith is fired, then we should fire Smith if we want the job finished.
10. In order to put on the show it will be necessary to find a substitute, if neither the leading lady nor her understudy recovers from the flu.

练习: 翻译 — Now it's your turn

11. If you make an appointment and do not keep it, then I shall be angry unless you have a good excuse.
12. Getting a 100 on the exam is necessary but not sufficient for getting an A.
13. I'll play tennis unless it is raining, in which case I'll play pingpong.
14. If it is sunny and not cold, I'll play tennis; otherwise, I'll play pingpong or go swimming.
15. You can pay by credit card or cheque, but not both.
16. Neither Sarah nor Peter was to blame for the mistake.
17. I want to buy either a new desktop computer or a laptop, but I have neither the cash nor the credit I need.
18. Even though the dark clouds veil the sky, You are by my side.

翻译 — 自然语言 vs 形式语言

1. 如果我进电梯, 电梯会坏; 如果你进电梯, 电梯会坏. (?)
2. 如果我进电梯, 并且你进电梯, 那么电梯会坏. (?)[4]

$$\frac{A \vee B \to C}{(A \to C) \wedge (B \to C)}$$

$$\frac{A \wedge B \to C}{(A \to C) \vee (B \to C)}$$

- 我 100 磅.
- 你 100 磅.
- 电梯容量 150 磅.

[4]经典逻辑里没法细分 $A \to B$; 若细分, 可以用 Linear Logic $A \otimes B \multimap C$ 表达.

Digression — 为什么要学一门 “形式语言”?

逻辑**仅仅**是一门语言吗?

- 语言是人与其他动物最大的区别.
- 阿拉伯数字 vs 汉语数字: 仅仅是语言的区别吗?
- 你甚至可以将 DNA 视为一种语言: 一种程序语言.

怎么以**简洁的**方式**准确地**表达我们的思想?

什么是理性的**范围与限度**?

“少于十八个汉字不可**定义**的最小自然数.”

— *Berry* 悖论



Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Meta-Theorems
- Boolean Algebra
- Application

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

真值赋值 Truth Assignment ♡

- 真值赋值是一个从原子公式到真、假的函数

$$\nu : \mathrm{Var} \to \{0, 1\}$$

- 真值赋值 ν 可以在满足下列条件的情况下递归地扩展到所有公式上 $\overline{\nu} : \mathrm{Wff} \to \{0, 1\}$.

1. $\overline{\nu}(p) = \nu(p)$ for $p \in \mathrm{Var}$
2. $\overline{\nu}(\neg A) = 1 - \overline{\nu}(A)$
3. $\overline{\nu}(A \wedge B) = \min\{\overline{\nu}(A), \overline{\nu}(B)\}$
4. $\overline{\nu}(A \vee B) = \max\{\overline{\nu}(A), \overline{\nu}(B)\}$
5. $\overline{\nu}(A \to B) = \max\{1 - \overline{\nu}(A), \overline{\nu}(B)\}$
6. $\overline{\nu}(A \leftrightarrow B) = 1 - \left|\overline{\nu}(A) - \overline{\nu}(B)\right|$

p	$\neg p$
0	1
1	0

$\wedge$	0	1
0	0	0
1	0	1

$\vee$	0	1
0	0	1
1	1	1

$\to$	0	1
0	1	1
1	0	1

$\leftrightarrow$	0	1
0	1	0
1	0	1

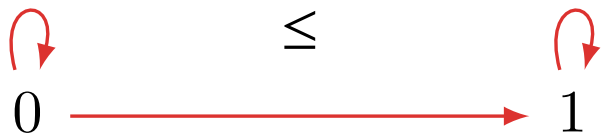
Remark: 在不引起歧义的情况下, 我们把 $\overline{\nu}$ 记作 ν.

真值表 Truth Table ♡

p	$\neg p$
0	1
1	0

p	q	$p \wedge q$	$p \vee q$	$p \to q$	$p \leftrightarrow q$
0	0	0	0	1	1
0	1	0	1	1	0
1	0	0	1	0	0
1	1	1	1	1	1

$$\frac{\nu(A \to B) = 1}{\nu(A) \leq \nu(B)}$$



Example:

- 如果 $0 = 1$, 那么罗素是上帝.
- 雪是白的当且仅当 $1 + 1 = 2$.

$$B^A$$

$$A \to B$$

充分 $\to$ 必要

Example

令 $v(p) = 1, v(q) = 0, v(r) = 1$, 求 $v(p \vee r \to \neg(p \to q))$.

p	q	r	$p \vee r$	$p \to q$	$\neg(p \to q)$	$p \vee r \to \neg(p \to q)$
1	0	1	1	0	1	1

Remark: 在不引起歧义的情况下, 我们经常省略 v 不写.

$$\begin{aligned} p \vee r \to \neg(p \to q) &= 1 \vee 1 \to \neg(1 \to 0) \\ &= 1 \to \neg 0 \\ &= 1 \to 1 \\ &= 1 \end{aligned}$$

Definition (Freely Generated)

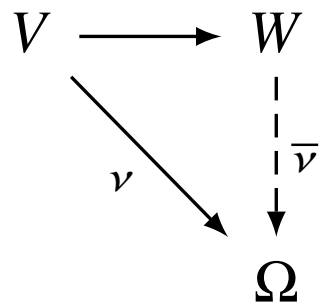
The set W is freely generated from V by a class of functions $\mathcal{F}$ iff in addition to the requirements for being generated, the following hold:

1. for every $f \in \mathcal{F}$: $f \restriction_W$ is injective.
2. the range of $f \restriction_W$ for all $f \in \mathcal{F}$, and the set V are pairwise disjoint.

Theorem (Recursion Theorem)

Assume that W is freely generated from V by $\mathcal{F}$, and for every n-ary $f \in \mathcal{F}$ we have $f^* : \Omega^n \to \Omega$. Then for every function $\nu : V \to \Omega$, there exists a unique function $\overline{\nu} : W \to \Omega$ such that

1. $\overline{\nu} \restriction_V = \nu$
2. *for all $f \in \mathcal{F}$ and all $x_1, \ldots, x_n \in W$:*
$$\overline{\nu}(f(x_1, \ldots, x_n)) = f^*\big(\overline{\nu}(x_1), \ldots, \overline{\nu}(x_n)\big)$$



Remark:

- ν tells you how to color the initial elements in V;
- f^* tells you how to convert the color of $\mathbf{x}$ into the color of $f(\mathbf{x})$.

Danger! f^* is saying "green" but F_g is saying "red" for the same point.

“实质蕴含” — 反直觉吗?

翻哪几张卡才能验证 “如果一面是偶数, 另一面肯定是红色”?



未满 18 岁禁止饮酒!

p	q	$p \to q$	如果 x 是有理数, 那么 x^2 是有理数
0	0	1	$x = \pi$
0	1	1	$x = \sqrt{2}$
1	0	0	
1	1	1	$x = \frac{1}{2}$

“实质蕴含”

Problem (辩护律师靠谱吗？)

某人因涉嫌参与盗窃而受审.

检察官 如果被告有罪, 那么他有搭档伙同作案.

辩护律师 这不是实话!

Example:

记者 如果你有一百万, 你愿意捐给国家吗?

农民 我愿意.

记者 如果你有一头牛, 你愿意捐给国家吗?

农民 我不愿意.

记者 为什么?

农民 我真的有一头牛!

Remark

- “如果我不来上课, 那么宇宙会爆炸.” — “反事实蕴含”
- 数学语境中的 “蕴含” 都是 “实质蕴含”.

Find the Error ☺

怎么解方程 $y + 2 = y$?

$$\begin{aligned} y + 2 = y &\implies (y+2)^2 = y^2 \\ &\implies y^2 + 4y + 4 = y^2 \\ &\implies 4y + 4 = 0 \\ &\implies y = -1 \end{aligned}$$

真值赋值、真值表[5]

如果雪是白的, 那么, 窦娥是杀人犯蕴含雪是白的.

	p	q	$q \to p$	$p \to q \to p$
v_1	0	0	1	1
v_2	0	1	0	1
v_3	1	0	1	1
v_4	1	1	1	1

n 个原子命题有 2^n 种真值赋值.

<table>
<tr><td>永真
重言
$A \to A$</td><td colspan="2">偶真/偶假
(综合命题?)
$A \to B$</td><td>永假
矛盾
$A \leftrightarrow \neg A$</td></tr>
<tr><td>有效</td><td colspan="3">无效</td></tr>
<tr><td colspan="3">可满足</td><td>不可满足</td></tr>
</table>

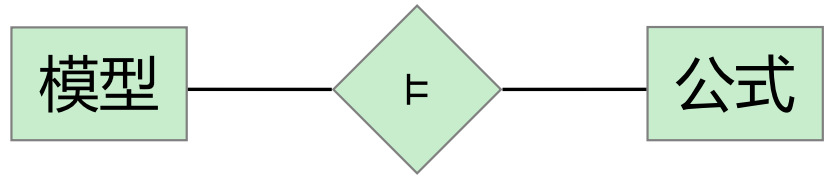
[5] Truth Table Generator

可满足、逻辑蕴含、有效式、逻辑等价 ♡

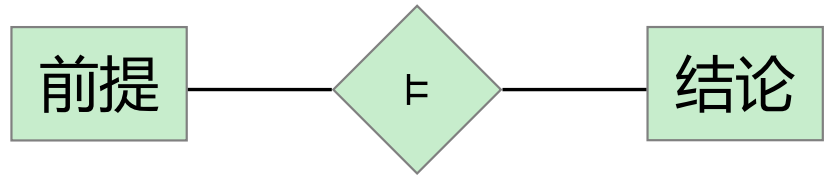
- Satisfiability 可满足: A is *satisfiable* iff there is some truth assignment ν s.t. $\nu(A) = 1$. (We also write $\nu \vDash A$ when $\nu(A) = 1$.)
- Entailment 逻辑蕴含: $\Gamma \vDash A$ iff for any truth assignment ν s.t. $(\text{for all } B \in \Gamma : \nu \vDash B) \implies \nu \vDash A$.
- Validity / Tautology 有效式/重言式: A is valid $\vDash A$ iff $\emptyset \vDash A$, in other words, for any truth assignment ν, $\nu \vDash A$.
- Logical Equivalence 逻辑等价: A and B are logically equivalent $A \equiv B$ iff $A \vDash B$ and $B \vDash A$.

逻辑蕴含	有效	可满足	逻辑等价
entailment	validity	satisfiability	logical equivalence
$A \vDash B$	$\vDash A \to B$	$A \wedge \neg B$ unsatisfiable	$A \to B \equiv \top$
$\top \vDash A$	$\vDash A$	$\neg A$ unsatisfiable	$A \equiv \top$
$A \nvDash \bot$	$\nvDash \neg A$	A satisfiable	$\neg A \not\equiv \top$
$A \vDash B$ and $B \vDash A$	$\vDash A \leftrightarrow B$	$A \leftrightarrow \neg B$ unsatisfiable	$A \equiv B$

满足 $\nu \vDash A$



逻辑蕴含 $\Gamma \vDash A$



直观理解	形式定义
可能世界/模型	真值赋值
连接词的语义	真值表
在所有可能世界上都真的命题	重言式/有效式
有效 ("保真" 的) 论证	逻辑蕴含

Remark

- 二值原则: 每个命题有且只有两个真值中的一个: 0 或 1
- 组合原则: 复合命题的真值由组成它的命题的真值以及这些命题的组合方式唯一确定.

$$\frac{\nu(A) = \mu(A)}{\nu(\neg A) = \mu(\neg A)} \qquad \frac{\nu(A) = \mu(A) \quad \nu(B) = \mu(B)}{\nu(A \to B) = \mu(A \to B)}$$

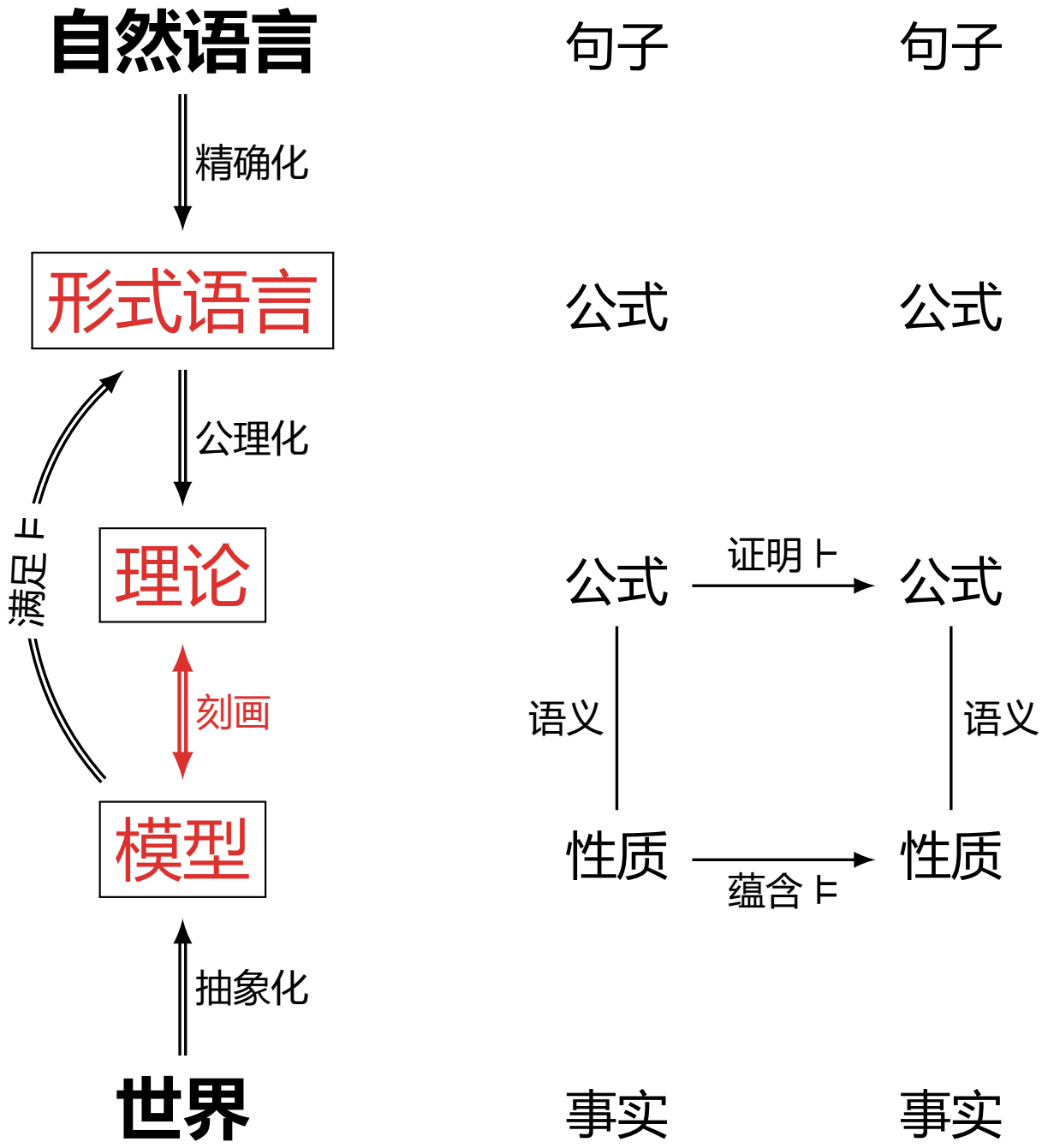
Example

证明 $p \wedge (\neg q \vee \neg r) \rightarrow (p \rightarrow \neg q)$ 既不是矛盾式也不是有效式.

p	q	r	$p \wedge (\neg q \vee \neg r) \rightarrow (p \rightarrow \neg q)$
0	0	0	1
1	1	0	0

$$
\begin{aligned}
p \wedge (\neg q \vee \neg r) \rightarrow (p \rightarrow \neg q) &= 0 \wedge (\neg 0 \vee \neg 0) \rightarrow (0 \rightarrow \neg 0) \\
&= 0 \wedge (1 \vee 1) \rightarrow (0 \rightarrow 1) \\
&= 0 \wedge 1 \rightarrow 1 \\
&= 0 \rightarrow 1 \\
&= 1
\end{aligned}
$$

$$
\begin{aligned}
p \wedge (\neg q \vee \neg r) \rightarrow (p \rightarrow \neg q) &= 1 \wedge (\neg 1 \vee \neg 0) \rightarrow (1 \rightarrow \neg 1) \\
&= 1 \wedge (0 \vee 1) \rightarrow (1 \rightarrow 0) \\
&= 1 \wedge 1 \rightarrow 0 \\
&= 1 \rightarrow 0 \\
&= 0
\end{aligned}
$$


自然语言
精确化
形式语言
公理化
理论
满足
刻画
模型
抽象化
世界
句子
句子
公式
公式
公式
证明⊢
公式
语义
语义
性质
蕴含⊨
性质
事实
事实

题外话

维特根斯坦 — 逻辑原子主义哲学

- 世界是事实的总和. 语言是命题的总和.
- 事实是事态的存在. 原子事态是对象的结合. 原子事态彼此独立.
- 命题是事实的图像.
  — 原子命题对应原子事态. 复合命题对应复杂事态.
- 复合命题是原子命题的真值函数.
- 图像与其描绘的实在拥有共同的逻辑结构.
  — 唱片、音乐思想、乐谱、声波, 彼此处在图示的内在关系中, 这也是语言和世界的关系.
- **理解一个命题, 意味着知道其为真的情形.**
- 真命题的总和是世界的图像.

Remark: 在翻译自然语言时, 我们根据需要来设定原子命题, 尽量保证原子命题彼此独立.

命题、真值赋值、可能世界

每个命题对应
一类可能世界

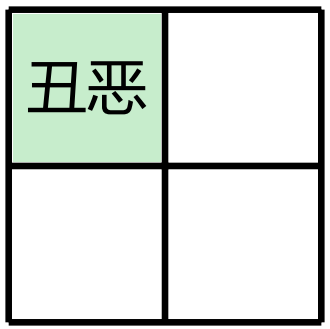
$v : \mathrm{Var} \to \{0, 1\}$

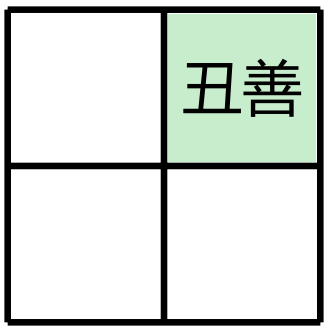
$[\![A]\!] = \{v : v \vDash A\}$

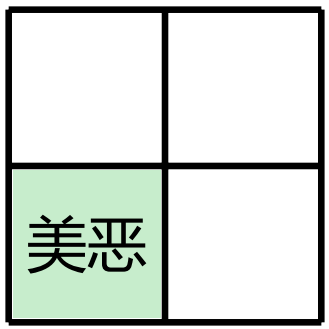
- p: 人美
- q: 心善

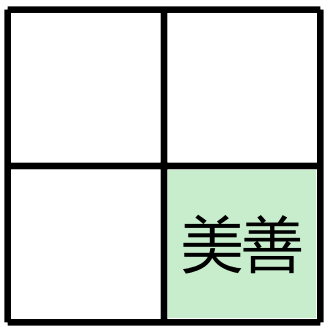
p \ q	0	1
0	v_1	v_2
1	v_3	v_4

	p	q
v_1	0	0
v_2	0	1
v_3	1	0
v_4	1	1

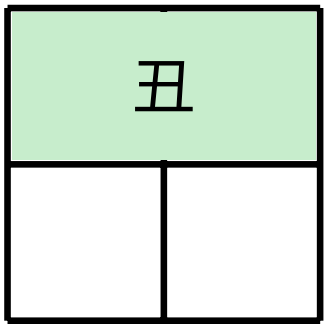




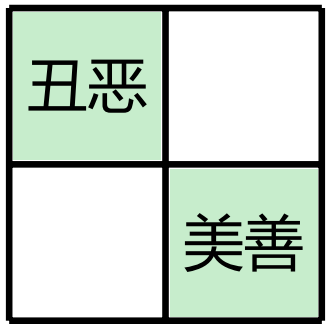




$$\frac{(\neg p \wedge \neg q) \vee (\neg p \wedge q)}{\neg p}$$



$p \leftrightarrow q$



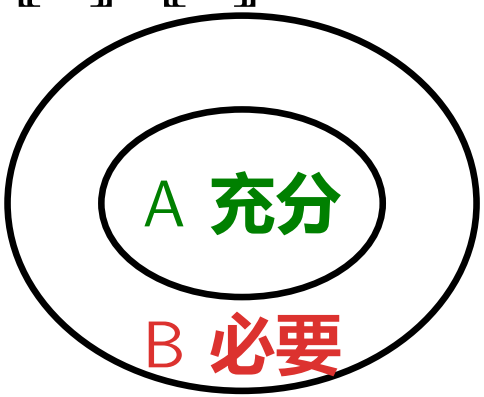
Truth Assignment — set-based version

真值赋值 $\nu : \mathrm{Var} \to \{0,1\}$ 可以在满足下列条件的情况下递归地扩展到所有公式上:

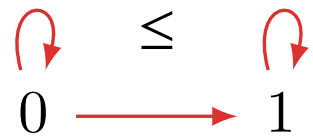
1. $[\![p]\!] := \{\nu : \nu(p) = 1\}$
2. $[\![\top]\!] := \{0,1\}^{\mathrm{Var}}$
3. $[\![\bot]\!] := \emptyset$
4. $[\![\neg A]\!] := \overline{[\![A]\!]}$
5. $[\![A \wedge B]\!] := [\![A]\!] \cap [\![B]\!]$
6. $[\![A \vee B]\!] := [\![A]\!] \cup [\![B]\!]$
7. $[\![A \to B]\!] := \overline{[\![A]\!]} \cup [\![B]\!]$
8. $[\![A \leftrightarrow B]\!] := \overline{[\![A]\!] \Delta [\![B]\!]}$

1. $[\![p]\!] := \nu(p)$
2. $[\![\top]\!] := 1$
3. $[\![\bot]\!] := 0$
4. $[\![\neg A]\!] := 1 - [\![A]\!]$
5. $[\![A \wedge B]\!] := \min\{[\![A]\!], [\![B]\!]\}$
6. $[\![A \vee B]\!] := \max\{[\![A]\!], [\![B]\!]\}$
7. $[\![A \to B]\!] := \max\{1 - [\![A]\!], [\![B]\!]\}$
8. $[\![A \leftrightarrow B]\!] := 1 - \left|[\![A]\!] - [\![B]\!]\right|$

$$\frac{\dfrac{\models A \to B}{A \models B}}{[\![A]\!] \subset [\![B]\!]}$$



$$\frac{\dfrac{\models A \to B}{A \models B}}{[\![A]\!] \leq [\![B]\!]}$$



你的论证有效 (“保真”) 吗? ♡

问: 一个从前提集 $\{A_1, \ldots, A_n\}$ 到结论 B 的论证

$$\frac{A_1, \ldots, A_n}{B}$$

怎么才算有效?

答: 逻辑蕴含即有效

$$A_1, \ldots, A_n \vDash B$$

1. 如果前提 $A_1, \ldots, A_n$ 为真, 那么结论 B 必为真
2. 如果结论 B 为假, 那么至少有一个前提 A_i 为假

$$\frac{\text{如果猪会飞, 那么这瓜保熟} \quad \text{猪会飞}}{\text{这瓜保熟}}$$ 论证保真



你的论证有效 (“保真”) 吗? — Example

整数 x 或是奇数或是偶数
如果 x 是奇数, 那么 $x + x$ 是偶数
如果 x 是偶数, 那么 $x + x$ 是偶数

$x + x$ 是偶数

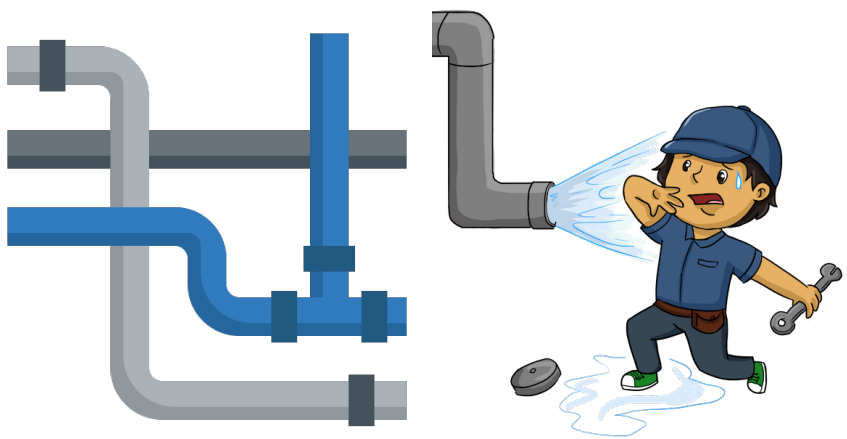
$$\frac{A \vee B \quad A \to C \quad B \to C}{C}$$

$$A \vee B, A \to C, B \to C \vDash C$$

兴, 百姓苦; 亡, 百姓苦.

逻辑学家是干啥的? — 水管工!

组装密封水管, 保证干净的水能到达我们想要的地方.



有效论证 — Example ☹

- 克莱因是教授，因此，克莱因或者是教授或者是连环杀手.

$$\frac{A}{A \vee B}$$

- 琳达，31 岁，单身，一位直率又聪明的女士，大学主修哲学. 在学生时代，她就关心种族歧视和社会公正问题，还参加了反核示威游行.
  1. 琳达是银行出纳.
  2. 琳达是银行出纳，还是女权主义者.

$$\frac{A \wedge B}{A}$$

有效性判定 — 真值表

$$\vDash (p \to q \to r) \to (p \to q) \to p \to r$$

p	q	r	$q \to r$	$p \to q \to r$	$p \to q$	$p \to r$	$(p \to q) \to p \to r$	$(p \to q \to r) \to (p \to q) \to p \to r$
0	0	0						1
0	0	1						1
0	1	0						1
0	1	1						1
1	0	0						1
1	0	1						1
1	1	0						1
1	1	1						1

$$p \vee q, p \to r, q \to r \vDash r$$

p	q	r	$p \vee q$	$p \to r$	$q \to r$	$(p \vee q) \wedge (p \to r) \wedge (q \to r) \to r$
0	0	0		1	1	1
0	0	1		1	1	1
0	1	0	1	1		1
0	1	**1**	1	1	1	1
1	0	0	1		1	1
1	0	**1**	1	1	1	1
1	1	0	1			1
1	1	1	1	1	1	1

有效式的判定 — 归谬赋值

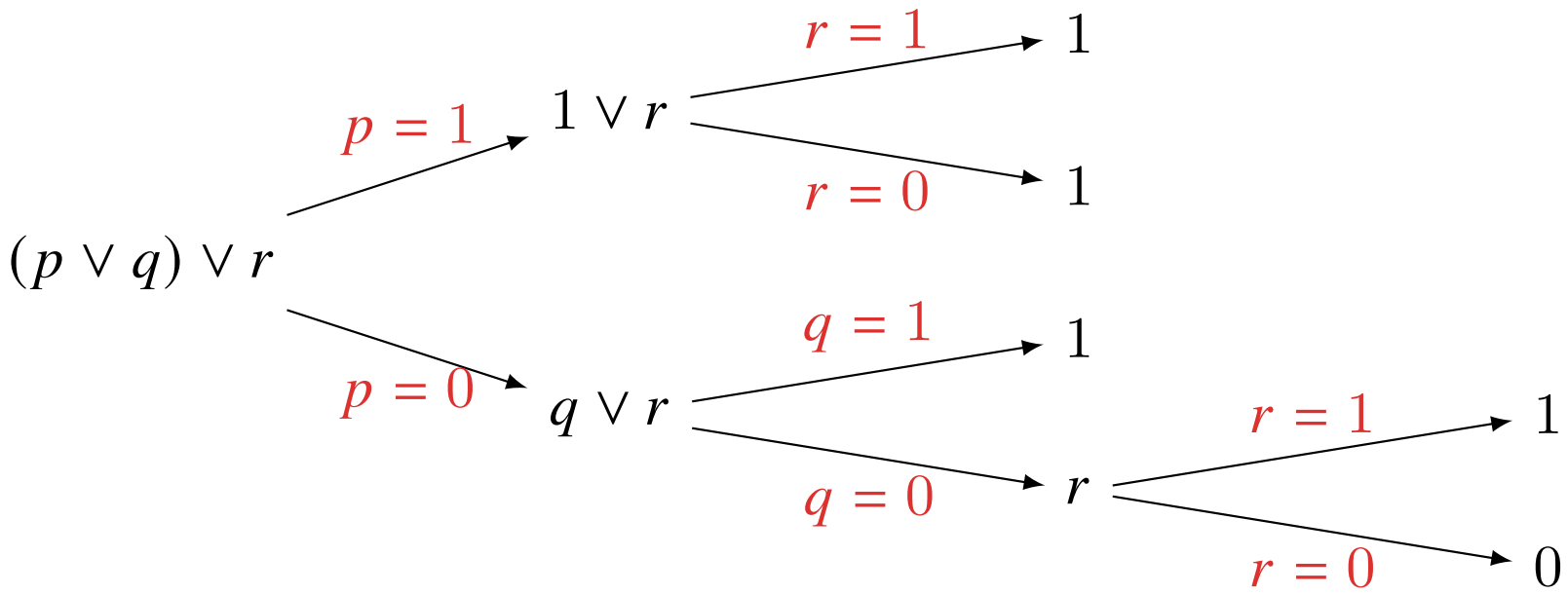
$$(p \to q \to r) \to (p \to q) \to p \to r$$

						0								
	1					**0**					**0**			
	1					**0**		1			**0**		**0**	
1	1			0		**0**	1	**1**			**0**	1	**0**	0
1	**1**	1		0		**0**	1	**1**	1		**0**	1	**0**	0
1	**1**	1	**1**	0		**0**	1	**1**	1		**0**	1	**0**	0
			×											

有效式的判定 — 代数计算

$$\neg 1 = 0 \qquad \neg 0 = 1$$
$$1 \vee A = A \vee 1 = 1 \qquad 0 \vee A = A \vee 0 = A$$
$$1 \wedge A = A \wedge 1 = A \qquad 0 \wedge A = A \wedge 0 = 0$$
$$1 \to 0 = 0 \qquad A \to 1 = 0 \to A = 1$$

1. 对公式 A, 将第一个命题变元分别代为 1 和 0, 计算 $A(1), A(0)$;
2. 重复上述操作直到所得结果不含命题变元;
3. 如果所得不含命题变元的结果都为 1, 则 A 有效, 否则无效.



有效式的判定 — 合取范式 CNF

$$
\begin{aligned}
A \leftrightarrow B &\equiv (\neg A \vee B) \wedge (A \vee \neg B) \\
A \rightarrow B &\equiv \neg A \vee B \\
\neg(A \vee B) &\equiv \neg A \wedge \neg B \\
\neg(A \wedge B) &\equiv \neg A \vee \neg B \\
\neg\neg A &\equiv A \\
(A \wedge B) \vee C &\equiv (A \vee C) \wedge (B \vee C) \\
C \vee (A \wedge B) &\equiv (C \vee A) \wedge (C \vee B)
\end{aligned}
$$

Example:

$$
\begin{aligned}
p \vee q \rightarrow q \vee r &\equiv \neg(p \vee q) \vee (q \vee r) && \text{Elim } \rightarrow \\
&\equiv (\neg p \wedge \neg q) \vee (q \vee r) && \text{Push } \neg \text{ in} \\
&\equiv (\neg p \vee q \vee r) \wedge (\neg q \vee q \vee r) && \text{Push } \vee \text{ in} \\
&\equiv \neg p \vee q \vee r && \text{Simplify}
\end{aligned}
$$

令 $p = 1, q = 0, r = 0$, 则 $p \vee q \rightarrow q \vee r = 0$, 故无效.

练习: 有效式的判定 — Now it's your turn ⚡

1.	$\neg\neg A \to A$	双重否定消去
2.	$A \to \neg\neg A$	双重否定引入
3.	$A \vee \neg A$	排中律
4.	$\neg(A \wedge \neg A)$	无矛盾律
5.	$A \wedge \neg A \to B$	爆炸律
6.	$(A \to B) \wedge (\neg A \to B) \to B$	二难推理
7.	$(A \to B) \wedge (A \to \neg B) \to \neg A$	归谬法
8.	$(\neg A \to B) \wedge (\neg A \to \neg B) \to A$	反证法
9.	$(\neg A \to \bot) \to A$	
10.	$(\neg A \to A) \to A$	
11.	$((A \to B) \to A) \to A$	Peirce 律

Problem (阿基米德有支点吗?)

大力士 听说 "如果你有一个支点, 就能翘起地球", 这是真的吗?

阿基米德 如果这是真的, 那么我有一个支点.

有效论证的判定 — Example: $\dfrac{A \vee B \quad A \to C \quad B \to C}{C}$

Proof.1.

画真值表. □

Proof.2.

- 往证: 对任意 ν, 若 $\nu \vDash A \vee B$ 且 $\nu \vDash A \to C$ 且 $\nu \vDash B \to C$, 则 $\nu \vDash C$.
- 假设 $\nu \vDash A \vee B$, 则 $\nu \vDash A$ 或 $\nu \vDash B$.
- 若 $\nu \vDash A$, 则由 $\nu \vDash A \to C$ 可得 $\nu \vDash C$.
- 若 $\nu \vDash B$, 则由 $\nu \vDash B \to C$ 可得 $\nu \vDash C$.

□

Proof.3.

$$
\begin{aligned}
(A \vee B) \wedge (A \to C) \wedge (B \to C) \to C &\equiv \neg[(A \vee B) \wedge (\neg A \vee C) \wedge (\neg B \vee C)] \vee C \\
&\equiv \neg(A \vee B) \vee \neg(\neg A \vee C) \vee \neg(\neg B \vee C) \vee C \\
&\equiv (\neg A \wedge \neg B) \vee (A \wedge \neg C) \vee (B \wedge \neg C) \vee C \\
&\equiv \dots \\
&\equiv \top
\end{aligned}
$$

练习: 有效论证的判定 — Now it's your turn ↯

$$\frac{A \vee B}{\neg A \to B} \qquad \frac{A \wedge B}{\neg(A \to \neg B)} \qquad \frac{A \leftrightarrow B}{(A \to B) \wedge (B \to A)} \qquad \frac{A \vee B}{(A \to B) \to B}$$

$$\frac{A \vee (B \vee C)}{(A \vee B) \vee C} \quad \frac{A \wedge (B \wedge C)}{(A \wedge B) \wedge C}$$ 结合律 $$\frac{A \vee B}{B \vee A} \quad \frac{A \wedge B}{B \wedge A}$$ 交换律

$$\frac{\neg(A \vee B)}{\neg A \wedge \neg B} \quad \frac{\neg(A \wedge B)}{\neg A \vee \neg B}$$ 德摩根律 $$\frac{A \vee (A \wedge B)}{A} \quad \frac{A \wedge (A \vee B)}{A}$$ 吸收律

$$\frac{A \wedge (B \vee C)}{(A \wedge B) \vee (A \wedge C)} \quad \frac{A \vee (B \wedge C)}{(A \vee B) \wedge (A \vee C)}$$ 分配律

$$\frac{A \vee B \quad \neg B \vee C}{A \vee C} \qquad \frac{\Gamma, A \vDash B}{\Gamma \vDash A \to B}$$

无效论证

$$\frac{A \to B \quad \neg A}{\neg B}$$

我思故我在
我不思
故我不在

$$\frac{A \to B \quad B}{A}$$

葡萄酸故我不吃
我不吃
故葡萄酸

$$\frac{A \vee B \quad A}{\neg B}$$

此论证或是有效的或是无效的
此论证是有效的
此论证不是无效的

$$\frac{\neg(A \wedge B) \quad \neg A}{B}$$

- 家里有一头猪和一头驴，你说我是杀猪呢，还是杀驴呢？
- 杀驴.
- 猪也是这么想的.

妈妈: “你只要把西蓝花吃了就可以去吃冰激凌了.”
儿子: “我把猪蹄吃了可以去吃冰激凌吗?”
妈妈: “你只有把西蓝花吃了才能去吃冰激凌!”

By all means marry; if you get a good wife, you'll be happy. If you get a bad one, you'll become a philosopher. — *Socrates*

$$\frac{\begin{array}{r}\text{Marry} \leftrightarrow \text{GoodWife} \vee \text{BadWife}\\ \text{GoodWife} \rightarrow \text{HappyLife}\\ \text{BadWife} \rightarrow \text{Philosopher}\\ \text{HappyLife} \rightarrow \text{PerfectEnding}\\ \text{Philosopher} \rightarrow \text{PerfectEnding}\\ \text{PerfectEnding}\end{array}}{\text{Marry}}$$

Let Marry = GoodWife = BadWife = 0 and
HappyLife = Philosopher = PerfectEnding = 1.

Then
$$\begin{array}{r}\text{Marry} \leftrightarrow \text{GoodWife} \vee \text{BadWife} = 1\\ \text{GoodWife} \rightarrow \text{HappyLife} = 1\\ \text{BadWife} \rightarrow \text{Philosopher} = 1\\ \text{HappyLife} \rightarrow \text{PerfectEnding} = 1\\ \text{Philosopher} \rightarrow \text{PerfectEnding} = 1\\ \text{PerfectEnding} = 1\end{array}$$
but $\text{Marry} = 0$

梁实秋 vs 鲁迅

说我是资本家的走狗，是哪一个资本家，还是所有的资本家？我还不知道我的主子是谁.

— 梁实秋《资本家的走狗》

凡走狗，虽或为一个资本家所豢养，其实是属于所有的资本家的，所以它遇见所有的阔人都驯良，遇见所有的穷人都狂吠.
不知道谁是它的主子，正是它遇见所有阔人都驯良的原因，也就是属于所有的资本家的证据.
即使无人豢养，饿的精瘦，变成野狗了，但还是遇见所有的阔人都驯良，遇见所有的穷人都狂吠的，不过这时它就愈不明白谁是主子了.

— 鲁迅《丧家的资本家的乏走狗》

有主子豢养 → 是走狗

无主子豢养

不是走狗 ×

两小儿辩日

《列子 · 汤问》

孔子东游, 见两小儿辩斗, 问其故.

- 一儿曰: “我以日始出时去人近, 而日中时远也.”
- 一儿曰: “我以日初出远, 而日中时近也.”
- 一儿曰: “日初出大如车盖, 及日中则如盘盂, 此不为远者小而近者大乎?”
- 一儿曰: “日初出沧沧凉凉, 及其日中如探汤, 此不为近者热而远者凉乎?”
- 孔子不能决也.
- 两小儿笑曰: “孰为汝多知乎?”

$$\frac{\begin{array}{c}\text{日出} \to \text{大},\ \text{日中} \to \text{小} \\ \text{大者} \to \text{近},\ \text{小者} \to \text{远}\end{array}}{\text{日出} \to \text{近},\ \text{日中} \to \text{远}} \qquad \frac{\begin{array}{c}\text{日出} \to \text{凉},\ \text{日中} \to \text{热} \\ \text{凉者} \to \text{远},\ \text{热者} \to \text{近}\end{array}}{\text{日出} \to \text{远},\ \text{日中} \to \text{近}}$$

Example

明·浮白斋主人《雅谑》

叶衡罢相归，一日病，问诸客曰："我且死，但未知死后佳否？"一士曰："甚佳"．叶惊问曰："何以知之？"答曰：**"使死而不佳，死者皆逃回矣．一死不返，以是知其佳也．"**

$$\frac{\neg 佳 \to 回 \qquad \neg 回}{佳}$$

好货不贱，贱货不好．

$$\frac{好 \to \neg 贱}{\overline{贱 \to \neg 好}}$$

痞子蔡《第一次的亲密接触》

1. 如果把整个太平洋的水倒出，也浇不灭我对你爱情的火焰．整个太平洋的水倒得出吗？不行．所以，我不爱你．
2. 如果把整个浴缸的水倒出，也浇不灭我对你爱情的火焰．整个浴缸的水倒得出吗？可以．所以，是的，我爱你．

二难推理

- 如果你工作，就能挣钱；如果你赋闲在家，就能悠然自在．你或者工作或者赋闲，总之，你或者能挣钱或者能悠然自在．
- 如果你工作，就不能悠然自在；如果你赋闲在家，就不能挣钱．你或者工作或者赋闲，总之，你或者不能悠然自在或者不能挣钱．

$$\frac{A \to C \quad B \to D \quad A \vee B}{C \vee D} \qquad \frac{A \to \neg D \quad B \to \neg C \quad A \vee B}{\neg D \vee \neg C}$$

- 老婆婆有俩儿子，老大卖阳伞，老二卖雨伞，晴天雨伞不好卖，雨天阳伞不好卖……
- 被困失火的高楼，走楼梯会被烧死，跳窗会摔死……

蒋介石

反腐，亡党；不反，亡国．

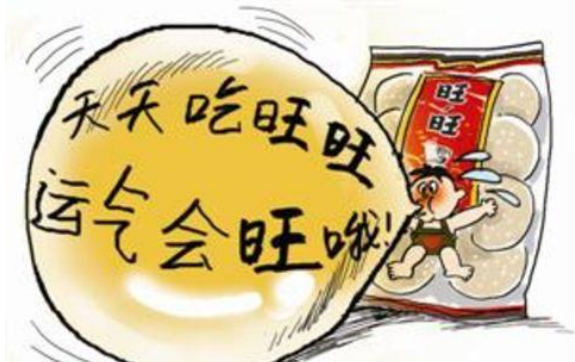
¬ 上帝万能

上帝能否创造一块自己举不起来的石头？

诉讼悖论

- 曾有师生签订合同：上学期间不收费，学生毕业打赢第一场官司后交学费.
- 可学生毕业后并未从事律师职业，于是老师威胁起诉学生.
- 老师说：如果我赢了，根据法庭判决，你必须交学费；如果你赢了，根据合同，你也必须交学费. 要么我赢要么你赢，你都必须交学费.
- 学生说：如果我赢了，根据法庭判决，我不用交学费；如果你赢了，根据合同，我不用交学费. 要么我赢要么你赢，我都不用交学费.

$$\frac{W \to P \quad \neg W \to P \quad W \vee \neg W}{P} \qquad \frac{W \wedge J \to P \quad \neg W \wedge C \to P \quad W \vee \neg W}{P} \; ? \qquad \frac{\neg W \wedge J \to \neg P \quad W \wedge C \to \neg P \quad W \vee \neg W}{\neg P} \; ? \qquad \frac{W \wedge J \to P \quad \neg W \wedge C \to P \quad (W \wedge J) \vee (\neg W \wedge C)}{P}$$



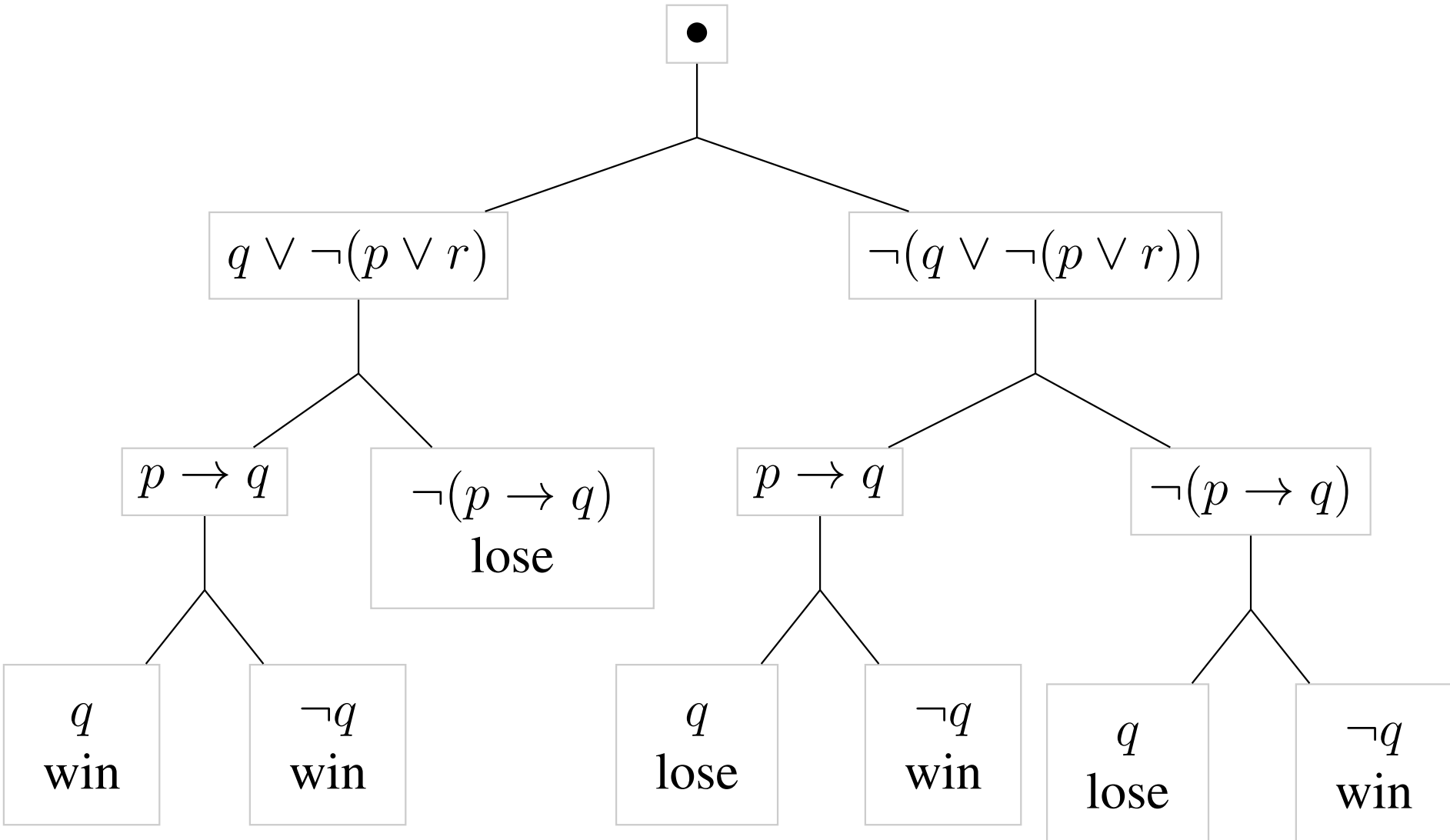
中世纪的逻辑考试

- 老师通过一种 n 轮的游戏来测试学生的逻辑水平.
- 每一轮中老师会给出一个命题 A_i.
- 学生必须选择 “接受” 或 “拒绝” 该命题.
- 如果接受该命题, 则将 A_i 放入已有命题的集合, 否则将 $\neg A_i$ 放入.
- 如果放入后的命题集合矛盾了, 则学生失败. 如果 n 轮后得到命题集合仍然没有矛盾, 学生通过测试.

Remark: 如果学生看到了老师的试题库 $\{A_1, A_2, \ldots, A_n\}$, 怎么提前备考?

比如: 假设老师的出题顺序是

1. $q \vee \neg(p \vee r)$
2. $p \to q$
3. q



反解真值表

Problem ("君子" 与 "小人")

一个岛上有 "君子"、"小人" 两类人. "君子" 只说真话, "小人" 只说假话.

1. 什么话 "君子" 可以说, 但 "小人" 不可以说?
2. 什么话 "小人" 可以说, 但 "君子" 不可以说?
3. 什么话 "君子" 和 "小人" 都可以说?
4. 什么话 "君子" 和 "小人" 都不可以说?

- p: 我是君子

身份	p	可说 x 吗?	x
小人	0	0	$\nu_1(x) = 1$
君子	1	0	$\nu_2(x) = 0$

$$x = \neg p$$

反解真值表

Problem (招驸马)

一个岛上有“君子”、“小人”两类人. “君子”只说真话, “小人”只说假话. 他们有的富有的穷. 国王现在欲招一位“穷君子”做驸马. “穷君子”可以说一句什么话证明自己的身份?

- p: 我是穷人
- q: 我是君子

身份	p	q	可说 x 吗?	x
富小人	0	0	0	$\nu_1(x) = 1$
富君子	0	1	0	$\nu_2(x) = 0$
穷小人	1	0	0	$\nu_3(x) = 1$
穷君子	1	1	1	$\nu_4(x) = 1$

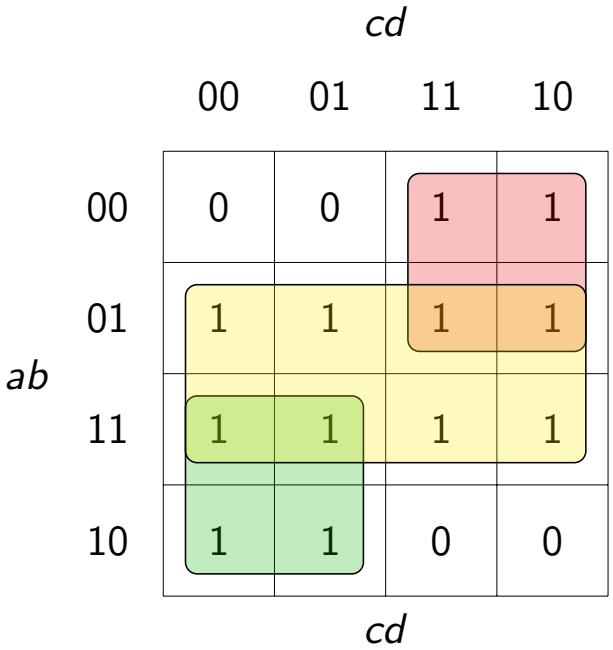
p \ q	0	1
0	1	0
1	1	1

$$x = \neg(\neg p \wedge q) \equiv p \vee \neg q \equiv q \to p \equiv \cdots$$

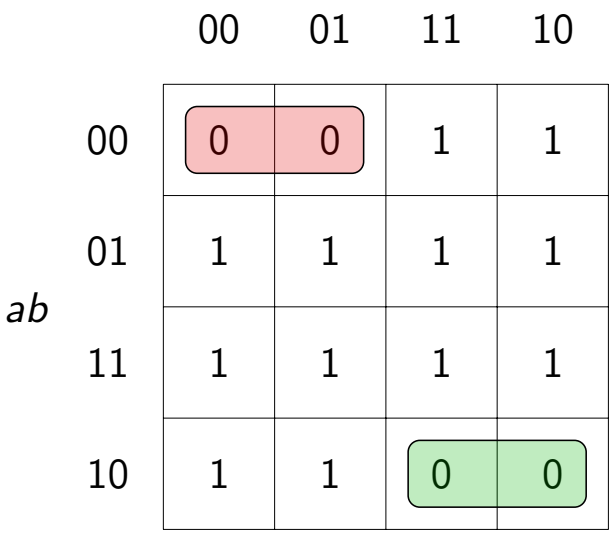
Remark: 析取/合取范式 $x = \bigvee_{k:\nu_k(x)=1} \bigwedge_{i=1}^{n} p_i^{\nu_k(p_i)} \equiv \bigwedge_{k:\nu_k(x)=0} \bigvee_{i=1}^{n} p_i^{1-\nu_k(p_i)}$

where $p^1 := p$, $p^0 := \neg p$.

可能世界集 vs 语言



$$(\neg a \wedge c) \vee (a \wedge \neg c) \vee b$$



$$\neg((\neg a \wedge \neg b \wedge \neg c) \vee (a \wedge \neg b \wedge c))$$

卡诺图: 在环面上尽量画大圈 (只含 2^n 个相邻项), 圈的个数尽量少.

Problem (鳄鱼困境)

I will return your child iff you can correctly predict what I will do next.

$$x =? \implies x \leftrightarrow r \vDash r$$

r	x	$x \leftrightarrow r$	$(x \leftrightarrow r) \rightarrow r$
0	$\nu_1(x) = ?$	?	1
1	$\nu_2(x) = ?$	?	1

r	x	$x \leftrightarrow r$	$(x \leftrightarrow r) \rightarrow r$
0	$\nu_1(x) = ?$	0	1
1	$\nu_2(x) = ?$	0/1	1

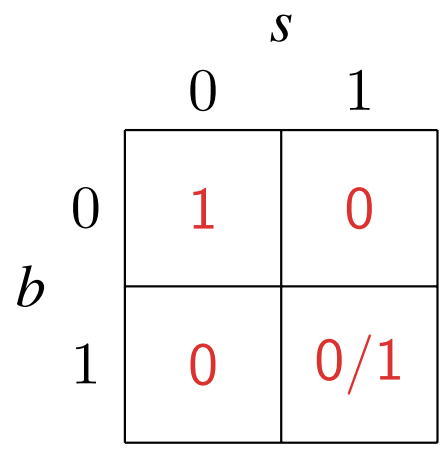
r	x	$x \leftrightarrow r$	$(x \leftrightarrow r) \rightarrow r$
0	$\nu_1(x) = 1$	0	1
1	$\nu_2(x) = 0/1$	0/1	1

$$x = \begin{cases} \neg r & \text{if } \nu_2(x) = 0 \\ r \vee \neg r & \text{if } \nu_2(x) = 1 \end{cases}$$

Remark: 若要求 $x \leftrightarrow r$ 可满足, 则需舍弃 $x = \neg r$.

Problem (怎么大奖小奖全都拿？)

- 说真话得一个大奖或一个小奖.
- 说假话不得奖.



- b: 我会得大奖.
- s: 我会得小奖.

$$x =? \implies x \leftrightarrow b \vee s \vDash b \wedge s$$

b	s	x	$b \vee s$	$x \leftrightarrow b \vee s$	$b \wedge s$	$(x \leftrightarrow b \vee s) \rightarrow b \wedge s$
0	0	$v_1(x) = 1$	0	0	0	1
0	1	$v_2(x) = 0$	1	0	0	1
1	0	$v_3(x) = 0$	1	0	0	1
1	1	$v_4(x) = 0/1$	1	0/1	1	1

$$x = \begin{cases} \neg b \wedge \neg s & \text{if } v_4(x) = 0 \\ b \leftrightarrow s & \text{if } v_4(x) = 1 \end{cases}$$

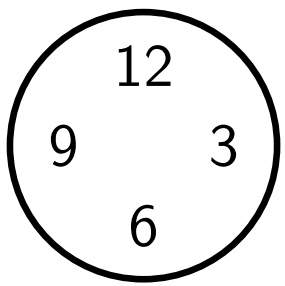
Remark: 若要求 $x \leftrightarrow b \vee s$ 可满足, 则需舍弃 $x = \neg b \wedge \neg s$.

逻辑等价

- Logical equivalence is an equivalence relation between formulas.
  1. reflexive Rxx
  2. symmetric $Rxy \to Ryx$
  3. transitive $Rxy \wedge Ryz \to Rxz$
- Logical equivalence is compatible with operators.

$$\frac{A \equiv A'}{\neg A \equiv \neg A'} \qquad \frac{\begin{array}{c} A \equiv A' \\ B \equiv B' \end{array}}{A \star B \equiv A' \star B'} \text{ where } \star \in \{\wedge, \vee, \to, \leftrightarrow\}$$

- Equivalence relation + Compatible with Operators = Congruence relation



$$2 + 3 \equiv 14 + 15 \pmod{12}$$

$$2 \times 3 \equiv 14 \times 15 \pmod{12}$$

代入

$$p_i[C/p] := \begin{cases} C & \text{if } p_i = p \\ p_i & \text{otherwise} \end{cases}$$

$$(\neg A)[C/p] := \neg A[C/p]$$

$$(A \to B)[C/p] := A[C/p] \to B[C/p]$$

Theorem (Equivalent Substitution)

$$\frac{\Gamma \vDash B \leftrightarrow C}{\Gamma \vDash A[B/p] \leftrightarrow A[C/p]}$$

Proof.

- $A = p_i$
- $A = \neg A_1$
- $A = A_1 \to A_2$

□

代入

$$p[C_1/p_1, \ldots, C_n/p_n] \coloneqq \begin{cases} C_i & \text{if } p = p_i \text{ for some } 1 \leq i \leq n \\ p & \text{otherwise} \end{cases}$$

$$(\neg A)[C_1/p_1, \ldots, C_n/p_n] \coloneqq \neg A[C_1/p_1, \ldots, C_n/p_n]$$

$$(A \to B)[C_1/p_1, \ldots, C_n/p_n] \coloneqq A[C_1/p_1, \ldots, C_n/p_n] \to B[C_1/p_1, \ldots, C_n/p_n]$$

Theorem

Consider a formula A and a sequence $C_1, \ldots, C_n$ of formulas.

1. *Let ν be a truth assignment for the set of all atomic propositions. Define μ to be the truth assignment for which $\mu(p_i) = \nu(C_i)$. Then $\mu(A) = \nu\left(A[C_1/p_1, \ldots, C_n/p_n]\right)$.*
2. $\vDash A \implies \vDash A[C_1/p_1, \ldots, C_n/p_n]$

Example: $\vDash p \vee \neg p \implies \vDash (p \wedge \neg p) \vee \neg(p \wedge \neg p)$

Digression — 何谓 “形式逻辑”?

- 何谓 “形式逻辑”? 对任意代入 $[C/p]$:

$$\frac{\Gamma \vDash A}{\Gamma[C/p] \vDash A[C/p]}$$

- 给定一个命题形式, 对相同的变元处处以相同的公式代入, 得到的公式是这个命题形式的特例.

$$p \to p \vee \neg q$$

$$A \to A \vee \neg A$$

$$(A \wedge B) \to (A \wedge B) \vee \neg (A \to B)$$

- 一个公式 A 的**命题形式**是一个命题逻辑公式 B, 使得 A 就是对 B 中相同的命题变元 p_i 处处以相同的公式 C_i 代入的结果.

$$A = B[C_1/p_1, \ldots, C_n/p_n]$$

Digression — 形式与抽象

$$\nabla(\odot\cdot\odot) = \odot\nabla\odot + \odot\nabla\odot$$

夏目漱石《梦十夜》

- “他怎能那样行云流水，凿刀所到之处，自然地雕琢出内心所想的眉毛、鼻子的样子?”
- “不难啊，那不是凿刻出眉毛、鼻子，而是眉毛、鼻子本来就埋藏在木头中，他只是用锤子、凿子将其挖出来而已.”

- 雕刻家凿掉无关的木头，使得雕像显现.
- 数学家抽象掉无关的细节，使得模式显现.

Digression — 形式与抽象 — 忽略一切无关细节!

Problem (我父亲年龄多大?)

我父亲的年龄是我的年龄的两倍, 十年前, 他的年龄是我的年龄的三倍.

Problem (这个袋子里有几个苹果?)

这个袋子里的苹果数量是那个袋子的两倍, 如果从两个袋子里各取出十个, 那么, 这个袋子里剩下的苹果数量是那个袋子的三倍.

$$x = 2y$$
$$x - 10 = 3(y - 10)$$

$$\begin{aligned} x - 2y &= 0 \\ x - 3y &= -20 \end{aligned} \qquad \left[\begin{array}{cc|c} 1 & -2 & 0 \\ 1 & -3 & -20 \end{array}\right]$$

$$\begin{bmatrix} 1 & -2 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ -20 \end{bmatrix} \qquad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & -3 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ -20 \end{bmatrix} = \begin{bmatrix} 40 \\ 20 \end{bmatrix}$$

Equivalent Replacement

Theorem (等价替换定理)

假设 B 是 A 的子公式, $A[C/\!\!/B]$ 是用 C 替换 B 在 A 中的一些出现后得到的结果. 则

$$\frac{\Gamma \vDash B \leftrightarrow C}{\Gamma \vDash A \leftrightarrow A[C/\!\!/B]}$$

Proof.

归纳证明. □

Example ☹

1. A logician's wife is having a baby.
2. The doctor immediately hands the newborn to the dad.
3. His wife asks impatiently: "So, is it a boy or a girl"?
4. The logician replies: "yes".

- wife.

$$B?$$

- logician.
  1. $A = B \vee G \qquad A?$
  2. $G \leftrightarrow \neg B$
  3. $A[\neg B /\!\!/ G] = B \vee \neg B \qquad \checkmark$

Duality

Theorem

Let A be a formula whose only connectives are $\neg, \wedge, \vee$. Let A^* be the result of interchanging $\wedge$ and $\vee$ and replacing each atomic proposition by its negation. Then $\neg A \equiv A^*$.

Proof.

Prove by induction.

- $A = p_i$
- $A = \neg B$
- $A = B \wedge C$
- $A = B \vee C$

□

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Meta-Theorems
- Boolean Algebra
- Application

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

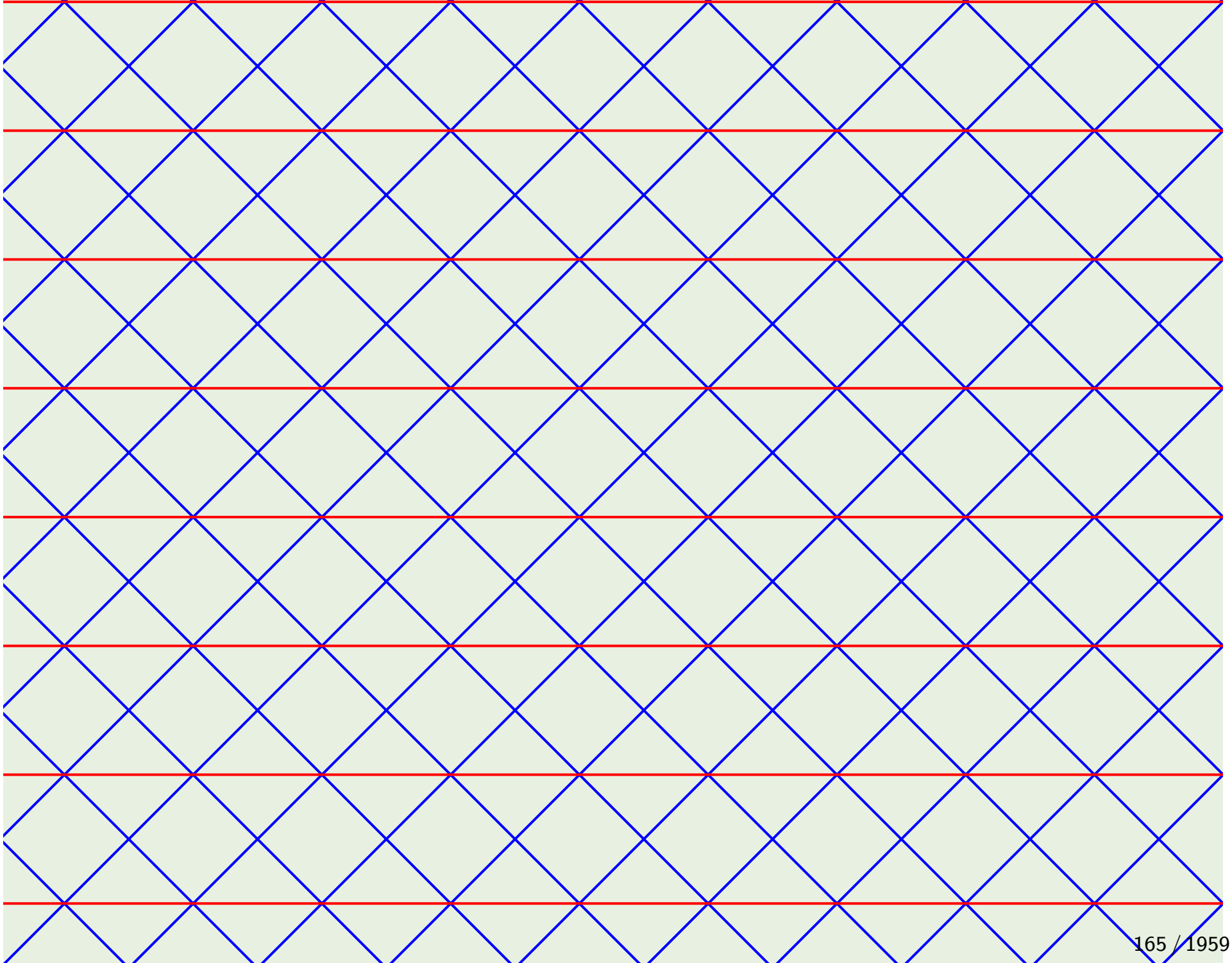
References2039

Why Study Formal System?

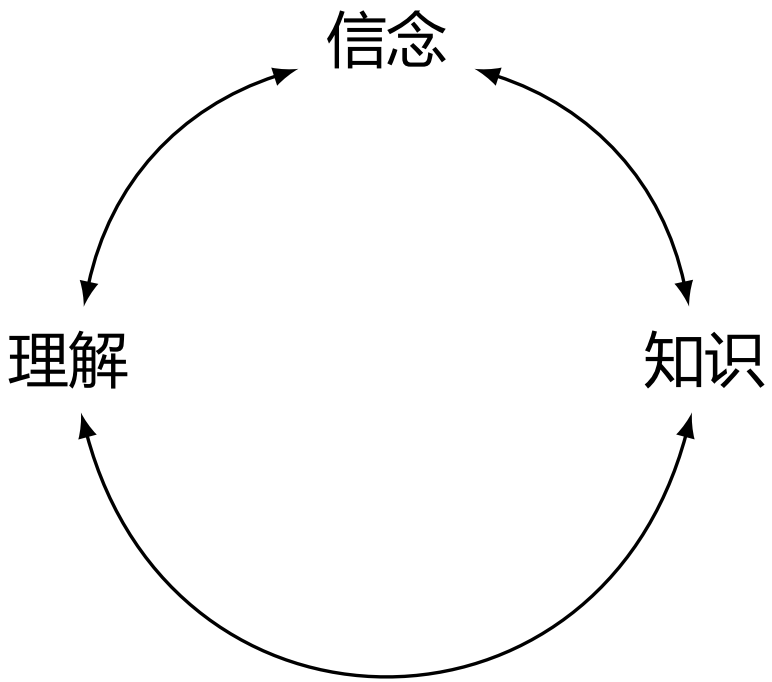
Why truth tables are not sufficient?

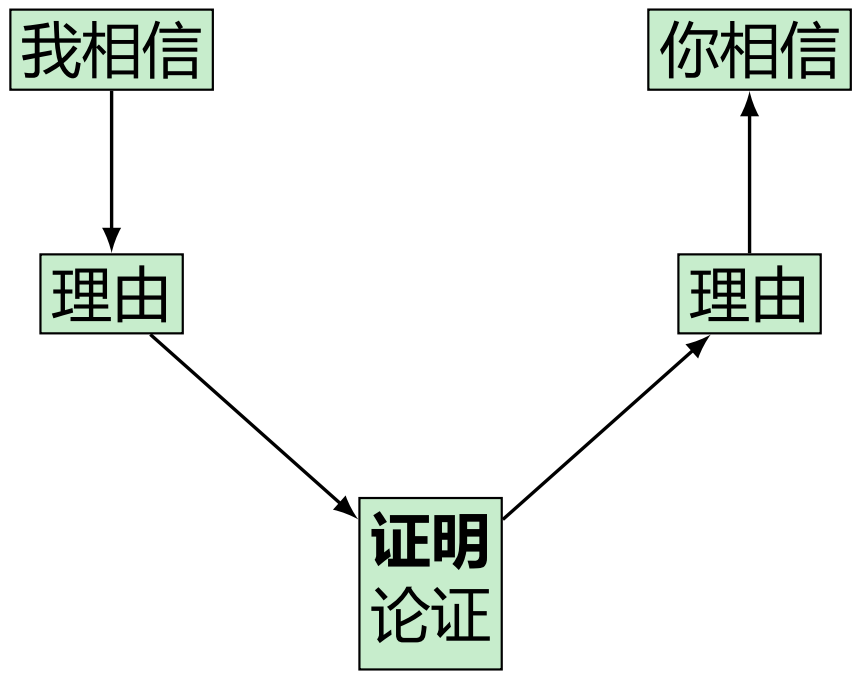
- Exponential size
  - How many times would you have to fold a piece of paper($0.1mm$) onto itself to reach the Moon?
  - Common Ancestors of All Humans
    (1) Someone alive $1000BC$ is an ancestor of everyone alive today;
    (2) Everyone alive $2000BC$ is either an ancestor of nobody alive today or of everyone alive today;
    (3) Most of the people you are descended from are no more genetically related to you than strangers are.
    (4) Even if everyone alive today had exactly the same set of ancestors from $2000BC$, the distribution of one's ancestors from that population could be very different.
- Inapplicability beyond Boolean connectives.

42

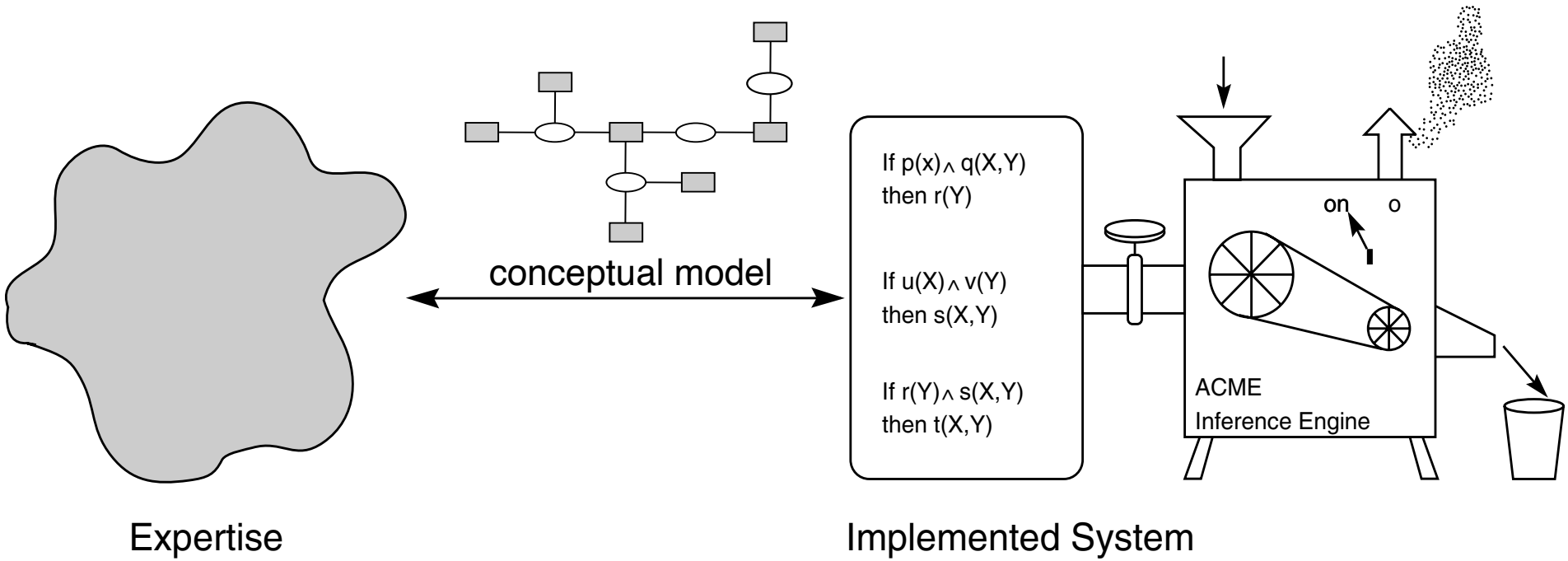


理解、信念、知识与证明



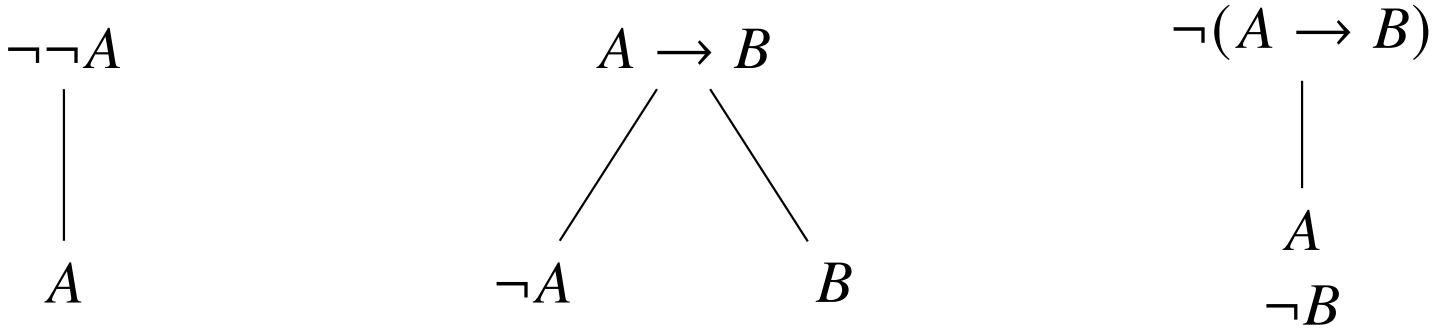


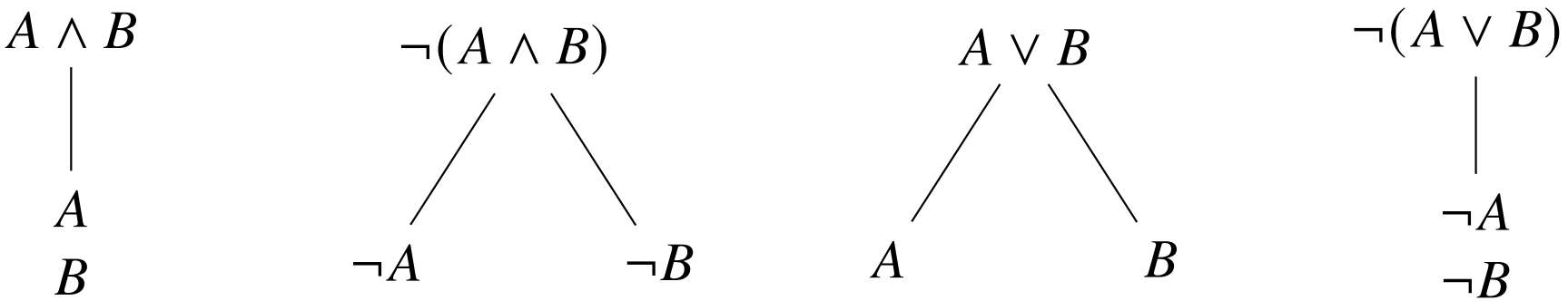
Formal Systems



- Tree Method
- Natural Deduction
- Sequent Calculus
- Hilbert System
- Resolution
- ...

命题逻辑的树形方法 ♡



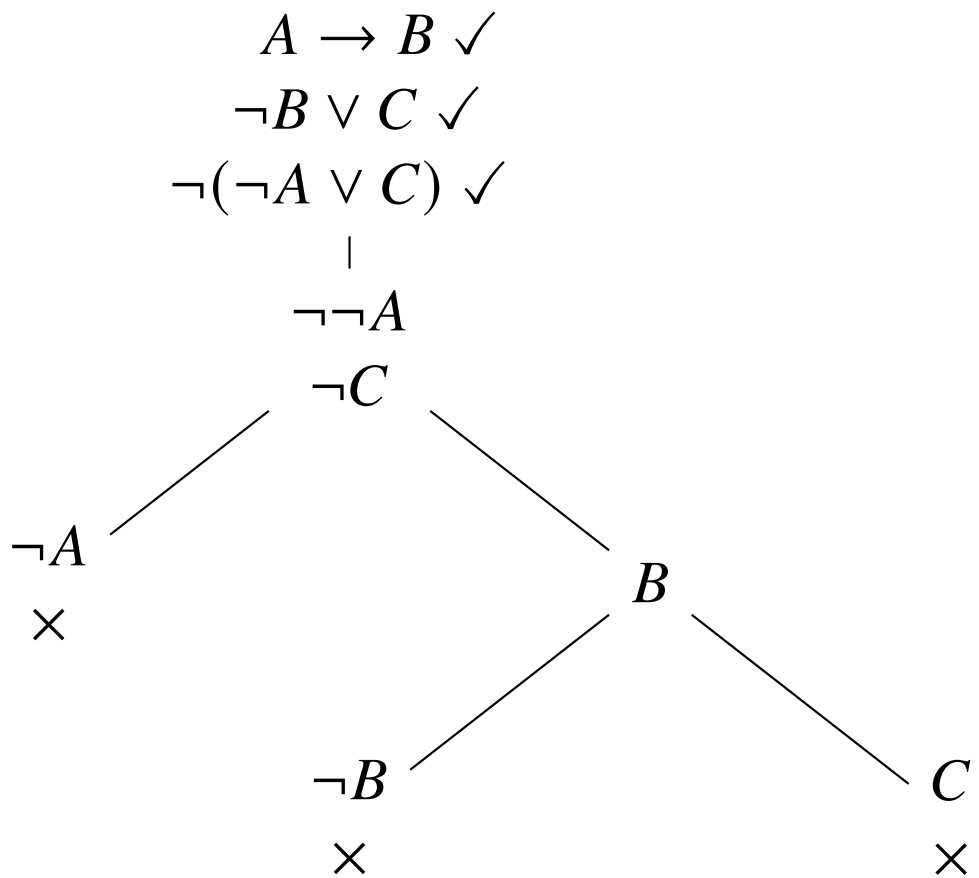




✓

证明 = 闭树

$$\frac{A \to B \quad \neg B \vee C}{\neg A \vee C}$$



树形方法指南[6] ♡

文字 一个原子公式或原子公式的否定.

闭枝 一枝上出现了某个公式及其否定.

闭树 所有枝都是闭枝的树.

1. 以前提和结论的否定作为根节点画树.
2. 检查每一个开路径, 如果在其上发现了矛盾, 则闭掉这个路径 ×.
3. 如果一个非文字公式在所有开路径上都拆过了, 则用 ✓ 标记.
4. 如果在所有开路径上都没有没标记过的非文字公式, 则停止画树!
5. 否则的话, 在开路径上选一个没标记过的非文字公式继续拆.
6. Goto 2.

[6] Tree Proof Generator: https://www.umsu.de/trees/

什么是 “证明”? ♡

Definition (证明)

$A_1, \ldots, A_n \vdash B$ 当且仅当, 存在一棵从 $\{A_1, \ldots, A_n, \neg B\}$ 开始的闭树.

小技巧

1. 少生枝节: 既能生一枝又能生两枝时, 优先生一枝.
2. 能闭掉的枝尽早闭掉.

Theorem (可靠性定理 & 完备性定理)

$$\frac{A_1, \ldots, A_n \vdash B}{A_1, \ldots, A_n \vDash B}$$

$\vdash$ *captures* $\vDash$
No more, no less

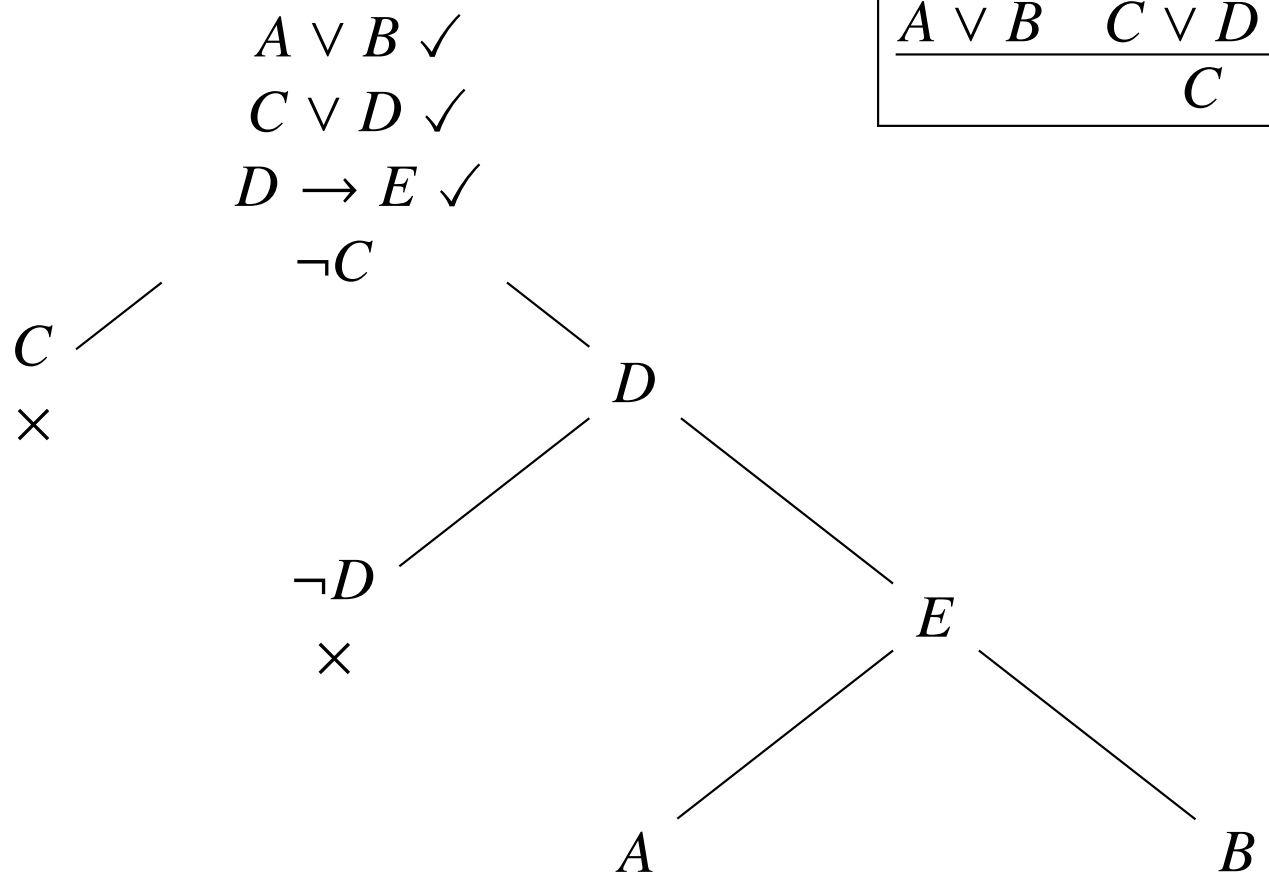
可靠 $\vdash \Longrightarrow \vDash$ 不多: 所有证明出来的论证都是有效的
完备 $\vDash \Longrightarrow \vdash$ 不少: 所有有效的论证都能够证明出来

Remark: 如果一个命题逻辑的论证不是有效的, 那么, 至少有一枝闭不掉. 通过闭不掉的开枝可以构造使得论证无效的反模型.

Remark: 闭不掉的开枝给出满足根节点的模型.

若无效，“开枝”给出“反模型”

$$\frac{A \vee B \quad C \vee D \quad D \to E}{C} \ ?$$

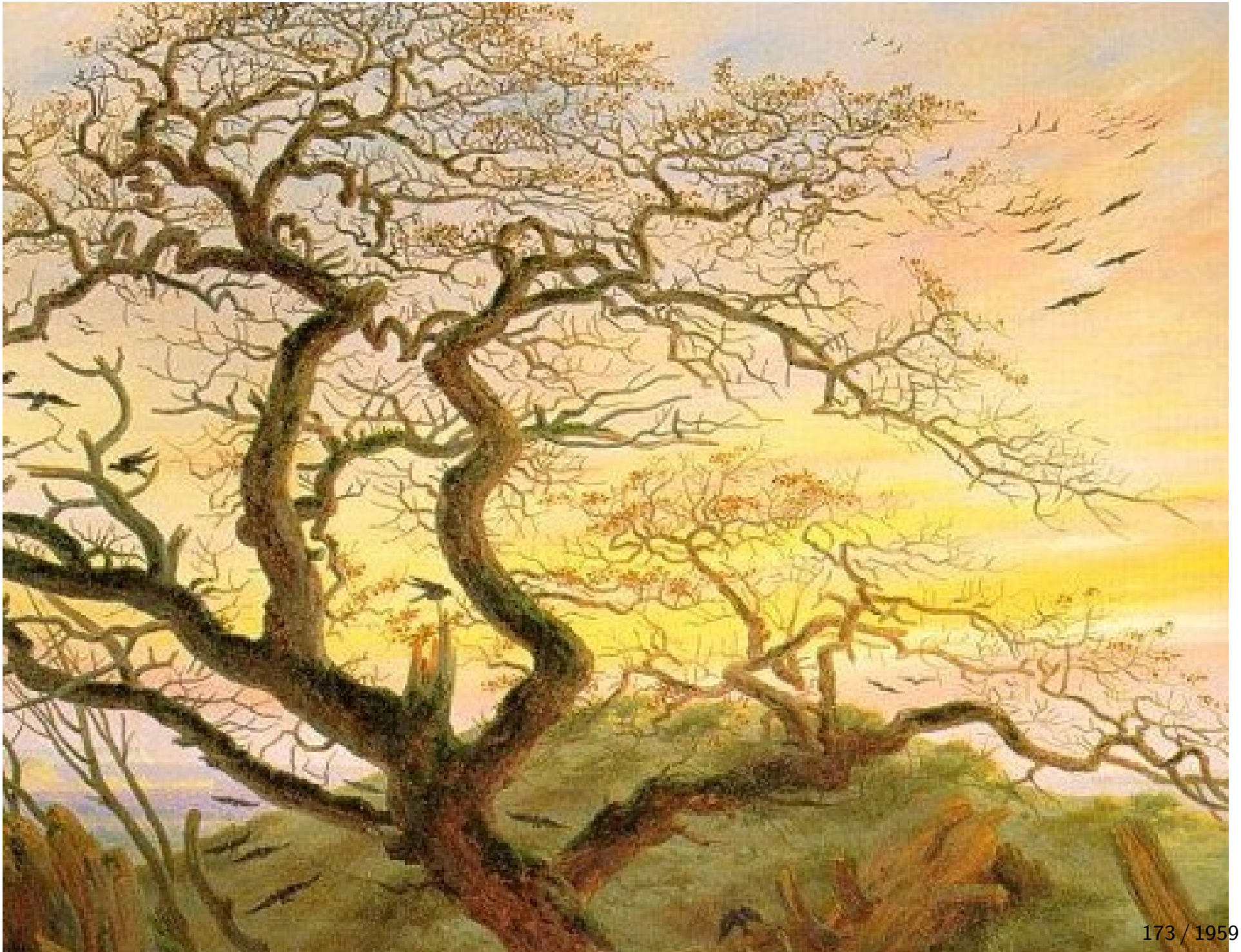


令 $A = 1, \quad E = 1, \quad D = 1 \quad C = 0 \quad B = 1 \ (\text{or } 0)$

$\quad B = 1, \quad E = 1, \quad D = 1 \quad C = 0 \quad A = 1 \ (\text{or } 0)$

则 $A \vee B = 1, \quad C \vee D = 1, \quad D \to E = 1$

但 $C = 0$



淘金客

这个岛上有金子吗?

一个岛上有“君子”、“小人”两类人. “君子”只说真话, “小人”只说假话. 你是一个淘金客, 来岛上遇到了一个土著.

- 你: “这个岛上有金子吗?”
- 土著: “如果我是君子, 那么这个岛上有金子.”

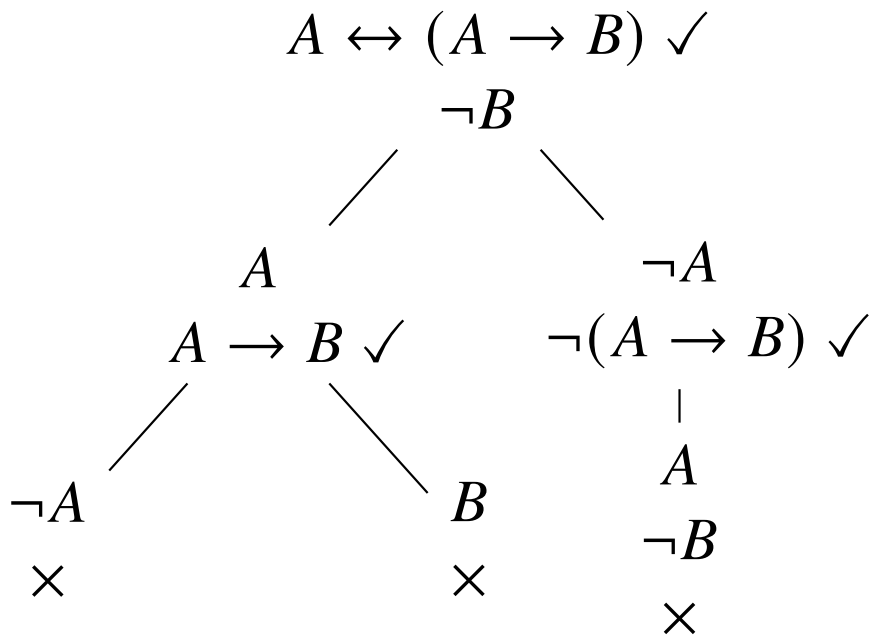
淘金客

这个岛上有金子吗?

一个岛上有“君子”、“小人”两类人. “君子”只说真话, “小人”只说假话. 你是一个淘金客, 来岛上遇到了一个土著.

- 你: “这个岛上有金子吗?”
- 土著: “如果我是君子, 那么这个岛上有金子.”

$$\frac{A \leftrightarrow (A \to B)}{B}$$



Curry's Paradox ☹

如果这句话是真的, 那么上帝存在.

$$\frac{A \leftrightarrow (A \to B)}{B}$$

Proof.

1. $A \leftrightarrow (A \to B)$
2. $A \to A \to B$
3. $(A \to A) \to A \to B$
4. $A \to B$
5. A
6. B

□

1. 甲: 如果我没说错, 那么上帝存在.
2. 乙: **如果你没说错**, 那么上帝存在.
3. 甲: 你承认我没说错了?
4. 乙: 当然.
5. 甲: 可见我没说错. 你已经承认: **如果我没说错, 那么上帝存在.** 所以, 上帝存在.

这句话是假的, 并且, 上帝不存在.

$$\frac{A \leftrightarrow (\neg A \wedge \neg B)}{B}$$

Smullyan 实力撩妹 ♡◎♡

1. "我说一句话, 如果它是真的, 可以给我签个名吗?"
2. "没问题呀."
3. "如果它是假的, 就不要给我签名了."
4. "好的."
5. 然后 Smullyan 说了一句话. 美女发现必须给他一个吻!

$$x =? \implies a \leftrightarrow x \vDash k$$

Hi 美女, 问你个问题呗

如果我问你"你能做我女朋友吗", 那么你的答案和**这个问题**的答案是一样的吗?

$$\frac{Q \leftrightarrow (G \leftrightarrow Q)}{G}$$



Don't just read it; fight it!

Ask your own questions,
look for your own examples,
discover your own proofs.
Is the hypothesis necessary?
Is the converse true?
What happens in the classical special case?
What about the degenerate cases?
Where does the proof use the hypothesis?

练习: 树形证明 — Now it's your turn ↯

$$\frac{}{(A \to B) \vee (B \to C)} \qquad \frac{A \to C \quad B \to D}{A \wedge B \to C \wedge D} \qquad \frac{A \to C \quad B \to D}{\neg C \vee \neg D \to \neg A \vee \neg B}$$

$$\frac{A \to B \quad \neg A \to B}{B} \qquad \frac{A \to C \quad B \to D}{A \vee B \to C \vee D} \qquad \frac{(A \to C) \wedge (B \to C)}{A \vee B \to C}$$

$$\frac{(A \to B) \to C}{B \to C} \qquad \frac{(A \to C) \vee (B \to C)}{A \wedge B \to C} \qquad \frac{C \to A \quad C \to B}{C \to A \wedge B}$$

$$\frac{A \to \neg B \to B}{A \to B} \qquad \frac{\neg(A \leftrightarrow B)}{\neg A \leftrightarrow B} \qquad \frac{A \leftrightarrow B}{A \vee B \to A \wedge B}$$

$$\frac{A \wedge (B \vee C)}{(A \wedge B) \vee (A \wedge C)} \qquad \frac{A \vee (\neg A \wedge B)}{A \vee B} \qquad \frac{A \to B \to C}{A \wedge B \to C}$$

练习: 有效性判定 — Now it's your turn ⚡

以下推理是否有效? 若无效, 请构造反模型.

$$\frac{B \to C}{(A \to B) \to C} \; ? \qquad \frac{A \vee (B \wedge C)}{(A \vee B) \wedge C} \; ? \qquad \frac{(A \leftrightarrow B) \to C}{A \leftrightarrow (B \to C)} \; ?$$

$$\frac{(A \to C) \to C}{(A \to B) \to C} \; ? \qquad \frac{A \to B \wedge C \quad \neg(A \vee B \to C)}{A} \; ?$$

$$\frac{(A \to C) \vee (B \to C)}{A \vee B \to C} \; ? \qquad \frac{A \to B}{(A \to C) \to (B \to C)} \; ?$$

1. 如果你不愿意做我女朋友, 那么 "如果我表白, 你答应" 是不可能的. 我不表白. 所以, 你愿意.
2. If you concentrate only if you are threatened, then you will not pass unless you are threatened — provided that concentrating is a necessary condition for passing.

自然推演 Natural Deduction

$$\frac{A \quad B}{A \wedge B}\ \wedge^{+} \qquad \frac{A \wedge B}{A}\ \wedge^{-} \qquad \frac{A \wedge B}{B}\ \wedge^{-}$$

$$\frac{A}{A \vee B}\ \vee^{+} \qquad \frac{B}{A \vee B}\ \vee^{+} \qquad \frac{\begin{array}{ccc} A \vee B & \overset{[A]^n}{\vdots}\ C & \overset{[B]^n}{\vdots}\ C \end{array}}{C}\ \vee^{-n}$$

$$\frac{\overset{[A]^n}{\vdots}\ B}{A \to B}\ \to^{+n} \qquad \frac{A \to B \quad A}{B}\ \to^{-}$$

$$\frac{\overset{[A]^n}{\vdots}\ \bot}{\neg A}\ \neg^{+n} \qquad \frac{\overset{[\neg A]^n}{\vdots}\ \bot}{A}\ \neg^{-n} \qquad \frac{\neg A \quad A}{\bot}\ \bot^{+} \qquad \frac{\bot}{A}\ \bot^{-}$$

试给出关于命题连接词 $\leftrightarrow$ 的引入和消去规则.

Examples

$$\frac{A \vee B \quad \neg B}{A}$$

Proof.

$$\dfrac{A \vee B \qquad \overset{[A]^1}{\vdots}\ A \qquad \dfrac{\dfrac{[B]^1 \quad \neg B}{\bot}\,\bot^-}{A}\,\bot^+}{A}\,\vee^{-1}$$

□

$$\frac{A \to B \quad \neg B}{\neg A}$$

Proof.

$$\dfrac{\dfrac{A \to B \quad [A]^1}{B}\,\to^- \qquad \neg B}{\bot}\,\bot^+$$
$$\frac{\bot}{\neg A}\,\neg^{+1}$$

Examples

$$\boxed{\frac{\neg\neg A}{A}}$$

Proof.

$$\dfrac{\dfrac{\neg\neg A \quad [\neg A]^1}{\bot}\,\neg^-}{A}\,{}^{1}\,\bot^+$$

□

$$\boxed{\frac{A}{\neg\neg A}}$$

Proof.

$$\dfrac{\dfrac{A \quad [\neg A]^1}{\bot}\,\bot^+}{\neg\neg A}\,\neg^+{}^{1}$$

□

$$\boxed{\frac{\neg A \to A}{A}}$$

Proof.

$$\dfrac{\dfrac{\neg A \to A \quad [\neg A]^1}{A}\,\to^- \qquad [\neg A]^1}{\bot}\,\bot^+$$

$$\dfrac{\bot}{A}\,\neg^-{}^{1}$$

□

Examples

$$\frac{A \vee B \to C}{(A \to C) \wedge (B \to C)}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{[A]^1}{A \vee B}\vee^+ \quad A \vee B \to C}{C}\to^- }{A \to C}\to^{+1} \quad \dfrac{\dfrac{\dfrac{[B]^2}{A \vee B}\vee^+ \quad A \vee B \to C}{C}\to^- }{B \to C}\to^{+2}}{(A \to C) \wedge (B \to C)}\wedge^+$$

□

$$\frac{(A \to C) \wedge (B \to C)}{A \vee B \to C}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{(A \to C) \wedge (B \to C)}{A \to C}\wedge^- \quad [A]^1}{C}\to^- \quad \dfrac{\dfrac{\dfrac{(A \to C) \wedge (B \to C)}{B \to C}\wedge^- \quad [B]^1}{C}\to^- \quad [A \vee B]^2}{C}\vee^{-1}}{A \vee B \to C}\to^{+2}$$

Examples

$$\boxed{\dfrac{\neg A \to \bot}{A}}$$

Proof.

$$\dfrac{\dfrac{\neg A \to \bot \qquad [\neg A]^1}{\bot}\ {\to}^{-}}{A}\ \neg^{-1}$$

□

$$\boxed{\dfrac{A \wedge \neg A}{B}}$$

Proof.

$$\dfrac{\dfrac{A \wedge \neg A}{A}}{A \vee B} \qquad \dfrac{A \wedge \neg A}{\neg A}$$
$$B$$

□

$$\boxed{\dfrac{}{A \vee \neg A}}$$

Proof.

$$\dfrac{\dfrac{[\neg(A \vee \neg A)]^2 \qquad \dfrac{[A]^1}{A \vee \neg A}\ \vee^{+}}{\bot}\ \bot^{+}}{\neg A}\ \neg^{+1}$$
$$\dfrac{[\neg(A \vee \neg A)]^2 \qquad \dfrac{\neg A}{A \vee \neg A}\ \vee^{+}}{\bot}\ \bot^{+}$$
$$A \vee \neg A\ \neg^{-2}$$

□

Examples

$$\frac{\neg A \vee \neg B}{\neg(A \wedge B)}$$

$$\dfrac{\dfrac{\neg A \vee \neg B \qquad \dfrac{[A \wedge B]^1}{A}}{\neg B} \qquad \dfrac{[A \wedge B]^1}{B}}{\dfrac{\bot}{\neg(A \wedge B)}\ \neg^{+1}}$$

$$\frac{\neg(A \wedge B)}{\neg A \vee \neg B}$$

$$\dfrac{\neg(A \wedge B) \qquad \dfrac{\dfrac{[\neg(\neg A \vee \neg B)]^3 \qquad \dfrac{[\neg A]^1}{\neg A \vee \neg B}}{\bot}\ \neg^{-1} \qquad \dfrac{\dfrac{[\neg(\neg A \vee \neg B)]^3 \qquad \dfrac{[\neg B]^2}{\neg A \vee \neg B}}{\bot}\ \neg^{-2}}{B}}{A \wedge B}}{\dfrac{\bot}{\neg A \vee \neg B}\ \neg^{-3}}$$

Example — 福尔摩斯《血字的研究》

- 这起谋杀的目的不是抢劫 $\neg R$, 因为死者身上的东西没有少 $\neg S$.
- 不是抢劫, 那么是政治暗杀 P 呢? 还是情杀 Q 呢?
- 我倾向后者 Q.
- 因为在政治暗杀中 P, 凶手一经得手势必立即逃离现场 L.
- 而在这起谋杀案中, 凶手没有立即离开现场 $\neg L$, 因为在屋子里到处留下了足迹 F.

$$\boxed{\neg S, \neg S \to \neg R, \neg R \to P \vee Q, P \to L, F \to \neg L, F \vdash Q}$$

$$\dfrac{\dfrac{\dfrac{\neg S \quad \neg S \to \neg R}{\neg R} \qquad \neg R \to P \vee Q}{P \vee Q} \qquad \dfrac{P \to L \qquad \dfrac{F \to \neg L \quad F}{\neg L}}{\neg P}}{Q}$$



Example — 谁是盗贼?

谁偷了《白玉美人》?

1. 《白玉美人》是白展堂或楚留香偷的.
2. 如果是白展堂偷的, 则偷窃时间不会在午夜前.
3. 如果楚留香的证词正确, 则午夜时烛光未灭.
4. 如果楚留香的证词不正确, 则偷窃发生在午夜前.
5. 午夜时没有烛光.

1. $B \vee C$
2. $B \to \neg M$
3. $T \to L$
4. $\neg T \to M$
5. $\neg L$

$$\dfrac{\dfrac{B \vee C \qquad \dfrac{B \to \neg M \qquad \dfrac{\neg T \to M \qquad \dfrac{T \to L \quad \neg L}{\neg T}}{M}}{\neg B}}{C}}$$

Example — 自然推演

如果在有限长的线段 L 上有无穷多个点的话，那么，如果这些点都有长度，则 L 将无限长；如果这些点都没有长度，则 L 将没有长度．而一个有限长的线段不可能无限长，也不可能没有长度．因此，在有限长的线段上不可能有无穷多个点．

$$\frac{\begin{array}{c} D \to (L \to I) \wedge (\neg L \to N) \\ \neg I \\ \neg N \end{array}}{\neg D}$$

$$\dfrac{\dfrac{\dfrac{[D]^1 \quad D \to (L \to I) \wedge (\neg L \to N)}{(L \to I) \wedge (\neg L \to N)}}{L \to I} \qquad \dfrac{\dfrac{\dfrac{[D]^1 \quad D \to (L \to I) \wedge (\neg L \to N)}{(L \to I) \wedge (\neg L \to N)}}{\neg L \to N} \qquad \neg N}{L}}{I} \qquad \neg I}{\dfrac{\bot}{\neg D}\ \neg^{+1}}$$

Valid Arguments & Proof Methods

1. **Direct Proof(直接证明):**

$$\frac{\begin{array}{l} A \to B \\ A \end{array}}{B}$$

   Example: The sum of two rational numbers is rational.

2. **Backward Reasoning(反向推理):** to prove B, find A and $A \to B$.
   Example: If x and y are non-negative real numbers, then

$$\frac{x+y}{2} \geq \sqrt{xy}$$

3. **Proof by Contraposition(逆否证明):**

$$\frac{A \to B}{\neg B \to \neg A}$$

   Example: If n is an integer and $3n+2$ is odd, then n is odd.

4. **Proof by Cases(分情况证明):**

$$\frac{A \vee B \to C}{(A \to C) \wedge (B \to C)}$$

Example: If n is an integer, then $n(n+1)$ is even.

5. **Proof by Elimination(排除法):**

$$\frac{A \to B \vee C}{A \wedge \neg B \to C}$$

Example: If $a + b\mathrm{i}$ and $c + d\mathrm{i}$ are complex numbers for which $(a + b\mathrm{i})(c + d\mathrm{i}) = 1$, then $a \neq 0$ or $b \neq 0$.

6. **Proof by Contradiction(反证法):**

$$\frac{\neg A \to B \qquad \neg A \to \neg B}{A}$$

Example: Euclid's theorem — There are infinitely many primes.

7. **Reductio Ad Absurdum (归谬法):**

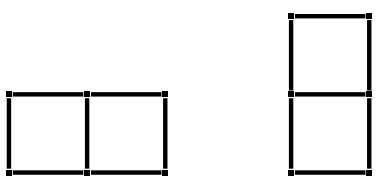
$$\frac{A \to B \qquad A \to \neg B}{\neg A}$$

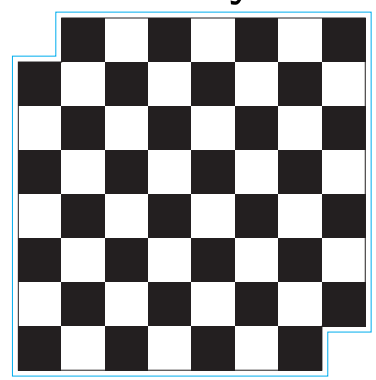
Example: There is no largest natural number.

Example: $\sqrt{2}$ is irrational.

Example: Russell Paradox.

Example: The following board cannot be tiled by the dominos.





Hilbert Formal System = Axiom + Inference Rule

公理模式

1. $A \to B \to A$
2. $(A \to B \to C) \to (A \to B) \to A \to C$
3. $(\neg A \to \neg B) \to (\neg A \to B) \to A$

推理规则

$$\frac{A \quad A \to B}{B}\ \text{MP}$$

什么是 “证明”?

Definition (证明 $\Gamma \vdash A$)

$\Gamma \vdash A$ 当且仅当, 存在以 A 结尾的有穷公式序列 $(C_1, \ldots, C_n)$, 其中 $C_n = A$, 使得序列中的每个公式 $C_k, k \leq n$:

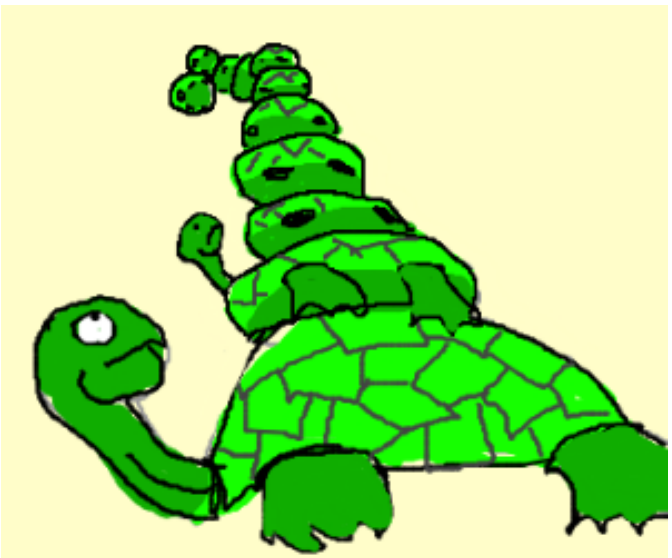
1. 或者是公理;
2. 或者在 Γ 里;
3. 或者由前面的公式通过推理规则得到.

当 $\Gamma = \emptyset$ 时, A 是定理 $\vdash A$.

A Joke ☺

数学家的房子起火了. 他的妻子用水将火扑灭了. 然后发生了煤气泄漏. 数学家点燃了它.

A 与同一物相等的两物彼此相等.
B 这个三角形的两边与同一物相等.
C 如果 A、B 为真, 那么 Z 必为真.
D 如果 A、B、C 为真, 那么 Z 必为真.
⋮
⋮
⋮
Z 这个三角形的两边彼此相等.



Example

Theorem

$$\vdash A \to A$$

Proof.

1. $A \to (A \to A) \to A$ — A1
2. $(A \to (A \to A) \to A) \to (A \to A \to A) \to A \to A$ — A2
3. $(A \to A \to A) \to A \to A$ — 1,2 MP
4. $A \to A \to A$ — A1
5. $A \to A$ — 3,4 MP

□

Remark☺ Logic is like love; a simple idea, but it can get complicated.

- 这 TM 也用证?
- 这 TM 也能证?

Example

Theorem

$$\vdash (\neg A \to A) \to A$$

Proof.

1. $(\neg A \to \neg A) \to (\neg A \to A) \to A$ A3
2. $\neg A \to \neg A$
3. $(\neg A \to A) \to A$ 1,2 MP

□

Example

Theorem

$$A \to B, B \to C \vdash A \to C$$

Proof.

1. $(B \to C) \to (A \to B \to C)$ — A1
2. $B \to C$ — Premise
3. $A \to B \to C$ — 1,2 MP
4. $(A \to B \to C) \to (A \to B) \to A \to C$ — A2
5. $(A \to B) \to A \to C$ — 4,3 MP
6. $A \to B$ — Premise
7. $A \to C$ — 5,6 MP

□

演绎定理 “$\vdash$” vs “$\rightarrow$”

Theorem (Deduction Theorem)

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B}$$

Proof.

Prove by induction on the length of the deduction sequence $(C_1, \ldots, C_n)$ of B from $\Gamma \cup \{A\}$.

Base step $n = 1$:

case1. B is an axiom. (use Axiom1.)

case2. $B \in \Gamma$.

case3. $B = A$.

Inductive step $n > 1$:

case1. B is either an axiom, or $B \in \Gamma$, or $B = A$.

case2. $C_i = C_j \rightarrow B$

$\Gamma, A \vdash C_j \implies \Gamma \vdash A \rightarrow C_j$

$\Gamma, A \vdash C_j \rightarrow B \implies \Gamma \vdash A \rightarrow C_j \rightarrow B$

$\Gamma \vdash A \rightarrow B$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Meta-Theorems
- Boolean Algebra
- Application

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Independence

Definition (Independence)

An axiom A in Γ is independent iff $\Gamma \setminus \{A\} \nvdash A$.

Find some property that makes the axiom false and the propositions deduced from the other axioms true.

- $\nvDash A$
- for all B, $\Gamma \setminus \{A\} \vdash B \implies \vDash B$

Theorem

Axiom3 is independent of Axiom1 and Axiom2.

p	$\neg p$
0	0
1	0

$\rightarrow$	0	1
0	1	1
1	0	1

Let $p = 0$ and $q = 1$, then $\nvdash (\neg p \rightarrow \neg q) \rightarrow (\neg p \rightarrow q) \rightarrow p$.

Independence

Axiom1 and Axiom2 axiomatizes the conditional ($\to$) fragment of intuitionistic propositional logic. To axiomatize the conditional fragment of classical logic, we also need *Peirce's law*: $((p \to q) \to p) \to p$.

Theorem

Peirce's law is independent of Axiom1 and Axiom2.

$\to$	0	u	1
0	1	1	1
u	0	1	1
1	0	u	1

Here we interpret 1 as "true", 0 as "false", and u as "maybe".
Let $p = u$ and $q = 0$, then $((p \to q) \to p) \to p = u$.

什么是 “理论”?

- $\mathrm{Mod}(A) \coloneqq \{v : v \vDash A\}$
- $\mathrm{Mod}(\Gamma) \coloneqq \bigcap_{A \in \Gamma} \mathrm{Mod}(A)$
- $\mathrm{Th}(v) \coloneqq \{A : v \vDash A\}$
- $\mathrm{Th}(\mathcal{K}) \coloneqq \bigcap_{v \in \mathcal{K}} \mathrm{Th}(v)$
- $\mathrm{Cn}(\Gamma) \coloneqq \{A : \Gamma \vDash A\}$

$$A \vDash B \iff \mathrm{Mod}(A) \subset \mathrm{Mod}(B)$$

What is “theory”?

- A set Γ of sentences is a theory iff $\Gamma = \mathrm{Cn}(\Gamma)$.
- A theory Γ is complete iff for every sentence A, either $A \in \Gamma$ or $\neg A \in \Gamma$.
- A theory Γ is axiomatizable iff there is a decidable set Σ of sentences s.t. $\Gamma = \mathrm{Cn}(\Sigma)$.

Consistency & Satisfiability

- Γ is consistent iff $\Gamma \nvdash \bot$.
- Γ is Post-consistent iff there is some formula $A : \Gamma \nvdash A$.

$$\boxed{\Gamma \text{ is consistent iff it is Post-consistent.}}$$

- Γ is maximal iff for every formula A, either $A \in \Gamma$ or $\neg A \in \Gamma$.
- Γ is maximal consistent iff it is both consistent and maximal.
- Γ is satisfiable iff $\mathrm{Mod}(\Gamma) \neq \emptyset$.
- Γ is finitely satisfiable iff every finite subset of Γ is satisfiable.
- If Γ is consistent and $\Gamma \vdash A$, then $\Gamma \cup \{A\}$ is consistent.
- $\Gamma \cup \{\neg A\}$ is inconsistent iff $\Gamma \vdash A$.
- If Γ is maximal consistent, then $A \notin \Gamma \implies \Gamma \cup \{A\}$ is inconsistent.
- The set $\mathrm{Th}(\nu) = \{A : \nu \vDash A\}$ is maximal consistent.

Soundness Theorem

Theorem (Soundness Theorem)

$$\Gamma \vdash A \implies \Gamma \vDash A$$

Proof.

Prove by induction on the length of the deduction sequence.
Case1: A is an axiom. (truth table)
Case2: $A \in \Gamma$
Case3:

$$\left.\begin{array}{l} \Gamma \vDash C_j \\ \Gamma \vDash C_j \to A \end{array}\right\} \implies \Gamma \vDash A$$

□

Corollary

Any satisfiable set of formulas is consistent.

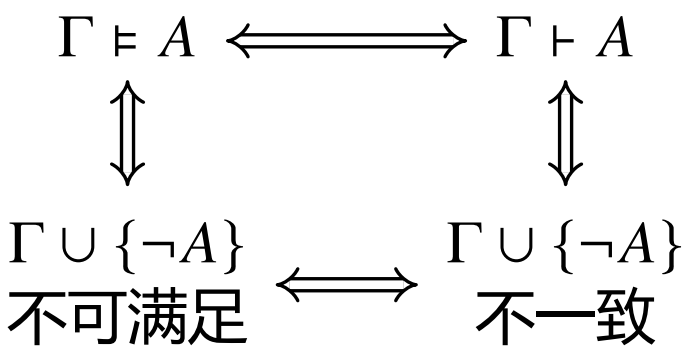
Completeness Theorem

Theorem (Completeness Theorem — Post1921)

$$\Gamma \vDash A \implies \Gamma \vdash A$$

Corollary

Any consistent set of formulas is satisfiable.



$\vdash$ captures $\vDash$
No more, no less

Corollary (Compactness Theorem)

A set of formulas is satisfiable iff it is finitely satisfiable.

Proof of Completeness Theorem

Proof.

step1. Extend the consistent set Γ to a maximal consistent set Δ.
Let $\langle A_i : i \in \mathbb{N} \rangle$ be a fixed enumeration of the formulas.

$$\Delta_0 := \Gamma$$

$$\Delta_{n+1} := \begin{cases} \Delta_n \cup \{A_n\} & \text{if } \Delta_n \cup \{A_n\} \text{ is consistent} \\ \Delta_n \cup \{\neg A_n\} & \text{otherwise} \end{cases}$$

$$\Delta := \bigcup_{n \in \mathbb{N}} \Delta_n$$

step2. Define a truth assignment that satisfies Γ.

$$v(p) := \begin{cases} 1 & \text{if } p \in \Delta \\ 0 & \text{otherwise} \end{cases} \implies \big(v \vDash A \iff A \in \Delta\big)$$

□

Compactness Theorem

Theorem (Compactness Theorem)

A set of formulas is satisfiable iff it is finitely satisfiable.

Corollary

If $\Gamma \models A$, *then there is a finite* $\Gamma_0 \subset \Gamma$ *s.t.* $\Gamma_0 \models A$.

Proof.

$$\begin{aligned} \Gamma_0 \not\models A \text{ for any } \Gamma_0 \subset \Gamma &\implies \Gamma_0 \cup \{\neg A\} \text{ is satisfiable for any } \Gamma_0 \subset \Gamma \\ &\implies \Gamma \cup \{\neg A\} \text{ is satisfiable} \\ &\implies \Gamma \not\models A \end{aligned}$$

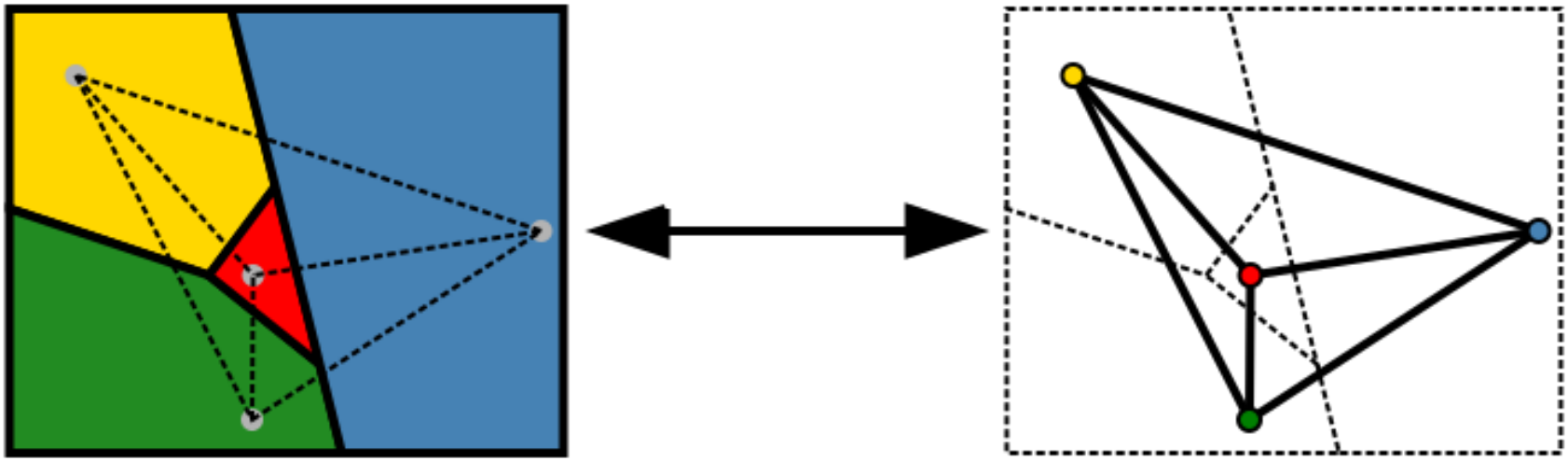
□

Remark: Γ is consistent iff every finite subset of Γ is consistent.

Remark: Compactness does not hold in a language with infinite disjunctions.

$$\left\{ \bigvee_{i=1}^{\infty} p_i, \neg p_1, \neg p_2, \dots \right\}$$

Applications of Compactness



An infinite graph (V, E) is n-colorable iff every finite subgraph of (V, E) is n-colorable.

Proof.

Take $\left\{p_v^i : v \in V, 1 \leq i \leq n\right\}$ as the set of atoms.

$\Gamma := \left\{p_v^1 \vee \cdots \vee p_v^n : v \in V\right\} \cup \left\{\neg\left(p_v^i \wedge p_v^j\right) : v \in V, 1 \leq i < j \leq n\right\} \cup$

$\left\{\neg\left(p_v^i \wedge p_w^i\right) : (v, w) \in E, 1 \leq i \leq n\right\}$ □

Proof of Compactness Theorem

Proof.

part1. Extend the finitely satisfiable set Γ to a maximal finitely satisfiable set Δ.

Let $\langle A_i : i \in \mathbb{N} \rangle$ be a fixed enumeration of the formulas.

$$\Delta_0 := \Gamma$$

$$\Delta_{n+1} := \begin{cases} \Delta_n \cup \{A_n\} & \text{if } \Delta_n \cup \{A_n\} \text{ is finitely satisfiable} \\ \Delta_n \cup \{\neg A_n\} & \text{otherwise} \end{cases}$$

$$\Delta := \bigcup_{n \in \mathbb{N}} \Delta_n$$

part2. Define a truth assignment that satisfies Γ.

$$v(p) := \begin{cases} 1 & \text{if } p \in \Delta \\ 0 & \text{otherwise} \end{cases} \implies \big(v \vDash A \iff A \in \Delta\big)$$

Weak Completeness Theorem

Lemma

Let A be a formula whose only atomic propositions are $p_1, \ldots, p_n$. Let

$$p_i^\nu := \begin{cases} p_i & \text{if } \nu \vDash p_i \\ \neg p_i & \text{otherwise} \end{cases} \qquad A^\nu := \begin{cases} A & \text{if } \nu \vDash A \\ \neg A & \text{otherwise} \end{cases}$$

then $p_1^\nu, \ldots, p_n^\nu \vdash A^\nu$.

Weak Completeness Theorem $\quad \vDash A \implies \vdash A$

$$\mu(p) := \begin{cases} 1 - \nu(p) & \text{if } p = p_n \\ \nu(p) & \text{otherwise} \end{cases}$$

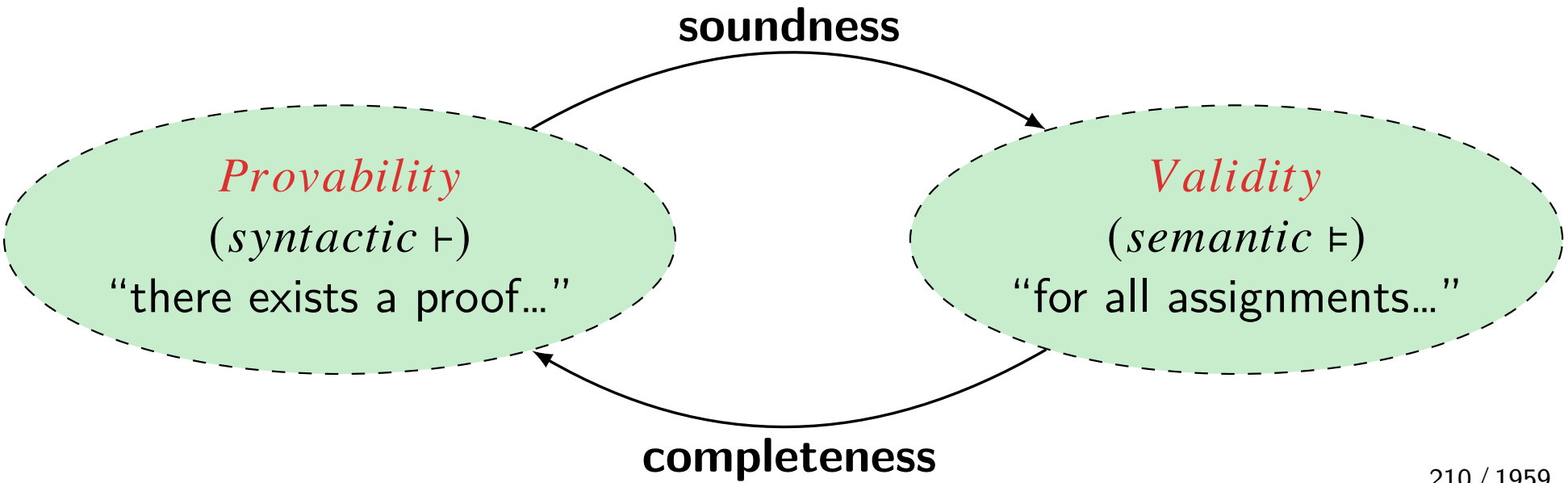
$$\left.\begin{array}{l} p_1^\nu, \ldots, p_{n-1}^\nu, p_n^\nu \vdash A \\ p_1^\mu, \ldots, p_{n-1}^\mu, p_n^\mu \vdash A \end{array}\right\} \implies p_1^\nu, \ldots, p_{n-1}^\nu \vdash A$$

$$\begin{array}{c} \vDash A \iff \vdash A \\ + \\ \text{Compactness Theorem} \\ \Downarrow \\ \Gamma \vDash A \iff \Gamma \vdash A \end{array}$$

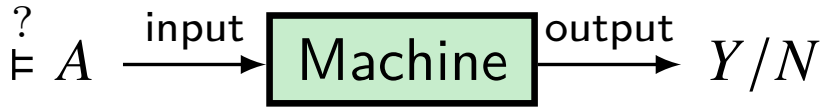
波斯特 Emil Post 1897-1954



- Truth table
- Completeness of propositional logic
- Post machine
- Post canonical system
- Post correspondence problem
- Post problem



Decidability



Theorem

There is an effective procedure that, given any expression, will decide whether or not it is a formula.

Theorem (Decidability — Post1921)

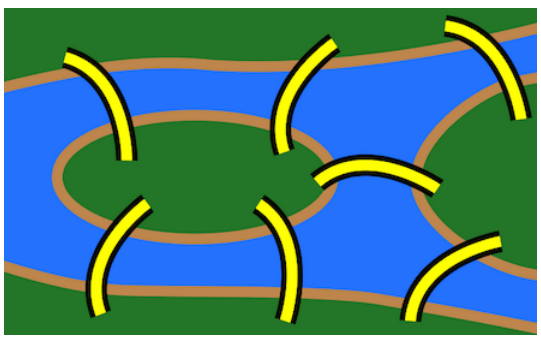
There is an effective procedure that, given a finite set $\Gamma \cup \{A\}$ of formulas, will decide whether or not $\Gamma \vDash A$.

Theorem

If Γ is a decidable set of formulas, then the set of logical consequences of Γ is recursively enumerable.

Model Checking & Satisfiability Checking & Validity Checking[7]

- Given a model ν and a formula A. Is $\nu \vDash A$? —P
- Given a formula A. Is there a model ν s.t. $\nu \vDash A$? —NP
- Given a formula A. Is $\vDash A$? —co-NP



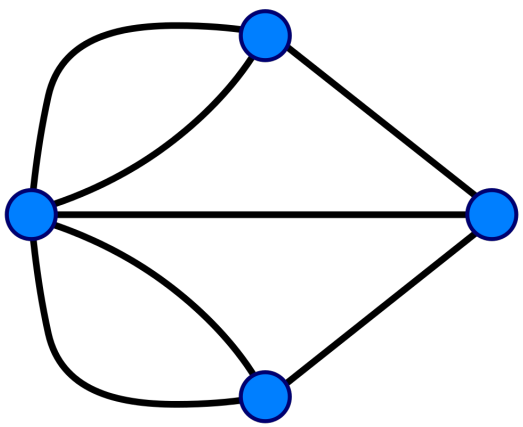


Figure: Eulerian Circle(P)

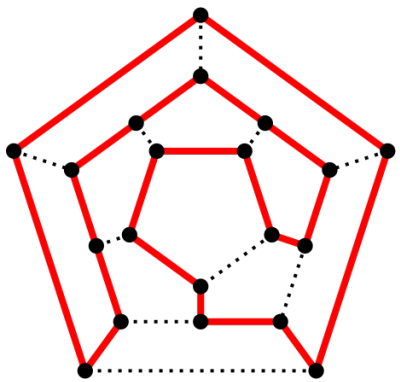


Figure: Hamiltonian Circle(NPC)

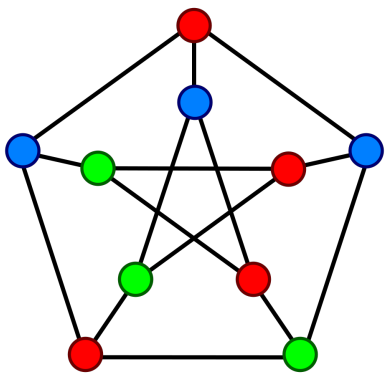
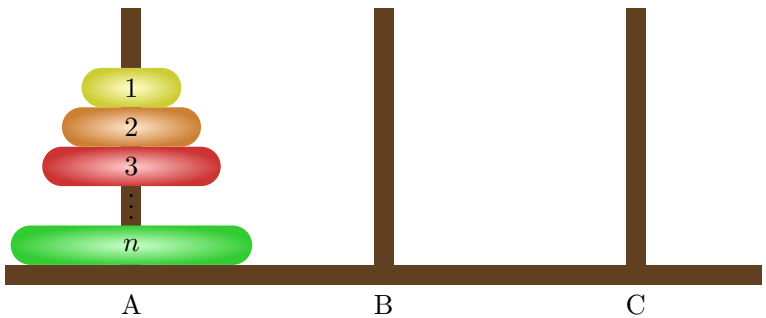


Figure: Graph Coloring(NPC)



[7] Aaronson: Why philosophers should care about computational complexity.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Meta-Theorems
- Boolean Algebra
- Application

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

布尔代数

$\bot$	$\top$	$\vee$	$\wedge$	$\neg$
0	1	$+$	$\cdot$	$^-$

$$(B, 0, 1, +, \cdot, ^-)$$

- $x + (y + z) = (x + y) + z$
  $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
- $x + y = y + x \quad x \cdot y = y \cdot x$
- $x + (x \cdot y) = x \quad x \cdot (x + y) = x$
- $x + (y \cdot z) = (x + y) \cdot (x + z)$
  $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$
- $\overline{\overline{x}} = x$
- $\overline{x + y} = \overline{x} \cdot \overline{y} \quad \overline{x \cdot y} = \overline{x} + \overline{y}$
- $x + \overline{x} = 1 \quad x \cdot \overline{x} = 0 \quad 0 \neq 1$
- $x + 0 = x \quad x \cdot 0 = 0$
  $x + 1 = 1 \quad x \cdot 1 = x$

$$x \leq y := x\overline{y} = 0$$

1. 解方程: $ax + b\overline{x} + c = 0$

$$ax = 0$$
$$b\overline{x} = 0$$
$$c = 0$$

x 的解集为:

$$b \leq x \leq \overline{a}$$

且满足有解的条件: $c = 0$ & $b \leq \overline{a}$.

2. 解方程: $ax = b$

$$ax = b \iff a\overline{b}x + b\overline{x} + \overline{a}b = 0$$
$$a\overline{b}x = 0$$
$$b\overline{x} = 0$$
$$\overline{a}b = 0$$

x 的解集为:

$$b \leq x \leq \overline{a} + b$$

且满足有解的条件: $\overline{a}b = 0$.

更多布尔运算和性质

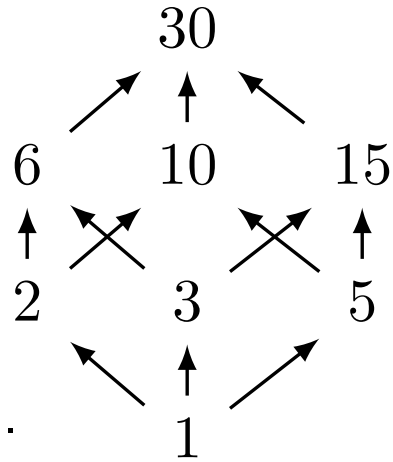
- $x - y := x\overline{y}$
- $\overline{x} = 1 - x$
- $x(y - z) = xy - xz$
- $x \le y \iff \overline{x} + y = 1 \iff x - y = 0 \iff xy = x \iff x + y = y$
- $x \oplus y := x\overline{y} + \overline{x}y$
- $x = y \iff x \oplus y = 0$
- $\overline{x} = 1 \oplus x$
- $x + x = x$
- $xx = x$ **Remark:** $x = x^2$ 意味着无矛盾律.

$$x = x^2 \implies x - x^2 = 0 \implies x(1 - x) = 0$$

- $x + y = 1 \ \& \ xy = 0 \implies y = \overline{x}$
- $x + \overline{x}y = x + y$
- $xy + \overline{x}z + yzw = xy + \overline{x}z = (x + z)(\overline{x} + y)$
- $(x + y)(\overline{x} + z)(y + z + w) = (x + y)(\overline{x} + z) = xz + \overline{x}y$

布尔代数 $(B, 0, 1, +, \cdot, \overline{\ })$ 的例子

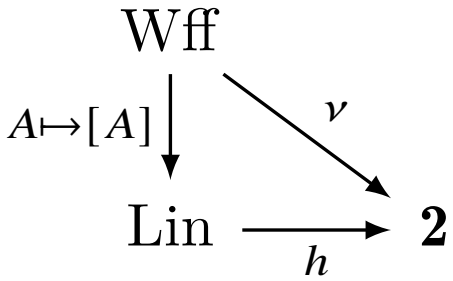
- $\left(\mathrm{P}(X), \emptyset, X, \cup, \cap, \overline{\ }\right)$
- $\left(\left\{k : k \mid n\right\}, 1, n, \mathrm{lcm}, \gcd, \frac{n}{\ }\right)$ where n is square-free.



$n = 20$ 则不成, $\overline{2} = \frac{20}{2} = 10$, $2 \cdot \overline{2} = \gcd(2, 10) = 2 \neq 1$.

- $\mathbf{2} := \left(\left\{0, 1\right\}, 0, 1, \max, \min, 1 - \right)$
- $\mathrm{Lin} := \left(\left\{[A] : A \in \mathrm{Wff}\right\}, 0, 1, +, \cdot, \overline{\ }\right)$ where $[A] := \left\{B : \vdash A \leftrightarrow B\right\}$

$$
\begin{aligned}
0 &:= [\bot] \\
1 &:= [\top] \\
[A] + [B] &:= [A \vee B] \\
[A] \cdot [B] &:= [A \wedge B] \\
\overline{[A]} &:= [\neg A]
\end{aligned}
$$



命题逻辑的布尔代数语义

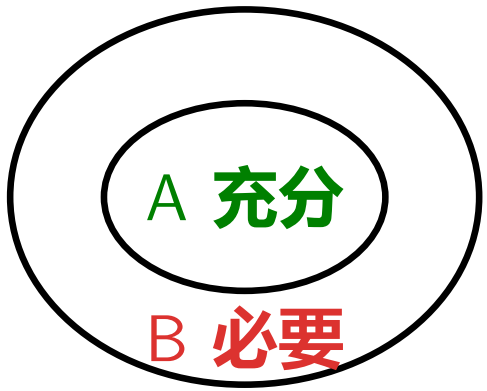
给定布尔代数 $\mathcal{M} := (M, 0, 1, +, \cdot, \bar{}, \leq)$, $\mathcal{M}$ 上的赋值 $\nu : \mathrm{Var} \to M$ 可以在满足下列条件的情况下递归地扩展到所有公式上:

1. $[\![p]\!] := \nu(p)$
2. $[\![\top]\!] := 1$
3. $[\![\bot]\!] := 0$
4. $[\![\neg A]\!] := \overline{[\![A]\!]}$
5. $[\![A \wedge B]\!] := [\![A]\!] \cdot [\![B]\!]$
6. $[\![A \vee B]\!] := [\![A]\!] + [\![B]\!]$

当 $[\![A]\!] = 1$ 时, 我们记 $\mathcal{M}, \nu \vDash A$. 若对任意 ν 都成立, 则记 $\mathcal{M} \vDash A$.

- 通常的真值赋值即取 $\mathcal{M} := \mathbf{2} := \left(\{0,1\}, 0, 1, \max, \min, 1-, \leq\right)$
- 若取 $\mathcal{M} := \left(\mathrm{P}(X), \emptyset, X, \cup, \cap, \bar{}, \subset\right)$, 则

$$\dfrac{\dfrac{\dfrac{\vDash A \to B}{A \vDash B}}{[\![A]\!] \leq [\![B]\!]}}{[\![A]\!] \subset [\![B]\!]}$$



Example

Theorem

给定幂集布尔代数 $(\mathtt{P}(X), \emptyset, X, \cup, \cap, ^{-}, \subset)$, 任取 $x \in X$, 任给真值赋值 $\nu : \mathrm{Var} \to \{0,1\}$, 任给集合赋值 $[\![\,]\!] : \mathrm{Var} \to \mathtt{P}(X)$, 使得对任意 $p \in \mathrm{Var}$, $\nu(p) = 1 \iff x \in [\![p]\!]$, 则对任意公式 A 有

$$\nu(A) = 1 \iff x \in [\![A]\!]$$

Corollary: $\vDash A \iff \mathbf{SET} \vDash A$, 其中 **SET** 是集合代数的类.

Example: 令 $\mathrm{Var} = \{p, q, r\}$, $X = \{4, 5, 6, 7\}$, 取特指值 x 为 4.

$\{0,1\}$	Var	$\mathtt{P}(\{4,5,6,7\})$
1	p	$\{4,5\}$
1	q	$\{4,6\}$
0	r	$\{5,7\}$
$\{0,1\}$	Wff	$\mathtt{P}(\{4,5,6,7\})$
0	$\neg p$	$\{6,7\}$
1	$p \wedge q$	$\{4\}$
1	$p \vee r$	$\{4,5,7\}$

命题逻辑的完全性

由

$$\vdash A \iff \mathrm{Lin} \vDash A$$

易证

$$\vdash A \iff \mathbf{BA} \vDash A$$

其中 $\mathbf{BA}$ 是布尔代数的类.
又因为已知

$$\vDash A \iff \mathbf{SET} \vDash A$$

所以, 欲证命题逻辑的完全性:

$$\vdash A \iff \vDash A$$

我们只需要

$$\mathbf{SET} \vDash A \iff \mathbf{BA} \vDash A$$

因为每个集合代数都是布尔代数, 而由 Stone 表示定理, 每个布尔代数都同构于某个集合代数. 显然成立.

命题逻辑 vs 布尔代数

- For a Boolean algebra $(B, 0, 1, +, \cdot, \bar{}, \leq)$, build language $\mathscr{L}$ by having a propositional constant P_x for each $x \in B$.
- Construct propositional theory T_B in $\mathscr{L}$ by adding for $x, y \in B$ an axiom

$$P_x \to P_y \quad \text{if } x \leq y$$

and axioms

$$P_x \wedge P_y \leftrightarrow P_{x \cdot y}$$
$$P_x \vee P_y \leftrightarrow P_{x+y}$$
$$\neg P_x \leftrightarrow P_{\overline{x}}$$

- Then Boolean algebra B is isomorphic to the Lindenbaum algebra $\mathrm{Lin}_{\mathrm{T}_B}$ of its theory T_B.

$$B \cong \mathrm{Lin}_{\mathrm{T}_B}$$

- We call two propositional theories equivalent if their Lindenbaum algebras are isomorphic. Then for a propositional theory T,

$$\mathrm{T} \equiv \mathrm{T}_{\mathrm{Lin}_{\mathrm{T}}}$$

命题逻辑 vs 布尔代数

$$\frac{x \vDash y}{x \leq y}$$

$$\frac{a \vee b, a \rightarrow c, b \rightarrow d \vDash c \vee d}{(a+b)(\overline{a}+c)(\overline{b}+d) \leq c+d}$$

$$(a+b)(\overline{a}+c)(\overline{b}+d)\overline{c}\,\overline{d} = 0$$

$$\begin{aligned}
&(a+b)(\overline{a}+c)(\overline{b}+d)\overline{c}\,\overline{d} \\
&= (a+b)(\overline{a}+c)\overline{c}(\overline{b}+d)\overline{d} \\
&= (a+b)\,\overline{a}\,\overline{c}\,\overline{b}\,\overline{d} \\
&= b\,\overline{a}\,\overline{c}\,\overline{b}\,\overline{d} \\
&= 0
\end{aligned}$$

布尔方程的通解

Theorem (布尔方程的通解)

Let $f : B \to B$ be a Boolean function for which $f(0) \cdot f(1) = 0$. Then

$$f(x) = 0$$

$$\Updownarrow$$

$$f(0) \leq x \leq \overline{f(1)}$$

$$\Updownarrow$$

$$x = f(0) + \theta \cdot \overline{f(1)} \quad \text{for } \theta \in B$$

Example: $x =? \implies \vDash (a \leftrightarrow x) \to k$

$$f(x) = \overline{\overline{(a \cdot x + \overline{a} \cdot \overline{x})} + k} = 0$$

$$\overline{a} \cdot \overline{k} \leq x \leq \overline{a} + k$$

$$x = \overline{a} \cdot \overline{k} + \theta \cdot (\overline{a} + k)$$

君子/小人

Problem (谁是君子? 谁是小人?)

有 *Ari, Benny, Carly, Darcy* 四个人.

- *Ari* 说: *Beny* 说 *Carly* 是小人.
- *Beny* 说: *Ari* 是小人或 *Carly* 是君子.
- *Carly* 说: *Beny, Darcy* 都是君子.
- *Darcy* 说: *Beny* 是君子.

1. $\overline{a} \oplus (\overline{b} \oplus \overline{c})$
2. $\overline{b} \oplus (\overline{a} + c)$
3. $\overline{c} \oplus bd$
4. $\overline{d} \oplus b$

Problem (天堂之门)

1. 你面前有左右两护卫镇守左右两门.
2. 一人只说真话, 一人只说假话.
3. 一门通天堂, 一门通地狱.
4. 你只能向其中一人提一个 "是/否" 的问题.
5. 怎么问出去天堂的门?

- 令 t 表示: 第一个人说真话.
- 设想你问第一个人问题 x, 第一个人对问题 x 的回答:

$$a := tx + \bar{t}\,\bar{x}$$

- 设想你问第二个人问题 Q: 如果我问另一个人问题 "x", 他会说 "是" 吗?
- 第二个人对问题 Q 的回答:

$$t\,\bar{a} + \bar{t}a = t\,\overline{tx + \bar{t}\,\bar{x}} + \bar{t}\,(tx + \bar{t}\,\bar{x}) = \bar{x}$$

- 令 x 为 "左门通天堂", 然后问 Q, 根据其回答, 选择相反的门.

Problem (天堂之门)

1. 你面前有左右两护卫镇守左右两门.
2. 一人只说真话, 一人只说假话.
3. 一门通天堂, 一门通地狱.
4. 你只能向其中一人提一个 “是/否” 的问题.
5. 怎么问出去天堂的门?

t \ h	0	1
0	1	0
1	0	1

- t: 你说真话
- h: 左门通天堂

$$x =? \implies \vDash (t \to (x \leftrightarrow h)) \wedge (\neg t \to (x \leftrightarrow \neg h))$$

$$f(x) = t(x\bar{h} + \bar{x}h) + \bar{t}(xh + \bar{x}\bar{h}) = 0 \iff x = th + \bar{t}\bar{h}$$

t	h	x	$x \leftrightarrow h$	$x \leftrightarrow \neg h$	$t \to (x \leftrightarrow h)$	$\neg t \to (x \leftrightarrow \neg h)$	A	守卫
0	0	$\nu_1(x) = 1$	0/1	1	1	1	1	N
0	1	$\nu_2(x) = 0$	0/1	1	1	1	1	Y
1	0	$\nu_3(x) = 0$	1	0/1	1	1	1	N
1	1	$\nu_4(x) = 1$	1	0/1	1	1	1	Y

$$x = t \leftrightarrow h$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Meta-Theorems
- Boolean Algebra
- Application

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

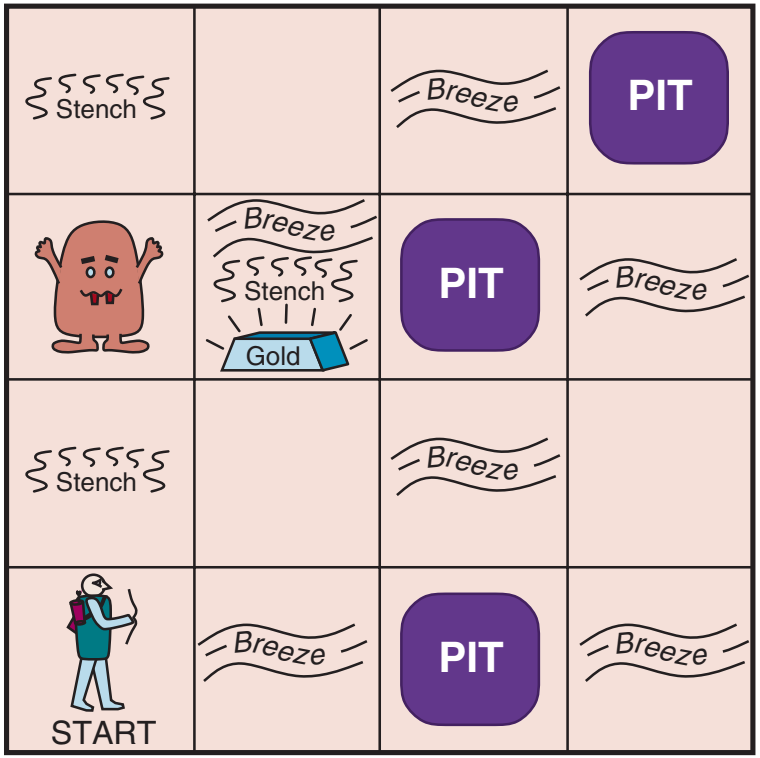
Category Theory

Quantum Computing

Answers to the Exercises

References2039

Wumpus 游戏



- squares adjacent to wumpus are smelly
- squares adjacent to pit are breezy
- glitter iff gold is in the same square
- shooting kills wumpus if you are facing it
- shooting uses up the only arrow
- grabbing picks up gold if in same square
- releasing drops the gold in same square

KB = wumpus-world rules + observations

Example: $B_{21} \leftrightarrow P_{11} \vee P_{22} \vee P_{31}$

Automated Theorem Prover: Prover9 is an automated theorem prover for first-order and equational logic, and Mace4 searches for finite models and counter-examples.[8] Vampire is more powerful.

[8] Adrian Groza: Modelling Puzzles in First Order Logic.

数独游戏

	8	6				2	9	
4			1		5			8
7				9				4
1								9
	5						1	
		8				3		
			5		9			
				2				

$p(i, j, n)$:= the cell in row i and column j contains the number n

- Every row/column contains every number.

$$\bigwedge_{i=1}^{9}\bigwedge_{n=1}^{9}\bigvee_{j=1}^{9} p(i,j,n) \qquad \bigwedge_{j=1}^{9}\bigwedge_{n=1}^{9}\bigvee_{i=1}^{9} p(i,j,n)$$

- Every 3×3 block contains every number.

$$\bigwedge_{r=0}^{2}\bigwedge_{s=0}^{2}\bigwedge_{n=1}^{9}\bigvee_{i=1}^{3}\bigvee_{j=1}^{3} p(3r+i, 3s+j, n)$$

- No cell contains more than one number. for all $1 \le i, j, n, n' \le 9$ and $n \ne n'$:

$$p(i,j,n) \rightarrow \neg p(i,j,n')$$

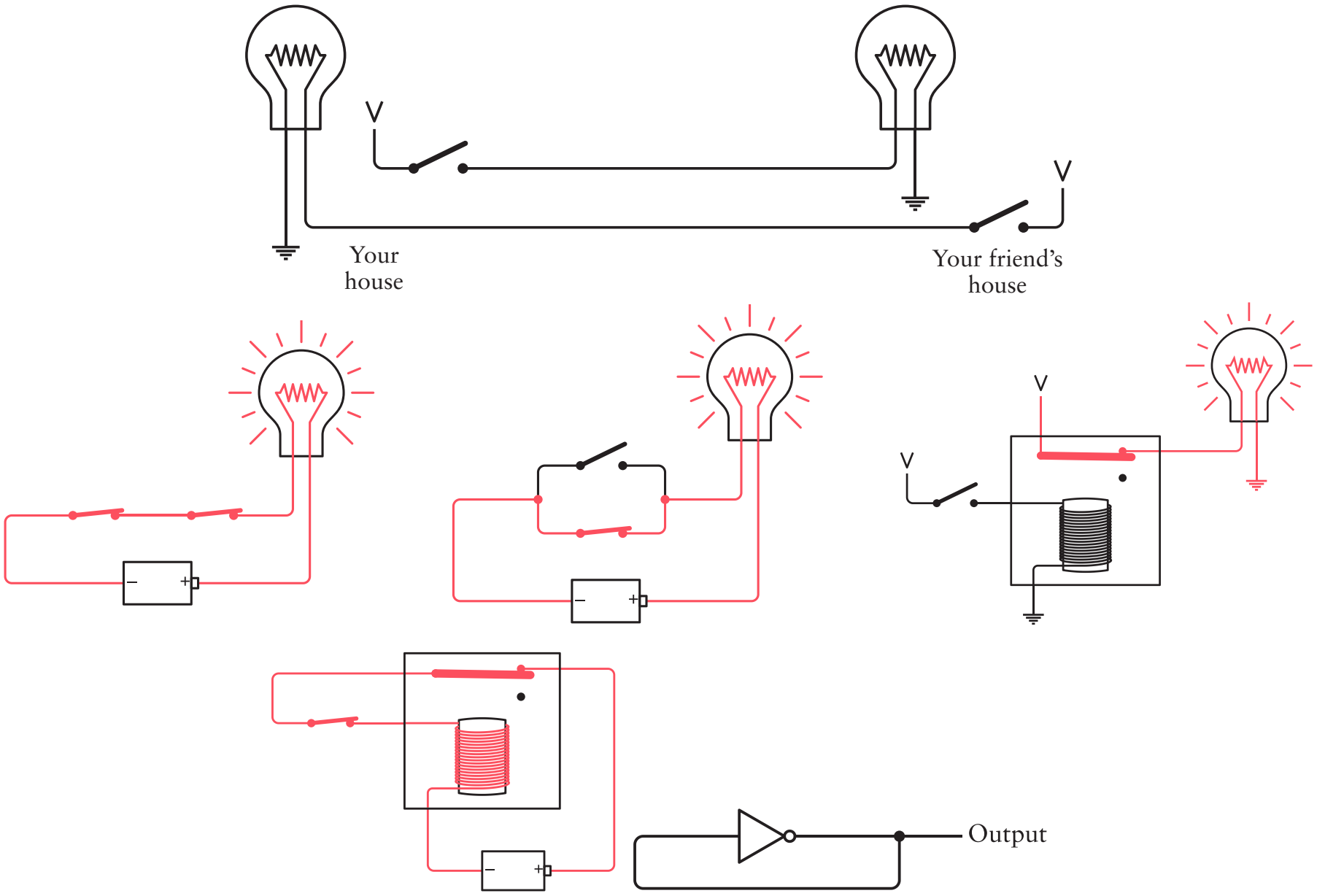
- 一个数独可以看作一个理论.
- 如果有格子填不了任何数字, 则 "公理" 不一致.
- 如果有格子有两种或两种以上的方式可以填数字, 则 "公理" 不完备.

化学反应

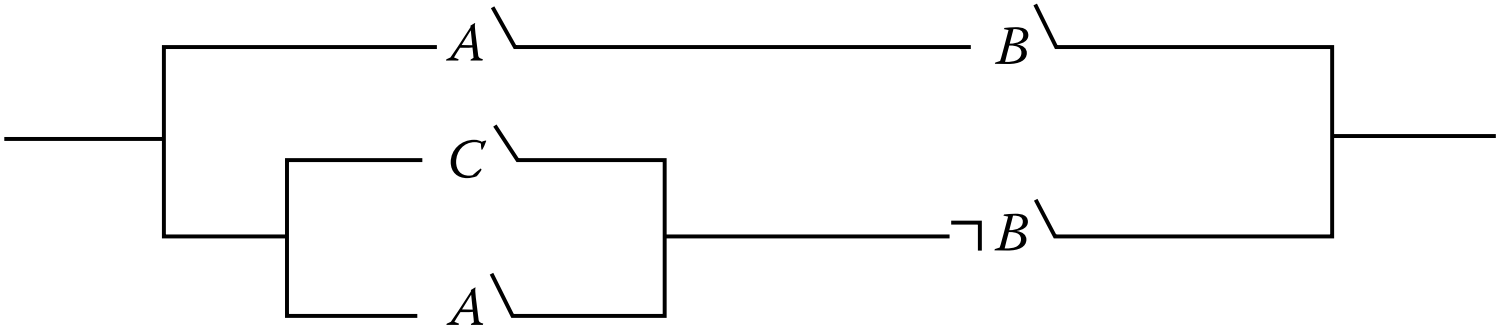
$$\frac{\begin{array}{c}\mathrm{C} \wedge \mathrm{O_2} \rightarrow \mathrm{CO_2} \\ \mathrm{H_2O} \wedge \mathrm{CO_2} \rightarrow \mathrm{H_2CO_3} \\ \mathrm{HCl} \wedge \mathrm{NaOH} \rightarrow \mathrm{NaCl} \wedge \mathrm{H_2O}\end{array}}{\mathrm{HCl} \wedge \mathrm{NaOH} \wedge \mathrm{C} \wedge \mathrm{O_2} \rightarrow \mathrm{NaCl} \wedge \mathrm{H_2CO_3}}$$

$$\dfrac{\dfrac{\underline{\mathrm{HCl}\quad \mathrm{NaOH}}}{\mathrm{NaCl}} \qquad \dfrac{\underline{\mathrm{HCl}\quad \mathrm{NaOH}} \qquad \dfrac{\underline{\mathrm{C}\quad \mathrm{O_2}}}{\mathrm{CO_2}}}{\mathrm{H_2CO_3}}}{\mathrm{NaCl} \wedge \mathrm{H_2CO_3}}$$

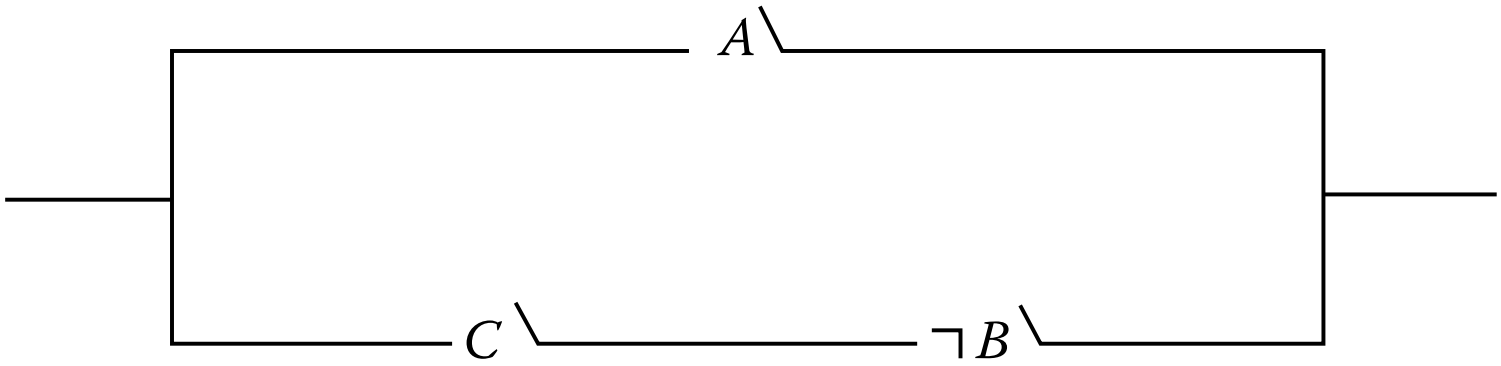
电源、电线、开关、灯泡能用来做什么?



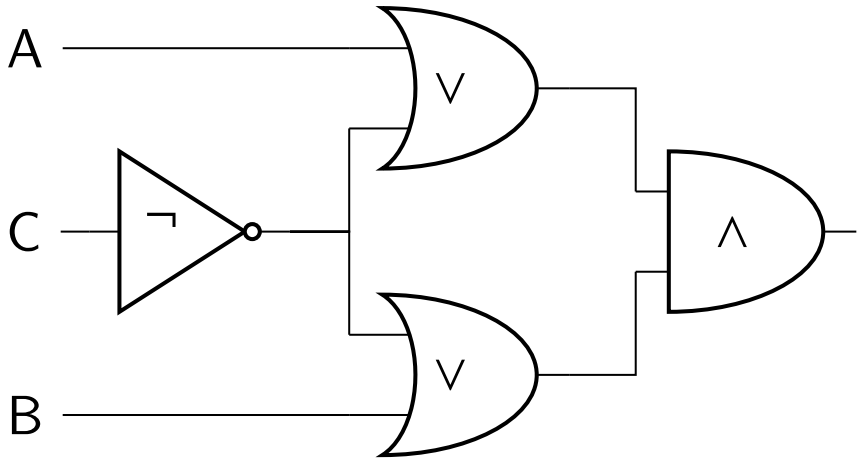
用逻辑做电路优化

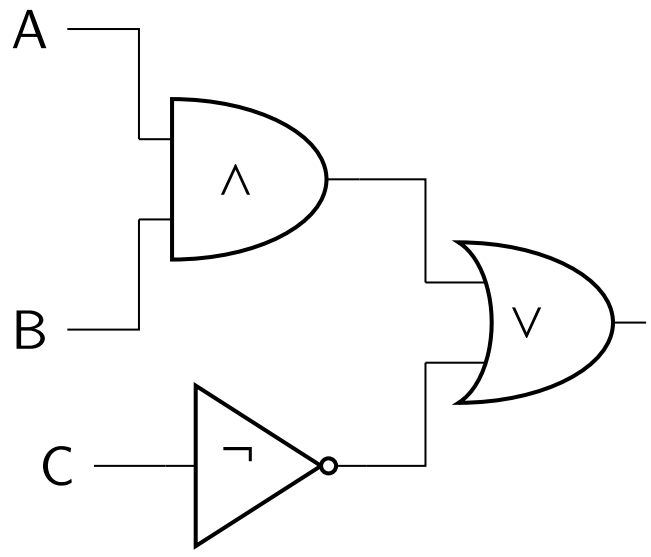


$$\frac{(A \wedge B) \vee ((C \vee A) \wedge \neg B)}{A \vee (C \wedge \neg B)}$$



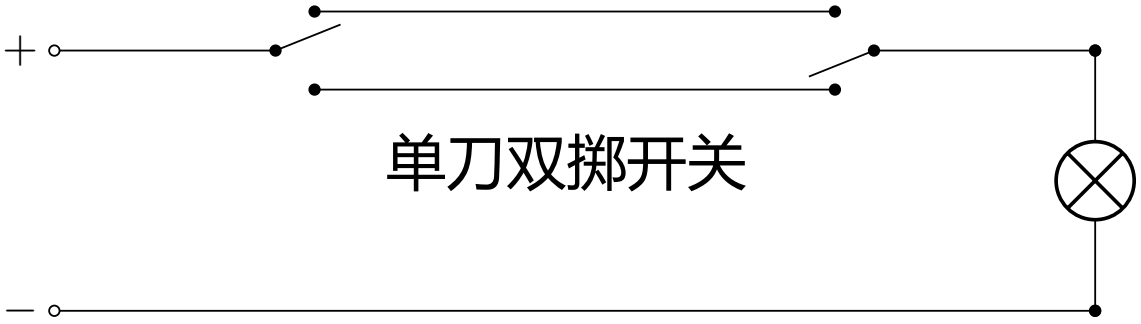
用逻辑做电路优化





$$\frac{(A \vee \neg C) \wedge (B \vee \neg C)}{(A \wedge B) \vee \neg C}$$

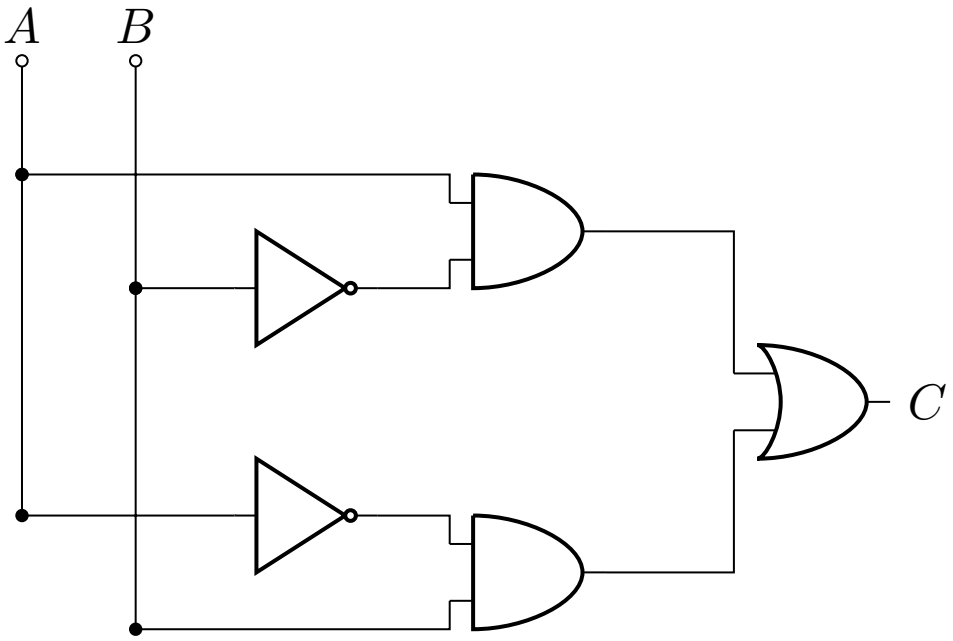
用逻辑做电路设计



单刀双掷开关

- 怎么用两个 “单刀单掷开关” 独立地控制一盏灯?

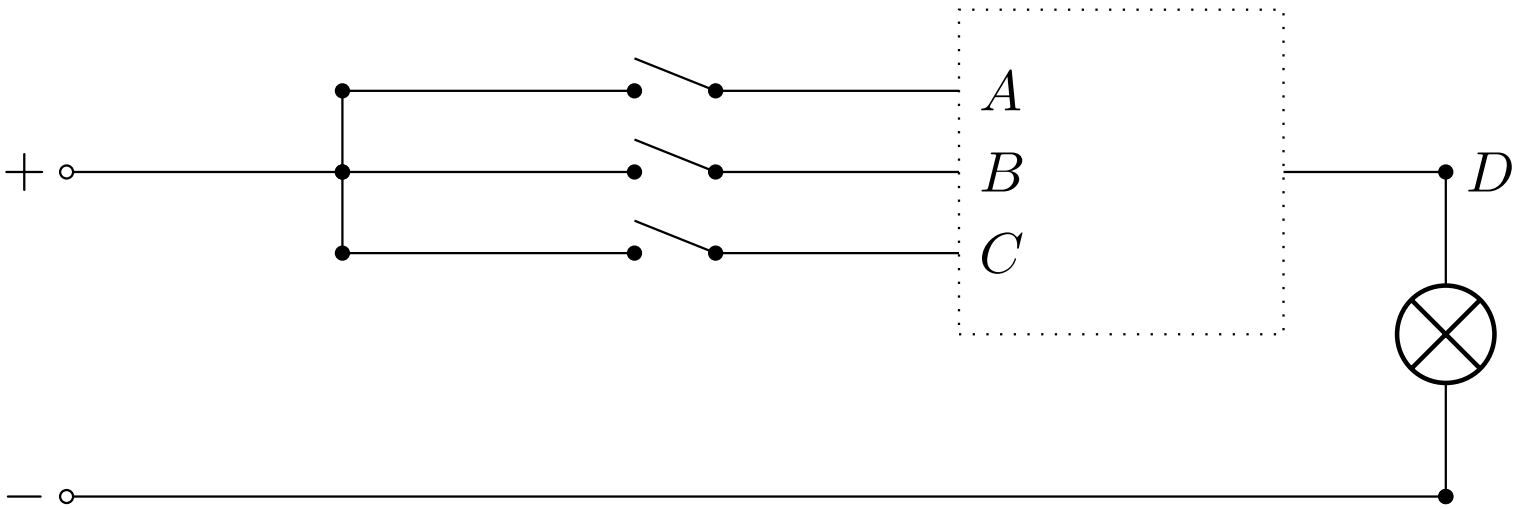
$$(A \wedge \neg B) \vee (\neg A \wedge B)$$



用逻辑做电路设计

- 怎么用三个 “单刀单掷开关” 独立地控制一盏灯?

$$(A \wedge \neg B \wedge \neg C) \vee (\neg A \wedge B \wedge \neg C) \vee (\neg A \wedge \neg B \wedge C) \vee (A \wedge B \wedge C)$$



Problem

设计一个配备三把钥匙的保险箱, 要求至少同时使用其中的两把钥匙才能开启.

用逻辑做算术计算

- 二进制自然数:
  $0, 1, 10, 11, 100, 101, 110, 111, 1000, 1001, 1010, 1011, 1100, \ldots$
- 二进制加法

$$\begin{array}{cccc} & 1 & 0 & 1 \\ & \bullet & 1\bullet & 1 \\ \hline 1 & 0 & 0 & 0 \end{array}$$

- 对于某一位上的加法, 我们给定的输入是被加数的值 a, 加数的值 b, 以及上一位的进位 c.
  1. 这一位需要进位当且仅当 a, b, c 中至少有两个 1

$$(a \wedge b) \vee (b \wedge c) \vee (a \wedge c)$$

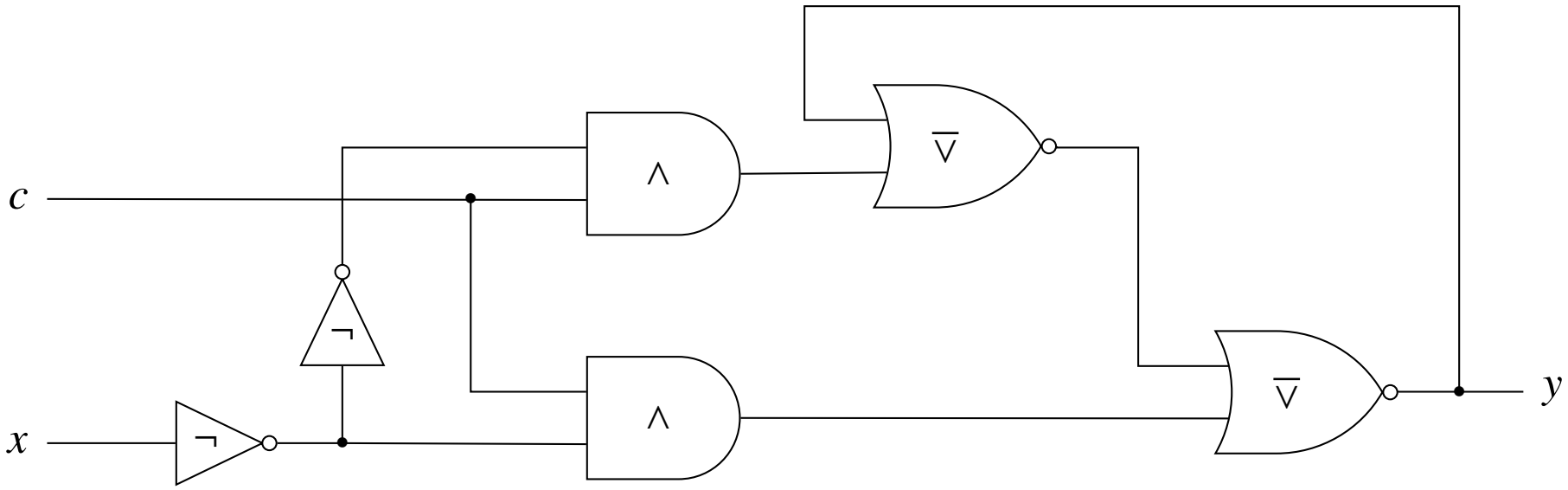
  2. 这一位实际输出为 1 当且仅当 a, b, c 中恰好有一个 1 或三个 1

$$a \oplus b \oplus c$$

Remark: 乘法呢?

- 整数因式分解问题可以归约到 SAT 问题.
- 整数因式分解是 RSA 密码的基础. 若 P = NP, 则 RSA 将失效.

用逻辑做记忆存储



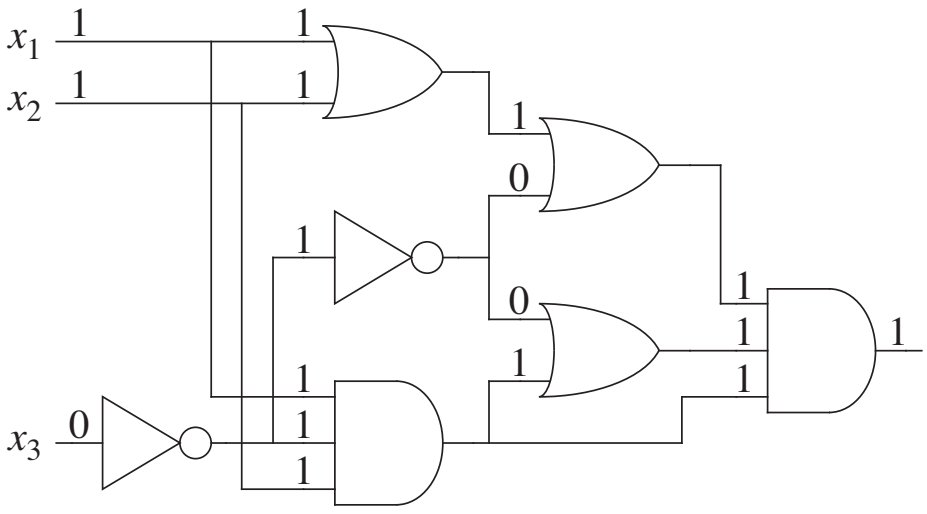
$$y \equiv (c \wedge \neg x) \,\overline{\vee}\, \left[(c \wedge x) \,\overline{\vee}\, y\right]$$

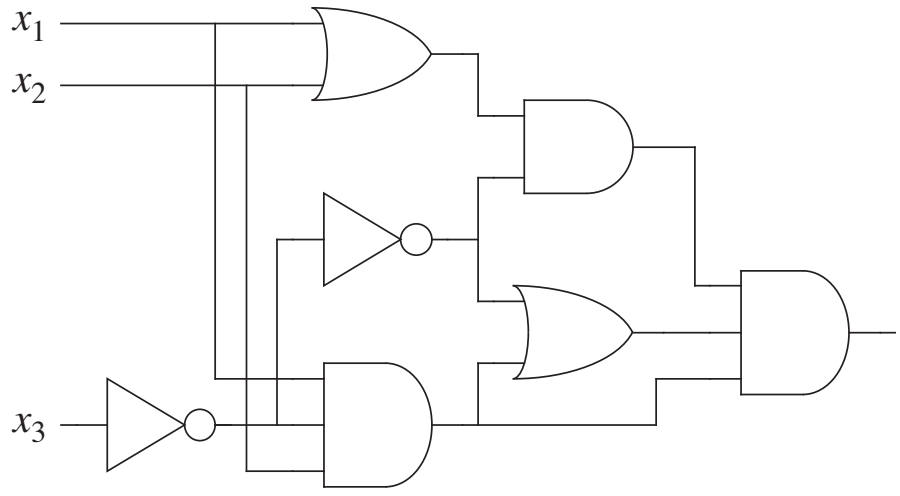
$$y = \overline{c\bar{x} + \overline{cx + y}}$$

$$y = \begin{cases} x & \text{if } c = 1 \\ y & \text{if } c = 0 \end{cases}$$

Remark: 锁存器. 当 $c = 1$ 时, 数据 x 会被 y 记住, 当 $c = 0$ 时, y 不会被修改.

电路可满足性问题 Circuit-Satisfiability Problem (CSAT)

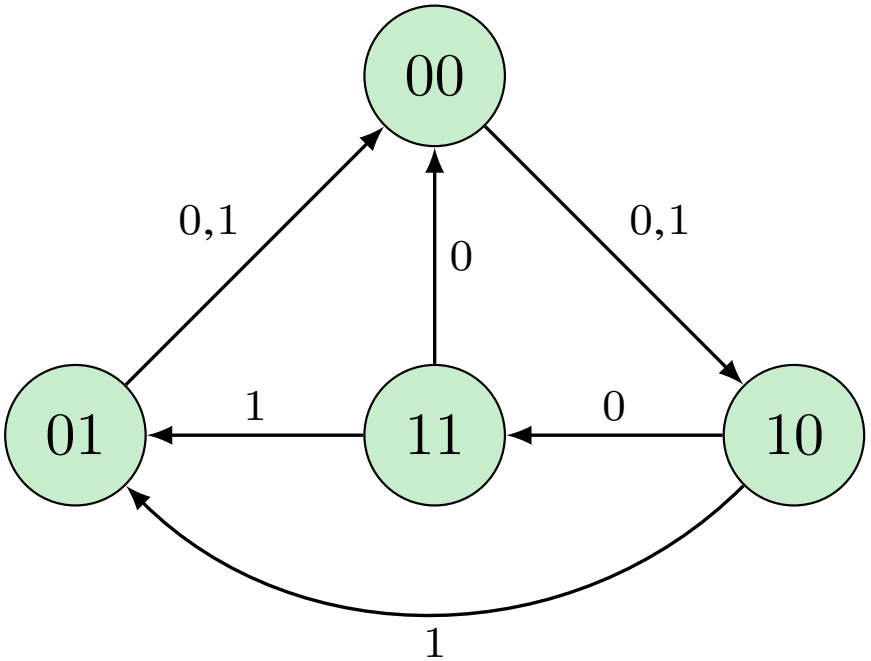




给定一个布尔电路, 是否存在一组输入使得输出为 1?

- CSAT 是 NP 完全的.
- CSAT 可以归于到 SAT; SAT 是 NP 完全的.
- 3SAT 是 NP 完全的. (3SAT 是每个子句只含有 3 个文字的合取范式的满足性问题)

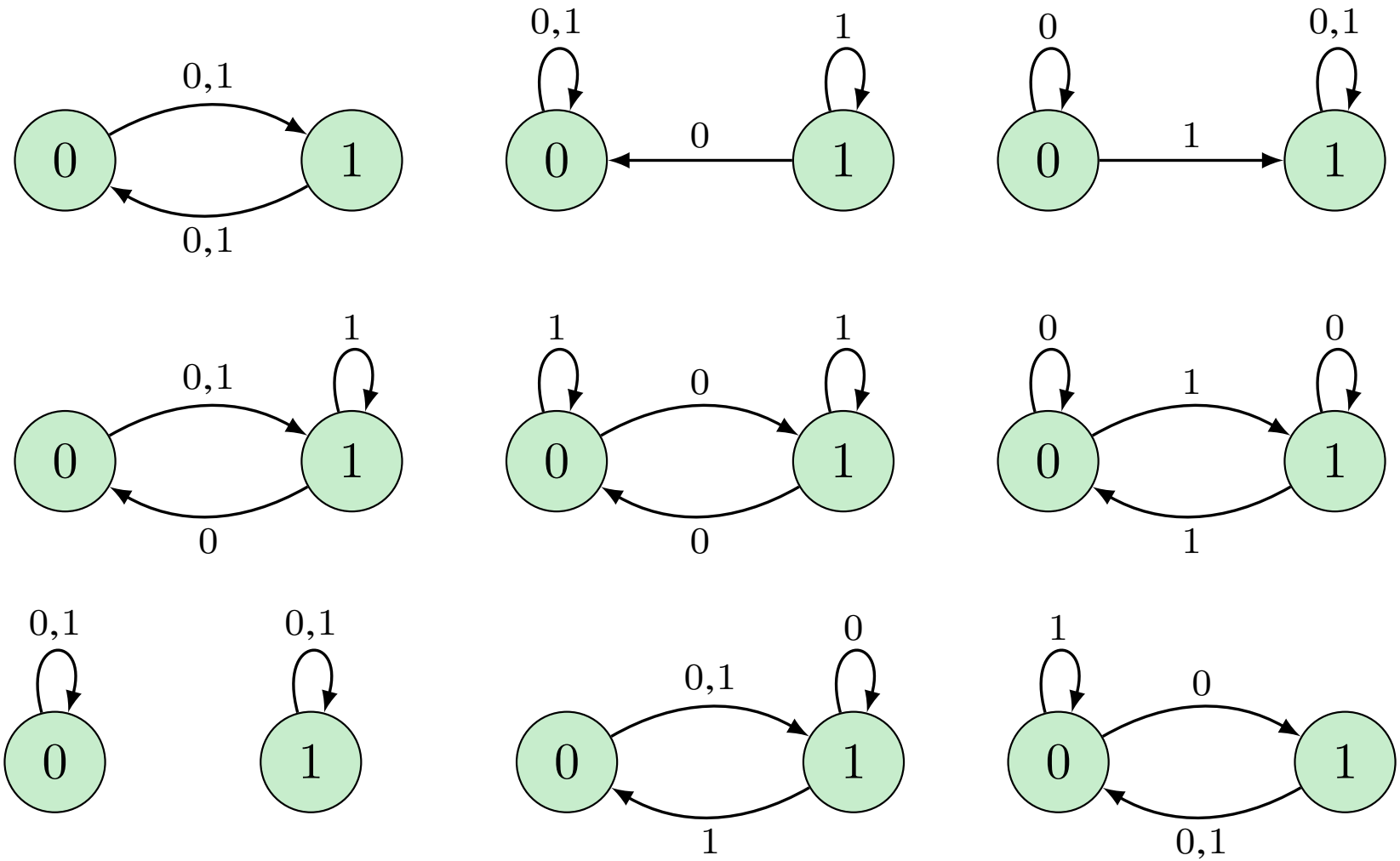
有穷状态自动机



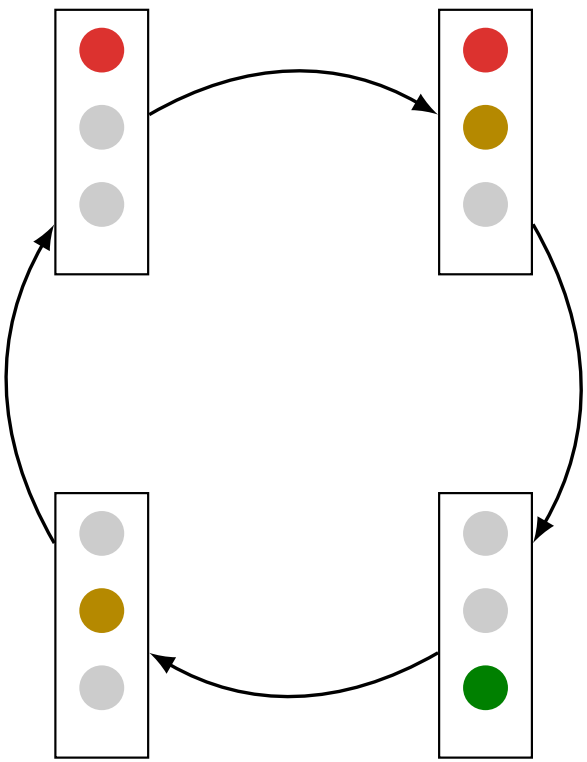
y_1	y_2	x	y_1'	y_2'
0	0	0	1	0
0	0	1	1	0
0	1	0	0	0
0	1	1	0	0
1	0	0	1	1
1	0	1	0	1
1	1	0	0	0
1	1	1	0	1

$$y_1' = \overline{y}_1\overline{y}_2 + \overline{y}_2\overline{x}$$
$$y_2' = y_1\overline{y}_2 + y_1 x$$

$\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \oplus, \bot, \top, \overline{\wedge}, \overline{\vee}$



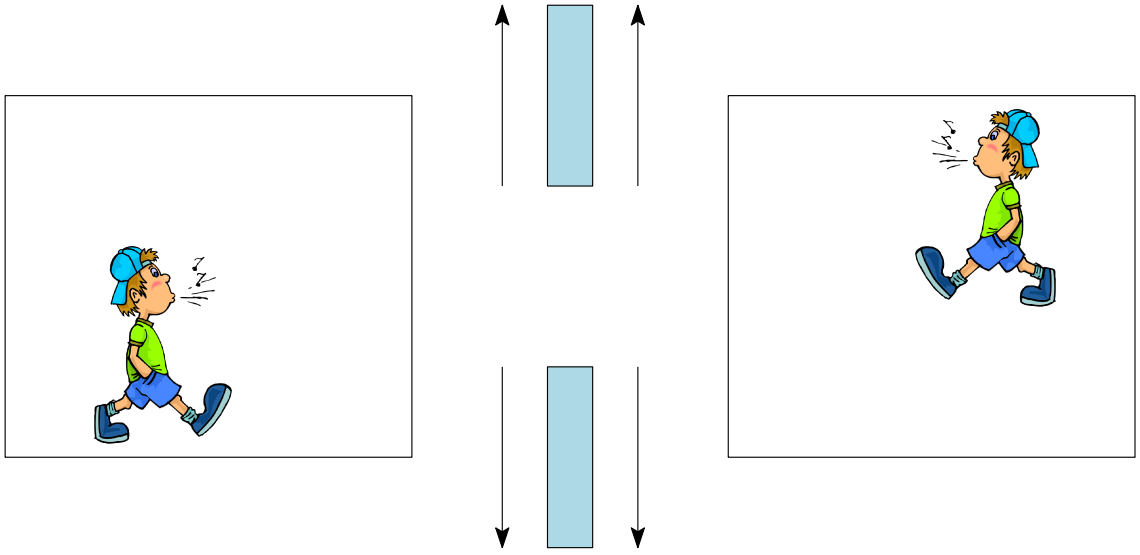
红绿灯



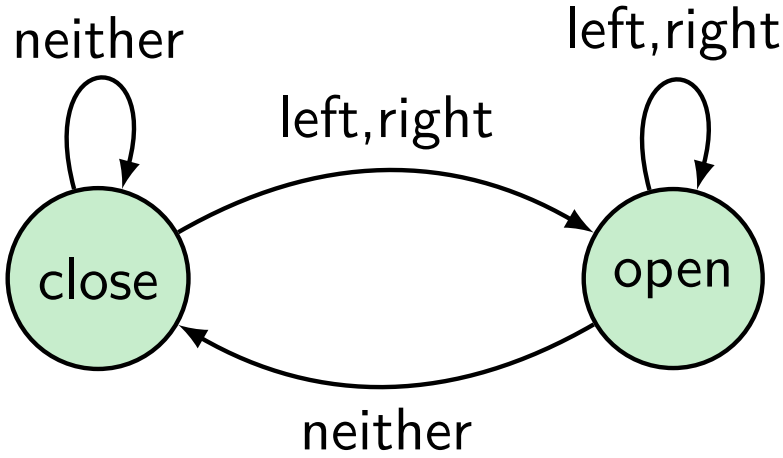
r	a	g	r'	a'	g'
1	0	0	1	1	0
1	1	0	0	0	1
0	0	1	0	1	0
0	1	0	1	0	0

$$r' = r \oplus a$$
$$a' = \overline{a}$$
$$g' = ra$$

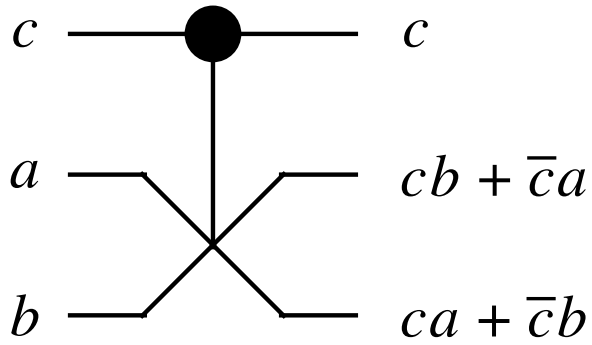
自动门

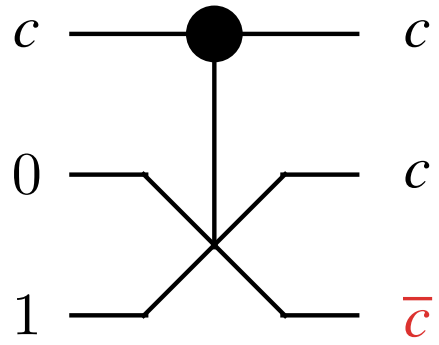


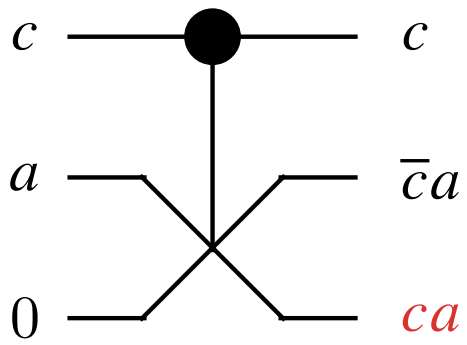
- States = {close, open}
- Input = {left, right, neither}



可逆计算 — Fredkin Gate: CSWAP







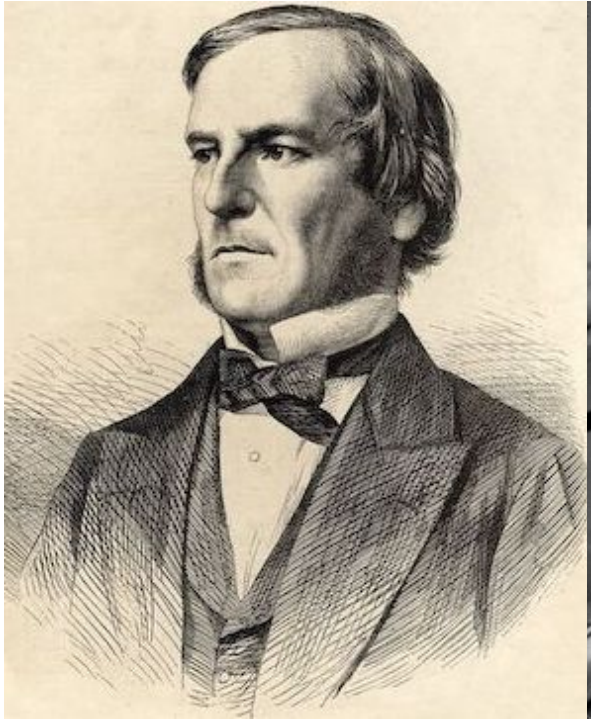
控制位 c 保持不变; 当且仅当 $c = 1$ 时, a 和 b 交换.

c	a	b	c	x	y
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	1	0
1	1	0	1	0	1
1	1	1	1	1	1

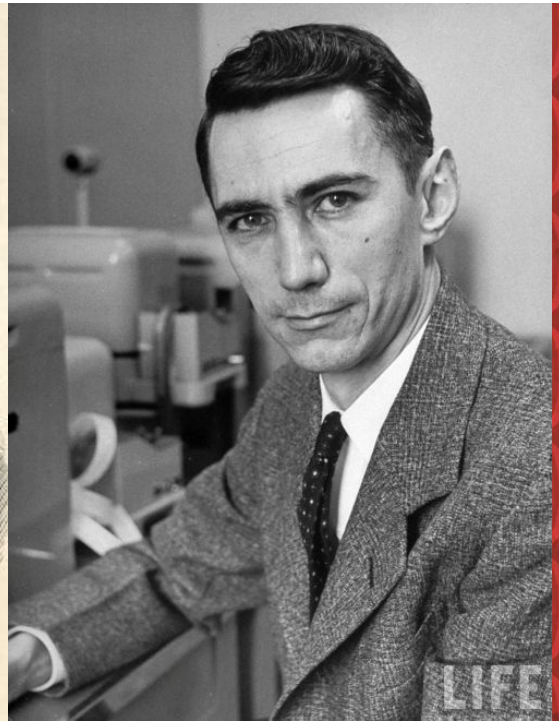
$$f : (c, a, b) \mapsto (c, cb + \bar{c}a, ca + \bar{c}b)$$

- Fredkin 门可模拟任何经典逻辑门.
- 在可逆计算中, 没有信息被擦除, 原则上没有能量耗散, 没有熵增.
- Landauer 原理: 擦除 1 比特信息会向环境中耗散至少 $kT \ln 2$ 的热量.

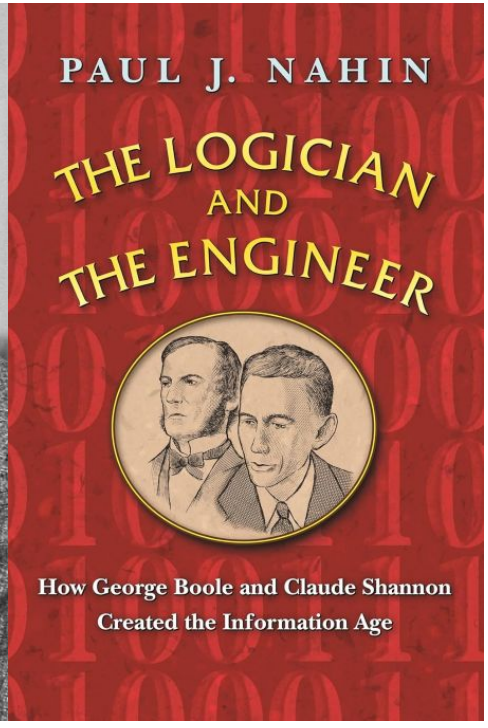
- 真值表 (表示唯一)
- 布尔表达式 (方便变换)
- 电路图 (方便应用)



(a) 布尔

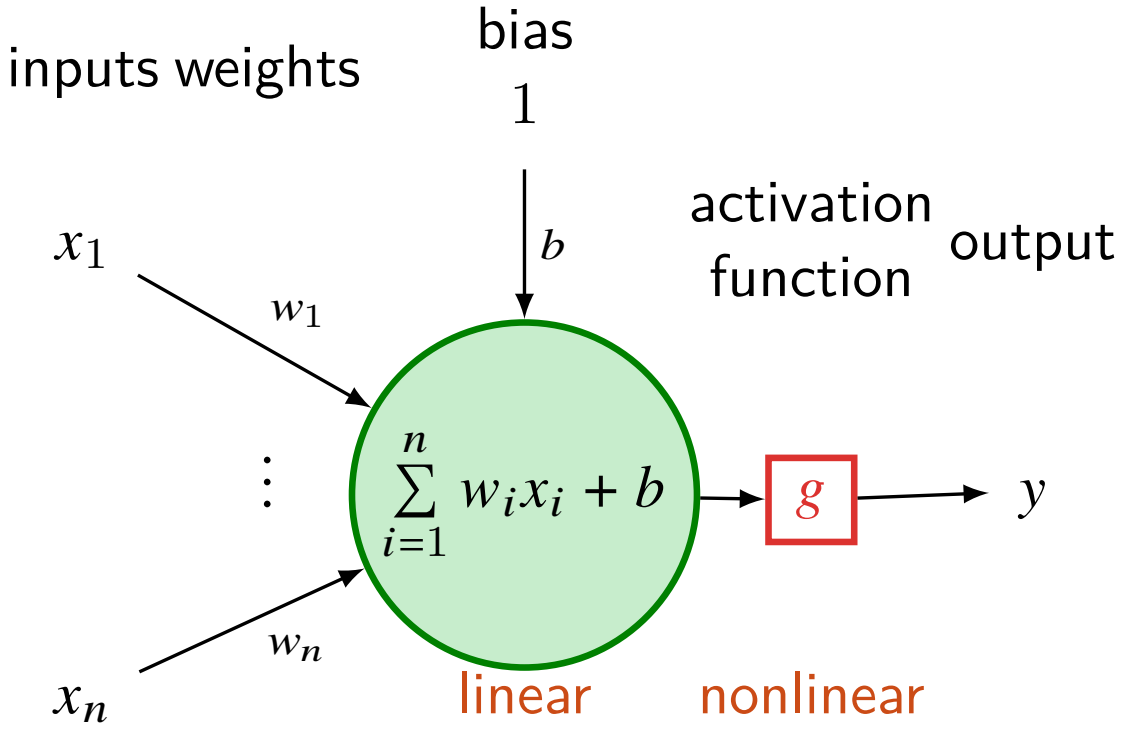


(b) 香农

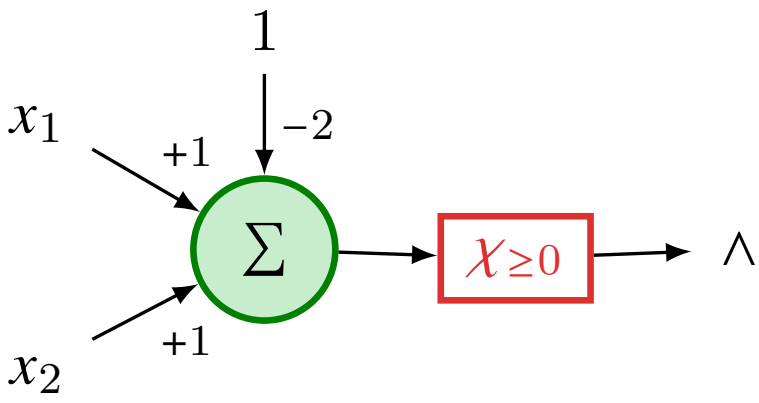


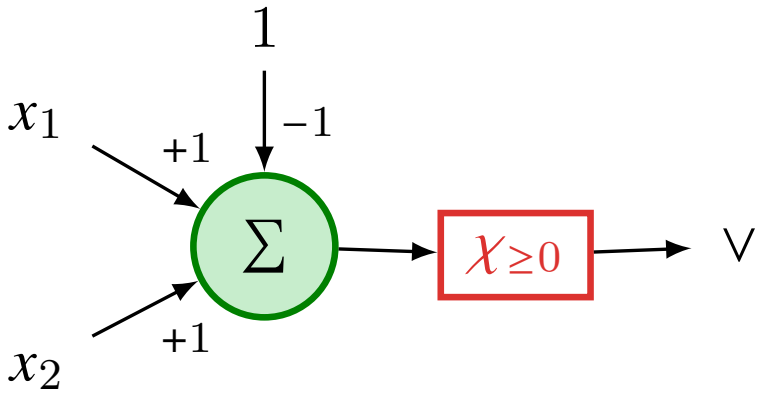
(c) 布尔 & 香农

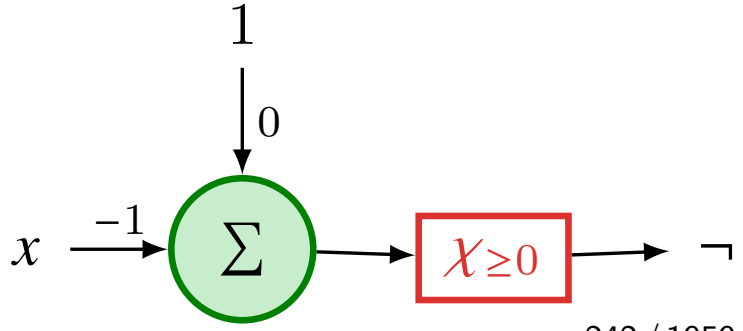
McCulloch-Pitts 人工神经网络 (神经的逻辑演算)

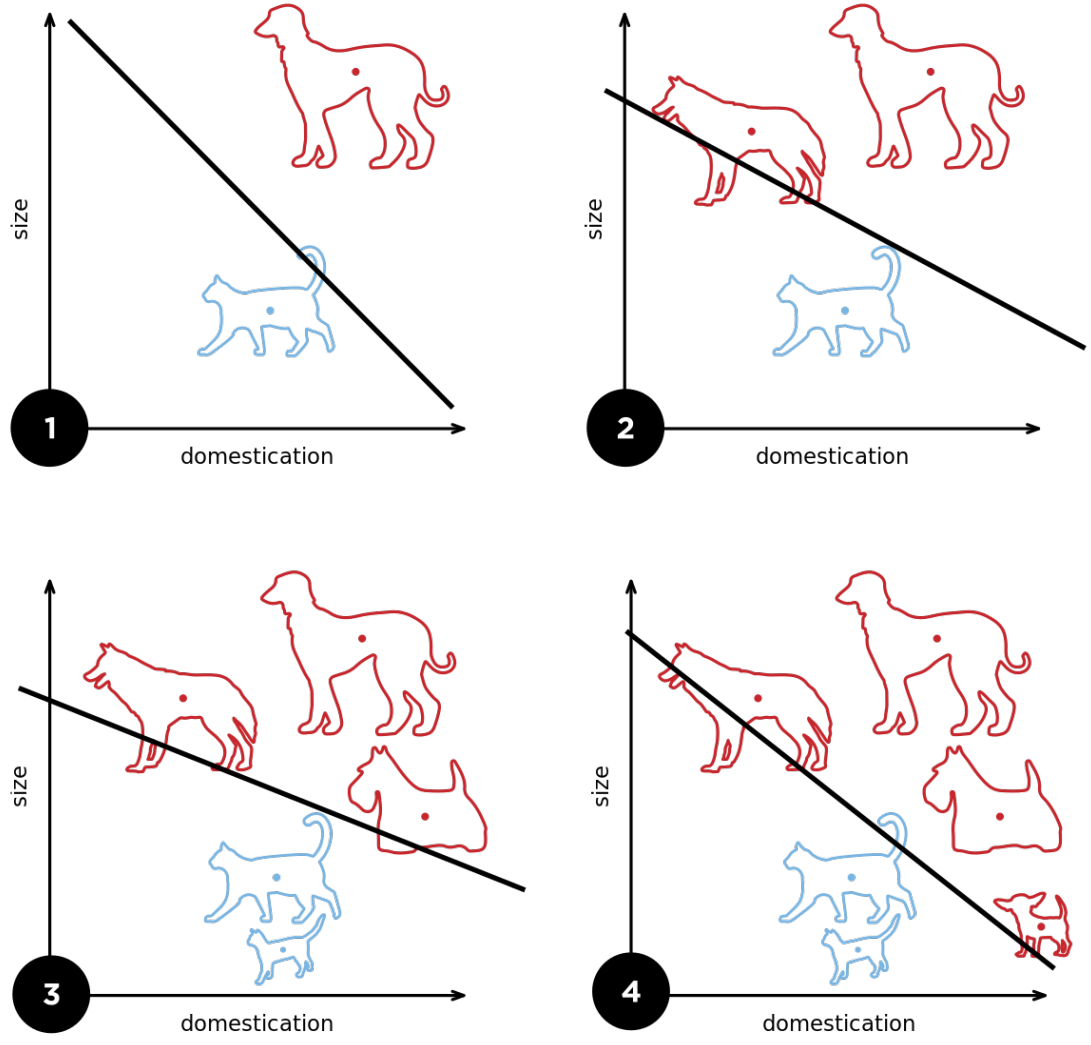


$$y = g\left(\sum_{i=1}^{n} w_i x_i + b\right)$$









1-layer NN $$y = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} w_i x_i + b \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

线性不可分问题

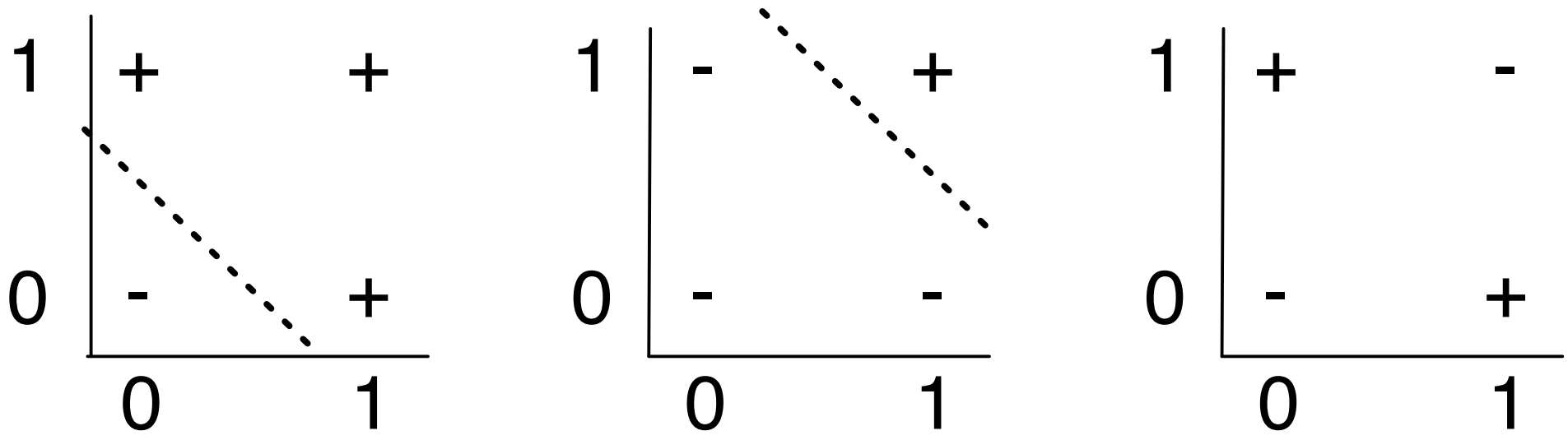


Figure: $\vee, \wedge, \oplus$

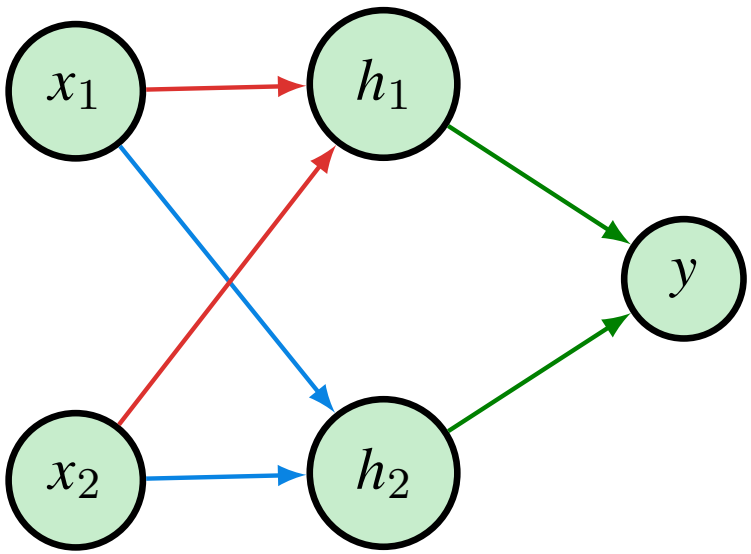
x_1	x_2	$x_1 \oplus x_2$
0	0	0
0	1	1
1	0	1
1	1	0

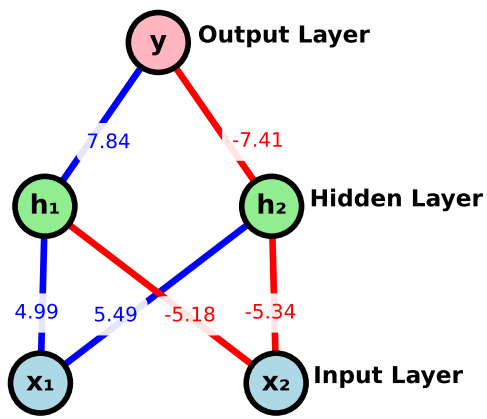
$$w_1 0 + w_2 0 + b < 0 \qquad b < 0$$
$$w_1 0 + w_2 1 + b \geq 0 \qquad w_2 + b \geq 0$$
$$w_1 1 + w_2 0 + b \geq 0 \qquad w_1 + b \geq 0$$
$$w_1 1 + w_2 1 + b < 0 \qquad w_1 + w_2 + b < 0$$

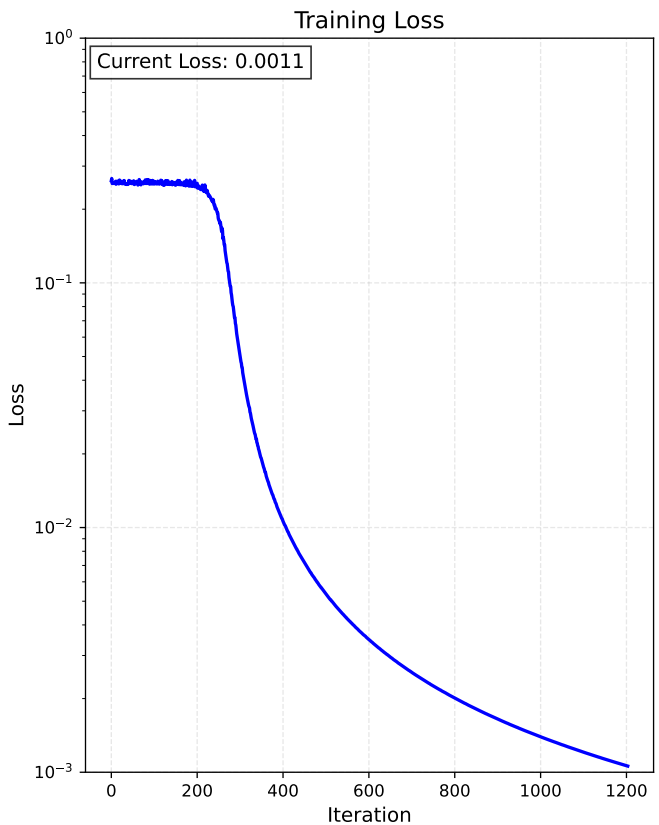
单层感知机无法解决线性不可分问题 (比如异或问题).

异或问题

$$\underbrace{x_1 \oplus x_2}_{y} \equiv \underbrace{(\neg x_1 \wedge x_2)}_{h_1} \vee \underbrace{(x_1 \wedge \neg x_2)}_{h_2}$$







《三体》— 人列计算机[9]

- 秦始皇：朕当然需要预测太阳的运行，但你们让我集结三千万大军，至少要首先向朕演示一下这种计算如何进行吧？
- 冯诺依曼：陛下，请给我三个士兵，我将为您演示．… 我们组建一千万个这样的门部件，再将这些部件组合成一个系统，这个系统就能进行我们所需要的运算，解出那些预测太阳运行的微分方程．



$$F = G\frac{m_1 m_2}{r^2} \longrightarrow a = \frac{F}{m}$$

$$\uparrow \qquad\qquad \downarrow$$

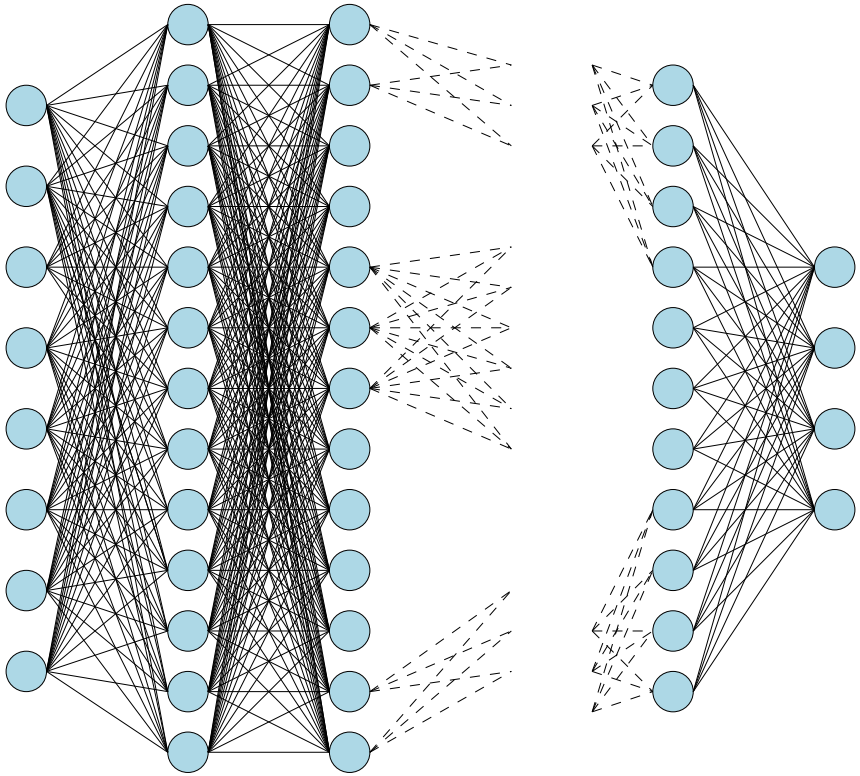
$$\Delta r = v\Delta t \longleftarrow \Delta v = a\Delta t$$

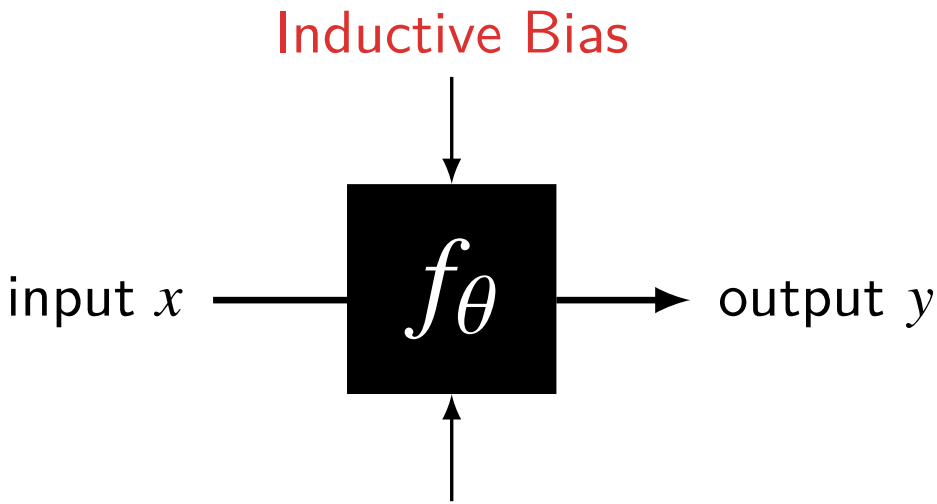
$$\frac{\mathrm{d}^2\mathbf{r_i}}{\mathrm{d}t^2} = -\sum_{j\neq i} Gm_j \frac{\mathbf{r_i} - \mathbf{r_j}}{|\mathbf{r_i} - \mathbf{r_j}|^3} \quad i = 1, 2, 3$$

用连续信号模拟离散信号会怎样？

[9]牛顿解决了二体问题．但发现 $n \geq 3$ 体不是稳定系统．他曾希望求解出太阳系崩溃的时间，如果到时没崩溃，则说明有超自然力量的干预，从而证明上帝存在 (牛顿的上帝改 Bug 显神迹，莱布尼茨的上帝程序完美前定和谐 ☺)．事实上，太阳系是亚稳态系统．

深度神经网络





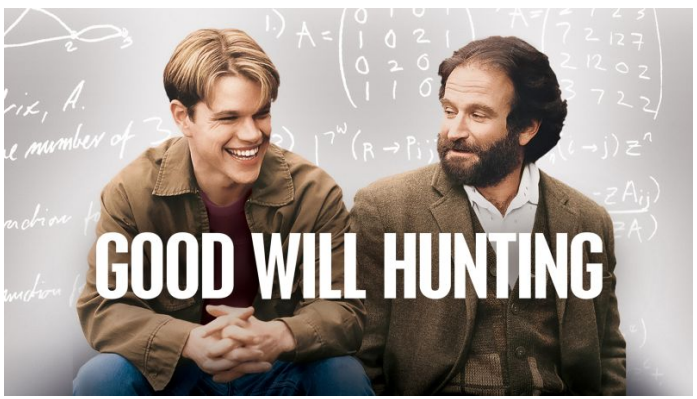




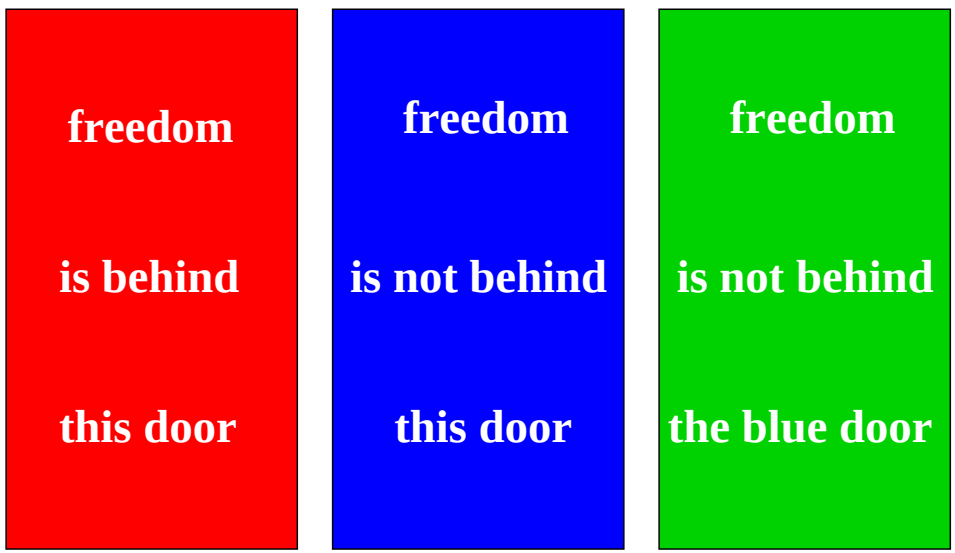


Figure: 电影《心灵捕手》& Walter Pitts & Jeff Hinton

自由之门

Problem (哪扇门通往自由?)

1. 三扇门中只有一扇是自由之门, 另外两扇是死亡之门.
2. 门上的话至少有一句是真的.
3. 门上的话至少有一句是假的.



1. $(r \wedge \neg b \wedge \neg g) \vee (\neg r \wedge b \wedge \neg g) \vee (\neg r \wedge \neg b \wedge g)$
2. $r \vee \neg b \vee \neg b$
3. $\neg r \vee \neg\neg b \vee \neg\neg b$

Solution

r	b	g	1	2	3
0	0	0	0	1	1
0	0	1	1	1	1
0	1	0	1	0	1
0	1	1	0	0	1
1	0	0	1	1	0
1	0	1	0	1	0
1	1	0	0	1	1
1	1	1	0	1	1

r	b	g	
0	0	0	1×
0	0	1	
0	1	0	2×
0	1	1	1×
1	0	0	3×
1	0	1	1×
1	1	0	1×
1	1	1	1×

$$(r \wedge \neg b \wedge \neg g) \vee (\neg r \wedge b \wedge \neg g) \vee (\neg r \wedge \neg b \wedge g), r \vee \neg b, \neg r \vee b \vdash g$$

$$\dfrac{\dfrac{\dfrac{\neg r \vee b}{\neg r \vee b \vee g}}{\neg(r \wedge \neg b \wedge \neg g)} \quad \dfrac{\dfrac{r \vee \neg b}{r \vee \neg b \vee g}}{\neg(\neg r \wedge b \wedge \neg g)} \quad (r \wedge \neg b \wedge \neg g) \vee (\neg r \wedge b \wedge \neg g) \vee (\neg r \wedge \neg b \wedge g)}{\neg r \wedge \neg b \wedge g}$$

$$(r\overline{b}\,\overline{g} + \overline{r}b\overline{g} + \overline{r}\,\overline{b}g)(r + \overline{b})(\overline{r} + b) = \overline{r}\,\overline{b}g$$

Problem (宝藏在哪里?)

你面前有三扇门, 只有一扇门后是宝藏. 门上各有一句话, 只有一扇门上的是真话.

① 宝藏不在这儿.

② 宝藏不在这儿.

③ 宝藏在②号门.

Solution

- ① $\neg t_1$; ② $\neg t_2$; ③ t_2

1. 只有一扇门上的是真话.
   $(\neg t_1 \wedge \neg\neg t_2 \wedge \neg t_2) \vee (\neg\neg t_1 \wedge \neg t_2 \wedge \neg t_2) \vee (\neg\neg t_1 \wedge \neg\neg t_2 \wedge t_2)$
2. 只有一扇门后是宝藏.
   $(t_1 \wedge \neg t_2 \wedge \neg t_3) \vee (\neg t_1 \wedge t_2 \wedge \neg t_3) \vee (\neg t_1 \wedge \neg t_2 \wedge t_3)$

$$\bar{t}_1\bar{t}_2\bar{t}_2 + \bar{\bar{t}}_1\bar{t}_2 + \bar{\bar{t}}_1\bar{\bar{t}}_2 t_2 = 1 \implies t_1 = 1$$

$$t_1\bar{t}_2\bar{t}_3 + \bar{t}_1 t_2\bar{t}_3 + \bar{t}_1\bar{t}_2 t_3 = 1 \implies t_2 = t_3 = 0$$

哪条大路通罗马?

Problem (哪条大路通罗马?)

你面前有左、中、右三条大路. 三条路分别被三个**说谎的**守卫守护着.

1. 左路护卫: "左路通罗马. 如果右路通罗马的话, 中路也通罗马".
2. 中路护卫: "左路、右路都不通罗马."
3. 右路护卫: "左路通罗马, 中路不通."

Solution

1. $\neg(x \wedge (z \to y))$
2. $\neg(\neg x \wedge \neg z)$
3. $\neg(x \wedge \neg y)$

$$\overline{x(\overline{z}+y)}\;\overline{x}\,\overline{z}\;\overline{x\overline{y}} = 1$$

$$x(\overline{z}+y)+\overline{x}\,\overline{z}+x\overline{y} = 0$$

$$\overline{z}+x = 0$$

$$z = 1, x = 0$$

竞赛排名

Problem (每人的名次是多少?)

A, B, C, D 四人参加了一场竞赛.

- A 猜测: C 第 1, B 第 2.
- B 猜测: C 第 2, D 第 3.
- C 猜测: D 第 4, A 第 2.

结果每人都只猜对了一半, 求每人的名次.

Solution

令 X_i 表示 X 是第 i 名.
显然, 当 $i \neq j$ 时, $X_i X_j = 0$. 当 $X \neq Y$ 时, $X_i Y_i = 0$.

$$(C_1\overline{B}_2 + \overline{C}_1 B_2)(C_2\overline{D}_3 + \overline{C}_2 D_3)(D_4\overline{A}_2 + \overline{D}_4 A_2) = 1$$

$$\Downarrow$$

$$C_1\overline{C}_2 A_2 D_3\overline{D}_4\overline{B}_2 = 1$$

重贴标签

Problem (只从一个盒子里取一个水果观察, 能否把标签重贴正确?)

- 三个盒子. 一个只装苹果, 一个只装橙子, 一个两样都装.
- 标签都是错的.

Apple **Both** **Orange**

令 Px 表示标签为 P 的盒子装的是 x.

1. $Ao \vee Ab$
2. $Ba \vee Bo$
3. $Oa \vee Ob$
4. $Px \to \neg Py$ 当 $x \neq y$ 时. 其中 $x \in \{a, b, o\}$.
5. $Px \to \neg Qx$ 当 $P \neq Q$ 时. 其中 $P, Q \in \{A, B, O\}$.

A	B	O	
a	b	o	×
a	o	b	×
b	a	o	×
b	o	a	✓
o	a	b	✓
o	b	a	×

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{Bo \quad Bo \to \neg Ao}{\neg Ao} \qquad Ao \vee Ab}{Ab} \qquad Ab \to \neg Ob}{\neg Ob} \qquad Oa \vee Ob}{Oa}$$

自指问题

Problem (下面哪一项是这个问题的正确答案?)

1. 下面所有项.
2. 下面没有一项.
3. 上面所有项.
4. 上面某一项.
5. 上面没有一项.
6. 上面没有一项.

1. $p_1 = p_2 p_3 p_4 p_5 p_6$
2. $p_2 = \overline{p}_3 \overline{p}_4 \overline{p}_5 \overline{p}_6$
3. $p_3 = p_1 p_2$
4. $p_4 = p_1 \overline{p}_2 \overline{p}_3 + \overline{p}_1 p_2 \overline{p}_3 + \overline{p}_1 \overline{p}_2 p_3$
5. $p_5 = \overline{p}_1 \overline{p}_2 \overline{p}_3 \overline{p}_4$
6. $p_6 = \overline{p}_1 \overline{p}_2 \overline{p}_3 \overline{p}_4 \overline{p}_5$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{2 \quad 3}{p_3 = 0} \quad 1}{p_1 = 0} \quad 4}{p_4 = p_2} \quad 2}{p_2 = p_4 = 0} \quad 5}{\dfrac{p_5 = 1 \quad 6}{p_6 = 0}}$$

君子/小人

Problem (谁是君子? 谁是小人?)

一个岛上有“君子”、“小人”两类人. “君子”只说真话, “小人”只说假话. 你来岛上遇到了小艾、小白、小菜三个土著.

1. 小艾: “如果小菜说谎, 我或小白说的就是真话”.
2. 小白: “如果小艾或小菜说真话, 那么, 我们三人中有且只有一人说真话是不可能的”.
3. 小菜: “小艾或小白说谎当且仅当小艾或我说真话”.

1. $a = c + a + \underline{b}$
2. $b = \overline{a + c} + a\overline{b}\,\overline{c} + \overline{a}b\overline{c} + \overline{a}\,\overline{b}c$
3. $c = (\overline{a} + \overline{b})(a + c) + ab\,\overline{a + c}$

$$c = (\overline{a} + \overline{b})(a + c) + ab\overline{a}\,\overline{c} = (\overline{a} + \overline{b})(a + c) + 0 = (\overline{a} + \overline{b})a = a\overline{b}$$

$$b = \overline{a + c} + \overline{c\overline{c} + \overline{c + a + b}\; b\overline{c} + \overline{c + a + b}\; \overline{b}c} = \overline{a + c} + \overline{0 + 0 + 0} = 1$$

$$a = c + a + 1 = 1$$

$$c = a\overline{b} = 0$$

君子/小人/凡夫

Problem (证明两人中必有讲真话的凡夫)

一个岛上有三类人. 君子只说真话, 小人只说假话, 凡夫有时真有时假.

- A 说: B 不是君子.
- B 说: A 是君子.

A, B 中必有凡夫, 否则矛盾.

$$\frac{A \leftrightarrow \neg B \quad B \leftrightarrow A}{\bot}$$

易证 B 不是君子.
因此 A 讲的是真话.
由此, A 或是君子, 或是讲真话的凡夫.
若 A 是君子, 则 B 是讲真话的凡夫.

练习 — Now it's your turn

Problem (我在做啥?)

1. 如果我不在打网球, 那就在看网球.
2. 如果我不在看网球, 那就在读网球杂志.
3. 但我不能同时做两件或两件以上的事.

Problem (你想下厨为全家准备晚餐. 做点儿什么呢?)

1. 香菇青菜、红烧鱼、千页豆腐三个菜里你希望至少做出两个来.
2. 爸爸红烧鱼和千页豆腐两样中必须吃一样, 但不同时吃.
3. 妈妈除非同时吃香菇青菜, 否则不吃千页豆腐, 而且, 只要吃红烧鱼就不吃香菇青菜.

Problem (A, B, C, D 四个嫌疑人, 谁有罪?)

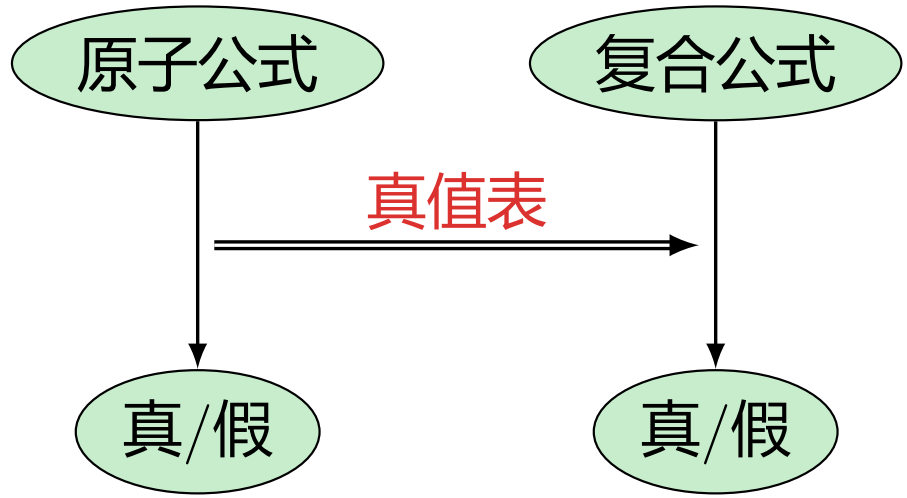
1. 如果 A 有罪, 那么 B 也有罪.
2. 如果 B 有罪, 那么或者 A 无罪或者 C 有罪.
3. 如果 D 有罪, 那么 A 也有罪.
4. 如果 D 无罪, 那么 A 有罪而 C 无罪.

总结

- 语法



- 语义



- 形式系统



- 表达力 / 简洁性
- 可满足性 / 有效性
- 可靠性 / 完备性 / 紧致性
- 可判定性 / 计算复杂性

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Meta-Theorems

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

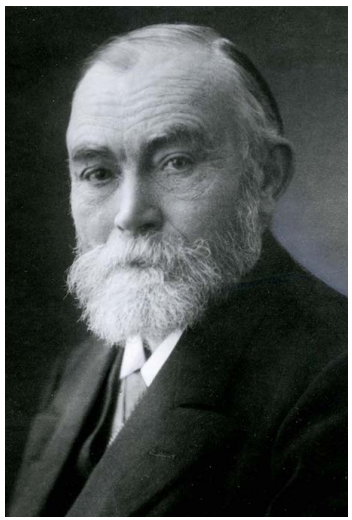
弗雷格 Gottlob Frege 1848-1925

皮尔士 Charles Peirce 1839-1914

- 《概念文字: 一种模仿算术语言构造的纯思维的形式语言》*1879.*
- 逻辑主义 数学可还原为逻辑.[a]
- 谓词逻辑之父
  (关系 & 量词)
  (Every boy loves some girl.)

$$\frac{\text{subject}}{\text{predicate}} \approx \frac{\text{argument}}{\text{function}}$$

- 语言哲学
  The evening star is the morning star.[b]





[a] Frege: The Foundations of Arithmetic. 1884.

[b] Frege: On Sense and Reference. 1892.

Why Study Predicate Logic?

- 命题逻辑预设世界由事实构成
- 谓词逻辑预设世界包含
  1. 个体: 人、狗、书、自然数、实数、城市、国家 ...
  2. 关系: 红的、圆的、大于、爱上、父子、朋友、老师 ...
  3. 函数: 平方、加法、母亲、老婆、最好的朋友、导师 ...
- 谓词逻辑的表达力更强

$$\frac{\text{Father}(\text{Father}(\text{alice})) = \text{Father}(\text{Mother}(\text{bob}))}{\text{Cousin}(\text{alice}, \text{bob})}$$

语言	本体论承诺	认识论承诺
Propositional Logic	facts	true/false/unknown
Predicate Logic	facts, objects, relations	true/false/unknown
Temporal Logic	facts, objects, relations, times	true/false/unknown
Probability Theory	facts	degree of belief $[0, 1]$
Fuzzy Logic	facts with degree of truth $[0, 1]$	known interval value

Example ☹

What will a logician choose: philosophy or money?

Philosophy! Because Nothing is better than Money, and Philosophy is better than Nothing.

$$p > 0 > m \implies p > m$$
$$\neg\exists x(x > m) \implies 0 \ngtr m$$

"The Nothing itself nothings."

— *Heidegger*

A cat has nine lives ☹

No cat has eight lives. A cat has one more life than no cat.

爱丽丝镜中奇遇 — Lewis Carroll

- "I see nobody on the road." said Alice.
- "I only wish I had such eyes," the King remarked in a fretful tone. "To be able to see Nobody! And at that distance too!"

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Meta-Theorems

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

什么是 “量词”?

- All students work hard.
- Some students are asleep.
- At least six students are awake.
- Exactly five students pass the exam.
- Eight out of ten students are good at logic.
- Nobody knows logic better than Donald Trump.
- The present King is smart.
- Most/Many/Few students love logic.
- There are finitely/infinitely/countably/uncountably many elements such that...
- For all but finitely many elements...
- There are more enemies than friends.
- **Some** girls admire only one another.

语法 ♡

Language

$$\mathscr{L}^1 := \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \forall, \exists, =, (,)\} \cup \mathrm{Var} \cup \overbrace{\mathrm{Cst} \cup \mathrm{Fun} \cup \mathrm{Pred}}^{\text{signature}}$$

where

$$\mathrm{Var} := \{x_i : i \in \mathbb{N}\}$$

$$\mathrm{Cst} := \{c_i : i \in \mathbb{N}\}$$

$$\mathrm{Fun} := \bigcup_{n \in \mathbb{N}} \mathrm{Fun}^n \quad \mathrm{Fun}^n := \left\{f_1^n, f_2^n, f_3^n, \dots\right\}$$

$$\mathrm{Pred} := \bigcup_{n \in \mathbb{N}} \mathrm{Pred}^n \quad \mathrm{Pred}^n := \left\{P_1^n, P_2^n, P_3^n, \dots\right\}$$

- c is a constant symbol.
- f^n is an n-place function symbol.
- P^n is an n-place predicate symbol.

项 & 公式 ♡

Definition (Term)

$$t \coloneqq x \mid c \mid f(t, \ldots, t)$$

where $x \in \mathrm{Var}$, $c \in \mathrm{Cst}$ and $f \in \mathrm{Fun}$.

- Term is freely generated from Var by Fun.

Definition (Well-Formed Formula Wff)

$$A \coloneqq \overbrace{t = t \mid P(t, \ldots, t)}^{\text{atomic formula}} \mid \neg A \mid A \wedge A \mid A \vee A \mid A \to A \mid A \leftrightarrow A \mid \forall x A \mid \exists x A$$

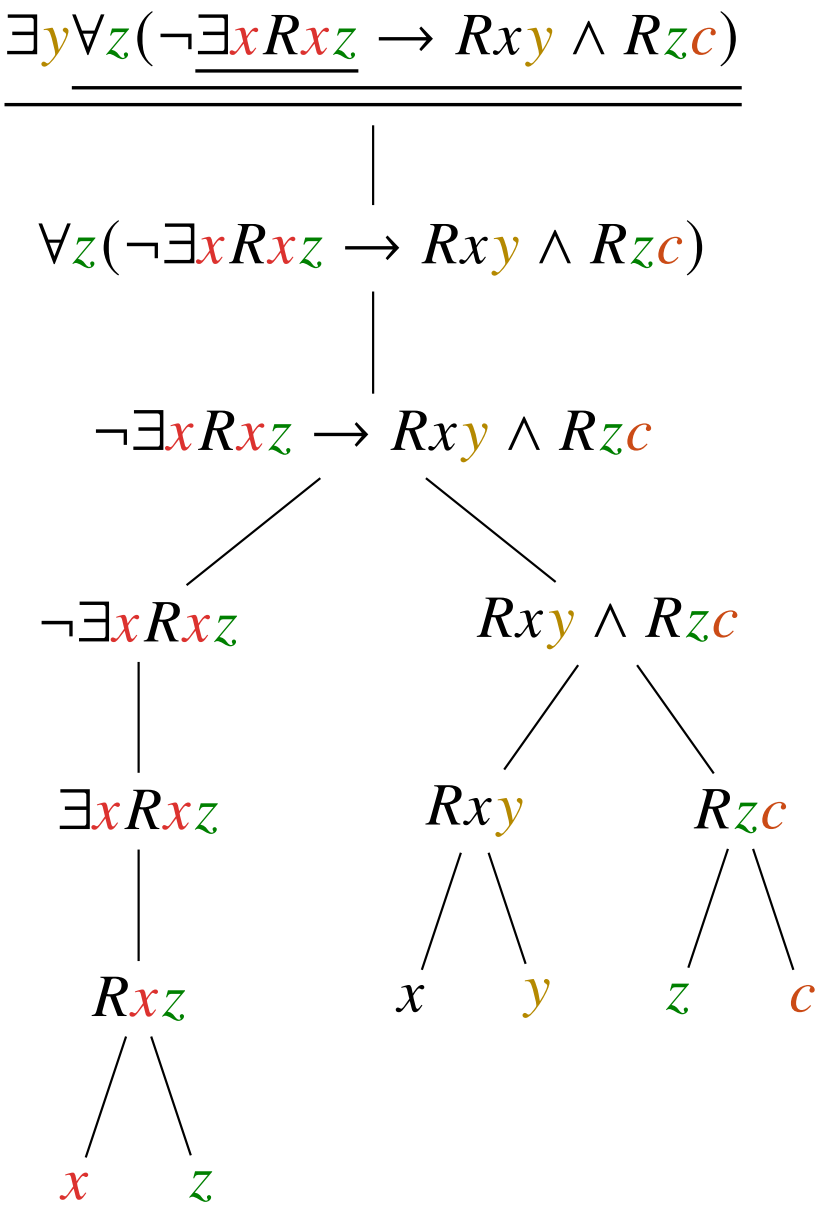
where $t \in \mathrm{Term}$ and $P \in \mathrm{Pred}$.

- Wff is freely generated from atomic formulas by connectives and quantifier operators.

语法 ♡

- $A \wedge B := \neg(A \to \neg B)$
- $A \vee B := \neg A \to B$
- $A \leftrightarrow B := (A \to B) \wedge (B \to A)$
- $\exists x A := \neg \forall x \neg A$
- $\bot := A \wedge \neg A$
- $\top := \neg \bot$
- Bottom up and Top down definitions of terms, subterms, formulas and subformulas.
- Induction Principle for terms and formulas.
- Unique readability theorem for terms and formulas.
- 省略括号的约定:
  1. 公式最外层的括号可以省略
  2. $\neg, \forall, \exists, \wedge, \vee, \to, \leftrightarrow$ 的组合强度依次减弱
  3. 同一连接词相邻出现时, 右边的组合力更强

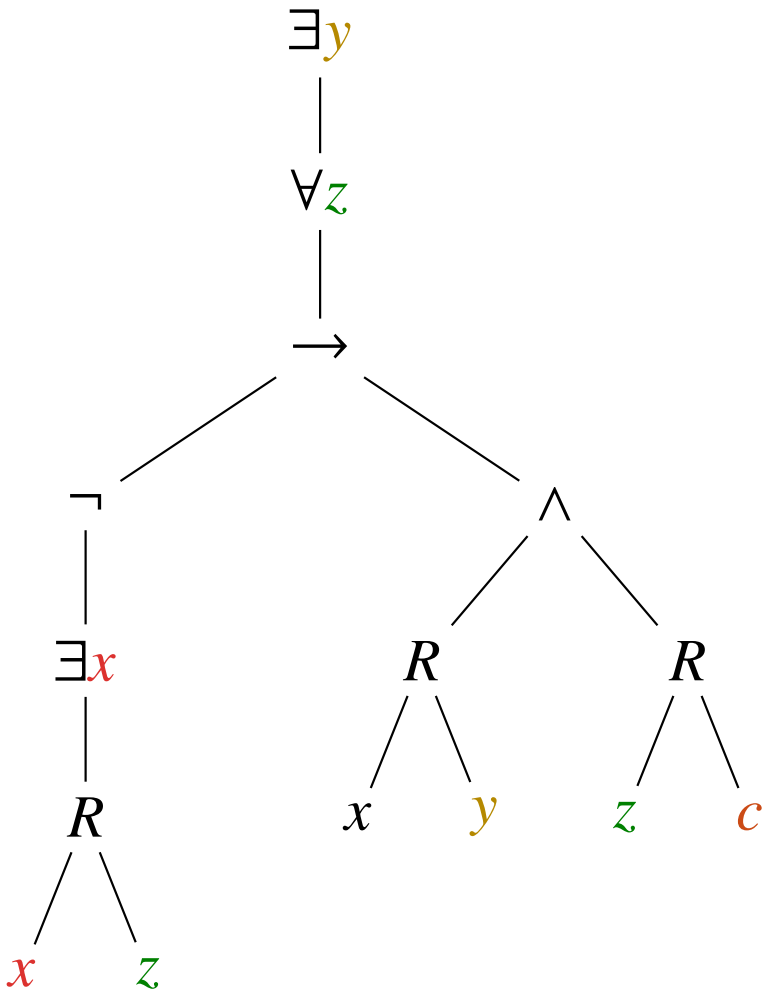
约束变元 & 自由变元 ♡



- In the formula $\forall x A$, we say that the scope(辖域) of this quantifier $\forall$ is A.
- An occurrence of a variable x is bound(约束的) if it occurs within the scope of $\forall x$, or it is the x in $\forall x$.
- An occurrence of a variable is free(自由的) if it is not bound.
- The term t is a closed/ground term(闭项) iff $\mathrm{Var}(t) = \emptyset$.
- A is a closed formula/sentence(闭公式/句子) iff $\mathrm{Fv}(A) = \emptyset$.

约束变元 & 自由变元

$$\underline{\underline{\exists y \forall z (\neg \underline{\exists x R x z} \to R x y \wedge R z c)}}$$



$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \in \mathbb{P}} \frac{1}{1 - \frac{1}{p^s}}$$

$$e^{i\theta} = \cos\theta + i\sin\theta$$

$$\frac{d}{dx} \int_a^x f(t)\,dt = f(x)$$

$$P_k(x) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi}\,dx$$

Variable, Free Variable, Bound Variable

$$\mathrm{Var}(t) := \begin{cases} \{x\} & \text{if } t = x \\ \emptyset & \text{if } t = c \\ \mathrm{Var}(t_1) \cup \cdots \cup \mathrm{Var}(t_n) & \text{if } t = f(t_1, \ldots, t_n) \end{cases}$$

$$\mathrm{Fv}(A) := \begin{cases} \mathrm{Var}(t_1) \cup \mathrm{Var}(t_2) & \text{if } A = t_1 = t_2 \\ \mathrm{Var}(t_1) \cup \cdots \cup \mathrm{Var}(t_n) & \text{if } A = P(t_1, \ldots, t_n) \\ \mathrm{Fv}(B) & \text{if } A = \neg B \\ \mathrm{Fv}(B) \cup \mathrm{Fv}(C) & \text{if } A = B \to C \\ \mathrm{Fv}(B) \setminus \{x\} & \text{if } A = \forall x B \end{cases}$$

$$\mathrm{Bv}(A) := \begin{cases} \emptyset & \text{if } A = t_1 = t_2 \\ \emptyset & \text{if } A = P(t_1, \ldots, t_n) \\ \mathrm{Bv}(B) & \text{if } A = \neg B \\ \mathrm{Bv}(B) \cup \mathrm{Bv}(C) & \text{if } A = B \to C \\ \mathrm{Bv}(B) \cup \{x\} & \text{if } A = \forall x B \end{cases}$$

Remark

- **命题**是具有真值的表达式，是“真值承担者”，或为真或为假，但不能既真又假.
  1. 简奥斯汀是《傲慢与偏见》的作者. ✓
  2. 爱因斯坦是《傲慢与偏见》的作者. ×
  3. 谁是《傲慢与偏见》的作者？？
  4. 因为这个论证有效，所以这个论证无效. ？
- 闭公式是命题.
  1. 某人是《傲慢与偏见》的作者. ✓
  2. 所有人都是《傲慢与偏见》的作者. ×
  3. 某人是某书的作者. ✓
  4. 某人是所有书的作者. ×
- 包含自由变元的公式是**命题函数**，给定自由变元的取值后变为命题.
  1. x 是 y 的作者. ？
  2. x 是《傲慢与偏见》的作者. ？
  3. 简奥斯汀是 y 的作者. ？
  4. 简奥斯汀是《傲慢与偏见》的作者. ✓

翻译 — 学说 “逻辑语”

A	所有学生都漂亮.	$\forall x(Sx \to Px)$
E	没有学生漂亮.	$\forall x(Sx \to \neg Px)$
I	有些学生漂亮.	$\exists x(Sx \wedge Px)$
O	有些学生不漂亮.	$\exists x(Sx \wedge \neg Px)$

1. 有些学生爱所有漂亮的学生.

$$\exists x(Sx \wedge \forall y(Sy \wedge Py \to Lxy))$$

2. 愚蠢的人自信满满, 而聪明的人却充满怀疑. —罗素

$$\forall x(Sx \to Cx) \wedge \forall x(Ix \to Dx)$$

3. 敌人的敌人是朋友.

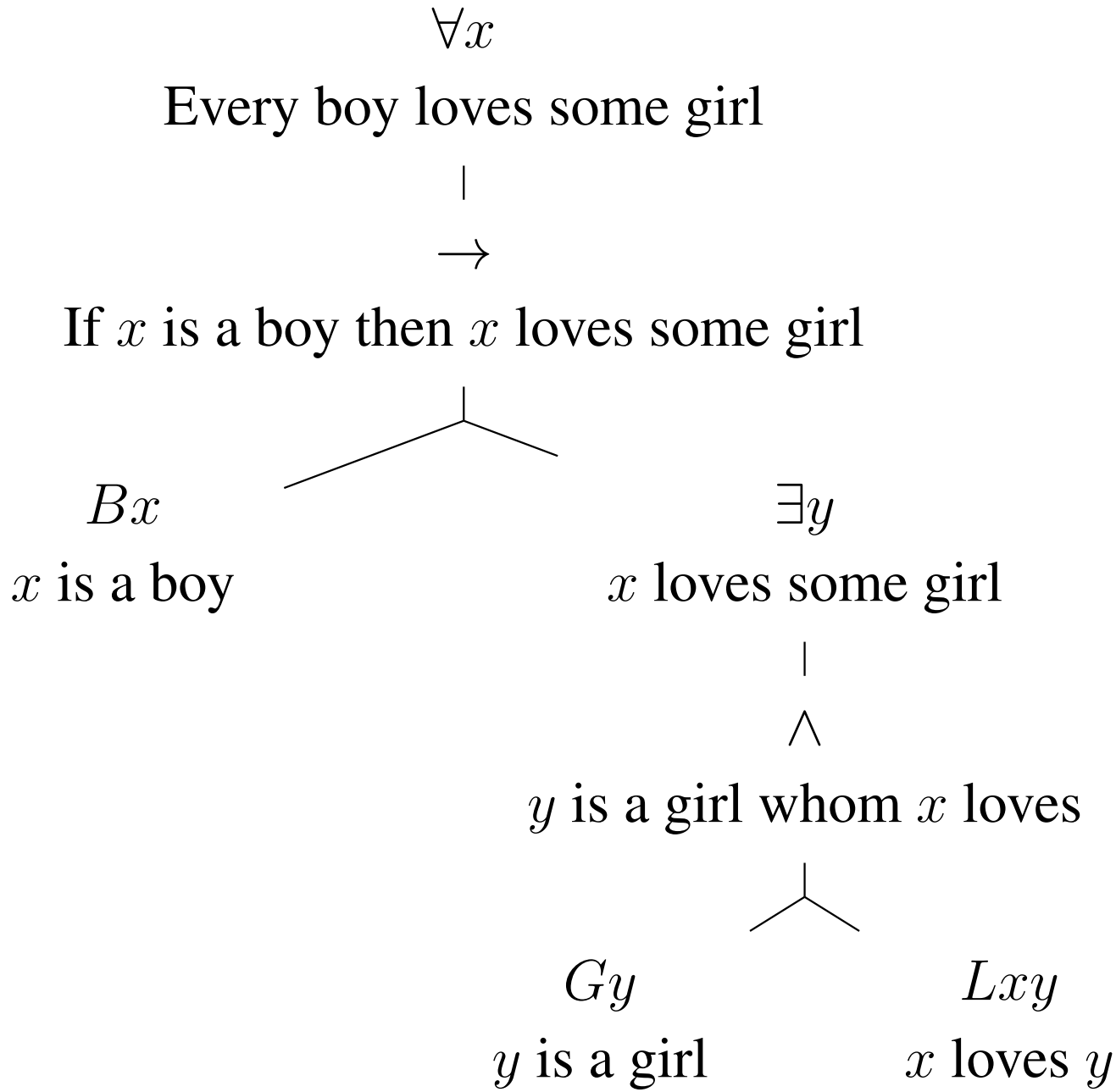
$$\forall x \forall y \forall z(Exy \wedge Eyz \to Fxz)$$

4. 爷爷是爸爸的爸爸.

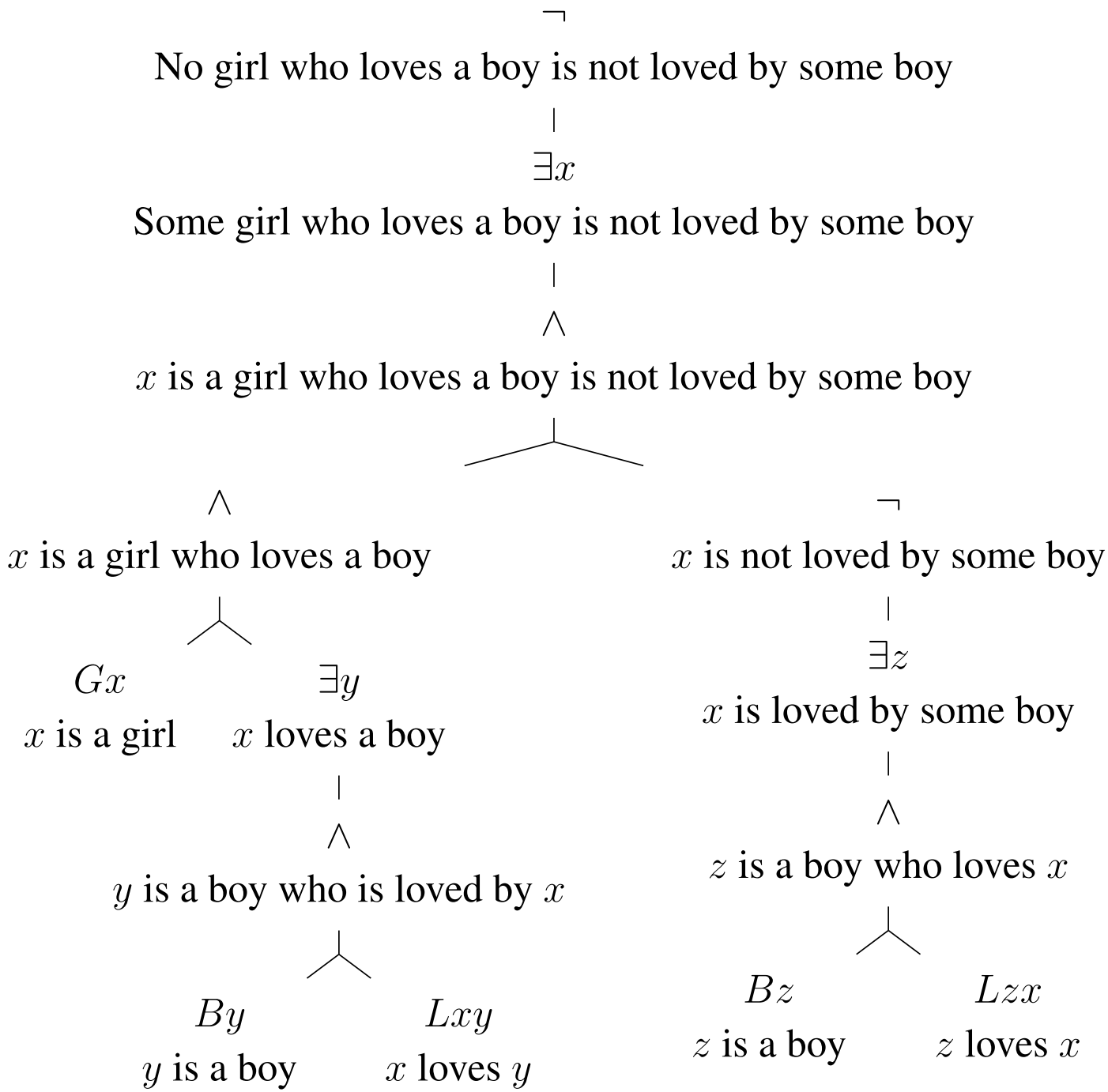
$$\forall x \forall y\big(Gxy \leftrightarrow \exists z(Fxz \wedge Fzy)\big) \qquad \forall x\big(g(x) = f(f(x))\big)$$

5. 有爹就有娘.

$$\forall x(\exists y Fyx \to \exists y Myx)$$



- Every boy loves some girl. $\forall x(Bx \rightarrow \exists y(Gy \wedge Lxy))$
- There is a girl whom every boy loves. $\exists x(Gx \wedge \forall y(By \rightarrow Lyx))$



$$\neg\exists x(Gx \wedge \exists y(By \wedge Lxy) \wedge \neg\exists z(Bz \wedge Lzx))$$

翻译 — 学说 “逻辑语”

1. 谁与你的敌人为友就是与我为敌.

$$\forall x \forall y \big(Eyu \wedge Fxy \to Exm \big)$$

2. 既没有朋友又没有敌人是寂寞的.

$$\forall x (\neg \exists y Fxy \wedge \neg \exists z Exz \to Lx)$$

3. 朋友之间要么都抽烟要么都不抽烟.

$$\forall x \forall y \big(Fxy \to (Sx \leftrightarrow Sy) \big)$$

4. 最可怕的敌人是最亲密的朋友.

$$\forall x \forall y \big(Exy \wedge \forall z (Exz \to Tyz) \to Fxy \wedge \forall z (Fxz \to Cyz) \big)$$

5. 名, 可名, 非常名.

$$\forall x \Big(\exists y \mathsf{Name}(x, y) \wedge \exists z \mathsf{Name}(z, x) \to \neg \mathsf{Unchanging}(x) \Big) ?$$

翻译 — 学说 “逻辑语”

1. If bad things happen to good people, then God is either not omnipotent or not benevolent.

$$\exists x \exists y (\mathrm{Bad}(x) \wedge \mathrm{Good}(y) \wedge \mathrm{Happen}(x, y)) \to \neg \mathrm{Omnip}(g) \vee \neg \mathrm{Benev}(g)$$

2. Every book that Alice lends to Bob she steals from Chris.

$$\forall x \big(Bx \wedge Lxab \to Sxac \big)$$

3. For every professor, there is a student who likes every book the professor recommends to the student.

$$\forall x \Big(Px \to \exists y \big(Sy \wedge \forall z (Bz \wedge Rzxy \to Lyz) \big) \Big)$$

4. When a mathematical or philosophical author writes with a misty profundity, he is talking nonsense. — *Alfred Whitehead*

$$\forall x \Big(\big(\mathrm{Math}(x) \vee \mathrm{Pilo}(x) \big) \wedge \mathrm{Write}(x) \to \mathrm{TalkNonsense}(x) \Big)$$

5. No dolphin sings unless it jumps.

$$\forall x \Big(\mathrm{Dolphin}(x) \to \neg \mathrm{Jump}(x) \to \neg \mathrm{Sing}(x) \Big)$$

翻译 — 学说 “逻辑语”

1. If all dancers have knee injuries, then *they* will be frustrated.

$$\forall x\Big(\text{Dancer}(x) \to \text{Knee}(x)\Big) \to \forall x\Big(\text{Dancer}(x) \to \text{Frustrated}(x)\Big)$$

2. If all dancers have knee injuries, then *some of them* will be frustrated.

$$\forall x\Big(\text{Dancer}(x) \to \text{Knee}(x)\Big) \to \exists x\Big(\text{Dancer}(x) \wedge \text{Frustrated}(x)\Big)$$

3. If a student takes a course and the course covers a concept, then the student knows that concept.

$$\forall x \forall y \forall z\Big(\text{Student}(x) \wedge \text{Course}(y) \wedge \text{Take}(x,y) \wedge \text{Concept}(z) \wedge \text{Cover}(y,z) \to \text{Know}(x,z)\Big)$$

4. Alice is the first to have completed the test.

$$Cat \wedge \forall x(Cxt \wedge x \neq a \to Baxt)$$

   Alice is the oldest daughter of the King.

5. Alice is the second to have completed the test.

$$Cat \wedge \exists x\Big(Cxt \wedge x \neq a \wedge Bxat \wedge \forall y(Cyt \wedge y \neq a \wedge y \neq x \to Bayt)\Big)$$

Example — 究竟有几个神?

有神论 至少有一个神

$$\exists x Gx$$

无神论 没有神

$$\neg \exists x Gx$$

泛神论 万物皆神

$$\forall x Gx$$

不可知论 至多有一个神

$$\forall x(Gx \to \forall y(Gy \to y = x))$$

一神论 有且仅有一个神

$$\exists x(Gx \wedge \forall y(Gy \to y = x))$$

二元神论 有且仅有两个神

$$\exists x(Gx \wedge \exists y(Gy \wedge x \neq y) \wedge \forall z(Gz \to z = x \vee z = y))$$

至多/至少/恰好

1. 至多有 1 个元素.

$$\forall x \forall y (x = y)$$

2. 至少有 2 个元素.

$$\exists x \exists y (x \neq y)$$

3. 有且仅有 2 个元素.

$$\exists x \exists y \big(x \neq y \wedge \forall z (z = x \vee z = y) \big)$$

4. 有且仅有 n 个元素.

$$\exists x_1 \ldots \exists x_n \left(\bigwedge_{1 \leq i < j \leq n} x_i \neq x_j \wedge \forall y \left(\bigvee_{i=1}^{n} y = x_i \right) \right)$$

5. 有且仅有 1 个元素有 P 性质. (缩写为: $\exists! x P x$)

$$\exists x P x \wedge \forall y \forall z (P y \wedge P z \to y = z)$$

$$\exists x \big(P x \wedge \forall y (P y \to y = x) \big)$$

$$\exists x \forall y (P y \leftrightarrow y = x)$$

爱丽丝镜中奇遇 — Lewis Carroll

- "You are sad," the Knight said in an anxious tone: "let me sing you a song to comfort you."
- "Is it very long?" Alice asked.
- "It's long," said the Knight, "but it's very, very beautiful. Everybody that hears me sing it — either it brings the tears into their eyes, or else —"

$$\exists x \forall y((\mathsf{Song}(x) \leftrightarrow y = x) \wedge \mathsf{Long}(x) \wedge \mathsf{Beautiful}(x))$$

- "Or else what?" said Alice.
- "Or else it doesn't, you know. The name of the song is called 'Haddocks' Eyes.'"

$$\exists! x(\mathsf{Song}(x) \wedge \forall y(\mathsf{Hear}(y, x) \rightarrow \mathsf{Tear}(x, y) \vee \neg\mathsf{Tear}(x, y)))$$

$$\mathsf{callname}(\mathsf{name}(\text{the-song})) = \text{Haddocks' Eyes}$$

- “Oh, that’s the name of the song, is it?” said Alice.

$$\text{name(the-song)} = \text{Haddocks' Eyes}$$

- “No, you don’t understand,” the Knight said. “That’s what the name is called. The name really is ‘The Aged Aged Man.”’

$$\text{name(the-song)} = \text{The Aged Aged Man}$$

- “Then I ought to have said ‘That’s what the song is called’?” Alice corrected herself.

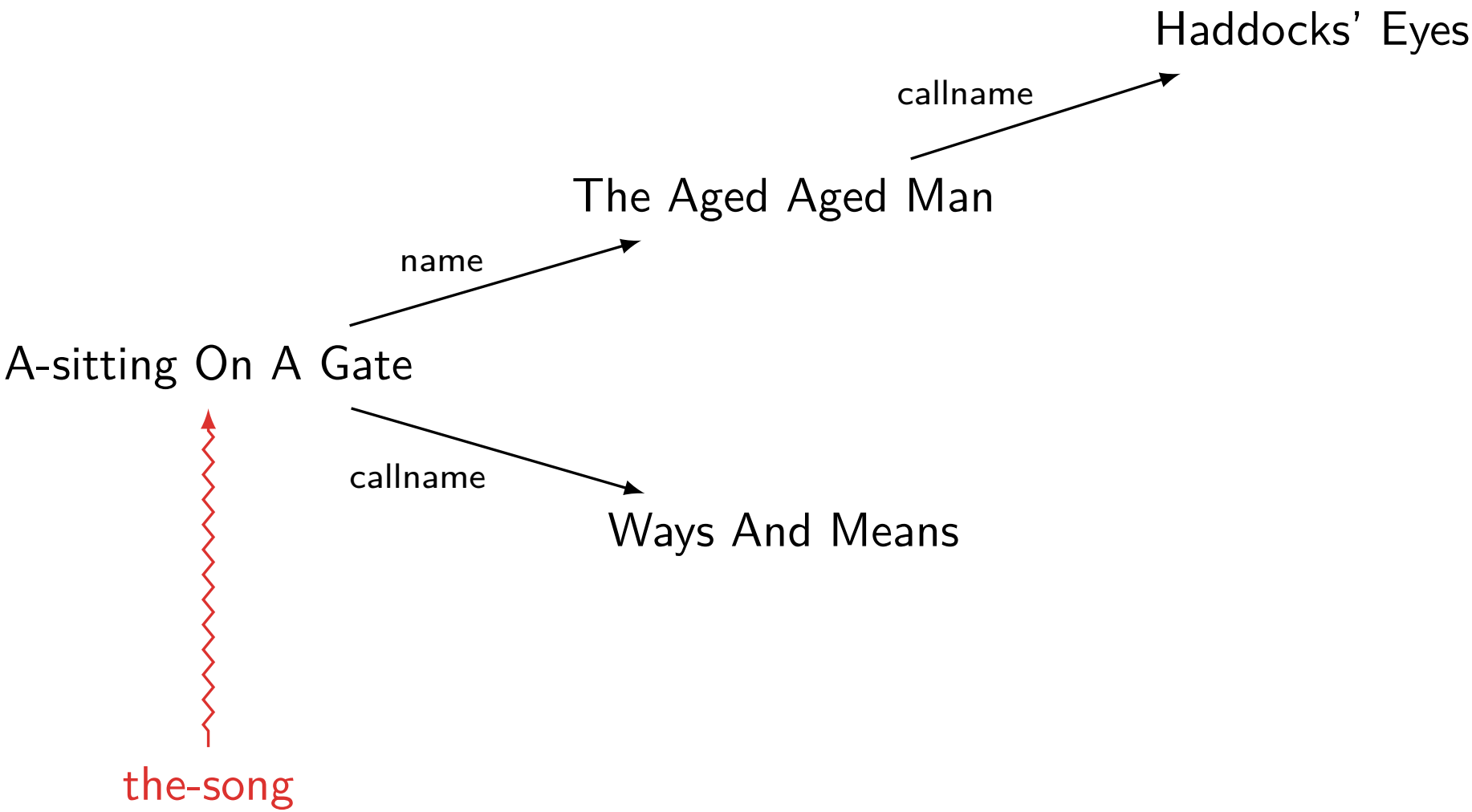
$$\text{callname(the-song)} = \text{The Aged Aged Man}$$

- “No, you oughtn’t: that’s quite another thing! The song is called ‘Ways And Means’: but that’s only what it’s called, you know!”

$$\text{callname(the-song)} = \text{Ways And Means}$$

- “Well, what is the song, then?” said Alice.
- “I was coming to that.” the Knight said. “The song really is ‘A-sitting On A Gate’.”

$$\text{the-song} = \text{A-sitting On A Gate}$$


Haddocks' Eyes
callname
The Aged Aged Man
name
A-sitting On A Gate
callname
Ways And Means
the-song

1. No dogs must be brought to this Park except on a lead.

$$\forall d \forall t (Bdt \rightarrow Ldt)$$

   Objection: this is consistent with $\exists d \exists t (Pdt \wedge \neg Ldt)$.
2. Dogs are not allowed in this Park without leads.

$$\forall d \forall t (Pdt \rightarrow Ldt)$$

   Objection? The order should be addressed to the owners, not to the dogs.
3. Owners of dogs are not allowed in this Park unless they keep them on leads.

$$\forall p \forall d \forall t (Opd \wedge Ppt \rightarrow Lpdt)$$

   Objection: doesn't allow $\exists t \exists p \exists d (Opd \wedge Ppt \wedge \neg Pdt \wedge \neg Lpdt)$.
4. Nobody without his dog on a lead is allowed in this Park.

$$\forall p \forall t (Ppt \rightarrow \exists d (Opd \wedge Lpdt))$$

   Objection: doesn't allow $\exists p \exists t (Ppt \wedge \neg \exists d Opd)$.
5. Dogs must be led in this Park.

$$\forall d \forall t (Ldt \wedge Pdt) \qquad \forall d \forall t (Ldt \rightarrow Pdt)$$

6. All dogs in this Park must be kept on the lead.

$$\forall d \forall t (Pdt \rightarrow Ldt)$$
$$\forall d \forall t \forall p (Pdt \wedge Opd \rightarrow Lpdt)$$

翻译 — 学说 “逻辑语”

1. He who learns but does not think, is lost. He who thinks but does not learn is in danger. — *Confucius*

$$\forall x(\mathsf{Learn}(x) \wedge \neg\mathsf{Think}(x) \rightarrow \mathsf{Lost}(x)) \wedge$$

$$\forall x(\mathsf{Think}(x) \wedge \neg\mathsf{Learn}(x) \rightarrow \mathsf{InDanger}(x))$$

2. He who can, does. He who cannot, teaches. — *Bernard Shaw*

$$\forall xy(\mathsf{Can}(x,y) \rightarrow \mathsf{Do}(x,y)) \wedge \forall xy(\neg\mathsf{Can}(x,y) \rightarrow \mathsf{Teach}(x,y))$$

3. Don't interrupt me, while I'm interrupting. — *Churchill*

$$\exists x\mathsf{Interrupt}(i,x) \rightarrow \forall y\neg\mathsf{Interrupt}(y,i)$$

4. There are no shortcuts to any place worth going. — *Beverly Sills*

$$\forall x(\mathsf{Place}(x) \wedge \mathsf{WorthGo}(x) \rightarrow \neg\exists y\mathsf{Shortcut}(y,x))$$

5. The old believe everything; the middle-aged suspect everything; the young know everything. — *Oscar Wilde*

$$\forall x(\mathsf{Old}(x) \rightarrow \forall y\mathsf{Believe}(x,y)) \wedge$$

$$\forall x(\mathsf{Middle}(x) \rightarrow \forall y\mathsf{Suspect}(x,y)) \wedge \forall x(\mathsf{Young}(x) \rightarrow \forall y\mathsf{Know}(x,y))$$

6. No married man is ever attractive except to his wife. — *Oscar Wilde*

$$\forall x(\mathsf{Man}(x) \wedge \mathsf{Married}(x) \rightarrow \forall y(\mathsf{Attractive}(x,y) \rightarrow y = \mathsf{wife}(x)))$$

翻译 — 学说 “逻辑语”

7. Behind every great fortune there is a crime. — *Balzac*

$$\forall x(\text{GreatFortune}(x) \rightarrow \exists y(\text{Crime}(y) \wedge \text{Behind}(y, x)))$$

8. Always two there are: a Master and an Apprentice. — *Yoda*

$$\exists xy(x \neq y \wedge \forall z(z = x \vee z = y) \wedge \text{Master}(x) \wedge \text{Apprentice}(y))$$

9. There are two ways to live your life. One is as though nothing is a miracle. The other is as though everything is a miracle. — *Einstein*

$$\exists xy(x \neq y \wedge \forall z(\text{WayLive}(z) \leftrightarrow z = x \vee z = y) \wedge$$
$$(\text{WayLive}(x) \rightarrow \neg\exists z\text{Miracle}(z)) \wedge (\text{WayLive}(y) \rightarrow \forall z\text{Miracle}(z)))$$

10. The weakest link in a chain is the strongest because it can break it. — *Stanislaw J. Lec*

$$\forall xy((\text{Link}(x) \wedge \text{Chain}(y) \wedge \text{In}(x, y) \wedge \text{Weakest}(x) \rightarrow \text{Break}(x, y))$$
$$\rightarrow \text{Strongest}(x))$$

11. A man with only a hammer sees every problem as a nail. Our age's greatest hammer is the algorithm. — *Poundstone*

$$\forall x(\text{Man}(x) \wedge \exists! y(\text{Hammer}(y) \wedge \text{With}(x, y)) \rightarrow \forall z(\text{Problem}(z) \rightarrow$$
$$\exists n(\text{Nail}(n) \wedge \text{Seeas}(x, z, n)))) \quad \wedge \quad \exists x(\text{Hammer}(x) \wedge \text{Age}(x, \text{we}) \wedge$$
$$\forall y(\text{Hammer}(y) \wedge \text{Age}(y, \text{we}) \rightarrow \text{Greater}(x, y)) \wedge x = \text{algorithm}))$$

1. $\mathsf{Think}(i) \to \exists x(x = i)$ *Descartes*
2. $\exists x(x = i) \vee \neg\exists x(x = i)$ *Shakespeare*
3. $\forall x\big(\mathsf{Month}(x) \to \mathsf{Crueler}(\mathsf{april}, x)\big)$ *Eliot*
4. $\forall x(\neg\mathsf{Weep}(x) \to \neg\mathsf{See}(x))$ *Hugo*
5. $\forall x\big(\mathsf{Time}(x) \to \mathsf{Better}(t, x)\big) \wedge \forall x\big(\mathsf{Time}(x) \to \mathsf{Better}(x, t)\big)$ *Dickens*
6. $\exists x\Big(\mathsf{Child}(x) \wedge \neg\mathsf{Growup}(x) \wedge \forall y\big(\mathsf{Child}(y) \wedge y \neq x \to \mathsf{Growup}(y)\big)\Big)$ *Barrie*
7. $\forall x\forall y\big(Fx \wedge Fy \to (Hx \wedge Hy \to Axy) \wedge (\neg Hx \wedge \neg Hy \to \neg Axy)\big)$ *Tolstoi*
8. $\forall x(Px \to \exists y(Ty \wedge Fuxy)) \wedge \exists x(Px \wedge \forall y(Ty \to Fuxy)) \wedge \neg\forall x(Px \to \forall y(Ty \to Fuxy))$ *Lincoln*
9. $\forall x\big(\mathsf{Problem}(x) \wedge \mathsf{Philo}(x) \wedge \mathsf{Serious}(x) \leftrightarrow x = \mathsf{suicide}\big)$ *Camus*
10. $\forall x\big(\mathsf{Feather}(x) \wedge \mathsf{Perch}(x, \mathsf{soul}) \leftrightarrow x = \mathsf{hope}\big)$ *Dickinson*
11. $\forall x\Big(\mathsf{Enter}(x) \to \forall y\big(\mathsf{Hope}(y) \to \mathsf{Abandon}(x, y)\big)\Big)$ *Dante*
12. $\exists x\forall y\big(\mathsf{For}(y, x) \wedge \mathsf{For}(x, y)\big)?$ $\forall x\forall y\big(\mathsf{For}(y, x) \leftrightarrow \mathsf{For}(x, y)\big)?$ *Dumas*
13. $\forall x\big(\mathsf{Fear}(\mathsf{we}, x) \leftrightarrow x = \mathsf{fear}\big)?$ *Roosevelt*
14. $\forall x\forall y(Ax \wedge Ay \to Exy) \wedge \exists x\exists y\big(Ax \wedge Ay \wedge [\![Exx]\!] > [\![Eyy]\!]\big)?$ *Orwell*

1. I think, therefore I am. *Descartes*
2. To be or not to be. *Shakespeare*
3. April is the cruellest month. *Eliot*
4. Those who do not weep, do not see. *Hugo*
5. It was the best of times, it was the worst of times. *Dickens*
6. All Children, except one, grow up. *Barrie*
7. All happy families are alike; each unhappy family is unhappy in it's own way. *Tolstoi*
8. You can fool all the people some of the time, and some of the people all the time, but you can't fool all the people all the time. *Lincoln*
9. There is but one truly serious philosophical problem and that is suicide. *Camus*
10. Hope is the thing with feathers that perches in the soul. *Dickinson*
11. All hope abandon, all you who enter here. *Dante*
12. One for all, all for one. *Dumas*
13. The only thing we have to fear is fear itself. *Roosevelt*
14. All animals are equal, but some animals are more equal than others. *Orwell*

翻译 — “逻辑语言” 到 “自然语言”

1. $\mathsf{Love}(\mathsf{alice}, \mathsf{bob}) \wedge \exists x(\mathsf{Girl}(x) \wedge x \neq \mathsf{alice} \wedge \mathsf{Love}(\mathsf{bob}, x))$
2. $\forall x\big(\mathsf{Buy}(\mathsf{alice}, x) \rightarrow \mathsf{Buy}(\mathsf{benny}, x)\big)$
3. $\forall x\big(\exists y \mathsf{Buy}(y, x) \vee \exists z \mathsf{Sell}(z, x) \rightarrow \neg \mathsf{Human}(x)\big)$
4. $\forall x\Big(\big(\mathsf{Petted}(x) \rightarrow \mathsf{Jump}(x)\big) \rightarrow \mathsf{Dog}(x) \vee \mathsf{Dolphin}(x)\Big)$
5. $\forall x\Big(\mathsf{Girl}(x) \wedge \mathsf{Love}(\mathsf{bob}, x) \rightarrow \forall y\big(\mathsf{PhilosophyBook}(y) \rightarrow \mathsf{Read}(x, y)\big)\Big)$

- One driver dies in car accident every minute.
- Every minute one driver dies in car accident.
- $\exists x\Big(\mathsf{Driver}(x) \wedge \forall y\big(\mathsf{Minite}(y) \rightarrow \mathsf{Die}(x, y)\big)\Big)$
- $\forall x\Big(\mathsf{Minute}(x) \rightarrow \exists y\big(\mathsf{Driver}(y) \wedge \mathsf{Die}(y, x)\big)\Big)$

翻译 — “逻辑语言” 到 “自然语言”

1. $\exists x\Big(Gx \wedge \forall y\big(By \wedge \forall z(Gz \wedge z \neq x \to \neg Lzy) \to Lxy\big)\Big) \to \forall x\Big(Bx \to \exists y\big(Gy \wedge Lyx\big)\Big)$
2. $\forall x \forall y\Big(\big(Gx \wedge \forall y(By \to \neg Lxy)\big) \wedge \big(Gy \wedge \exists x(Bx \wedge Lyx)\big) \to \neg Lxy\Big)$

相信我，我肯定能找到一种你不屑于理解的语言来试图跟你对（zhuang）话（B）的。

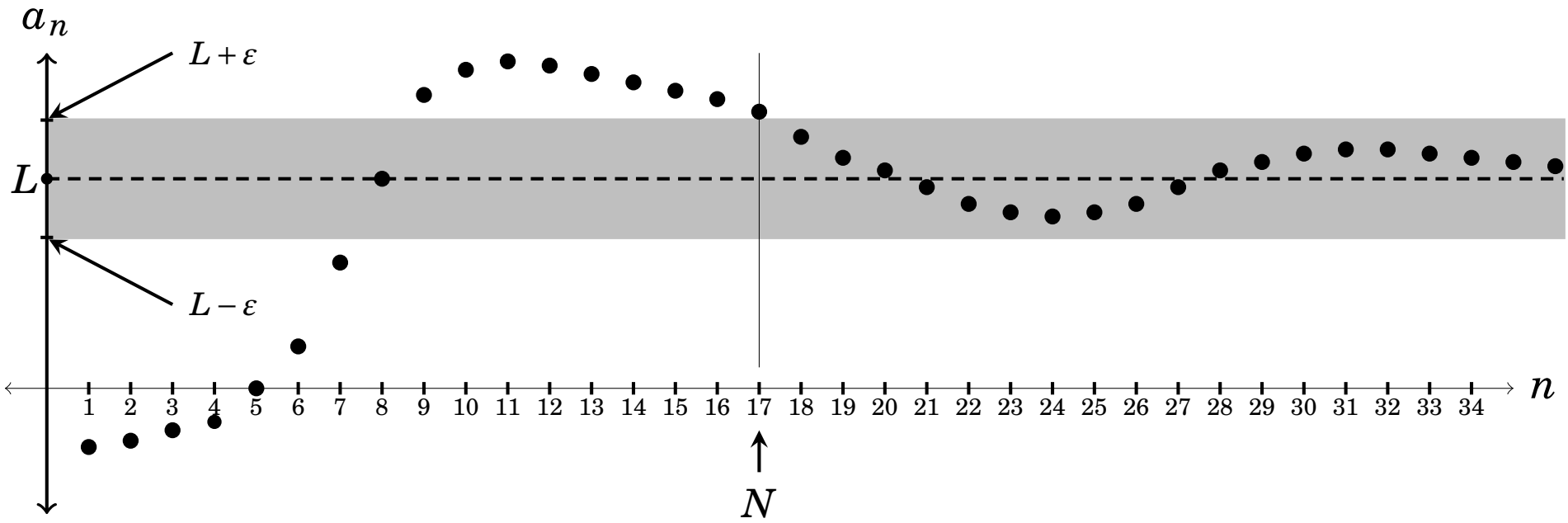


安得圣母爱剩男，大庇天下雄性有红颜

女同不爱女异

$$\forall x \forall y(((Gx \wedge \forall v(Bv \to \neg Lxv)) \wedge (Gy \wedge \exists z(Bz \wedge Lyz))) \to \neg Lxy)$$

翻译 — “自然语言” 到 “逻辑语言”



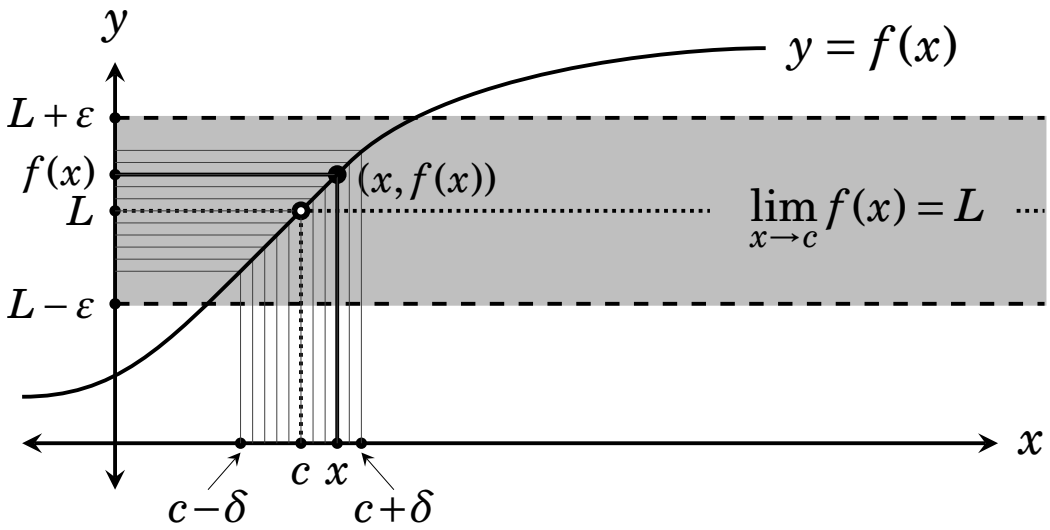
Definition (Convergence)

A sequence $\{a_n\}$ converges to a number $L \in \mathbb{R}$ provided that for any real number $\varepsilon > 0$ there is an $N \in \mathbb{N}$ for which $n > N$ implies $|a_n - L| < \varepsilon|$.

$$\lim_{n\to\infty} a_n = L \iff \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N \big(|a_n - L| < \varepsilon\big)$$

$$\lim_{n\to\infty} a_n = \infty \iff \forall L \in \mathbb{R} \exists N \in \mathbb{N} \forall n \geq N \big(a_n > L\big)$$

翻译 — “自然语言” 到 “逻辑语言”



- Informal definition of a limit: $f(x)$ is arbitrarily close to L provided that x is sufficiently close to c.
- Precise definition: for any real number $\varepsilon > 0$, there is a real number $\delta > 0$ for which $|f(x) - L| < \varepsilon$ provided that $0 < |x - c| < \delta$.
- Formal defnition:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \big(0 < |x - c| < \delta \rightarrow |f(x) - L| < \varepsilon\big)$$

- Divergence

$$\lim_{x \to c} f(x) \uparrow \iff \forall L \in \mathbb{R} \exists \varepsilon > 0 \forall \delta > 0 \exists x \big(0 < |x - c| < \delta \wedge |f(x) - L| \geq \varepsilon\big)$$

翻译 — “自然语言” 到 “逻辑语言”

- Continuity

$$\forall x \in I \forall \varepsilon > 0 \exists \delta > 0 \forall y \in I \big(|x - y| < \delta \rightarrow |f(x) - f(y)| < \varepsilon \big)$$

$$\forall (x_n)_{n \in \mathbb{N}} \subset I : \lim_{n \to \infty} f(x_n) = f(\lim_{n \to \infty} x_n)$$

- Uniform Continuity

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in I \forall y \in I \big(|x - y| < \delta \rightarrow |f(x) - f(y)| < \varepsilon \big)$$

- Any two distinct points determine a unique line.

$$\forall x \forall y \Big(\mathrm{Point}(x) \wedge \mathrm{Point}(y) \wedge x \neq y \rightarrow \exists! z \mathrm{Line}(z) \wedge \mathrm{On}(x, z) \wedge \mathrm{On}(y, z) \Big)$$

- There are infinitely many Prime numbers.

$$\forall x \exists y \big(y > x \wedge \mathrm{Prime}(y) \big)$$

- There are only finitely many Even Prime numbers.

$$\exists x \forall y \big(\mathrm{Even}(y) \wedge \mathrm{Prime}(y) \rightarrow y < x \big)$$

练习: 翻译 — Now it's your turn

1. Everyone hears only what he understands. — Goethe
2. If you can't solve a problem, then there is an easier problem that you can't solve. — Polya
3. Men *and* women are welcome to apply.
4. Alice can't do *every* job right.
5. Alice can't do *any* job right.
6. 除了熟香蕉都不可以吃.
7. 只有苏格拉底和柏拉图是人.
8. 除了苏格拉底和柏拉图都是人.
9. 每个男孩爱着至少两个女孩.
10. 闪光的不都是金子.
11. 同一个城市的人的身份证号的第一位数字相同.
12. 有女朋友的男孩都幸福.
13. 有女朋友的男孩都宠她.
14. 所有的偶数都可以被 2 整除, 但只有一些可以被 4 整除.

练习: 翻译 — Now it's your turn

15. 喜欢所有女生的男生没有女生喜欢.
16. 每个班都有一个同学被所有同班同学喜欢.
17. 如果狗是动物, 那么狗的头就是动物的头.
18. 苏格拉底的老婆的脸只有她亲妈才不嫌弃.
19. 只有段正淳才会喜欢两个或两个以上的女孩.
20. 黄药师鄙视所有不自我鄙视的人.
21. 乔峰只爱也爱他的那个女孩.
22. 爱着王语嫣的那个人不是王语嫣所爱的那个人.
23. 最高的男孩爱着最矮的女孩.
24. If a clown enters the room, then it will be displeased if no person is surprised.
25. Everyone alive $2000BC$ is either an ancestor of nobody alive today or of everyone alive today.
26. No man loves children unless he has his own.
27. Alice and Bob have the same maternal grandmother. Mother(x, y)
28. Someone *other than the girl* who loves Bob is stupid.

翻译

- 团结既不是胜利的充分条件也不是胜利的必要条件.

$$\neg(U \to V) \wedge \neg(V \to U)?$$

Remark: 严格来说, 命题逻辑不足以表达充分和必要条件.

$$\neg\forall x(Px \wedge Ux \to Vx) \wedge \neg\forall x(Px \wedge Vx \to Ux)$$

Remark: 存在一些可能情境, 团结了但没胜利; 也存在一些可能情境, 胜利了但不团结.

- 每个人的自由发展是一切人的自由发展的必要条件.

— 马克思

$$\forall x(Hx \to Fx) \to \forall x(Hx \to Fx)?$$

$$\frac{\forall x(Hx \to Fx \to \forall y(Hy \to Fy))}{\exists x(Hx \wedge Fx) \to \forall y(Hy \to Fy)}$$

代入 $[t/x]$: 用 t 代入 x 的所有自由出现

Definition (Substitution in a term)

$$s[t/x] := \begin{cases} c & \text{if } s = c \\ t & \text{if } s = x \\ y & \text{if } s = y \ \& \ y \neq x \\ f(t_1[t/x], \ldots, t_n[t/x]) & \text{if } s = f(t_1, \ldots, t_n) \end{cases}$$

Definition (Substitution in a formula)

$$A[t/x] := \begin{cases} t_1[t/x] = t_2[t/x] & \text{if } A = t_1 = t_2 \\ P(t_1[t/x], \ldots, t_n[t/x]) & \text{if } A = P(t_1, \ldots, t_n) \\ \neg B[t/x] & \text{if } A = \neg B \\ B[t/x] \to C[t/x] & \text{if } A = B \to C \\ \begin{cases} \forall y B[t/x] & \text{if } y \neq x \\ \forall y B & \text{if } y = x \end{cases} & \text{if } A = \forall y B \end{cases}$$

代入 vs 可代入[10]

Definition: 代入 $[t/x]$ 用项 t 代入变元 x 的所有自由出现.

Definition (可代入)

项 t 对变元 x 在公式 A 中可代入, 当且仅当, 用 t 替换 x 在 A 中的所有自由出现后, t 中的变元不会被 A 中的量词所约束.

Remark: 我们写 $A[t/x]$ 时通常默认 t 对 x 在 A 中可代入, 有时会简写为 $A(t)$.

Counter-example: $A = \exists y(y \neq x) \quad t = y \quad A[t/x]? \quad \forall x A \vdash A(t)$

---

[10] A term t is substitutable for x in

- any atomic formula A.
- $\neg A$ if it is substitutable for x in A.
- $A \to B$ if it is substitutable for x in A and B.
- $\forall y A$ if either
  1. $x \notin \mathrm{Fv}(A)$ or
  2. $y \notin \mathrm{Var}(t)$ and t is substitutable for x in A.

Contents

Meta-Theorems

Introduction

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Term Logic

Recursion Theory

Propositional Logic

Equational Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

What does "true" mean?

Correspondence Theory	corresponding to the facts
Coherence Theory	cohering with our other beliefs
Pragmatist Theory	being useful to believe
Verification Theory	being verified
Ideal Consensus Theory	being what we'd agree to under cognitively ideal conditions
Redundancy Theory	"It's true that A" is just a wordy way to assert A

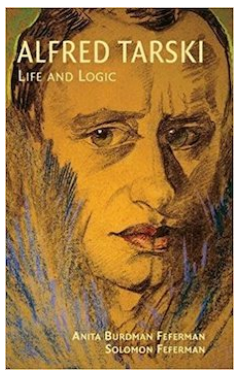
什么是 “真”?

- No entity without identity. — Quine's standards of ontological admissibility
- To be is to be the value of a bound variable. — Quine's criterion of ontological commitments
- To be is to be constructed by intuition. — Brouwer
- To be true is to be provable. — Kolmogorov
- ⌜snow is white⌝ is true iff snow is white. — Tarski's "T-schema"

What is "truth"? — Are all truths knowable?

1. *formally correct* $\forall x\big(\mathrm{True}(x) \leftrightarrow \varphi(x)\big)$
2. *materially adequate* $\varphi(\ulcorner p \urcorner) \leftrightarrow p^*$
   where $\ulcorner p \urcorner$ is the name of a sentence p of $\mathscr{L}$, and p^* is the translation of p in the meta-language $\mathscr{L}'$.

$$\mathrm{True}(\ulcorner W(s) \urcorner) \leftrightarrow \mathrm{White}(\mathrm{snow})$$



塔斯基 Alfred Tarski 1901-1983

什么是 “真”?

「雪是白的」是真的, 当且仅当, 雪是白的.

“I am not true.”[11]

模型论



Undefinability of truth Theorem

Arithetical truth can't be defined in arithmetic.

The theory of real closed fields / elementary geometry is complete and decidable.

Banach-Tarski Paradox

[11] Tarski: On the Concept of Truth in Formalized Languages. 1933.
Tarski: The Semantic Conception of Truth and the Foundations of Semantics. 1944.

语法 vs 语义

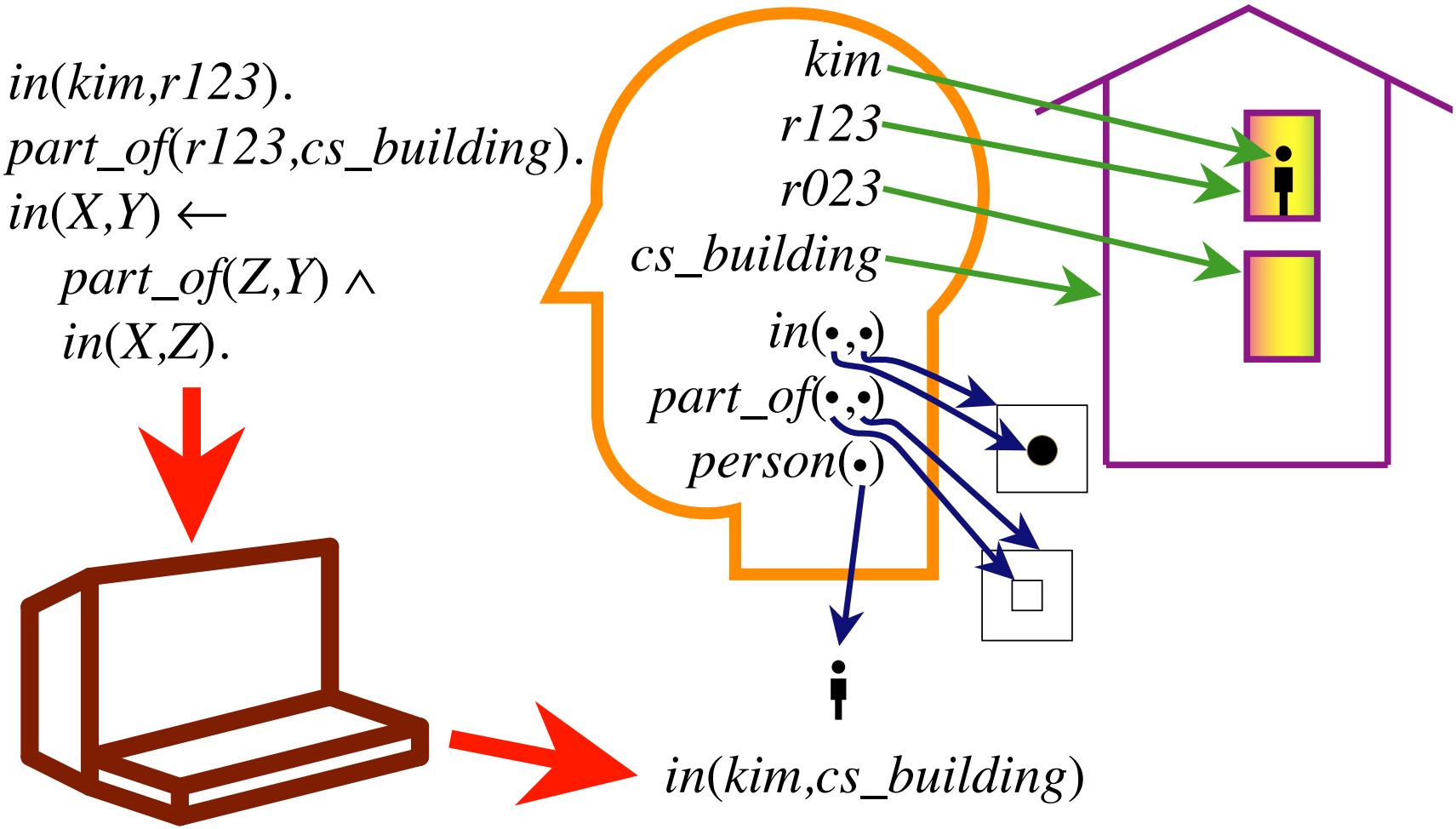


Figure: The computer takes in symbols and outputs symbols. The meaning of the symbols are in the user's head.

语言 — 思想 — 世界

- What is the meaning of 'meaning'? (symbol grounding problem)
- How do words relate to objects? thought?
- What makes a sentence true/false?

proper name	predicate	sentence
↓	↓	↓
sense of the proper name	property	proposition
↓	↓	↓
object	concept	truth-value
	↓	
	objects that fall under the concept	

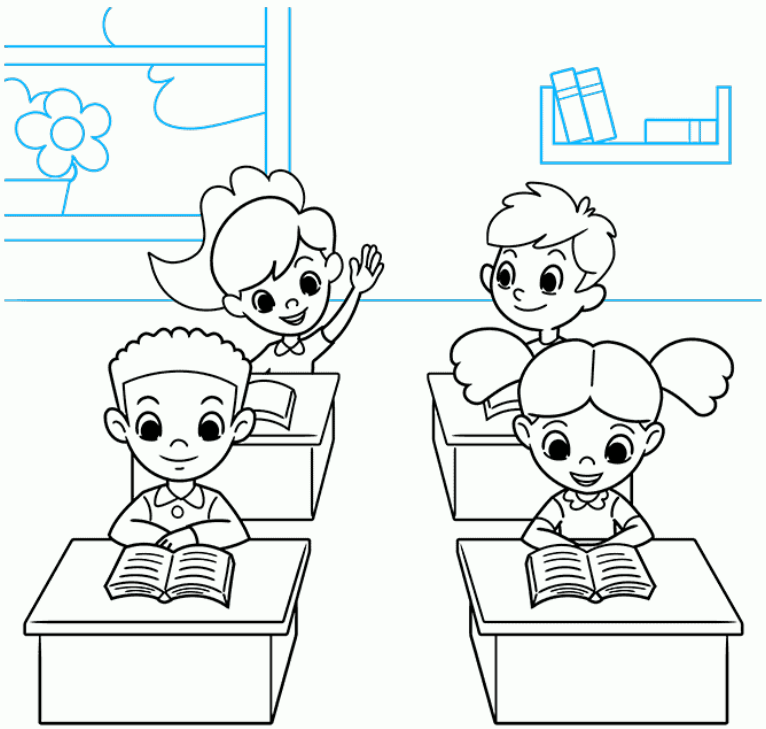
Table: Frege: symbol, sense & reference

Frege: The meaning of a term is a function/algorithm which computes its denotation.

Wittgenstein: The limits of my language means the limits of my world.

结构! 结构! 什么是 “结构”?

结构



语言

- 班里有四名同学
- 两名男同学
- 两名女同学
- 有个女同学叫小艾
- 小艾在举手
- 小白在看小艾
- 小艾的大姐在小白前面
- 某男同学在小艾前面
- 小艾前面是一头粉色大象

$$(\mathbb{Z}, 0, 1, +, \cdot, <) \vDash \exists x \neq 0 \exists y \neq 0 \exists z \neq 0 : x^2 + y^2 = z^2$$

$$(\mathbb{Z}, 0, 1, +, \cdot, <) \nvDash \exists x \neq 0 \exists y \neq 0 \exists z \neq 0 : x^3 + y^3 = z^3$$

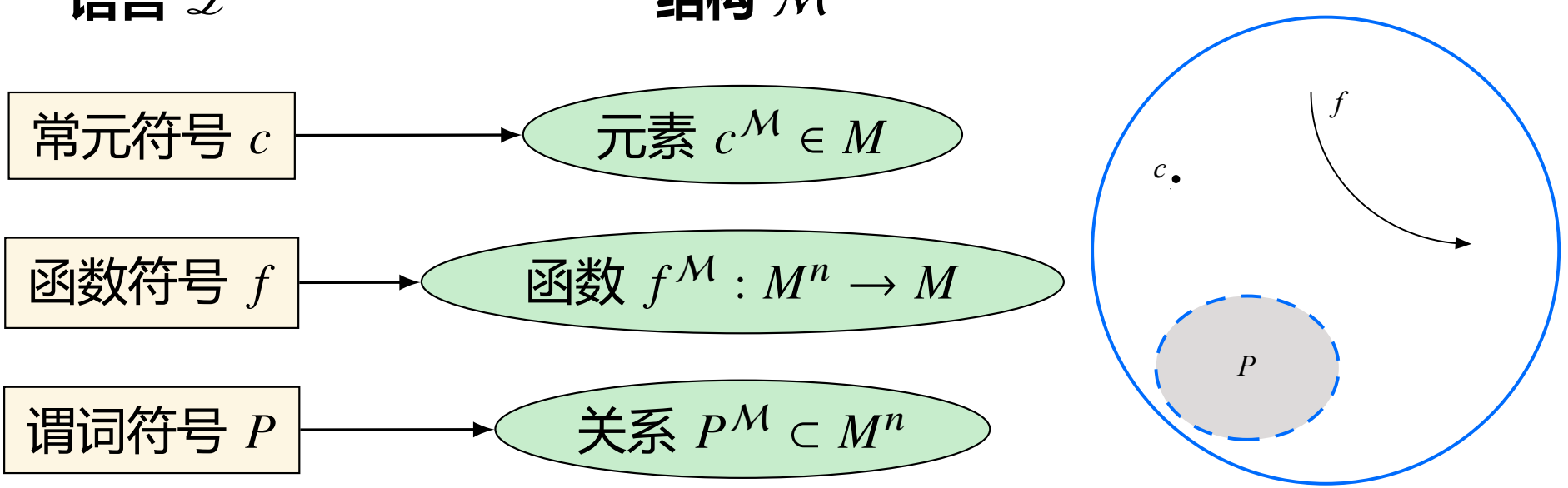
结构 ♡

Definition (结构)

结构 $\mathcal{M} := (M, [\![\,]\!])$ 由非空集合 M 及其上的映射 $[\![\,]\!]$ 构成, 其中 $[\![\,]\!]$ 把

- 常元符号 c 映射为元素 $[\![c]\!] \in M$
- n 元函数符号 f 映射为函数 $[\![f]\!] : M^n \to M$
- n 元谓词符号 P 映射为关系 $[\![P]\!] \subset M^n$

Remark: 为了方便, 我们通常把结构 $\mathcal{M}$ 简记为 $(M, c^{\mathcal{M}}, f^{\mathcal{M}}, P^{\mathcal{M}})$.



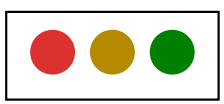
Remarks

荀子: 制 “名” 以指 “实”

$$[\![\text{名}]\!] = \text{实}$$

$$\text{名}^{\mathcal{M}} = \text{实}$$

红灯停; 绿灯行; 黄灯表示警示的功能.

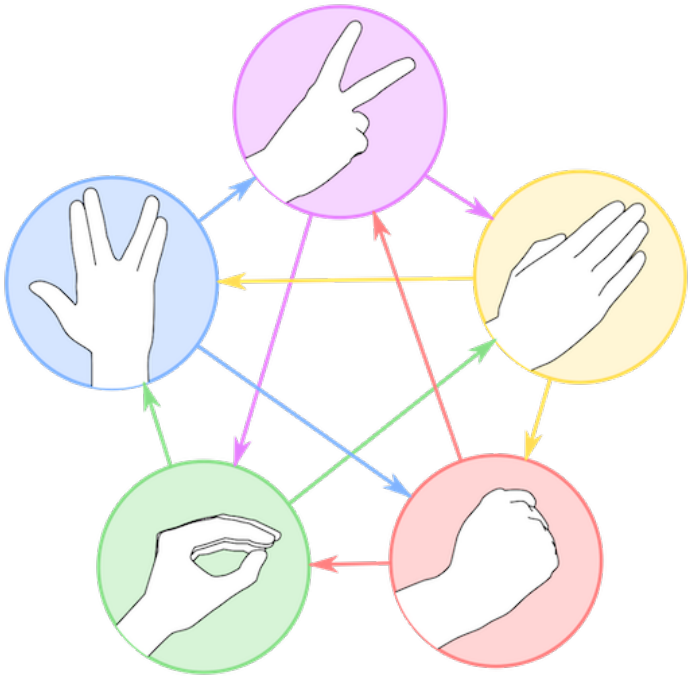


$$[\![\bullet]\!] = \text{Stop}$$

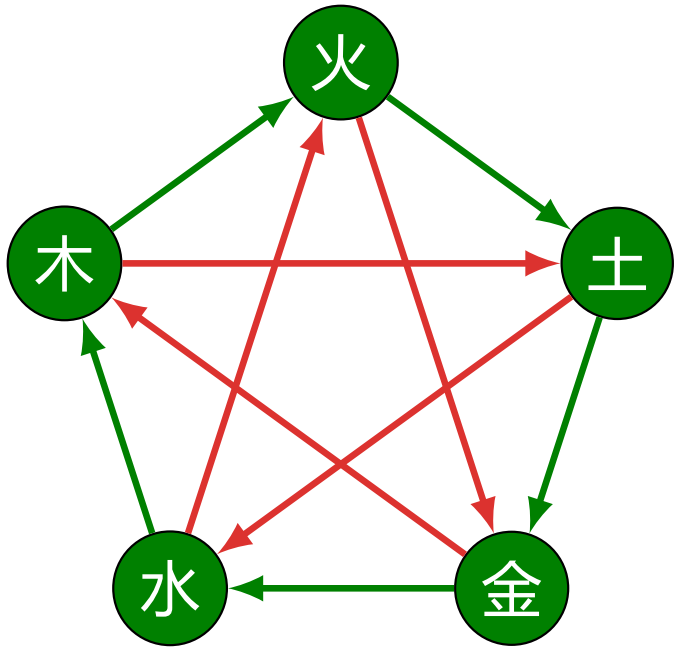
$$[\![\bullet]\!] = \text{Caution}$$

$$[\![\bullet]\!] = \text{Go}$$

Example — 结构



同构吗?
$\cong$



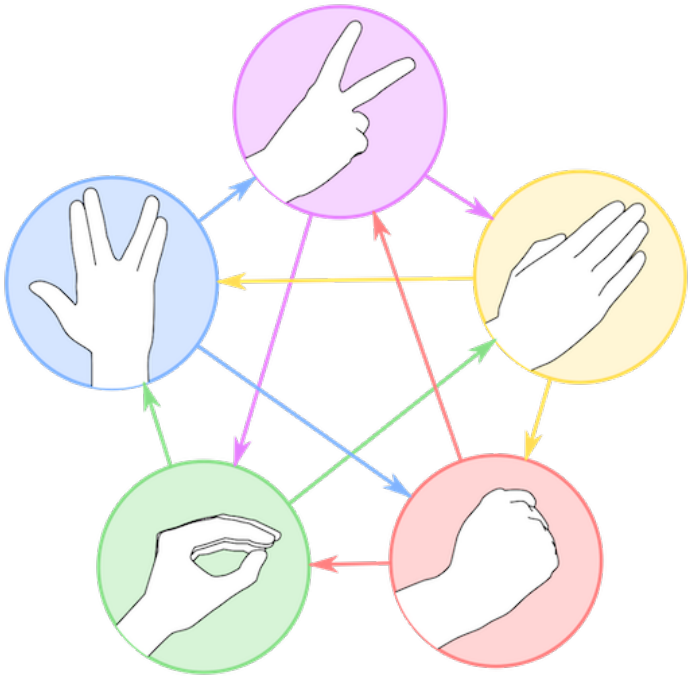
$$M = \{金, 水, 木, 火, 土\}$$

$$c^{\mathcal{M}}_{金} = 金, \quad c^{\mathcal{M}}_{水} = 水, \quad c^{\mathcal{M}}_{木} = 木, \quad c^{\mathcal{M}}_{火} = 火, \quad c^{\mathcal{M}}_{土} = 土$$

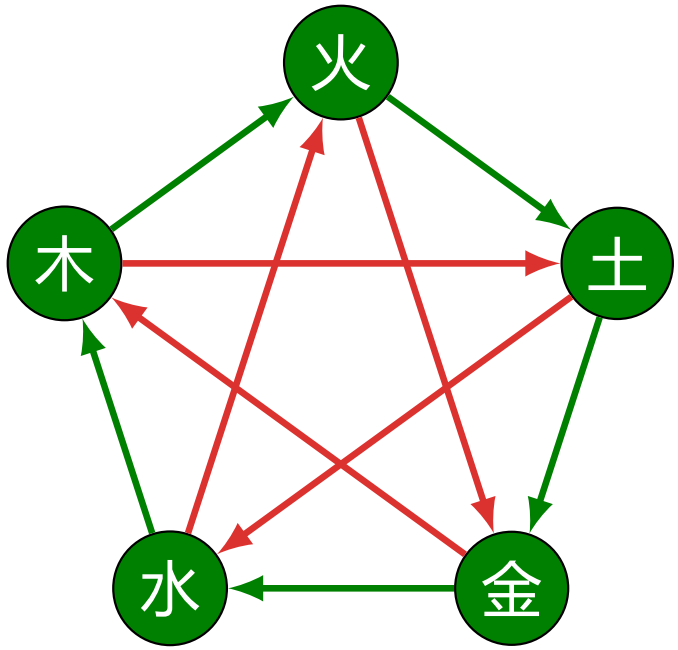
$$R^{\mathcal{M}}_{生} = \{(金, 水), (水, 木), (木, 火), (火, 土), (土, 金)\}$$

$$R^{\mathcal{M}}_{克} = \{(金, 木), (木, 土), (土, 水), (水, 火), (火, 金)\}$$

Example — 结构



同构吗?
$\cong$

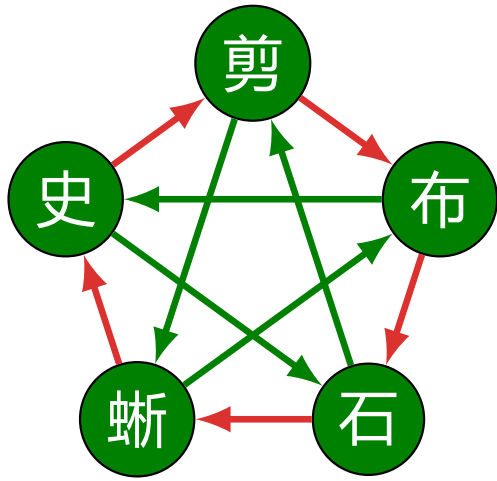


$$M = \{金, 水, 木, 火, 土\}$$

$$c^{\mathcal{M}}_{金} = 金, \quad c^{\mathcal{M}}_{水} = 水, \quad c^{\mathcal{M}}_{木} = 木, \quad c^{\mathcal{M}}_{火} = 火, \quad c^{\mathcal{M}}_{土} = 土$$

$$R^{\mathcal{M}}_{生} = \{(金, 水), (水, 木), (木, 火), (火, 土), (土, 金)\}$$

$$R^{\mathcal{M}}_{克} = \{(金, 木), (木, 土), (土, 水), (水, 火), (火, 金)\}$$

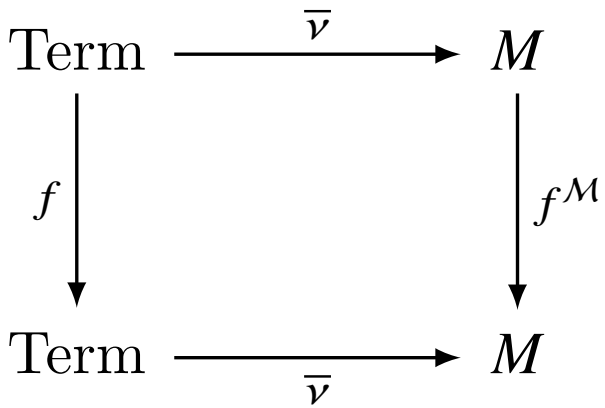


项的解释: “名” 副其 “实” ♡

Definition (变元赋值)

结构 $\mathcal{M}$ 上的变元赋值是一个函数 $\nu : \mathrm{Var} \to M$.
变元赋值 ν 可以递归地扩展到所有项上 $\overline{\nu} : \mathrm{Term} \to M$:

- $\overline{\nu}(x) := \nu(x)$
- $\overline{\nu}(c) := c^{\mathcal{M}}$
- $\overline{\nu}(f(t_1, \ldots, t_n)) := f^{\mathcal{M}}(\overline{\nu}(t_1), \ldots, \overline{\nu}(t_n))$



Remark: 在不引起歧义的情况下, 我们把 $\overline{\nu}$ 写作 ν.

公式的解释: "满足关系" ♡

Definition (满足关系 $\mathcal{M}, \nu \vDash A$)

- $\mathcal{M}, \nu \vDash t_1 = t_2$ iff $\nu(t_1) = \nu(t_2)$
- $\mathcal{M}, \nu \vDash P(t_1, \ldots, t_n)$ iff $(\nu(t_1), \ldots, \nu(t_n)) \in P^{\mathcal{M}}$
- $\mathcal{M}, \nu \vDash \neg A$ iff $\mathcal{M}, \nu \nvDash A$
- $\mathcal{M}, \nu \vDash A \wedge B$ iff $\mathcal{M}, \nu \vDash A$ and $\mathcal{M}, \nu \vDash B$
- $\mathcal{M}, \nu \vDash A \vee B$ iff $\mathcal{M}, \nu \vDash A$ or $\mathcal{M}, \nu \vDash B$
- $\mathcal{M}, \nu \vDash A \to B$ iff $\mathcal{M}, \nu \nvDash A$ or $\mathcal{M}, \nu \vDash B$
- $\mathcal{M}, \nu \vDash A \leftrightarrow B$ iff $\mathcal{M}, \nu \vDash A \iff \mathcal{M}, \nu \vDash B$
- $\mathcal{M}, \nu \vDash \forall x A$ iff for every $a \in M : \mathcal{M}, \nu(a/x) \vDash A$

$$\text{where} \quad \nu(a/x)(y) := \begin{cases} a & \text{if } y = x \\ \nu(y) & \text{otherwise} \end{cases}$$

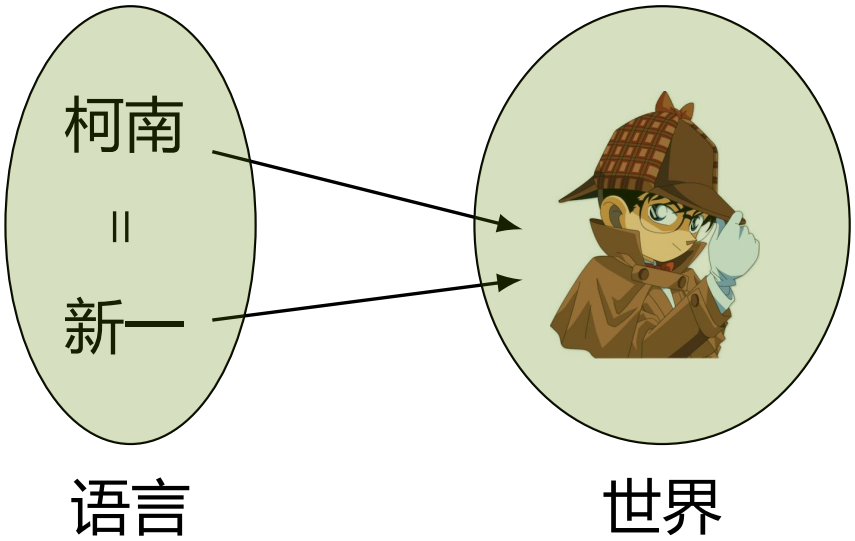
- $\mathcal{M}, \nu \vDash \exists x A$ iff there is an $a \in M : \mathcal{M}, \nu(a/x) \vDash A$

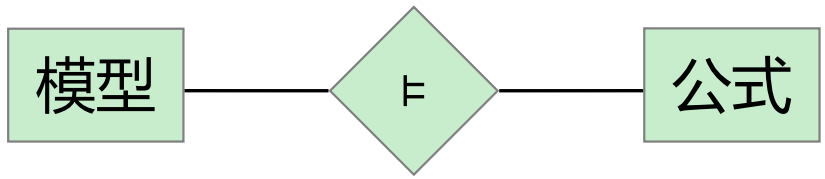
Remark: $\nu(a/x)$ 就是改变 ν 对 x 的赋值为 a, 其它不变.

$\mathcal{M}, \nu \vDash \forall x A$ iff for all $\nu' \sim_x \nu : \mathcal{M}, \nu' \vDash A$. where $\nu' \sim_x \nu$ iff for all $y \neq x : \nu'(y) = \nu(y)$.

公式的解释: 从 “满足关系” 到 “真” ♡

- $\mathcal{M}, \nu \vDash t_1 = t_2$ iff $\nu(t_1) = \nu(t_2)$





- 满足关系 $\mathcal{M}, \nu \vDash A$
  Remark: 先把符号串 A 里的常元符号、函数符号、谓词符号按照结构 $\mathcal{M}$ 的规定来解释, 把量词的论域限定在集合 M 上, 把自由变元 x 解释成它的赋值 $\nu(x)$, 从而把公式 A 翻译成一个元语言中关于结构 $\mathcal{M}$ 的命题, 而利用结构 $\mathcal{M}$ 的知识, 就可以知道命题 A 是否成立.
- $\mathcal{M} \vDash A$ 当且仅当, 对任意变元赋值 ν 都有: $\mathcal{M}, \nu \vDash A$. (真)
  Remark: 此时, 我们说, A 在 $\mathcal{M}$ 上为真, 或 $\mathcal{M}$ 是 A 的模型.

Notation

当项 t 和公式 A 中的自由变元都在 $\{x_1, \ldots, x_n\}$ 里时, 如果变元赋值 $v(x_i) = a_i$ for $1 \leq i \leq n$, 那么[12], 我们有时会把

$$v(t) \quad \text{写成} \quad t^{\mathcal{M}}[a_1, \ldots, a_n]$$

$$\mathcal{M}, v \vDash A \quad \text{写成} \quad \mathcal{M} \vDash A[a_1, \ldots, a_n]$$

Remark: When $\mathrm{Fv}(A) = \{x\}$,

$$\mathcal{M} \vDash \forall x A \iff \text{for every } a \in M : \mathcal{M} \vDash A[a]$$

$$\mathcal{M} \vDash \exists x A \iff \text{for some } a \in M : \mathcal{M} \vDash A[a]$$

---

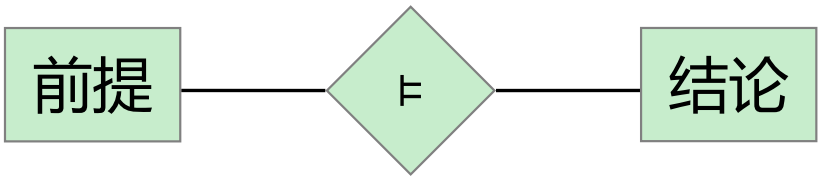
[12]这一记号的合理性由 Coincidence Lemma 保证.

Remark: 借助 Substitution Lemma, $\mathcal{M} \vDash A[a_1, \ldots, a_n]$ 相当于: 把论域 M 里的个体 a_i 命名为 a_i, 然后把名字 a_i 代入到 $A(x_1, \ldots, x_n)$ 中的自由变元 x_i, 得到的句子 $A(a_1, \ldots, a_n)$ 在结构 $\mathcal{M}$ 上是真的.

Remark: 无名, 天地之始; 有名, 万物之母. — 老子 ☺

可满足、逻辑蕴含、有效式 ♡

- $\mathcal{M}, \nu \vDash \Gamma$ iff for all $A \in \Gamma : \mathcal{M}, \nu \vDash A$.
- Γ is satisfiable iff there exists $\mathcal{M}, \nu$ s.t. $\mathcal{M}, \nu \vDash \Gamma$. (可满足)
- $\mathcal{M} \vDash \Gamma$ iff for all $A \in \Gamma : \mathcal{M} \vDash A$.
- $\Gamma \vDash A$ iff for all $\mathcal{M}, \nu : \mathcal{M}, \nu \vDash \Gamma \implies \mathcal{M}, \nu \vDash A$. (逻辑蕴含)[13]



- A is valid $\vDash A$ iff $\emptyset \vDash A$. (有效式)

说是者为非，或非者为是，是假的；说是者为是，或非者为非，是真的.
— 亚里士多德

[13] If we define $\Gamma \vDash^* A$ iff for all $\mathcal{M} : \mathcal{M} \vDash \Gamma \implies \mathcal{M} \vDash A$, then we have $Px \vDash^* \forall x Px$.

- 逻辑蕴含 $\vDash$ 是 “保赋值” 的，而 $\vDash^*$ 是 “保真” 或 “保模型” 的，我们还可以定义 “保有效”: $\vDash \Gamma \implies \vDash A$.
- 显然，“保赋值” 强于 “保真” 强于 “保有效”.

Interpretation of Formulas — another version

- $\nu(t_1 = t_2) := \begin{cases} 1 & \text{if } \nu(t_1) = \nu(t_2) \\ 0 & \text{otherwise} \end{cases}$
- $\nu(P(t_1, \ldots, t_n)) := \begin{cases} 1 & \text{if } \big(\nu(t_1), \ldots, \nu(t_n)\big) \in P^{\mathcal{M}} \\ 0 & \text{otherwise} \end{cases}$
- $\nu(\neg A) := 1 - \nu(A)$
- $\nu(A \wedge B) := \min\big\{\nu(A), \nu(B)\big\}$
- $\nu(A \vee B) := \max\big\{\nu(A), \nu(B)\big\}$
- $\nu(A \to B) := \max\big\{1 - \nu(A), \nu(B)\big\}$
- $\nu(A \leftrightarrow B) := 1 - \big|\nu(A) - \nu(B)\big|$
- $\nu(\forall x A) := \min\limits_{a \in M} \big\{\nu(a/x)(A)\big\}$
- $\nu(\exists x A) := \max\limits_{a \in M} \big\{\nu(a/x)(A)\big\}$

$$\mathcal{M}, \nu \vDash A \iff \nu(A) = 1$$

Interpretation of Formulas — set-based version

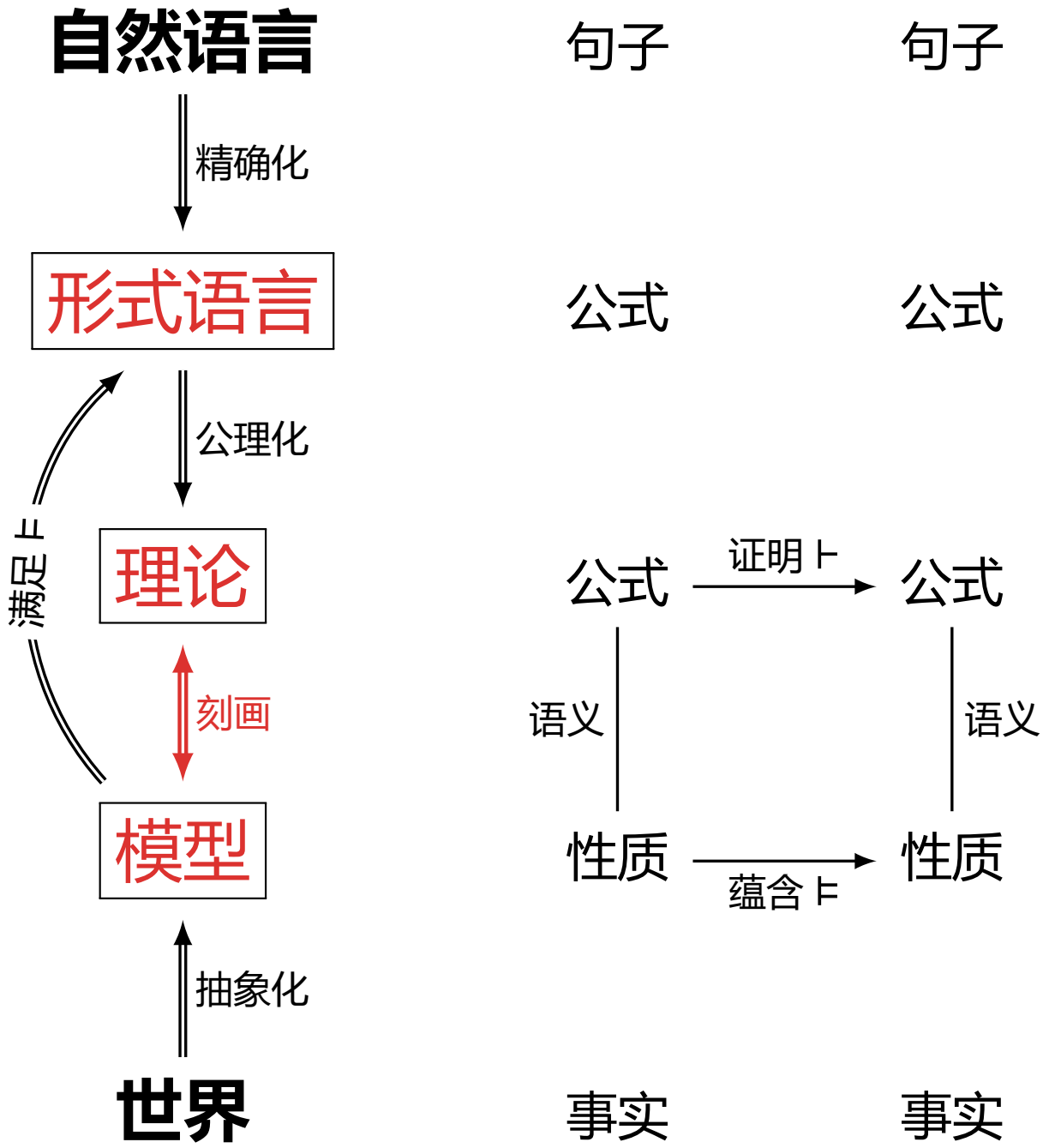
Let $[\![\,]\!]$ map atomic formulas to variable assignments $\mathtt{P}(M^{\mathrm{Var}})$.

- $[\![t_1 = t_2]\!] := \{\nu : \nu(t_1) = \nu(t_2)\}$
- $[\![P(t_1, \ldots, t_n)]\!] := \{\nu : (\nu(t_1), \ldots, \nu(t_n)) \in P^{\mathcal{M}}\}$

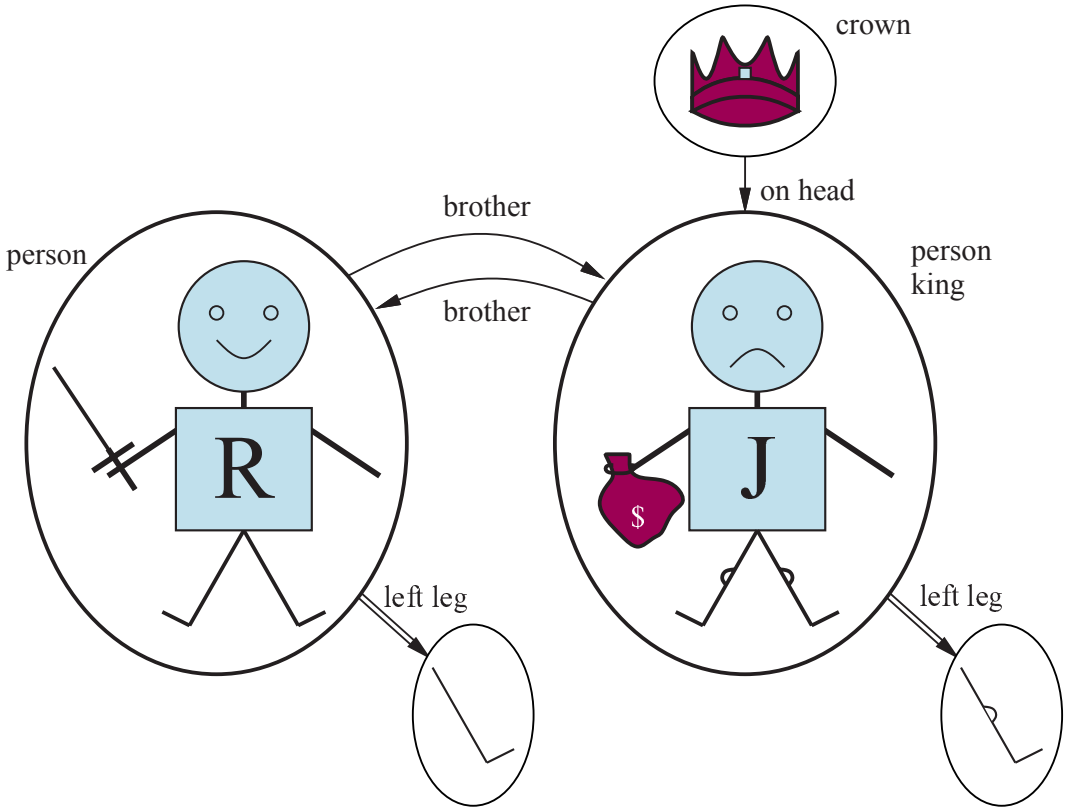
We extend $[\![\,]\!]$ to all formulas by recursion as follows:

1. $[\![\neg A]\!] := \overline{[\![A]\!]}$
2. $[\![A \wedge B]\!] := [\![A]\!] \cap [\![B]\!]$
3. $[\![A \vee B]\!] := [\![A]\!] \cup [\![B]\!]$
4. $[\![A \to B]\!] := \overline{[\![A]\!]} \cup [\![B]\!]$
5. $[\![A \leftrightarrow B]\!] := \overline{[\![A]\!] \Delta [\![B]\!]}$
6. $[\![\forall x A]\!] := \bigcap_{a \in M} \{\nu : \nu(a/x) \in [\![A]\!]\}$
7. $[\![\exists x A]\!] := \bigcup_{a \in M} \{\nu : \nu(a/x) \in [\![A]\!]\}$

$$\mathcal{M}, \nu \vDash A \iff \nu \in [\![A]\!]$$


自然语言
精确化
形式语言
公理化
理论
满足
刻画
模型
抽象化
世界
句子
句子
公式
公式
公式
证明⊢
公式
语义
语义
性质
蕴含⊢
性质
事实
事实

Example — 模型



1. $\mathsf{Brother}(r, j)$
2. $\neg\mathsf{King}(r) \rightarrow \mathsf{King}(j)$
3. $\mathsf{King}(j) \wedge \exists x \big(\mathsf{Crown}(x) \wedge \mathsf{OnHead}(x, j)\big)$
4. $\exists x(\mathsf{Person}(x) \wedge \exists y(\mathsf{Crown}(y) \wedge \mathsf{OnHead}(y, x)))$
5. $\neg\mathsf{Brother}\big(\mathsf{leftLeg}(r), j\big)$
6. $\forall x \big(\mathsf{King}(x) \rightarrow \mathsf{Person}(x)\big)$

Example — 模型

$$M = \{\text{舔狗}, \text{男神}, \text{女神}, \text{剩女}\}$$

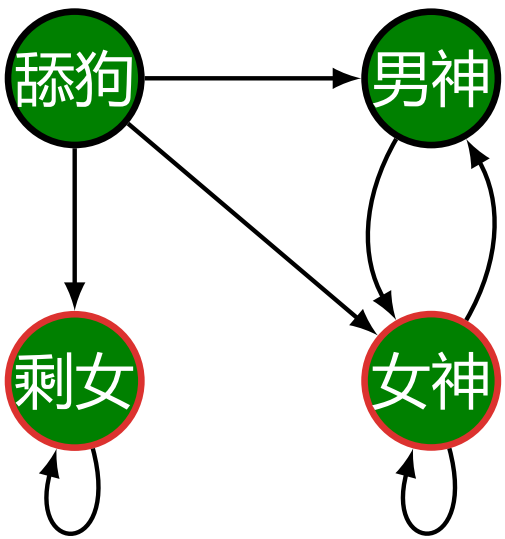
$$c^{\mathcal{M}} = \text{女神}$$

$$B^{\mathcal{M}} = \{\text{舔狗}, \text{男神}\}$$

$$G^{\mathcal{M}} = \{\text{剩女}, \text{女神}\}$$

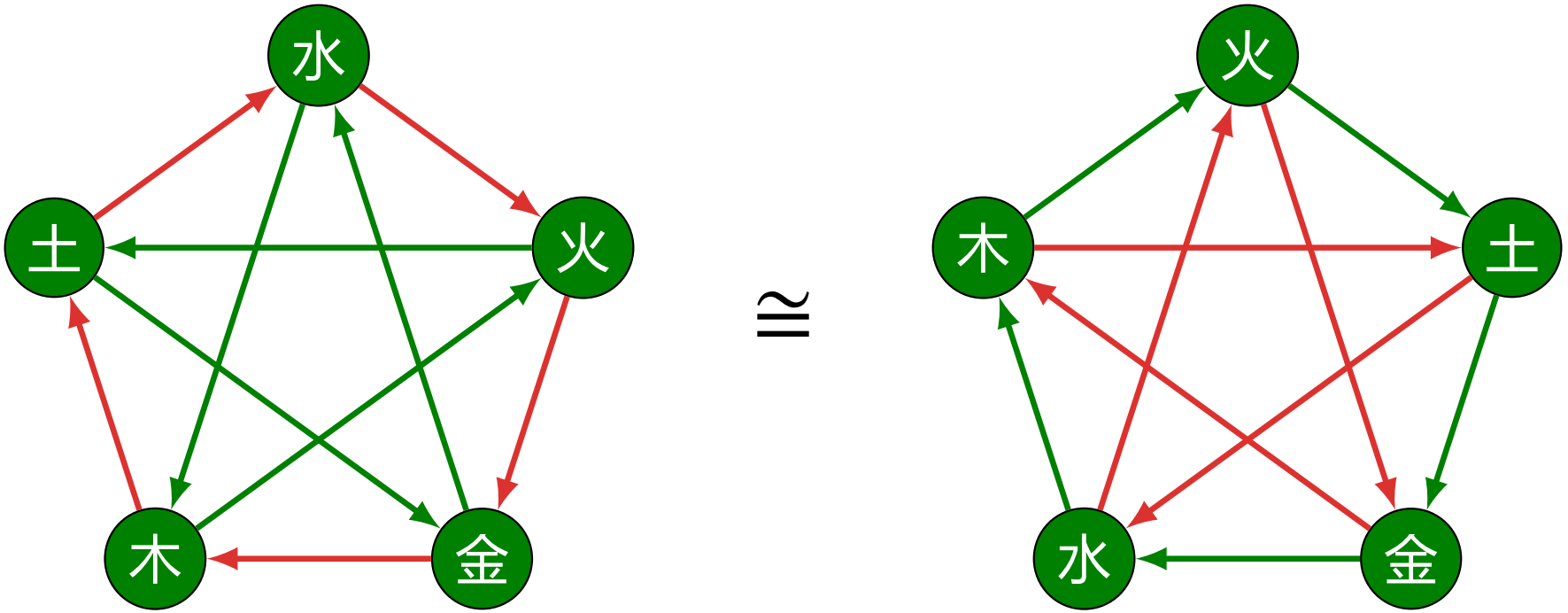
$$L^{\mathcal{M}} = \{(\text{舔狗}, \text{男神}), (\text{舔狗}, \text{女神}), (\text{女神}, \text{女神}) \cdots\cdots\}$$

$(M, c^{\mathcal{M}}, B^{\mathcal{M}}, G^{\mathcal{M}}, L^{\mathcal{M}})$



1. $\mathcal{M} \overset{?}{\models} Bc$
2. $\mathcal{M} \overset{?}{\models} Bc \vee Lcc$
3. $\mathcal{M} \overset{?}{\models} \exists x \neg Lxc$
4. $\mathcal{M} \overset{?}{\models} \forall x(Bx \vee Lxx)$
5. $\mathcal{M} \overset{?}{\models} \forall x \exists y Lxy$
6. $\mathcal{M} \overset{?}{\models} \exists x \forall y(y = x \vee Lxy)$
7. $\mathcal{M} \overset{?}{\models} \exists x(Bx \wedge \forall y(Gy \rightarrow \neg Lyx))$

Example — 模型



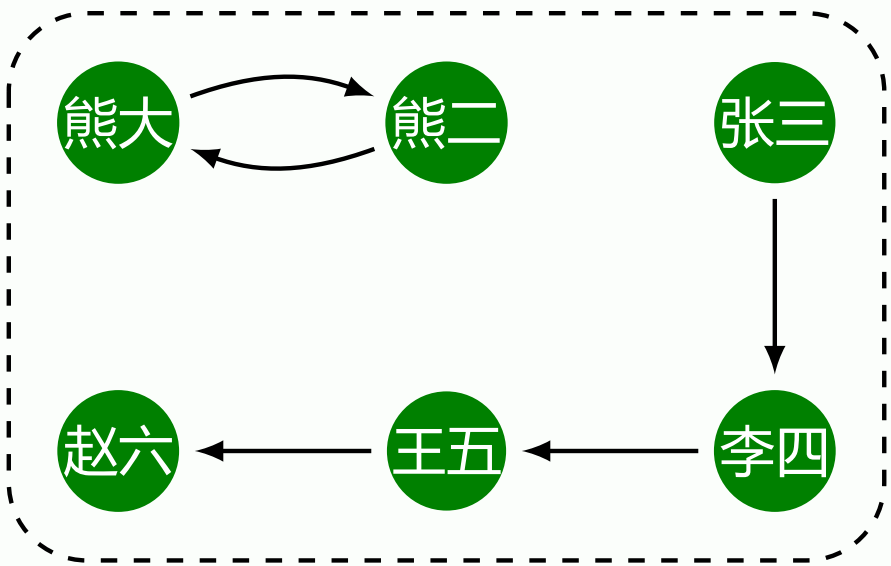
1. $\forall x \exists y R_{生}(x, y)$
2. $\forall x \exists y R_{克}(x, y)$
3. $\forall x \exists y R_{生}(y, x)$
4. $\forall x \exists y R_{克}(y, x)$
5. $\forall x \exists y_1 y_2 y_3 y_4 \left[R_{生}(x, y_1) \wedge R_{生}(y_1, y_2) \wedge R_{生}(y_2, y_3) \wedge R_{生}(y_3, y_4) \wedge R_{生}(y_4, x) \right]$
6. $\forall x \exists y_1 y_2 y_3 y_4 \left[R_{克}(x, y_1) \wedge R_{克}(y_1, y_2) \wedge R_{克}(y_2, y_3) \wedge R_{克}(y_3, y_4) \wedge R_{克}(y_4, x) \right]$

Example — 模型

Transitive Relation

$$\forall xyz(Rxy \wedge Ryz \rightarrow Rxz)$$

在下图上为真吗? 如果不为真, 添加尽量少的箭头使之为真.



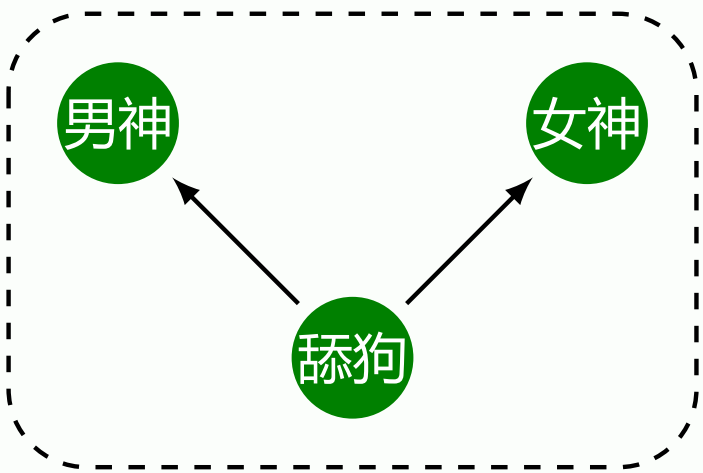
如果是删箭头呢?

Example — 模型

Euclidean Relation

$$\forall xyz(Rxy \wedge Rxz \rightarrow Ryz)$$

在下图上为真吗? 如果不为真, 添加尽量少的箭头使之为真.



Example — 模型

$\exists x \forall y Rxy$

R	1	2	3	4
1	✓	✓	✓	✓
2				
3				
4				

$\forall y \exists x Rxy$

R	1	2	3	4
1	✓	✓		
2			✓	
3				✓
4				

$\exists y \forall x Rxy$

R	1	2	3	4
1				✓
2				✓
3				✓
4				✓

$\forall x \exists y Rxy$

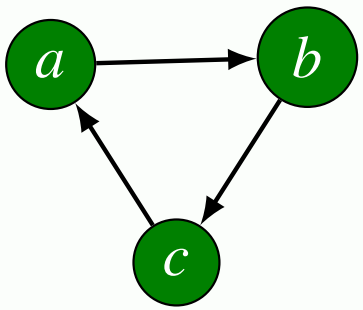
R	1	2	3	4
1		✓		
2			✓	
3				✓
4				✓

Example — 反模型 ♡

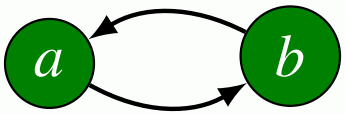
若 $A \not\vdash B$, 则存在反模型$\mathcal{M}$ 使得 $\mathcal{M} \vDash A$ 但 $\mathcal{M} \not\vDash B$

构造反模型

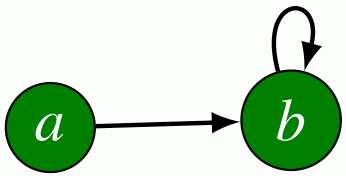
$\dfrac{\forall x \exists y Lxy}{\exists x \forall y Lxy}$ ×



$\dfrac{\forall x \exists y Lxy}{\exists y \forall x Lxy}$ ×



$\dfrac{\exists y \forall x Lxy}{\forall y \exists x Lxy}$ ×



$$\cdots \to -2 \to -1 \to 0 \to 1 \to 2 \to 3 \to 4 \to 5 \to \cdots$$

$$(\mathbb{Z}, <)$$

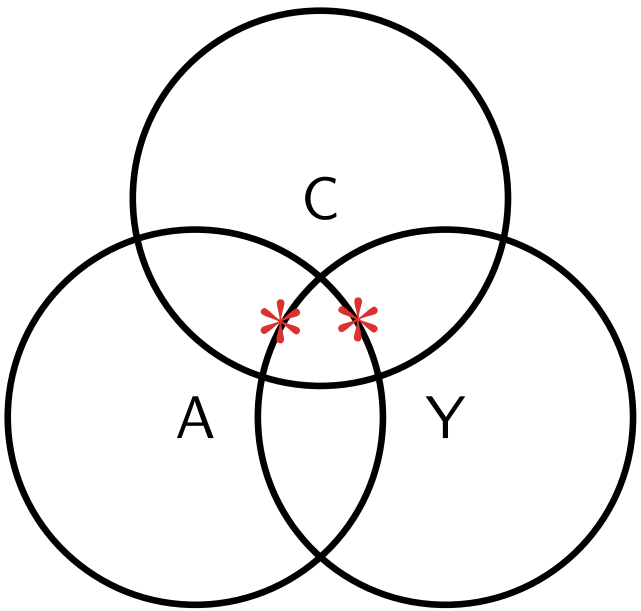
Example — 反模型 ♡

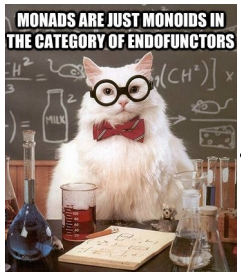
有些动物是猫
有些猫是黄色的

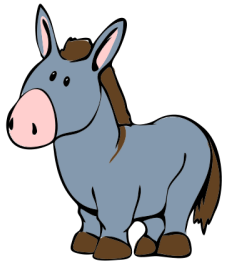
有些动物是黄色的

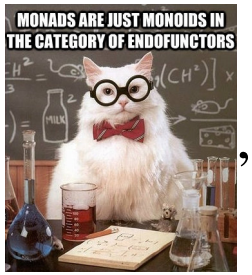
$$\exists x(Ax \wedge Cx)$$
$$\exists x(Cx \wedge Yx)$$

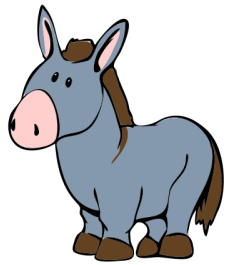
$$\exists x(Ax \wedge Yx)$$

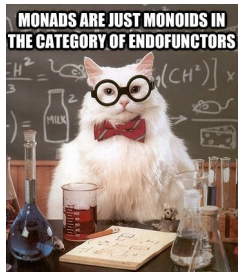


$$M = \left\{ \text{  }, \text{ 

}, \text{  } \right\}$$

$$A^M = \left\{ \text{ 

}, \text{  } \right\}$$

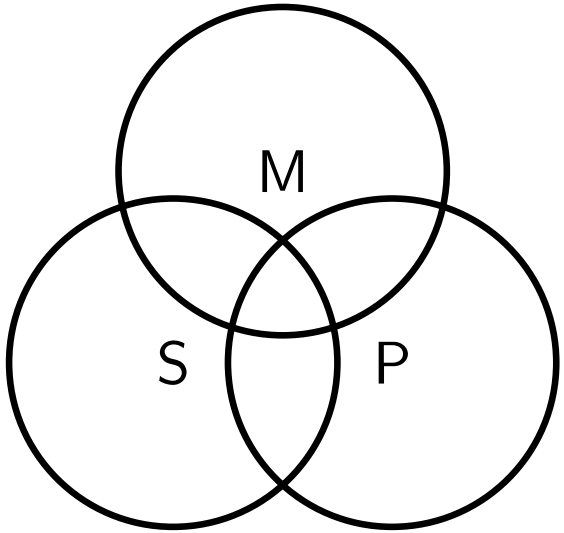
$$C^M = \left\{ \text{  }, \text{ 

} \right\}$$

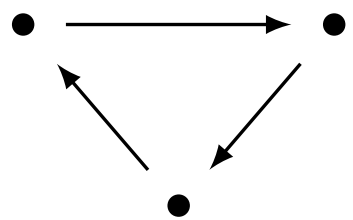
$$Y^M = \left\{ \text{  } \right\}$$

Remark

- 只含有一元谓词的公式集对应欧拉图



- 只含有一个二元谓词的公式集对应有向图



Theorems come and theorems go, but the example remains forever. Explain this to me on a simple example; the difficult example I will be able to do on my own. — *Israel. M. Gelfand*

Example — 反模型

$$\frac{\begin{array}{c}\text{Everybody loves somebody}\\ \text{Everybody loves all persons who are loved by his loved ones}\end{array}}{\text{There is at least a pair of persons who love each other}}$$

$$\frac{\begin{array}{c}\text{万物皆有因}\\ \text{原因的原因也是原因}\end{array}}{\text{某物是自身的原因}}$$

$$\frac{\begin{array}{c}\forall x\exists y Rxy\\ \forall xyz(Rxy \wedge Ryz \to Rxz)\end{array}}{\exists xy(Rxy \wedge Ryx)} \times \qquad \frac{\begin{array}{c}\forall x\exists y Ryx\\ \forall xyz(Rxy \wedge Ryz \to Rxz)\end{array}}{\exists x Rxx} \times$$

$$\cdots \to -2 \to -1 \to 0 \to 1 \to 2 \to 3 \to 4 \to 5 \to \cdots$$

$$(\mathbb{Z}, <)$$

Remark: 存在有穷反模型吗?

Problem (写一个只能被无穷结构满足的公式)

$$\forall x \exists y Rxy \ \wedge \ \forall xyz(Rxy \wedge Ryz \rightarrow Rxz) \ \wedge \ \forall x \neg Rxx$$

- 假设有满足它的有穷结构 $\mathcal{M}$, 不妨假设 $|M| = n$.
- 由 $\mathcal{M} \vDash \forall x \exists y Rxy$, 存在序列 $(a_1, a_2, \ldots, a_{n+1})$ 使得对任意 $1 \leq i \leq n+1$ 都有 $R^{\mathcal{M}} a_i a_{i+1}$.
- 又由 $\mathcal{M} \vDash \forall xyz(Rxy \wedge Ryz \rightarrow Rxz)$, 对任意 $1 \leq i < j \leq n+1$ 都有 $R^{\mathcal{M}} a_i a_j$.
- 因为 $|M| = n$, 所以存在 $1 \leq i < j \leq n+1$ 使得 $a_i = a_j$, 这意味着 $R^{\mathcal{M}} a_i a_i$. 这与 $\mathcal{M} \vDash \forall x \neg Rxx$ 矛盾.

Examples — 模型 & 反模型

$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$

1. $\forall x \exists y : x + y = 0$

$$(\mathbb{N}, 0, 1, +, \cdot, <) \not\models \forall x \exists y : x + y = 0$$
$$(\mathbb{Z}, 0, 1, +, \cdot, <) \models \forall x \exists y : x + y = 0$$

2. $\forall x \neq 0 \exists y : x \cdot y = 1$

$$(\mathbb{Z}, 0, 1, +, \cdot, <) \not\models \forall x \neq 0 \exists y : x \cdot y = 1$$
$$(\mathbb{Q}, 0, 1, +, \cdot, <) \models \forall x \neq 0 \exists y : x \cdot y = 1$$

3. $\forall x \geq 0 \exists y : x = y \cdot y$

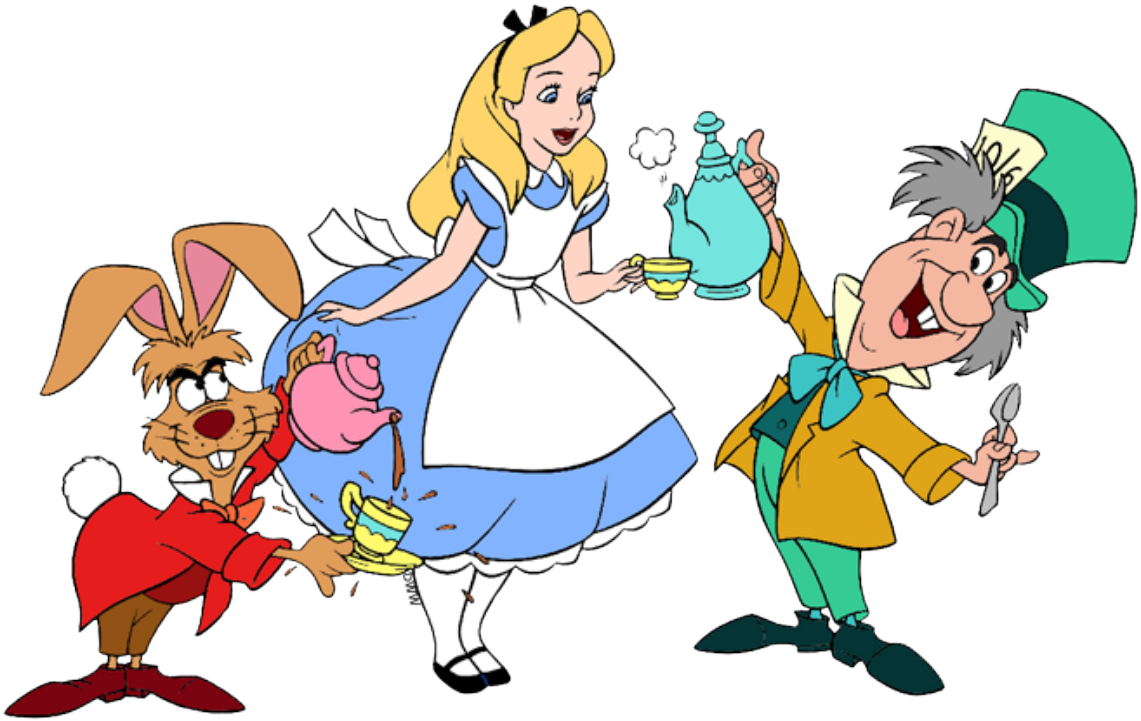
$$(\mathbb{Q}, 0, 1, +, \cdot, <) \not\models \forall x \geq 0 \exists y : x = y^2$$
$$(\mathbb{R}, 0, 1, +, \cdot, <) \models \forall x \geq 0 \exists y : x = y^2$$

4. $\forall x \exists y : x = y \cdot y$

$$(\mathbb{R}, 0, 1, +, \cdot, <) \not\models \forall x \exists y : x = y^2$$
$$(\mathbb{C}, 0, 1, +, \cdot, <) \models \forall x \exists y : x = y^2$$

数学的本质在于它的自由. — 康托尔

爱丽丝梦游仙境 — Lewis Carroll



$$\frac{\forall x(A \to B)}{\forall x(B \to A)} \times$$

- 三月兔: 那你怎么想就怎么说.
- 爱丽丝: 我说的就是我想的 — 这是一回事.
- 疯帽匠: 不是一回事, 那样的话你可以说 — “凡是我吃的东西我都能看见” 和 “凡是我能看见的东西我都吃” 是一回事.
- 三月兔: 那样的话你可以说 — “凡是我拥有的东西我都喜欢” 和 “凡是我喜欢的东西我都拥有” 是一回事.

易犯的 “错误” ♡

$$\forall x(Bx \to Sx)$$

$$\exists x(Bx \wedge Sx)$$

- $\forall x(Bx \wedge Sx) \equiv \forall xBx \wedge \forall xSx$
  万物皆男孩, 并且, 万物皆聪明.
- $\exists x(Bx \to Sx) \equiv \exists x\neg Bx \vee \exists xSx$
  只要有不是男孩的东西, 这句话就为真.

下面这句话为真吗?

All the elephants in this room are purple.

$$\forall x(\mathsf{Elephant}(x) \wedge \mathsf{In}(x, \mathsf{room}) \to \mathsf{Purple}(x))$$

$$\forall x(\mathsf{Elephant}(x) \wedge \mathsf{In}(x, \mathsf{room}) \wedge \mathsf{Purple}(x))$$

Coincidence Lemma

- 赋值 $\nu_1, \nu_2 : \mathrm{Var} \to M$ 在项 t 上一致, 记为 $\nu_1 \equiv \nu_2 \pmod{t}$, 当且仅当, 对任意 $x \in \mathrm{Var}(t)$: $\nu_1(x) = \nu_2(x)$.
- 赋值 $\nu_1, \nu_2 : \mathrm{Var} \to M$ 在公式 A 上一致, 记为 $\nu_1 \equiv \nu_2 \pmod{A}$, 当且仅当, 对任意 $x \in \mathrm{Fv}(A)$: $\nu_1(x) = \nu_2(x)$.

Lemma (Coincidence Lemma)

- *If* $\nu_1 \equiv \nu_2 \pmod{t}$, *then*

$$\nu_1(t) = \nu_2(t)$$

- *If* $\nu_1 \equiv \nu_2 \pmod{A}$, *then*

$$\mathcal{M}, \nu_1 \vDash A \iff \mathcal{M}, \nu_2 \vDash A$$

Remark:

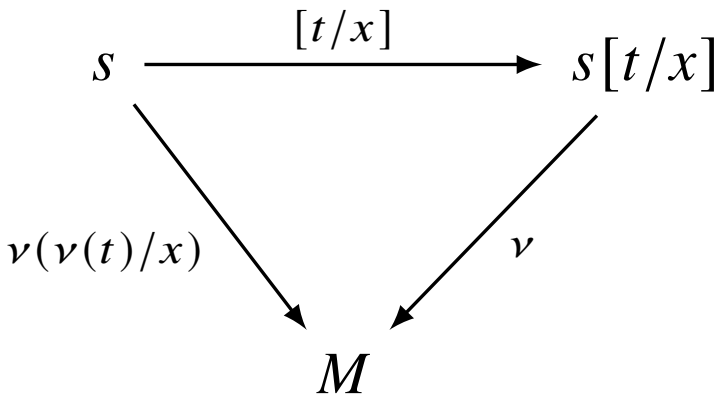
- If A is a closed formula, then either $\mathcal{M} \vDash A$ or $\mathcal{M} \vDash \neg A$.
- $\mathcal{M} \vDash A \implies \mathcal{M} \vDash \forall x A$
- $\mathcal{M}, \nu \vDash \forall x A$ iff for all $\nu' \equiv \nu \pmod{\forall x A} : \mathcal{M}, \nu' \vDash A$.
- $\mathcal{M}, \nu \vDash \exists x A$ iff for some $\nu' \equiv \nu \pmod{\exists x A} : \mathcal{M}, \nu' \vDash A$.

Substitution Lemma

Lemma (Substitution Lemma)

- $\nu\,(s[t/x]) = \nu(\nu(t)/x)(s)$
- *If the term t is substitutable for the variable x in the formula A, then*

$$\mathcal{M}, \nu \vDash A[t/x] \iff \mathcal{M}, \nu(\nu(t)/x) \vDash A$$



Remark:

$$\mathscr{L}_M := \mathscr{L} \cup \{c_a : a \in M\} \quad \text{and} \quad c_a^{\mathcal{M}_M} = a \quad \text{and} \quad \mathcal{M}_M := (\mathcal{M}, a)_{a \in M}$$

$$\mathcal{M}_M, \nu \vDash A[c_a/x] \iff \mathcal{M}, \nu(a/x) \vDash A$$

Equivalent Replacement

Theorem (等价替换定理)

假设 B 是 A 的子公式, $A[C/\!\!/B]$ 是用 C 替换 B 在 A 中的一些出现后得到的结果. 则

$$\frac{\Gamma \vDash B \leftrightarrow C}{\Gamma \vDash A \leftrightarrow A[C/\!\!/B]}$$

Alphabetic Variant

Definition (Alphabetic Variant)

If $y \notin \mathrm{Fv}(A)$, and y is substitutable for x in A, we say that $\forall y A[y/x]$ is an alphabetic variant of $\forall x A$.

Theorem

If $\forall y A[y/x]$ is an alphabetic variant of $\forall x A$, then

$$\vDash \forall x A \leftrightarrow \forall y A[y/x]$$

If $y \notin \mathrm{Fv}(A)$, then $A[y/x][x/y] = A$.

- **Convention**: When we write $A[t/x]$ we assume that t is substitutable for x in A. — *For any formula A and a finite number of variables $y_1, \ldots, y_n$ (occurring in t), we can always find a logically equivalent alphabetic variant A^* of A s.t. $y_1, \ldots, y_n$ do not occur bound in A^*.*

Equality & Equivalence

Lemma

Suppose $\mathrm{Var}(s) \cup \mathrm{Var}(t) \subset \{x_1, \ldots, x_n\}$, *and* A^* *arises from the formula* A *by replacing one occurrence of* s *in* A *by* t. *Then*

$$\vDash \forall x_1 \ldots x_n (s = t) \to (A \leftrightarrow A^*)$$

$$\mathcal{M} \vDash s = t \implies \mathcal{M} \vDash A \leftrightarrow A^*$$

Lemma

Suppose $\mathrm{Fv}(B) \cup \mathrm{Fv}(C) \subset \{x_1, \ldots, x_n\}$, *and* A^* *arises from the formula* A *by replacing one occurrence of* B *in* A *by* C. *Then*

$$\vDash \forall x_1 \ldots x_n (B \leftrightarrow C) \to (A \leftrightarrow A^*)$$

$$\mathcal{M} \vDash B \leftrightarrow C \implies \mathcal{M} \vDash A \leftrightarrow A^*$$

Remark:

- $\vDash \forall x \big(Px \leftrightarrow Qx\big) \to \big(\forall x Px \leftrightarrow \forall x Qx\big)$
  $\nvDash \big(Px \leftrightarrow Qx\big) \to \big(\forall x Px \leftrightarrow \forall x Qx\big)$
- $\mathcal{M}, \nu \vDash s = t \implies \mathcal{M}, \nu \vDash A \leftrightarrow A^*$
- $\mathcal{M}, \nu \vDash B \leftrightarrow C \implies \mathcal{M}, \nu \vDash A \leftrightarrow A^*$

$$B = Px, \quad C = Py, \quad A = \forall x Px, \quad A^* = \forall x Py$$

How to Check Validity? — Example

$$\boxed{Ay \to \forall x Ay \quad \text{where } x \neq y}$$

Proof.

Assume $\mathcal{M}, v \vDash Ay$.

Since for any $a \in M$, $v(a/x)(y) = v(y)$, then

$$\mathcal{M}, v(a/x) \vDash Ay$$

$$\mathcal{M}, v \vDash \forall x Ay$$

$$\mathcal{M}, v \vDash Ay \to \forall x Ay$$

□

How to Check Validity? — Example

$$\boxed{\exists x \forall y Rxy \rightarrow \forall y \exists x Rxy}$$

Proof.

Assume $\mathcal{M}, v \vDash \exists x \forall y Rxy$.
Then there exists $a \in M$ s.t.

$$\mathcal{M}, v(a/x) \vDash \forall y Rxy$$

For all $b \in M$,

$$\mathcal{M}, v(a/x)(b/y) \vDash Rxy$$

Since

$$v(a/x)(b/y) = v(b/y)(a/x)$$

it follows that

$$\mathcal{M}, v(b/y) \vDash \exists x Rxy$$

Therefore $\mathcal{M}, v \vDash \forall y \exists x Rxy$. □

How to Check Validity? — Example

$$\boxed{\forall x A \to A[t/x]}$$

$\mathcal{M}, \nu \vDash \forall x A \implies$ for all $a \in M : \mathcal{M}, \nu(a/x) \vDash A \implies \mathcal{M}, \nu(\nu(t)/x) \vDash A$
According to Substitution Lemma, $\mathcal{M}, \nu \vDash A[t/x]$.

$$\boxed{\forall x(B \to A) \to (\exists x B \to A) \quad \text{where } x \notin \mathrm{Fv}(A)}$$

Assume $\mathcal{M}, \nu \vDash \exists x B$ and $\mathcal{M}, \nu \nvDash A$. Then there exists $a \in M$ s.t. $\mathcal{M}, \nu(a/x) \vDash B$. According to Coincidence Lemma and $x \notin \mathrm{Fv}(A)$, we have $\mathcal{M}, \nu(a/x) \nvDash A$. Therefore $\mathcal{M}, \nu(a/x) \nvDash B \to A$. This contradicts $\mathcal{M}, \nu \vDash \forall x(B \to A)$.

$$\boxed{(\exists x B \to A) \to \forall x(B \to A) \quad \text{where } x \notin \mathrm{Fv}(A)}$$

$\mathcal{M}, \nu \vDash \exists x B \to A \implies \mathcal{M}, \nu \nvDash \exists x B$ or $\mathcal{M}, \nu \vDash A$.
If $\mathcal{M}, \nu \nvDash \exists x B$, then for all $a \in M$, $\mathcal{M}, \nu(a/x) \nvDash B$. It follows that $\mathcal{M}, \nu(a/x) \vDash B \to A$. Therefore $\mathcal{M}, \nu \vDash \forall x(B \to A)$.
If $\mathcal{M}, \nu \vDash A$, then according to Coincidence Lemma and $x \notin \mathrm{Fv}(A)$, for all $a \in M$, $\mathcal{M}, \nu(a/x) \vDash A$. It follows that $\mathcal{M}, \nu(a/x) \vDash B \to A$.
Therefore $\mathcal{M}, \nu \vDash \forall x(B \to A)$.

How to Check Validity? — Example

$$A[t/x] \leftrightarrow \forall x(x = t \to A) \quad \text{where } x \notin \mathrm{Var}(t)$$

$$\boxed{A[t/x] \to \forall x(x = t \to A) \quad \text{where } x \notin \mathrm{Var}(t)}$$

$\mathcal{M}, \nu \vDash A[t/x] \implies \mathcal{M}, \nu(\nu(t)/x) \vDash A$
Assume $\nu(t) = b$. Then for all $a \in M$, ether $a = b$ or $a \neq b$.
If $a = b$, then $\mathcal{M}, \nu(a/x) \vDash A$.
If $a \neq b$, then $\nu(a/x)(x) = a \neq b = \nu(t) = \nu(a/x)(t)$. So $\mathcal{M}, \nu(a/x) \nvDash x = t$.
Therefore we have $\mathcal{M}, \nu(a/x) \vDash x = t \to A$ for all $a \in M$.

$$\boxed{\forall x(x = t \to A) \to A[t/x] \quad \text{where } x \notin \mathrm{Var}(t)}$$

$\mathcal{M}, \nu \vDash \forall x(x = t \to A) \implies$ for all $a \in M : \mathcal{M}, \nu(a/x) \vDash x = t \to A$
Let $\nu(t) = b$. Then $\mathcal{M}, \nu(b/x) \vDash x = t \to A$.
$\nu(b/x)(x) = b = \nu(t) = \nu(b/x)(t) \implies \mathcal{M}, \nu(b/x) \vDash x = t$
Therefore $\mathcal{M}, \nu(b/x) \vDash A$. By Substitution Lemma, $\mathcal{M}, \nu \vDash A[t/x]$.

谓词逻辑里的命题逻辑

- 一个公式 A 的**命题形式**是一个命题逻辑公式 B, 使得 A 就是对 B 中相同的命题变元 p_i 处处以相同的公式 C_i 代入的结果.

$$A = B[C_1/p_1, \ldots, C_n/p_n]$$

- 公式的命题形式不唯一.

$$(\forall x Px \to \neg\exists y Qy) \vee \neg(\forall x Px \to \neg\exists y Qy)$$

$$(A \to B) \vee \neg(A \to B) \qquad A \vee \neg A$$

Theorem (重言式代入定理)

如果 B 是谓词逻辑公式 A 的命题形式, 则

$$\vDash B \implies \vDash A$$

Proof.

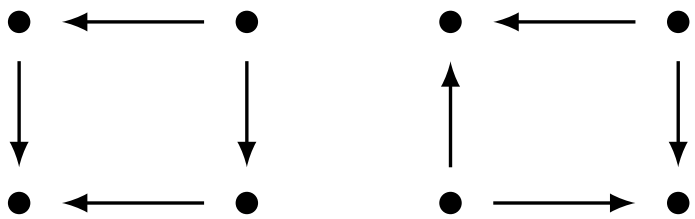
B 是 A 的命题形式, $A = B[C_1/p_1, \ldots, C_n/p_n]$. 任给 $\mathcal{M}, v$, 定义真值赋值 $\overline{v}$ 为 $\overline{v}(p_i) = 1 \iff \mathcal{M}, v \vDash C_i$. 归纳证明 $\mathcal{M}, v \vDash A \iff \overline{v} \vDash B$. □

练习: 模型 & 反模型 — Now it's your turn ⚡

1. 公式 $\exists x \forall y Rxy$ 在此图上为真吗? 如果不为真, 添加尽量少的箭头使之为真.



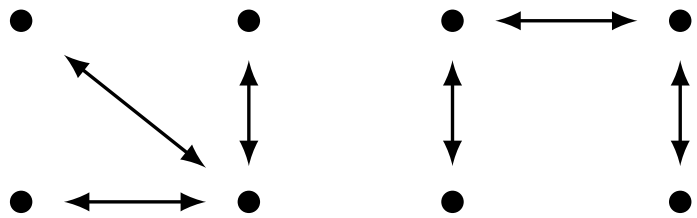
2. 写一个公式, 使其在左图上为真, 右图上为假.



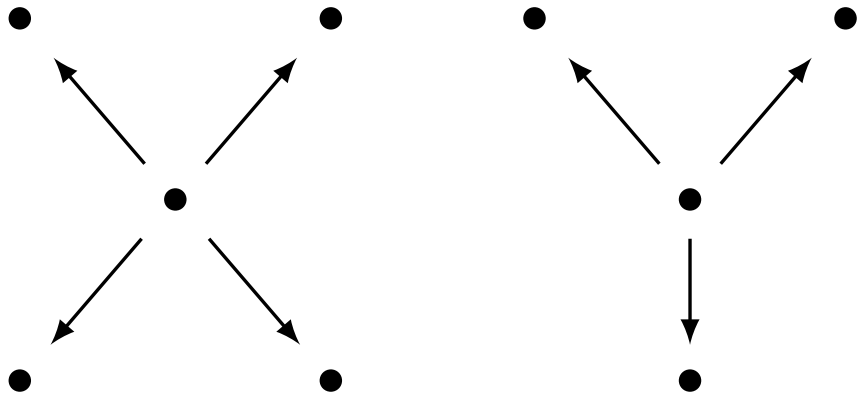
3. 画两幅图, 使公式在一幅图上为真, 另一幅图上为假.
   - 3.1 $\forall xy(Rxy \to Ryy)$
   - 3.2 $\forall x \exists y Rxy$
   - 3.3 $\forall xyz(Rxy \wedge Rxz \to Ryz \vee Rzy)$
   - 3.4 $\forall x(\exists y Ryx \to \forall z Rzx)$
   - 3.5 $\forall xyz(Rxy \wedge Rxz \to \exists w(Ryw \wedge Rzw))$

练习: 模型 & 反模型 — Now it's your turn ↯

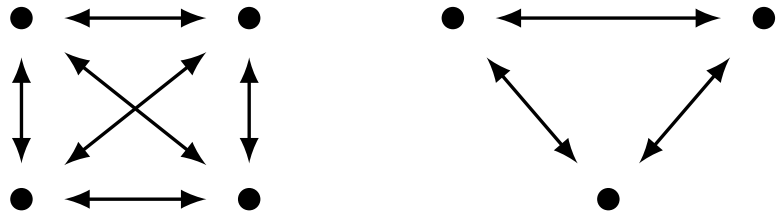
1. 写一个公式, 使其在左图上为真, 右图上为假.



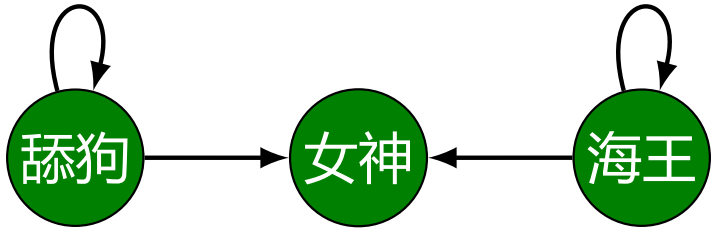
2. 写一个公式, 使其在左图上为真, 右图上为假.



3. 写一个公式, 使其在左图上为真, 右图上为假.



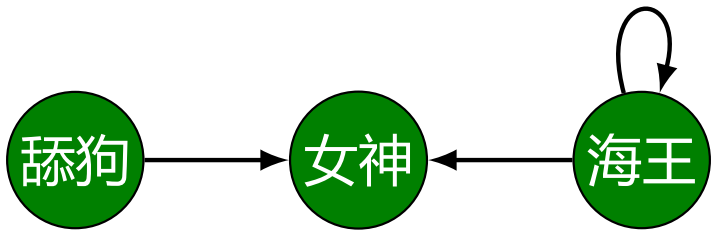
思考一下



假设 $\mathscr{L} = \{L\}$.

1. "女神" 可定义吗? $x =$ 女神 $\iff \mathcal{M} \vDash \forall y(y \neq x \to Lyx)$
2. $\{$舔狗, 海王$\}$ 可定义吗?
   $x \in \{$舔狗, 海王$\} \iff \mathcal{M} \vDash \neg\exists y(y \neq x \wedge Lyx)$
3. "舔狗" 可定义吗?

对称 $\implies$ 不可区分 $\implies$ 不可定义



现在的 $\mathcal{M}$ 有什么特点?

可定义 $\implies$ 可区分 $\implies$ 不对称

1. 是否存在非平凡的自同构?
2. 是否任意两个个体都可区分? $\mathcal{M} \vDash A[a] \ \& \ \mathcal{M} \nvDash A[b]$ $(\mathbb{R}^{\geq 0}, <)$
3. 是否每个个体都可定义? $x = a \iff \mathcal{M} \vDash A(x)$ $(\mathbb{R}, 0, 1, +, \cdot, <)$

- 能否找一个句子, 使其在一个结构上为真、另一个结构上为假?

思考一下 

Problem

- 一个图是完全的，当且仅当，图中任意两个不同节点之间至少有一个箭头.
- 一个节点是社牛，当且仅当，它可以通过至多两个箭头指到图中任何一个节点.
- 问题：怎么用公式表达“完全图”和“社牛节点”?

$$\mathsf{Complete}(G) \iff G \vDash A$$

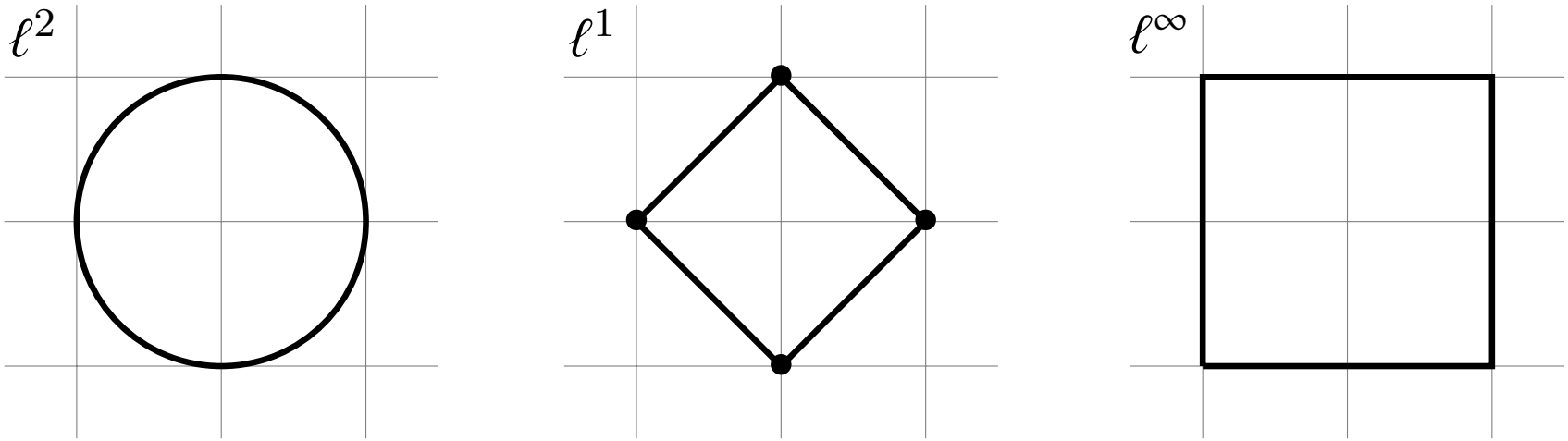
$$\mathsf{Social}(x) \iff G \vDash A(x)$$

- 证明：任何有穷的、自反的完全图都有一个社牛.
- 无穷图呢?

$(\mathbb{N}, \geq)$

$$\cdots \longrightarrow n+1 \longrightarrow n \longrightarrow n-1 \longrightarrow \cdots$$

什么是圆?



- 什么是圆? — 到定点 $\mathbf{c}$ 的距离等于定长 r 的点的集合.

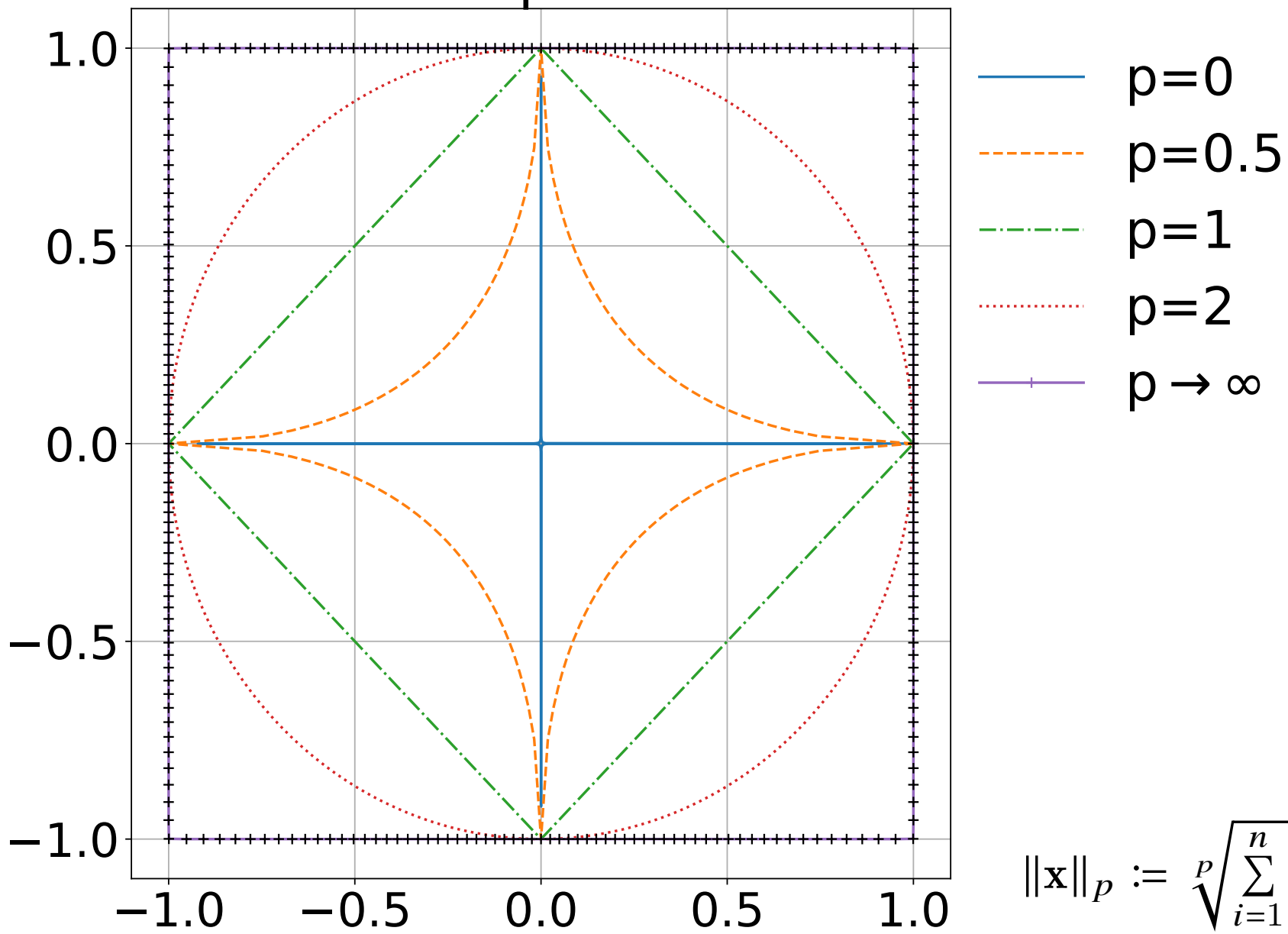
$$\{\mathbf{x} : \mathcal{M} \vDash d(\mathbf{x}, \mathbf{c}) = r\}$$

- 什么是半径为 1 的圆? — 基于什么距离? $\ell^2, \ell^1, \ell^\infty$ norm?

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_i |x_i - y_i|^2} \quad d_1(\mathbf{x}, \mathbf{y}) = \sum_i |x_i - y_i| \quad d_\infty(\mathbf{x}, \mathbf{y}) = \max_i |x_i - y_i|$$

- 什么是圆周率 π? 在 "出租车世界" $(\mathbb{R}^2, d_1)$, $\pi = \frac{\text{周长}}{\text{直径}} = \frac{8}{2} = 4$

"We are often interested not just in whether or not something is true, but in where it is true." — *John Baez*



$$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

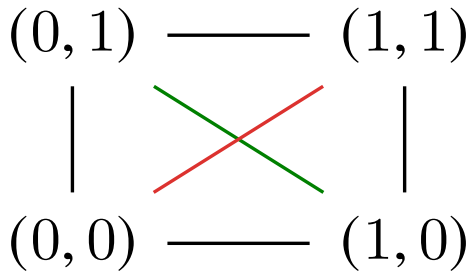
宇宙不仅比我们想象的更古怪，甚至比我们能够想象的更古怪.

$\mathbb{F}_2^2$ 平面上的圆

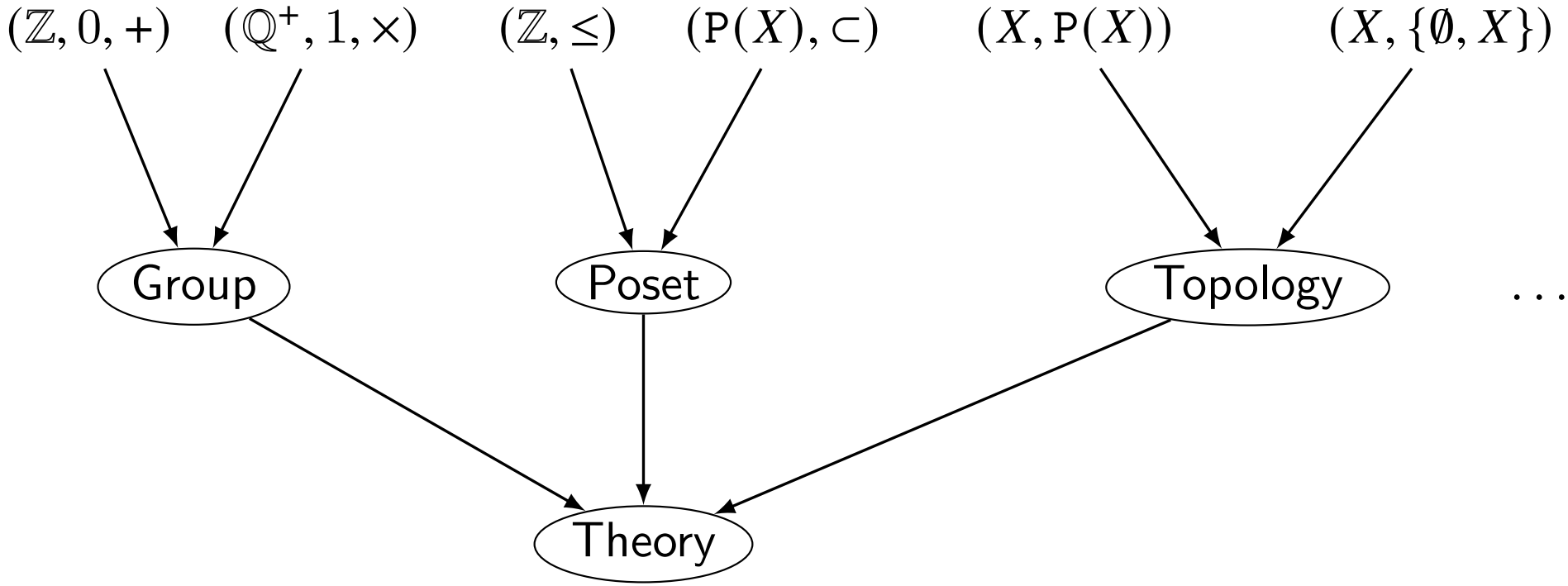
- $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z} = (\{0,1\}, 0, 1, +, \cdot)$

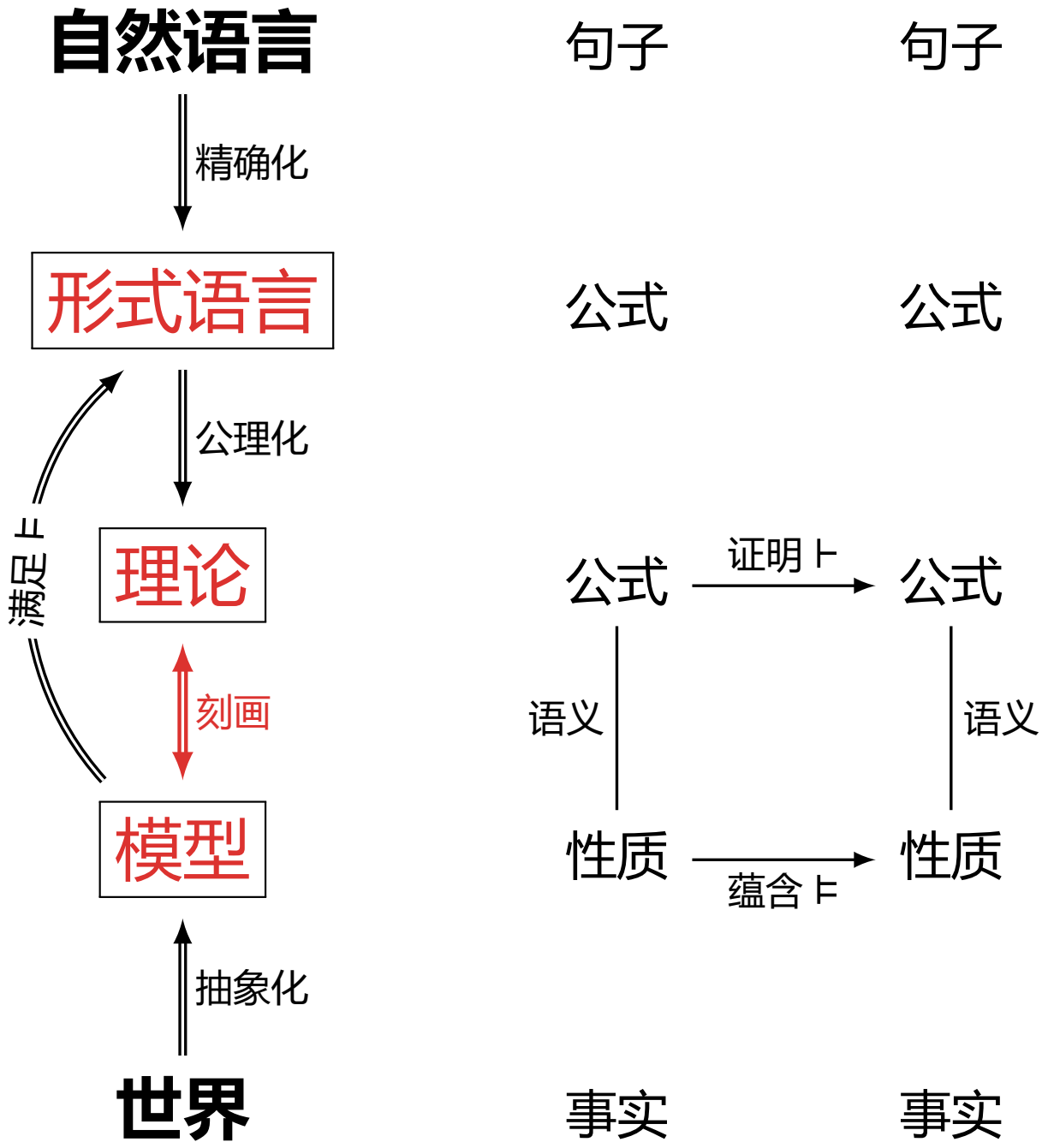
$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1



- $\mathbb{F}_2^2$ 平面只有四个点: $(0,0), (0,1), (1,0), (1,1)$
- $\mathbb{F}_2^2$ 平面有多少条直线? $ax + by + c = 0 \ \ (a, b, c \in \mathbb{F}_2)$, a, b 不同为 0
- $\mathbb{F}_2^2$ 平面有多少个圆? $(x-a)^2 + (y-b)^2 = r^2 \ \ (a, b, r \in \mathbb{F}_2)$
- 因为 $\forall x \in \mathbb{F}_2 : x^2 = x$, 所以
  $(x-a)^2 + (y-b)^2 = r^2 \iff x - a + y - b = r$
  即 $x + y = 0$ 或 $x + y = 1$.
  因此, 有且仅有 2 个圆, 都是直线!
- 例: $x + y = 0 \iff (x-0)^2 + (y-0)^2 = 0 \iff (x-1)^2 + (y-1)^2 = 0 \iff (x-0)^2 + (y-1)^2 = 1 \iff (x-1)^2 + (y-0)^2 = 1$
- 该圆周上有两个点: $(0,0), (1,1)$. 两个点都是圆心, 半径是 0.
- 该圆周外的两个点 $(0,1), (1,0)$ 也是圆心, 此时半径为 1.
- 因此, $\mathbb{F}_2^2$ 平面有 2 个圆, 每个圆都有 4 个圆心, 2 个半径.


$(\mathbb{Z}, 0, +)$
$(\mathbb{Q}^+, 1, \times)$
$(\mathbb{Z}, \leq)$
$(\mathtt{P}(X), \subset)$
$(X, \mathtt{P}(X))$
$(X, \{\emptyset, X\})$
Group
Poset
Topology
...
Theory


自然语言
精确化
形式语言
公理化
理论
满足
刻画
模型
抽象化
世界
句子
句子
公式
公式
公式
证明⊢
公式
语义
语义
性质
蕴含⊨
性质
事实
事实

Contents

Meta-Theorems

Introduction

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Term Logic

Recursion Theory

Propositional Logic

Equational Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Homotopy Type Theory

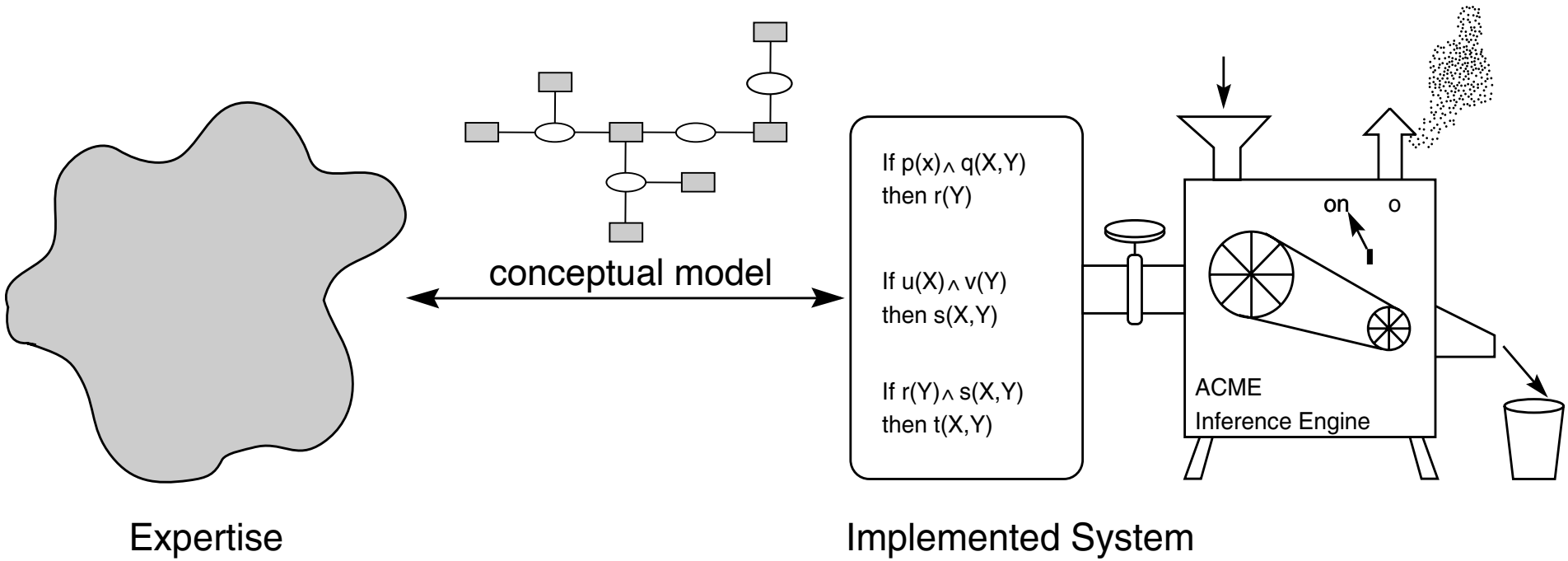
Category Theory

Quantum Computing

Answers to the Exercises

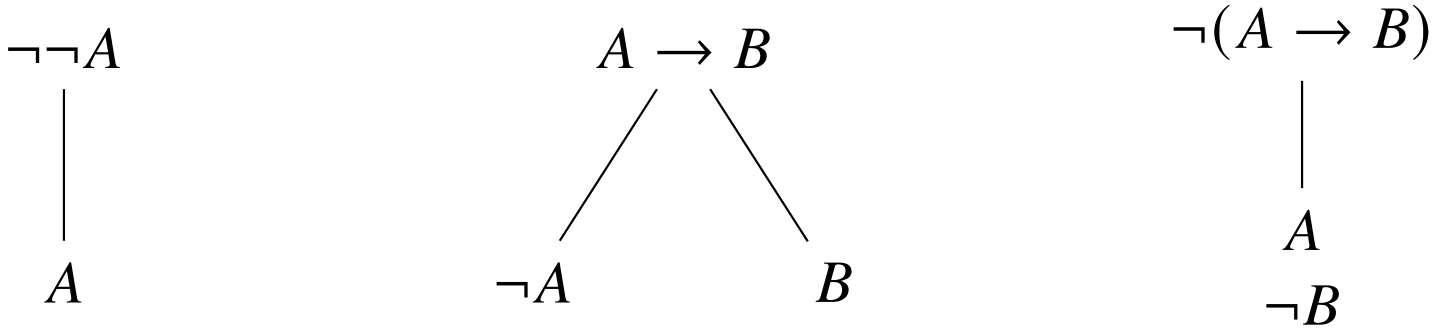
References2039

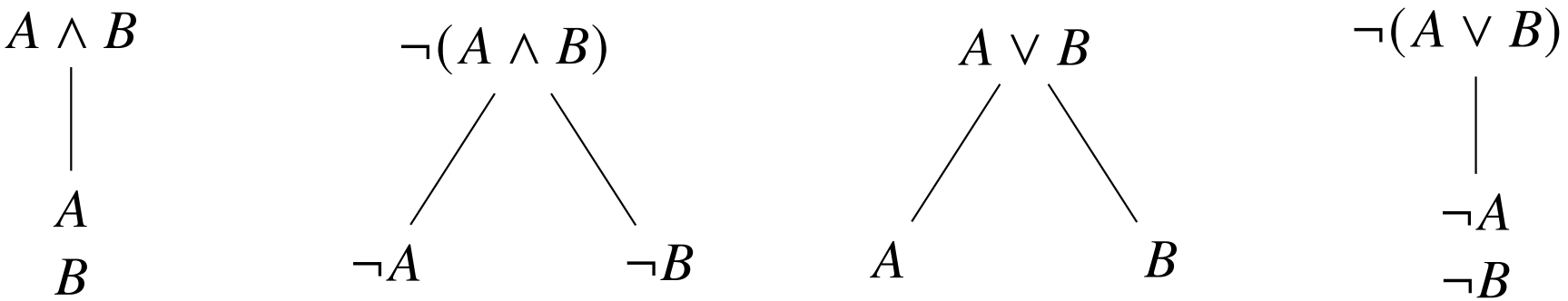
Formal Systems



- Tree Method
- Natural Deduction
- Sequent Calculus
- Hilbert System
- Resolution
- ...

命题逻辑的树形方法 ♡







✓

谓词逻辑的树形方法 ♡

$$\forall x A \qquad\qquad \exists x A \ \checkmark$$
$$\Big| \qquad\qquad \Big|$$
$$A[t/x] \qquad\qquad A(a)$$

where a is a new constant which does not appear on the same branch as $\exists x A$.

$$\neg\forall x A \ \checkmark \qquad\qquad \neg\exists x A \ \checkmark$$
$$\Big| \qquad\qquad \Big|$$
$$\exists x \neg A \qquad\qquad \forall x \neg A$$

$$\vdots \qquad\qquad s = t$$
$$\Big| \qquad\qquad A[s/x]$$
$$t = t \qquad\qquad \Big|$$
$$A[t/x]$$

什么是 “证明”? ♡

Definition (证明)

$A_1, \ldots, A_n \vdash B$ 当且仅当, 存在一棵从 $\{A_1, \ldots, A_n, \neg B\}$ 开始的闭树.

小技巧

1. 少生枝节: 既能生一枝又能生两枝时, 优先生一枝.
2. 能闭掉的枝尽早闭掉.
3. 肯定量词和否定量词同时出现时, 优先处理否定量词.
4. 存在量词和全称量词同时出现时, 优先处理存在量词.

可靠性 & 完备性 ♡

Theorem (可靠性定理)

If the tree closes, the set is unsatisfiable.

Theorem (完备性定理)

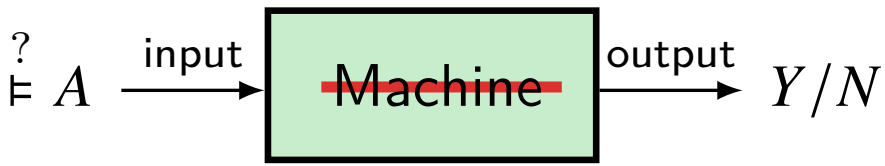
If a set is unsatisfiable, there exists a closed tree from it.

$$\frac{A_1, \ldots, A_n \vdash B}{A_1, \ldots, A_n \vDash B}$$

$\vdash$ captures $\vDash$
No more, no less

可靠 $\vdash \Longrightarrow \vDash$ 不多: 所有证明出来的论证都是有效的

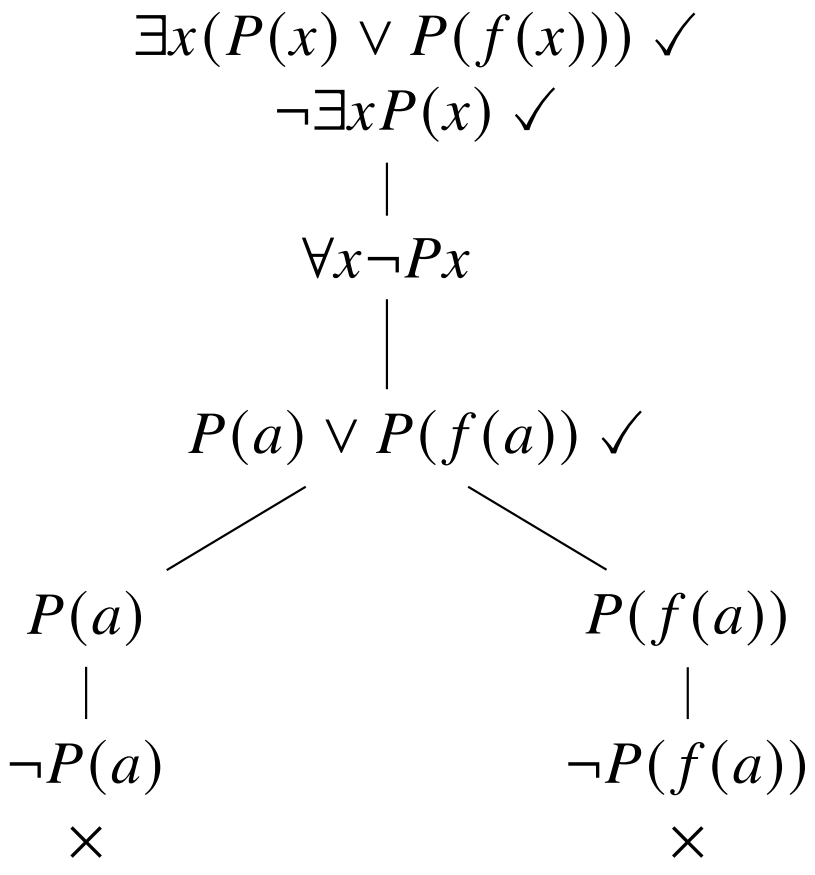
完备 $\vDash \Longrightarrow \vdash$ 不少: 所有有效的论证都能够证明出来



Remark: 如果一个谓词逻辑的论证不是有效的, 且其反模型是无穷的, 那么, 我们无法通过树形方法找到它.

Example

公式集 $\left\{\exists x(P(x) \vee P(f(x))), \neg\exists x P(x)\right\}$ 不可満足



$$\frac{\exists x(P(x) \vee P(f(x)))}{\exists x Px}$$

Theorem (有一个人, 如果他酗酒, 那么所有人都酗酒.)

$$\vdash \exists x(A \to \forall x A)$$

Theorem

存在一个大于 2 的偶数 k, 使得如果 k 能被写成两个素数的和, 那么所有偶数都能写成两个素数的和.

Proof.

- 要么所有大于 2 的偶数都能被写成两个素数的和, 要么有大于 2 的偶数不能被写成两个素数的和.
- 如果是前者, 那么任意一个大于 2 的偶数都能满足这个定理.
- 如果是后者, 那么令 k 为最小的不能被写成两个素数的和的偶数, 则如果 k 能被写成两个素数的和, 那么所有偶数都能写成两个素数的和.

□

Example: 有一个人, 如果他酗酒, 那么所有人都酗酒

$\vdash \exists x(A \to \forall x A)$

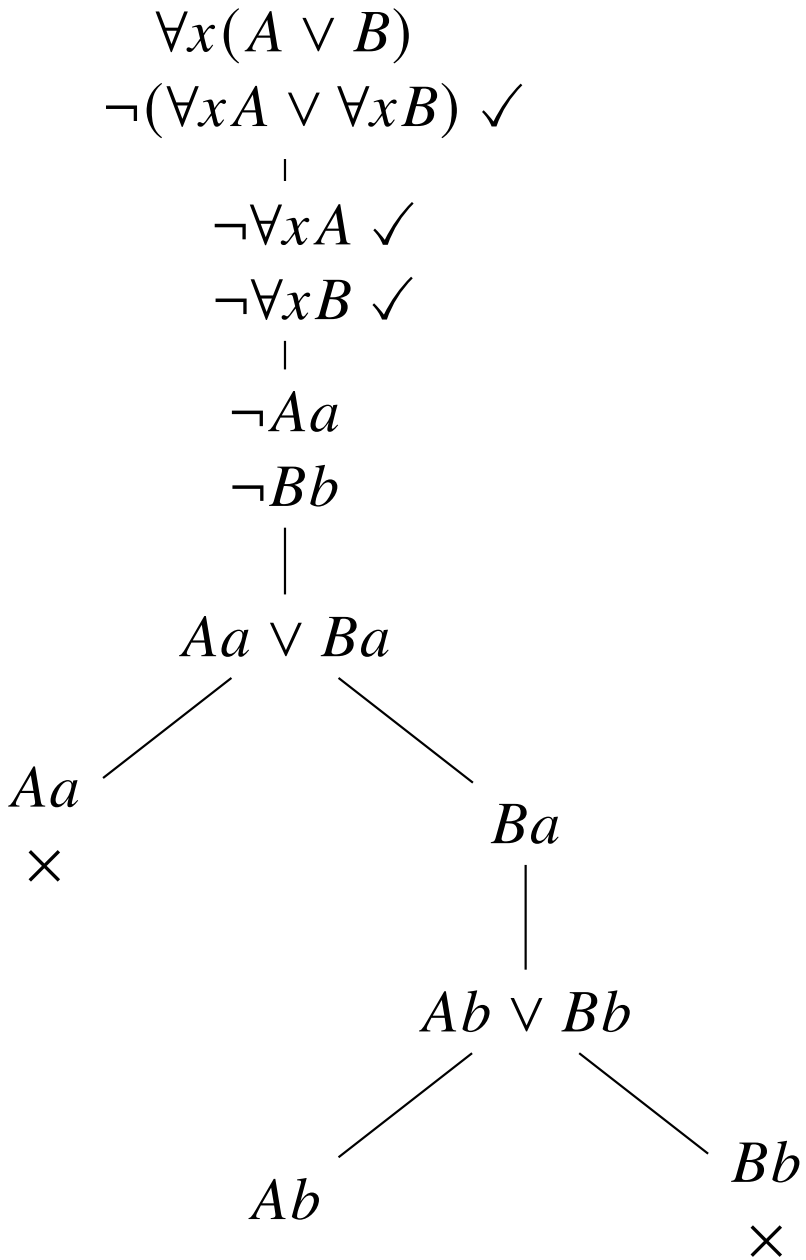
$$\neg\exists x(A \to \forall x A)$$
$$\neg(Aa \to \forall x A)\ \checkmark$$
$$Aa$$
$$\neg\forall x A\ \checkmark$$
$$\neg Ab$$
$$\neg(Ab \to \forall x A)\ \checkmark$$
$$Ab$$
$$\neg\forall x A$$
$$\times$$

Remark: $\nvdash \exists x A \to \forall x A$

反模型

- 在谓词逻辑里, 当一个论证无效时, 通常无法通过闭不掉的树证明其无效.
- 偶尔, 可以借助 "饱和开枝" 构造反模型.
- **Saturated open branch:** An open branch is called saturated iff every non-literal has been analyzed at least once and, additionally, every $\forall$-formula has been instantiated with every term we can construct using the function symbols on the branch.
- We can construct a countermodel from a saturated open branch.

构造反模型



$$\frac{\forall x(A \vee B)}{\forall xA \vee \forall xB} \; ?$$

Let $\mathcal{M} = \left(M, A^{\mathcal{M}}, B^{\mathcal{M}}\right)$

$$M = \{a, b\}$$
$$A^{\mathcal{M}} = \{b\}$$
$$B^{\mathcal{M}} = \{a\}$$

then

$$\mathcal{M} \vDash \forall x(A \vee B)$$

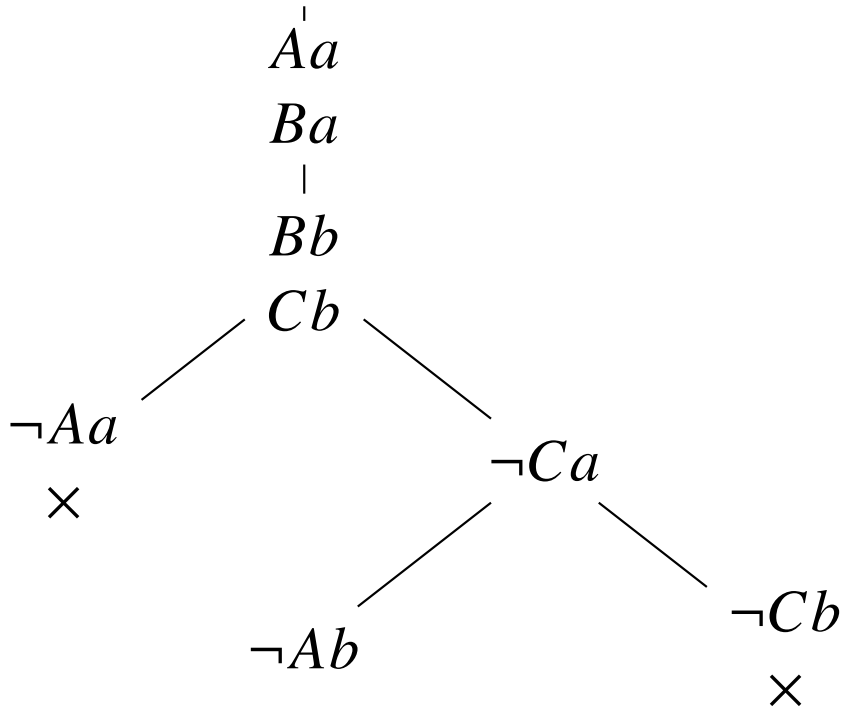
but

$$\mathcal{M} \nvDash \forall xA \vee \forall xB$$

构造反模型

$$\frac{\exists x(A \wedge B) \quad \exists x(B \wedge C)}{\exists x(A \wedge C)} \; ?$$

$\exists x(A \wedge B)$ ✓
$\exists x(B \wedge C)$ ✓
$\neg\exists x(A \wedge C)$



Let $\mathcal{M} = \left(M, A^{\mathcal{M}}, B^{\mathcal{M}}, C^{\mathcal{M}}\right)$

$$M = \{a, b\}$$
$$A^{\mathcal{M}} = \{a\}$$
$$B^{\mathcal{M}} = \{a, b\}$$
$$C^{\mathcal{M}} = \{b\}$$

then

$$\mathcal{M} \vDash \exists x(A \wedge B)$$

$$\mathcal{M} \vDash \exists x(B \wedge C)$$

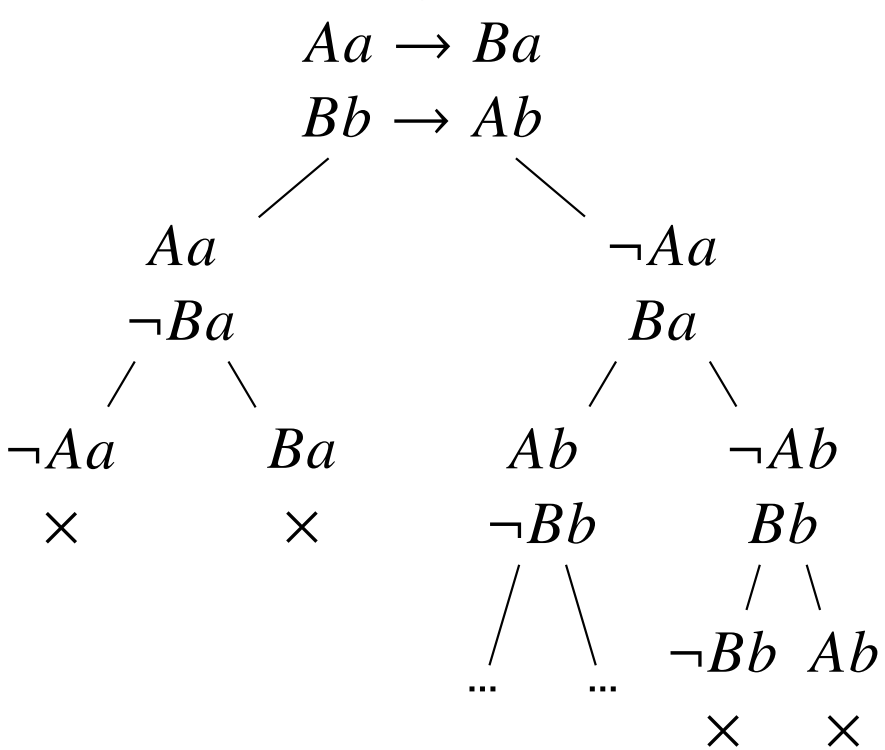
but

$$\mathcal{M} \nvDash \exists x(A \wedge C)$$

构造反模型

$$\frac{\exists x(A \to B) \quad \exists x(B \to A)}{\exists x(A \leftrightarrow B)} \; ?$$

$\exists x(A \to B)$ ✓
$\exists x(B \to A)$ ✓
$\neg\exists x(A \leftrightarrow B)$



Let $\mathcal{M} = \left(M, A^{\mathcal{M}}, B^{\mathcal{M}}\right)$

$$M = \{a, b\}$$
$$A^{\mathcal{M}} = \{b\}$$
$$B^{\mathcal{M}} = \{a\}$$

then

$$\mathcal{M} \vDash \exists x(A \to B)$$

$$\mathcal{M} \vDash \exists x(B \to A)$$

but

$$\mathcal{M} \nvDash \exists x(A \leftrightarrow B)$$

构造反模型

$$\frac{\exists y \forall x Lxy}{\forall y \exists x Lxy} \; ?$$

$\exists y \forall x Lxy$ ✓
$\neg \forall y \exists x Lxy$ ✓
|
$\exists y \neg \exists x Lxy$ ✓
|
$\neg \exists x Lxa$
|
$\forall x Lxb$
|
Lab
Lbb
|
$\neg Laa$
$\neg Lba$



Let $\mathcal{M} = (M, L^{\mathcal{M}})$

$$M = \{a, b\}$$
$$L^{\mathcal{M}} = \{(a, b), (b, b)\}$$

then

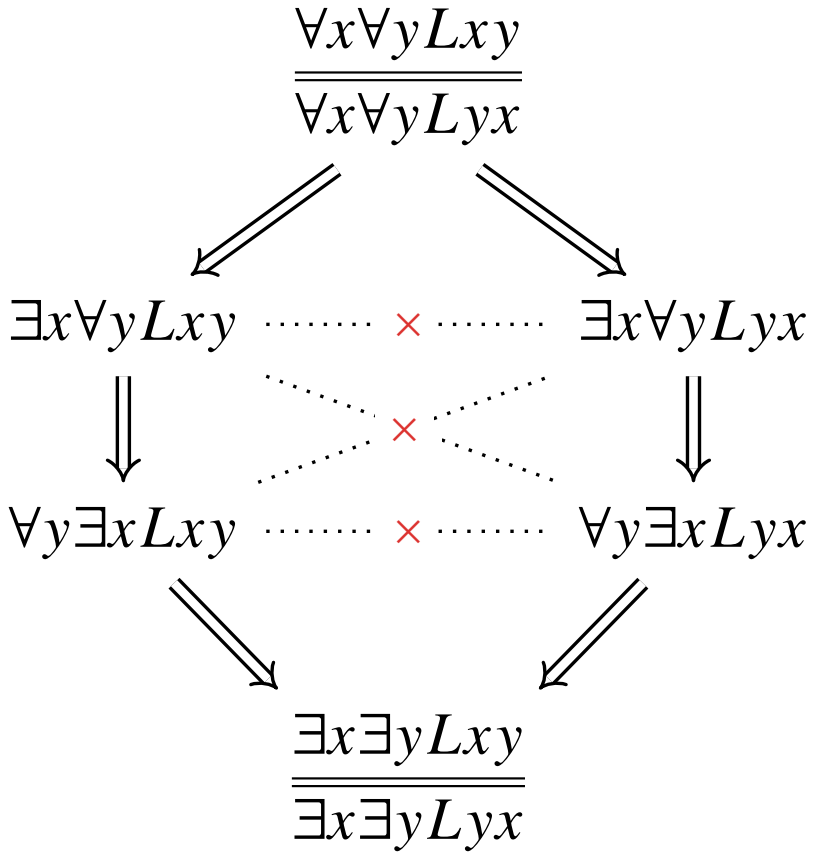
$$\mathcal{M} \vDash \exists y \forall x Lxy$$

but

$$\mathcal{M} \nvDash \forall y \exists x Lxy$$

Examples

1. Everybody loves everybody
2. Everybody is loved by everybody
3. Somebody loves everybody
4. Somebody is loved by everybody
5. Everybody is loved by somebody
6. Everybody loves somebody
7. Somebody loves somebody
8. Somebody is loved by somebody



练习: 有效性判定 ↯

$$\frac{\forall x A}{A[t/x]} \qquad \frac{A[t/x]}{\exists x A} \qquad \frac{\neg\forall x A}{\exists x \neg A} \qquad \frac{\neg\exists x A}{\forall x \neg A}$$

$$\frac{\forall x(A \wedge B)}{\forall x A \wedge \forall x B} \qquad \frac{\forall x A \vee \forall x B}{\forall x(A \vee B)} \qquad \frac{\exists x(A \vee B)}{\exists x A \vee \exists x B} \qquad \frac{\exists x(A \wedge B)}{\exists x A \wedge \exists x B}$$

$$\frac{\forall x(A \to B)}{\forall x A \to \forall x B} \qquad \frac{\forall x(A \to B)}{\exists x A \to \exists x B} \qquad \frac{\forall x(A \leftrightarrow B)}{\forall x A \leftrightarrow \forall x B}$$

$$\frac{\forall x \forall y A x y}{\forall y \forall x A x y} \qquad \frac{\exists x \exists y A x y}{\exists y \exists x A x y} \qquad \frac{\exists x \forall y A x y}{\forall y \exists x A x y}$$

$$\frac{\forall x A \to \exists x B}{\exists x(A \to B)} \qquad \frac{\exists x A \to \forall x B}{\forall x(A \to B)} \qquad \frac{}{\exists x(A \to \forall x A)}$$

练习: 有效性判定

$x \notin \mathrm{Fv}(A)$:

$$\frac{A}{\forall x A} \qquad \frac{A}{\exists x A}$$

$$\frac{\forall x(A \vee B)}{A \vee \forall x B} \qquad \frac{\exists x(A \vee B)}{A \vee \exists x B}$$

$$\frac{\forall x(A \wedge B)}{A \wedge \forall x B} \qquad \frac{\exists x(A \wedge B)}{A \wedge \exists x B}$$

$$\frac{\forall x(A \to B)}{A \to \forall x B} \qquad \frac{\exists x(A \to B)}{A \to \exists x B}$$

$$\frac{\exists x B \to A}{\forall x(B \to A)} \qquad \frac{\forall x B \to A}{\exists x(B \to A)} \qquad \frac{\exists x(A \to B) \quad \exists x(B \to A)}{\exists x(A \leftrightarrow B)}$$

Exercise

$$\frac{\forall x B \to A}{\exists x (B \to A)} \; x \notin \mathrm{Fv}(A)$$

$$\mathsf{diam}(X) := \sup \left\{ |x - y| : x, y \in X \right\}$$

$$\frac{\left(\forall x \in X \; |x| \leq 1\right) \to \mathsf{diam}(X) \leq 2}{\exists x \in X \left(|x| \leq 1 \to \mathsf{diam}(X) \leq 2 \right)?} \; ?$$

1. 上述两公式是否等价?
2. 第二个公式是否成立?

$$\forall x (Cx \to Bx) \to A \quad \text{vs} \quad \exists x (Cx \wedge (Bx \to A))$$

有效性判定 ♡

$$\frac{}{t = t} \qquad \frac{}{s = t \to t = s} \qquad \frac{}{r = s \to s = t \to r = t}$$

$$\frac{s_1 = t_1, \ldots, s_n = t_n}{f(s_1, \ldots, s_n) = f(t_1, \ldots, t_n)} \qquad \frac{s_1 = t_1, \ldots, s_n = t_n}{P(s_1, \ldots, s_n) \leftrightarrow P(t_1, \ldots, t_n)}$$

$$\frac{s = t}{r[s/x] = r[t/x]} \qquad \frac{s = t}{A[s/x] \leftrightarrow A[t/x]}$$

有效性判定 ♡

$x \notin \mathrm{Var}(t)$:

$$\frac{}{\exists x(x = t)} \qquad \frac{A[t/x]}{\exists x(x = t \wedge A)} \qquad \frac{A[t/x]}{\forall x(x = t \to A)}$$

$y \notin \mathrm{Var}(A)$:

$$\frac{}{\forall x(A \leftrightarrow \exists y(y = x \wedge A[y/x]))} \qquad \frac{}{\forall x(A \leftrightarrow \forall y(y = x \to A[y/x]))}$$

构造性证明 vs 存在性证明

$$\frac{\Gamma \vdash A[t/x]}{\Gamma \vdash \exists x A} \qquad \frac{\Gamma, \forall x \neg A \vdash \bot}{\Gamma \vdash \exists x A}$$

Example (There exist two irrational numbers x, y s.t. x^y is rational.)

$$\begin{cases} x := \sqrt{2} \\ y := \sqrt{2} \end{cases} \qquad \begin{cases} x := \sqrt{2}^{\sqrt{2}} \\ y := \sqrt{2} \end{cases}$$

- $R(x)$: x is rational
- $a = \sqrt{2}$
- $f(x, y) = x^y$

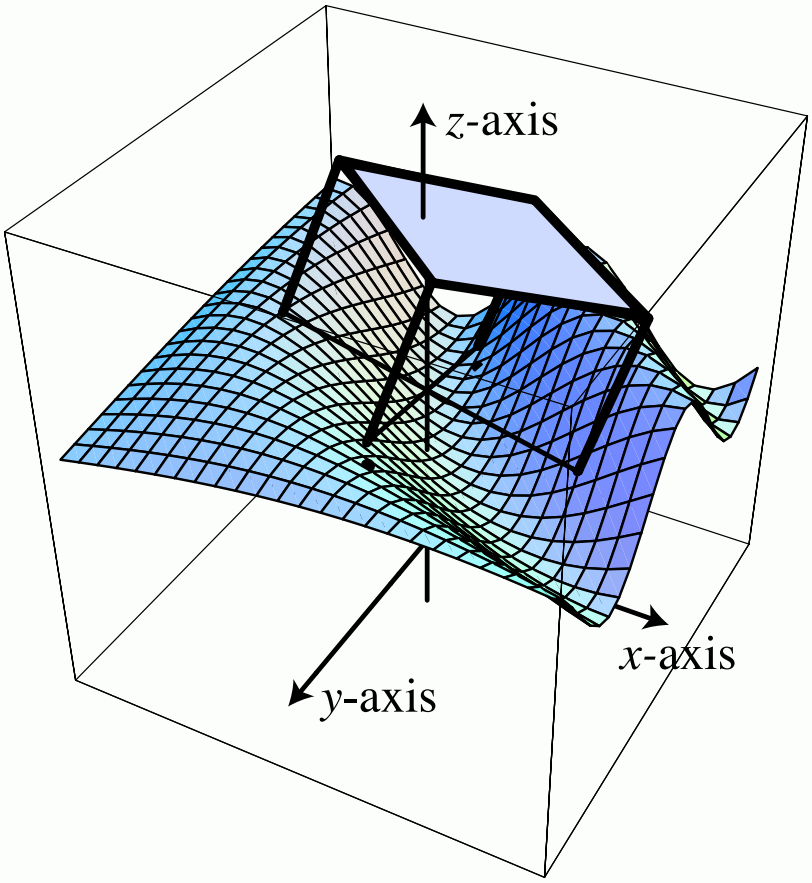
逻辑推理:

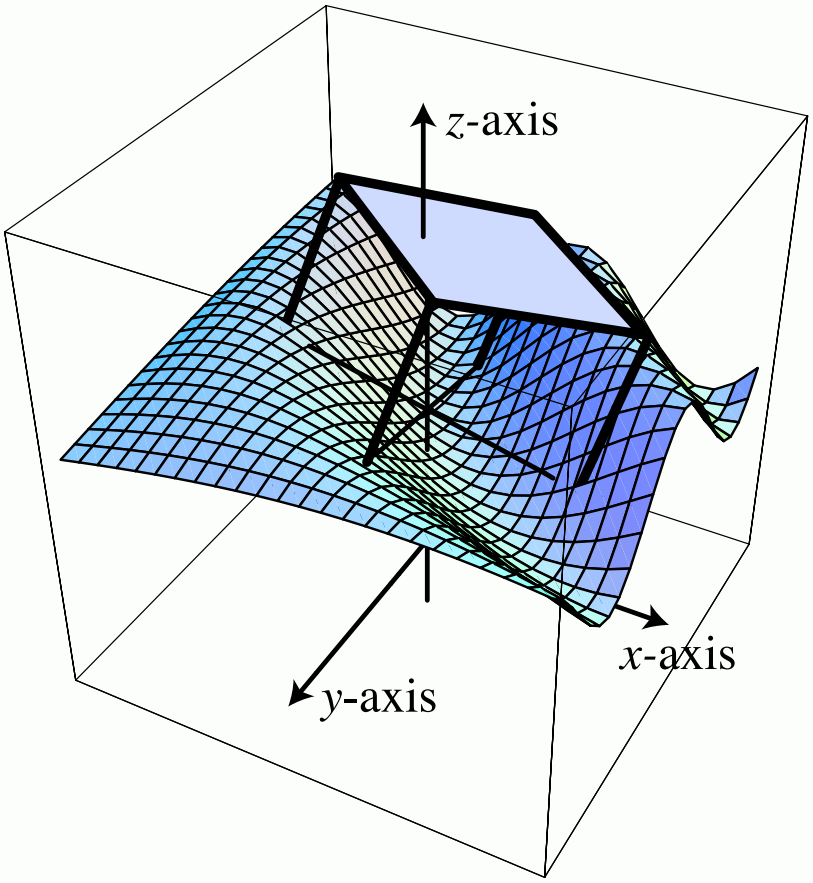
$$\frac{\neg R(a) \quad R(f(f(a,a),a))}{\exists x y (\neg R(x) \wedge \neg R(y) \wedge R(f(x,y)))}$$

$$\begin{cases} x := \sqrt{2} \\ y := \log_2 9 \end{cases}$$

存在性证明

Example (连续起伏的地面上摇晃的四腿方桌可以通过旋转放平稳)





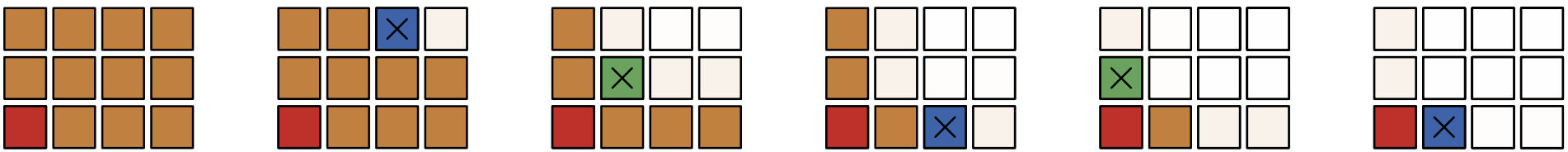
Example ($e+\pi$ 和 $e\pi$ 中至少有一个是超越数)

$$x^2 - (e+\pi)x + e\pi = 0$$

存在性证明

毒巧克力博弈

左下角格子的巧克力有毒. 玩家轮流选一个格子, 并吃掉格子里的以及格子右上方的所有巧克力.



Theorem

玩家 1 有必胜策略.

Proof.

玩家 1 开始可以选择最右上角的格子. 假如玩家 2 有必胜策略, 那么玩家 1 一开始就可以选择玩家 2 的动作. □

Application of Logic in Game Theory

Theorem (Zermelo's Theorem 1913)

在两人的、完美信息的、没有平局的、有穷博弈中, 必有一方有必胜策略.

Proof.

$$\exists x_1 \forall y_1 \ldots \exists x_n \forall y_n A \vee \forall x_1 \exists y_1 \ldots \forall x_n \exists y_n \neg A$$

where A states that a final position is reached where player 1 wins. □

Remark: 两个上帝下围棋, 只需要猜先, 不用落子, 猜先完毕游戏结束.

Proof.

First, color those end nodes black that are wins for player 1, and color the other end nodes white, being the wins for player 2. Then

- if player 1 is to move, and at least one child is black, color it black; if all children are white, color it white.
- if player 2 is to move, and at least one child is white, color it white; if all children are black, color it black.

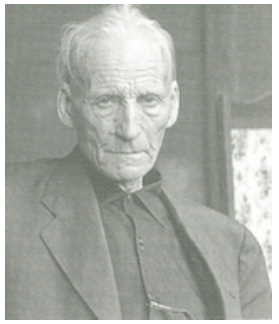
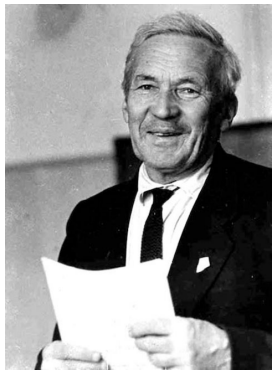
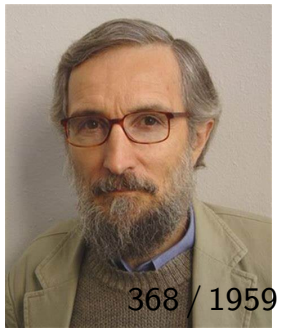
直觉主义

- 非直谓主义 (*Poincaré*, Russell)
  禁止恶性循环原则: 任何一个实体都不能仅仅通过它所属的整体来定义.
- 直觉主义 数学是心灵的构造.
  (Kronecker, *Brouwer*, Heyting, *Kolmogorov*, Weyl)
  - 潜无穷 vs 实无穷
  - "存在就是被 (直觉) 构造" — Brouwer
  - 非构造性证明 ×
  - 反证法 ×
  - 双重否定消去律 ×
  - 排中律 ×
  - 选择公理 ×

  (There exist two irrational numbers x and y s.t. x^y is rational.)

  $$\sqrt{2} \quad \log_2 9$$

  "上帝创造了整数, 其余的都是人造的." — Kronecker
- 构造主义数学 (Bishop, *Martin-Löf*)

Excluded Middle is sometimes fine

- We can sometimes use excluded middle.
- We just can't assume it without proving it.

Theorem

Every natural number is either equal to zero or not equal to zero.

Constructive proof.

By induction.

1. If $n = 0$, then $n = 0$.
2. Assume inductively that either $n = 0$ or $n \neq 0$.
   In either case, $n + 1 \neq 0$.

Thus, by induction, for all natural numbers n, either $n = 0$ or $n \neq 0$. □

练习: 树形证明 — Now it's your turn ↯

$$\frac{\exists x \forall y (P(y) \to y = x) \quad \forall x P(f(x))}{\exists x (f(x) = x)}$$

$$\frac{\exists x \big(Px \wedge \forall y (Py \to y = x) \wedge Qx\big)}{\exists x \forall y \big((Py \leftrightarrow y = x) \wedge Qx\big)}$$

$$\frac{\exists x \big(Px \wedge \forall y (Py \to y = x)\big) \quad \exists x \big(Qx \wedge \forall y (Qy \to y = x)\big) \quad \neg \exists x (Px \wedge Qx)}{\exists x y \big(x \neq y \wedge (Px \vee Qx) \wedge (Py \vee Qy) \wedge \forall z (Pz \vee Qz \to z = x \vee z = y)\big)}$$

$*\mathbf{54} \cdot \mathbf{43}$. $\vdash :. \alpha, \beta \in 1. \supset : \alpha \cap \beta = \Lambda. \equiv .\alpha \cup \beta \in 2$

Dem.

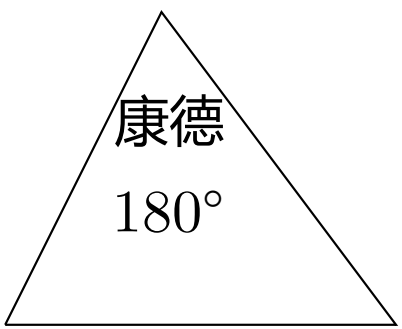
$$\begin{aligned}
&\vdash .*54 \cdot 26. \supset \vdash :. \alpha = \iota' x . \beta = \iota' y. \supset : \alpha \cup \beta \in 2. \equiv .x \neq y. \\
&[*51 \cdot 231] \qquad\qquad \equiv .\iota' x \cap \iota' y = \Lambda. \\
&[*13 \cdot 12] \qquad\qquad \equiv .\alpha \cap \beta = \Lambda \qquad (1) \\
&\vdash .(1). *11 \cdot 11 \cdot 35. \supset \\
&\qquad \vdash :. (\exists x, y). \alpha = \iota' x . \beta = \iota' y. \supset : \alpha \cup \beta \in 2. \equiv .\alpha \cap \beta = \Lambda \qquad (2) \\
&\vdash .(2). *11 \cdot 54. *52 \cdot 1. \supset \vdash .Prop
\end{aligned}$$

From this proposition it will follow, when arithmetical addition has been defined, that $\mathbf{1} + \mathbf{1} = \mathbf{2}$.

$$\exists_1 x P \wedge \exists_1 x Q \wedge \neg \exists x (P \wedge Q) \to \exists_2 x (P \vee Q)$$

数学是“先验综合判断”吗?

- 苏格拉底: “我只知道我什么也不知道.”
- 笛卡尔: “我思故我在”. 我们可以通过内省获得部分知识的确定性, 但关于外部世界的知识呢?
- 休谟: 放弃吧. (i) 关于外部世界的知识是因果陈述. (ii) 因果陈述是综合判断, 所以只能后验的获知. (iii) 我们既无法直接观察因果本身, 又无法非循环的论证未来与过去的相似性, 所以我们无法后验的获知因果陈述.
- 康德: 虽然因果陈述是综合的, 但我们可以先验的获知因果. 因为因果结构既是关于外部世界的原始材料的, 又是内嵌在人类心灵的认知模式的一部分. “人为自然立法.”



康德 数学是先验综合判断.

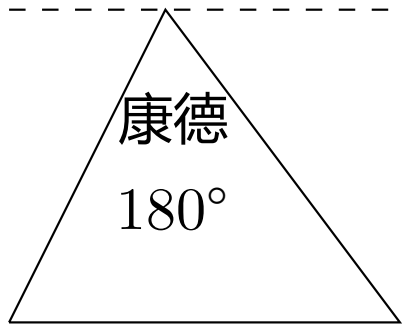
- 几何 — 空间的纯粹直观
- 算术 — 时间的纯粹直观

高斯 几何是后验的. (黎曼几何, 空间曲率)

弗雷格 算术是分析的.

$$\vdash 1 + 1 = 2$$

罗素[14]



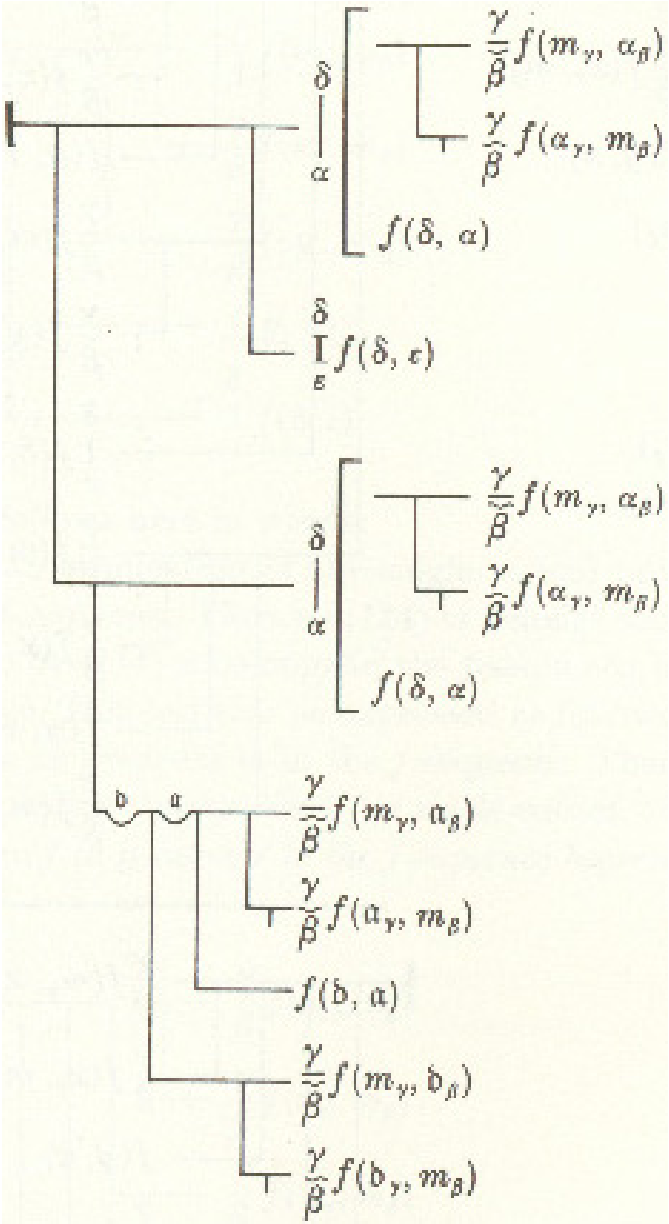
- 康德正确的认识到欧式几何的命题不能脱离图形的辅助直接从公理推出, 为此, 他不惜发明了一整套认识论.
- 康德发明的 “先验综合判断” 只不过是为了给一个有缺陷的逻辑 (亚里士多德三段论) 打补丁.
- 康德说休谟将他从 “教条的迷梦” 中唤醒, 但康德自己又发明了新的 “安眠药” 再次昏睡.
- 纯数学, 包括几何学, 不过是形式逻辑. 这对康德哲学是致命一击.

[14] Bertrand Russell: “Mathematics and the Metaphysicians”. Reprinted in *Mysticism and Logic, and Other Essays*.

Remark: 罗素的第一本书《论几何学的基础》贯彻的是康德的几何观, 把物质与空间分离, 用先验直观拒绝曲率不为 0 的空间, 结果被爱因斯坦打脸, 广义相对论恰恰连接了物质与时空曲率.

弗雷格

- 算术规律是分析判断，因此是先验的．算术不过是发展了的逻辑，算术定理是逻辑规律．
- 算术在自然科学上的应用不过是对观察现象的逻辑加工．计算即推理．
- 如果哲学的任务是破除语词对人类精神的支配，揭开由于语言的表达方式而造成的概念关系的假象，把思想从日常语言的迷雾中解放出来，那么，我的《概念文字》将成为哲学家手中的有用工具．
- 一个好的数学家至少是半个哲学家；一个好的哲学家至少是半个数学家．



	先验	后验
分析	数学	×
综合	×	物理

Table: 康德之前

	先验	后验
分析	逻辑	×
综合	算术 几何	

Table: 康德

	先验	后验
分析	×	×
综合	×	逻辑 算术 几何

Table: 蒯因、普特南

	先验	后验
分析	逻辑 算术 纯粹几何	×
综合	×	应用几何

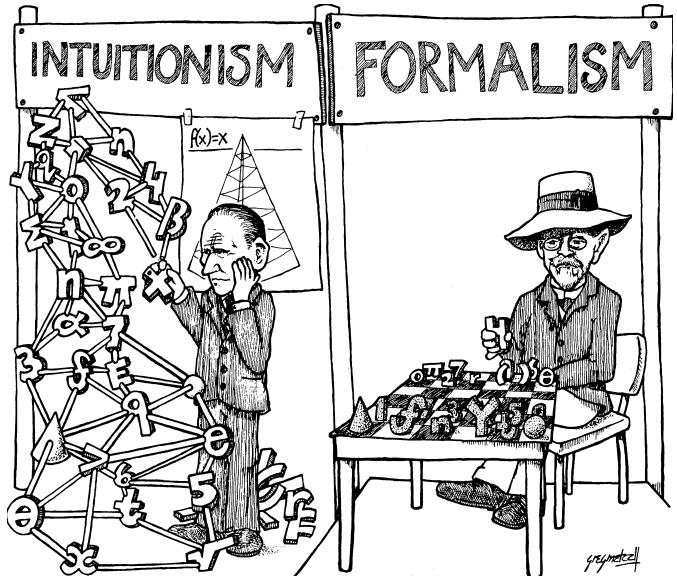
Table: 逻辑实证主义

	先验	后验
分析		×
综合	逻辑 算术 纯粹几何	应用几何

Table: Martin-Löf

数学哲学: 逻辑主义/直觉主义/形式主义





逻辑主义	直觉主义	形式主义
数学 / 逻辑	逻辑 / 数学 / 心灵构造	数学 / 符号游戏
实在论	概念论	唯名论

何物存在?

唯名论:
抽象事物不存在

极端唯名论:
即使具体事物也不存在

柏拉图主义:
抽象事物存在

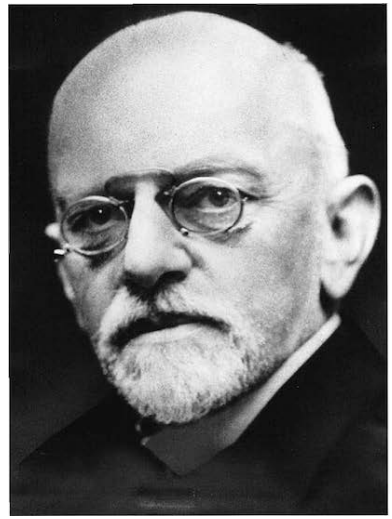
极端柏拉图主义:
就连具体事物也存在

- 名字叫 “罗素” 的那个人存在
- 人存在
- 电子存在
- 红存在
- 2 存在

“如果其他事物存在, 那数也存在.”

— 柏拉图《智者篇》

希尔伯特 David Hilbert 1862-1943



- 几何的**形式**公理化
  几何相对于算术的一致性
  (克莱因: 非欧几何相对于欧式几何的一致性)
  (natural/integer/rational/real/complex)
- 希尔伯特 $23/24$ 问题 $(1^{\text{st}}, 2^{\text{nd}}, 10^{\text{th}}, 24^{\text{th}}, 6^{\text{th}})$
- 元数学 — 证明论
- 形式主义 数学是符号游戏.
- 公理是初始概念的隐定义.
- 我们完全可以用 “桌子、椅子、啤酒瓶” 代替 “点、线、面” 而不影响推理的正确性.
- 数学是根据某些简单规则使用毫无意义的符号在纸上进行的游戏, 是制造快乐的游戏.
- 我们的内心响起了永恒的召唤: 那里有一个问题, 去找出它的答案! 你可以通过纯粹理性找到它 — 数学里没有不可知!
- 我们必须知道; 我们必将知道!

1. 形式化与公理化 (语言的丰富性): 形式化 — 形式与内涵分离; 公理化 — 公理化 (1) 逻辑 L, (2) 关于现实世界的 "有穷" 数学 F, (3) 关于理想世界的 "无穷" 数学 T. 所有数学命题都可以在 T 中表达.
2. 独立性: 证明公理之间彼此 "独立".
3. 完备性: (1) 所有有效的逻辑命题都可以在 L 中证明; (2) 所有真的数学命题都可以在 T 中证明. — 实现形式与内涵的统一.
4. **一致性**: 用 "有穷数学"F 的方法证明 "无穷数学"T 中推不出矛盾.
5. 保守性 ($\forall A \in \Pi_1 [\mathrm{T} \vdash A \implies \mathrm{F}, \mathrm{Con}_{\mathrm{T}} \vdash A]$ 一致性蕴含保守性): 任何关于 '实在对象' 的陈述, 如果可以在 T 中证明, 那么也可以在 F 中证明.
6. 可判定性 (能行性): 逻辑命题的有效性和数学命题的真理性都可以机械地判定.
7. 简单性: 证明某些证明的最大简单性.
8. 范畴性? 在同构的意义上, T 刻画了唯一一个结构.

[15]绕道理想世界迂回地获取关于现实世界的知识, 并基于现实世界证明理想世界的合理性.

希尔伯特规划

“当我们更仔细地考虑逻辑的公理化时，我们很快就会认识到，整数和集合的一致性问题不是孤立的，而是属于一个广阔的**认识论**领域[16]，这些问题具有特殊的数学色彩：例如，每个数学问题原则上是否可解的问题，结论是否可验证的问题，数学证明的简单性标准问题，数学与逻辑中形式与内涵的关系问题，以及一个数学问题是否在有限步骤内可判定的问题.”

— 希尔伯特

[16] Chaitin: “今天正在发生的计算机接管世界、数字化、信息化，都是希尔伯特在 20 世纪初提出的一个**哲学问题**的结果.”

Example — 皮亚诺算术理论 ☺

1. 我是一个神

$$0 \in \mathbb{N}$$

2. 每个神的兽也是一个神

$$\forall n \in \mathbb{N} : s(n) \in \mathbb{N}$$

3. 我不是任何神的兽

$$\forall n \in \mathbb{N} : s(n) \neq 0$$

4. 不同的神有不同的兽

$$\forall mn \in \mathbb{N} : s(m) = s(n) \to m = n$$

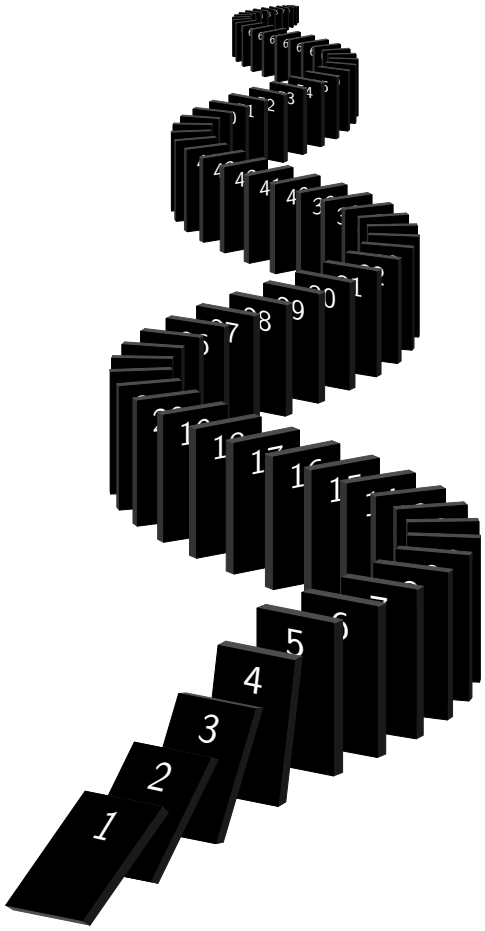
5. 如果我有 X, 且每个神都把 X 递送给它的兽, 那么所有的神都有 X

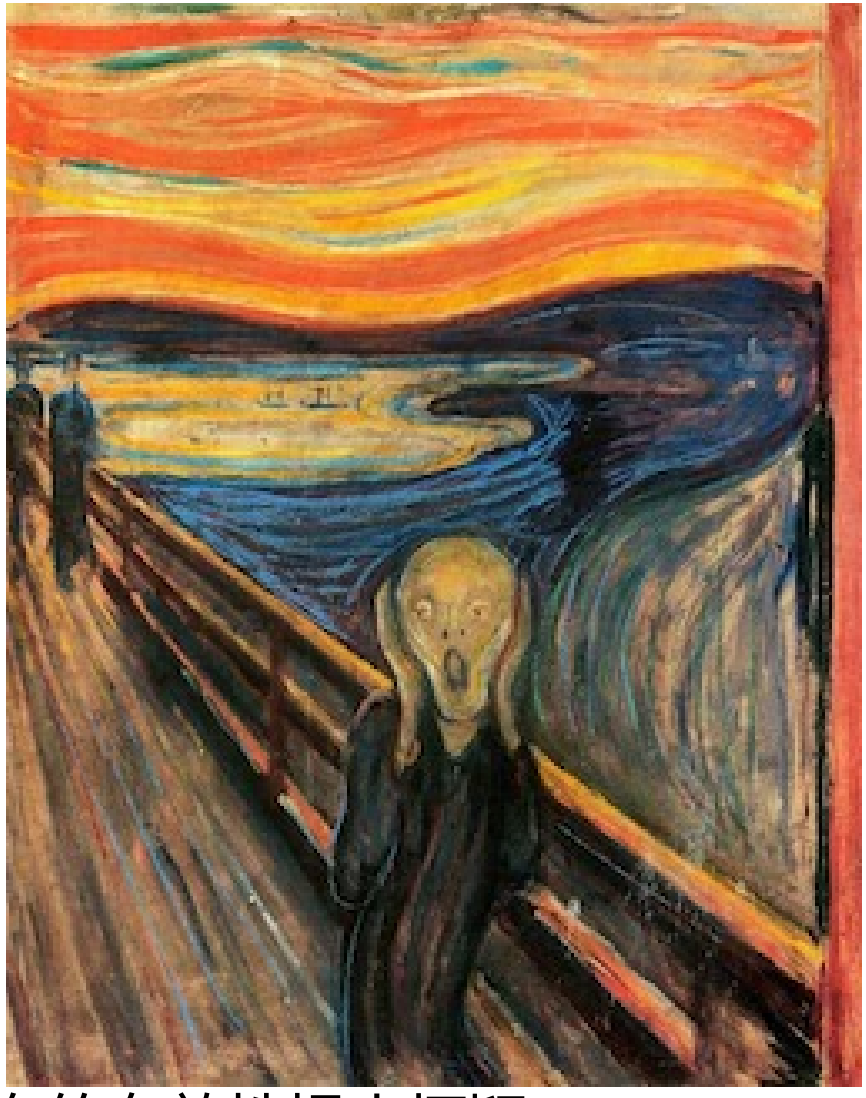
$$\forall X \left[0 \in X \land \forall n (n \in X \to s(n) \in X) \to \forall n \in \mathbb{N} (n \in X) \right]$$

$+, \cdot$ 可以递归定义:

- $\forall n \in \mathbb{N} : n + 0 = n$
- $\forall mn \in \mathbb{N} : m + s(n) = s(m + n)$
- $\forall n \in \mathbb{N} : n \cdot 0 = 0$
- $\forall mn \in \mathbb{N} : m \cdot s(n) = m \cdot n + m$

从莱布尼茨到希尔伯特到哥德尔 — 梦想破碎的声音...





理性可以对自身的有效性提出怀疑.

哥德尔句

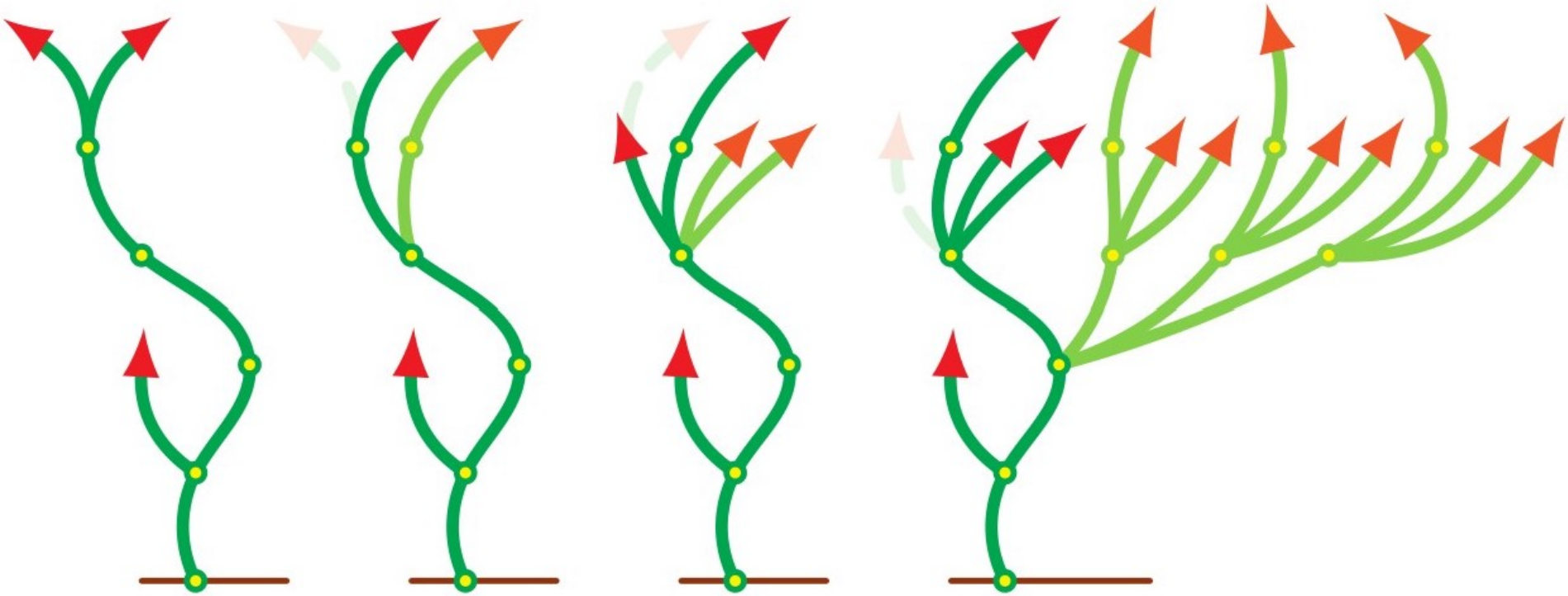
“我不可证”

Problem (哥德尔是什么人?)

- 一个岛上有“君子”、“小人”两类人. “君子”只说真话, “小人”只说假话.
- 岛上有人有身份证, 有人没有.
- 有身份证的都是君子.
- 你来岛上遇到了一个名字叫“哥德尔”的土著.
- 哥德尔说: “我没有身份证”.

Hydra 九头蛇游戏 — “自然的” 不可证命题

- 每一步你砍掉它一个头
- 在第 n 步, 如果它的一个非根部的头被砍掉, 则从被砍掉头的下方节点处长出 n 个副本



Goodstein Theorem 不管你怎么砍都能杀死 “九头蛇”

Kirby-Paris Theorem 但你无法通过 PA 证明这一点

数学哲学

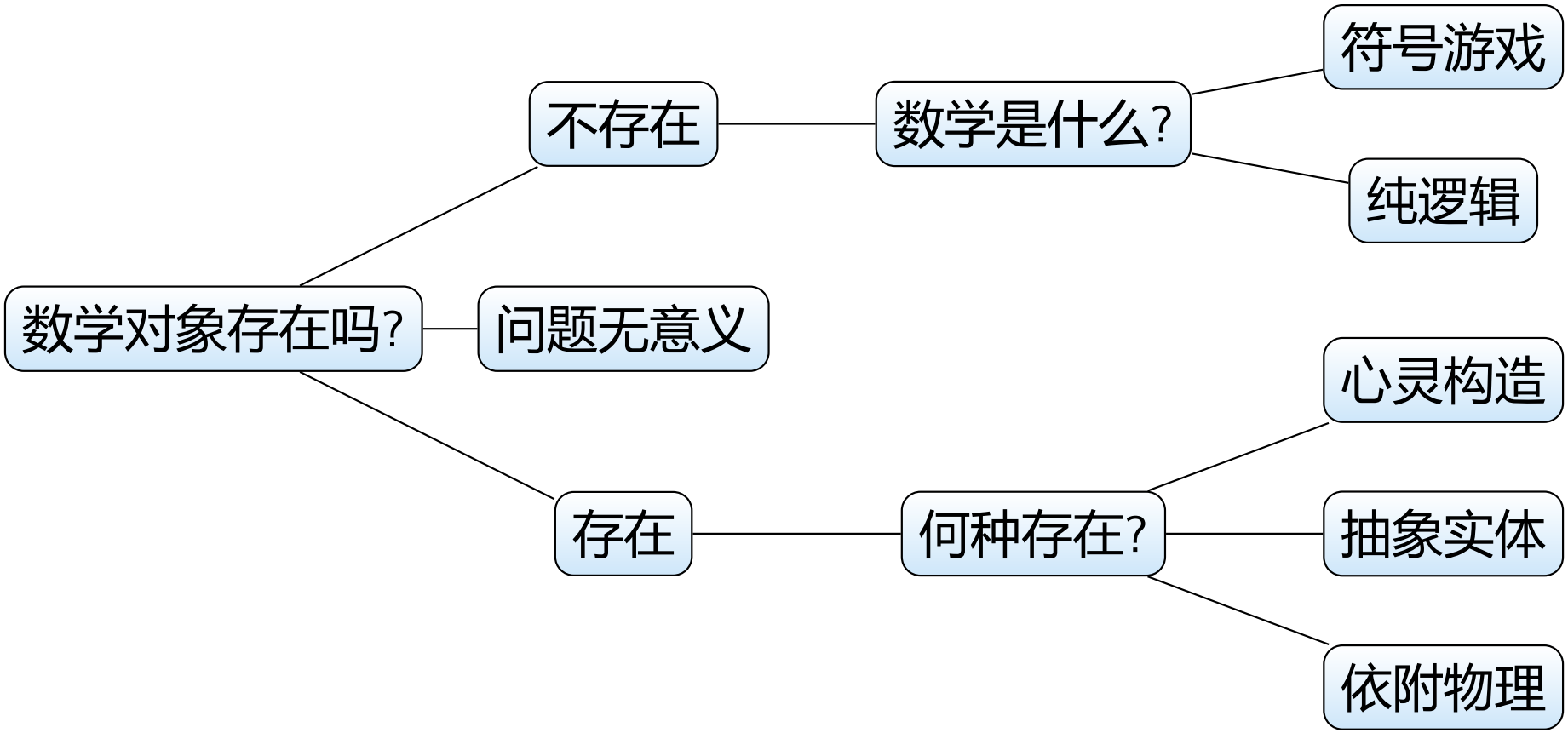


Figure: 形式主义/逻辑主义/直觉主义/柏拉图主义/物理主义

Is there more than one mathematical universe?[17]

[17] Penelope Maddy: What Do We Want a Foundation to Do?

有效论证?

$$\frac{\begin{array}{c}\text{Whatever begins to exist, has a cause of its existence}\\ \text{The universe began to exist}\end{array}}{\text{The universe has a cause of its existence}} \qquad \frac{\begin{array}{c}\forall x(Bx \to \exists y Cyx)\\ Bu\end{array}}{\exists y Cyu}$$

1. 如果上帝存在, 那么不是真诚地相信上帝存在的人不会得到救赎.
2. 如果上帝不存在, 那么没有人会得到救赎.
3. 如果一个人相信上帝存在仅仅是由于帕斯卡赌, 那就不算真诚地相信上帝存在.
4. 帕斯卡仅仅由于帕斯卡赌才相信上帝存在.
5. 因此, 帕斯卡不会得到救赎.

$$\frac{\begin{array}{c}Eg \to \forall x(\neg Bx \to \neg Sx)\\ \neg Eg \to \neg \exists x Sx\\ \forall x(Wx \to \neg Bx)\\ Wp\end{array}}{\neg Sp}$$

有效论证?

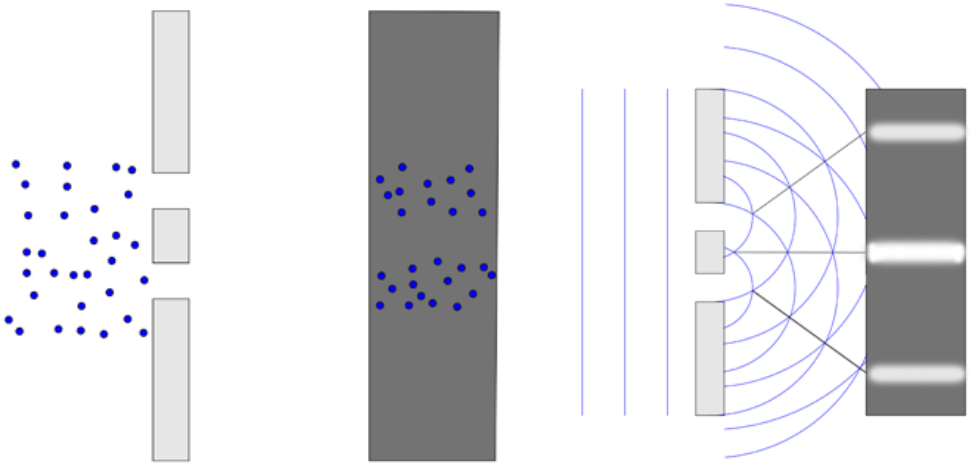
每个全善的东西都消灭了所有它能消灭的邪恶的东西
全能的东西能消灭所有东西
存在邪恶的东西
不存在被某个东西消灭了的东西

不存在全善且全能的东西

$$\forall x(Bx \rightarrow \forall y(Ey \wedge Cxy \rightarrow Dxy))$$
$$\forall x(Ox \rightarrow \forall yCxy)$$
$$\exists xEx$$
$$\neg\exists x\exists yDyx$$

$$\neg\exists x(Bx \wedge Ox)$$

有效论证?



1. 假如发射光是粒子, 那么, 探测屏上的撞击图案就是一系列闪光, 并且会留下至多两条条纹.
2. 假如发射光是波, 那么, 探测屏上的撞击图案就不是一系列闪光, 并且会留下两条以上的条纹.
3. 假如撞击图案是一系列连续的闪光, 但留下了两条以上的条纹, 那么, 发射光既不是粒子也不是波.

$$\frac{\forall x(Px \to Fx \wedge Sx) \quad \forall x(Wx \to \neg Fx \wedge \neg Sx)}{\forall x(Fx \wedge \neg Sx \to \neg Px \wedge \neg Wx)}$$

有效论证?

- 老公: “赌博时, 我押上了自己, 输了; 我又押上了你, 也输了.”
- 老婆: “当你输掉自己时, 你就已经不再拥有我了, 所以你无权押我.”

1. For all y there exists z such that z owns y.
2. For all x, y, z if z owns y and y owns x then z owns x.
3. For all x, y, z if y owns x and z owns x then $y = z$.

$$\frac{\begin{array}{c}\forall y \exists z O z y \\ \forall x y z (O z y \wedge O y x \rightarrow O z x) \\ \forall x y z (O y x \wedge O z x \rightarrow y = z)\end{array}}{\forall y (\neg O y y \rightarrow \neg \exists x O y x)}$$

有效论证?

It is a crime for an American to sell weapons to hostile nations. The country Nono, an enemy of America, has some missiles, and all of its missiles were sold to it by Colonel West, who is American. Theorefore, Colonel West is a criminal.

1. $\forall xyz(\text{American}(x) \land \text{Weapon}(y) \land \text{Hostile}(z) \land \text{Sell}(x, y, z) \rightarrow \text{Criminal}(x))$
2. $\exists x(\text{Own}(\text{nono}, x) \land \text{Missile}(x))$
3. $\forall x(\text{Missile}(x) \land \text{Own}(\text{nono}, x) \rightarrow \text{Sell}(\text{west}, x, \text{nono}))$
4. $\forall x(\text{Missile}(x) \rightarrow \text{Weapon}(x))$
5. $\forall x(\text{Enemy}(x, \text{america}) \rightarrow \text{Hostile}(x))$
6. $\text{American}(\text{west})$
7. $\text{Enemy}(\text{nono}, \text{america})$
8. $\text{Criminal}(\text{west})$

练习: 有效性判定 — Now it's your turn ϟ

1. 有些战争是正义的. 没有侵略战争是正义的. 因此, 有些战争是非侵略性的.
2. 所有中国学生和所有美国学生都爱 Anne. 中南大学只有中国学生和美国学生. 因此, 中南大学的学生都爱 Anne.
3. 非分析的有意义的命题均可证实或可证伪. 哲学命题既不是分析的又不可证实又不可证伪. 所以, 哲学命题无意义.
4. 如果狗是动物, 那么, 狗的头就是动物的头.
5. 唯有亚里士多德是哲学家. 柏拉图是抽烟的哲学家. 所以, 亚里士多德抽烟.
6. 小艾爱着小白, 小白爱着小菜, 小艾已婚, 小菜未婚. 因此, 某个已婚人士爱着某个未婚人士.
7. 人人都爱小艾. 小艾除我之外谁都不爱. 因此, 人人都爱我.
8. 凡爱人者人皆爱之. 罗密欧爱朱丽叶. 所以, 你爱我!
9. 没有理发师会给并且只给那些不给自己理发的人理发. *Russell*
10. 如果大鱼比小鱼游得快, 那么, 只要有最大的鱼就有游得最快的鱼.
11. 如果不知道爱情密码, 没有人能偷走小白的心. 小白的心被偷了. 只有小艾知道爱情密码. 所以, 小艾偷了小白的心.

12. 没有女孩会爱花心的人. 小艾是一个会爱上所有爱她的人的女孩. 小白爱小艾. 所以, 小白不花心.
13. 如果所有的思想都清楚, 那么没有思想需要解释; 如果所有的思想都不清楚, 那么没有思想能够解释清楚. 因此, 如果有的思想既需要解释又能够解释清楚, 那么说明有的思想清楚、有的思想不清楚.
14. 没有人相信并且只相信那些无法与人建立相互信任关系的人. *Quine*
15. 我是个玩摇滚的哲学家. 唯有哲学家才欣赏哲学家. 没有哲学家没有怪癖. 有怪癖的摇滚人都会有女孩欣赏. 有怪癖的人都自命不凡. 因此, 有女孩自命不凡.
16. 当今的那位皇上是秃子. 秃子都性感. 因此, 不管谁是当今的皇上, 都性感.
17. 只有小艾和小白欣赏小菜. 小艾和小白不是同一个人. 无论谁欣赏小菜都会爱上他, 也只有欣赏他的人才会爱他. 因此, 有且仅有两个人爱小菜.
18. If anyone speaks to anyone, then someone introduces them; no one introduces anyone to anyone unless he knows them both; everyone speaks to Alice; therefore everyone is introduced to Alice by someone who knows her.

Application — 扫雷



- There are exactly n mines in the game.
- If a cell contains the number 1, then there is exactly one mine in the adjacent cells.

  $\forall x(\text{Contain}(x, 1) \rightarrow$

  $\exists y(\text{Adjacent}(x, y) \wedge \text{Mine}(y) \wedge \forall z(\text{Adjacent}(x, z) \wedge \text{Mine}(z) \rightarrow z = y)))$

- ...

Application — “君子” 与 “抽烟”

Problem (“君子岛” 还是 “小人岛”? “抽烟” 还是 “不抽烟”?)

你想研究抽烟与说谎之间的相关性, 于是走访各个 “君子岛” 与 “小人岛”. “君子岛” 上的人只说真话, “小人岛” 上的人只说假话.

岛 *1:* 每个土著都说: “岛上君子都不抽烟”.

岛 *2:* 每个土著都说: “岛上有小人抽烟”.

岛 *3:* 每个土著都说: “如果我抽烟, 那么所有岛人都抽烟”.

岛 *4:* 每个土著都说: “如果有岛人抽烟, 那么我也抽烟”.

岛 *5:* 每个土著都说: “虽然有岛人抽烟, 但我不抽烟”.

岛 5:
$$\frac{\forall x(Tx \leftrightarrow \exists y Sy \wedge \neg Sx) \quad \forall x Tx \vee \forall x \neg Tx}{\forall x \neg Tx \wedge (\forall x Sx \vee \forall x \neg Sx)}$$

1. T None
2. F None
3. T All or None
4. T All or None
5. F All or None

Application — “人有来生”?

1. 人的一生有太多可能性没有实现.
2. 如果只有一生没有来生, 没有实现的可能性将永远不可实现.
3. 永远不可实现的可能性没有意义.
4. 如果宇宙是有意义的, 那么其包含的对象的可能性都是有意义的.
5. 可被学习理解的东西是有意义的.
6. 有秩序的东西是可被学习理解的.
7. 宇宙是有秩序的.

1. $\forall xy(\mathsf{Man}(x) \wedge \mathsf{Life}(y,x) \to \exists z(\mathsf{Possible}(z,x,y) \wedge \neg\mathsf{Realize}(z)))$
2. $\forall xy(\mathsf{Man}(x) \wedge \mathsf{Life}(y,x) \wedge \neg\exists y'(\mathsf{Life}(y',x) \wedge y' \neq y) \to \forall z(\mathsf{Possible}(z,x,y) \wedge \neg\mathsf{Realize}(z) \to \mathsf{UnRealizable}(z)))$
3. $\forall x(\mathsf{UnRealizable}(x) \to \neg\mathsf{Meaning}(x))$
4. $\mathsf{Meaning}(u) \to \forall xyz(\mathsf{Contain}(u,x) \wedge \mathsf{Possible}(z,x,y) \to \mathsf{Meaning}(z))$
5. $\forall x(\mathsf{Learnable}(x) \to \mathsf{Meaning}(x))$
6. $\forall x(\mathsf{Ordered}(x) \to \mathsf{Learnable}(x))$
7. $\mathsf{Ordered}(u)$
8. $\forall x(\mathsf{Man}(x) \to \mathsf{Contain}(u,x))$
9. $\forall xy(\mathsf{Man}(x) \wedge \mathsf{Life}(y,x) \to \exists y'(\mathsf{Life}(y',x) \wedge y' \neq y))$

怎么用一个 5 升和一个 7 升的桶去打 1 升的水?

- $S(x, y)$: x 和 y 分别是第一个和第二个桶中的水量.

Initial State $S(0, 0)$

Goal State $\exists x[S(1, x) \vee S(x, 1)]$

一共有 8 种可能的动作:

A_1. $\forall xy\big[S(x, y) \to S(5, y)\big]$ [Fill 1]

A_2. $\forall xy\big[S(x, y) \to S(0, y)\big]$ [Empty 1]

A_3. $\forall xy\big[S(x, y) \to S(x, 7)\big]$ [Fill 2]

A_4. $\forall xy\big[S(x, y) \to S(x, 0)\big]$ [Empty 2]

A_5. $\forall xy\big[S(x, y) \wedge x + y \leq 7 \to S(0, y + x)\big]$ [Empty 1 into 2]

A_6. $\forall xy\big[S(x, y) \wedge x + y > 7 \to S(x - (7 - y), 7)\big]$ [Pour 1 into 2]

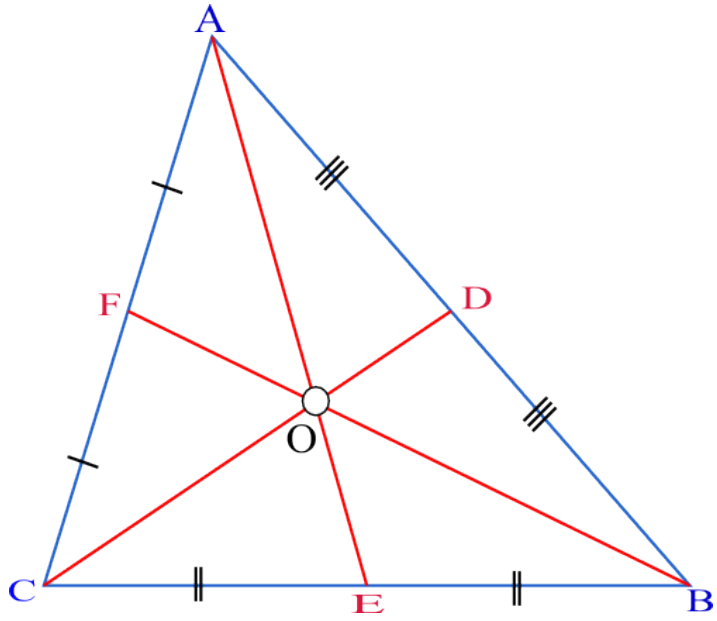
A_7. $\forall xy\big[S(x, y) \wedge x + y \leq 5 \to S(x + y, 0)\big]$ [Empty 2 into 1]

A_8. $\forall xy\big[S(x, y) \wedge x + y > 5 \to S(5, y - (5 - x))\big]$ [Pour 2 into 1]

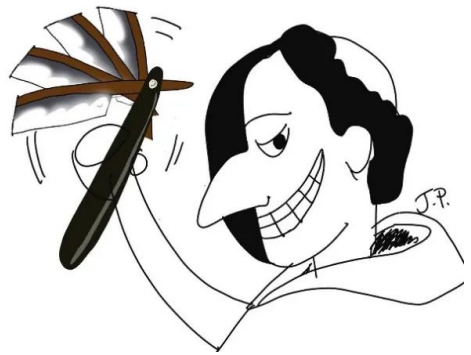
Initial State, $A_1, \ldots, A_8 \vdash$ Goal State

罗素的 “摹状词理论”

1. The substitution of identicals.
   “The morning star is the evening star.”
2. The law of the excluded middle.
   “The present King of France is bald.” or
   “The present King of France is not bald.”
3. The problem of negative existentials.
   “The flying horse does not exist.”







$$B(\iota_x A) := \exists x\big(Ax \wedge \forall y(Ay \to y = x) \wedge Bx\big)$$

- Oedipus did not know the woman he married was his mother.
  $\neg \mathsf{K}(\iota_x A = m(o))$?
  $\iota_x A = m(o) \to \mathsf{K}(m(o) = m(o)) \to \mathsf{K}(\iota_x A = m(o))$?
  $$\exists x\big(Ax \wedge \forall y(Ay \to y = x) \wedge \neg \mathsf{K}(x = m(o))\big)$$
  $$\neg \mathsf{K}\left[\exists x\big(Ax \wedge \forall y(Ay \to y = x) \wedge x = m(o)\big)\right]$$
- The present King of France is bald. $B(\iota_x K) \vee (\neg B)(\iota_x K)$?
  $$\exists x\big(Kx \wedge \forall y(Ky \to y = x) \wedge \neg Bx\big) \quad (\neg B)(\iota_x K)$$
  $$\neg\exists x\big(Kx \wedge \forall y(Ky \to y = x) \wedge Bx\big) \quad \neg B(\iota_x K)$$
- The flying horse does not exist. $\neg E(\iota_x Fx)$
  $$\exists x\big(Fx \wedge \forall y(Fy \to y = x) \wedge \neg Ex\big) \ ?$$
  $$\neg\exists x\big(Fx \wedge \forall y(Fy \to y = x)\big)$$
  $$Ex \stackrel{?}{:=} \exists P\big(Px \wedge \exists y \neg Py\big)$$
- Get rid of function symbols.
  $$\mathsf{Bald}(\iota_x \mathsf{Father}(x, \mathsf{alice})) \quad vs \quad \mathsf{Bald}(\mathsf{father}(\mathsf{alice}))$$
- Universal Instantiation. $\forall x B \to B(\iota_x A)$? $\quad \vdash \forall x B \to B(\iota_x^y A)$
  $$B(\iota_x^y A) := \big(\exists! x A \to \exists x(A \wedge B)\big) \wedge \big(\neg\exists! x A \to B[y/x]\big)$$

Translation

1. Every citizen of every country respects the King of that country.

$$\forall xy(Cy \wedge Zxy \rightarrow Rx\iota_z Kzy)$$

2. The daughter of the King of China is the person everyone respects.

$$\iota_y Dy\iota_x Kxc = \iota_x\big(Px \wedge \forall y(Py \rightarrow Ryx)\big)$$

3. The person everyone respects is a citizen of the country everyone respects.

$$Z\iota_x\big(Px \wedge \forall y(Py \rightarrow Ryx)\big)\iota_x\big(Cx \wedge \forall y(Py \rightarrow Ryx)\big)$$

罗素 Bertrand Russell 1872-1970

- 罗素悖论
  (第三次数学基础危机)
- 摹状词理论
  (当今的法国国王是秃子或不是秃子)
- 类型论
- 《数学原理》



没有理发师给并且只给那些不给自己理发的人理发. [18]

“哲学的特点：从简单得不值一提的东西开始，以荒谬得没人会相信的东西结束.”

— 罗素

[18] Russell: On denoting. 1905.

分析哲学与数学分析 — 摹状词理论的由来

- 一个陈述的**逻辑形式**可能不等于它的**语法形式**.
- 弗雷格的**语境原则**: 不要孤立的问一个词的意义, 一个词只有包含在上下文语境中才有意义. [19]
- 罗素的语境定义的思想隐含在 19 世纪的数学分析严格化之中.
- 贝克莱: 第二次数学基础危机. "无穷小是不是 0?"
  对 $f(x) = x^2$,

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{f(x+\mathrm{d}x) - f(x)}{\mathrm{d}x} = \frac{(x+\mathrm{d}x)^2 - x^2}{\mathrm{d}x} = \frac{2x\mathrm{d}x + (\mathrm{d}x)^2}{\mathrm{d}x} = 2x + \mathrm{d}x = 2x$$

- 维尔斯特拉斯: $\frac{\mathrm{d}f(x)}{\mathrm{d}x}$ 不是 $\mathrm{d}f(x)$ 与 $\mathrm{d}x$ 的商, $\frac{\mathrm{d}}{\mathrm{d}x}$ 作为微分运算作用于 $f(x)$,

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x} f(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

[19] Compositionality: The meaning of a whole (cf. sentence) should only depend on the meanings of its parts (cf. words) and how they are fitted together (cf. grammar).

三次数学基础危机

1. 万物皆 (有理) 数, $\sqrt{2}$ 是数吗?
2. 无穷小是数吗?
3. 什么是合法存在的 "集合"?

Remark

- 古巴比伦人和古埃及人的兴趣在于 5 个橘子而不是 "5".
- 正是古希腊人把数学变成一个抽象系统, 一种特殊的符号语言.
- 这使得人们不仅可以描述具体的现实世界, 而且可以解释它最深层次的模式和规律.
- 正因如此, 数学才得以从诸如 "如何辩护归纳法" 之类的迷宫问题中解放出来.

— 大卫·福斯特·华莱士

西方科学的发展是以两个伟大的成就为基础: 希腊哲学家发明的形式逻辑体系, 以及在文艺复兴时期发现通过系统的实验可能找出因果关系.

— 爱因斯坦

逻辑主义 & 逻辑实证主义

- 数学可以还原为逻辑.
- 科学可以还原为关于感觉材料的命题的逻辑复合.
- 只有可以被观察验证的或逻辑证明的命题才是有意义的.
- 弗雷格和罗素的新逻辑工具诱惑实证主义者做出了超出他们证明能力的猜想, 也让他们清楚的知道, 他们的猜想是不可证的.
- 在逻辑实证主义之前, 很少有哲学流派能够把自己的纲领陈述得足够清楚, 使得有可能看清其目标是不可达的.

卡尔纳普《通过语言的逻辑分析清除形而上学》

- 一个陈述的意义在于它的证实方法. 形而上学陈述不能被证实, 毫无意义.
- 那么留给哲学的还有什么呢? 一种方法: 逻辑分析法.
- 逻辑分析的消极应用是清除无意义的词和陈述, 积极应用是澄清有意义的概念和命题, 为经验科学和数学奠基.
- 形而上学家相信自己是在攸关真假的领域里前行, 却未断言任何东西. 他们只是试图表达一点儿人生态度.
- 艺术是表达人生态度的恰当手段. 抒情诗人并不企图在自己的诗里驳倒其他抒情诗人诗里的陈述, 但形而上学家却用论证维护他的陈述. 形而上学家是没有艺术才能的艺术家, 有的是在理论环境里工作的爱好, 却既不在科学领域里发挥这种爱好, 又不能满足艺术表达的要求, 倒是混淆了这两个方面, 创造出一种对知识既无贡献、对人生态度的表达又不相宜的东西.

休谟《人类理解研究》

如果我们拿起一本书，比如神学或经院哲学书，让我们问一下，其中包含着数和量方面的任何抽象推论么？没有．其中包含着事实和存在方面的任何经验推论么？没有．那就扔进火里吧，因为它所包含的没有别的，只有诡辩和幻想．



Russell's Theory of Descriptions & Church's λ-Abstraction

$$\nu(\iota_x A) = \begin{cases} a & \text{if there is a unique } a \in M : \mathcal{M}, \nu(a/x) \vDash A \\ \uparrow & \text{otherwise} \end{cases}$$

$$\begin{cases} \mathcal{M}, \nu \vDash (\lambda x.A)t \iff \mathcal{M}, \nu \vDash A[t/x] & \text{if } \nu(t) \downarrow \\ \mathcal{M}, \nu \nvDash (\lambda x.A)t & \text{if } \nu(t) \uparrow \end{cases}$$

The present King of France is not bald.

$$(\lambda x.\neg Bx)\iota_x Kx$$

It's not the case that the present King of France is bald.

$$\neg(\lambda x.Bx)\iota_x Kx$$

Crossing the street without looking is dangerous.

$$\mathbf{D}(\lambda x(Cx \wedge \neg Lx))$$

Every dog barks.

$$(\lambda P.\forall x(\mathrm{Dog}(x) \rightarrow P(x)))(\mathrm{Bark})$$

Expressive Limitation of First Order Language

- Most boys are funny.
- For every dog there is a cat. (There are more cats than dogs.)
- Some girls admire only one another.

$$\exists X\Big(\exists x Xx \wedge \forall x(Xx \to Gx) \wedge \forall x\forall y\big(Xx \wedge Axy \to Xy \wedge x \neq y\big)\Big)$$

- There are some gunslingers each of whom has shot the right foot of at least one of the others.

$$\exists X\Big(\exists x Xx \wedge \forall x(Xx \to Gx) \wedge \forall x\big(Xx \to \exists y(Xy \wedge y \neq x \wedge Sxy)\big)\Big)$$

- Least Number Principle.

$$\forall X\Big(\exists x Xx \wedge \forall x(Xx \to Nx) \to \exists x\big(Xx \wedge \forall y(Xy \wedge y \neq x \to x < y)\big)\Big)$$

- A linear order $(P, <)$ is *complete* iff every non-empty subset of P that is bounded above has a supremum in P.

$$\forall X\Big(\exists x Xx \wedge \exists y\forall x(Xx \to x \leq y) \to$$

$$\exists y\Big(\forall x(Xx \to x \leq y) \wedge \forall z\big(\forall x(Xx \to x \leq z) \to y \leq z\big)\Big)\Big)$$

Hilbert's Epsilon Calculus

$$\nu(\varepsilon_x A) = \Phi\left(\{a \in M : \mathcal{M}, \nu(a/x) \vDash A\}\right)$$

where $\Phi : \mathrm{P}(M) \to M :: \Phi(X) \in X$ whenever $X \neq \emptyset$ and $\Phi(\emptyset) \in M$.

Axiom ε

$$A(t) \to A(\varepsilon_x A)$$

where t is an arbitrary term.

ε-Extensionality Axiom

$$\forall x(A(x) \leftrightarrow B(x)) \to \varepsilon_x A = \varepsilon_x B$$

Hilbert's epsilon calculus is quantifier-free.
Predicate logic can be embedded in epsilon calculus.
Quantifiers can be defined as follows:

$$\exists x A(x) \equiv A(\varepsilon_x A)$$

$$\forall x A(x) \equiv A(\varepsilon_x \neg A)$$

Hilbert's Epsilon Calculus

- First ε-theorem: Suppose A is a quantifier-free and ε-free formula. Then $\vdash_{\varepsilon} A \implies \vdash_{\mathrm{QF}} A$ in quantifier-free predicate logic.
- Extended first ε-theorem: Suppose $\exists x_1 \ldots \exists x_n A(x_1, \ldots, x_n)$ is a purely existential formula containing only the bound variables $x_1, \ldots, x_n$. Then $\vdash_{\varepsilon} \exists x_1 \ldots \exists x_n A(x_1, \ldots, x_n) \implies \vdash \bigvee_i A(t_{i1}, \ldots, t_{in})$ for some terms t_{ij}.
- Second ε-theorem: Suppose A is an ε-free formula. Then $\vdash_{\varepsilon} A \implies \vdash A$.

Herbrand's Theorem is a corollary of the extended first ε-theorem.

Remark: Hilbert viewed the epsilon terms as representing the ideal elements that are added to finitistic reasoning to allow reasoning over infinite domains.

Tree Method with Unification

$\forall x A$ ✓

|

$A[x_i/x]$

where x_i is a new variable.

$\exists x A$ ✓

|

$A[f(x_1, \ldots, x_m)/x]$

where f is a new function and $\{x_1, \ldots, x_m\} = \mathrm{Fv}(\exists x A)$.

$\neg\forall x A$ ✓

|

$\exists x \neg A$

$\neg\exists x A$ ✓

|

$\forall x \neg A$

Tree Method with Unification

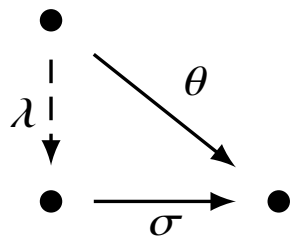
- when expanding a universally quantified formula, do not choose a specific term but a rigid variable as a placeholder.
- choose the term only when it is clear it allows closing a branch.

rigid variable=same value in the whole tree

- variables can assigned to closed terms, like $x_1 = a$.
- can also be assigned to unclosed terms, like $x_1 = f(x_2)$.
- make literals one the opposite of the other.
- using terms as unspecified as possible — Given literals A and $\neg B$ on the same branch, take the most general unifier of A and B.

Unifier

- A substitution σ is a *unifier* for a set Γ of formulas iff for every $A, B \in \Gamma : A\sigma = B\sigma$.
- A unifier σ is a *most general unifier* for Γ iff for each unifier θ there exists a substitution λ s.t. $\theta = \sigma\lambda$.



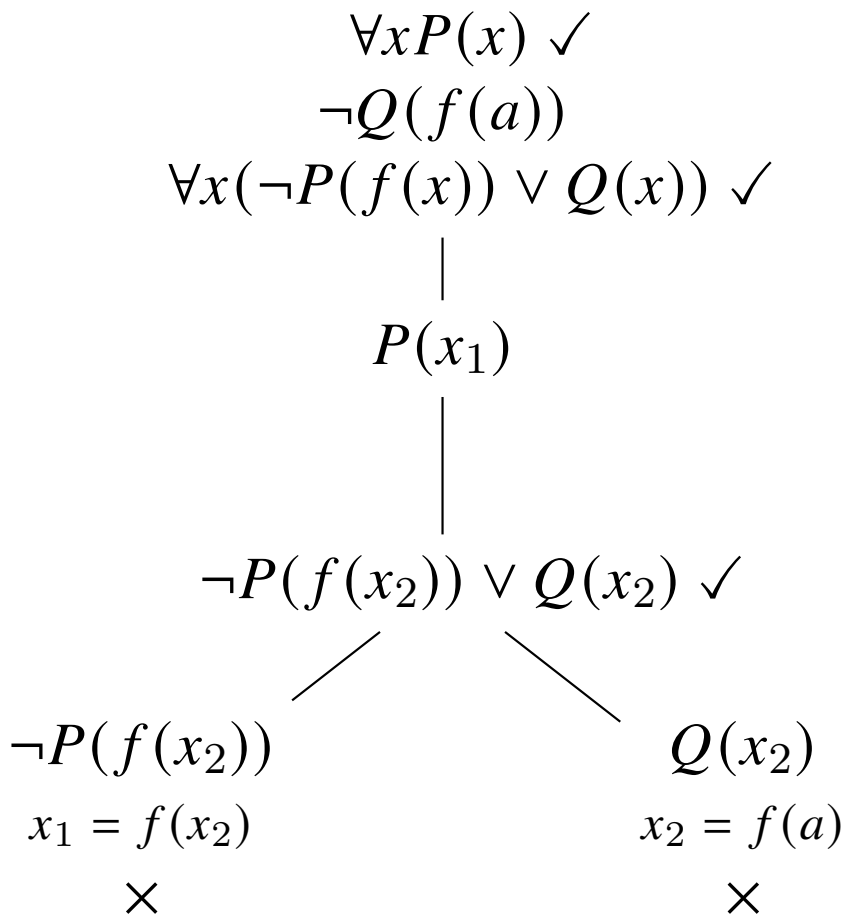
$$\sigma := \{t_1/x_1, \ldots, t_m/x_m\} \qquad \lambda := \{s_1/y_1, \ldots, s_n/y_n\}$$

$$\sigma\lambda = \{t_1\lambda/x_1, \ldots, t_m\lambda/x_m,\ s_1/y_1, \ldots, s_n/y_n\} \setminus \{s_i/y_i : y_i \in \{x_1, \ldots, x_m\}\}$$

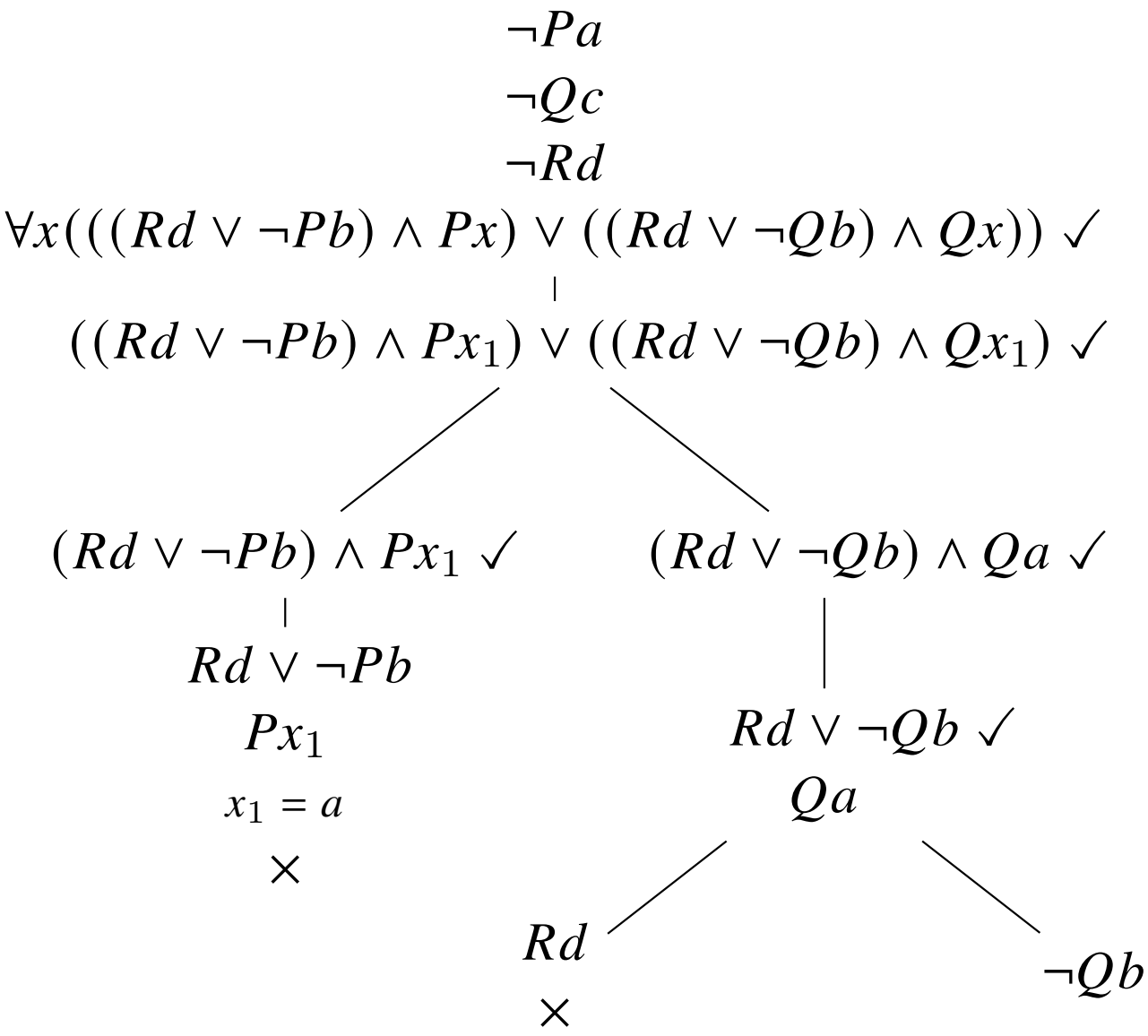
- $(A\sigma)\lambda = A(\sigma\lambda)$ and $(t\sigma)\lambda = t(\sigma\lambda)$
- $(\sigma\lambda)\theta = \sigma(\lambda\theta)$

Example — Tree Method with Unification

$\{\forall x P(x), \neg Q(f(a)), \forall x(\neg P(f(x)) \vee Q(x))\}$ is unsatisfiable.

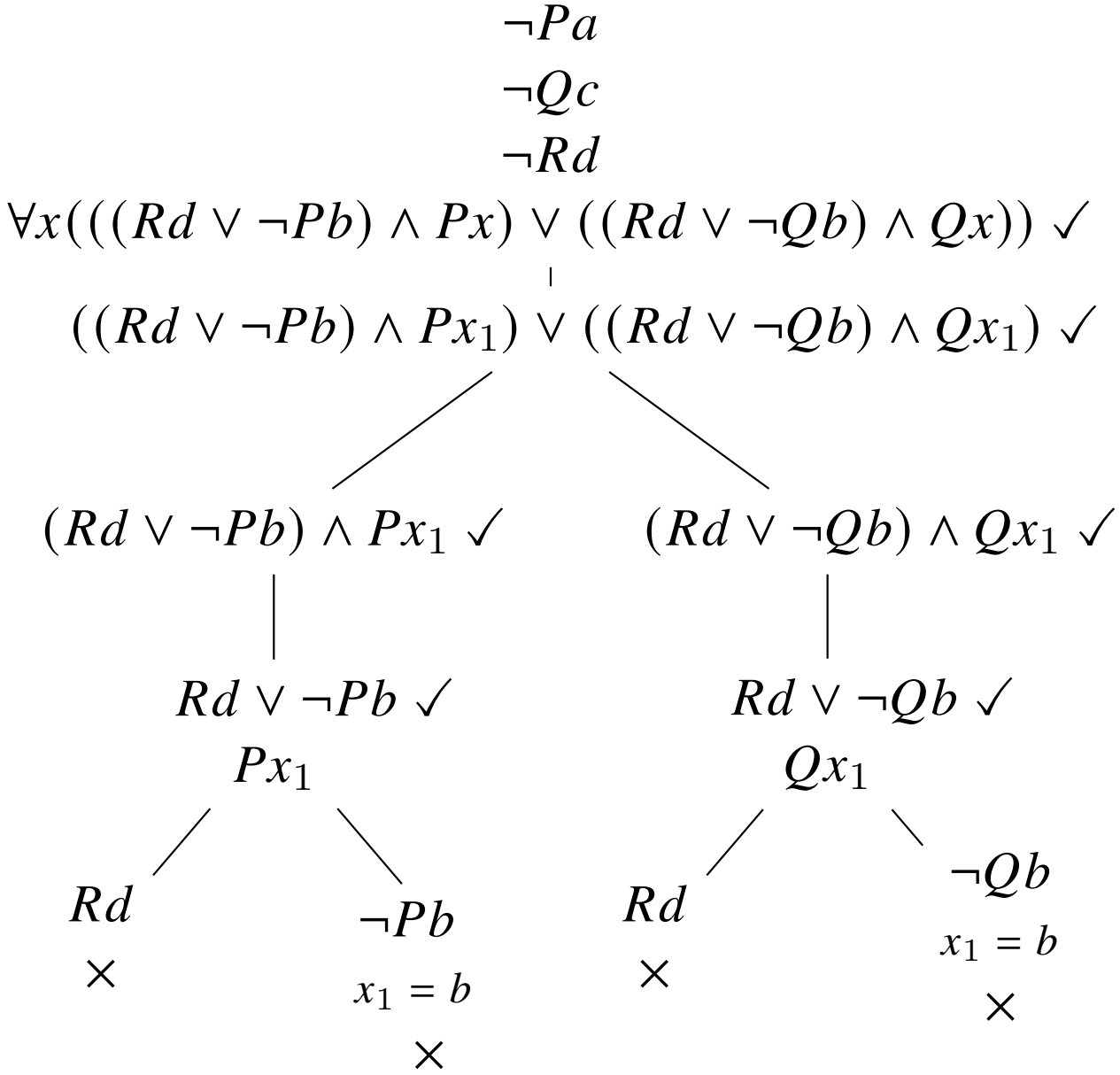


Unification — Greedy Unification (incomplete)



Applying unification as soon as a branch can be closed by lead to incompleteness.

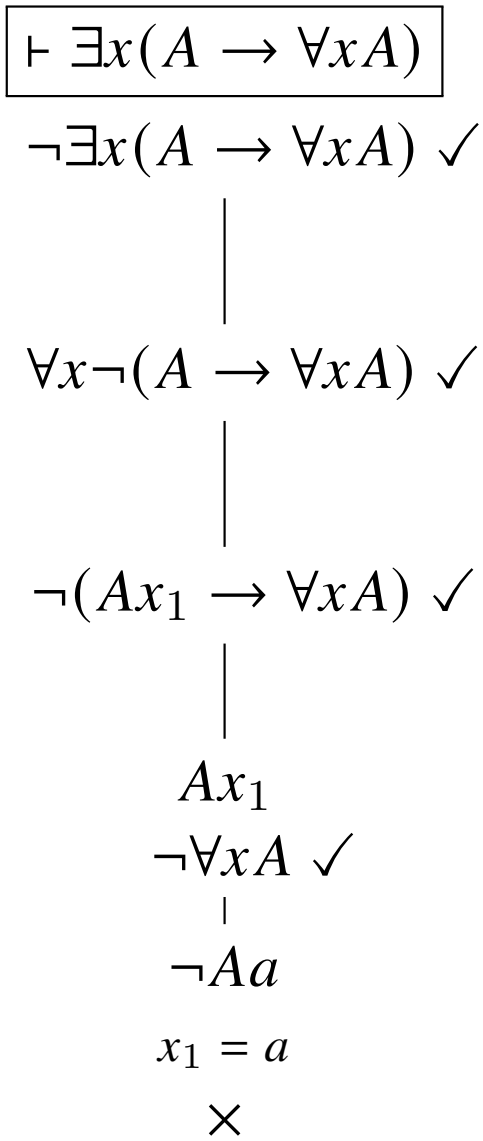
Unification — Final Closure

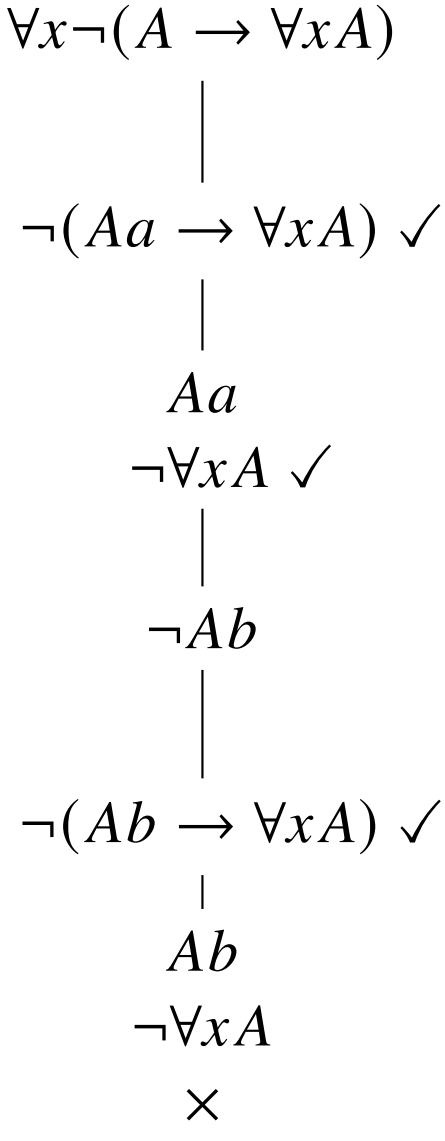


Unification is applied only when it closes all open branches at the same time.

Example — Unification vs Ground Tree

There is someone such that if he is drinking, then everyone is drinking.





Gentzen

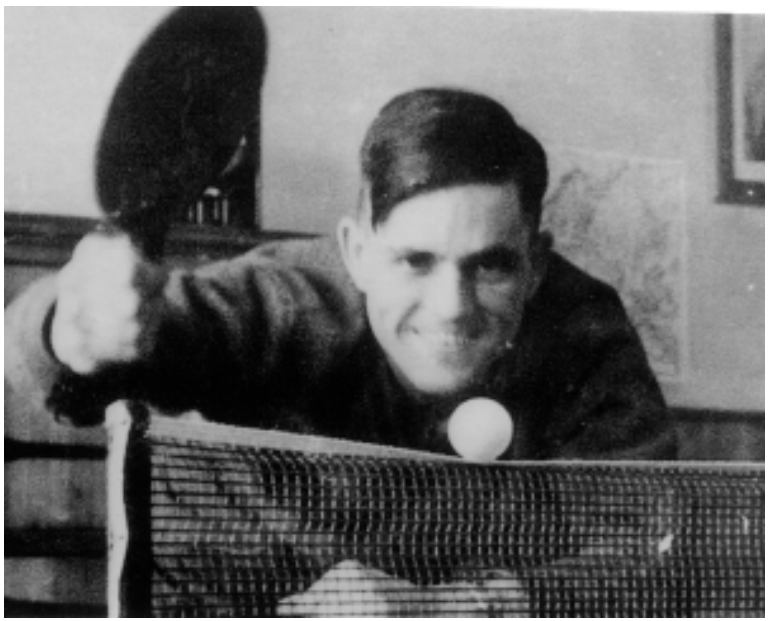


Figure: Gentzen 1909-1945

- Natural Deduction: one proposition on the right.
- Sequent Calculus: zero or more propositions on the right.

$$\Gamma \vdash \Delta \iff \vdash \bigwedge \Gamma \to \bigvee \Delta$$

- Consistency of PA
(proof-theoretical strength of PA)

Natural Deduction

$$\frac{A \quad B}{A \wedge B}\ \wedge^{+} \qquad \frac{A \wedge B}{A}\ \wedge^{-} \qquad \frac{A \wedge B}{B}\ \wedge^{-}$$

$$\frac{A}{A \vee B}\ \vee^{+} \qquad \frac{B}{A \vee B}\ \vee^{+} \qquad \frac{\begin{array}{ccc} A \vee B & \overset{[A]^n}{\vdots}\ C & \overset{[B]^n}{\vdots}\ C \end{array}}{C}\ \vee^{-n}$$

$$\frac{\begin{array}{c} [A]^n \\ \vdots \\ B \end{array}}{A \to B}\ \to^{+n} \qquad \frac{A \to B \quad A}{B}\ \to^{-}$$

$$\frac{\begin{array}{c} [A]^n \\ \vdots \\ \bot \end{array}}{\neg A}\ \neg^{+n} \qquad \frac{\begin{array}{c} [\neg A]^n \\ \vdots \\ \bot \end{array}}{A}\ \neg^{-n} \qquad \frac{\neg A \quad A}{\bot}\ \bot^{+} \qquad \frac{\bot}{A}\ \bot^{-}$$

Natural Deduction

$$\frac{A(a)}{\forall x A}\ \forall^+ \qquad\qquad \frac{\forall x A}{A[t/x]}\ \forall^-$$

where $a \notin \mathrm{Cst}(\forall x A)$, and a is not in any assumption which is uncanceled in the derivation ending with $A(a)$.

$$\frac{A[t/x]}{\exists x A}\ \exists^+ \qquad\qquad \frac{\exists x A \qquad \begin{array}{c}[A(a)]^n \\ \vdots \\ B\end{array}}{B}\ \exists^- n$$

where $a \notin \mathrm{Cst}(\exists x A, B)$, and a is not in any assumption which is uncanceled in the derivations ending with $\exists x A, B$ except in $A(a)$.

$$\frac{}{t = t}\ =^+ \qquad\qquad \frac{s = t \quad A[s/x]}{A[t/x]}\ =^-$$

Natural Deduction — another version — constant vs variable

$$\frac{A(a)}{\forall x A}\ \forall^+ \qquad \frac{A[y/x]}{\forall x A}\ \forall^+$$

where $a \notin \mathrm{Cst}(\forall x A)$, and a is not in any assumption which is uncanceled in the derivation ending with $A(a)$.

where $y \notin \mathrm{Fv}(\forall x A)$, and y is not free in any assumption which is uncanceled in the derivation ending with $A[y/x]$.

$$\frac{\exists x A \qquad \begin{array}{c}[A(a)]^n \\ \vdots \\ B\end{array}}{B}\ \exists^- n \qquad \frac{\exists x A \qquad \begin{array}{c}[A[y/x]]^n \\ \vdots \\ B\end{array}}{B}\ \exists^- n$$

where $a \notin \mathrm{Cst}(\exists x A, B)$, and a is not in any assumption which is uncanceled in the derivations ending with $\exists x A, B$ except in $A(a)$.

where $y \notin \mathrm{Fv}(\exists x A, B)$, and y is not free in any assumption which is uncanceled in the derivations ending with $\exists x A, B$ except in $A[y/x]$.

Remark: 怎么证明全称命题?

- 从给定集合中选择一个任意 (arbitrary) 对象, 并证明该对象具有所需的性质.

$$\frac{A(a)}{\forall x A}\ \forall^{+}$$

where $a \notin \mathrm{Cst}(\forall x A)$, and a is not in any assumption which is uncanceled in the derivation ending with $A(a)$.
Remark: a 不依赖任何特定的假设. 不预设任何信息确保了 a 的任意性.

- **"arbitrary" ≠ "random"**
  - 任意对象是一个占位符, 它可以代表任何其他对象.
  - 如果我们随机选择一个对象, 这意味着选择它的概率与选择任何其他对象的概率相等.

Remarks: 违反 “可代入” 条件导致的错误

$$\frac{\forall x \exists y Rxy}{\exists y Ryy}\ \forall^{-} \quad \textbf{incorrect} \quad \exists y Rxy[y/x] \times$$

$$\frac{\forall y Ryy}{\exists x \forall y Rxy}\ \exists^{+} \quad \textbf{incorrect} \quad \forall y Rxy[y/x] \times$$

$$\frac{x = y \quad \exists y Rxy}{\exists y Ryy}\ [=^{-}] \quad \textbf{incorrect} \quad \exists y Rxy[y/x] \times$$

Remarks: 违反 $\forall^+$ 的约束条件导致的错误

$$\frac{\forall x Rxx}{\forall x Rxa} \; ?$$

$$\dfrac{\dfrac{\forall x Rxx}{Raa}\,\forall^-}{\forall x Rxa}\,\forall^+ \qquad \textbf{incorrect}$$

$$\frac{}{\forall x(Ax \to \forall x Ax)} \; ?$$

$$\dfrac{\dfrac{\dfrac{[Aa]^1}{\forall x Ax}\,\forall^+}{Aa \to \forall x Ax}\,{\to}^{+1}}{\forall x(Ax \to \forall x Ax)}\,\forall^+ \qquad \textbf{incorrect}$$

Remarks: 违反 $\exists^-$ 的约束条件导致的错误

$$\boxed{\frac{\exists x A}{A a}\ ?}$$

$$\boxed{\frac{\forall x \exists y R x y}{\exists z R z z}\ ?}$$

$$\frac{\exists x A \qquad \begin{array}{c}[A a]^1 \\ \vdots \\ A a\end{array}}{A a}\ \exists^{-1} \qquad \textbf{incorrect}$$

$$\frac{\dfrac{\forall x \exists y R x y}{\exists y R a y}\ \forall^- \qquad \dfrac{[R a a]^1}{\exists z R z z}\ \exists^+}{\exists z R z z}\ \exists^{-1} \qquad \textbf{incorrect}$$

$$\boxed{\frac{\exists x A \quad \exists x B}{\exists x (A \wedge B)}\ ?}$$

$$\frac{\exists x B \qquad \dfrac{\exists x A \qquad \dfrac{\dfrac{[A a]^1 \quad [B a]^2}{A a \wedge B a}\ \wedge^+}{\exists x (A \wedge B)}\ \exists^+}{\exists x (A \wedge B)}\ \exists^{-1}}{\exists x (A \wedge B)}\ \exists^{-2} \qquad \textbf{incorrect}$$

Example

$$\boxed{\dfrac{\forall x(A \to B)}{\exists x A \to \exists x B}}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\forall x(A \to B)}{Aa \to Ba}\,\forall^{-} \qquad [Aa]^{1}}{Ba}\,{\to}^{-}}{\dfrac{[\exists x A]^{2} \qquad \exists x B}{\exists x B}\,\exists^{-1}}\,{\to}^{+2}\,\exists^{+}$$

□

Examples

$$\frac{\forall x(A \to B)}{A \to \forall x B} \quad \text{where } x \notin \mathrm{Fv}(A)$$

Proof.

$$\dfrac{\dfrac{\dfrac{\forall x(A \to B)}{A \to Ba}\,\forall^- \qquad [A]^1}{Ba}\,\to^-}{\dfrac{\forall x B}{A \to \forall x B}\,\forall^+}\,\to^{+1}$$

□

$$\frac{A \to \forall x B}{\forall x(A \to B)} \quad \text{where } x \notin \mathrm{Fv}(A)$$

Proof.

$$\dfrac{\dfrac{\dfrac{A \to \forall x B \qquad [A]^1}{\forall x B}\,\to^-}{Ba}\,\forall^-}{\dfrac{A \to Ba}{\forall x(A \to B)}\,\forall^+}\,\to^{+1}$$

□

Examples

$$\frac{\forall x(A \to B)}{\exists x A \to B} \quad \text{where } x \notin \mathrm{Fv}(B)$$

Proof.

$$\dfrac{\dfrac{[\exists x A]^1 \qquad \dfrac{\dfrac{[Aa]^2 \qquad \dfrac{\forall x(A \to B)}{Aa \to B}\,\forall^-}{B}\,\to^-}{B}\,\exists^{-2}}{\exists x A \to B}\,\to^{+3}$$

□

$$\frac{\forall x(A \to B)}{\forall x(A \to B)} \quad \text{where } x \notin \mathrm{Fv}(B)$$

Proof.

$$\dfrac{\dfrac{[Aa]^1}{\exists x A}\,\exists^+ \qquad \exists x A \to B}{\dfrac{\dfrac{B}{Aa \to B}\,\to^{+1}}{\forall x(A \to B)}\,\forall^+}\,\to^-$$

□

Examples

$$\boxed{\frac{\exists x \forall y Rxy}{\forall y \exists x Rxy}}$$

Proof.

$$\dfrac{\exists x \forall y Rxy \qquad \dfrac{\dfrac{\dfrac{[\forall y Ray]^1}{Rab}\,\forall^-}{\exists x Rxb}\,\exists^+}{\forall y \exists x Rxy}\,\forall^+}{\forall y \exists x Rxy}\,\exists^{-1}$$

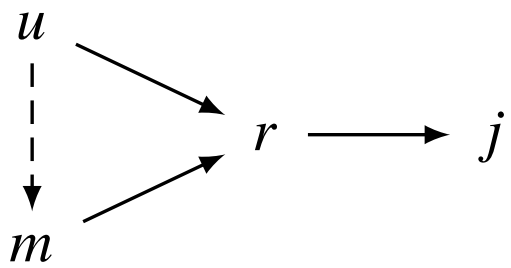
□

Examples

Everybody loves anybody who loves somebody

If somebody loves somebody, everybody loves everybody

$$\boxed{\frac{\forall xyz(Lyz \to Lxy)}{\exists xyLxy \to \forall xyLxy}}$$



Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{\forall xyz(Lyz \to Lxy)}{Lrj \to Lmr} \quad [Lrj]^1}{Lmr} \quad \dfrac{\forall xyz(Lyz \to Lxy)}{Lmr \to Lum}}{Lum}}{\dfrac{[\exists xyLxy]^2 \qquad Lum}{Lum}\ \exists^{-1}}$$

$$\dfrac{\dfrac{Lum}{\forall xyLxy}}{\exists xyLxy \to \forall xyLxy}\ \to^{+2}$$

□

Example

If anyone speaks to anyone, then someone introduces them
No one introduces anyone to anyone unless he knows them both
Everyone speaks to Alice

Everyone is introduced to Alice by someone who knows her

$$\frac{\begin{array}{c}\forall x\forall y(Sxy \to \exists z Izxy)\\ \forall x\forall y\forall z(Izxy \to Kzx \wedge Kzy)\\ \forall x Sxa\end{array}}{\forall x\exists y(Iyxa \wedge Kya)}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{\forall xy(Sxy \to \exists z Izxy)}{Sba \to \exists z Izba} \quad \dfrac{\forall x Sxa}{Sba}}{\exists z Izba} \qquad \dfrac{\dfrac{[Icba]^1 \quad \dfrac{\forall xyz(Izxy \to Kzx \wedge Izy)}{Icba \to Kcb \wedge Kca}}{Kca}}{Icba \wedge Kca}}{\exists y(Iyba \wedge Kya)}\,\exists^{-1}}{\exists y(Iyba \wedge Kya)}}{\forall x\exists y(Iyxa \wedge Kya)}$$

Example

$$\frac{\text{If dogs are animals}}{\text{Every head of a dog is the head of an animal}}$$

$$\boxed{\frac{\forall x(Dx \to Ax)}{\forall x(Dx \to \exists y(Ay \wedge hx = hy))}}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall x(Dx \to Ax)}{Da \to Aa} \quad [Da]^1}{Aa} \quad \dfrac{}{ha = ha}}{Aa \wedge ha = ha}}{\exists y(Ay \wedge ha = hy)}}{Da \to \exists y(Ay \wedge ha = hy)} \to^{+1}$$
$$\frac{}{\forall x(Dx \to \exists y(Ay \wedge hx = hy))}$$

□

Example

$$\frac{\text{如果大鱼比小鱼游得快}}{\text{只要有最大的鱼就有游得最快的鱼}}$$

$$\frac{\forall x \forall y (Fx \wedge Fy \wedge Bxy \to Sxy)}{\exists x (Fx \wedge \forall y (Fy \to Bxy)) \to \exists x (Fx \wedge \forall y (Fy \to Sxy))}$$

Proof.

$$\dfrac{\dfrac{[\exists x (Fx \wedge \forall y (Fy \to Bxy))]^3 \qquad \dfrac{\dfrac{\dfrac{\dfrac{[Fa \wedge \forall y (Fy \to Bay)]^1}{Fa} \qquad \dfrac{[Fa \wedge \forall y (Fy \to Bay)]^1}{\forall y (Fy \to Bay)}}{Fb \to Bab} \qquad [Fb]^2}{Bab}}{Fa \wedge Fb \wedge Bab} \qquad \dfrac{\forall x \forall y (Fx \wedge Fy \wedge Bxy \to Sxy)}{Fa \wedge Fb \wedge Bab \to Sab}}{Sab}}{Fb \to Sab} \to^{+2}}{\forall y (Fy \to Say)}}{Fa \wedge \forall y (Fy \to Say)}}{\exists x (Fx \wedge \forall y (Fy \to Sxy))} \exists^{-1}}{\exists x (Fx \wedge \forall y (Fy \to Sxy))}}{\exists x (Fx \wedge \forall y (Fy \to Bxy)) \to \exists x (Fx \wedge \forall y (Fy \to Sxy))} \to^{+3}$$

Example: 偶数的平方是偶数

$$\dfrac{\dfrac{\dfrac{[E(a)]^1}{\exists y(a = 2y)} \quad \dfrac{\dfrac{[a = 2b]^2}{a^2 = 2(2b^2)}}{\exists y(a^2 = 2y)}}{\exists y(a^2 = 2y)}\,\exists^{-2}}{\dfrac{E(a^2)}{E(a) \to E(a^2)}\,{\to}^{+1}}$$
$$\forall x(E(x) \to E(x^2))$$

归谬法 — 比萨斜塔思想实验

$$\dfrac{\dfrac{\dfrac{[\forall xy(m_x < m_y \to v_x < v_y)]^1 \quad \forall xy(v_x < v_y \to v_x < v_{x+y} < v_y)}{\vdots \qquad\qquad\qquad \vdots}}{v_{重} < v_{轻+重} \qquad\qquad v_{轻} < v_{轻+重} < v_{重}}}{\dfrac{\dfrac{v_{重} < v_{重}}{\bot}}{\neg\forall xy(m_x < m_y \to v_x < v_y)}\ \neg^{+1}}$$



哲学写在宇宙这本大书上.
这本书永远向我们敞开.
但除非先学会它的语言,
否则, 我们将一个字都理解不了,
从而只能在黑暗的迷宫中徒劳地游荡.
这本书是用数学语言写成的.

— 伽利略

$$P \wedge O \wedge (\neg L \to \neg U) \wedge (\neg U \to W) \wedge M$$

Natural Deduction — another version

$$\frac{}{\Gamma \vdash A}\ \mathsf{I}\ A \in \Gamma \qquad \frac{\Gamma \vdash A}{\Gamma' \vdash A}\ \mathsf{M}\ \Gamma \subset \Gamma'$$

$$\frac{\Gamma_1 \vdash A \quad \Gamma_2 \vdash B}{\Gamma_1, \Gamma_2 \vdash A \wedge B}\ \wedge^+ \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A}\ \wedge^- \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}\ \wedge^-$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B}\ \vee^+ \qquad \frac{\Gamma \vdash A}{\Gamma \vdash B \vee A}\ \vee^+ \qquad \frac{\Gamma_1 \vdash A \vee B \quad \Gamma_2, A \vdash C \quad \Gamma_3, B \vdash C}{\Gamma_1, \Gamma_2, \Gamma_3 \vdash C}\ \vee^-$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B}\ \to^+ \qquad \frac{\Gamma_1 \vdash A \to B \quad \Gamma_2 \vdash A}{\Gamma_1, \Gamma_2 \vdash B}\ \to^-$$

$$\frac{\Gamma, A \vdash \bot}{\Gamma \vdash \neg A}\ \neg^+ \qquad \frac{\Gamma, \neg A \vdash \bot}{\Gamma \vdash A}\ \neg^- \qquad \frac{\Gamma_1 \vdash \neg A \quad \Gamma_2 \vdash A}{\Gamma_1, \Gamma_2 \vdash \bot}\ \bot^+ \qquad \frac{\Gamma \vdash \bot}{\Gamma \vdash A}\ \bot^-$$

Natural Deduction — another version

$$\frac{\Gamma \vdash A(a)}{\Gamma \vdash \forall x A}\ \forall^+\ a \notin \mathrm{Cst}(\Gamma, \forall x A) \qquad \frac{\Gamma \vdash \forall x A}{\Gamma \vdash A[t/x]}\ \forall^-$$

$$\frac{\Gamma \vdash A[t/x]}{\Gamma \vdash \exists x A}\ \exists^+ \qquad \frac{\Gamma_1 \vdash \exists x A \qquad \Gamma_2, A(a) \vdash B}{\Gamma_1, \Gamma_2 \vdash B}\ \exists^-\ a \notin \mathrm{Cst}(\Gamma_1, \Gamma_2, \exists x A, B)$$

$$\frac{}{\Gamma \vdash t = t}\ =^+ \qquad \frac{\Gamma_1 \vdash s = t \quad \Gamma_2 \vdash A[s/x]}{\Gamma_1, \Gamma_2 \vdash A[t/x]}\ =^-$$

Natural Deduction — constant vs variable

$$\frac{\Gamma \vdash A(a)}{\Gamma \vdash \forall x A}\ \forall^{+}\ a \notin \mathrm{Cst}(\Gamma, \forall x A)$$

$$\frac{\Gamma \vdash A[y/x]}{\Gamma \vdash \forall x A}\ \forall^{+}\ y \notin \mathrm{Fv}(\Gamma, \forall x A)$$

$$\frac{\Gamma_1 \vdash \exists x A \qquad \Gamma_2, A(a) \vdash B}{\Gamma_1, \Gamma_2 \vdash B}\ \exists^{-}\ a \notin \mathrm{Cst}(\Gamma_1, \Gamma_2, \exists x A, B)$$

$$\frac{\Gamma_1 \vdash \exists x A \qquad \Gamma_2, A[y/x] \vdash B}{\Gamma_1, \Gamma_2 \vdash B}\ \exists^{-}\ y \notin \mathrm{Fv}(\Gamma_1, \Gamma_2, \exists x A, B)$$

Examples

$\boxed{\neg A \to \bot \vdash A}$

$$\dfrac{\dfrac{\dfrac{}{\neg A \to \bot \vdash \neg A \to \bot}\,\mathrm{I} \quad \dfrac{}{\neg A \vdash \neg A}\,\mathrm{I}}{\neg A \to \bot, \neg A \vdash \bot}\,{\to}^{-}}{\neg A \to \bot \vdash A}\,\neg^{-}$$

$\boxed{\text{If } x \notin \mathrm{Fv}(A), \text{ then } \vdash \forall x(A \to B) \to A \to \forall x B}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\forall x(A \to B) \vdash \forall x(A \to B)}\,\mathrm{I}}{\forall x(A \to B) \vdash A \to Ba}\,\forall^{-} \quad \dfrac{}{A \vdash A}\,\mathrm{I}}{\forall x(A \to B), A \vdash Ba}\,{\to}^{-}}{\forall x(A \to B), A \vdash \forall x B}\,\forall^{+}}{\forall x(A \to B) \vdash A \to \forall x B}\,{\to}^{+}}{\vdash \forall x(A \to B) \to A \to \forall x B}\,{\to}^{+}$$

Sequent Calculus

Axiom

$$\frac{}{A \vdash A}\ \mathsf{I}$$

Cut

$$\frac{\Gamma_1 \vdash \Delta_1, A \quad \Gamma_2, A \vdash \Delta_2}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\ \mathsf{Cut}$$

Left structural rules

$$\frac{\Gamma \vdash \Delta}{\Gamma, A \vdash \Delta}\ \mathsf{WL}$$

$$\frac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta}\ \mathsf{CL}$$

$$\frac{\Gamma_1, A, B, \Gamma_2 \vdash \Delta}{\Gamma_1, B, A, \Gamma_2 \vdash \Delta}\ \mathsf{PL}$$

Right structural rules

$$\frac{\Gamma \vdash \Delta}{\Gamma \vdash A, \Delta}\ \mathsf{WR}$$

$$\frac{\Gamma \vdash A, A, \Delta}{\Gamma \vdash A, \Delta}\ \mathsf{CR}$$

$$\frac{\Gamma \vdash \Delta_1, A, B, \Delta_2}{\Gamma \vdash \Delta_1, B, A, \Delta_2}\ \mathsf{PR}$$

Sequent Calculus

Left logical rules:

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta} \wedge L$$

$$\frac{\Gamma, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta} \wedge L$$

$$\frac{\Gamma_1, A \vdash \Delta_1 \quad \Gamma_2, B \vdash \Delta_2}{\Gamma_1, \Gamma_2, A \vee B \vdash \Delta_1, \Delta_2} \vee L$$

$$\frac{\Gamma_1 \vdash A, \Delta_1 \quad \Gamma_2, B \vdash \Delta_2}{\Gamma_1, \Gamma_2, A \to B \vdash \Delta_1, \Delta_2} \to L$$

Right logical rules:

$$\frac{\Gamma \vdash A, \Delta}{\Gamma \vdash A \vee B, \Delta} \vee R$$

$$\frac{\Gamma \vdash B, \Delta}{\Gamma \vdash A \vee B, \Delta} \vee R$$

$$\frac{\Gamma_1 \vdash A, \Delta_1 \qquad \Gamma_2 \vdash B, \Delta_2}{\Gamma_1, \Gamma_2 \vdash A \wedge B, \Delta_1, \Delta_2} \wedge R$$

$$\frac{\Gamma, A \vdash B, \Delta}{\Gamma \vdash A \to B, \Delta} \to R$$

Sequent Calculus

Left logical rules: Right logical rules:

$$\frac{\Gamma \vdash A, \Delta}{\Gamma, \neg A \vdash \Delta}\ \neg L \qquad \frac{\Gamma, A \vdash \Delta}{\Gamma \vdash \neg A, \Delta}\ \neg R$$

$$\frac{\Gamma, A[t/x] \vdash \Delta}{\Gamma, \forall x A \vdash \Delta}\ \forall L \qquad \frac{\Gamma \vdash A(a), \Delta}{\Gamma \vdash \forall x A, \Delta}\ \forall R\ a \notin \mathrm{Cst}(\Gamma, \forall x A, \Delta)$$

$$\frac{\Gamma, A(a) \vdash \Delta}{\Gamma, \exists x A \vdash \Delta}\ \exists L\ a \notin \mathrm{Cst}(\Gamma, \exists x A, \Delta) \qquad \frac{\Gamma \vdash A[t/x], \Delta}{\Gamma \vdash \exists x A, \Delta}\ \exists R$$

Sequent Calculus — constant vs variable

$$\frac{\Gamma \vdash A(a), \Delta}{\Gamma \vdash \forall x A, \Delta}\ \forall R\ a \notin \mathrm{Cst}(\Gamma, \forall x A, \Delta)$$

$$\frac{\Gamma \vdash A[y/x], \Delta}{\Gamma \vdash \forall x A, \Delta}\ \forall R\ y \notin \mathrm{Fv}(\Gamma, \forall x A, \Delta)$$

$$\frac{\Gamma, A(a) \vdash \Delta}{\Gamma, \exists x A \vdash \Delta}\ \exists L\ a \notin \mathrm{Cst}(\Gamma, \exists x A, \Delta)$$

$$\frac{\Gamma, A[y/x] \vdash \Delta}{\Gamma, \exists x A \vdash \Delta}\ \exists L\ y \notin \mathrm{Fv}(\Gamma, \exists x A, \Delta)$$

Examples

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}\,\mathsf{I}}{\vdash \neg A, A}\,\neg R}{\vdash \neg A \vee A, A}\,\vee R}{\vdash A, A \vee \neg A}\,\mathrm{PR}}{\vdash A \vee \neg A, A \vee \neg A}\,\vee R \qquad \vdash A \vee \neg A\,\mathrm{CR}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}\,\mathsf{I}}{A \vdash B, A}\,\mathrm{WR}}{A, \neg A \vdash B}\,\neg L \qquad \dfrac{\dfrac{}{B \vdash B}\,\mathsf{I}}{A, B \vdash B}\,\mathrm{WL}}{A, \neg A \vee B \vdash B}\,\vee L \qquad \neg A \vee B \vdash A \to B\,\to R$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{Aa \vdash Aa}\,\mathsf{I}}{\forall x A \vdash Aa}\,\forall L}{\forall x A, \neg Aa \vdash}\,\neg L}{\neg Aa \vdash \neg \forall x A}\,\neg R \qquad \exists x \neg A \vdash \neg \forall x A\,\exists L$$

$$\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}\,\mathsf{I}}{A \wedge B \vdash A}\,\wedge L}{A \wedge B, \neg A \vdash}\,\neg L \qquad \dfrac{\dfrac{\dfrac{}{B \vdash B}\,\mathsf{I}}{A \wedge B \vdash B}\,\wedge L}{A \wedge B, \neg B \vdash}\,\neg L}{A \wedge B, \neg A \vee \neg B \vdash}\,\vee L \qquad \neg A \vee \neg B \vdash \neg(A \wedge B)\,\neg R$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{Rab \vdash Rab}\,\mathsf{I}}{\forall x Rxb \vdash Rab}\,\forall L}{\forall x Rxb \vdash \exists y Ray}\,\exists R}{\exists y \forall x Rxy \vdash \exists y Ray}\,\exists L \qquad \exists y \forall x Rxy \vdash \forall x \exists y Rxy\,\forall R$$

Cut-Elimination Theorem

Theorem (Cut-Elimination Theorem — Gentzen1934)

If $\Gamma \vdash \Delta$ is provable, then it is provable without using the Cut Rule.

Corollary (The Subformula Property)

If $\Gamma \vdash \Delta$ is provable, then it has a deduction sequence all of whose formulas are subformulas of Γ and Δ.

Corollary (Consistency)

A contradiction, i.e. the empty sequent $\emptyset \vdash \emptyset$, is not deducible.

Corollary (Conservation)

Predicate logic is conservative over propositional logic.

Theorem (Cut-free Completeness Theorem)

Let Θ be a set of sentences. If Θ logically implies $\Gamma \vdash \Delta$, then there is a finite subset $\Sigma \subset \Theta$ s.t. $\Sigma, \Gamma \vdash \Delta$ has a cut-free proof.

Hilbert System = Axiom + Inference Rule

公理模式

1. $A \to B \to A$
2. $(A \to B \to C) \to (A \to B) \to A \to C$
3. $(\neg A \to \neg B) \to (\neg A \to B) \to A$
4. $\forall x(A \to B) \to \forall x A \to \forall x B$
5. $\forall x A \to A[t/x]$ where t is substitutable for x in A.
6. $A \to \forall x A$ where $x \notin \mathrm{Fv}(A)$.
7. $x = x$
8. $x = y \to A \to A[y/\!\!/x]$ where A is atomic and $A[y/\!\!/x]$ is obtained from A by replacing x in one or more places by y.
9. $\forall x_1 \ldots x_n A$ where $n \geq 0$ and A is any axiom of the preceding groups.

推理规则

$$\frac{A \quad A \to B}{B}\ \text{MP}$$

Example

Theorem

$$A \vdash \exists x A$$

Proof.

1.	$(\forall x \neg A \to \neg A) \to A \to \neg \forall x \neg A$	Tautology
2.	$\forall x \neg A \to \neg A$	A5
3.	$A \to \neg \forall x \neg A$	1,2 MP
4.	A	Premise
5.	$\neg \forall x \neg A$	3,4 MP
6.	$\exists x\, A$	Definition of $\exists$

□

演绎定理

Theorem (Deduction Theorem1)

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B}$$

Inference Rule

$$\frac{A}{\forall x A}\ \mathsf{G}$$

What if we remove Axiom9 and add the rule of generalization G[20] to Hilbert System?

Theorem (Deduction Theorem2)

If $\Gamma, A \vdash B$, where the rule of generalization is not applied to the free variables of A, then $\Gamma \vdash A \to B$.

[20] MP "保赋值", G "保真".

Meta-properties

- Tautology. For $A \in \mathscr{L}^0$, $B_1, \ldots, B_n \in \mathscr{L}^1$:
$$\frac{\vdash A}{\vdash A[B_1/p_1, \ldots, B_n/p_n]}$$
- reductio ad absurdum / proof by contradiction
$$\frac{\Gamma, A \vdash B \quad \Gamma, A \vdash \neg B}{\Gamma \vdash \neg A} \qquad \frac{\Gamma, \neg A \vdash B \quad \Gamma, \neg A \vdash \neg B}{\Gamma \vdash A}$$
- contraposition
$$\frac{\Gamma, A \vdash \neg B}{\Gamma, B \vdash \neg A}$$
- substitution
$$\frac{s = t}{r[s/x] = r[t/x]} \qquad \frac{s = t}{A[s/x] \leftrightarrow A[t/x]}$$
- equivalent replacement
$$\frac{\vdash B \leftrightarrow C}{\vdash A \leftrightarrow A[C/\!\!/B]}$$
where $A[C/\!\!/B]$ arises from A by replacing one or more occurrences of B in A by C.

Meta-properties

$$\frac{\Gamma, A[t/x] \vdash B}{\Gamma, \forall x A \vdash B}\ \forall L \qquad \frac{\Gamma \vdash A[t/x]}{\Gamma \vdash \exists x A}\ \exists R$$

$$\frac{\Gamma, A \vdash B}{\Gamma, \exists x A \vdash B}\ \exists L\ x \notin \mathrm{Fv}(\Gamma, B) \qquad \frac{\Gamma \vdash A}{\Gamma \vdash \forall x A}\ \forall R\ x \notin \mathrm{Fv}(\Gamma)$$

$$\frac{\Gamma, A[y/x] \vdash B}{\Gamma, \exists x A \vdash B}\ \exists L\ y \notin \mathrm{Fv}(\Gamma, \exists x A, B) \qquad \frac{\Gamma \vdash A[y/x]}{\Gamma \vdash \forall x A}\ \forall R\ y \notin \mathrm{Fv}(\Gamma, \forall x A)$$

$$\frac{\Gamma, A[a/x] \vdash B}{\Gamma, \exists x A \vdash B}\ \exists L\ a \notin \mathrm{Cst}(\Gamma, \exists x A, B) \qquad \frac{\Gamma \vdash A[a/x]}{\Gamma \vdash \forall x A}\ \forall R\ a \notin \mathrm{Cst}(\Gamma, \forall x A)$$

Theorem (Existence of Alphabetic Variants)

Let A be a formula, t a term, and x a variable. Then we can find a formula A^* which differs from A only in the choice of quantified variables s.t.

1. $\vdash A \leftrightarrow A^*$
2. *t is substitutable for x in A^*.*

Proof Tactics

$\rightarrow$
- $\Gamma \vdash A \rightarrow B \iff \Gamma, A \vdash B$

$\forall$
1. if $x \notin \mathrm{Fv}(\Gamma)$, $\Gamma \vdash \forall x A \iff \Gamma \vdash A$
2. if $x \in \mathrm{Fv}(\Gamma)$, $\Gamma \vdash \forall x A \iff \Gamma \vdash \forall y A[y/x] \iff \Gamma \vdash A[y/x]$ for some new y.

$\neg$
1. $(\neg\rightarrow)$ $\Gamma \vdash \neg(A \rightarrow B) \iff \Gamma \vdash A \ \& \ \Gamma \vdash \neg B$
2. $(\neg\neg)$ $\Gamma \vdash \neg\neg A \iff \Gamma \vdash A$
3. $(\neg\forall)$ $\Gamma \vdash \neg\forall x A \iff \Gamma \vdash \neg A[t/x]$

   Unfortunately this is not always possible. Try contraposition, reductio ad absurdum or prove by contradiction...

什么是一个理论? — 推演封闭的句子集 

一个公理化的理论由哪几部分组成?

一: 基本概念 (初始符号)

二: 逻辑公理 (有效式/普遍真理)

三: 非逻辑公理 (相对真理)

四: 推理规则

面对一个理论, 我们关心什么?

1. 语言表达力够丰富吗?
2. 可递归公理化吗?
3. 一致吗?
4. 完备吗?
5. 公理独立吗?
6. 能证明自身的一致性吗?
7. 可判定吗?
8. 有范畴性吗?

一些理论的例子:

- 欧几里得几何理论
- 皮亚诺算术理论
- 集合论
- 群论
- 概率论
- 牛顿力学
- 狭义相对论
- 广义相对论
- 量子力学
- 进化论
- 博弈论
- 马克思主义理论
- 五行生克理论

Example — 欧几里得几何理论

欧几里得《几何原本》

公理 1 从一点到另一点可以作一条直线.

公理 2 一条线段可以延伸成一条直线.

公理 3 以线段的一个端点为圆心, 该线段为半径, 可以作一个圆.

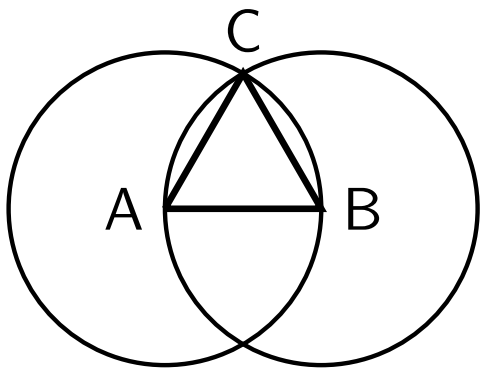
公理 4 所有直角都相等.

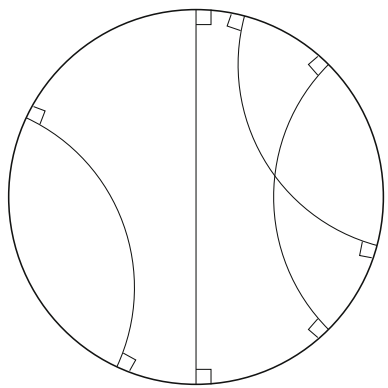
公理 5 过直线外一点, 有且只有一条直线与之平行.

定理

给定一条线段, 可以作一个以其为边的等边三角形.

1. $\odot A$ 公理 3
2. $\odot B$ 公理 3
3. AC 公理 1
4. BC 公理 1





Example — 群论

群 $\mathscr{L} = \{e, \cdot\}$

公理 1 $\forall xyz : x \cdot (y \cdot z) = (x \cdot y) \cdot z$

公理 2 $\forall x : e \cdot x = x \cdot e = x$

公理 3 $\forall x \exists y : x \cdot y = y \cdot x = e$

模型

- $(\mathbb{Z}, 0, +)$
- $(\mathbb{Q}^+, 1, \times)$
- Klein group: $(\{\diamond, \heartsuit, \spadesuit, \clubsuit\}, \diamond, \cdot)$

$\cdot$	$\diamond$	$\heartsuit$	$\spadesuit$	$\clubsuit$	permutation
$\diamond$	$\diamond$	$\heartsuit$	$\spadesuit$	$\clubsuit$	$\diamond$
$\heartsuit$	$\heartsuit$	$\diamond$	$\clubsuit$	$\spadesuit$	$(1,2)(3,4)$
$\spadesuit$	$\spadesuit$	$\clubsuit$	$\diamond$	$\heartsuit$	$(1,3)(2,4)$
$\clubsuit$	$\clubsuit$	$\spadesuit$	$\heartsuit$	$\diamond$	$(1,4)(2,3)$

双脚并拢, 跳过计算.
对运算按照复杂性而不是其表象加以群分类聚.

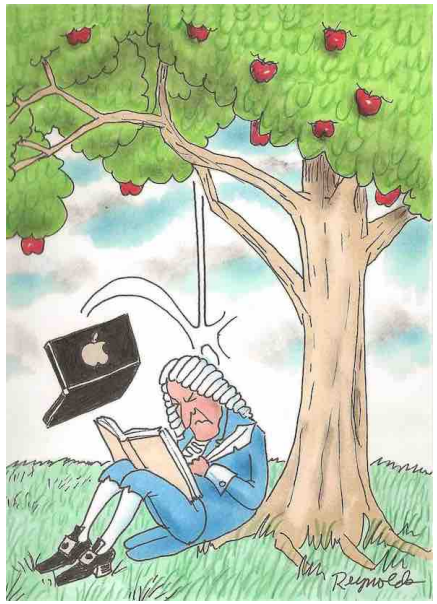
— 伽罗瓦

数学是给不同的事物起同一个名字的艺术.

— 庞加莱

牛顿力学: 苹果落地, 月亮咋就不落地?

"力是**保持**物体运动状态的原因."
— 亚里士多德



- 什么是 "运动"?
- 什么是 "运动状态"?
- 什么是 "物体"?
- 怎么 "保持"? 还是 "改变"?

牛顿力学理论

- 基本概念: 质量, 动量, 惯性, 外力
- 非逻辑公理:

公理 1 惯性定律
公理 2 动量变化定律
公理 3 作用与反作用定律
公理 4 万有引力定律

$$F = 0 \iff \frac{\mathrm{d}^2 s}{\mathrm{d}t^2} = 0$$

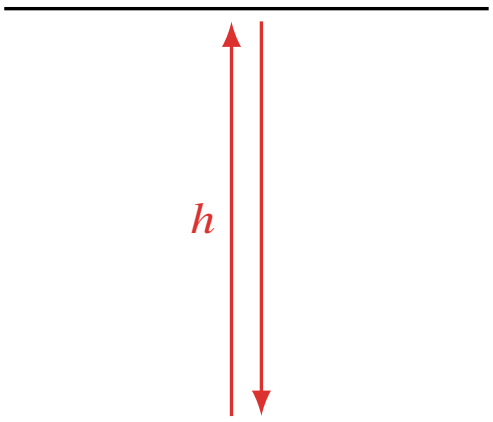
$$F = m\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}$$

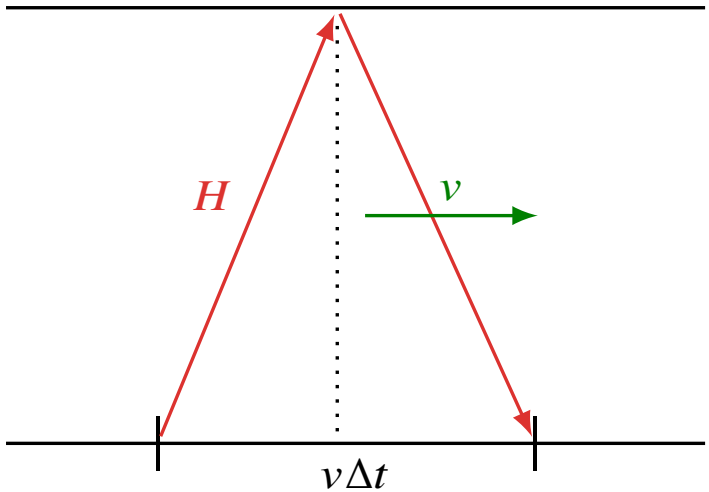
$$F_{12} = -F_{21}$$

$$F = \frac{G m_1 m_2}{r^2}$$

爱因斯坦的狭义相对论

- “力学的目的是描述物体在空间中的位置如何随时间变化.”
- 什么是 “位置”? 什么是 “空间”? 什么是 “时间”? 是绝对的吗?
- 不同惯性参照系的观察者对同一事件的描述相同吗?





$$\left.\begin{aligned} \Delta t' &= \frac{2h}{c} \\ \Delta t &= \frac{2H}{c} = \frac{2}{c}\sqrt{h^2 + \left(\frac{v\Delta t}{2}\right)^2} \end{aligned}\right\} \Longrightarrow \Delta t' = \Delta t\sqrt{1 - \frac{v^2}{c^2}}$$

- 假如以光速飞行, 你还能从镜子中看到自己的脸吗?
- 列车头尾同时被闪电劈中. 怎么判断 “同时”? 怎么测量 “时间”?

Definition (时钟同步)

在同一个惯性参照系中, 处在不同位置的处于静止状态的观察者各自拥有一个时钟 A 和 B. 时钟 A 和 B 同步, 当且仅当, 如果在 A 的时刻 t_A 从 A 朝 B 发出的光在 B 的时刻 t_B 到达 B 后立刻反射回 A, 并在 A 的时刻 t'_A 到达 A, 下述等式成立:

$$t_B - t_A = t'_A - t_B$$

Theorem

在同一个惯性参照系中, 处在不同位置的处于静止状态的时钟之间的同步关系是等价关系.

Remark: 同时性的相对性 — 在一个惯性参照系的观察者眼中同时发生的两个事件, 在相对匀速运动的另一个惯性参照系的观察者眼中不再是同时发生的.



站台上的观察者: "同时劈中." 列车中间的观察者: "先劈中头后劈中尾."

狭义相对论

公理 1 狭义相对性原理: 自然规律在所有惯性参照系中都相同.

公理 2 光速不变原理: 光在真空中的传播速度不变, 既与惯性参照系的选择无关, 也与光源的运动状态无关.

Remark: 光速上限假设. (因果作用只能在光锥内传播 $\Delta s \leq c\Delta t$)

洛伦兹变换:

$$\begin{bmatrix} x \\ ct \end{bmatrix} = \begin{bmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{bmatrix} \begin{bmatrix} x' \\ ct' \end{bmatrix} \quad t = \frac{t' + \frac{vx'}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \quad x = \frac{x' + vt'}{\sqrt{1-\frac{v^2}{c^2}}} \quad y = y' \quad z = z'$$

时空间隔与惯性系的选择无关.

$$c^2\mathrm{d}t^2 - (\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) = c^2\mathrm{d}t'^2 - (\mathrm{d}x'^2 + \mathrm{d}y'^2 + \mathrm{d}z'^2)$$

时间与长度的相对性 — 钟慢尺缩

$$\Delta t' = \Delta t\sqrt{1-\frac{v^2}{c^2}} \qquad l' = l\sqrt{1-\frac{v^2}{c^2}}$$

公孙大娘剑术精, 出刺迅捷如流星, 由于空间收缩性, 长剑变成短铁钉.

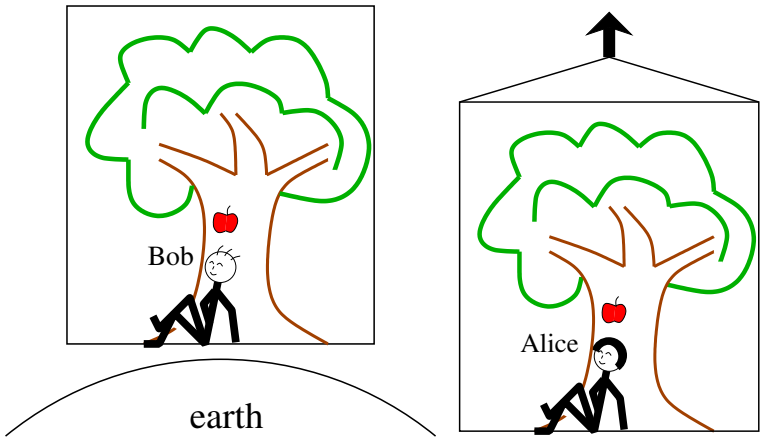
$$w = \frac{u+v}{1+\frac{uv}{c^2}} \qquad m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} \qquad p = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}v \qquad E = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}c^2$$

爱因斯坦的广义相对论

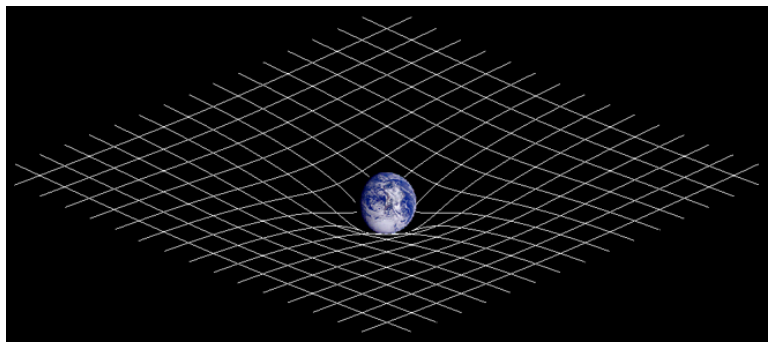
广义相对论

公理 1 广义相对性原理: 自然规律在所有参照系中都相同.

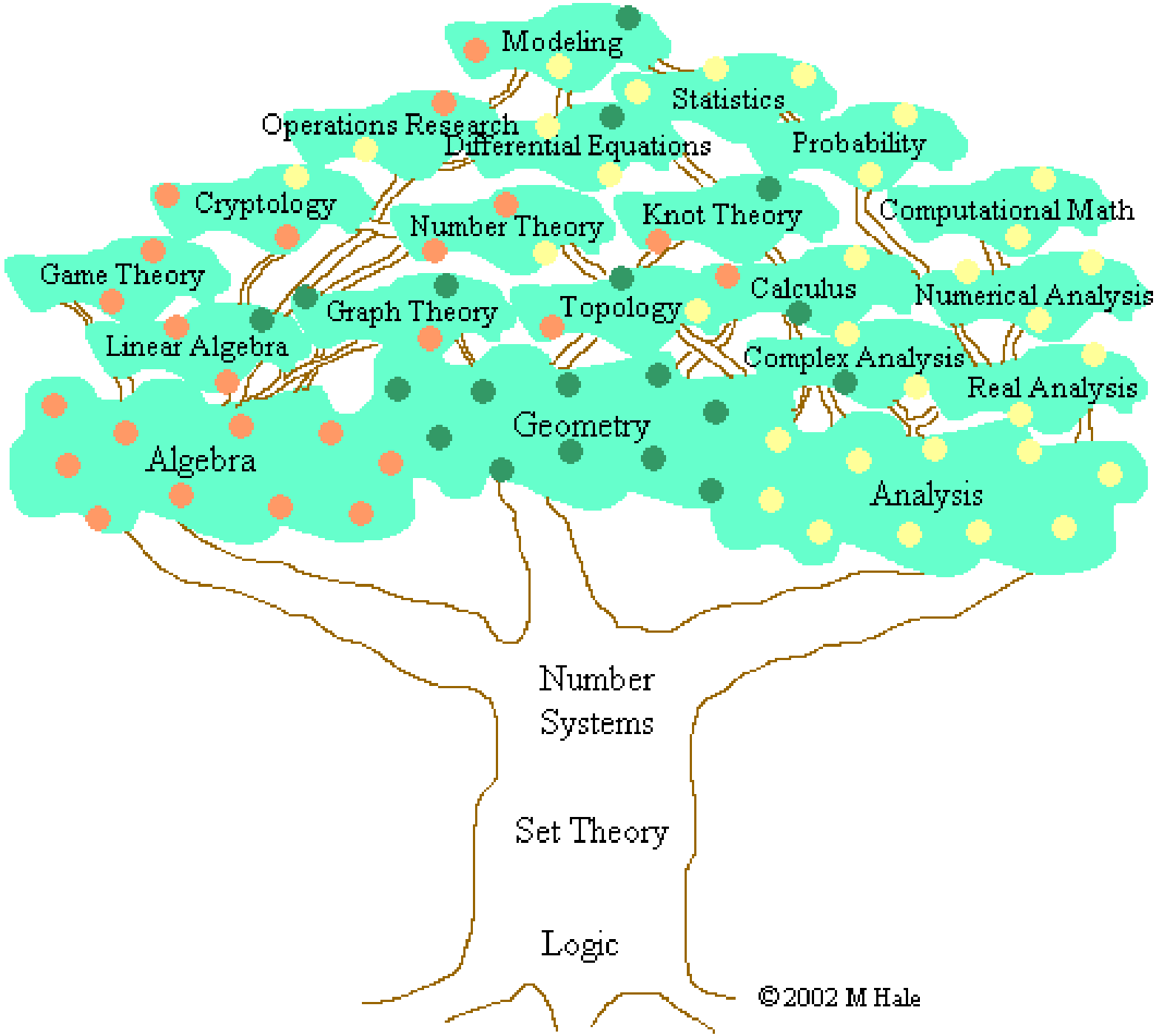
公理 2 等效原理: 引力场与惯性力场局域不可区分.

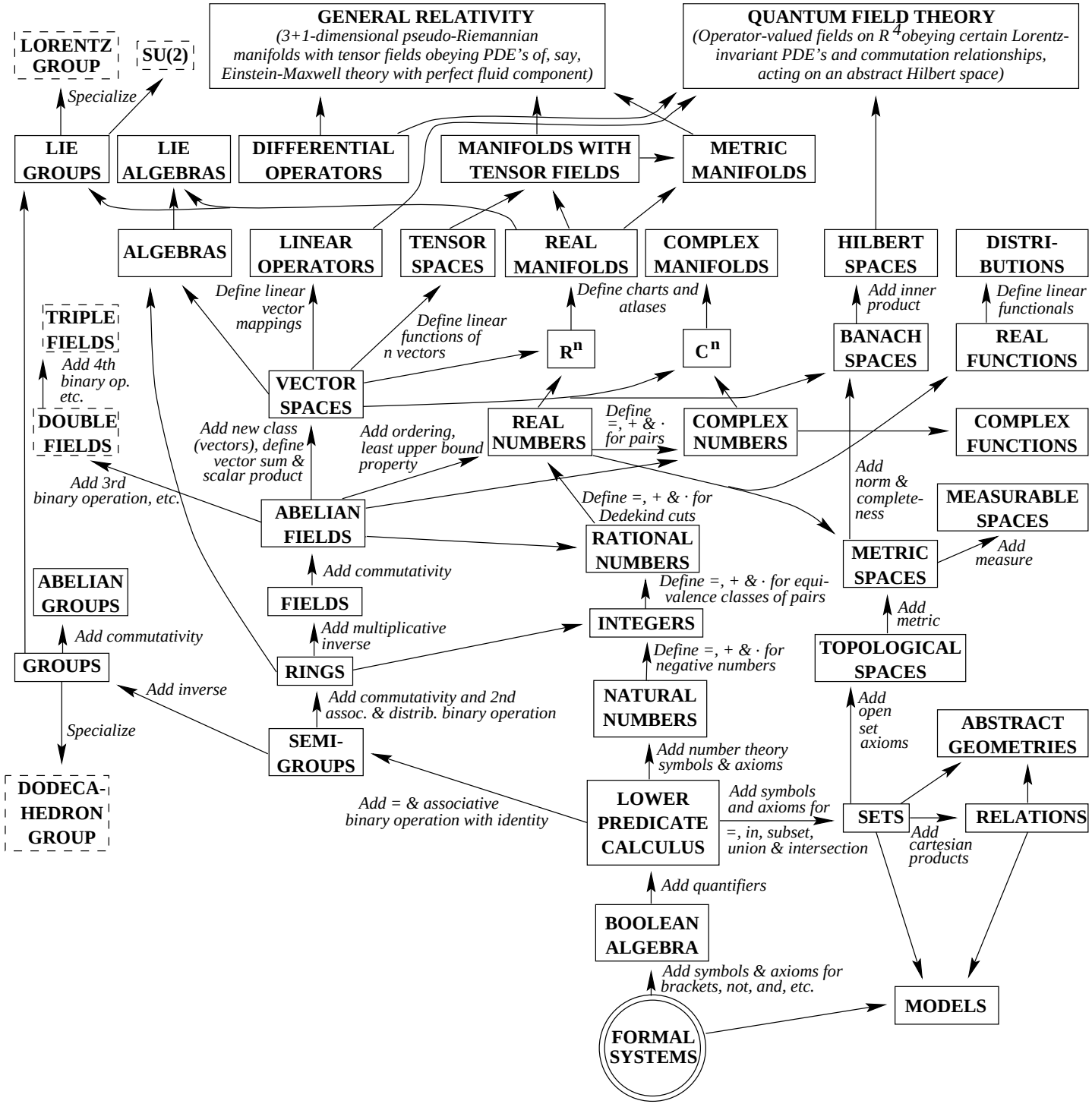


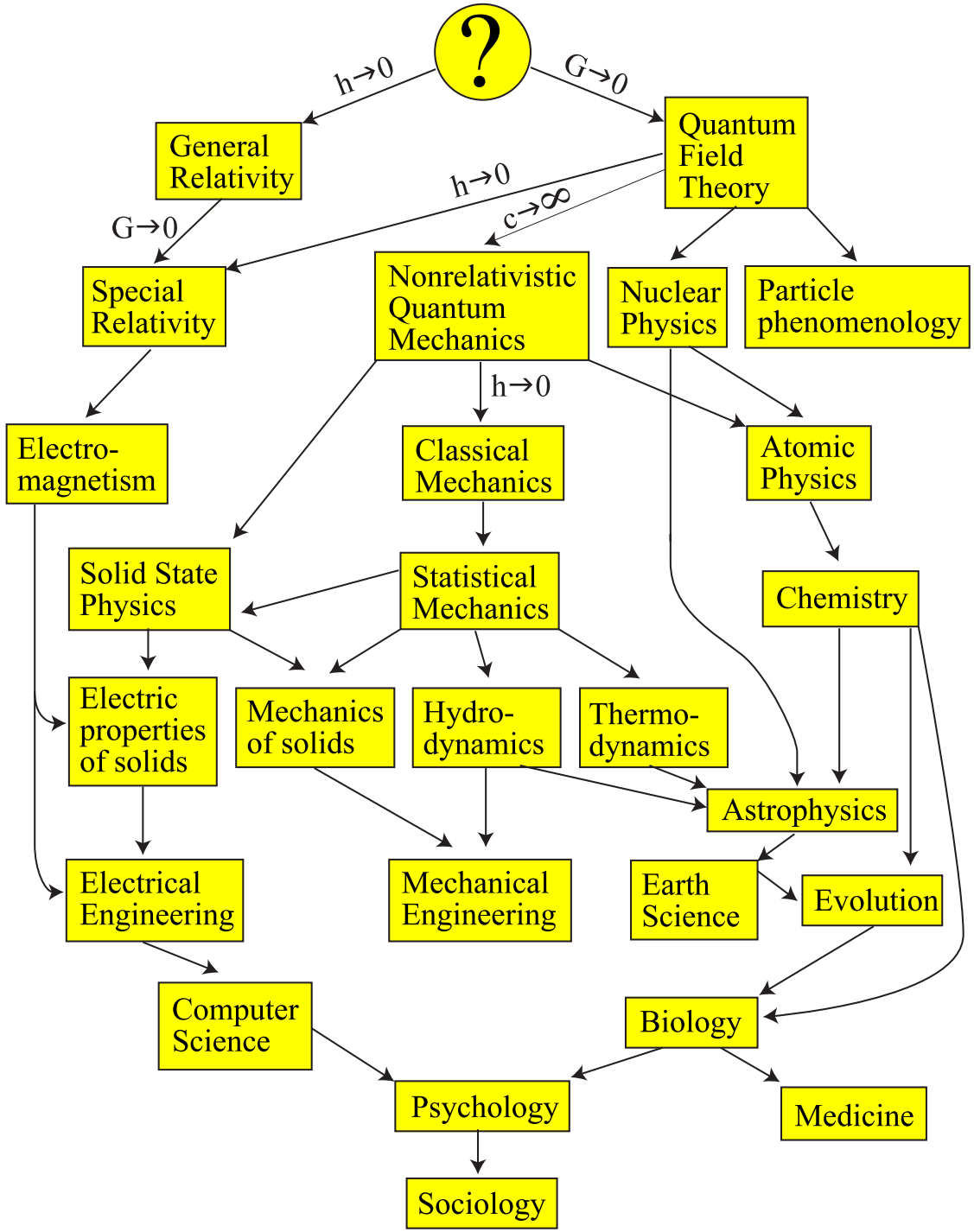
- 在大质量物体上感受到的引力, 与在加速参照系中感受到的惯性力局域不可区分.
- 惯性质量等于引力质量 $m_I = m_G$.
- 引力是由质量/能量分布不均导致的时空弯曲通过测地线原理作用的结果.
- 时空告诉物质如何运动; 物质告诉时空如何弯曲.



$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}$$


Modeling
Statistics
Operations Research
Differential Equations
Probability
Cryptology
Number Theory
Knot Theory
Computational Math
Game Theory
Graph Theory
Topology
Calculus
Numerical Analysis
Linear Algebra
Complex Analysis
Real Analysis
Geometry
Algebra
Analysis
Number Systems
Set Theory
Logic
©2002 M Hale


GENERAL RELATIVITY
(3+1-dimensional pseudo-Riemannian manifolds with tensor fields obeying PDE's of, say, Einstein-Maxwell theory with perfect fluid component)
QUANTUM FIELD THEORY
(Operator-valued fields on R^4 obeying certain Lorentz-invariant PDE's and commutation relationships, acting on an abstract Hilbert space)
LORENTZ GROUP
SU(2)
Specialize
LIE GROUPS
LIE ALGEBRAS
DIFFERENTIAL OPERATORS
MANIFOLDS WITH TENSOR FIELDS
METRIC MANIFOLDS
ALGEBRAS
LINEAR OPERATORS
TENSOR SPACES
REAL MANIFOLDS
COMPLEX MANIFOLDS
HILBERT SPACES
DISTRI-BUTIONS
Define linear vector mappings
Define linear functions of n vectors
Define charts and atlases
Add inner product
Define linear functionals
TRIPLE FIELDS
Add 4th binary op. etc.
DOUBLE FIELDS
VECTOR SPACES
R^n
C^n
BANACH SPACES
REAL FUNCTIONS
Add new class (vectors), define vector sum & scalar product
Add ordering, least upper bound property
REAL NUMBERS
Define =, + & · for pairs
COMPLEX NUMBERS
COMPLEX FUNCTIONS
Add 3rd binary operation, etc.
ABELIAN FIELDS
Define =, + & · for Dedekind cuts
Add norm & complete-ness
MEASURABLE SPACES
RATIONAL NUMBERS
METRIC SPACES
Add measure
ABELIAN GROUPS
Add commutativity
FIELDS
Define =, + & · for equi-valence classes of pairs
Add metric
Add multiplicative inverse
INTEGERS
TOPOLOGICAL SPACES
GROUPS
RINGS
Define =, + & · for negative numbers
Add inverse
Add commutativity and 2nd assoc. & distrib. binary operation
NATURAL NUMBERS
Add open set axioms
ABSTRACT GEOMETRIES
Specialize
SEMI-GROUPS
Add number theory symbols & axioms
DODECA-HEDRON GROUP
Add = & associative binary operation with identity
LOWER PREDICATE CALCULUS
Add symbols and axioms for =, in, subset, union & intersection
SETS
RELATIONS
Add cartesian products
Add quantifiers
BOOLEAN ALGEBRA
Add symbols & axioms for brackets, not, and, etc.
FORMAL SYSTEMS
MODELS


?
h→0
G→0
General Relativity
Quantum Field Theory
h→0
c→∞
G→0
Special Relativity
Nonrelativistic Quantum Mechanics
Nuclear Physics
Particle phenomenology
h→0
Electro-magnetism
Classical Mechanics
Atomic Physics
Solid State Physics
Statistical Mechanics
Chemistry
Electric properties of solids
Mechanics of solids
Hydro-dynamics
Thermo-dynamics
Astrophysics
Electrical Engineering
Mechanical Engineering
Earth Science
Evolution
Computer Science
Biology
Psychology
Medicine
Sociology

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Meta-Theorems

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Definability

What is "definability"?

Berry Paradox

The smallest positive integer not definable in fewer than twelve words.

Definition (Definability)

- $X \subset M^n$ is Y-definable over $\mathcal{M}$ (denoted by $X \in \mathrm{Def}(\mathcal{M}, Y)$) iff there is a formula A and $b_1, \ldots, b_m \in Y^m$ s.t.

$$X = \left\{(a_1, \ldots, a_n) : \mathcal{M} \vDash A[a_1, \ldots, a_n, b_1, \ldots, b_m]\right\}$$

- X is definable in $\mathcal{M}$ iff it is $\emptyset$-definable in $\mathcal{M}$.

A definition is acceptable only on condition that it implies no contradiction.

— *Poincaré*

Representability

What is "representability"?

Definition (Representable Functions)

A n-ary function $f : \mathbb{N}^n \to \mathbb{N}$ is representable in the theory T iff there is a formula $A(x_1, \ldots, x_n, y)$ s.t. for all $a_1, \ldots, a_n$,

$$\mathrm{T} \vdash \forall y \Big(A(\underline{a_1}, \ldots, \underline{a_n}, y) \leftrightarrow y = \underline{f(a_1, \ldots, a_n)} \Big)$$

Definition (Representable Relations)

A n-ary relation $R \subset \mathbb{N}^n$ is representable in the theory T iff there is a formula A s.t. for all $a_1, \ldots, a_n$,

$$(a_1, \ldots, a_n) \in R \implies \mathrm{T} \vdash A[a_1, \ldots, a_n]$$
$$(a_1, \ldots, a_n) \notin R \implies \mathrm{T} \vdash \neg A[a_1, \ldots, a_n]$$

A function/relation is representable in Robinson Q iff it is computable.

Example

- The interval $[0, \infty)$ is definable in $\mathcal{R} = (\mathbb{R}, 0, 1, +, \cdot)$, where the language is $\mathscr{L} = \{0, 1, +, \cdot\}$.

$$x \geq 0 \iff \mathcal{R} \vDash \exists y (x = y \cdot y)$$

- The ordering relation $<$ is definable in $\mathcal{N} = (\mathbb{N}, 0, s, +, \cdot)$, where the language is $\mathscr{L} = \{0, s, +, \cdot\}$.

$$\exists z \big(x + s(z) = y\big)$$

- The set of primes is definable in $\mathcal{N}$ by the formula

$$\exists y \big(x = s(0) + s(y)\big) \wedge \forall y z \big(x = y \cdot z \to y = s(0) \vee z = s(0)\big)$$

- Exponentiation $\big\{(m, n, p) : p = m^n\big\}$ is definable in $\mathcal{N}$. (use the Chinese remainder theorem)

Example

- The number 0 is definable in $(\mathbb{Z}, +)$, since

$$x = 0 \iff (\mathbb{Z}, +) \vDash x + x = x$$

No other elements are definable, because negation $x \mapsto -x$ is an automorphism.

- The number 1 is definable in $(\mathbb{Z}, +, \cdot)$, since

$$x = 1 \iff (\mathbb{Z}, +, \cdot) \vDash x \cdot x = x$$

- $\mathbb{N}$ is definable in $(\mathbb{Z}, +, \cdot)$ by

$$\exists y_1 y_2 y_3 y_4 \left(x = y_1^2 + y_2^2 + y_3^2 + y_4^2 \right) \quad \text{(Lagrange four-square theorem)}$$

- Every definable subsets of $(\mathbb{N}, <)$ is either finite or cofinite.

Example

- No point is definable in $(\mathbb{R}, <)$, since any two real numbers are automorphic by translation.
- The ordering relation $<$ is definable in $\mathcal{R} = (\mathbb{R}, 0, 1, +, \cdot)$.

$$x < y \iff \mathcal{R} \vDash \exists z(x + z \cdot z = y)$$

Every individual integer is definable.
Every rational number is definable.
Every algebraic number is definable.
But only algebraic numbers are definable.

Theorem (Tarski)

In $(\mathbb{R}, 0, 1, +, \cdot)$, every formula $A(x)$ is equivalent to a quantifier-free formula.

- π is definable in $(\mathbb{R}, 0, 1, +, \cdot, \sin)$.
  $\mathbb{Z}$ is definable.
  Every computable real number is definable.
  Every arithmetic real & every projective real is definable.

Pointwise Definable Model vs Leibnizian Model

Definition (Leibnizian Model)

A model $\mathcal{M}$ is Leibnizian if any two distinct objects can be distinguished by some property: for $a \neq b$, there is a formula A s.t.,

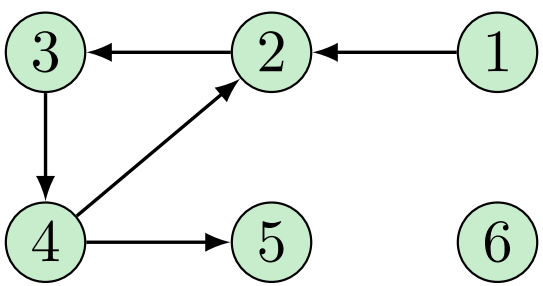
$$\mathcal{M} \vDash A[a] \ \ \& \ \ \mathcal{M} \nvDash A[b]$$

Definition (Pointwise Definable Model)

A model $\mathcal{M}$ is pointwise definable if every object of $\mathcal{M}$ is definable in $\mathcal{M}$ without parameters.

$$x = a \iff \mathcal{M} \vDash A(x)$$

Example:



- 1 is the node that points at some node but is not pointed at by any node.
- 2 is the node that is pointed at by a node that is not pointed at by any node.
- ...

Pointwise Definable Model vs Leibnizian Model

Problem: Are these notions the same?

- Every pointwise definable model is Leibnizian.
- The real ordered field $(\mathbb{R}, 0, 1, +, \cdot, <)$ is Leibnizian, but is not pointwise definable.
  - For any reals $x < y$, there is a rational $\frac{p}{q}$ between them, and x has the property that $x < \frac{p}{q}$, while y does not.
  - It is not pointwise definable, because there are only countably many definitions.

Homomorphism & Isomorphism

Definition (Homomorphism)

A homomorphism h of $\mathcal{M}$ into $\mathcal{N}$ is a function $h : M \to N$ s.t.

- For each n-place predicate symbol P and each n-tuple $(a_1, \ldots, a_n) \in M^n$,

$$(a_1, \ldots, a_n) \in P^{\mathcal{M}} \iff (h(a_1), \ldots, h(a_n)) \in P^{\mathcal{N}}$$

- For each n-place function symbol f and each n-tuple $(a_1, \ldots, a_n) \in M^n$,

$$h : f^{\mathcal{M}}(a_1, \ldots, a_n) \mapsto f^{\mathcal{N}}\big(h(a_1), \ldots, h(a_n)\big)$$

  In the case of a constant symbol c this becomes $h : c^{\mathcal{M}} \mapsto c^{\mathcal{N}}$.

- An isomorphism (monomorphism/epimorphism) is a bijective (injective/surjective) homomorphism. $\mathcal{M} \cong \mathcal{N}$
- An automorphism (endomorphism) is an isomorphism (homomorphism) from $\mathcal{M}$ to itself.

Leibnizian Model vs Rigid Model

Definition (Rigid Model)

A model $\mathcal{M}$ is rigid iff it has no automorphisms other than the trivial automorphism 1_M.

- Every Leibnizian model must be rigid, because automorphisms are truth-preserving.
- There are rigid models that are not Leibnizian.
  — Every well-order structure is rigid, but when an order is sufficiently large (larger than the continuum), then not every point can be characterized by its properties, because there aren't enough sets of formulas in the language to distinguish all the points.
- The complex field is not rigid, because it admits an automorphism $a + b\mathrm{i} \mapsto a - b\mathrm{i}$.
- The real number field is not definable in the complex field.
- Once we augment the complex field $\mathbb{C}$ with

$$\mathtt{Re}(a + b\mathrm{i}) = a \qquad \mathtt{Im}(a + b\mathrm{i}) = b$$

  then the expanded structure is Leibnizian.

Nonrigid Model as Reduct of Rigid Model

- The additive group of integers $(\mathbb{Z}, +)$ admits an automorphism by negation, but is made rigid with the multiplicative structure $(\mathbb{Z}, +, \cdot)$ or the order structure $(\mathbb{Z}, +, <)$.
- The rational order $(\mathbb{Q}, <)$ becomes rigid with the field structure $(\mathbb{Q}, +, \cdot, <)$.
- Every group G with at least three elements is nonrigid, but elements are distinguished when the group is given a particular presentation, such as by means of generators and relations or as permutations of a particular set.
- Every set is a subset of a transitive set, and every transitive set is rigid with respect to the $\in$ membership relation. Indeed, the set-theoretic universe $(V, \in)$ as a whole is rigid.

Homomorphism Theorem

Theorem (Homomorphism Theorem)

Let h be a homomorphism of $\mathcal{M}$ into $\mathcal{N}$, and $\nu : \mathrm{Var} \to M$.

1. *For any term t, $h\big(\nu(t)\big) = h \circ \nu(t)$*
2. *For any open formula A not containing $=$, $\mathcal{M}, \nu \vDash A \iff \mathcal{N}, h \circ \nu \vDash A$*
3. *If $h : M \rightarrowtail N$, we may delete the restriction "not containing $=$".*
4. *If $h : M \twoheadrightarrow N$, we may delete the restriction "open".*

Elementary Equivalence

Definition (Elementary Equivalence)

$\mathcal{M} \equiv \mathcal{N}$ iff for any sentence A:

$$\mathcal{M} \vDash A \iff \mathcal{N} \vDash A$$

- $(\mathbb{Z}^+, <) \equiv (\mathbb{N}, <)$
- $(\mathbb{Q}, \leq) \not\equiv (\mathbb{N}, \leq)$
- $(\mathbb{Q}, \leq) \not\equiv (\mathbb{R}, \leq)$ but $(\mathbb{Q}, \leq) \equiv (\mathbb{R}, \leq)$

Theorem

$$\mathcal{M} \cong \mathcal{N} \implies \mathcal{M} \equiv \mathcal{N}$$

Theorem

$\mathcal{M} \equiv \mathcal{N} \;\&\; |M| < \infty \implies \mathcal{M} \cong \mathcal{N}$

Proof.

Suppose $|M| = n$. Then $|N| = n$.
There are only finitely many functions $f_1, \ldots, f_m : M \to N$. Assume none of $f : M \to N$ is an isomorphism. For each $f_i, 1 \leq i \leq m$, there is a formula A_i s.t. $\mathcal{M} \vDash A_i(a_1, \ldots, a_n)$ but $\mathcal{N} \nvDash A_i(f_i(a_1), \ldots, f_i(a_n))$. Then we have

$$\mathcal{M} \vDash \bigwedge_{i=1}^{m} A_i(a_1, \ldots, a_n) \quad \& \quad \mathcal{M} \vDash \exists x_1 \ldots x_n \bigwedge_{i=1}^{m} A_i(x_1, \ldots, x_n)$$

Since $\mathcal{M} \equiv \mathcal{N}$, then $\mathcal{N} \vDash \exists x_1 \ldots x_n \bigwedge_{i=1}^{m} A_i(x_1, \ldots, x_n)$, and
$\mathcal{N} \vDash \bigwedge_{i=1}^{m} A_i(b_1, \ldots, b_n)$ for some $b_1, \ldots, b_n \in N$.
Let $f_j : a_i \mapsto b_i$. But $\mathcal{N} \nvDash A_j(b_1, \ldots, b_n)$. □

Definition (Substructure)

$\mathcal{M}$ is called a *substructure* of $\mathcal{N}$ ($\mathcal{M} \subset \mathcal{N}$) iff

- $M \subset N$
- 1. $P^{\mathcal{M}} = P^{\mathcal{N}} \cap M^n$ for any n-ary predicate symbol P.
  2. $f^{\mathcal{M}} = f^{\mathcal{N}} \restriction_{M^n}$ for any n-ary function symbol f.

Suppose $\mathcal{M} \subset \mathcal{N}$. Then

- for any term $t(x_1, \ldots, x_n)$, and any $a_1, \ldots, a_n \in M$,

$$t^{\mathcal{M}}[a_1, \ldots, a_n] = t^{\mathcal{N}}[a_1, \ldots, a_n]$$

- for any formula $A(x_1, \ldots, x_n)$, and any $a_1, \ldots, a_n \in M$,

$$\mathcal{M} \vDash A[a_1, \ldots, a_n] \iff \mathcal{N} \vDash A[a_1, \ldots, a_n]$$

Theorem

Let A be a Σ_1 sentence and B a Π_1 sentence, and $\mathcal{M} \subset \mathcal{N}$. Then

- $\mathcal{M} \vDash A \implies \mathcal{N} \vDash A$
- $\mathcal{N} \vDash B \implies \mathcal{M} \vDash B$

Remark: The theorem can be used to show that some class is not Π_1-axiomatizable.

- A strict partial order satisfying
  1. $\forall xyz : Rxy \wedge Ryz \rightarrow Rxz$
  2. $\forall x : \neg Rxx$
- A strict partial order is dense if it also satisfies the Π_2 sentence
  3. $\forall xy \exists z (Rxy \rightarrow Rxz \wedge Rzy)$
- Is it Π_1-axiomatizable?
- $([0,1], <)$ is a dense strict partial order. But its substructure $(\{0,1\}, <)$ is not dense! So the class of dense order is not Π_1-axiomatizable.

Example

- $\mathscr{L} = \{0, 1, +, \cdot\}, \mathcal{N} = (\mathbb{N}, 0, 1, +, \cdot), \mathcal{R} = (\mathbb{R}, 0, 1, +, \cdot)$

$$\mathcal{N} \subset \mathcal{R}$$

- $\mathscr{L} = \{<\}, \mathcal{M} = (\mathbb{N}, <), \mathcal{N} = (\{2n : n \in \mathbb{N}\}, <)$

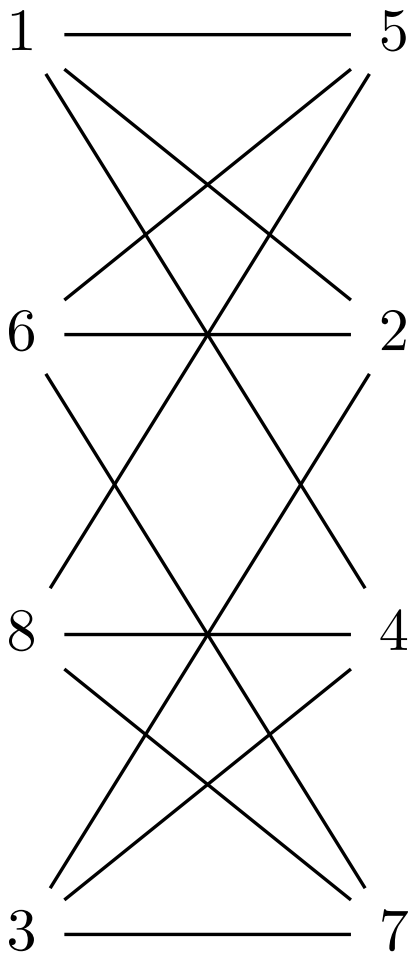
$$h : n \mapsto 2n, \quad h : \mathcal{M} \cong \mathcal{N}, \quad \text{but} \quad \mathcal{M} \not\subset \mathcal{N}$$

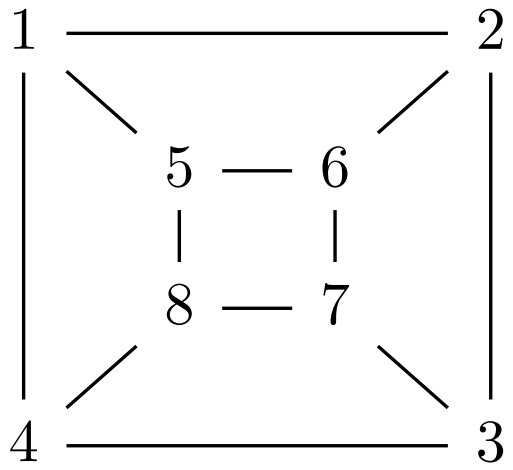
- $\mathscr{L} = \{0, +\}$

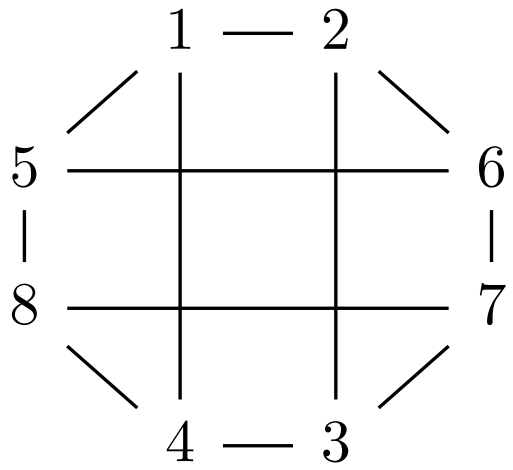
$$\mathcal{M} = \left(\mathbb{N}, 0^{\mathcal{M}}, +^{\mathcal{M}}\right), \qquad \text{where} \quad 0^{\mathcal{M}} = 0, \quad +^{\mathcal{M}}(a, b) = a + b$$
$$\mathcal{N} = \left(\{2^n : n \in \mathbb{N}\}, 0^{\mathcal{N}}, +^{\mathcal{N}}\right), \quad \text{where} \quad 0^{\mathcal{N}} = 1, \quad +^{\mathcal{N}}(a, b) = a \cdot b$$

$$h : n \mapsto 2^n, \quad h : \mathcal{M} \cong \mathcal{N} \quad \text{but} \quad \mathcal{M} \not\subset \mathcal{N}.$$

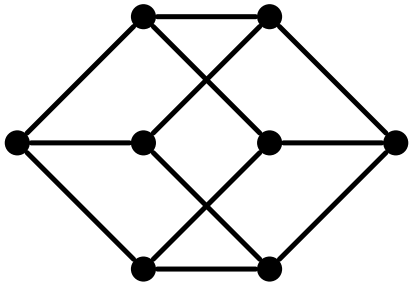
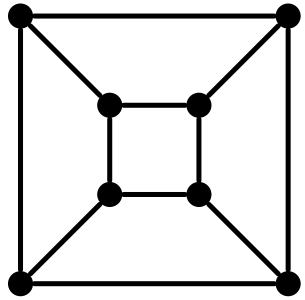
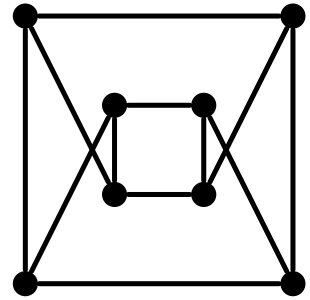
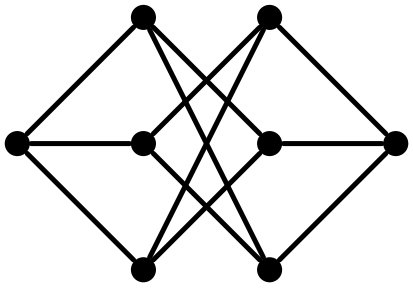
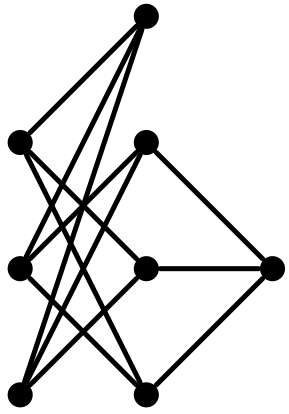
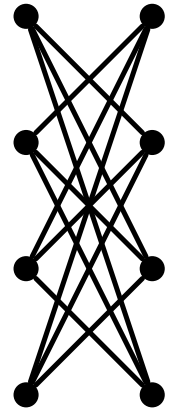
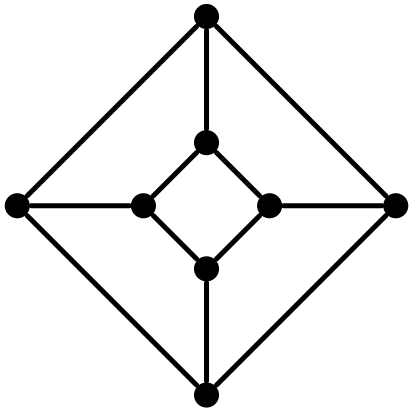
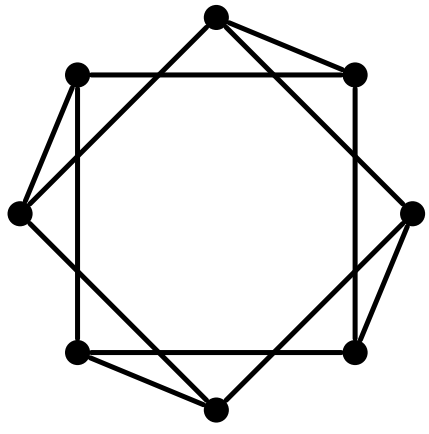
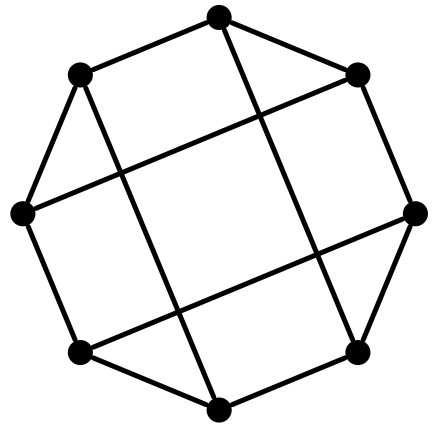
Example



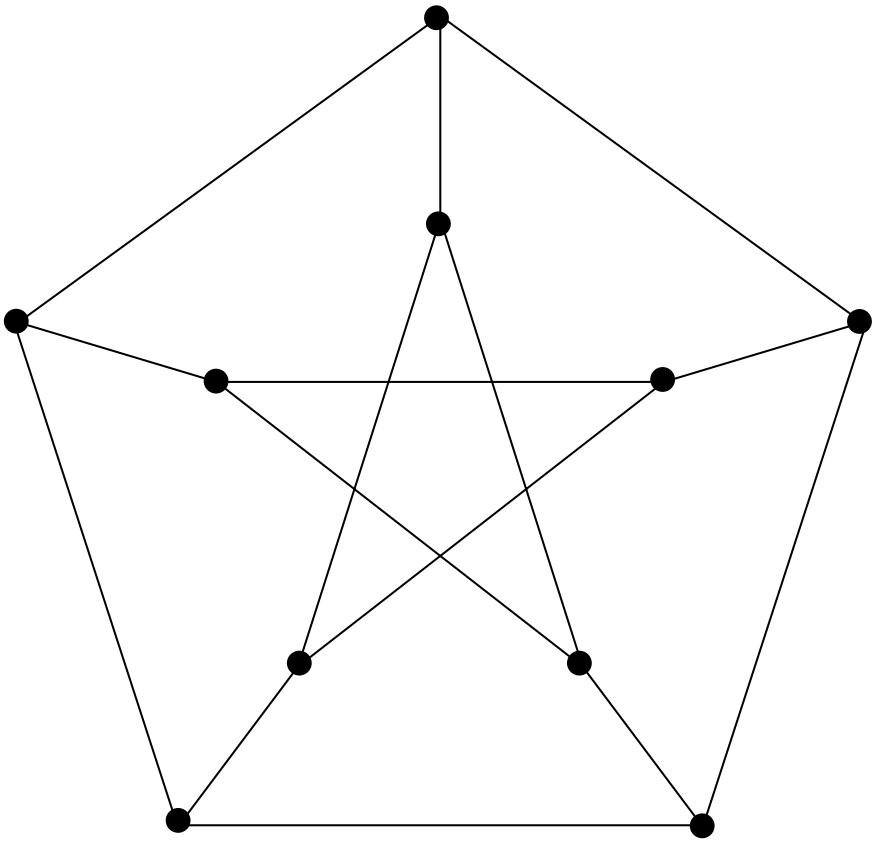


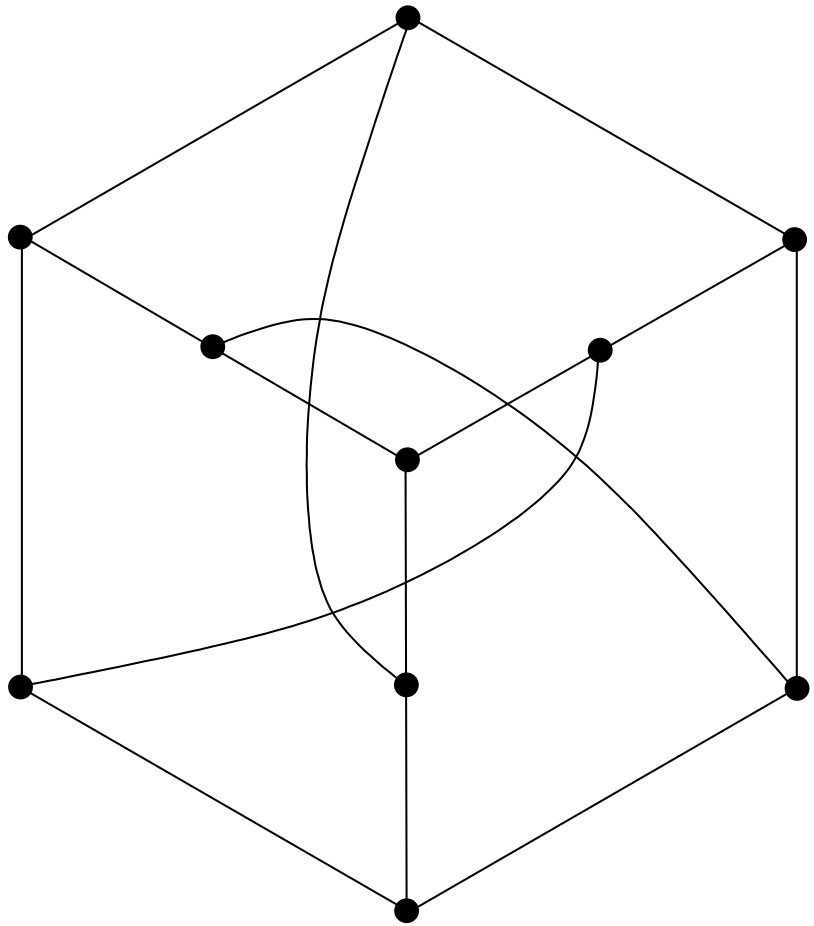


quasipolynomial $2^{O((\log n)^c)}$

Example





A Joke

"Let G_1 be the group …, and G_2 be the group … Prove that G_1 and G_2 are isomorphic."

One of the papers submitted had an answer "We will show that G_1 is isomorphic..." and some nonsense, followed by "Now we'll show that G_2 is isomorphic..." and more nonsense.

share cite

answered Jan 31 '11 at 18:53

community wiki
Asaf Karagila

---

86 I gave a homework problem, "Let G_1 be the group …, let G_2 be the group …. Are G_1 and G_2 isomorphic?" and was astonished to get the response, "G_1 is, but G_2 isn't." Are Asaf's story and mine isomorphic? – Gerry Myerson Jan 31 '11 at 22:39

165 @Gerry: Asaf's is, but yours isn't. – Nate Eldredge Feb 1 '11 at 1:20

Automorphism & Undefinability

Corollary

$$X \in \mathrm{Def}(\mathcal{M}, Y) \;\;\&\;\; h \in \mathrm{Aut}(\mathcal{M}/Y) \implies h(X) = X$$

i.e., $\forall a_1, \ldots, a_n \in M : [(a_1, \ldots, a_n) \in X \iff (h(a_1), \ldots, h(a_n)) \in X]$.

Remark: This corollary is sometimes useful in showing that a given set is not definable.

The set $\mathbb{N}$ is not definable in $(\mathbb{R}, <)$ where $\mathscr{L} = \{<\}$.

$h : a \mapsto a^3$ is an automorphism of $\mathbb{R}$.
It maps points outside of $\mathbb{N}$ into $\mathbb{N}$.

$$\boxed{\mathbb{N} \text{ is not definable in } (\mathbb{R}, 0, 1, +, \cdot, <)}$$

Natural numbers are not definable over the theory of real-closed fields.

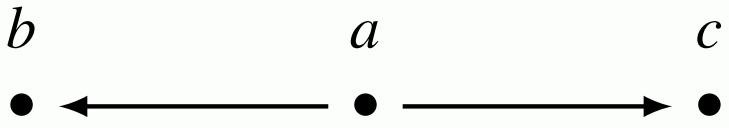
Theorem

Suppose $\mathcal{M}$ is finite, or countable and $\aleph_0$-categorical. If $X \subset M^n$ is invariant under all $h \in \mathrm{Aut}(\mathcal{M})$, then X is definable.

Example

Example

The model $\mathcal{M} \coloneqq \big(\{a, b, c\}, \{(a, b), (a, c)\}\big)$
where the language is $\mathscr{L} = \{R\}$.



- $\{b, c\}$ is definable in $\mathcal{M}$: $\quad \exists y R(y, x)$
- $\{b\}$ is not definable in $\mathcal{M}$.

Example

Consider the vector space $\mathcal{E} \coloneqq (E, +, f_r)_{r \in \mathbb{R}}$, where E is the universe, f_r is the scalar multiplication by r.

- $U \coloneqq \{\mathbf{x} \in E : |\mathbf{x}| = 1\}$ is not definable in $\mathcal{E}$.
- $h : \mathbf{x} \mapsto 2\mathbf{x}$ is an automorphism but it does not preserve U.

Ehrenfeucht-Fraïssé Game (EF Game)

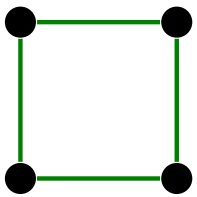
Spoiler and *Defender*, played on two models $\mathcal{M}$ and $\mathcal{N}$.
Each run of the game has n moves. In each move,

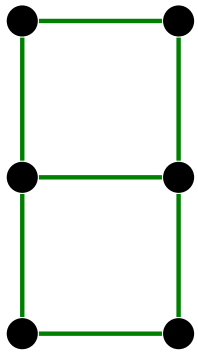
- *Spoiler* picks an element from $\mathcal{M}$ or from $\mathcal{N}$.
- *Defender* picks an element from $\mathcal{N}$ or from $\mathcal{M}$.
- *Defender* wins the run if $(a_i, b_i)_{i=1}^{n}$ is a partial isomorphism from $\mathcal{M}$ to $\mathcal{N}$.
- *Spoler* wins the run otherwise.
- $\mathcal{M} \sim_n \mathcal{N}$ iff *Defender* has a winning strategy in the n-move game.
- $\mathcal{M} \equiv_n \mathcal{N}$ iff $\mathcal{M} \vDash A \iff \mathcal{N} \vDash A$ for all sentences up to *quantifier depth* n.

Theorem

$$\mathcal{M} \sim_n \mathcal{N} \iff \mathcal{M} \equiv_n \mathcal{N}$$

Ehrenfeucht-Fraïssé Game (EF Game)





$$\mathcal{M} \sim_2 \mathcal{N} \qquad \mathcal{M} \nsim_3 \mathcal{N}$$

$$\mathcal{M} \models A \qquad \mathcal{N} \models \neg A$$

$$\forall xy \exists z (\neg Exy \rightarrow Exz \wedge Eyz)$$

Isomorphic Embedding, Elementary Embedding

Definition

- Isomorphic embedding $f : \mathcal{M} \subset \mathcal{N}$ iff there is $\mathcal{A} \subset \mathcal{N}$ s.t. $f : \mathcal{M} \cong \mathcal{A}$.
- Elementary embedding $f : \mathcal{M} \prec \mathcal{N}$ iff for any formula $A(x_1, \ldots, x_n)$ and any $a_1, \ldots, a_n \in M$,
$$\mathcal{M} \vDash A[a_1, \ldots, a_n] \iff \mathcal{N} \vDash A[f(a_1), \ldots, f(a_n)]$$
- Elementary substructure $\mathcal{M} \prec \mathcal{N}$ iff $M \subset N \ \& \ 1_M : \mathcal{M} \prec \mathcal{N}$.

Example

- $(\mathbb{N}^+, \leq) \subset (\mathbb{N}, \leq) \quad (\mathbb{N}^+, \leq) \cong (\mathbb{N}, \leq) \quad (\mathbb{N}^+, \leq) \not\prec (\mathbb{N}, \leq)$
$A(x) := \exists y(y \leq x \wedge y \neq x) \qquad (\mathbb{N}, \leq) \vDash A[1] \qquad (\mathbb{N}^+, \leq) \nvDash A[1]$
- $(2\mathbb{Z}, <) \subset (\mathbb{Z}, <) \qquad (2\mathbb{Z}, <) \cong (\mathbb{Z}, <) \qquad (2\mathbb{Z}, <) \not\prec (\mathbb{Z}, <)$
$A(x, y) := \exists z(x < z < y) \qquad (\mathbb{Z}, <) \vDash A[0, 2] \qquad (2\mathbb{Z}, <) \nvDash A[0, 2]$

$$f : \mathcal{M} \prec \mathcal{N} \iff \exists \mathcal{A}\big(f : \mathcal{M} \cong \mathcal{A} \prec \mathcal{N}\big) \iff \exists \mathcal{A}\big(\mathcal{M} \prec \mathcal{A} \cong \mathcal{N}\big)$$

Isomorphic Embedding, Elementary Embedding

$$(\mathbb{N}, <) \subset (\mathbb{Z}, <) \subset (\mathbb{Q}, <)$$

$$(\mathbb{N}, <) \not\prec (\mathbb{Z}, <) \not\prec (\mathbb{Q}, <)$$

$$(\mathbb{Q}, 0, 1, +, \cdot) \subset (\mathbb{R}, 0, 1, +, \cdot) \subset (\mathbb{C}, 0, 1, +, \cdot)$$

$$(\mathbb{Q}, 0, 1, +, \cdot) \not\prec (\mathbb{R}, 0, 1, +, \cdot) \not\prec (\mathbb{C}, 0, 1, +, \cdot)$$

$$(\mathbb{N}, 0, 1, +, \cdot, <) \subset (\mathbb{Z}, 0, 1, +, \cdot, <) \subset (\mathbb{Q}, 0, 1, +, \cdot, <) \subset (\mathbb{R}, 0, 1, +, \cdot, <)$$

$$(\mathbb{N}, 0, 1, +, \cdot, <) \not\prec (\mathbb{Z}, 0, 1, +, \cdot, <) \not\prec (\mathbb{Q}, 0, 1, +, \cdot, <) \not\prec (\mathbb{R}, 0, 1, +, \cdot, <)$$

$$(\mathbb{Q}, <) \prec (\mathbb{R}, <)$$

$$(2\mathbb{Z}, 0, +, -, <) \subset (\mathbb{Z}, 0, +, -, <)$$

$$(2\mathbb{Z}, 0, +, -, <) \equiv (\mathbb{Z}, 0, +, -, <)$$

$$(2\mathbb{Z}, 0, +, -, <) \not\prec (\mathbb{Z}, 0, +, -, <)$$

Isomorphic Embedding, Elementary Embedding

- If $f : \mathcal{M} \cong \mathcal{N}$, then for any term $t(x_1, \ldots, x_n)$, any formula $A(x_1, \ldots, x_n)$, and any $a_1, \ldots, a_n \in M$,

$$f(t^{\mathcal{M}}[a_1, \ldots, a_n]) = t^{\mathcal{N}}[f(a_1), \ldots, f(a_n)]$$

$$\mathcal{M} \vDash A[a_1, \ldots, a_n] \iff \mathcal{N} \vDash A[f(a_1), \ldots, f(a_n)]$$

- **Tarski-Vaught Test:** $f : \mathcal{M} \prec \mathcal{N}$ iff $f : \mathcal{M} \subset \mathcal{N}$ and for any formula $\exists x A(x_1, \ldots, x_n, x)$ and any $a_1, \ldots, a_n \in M$,

$$\mathcal{N} \vDash \exists x A[f(a_1), \ldots, f(a_n), x] \implies \exists a \in M : \mathcal{N} \vDash A[f(a_1), \ldots, f(a_n), f(a)]$$

$$\mathcal{M} \prec \mathcal{N} \iff \mathcal{M} \subset \mathcal{N} \ \& \ \forall X \in \mathrm{Def}(\mathcal{N}, M) : X \cap M \neq \emptyset$$

Let $M \subset \mathbb{R}$. $(M, <) \prec (\mathbb{R}, <)$ iff $(M, <)$ is a dense linear ordering without endpoints.

Isomorphic Embedding, Elementary Embedding

- Let $\mathcal{M}$ be a $\mathscr{L}$-model. $\mathcal{M}_M \coloneqq (\mathcal{M}, a)_{a \in M}$ is a $\mathscr{L}_M$-model by interpreting c_a by a.
- Let $\mathcal{N}$ be a $\mathscr{L}$-model, and $X \subset M \ \ \& \ \ f : X \to N$. $(\mathcal{N}, f(a))_{a \in X}$ is a $\mathscr{L}_X$-model by interpreting c_a by $f(a)$.

$\mathrm{diag}(\mathcal{M}) \coloneqq \{A : A \text{ is an atomic or negated atomic sentence of } \mathscr{L}_M \text{ and } \mathcal{M}_M \vDash A\}$

- $f : \mathcal{M} \subset \mathcal{N} \iff (\mathcal{N}, f(a))_{a \in M} \vDash \mathrm{diag}(\mathcal{M})$
- $f : \mathcal{M} \prec \mathcal{N} \iff (\mathcal{N}, f(a))_{a \in M} \vDash \mathrm{Th}(\mathcal{M}_M)$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Meta-Theorems

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

What is Logic?

- Arithmetic — the study of numbers.
- Geometry — the study of figures.
- Algebra — the study of mathematical symbols.
- Set Theory — the study of sets.
- Logic — the study of logical notions.

- What is a number?
- What is a line?
- What is a set?
- What is a logical notion?

What is Mathematics?

Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true.

— Bertrand Russell

Mathematics is the art of giving the same name to different things.

— Henri Poincaré

Not substance but invariant form is the carrier of the relevant mathematical information.

— F. William Lawvere

Mathematics is the analysis of invariants and of the transformations that preserve them (including the analysis of non-preservations, deformations and symmetry breakings).

What is Geometry? — Klein's Erlangen Program

What is Geometry?

The study of *invariants* under *a group of transformations*.

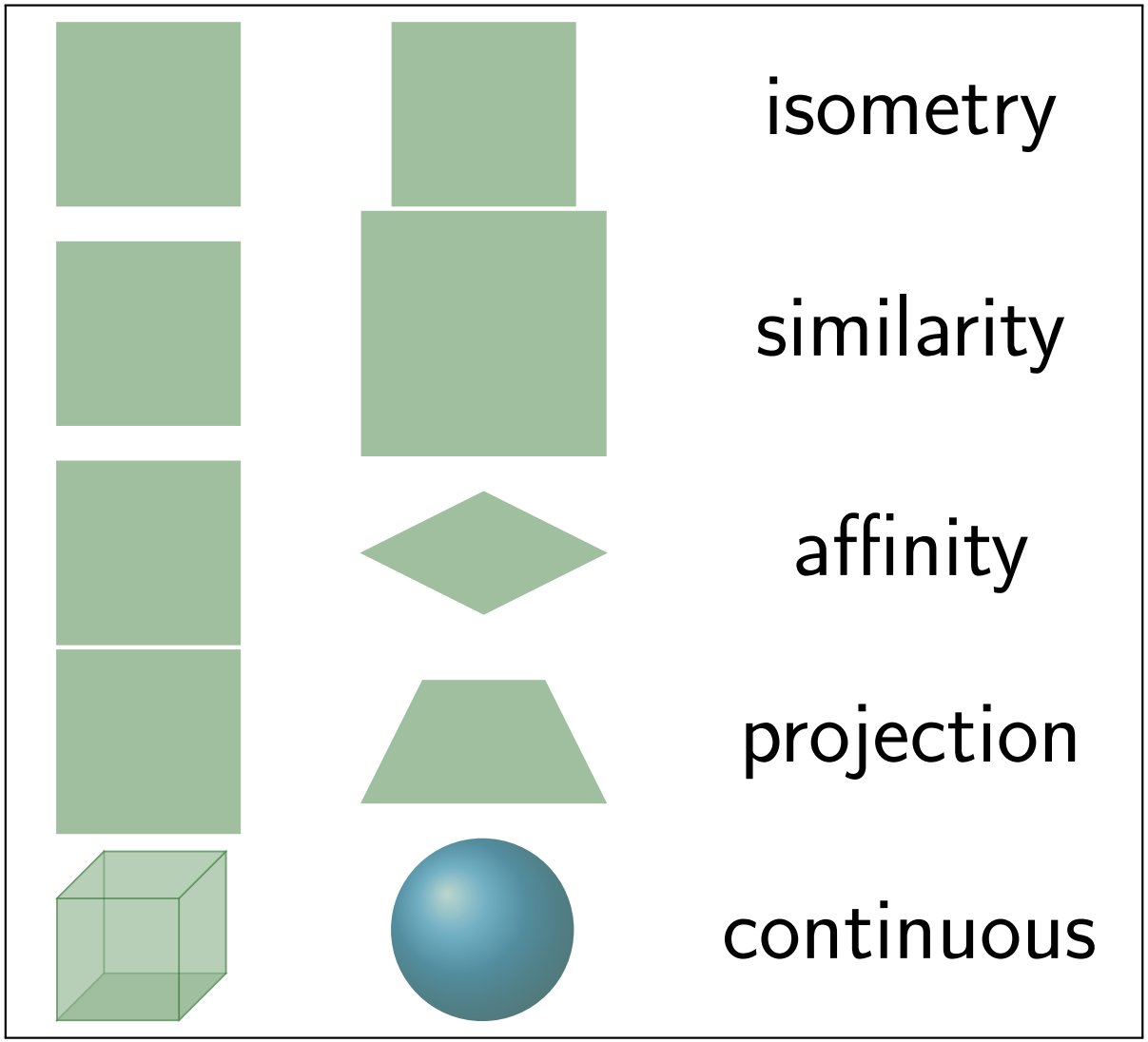
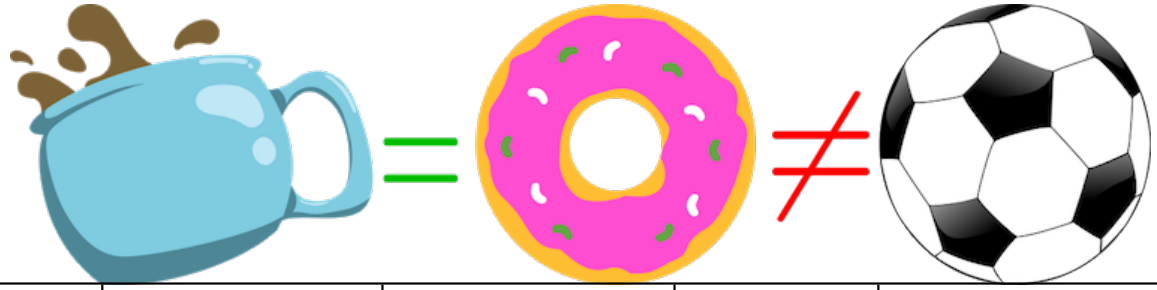




Figure: Felix Klein

What is Geometry? — Klein's Erlangen Program



	isometry	similarity	affine	projective	continuous
location					
length	✓				
area	✓				
perpendicularity	✓	✓			
parallelism	✓	✓	✓		
collinearity	✓	✓	✓	✓	
concurrence	✓	✓	✓	✓	
connectedness	✓	✓	✓	✓	✓

Given a manifold, and a transformation group acting on it, to study its invariants. — *Felix Klein*

What is Logic?[21]

Logic is the science that investigates the principles of valid reasoning.

what follows from what

The art of thinking and reasoning in strict accordance with the limitations and incapacities of the human misunderstanding. ◎ô◎

— *The Devil's Dictionary*

The study of invariants under all automorphisms (symmetries).

[21] Tarski: What are logical notions?

Logic as permutation-invariant theory

- 欧氏几何是关于保持欧氏距离、角度不变的线性变换群的不变量的研究.
- 狭义相对论是关于保持 $c^2t^2 - x^2 - y^2 - z^2$ 不变的洛伦兹变换群的不变量的研究.

Logic as permutation-invariant theory.

The study of invariants under all automorphisms (symmetries).

> *A notion is "logical" iff it is invariant under all possible one-one transformations of the universe of discourse onto itself.*
>
> — *Tarski*
>
> *Logic analyzes the meaning of the concepts common to all the sciences, and establishes the general laws governing the concepts.*
>
> — *Tarski*

Contents

Meta-Theorems

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Would we gain anything by adding more connectives to the language?

Exclusive Disjunction

$$
\begin{aligned}
v(p \oplus q) &= v(p) + v(q) \bmod 2 \\
&\Downarrow \\
p \oplus q &\equiv (\neg p \wedge q) \vee (p \wedge \neg q) \\
&\equiv (p \vee q) \wedge \neg(p \wedge q) \\
&\equiv \neg(p \leftrightarrow q) \\
&\equiv \neg p \leftrightarrow q \\
&\equiv p \leftrightarrow \neg q
\end{aligned}
$$

p	q	$p \oplus q$
0	0	0
0	1	1
1	0	1
1	1	0

$$\frac{p \oplus q \quad p}{\neg q} \qquad \frac{p \oplus q \quad \neg p}{q} \qquad \frac{p}{p \oplus q}\ ?$$

Nim Game

Given several piles of stones. Two players take turns moving. Each move consists of selecting one of the piles and removing any positive number of stones from it. The winner is the player who removes the last stone.

★★★

★★★★

★★★★★

3	0	1	1
4	1	0	0
5	1	0	1
2	0	1	0

$$\begin{array}{cccccc} & x & a_1 & \cdots & a_n \\ \oplus & y & b_1 & \cdots & b_n \\ \hline = & z & c_1 & \cdots & c_n \end{array}$$

where $c_i := a_i \oplus b_i$

$$x \oplus (y \oplus z) = (x \oplus y) \oplus z$$

$$x \oplus y = y \oplus x$$

$$x \oplus 0 = x$$

$$x \oplus x = 0$$

$$x \oplus y = x \oplus z \implies y = z$$

Nim Game

Theorem (Bouton Theorem)

In a Nim game with piles of size $x_1, \ldots, x_n$, the first player has a winning strategy iff $\bigoplus_{i=1}^{n} x_i \neq 0$.

1. If $\bigoplus_{i=1}^{n} x_i \neq 0$, it is possible to make a move $x_k - x'_k$ so that $x_1 \oplus \cdots x'_k \cdots \oplus x_n = 0$, where $x'_k := x_k \oplus \bigoplus_{i=1}^{n} x_i$.

   Here's how we construct such a move. Form the nim-sum as a column addition, and look at the leftmost column with an odd number of 1's. Choose the pile that have a 1 in that column.
2. If $\bigoplus_{i=1}^{n} y_i = 0$, and y_k is changed to $y'_k < y_k$, then $y_1 \oplus \cdots y'_k \cdots \oplus y_n \neq 0$, because otherwise the cancellation law would imply that $y_k = y'_k$.

Example

Example

Let $\#$ be a three-place proposition connective.
The interpretation of $\#$ is given by

$$\nu(\#(p,q,r)) = \left\lfloor \frac{\nu(p)+\nu(q)+\nu(r)}{2} \right\rfloor$$

then

$$\#(p,q,r) \equiv (p \wedge q) \vee (p \wedge r) \vee (q \wedge r)$$

p	q	r	$\#(p,q,r)$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Truth Table & Truth/Boolean Function

A truth assignment for $\mathscr{L}^0$ is a function $\nu : \mathrm{Var} \to \{0, 1\}$.

p	$\neg p$
0	1
1	0

p	q	$p \wedge q$	$p \vee q$	$p \to q$	$p \leftrightarrow q$	$\cdots$
0	0	0	0	1	1	$\cdots$
0	1	0	1	1	0	$\cdots$
1	0	0	1	0	0	$\cdots$
1	1	1	1	1	1	$\cdots$

A n-place truth/Boolean function is a function $F : \{0, 1\}^n \to \{0, 1\}$.

x	$F_\neg(x)$
0	1
1	0

x	y	$F_\wedge(x, y)$	$F_\vee(x, y)$	$F_\to(x, y)$	$F_\leftrightarrow(x, y)$	$\cdots$
0	0	0	0	1	1	$\cdots$
0	1	0	1	1	0	$\cdots$
1	0	0	1	0	0	$\cdots$
1	1	1	1	1	1	$\cdots$

There are 2^{2^n} distinct truth functions with n places.

Truth Table & Truth/Boolean Function

$$\nu : \mathrm{Var} \to \{0,1\}$$
$$F : \{0,1\}^n \to \{0,1\}$$

$\nu(p_1),\ldots,\nu(p_n)$		$x_1,\ldots,x_n$	$F(x_1,\ldots,x_n)$		$\nu(A)$
$\nu_1(p_1),\ldots,\nu_1(p_n)$	$=$	$0,\ldots,0$	$F(0,\ldots,0)$	$=$	$\nu_1(A)$
$\vdots$		$\vdots$	$\vdots$		$\vdots$
$\nu_{2^n}(p_1),\ldots,\nu_{2^n}(p_n)$	$=$	$1,\ldots,1$	$F(1,\ldots,1)$	$=$	$\nu_{2^n}(A)$

Theorem (Post1921)

任意 n 元布尔函数 $F : \{0,1\}^n \to \{0,1\}$ 都可以被某个命题逻辑公式表示. 即, 存在至多包含 n 个原子公式 $p_1,\ldots,p_n$ 的公式 A, 使得对任意赋值 $\nu : \mathrm{Var} \to \{0,1\}$ 都有

$$F(\nu(p_1),\ldots,\nu(p_n)) = \nu(A)$$

Theorem (Post1921)

Every truth function $F : \{0,1\}^n \to \{0,1\}$ can be represented by some formula whose only connectives are $\neg, \wedge, \vee$.

Proof.

$$p_i^{x_i} := \begin{cases} p_i & \text{if } x_i = 1 \\ \neg p_i & \text{otherwise} \end{cases}$$

Case1: $F(\mathbf{x}) = 0$ for all $\mathbf{x} \in \{0,1\}^n$.
Let $A := p \wedge \neg p$.
Case2:

$$A := \bigvee_{\mathbf{x}:F(\mathbf{x})=1} \bigwedge_{i=1}^{n} p_i^{x_i}$$

Case1: $F(\mathbf{x}) = 1$ for all $\mathbf{x} \in \{0,1\}^n$.
Let $B := p \vee \neg p$.
Case2:

$$B := \bigwedge_{\mathbf{x}:F(\mathbf{x})=0} \bigvee_{i=1}^{n} p_i^{1-x_i}$$

□

Normal Form

Corollary

Every formula which is not a contradiction is logically equivalent to a formula of disjunctive normal form (DNF):

$$\bigvee_{i=1}^{m} \bigwedge_{j=1}^{n} \pm p_{ij}$$

Corollary

Every formula which is not a tautology is logically equivalent to a formula of conjunctive normal form (CNF):

$$\bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} \pm p_{ij}$$

Proof.

$$\neg A \equiv \bigvee_{i=1}^{m} \bigwedge_{j=1}^{n} \pm p_{ij} \implies A \equiv \neg \left(\bigvee_{i=1}^{m} \bigwedge_{j=1}^{n} \pm p_{ij} \right) \equiv \bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} \mp p_{ij}$$

How many connectives are really necessary?

Definition (Adequate Sets of Connectives)

A set of connectives is adequate iff every truth function can be represented by a formula containing only connectives from that set.

- $\{\neg, \wedge, \vee\}$
- $\{\neg, \wedge\}; \{\neg, \vee\}; \{\neg, \rightarrow\}; \{\bot, \rightarrow\}$
- $\{\overline{\wedge}\}; \{\overline{\vee}\}$
- $\{\wedge, \vee, \rightarrow, \leftrightarrow\}; \{\neg, \leftrightarrow\}$ not adequate.

p	$\bot$
0	0
1	0

$$\bot \coloneqq p \wedge \neg p$$
$$p \mathbin{\overline{\wedge}} q \coloneqq \neg(p \wedge q)$$
$$p \mathbin{\overline{\vee}} q \coloneqq \neg(p \vee q)$$
$$\neg p \coloneqq p \mathbin{\overline{\wedge}} p$$
$$p \wedge q \coloneqq (p \mathbin{\overline{\wedge}} q) \mathbin{\overline{\wedge}} (p \mathbin{\overline{\wedge}} q)$$
$$p \vee q \coloneqq (p \mathbin{\overline{\wedge}} p) \mathbin{\overline{\wedge}} (q \mathbin{\overline{\wedge}} q)$$

p	q	$p \mathbin{\overline{\wedge}} q$	$p \mathbin{\overline{\vee}} q$
0	0	1	1
0	1	1	0
1	0	1	0
1	1	0	0

练习

Exercise

令 ℓ 是一个三元连接词: $\nu(\ell(p,q,r)) = 1 \iff \nu(p) + \nu(q) + \nu(r) = 1$. 证明: 不存在二元连接词 $\circ$ 和 $*$ 使得 $\ell(A,B,C) \equiv (A \circ B) * C$.

p	q	r	$\ell(p,q,r)$
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	0

提示: 任给二元布尔函数 $F_\circ, F_*$, 假设 $F_*(F_\circ(A,B),C) = F_{\ell(A,B,C)}$, 则 $F_\circ(1,1), F_\circ(1,0), F_\circ(0,0)$ 两两不相等.

3-valued Logics

p	$\neg p$
0	1
u	u
1	0

$\wedge$	0	u	1
0	0	u	0
u	u	u	u
1	0	u	1

$\vee$	0	u	1
0	0	u	1
u	u	u	u
1	1	u	1

$\rightarrow$	0	u	1
0	1	u	1
u	u	u	u
1	0	u	1

$\leftrightarrow$	0	u	1
0	1	u	0
u	u	u	u
1	0	u	1

Table: Bochvar: u as "meaningless"

p	$\neg p$
0	1
u	u
1	0

$\wedge$	0	u	1
0	0	0	0
u	0	u	u
1	0	u	1

$\vee$	0	u	1
0	0	u	1
u	u	u	1
1	1	1	1

$\rightarrow$	0	u	1
0	1	1	1
u	u	u	1
1	0	u	1

$\leftrightarrow$	0	u	1
0	1	u	0
u	u	u	u
1	0	u	1

Table: Kleene: u as "undefined"

p	$\neg p$
0	1
u	u
1	0

$\wedge$	0	u	1
0	0	0	0
u	0	u	u
1	0	u	1

$\vee$	0	u	1
0	0	u	1
u	u	u	1
1	1	1	1

$\rightarrow$	0	u	1
0	1	1	1
u	u	1	1
1	0	u	1

$\leftrightarrow$	0	u	1
0	1	u	0
u	u	1	u
1	0	u	1

Table: Lukasiewicz: u as "possible"

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Meta-Theorems

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

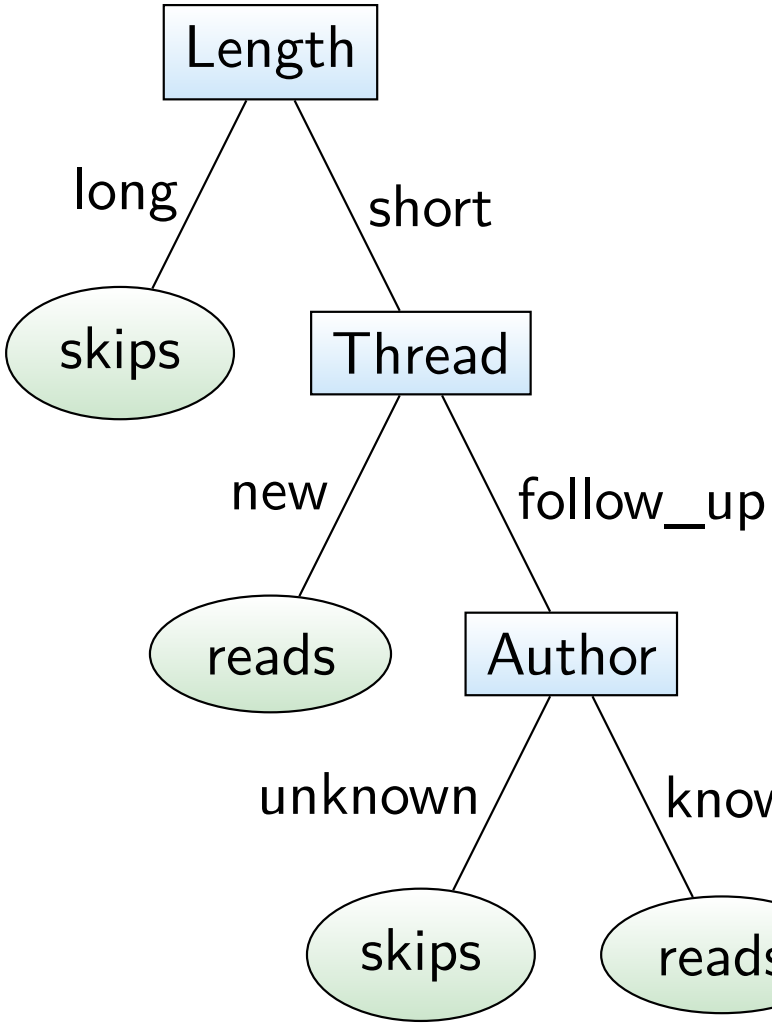
Answers to the Exercises

References2039

Normal Form

- A *literal* is an atomic formula or its negation.
- A formula is in negation normal form (NNF) iff it contains no other connectives than $\neg, \wedge, \vee$, and the negation sign $\neg$ appears in literals only.
- A clause is any formula of the form: $A_1 \vee A_2 \vee \cdots \vee A_n$, where $n \geq 1$ and $A_1, A_2, \ldots, A_n$ are literals.
- A Horn clause is a clause in which at most one literal is positive.
- A formula is in conjuctive normal form (CNF) iff it is a conjuction of one or more clauses.
- A formula is in disjunctive normal form (DNF) iff it is a disjunction of one or more conjunctions of one or more literals.
- A CNF formula is in full conjunctive normal form (FCNF) iff each of its variables appears exactly once in every clause. (similarly, full disjunctive normal form)

Decision Tree vs Horn Clause





skips ← Long

reads ← short ∧ new

reads ← short ∧ follow_up ∧ known

skips ← short ∧ follow_up ∧ unknown

- We want a small and efficient tree
- Ask the question which is most informative

NNF/CNF/DNF

subformula	replaced by
$A \leftrightarrow B$	$(\neg A \vee B) \wedge (A \vee \neg B)$
$A \rightarrow B$	$\neg A \vee B$
$\neg(A \vee B)$	$\neg A \wedge \neg B$
$\neg(A \wedge B)$	$\neg A \vee \neg B$
$\neg\neg A$	A
$\neg\forall x A$	$\exists x \neg A$
$\neg\exists x A$	$\forall x \neg A$

subformula	replaced by
$(A \wedge B) \vee C$	$(A \vee C) \wedge (B \vee C)$
$C \vee (A \wedge B)$	$(C \vee A) \wedge (C \vee B)$

subformula	replaced by
$(A \vee B) \wedge C$	$(A \wedge C) \vee (B \wedge C)$
$C \wedge (A \vee B)$	$(C \wedge A) \vee (C \wedge B)$

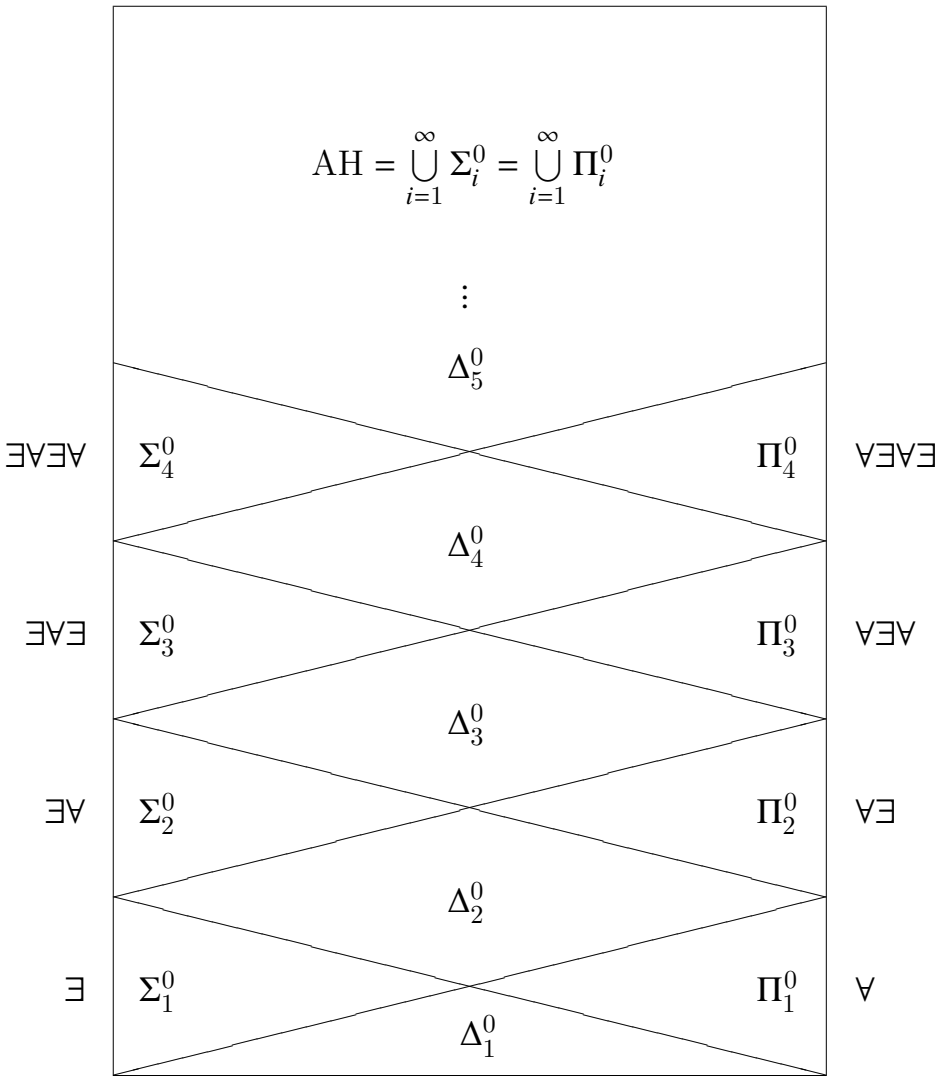
- Any formula can be equivalently transformed into NNF.
- Any quantifier-free formula can be equivalently transformed into CNF.
- Any quantifier-free formula can be equivalently transformed into DNF.

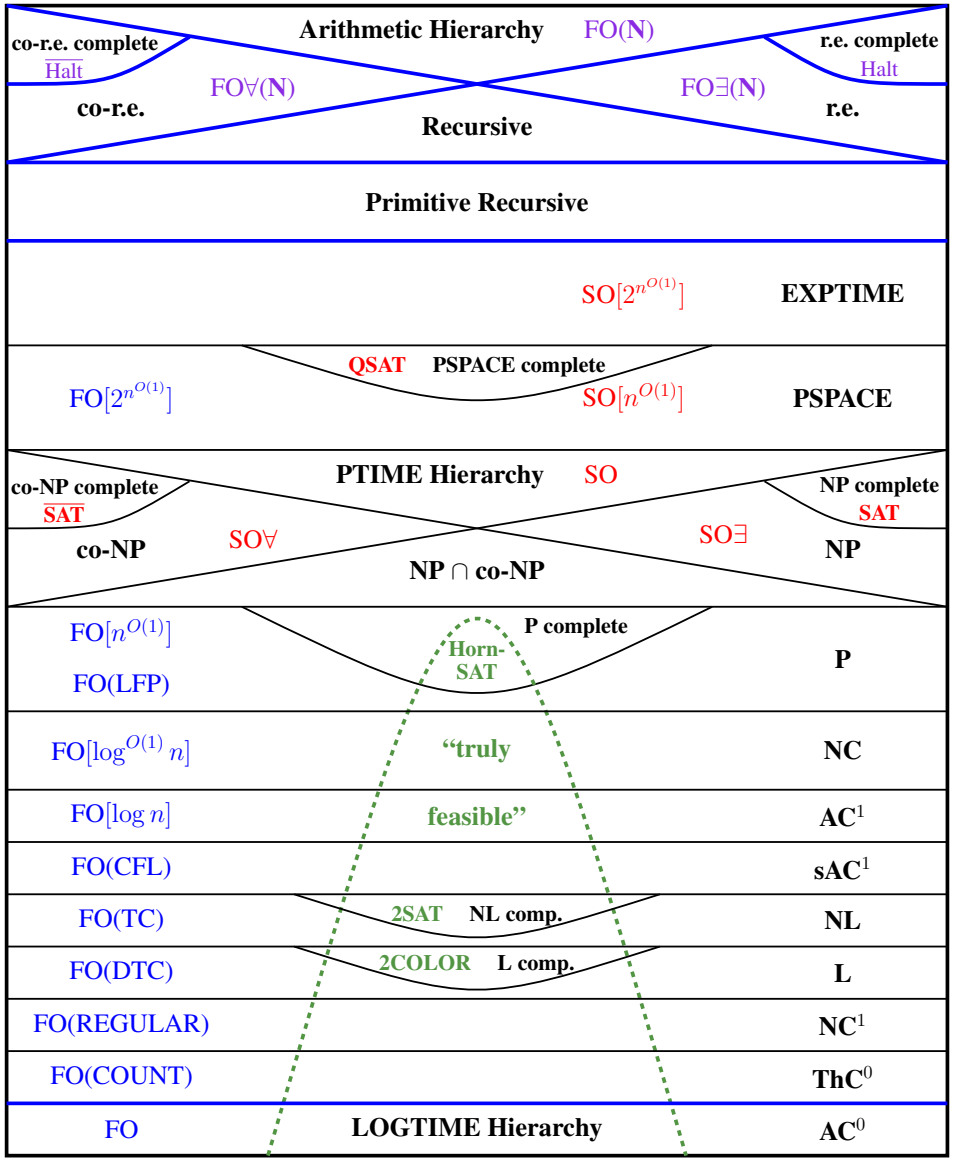
PNF

- A formula is in prenex normal form (PNF) iff all its quantifiers (if any) are in its prefix. (PCNF)
- Any formula can be equivalently transformed into PNF/PCNF.

subformula	replaced by	
$\neg\forall x A$	$\exists x \neg A$	
$\neg\exists x A$	$\forall x \neg A$	
$\forall x A(x) \wedge \forall x B(x)$	$\forall x (A(x) \wedge B(x))$	
$\exists x A(x) \vee \exists x B(x)$	$\exists x (A(x) \vee B(x))$	
$\forall x A(x)$	$\forall y A[y/x]$	where $y \notin \mathrm{Fv}(A) \cup \mathrm{Bv}(A)$
$A \vee Qx B$	$Qx(A \vee B)$	where $x \notin \mathrm{Fv}(A)$
$A \wedge Qx B$	$Qx(A \wedge B)$	where $x \notin \mathrm{Fv}(A)$
$A \to \forall x B$	$\forall x (A \to B)$	where $x \notin \mathrm{Fv}(A)$
$A \to \exists x B$	$\exists x (A \to B)$	where $x \notin \mathrm{Fv}(A)$
$\forall x A \to B$	$\exists x (A \to B)$	where $x \notin \mathrm{Fv}(A)$
$\exists x A \to B$	$\forall x (A \to B)$	where $x \notin \mathrm{Fv}(A)$

PNF & Arithmetical Hierarchy





Skolem Normal Form

- A formula is in *Skolem normal form* (SNF) iff it is in PNF (PCNF) without existential quantifiers.
- Skolemization: Replace $\forall x_1 \ldots \forall x_n \exists y A$ by $\forall x_1 \ldots \forall x_n A[f(x_1, \ldots, x_n)/y]$, where f is a new function symbol. If there are no universal quantifiers preceding $\exists$, replace $\exists x A$ by $A[c/x]$, where c is a new constant. Given A, in finitely many steps we can obtain its *Skolem normal form* A^{SNF} without existential quantifiers.

Theorem (Skolem Normal Form)

Any formula A can be equisatisfiably transformed into a SNF A^{SNF}, and we have $A^{\mathrm{SNF}} \models A$.

Remark: $A^{\mathrm{SNF}} \models A$ but the converse is not true in general.

Exercise (Transform the following sentence into SNF)

Who loves all animals, is in turn loved by someone.

Herbrand Normal Form

A is satisfiable iff A^{SNF} is satisfiable.

$$A^{\mathrm{HNF}} := \left(\neg(\neg A)^{\mathrm{SNF}}\right)^{\mathrm{NNF}}$$

$$\vDash A \iff \vDash A^{\mathrm{HNF}}$$

Example:

$$(\forall x \exists y \forall z A(x,y,z))^{\mathrm{SNF}} = \forall x z A(x, f(x), z)$$

$$(\forall x \exists y \forall z A(x,y,z))^{\mathrm{HNF}} = \exists y A(a, y, f(y))$$

Herbrand Universe

Definition (Herbrand Universe)

Given a sentence A in Skolem normal form,

- $H_0 \coloneqq \{\text{all constants in } A\}$. If no constant in A then $H_0 \coloneqq \{a\}$ for a new constant a.
- $H_{i+1} \coloneqq \left\{f(t_1, \ldots, t_n) : f \text{ in } A \text{ and } t_j \in H_i,\ j = 1, \ldots, n\right\}$
- $H_A \coloneqq \bigcup_{i \in \omega} H_i$

The Herbrand universe of a language $\mathscr{L}$ is the set of all closed terms of $\mathscr{L}$. If no constant in $\mathscr{L}$, then add a new constant to $\mathscr{L}$.

Herbrand Model

Definition (Herbrand Model)

A Herbrand model for $\mathscr{L}$ is (H, I) s.t.

- H is the Herbrand universe of $\mathscr{L}$.
- for every closed term t, $I(t) = t$.

Theorem

A formula A is satisfiable iff there is a Herbrand model satisfying it.

Proof.

Assume A is in Skolem normal form, and it is satisfied by some model (M, I). Then Herbrand model $(H_A, J) \vDash A$, where for each closed term $t : J(t) = t$, and for each predicate symbol P,
$J(P) = \{(t_1, \ldots, t_n) \in H_A : M, I \vDash P(t_1, \ldots, t_n)\}$. □

Herbrand's Theorem

For a quantifier-free formula $A(x_1, \ldots, x_n)$, the Herbrand expansion over a set D of closed terms is $\mathcal{E}(A, D) \coloneqq \{A(t_1, \ldots, t_n) : t_i \in D\}$.

Theorem (Herbrand's Theorem)

A sentence $\forall \mathbf{x} A(\mathbf{x})$ in Skolem normal form is unsatisfiable iff some finite subset $K \subset \mathcal{E}(A, H_A)$ is unsatisfiable.

Theorem (Herbrand's Theorem)

Suppose A is a sentence. Then

$$\vdash A \iff \vdash \bigvee_{i=1}^{m} A'(t_{i1}, \ldots, t_{in})$$

for some $m > 0$ and a finite sequence of terms t_{ij} with $1 \leq i \leq m$ and $1 \leq j \leq n$, where A' is obtained from A^{HNF} by dropping the quantifiers.

Remark: If a sentence is entailed by an FOL KB, it is entailed by a finite subset of the propositional KB. If not, may continue for ever. semidecidable

Resolution — Example

Given clauses $A = P(f(x)) \vee Q(x)$ and $B = \neg P(x) \vee \neg P(y)$.

1. Separate their variables by standard substitutions $\{x_1/x\}$ and $\{y_1/x, y_2/y\}$.

$$A' = \{P(f(x_1)), Q(x_1)\} \quad B' = \{\neg P(y_1), \neg P(y_2)\}$$

2. Pick a subset $C \subset A'$ containing literals all of the same sign, and a subset $D \subset B'$ containing literals all of the opposite sign of C such that $|C \cup D|$ is unifiable.

$$C = \{P(f(x_1))\} \quad D = \{\neg P(y_1), \neg P(y_2)\}$$

$$|C \cup D| = \{P(f(x_1)), P(y_1), P(y_2)\}$$

3. A most general unifier for $|C \cup D|$ is

$$\sigma = \{f(x_1)/y_1, f(x_1)/y_2\}$$

4. A resolvent for A and B is:

$$R = (A'\sigma \setminus C\sigma) \cup (B'\sigma \setminus D\sigma) = \{Q(x_1)\}$$

Resolution

Theorem (Soundness)

If R is a resolvent of A and B, then any model satisfying both A and B will also satisfy R.

- For a formula A, $\mathcal{R}(A)$ is A extended with all resolvents to clauses of A.
- The successive application of the resolution rule yields a complete proof procedure.

Theorem (Completeness)

If A is unsatisfiable, then $\mathcal{R}^n(A)$ will contain the empty clause for some n.

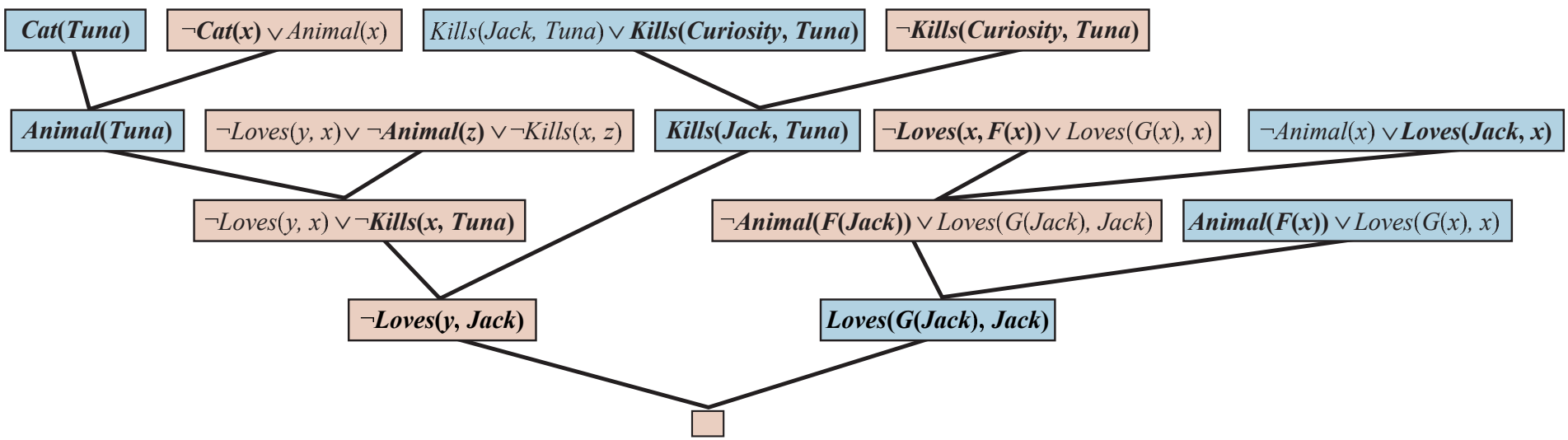
Resolution — Example

- Everyone who loves all animals is loved by someone.
- Anyone who kills an animal is loved by no one.
- Jack loves all animals.
- Either Jack or Curiosity killed the cat, who is named Tuna.
- Did Curiosity kill the cat?

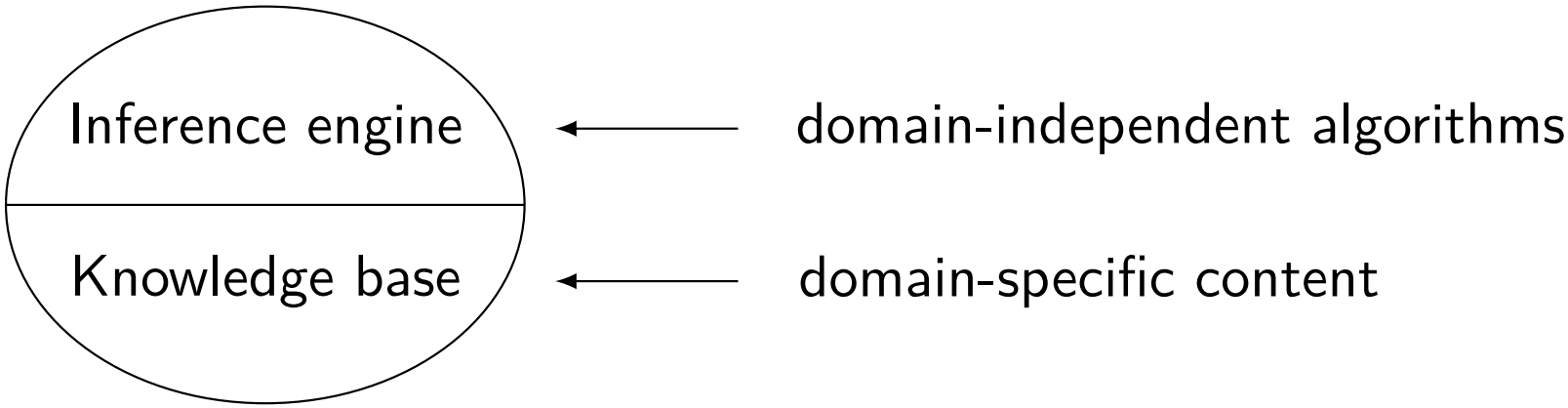
1. $\forall x(\forall y(\text{Animal}(y) \rightarrow \text{Love}(x, y)) \rightarrow \exists z \text{Love}(z, x))$
2. $\forall x((\exists z \text{Animal}(z) \land \text{Kill}(x, z)) \rightarrow \forall y \neg \text{Love}(y, x))$
3. $\forall x(\text{Animal}(x) \rightarrow \text{Love}(\text{Jack}, x))$
4. $\text{Kill}(\text{Jack}, \text{Tuna}) \lor \text{Kill}(\text{Curiosity}, \text{Tuna})$
5. $\text{Cat}(\text{Tuna})$
6. $\forall x(\text{Cat}(x) \rightarrow \text{Animal}(x))$
7. $\neg \text{Kill}(\text{Curiosity}, \text{Tuna})$

Resolution — Example

1. $\Big(\text{Animal}(F(x)) \vee \text{Love}(G(x), x)\Big) \wedge \Big(\neg\text{Love}(x, F(x)) \vee \text{Love}(G(x), x)\Big)$
2. $\neg\text{Love}(y, x) \vee \neg\text{Animal}(z) \vee \neg\text{Kill}(x, z)$
3. $\neg\text{Animal}(x) \vee \text{Love}(\text{Jack}, x)$
4. $\text{Kill}(\text{Jack}, \text{Tuna}) \vee \text{Kill}(\text{Curiosity}, \text{Tuna})$
5. $\text{Cat}(\text{Tuna})$
6. $\neg\text{Cat}(x) \vee \text{Animal}(x)$
7. $\neg\text{Kill}(\text{Curiosity}, \text{Tuna})$



Logical Agent: Knowledge-Based Agent



A knowledge-based agent uses its knowledge base to

- represent its background knowledge: states, actions, etc.
- incorporate new percepts
- update internal representations of the world
- deduce hidden properties of the world
- deduce appropriate actions

Reducing First Order Inference to Propositional Inference

- Universal Instantiation

$$\frac{\forall x A}{A[t/x]} \text{ where } t \text{ is a closed term}$$

  - can be applied several times to add new sentences
  - the new KB is logically equivalent to the old

- Existential Instantiation

$$\frac{\exists x A}{A(a)} \text{ where } a \text{ is a new constant}$$

  - can be applied once to replace the existential sentence
  - the new KB is not equivalent to the old
  - but is satisfiable iff the old KB was satisfiable

Reducing First Order Inference to Propositional Inference

- Claim: a sentence is entailed by the new KB iff it is entailed by the original KB
- Claim: every FOL KB can be propositionalized so as to preserve entailment

Theorem (Herbrand's Theorem)

If a sentence A is entailed by an FOL KB, it is entailed by a finite subset of the propositional KB.

- Idea: propositionalize KB and query, apply resolution, return result

for $n = 0$ to ∞ do
  create a propositional KB by instantiating with depth-n terms see if A is entailed by this KB

- Problem: works if A is entailed, loops if A is not entailed
- Theorem (Turing, Church): entailment in FOL is semidecidable.

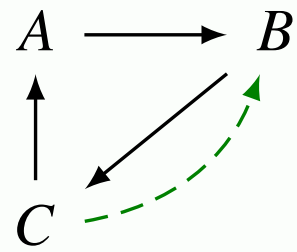
GNF & Self-Reference Paradox

Example

A. The next statement is false. $A \leftrightarrow \neg B$

B. The next statement is false. $B \leftrightarrow \neg C$

C. The first statement is false. $C \leftrightarrow \neg A$



- The above odd cycle is a generalization of the liar paradox $\bullet \circlearrowright$
- To remove the paradox, we can add the edge C to B, i.e., $C \leftrightarrow \neg A \wedge \neg B$.
- Then $A = C = 0, B = 1$.

Definition (Graph Normal Form)

- A *basic formula*, over alphabet Σ, is an equivalence $x \leftrightarrow \bigwedge_{i \in I} \neg x_i$.
- A *theory in GNF* is a set of basic formulas with each $x \in \Sigma$ occurs exactly once on the left of such an equivalence.

- For a theory Δ in GNF, we form a graph $G(\Delta) = (W, R)$, with the vertex set W and an edge from each variable occurring on the left of $\leftrightarrow$ in a basic formula to all the variables occurring on the right of $\leftrightarrow$.
- Conversely, given a directed graph $F = (W, R)$, we can form a theory in GNF $D(F) := \left\{ x \leftrightarrow \bigwedge_{y:Rxy} \neg y : x \in W \right\}$.
- Obviously, we have $D(G(\Delta)) = \Delta$ and $G(D(F)) = F$.
- We say that $\nu \vDash (W, R)$ iff ν assignes truth-values to W so that
$$\nu(x) = 1 \iff \forall y : Rxy \to \nu(y) = 0$$
- Obviously, for a graph F: $\nu \vDash F \iff \nu \vDash D(F)$.
For a theory Δ in GNF: $\nu \vDash \Delta \iff \nu \vDash G(\Delta)$.

A directed graph (W, R) is *finitary* iff $|\{y : Rxy\}| < \infty$ for all $x \in W$.

Theorem (Richardson's Theorem)

- *A finitary graph with no odd cycle is satisfiable.*
- *If every odd cycle has at least two symmetrical arcs, then the graph is satisfiable.*

Remark: Sinks (vertices with no outgoing edges) must be assigned 1.
Vertices with edges to sinks must be assigned 0....

Contents

Meta-Theorems

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic
- Introduction
- Syntax
- Semantics
- Formal System
- Definability & Isomorphism
- What is Logic?
- Connectives
- Normal Forms

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Model & Theory

- $\mathrm{Mod}(A) \coloneqq \{\mathcal{M} : \mathcal{M} \vDash A\}$
- $\mathrm{Mod}(\Gamma) \coloneqq \bigcap_{A \in \Gamma} \mathrm{Mod}(A)$
- $\mathrm{Th}(\mathcal{M}) \coloneqq \{A : \mathcal{M} \vDash A\}$
- $\mathrm{Th}(\mathcal{K}) \coloneqq \bigcap_{\mathcal{M} \in \mathcal{K}} \mathrm{Th}(\mathcal{M})$
- $\mathrm{Cn}(\Gamma) \coloneqq \{A : \Gamma \vDash A\}$

- $\Gamma \subset \Gamma' \implies \mathrm{Mod}(\Gamma') \subset \mathrm{Mod}(\Gamma)$
- $\mathcal{K} \subset \mathcal{K}' \implies \mathrm{Th}(\mathcal{K}') \subset \mathrm{Th}(\mathcal{K})$
- $\Gamma \subset \mathrm{Th}(\mathrm{Mod}(\Gamma))$
- $\mathcal{K} \subset \mathrm{Mod}(\mathrm{Th}(\mathcal{K}))$
- $\mathrm{Mod}(\Gamma) = \mathrm{Mod}(\mathrm{Th}(\mathrm{Mod}(\Gamma)))$
- $\mathrm{Th}(\mathcal{K}) = \mathrm{Th}(\mathrm{Mod}(\mathrm{Th}(\mathcal{K})))$
- $\mathrm{Cn}(\Gamma) = \mathrm{Th}(\mathrm{Mod}(\Gamma))$
- $\Gamma \subset \Gamma' \implies \mathrm{Cn}(\Gamma) \subset \mathrm{Cn}(\Gamma')$
- $\mathrm{Cn}(\mathrm{Cn}(\Gamma)) = \mathrm{Cn}(\Gamma)$

Theory & Axiomatization

- A set Γ of sentences is a theory iff $\Gamma = \mathrm{Cn}(\Gamma)$.
- A theory Γ is complete iff for every sentence A, either $A \in \Gamma$ or $\neg A \in \Gamma$.
- A theory Γ is finitely axiomatizable iff $\Gamma = \mathrm{Cn}(\Sigma)$ for some finite set Σ of sentences.
- A theory Γ is axiomatizable iff there is a decidable set Σ of sentences s.t. $\Gamma = \mathrm{Cn}(\Sigma)$.
- A class $\mathcal{K}$ of models is an elementary class (EC) iff $\mathcal{K} = \mathrm{Mod}(A)$ for some sentence A.
- A class $\mathcal{K}$ of models is an elementary class in wider sense (EC_Δ) iff $\mathcal{K} = \mathrm{Mod}(\Sigma)$ for some set Σ of sentences.

Definitional Extension

Definition (Definitional Extension)

$\mathscr{L}$ is a first order language, and T is a $\mathscr{L}$ theory.

- An explicit definition of a predicate symbol $P \in \mathscr{L}^+$ in $\mathscr{L}$ is a sentence $\forall \mathbf{x}(P(\mathbf{x}) \leftrightarrow A(\mathbf{x}))$, where $A(\mathbf{x})$ is a $\mathscr{L}$-formula.
- An explicit definition of an n-ary function symbol $f \in \mathscr{L}^+$ in $\mathscr{L}$ is a sentence $\forall \mathbf{x}(f(\mathbf{x}) = y \leftrightarrow A(\mathbf{x}, y))$, where $A(\mathbf{x}, y)$ is a $\mathscr{L}$-formula and $\mathrm{T} \vdash \forall \mathbf{x} \exists! y A(\mathbf{x}, y)$.
- An explicit definition of a constant symbol $c \in \mathscr{L}^+$ in $\mathscr{L}$ is a sentence $\forall x(x = c \leftrightarrow A(x))$, where $A(\mathbf{x})$ is a $\mathscr{L}$-formula and $\mathrm{T} \vdash \exists! x A(x)$.

A *definitional extension* of a $\mathscr{L}$-theory T to $\mathscr{L}^+$ is a T^+-theory

$$\mathrm{T}^+ = \mathrm{T} \cup \{\delta_s : s \in \mathscr{L}^+ \setminus \mathscr{L}\}$$

where the sentence δ_s is an explicit definition of s in $\mathscr{L}$.

Definitionally Equivalent

Let T^+ be a definitional extension of T. Define the inclusion translation map $I : \mathrm{T} \to \mathrm{T}^+$ and the reduction map $R : \mathrm{T}^+ \to \mathrm{T}$: for each symbol $s \in \mathscr{L}^+ \setminus \mathscr{L}$, let $Rs = \delta_s$, where δ_s is the explicit definition of s. For $s \in \mathscr{L}$, $Rs = s$.

Theorem

If T^+ is a definitional extension of T, then

- $\mathrm{T} \vdash A \leftrightarrow RIA$ *for any $\mathscr{L}$-formula A.*
- $\mathrm{T}^+ \vdash A \leftrightarrow IRA$ *for any $\mathscr{L}^+$-formula A.*
- T^+ *is a conservative extension of* T.

Definition (Definitionally Equivalent)

Let T_1 be a $\mathscr{L}_1$-theory and T_2 be a $\mathscr{L}_2$-theory. Then T_1 and T_2 are said to be *definitionally equivalent* if there is a definitional extension T_1^+ of T_1 to $\mathscr{L}_1 \cup \mathscr{L}_2$ and a definitional extension T_2^+ of T_2 to $\mathscr{L}_1 \cup \mathscr{L}_2$ such that $\mathrm{Cn}(\mathrm{T}_1^+) = \mathrm{Cn}(\mathrm{T}_2^+)$.

Consistency & Satisfiability

- Γ is consistent iff $\Gamma \nvdash \bot$.
- Γ is maximal iff for every formula A, either $A \in \Gamma$ or $\neg A \in \Gamma$.
- Γ is maximal consistent iff it is both consistent and maximal.
- Γ is satisfiable iff $\mathrm{Mod}(\Gamma) \neq \emptyset$.
- Γ is finitely satisfiable iff every finite subset of Γ is satisfiable.

Soundness Theorem

Theorem (Soundness Theorem)

$$\Gamma \vdash A \implies \Gamma \models A$$

Proof.

by induction on derivation lengths. □

Truth in a model is preserved under making deductions.

Kurt Gödel 1906-1978

"I am unprovable."[22]

- Completeness.
  I think (consistently), therefore I am.
  (Consistency implies existence.)
- Incompleteness.
  1. provable < true
  2. un-self-aware
- Consistency of AC and CH.





[22] Gödel: On formally undecidable propositions of Principia Mathematica and related systems.

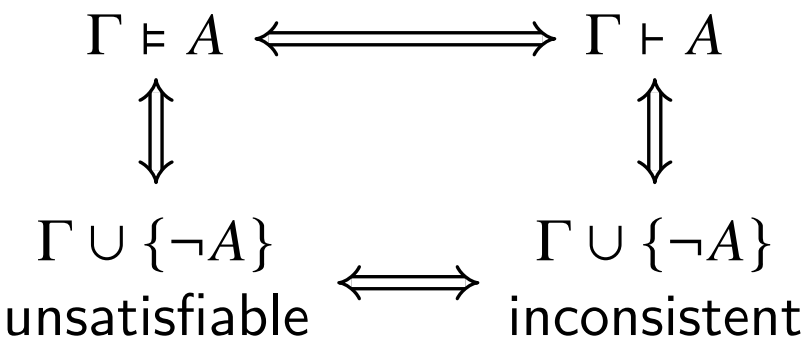
Completeness Theorem

Theorem (Completeness Theorem — Gödel1930)

$$\Gamma \vDash A \implies \Gamma \vdash A$$

Corollary

Any consistent set of formulas is satisfiable.



$\vdash$ captures $\vDash$
No more, no less

Proof Sketch

1. 给定一致的公式集 Γ
2. 向语言 $\mathscr{L}$ 中添加新常元 $c_1, c_2, c_3, \ldots$ 得到语言 $\mathscr{L}^+$, 一致的公式集 Γ 在新语言 $\mathscr{L}^+$ 中保持一致 — 验证一致性
3. 对于每个变元 x, 枚举所有以 x 为自由变元的公式: $A_1(x), A_2(x), A_3(x), \ldots$
4. 对于每个公式 $A_i(x)$, 添加形如 $\exists x A_i(x) \to A_i[c/x]$ 的公式, 得到一个扩张的一致公式集 $\Theta \supset \Gamma$ — 验证一致性
5. 应用 Lindenbaum Lemma, 将 Θ 扩张为极大一致集 Δ, 使得对新语言 $\mathscr{L}^+$ 中的每个公式 A 都有: $A \in \Delta$ 或 $\neg A \in \Delta$
6. 应用 Term Models Lemma, 为极大一致集 Δ 定义一个解释 $\mathcal{M}, \nu$
7. 验证: $\mathcal{M}, \nu \vDash A \iff A \in \Delta$
8. 将解释限制到原来的语言 $\mathscr{L}$ 上, 由于 $\Gamma \subset \Delta$, 所以: $\mathcal{M}, \nu \vDash \Gamma$

Proof of Completeness Theorem — step1

Lemma (Lindenbaum Lemma)

Any consistent set Θ of sentences can be extended to a maximal consistent set Δ of sentences of the same language.

Proof.

Arrange all the sentences in a sequence $\langle A_\xi : \xi < \kappa \rangle$.

$$\Theta_0 := \Theta$$

$$\Theta_{\xi+1} := \begin{cases} \Theta_\xi \cup \{A_\xi\} & \text{if } \Theta_\xi \cup \{A_\xi\} \text{ is consistent} \\ \Theta_\xi \cup \{\neg A_\xi\} & \text{otherwise} \end{cases}$$

$$\Theta_\xi := \bigcup_{\alpha<\xi} \Theta_\alpha \qquad \text{if } \xi \text{ is a limit ordinal}$$

$\Delta := \bigcup_{\xi<\kappa} \Theta_\xi$ is maximal consistent. □

Proof of Completeness Theorem — step2

A set Δ is Henkin iff Δ is maximally consistent and for any formula of the form $\exists x A$ there exists a closed term t s.t. $\exists x A \to A[t/x] \in \Delta$.

Lemma (Closure Lemma)

If Γ is consistent, then there is a Henkin set $\Delta \supset \Gamma$.

Proof.

Let C be a set of new constants. $\mathscr{L}^+ := \mathscr{L} \cup C$, $\mathscr{L} \cap C = \emptyset$, $|C| = |\mathscr{L}|$. Assume $|C| = \kappa$, and $C = \{c_\xi : \xi < \kappa\}$. Arrange all formulas of $\mathscr{L}^+$ with at most one free variable in a sequence $\langle A_\xi : \xi < \kappa \rangle$. Let

$$\Gamma_0 := \Gamma$$

$$\Gamma_{\xi+1} := \Gamma_\xi \cup \{\exists x A_\xi(x) \to A_\xi[c_\beta/x]\}$$

where c_β is the first new constant not occurring in $\Gamma_\xi \cup \{A_\xi\}$.

$$\Gamma_\xi := \bigcup_{\alpha<\xi} \Gamma_\alpha \quad \text{if } \xi \text{ is a limit ordinal}$$

Then $\Theta := \bigcup_{\xi<\kappa} \Gamma_\xi$ is consistent, and we can extend Θ to a maximal consistent set $\Delta \supset \Theta$ by Lindenbaum lemma. □

Proof of Completeness Theorem — step3

Lemma (Term Models Lemma)

If Δ is maximal consistent, then there is an interpretation $\mathcal{M}, v$ s.t.

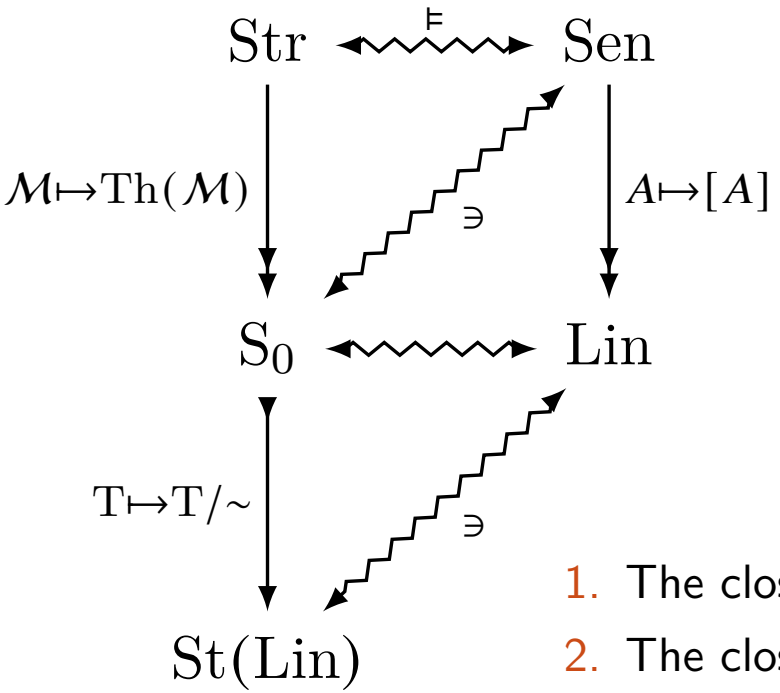
$$\mathcal{M}, v \vDash A \iff A \in \Delta$$

Proof.

- $M := \{[t] : t \in \mathrm{Term}\}$ where Term is the set of terms of $\mathscr{L}^+$, $[t] := \{s : s \sim t\}$, $s \sim t := s = t \in \Delta$
- $([t_1], \ldots, [t_n]) \in P^{\mathcal{M}} := P(t_1, \ldots, t_n) \in \Delta$
- $f^{\mathcal{M}}([t_1], \ldots, [t_n]) := [f(t_1, \ldots, t_n)]$
- $c^{\mathcal{M}} = [c]$
- $v(x) := [x]$

□

Stone Space of Lindenbaum Algebra



- Str: the class of structures.
- Sen: the set of sentences.
- S_0: the set of complete theories.
- $\mathrm{Lin} = \mathrm{Sen}/\sim$
- $\mathrm{St}(M)$: the set of ultrafilters of M.

1. The closed subsets of Sen are precisely the theories.
2. The closed subsets of S_0 compose a Hausdorff topology.
3. $\mathcal{M} \mapsto \mathrm{Th}(\mathcal{M}) : \mathrm{Str} \to \mathrm{S}_0$ is continuous, and $[\mathcal{M}] \mapsto \mathrm{Th}(\mathcal{M})$ is a homomorphism, S_0 is a Kolmogorov quotient $\mathrm{Str}/\equiv$.
4. For every theory T, $\mathrm{T}/\sim$ is a filter of Lin.
5. For every complete theory T, $\mathrm{T}/\sim$ is an ultrafilter of Lin.
6. $\mathrm{T} \mapsto \mathrm{T}/\sim : \mathrm{S}_0 \rightarrowtail \mathrm{St}(\mathrm{Lin})$ is a homomorphism.
7. The image is dense in $\mathrm{St}(\mathrm{Lin})$.
8. The image is a closed subspace of $\mathrm{St}(\mathrm{Lin})$.
9. $\mathrm{T} \mapsto \mathrm{T}/\sim : \mathrm{S}_0 \twoheadrightarrow \mathrm{St}(\mathrm{Lin})$ iff the topology on S_0 is compact.

Compactness Theorem

Theorem (Compactness Theorem)

A set of formulas is satisfiable iff it is finitely satisfiable.

Corollary

If $\Gamma \vDash A$*, then there is a finite* $\Gamma_0 \subset \Gamma$ *s.t.* $\Gamma_0 \vDash A$*.*

Corollary

There is a countable model $\mathcal{M} \equiv \mathcal{N}$ *but* $\mathcal{M} \ncong \mathcal{N}$*, where* $\mathcal{N} = (\mathbb{N}, 0, S, +, \cdot)$ *is the standard model of arithmetic.*

Proof.

By compactness theorem, $\mathrm{Th}(\mathcal{N}) \cup \{c \neq 0, c \neq S0, c \neq SS0, \dots\}$ has a model $\mathcal{M}$. □

Ultraproduct & Łoś Theorem

Definition (Ultraproduct)

Suppose $\{\mathcal{M}_i : i \in I\}$ is a set of models, and U is an ultrafilter on I. Define $\mathcal{N} := \prod_{i \in I} \mathcal{M}_i / U$ as follows:

- $N := \prod_{i \in I} M_i / \sim = \left\{ [f] : f \in \prod_{i \in I} M_i \right\}$ where
  $f \sim g := \left\{ i \in I : f(i) = g(i) \right\} \in U$
- $P^{\mathcal{N}}\big([f_1], \ldots, [f_n]\big) := \left\{ i \in I : P^{\mathcal{M}_i}(f_1(i), \ldots, f_n(i)) \right\} \in U$
- $F^{\mathcal{N}}\big([f_1], \ldots, [f_n]\big) := [f]$ where $f(i) := F^{\mathcal{M}_i}\big(f_1(i), \ldots, f_n(i)\big)$

Theorem (Łoś Theorem)

$$\prod_{i \in I} \mathcal{M}_i / U \models A\left([f_1], \ldots, [f_n]\right) \iff \left\{ i \in I : \mathcal{M}_i \models A\big(f_1(i), \ldots, f_n(i)\big) \right\} \in U$$

Ultrapower

Let $j : a \mapsto [f_a]$, where $f_a(i) = a$ for $i \in I$. Then

$$j : \mathcal{M} \prec \prod_{i \in I} \mathcal{M}/U$$

Theorem (Kiesler-Shelah)

$\mathcal{M} \equiv \mathcal{N}$ iff for some I and an ultrafilter U on I, $\prod_{i \in I} \mathcal{M}/U \cong \prod_{i \in I} \mathcal{N}/U$.

Compactness Theorem

Corollary (Compactness Theorem)

A set Γ of formulas is satisfiable iff it is finitely satisfiable.

Proof.

Let $I := \{\Delta \subset \Gamma : |\Delta| < \infty\}$.
Then $\forall \Delta \in I \; \exists \mathcal{M}_\Delta : \mathcal{M}_\Delta \vDash \Delta$.
Let $\hat{A} := \{\Delta \in I : A \in \Delta\}$.
Then $F := \{\hat{A} : A \in \Gamma\}$ has the finite intersection property because

$$\{A_1, \ldots, A_n\} \in \hat{A}_1 \cap \cdots \cap \hat{A}_n$$

By the ultrafilter theorem, F can be extended to an ultrafilter U on I.
For $A \in \Gamma$,

$$\hat{A} \in U \;\; \& \;\; \hat{A} \subset \{\Delta \in I : \mathcal{M}_\Delta \vDash A\} \implies \{\Delta \in I : \mathcal{M}_\Delta \vDash A\} \in U$$
$$\implies \prod_{\Delta \in I} \mathcal{M}_\Delta / U \; \vDash A$$

Compactness and Compactification

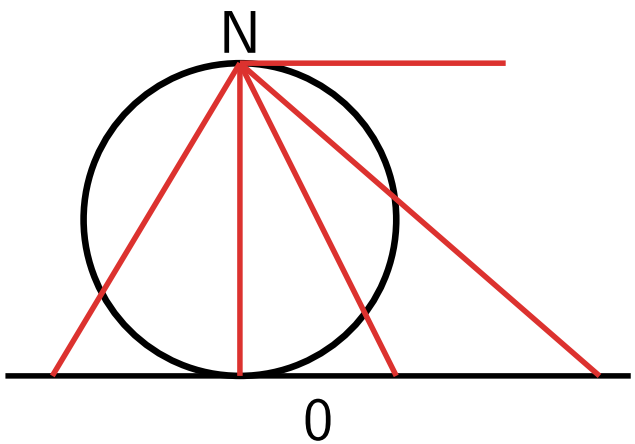
- Extreme value theorem: A continuous real-valued function on a compact space is bounded and attains its maximum and minimum values.
- A subset of a topological space is *compact* iff every open cover of it has a finite subcover.
- Heine-Borel Theorem: A subset of $\mathbb{R}$ is compact iff it is closed and bounded.
- Cantor's Intersection Theorem: A decreasing nested sequence of non-empty, closed and bounded subsets of $\mathbb{R}$ has a non-empty intersection.
- Bolzano-Weierstrass Theorem: Every bounded sequence of real numbers has a convergent subsequence.

Compactness

$$\text{finite} \implies \text{infinite}$$
$$\text{local} \implies \text{global}$$

Compactification

$$\mathbb{R} \implies \mathbb{R} \cup \{-\infty, +\infty\}$$



$$x \mapsto \left(\frac{x}{1+x^2}, \frac{x^2}{1+x^2} \right)$$

Theorem

There is a model of reals with infinite numbers.

Proof.

$$\mathrm{Th}(\mathcal{R}) \cup \{c > n : n \in \mathbb{N}\}$$

□

Theorem

There is a model of reals with infinitesimals.

Proof.

$$\mathrm{Th}(\mathcal{R}) \cup \left\{0 < c < \frac{1}{n+1} : n \in \mathbb{N}\right\}$$

□



Figure: Robinson

Let U be a nonprinciple ultrafilter on $\mathbb{N}$.

$$\prod_{i \in \mathbb{N}} \mathcal{R}/U \vDash \mathrm{Th}(\mathcal{R})$$

Let $\varepsilon := \left[(1, \frac{1}{2}, \frac{1}{3}, \ldots)\right] \in \mathbb{R}$.
For any $n \in \mathbb{N}$,

$$\prod_{i \in \mathbb{N}} \mathcal{R}/U \vDash \varepsilon < \frac{1}{n}$$

Problem

Show that there is an alternative model of the rational numbers $\mathbb{Q}$ in which $\sqrt{2}$ is rational.

Take the language of the theory $\mathbb{Q}$ of rational numbers and extend it with the constant q. Consider the following axioms:

$$A_n := 1.d_1 \dots d_n \leq q \leq 1.d_1 \dots d_n + 10^{-n}$$

for every $n \in \mathbb{N}$ with $1.d_1 d_2 \dots$ the infinite decimal expansion of $\sqrt{2}$.
Then $\mathbb{Q} \cup \{A_n\}_{n \in \mathbb{N}}$ is finitely satisfiable, and has a model by the Compactness Theorem.
In this model $(1.d_1 \dots d_n)^2 \leq q^2 \leq (1.d_1 \dots d_n + 10^{-n})^2$ for every $n \in \mathbb{N}$.
By definition $(1.d_1 \dots d_n)^2 \leq 2 \leq (1.d_1 \dots d_n + 10^{-n})^2$ for every $n \in \mathbb{N}$.
Thus q^2 cannot be distinguished from 2 inside the model.

Theorem

- *$\mathcal{K}$ is EC_Δ iff $\mathcal{K}$ is closed under ultraproducts and elementary equivalence.*
- *$\mathcal{K}$ is EC iff both $\mathcal{K}$ and the complement of $\mathcal{K}$ are closed under ultraproducts and elementary equivalence.*

Application of Compactness — Limitation of FOL

"有无穷多头猪."

Theorem

令 P 是一个一元谓词. 不存在一个一阶逻辑句子 A 使得对任何可以解释 A 中使用的非逻辑符号以及谓词 P 的结构 $\mathcal{M}$,

$$\mathcal{M} \vDash A \iff P^{\mathcal{M}}\text{是无穷的}$$

Proof.

假设有一个一阶句子 A 表达了 "有无穷多个 P".
考虑扩张的语言 $\mathscr{L}^+ := \mathscr{L} \cup \{c_i : i \in \mathbb{N}\}$.
考虑公式集 $\Gamma := \{\neg A, \ldots, c_i \neq c_j, \ldots, P(c_i), \ldots\}(i \neq j)$.
对任意 Γ 的有穷子集 Γ_0 都有 $\Gamma_0 \nvdash \bot$, 所以 $\Gamma_0 \nvdash \bot$, 所以 $\Gamma \nvdash \bot$. 但 $\Gamma \vDash \bot$. □

Application of Compactness — Limitation of FOL

Theorem

If a set Γ of sentences has arbitrarily large finite models, then it has an infinite model.

Proof.

Consider the set $\Gamma \cup \{\varphi^{\geq n} : n \in \mathbb{N}\}$, where $\varphi^{\geq n} = \exists x_1 \dots x_n \bigwedge_{1 \leq i < j \leq n} x_i \neq x_j$.

By hypothesis any finite subset has a model. By Compactness Theorem the entire set has a model, which is infinite. □

Applications of Compactness

- The class of all finite models is not EC_Δ. (Model finiteness is undefinable even by a set of formulas.)
- The class of all infinite models is EC_Δ but not EC. (Model infiniteness is definable by a set of formulas but undefinable by a single formula.)
- The class of graphs / groups / rings / fields / ordered fields / n-dimensional vector spaces / fields of characteristic p is EC; the class of infinite groups / divisible groups / torsion-free groups / infinite rings / infinite-dimensional vector spaces / fields of characteristic 0 is EC_Δ but not EC; the class of all connected graphs / finite graphs / finite groups / finite rings / finite fields / algebraically closed fields / torsion groups / finite-dimensional vector spaces / noetherian commutative rings is not EC_Δ.

From “find-a-word” to Conspiracy Theory ◎ô◎

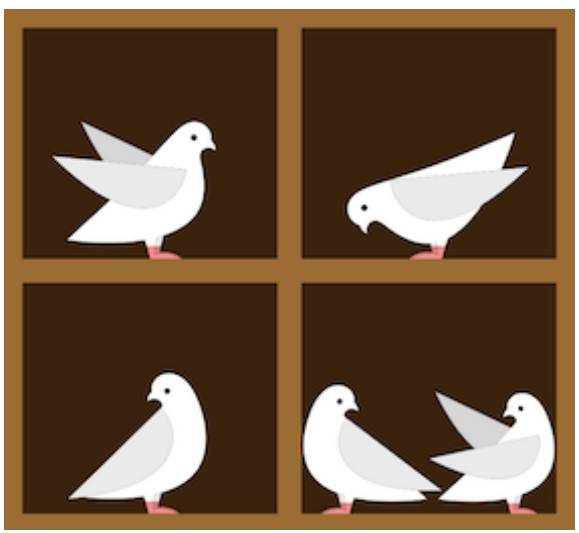
m	b	o
j	d	a
l	l	w

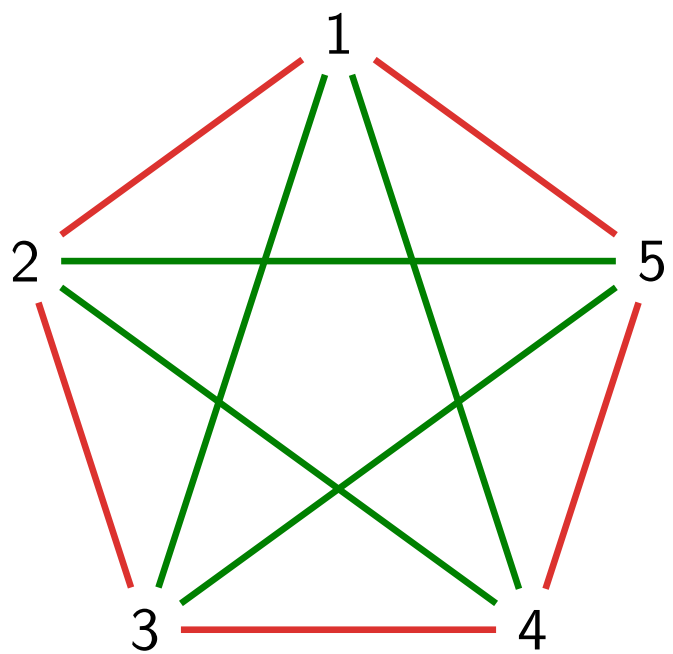
m	b	o	u	n
j	d	a	d	o
l	l	w	d	z
n	n	v	b	e
c	i	l	s	d

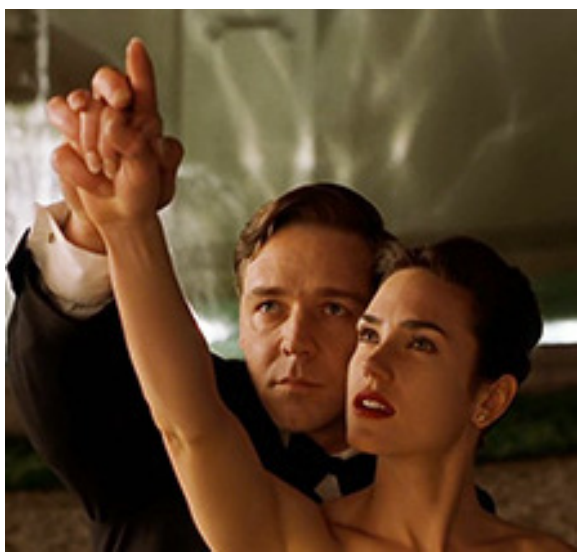
Ramsey in the Dining Room

Problem (Complete Disorder is Impossible!)

- *How many people do you need to invite in a party in order to have that either at least n of them are mutual strangers or at least n of them are mutual acquaintances?*
- *How may we know that such number exists for any n?*







Complete Disorder is Impossible!

Theorem (Infinite Ramsey Theorem)

If (V, E) is a graph with infinitely many vertices, then it has an infinite clique or an infinite independent set.

Infinite Ramsey Theorem $\xrightarrow{\text{Compactness Theorem}}$ Finite Ramsey Theorem

Theorem (Finite Ramsey Theorem)

For every $m, n \geq 1$ there is an integer $R(m, n)$ s.t. any graph with at least $R(m, n)$ vertices has a clique with m vertices or an independent set with n vertices.

$$R(m, n) \leq R(m-1, n) + R(m, n-1)$$

$$R(m, n) \leq \binom{m+n-2}{m-1}$$

Ramsey Number

m, n	1	2	3	4	5	6	7	8	9	10
1	**1**	1	1	1	1	1	1	1	1	1
2	1	**2**	3	4	5	6	7	8	9	10
3	1	3	**6**	9	14	18	23	28	36	40 – 42
4	1	4	9	**18**	25	36 – 41	49 – 61	59 – 84	73 – 115	92 – 149
5	1	5	14	25	**43 – 48**	58 – 87	80 – 143	101 – 216	133 – 316	149 – 442
6	1	6	18	36 – 41	58 – 87	**102 – 165**	115 – 298	134 – 495	183 – 780	204 – 1171
7	1	7	23	49 – 61	80 – 143	115 – 298	**205 – 540**	217 – 1031	252 – 1713	292 – 2826
8	1	8	28	59 – 84	101 – 216	134 – 495	217 – 1031	**282 – 1870**	329 – 3583	343 – 6090
9	1	9	36	73 – 115	133 – 316	183 – 780	252 – 1713	329 – 3583	**565 – 6588**	581 – 12677
10	1	10	40 – 42	92 – 149	149 – 442	204 – 1171	292 – 2826	343 – 6090	581 – 12677	**798 – 23556**



Figure: Ramsey 1903-1930



Figure: Erdős 1913-1996

Ramsey Number — Probabilistic Method

Theorem

$$\forall k \geq 2 : R(k,k) \geq 2^{\frac{k}{2}}$$

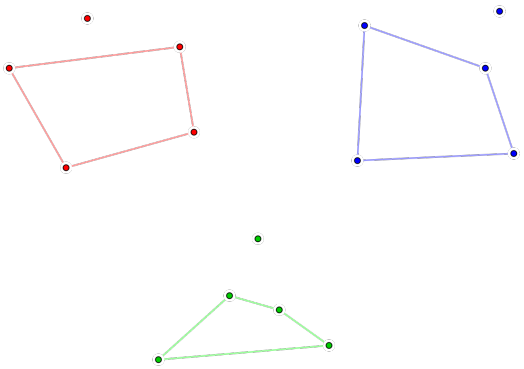
Proof.

$R(2,2) = 2, R(3,3) = 6$. Assume $k \geq 4$. Suppose $N < 2^{\frac{k}{2}}$, and consider all random red-blue colorings. Let A be a set of vertices of size k. The probability of the event A_R that the edges in A are all colored red is then $2^{-\binom{k}{2}}$. Hence the probability p_R for some k-set to be colored all red is bounded by

$$p_R = P\left(\bigcup_{|A|=k} A_R\right) \leq \sum_{|A|=k} P(A_R) = \binom{N}{k} 2^{-\binom{k}{2}} < \frac{1}{2}$$

By symmetry, $p_B < \frac{1}{2}$. So $p_R + p_B < 1$ for $N < 2^{\frac{k}{2}}$. □

Happy Ending Problem



Any set of 5 points in the plane in general position has a subset of 4 points that form the vertices of a convex quadrilateral, where general position means that no two points coincide and no three points are collinear.

Theorem (Erdös & Szekeres 1935)

For $N \in \mathbb{N}$, any sufficiently large finite set of points in the plane in general position has a subset of N points that form the vertices of a convex polygon.

Let $f(N)$ denote the minimum M for which any set of M points in general position must contain a convex N-gon. It is known that

- $f(3) = 3$
- $f(4) = 5$
- $f(5) = 9$
- $f(6) = 17$
- $1 + 2^{N-2} \leq f(N) \leq 2^{N+o(N)}$

Complete Disorder is Impossible!

Theorem (Hales-Jewett Theorem)

For every $k, n \in \mathbb{N}^+$, there is $d \in \mathbb{N}^+$ s.t. if the unit hypercubes in a d-dimensional hypercube n^d are colored in k colors, then there exists at least one row, column or diagonal of n squares, all of the same color.

Theorem (van der Waerden Theorem)

For every $k, m \in \mathbb{N}^+$, there is $n \in \mathbb{N}^+$ s.t. if the numbers from 1 to n are colored in k colors, then there exists at least m numbers in arithmetic progression, all of the same color.

Theorem (Green-Tao Theorem)

A subset of prime numbers A with $\limsup_{n\to\infty} \frac{|A\cap[1,n]|}{\pi(n)} > 0$ contains arbitrarily long arithmetic progressions, where $\pi(n)$ is the number of primes $\leq n$.

Complete Disorder is Impossible!

Theorem (Szemerédi Theorem)

A set $A \subset \mathbb{N}$ *with* $\limsup_{n\to\infty} \frac{|A\cap[1,n]|}{n} > 0$ *contains arbitrarily long arithmetic progressions.*

Theorem (Furstenberg Multiple Recurrence Theorem)

Let $(X, \mathcal{B}, \mu, T)$ *be a measure-preserving system and* $A \in \mathcal{B}$ *with* $\mu(A) > 0$. *Then,*

$$\forall k \in \mathbb{N}: \ \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu\left(\bigcap_{j=0}^{k} T^{-jn} A\right) > 0$$

Szemerédi Theorem $\iff$ Furstenberg Multiple Recurrence Theorem

Complete Disorder is Impossible!

Theorem (Poincaré Recurrence Theorem)

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and $A \in \mathcal{B}$ with $\mu(A) > 0$. Then almost every $x \in A$ returns infinitely often to A.

$$\mu\left(\{x \in A : \exists N \forall n > N : T^n x \notin A\}\right) = 0$$

Lemma (Kac's Lemma)

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and $A \in \mathcal{B}$ with $\mu(A) > 0$. Then the recurrence time $\tau_A(x) := \min\left\{k \geq 1 : T^k x \in A\right\}$ satisfies

$$\int_A \tau_A(x)\,\mathrm{d}\mu(x) = 1$$

Equivalently, the mean recurrence time $\langle \tau_A \rangle := \frac{1}{\mu(A)} \int_A \tau_A(x)\,\mathrm{d}\mu(x) = \frac{1}{\mu(A)}$.

Correlation Supersedes Causation?

- The average recurrence time to a subset A in Poincaré recurrence theorem is the inverse of the probability of A. The probability decrease exponentially with the size (dimension) of the phase space (observables and parameters) and the recurrence time increases exponentially with that size. One can't reliably predict by "analogy" with the past, even in deterministic systems, chaotic or not.
- Given any arbitrary correlation on sets of data, there exists a large enough number such that any data set larger than that size realizes that type of correlation. Every large set of numbers, points or objects necessarily contains a highly regular pattern.
- There is no true randomness. Randomness means unpredictability with respect to some fixed theory.

Correlation Supersedes Causation?

- How to distinguish correlation from causation?
- How to distinguish content-correlations from Ramsey-type correlations?
- Ramsey-type correlations appear in all large enough databases.
- A correlation is *spurious* iff it appears in a "randomly" generated database.
- How "large" is the set of spurious correlations?
- Most strings are algorithmically random.

$$P\left(\left\{x \in \mathcal{X}^n : \frac{K(x)}{n} < 1 - \delta\right\}\right) < 2^{-\delta n}$$

- Most correlations are spurious.
- It may be the case that our part of the universe is an oasis of regularity in a maximally random universe.

Complete Disorder is Impossible!

For sufficiently large n and any $x \in \mathcal{X}^n$, if $C(x) \geq n - \delta(n)$, then each block of length $\log n - \log\log n - \log(\delta(n) + \log n) - O(1)$ occurs at least once in x.

Remark: 混沌 vs 秩序

1. 出自秩序的秩序

$$1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2$$

2. 出自秩序的混沌

$$\pi = 3.14159265358979323384\ldots$$

3. 出自混沌的混沌

$$53278 \times 2147 = 114387866$$

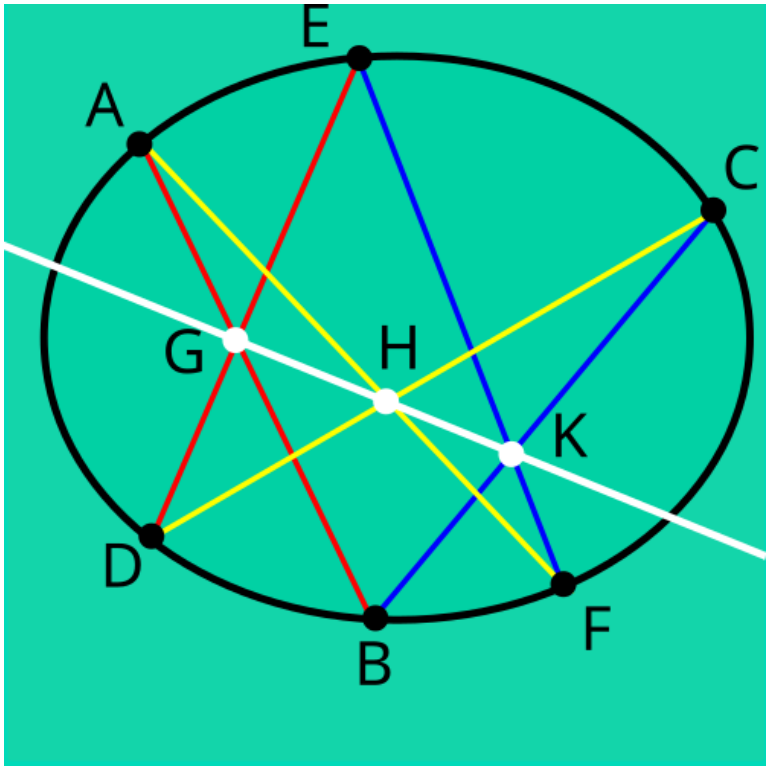
4. 出自混沌的秩序

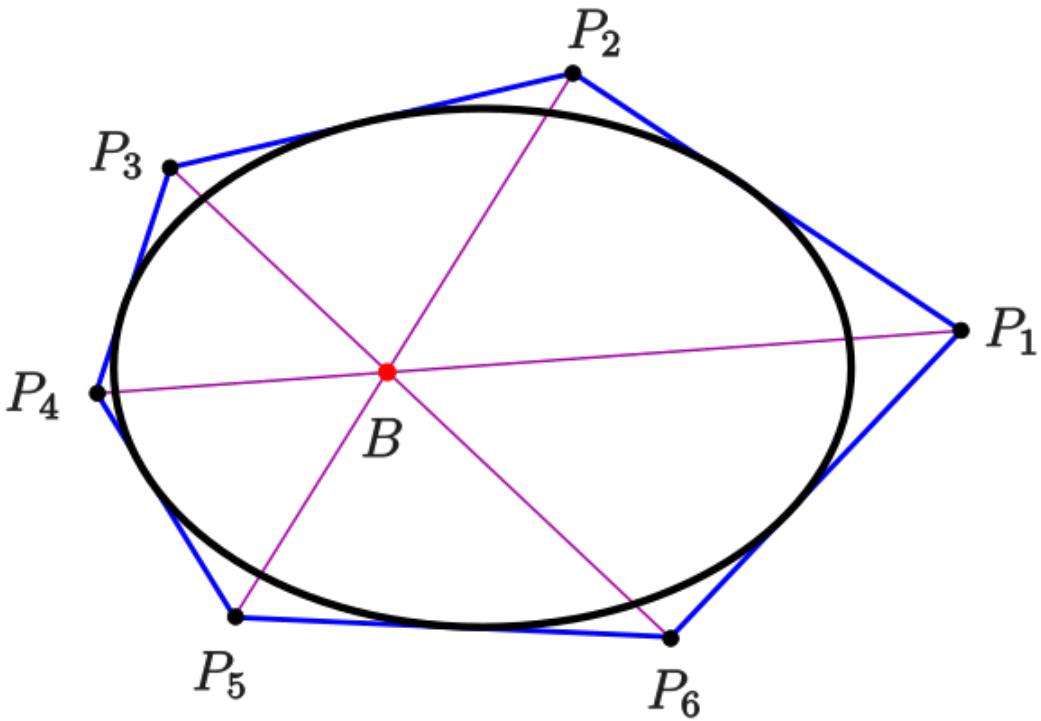
$$\lim_{n\to\infty} \frac{\pi(n)}{\frac{n}{\ln n}} = 1$$

李生素数定理, Ramsey 定理, 帕斯卡定理...

Remark: 最后一种最值得珍视. 数学家和艺术家一样, 都是从混沌中创造秩序. 画家用线条, 作曲家用声音, 诗人用文字, 而数学家用理念.

帕斯卡定理 (又名 “神秘六角星定理”)





- Pascal 定理: 圆锥曲线 (可以是适当的仿射平面中的椭圆、抛物线或双曲线) 的内接六边形的三组对边的交点共线.
- 其对偶是 Brianchon 定理: 圆锥曲线的外切六边形的主对角线共点.

Löwenheim-Skolem Theorem

Theorem (Downward Löwenheim-Skolem Theorem)

A consistent set of sentences in a language of cardinality λ has a model of cardinality $\leq \lambda$.

Theorem (Upward Löwenheim-Skolem Theorem)

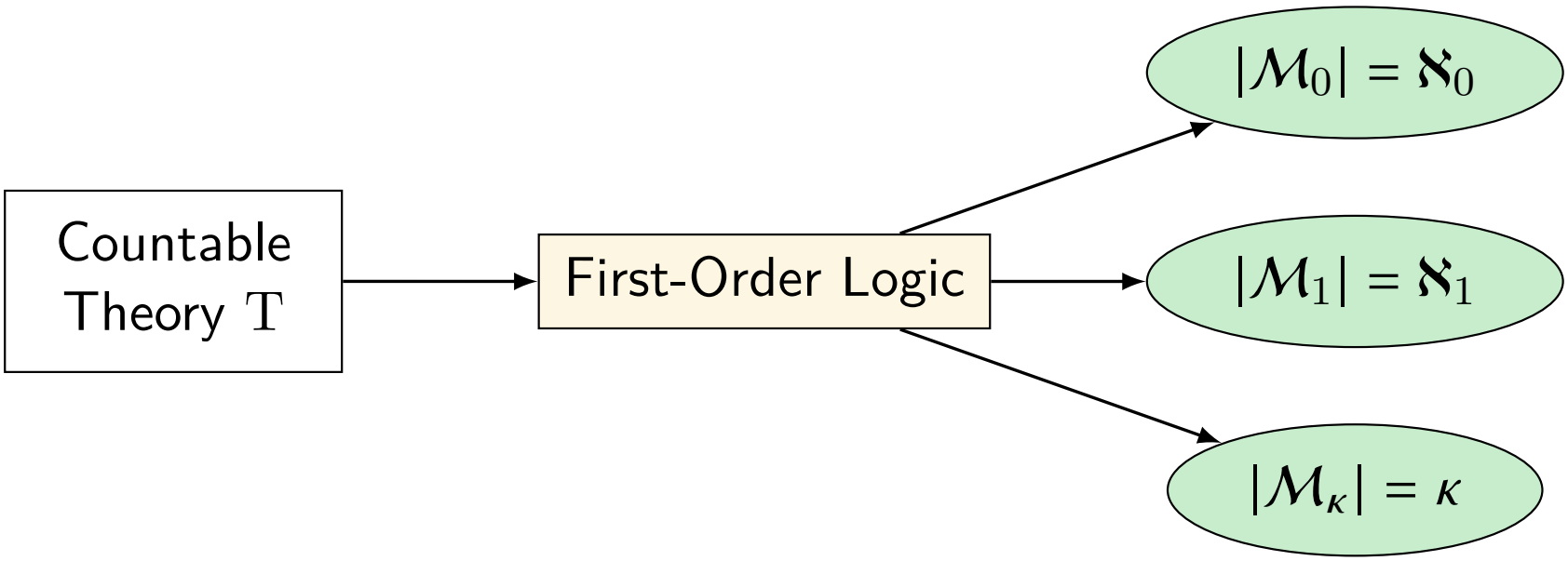
If a set of sentences in a language of cardinality λ has an infinite model, then it has models of every cardinality $\geq \lambda$.

Proof.

Add $\{c_\xi\}_{\xi<\lambda}$ to $\mathscr{L}$. Let $\mathrm{T}' := \mathrm{T} \cup \{c_\xi \neq c_\eta : \xi < \eta < \lambda\}$.
Every finite subset $\Sigma \subset \mathrm{T}'$ will involve at most a finite number of c_ξ.
Hence any infinite model of T can be expanded to a model of Σ.
By compactness and the downward Löwenheim-Skolem theorem, T' has a model $\mathcal{M}$ of cardinality $\leq \lambda$.
On the other hand, the interpretations of c_ξ in $\mathcal{M}$ must give different elements. So $\lambda \leq |\mathcal{M}| \leq \lambda$. □

Upward Löwenheim-Skolem Theorem

Syntactic World **Semantic World**



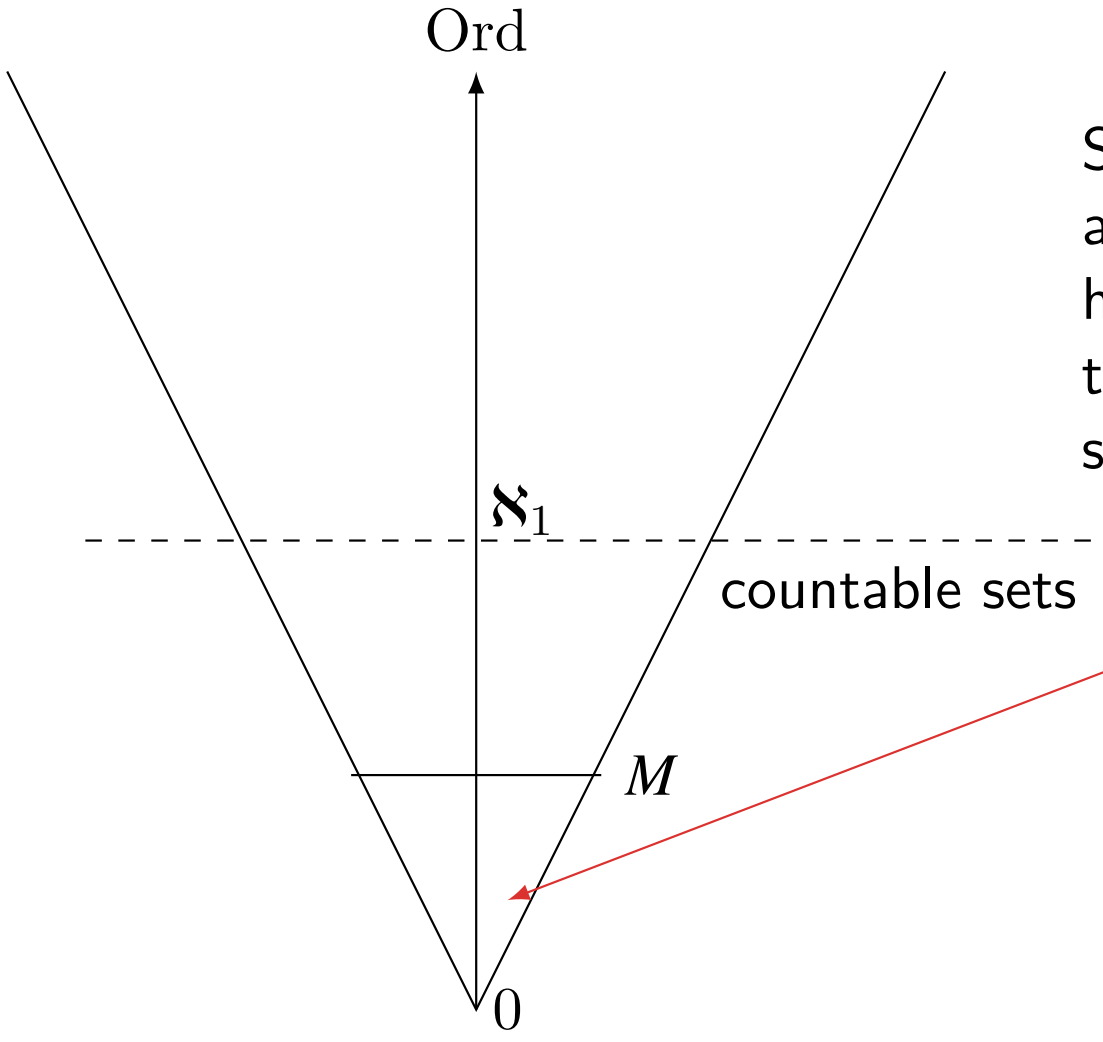
The countable theory T, when viewed through the "lens" of first-order logic, satisfies models of all infinite cardinalities. The logic cannot "see" the difference between $\aleph_0, \aleph_1, \ldots, \kappa$.

- 根据 LST 定理, 有与 $\mathcal{N}$ 初等等价但不同构的算术模型.
- 根据第一不完备性定理, 有与 $\mathcal{N}$ 不初等等价的算术模型.

Skolem Paradox? — Models and Reality

- Cantor: $\mathtt{P}(\mathbb{N})$ is uncountable.
- There is a countable model $\mathcal{M} \vDash \mathrm{ZF} \vdash$"$\mathtt{P}(\mathbb{N})$ is uncountable".
- The statement "$\mathtt{P}(\mathbb{N})$ is uncountable" is interpreted in $\mathcal{M}$ as — within $\mathcal{M}$, there is a set M_1 that looks like $\mathtt{P}(\mathbb{N})$ and M_2 that looks like $\mathbb{N}$, but there is no set corresponding to the set of pairs of members of M_1 and M_2."
- Outside of $\mathcal{M}$, we can see that all $\mathcal{M}$-sets are really only countable. The $\mathcal{M}$-set M_1 that $\mathcal{M}$ says is $\mathtt{P}(\mathbb{N})$ really isn't — outside $\mathcal{M}$, M_1 and $\mathbb{N}$ can be paired, but this requires the existence of a "pairing" set that isn't in $\mathcal{M}$.
- What we think are uncountable sets in our hierarchy may really be countable $\mathcal{M}'$-sets in the larger hierarchy.
- There is no absolute notion of countability. A set can only be said to be countable or uncountable relative to an interpretation of ZF.

Skolem Paradox? — Models and Reality



Skolem paradox: How can an uncountable set way up here in the Hierarchy be thought of as a countable set M way down here?

Craig's Interpolation Theorem

Theorem (Craig's Interpolation Theorem)

If $\vDash A \to B$, *then there is a sentence* C *s.t.* $\vDash A \to C$ *and* $\vDash C \to B$, *and* C *contains no non-logical symbols except such as are both in* A *and in* B.

Beth's Definability Theorem

Definition (Explicit Definition)

Suppose $\mathscr{L}$ is a language not containing the predicate symbol P. A set $\Sigma(P)$ of sentences of $\mathscr{L} \cup \{P\}$ *explicitly defines* P iff there is a formula $A(x_1, \ldots, x_n)$ of $\mathscr{L}$ s.t.

$$\Sigma(P) \vDash \forall x_1 \ldots \forall x_n (P(x_1, \ldots, x_n) \leftrightarrow A(x_1, \ldots, x_n))$$

Definition (Implicit Definition)

Suppose $\mathscr{L}$ is a language not containing the predicate symbol P and P'. A set $\Sigma(P)$ of sentences of $\mathscr{L} \cup \{P\}$ *implicitly defines* P iff

$$\Sigma(P) \cup \Sigma(P') \vDash \forall x_1 \ldots \forall x_n (P(x_1, \ldots, x_n) \leftrightarrow P'(x_1, \ldots, x_n))$$

where $\Sigma(P')$ is the result of uniformly replacing P with P' in $\Sigma(P)$.

Theorem (Beth's Definability Theorem)

$\Sigma(P)$ *implicitly defines* P *iff* $\Sigma(P)$ *explicitly defines* P.

Philosophy question: Is supervenience equivalent to reducibility?

Proof.

Assume that $\Sigma(P)$ implicitly defines P. First, we add constant c_1, ..., c_n to $\mathscr{L}$. Then $\Sigma(P) \cup \Sigma(P') \vDash P(c_1, \ldots, c_n) \to P'(c_1, \ldots, c_n)$. By compactness, there are $A_1(P), \ldots, A_m(P) \in \Sigma(P)$ and $B_1(P'), \ldots, B_n(P') \in \Sigma(P')$ s.t. $A_1(P) \wedge \cdots \wedge A_m(P) \wedge B_1(P') \wedge \cdots \wedge B_n(P') \vDash P(c_1, \ldots, c_n) \to P'(c_1, \ldots, c_n)$.
Let $\varphi(P) \coloneqq A_1(P) \wedge \cdots \wedge A_m(P) \wedge B_1(P) \wedge \cdots \wedge B_n(P)$. Then $\varphi(P) \wedge \varphi(P') \vDash P(c_1, \ldots, c_n) \to P'(c_1 \ldots c_n)$, from which $\varphi(P) \wedge P(c_1, \ldots, c_n) \vDash \varphi(P') \to P'(c_1, \ldots, c_n)$.
By Craig's Interpolation theorem there is a sentence $C(c_1, \ldots, c_n)$ not containing P or P' such that:

$$\varphi(P) \wedge P(c_1, \ldots, c_n) \vDash C(c_1, \ldots, c_n)$$
$$C(c_1, \ldots, c_n) \vDash \varphi(P') \to P'(c_1, \ldots, c_n)$$

Since an $\mathscr{L} \cup \{P\}$-model $\mathcal{M} \vDash A(P)$ iff the corresponding $\mathscr{L} \cup \{P'\}$-model $\mathcal{M} \vDash A(P')$, we have $C(c_1, \ldots, c_n) \vDash \varphi(P) \to P(c_1, \ldots, c_n)$.
Putting them together, $\varphi(P) \vDash P(c_1, \ldots, c_n) \leftrightarrow C(c_1, \ldots, c_n)$.
Therefore $\Sigma(P) \vDash \forall x_1 \ldots \forall x_n \big(P(x_1, \ldots, x_n) \leftrightarrow C(x_1, \ldots, x_n)\big)$. □

Robinson's Joint Consistency Theorem

Theorem (Robinson's Joint Consistency Theorem)

Let $\mathscr{L}_0, \mathscr{L}_1, \mathscr{L}_2$ be languages, with $\mathscr{L}_0 = \mathscr{L}_1 \cap \mathscr{L}_2$. Let T_i be a theory in $\mathscr{L}_i$ for $i = 1, 2$. If T_1 and T_2 are consistent and if there is no formula A of $\mathscr{L}_0$ s.t. $\mathrm{T}_1 \vdash A$ and $\mathrm{T}_2 \vdash \neg A$, then the union $\mathrm{T}_1 \cup \mathrm{T}_2$ is consistent.

Proof.

Suppose $\mathrm{T}_1 \cup \mathrm{T}_2$ is inconsistent.
Then there exists finite subsets $\Sigma_1 \subset \mathrm{T}_1$ and $\Sigma_2 \subset \mathrm{T}_2$ such that $\Sigma_1 \cup \Sigma_2$ is inconsistent.
Let $A := \bigwedge \Sigma_1$ and $B := \bigwedge \Sigma_2$.
It follows that $A \vDash \neg B$.
By Craig's interpolation theorem, there is a sentence C of $\mathscr{L}_0$ such that $A \vDash C$ and $C \vDash \neg B$.
Then we have $\mathrm{T}_1 \vdash C$ and $\mathrm{T}_2 \vdash \neg C$. Contradiction. □

Abstract Logics

Definition (Abstract Logic)

An *abstract logic* is a pair $\mathcal{L} := (\mathcal{S}, \vDash_{\mathcal{L}})$, where $\mathcal{S}$: signatures $\to$ sets assigns to signature τ a set $\mathcal{S}(\tau)$ of sentences, and $\vDash_{\mathcal{L}}$ is a relation between τ-models and elements of $\mathcal{S}(\tau)$ s.t.

1. (*Monotony*) $\tau \subset \tau' \implies \mathcal{S}(\tau) \subset \mathcal{S}(\tau')$
2. (*Isomorphism*) $\mathcal{M} \vDash_{\mathcal{L}} A \;\&\; \mathcal{M} \cong \mathcal{N} \implies \mathcal{N} \vDash_{\mathcal{L}} A$
3. (*Expansion*) If $\tau \subset \tau', A \in \mathcal{S}(\tau)$, and $\mathcal{M}$ is an τ'-model, then $\mathcal{M} \vDash_{\mathcal{L}} A \iff \mathcal{M}{\restriction_\tau} \vDash_{\mathcal{L}} A$

$$\mathrm{Mod}^{\tau}_{\mathcal{L}}(A) := \left\{\mathcal{M} \in \tau\text{-models} : \mathcal{M} \vDash_{\mathcal{L}} A\right\}$$

Example — $\mathcal{L}_{\kappa\lambda}$

For $\kappa \geq \lambda$, define the $\mathcal{L}_{\kappa\lambda}$ formulas as for first order logic, plus:

- Given a set of formulas $\{A_i : i \in I\}, |I| < \kappa$, then $\bigwedge_{i \in I} A_i$ and $\bigvee_{i \in I} A_i$ are formulas.
- Given a set of variables $\{x_i : i \in J\}, |J| < \lambda$ and a formula A, then $\exists (x_i : i \in J) A$ and $\forall (x_i : i \in J) A$ are formulas.

Satisfaction relation:

- $\mathcal{M} \vDash_{\mathcal{L}_{\kappa\lambda}} \bigwedge_{i \in I} A_i$ iff $\mathcal{M} \vDash_{\mathcal{L}_{\kappa\lambda}} A_i$ for all $i \in I$.
- $\mathcal{M} \vDash_{\mathcal{L}_{\kappa\lambda}} \exists (x_i : i \in J) A$ iff $\mathcal{M} \vDash_{\mathcal{L}_{\kappa\lambda}} A[a_i : i \in J]$ for some $\{a_i : i \in J\} \subset M$.

Note: $\mathcal{L}_{\omega\omega}$ is classical first order logic.

Definition (Regular Abstract Logic)

An abstract logic $\mathcal{L}$ is *regular* iff it satisfies:

1. (*Bool*) For $A \in \mathcal{S}(\tau)$ there is $B \in \mathcal{S}(\tau)$ s.t. $\mathcal{M} \models_{\mathcal{L}} B \iff \mathcal{M} \not\models_{\mathcal{L}} A$; and $\forall A, B \in \mathcal{S}(\tau)\ \exists C \in \mathcal{S}(\tau) : \mathcal{M} \models_{\mathcal{L}} C \iff \mathcal{M} \models_{\mathcal{L}} A \ \&\ \mathcal{M} \models_{\mathcal{L}} B$.
2. (*Quantifier*) $\forall c \in \tau \forall A \in \mathcal{S}(\tau)\ \exists B \in \mathcal{S}(\tau)$:

$$\mathrm{Mod}_{\mathcal{L}}^{\tau\setminus\{c\}}(B) = \left\{\mathcal{M} : (\mathcal{M}, a) \in \mathrm{Mod}_{\mathcal{L}}^{\tau}(A) \text{ for some } \ a \in M\right\}$$

   where $(\mathcal{M}, a)$ is the expansion of $\mathcal{M}$ to τ assigning a to c.
3. (*Renaming*) Let $\pi : \tau \to \tau'$ be a bijection which respects arity, and we extend π in a canonical way to $\hat{\pi}$: τ-models $\to \tau'$-models. Then

$$\forall A \in \mathcal{S}(\tau)\ \exists A' \in \mathcal{S}(\tau) : \mathcal{M} \models_{\mathcal{L}} A \iff \hat{\pi}(\mathcal{M}) \models_{\mathcal{L}} A'$$

4. (*Relativization*) Given $A \in \mathcal{S}(\tau)$ and symbols R, $c_1, \ldots, c_n \notin \tau$, there is $B \in \mathcal{S}(\tau \cup \{R, c_1, \ldots, c_n\})$ called the *relativization* of A to $R(x, c_1, \ldots c_n)$, s.t. for $\mathcal{M} : (\mathcal{M}, X, b_1, \ldots, b_n) \models_{\mathcal{L}} B \iff \mathcal{N} \models_{\mathcal{L}} A$ where $\mathcal{N} \subset \mathcal{M}$ with $N = \left\{a \in M : R^{\mathcal{M}}(a, b_1, \ldots, b_n)\right\}$, and $(\mathcal{M}, X, b_1, \ldots, b_n)$ is the expansion of $\mathcal{M}$ interpreting R, $c_1, \ldots, c_n$ by X, $b_1, , \ldots, b_n$ (with $X \subset M^{n+1}$).

Lindström's Theorem

Definition (Expressive Power)

$\mathcal{L}_2$ is *at least as expressive* as $\mathcal{L}_1$ ($\mathcal{L}_1 \leq \mathcal{L}_2$) iff for each signature τ and $A \in \mathcal{S}_1(\tau)$ there is $B \in \mathcal{S}_2(\tau)$ s.t.

$$\mathrm{Mod}^{\tau}_{\mathcal{L}_1}(A) = \mathrm{Mod}^{\tau}_{\mathcal{L}_2}(B)$$

$$\mathcal{L}_1 \sim \mathcal{L}_2 := \mathcal{L}_1 \leq \mathcal{L}_2 \ \& \ \mathcal{L}_2 \leq \mathcal{L}_1$$

Theorem (Lindström's Theorem)

If a regular abstract logic $\mathcal{L}$ has the Countable Compactness and the Downward Löwenheim-Skolem Properties, then $\mathcal{L} \sim \mathcal{L}_{\omega\omega}$.

1. The set of *Horn formulas* is the smallest set containing the set of atomic formulas and closed under $\top, \wedge$.
2. The set of *regular formulas* is the smallest set containing the set of atomic formulas and closed under $\top, \wedge, \exists$.
3. The set of *coherent formulas* is the smallest set containing the set of atomic formulas and closed under $\top, \bot, \wedge, \vee, \exists$.
4. The set of *first order formulas* is the smallest set containing the set of atomic formulas and closed under $\top, \bot, \neg, \wedge, \vee, \longrightarrow, \exists, \forall$.
5. The class of *geometric formulas* over is the smallest class containing the class of atomic formulas and closed under $\top, \bot, \wedge, \vee, \exists$ and infinitary disjunction.
6. The class of *infinitary first order formulas* is the smallest class containing the class of atomic formulas and closed under $\top, \bot, \neg, \wedge, \vee, \longrightarrow, \exists, \forall$ and infinitary conjunction and infinitary disjunction.

- $\mathbb{T}$ is an algebraic theory iff its signature has no relation symbols and its axioms are all of the form $\top \vdash_{\mathbf{x}} A$ where A is an atomic formula of the form $s = t$ and $\mathbf{x}$ its canonical context.
- $\mathbb{T}$ is a Horn (resp. regular, coherent, geometric) theory iff all the sequents in $\mathbb{T}$ are Horn (resp. regular, coherent, geometric).
- $\mathbb{T}$ is a universal Horn theory iff its axioms are all of the form $A \vdash_{\mathbf{x}} B$, where A is a finite conjunction of atomic formulas and B is an atomic formula or the formula $\bot$.
- $\mathbb{T}$ is a propositional theory iff it only consists of 0-ary relation symbols.

Identity Axiom $A \vdash_{\mathbf{x}} A$

Equality $\top \vdash_x x = x$ and $\mathbf{x} = \mathbf{y} \wedge A \vdash_{\mathbf{z}} A[\mathbf{y}/\mathbf{x}]$ where $\mathrm{Fv}(\mathbf{x}, \mathbf{y}, A) \subset \mathbf{z}$.

Substitution $\dfrac{A \vdash_{\mathbf{x}} B}{A[\mathbf{t}/\mathbf{x}] \vdash_{\mathbf{y}} B[\mathbf{t}/\mathbf{x}]}$ where $\mathrm{Var}(\mathbf{t}) \subset \mathbf{y}$.

Cut $\dfrac{A \vdash_{\mathbf{x}} B \quad B \vdash_{\mathbf{x}} C}{A \vdash_{\mathbf{x}} C}$

Conjunction $A \vdash_{\mathbf{x}} \top$ $\quad A \wedge B \vdash_{\mathbf{x}} A$ $\quad A \wedge B \vdash_{\mathbf{x}} B$ $\quad \dfrac{C \vdash_{\mathbf{x}} A \quad C \vdash_{\mathbf{x}} B}{C \vdash_{\mathbf{x}} A \wedge B}$

Disjunction $\bot \vdash_{\mathbf{x}} A$ $\quad A \vdash_{\mathbf{x}} A \vee B$ $\quad B \vdash_{\mathbf{x}} A \vee B$ $\quad \dfrac{A \vdash_{\mathbf{x}} C \quad B \vdash_{\mathbf{x}} C}{A \vee B \vdash_{\mathbf{x}} C}$

Implication $\dfrac{A \wedge B \vdash_{\mathbf{x}} C}{A \vdash_{\mathbf{x}} B \to C}$

Existential Quantification $\dfrac{A \vdash_{\mathbf{x}y} B}{\exists y A \vdash_{\mathbf{x}} B}$

Universal Quantification $\dfrac{A \vdash_{\mathbf{x}y} B}{A \vdash_{\mathbf{x}} \forall y B}$

Distributive Axiom $A \wedge (B \vee C) \vdash_{\mathbf{x}} (A \wedge B) \vee (A \wedge C)$

Frobenius Axiom $A \wedge \exists y B \vdash_{\mathbf{x}} \exists y (A \wedge B)$ where $y \notin \mathbf{x}$.

Law of Excluded Middle $\top \vdash_{\mathbf{x}} A \vee \neg A$

Fragments of First Order Logic

In addition to the usual structural rules (Identity axiom, Equality rules, Substitution rule and Cut rule), our deduction systems consist of the following rules:

Algebraic logic	No additional rule
Horn logic	Finite conjunction
Regular logic	Finite conjunction, existential quantification and Frobenius axiom
Coherent logic	Finite conjunction, finite disjunction, existential quantification, distributive axiom and Frobenius axiom
Geometric logic	Finite conjunction, infinitary disjunction, existential quantification, 'infinitary' distributive axiom, Frobenius axiom
Intuitionistic FOL	All the finitary rules except for the law of excluded middle
Classical FOL	All the finitary rules

Intuitionistic Propositional Logic vs Heyting Algebra

A Heyting algebra $(H, \bot, \top, \wedge, \vee, \rightarrow, \leq)$ is a bounded lattice $(H, \bot, \top, \wedge, \vee, \leq)$ equipped with $\rightarrow$ s.t. for all $a, b, c \in H$:

1. $a \leq \top$
2. $a \wedge b \leq a$
3. $a \wedge b \leq b$
4. $c \leq a \ \ \& \ \ c \leq b \implies c \leq a \wedge b$
5. $\bot \leq a$
6. $a \leq a \vee b$
7. $b \leq a \vee b$
8. $a \leq c \ \ \& \ \ b \leq c \implies a \vee b \leq c$
9. $a \leq b \rightarrow c \iff a \wedge b \leq c$

Define $\neg a := a \rightarrow \bot$.

Example — Heyting Algebra

$$\left(\left\{0, \frac{1}{2}, 1\right\}, 0, 1, \wedge, \vee, \rightarrow, \leq\right)$$

$\wedge$	0	$\frac{1}{2}$	1
0	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	1

$\vee$	0	$\frac{1}{2}$	1
0	0	$\frac{1}{2}$	1
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
1	1	1	1

$\rightarrow$	0	$\frac{1}{2}$	1
0	1	1	1
$\frac{1}{2}$	0	1	1
1	0	$\frac{1}{2}$	1

a	$\neg a$
0	1
$\frac{1}{2}$	0
1	0

$$\frac{\text{Intuitionistic}}{\text{Classical}} \approx \frac{\text{Logic of Knowledge}}{\text{Logic of Being}}$$

Intuitionistic truth of A: knowing how to **prove/solve** A

- There is no proof/solution of $\bot$.
- A proof of $A \wedge B$ is a pair $\langle m, n \rangle$ where m is a proof of A and n is a proof of B.
- A proof of $A \vee B$ is a pair $\langle m, n \rangle$ where m is 0 and n is a proof of A, or m is 1 and n is a proof of B.
- A proof of $A \to B$ is a function f that converts a proof m of A into a proof $f(m)$ of B.
- A proof of $\forall x A(x)$ is a function f that converts an element a of the domain of definition into a proof $f(a)$ of $A(a)$.
- A proof of $\exists x A(x)$ is a pair $\langle a, n \rangle$ where a is an element of the domain of definition, and n is a proof of $A(a)$.

Logic of Problems by Medvedev

- Medvedev formalized Kolmogorov's interpretation in terms of problems and solutions.
- A problem is a pair of finite sets (X, Y) s.t. $Y \subset X$ and $X \neq \emptyset$.
- X and Y are called respectively possible solutions and actual solutions.
- Operations between problems are defined as

$$(X_1, Y_1) \wedge (X_2, Y_2) := (X_1 \times X_2, Y_1 \times Y_2)$$

$$(X_1, Y_1) \vee (X_2, Y_2) := \left(X_1 \coprod X_2, Y_1 \coprod Y_2\right)$$

$$(X_1, Y_1) \to (X_2, Y_2) := \left(X_2^{X_1}, \left\{f \in X_2^{X_1} : f(Y_1) \subset Y_2\right\}\right)$$

- An interpretation is a map j : atomic propositions $\to$ problems.
  e.g. $j(p) = (X, Y)$, and we denote $j_1(p) = X, j_2(p) = Y$.
  Specifically, $j(\bot) := (\{\emptyset\}, \emptyset)$.
- Two interpretations j, j' are equivalent $j \sim j'$ iff $j_1(p) = j'_1(p)$ for any atomic p.
- A formula A is true under j iff $\bigcap_{j' \sim j} j'_2(A) \neq \emptyset$ iff $\exists x \forall j' \sim j : x \in j'_2(A)$.
- A formula A is Medvedev valid iff it is true under any interpretation.

Kleene's Realizability Interpretation

- An *assembly* $\mathbf{M} = (M, \vDash_{\mathbf{M}})$ is a set M with a *realizability* relation $\vDash_{\mathbf{M}} \subset \mathbb{N} \times M$ such that $\forall x \in M \exists n \in \mathbb{N} : n \vDash_{\mathbf{M}} x$.
- An *assembly morphism* $f : \mathbf{X} \to \mathbf{Y}$ is a function $f : X \to Y$ for which there exists $n \in \mathbb{N}$ such that

$$\forall x \in X \forall m \in \mathbb{N} : m \vDash_{\mathbf{X}} x \implies \varphi_n(m) \downarrow \ \& \ \varphi_n(m) \vDash_{\mathbf{Y}} f(x)$$

$n \vDash \top$	for every $n \in \mathbb{N}$
$n \vDash \bot$	for no $n \in \mathbb{N}$
$n \vDash s = t$	if $s^{\mathcal{N}} = t^{\mathcal{N}}$
$\langle m, n \rangle \vDash A \wedge B$	if $m \vDash A$ and $n \vDash B$
$\langle m, n \rangle \vDash A \vee B$	if $m = 0$ and $n \vDash A$, or $m = 1$ and $n \vDash B$
$n \vDash A \to B$	if $m \vDash A$ implies $\varphi_n(m) \vDash B$
$n \vDash \forall x A$	if $a \in M$ and $m \vDash a$ implies $\varphi_n(m) \vDash A(a)$
$\langle m, n \rangle \vDash \exists x A$	if there is $a \in M$ such that $m \vDash a$ and $n \vDash A(a)$

Beth-Kripke Semantics for Intuitionistic Logic

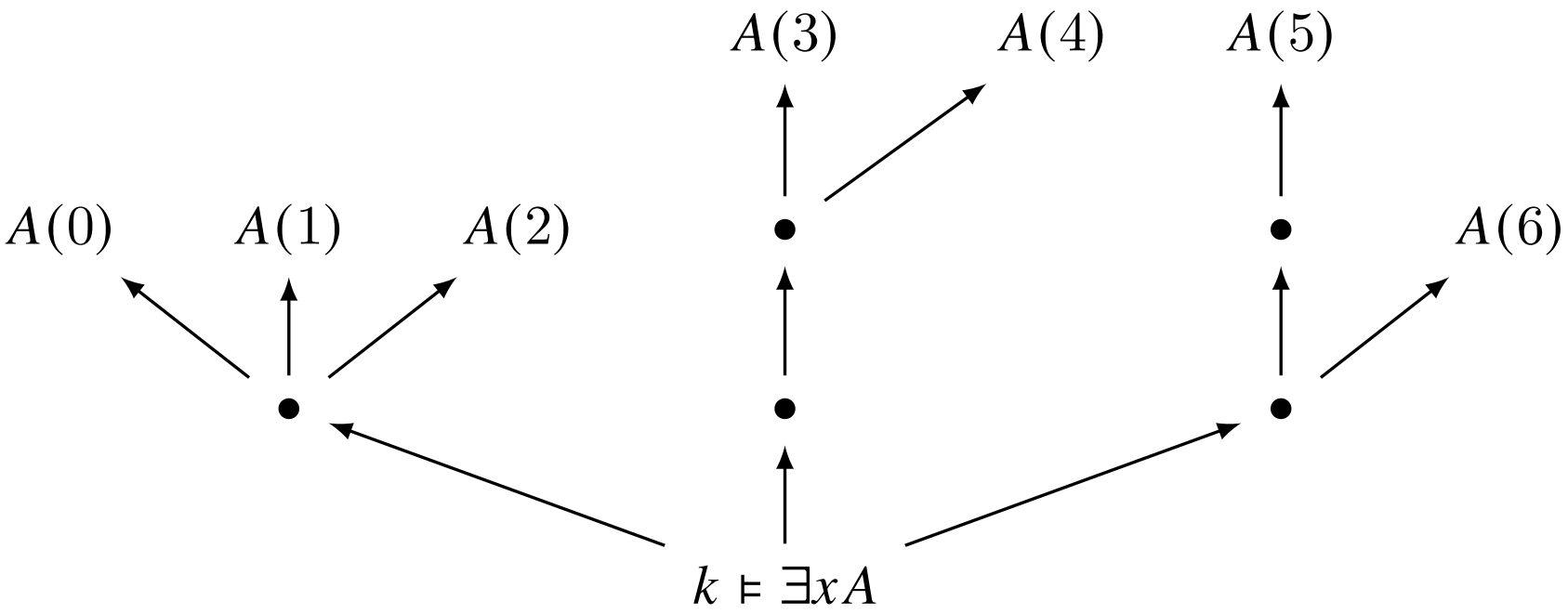
Definition (Beth-Kripke Model)

A Beth-Kripke model $\mathcal{M} = (\mathcal{I}, M, V, \vDash)$ consists of a partially ordered set $\mathcal{I}$, a domain function M assigning to $k \in \mathcal{I}$ an inhabited set M_k s.t. $k \leq l \implies M_k \subset M_l$, and an assignment function V assigning to $k \in \mathcal{I}$ a set $V(k)$ of atomic sentences s.t. $k \leq l \implies V(k) \subset V(l)$.

- $k \vDash A$ for atomic A iff every branch passing through k contains a node $l \geq k$ s.t. $A \in V(l)$
- $k \vDash A \wedge B$ iff $k \vDash A$ and $k \vDash B$
- $k \vDash A \vee B$ iff every branch passing through k contains a node $l \geq k$ s.t. $l \vDash A$ or $l \vDash B$
- $k \vDash A \to B$ iff for all $l \geq k$, if $l \vDash A$ then $l \vDash B$
- $k \vDash \neg A$ iff for all $l \geq k$, $l \nvDash A$
- $k \vDash \forall x A$ iff for all $l \geq k$ and all $a \in M_l$, $l \vDash A(a)$
- $k \vDash \exists x A$ iff every branch passing through k contains a node $l \geq k$ s.t. $l \vDash A(a)$ for some $a \in M_l$.

Remark: $k \vDash A$: at k, it is known that A will be proved.

Example



Theorem (Soundness and Completeness)

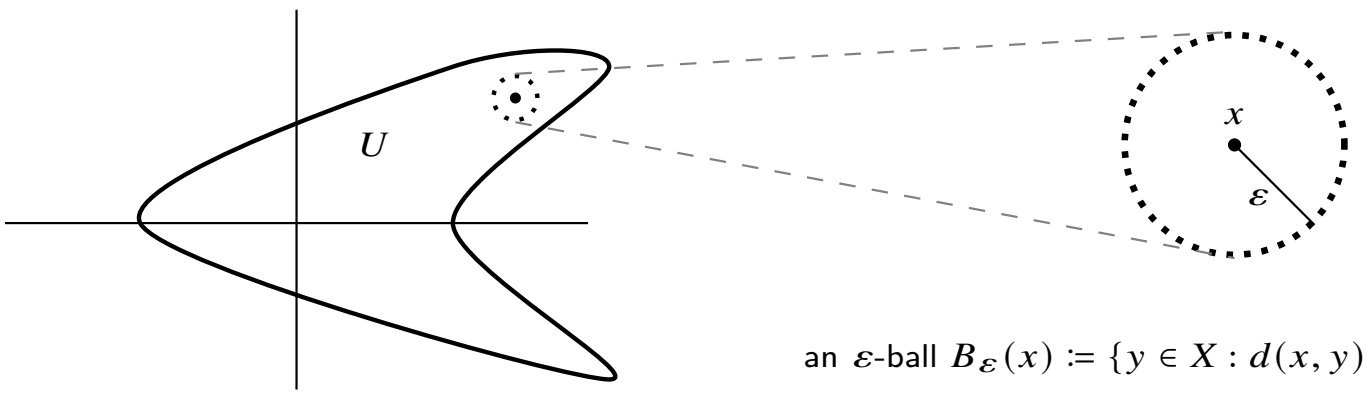
$$\Gamma \vdash A \iff \Gamma \Vdash A$$

Continuity, Metric and Topology

- A function $f : X \to Y$ between metric spaces is *continuous* at point $a \in X$ iff $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in X : d_X(x, a) < \delta \to d_Y(f(x), f(a)) < \varepsilon$.
- A *metric space* (X, d) is a set X with a metric $d : X \times X \to \mathbb{R}$ s.t. for all $x, y, z \in X$:
  1. $d(x, y) = 0 \leftrightarrow x = y$
  2. $d(x, y) = d(y, x)$
  3. $d(x, z) \leq d(x, y) + d(y, z)$

$$e.g. \qquad d(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}} \qquad d(x, y) = \max_{1 \leq i \leq n} |x_i - y_i|$$

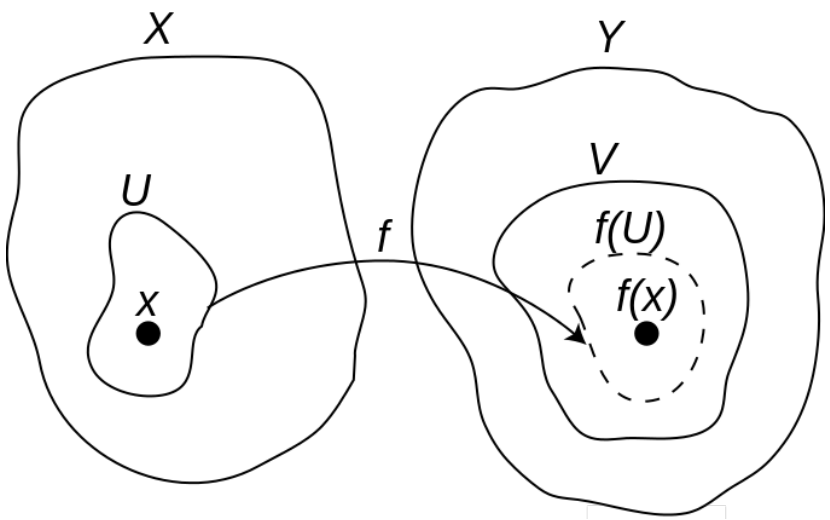
- A set $U \subset X$ is *open* iff $\forall x \in U \exists \varepsilon > 0 : B_\varepsilon(x) \subset U$.



an ε-ball $B_\varepsilon(x) := \{y \in X : d(x, y) < \varepsilon\}$ centered at x

Continuity, Metric and Topology

- For the definition of continuity ("nearby" points in U go into nearby points in V), the notion of 'open set' is more **intrinsic** than that of distance.
- A function $f : X \longrightarrow Y$ between topological spaces is *continuous*, iff, the inverse image $f^{-1}(V)$ of any open subset $V \subset Y$ is an open subset of X, iff, $f^{-1}(B^\circ) \subset \left(f^{-1}(B)\right)^\circ$ for any $B \subset Y$, iff, $f(\mathrm{cl}(A)) \subset \mathrm{cl}(f(A))$ for any $A \subset X$.
- Equivalently, f is *continuous* at point $x \in X$, iff, to each neighborhood V of $f(x)$ there is a neighborhood U of x for which $f(U) \subset V$.



$$\begin{array}{ccc} U & \hookrightarrow & f^{-1}(V) \\ {\scriptstyle f}\downarrow & & \uparrow{\scriptstyle f^{-1}} \\ f(U) & \hookrightarrow & V \end{array} \qquad \begin{array}{ccc} \mathcal{O}_X & \hookrightarrow & \mathrm{P}(X) \\ \uparrow & & \uparrow{\scriptstyle f^{-1}} \\ \mathcal{O}_Y & \hookrightarrow & \mathrm{P}(Y) \end{array}$$

Intrinsic vs. Extrinsic Properties of Functions

A property of functions $f : A \to B$ is *intrinsic* if it can be defined purely in terms of f (as a set of input-output pairs) and any structure pertaining to A and B (using only bounded quantification over the structures A and B.).

Example:

- G and H are groups. $f : G \to H$ is a group homomorphism if

$$f(x \cdot_G y) = f(x) \cdot_H f(y) \quad f(e_G) = e_H$$

- X and Y are topological spaces. $f : X \to Y$ is continuous if $f^{-1}(V)$ is open in the topology on X for every open subset V of Y.
- Computability is not an intrinsic property of partial functions $f : \mathbb{N} \rightharpoonup \mathbb{N}$. To say f is computable we need to refer to some external algorithmic process which computes f.

Topological Semantics of Intuitionistic Propositional Logic

- A *topological space* $(X, \mathcal{O}_X)$ is a set X with a family $\mathcal{O}_X \subset \mathrm{P}(X)$ of subsets of X which contains $\emptyset$ and X, and is closed under finite intersections and arbitrary unions.
- The topological *interior* of a subset $S \subset X$ is

$$S^\circ := \bigcup \{U \in \mathcal{O}_X : U \subset S\}$$

- Define for $A, B \in \mathcal{O}_X$ the open set

$$A \to B := \left(\overline{A} \cup B\right)^\circ = \bigcup \{U \in \mathcal{O}_X : U \cap A \subset B\}$$

by definition, for all $U \in \mathcal{O}_X$,

$$U \subset A \to B \iff U \cap A \subset B$$

Define $\neg A := A \to \bot$. Thus

$$\neg A = \left(\overline{A}\right)^\circ \quad \text{and} \quad A \vee \neg A = A \cup \left(\overline{A}\right)^\circ = \overline{\partial A}$$

- $A := (0,1) \cup (1,\infty), \neg A = (-\infty, 0), \neg\neg A = (0,\infty), A \cup \neg A \neq \mathbb{R}, A \subsetneq \neg\neg A$.
- A topological model of intuitionistic propositional logic is $(X, \mathcal{O}_X, [\![\,]\!])$ where $[\![\,]\!] : \mathrm{Var} \to \mathcal{O}_X$.

Topological Semantics of Intuitionistic Logic

$$[\![A \wedge B]\!] := [\![A]\!] \cap [\![B]\!]$$
$$[\![A \vee B]\!] := [\![A]\!] \cup [\![B]\!]$$
$$[\![A \to B]\!] := \left(\overline{[\![A]\!]} \cup [\![B]\!]\right)^\circ$$
$$[\![\bot]\!] := \emptyset$$
$$[\![\neg A]\!] := \left(\overline{[\![A]\!]}\right)^\circ$$
$$[\![\exists x A]\!] := \bigcup_{a \in M} [\![A[a]]\!]$$
$$[\![\forall x A]\!] := \left(\bigcap_{a \in M} [\![A[a]]\!]\right)^\circ$$

$$\mathcal{O}_X \vDash A := [\![A]\!] = X$$
$$\Gamma \vDash A := \left(\bigcap_{B \in \Gamma} [\![B]\!]\right)^\circ \subset [\![A]\!]$$

Remark: Classical logic appears as a special case when $\mathcal{O}_X = \mathrm{P}(X)$.

Remark: The algebra of open subsets of a topological space is a special case of a Heyting algebra.

Formal Systems

Axiom Schema

1. $A \to B \to A$
2. $(A \to B \to C) \to (A \to B) \to A \to C$
3. $A \wedge B \to A$
4. $A \wedge B \to B$
5. $(C \to A) \to (C \to B) \to C \to A \wedge B$
6. $A \to A \vee B$
7. $B \to A \vee B$
8. $(A \to C) \to (B \to C) \to A \vee B \to C$
9. $(A \to \neg B) \to (A \to B) \to \neg A$
10. $\neg A \to A \to B$
11. $\neg\neg A \to A$

Reference Rule

$$\frac{A \quad A \to B}{B}\ \text{MP}$$

- 1–8+MP=**P**ositive Calculus
- **P**+9=**M**inimal Calculus
- **M**+10=**I**ntuitionistic Calculus
- **I**+11=**C**lassical Calculus

Gödel's ¬¬ Translation

$$
\begin{aligned}
p' &:= \neg\neg p \\
(A \wedge B)' &:= A' \wedge B' \\
(A \vee B)' &:= \neg(\neg A' \wedge \neg B') \\
(A \to B)' &:= A' \to B' \\
(\forall x A)' &:= \forall x A' \\
(\exists x A)' &:= \neg\forall x \neg A'
\end{aligned}
$$

- $\Gamma \vdash_{\mathrm{C}} A \iff \Gamma' \vdash_{\mathrm{I}} A'$ where $\Gamma' := \{A' : A \in \Gamma\}$.
- $\vdash_{\mathrm{C}} A \iff \vdash_{\mathrm{I}} \neg\neg A$

Remark:

- Each classical theory can be translated into intuitionistic logic, yielding a classically equivalent result.
- Intuitionistic logic is more expressive than classical logic since it allows to distinguish formulas which are classically equivalent.

Intuitionistic Logic vs Classical Logic

- The difference between intuitionistic and classical logic is in the criteria for truth, i.e. what evidence must be provided before a statement is accepted as true. Speaking vaguely, intuitionistic logic demands positive evidence, while classical logic is happy with lack of negative evidence.
- $\neg\neg A$ holds if it is contradictory to assume that it is contradictory to assume A. In essence, it says that $\neg\neg A$ is accepted when there is no evidence against it. In other words, $\neg\neg A$ means something like "A cannot be falsified" or "A is potentially true".

- 宇宙之外一无所有.
  — 宇宙是因果封闭的.
  — 适用于整个宇宙的理论是背景独立的.
  — 宇宙不是由物质构成, 而是由事件的因果过程构成. 宇宙是一个因果关系的网络. 其中每一部分的性质由其与其它部分的关系决定.
  — 空间是事件之间关系网络的近似描述.
  — 时间是事件关系网络变化的度量.
  — 观察者是宇宙的一部分.
  — 观察者只能观测到宇宙的部分信息. 未来的观察者可以看到更多信息.
  — 逻辑与观察者有关. 不能假定每个陈述非真即假. 至少有三类: 一类可判断为真, 一类可判断为假, 一类无法判断.

Theorem

The axiom of choice implies excluded middle.

Proof.

Consider an arbitrary proposition P. To decide P, define

$$A \coloneqq \{x \in \{0,1\} : P \vee x = 0\}$$
$$B \coloneqq \{x \in \{0,1\} : P \vee x = 1\}$$

Every member of $\{A, B\}$ is inhabited because $0 \in A$ and $1 \in B$. By the axiom of choice there is a function $f : \{A, B\} \to A \cup B$ s.t. $f(A) \in A$ and $f(B) \in B$. Then

1. $f(A) = 1 \implies P$
2. $f(B) = 0 \implies P$
3. $f(A) = 0 \;\&\; f(B) = 1 \implies \neg P$, otherwise
   $P \implies A = B = \{0,1\} \implies 0 = f(A) = f(B) = 1.$

In each case we decided whether P or $\neg P$ holds. □

Categoricity

- A theory is categorical iff all models of it are isomorphic.
- In such a case, the theory completely captures the structural essence of what it is trying to describe.
- Any two models of Dedekind arithmetic are isomorphic.
- All complete ordered fields are isomorphic.
- The complex field is also characterized up to isomorphism as the unique algebraically closed field of characteristic 0 having size continuum.
- Because of the upward Löwenheim-Skolem theorem, a first-order theory can't provide a categorical characterization of an infinite model. The categorical characterizations use second-order logic.
- Categoricity is central to the philosophy of structuralism.

κ-Categoricity

Definition (categoricity / κ-categoricity)

- A theory is *categorical* iff it has exactly one model.
- A theory is *κ-categorical* iff it has exactly one model of cardinality κ.

Theorem

A theory T is complete iff for all $\mathcal{M} \models \mathrm{T}$ and $\mathcal{N} \models \mathrm{T}$: $\mathcal{M} \equiv \mathcal{N}$.

Theorem

If a theory is categorical, then it is complete.

Theorem

For a theory with finite models, it is complete iff it is categorical.

Decidability

- Given any computably enumerable set of axioms, the set of theorems is also computably enumerable.
- If T is a complete theory with a computably enumerable set of axioms, then T is decidable.

Theorem (Łoś-Vaught Test)

If a theory with no finite model is κ-categorical, then it is complete.

Dense Linear Ordering without Endpoints DLO

1. $\forall x(x \not< x)$
2. $\forall x \forall y(x < y \to y < z \to x < z)$
3. $\forall x \forall y(x < y \vee x = y \vee y < x)$
4. $\forall x \forall y(x < y \to \exists z(x < z < y))$
5. $\forall x \exists y \exists z(y < x < z)$

Definition (κ-categoricity)

A theory is κ-categorical iff it has a unique model of cardinality κ.

Theorem (Cantor)

- *Any countable model of* DLO *is isomorphic to* $(\mathbb{Q}, <)$.
- *The theory* DLO *is* $\aleph_0$*-categorical.*
- *The theory* DLO *is complete and decidable.*
- $(\mathbb{R}, <)$ *is the unique complete linear ordering that has a countable dense subset isomorphic to* $(\mathbb{Q}, <)$.

Quantifier Elimination for DLO

- Every sentence can be rewritten as a collection of $<$s and $\leq$s joined by $\bigwedge$s and $\bigvee$s.
- Since $\exists x(A \vee B) \equiv \exists x A \vee \exists x B$, it suffices to eliminate quantifiers from a sentence of the form

$$\exists x \left(\bigwedge_i y_i \leq x \wedge \bigwedge_j y_j < x \wedge \bigwedge_k x \leq z_k \wedge \bigwedge_l x < z_l \right)$$

- This is equivalent to

$$\bigwedge_{i,k} y_i \leq z_k \wedge \bigwedge_{i,l} y_i < z_l \wedge \bigwedge_{j,k} y_j < z_k \wedge \bigwedge_{j,l} y_j < z_l$$

The Random Graph 随机图理论

G_1 $\forall x : \neg Rxx$

G_2 $\forall xy : Rxy \to Ryx$

G_3 $\exists xy : x \neq y$

ψ_n $\forall x_1 \dots x_n \forall y_1 \dots y_n : \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{n} x_i \neq x_j \to \exists z \bigwedge_{i=1}^{n} (Rx_i z \to \neg Ry_i z)$

无向图满足 G_1 和 G_2. 随机图 RG 还需满足 G_3 和 $\{\psi_n : n = 1, 2, \dots\}$.

- RG 不是不可数无穷范畴的.
- RG 是 $\aleph_0$ 范畴的. 也是完全的, 可判定的.

令 $\mathcal{G}_N$ 表示由论域为 $\{1, \dots, N\}$ 且满足 G_1 和 G_2 的无向图组成的集合. 对于 $\mathscr{L}_R$ 句子 A, 其在大小为 N 的图上为真的可能性为:

$$P_N(A) := \frac{|\{G \in \mathcal{G}_n : G \vDash A\}|}{|\mathcal{G}_N|}$$

Theorem (图上的 $0-1$ 律)

对任意 $\mathscr{L}_R$ 句子 A, $\lim_{N\to\infty} P_N(A) = 0$ 或 $\lim_{N\to\infty} P_N(A) = 1$.

而且, RG 公理化了几乎为真的理论 $\mathrm{Cn}(\mathrm{RG}) = \left\{A : \lim_{N\to\infty} P_N(A) = 1\right\}$.

Theorem

The theory ACF_p of algebraically closed fields of characteristic p (for p prime or 0) is κ-categorical for all uncountable cardinals κ.

Corollary

For $p \in \mathbb{P}$ or $p = 0$, ACF_p is complete and decidable.

Remark: 这并不意味着可以用计算机代替实分析学家的工作. 一些关于 $(\mathbb{R}, 0, 1, +, \cdot, \leq)$ 的陈述不是一阶的, 比如 $\mathbb{R}$ 的完备性.

Theorem (Morley's Categoricity Theorem)

If a theory is κ-categorical for some $\kappa \geq |\mathscr{L}|$, then it is categorical in all cardinalities $\geq |\mathscr{L}|$.

Problem

1. *Which view is the more plausible — that theories are the better the more nearly they are categorical, or that theories are the better the more they give rise to significant non-isomorphic interpretations?*
2. *Or is it rather the case that categoricity is a virtue in some theories but not in others?*

- The field of real numbers is not κ-categorical for any uncountable κ.
- The field of complex numbers is κ-categorical for any uncountable κ.
- Euclidean geometry studies spaces with real co-ordinates, smooth trajectories in these spaces and fits best with Newtonian physics.
- Algebraic geometry studies spaces with complex co-ordinates, or more generally, co-ordinates over algebraically closed fields in which all polynomial equations have solutions.

Remark: Categoricity versus Algorithmic Compressibility. If physical universe is co-ordinatizable in terms of complex numbers, then this explains that the comprehensive physical science is possible.

Lefschetz's Transfer Principle

Theorem (Lefschetz's Transfer Principle)

For a sentence A in the language of fields, the following are equivalent:

1. $\mathbb{C} \vDash A$
2. $\mathrm{ACF}_0 \vDash A$
3. $\mathrm{ACF}_p \vDash A$ *for all sufficiently large primes* p.
4. $\mathrm{ACF}_p \vDash A$ *for infinitely many primes* p.

Proof.

$(1 \leftrightarrow 2)$ follows from the completeness of ACF_0.
$(2 \to 3)$ assume $\mathrm{ACF}_0 \vDash A$, since the deduction $\mathrm{ACF}_0 \vdash A$ only use finitely many instances of $\overbrace{1+1+\cdots+1}^{n \text{ times}} \neq 0$, then for some finite $\Delta \subset \mathrm{ACF}_0 : \Delta \vdash A$, and $\mathrm{ACF}_p \vDash \Delta$ for all sufficiently large primes p.
$(3 \to 4)$ is trivial.
$(4 \to 2)$ $\mathrm{ACF}_0 \nvdash A \implies \mathrm{ACF}_0 \vdash \neg A \implies \mathrm{ACF}_p \vdash \neg A$ for all sufficiently large primes p. □

Ax-Grothendieck Theorem

- An *affine variety* is a set $V \subset \mathbb{C}^n$ s.t. $V = \{(x_1, \ldots, x_n) : f_i(x_1, \ldots, x_n) = 0, 1 \leq i \leq k\}$ with $f_i \in \mathbb{C}[x_1, \ldots, x_n]$.
- For any field K a map $f : K^n \to K^n$ is *polynomial* iff $f(x_1, \ldots, x_n) = \big(f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)\big)$ with $f_i \in K[x_1, \ldots, x_n]$.

Theorem (Ax-Grothendieck Theorem)

Let $f : V \to V$ be a polynomial map of an affine variety in $\mathbb{C}^n$. If f is injective, then it is surjective.

Ax-Grothendieck Theorem

Every injective polynomial map f: $\mathbb{C}^n \to \mathbb{C}^n$ is surjective.

Let $A :=$"Every injective polynomial map f of degree d is surjective".
By Lefschetz's transfer principle, we just have to show for all primes p, $\mathrm{ACF}_p \vdash A$.
Moreover, for each p, by completeness of ACF_p, we only need to show A is true in *some* model of ACF_p.
Consider the algebraic closure $F := \overline{\mathbb{F}}_p$ of the prime field $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$. We have $F = \bigcup_{n=1}^{\infty} \mathbb{F}_{p^n}$.
Let $\{a_1, \cdots, a_k\}$ be the set of coefficients appearing in $f : F^n \to F^n$.
For $(b_1, \cdots, b_n) \in F^n$, let K be the subfield of F generated by $\{a_1, \cdots, a_k, b_1, \cdots, b_n\}$.
Since K is finitely generated and $\exists N : K \subset \bigcup_{n=1}^{N} \mathbb{F}_{p^n}$, hence it is finite.
So $f\restriction_{K^n} : K^n \to K^n$ that is injective must be surjective.
Hence $f : F^n \to F^n$ is surjective.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Readings

1. J. van Benthem: Modal Logic for Open Minds
2. P. Blackburn, M. de Rijke, Y. Venema: Modal Logic

It is possible that it is raining.
It is certain that we will get wet if it is raining.

It is possible that we will get wet.

$$\frac{\Diamond R \qquad \Box(R \rightarrow W)}{\Diamond W}$$

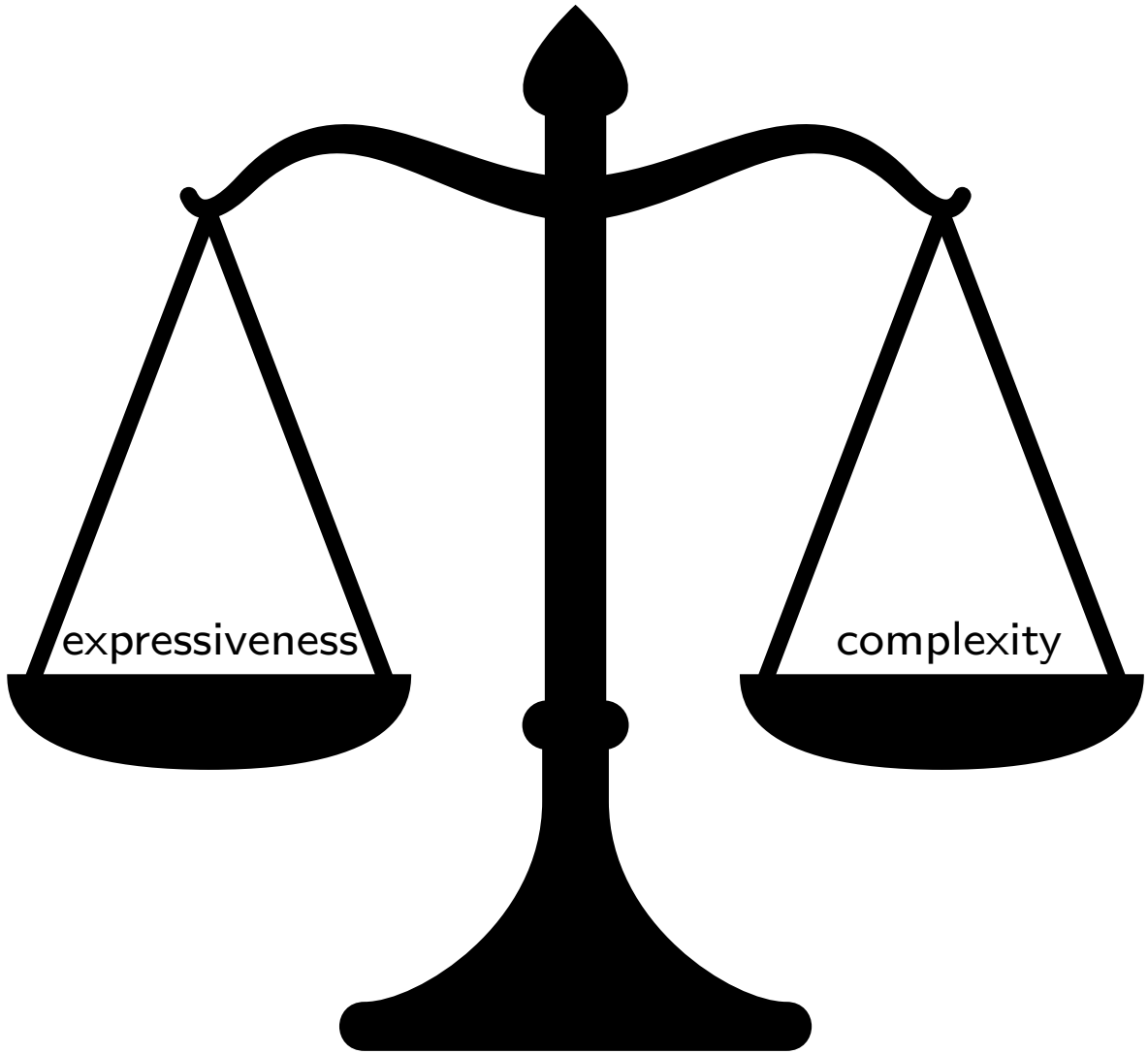
Paradox of Material Implication

- If God does not exist, then it's not the case that if I believe in God, I will have eternal life;
- and I don't believe in God;
- so God exists!

$$\frac{\neg G \rightarrow \neg(B \rightarrow E) \qquad \neg B}{G} \qquad \frac{\neg G \rightarrow \neg\Diamond(B \rightarrow E) \qquad \neg B}{G}\ ?$$

Why Study Modal Logic?

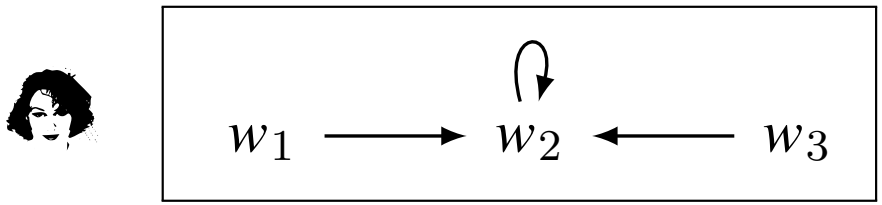
- Modal languages are simple yet expressive languages for talking about relational models.
- Modal languages are not isolated formal systems.
  - Modal vs classical (FOL,SOL), internal vs external perspective.
    In FOL, models are described from the top point of view. Each object and relation can be named. In modal logic, relational models are described from an internal perspective, there is no way to mention objects and relations.
  - Relational models vs Boolean algebra with operators.
    (Jónsson and Tarski's representation theorem.)
- Decidability.
  (seeking a balance between expressiveness and efficiency/complexity)



The more you can say, the less you can effectively/efficiently do.

Relational structures in first order and modal logic

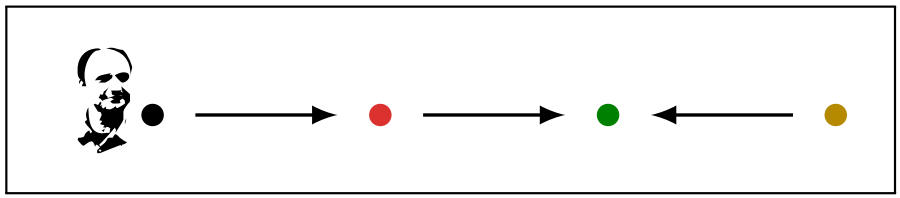
- In first order logic, relational structures are described from the top point of view. Each point of W and the relation R can be named.



Alice (in first order logic):

$$Rw_1w_2, Rw_2w_2, Rw_3w_2, \neg Rw_1w_3, \neg Rw_2w_1, \neg Rw_2w_3, \neg Rw_1w_1, \ldots$$

- In modal logic, relational structures are described from an internal, local perspective.



Bob (in modal logic): "I can reach a green point in 2 steps. There is no yellow point I can reach."

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Counterfactual Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Logics about Modalities

Alice _____ married.

- is possibly (basic modal logic)
- will be (temporal logic)
- is permitted to be (deontic logic)
- is known/believed to be (epistemic logic)
- is proved to be (provability logic)
- will be (after certain procedure) (dynamic logic)
- can be ensured (by her parents) to be (coalition logic)



Figure: Saul Kripke: 1920-2022

Syntax

Language

$$\mathscr{L} := \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \Box, \Diamond, (,)\} \cup \mathrm{Var}$$

where $\mathrm{Var} := \{p_1, p_2, p_3, \dots\}$.

Definition (Well-Formed Formula Wff)

$$A := p \mid \neg A \mid A \wedge A \mid A \vee A \mid A \rightarrow A \mid A \leftrightarrow A \mid \Box A \mid \Diamond A$$

- It will always be A. $\mathsf{G}A$
- You ought to do A. $\mathsf{O}A$
- I know A. $\mathsf{K}_i A$
- I believe A. $\mathsf{B}_i A$
- A is provable in T. $\Box_{\mathrm{T}} A$
- After the execution of the program α, A holds. $[\alpha]A$

Examples

1. "Ought" implies "can", but it does not imply "is".
$$(\Box p \to \Diamond p) \wedge \neg(\Box p \to p)$$
2. What must be is, and what is, is possible.
$$(\Box p \to p) \wedge (p \to \Diamond p)$$
3. Just because it happened that doesn't make it acceptable.
$$\neg(p \to \Diamond p)$$
4. Just because it happened that doesn't mean it ought to be permitted.
$$\neg(p \to \Box\Diamond p)$$
5. If it is raining, it is necessarily possible that it is raining.
$$p \to \Box\Diamond p$$
6. If it is possible that it is raining then it is necessarily possible that the corners are slippery.
$$\Diamond p \to \Box\Diamond q$$
7. Necessarily, if it is raining then it is possible that the corners are slippery.
$$\Box(p \to \Diamond q)$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Counterfactual Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Kripke Possible World Semantics[23]

A Kripke frame is a pair $\mathcal{F} := (W, R)$, where

- $W \neq \emptyset$
- $R \subset W \times W$

A Kripke model is $\mathcal{M} := (\mathcal{F}, V) = (W, R, V)$, where $V : \mathrm{Var} \to \mathtt{P}(W)$.

- $\mathcal{M}, w \vDash p$ iff $w \in V(p)$
- $\mathcal{M}, w \vDash \neg A$ iff $\mathcal{M}, w \nvDash A$
- $\mathcal{M}, w \vDash A \wedge B$ iff $\mathcal{M}, w \vDash A$ and $\mathcal{M}, w \vDash B$
- $\mathcal{M}, w \vDash \Box A$ iff $\forall v \in W \Big(Rwv \implies \mathcal{M}, v \vDash A \Big)$
- $\mathcal{M}, w \vDash \Diamond A$ iff $\exists v \in W \Big(Rwv \ \ \& \ \ \mathcal{M}, v \vDash A \Big)$

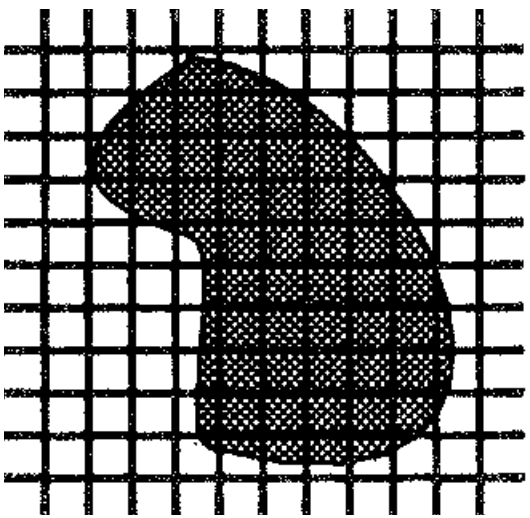
[23] Leibniz:

$$\mathcal{M}, w \vDash \Box A \text{ iff } \forall v \in W : \mathcal{M}, v \vDash A$$
$$\mathcal{M}, w \vDash \Diamond A \text{ iff } \exists v \in W : \mathcal{M}, v \vDash A$$

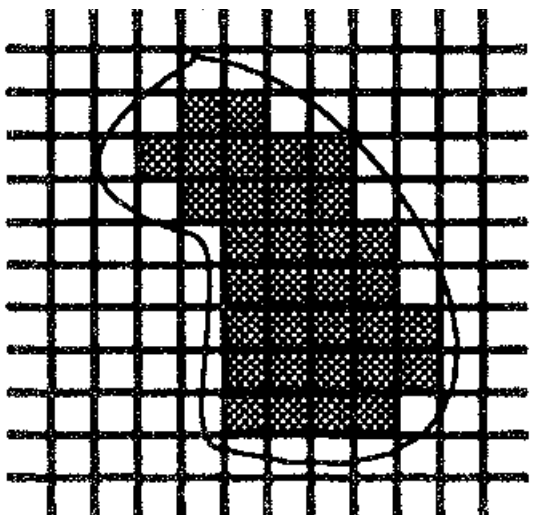
Possible World Semantics

We extend V to $[\![\,]\!] : \mathrm{Wff} \to \mathrm{P}(W)$ by recursion as follows:

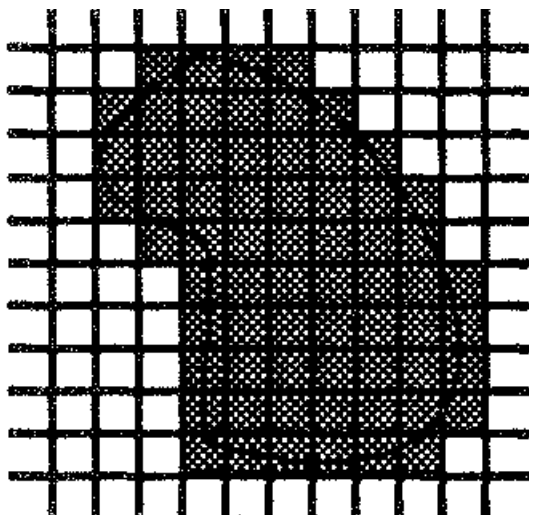
- $[\![p]\!] := V(p)$
- $[\![\neg A]\!] := \overline{[\![A]\!]}$
- $[\![A \wedge B]\!] := [\![A]\!] \cap [\![B]\!]$
- $[\![A \vee B]\!] := [\![A]\!] \cup [\![B]\!]$
- $[\![A \to B]\!] := \overline{[\![A]\!]} \cup [\![B]\!]$
- $[\![\Box A]\!] := \Box_R [\![A]\!]$ where $\Box_R X := \{w \in W : R(w) \subset X\}$
- $[\![\Diamond A]\!] := \Diamond_R [\![A]\!]$ where $\Diamond_R X := \{w \in W : R(w) \cap X \neq \emptyset\}$



A



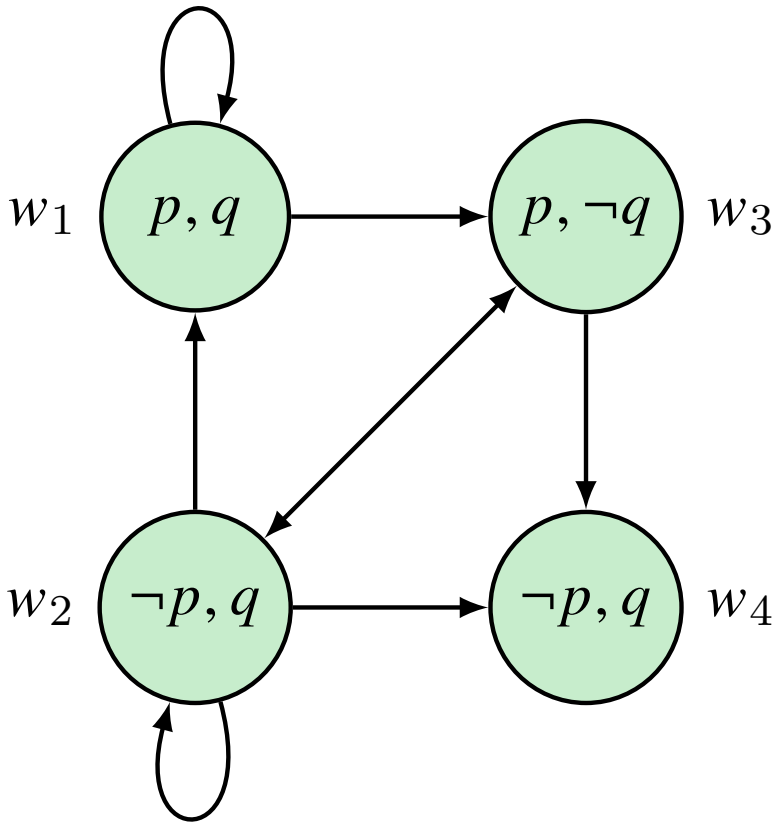
$\Box A$



$\Diamond A$

$$\mathcal{M}, w \vDash A \iff w \in [\![A]\!]$$

Example



$\mathcal{M}, w_1 \vDash p \wedge \Box p$

$\mathcal{M}, w_1 \vDash q \wedge \Diamond q$

$\mathcal{M}, w_1 \vDash \neg \Box q$

$\mathcal{M}, w_2 \vDash q \wedge \Diamond \neg q$

$\mathcal{M}, w_3 \vDash p$

$\mathcal{M}, w_3 \vDash \Box \neg p$

$\mathcal{M}, w_4 \vDash \Box p \wedge \neg \Diamond p$

Satisfiability & Validity

- A is satisfiable at $\mathcal{M}, w$ iff $\mathcal{M}, w \vDash A$.
- A is true in $\mathcal{M}$ ($\mathcal{M} \vDash A$) iff $\forall w \in W : \mathcal{M}, w \vDash A$
- A is valid in a pointed frame $\mathcal{F}, w$ ($\mathcal{F}, w \vDash A$) iff $\mathcal{M}, w \vDash A$ for every model $\mathcal{M}$ based on $\mathcal{F}$.
- A is valid in $\mathcal{F}$ ($\mathcal{F} \vDash A$) iff $\mathcal{M} \vDash A$ for every model $\mathcal{M}$ based on $\mathcal{F}$.
- A is valid in $\mathcal{C}$ ($\mathcal{C} \vDash A$) iff $\mathcal{F} \vDash A$ for every frame $\mathcal{F} \in \mathcal{C}$.
- $\vDash A$ iff $\mathcal{F} \vDash A$ for every $\mathcal{F}$.

Truth is in the eye of the beholder.

Example

$$\vDash \Box(A \to B) \to \Box A \to \Box B$$

逻辑蕴含 Entailment

- 保点

$$\Gamma \vDash A \ := \ \forall \mathcal{M} \forall w \in W \Big(\mathcal{M}, w \vDash \Gamma \implies \mathcal{M}, w \vDash A \Big)$$

- 保模型

$$\Gamma \vDash_M A \ := \ \forall \mathcal{M} \Big(\mathcal{M} \vDash \Gamma \implies \mathcal{M} \vDash A \Big)$$

- 保框架

$$\Gamma \vDash_F A \ := \ \forall \mathcal{F} \Big(\mathcal{F} \vDash \Gamma \implies \mathcal{F} \vDash A \Big)$$

- 保有效

$$\vDash \Gamma \implies \vDash A$$

Remark: “保点” 强于 “保模型” 强于 “保框架” 强于 “保有效”.
分离规则保点有效, 必然化规则保模型有效, 代入规则保框架有效.

Example

- $p \nvDash \Box p$
- $p \vDash_M \Box p$

Material Implication vs C. I. Lewis' Strict Implication

$$p \dashv\!\!\!\supset q \coloneqq \Box(p \rightarrow q)$$

- $\not\vDash A \dashv\!\!\!\supset B \dashv\!\!\!\supset A$?
- $\not\vDash (A \dashv\!\!\!\supset B) \vee (B \dashv\!\!\!\supset C)$?
- $\not\vDash \neg(A \dashv\!\!\!\supset B) \dashv\!\!\!\supset (A \wedge \neg B)$?
- $\vDash (A \wedge \neg A) \dashv\!\!\!\supset B$
- $\vDash A \dashv\!\!\!\supset (B \vee \neg B)$
- $\neg\Diamond A \vDash A \dashv\!\!\!\supset B$
- $\Box A \vDash B \dashv\!\!\!\supset A$

Accessibility

serial	$\forall x \exists y : Rxy$
reflexive	$\forall x : Rxx$
symmetric	$\forall xy : Rxy \to Ryx$
transitive	$\forall xyz : Rxy \wedge Ryz \to Rxz$
euclidean	$\forall xyz : Rxy \wedge Rxz \to Ryz$
total	$\forall xy : Rxy \vee Ryx$
isolation	$\exists x \forall y : \neg Rxy \wedge \neg Ryx$
successor reflexive	$\forall x \exists y : Rxy \wedge Ryy$
asymmetric	$\forall xy : Rxy \to \neg Ryx$
antisymmetric	$\forall xy : Rxy \wedge Ryx \to x = y$

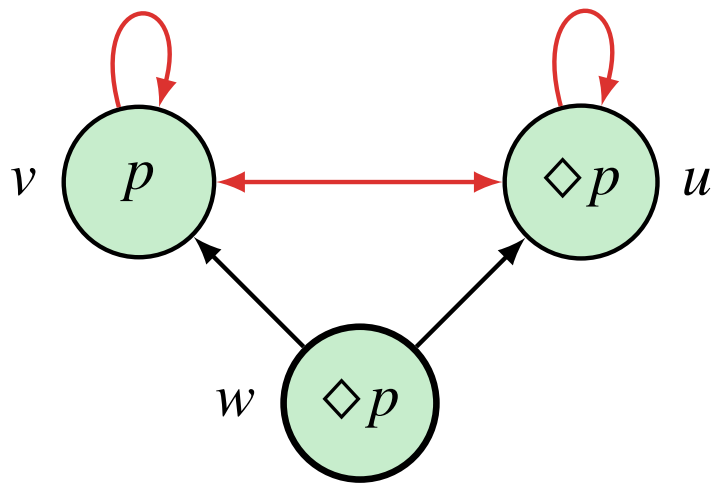
Correspondence Theorem

Theorem (Correspondence Theorem)

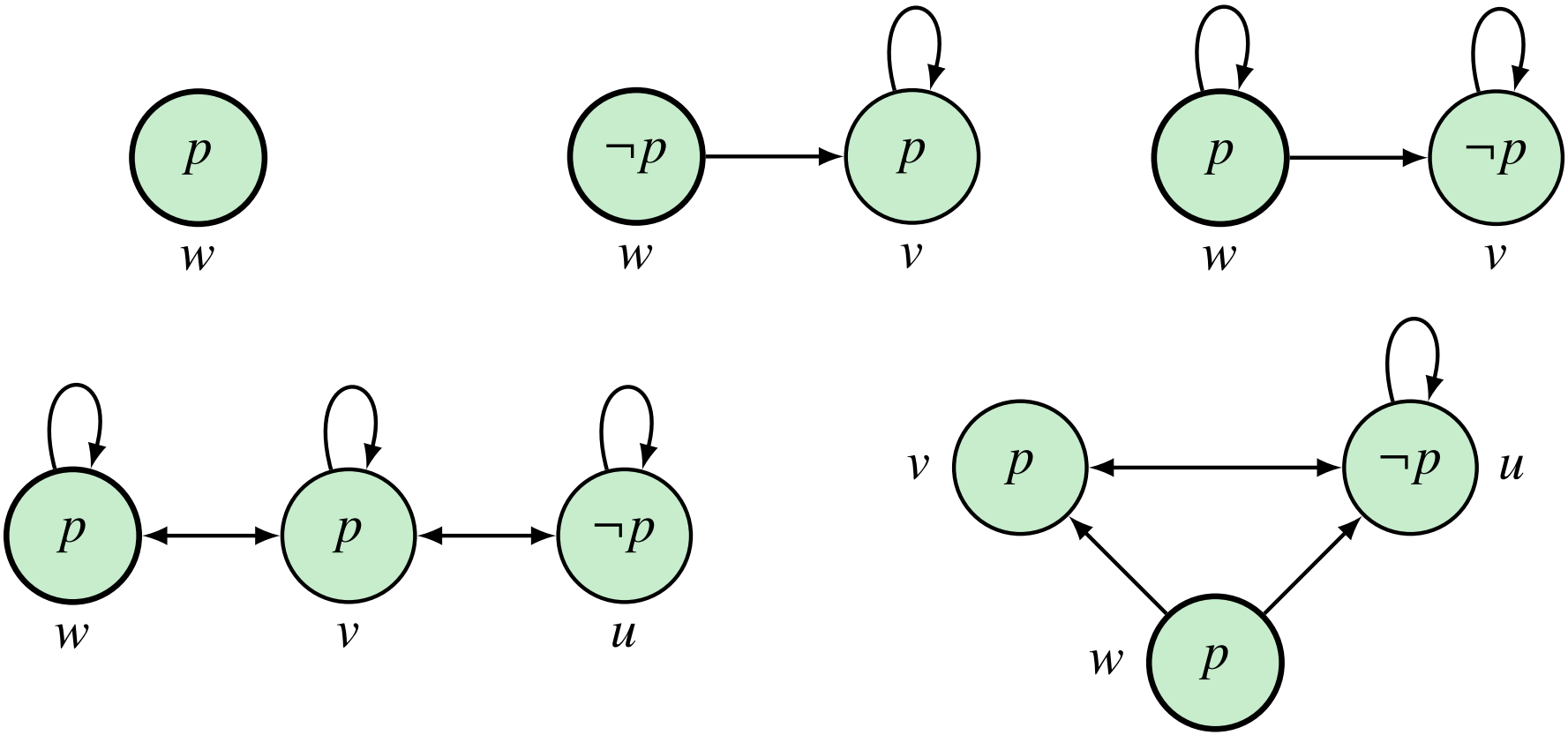
D	$W, R \vDash \Box p \to \Diamond p$	$\iff$	R *is serial*	$\forall x \exists y : Rxy$
T	$W, R \vDash \Box p \to p$	$\iff$	R *is reflexive*	$\forall x : Rxx$
B	$W, R \vDash p \to \Box \Diamond p$	$\iff$	R *is symmetric*	$\forall xy : Rxy \to Ryx$
4	$W, R \vDash \Box p \to \Box \Box p$	$\iff$	R *is transitive*	$\forall xyz : Rxy \wedge Ryz \to Rxz$
5	$W, R \vDash \Diamond p \to \Box \Diamond p$	$\iff$	R *is euclidean*	$\forall xyz : Rxy \wedge Rxz \to Ryz$

Remark: B is for Brouwer: if we regard "negation" as "necessarily negative", then we get $p \to \Box \neg \Box \neg p$ from $p \to \neg \neg p$.

Proof 5 ($\Longleftarrow$): Assume $W, R \not\vDash 5$.
Then $w \vDash \Diamond p$ but $w \not\vDash \Box \Diamond p$.
Now $w \vDash \Diamond p \implies \exists v : Rwv \ \& \ v \vDash p$,
and $w \not\vDash \Box \Diamond p \implies w \vDash \Diamond \neg \Diamond p \implies \exists u :$
$Rwu \ \& \ u \vDash \neg \Diamond p$.
However, $Ruv \implies u \vDash \Diamond p$.
So we have $Rwv \wedge Rwu$ but $\neg Ruv$.



Proof ($\Longrightarrow$): Counter-model for $D, T, B, 4, 5$



More Correspondence

D	$W, R \vDash \Box p \to \Diamond p$	$\iff$	R is serial	$\forall x \exists y : Rxy$
T	$W, R \vDash \Box p \to p$	$\iff$	R is reflexive	$\forall x : Rxx$
B	$W, R \vDash p \to \Box \Diamond p$	$\iff$	R is symmetric	$\forall xy : Rxy \to Ryx$
4	$W, R \vDash \Box p \to \Box \Box p$	$\iff$	R is transitive	$\forall xyz : Rxy \wedge Ryz \to Rxz$
5	$W, R \vDash \Diamond p \to \Box \Diamond p$	$\iff$	R is euclidean	$\forall xyz : Rxy \wedge Rxz \to Ryz$
CD	$W, R \vDash \Diamond p \to \Box p$	$\iff$	R is partially functional	$\forall xyz : Rxy \wedge Rxz \to y = z$
	$W, R \vDash \Diamond p \leftrightarrow \Box p$	$\iff$	R is functional	$\forall x \exists! y : Rxy$
$\Box M$	$W, R \vDash \Box(\Box p \to p)$	$\iff$	R is shift reflexive	$\forall xy : Rxy \to Ryy$
$C4$	$W, R \vDash \Box \Box p \to \Box p$	$\iff$	R is dense	$\forall xy : Rxy \to \exists z : Rxz \wedge Rzy$
C	$W, R \vDash \Diamond \Box p \to \Box \Diamond p$	$\iff$	R is convergent	$\forall xyz : Rxy \wedge Rxz \to \exists w : Ryw \wedge Rzw$
	$W, R \vDash \Box(\Box p \to q) \vee \Box(\Box q \to p)$	$\iff$	R is connected	$\forall xyz : Rxy \wedge Rxz \to Ryz \vee Rzy$
	$W, R \vDash \Box(p \wedge \Box p \to q) \vee \Box(q \wedge \Box q \to p)$	$\iff$	R is weakly connected	$\forall xyz : Rxy \wedge Rxz \to Ryz \vee y = z \vee Rzy$
	$W, R \vDash p \to \Box p$	$\iff$		$\forall xy : Rxy \to y = x$
	$W, R \vDash \Box \bot$	$\iff$		$\forall x \neg \exists y : Rxy$
	$W, R \vDash \Diamond p \to p \vee \Box \Diamond p$	$\iff$		$\forall xyz : Rxy \wedge Rxz \to y = x \vee Ryz$

Standard Translation

Definition (Standard Translation)

$$T_x(p) = P(x) \qquad T_y(p) = P(y)$$
$$T_x(\neg A) = \neg T_x(A) \qquad T_y(\neg A) = \neg T_y(A)$$
$$T_x(A \wedge B) = T_x(A) \wedge T_x(B) \qquad T_y(A \wedge B) = T_y(A) \wedge T_y(B)$$
$$T_x(\Box A) = \forall y(Rxy \rightarrow T_y(A)) \qquad T_y(\Box A) = \forall x(Ryx \rightarrow T_x(A))$$

Theorem (Correspondence on Models)

$$\mathcal{M}, w \vDash A \iff \mathcal{M} \vDash T_x(A)[w]$$
$$\mathcal{M} \vDash A \iff \mathcal{M} \vDash \forall x T_x(A)$$
$$\mathcal{F}, w \vDash A \iff \mathcal{F} \vDash \forall P_1, \ldots, P_n T_x(A)[w]$$
$$\mathcal{F} \vDash A \iff \mathcal{F} \vDash \forall P_1, \ldots, P_n \forall x T_x(A)$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Counterfactual Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Tree Method for Modal Logic

$w \vDash \Box A$

|

$v \vDash A$

if Rwv is already in the branch.

$w \nvDash \Box A$

|

Rwv

$v \nvDash A$

where v is new in the branch.

$w \vDash \Diamond A$

|

Rwv

$v \vDash A$

where v is new in the branch.

$w \nvDash \Diamond A$

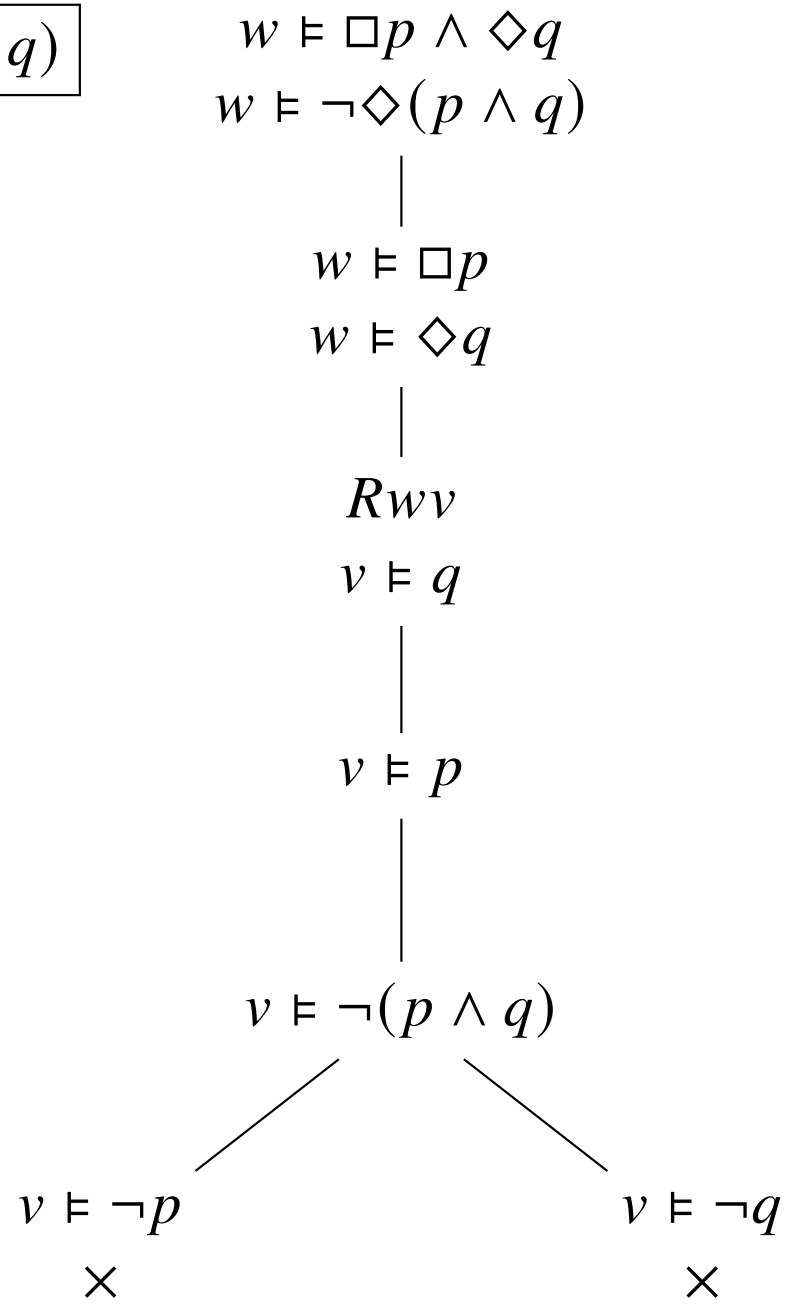
|

$v \nvDash A$

if Rwv is already in the branch.

Example — Tree Method for Modal Logic

$\vDash \Box p \wedge \Diamond q \rightarrow \Diamond(p \wedge q)$



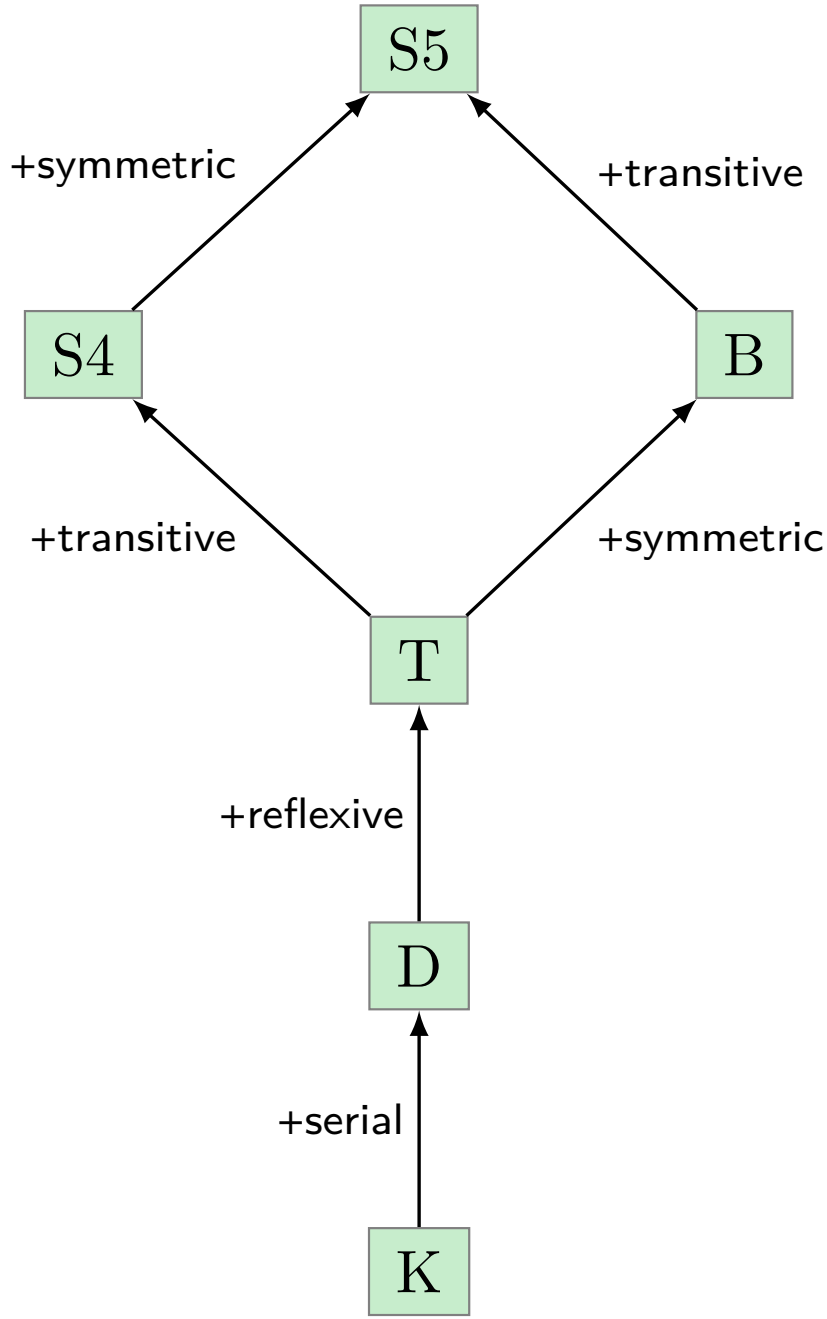
Formal System = Axiom + Inference Rule

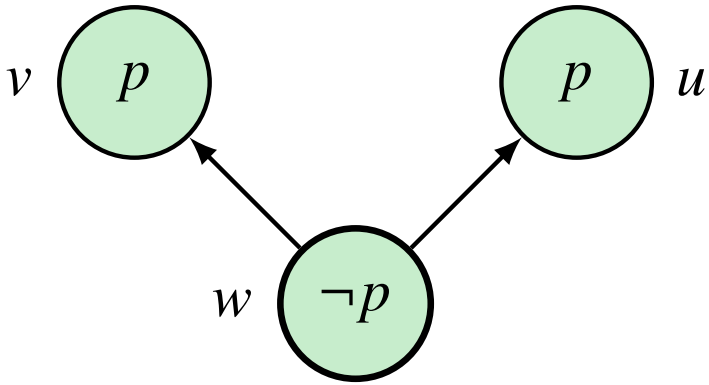
Axiom Schema

	tautologies
Dual	$\Diamond A \leftrightarrow \neg\Box\neg A$
K	$\Box(A \to B) \to \Box A \to \Box B$
D	$\Box A \to \Diamond A$
T	$\Box A \to A$
B	$A \to \Box\Diamond A$
4	$\Box A \to \Box\Box A$
5	$\Diamond A \to \Box\Diamond A$
L	$\Box(\Box A \to A) \to \Box A$

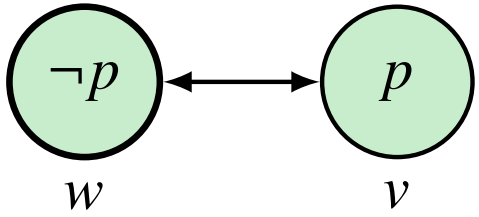
Inference Rule

$$\frac{A \quad A \to B}{B}\ \text{MP} \qquad\qquad \frac{\vdash A}{\vdash \Box A}\ \text{N}$$


S5
+symmetric
+transitive
S4
B
+transitive
+symmetric
T
+reflexive
D
+serial
K

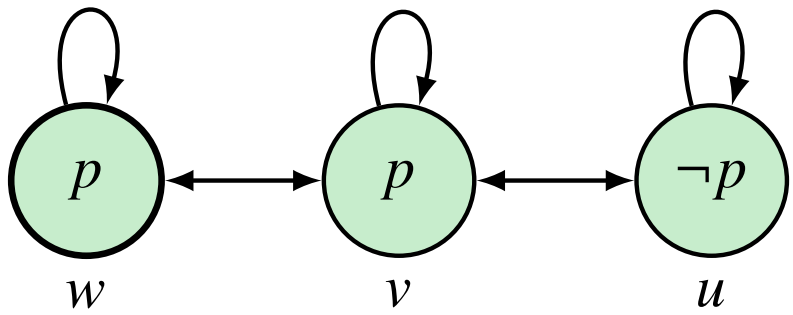


$p \rightarrow \Diamond p \not\models \Box p \rightarrow p$



KB $\not\vdash 4$

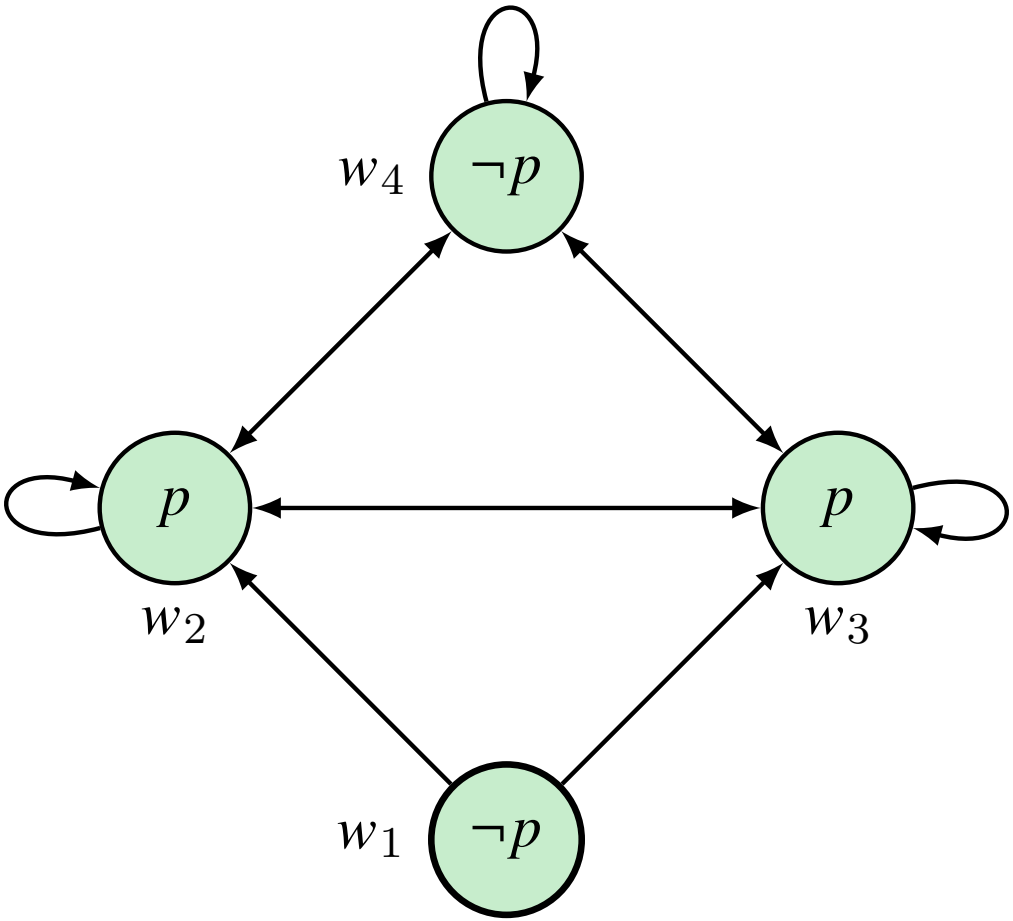
$w \not\models \Box p \rightarrow \Box\Box p$



KTB $\not\vdash 4$ $\quad w \not\models \Box p \rightarrow \Box\Box p$

KTB $\not\vdash 5$ $\quad v \not\models \Diamond\neg p \rightarrow \Box\Diamond\neg p$

KD5 $\nvdash$ 4



$w_1 \nvDash \Box p \rightarrow \Box\Box p$

Intuitionistic Logic vs Modal Logic

$$\mathrm{S4} := K + T + 4$$

$$\mathrm{Grz} := \mathrm{S4} + \Box(\Box(A \to \Box A) \to A) \to A$$

$$p^* := \Box p$$
$$(\neg A)^* := \Box \neg A^*$$
$$(A \wedge B)^* := A^* \wedge B^*$$
$$(A \vee B)^* := A^* \vee B^*$$
$$(A \to B)^* := \Box(A^* \to B^*)$$

$$\vdash_{\mathrm{I}} A \iff \vdash_{\mathrm{S4}} A^* \iff \vdash_{\mathrm{Grz}} A^*$$

$$\mathrm{GL} := K + L$$

$$p' := p$$
$$(\neg A)' := \neg A'$$
$$(A \wedge B)' := A' \wedge B'$$
$$(A \vee B)' := A' \vee B'$$
$$(A \to B)' := A' \to B'$$
$$(\Box A)' := A' \wedge \Box A'$$

$$\vdash_{\mathrm{Grz}} A \iff \vdash_{\mathrm{GL}} A'$$
$$\vdash_{\mathrm{I}} A \iff \vdash_{\mathrm{GL}} (A^*)'$$

$$\boxed{\Box(\Box A \to A) \to \Box A \vdash_{\mathrm{GL}} \Box A \to \Box\Box A}$$

- $A \wedge \Box A \wedge \Box\Box A \to A \wedge \Box A$
- $A \to \Box(A \wedge \Box A) \to A \wedge \Box A$ $\qquad\qquad \Box(p \wedge q) \leftrightarrow \Box p \wedge \Box q$
- $\Box(\Box(A \wedge \Box A) \to A \wedge \Box A) \to \Box(A \wedge \Box A)$ $\qquad\qquad$ L
- $\Box A \to \Box(A \wedge \Box A)$
- $\Box A \to \Box\Box A$

Theorem

$W, R \vDash \Box(\Box A \to A) \to \Box A \iff R$ *is transitive &* R *is reverse well-founded: there are no chains* $w_0 R w_1 R w_2 \dots$.

Proof.

Assume Rw_0w_1 and Rw_1w_2, but not Rw_0w_2. Setting $V(p) := W \setminus \{w_1, w_2\}$ makes L false at w_0.
Assume R is transitive, and there is an ascending sequence $w_0 R w_1 R w_2 \dots$. Then $V(p) := W \setminus \{w_0, w_1, w_2, \dots\}$ refutes L at w_0.
Conversely, if L fails at w_0, there must be an infinite upward sequence of $\neg p$-worlds. This arises by taking any successor of w_0 where p fails, and repeatedly applying the truth of $\Box(\Box p \to p)$ — using the transitivity of the frame. □

Remark: transitivity is definable in first order logic, but well-foundedness can't be defined in first order logic. Frame truth is a second order notion.

Theorem (Craig Interpolation)

If $\mathrm{GL} \vdash A \to B$*, then there is a* C *with* $\mathrm{Var}(C) \subset \mathrm{Var}(A) \cap \mathrm{Var}(B)$ *s.t.*

$$\mathrm{GL} \vdash A \to C \quad \textit{and} \quad \mathrm{GL} \vdash C \to B$$

Corollary (Beth Definability)

Assume $\mathrm{GL} \vdash A(p) \wedge A(q) \to (p \leftrightarrow q)$ *where* $q \notin \mathrm{Var}(A)$ *and* $A(q)$ *is obtained from* $A(p)$ *by replacing all occurrences of* p *by* q*. Then there exists a formula* B *with* $\mathrm{Var}(B) \subset \mathrm{Var}(A) \setminus \{p\}$ *s.t.*

$$\mathrm{GL} \vdash A(p) \to (p \leftrightarrow B)$$

Proof.

Let B be an interpolant for $\mathrm{GL} \vdash A(p) \wedge p \to (A(q) \to q)$. □

Theorem (Uniqueness of Fixpoint)

If p occurs only boxed in $A(p)$ and $q \notin \mathrm{Var}(A)$, then

$$\mathrm{GL} \vdash \boxdot\Big(\big(p \leftrightarrow A(p)\big) \wedge \big(q \leftrightarrow A(q)\big)\Big) \rightarrow (p \leftrightarrow q)$$

where $\boxdot A := A \wedge \Box A$.

Corollary

If p occurs only boxed in $A(p)$, then

$$\mathrm{GL} \vdash B \leftrightarrow A(B) \;\;\&\;\; \mathrm{GL} \vdash C \leftrightarrow A(C) \implies \mathrm{GL} \vdash B \leftrightarrow C$$

Theorem (Existence of Fixpoint)

If p occurs only boxed in $A(p)$, then there exists a formula B with $\mathrm{Var}(B) \subset \mathrm{Var}(A) \setminus \{p\}$ s.t.

$$\mathrm{GL} \vdash B \leftrightarrow A(B)$$

Uniqueness of Fixpoint + Beth Definability $\implies$ Existence of Fixpoint

$$\mathrm{GL} \vdash \neg\Box\bot \leftrightarrow \neg\Box(\neg\Box\bot) \qquad \mathrm{GL} \vdash \top \leftrightarrow \Box\top$$

Soundness & Completeness

Definition (Theorem & Local Syntactic Deduction)

- $\vdash_S A$
- $\Gamma \vdash_S A$ iff $\vdash_S \bigwedge_{i=1}^{n} B_i \to A$ for some finite subset $\{B_1, \ldots, B_n\} \subset \Gamma$.

Theorem (Soundness & Completeness)

Let S be the normal system $KX_1 \ldots X_n$ and $C = \bigcap_{i=1}^{n} C_i$ where each C_i is the corresponding class of frames for axiom schema X_i.

$$\Gamma \vdash_S A \iff \Gamma \vDash_C A$$

时空逻辑

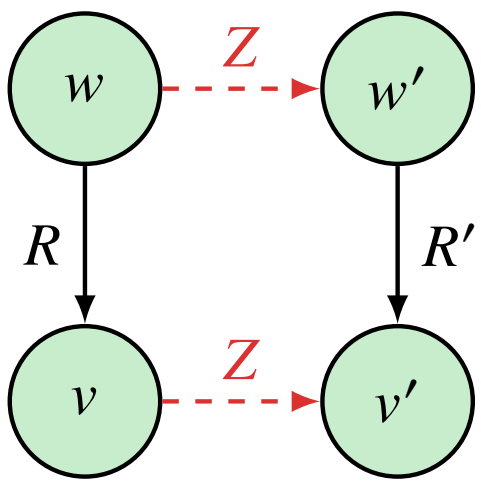
- 相对论下四维时空中的点: (s, t)
- 定义 $(s, t) \leq (s', t')$, 当且仅当, (s', t') 在 (s, t) 的因果未来中 $(s - s')^2 \leq (t - t')^2$ (即: 因果作用只能在光锥内传播).
- $\leq$ 是自反、传递、有向 (directed) 的关系.
- 公理系统 $\mathrm{S4.2} := \mathrm{S4} + \Diamond\Box p \to \Box\Diamond p$
- $\mathrm{S4.2}$ 对自反、传递、有向的框架是可靠且完备的.
- 假设时间有终点:
  — 如果一个自反、传递、有向的框架有极大元, 那么它有唯一的最大元.
  — $\mathrm{S4.2} + \Box\Diamond p \to \Diamond\Box p$ 对有最大元的自反、传递、有向的框架是可靠且完备的.

Bisimulation

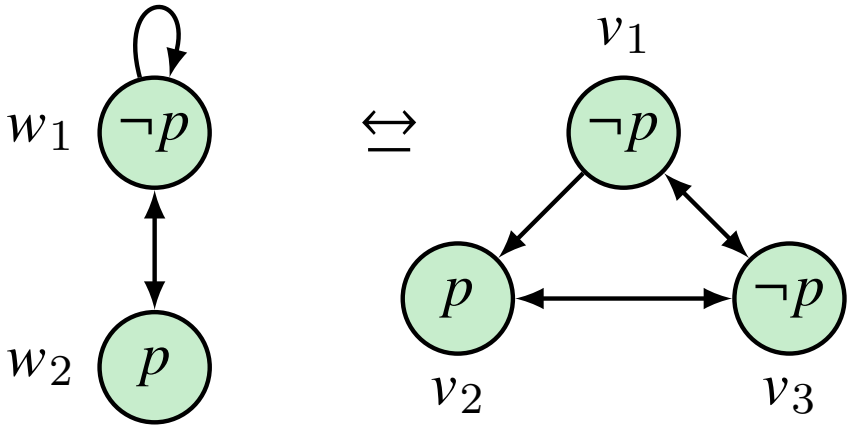
Definition (Bisimulation)

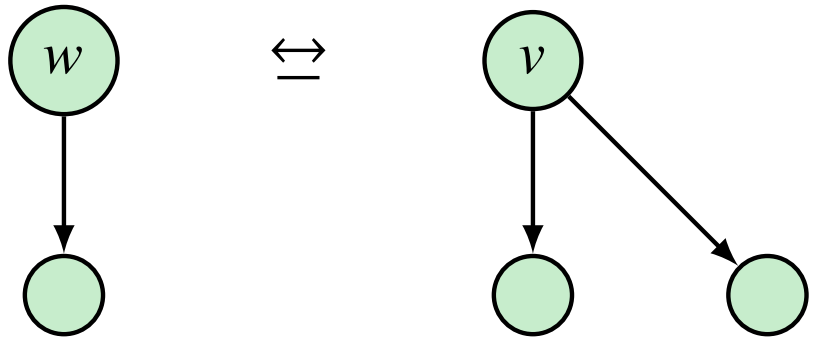
A bisimulation $Z : \mathcal{M} \leftrightarrows \mathcal{M}'$ between Kripke models $\mathcal{M} = (W, R, V)$ and $\mathcal{M}' = (W', R', V')$ is a binary relation $Z \subset W \times W'$ such that, $w \xrightarrow{Z} w'$ implies that

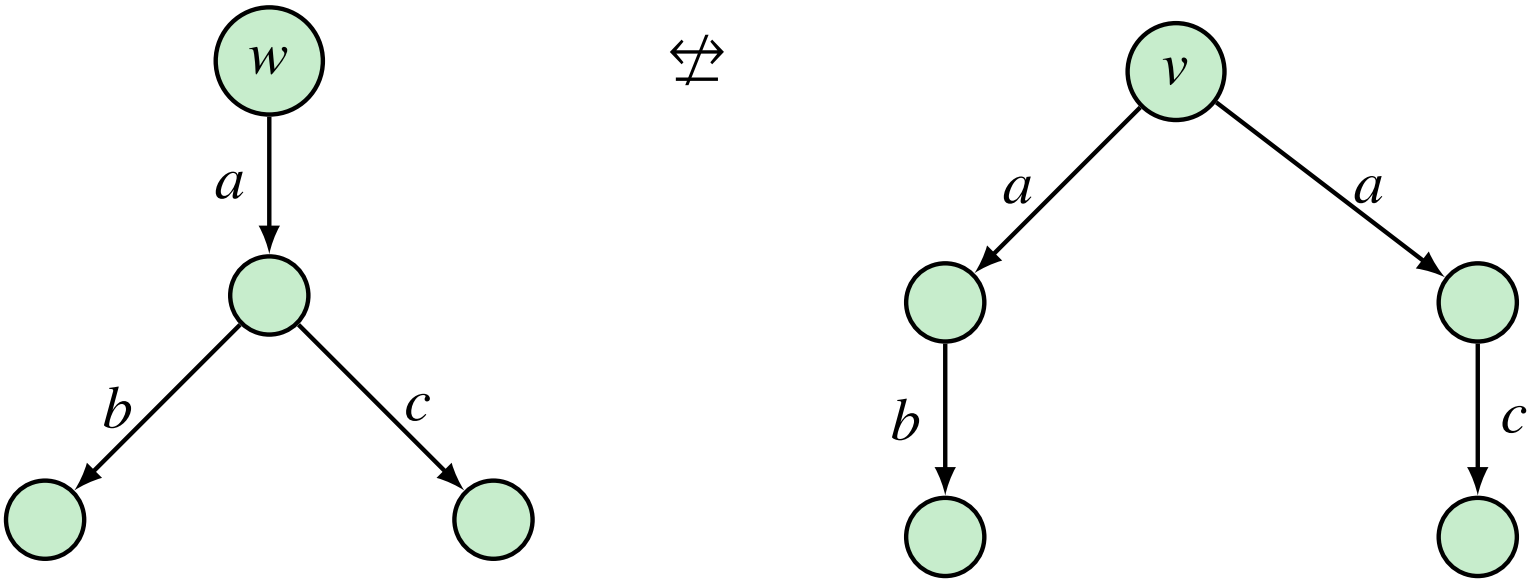
1. $w \in V(p) \iff w' \in V'(p)$ for all proposition letters p.
2. $w \xrightarrow{R} v \implies \exists v' \in W' : \; v \xrightarrow{Z} v' \;\&\; w' \xrightarrow{R'} v'$
3. $w' \xrightarrow{R'} v' \implies \exists v \in W : \; v \xrightarrow{Z} v' \;\&\; w \xrightarrow{R} v$



Bisimulation — Example







$\Box_a(\Diamond_b \top \wedge \Diamond_c \top)$ or $\Box_a \Diamond_b \top$

Bisimulation Game

Definition (n-round Bisimulation Game)

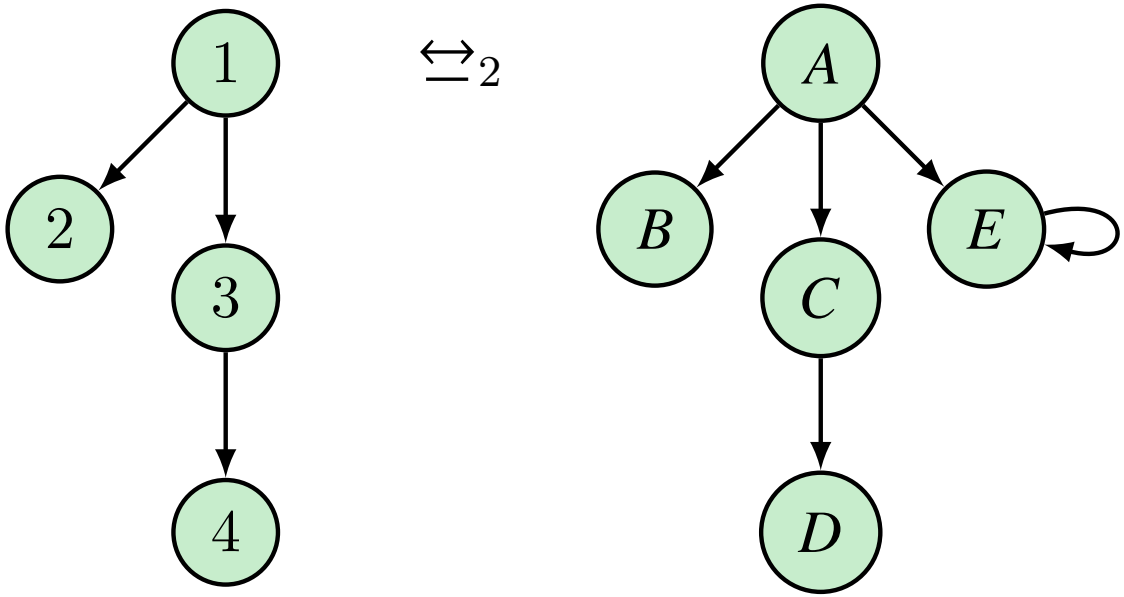
An n-round bisimulation game G_n between $(\mathcal{M}, w)$ and $(\mathcal{N}, v)$ is a two player game based on the configurations in $W_{\mathcal{M}} \times W_{\mathcal{N}}$. The initial configuration is (w, v) and the players, Spoiler and Defender, play in rounds. At each configuration (s, t):

1. Spoiler chooses either model $\mathcal{M}$ or $\mathcal{N}$, and an successor x of s or t.
2. Defender must select a successor y in the other model. If Defender cannot find a matching successor, or (x, y) differ in their atomic properties, Spoiler wins.

Theorem

1. $\mathcal{M}, w \leftrightarrows_n \mathcal{N}, v$ *iff Defender has a winning strategy in* G_n.
2. $\mathcal{M}, w \leftrightarrows_\omega \mathcal{N}, v$ *iff Defender has a winning strategy in* G_n *for each* n.
3. $\mathcal{M}, w \leftrightarrows \mathcal{N}, v$ *iff Defender has a winning strategy in* G_∞.

Bisimulation — Example



Defender can win k-round games ($k < 3$).

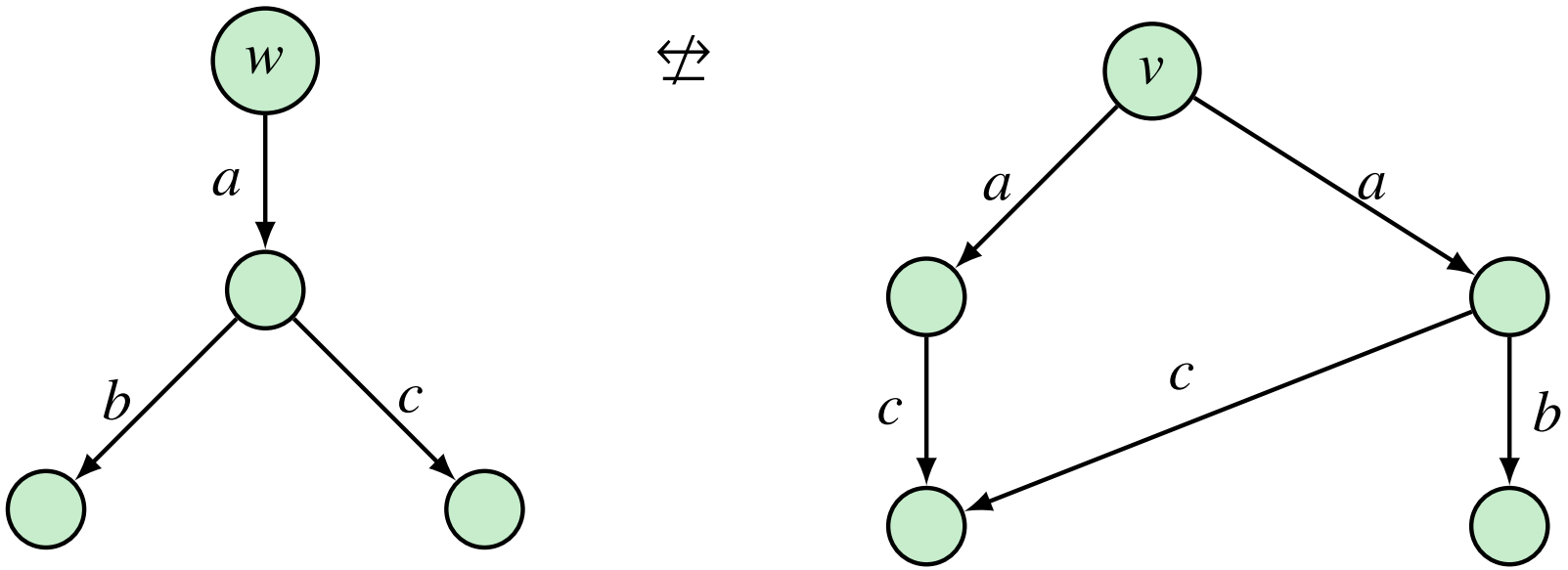
$\leftrightarrow_0 = \{1,2,3,4\} \times \{A,B,C,D,E\}$

$\leftrightarrow_1 = \{(2,B),(2,D),(4,B),(4,D),(1,A),(1,C),(1,E),(3,A),(3,C),(3,E)\}$

$\leftrightarrow_2 = \{(2,B),(2,D),(4,B),(4,D),(1,A),(3,C)\}$

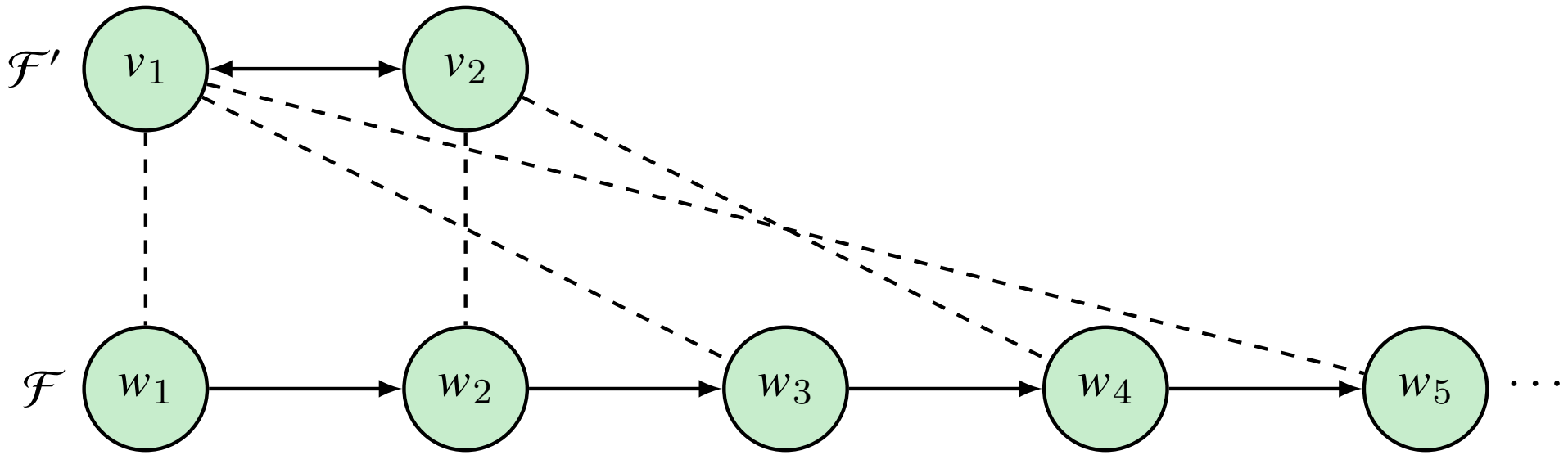
$\leftrightarrow_3 = \{(2,B),(2,D),(4,B),(4,D),(3,C)\}$ but $1 \not\leftrightarrow_3 A$

Bisimulation Game — Example



Spoiler can win the 2-round bisimulation game. It is important to switch models!

反对称性不可模态定义



假设公式 A 定义反对称性, 则对任意框架 $\mathcal{F}$: $\mathcal{F} \vDash A \iff \mathcal{F}$ 是反对称的. 考虑如上两个框架: $\mathcal{F} \vDash A, \mathcal{F}' \nvDash A$, 这意味着 $\mathcal{F}'$ 有赋值 V' 和点 v, 使得 $\mathcal{F}', V', v \nvDash A$, 但是按照虚线我们可以把这个赋值迁移到 $\mathcal{F}$ 上 (记为 V), 并使虚线是互模拟, 所以 $\mathcal{F}', V', v \leftrightarrows \mathcal{F}, V, w$, 因而 $\mathcal{F}, V, w \nvDash A$ 与 $\mathcal{F} \vDash A$ 矛盾.

Bisimulation

Theorem (van Benthem Characterization Theorem1976)

Let $A(x)$ be a first order formula. Then $A(x)$ is bisimulation invariant iff it is (equivalent to) the standard translation of a modal formula.

Remark: Modal logic is the bisimulation-invariant fragment of first-order logic!

Theorem (van Benthem2007)

An abstract modal logic extending basic modal logic and satisfying compactness and bisimulation invariance is equally expressive as the basic modal logic K.

Topological Semantics for Modal Logic S4

$$\begin{aligned}
[\![\neg A]\!] &:= \overline{[\![A]\!]} \\
[\![A \wedge B]\!] &:= [\![A]\!] \cap [\![B]\!] \\
[\![A \vee B]\!] &:= [\![A]\!] \cup [\![B]\!] \\
[\![A \to B]\!] &:= \overline{[\![A]\!]} \cup [\![B]\!] \\
[\![\bot]\!] &:= \emptyset \\
[\![\top]\!] &:= W \\
[\![\Box A]\!] &:= [\![A]\!]^\circ \\
[\![\Diamond A]\!] &:= \mathrm{cl}([\![A]\!])
\end{aligned}$$

where $^\circ$ is the interior operator $A^\circ := \bigcup\{U \in \mathcal{O}_W : U \subset A\}$, and cl is the closure operator $\mathrm{cl}(X) := \overline{\left(\overline{X}\right)^\circ}$.

$$\mathcal{O}_W \vDash A := [\![A]\!] = W$$

Theorem

$A \vdash_{\mathrm{S4}} B \iff [\![A]\!] \subset [\![B]\!]$ *for every topological interpretation* $(W, \mathcal{O}_W, [\![\,]\!])$

Topological Semantics for Modal Logic — Another Version

$$[\![\diamond A]\!] := [\![A]\!]'$$

where $'$ is the derivative operator

$$A' := \Big\{ w \in W : \forall U \in \mathcal{O}_W \big(w \in U \implies A \cap (U \setminus \{w\}) \neq \emptyset \big) \Big\}.$$

$$\mathcal{M}, w \vDash \boxdot A \iff \forall U \ni w \forall v \in U \setminus \{w\} : \mathcal{M}, v \vDash A$$
$$\mathcal{M}, w \vDash \diamond A \iff \forall U \ni w \exists v \in U \setminus \{w\} : \mathcal{M}, v \vDash A$$

Example: Suppose $\mathcal{M} = (\mathbb{R}, \mathcal{O}_{\mathbb{R}}, V)$, where $\mathcal{O}_{\mathbb{R}}$ is the set of the union of open intervals. Let $V : \mathrm{Var} \to \mathrm{P}(W)$ with $V(p) = \{\frac{1}{n} : n \geq 1\}$. Then $\mathcal{M}, 0 \vDash \diamond p$. But $[\![\diamond p]\!]' = \emptyset$. Therefore $\mathcal{M}, 0 \nvDash \diamond\diamond p$.

Neighborhood Semantics for Modal Logic

Definition (Neighborhood Model)

A neighborhood model is $\mathcal{M} = (W, N, V)$, where $W \neq \emptyset$, $N : W \to \mathtt{P}(\mathtt{P}(W))$ is a neighborhood function, and $V : \mathrm{Var} \to \mathtt{P}(W)$ is a assignment function.

Definition (Truth)

$$\mathcal{M}, w \vDash p \iff w \in V(p)$$
$$\mathcal{M}, w \vDash \neg A \iff \mathcal{M}, w \nvDash A$$
$$\mathcal{M}, w \vDash A \wedge B \iff \mathcal{M}, w \vDash A \ \ \& \ \ \mathcal{M}, w \vDash B$$
$$\mathcal{M}, w \vDash \Box A \iff [\![A]\!] \in N(w)$$
$$\mathcal{M}, w \vDash \Diamond A \iff \overline{[\![A]\!]} \notin N(w)$$

where $[\![A]\!] := \{w : \mathcal{M}, w \vDash A\}$.

Remark:

$$[\![\Box A]\!] = \Box_N [\![A]\!] \quad \text{where} \quad \Box_N X := \left\{w \in W : X \in N(w)\right\}$$
$$[\![\Diamond A]\!] = \Diamond_N [\![A]\!] \quad \text{where} \quad \Diamond_N X := \left\{w \in W : \overline{X} \notin N(w)\right\}$$

Topological Semantics vs Neiborhood Semantics

Definition (Topological Space Definition via Neighborhoods)

(W, N) is a topological space iff $N : W \to P(P(W))$ satisfies

1. $U \in N(w) \implies w \in U$
2. $U, V \in N(w) \implies U \cap V \in N(w)$
3. $U \in N(w) \ \& \ U \subset V \implies V \in N(w)$
4. $U \in N(w) \implies \exists V \in N(w) : V \subset U \ \& \ \forall v \in V : U \in N(v)$

Topological space via open sets $(W, \mathcal{O}_W)$ and topological space via neighborhoods (W, N).
We know that $U \subset W$ is open iff $\forall w \in U : U \in N(w)$.
Conversely, $V \in N(w)$ is a neighborhood of w iff $\exists U \in \mathcal{O}_W : w \in U \subset V$.
Remark: Topological semantics is subsumed by neighborhood semantics, being just neighborhood semantics with the above conditions in definition243.

$$[\![A]\!] \in N(w) \iff \exists U \in \mathcal{O}_W : w \in U \subset [\![A]\!] \iff w \in [\![A]\!]^\circ$$

Intension vs Extension

$$[\![P]\!] = \lambda w.[\![P]\!]^w$$

Carnap: Intension as a function from possible worlds to extensions

- The intension of a name is a function from possible worlds to objects.

$$[\![\text{alice}]\!]^w = \text{Alice}$$

- The intension of a predicate is a function from possible worlds to functions from objects to truth-values.

$$[\![\text{smoke}]\!]^w = \lambda x.[\![x \text{ smokes in } w]\!]$$

- The intension of a sentence is a function from possible worlds to truth-values.

	w_1	w_2
$[\![a = a]\!]$	1	1
$[\![a = b]\!]$	1	0

内涵 vs 外延

- 外延同一替换失败:

$$\frac{\begin{array}{c}\text{小明在寻找晨星}\\ \text{晨星} = \text{昏星}\end{array}}{\text{小明在寻找昏星}}$$

- 逻辑等价替换失败:

$$\frac{\begin{array}{c}\text{祖冲之相信 } 1+1=2\\ 1+1=2 \iff \mathrm{e}^{\mathrm{i}\pi}+1=0\end{array}}{\text{祖冲之相信 } \mathrm{e}^{\mathrm{i}\pi}+1=0}$$

需要放弃哪些内涵原则?

1. 复合命题的内涵是其部分之内涵的函数.
2. 如果句子的真值不同, 则其内涵也不同.
3. 如果句子在所有情况下的真值都相同, 则其内涵相同.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Counterfactual Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

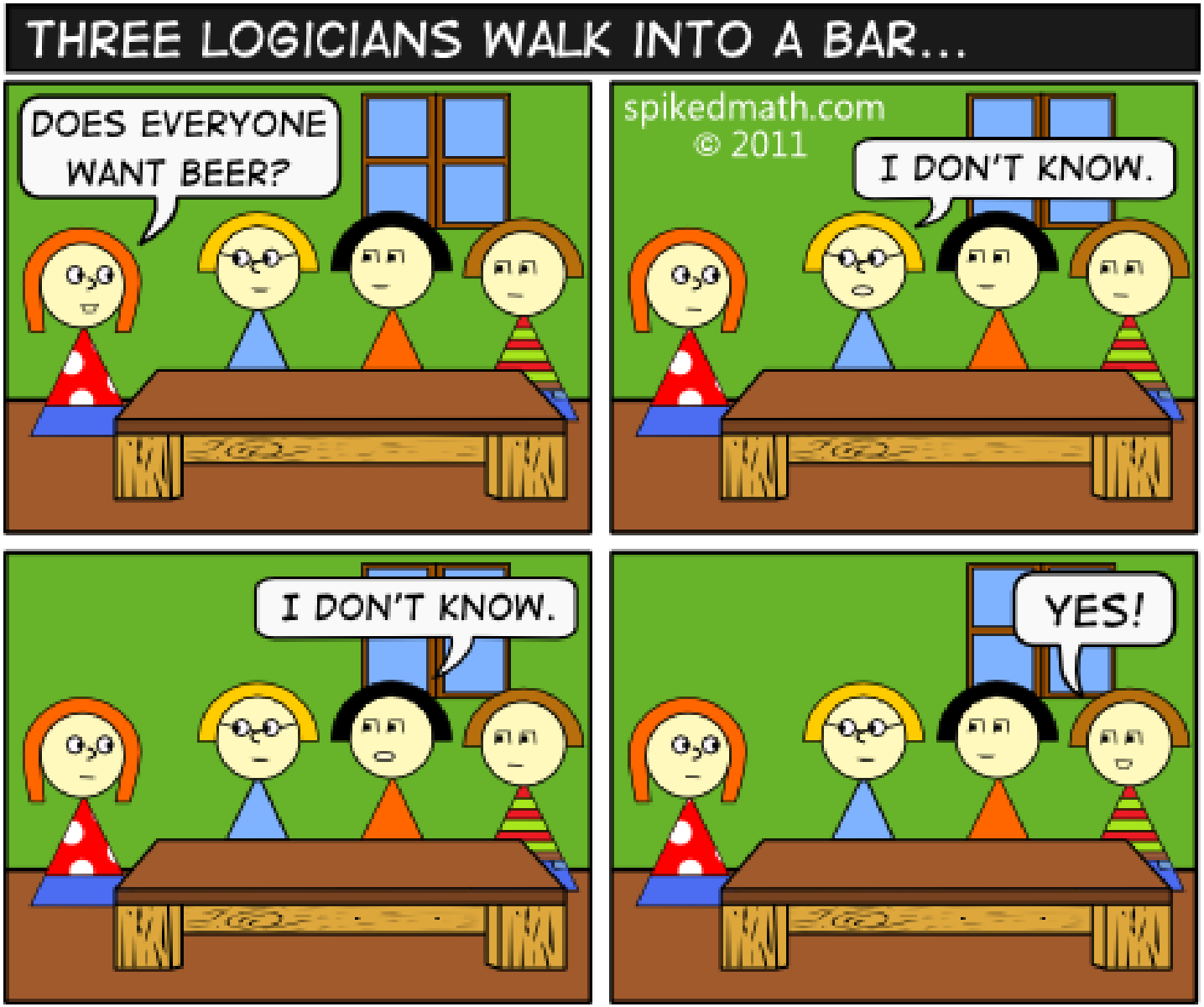
References2039

Jaakko Hintikka



Figure: Jaakko Hintikka: 1929–2015. Epistemic Logic / Game Semantics / Independence-Friendly Logic

Information Update



知识的视角看问题

- 什么是密码？你知我知，但别人不知.
- 微信群是干嘛的？制造公共知识.
- 邮件密送是干嘛的？你知他知，他不知你知，且这是你我的公共知识.
- “代我问他好”是干嘛的？让你知道我尊重他.
- 送什么礼物给太太？我知道她也知道对她有用的.
- 广告语的意义？制造带意义的动作传递知识.
- 《三体》中的黑暗森林法则：爱好和平的公共知识难以达成.
- 如何建设健康学术环境：让他知道你知道学术规范.
- 狼人杀？理性利用别人的不理性.
- 付费知识分享平台：让你相信你知道.
- Would you like to come up to my apartment to see my etchings?
  阿 Q：我想和你困觉!
  Nash: Could we just go straight to the sex? ✓

测试

从 $0 \sim 100$ 之间选一个自然数, 谁的数字最接近平均数的 $\frac{2}{3}$ 谁赢.

从逻辑的视角看 “公共知识”

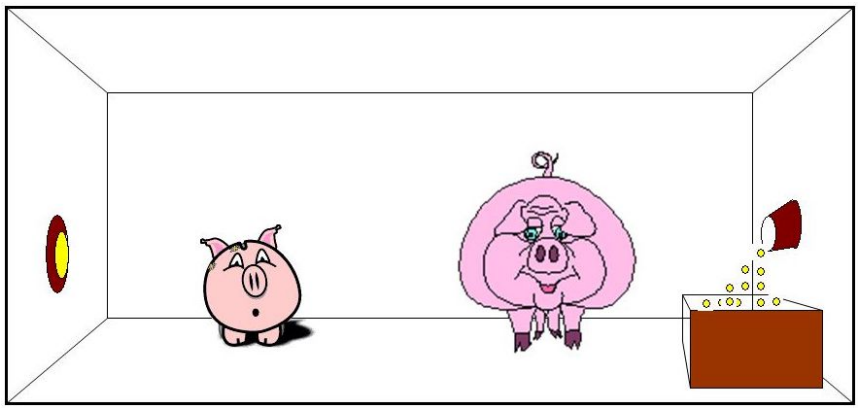
Problem (大女子主义村出轨问题)

1. 大女子主义村村规: 每个发现其老公出轨的女子必须当天枪杀之!
2. 村里住着 100 对夫妇.
3. 男人全都出轨了.
4. 虽然每个女子都知道别人的老公是否出轨, 但不知道自己老公是否出轨. 村里生活和谐幸福.
5. 某天, 女王来访, 临别提醒: “村里至少有一个男人出轨了”.



After a date, one says to the other: “Would you like to come up to my apartment to see my etchings?”

智猪博弈 — 反复剔除劣策略均衡 $\neq$ 占优均衡



- 小猪是理性的

$$\mathsf{Rational}(s)$$

- 大猪知道小猪是理性的

$$\mathsf{K}_b\mathsf{Rational}(s)$$

	按	等
按	$4,0$	$3,3$
等	$7,-1$	$0,0$

	按	等
按		$3,3$
等		$0,0$

	按	等
按		$3,3$
等		

反复剔除劣策略与 n-阶理性

	c_1	c_2	c_3	c_4
r_1	$5, 10$	$0, 11$	$1, 20$	$10, 10$
r_2	$4, 0$	$1, 1$	$2, 0$	$20, 0$
r_3	$3, 2$	$0, 4$	$4, 3$	$50, 1$
r_4	$2, 93$	$0, 92$	$0, 91$	$100, 90$

	c_1	c_2	c_3
r_1	$5, 10$	$0, 11$	$1, 20$
r_2	$4, 0$	$1, 1$	$2, 0$
r_3	$3, 2$	$0, 4$	$4, 3$
r_4	$2, 93$	$0, 92$	$0, 91$

	c_1	c_2	c_3
r_1	$5, 10$	$0, 11$	$1, 20$
r_2	$4, 0$	$1, 1$	$2, 0$
r_3	$3, 2$	$0, 4$	$4, 3$

	c_2	c_3
r_1	$0, 11$	$1, 20$
r_2	$1, 1$	$2, 0$
r_3	$0, 4$	$4, 3$

	c_2	c_3
r_2	$1, 1$	$2, 0$
r_3	$0, 4$	$4, 3$

	c_2
r_2	$1, 1$
r_3	$0, 4$

	c_2
r_2	$1, 1$

0-order $c_4\times$ $\mathsf{Rational}(c)$

1-order $r_4\times$ $\mathsf{K}_r\mathsf{Rational}(c)$

2-order $c_1\times$ $\mathsf{K}_c\mathsf{K}_r\mathsf{Rational}(c)$

3-order $r_1\times$ $\mathsf{K}_r\mathsf{K}_c\mathsf{K}_r\mathsf{Rational}(c)$

4-order $c_3\times$ $\mathsf{K}_c\mathsf{K}_r\mathsf{K}_c\mathsf{K}_r\mathsf{Rational}(c)$

5-order $r_3\times$ $\mathsf{K}_r\mathsf{K}_c\mathsf{K}_r\mathsf{K}_c\mathsf{K}_r\mathsf{Rational}(c)$

诸葛亮的《空城计》何以有效?



- 诸葛亮谨慎. $C(z)$
- 司马懿知道诸葛亮谨慎. $\mathsf{K}_s C(z)$
- 诸葛亮知道司马懿知道诸葛亮谨慎. $\mathsf{K}_z \mathsf{K}_s C(z)$
- 司马懿不知道诸葛亮知道司马懿知道诸葛亮谨慎. $\neg \mathsf{K}_s \mathsf{K}_z \mathsf{K}_s C(z)$

老狐狸与小狐狸

- 老狐狸看到满载而归的渔夫驾车经过，于是躺在路边装病.
- 渔夫看老狐狸可怜，就让它搭个便车.
- 老狐狸悄悄地把鱼一条一条地扔到路边草丛里，然后一跃而下吃鱼去了.
- 小狐狸问老狐狸是如何得到这么多鱼的，老狐狸“如实相告”.
- 第二天，小狐狸也学着躺在路边装病，渔夫愤怒地把它打死了.

- $\neg \mathsf{K}_{渔} A$
- $\mathsf{K}_{老} \neg \mathsf{K}_{渔} A$
- $[A] \mathsf{K}_{渔} A$
- $\mathsf{K}_{老} [A] \mathsf{K}_{渔} A$
- $\neg \mathsf{K}_{小} A$
- $[A] \mathsf{K}_{小} A$
- $\mathsf{K}_{老} [A] \mathsf{K}_{小} A$
- $\neg \mathsf{K}_{小} [A] \mathsf{K}_{渔} A$
- $\mathsf{K}_{老} [A] \neg \mathsf{K}_{小} [A] \mathsf{K}_{渔} A$

知识就是力量

- Knowledge is power: act properly to achieve goals;
- Knowledge is time: to make decisions more efficiently;
- Knowledge is money: can be traded;
- Knowledge is responsibility: to prove someone is guilty;
- Knowledge is you: to identify oneself;
- Knowledge is an immune system: to protect you;
- Knowledge satisfies our curiosity.

The only good is knowledge and the only evil is ignorance.

— *Socrates*

ALL men by nature desire to know.

— *Aristotle*

The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge.

— *Stephen Hawking*

知之为知之

know the unknown from the known

- There are things we know we know. There are things we know we don't know. There are things we don't know we don't know.

$$\exists x \mathsf{K}\mathsf{K}x \wedge \exists x \mathsf{K}\neg\mathsf{K}x \wedge \exists x \neg\mathsf{K}\neg\mathsf{K}x$$

$$\neg\mathsf{K}\neg\mathsf{K}p \rightsquigarrow \mathsf{K}\neg\mathsf{K}p \rightsquigarrow \neg\mathsf{K}\mathsf{K}p \rightsquigarrow \mathsf{K}\mathsf{K}p$$

- 知之为知之, 不知为不知, 是知也.

$$\mathsf{K}p \to \mathsf{K}\mathsf{K}p \quad \& \quad \neg\mathsf{K}p \to \mathsf{K}\neg\mathsf{K}p$$

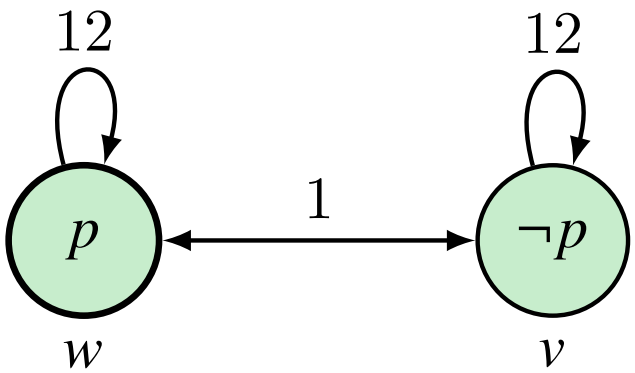
"Real knowledge is to know the extent of one's ignorance."

— 孔子

- 知其然, 知其所以然

$$\mathsf{K}\exists x Ax \to \exists x \mathsf{K}Ax$$

$$\mathcal{M}, w \vDash \mathsf{K}_i A \quad \text{iff} \quad \forall v \in W \Big(R_i w v \implies \mathcal{M}, v \vDash A \Big)$$



$w \vDash p$

$w \vDash \neg \mathsf{K}_1 p$

$v \vDash \neg \mathsf{K}_1 p$

$w \vDash \mathsf{K}_2 p$

$v \vDash \mathsf{K}_2 \neg p$

$w \vDash \mathsf{K}_1 (\mathsf{K}_2 p \vee \mathsf{K}_2 \neg p)$

$w \vDash \mathsf{K}_2 \neg \mathsf{K}_1 p$

$v \vDash \mathsf{K}_2 \neg \mathsf{K}_1 p$

$w \vDash \mathsf{K}_1 \mathsf{K}_2 \neg \mathsf{K}_1 p$

$$w \vDash p \wedge \neg \mathsf{K}_1 p \wedge \mathsf{K}_2 p \wedge \mathsf{K}_1 (\mathsf{K}_2 p \vee \mathsf{K}_2 \neg p) \wedge \mathsf{K}_2 \neg \mathsf{K}_1 p \wedge \mathsf{K}_1 \mathsf{K}_2 \neg \mathsf{K}_1 p$$

假设纽约在下雨 (p), 但 1 不知道, 而 2 知道, 不过 1 知道 2 知道纽约是否在下雨 (因为 2 住在纽约).....

- Mutual Knowledge:

  everybody in G knows p.

- Distributed Knowledge:

  everybody in G would know p
  if agents in G shared all their knowledge.

- Common Knowledge:

  everybody in G knows p,
  everybody knows that everybody knows,
  and so on.

Mutual Knowledge

Suppose a group $G \subset \{1 \dots n\}$ of agents, everyone in G knows A:

$$\mathsf{E}_G A \;:=\; \bigwedge_{i \in G} \mathsf{K}_i A$$

$$R_E \;:=\; \bigcup_{i \in G} R_i$$

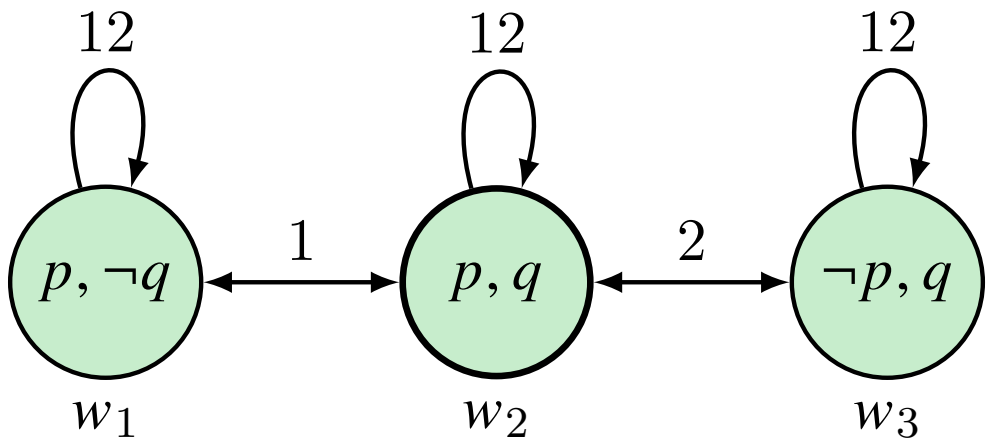
$$\mathcal{M}, w \vDash \mathsf{E}_G A \quad \text{iff} \quad \forall v \in W \Big(R_E w v \implies \mathcal{M}, v \vDash A \Big)$$

Distributed Knowledge

Intuition: what we know if we put all of our knowledge together.

$$R_D := \bigcap_{i \in G} R_i$$

$$\mathcal{M}, w \vDash \mathsf{D}_G A \quad \text{iff} \quad \forall v \in W \Big(R_D w v \implies \mathcal{M}, v \vDash A \Big)$$



$$w_2 \vDash \mathsf{K}_1 p \wedge \neg \mathsf{K}_1 q \wedge \mathsf{K}_2 q \wedge \neg \mathsf{K}_2 p \wedge \mathsf{D}_{\{1,2\}}(p \wedge q)$$

$$\mathsf{D}_G A \not\equiv \bigvee_{i \in G} \mathsf{K}_i A$$

Common Knowledge

$$\mathsf{E}_G^1 A := \mathsf{E}_G A$$
$$\mathsf{E}_G^{k+1} A := \mathsf{E}_G \mathsf{E}_G^k A$$
$$\mathsf{C}_G A := \bigwedge_{k=1}^{\infty} \mathsf{E}_G^k A$$

$$R^1 := R$$
$$R^{k+1} := R \circ R^k$$
$$R \circ S := \{(x, y) : \exists z (Rxz \wedge Szy)\}$$
$$R^* := \bigcup_{k=1}^{\infty} R^k$$

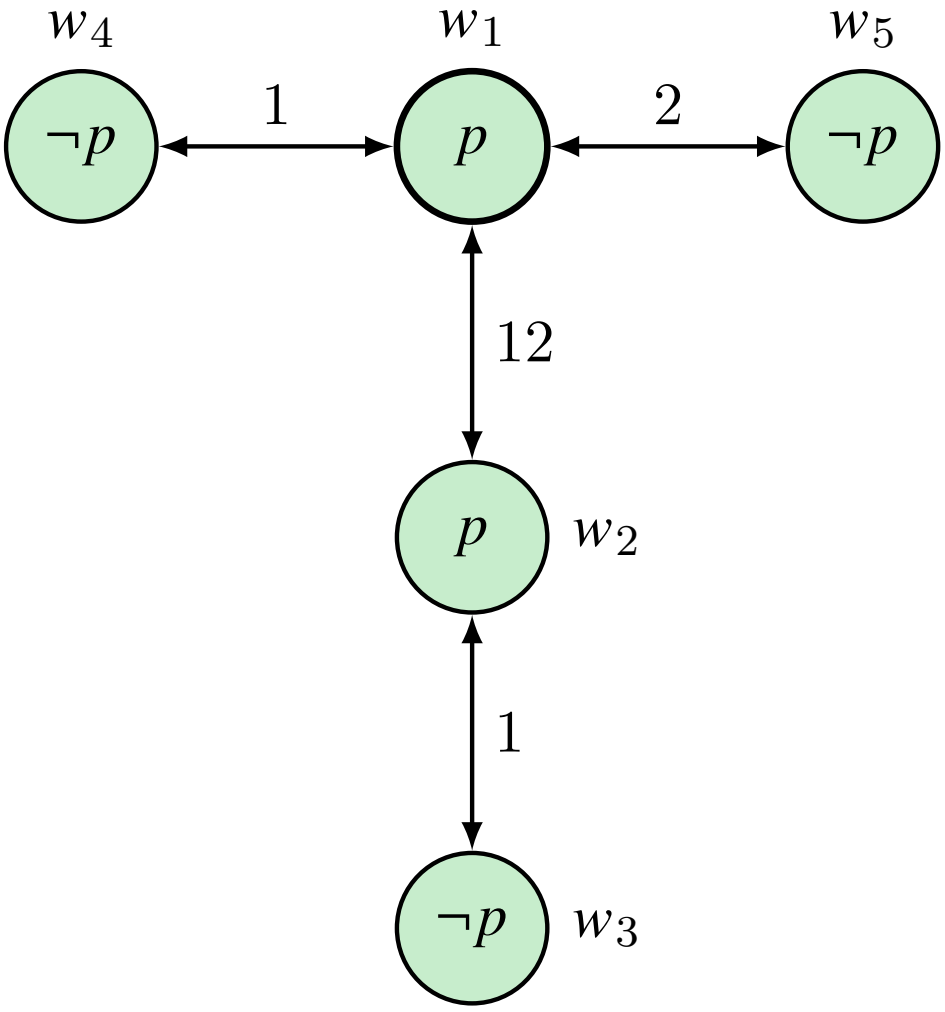
$$R_C := \left(\bigcup_{i \in G} R_i \right)^*$$

$$\mathcal{M}, w \vDash \mathsf{C}_G A \text{ iff } \forall v \in W \left(R_C w v \implies \mathcal{M}, v \vDash A \right)$$

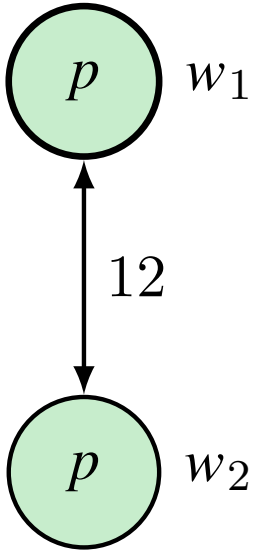
A Hierarchy of States of Knowledge

$$\mathsf{C}_G A \implies \cdots \mathsf{E}_G^k A \implies \cdots \mathsf{E}_G A \implies \bigvee_{i \in G} \mathsf{K}_i A \implies \mathsf{D}_G A \implies A$$

分布式知识的定义所面临的问题[24]



$w_1 \vDash \mathsf{D}_{12}(p \wedge \neg \mathsf{K}_1 p)$



"Communication Core"

$w_1 \vDash \mathsf{D}_{12}(p \wedge \mathsf{K}_1 p)$

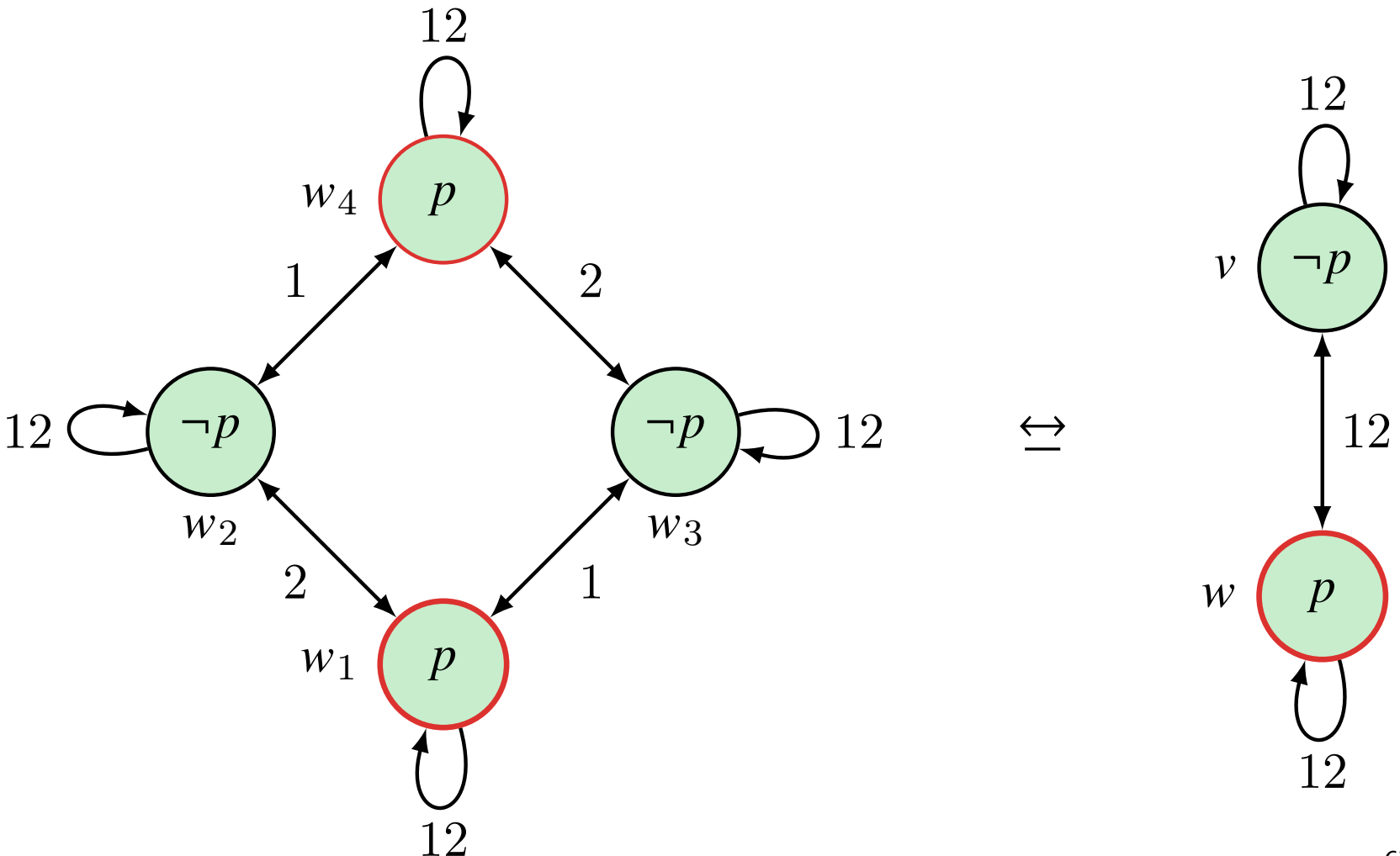
Remark: Communication turns distributed knowledge into common knowledge.

[24] "Communication Core" consists of the actual world plus all worlds linked to it by the intersection of all uncertainty relations.

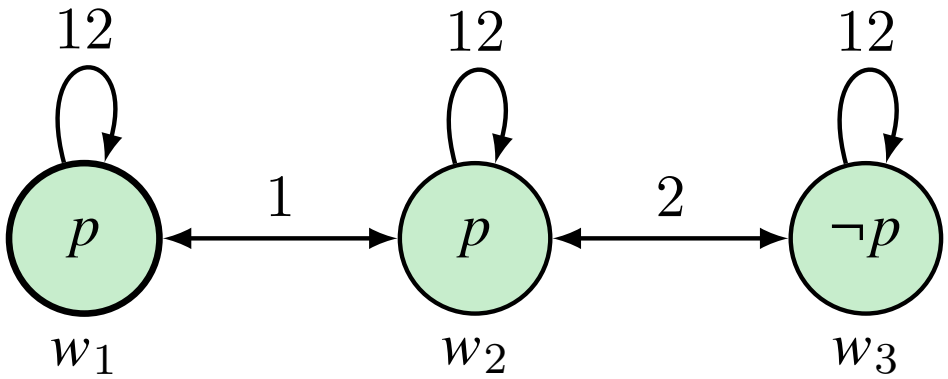
分布式知识的定义所面临的问题

- It is not invariant under bisimulation.
- It is not the case:

$$\{B : \mathcal{M}, w \vDash \mathsf{K}_i B \text{ for some } i \in G\} \vDash A \iff \mathcal{M}, w \vDash \mathsf{D}_G A$$



Can we easily have full common knowledge?



$$w_1 \vDash \mathsf{E}_{\{1,2\}} p \wedge \neg \mathsf{C}_{\{1,2\}} p$$

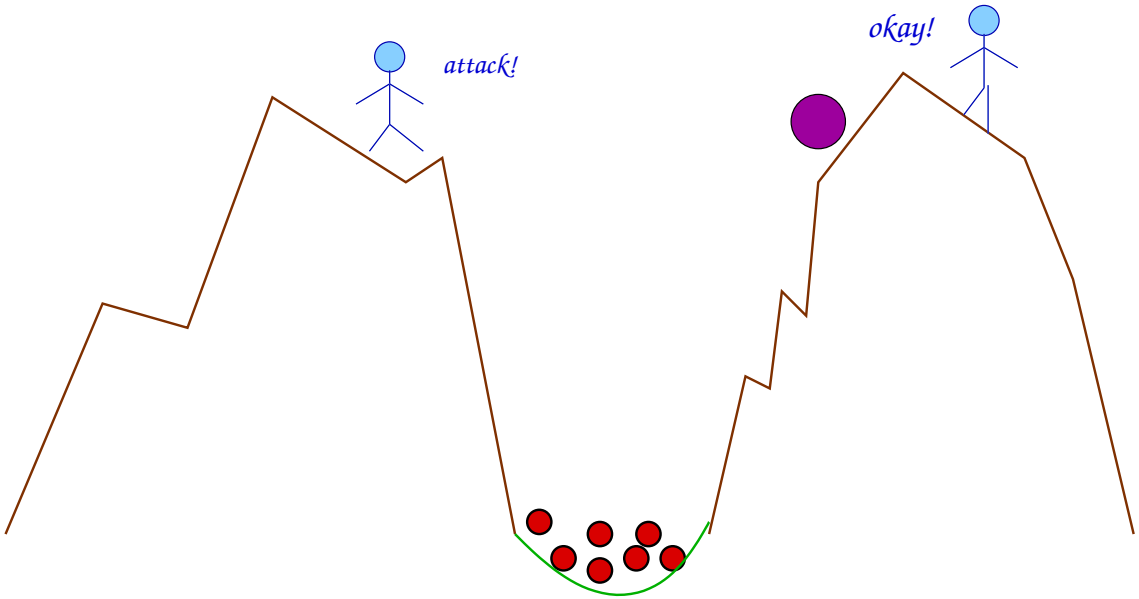
假设你秘密的分别给了 a 和 b 两个数字 2 和 3, 你只告诉他们俩这两个数字是相邻的自然数. 令 p 为 "两数字之和小于一千万", 请问 p 是 a 和 b 的公共知识么?

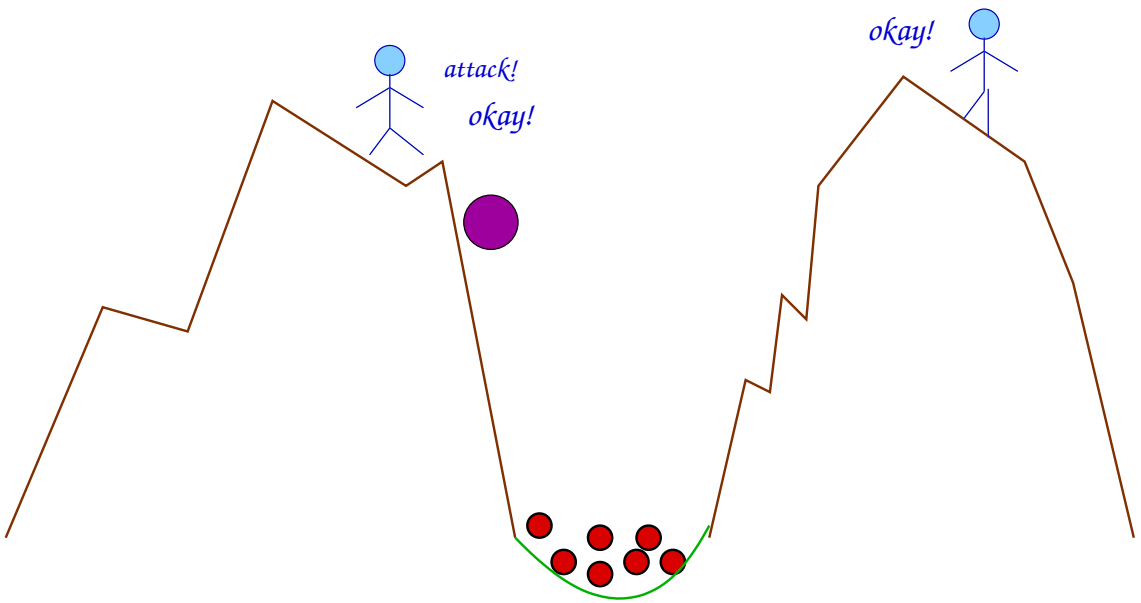
$$(0,1) \overset{b}{\longleftrightarrow} (2,1) \overset{a}{\longleftrightarrow} \underline{(2,3)} \overset{b}{\longleftrightarrow} (4,3) \overset{a}{\longleftrightarrow} (4,5) \overset{b}{\longleftrightarrow} (6,5) \cdots$$

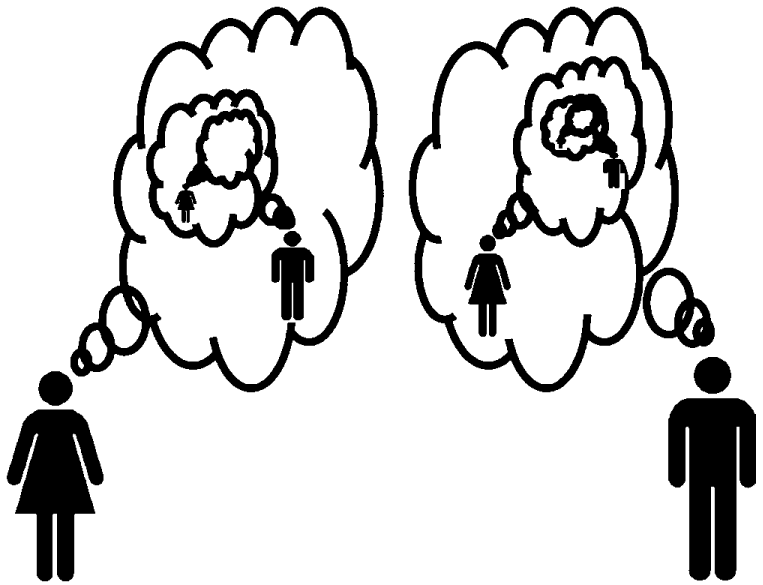
$$(2,3) \vDash \neg \mathsf{K}_b \mathsf{K}_a \mathsf{K}_b (x + y \leq 10)$$

a and b commonly know that b's number is odd.

拜占庭将军协同进攻问题 Coordinated Attack



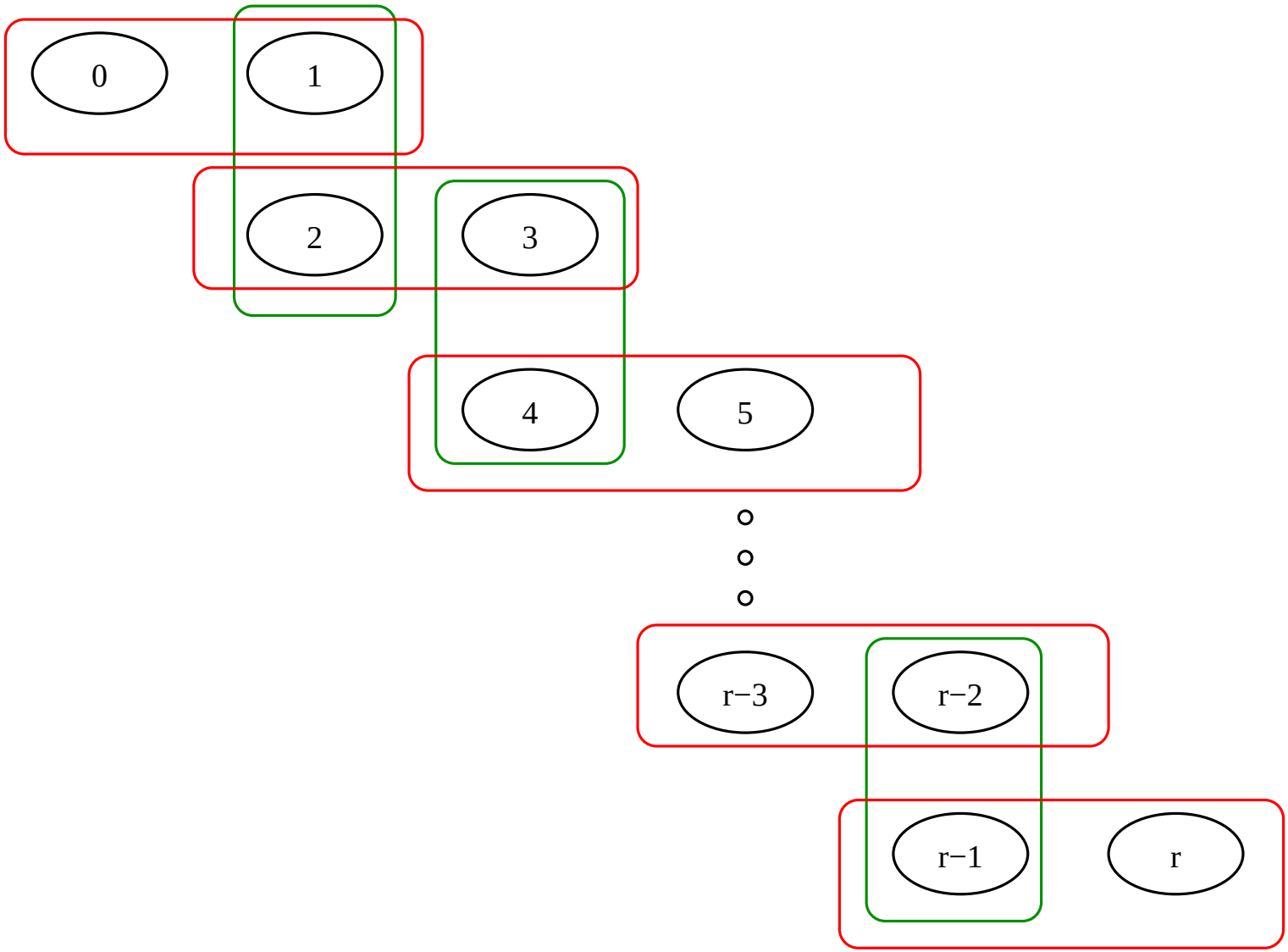




A: 如果我知道 B 一定进攻, 我就进攻!

B: 如果我知道 A 一定进攻, 我就进攻!

拜占庭将军协同进攻问题 Coordinated Attack



《三体》— 黑暗森林

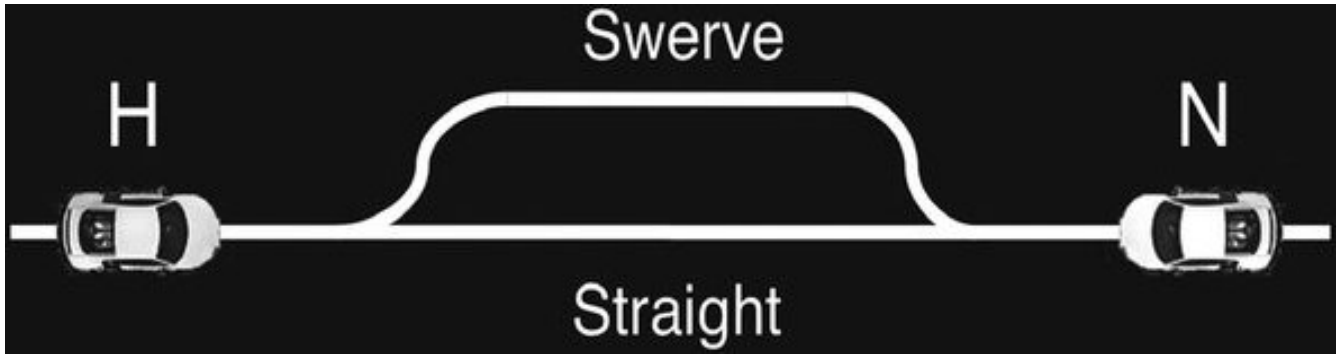
- 公理一：生存是文明的第一需要.
- 公理二：文明不断增长和扩张，但宇宙中的物质总量保持不变.

猜疑链

如果你认为我是善意的，这并不是你感到安全的理由，因为按照第一条公理，善意文明并不能预先把别的文明也想成善意的，所以，你现在还不知道我是怎么认为你的，你不知道我认为你是善意还是恶意；进一步，即使你知道我把你也想象成善意的，我也知道你把我想象成善意的，但是我不知道你是怎么想我怎么想你怎么想我的……

黑暗森林

宇宙就是一座黑暗森林，每个文明都是带枪的猎人，像幽灵般潜行于林间……如果他发现了别的生命，不管是不是猎人，不管是天使还是魔鬼，不管是娇嫩的婴儿还是步履蹒跚的老人，也不管是天仙般的少女还是天神般的男神，能做的只有一件事：开枪消灭之．在这片森林中，他人就是地狱，就是永恒的威胁 (技术爆炸)，任何暴露自己存在的生命都将很快被消灭．这就是宇宙文明的图景，这就是对费米悖论的解释.

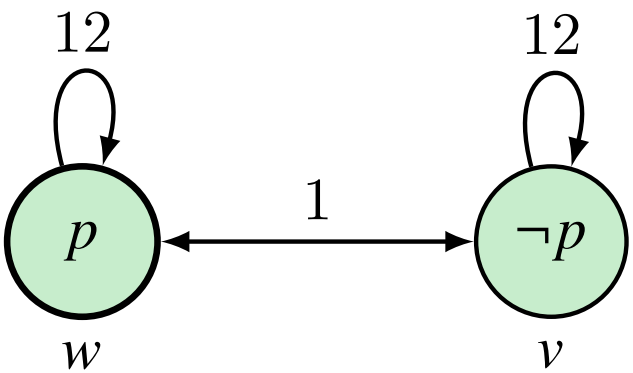


p: "Every driver must drive on the right."

- What kind of knowledge is enough to let people feel safe in driving on the right?
- Is 'everybody knows that p' enough ($\mathsf{E}p$)?
- What about everybody knows that everybody knows that p ($\mathsf{EE}p$)?
- No, since I consider possible that You consider possible that I do not know ($\neg \mathsf{K}_i \mathsf{K}_j \mathsf{K}_i p$), and thus You may drive on the left.

$$\mathsf{C}A = \mathsf{E}A \wedge \mathsf{EE}A \wedge \mathsf{EEE}A \ldots$$

Example — A Question/Answer Scenario



$w \models p$

$w \models \mathsf{K}_2 p$

$w \models \neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1 \neg p$

$w \models \mathsf{K}_1(\mathsf{K}_2 p \vee \mathsf{K}_2 \neg p)$

$w \models \mathsf{K}_2(\mathsf{K}_2 p \vee \mathsf{K}_2 \neg p)$

$w \models \mathsf{E}_{\{1,2\}}(\mathsf{K}_2 p \vee \mathsf{K}_2 \neg p)$

$w \models \mathsf{K}_1(\neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1 \neg p)$

$w \models \mathsf{K}_2(\neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1 \neg p)$

$w \models \mathsf{E}_{\{1,2\}}(\neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1 \neg p)$

$w \models \mathsf{C}_{\{1,2\}}(\mathsf{K}_2 p \vee \mathsf{K}_2 \neg p)$

$w \models \mathsf{C}_{\{1,2\}}(\neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1 \neg p)$

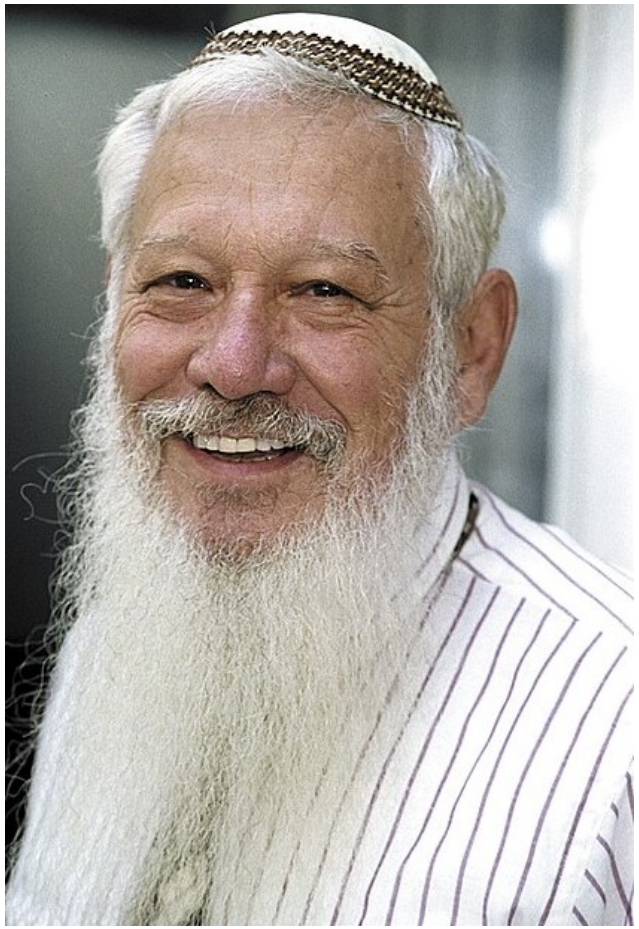
Remark: This is an excellent situation for 1 to ask 2 whether p is the case.

Aumann's Agreement Theorem

Theorem (Aumann's Agreement Theorem)

If two Bayesian rationalists have the same priors, and their posteriors are common knowledge, then these posteriors are equal.

如果两个人的先验知识相同, 则他们不可能对有分歧的后验 (经过各自的实验获取私人证据) 形成公共知识. 经过足够的交流和更新 (交流后验信念本身, 而不是其支撑证据), 大家不可能保留有分歧的后验 (agree to disagree), 最终总会达成共识.



Aumann's Agreement Theorem

$(W, \{\mathcal{I}_i\}_{i \in G}, \{\mathsf{K}_i\}_{i \in G})$.

- W is a non-empty set of worlds.
- $\mathcal{I}_i$ is agent i's partition of W. $\mathcal{I}_i(w)$ is the element of the partition that contains w.
- $\mathsf{K}_i : \mathsf{P}(W) \to \mathsf{P}(W)$ is agent i's knowledge operator. $\mathsf{K}_i(A) = \{w : \mathcal{I}_i(w) \subset A\}$.
- Mutual Knowledge $\mathsf{E}_G(A) := \bigcap_{i \in G} \mathsf{K}_i(A)$.
- Common knowledge $\mathsf{C}_G(A) := \bigcap_{n=1}^{\infty} \mathsf{E}_G^n(A)$.

1. $\mathsf{K}(W) = W$
2. $A \subset B \implies \mathsf{K}(A) \subset \mathsf{K}(B)$
3. $\mathsf{K}(A) \cap \mathsf{K}(B) = \mathsf{K}(A \cap B)$
4. $\mathsf{K}(A) \subset A$
5. $\mathsf{K}(A) \subset \mathsf{K}(\mathsf{K}(A))$
6. $\overline{\mathsf{K}(A)} \subset \mathsf{K}(\overline{\mathsf{K}(A)})$

Aumann's Agreement Theorem

Lemma

If $\mathsf{C}_G(A) \neq \emptyset$, *then* $\forall i \in G \exists \mathcal{D}_i \subset \mathcal{I}_i : \mathsf{C}_G(A) = \bigcup \mathcal{D}_i$.

Proof.

$w \in \mathsf{C}_G(A) \implies \forall i \forall n : w \in \mathsf{K}_i \mathsf{E}_G^n(A) \implies \forall i \forall n : \mathcal{I}_i(w) \subset \mathsf{E}_G^n(A) \implies \forall i : \mathcal{I}_i(w) \subset \mathsf{C}_G(A)$ □

Theorem (Aumann's Agreement Theorem)

Let P *be the common prior belief, and* $B := \bigcap_{i \in G} \{w : P(A \mid \mathcal{I}_i(w)) = q_i\}$. *If* $P(\mathsf{C}_G(B)) > 0$, *then* $\forall i \in G : q_i = P(A \mid \mathsf{C}_G(B))$.

Proof.

$$P(A \mid \mathsf{C}_G(B)) = \frac{P(A \cap \bigcup \mathcal{D}_i)}{\sum_{D \in \mathcal{D}_i} P(D)} = \frac{\sum_{D \in \mathcal{D}_i} P(A \mid D) P(D)}{\sum_{D \in \mathcal{D}_i} P(D)} = \frac{\sum_{D \in \mathcal{D}_i} q_i P(D)}{\sum_{D \in \mathcal{D}_i} P(D)} = q_i$$

□

Epistemic Logic — Formal Systems

- Knowledge $\mathrm{S5} := K + T + 4 + 5$

  知之为知之(4), 不知为不知(5), 是知也

- Belief $K + D + 4 + 5$
- Common Knowledge

$$\mathrm{S5} \quad \text{for} \quad \mathsf{C}_G$$

$$+$$

$$\mathsf{C}_G A \leftrightarrow A \wedge \mathsf{E}_G \mathsf{C}_G A$$

$$+$$

$$A \wedge \mathsf{C}_G (A \rightarrow \mathsf{E}_G A) \rightarrow \mathsf{C}_G A$$

- Distributed Knowledge

$$\mathrm{S5} \quad \text{for} \quad \mathsf{D}_G$$

$$+$$

$$\mathsf{K}_i A \rightarrow \mathsf{D}_G A \quad \text{when} \quad i \in G$$

$$+$$

$$\mathsf{D}_G A \rightarrow \mathsf{D}_{G'} A \quad \text{when} \quad G \subset G'$$

知识 vs 信念 Knowledge vs Belief

1. $\mathsf{K}(p \to q) \to \mathsf{K}p \to \mathsf{K}q$
2. $\mathsf{K}p \to p$
3. $\mathsf{K}p \to \mathsf{KK}p$
4. $\neg\mathsf{K}p \to \mathsf{K}\neg\mathsf{K}p$

1. $\mathsf{K}p \to \mathsf{B}p$
2. $\mathsf{B}p \to \mathsf{KB}p$
3. $\neg\mathsf{B}p \to \mathsf{K}\neg\mathsf{B}p$
4. $\mathsf{B}p \to \mathsf{BK}p$ (strong belief)

1. $\mathsf{B}(p \to q) \to \mathsf{B}p \to \mathsf{B}q$
2. $\mathsf{B}p \to \neg\mathsf{B}\neg p$
3. $\mathsf{B}p \to \mathsf{BB}p$
4. $\neg\mathsf{B}p \to \mathsf{B}\neg\mathsf{B}p$

$$\mathsf{B}p \overset{?}{\leftrightarrow} \neg\mathsf{K}\neg\mathsf{K}p$$

摩尔悖论 Moore's Paradox

Are all truths knowable?

It's rainning but I don't know it's raining.

$$\vdash \neg \mathsf{K}(p \wedge \neg \mathsf{K}p)$$

1. $\mathsf{K}(p \wedge \neg \mathsf{K}p)$ Assumption
2. $\mathsf{K}p$
3. $\mathsf{K}\neg \mathsf{K}p$
4. $\neg \mathsf{K}p$ $\mathsf{K}p \rightarrow p$
5. $\mathsf{K}p \wedge \neg \mathsf{K}p$
6. $\neg \mathsf{K}(p \wedge \neg \mathsf{K}p)$

摩尔悖论 Moore's Paradox

Are all truths believable?

It's raining but I don't believe it's raining.

$$\vdash \neg \mathsf{B}(p \wedge \neg \mathsf{B}p)$$

1.	$\mathsf{B}(p \wedge \neg \mathsf{B}p)$	Assumption
2.	$\mathsf{B}p \wedge \mathsf{B}\neg \mathsf{B}p$	$\mathsf{B}(p \wedge q) \rightarrow \mathsf{B}p \wedge \mathsf{B}q$
3.	$\mathsf{B}p$	
4.	$\mathsf{BB}p$	$\mathsf{B}p \rightarrow \mathsf{BB}p$
5.	$\mathsf{B}\neg \mathsf{B}p$	
6.	$\mathsf{BB}p \wedge \mathsf{B}\neg \mathsf{B}p$	
7.	$\neg(\mathsf{BB}p \wedge \mathsf{B}\neg \mathsf{B}p)$	$(\mathsf{B}p \rightarrow \neg \mathsf{B}\neg p) \leftrightarrow \neg(\mathsf{B}p \wedge \mathsf{B}\neg p)$
8.	$\neg \mathsf{B}(p \wedge \neg \mathsf{B}p)$	

Against Negative Introspection?

1. $\neg p \wedge \mathsf{BK}p$ — suppose you falsely believes that you know p
2. $\neg \mathsf{K}p$ — knowledge implies truth
3. $\mathsf{K}\neg \mathsf{K}p$ — negative introspection
4. $\mathsf{B}\neg \mathsf{K}p$ — knowledge implies belief
5. $\mathsf{B}\bot$

$$\neg \mathsf{K}p \rightarrow \mathsf{K}\neg \mathsf{K}p \ ? \qquad \text{(Ax 5)}$$

$$\neg \mathsf{K}\neg \mathsf{K}p \rightarrow \mathsf{K}p \ ? \qquad \text{(Ax 5)}$$

$$\neg \mathsf{K}\neg \mathsf{K}p \rightarrow \mathsf{K}\neg \mathsf{K}\neg p \ ? \qquad \text{(Ax 4.2)}$$

Margin for Error and Against Positive Introspection?

p_n : n grains is a heap.

1. $\mathsf{K}\neg p_n$ — Assumption
2. $\mathsf{K}(p_{n+1} \to \neg\mathsf{K}\neg p_n)$ — Margin for Error
3. $\mathsf{K}(\mathsf{K}\neg p_n \to \neg p_{n+1})$
4. $\mathsf{K}\neg p_n \to \mathsf{K}\mathsf{K}\neg p_n$ — Positive Introspection
5. $\mathsf{K}\neg p_{n+1}$ — ?

$$\mathsf{K}p \to \mathsf{K}\mathsf{K}p \ ?$$

Remark: 谷堆悖论是连锁悖论, 下面这个是连锁悖论吗: 不存在哺乳动物, 因为哺乳动物的母亲也必须是哺乳动物.

逻辑全知 vs 怀疑论

$$\frac{A}{\mathsf{K}A} \qquad \frac{A \to B}{\mathsf{K}A \to \mathsf{K}B} \qquad \frac{A \leftrightarrow B}{\mathsf{K}A \leftrightarrow \mathsf{K}B} \qquad \frac{\mathsf{K}(A \to B) \quad \mathsf{K}A}{\mathsf{K}B}$$

S_1 I know "I have hands". $\mathsf{K}p$

S_2 I know if "I have hands" then "I am not a brain in the vat". $\mathsf{K}(p \to q)$

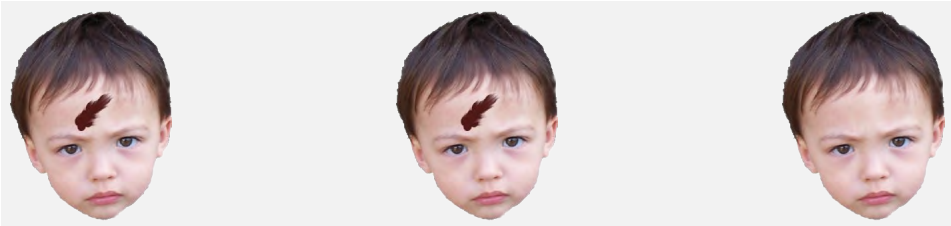
S_3 I don't know that "I am not a brain in the vat". $\neg\mathsf{K}q$

Possible "solutions":

- Impossible worlds: inconsistent alternatives
- Awareness: K = awareness + implicit knowledge
- Algorithmic knowledge: K = answer by algorithm
- Timed knowledge: reasoning takes time
- Neighbourhood semantics: still problematic
- Counterfactual implication:
  - Nozick: $\mathsf{K}p := p \wedge \mathsf{B}p \wedge (\neg p \mathrel{\Box\!\!\to} \neg \mathsf{B}p) \wedge (p \mathrel{\Box\!\!\to} \mathsf{B}p)$
  - Sosa: $\mathsf{K}p := p \wedge \mathsf{B}p \wedge (\mathsf{B}p \mathrel{\Box\!\!\to} p) \wedge (p \mathrel{\Box\!\!\to} \mathsf{B}p)$

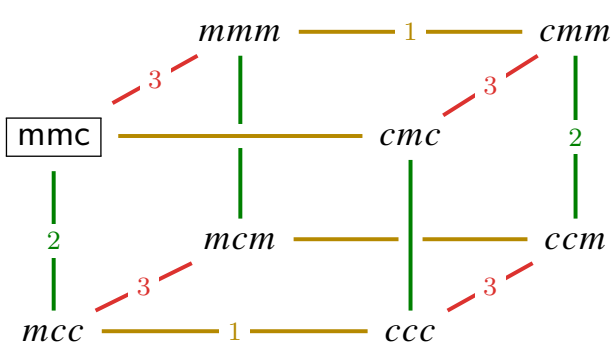
knowing-whether	$\mathsf{K}A \vee \mathsf{K}\neg A$	我知道股票是否会涨
knowing-what	$\exists x \mathsf{K}(A \to x = c)$	我知道彩票的中奖号码是什么
knowing-how	$\exists \alpha \mathsf{K}[\alpha]A$	我知道如何实现一个小目标
knowing-why	$\exists t \mathsf{K}(t : A)$	我知道为什么她生气了

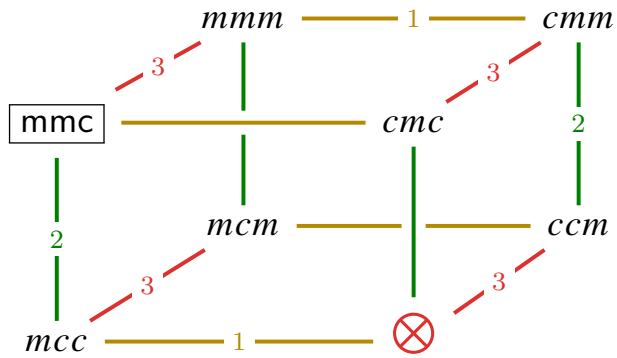
Information Update — Muddy Children Problem

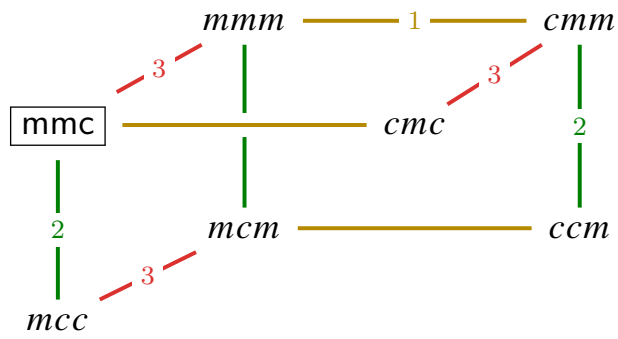


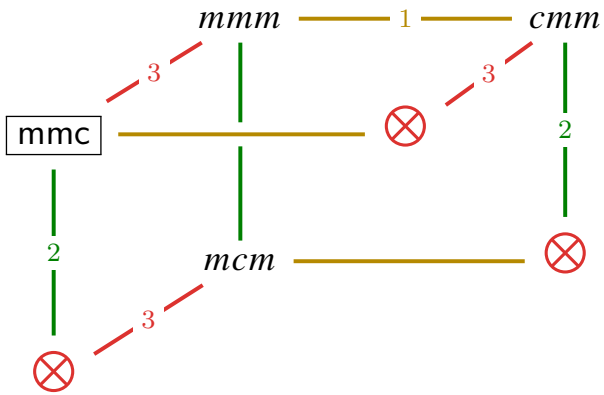
Problem (泥孩问题)

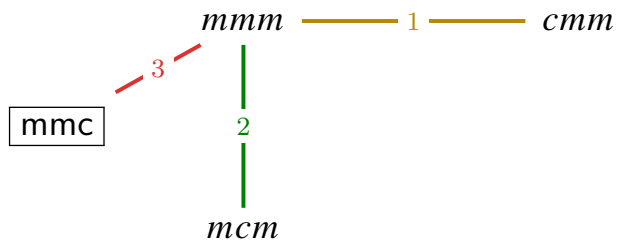
- n 个小朋友里的 k 个脸上弄上了泥巴.
- 不过他们只能看到别人的脸, 而不知道自己脸上是否有泥巴.
- 老师看到了很生气: "你们中间有人把泥巴弄到脸上了!"
- 然后老师命令道: "知道自己脸上有泥巴的给我站出来!"
- 没有人站出来.
- 老师重复道: "知道自己脸上有泥巴的给我站出来!"
- .....

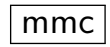












1. “有人把泥巴弄到脸上了！知道自己脸上有泥巴的给我站出来!”

$$\neg \mathsf{K}_1 m_1 \wedge \neg \mathsf{K}_1 \neg m_1 \wedge \neg \mathsf{K}_2 m_2 \wedge \neg \mathsf{K}_2 \neg m_2 \wedge \neg \mathsf{K}_3 m_3 \wedge \neg \mathsf{K}_3 \neg m_3$$

2. 没有人站出来. “知道自己脸上有泥巴的给我站出来!”

$$\mathsf{K}_1 m_1 \wedge \mathsf{K}_2 m_2 \wedge \neg \mathsf{K}_3 m_3 \wedge \neg \mathsf{K}_3 \neg m_3$$

3. 1 和 2 站了出来.

$$\mathsf{K}_3 \neg m_3$$

公开宣告逻辑 Public Announcement Logic (PAL)

$$A := p \mid \neg A \mid A \wedge A \mid \mathsf{K}_i A \mid [A]A$$

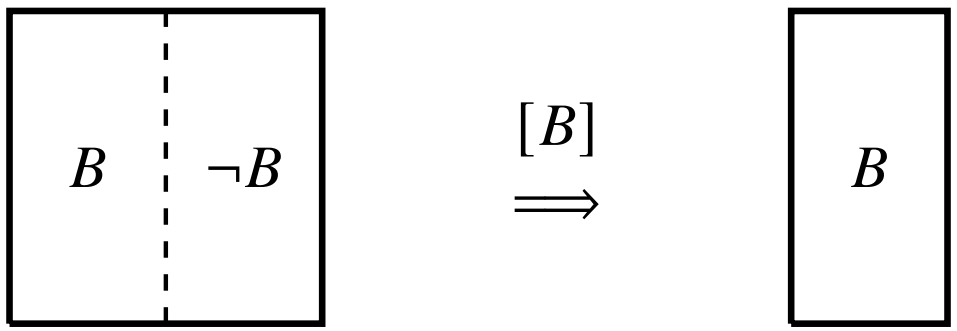
$$\mathcal{M}, w \vDash [B]A \;\text{ iff }\; \mathcal{M}, w \vDash B \implies \mathcal{M}{\restriction_B}, w \vDash A$$
$$\mathcal{M}, w \vDash \langle B\rangle A \;\text{ iff }\; \mathcal{M}, w \vDash B \;\&\; \mathcal{M}{\restriction_B}, w \vDash A$$

where

$$\mathcal{M}{\restriction_B} := \left(W', \{R'_i\}_{i\in G}, V'\right)$$

and

$$W' := \{w \in W : \mathcal{M}, w \vDash B\} \quad R'_i := R_i{\restriction_{W'\times W'}} \quad V'(p) := V(p) \cap W'$$



Remark: the meaning of an action is the change it brings to the states!

泥孩问题

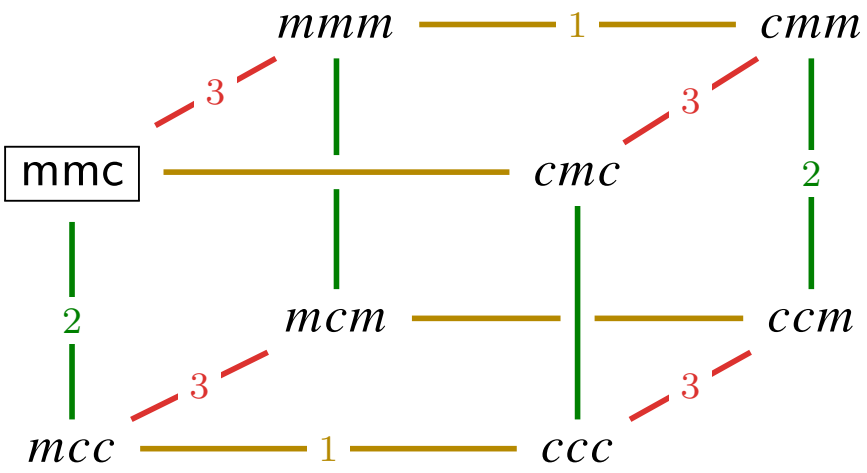
$\mathcal{M}, mmc \vDash m_1 \wedge m_2 \wedge \neg m_3$

$\mathcal{M}, mmc \vDash \mathsf{E}_{\{1,2,3\}} P$

$\mathcal{M}, mmc \vDash \neg \mathsf{C}_{\{1,2,3\}} P$

$\mathcal{M}, mmc \vDash \neg \mathsf{K}_1 m_1 \wedge \mathsf{K}_1 m_2$

$\mathcal{M}, mmc \vDash \mathsf{K}_1 \mathsf{K}_3 m_2 \wedge \mathsf{K}_1 \neg \mathsf{K}_2 m_2$



$\mathcal{M}\restriction_P, mcc \vDash \mathsf{K}_1 m_1$

$\mathcal{M}\restriction_P, mmc \vDash \langle \neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3 \rangle Q_1 \vee Q_2 \vee Q_3$

$\mathcal{M}\restriction_P, mmm \vDash \langle \neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3 \rangle \neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3$

$\mathcal{M}\restriction_P \restriction_{\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3}, mmm \vDash \langle \neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3 \rangle Q_1 \vee Q_2 \vee Q_3$

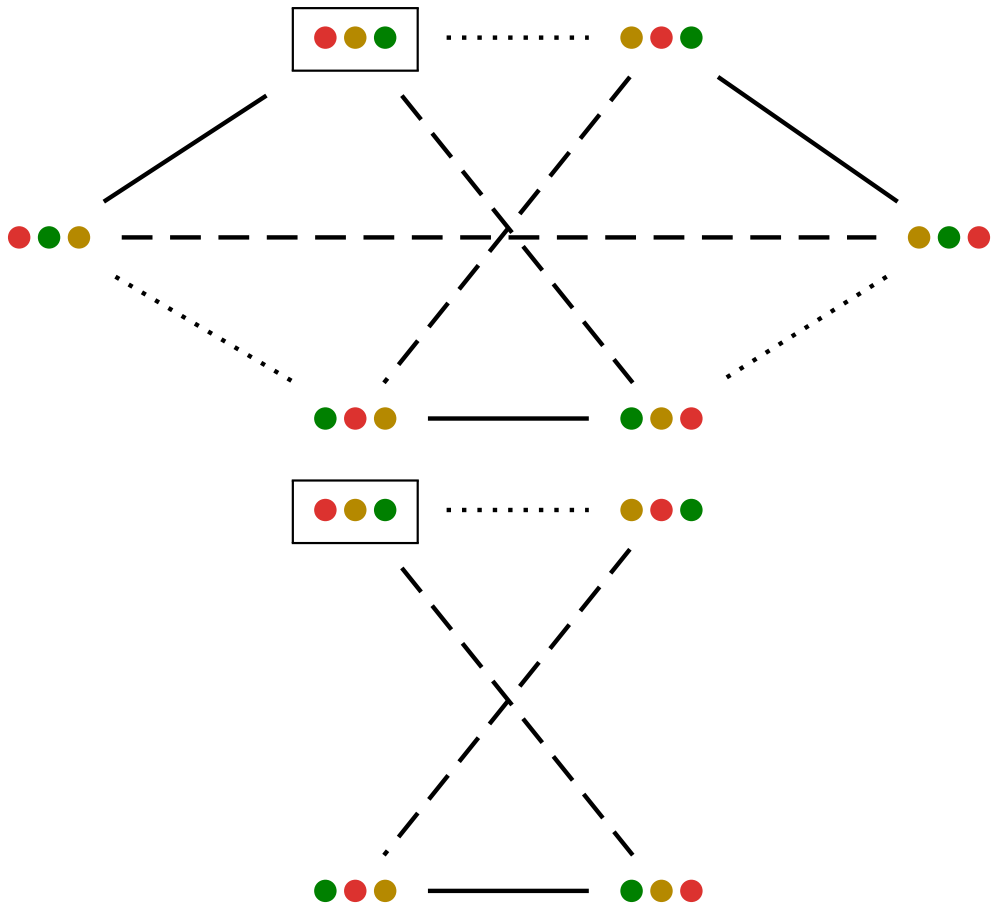
1. “有人把泥巴弄到脸上了!” $P := m_1 \vee m_2 \vee m_3$
2. “知道...站出来!” $\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3$ where $Q_i := \mathsf{K}_i m_i \vee \mathsf{K}_i \neg m_i$
3. “知道...站出来!” $Q_1 \wedge Q_2 \wedge \neg Q_3$

$\mathcal{M}, mmc \vDash [P][\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3][Q_1 \wedge Q_2 \wedge \neg Q_3](\mathsf{K}_1 m_1 \wedge \mathsf{K}_2 m_2 \wedge \mathsf{K}_3 \neg m_3)$

$\mathcal{M}, mmm \vDash [P][\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3][\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3](\mathsf{K}_1 m_1 \wedge \mathsf{K}_2 m_2 \wedge \mathsf{K}_3 m_3)$

三色卡

- “红”、“黄”、“绿” 三色卡依次被分给了 $1, 2, 3$ 三个小朋友.
- 每个小朋友只能看到自己手里的卡.
- 2 问 1: “你有绿卡吗?”
- 1: “没有!”





孩子们的年龄分别多大?

一个人口普查员向一位母亲询问她的孩子的情况.

- 母亲: "我有三个孩子, 他们的年龄乘起来是 36, 加起来是今天的日期."
- 普查员: "我还是不知道你孩子多大啊."
- 母亲: "不好意思, 忘了说了, 我们家老大喜欢喝咸豆脑."
- 普查员: "那我知道了!"

$$\begin{array}{ccc|c} 1 & 1 & 36 & \cancel{38} \\ 1 & 2 & 18 & 21 \\ 1 & 4 & 9 & 14 \\ 1 & 6 & 6 & 13 \\ 2 & 2 & 9 & 13 \\ 2 & 3 & 6 & 11 \\ 3 & 3 & 4 & 10 \end{array} \Longrightarrow \begin{array}{ccc|c} 1 & 6 & 6 & 13 \\ 2 & 2 & 9 & 13 \end{array} \Longrightarrow \begin{array}{ccc|c} 2 & 2 & 9 & 13 \end{array}$$

小菜的生日是哪天?

小艾和小白都想知道小菜的生日.
小菜给了他们 10 个可能的候选:

5.15	5.16	5.19
6.17	6.18	
7.14	7.16	
8.14	8.15	8.17

然后小菜分别告诉了小艾和小白她的生日的月份和日子.

- 小艾: "我不知道小菜的生日, 但我知道小白也不知道."
- 小白: "之前我不知道, 但现在我知道了."
- 小艾: "那我也知道了."

小菜的生日是哪天?

小艾和小白都想知道小菜的生日.
小菜给了他们 10 个可能的候选:

5.15	5.16	5.19
6.17	6.18	
7.14	7.16	
8.14	8.15	8.17

然后小菜分别告诉了小艾和小白她的生日的月份和日子.

- 小艾: "我不知道小菜的生日, 但我知道小白也不知道."
- 小白: "之前我不知道, 但现在我知道了."
- 小艾: "那我也知道了."

5.15	5.16	5.19									
6.17	6.18		$\Longrightarrow$	7.14	7.16		$\Longrightarrow$	7.16		$\Longrightarrow$	7.16
7.14	7.16			8.14	8.15	8.17		8.15	8.17		
8.14	8.15	8.17									

俄罗斯纸牌

俄罗斯纸牌

- 从七张纸牌 “0123456” 里小艾和小白分别被秘密的发放了三张纸牌, 小菜拿走了剩下的一张.
- **问题:** 小艾和小白有没有可能通过公开的宣告使得彼此都知道对方的牌是什么, 但同时小菜还是不知道任何一张不在手里的牌的归属?

Solution

- 假设小艾手里的是 012, 剩下的几张是 3456. 从 012 里选一张, 然后从 3456 里选两张. 凑成 $034, 056, 135, 146, 236, 245$. 小艾公开宣告拥有七种组合里的某一种:

$$\begin{array}{ll} 012 & \\ 034 & 056 \\ 135 & 146 \\ 236 & 245 \end{array}$$

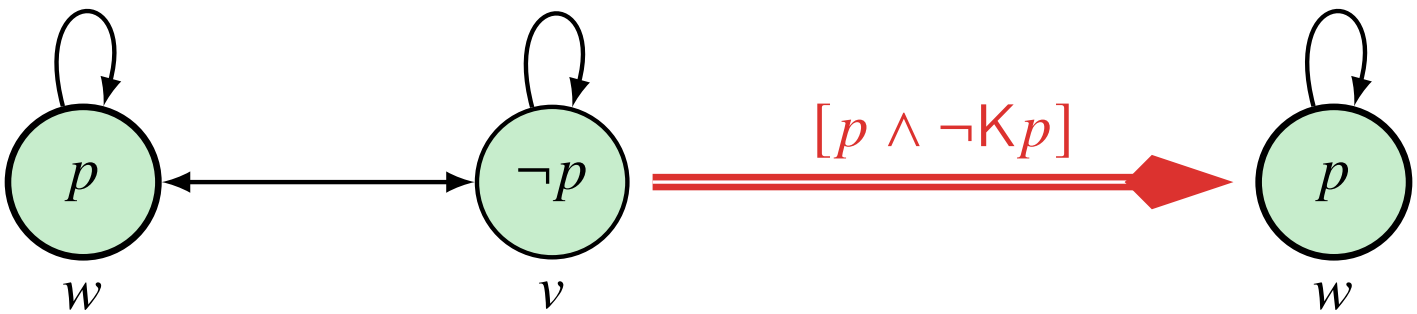
- 小白公开宣告小菜拥有的那张牌.

Unsuccessful Update

$$\vDash [p]\mathsf{C}_G p$$

$$\vDash [\mathsf{C}_G A]\mathsf{C}_G A$$

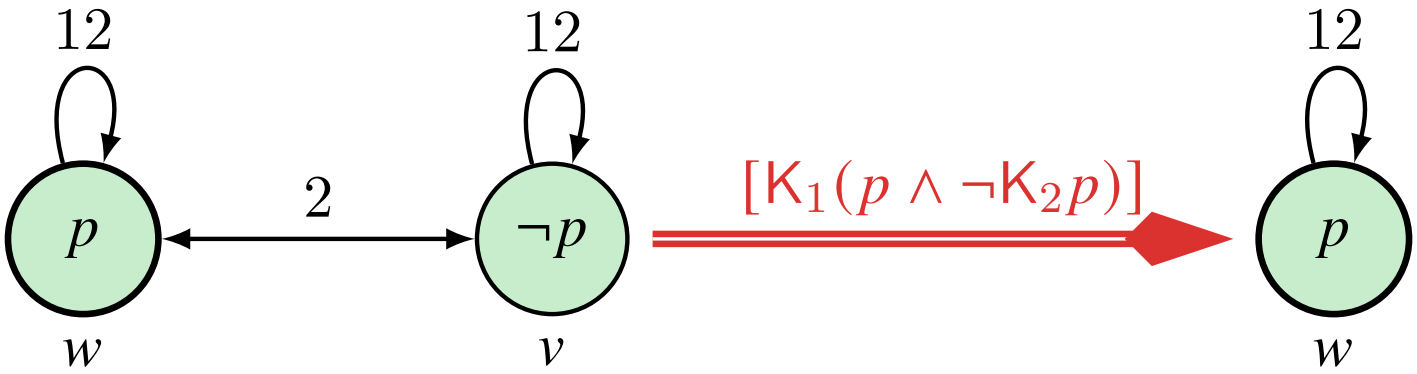
$$\overset{?}{\vDash} [A]\mathsf{C}_G A \qquad \overset{?}{\vDash} [A]\mathsf{K}A \qquad \overset{?}{\vDash} [A]A$$



$$\mathcal{M}, w \vDash (p \wedge \neg\mathsf{K}p) \wedge [p \wedge \neg\mathsf{K}p]\mathsf{K}p$$

Remark: If the goal of the announcing person was to "spread the truth of this formula," then this attempt was clearly unsuccessful.

Unsuccessful Update



$$\mathcal{M}, w \vDash (p \wedge \neg \mathsf{K}_2 p) \wedge \mathsf{K}_1(p \wedge \neg \mathsf{K}_2 p) \wedge [\mathsf{K}_1(p \wedge \neg \mathsf{K}_2 p)]\mathsf{K}_1 \mathsf{K}_2 p$$

1. 哪些 “公开宣告” 可以使 “真” 的变成 “假” 的?
2. 哪些 “公开宣告” 可以使 “假” 的变成 “真” 的?
3. 哪些 “公开宣告” 可以保 “真”?
4. 哪些 “公开宣告” 可以被 “知道”?
5. 哪些 “公开宣告” 可以变成 “公共知识”?

Valid?

1. $\langle B \rangle A \leftrightarrow B \wedge [B]A$
2. $[B](A \rightarrow C) \leftrightarrow ([B]A \rightarrow [B]C)$
3. $[B]p \leftrightarrow (B \rightarrow p)$
4. $[B]\neg A \leftrightarrow (B \rightarrow \neg A)$?
5. $[B]\neg A \leftrightarrow \neg[B]A$?
6. $[B]\neg A \leftrightarrow (B \rightarrow \neg[B]A)$
7. $[B]\mathsf{K}_i A \leftrightarrow (B \rightarrow \mathsf{K}_i(B \rightarrow [B]A))$
8. $[B]\mathsf{K}_i A \leftrightarrow (B \rightarrow \mathsf{K}_i[B]A)$
9. $[A]\mathsf{K}_i A$?
10. $[B][C]A \leftrightarrow [B \wedge C]A$?
11. $[B][C]A \leftrightarrow [B \wedge [B]C]A$
12. $\dfrac{A}{[B]A}$ $\quad$ $\dfrac{A(p)}{A[B/p]}$? $\quad$ $\dfrac{A \leftrightarrow B}{[A]C \leftrightarrow [B]C}$ $\quad$ $\dfrac{A \leftrightarrow B}{[C]A \leftrightarrow [C]B}$

Public Announcement Logic (PAL)

Axiom Schema

1. Tautologies
2. $\mathsf{K}_i(A \to B) \to \mathsf{K}_i A \to \mathsf{K}_i B$
3. $[B]p \leftrightarrow (B \to p)$
4. $[B]\neg A \leftrightarrow (B \to \neg[B]A)$
5. $[B](A \wedge C) \leftrightarrow [B]A \wedge [B]C$
6. $[B]\mathsf{K}_i A \leftrightarrow (B \to \mathsf{K}_i[B]A)$
7. $[B][C]A \leftrightarrow [B \wedge [B]C]A$

Inference Rule

$$\frac{A \quad A \to B}{B}\ \text{MP} \qquad\qquad \frac{A}{\mathsf{K}_i A}\ \text{N}$$

Expressive Power

Theorem

PAL *is equally expressive as basic modal logic.*

Proof.

$$t(\top) = \top$$
$$t(p) = p$$
$$t(\neg A) = \neg t(A)$$
$$t(A \wedge B) = t(A) \wedge t(B)$$
$$t(\mathsf{K}_i A) = \mathsf{K}_i t(A)$$

$$t([B]\top) = t(B \to \top)$$
$$t([B]p) = t(B \to p)$$
$$t([B]\neg A) = t(B \to \neg[B]A)$$
$$t([B](A \wedge C)) = t([B]A \wedge [B]C)$$
$$t([B]\mathsf{K}_i A) = t(B \to \mathsf{K}_i[B]A)$$
$$t([B][C]A) = t([B \wedge [B]C]A)$$

$$\vDash A \leftrightarrow t(A)$$

□

Succinctness

Theorem

PAL *is complete w.r.t. the standard semantics of Public Announcement Logic.*

Proof.

$$\vDash A \implies \vDash t(A) \implies \vdash_{\mathrm{K}} t(A) \implies \vdash_{\mathrm{PAL}} t(A) \implies \vdash_{\mathrm{PAL}} A$$

□

Theorem

PAL *is exponentially more succinct than modal logic on arbitrary models.*

$$A_0 \coloneqq \top$$
$$A_{n+1} \coloneqq \langle\langle A_n \rangle\rangle \Diamond_1 \top \rangle\rangle \Diamond_2 \top$$

where $\Diamond_i A \coloneqq \neg \mathsf{K}_i \neg A$ and $\langle B \rangle A \coloneqq \neg [B] \neg A$.

Announcement & Common Knowledge

$$\frac{P \to [Q]A \quad P \wedge Q \to \mathsf{E}_G P}{P \to [Q]\mathsf{C}_G A}$$

'Common knowledge induction' is a special case.
Take $P := A$ and $Q := \top$.

$$\mathsf{C}_G(A \to \mathsf{E}_G A) \to A \to \mathsf{C}_G A$$

Propositional Dynamic Logic

$$A \coloneqq \top \mid p \mid \neg A \mid A \wedge A \mid [\alpha]A$$
$$\alpha \coloneqq a \mid A? \mid \alpha;\alpha \mid \alpha \cup \alpha \mid \alpha^*$$

$$\mathcal{M}, w \vDash [\alpha]A \;\text{ iff }\; \forall v \in W\Big(R_\alpha wv \implies \mathcal{M}, v \vDash A\Big)$$
$$\mathcal{M}, w \vDash \langle\alpha\rangle A \;\text{ iff }\; \exists v \in W\Big(R_\alpha wv \;\&\; \mathcal{M}, v \vDash A\Big)$$

where

$$\begin{aligned} R_{A?} &\coloneqq \{(w,w) : \mathcal{M}, w \vDash A\} \\ R_{\alpha;\beta} &\coloneqq \big\{(w,v) : \exists u\big(R_\alpha wu \wedge R_\beta uv\big)\big\} \\ R_{\alpha\cup\beta} &\coloneqq R_\alpha \cup R_\beta \\ R_{\alpha^*} &\coloneqq \bigcup_{n=0}^{\infty} R_{\alpha^n} \end{aligned}$$

Propositional Dynamic Logic (PDL)

Axiom Schema

1. Tautologies
2. $[\alpha](A \to B) \to [\alpha]A \to [\alpha]B$
3. $[B?]A \leftrightarrow (B \to A)$
4. $[\alpha;\beta]A \leftrightarrow [\alpha][\beta]A$
5. $[\alpha \cup \beta]A \leftrightarrow [\alpha]A \wedge [\beta]A$
6. $[\alpha^*]A \leftrightarrow A \wedge [\alpha][\alpha^*]A$
7. $A \wedge [\alpha^*](A \to [\alpha]A) \to [\alpha^*]A$

Inference Rule

$$\frac{A \quad A \to B}{B}\ \text{MP} \qquad\qquad \frac{A}{[\alpha]A}\ \text{N}$$

PDL is sound and weak complete.

PDL is not compact. $\{\langle a^* \rangle p, \neg p, \neg\langle a \rangle p, \neg\langle a;a \rangle p, \neg\langle a;a;a \rangle p, \dots\}$

Its satisfiability is decidable (in EXPTIME).

一阶动态逻辑 First Order Dynamic Logic

Axiom Schema

1. FOL
2. PDL
3. $\langle x := t \rangle A \leftrightarrow A[t/x]$

Inference Rule

$$\frac{A \quad A \to B}{B}\ \text{MP} \qquad \frac{A}{[\alpha]A}\ \text{N}$$

$$\frac{A \to [\alpha^n]B, \quad n \in \omega}{A \to [\alpha^*]B}\ \text{IC} \qquad \frac{A}{\forall x A}\ \text{G}$$

Application — Program Analysis

$$\begin{aligned}
\textbf{skip} &:= \top? \\
\textbf{fail} &:= \bot? \\
\textbf{if } B \textbf{ then } \alpha \textbf{ else } \beta &:= B?;\alpha \cup \neg B?;\beta \\
\textbf{while } B \textbf{ do } \alpha &:= (B?;\alpha)^*;\neg B? \\
\textbf{repeat } \alpha \textbf{ until } B &:= \alpha;(\neg B?;\alpha)^*;B? \\
\{A\}\ \alpha\ \{B\} &:= A \rightarrow [\alpha]B
\end{aligned}$$

Algorithm GCD

while $x \neq y$ **do**
  if $x > y$ **then**
    $x \leftarrow x - y$
  else
    $y \leftarrow y - x$
  end if
end while

$$[(x = m \wedge y = n)?]\left\langle (x \neq y?;(x > y?;x \leftarrow x - y) \cup (x < y?;y \leftarrow y - x))^*;x = y?\right\rangle x = \gcd(m,n)$$

Hoare Logic

$$\frac{}{\{P\}\ \mathbf{skip}\ \{P\}}$$

$$\frac{}{\{P[t/x]\}\ x := t\ \{P\}}$$

$$\frac{\{P\}\ \alpha\ \{Q\} \quad \{Q\}\ \beta\ \{R\}}{\{P\}\ \alpha;\beta\ \{R\}}$$

$$\frac{\{B \wedge P\}\ \alpha\ \{Q\} \quad \{\neg B \wedge P\}\ \beta\ \{Q\}}{\{P\}\ \mathbf{if}\ B\ \mathbf{then}\ \alpha\ \mathbf{else}\ \beta\ \{Q\}}$$

$$\frac{P_1 \to P_2 \quad \{P_2\}\ \alpha\ \{Q_2\} \quad Q_2 \to Q_1}{\{P_1\}\ \alpha\ \{Q_1\}}$$

$$\frac{\{P \wedge B\}\ \alpha\ \{P\}}{\{P\}\ \mathbf{while}\ B\ \mathbf{do}\ \alpha\ \{\neg B \wedge P\}}$$

Theorem

The rules of Hoare Logic are derivable in Dynamic Logic.

$\{x = 4 \wedge y = 3\}\ \mathbf{if}\ x < y\ \mathbf{then}\ z := x; y := y + 1\ \mathbf{else}\ z := y; z := z + 1\ \{x = 4 \wedge y = 3 \wedge z = 4\}$

Temporal Logic

some past / finally future / all past / globally future / next / since / until

$$A := p \mid \bot \mid \neg A \mid A \wedge A \mid \mathsf{P}A \mid \mathsf{F}A \mid \mathsf{H}A \mid \mathsf{G}A \mid \mathsf{X}A \mid A\mathsf{S}A \mid A\mathsf{U}A$$

- $\mathcal{M}, n \vDash \mathsf{P}A$ iff $\exists m < n \big(\mathcal{M}, m \vDash A\big)$
- $\mathcal{M}, n \vDash \mathsf{F}A$ iff $\exists m > n \big(\mathcal{M}, m \vDash A\big)$
- $\mathcal{M}, n \vDash \mathsf{H}A$ iff $\forall m < n \big(\mathcal{M}, m \vDash A\big)$
- $\mathcal{M}, n \vDash \mathsf{G}A$ iff $\forall m > n \big(\mathcal{M}, m \vDash A\big)$
- $\mathcal{M}, n \vDash \mathsf{X}A$ iff $\mathcal{M}, n+1 \vDash A$
- $\mathcal{M}, n \vDash A\mathsf{S}B$ iff $\exists m < n \Big[\mathcal{M}, m \vDash B \;\; \& \;\; \forall t \big(m < t < n \implies \mathcal{M}, t \vDash A\big)\Big]$
- $\mathcal{M}, n \vDash A\mathsf{U}B$ iff $\exists m > n \Big[\mathcal{M}, m \vDash B \;\; \& \;\; \forall t \big(n < t < m \implies \mathcal{M}, t \vDash A\big)\Big]$

$$\mathsf{H}A \equiv \neg\mathsf{P}\neg A \qquad \mathsf{G}A \equiv \neg\mathsf{F}\neg A \qquad \mathsf{P}A \equiv \top\mathsf{S}A \qquad \mathsf{F}A \equiv \top\mathsf{U}A$$

Yablo Paradox in Linear Temporal Logic

Yablo Paradox

A_1: for all $k > 1$, A_k is false.

A_2: for all $k > 2$, A_k is false.

A_3: for all $k > 3$, A_k is false.

$\vdots$

$$\not\models A \leftrightarrow \mathsf{G}\neg A$$

Other Versions of Yablo Paradox

always	$A_n \leftrightarrow \forall m > n : \neg A_m$	$\models \neg\mathsf{G}(A \leftrightarrow \mathsf{G}\neg A)$
sometimes	$A_n \leftrightarrow \exists m > n : \neg A_m$	$\models \neg\mathsf{G}(A \leftrightarrow \mathsf{F}\neg A)$
almost always	$A_n \leftrightarrow \exists k > n \forall m > k : \neg A_m$	$\models \neg\mathsf{G}(A \leftrightarrow \mathsf{FG}\neg A)$
infinitely often	$A_n \leftrightarrow \forall k > n \exists m > k : \neg A_m$	$\models \neg\mathsf{G}(A \leftrightarrow \mathsf{GF}\neg A)$

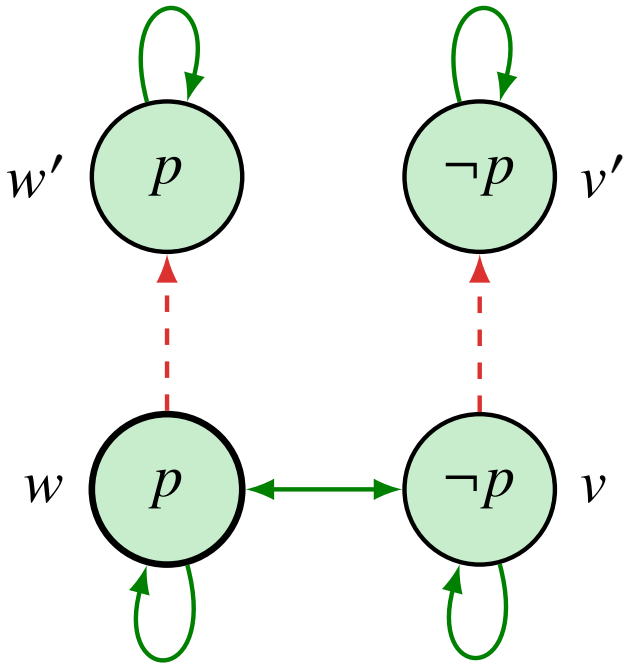
Temporal Logic — Formal System

Axiom Schema

- Tautologies
- $\mathsf{G}(A \to B) \to \mathsf{G}A \to \mathsf{G}B$
- $\mathsf{H}(A \to B) \to \mathsf{H}A \to \mathsf{H}B$
- $A \to \mathsf{GP}A$
- $A \to \mathsf{HF}A$

Inference Rule

$$\frac{A \quad A \to B}{B} \qquad \frac{A}{\mathsf{H}A} \qquad \frac{A}{\mathsf{G}A}$$



$$w \vDash \neg \mathsf{K} p \wedge \mathsf{FK} p$$

Remark: two-dimensional semantics

宿命论论证

1. 如果明天有海战为真，那么明天有海战就不可能为假，即明天必然有海战.
2. 如果明天有海战为假，那么明天有海战就不可能为真，即明天必然没有海战.
3. 因此，要么明天必然有海战，要么明天必然没有海战.

$$\frac{\begin{array}{c} p \to \Box p \\ \neg p \to \Box \neg p \end{array}}{\Box p \vee \Box \neg p}$$

$$\frac{\begin{array}{c} \Box(p \to p) \\ \Box(\neg p \to \neg p) \end{array}}{\Box p \vee \Box \neg p} \times$$

庄子《秋水》

庄子与惠子游于濠梁之上.

1. 庄子: 鲦鱼出游从容, 是鱼之乐也.
2. 惠子: 子非鱼, 安知鱼之乐?

$$\forall xy(\mathsf{K}_x Hy \vee \mathsf{K}_x \neg Hy \to Fy \to Fx) \quad ?$$

$$\forall x(\mathsf{K}_x Hf \vee \mathsf{K}_x \neg Hf \to x = f) \quad \checkmark$$

3. 庄子: 子非我, 安知我不知鱼之乐?

$$\forall xy(\mathsf{K}_x \mathsf{K}_y Hf \vee \mathsf{K}_x \neg \mathsf{K}_y Hf \to x = y)$$

4. 惠子: 我非子, 固不知子矣; 子固非鱼也, 子之不知鱼之乐, 全矣.

For any **'subjective'** formula A,

$$\frac{\forall xy\big(\mathsf{K}_x Ay \vee \mathsf{K}_x \neg Ay \to x = y\big) \quad z \neq f \quad h \neq z}{\neg \mathsf{K}_z Hf \wedge \neg \mathsf{K}_z \neg Hf \wedge \neg \mathsf{K}_h \mathsf{K}_z Hf \wedge \neg \mathsf{K}_h \neg \mathsf{K}_z Hf}$$ Moore's Paradox?

5. 庄子: 请循其本. 子曰 '汝安知鱼乐' 云者, 既已知吾知之而问我. 我知之濠上也.

上帝存在的本体论论证?

1. God is a being which has every perfection.
2. Existence is a perfection.
3. Hence God exists.

$$E\left(\iota_x\left(\bigwedge_{i\in I} P_i x \wedge Ex\right)\right)$$

Problem: 这是否是循环论证?

上帝存在的本体论论证

- R: reality
- M: mind
- P: the positive qualities
- $x \in y$: x belongs to y

1. There exists a thing belonging to mind that has all the positive qualities and no negative quality.

$$\exists x[x \in M \wedge \forall y(y \in P \leftrightarrow x \in y)]$$

2. "Being real" is a positive quality.

$$R \in P$$

3. Two things belonging to mind that have exactly the same qualities are identical.

$$\forall xy[x \in M \wedge y \in M \rightarrow \forall z(x \in z \leftrightarrow y \in z) \rightarrow x = y]$$

4. God belongs to reality.

$$\iota_x[x \in M \wedge \forall y(y \in P \leftrightarrow x \in y)] \in R$$

Gödel's Proof of God's Existence

Ax.1 Either a property or its negation is positive, but not both. $\forall X[P(\neg X) \leftrightarrow \neg P(X)]$

Ax.2 A property necessarily implied by a positive property is positive.
$\forall X \forall Y\,[P(X) \wedge \Box \forall x[X(x) \rightarrow Y(x)] \rightarrow P(Y)]$

Th.1 Positive properties are possibly exemplified. $\forall X[P(X) \rightarrow \Diamond \exists x X(x)]$

Df.1 A *God-like* being possesses all positive properties. $G(x) \coloneqq \forall X[P(X) \rightarrow X(x)]$

Ax.3 The property of being God-like is positive. $P(G)$

Th.2 Possibly, God exists. $\Diamond \exists x G(x)$

Ax.4 Positive properties are necessarily positive. $\forall X[P(X) \rightarrow \Box P(X)]$

Df.2 An *essence* of an individual is a property necessarily implying any of its properties.
$E(X,x) \coloneqq X(x) \wedge \forall Y(Y(x) \rightarrow \Box \forall y(X(y) \rightarrow Y(y)))$

Th.3 Being God-like is an essence of any God-like being. $\forall x[G(x) \rightarrow E(G,x)]$

Df.3 *Necessary existence* of an individual is the necessary exemplification of all its essences. $N(x) \coloneqq \forall X[E(X,x) \rightarrow \Box \exists y X(y)]$

Ax.5 Necessary existence is a positive property. $P(N)$

Th.4 Necessarily, God exists. $\Box \exists x G(x)$

$$\frac{\text{任何经验都可能不可靠}}{\text{可能所有经验都不可靠}}\ ?$$

- $\forall x \Diamond A \rightarrow \Diamond \forall x A$?
- $\Diamond \forall x A \rightarrow \forall x \Diamond A$?

1. Barcan Formula

$$\forall x \Box A \rightarrow \Box \forall x A$$

holds in frame iff

$$Rwv \rightarrow D_v \subset D_w$$

2. The Converse of Barcan Formula

$$\Box \forall x A \rightarrow \forall x \Box A$$

holds in frame iff

$$Rwv \rightarrow D_w \subset D_v$$

Propositional Quantifiers

- I believe that everything I believe is true: $\mathsf{B}\forall p(\mathsf{B}p \to p)$.
- I know that there's a truth I don't know: $\mathsf{K}\exists p(p \wedge \neg \mathsf{K}p)$.
- a knows that b knows everything a knows: $\mathsf{K}_a\forall p(\mathsf{K}_a p \to \mathsf{K}_b p)$.
- There is a true proposition that necessarily implies every true proposition: $\exists p(p \wedge \forall q(q \to \Box(p \to q)))$.

$$[\![\forall p\varphi]\!]_V = \bigcap_{X \subset W} [\![\varphi]\!]_{V[X/p]}$$

- $[\![\forall p(\Box p \to p)]\!] = \{w \in W : Rww\}$
- $[\![\exists p(p \wedge \neg\Box p)]\!] = \{w \in W : \exists v \neq w : Rwv\}$
- $[\![\exists p(p \wedge \forall q(q \to \Box(p \to q)))]\!] = W$

Fitch's Paradox of Knowability

Theorem

If all truths are knowable, then all truths are known.

$$\forall p(p \to \Diamond \mathsf{K} p) \vdash \forall p(p \to \mathsf{K} p)$$

1. $\neg \mathsf{K}(p \wedge \neg \mathsf{K} p)$ 摩尔悖论
2. $\Box \neg \mathsf{K}(p \wedge \neg \mathsf{K} p)$ $\dfrac{\vdash A}{\vdash \Box A}$
3. $\neg \Diamond \mathsf{K}(p \wedge \neg \mathsf{K} p)$
4. $p \wedge \neg \mathsf{K} p$ 假设
5. $\Diamond \mathsf{K}(p \wedge \neg \mathsf{K} p)$ 前提 $p \to \Diamond \mathsf{K} p$
6. $\neg(p \wedge \neg \mathsf{K} p)$
7. $p \to \mathsf{K} p$

$$\text{Is } \Diamond(\mathsf{K} p \vee \mathsf{K} \neg p)?$$

翻译 — 学说 “逻辑语”

- He who refuses to do arithmetic is doomed to talk nonsense.

  — *John McCarthy*

$$\forall x(\text{Refuse}(x, \text{arithmetic}) \rightarrow \Box\text{TalkNonsense}(x))$$

- It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife.

  — *Jane Austen*

$$\mathsf{C}_{\text{Human}}\Big(\forall x\Big(\text{Man}(x) \wedge \text{Single}(x) \wedge \text{Fortune}(x) \rightarrow \mathsf{Want}_x\Big(\exists y\big(\text{Woman}(y) \wedge \text{Wife}(y, x)\big)\Big)\Big)\Big)$$

- 我们知道他们在说谎, 他们也知道自己在说谎, 他们也知道我们知道他们在说谎, 我们也知道他们知道我们知道他们说谎, 但是他们依然在说谎.

  — 索尔仁尼琴

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Counterfactual Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

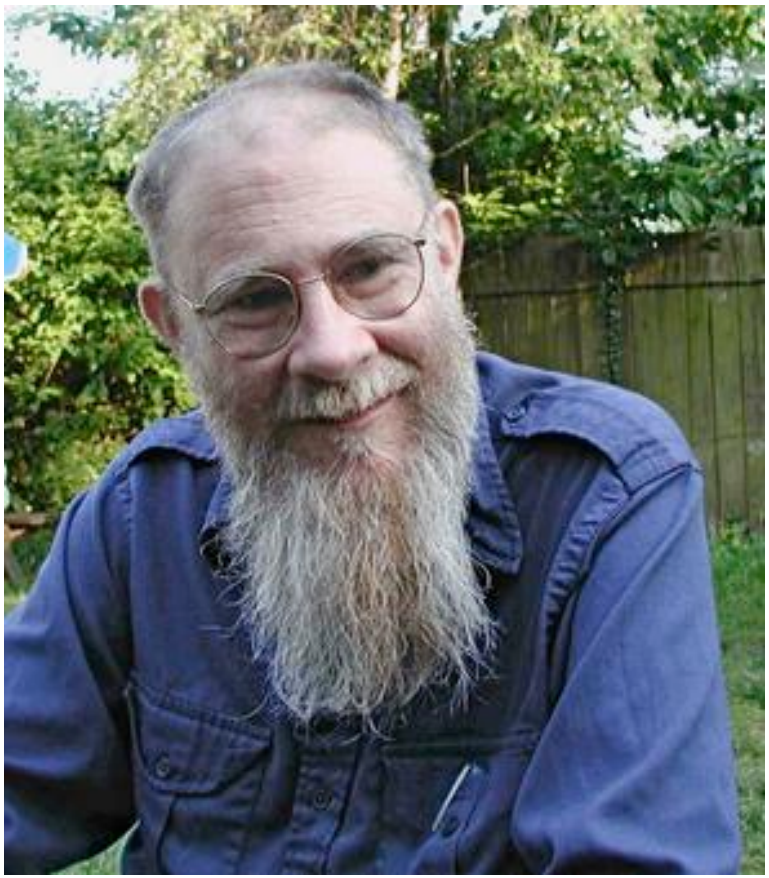
Quantum Computing

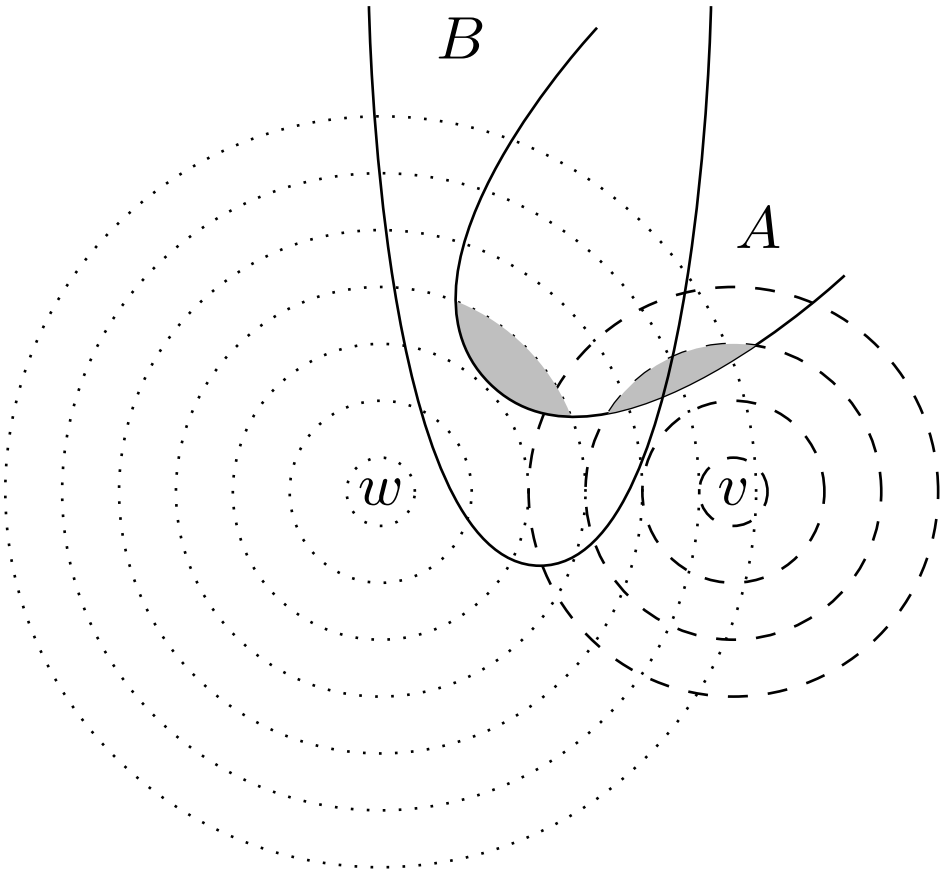
Answers to the Exercises

References2039

David Lewis 1941–2001

Counterfactual Causation





$$w \vDash A \boxdot\!\!\rightarrow B$$

$$v \nvDash A \boxdot\!\!\rightarrow B$$

Indicative vs Counterfactual Conditional

- If Oswald did not kill Kennedy, someone else did.
- If Oswald had not killed Kennedy, someone else would have.

Antecedents and Consequents

$$\frac{\neg A}{A \rightarrow B}\ \checkmark \qquad \frac{\neg A}{A \mathrel{\square\!\!\rightarrow} B}\ \times$$

$$\frac{\text{I did not strike the match}}{\text{if I had struck the match, it would have turned into a feather}}\ ?$$

$$\frac{B}{A \rightarrow B}\ \checkmark \qquad \frac{B}{A \mathrel{\square\!\!\rightarrow} B}\ \times$$

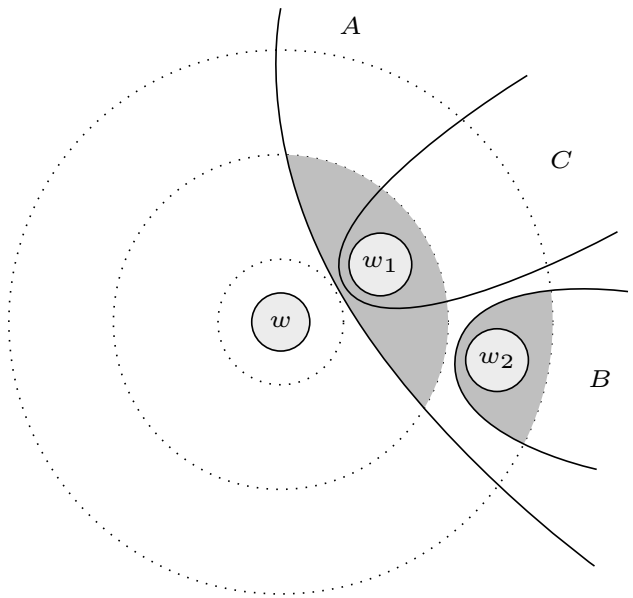
$$\frac{\text{George W. Bush won the } 2004 \text{ United States presidential election}}{\begin{array}{l}\text{If the newspapers had discovered beforehand that}\\ \text{Bush had an affair with Al Gore, he would still have won}\end{array}}\ ?$$

Antecedent Strengthening

$$\frac{A \rightarrow C}{A \wedge B \rightarrow C}\ \checkmark \qquad \frac{A \mathrel{\Box\!\!\rightarrow} C}{A \wedge B \mathrel{\Box\!\!\rightarrow} C}\ \times$$

If I went to the beach, I would have a good time

If I went to the beach and got attacked by a shark, I would have a good time



The closest world where I go to the beach and get attacked by a shark is much further removed from the actual world than the closest world where I go to the beach is.

Transitivity

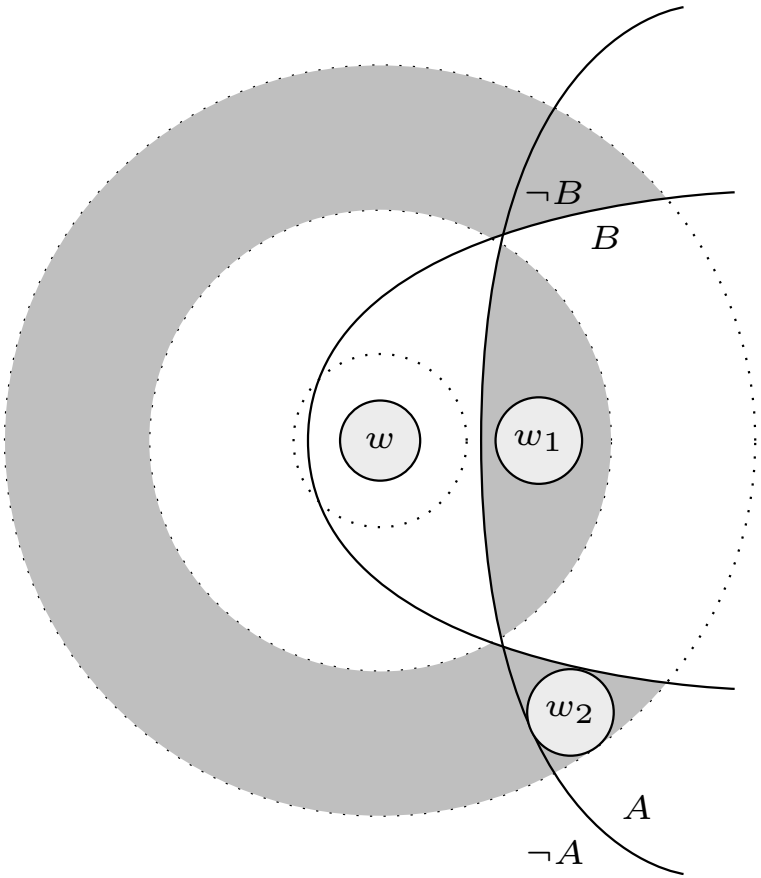
$$\frac{A \to B \qquad B \to C}{A \to C} \checkmark \qquad \frac{A \boxrightarrow B \qquad B \boxrightarrow C}{A \boxrightarrow C} \times \qquad \frac{A \boxrightarrow B \qquad A \wedge B \boxrightarrow C}{A \boxrightarrow C} \checkmark$$

$$\frac{\text{If I were king, I would wear a crown} \qquad \text{If I wore a crown, people would find me ridiculous}}{\text{If I were king, people would find me ridiculous}} \ ?$$

Contraposition

$$\frac{A \rightarrow B}{\neg B \rightarrow \neg A} \checkmark \qquad \frac{A \mathrel{\Box\!\!\rightarrow} B}{\neg B \mathrel{\Box\!\!\rightarrow} \neg A} \times$$

$$\frac{\text{If Goethe hadn't died in 1832, he would (still) be dead now}}{\text{If Goethe weren't dead now, he would have died in 1832}} \ ?$$



Minimal Change Semantics

Definition (Sphere Model)

A *sphere model* is a triple $\mathcal{M} = \{W, O, V\}$, where $W \neq \emptyset$, $V : \text{Atom} \longrightarrow \mathtt{P}(W)$, and $O : W \longrightarrow \mathtt{P}(\mathtt{P}(W))$ assigns to each world w a *system of spheres* O_w. For each w, O_w is a set of sets of worlds such that:

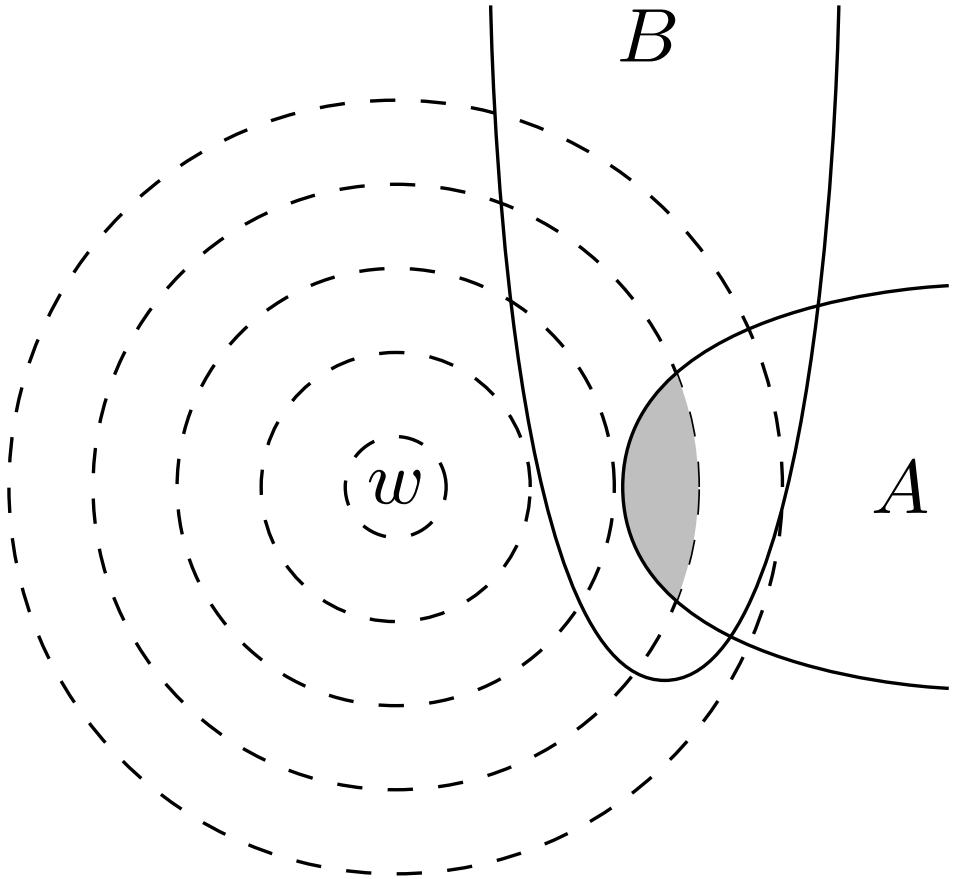
1. O_w is *centered* on w: $\{w\} \in O_w$.
2. O_w is *nested*: whenever $S_1, S_2 \in O_w$, $S_1 \subset S_2$ or $S_2 \subset S_1$.
3. O_w is closed under non-empty unions.
4. O_w is closed under non-empty intersections.

Definition (Counterfactual Conditional)

$\mathcal{M}, w \vDash A \boxdot\!\!\rightarrow B$ iff either

1. for all $v \in \bigcup O_w : \mathcal{M}, v \nvDash A$, or
2. for some $S \in O_w$,
   - 2.1 $\mathcal{M}, v \vDash A$ for some $v \in S$, and
   - 2.2 for all $v \in S : \mathcal{M}, v \vDash A \longrightarrow B$.

Minimal Change Semantics



Problem

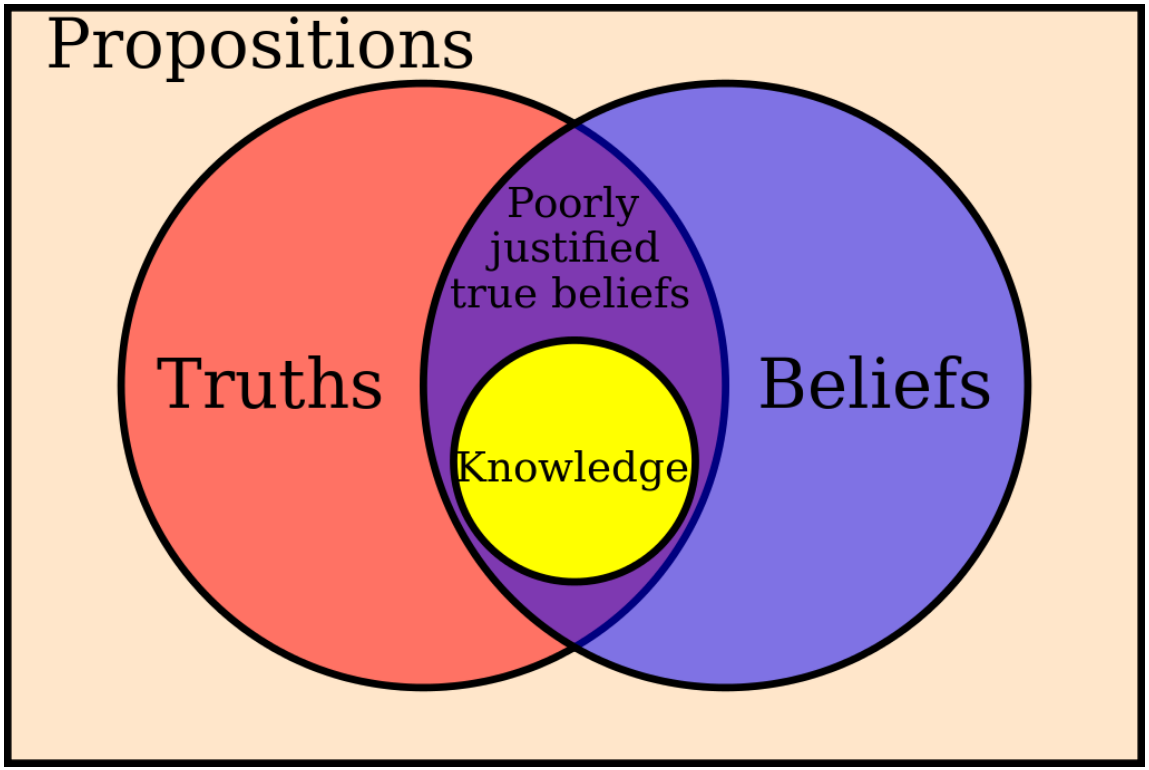
How do humans represent "possible worlds" in their minds and compute the closest one, when the number of possibilities is far beyond the capacity of the human brain?

Formal System

$$\frac{\bigwedge_{i=1}^{n} B_i \to C}{\bigwedge_{i=1}^{n} (A \boxdot\!\!\to B_i) \to (A \boxdot\!\!\to C)}$$

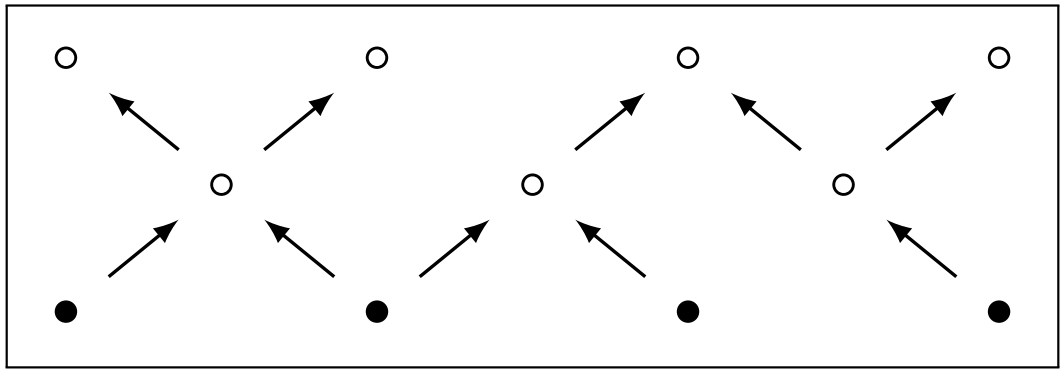
- $A \boxdot\!\!\to A$
- $(A \boxdot\!\!\to B) \wedge (B \boxdot\!\!\to A) \to (A \boxdot\!\!\to C) \leftrightarrow (B \boxdot\!\!\to C)$
- $((A \vee B) \boxdot\!\!\to A) \vee ((A \vee B) \boxdot\!\!\to B) \vee (((A \vee B) \boxdot\!\!\to C) \leftrightarrow (A \boxdot\!\!\to C) \wedge (B \boxdot\!\!\to C))$
- $(\neg A \boxdot\!\!\to A) \to (B \boxdot\!\!\to A)$
- $(A \boxdot\!\!\to \neg B) \vee ((A \wedge B \boxdot\!\!\to C) \leftrightarrow (A \boxdot\!\!\to (B \to C)))$
- $(A \boxdot\!\!\to B) \to (A \to B)$
- $A \wedge B \to (A \boxdot\!\!\to B)$

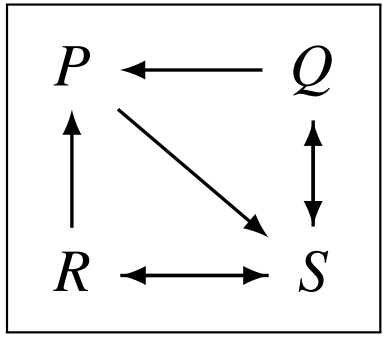
Gettier Problem: Is Justified True Belief Knowledge?



1. Smith believes **"Bob owns a Ford"**.
2. He was told this by Bob.
3. Bob then sells his Ford.
4. Meanwhile Bob wins a Ford in a raffle.

Epistemic Justification





- Internalist theory of epistemic justification.
  - Foundationalism
    — S's belief p is justified iff either
    (a) p is a basic belief; or
    (b) p is a non-basic belief justified inferentially by S's basic beliefs.
  - Coherentism
    — S's belief p is justified iff S has a coherent set of beliefs which includes p.
- Externalist theory of epistemic justification.
  - Reliabilism
    — S's belief p is justified iff p is produced by a reliable cognitive process.

证成难题

1. 一个信念只能由另一个信念证成.
2. 不允许循环证成.
3. 证成链条是有穷的.
4. 有被证成的信念.

- 上面四个命题不相容.
- 基础论者拒斥 1.
- 融贯论者拒斥 2.
- 皮尔士拒斥 3.
- 取消论者拒斥 4.

Nozick's Truth-Tracking Condition

Nozick's Truth-Tracking Condition

S knows that p iff:

1. p is true;
2. S believes that p;
3. if p were not true, S would not believe that p;
4. if p were true, S would believe that p.

$$\mathsf{K}p := p \wedge \mathsf{B}p \wedge (\neg p \mathrel{\Box\!\!\rightarrow} \neg \mathsf{B}p) \wedge (p \mathrel{\Box\!\!\rightarrow} \mathsf{B}p)$$

Kripke's counter-example — Fake Barn Country

- Smith is driving in a country containing fake barns.
- The fake barns are painted green.
- In the midst of these fake barns is one real barn, which is painted red.
  1. Smith looks up and happens to see the real barn. **"I see a red barn."**
  2. What if Smith looks up and forms the belief **"I see a barn"**?
- Smith knows "there is a red barn", but doesn't know "there is a barn".

怀疑论 vs 反事实

S_1 I know "I have hands". $\mathsf{K}p$

S_2 I know if "I have hands" then "I am not a brain in the vat". $\mathsf{K}(p \to q)$

S_3 I don't know that "I am not a brain in the vat". $\neg\mathsf{K}q$

Remark: I fail to know that "I am not a brain in the vat" $\neg\mathsf{K}q$, since I would falsely believe "I was not a brain in the vat" in the closest world in which I am a brain in the vat.

- Nozick's definition of Knowledge

$$\mathsf{K}p := p \wedge \mathsf{B}p \wedge (\neg p \boxdot\!\!\!\rightarrow \neg \mathsf{B}p) \wedge (p \boxdot\!\!\!\rightarrow \mathsf{B}p)$$

  then K is not closed under known entailment.

$$\mathsf{K}p, \mathsf{K}(p \to q) \nvDash \mathsf{K}q$$

- By Sosa's definition,

$$\mathsf{K}p := p \wedge \mathsf{B}p \wedge (\mathsf{B}p \boxdot\!\!\!\rightarrow p) \wedge (p \boxdot\!\!\!\rightarrow \mathsf{B}p)$$

  then

$$\mathsf{K}p, \mathsf{K}(p \to q) \vDash \mathsf{K}q$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Readings

1. H. Enderton: Elements of Set Theory
2. T. Jech: Set Theory
3. K. Kunen: Set Theory
4. A. Kanamori: The Higher Infinite
5. R. Schindler: Set Theory
6. F. W. Lawvere, R. Rosebrugh: Sets for Mathematics

- Without a consistent theory of the mathematical infinite there is no theory of irrationals;
- Without a theory of irrationals there is no mathematical analysis;
- Without analysis the major part of mathematics — including geometry and most of applied mathematics — as it now exists would cease to exist.
- The most important task confronting mathematicians would therefore seem to be the construction of a satisfactory theory of the infinite.
- ZFC provides a common language and a powerful basic toolset.
- ZFC is rich enough to encode all of mathematics.
- All of mathematics is consistent relative to set theory.
- It supplies a consistent ontology for mathematics, and a context in which to ask metamathematical questions.

康托尔 Georg Cantor 1845-1918

- Mathematics $\rightsquigarrow$ Set Theory.
- Diagonalization.
- There are many different levels of infinity.
- Cantor set.
- Continuum Hypothesis (CH).
  How many points on the line?



"I don't know what predominates in Cantor's theory — philosophy or theology, but I am sure that there is no mathematics there."

— Kronecker

Welcome to Cantor's Paradise



Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory
- Axioms of ZFC
- Ordinal Numbers
- Cardinal Numbers
- Axiom of Choice

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

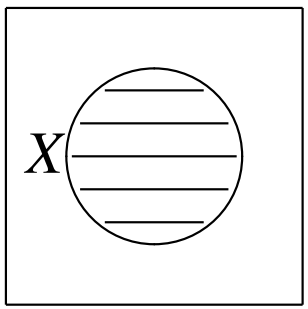
Quantum Computing

Answers to the Exercises

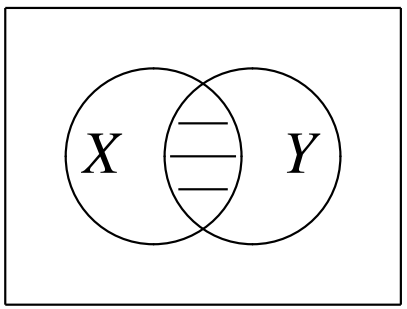
References2039

一切皆集合

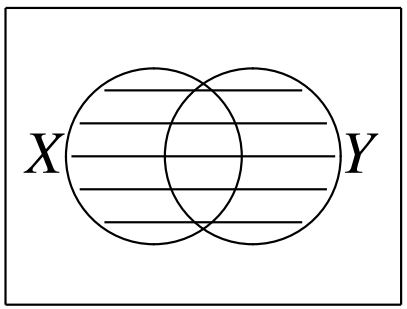
X



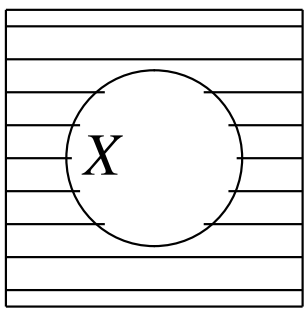
$X \cap Y$



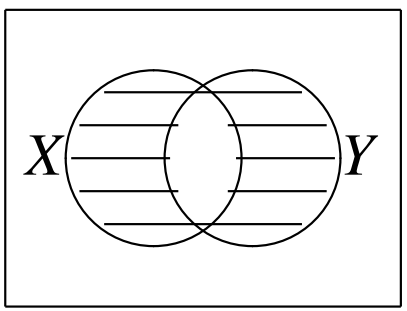
$X \cup Y$



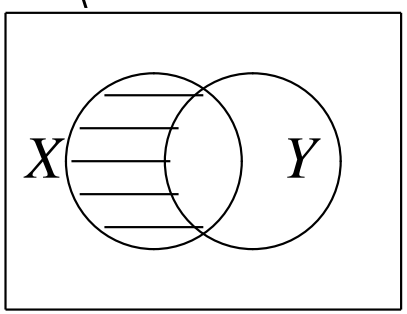
$\overline{X}$



$X \Delta Y$



$X \setminus Y$



一切皆集合

- 圆是到定点的距离等于定长的点集.
- 抛物线是到定点和定直线等距的点集.
- 椭圆是到两个定点的距离之和为常数 (大于两定点的距离) 的点集.
- 双曲线是到两个定点距离之差的绝对值为常数 (小于两定点的距离) 的点集.

ZFC — Axioms

- $a \in A$ reads: a is an element of A. (Definition? No!)
- **Extensionality**.
$$X = Y \leftrightarrow \forall u(u \in X \leftrightarrow u \in Y)$$
- **Axiom Schema of Comprehension**. (×)
For any formula A, there exists a set $Y = \{x : A(x)\}$.

$$R := \{x : x \notin x\} \quad R \in R? \qquad \text{(Russell Paradox)}$$

- **Separation Schema**.
For any formula A, for any X, there exists a set $Y = \{u \in X : A(x)\}$.

$$\forall X \exists Y \forall u(u \in Y \leftrightarrow u \in X \wedge A(x))$$

The Hyper-game Paradox

- Let G be the collection of all finite games which can be played by two players by making successive alternate moves.
- Define the hyper-game as the game in which the first player chooses a finite game $g \in G$, and then the second player starts playing g.
- **Question:** Is the hyper-game finite?
- The problem here is that the collection of all finite games G is a class and we define the hyper-game as a particular element which depends on the whole class.

Curry's Paradox

- $X \coloneqq \{x : x \in x \to A\}$
- $X \in X \iff X \in X \to A$
- A

ZFC — Axioms

- **Pairing**. For any a and b there exists a set $c = \{a, b\}$.

$$\forall ab \exists c \forall x (x \in c \leftrightarrow x = a \vee x = b)$$

- **Power**. For any X there exists a set $Y = \mathrm{P}(X) \coloneqq \{u : u \subset X\}$.

$$\forall X \exists Y \forall u (u \in Y \leftrightarrow \forall z (z \in u \to z \in X))$$

- **Union**. For any X there exists a set $Y = \bigcup X$.

$$\forall X \exists Y \forall u (u \in Y \leftrightarrow \exists z (z \in X \wedge u \in z))$$

$$\bigcup_{i \in I} X_i \coloneqq \bigcup \{X_i : i \in I\}$$

$$\bigcap X \coloneqq \{u : \forall z (z \in X \to u \in z)\}$$

Relation

- ordered pair.

$$(a, b) \coloneqq \{\{a\}, \{a, b\}\}$$
$$(a_1, \ldots, a_{n+1}) \coloneqq ((a_1, \ldots, a_n), a_{n+1})$$

$$(a_1, \ldots, a_n) = (b_1, \ldots, b_n) \to a_i = b_i \quad \text{for} \quad 1 \leq i \leq n$$

$$X \subsetneq Y \quad X \cup Y \quad X \cap Y \quad X \setminus Y \quad X \Delta Y \quad X \times Y \quad \prod_{i=1}^{n} X_i \quad X^n$$

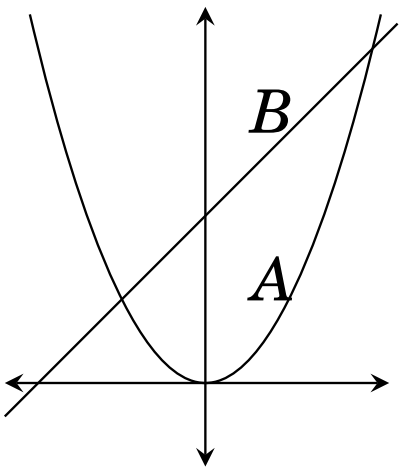
- n-ary relation R on $X_1, \ldots, X_n$.

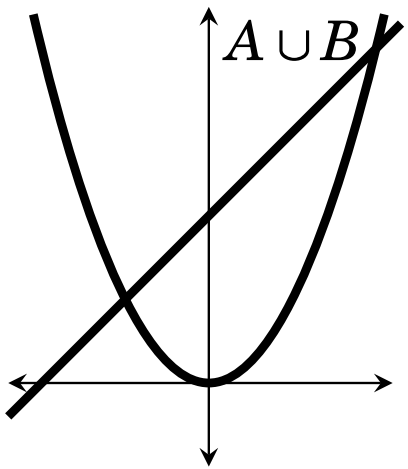
$$R \subset \prod_{i=1}^{n} X_i$$

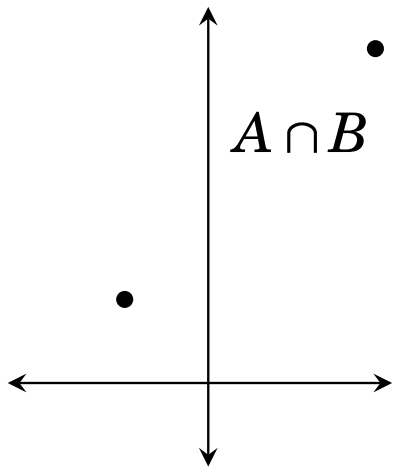
$$R(x_1, \ldots, x_n) \coloneqq (x_1, \ldots, x_n) \in R$$

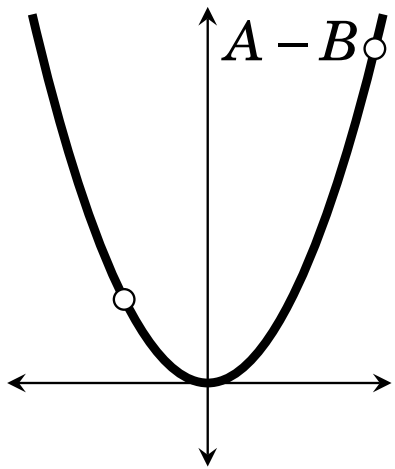
Union & Intersection & Difference

$$A := \{(x, x^2) : x \in \mathbb{R}\} \qquad y = x^2$$
$$B := \{(x, x+2) : x \in \mathbb{R}\} \qquad y = x + 2$$





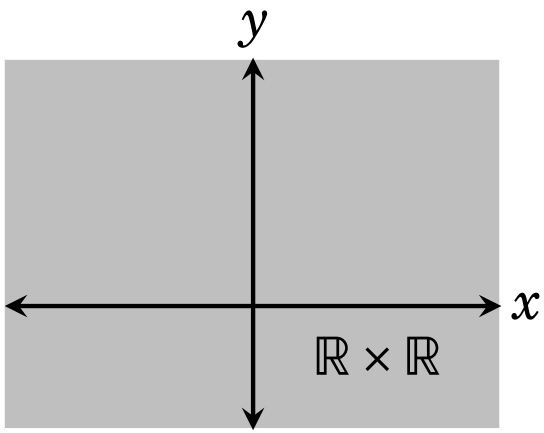


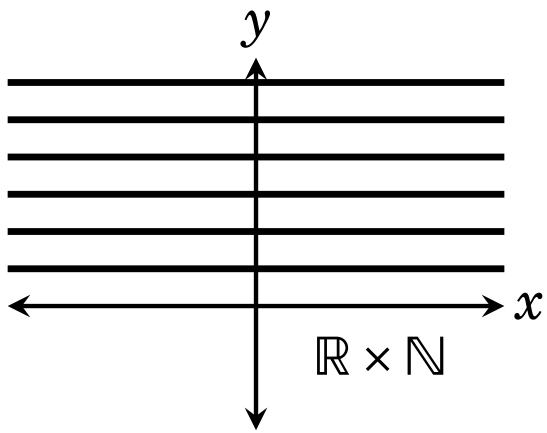


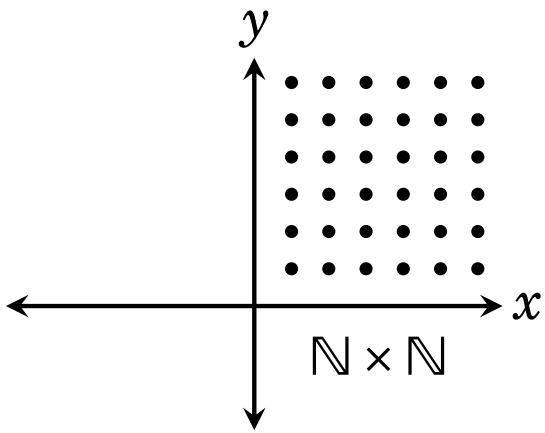
$$A \cup B = \{(x, y) : x \in \mathbb{R}, y = x^2 \vee y = x + 2\}$$
$$A \cap B = \{(-1, 1), (2, 4)\}$$
$$A \setminus B = \{(x, x^2) : x \in \mathbb{R} \setminus \{-1, 2\}\}$$

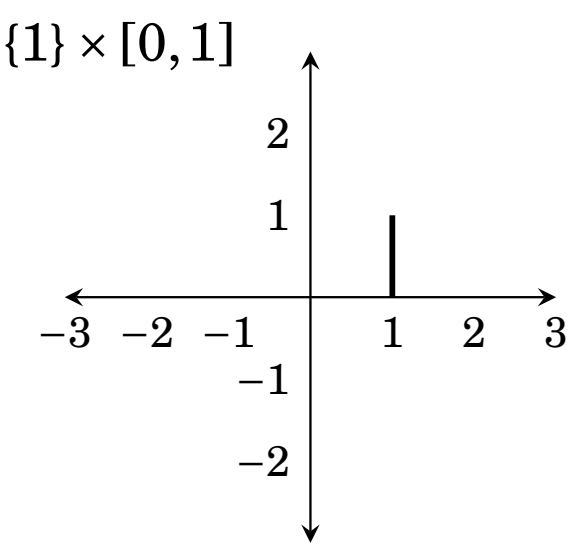
Cartesian Product

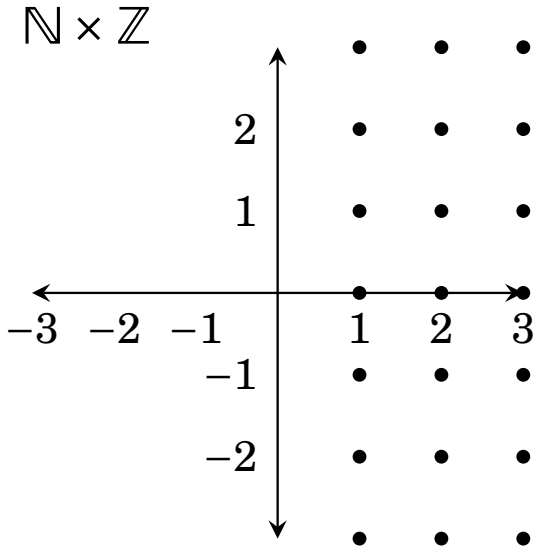
$$X \times Y \coloneqq \{(x, y) : x \in X \wedge y \in Y\}$$

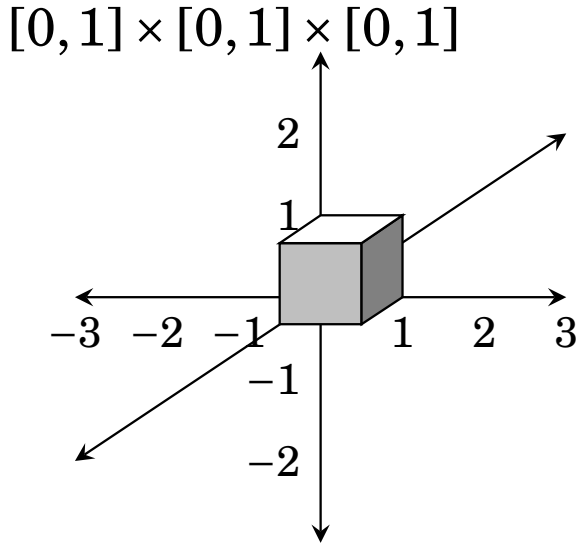










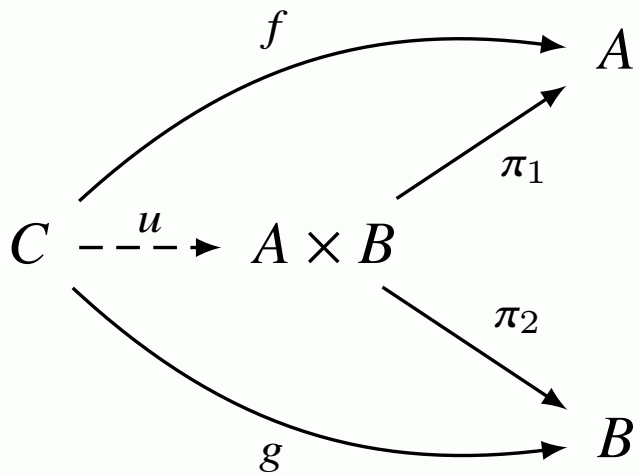


Cartesian Product — The Product in $\mathbf{Set}$

Definition (Product)

A *product* of A and B is an object $A \times B$ with a pair of morphisms $A \xleftarrow{\pi_1} A \times B \xrightarrow{\pi_2} B$ s.t

$$\forall C \forall f g : A \xleftarrow{f} C \xrightarrow{g} B \; \exists! u : C \to A \times B \big[f = \pi_1 \circ u \;\; \& \;\; g = \pi_2 \circ u \big]$$



- $A \times B$ is the set of $\{\{a\}, \{a, b\}\}$ with $a \in A, b \in B$, which happens to have that universal property.
- The universal property is the essense behind the idea of product, while $\{\{a\}, \{a, b\}\}$ is just one of many ways to make it work.

Equivalence Relation, Quotient, Partition

- $x \sim x$ (Reflexivity)
- $x \sim y \to y \sim x$ (Symmetry)
- $x \sim y \wedge y \sim x \to x \sim z$ (Transitivity)
- equivalence class: $[x] \coloneqq \{y \in X : x \sim y\}$
- quotient set: $X/\sim \coloneqq \{[x] : x \in X\}$
- we say $\mathcal{E} \subset \mathrm{P}(X)$ is a partition of X iff
  1. $\forall xy \in \mathcal{E} : x \neq y \to x \cap y = \emptyset$
  2. $\bigcup \mathcal{E} = X$
- $X/\sim$ is a partition of X.
- $R \subset X^2$ is an equivalence relation iff there is a partition $\mathcal{E}$ of X s.t $R(x,y) \iff \exists A \in \mathcal{E} : x, y \in A$.

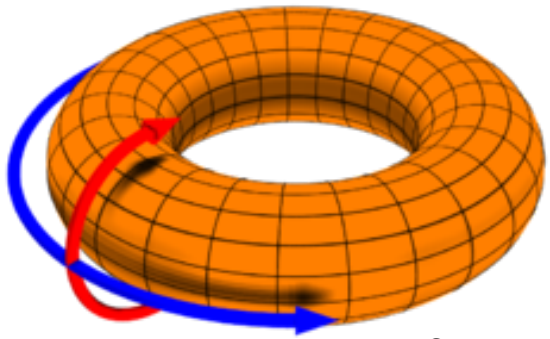
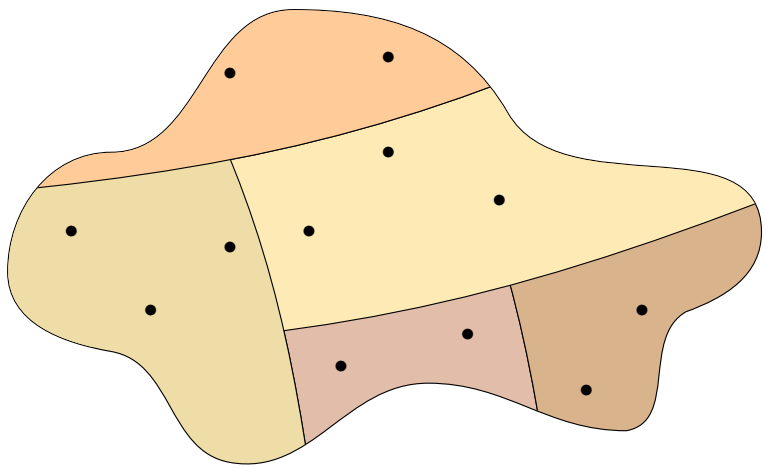


Figure: torus $\mathbb{R}^2/\sim$

$$(x,y) \sim (x',y') \coloneqq (x-x', y-y') \in \mathbb{Z}^2$$



Equivalence Relation — Exercise

Definition (Ideal)

A subset $I \subset \mathrm{P}(X)$ is an ideal, iff

1. $I \neq \emptyset$
2. $A \subset B \in I \rightarrow A \in I$
3. $A, B \in I \rightarrow A \cup B \in I$

Exercise

If I is an ideal, then $R := \{(A, B) : A \Delta B \in I\}$ is an equivalence relation.

Function

- A n-ary operation $f : \prod_{i=1}^{n} X_i \to Y$ is a function iff

$$(\mathbf{x}, y) \in f \wedge (\mathbf{x}, z) \in f \to y = z$$

- injection (one-to-one). $f : X \rightarrowtail Y$

$$f(x) = f(y) \to x = y$$

- surjection (onto). $f : X \twoheadrightarrow Y$.

$$\forall y \in Y \exists x \in X (f(x) = y)$$

- bijection. $f : X \rightarrowtail\!\!\!\!\twoheadrightarrow Y$
- restriction. composition. image. inverse image. inverse function.

$$f\restriction_A := \{(x, y) \in f : x \in A\} \quad (g \circ f)(x) := g(f(x))$$

$$f(A) := \{f(x) : x \in A\} \quad f^{-1}(A) := \{x : f(x) \in A\}$$

Remark

- 这里讲的从集合到集合的函数通常称为映射 (mapping).
- 通常讲的函数 (function) 是数域到数域的映射.
- 通常讲的泛函 (functional) 是函数空间到数域的映射.

$$f \mapsto \int_a^b f(t)\,\mathrm{d}t$$

$$\|\mathbf{x}\|_p = \sqrt[p]{\int_a^b |f(x)|^p\,\mathrm{d}x}$$

- 通常讲的算子 (operator) 是函数空间到函数空间的映射.

$$\frac{\mathrm{d}}{\mathrm{d}x} : f \mapsto \frac{\mathrm{d}}{\mathrm{d}x} f$$

$$V(f)(x) = \int_a^x f(t)\,\mathrm{d}t$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Image & Inverse Image

- Let $f : X \to Y$ be a function, and let $(A_i)_{i \in I}$ be a family of subsets of X. Then

$$f\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} f(A_i) \quad \text{and} \quad f\left(\bigcap_{i \in I} A_i\right) \subset \bigcap_{i \in I} f(A_i)$$

  The direct image function preserves only unions.

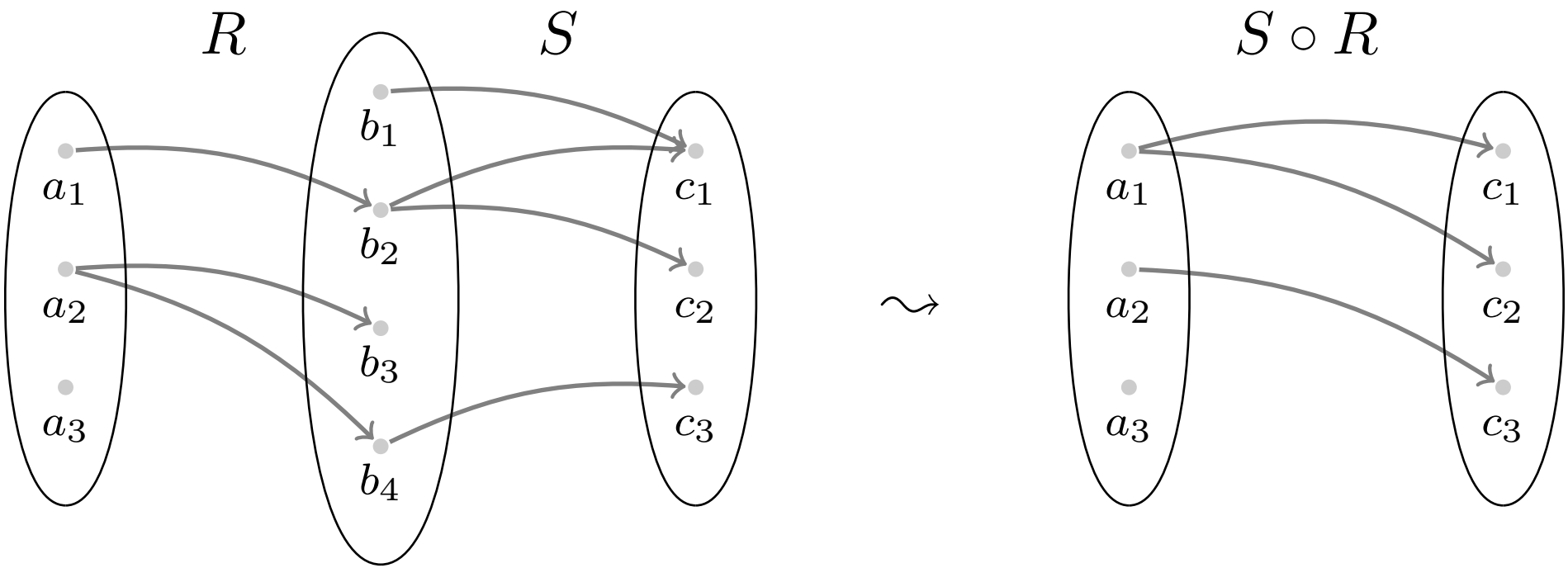
- Let $f : X \to Y$ be a function, and let $(B_i)_{i \in I}$ be a family of subsets of Y. Then

$$f^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f^{-1}(B_i) \quad \text{and} \quad f^{-1}\left(\bigcap_{i \in I} B_i\right) = \bigcap_{i \in I} f^{-1}(B_i)$$

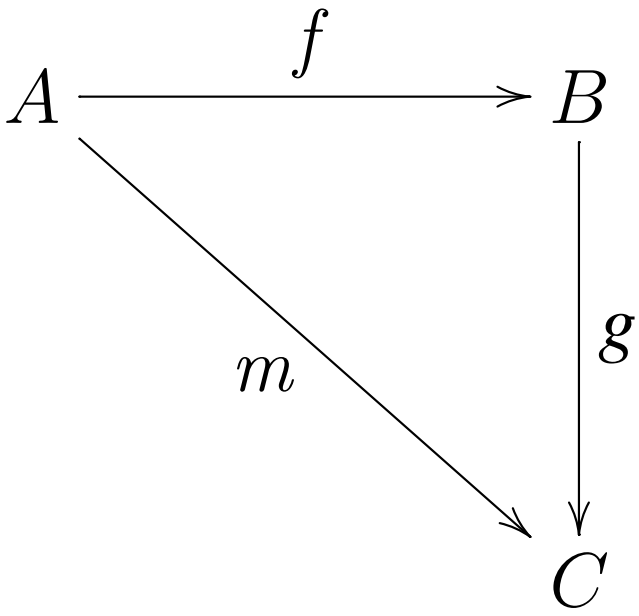
  The inverse image function preserves both unions and intersections.

- $A \subset f^{-1}(f(A))$ for all $A \subset X$, and $f(f^{-1}(B)) \subset B$ for all $B \subset Y$.

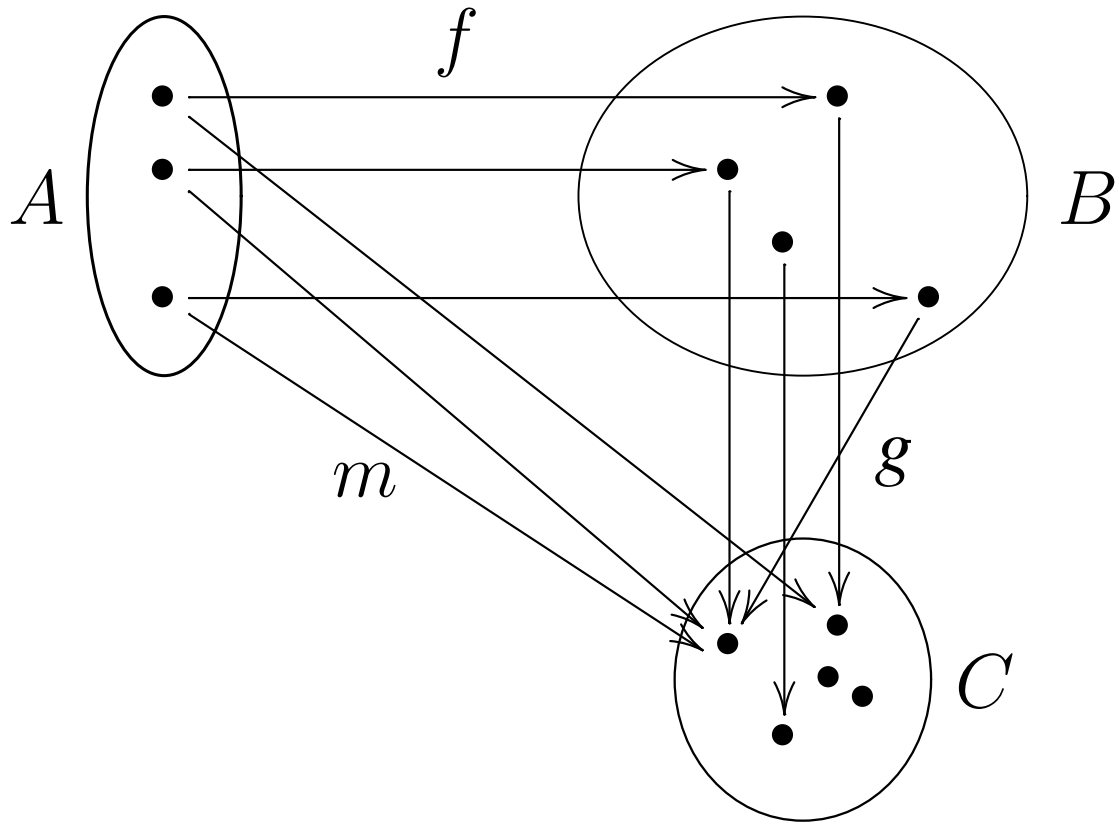
Relation Composition



$$S \circ R := \{(a, c) : \exists b(Rab \wedge Sbc)\}$$



External Diagram



Internal Diagram

用矩阵表示关系

- 关系 $R \subset X \times Y$ 可以用矩阵 $M_R = [m_{ij}]$ (或 $M_R : X \times Y \to \mathbb{Z}_2$) 表示, 其中

$$m_{ij} = [\![(x, y) \in R]\!]$$

- 关系的并、交、复合的矩阵

$$M_{R \cup S} = M_R \vee M_S \qquad M_{R \cap S} = M_R \wedge M_S \qquad M_{S \circ R} = M_R \odot M_S$$

其中

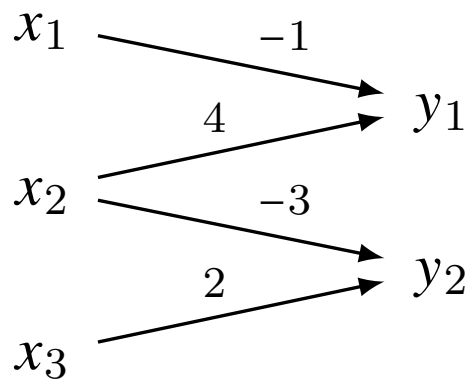
$$M \vee N = [m_{ij} \vee n_{ij}] \qquad M \wedge N = [m_{ij} \wedge n_{ij}]$$

$$M \odot N = \left[\bigvee_k (m_{ik} \wedge n_{kj}) \right]$$

关系、矩阵、概率、图

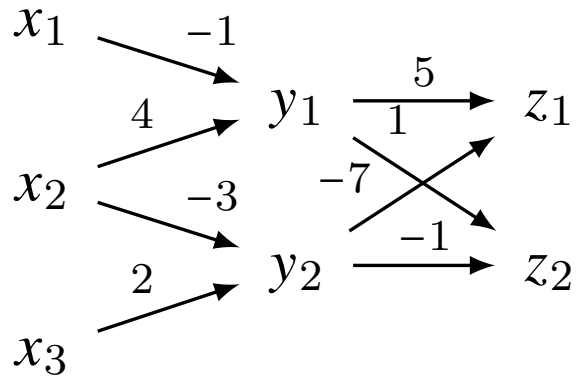
- Every matrix corresponds to a weighted bipartite graph.

$$M : X \times Y \to \mathbb{R} \qquad \begin{bmatrix} -1 & 0 \\ 4 & -3 \\ 0 & 2 \end{bmatrix}$$



- Symmetric matrices correspond to symmetric graphs.
- Matrix multiplication $M : X \times Y \to \mathbb{R}$ and $N : Y \times Z \to \mathbb{R}$ corresponds to traveling along paths.

$$\begin{bmatrix} -1 & 0 \\ 4 & -3 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ -7 & -1 \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ 41 & 7 \\ -14 & -2 \end{bmatrix}$$

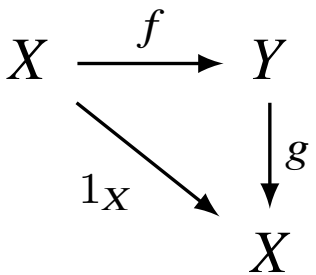


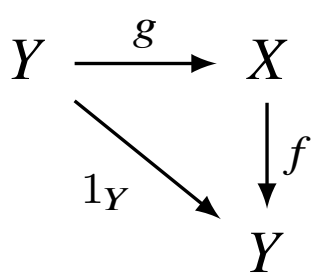
- Joint Probability $P(x_i, y_j)$, Marginal Probability $P(x_i)$ or $P(y_j)$

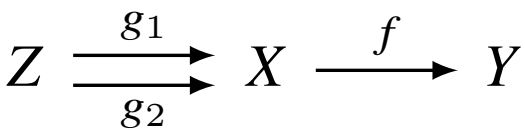
Exercises

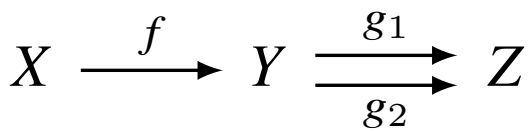
In **Set**,

$$f : X \rightarrowtail Y \quad \text{iff} \quad \exists g : Y \to X : gf = 1_X$$
$$\text{iff} \quad \forall Z \forall g_1 g_2 : Z \to X : fg_1 = fg_2 \implies g_1 = g_2$$
$$f : X \twoheadrightarrow Y \quad \text{iff} \quad \exists g : Y \to X : fg = 1_Y$$
$$\text{iff} \quad \forall Z \forall g_1 g_2 : Y \to Z : g_1 f = g_2 f \implies g_1 = g_2$$

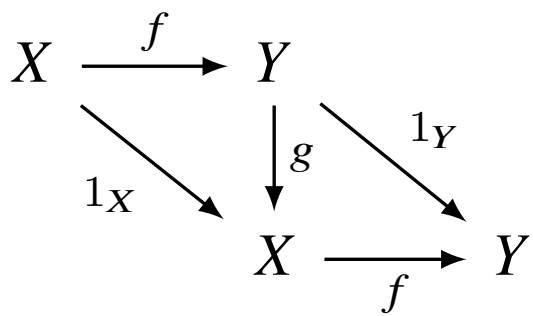








$f : X \rightarrowtail\!\!\!\!\twoheadrightarrow Y$ iff the diagram commutes:



Relation & Function

A relation $R \subset X \times Y$ is

- total-valued iff $1_X \subset R^{-1} \circ R$
- single-valued iff $R \circ R^{-1} \subset 1_Y$
- injective iff $R^{-1} \circ R \subset 1_X$
- surjective iff $1_Y \subset R \circ R^{-1}$

A relation $R \subset X \times X$ is

- reflexive iff $1_X \subset R$
- symmetric iff $R = R^{-1}$
- transitive iff $R \circ R \subset R$
- anti-symmetric iff $R \cap R^{-1} \subset 1_X$
- total iff $R \cup R^{-1} = X \times X$

- A partial function is a relation that is single-valued.
- A function is a partial function that is total-valued, i.e., a relation that is total-valued and single-valued.

Equivalence Relation / Partition / Transformation Group

Definition (Transformation Group)

A *transformation group* G on some set X is a set of invertible functions $f : X \to X$ which is closed under inversion and composition.

1. if $f \in G$, then its inverse $f^{-1} \in G$.
2. if $f, g \in G$, then their coposition $g \circ f \in G$.

Obviously, the identity transformation $1_X \in G$.

i. The effect of composition $g \circ f$ is

$$\text{first do } f\text{, then do } g$$

ii. To undo the effect of $g \circ f$,

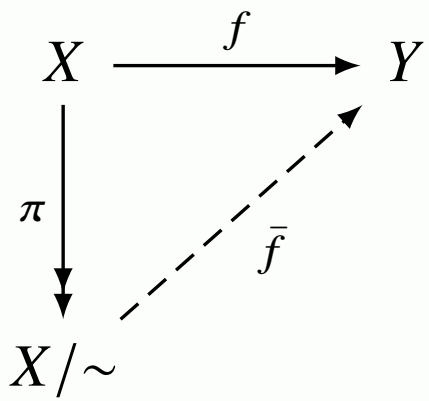
$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

$$\text{first do } g^{-1}\text{, then do } f^{-1}$$

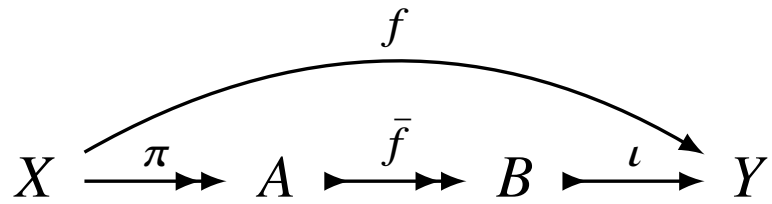
- Let $R(x, y) := \exists f \in G : y = f(x)$. Then R is an equivalence relation.
- Conversely, suppose R is an equivalent relation, then there is a group G s.t. $R(x, y) \iff \exists f \in G : y = f(x)$.

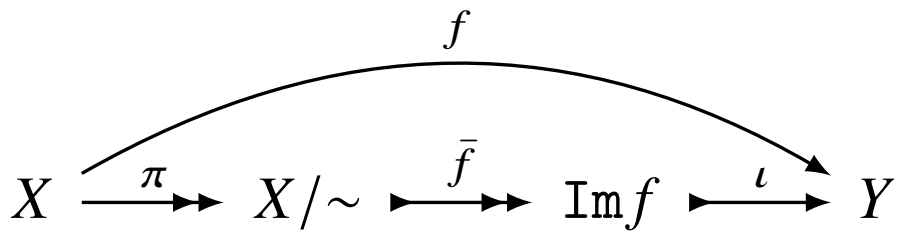
Theorem (Universal Property of Quotient Set)

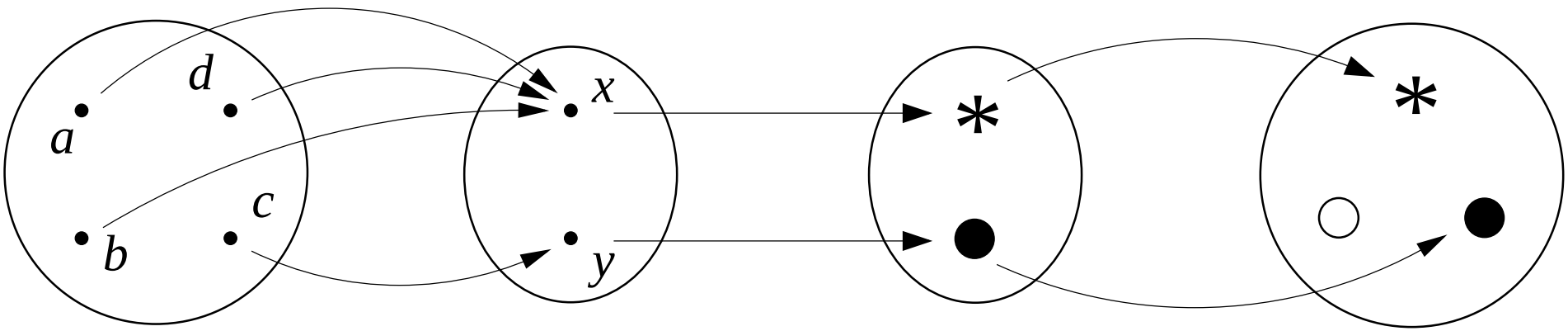
Suppose $f : X \to Y$ *is a surjection. Define an equivalence relation* $\sim$ *on* X *as* $x \sim x' \iff fx = fx'$. *Then there is a unique map* $\bar{f} : X/\sim \to Y$ *s.t.* $\bar{f} \circ \pi = f$, *where* π *is the canonical map* $x \mapsto [x]$.



For any function $f : X \to Y$, there exist two sets A and B, along with a surjection $\pi : X \twoheadrightarrow A$, and an injection $\iota : B \rightarrowtail Y$, and a bijection $\bar{f} : A \rightarrowtail\!\!\!\!\twoheadrightarrow B$, such that $f = \iota \circ \bar{f} \circ \pi$.

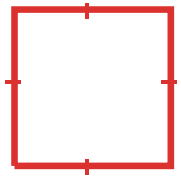




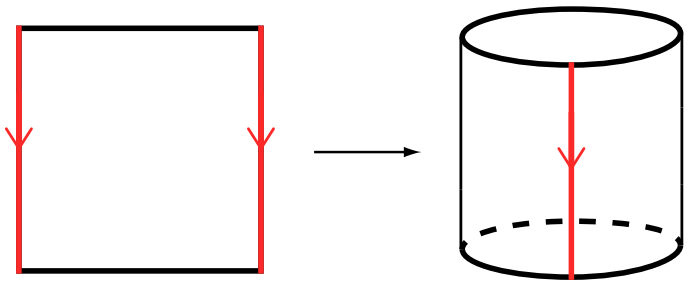


商

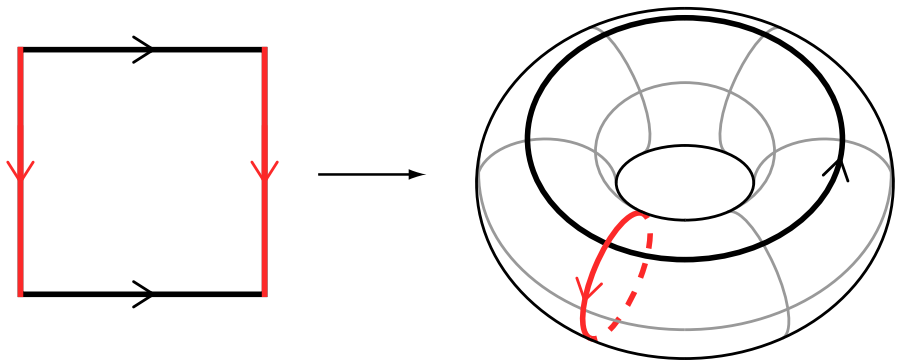
- 圆周 S^1: $\mathbb{R}/\mathbb{Z} = \{[x] : x \in [0,1)\} \cong S^1 \cong [0,1]/\{0,1\}$
- 二维球面 $S^2 \cong D^2/S^1$: $[0,1]\times[0,1]/(\{0,1\}\times[0,1] \cup [0,1]\times\{0,1\})$



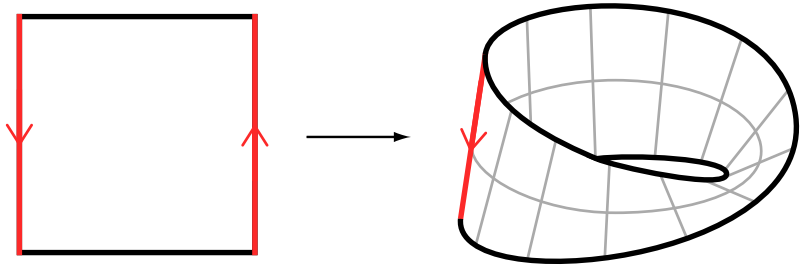
- 圆柱面: $[0,1]\times[0,1]/\{(0,y)\sim(1,y) : 0 \le y \le 1\} \cong S^1 \times [0,1]$



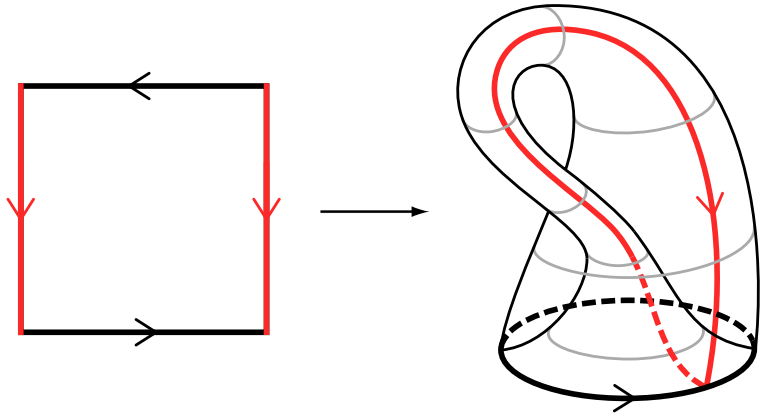
- 环面: $[0,1]\times[0,1]/\{(x,0)\sim(x,1),(0,y)\sim(1,y) : 0 \le x,y \le 1\} \cong S^1 \times S^1 \cong \mathbb{R}/\mathbb{Z}\times\mathbb{R}/\mathbb{Z} = (\mathbb{R}\times\mathbb{R})/(\mathbb{Z}\times\mathbb{Z})$



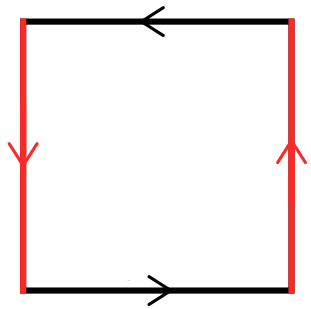
- 莫比乌斯带: $[0,1]\times[0,1]/\{(0,y)\sim(1,1-y):0\le y\le 1\}$



- 克莱因瓶:
  $[0,1]\times[0,1]/\{(x,0)\sim(1-x,1),(0,y)\sim(1,y):0\le x,y\le 1\}$



- 射影平面 $\mathbb{R}P^2$:
  $[0,1]\times[0,1]/\{(x,0)\sim(1-x,1),(0,y)\sim(1,1-y):0\le x,y\le 1\}$



Digression

- 将两条莫比乌斯带沿着边缝在一起，即是克莱因瓶.
- 将莫比乌斯带沿着中线剪开，会得到一条全转的带子，长度是原来的两倍.
- 全转的带子与圆柱面带子拓扑等价.
- 将圆柱面带子沿着中线剪开，会得到跟原来一样长的两条分开的圆柱面带子.
- 将全转的带子沿着中线剪开，会得到跟原来一样长的两条扣在一起的全转的带子.
- 既然全转的带子与圆柱面带子拓扑等价，为什么剪开后的结果却有差别？— 差别源自带子被嵌入三维空间的方式.

Order

- preorder: $x \leq x, x \leq y \wedge y \leq z \rightarrow x \leq z$.
- partial order: preorder with $x \leq y \wedge y \leq x \rightarrow x = y$.
- strict partial order: $x \not< x, x < y \wedge y < z \rightarrow x < z$.
- total (or linear) order: partial order with $x \leq y \vee y \leq x$.
- A total order of P is a *well order* iff every non-empty subset of P has a least element.

Definition

If $(P, \leq)$ is a partially ordered set, $X \subset P$, and $a \in P$, then:

- a is a *maximal* element of X iff $a \in X \wedge \forall x \in X(a \leq x \rightarrow a = x)$;
- a is a *greatest* element of X iff $a \in X \wedge \forall x \in X(x \leq a)$;
- a is an *upper bound* of X iff $\forall x \in X(x \leq a)$;
- a is the *supremum* of X iff a is the least upper bound of X.

Remark

- “greatest” $\implies$ “maximal”
- “upper bound” may not exist.
- when “greatest” exists,

$$\text{“greatest”} = \text{“supremum”}$$

- If X is total and “maximal” exists,

$$\text{“greatest”} = \text{“maximal”} = \text{“supremum”}$$

ZFC — Axioms

- **Replacement Schema**.
  If a class F is a function, then for every set X, $F(X)$ is a set.

$$\forall xyz(A(x,y) \wedge A(x,z) \rightarrow y = z) \rightarrow \forall X \exists Y \forall y(y \in Y \leftrightarrow \exists x \in X A(x,y))$$

- **Axiom of Foundation**. Every non-empty set has an $\in$-minimal element.

$$\forall X(X \neq \emptyset \rightarrow \exists x(x \in X \wedge X \cap x = \emptyset))$$

- **Axiom of Infinity**.

$$\exists X(\emptyset \in X \wedge \forall x(x \in X \rightarrow x \cup \{x\} \in X))$$

- **Axiom of Choice** (AC). For any set X of non-empty sets, there exists a choice function f defined on X.

$$\forall X \left[\emptyset \notin X \rightarrow \exists f : X \rightarrow \bigcup X \;\; \forall A \in X \, (f(A) \in A)\right]$$

集合 vs 类

- 有些类是集合, 有些 “太大” 或 “太不规则” 以至于不是集合.
- 不是集合的类, 称为 “真类”.
- 哪些类是集合? — 一个类是集合, 当且仅当, 它可以作为另一个类的元素.

对于 $\mathscr{L}_{\in}$ 中的任意公式 A, $\{x : A(x)\}$ 是一个满足 A 的对象组成的类. 类符号是元语言中的符号, 被视作公式的缩写. 关于类的表述可以被递归地消去:

- “$y \in \{x : A(x)\}$” 即为 $A[y/x]$
- “$\{x : A(x)\} = \{x : B(x)\}$” 即为 $\forall x(A(x) \leftrightarrow B(x))$
- “$z = \{x : A(x)\}$” 即为 $\forall y(y \in z \leftrightarrow A[y/x])$
- “$\{x : A(x)\} \in \{x : B(x)\}$” 即为 $\exists y(y \in \{x : B(x)\} \wedge y = \{x : A(x)\})$
- “$\{x : A(x)\} \in y$” 即为 $\exists z(z \in y \wedge z = \{x : A(x)\})$
- “$\{x : A(x)\}$ 是一个集合” 即为 $\exists y(y = \{x : A(x)\})$

Set vs Well-Founded Rigid Accessible Pointed Tree

- A graph is **pointed** if it is equipped with a specified node called the root.
- A pointed graph is **accessible** if for every node x there exists a directed path to the root.
- A set S of nodes in a graph is **inductive** if whenever all children of some node x are in S, then x itself is in S.
- A graph is **well-founded** if the only inductive set of nodes is the set of all nodes.
- A **tree** is a pointed graph that every node admits a unique directed path to the root.[25]

1. Extensionality: The rooted tree is rigid.
2. Well-Foundedness: Every branch is finite.

[25]集合论里的 tree 是有向的. 而图论里的 tree 指无向的连通无环图.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory
- Axioms of ZFC
- Ordinal Numbers
- Cardinal Numbers
- Axiom of Choice

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Ordinal vs Cardinal

- ordinal. ("length") A set is an ordinal iff it is transitive and well-ordered by $\in$. or equivalently,

$$\mathrm{Ord}(x) \;:=\; \bigcup x \subset x \wedge \forall yz(y \in x \wedge z \in x \to y \in z \vee y = z \vee z \in y)$$

- cardinal. ("size") $\mathrm{Card}(x) \;:=\; \mathrm{Ord}(x) \wedge \forall y \in x(|y| \neq |x|)$

$$|M| = |N| \;:=\; \exists f : M \twoheadrightarrow N \qquad |M| \leq |N| \;:=\; \exists f : M \rightarrowtail N$$

$$|M| := \min\{\alpha \in \mathrm{Ord} : |\alpha| = |M|\}$$

The infinite ordinal numbers that are cardinals are called alephs.

Ordinal

$$\alpha < \beta := \alpha \in \beta$$

- $\emptyset$ is an ordinal.
- If α is an ordinal and $\beta \in \alpha$, then β is an ordinal.
- If $\alpha \neq \beta$ are ordinals and $\alpha \subset \beta$, then $\alpha \in \beta$.
- If α, β are ordinals, then either $\alpha \subset \beta$ or $\beta \subset \alpha$.
- $<$ is a linear ordering of the class *Ord*.
- For each α, $\alpha = \{\beta : \beta < \alpha\}$.
- If C is a non-empty class of ordinals, then $\bigcap C$ is an ordinal, $\bigcap C \in C$ and $\bigcap C = \inf C$.
- If X is a non-empty set of ordinals, then $\bigcup X$ is an ordinal, $\bigcup X = \sup X$.
- For every α, $\alpha \cup \{\alpha\}$ is an ordinal and $\alpha \cup \{\alpha\} = \inf\{\beta : \beta > \alpha\}$.

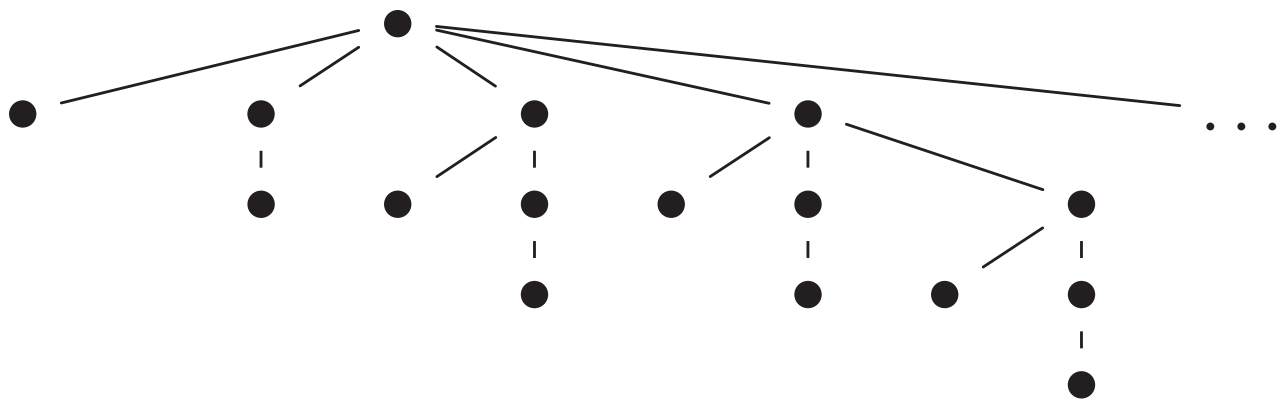
Natural Number $\mathbb{N}$

What is "number"? What is "infinity"? What is beyond "infinity"?

$$\alpha + 1 := \alpha \cup \{\alpha\}$$

$$0 := \emptyset, \quad 1 := 0 + 1, \quad 2 := 1 + 1, \quad 3 := 2 + 1, \ldots$$

$$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots\}$$



- A set A is inductive iff $\emptyset \in A$ and $\forall x \in A : x + 1 \in A$.
- A natural number is a set that belongs to every inductive set.

$$\mathbb{N} := \left\{ n : \forall A \left(\emptyset \in A \wedge \forall x \in A (x + 1 \in A) \rightarrow n \in A \right) \right\}$$

- A set A is finite iff $\exists n \in \mathbb{N} : |A| = n$.
- A set A is countable iff $|A| \leq |\mathbb{N}|$.

Integer $\mathbb{Z}$

$$(m,n) \sim (a,b) \coloneqq m +_{\mathbb{N}} b = a +_{\mathbb{N}} n$$

$$\mathbb{Z} \coloneqq \mathbb{N} \times \mathbb{N} / \sim$$

$$0_{\mathbb{Z}} \coloneqq [(0,0)]$$

$$[(m,n)] \leq_{\mathbb{Z}} [(a,b)] \coloneqq m +_{\mathbb{N}} b \leq_{\mathbb{N}} a +_{\mathbb{N}} n$$

$$[(m,n)] +_{\mathbb{Z}} [(a,b)] \coloneqq [(m +_{\mathbb{N}} a, n +_{\mathbb{N}} b)]$$

$$[(m,n)] \cdot_{\mathbb{Z}} [(a,b)] \coloneqq [(m \cdot_{\mathbb{N}} a + n \cdot_{\mathbb{N}} b, m \cdot_{\mathbb{N}} b + n \cdot_{\mathbb{N}} a)]$$

$$- [(m,n)] \coloneqq [(n,m)]$$

$$\mathbb{Z}^{+} \coloneqq \{x \in \mathbb{Z} : x >_{\mathbb{Z}} 0_{\mathbb{Z}}\}$$

$$\exists f : \mathbb{N} \to \mathbb{Z} \quad n \mapsto [(n,0)]$$

Rational Number $\mathbb{Q}$

$$(m, n) \sim (a, b) := m \cdot_{\mathbb{Z}} b = a \cdot_{\mathbb{Z}} n$$

$$\mathbb{Q} := \mathbb{Z} \times \mathbb{Z}^{+} / \sim$$

$$0_{\mathbb{Q}} := [(0_{\mathbb{Z}}, x)]$$

$$1_{\mathbb{Q}} := [(x, x)]$$

$$[(m, n)] \leq_{\mathbb{Q}} [(a, b)] := m \cdot_{\mathbb{Z}} b \leq_{\mathbb{Z}} a \cdot_{\mathbb{Z}} n$$

$$[(m, n)] +_{\mathbb{Q}} [(a, b)] := [(m \cdot_{\mathbb{Z}} b +_{\mathbb{Z}} a \cdot_{\mathbb{Z}} n, n \cdot_{\mathbb{Z}} b)]$$

$$[(m, n)] \cdot_{\mathbb{Q}} [(a, b)] := [(m \cdot_{\mathbb{Z}} a, n \cdot_{\mathbb{Z}} b)]$$

$$- [(m, n)] := [(-m, n)]$$

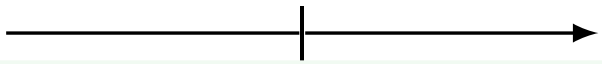
$$\exists f : \mathbb{Z} \to \mathbb{Q} \quad x \mapsto [(x, 1)]$$

Dedekind's Construction of Real Numbers $\mathbb{R}$

Definition (Dedekind Cut)

$\mathbb{R}$ is the set of all $x \in \mathsf{P}(\mathbb{Q})$ s.t.

- $x \neq \emptyset,\ x \neq \mathbb{Q}$
- $\forall a \in x \exists b \in x : a < b$
- $\forall ab \in x : a \in x \wedge b < a \rightarrow b \in x$



Theorem (Least-upper-bound)

Any bounded non-empty subset of $\mathbb{R}$ has a least upper bound.

$$\exists f : \mathbb{Q} \rightarrow \mathbb{R} \quad a \mapsto \{b \in \mathbb{Q} : b < a\}$$

$$x \leq_{\mathbb{R}} y := x \subset y$$

$$x +_{\mathbb{R}} y := \{a +_{\mathbb{Q}} b : a \in x \ \& \ b \in y\}$$

$$-x := \{a - b : a < 0 \ \& \ b \in \mathbb{Q} \setminus x\}$$

$$|x| := x \cup -x$$

$$x \cdot_{\mathbb{R}} y := \begin{cases} \{c : c \leq a \cdot_{\mathbb{Q}} b \ \& \ a \in x \ \& \ b \in y\} & \text{if } x > 0, y > 0 \\ 0 & \text{if } x = 0 \text{ or } y = 0 \\ |x| \cdot_{\mathbb{R}} |y| & \text{if } x < 0, y < 0 \\ -(|x| \cdot_{\mathbb{R}} |y|) & \text{if } x < 0, y > 0 \text{ or } x > 0, y < 0 \end{cases}$$

Cantor's Construction of Real Numbers $\mathbb{R}$

- Cauchy sequence: a sequence of rational numbers (a_i) is Cauchy iff $\forall \varepsilon > 0 \exists N \forall mn > N : |a_m - a_n| < \varepsilon$.
- Two Cauchy sequences are equivalent $(a_i) \sim (b_i)$ iff $\lim_{n\to\infty} |a_n - b_n| = 0$.
- $+, \cdot, <$ can be defined as follows:

$$(a_i) + (b_i) = (a_i + b_i)$$
$$(a_i) \times (b_i) = (a_i \times b_i)$$
$$(a_i) < (b_i) \iff \exists \varepsilon > 0 \exists N \forall n > N : a_i + \varepsilon < b_i$$

- A real number is an equivalence class of Cauchy sequences.

Remark: The usual decimal notation can be translated to Cauchy sequences in a natural way.
The equation $0.999\cdots = 1$ states that the sequences $(0, 0.9, 0.99, 0.999, \dots)$ and $(1, 1, 1, 1, \dots)$ are equivalent.

Theorem

A sequence of real numbers is convergent (in the reals) iff it is Cauchy.

$\mathbb{R}$ is Cauchy sequences (in $\mathbb{Q}^{\mathbb{N}}$) or Dedekind cuts (in $2^{\mathbb{Q}}$).

Hilbert's Axiomatic Definition of Real Numbers $\mathbb{R}$

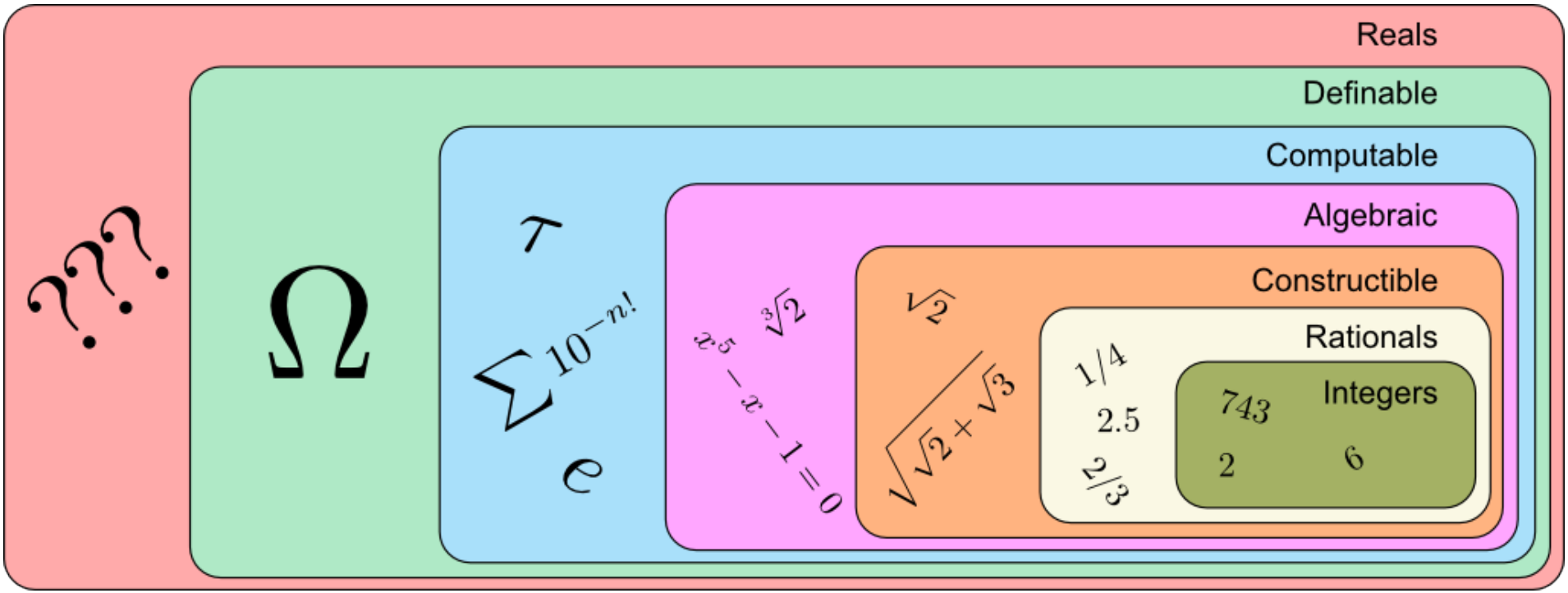
- $(\mathbb{R}, 0, 1, +, \cdot)$ is a field.
- $(\mathbb{R}, \leq)$ is a total order.
- Preservation of order under addition and multiplication.

$$\forall xyz(x \leq y \rightarrow x + z \leq y + z)$$
$$\forall xy(0 \leq x \wedge 0 \leq y \rightarrow 0 \leq x \cdot y)$$

- $(\mathbb{R}, 0, 1, +, \cdot, \leq)$ is complete: every non-empty subset of $\mathbb{R}$ that is bounded above has a least upper bound.

Theorem

The complete totally ordered field is unique up to isomorphisms.



$$\Omega := \sum_{p:U(p)\downarrow} 2^{-\ell(p)}$$

Principle of Continuous Induction

Theorem (Principle of Continuous Induction)

If A is a set of nonnegative real numbers such that

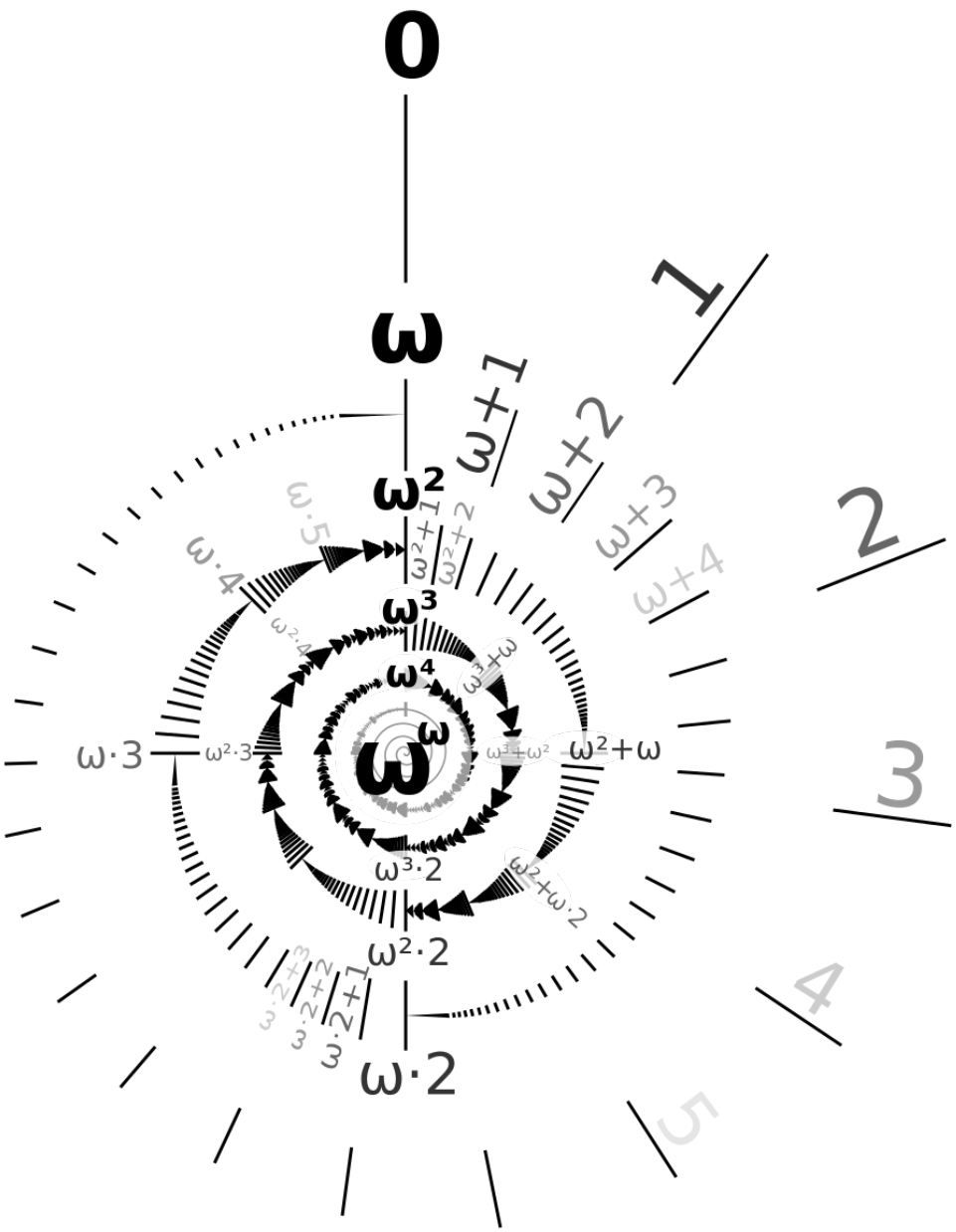
1. $0 \in A$;
2. *If $x \in A$, then there is $\delta > 0$ with $[x, x + \delta) \subset A$;*
3. *If x is a real number and $[0, x) \subset A$, then $x \in A$.*

Then A is the set of all nonnegative real numbers.

Proof.

It can be proved from the least-upper-bound principle.
Suppose there is some nonnegative real $r \notin A$.
Let x be the supremum of the set of numbers a for which $[0, a] \subset A$.
Since $0 \in A$, we know that $x \geq 0$.
Also, $[0, x) \subset A$, and so $x \in A$.
Then there is some $\delta > 0$ with $[x, x + \delta) \subset A$.
But in this case, the supremum x would be have been bigger. □

Ordinal



$0, 1, 2, 3, \ldots$

$\omega, \omega + 1, \omega + 2, \ldots$

$\omega \cdot 2, (\omega \cdot 2) + 1, (\omega \cdot 2) + 2, \ldots$

$\vdots$

$\omega^2, \omega^2 + 1, \omega^2 + 2, \ldots$

$\vdots$

$\omega^\omega, \omega^\omega + 1, \omega^\omega + 2, \ldots$

$\vdots$

$\omega^{\omega^\omega}, \ldots$

$\vdots$

$\omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}}, \ldots$

$\vdots$

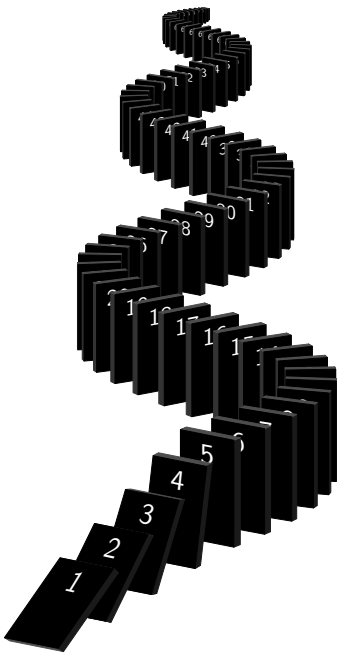
Ordinal and Induction

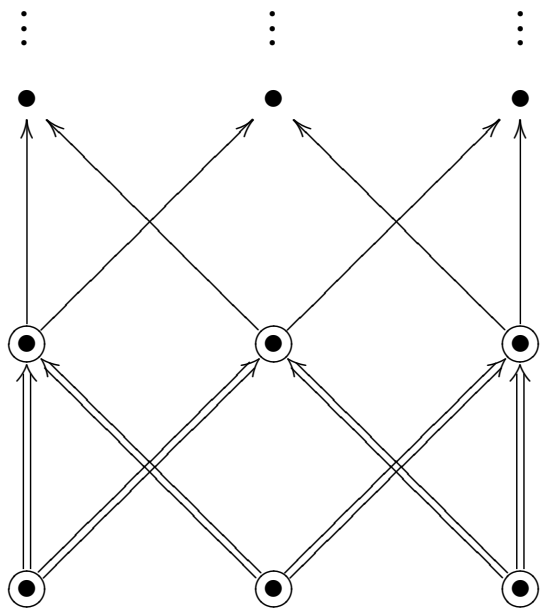
Theorem (Transfinite Induction Theorem)

Given a well ordered set A, let P be a property. Then

$$P(\min(A)) \wedge \forall x \in A[\forall y < x\, P(y) \rightarrow P(x)] \rightarrow \forall x \in A\, P(x)$$

$$\forall P[P(0) \wedge \forall k \in \mathbb{N}(P(k) \rightarrow P(k+1)) \rightarrow \forall n \in \mathbb{N} P(n)]$$



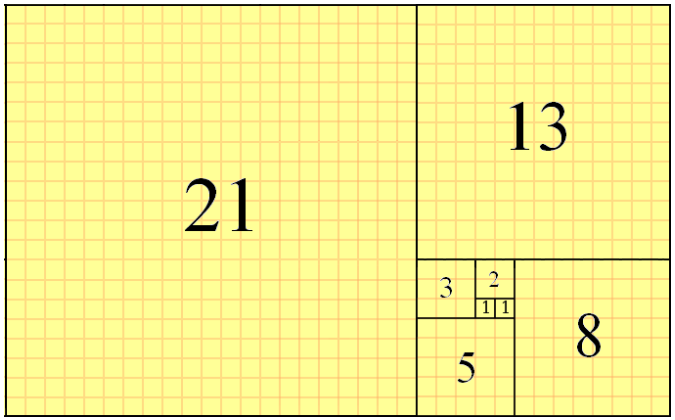


Induction — Example

Theorem

$$\sum_{i=0}^{n} F_i^2 = F_n F_{n+1}$$

where $F_0 = 1, F_1 = 1, F_n = F_{n-1} + F_{n-2}$.



Proof.

Base step:

$$F_0^2 = 1^2 = 1 = 1 \times 1 = F_0 F_1$$

Inductive step:

$$\sum_{i=0}^{n+1} F_i^2 = \sum_{i=0}^{n} F_i^2 + F_{n+1}^2 = F_n F_{n+1} + F_{n+1}^2 = F_{n+1}(F_n + F_{n+1}) = F_{n+1} F_{n+2}$$

□

Induction — Example

Theorem

$$F_n^2 + F_{n+1}^2 = F_{2n+2}$$

Proof.

Strengthen the original statement to

$$F_n^2 + F_{n+1}^2 = F_{2n+2} \quad \text{and} \quad F_{n+1}^2 + 2F_nF_{n+1} = F_{2n+3}$$

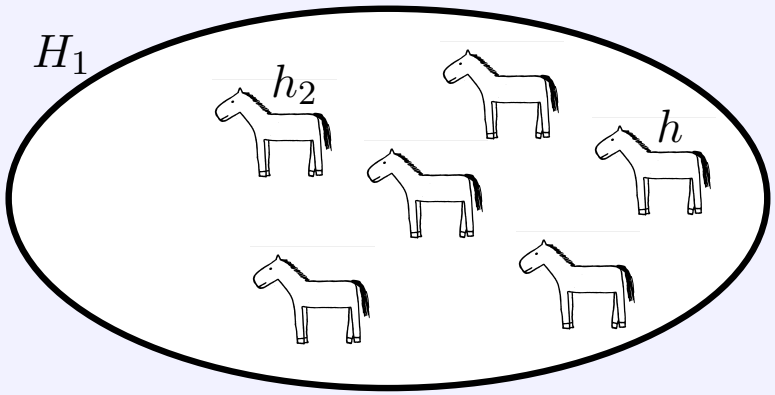
□

Problem

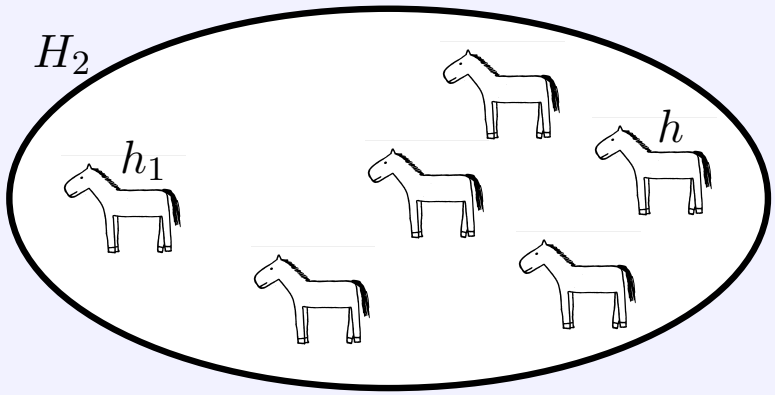
For all natural numbers n, $n \not< n$.

Induction — Example ☹

All horses have the same color ☹



All of the horses in H_1 have the same colour, so they all have the same colour as horse h.



All of the horses in H_2 have the same colour, so they all have the same colour as horse h.

Since every horse is in either H_1 or H_2, this implies that all of the horses have the same colour as h. So all of the horses have the same colour.

$$P(1) \overset{?}{\rightarrow} P(2)$$

All positive integers are interesting ☹

Assume the contrary. Then there is a lowest non-interesting positive integer. But that's pretty interesting!

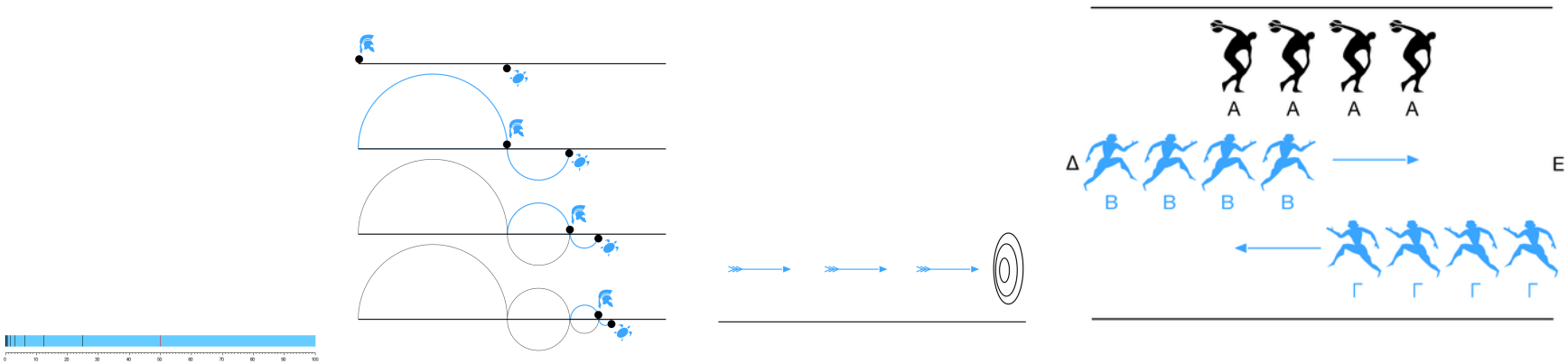
Induction — Example ☹

Surprise Exam Paradox ☹

A teacher announces in class: "next week you are going to have an exam, but you will not be able to know on which day of the week the exam is held until that day".
The students argue that a surprise exam can't occur:

- The exam can't be held on the last day, because otherwise, the night before the students will know that the exam is going to be held the next day.
- Since the last day has already been eliminated, the same logic applies to the day before the last day.
- Similarly, all the days can be removed from the list.
- So the teacher can't give a surprise exam at all.

Zeno Paradox



1. 二分悖论: 空间是连续的 $\Longrightarrow$ 运动是不可能的
2. 阿喀琉斯追不上乌龟: 空间是连续的 $\Longrightarrow$ 运动是不可能的
3. 飞矢不动: 时间是连续的 $\Longrightarrow$ 运动是不可能的
4. 运动场悖论: 时间和空间是离散的 $\Longrightarrow$ 矛盾



Ross-Littlewood Paradox — Hilbert's Train

- Suppose a train is empty at 1 minute before noon.
- At 2^{-n} minutes before noon, 10 passenger get on, and 1 gets off.
  1. the first gets off?
  2. the last gets off?
  3. randomly gets off?
- How many passengers are on the train at noon?

Proof.

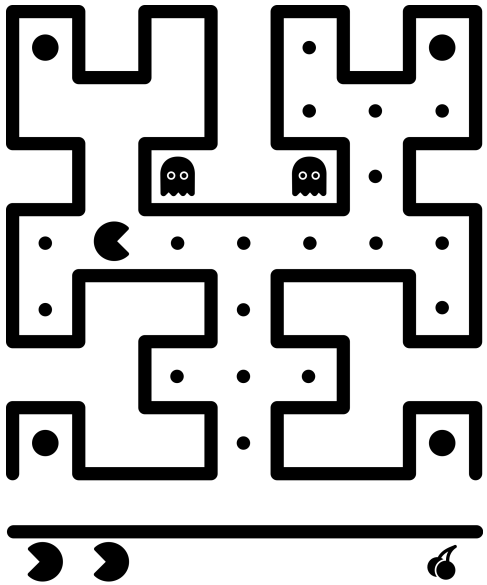
Define E_n to be the event that passenger 1 is still on the train after the first n station, and F_i the event that passenger i is on the train at noon.

$$P(F_1) = P\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} P(E_n) = \prod_{i=1}^{\infty} \frac{9n}{9n+1} = 0$$

$$\forall i : P(F_i) = 0 \implies P\left(\bigcup_{i=1}^{\infty} F_i\right) \leq \sum_{i=1}^{\infty} P(F_i) = 0$$

□

无穷吃豆人游戏



- 你有自然数 (实数) 那多么颗互相可区分的豆子.
- 12 点前 1 分钟, 桌上没有豆子.
- 在 12 点前 2^{-n} 分钟, 你可以往桌上添加任意有穷 (可数) 数量的豆子, 然后吃豆人选择桌上的某一颗豆子吃掉.
- 游戏在 12 点整结束. 桌面上如果还有豆子, 你胜; 如果没有豆子, 吃豆人胜.

Remark: 如果吃豆人记性不好, 不记得桌上的豆子是在哪一回合被你放上去的呢?

The Delayed Heaven Paradox

Problem (The Delayed Heaven Paradox)

- *Heaven:* 1 *every day for eternity.*
- *Hell:* -1 *every day for eternity.*
- *Limbo:* 0 *every day for eternity.*

God offers you the chance

1. *to go straight to Limbo, or*
2. *to take one day in Hell, followed by two days in Heaven, followed by the rest of eternity in Limbo.*

Suppose you die and the devil offers to play a game of chance. If you win, you can go to heaven. If you lose, you'll stay in hell forever. If you play today, you have $1/2$ chance of winning. Tomorrow $2/3$. Then $3/4, 4/5, 5/6, 6/7 \ldots$ Will you stay forever in hell in order to increase the chance of leaving it?

Transfinite Recursion Theorem

Theorem (Recursion Theorem)

Given a function $G : V^{<\omega} \to V$*, there exists a unique function* $F : \omega \to V$ *s.t.*

$$F_n = G(F\restriction_n) = G(F_0, \ldots, F_{n-1})$$

Theorem (Transfinite Recursion Theorem)

Given a class function $G : V \to V$*, there exists a unique function* $F : \mathrm{Ord} \to V$ *s.t.*

$$F(\alpha) = G(F\restriction_\alpha)$$

for each α*.*

Limit of a Sequence

Definition (Limit of a Sequence)

Let $\alpha > 0$ be a limit ordinal and let $\{\gamma_\xi : \xi < \alpha\}$ be a nondecreasing sequence of ordinals. We define the limit of the sequence by

$$\lim_{\xi \to \alpha} \gamma_\xi := \sup \{\gamma_\xi : \xi < \alpha\}$$

Definition (Normal Sequence)

A sequence of ordinals $\{\gamma_\xi : \xi \in \mathrm{Ord}\}$ is normal if it is increasing and continuous, i.e., for every limit γ_α:

$$\gamma_\alpha = \lim_{\xi \to \alpha} \gamma_\xi$$

Ordinal Addition

Let α and β be ordinals then $\alpha + \beta$ is the unique ordinal such that there is $h : S \to \alpha + \beta$ bijective and monotone, i.e., such that $x \leq y$ in S implies $hx \leq hy$ in $\alpha + \beta$ where $S = (\alpha \coprod \beta, \leq)$, the disjoint union of α and β, and $x \leq y$ in S iff $x \leq y$ in α, or $x \leq y$ in β, or $x \in \alpha$ and $y \in \beta$.

Remark: The intuition of $\alpha + \beta$: α followed by β.

Ordinal Multiplication

Let α and β be ordinals then $\alpha \cdot \beta$ is the unique ordinal such that there is $h : S \to \alpha \cdot \beta$ bijective and monotone, i.e., such that $x \leq y$ in S implies $hx \leq hy$ in $\alpha \cdot \beta$ where $S = (\coprod_{i \in \beta} \alpha, \leq)$, and $x \leq y$ in S iff $x \leq y$ in α, or $x \in \alpha_i$ and $y \in \alpha_j$ with $i < j$.

Remark: The intuition of $\alpha \cdot \beta$: β copies of α.

$$(A \otimes B, \leq)$$

$$(a_0, b_0) \leq (a_1, b_1) \iff b_0 < b_1 \text{ or, } b_0 = b_1 \ \wedge \ a_0 < a_1$$

Ordinal Arithmetic

Definition (Addition)

1. $\alpha + 0 = \alpha$
2. $\alpha + (\beta + 1) = \alpha + \beta + 1$
3. $\alpha + \beta = \lim_{\xi \to \beta} (\alpha + \xi)$ for limit $\beta > 0$

Definition (Multiplication)

1. $\alpha \cdot 0 = 0$
2. $\alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$
3. $\alpha \cdot \beta = \lim_{\xi \to \beta} \alpha \cdot \xi$ for limit $\beta > 0$

Definition (Exponentiation)

1. $\alpha^0 = 1$
2. $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$
3. $\alpha^\beta = \lim_{\xi \to \beta} \alpha^\xi$ for limit $\beta > 0$

$$
\begin{aligned}
\omega &= 0 < 1 < 2 < 3 < \cdots \\
\omega + 1 &= 0 < 1 < 2 < 3 < \cdots < \omega \\
1 + \omega &= \bullet < 0 < 1 < 2 < 3 < \cdots
\end{aligned}
$$

- $1 + \omega = \omega \neq \omega + 1$
- $2 \cdot \omega = \omega \neq \omega \cdot 2 = \omega + \omega$
- $(\omega + 1) \cdot 2 \neq \omega \cdot 2 + 1 \cdot 2$
- $(\omega \cdot 2)^2 \neq \omega^2 \cdot 2^2$
- $\beta < \gamma \to \alpha + \beta < \alpha + \gamma$
- $\alpha < \beta \to \exists \delta (\alpha + \delta = \beta)$
- $\beta < \gamma \wedge \alpha > 0 \to \alpha \cdot \beta < \alpha \cdot \gamma$
- $\beta < \gamma \wedge \alpha > 1 \to \alpha^\beta < \alpha^\gamma$
- $\alpha > 0 \to \forall \gamma \exists! \beta \exists! \rho < \alpha (\gamma = \alpha \cdot \beta + \rho)$
- $\alpha < \beta \to \alpha + \gamma \leq \beta + \gamma$
- $\alpha < \beta \to \alpha \cdot \gamma \leq \beta \cdot \gamma$
- $\alpha < \beta \to \alpha^\gamma \leq \beta^\gamma$
- $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$
- $\alpha^{\beta+\gamma} = \alpha^\beta \cdot \alpha^\gamma$
- $(\alpha^\beta)^\gamma = \alpha^{\beta \cdot \gamma}$

Ordinal Arithmetic

- $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$
- $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$
- $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$
- $\alpha^{\beta+\gamma} = \alpha^{\beta} \cdot \alpha^{\gamma}$
- $\alpha + \beta \neq \beta + \alpha$
- $\alpha \cdot \beta \neq \beta \cdot \alpha$
- $(\alpha + \beta) \cdot \gamma \neq \alpha \cdot \gamma + \beta \cdot \gamma$
- $(\alpha \cdot \beta)^{\gamma} \neq \alpha^{\gamma} \cdot \beta^{\gamma}$

Example

- $(\omega + 1) \cdot 2 = (\omega + 1) + (\omega + 1) = (\omega + (1 + \omega)) + 1 = (\omega + \omega) + 1 = \omega \cdot 2 + 1 \neq \omega \cdot 2 + 1 \cdot 2$
- $(\omega \cdot 2)^2 = (\omega \cdot 2) \cdot (\omega \cdot 2) = (\omega \cdot (2 \cdot \omega)) \cdot 2 = (\omega \cdot \omega) \cdot 2 = \omega^2 \cdot 2 \neq \omega^2 \cdot 2^2$

Cantor's Normal Form Theorem

Theorem (Cantor's Normal Form Theorem)

Every ordinal $\alpha > 0$ can be represented uniquely in the form

$$\alpha = \omega^{\beta_1} \cdot k_1 + \cdots + \omega^{\beta_n} \cdot k_n$$

where $n \geq 1, \alpha \geq \beta_1 > \cdots > \beta_n$, and $k_1, \ldots, k_n \in \mathbb{N}^+$.

Prime Ordinal

Definition

A ordinal $\alpha > 1$ is prime iff there are no ordinals $\beta, \gamma < \alpha$ s.t.

$$\alpha = \beta \cdot \gamma$$

There are three sorts of prime ordinals:

1. $2, 3, 5, 7, 11, \ldots$ (finite primes)
2. ω^{ω^α} for any $\alpha \in \mathrm{Ord}$. (limit primes)
3. $\omega^\alpha + 1$ for any $\alpha \in \mathrm{Ord} \setminus \{0\}$. (infinite successor primes)

Now I Know!

1. **C:** Hello **A** and **B**! I have given you each a different natural number. Who of you has the larger number?
2. **A:** I don't know.
3. **B:** Neither do I.
4. **A:** Even though you say that, I still don't know.
5. **B:** Still neither do I.
6. **A:** Alas, even now I do not know.
7. **B:** I regret that I also do not know.
8. **A:** Yet, I still do not know.
9. **B:** Aha! Now I know which has the larger number.
10. **A:** Then I know both our numbers.
11. **B:** Well, now I also know them.

Now I Know! — transfinite

1. **C:** I have given you each a different ordinal. Who has the larger one?
2. **A:** I don't know.
3. **B:** Neither do I.
4. **A:** I still don't know.
5. **B:** Still neither do I.
6. **C:** Well, I can tell you that no matter how long you continue that back-and-forth, you shall not come to know who has the larger number.
7. **A:** What interesting new information! But I still do not know.
8. **B:** And still neither do I.
9. **A:** Alas, even now I do not know!
10. **B:** I regret that I also do not know.
11. **C:** Let me say once again that no matter how long you continue that back-and-forth, you will not know who has the larger number.
12. **A:** Yet, I still do not know.
13. **B:** Aha! Now I know who has the larger ordinal.
14. **A:** Then I know both our ordinals.
15. **B:** Well, now I also know them.

Now I Know! — transfinite

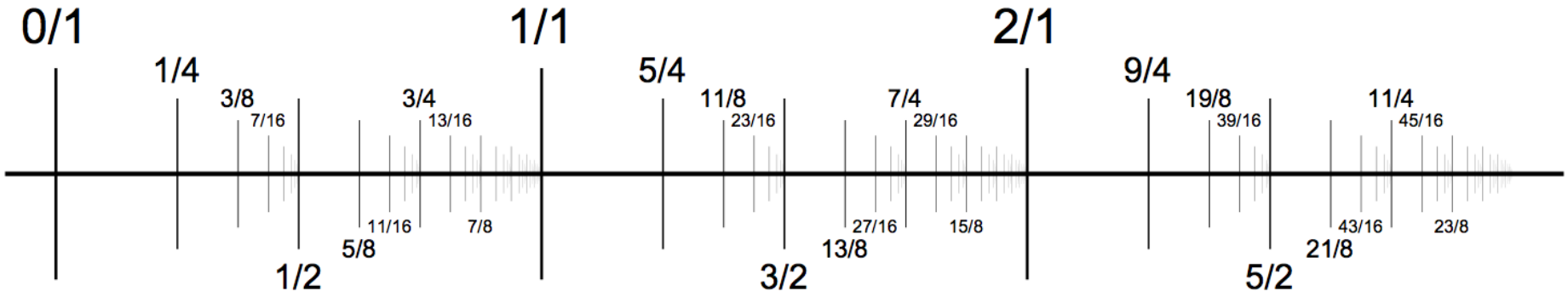
1. **C:** I have given you each a different rational number of the form
$$n - \frac{1}{2^k} - \frac{1}{2^{k+r}}$$
where $n, k \in \mathbb{N}^+$ and $r \in \mathbb{N}$. Who of you has the larger number?
2. **A:** I don't know.
3. **B:** Neither do I.
4. **A:** I still don't know.
5. **B:** Still neither do I.
6. **C:** Well, I can tell you that no matter how long you continue that back-and-forth, you shall not come to know who has the larger number.
7. **A:** What interesting new information! But I still do not know.
8. **B:** And still neither do I.
9. **A:** Alas, even now I do not know!
10. **B:** I regret that I also do not know.
11. **C:** Let me say once again that no matter how long you continue that back-and-forth, you will not know who has the larger number.

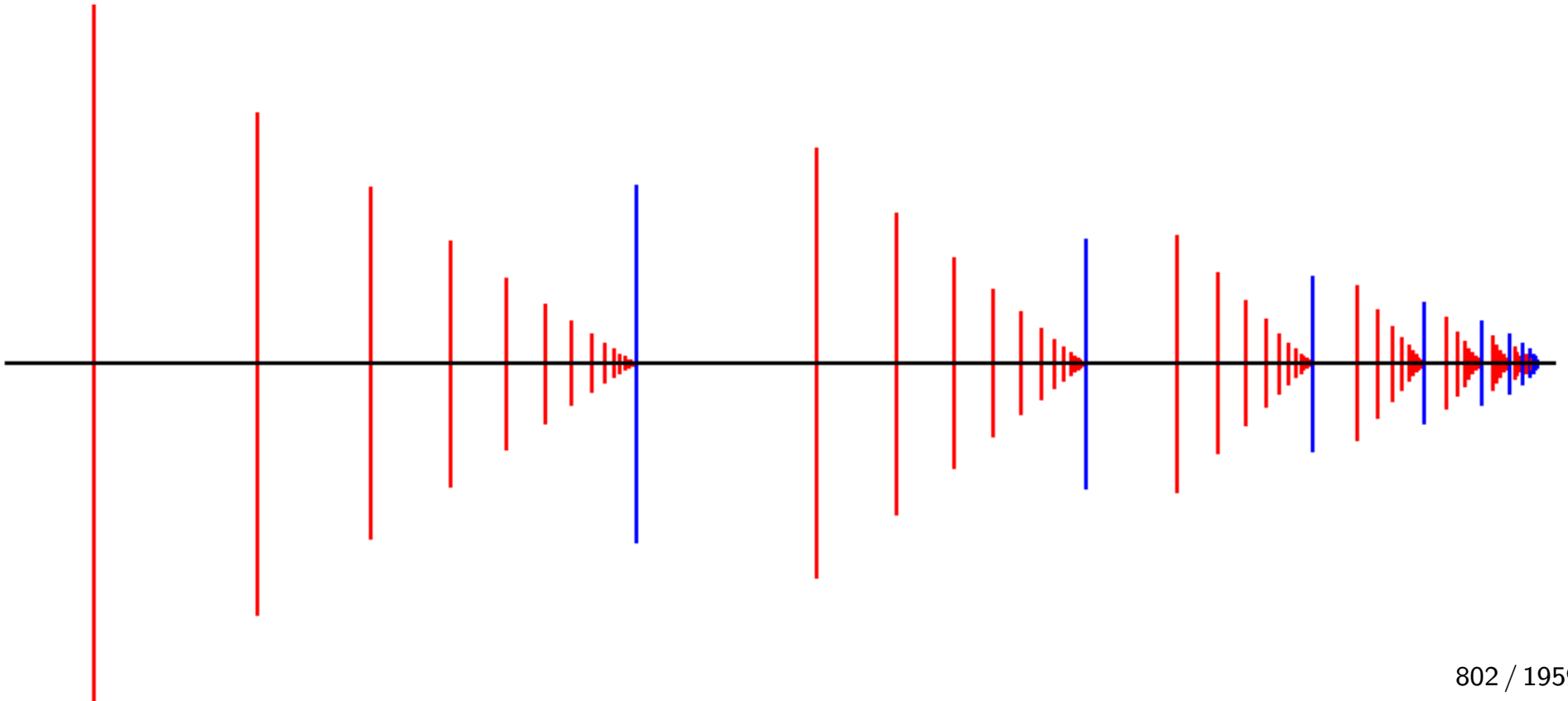
Now I Know! — transfinite — continued

12. **A:** Yet, I still do not know.
13. **B:** And also I remain in ignorance. However shall we come to know?
14. **C:** Well, in fact, no matter how long we three continue from now in the pattern we have followed so far — namely, the pattern in which you two state back-and-forth that still you do not yet know whose number is larger and then I tell you yet again that no further amount of that back-and-forth will enable you to know — then still after as much repetition of that pattern as we can stand, you will not know whose number is larger! Furthermore, I could make that same statement a second time, even after now that I have said it to you once, and it would still be true!
15. **A:** Such powerful new information! But I still do not know.
16. **B:** And also I do not know.
17. **A:** Aha! Now I know who has the larger number!
18. **B:** Then I know both our numbers!
19. **A:** Well, now I also know them!

Now I Know! — Solution

$(7, 6)$ $(\omega \cdot 2 + 1, \omega \cdot 2)$ $(\frac{19}{8}, \frac{39}{16})$





Well-Founded Relation

- $R \subset A^2$ is *set-like* iff $\mathrm{ext}_R(x) := \{y \in A : Ryx\}$ is a set for every $x \in A$.
- $y \in A$ is *R-minimal* in A iff $\neg\exists z(z \in A \wedge Rzy)$.
- A set-like relation R is *well-founded* on A iff every non-empty set $X \subset A$ has a R-minimal element.
- R *well-orders* A iff R totally orders A strictly and R is well-founded on A.



Figure: Noether

Well-Founded Induction/Recursion

Theorem (Well-Founded/Noetherian Induction)

Let R be a well-founded relation on A. Let P be a property.

$$\forall x \in \min(A)\, P(x) \wedge \forall x \in A\, [\forall y \in A\, (Ryx \to P(y)) \to P(x)] \to \forall x \in A\, P(x)$$

Theorem (Well-Founded Recursion)

Let R be a well-founded relation on A. Let G be a function. Then there is a unique function F on A s.t. for every $x \in A$,

$$F(x) = G(x, F\restriction_{\mathrm{ext}_R(x)})$$

Perfect Set

- x is a *limit point* of A iff every neighborhood U of x contains a point of A other than x itself, i.e., $(U \setminus \{x\}) \cap A \neq \emptyset$.
- x is a *adherent point* of A iff every neighbourhood of x contains a point of A, i.e., $U \cap A \neq \emptyset$.
- x is an *interior point* of A iff there is an open set U s.t. $x \in U \subset A$.
- The *derived set* $A' := \left\{x : x \text{ is a limit point of } A\right\}$.
- The *closure* of A is $\mathrm{cl}(A) := \left\{x : x \text{ is a adherent point of } A\right\}$.

$$A^\circ = \overline{\mathrm{cl}\left(\overline{A}\right)} \quad \text{and} \quad \mathrm{cl}(A) = \overline{\left(\overline{A}\right)^\circ} \quad \text{and} \quad \mathrm{cl}(A) = A \cup A'$$

- The *interior* of A is $A^\circ := \left\{x : x \text{ is an interior point of } A\right\}$.
- The *boundary* of A is $\partial A := \mathrm{cl}(A) \setminus A^\circ$. Equivalently, $\partial A = \mathrm{cl}(A) \cap \mathrm{cl}(\overline{A})$.
- x is an *isolated point* of A iff $x \in A \setminus A'$.
- Example: if $A := (0, 1] \cup \{2\}$, then $A' = [0, 1]$, $\mathrm{cl}(A) = [0, 1] \cup \{2\}$, $A^\circ = (0, 1)$, and $\partial A = \{0, 1, 2\}$.
- A is *closed* iff $\mathrm{cl}(A) = A$. Equivalently, $A' \subset A$.
- A is *perfect* iff $A' = A$ iff it is closed and has no isolated points.
- Theorem: Every perfect set has cardinality $2^{\aleph_0}$.

Cantor-Bendixson Rank

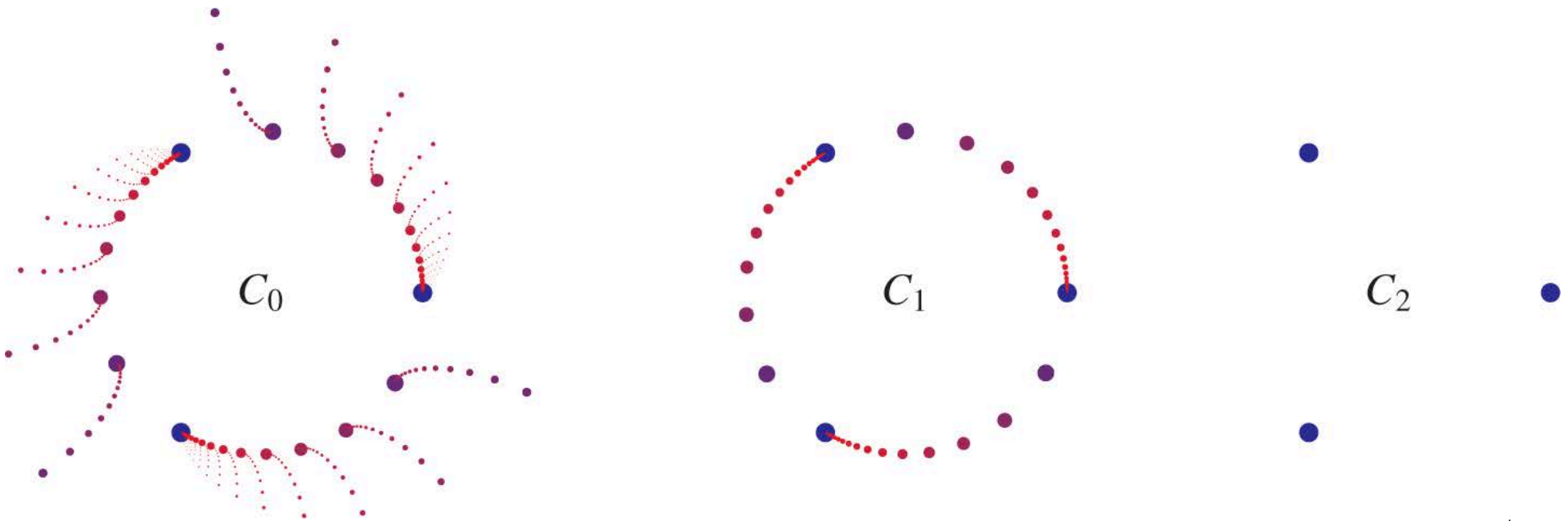
$$A_0 := A$$

$$A_{\alpha+1} := A'_\alpha \text{ the derived set of } A_\alpha$$

$$A_\alpha := \bigcap_{\gamma<\alpha} A_\gamma \text{ if } \alpha \text{ is a limit ordinal}$$

$$\mathrm{rank}(A) := \mu\alpha[A_\alpha = A_{\alpha+1}]$$

Example: a set with rank 3



Cantor-Bendixson Theorem

Theorem (Cantor-Bendixson Theorem)

If A is an uncountable closed set, then $A = P \cup S$, where P is perfect and S is countable.

Proof.

Let $P \coloneqq A_{\mathrm{rank}(A)}$. Then

$$A \setminus P = \bigcup_{\alpha < \mathrm{rank}(A)} \left(A_\alpha \setminus A'_\alpha\right)$$

Let $\langle J_k : k \in \mathbb{N} \rangle$ be an enumeration of rational intervals.
Hence for $a \in A \setminus P$, there is a unique α s.t. a is an isolated point of A_α.
Let $f(a) \coloneqq \mu k[A_\alpha \cap J_k = \{a\}]$. Then $f : A \setminus P \to \mathbb{N}$ is injective. □

Trigonometric Expansion

Definition (Trigonometric Expansion)

A function $f : \mathbb{R} \to \mathbb{C}$ *admits a trigonometric expansion* iff

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos nx + b_n \sin nx \right)$$

For example, a continuously differentiable function admits a trigonometric expansion, where a_n, b_n can be computed by Fourier formulas

$$a_n := \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx \, \mathrm{d}x = \frac{1}{\sqrt{\pi}} \left\langle f(x) \middle| \frac{\cos nx}{\sqrt{\pi}} \right\rangle$$

$$b_n := \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx \, \mathrm{d}x = \frac{1}{\sqrt{\pi}} \left\langle f(x) \middle| \frac{\sin nx}{\sqrt{\pi}} \right\rangle$$

Remark: $\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \ldots \right\}$ forms an orthonormal basis for Hilbert space $L^2(0, 2\pi)$, where $\langle f | g \rangle = \int_0^{2\pi} f(x) \overline{g(x)} \, \mathrm{d}x$.

Cantor-Lebesgue Theorem

- Characterization: which functions admit a trigonometric expansion?
- Coefficient: How to "compute" the coefficients of the expansion?
- **Uniqueness**: Is such an expansion unique?

Theorem (Cantor-Lebesgue Theorem)

For any $A \subset \mathbb{R}$, *if* $A_{\operatorname{rank}(A)} = \emptyset$, *then*

$$\forall x \in \mathbb{R} \setminus A \left(\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos nx + b_n \sin nx \right) = 0 \right) \implies \forall n \in \mathbb{N} \left(a_n = b_n = 0 \right)$$

If f is continuous at all but countable points, then it admits a unique trigonometric expansion.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory
- Axioms of ZFC
- Ordinal Numbers
- Cardinal Numbers
- Axiom of Choice

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

How do we count a finite set?

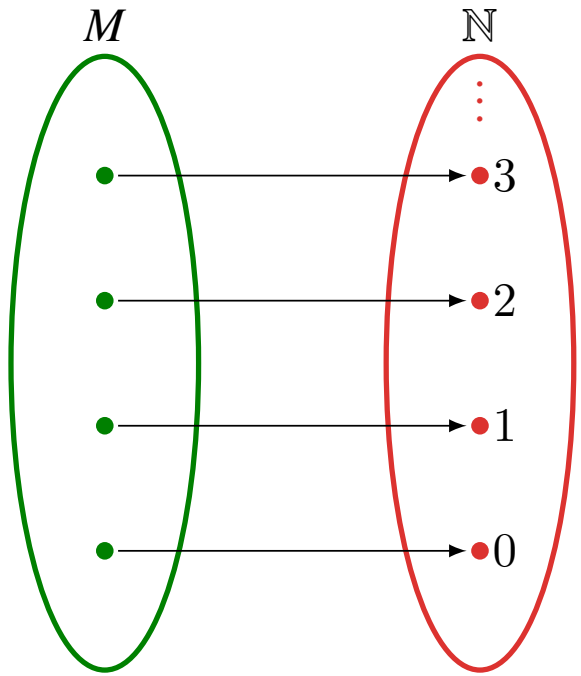
$$M := \{\text{apple}, \text{orange}, \text{banana}, \text{grape}\}$$

What does $|M| = 4$ mean?

There is a bijection between M and $N := \{1, 2, 3, 4\}$.

$$\begin{aligned} \text{apple} &\longleftrightarrow 1 \\ \text{orange} &\longleftrightarrow 2 \\ \text{banana} &\longleftrightarrow 3 \\ \text{grape} &\longleftrightarrow 4 \end{aligned}$$

$$|M| = |N| \;:=\; \exists f : M \rightarrowtail\!\!\!\!\twoheadrightarrow N$$



A set A is finite iff $\exists n \in \mathbb{N} : |A| = n$.

Does renaming the elements of a set change its size? No!

Bijection is nothing more than renaming.

How do we compare the sizes of finite sets?

$M := \{\text{apple}, \text{orange}, \text{banana}, \text{grape}\}$

$N := \{\text{John}, \text{Peter}, \text{Bell}, \text{Emma}, \text{Sam}\}$

What does $|M| \leq |N|$ mean?

$$\begin{array}{lcccl} \text{apple} & \longleftrightarrow & 1 & \longleftrightarrow & \text{John} \\ \text{orange} & \longleftrightarrow & 2 & \longleftrightarrow & \text{Peter} \\ \text{banana} & \longleftrightarrow & 3 & \longleftrightarrow & \text{Bell} \\ \text{grape} & \longleftrightarrow & 4 & \longleftrightarrow & \text{Emma} \\ & & 5 & \longleftrightarrow & \text{Sam} \end{array}$$

$$\begin{array}{lcl} \text{apple} & \longrightarrow & \text{John} \\ \text{orange} & \longrightarrow & \text{Peter} \\ \text{banana} & \longrightarrow & \text{Bell} \\ \text{grape} & \longrightarrow & \text{Emma} \\ & & \text{Sam} \end{array}$$

$$|M| \leq |N| \;:=\; \exists f : M \rightarrowtail N$$

$$|M| \leq |N| \;:=\; \exists f : N \twoheadrightarrow M$$

$$\begin{array}{lcl} \text{apple} & \longleftarrow & \text{John} \\ \text{orange} & \longleftarrow & \text{Peter} \\ \text{banana} & \longleftarrow & \text{Bell} \\ \text{grape} & \longleftarrow & \text{Emma} \\ & \nwarrow & \text{Sam} \end{array}$$

The way of comparing the size of finite sets generalizes to infinite sets!

$$|\mathbb{N}| = |\mathbb{Z}|$$

$$\begin{array}{ccc} 0 & \longleftrightarrow & 0 \\ 1 & \longleftrightarrow & 1 \\ 2 & \longleftrightarrow & -1 \\ 3 & \longleftrightarrow & 2 \\ 4 & \longleftrightarrow & -2 \\ 5 & \longleftrightarrow & 3 \\ 6 & \longleftrightarrow & -3 \\ 7 & \longleftrightarrow & 4 \\ 8 & \longleftrightarrow & -4 \\ \vdots & & \vdots \end{array}$$

Dedekind-Infinite

A set A is Dedeking-infinite iff some proper subset $B \subsetneq A$ is equinumerous to A.



Figure: Dedekind

Countable?

$$\{0, 1, 2, 3, 4, \dots\}$$

$$\{1, 3, 5, 7, 9, \dots\}$$

$$\{0, 2, 4, 6, 8, \dots\}$$

$$\{0, 1, 4, 9, 16, \dots\}$$

$$\{2, 3, 5, 7, 11, \dots\}$$

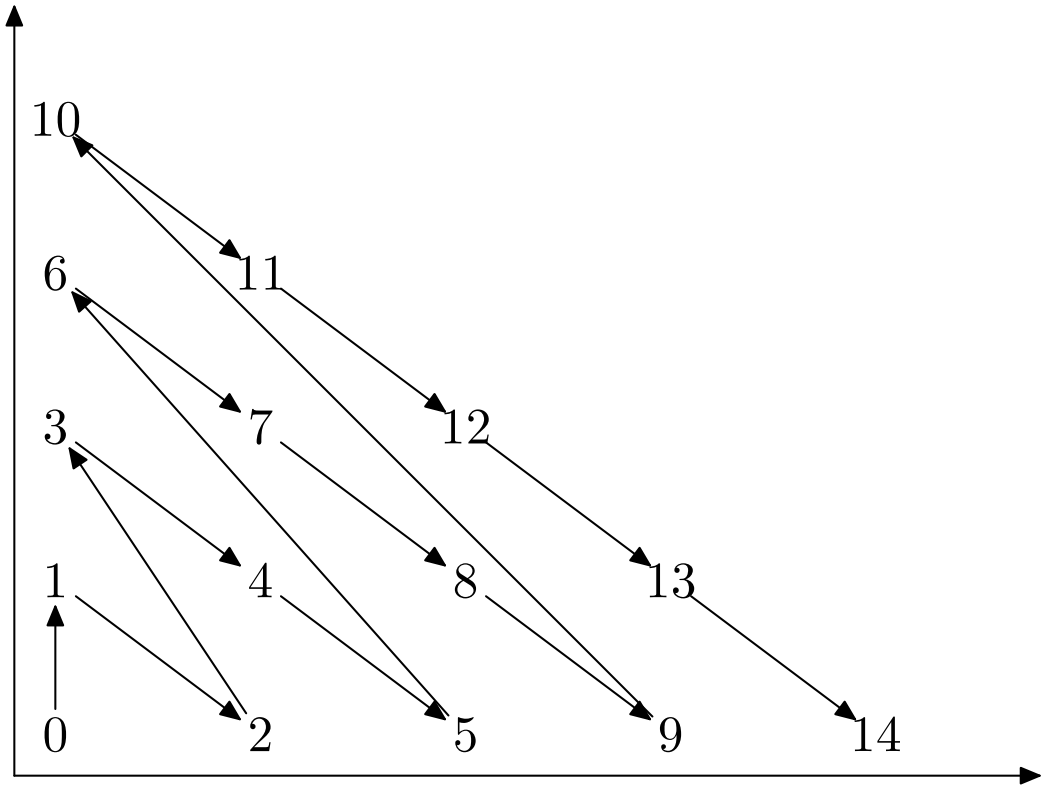
- A set A is countable iff $|A| \leq |\mathbb{N}|$.
- Is it possible that A is infinite, but $|A| < |\mathbb{N}|$?
- A set A is countably infinite iff $|A| = |\mathbb{N}|$.

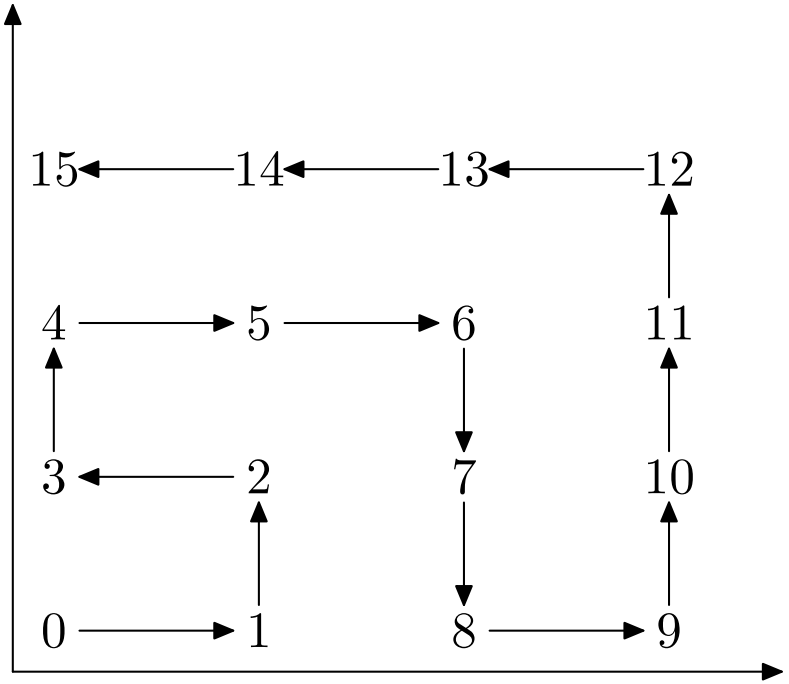
$$|\mathbb{N}| = |2^{<\omega}|$$

$$\begin{array}{rcr} 0 & \longleftrightarrow & \epsilon \\ 1 & \longleftrightarrow & 0 \\ 2 & \longleftrightarrow & 1 \\ 3 & \longleftrightarrow & 00 \\ 4 & \longleftrightarrow & 01 \\ 5 & \longleftrightarrow & 10 \\ 6 & \longleftrightarrow & 11 \\ 7 & \longleftrightarrow & 000 \\ 8 & \longleftrightarrow & 001 \\ \vdots & & \vdots \end{array}$$

Countable?

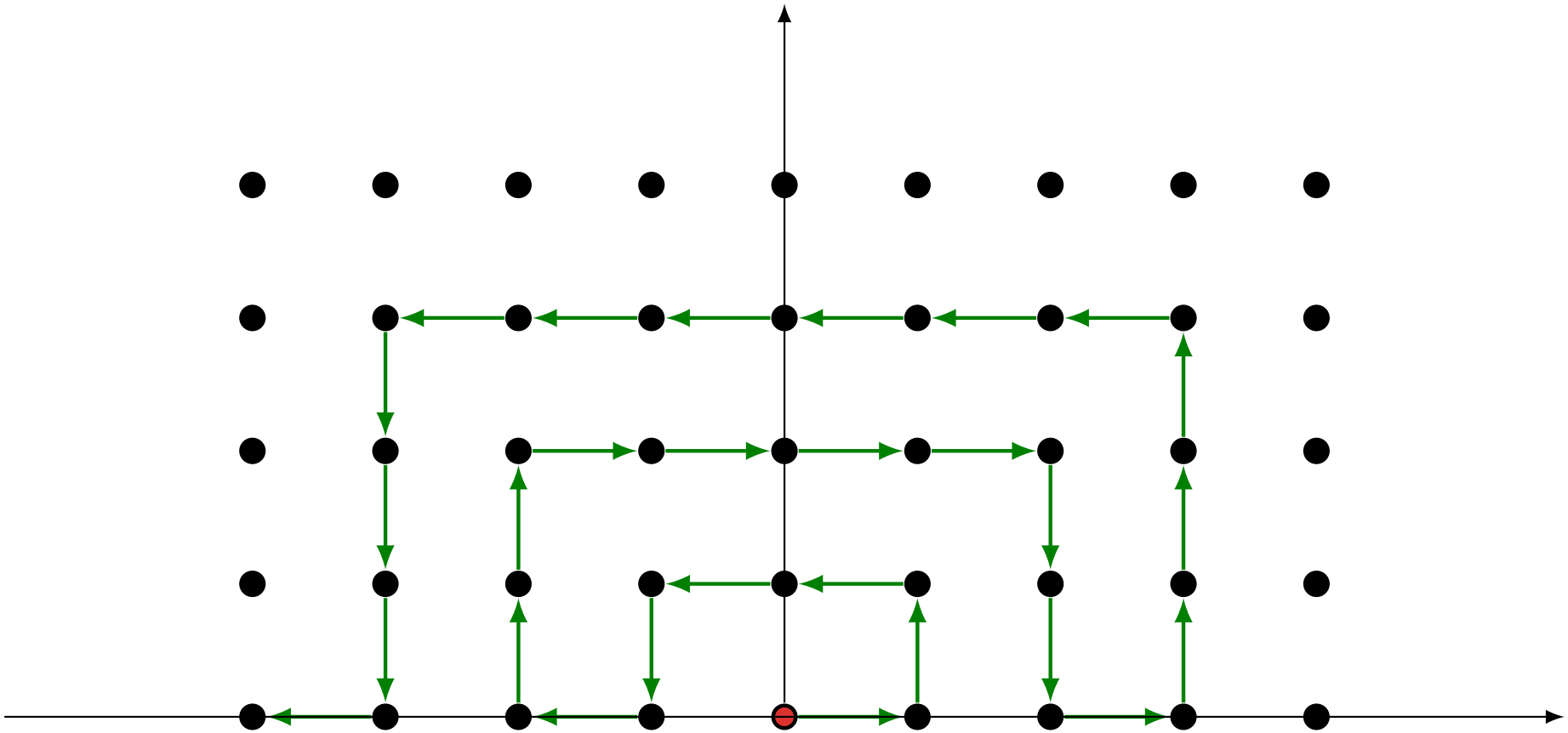
$$|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$$





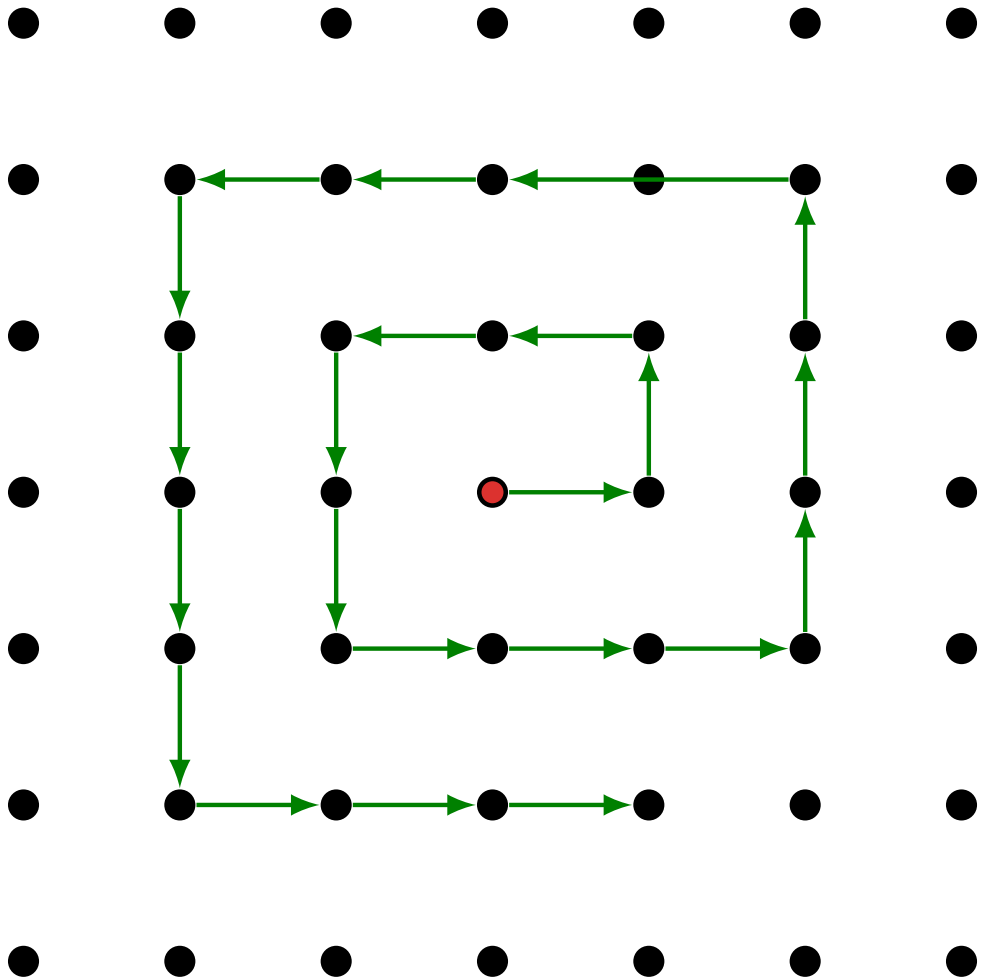
Countable?

$$|\mathbb{N}| = |\mathbb{Z} \times \mathbb{N}|$$

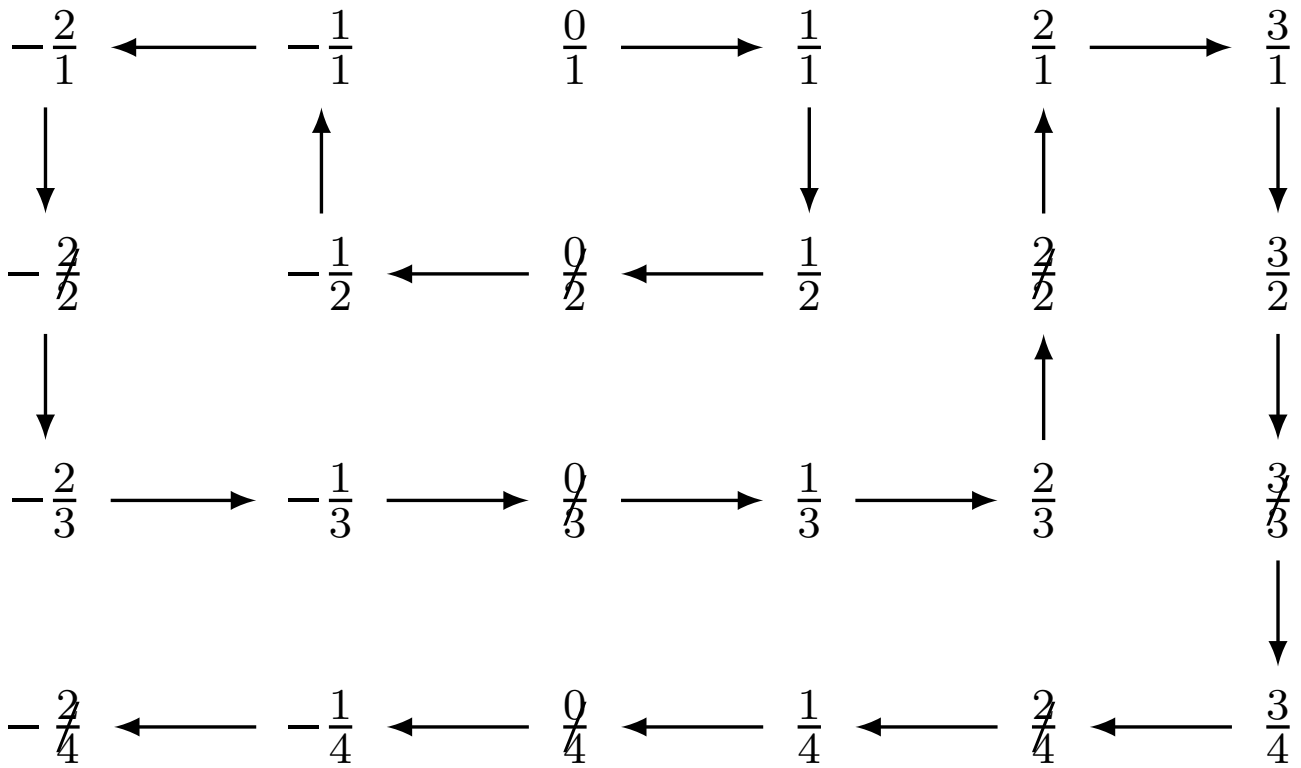


Countable?

$$|\mathbb{N}| = |\mathbb{Z} \times \mathbb{Z}|$$

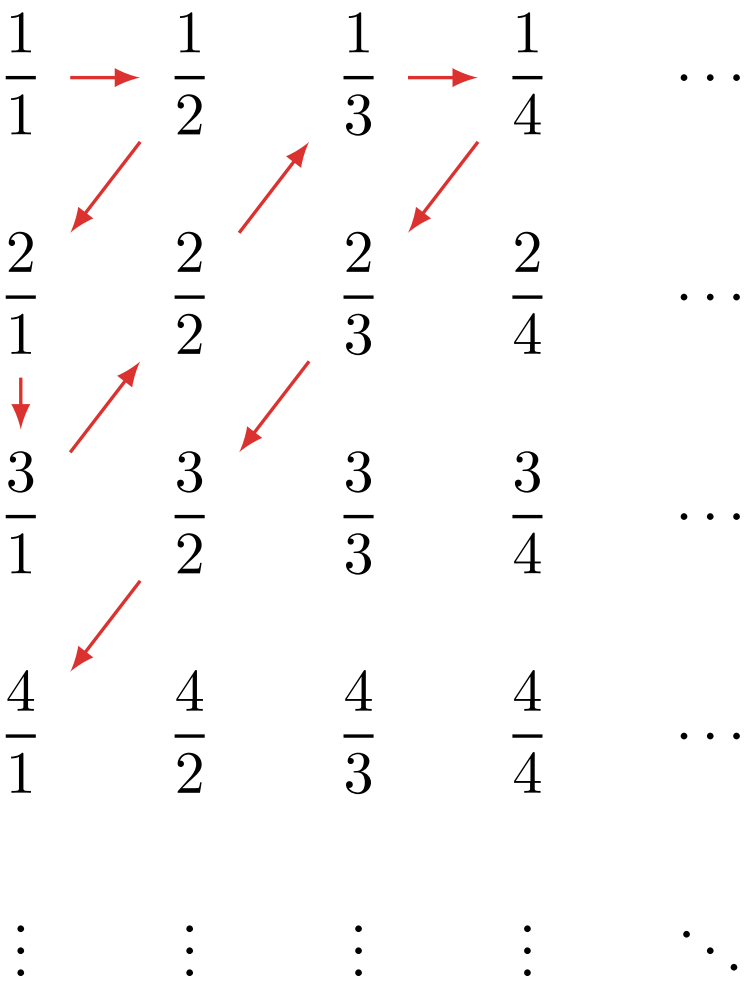


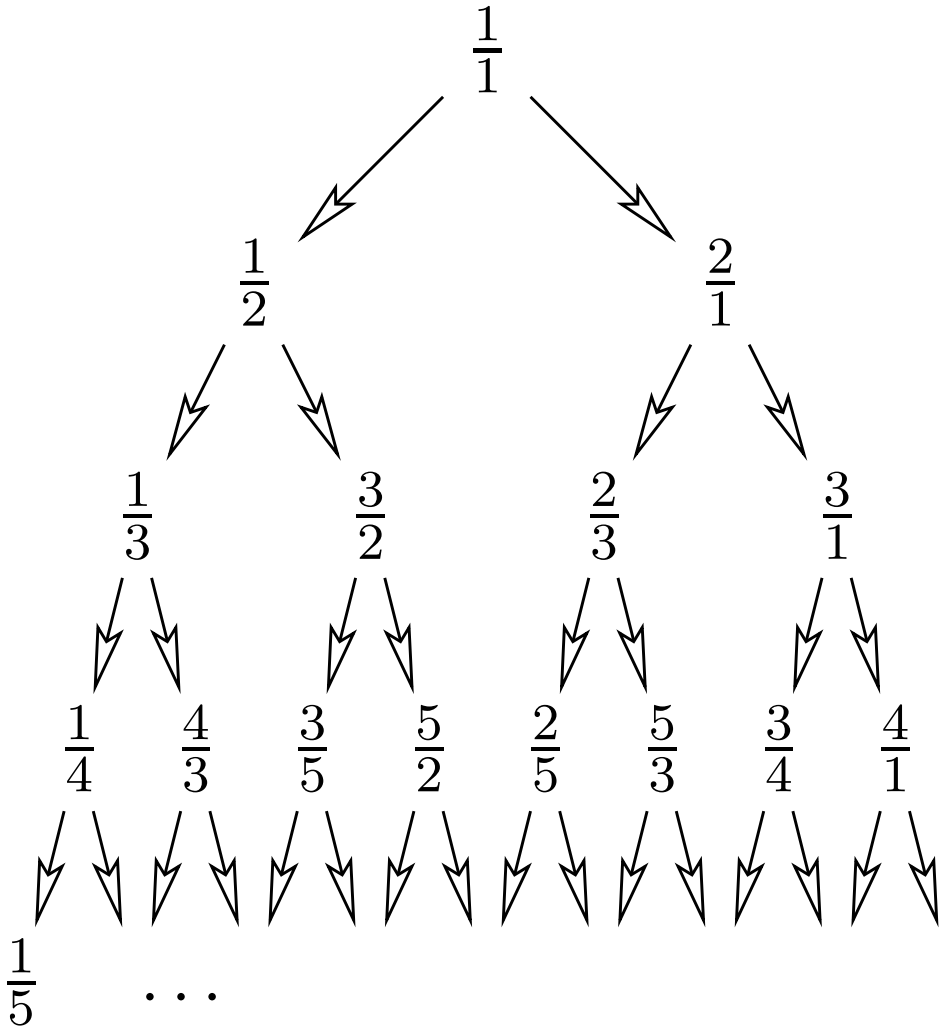
$|\mathbb{N}| = |\mathbb{Q}|$



$$|\mathbb{N}| = |\mathbb{Z}| = |2^{<\omega}| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{Z} \times \mathbb{N}| = |\mathbb{Z} \times \mathbb{Z}| = |\mathbb{Q}|$$

$|\mathbb{N}| = |\mathbb{Q}^+|$





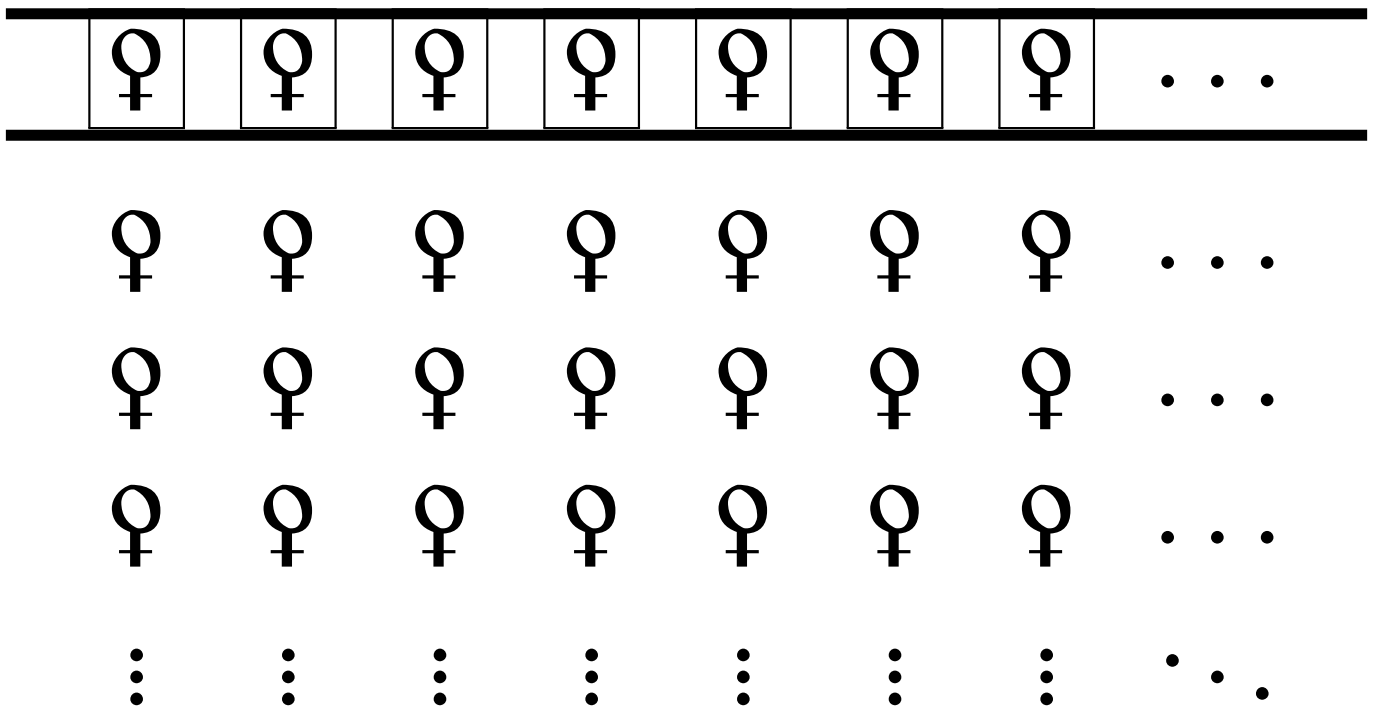
$$x \mapsto \frac{1}{\lfloor x \rfloor + 1 - \{x\}}$$

Hilbert's Hotel

Problem (Hilbert's Hotel)

Consider a hypothetical hotel with a countably infinite number of rooms, all of which are occupied.

1. *Finitely many new guests.*
2. *Infinitely many new guests.*
3. *Infinitely many buses with infinitely many guests each.*



Hilbert's Hotel

$$\aleph_0 + n = \aleph_0$$
$$\aleph_0 \cdot n = \aleph_0$$
$$\aleph_0 \cdot \aleph_0 = \aleph_0$$

Theorem

Let A be a countable set. Then the set of all finite sequences of members of A is also countable.

Proof.

$$f : A \rightarrowtail \mathbb{N} \implies \exists g : \bigcup_{n\in\mathbb{N}} A^{n+1} \rightarrowtail \mathbb{N} \quad (a_1, \ldots, a_n) \mapsto \prod_{i=1}^{n} p_i^{f(a_i)+1}$$

□

The set of real numbers is uncountable

Is every set countable?

Theorem (Cantor)

$$|\mathbb{R}| > |\mathbb{N}|$$

Proof.

$$\begin{array}{lccccc} 0. & r_{11} & r_{12} & r_{13} & r_{14} & \dots \\ 0. & r_{21} & r_{22} & r_{23} & r_{24} & \dots \\ 0. & r_{31} & r_{32} & r_{33} & r_{34} & \dots \\ 0. & r_{41} & r_{42} & r_{43} & r_{44} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Let $d = 0.d_1 d_2 \dots$ where

$$d_n = 9 - r_{nn}$$

□

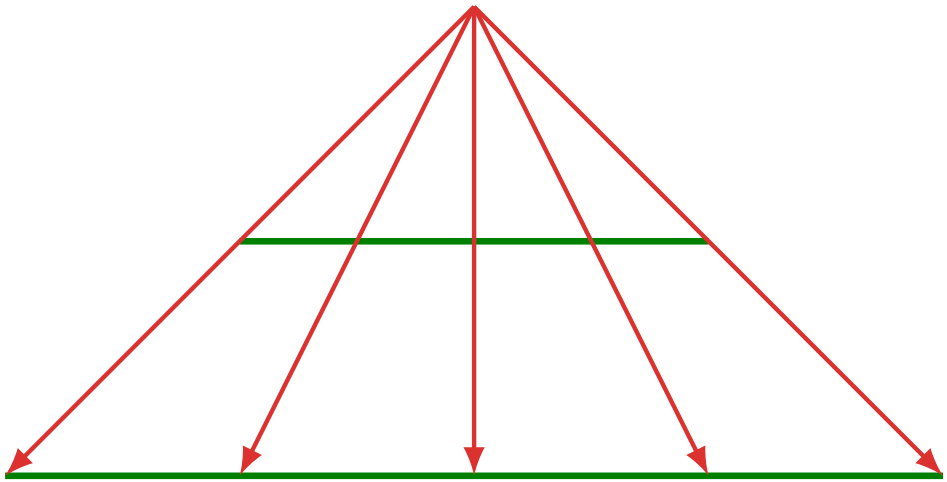
The set of real numbers is uncountable — another proof

Proof.

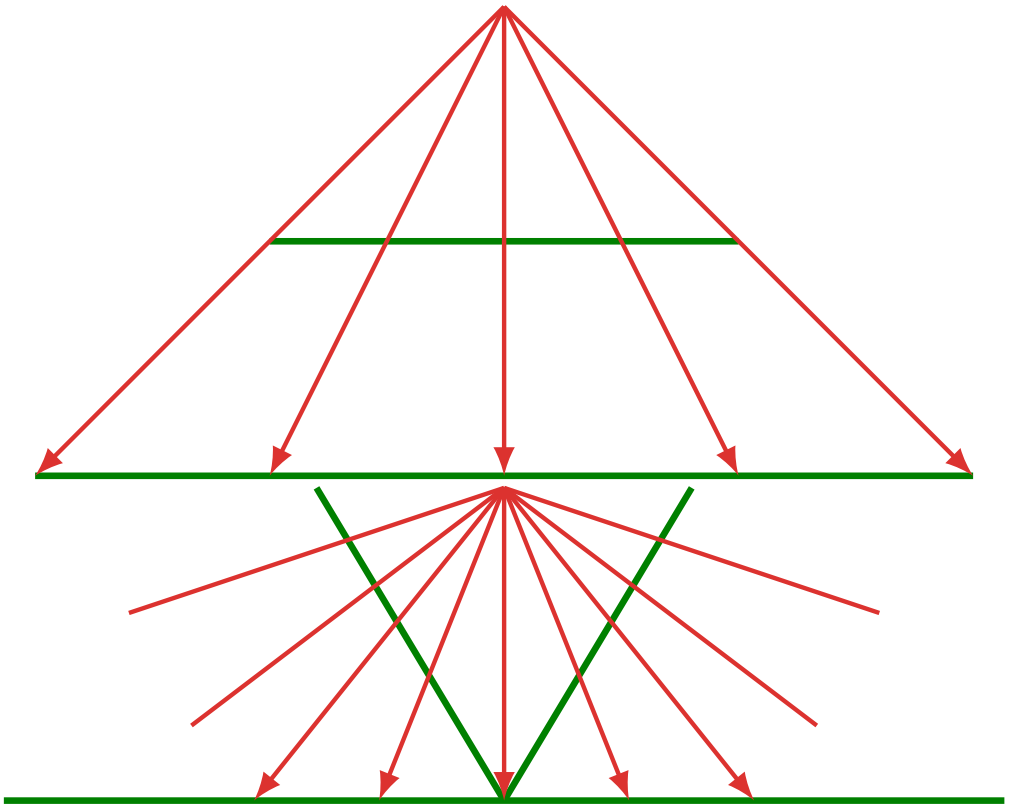
- **Game**. Fix $S \subset [0,1]$. Let $a_0 = 0, b_0 = 1$. In round $n \geq 1$, Alice chooses a_n s.t. $a_{n-1} < a_n < b_n$, then Bob chooses b_n s.t. $a_n < b_n < b_{n-1}$. Since a monotonically increasing sequence of real numbers bounded above has a limit, $\alpha = \lim_{n\to\infty} a_n$ is well-defined. Alice wins if $\alpha \in S$, otherwise Bob wins.
- Assume S is countable, $S = \{s_1, s_2, \dots\}$. On move $n \geq 1$, Bob chooses $b_n = s_n$ if this is a legal move, otherwise he randomly chooses any allowable number for b_n. Bob always wins with this strategy!
- But when $S = [0,1]$, Alice can't lose!

□

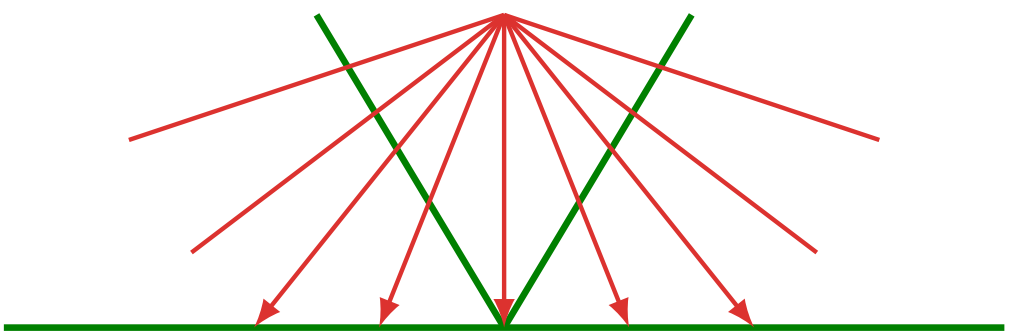
Continuum

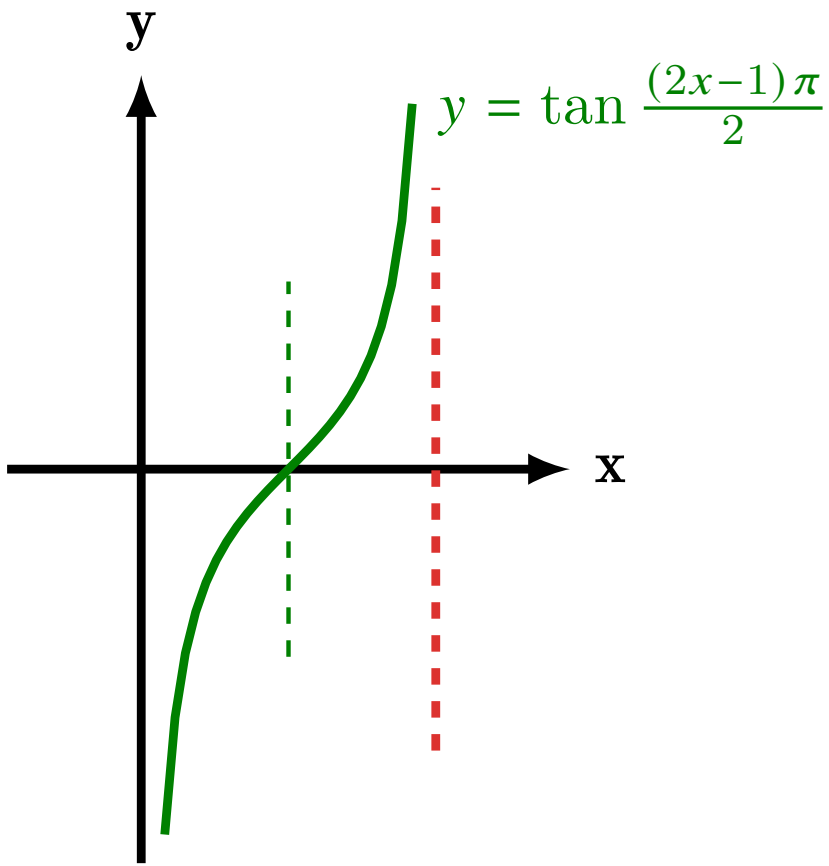


Continuum



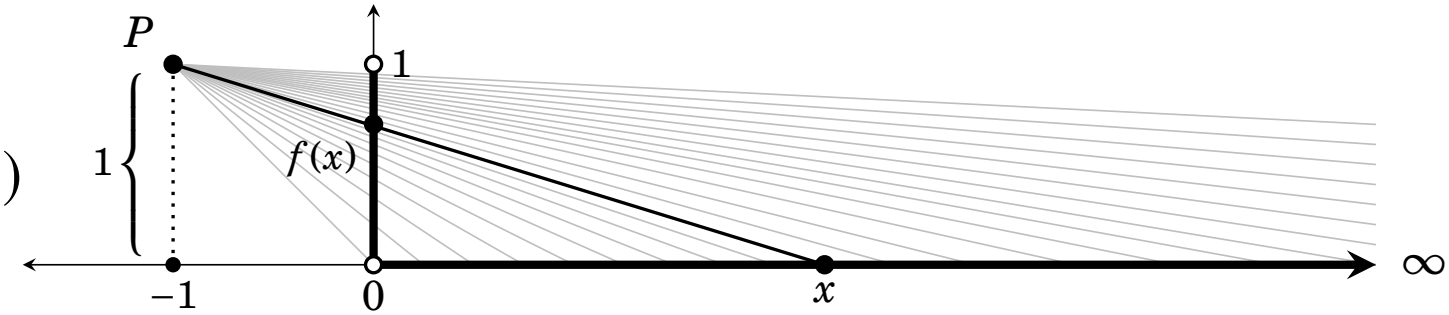
Continuum





Continuum

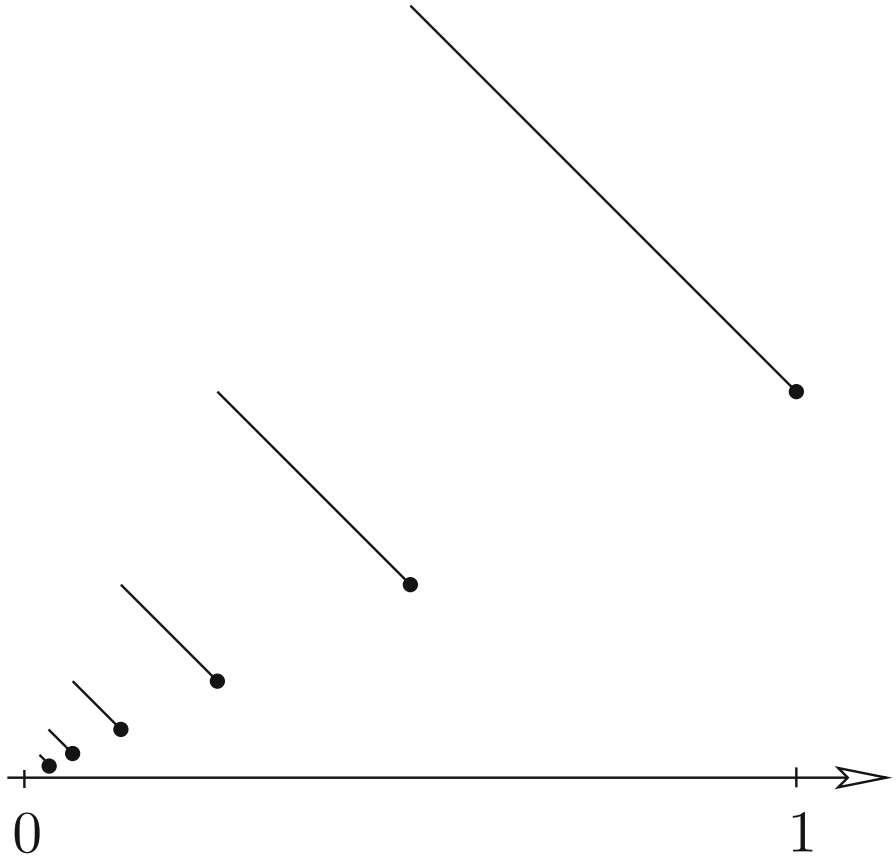
$f : (0,1) \twoheadrightarrow (0,\infty)$



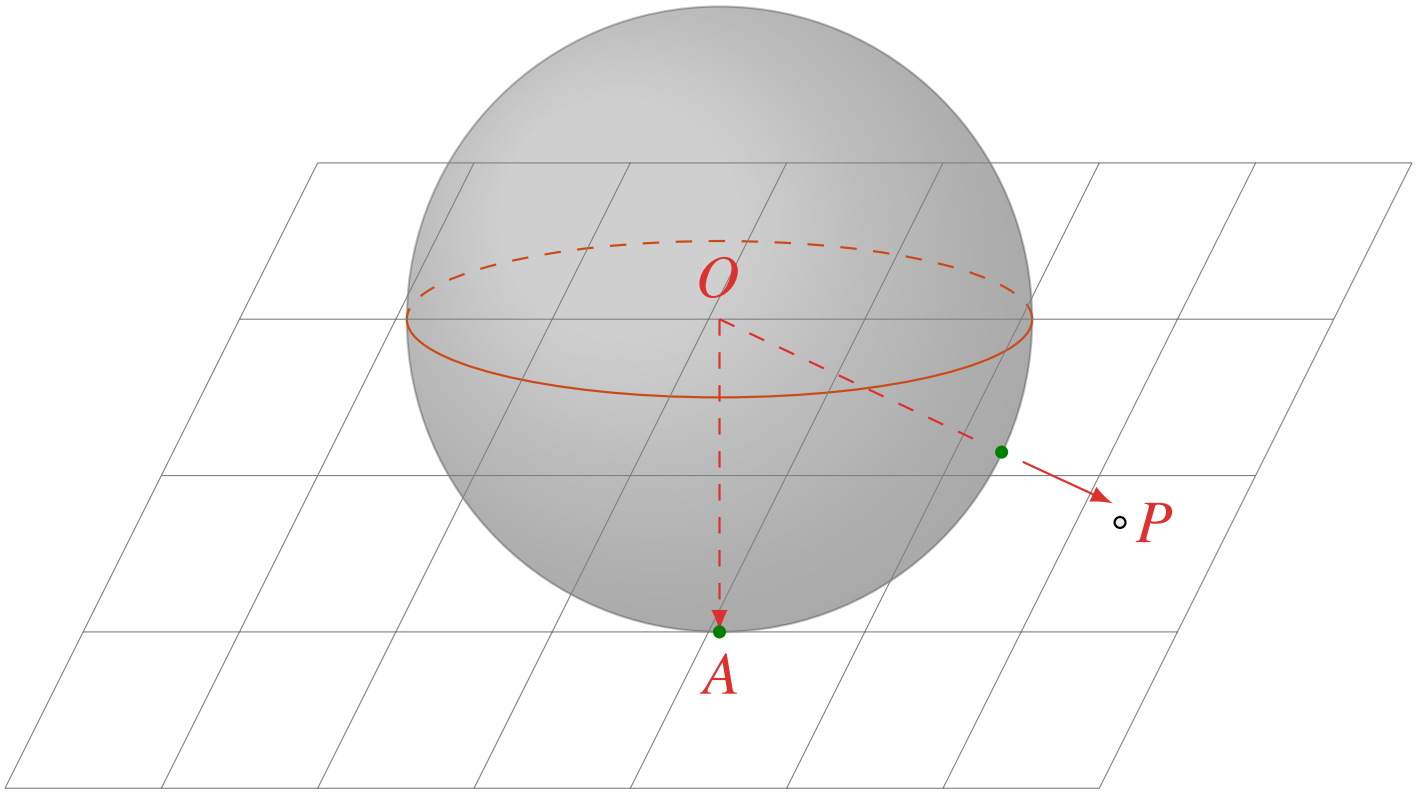
$f : (0,1] \twoheadrightarrow (0,1)$

$$f(x) := \begin{cases} \frac{3}{2} - x & \text{for } \frac{1}{2} < x \leq 1 \\ \frac{3}{4} - x & \text{for } \frac{1}{4} < x \leq \frac{1}{2} \\ \frac{3}{8} - x & \text{for } \frac{1}{8} < x \leq \frac{1}{4} \\ \vdots & \end{cases}$$

$$f(x) := \begin{cases} \frac{x}{x+1} & \text{if } \exists n \in \mathbb{N} : x = \frac{1}{n} \\ x & \text{otherwise} \end{cases}$$



Continuum



Continuum

Theorem

$$|\mathbb{R}| = |\mathbb{R} \times \mathbb{R}|$$

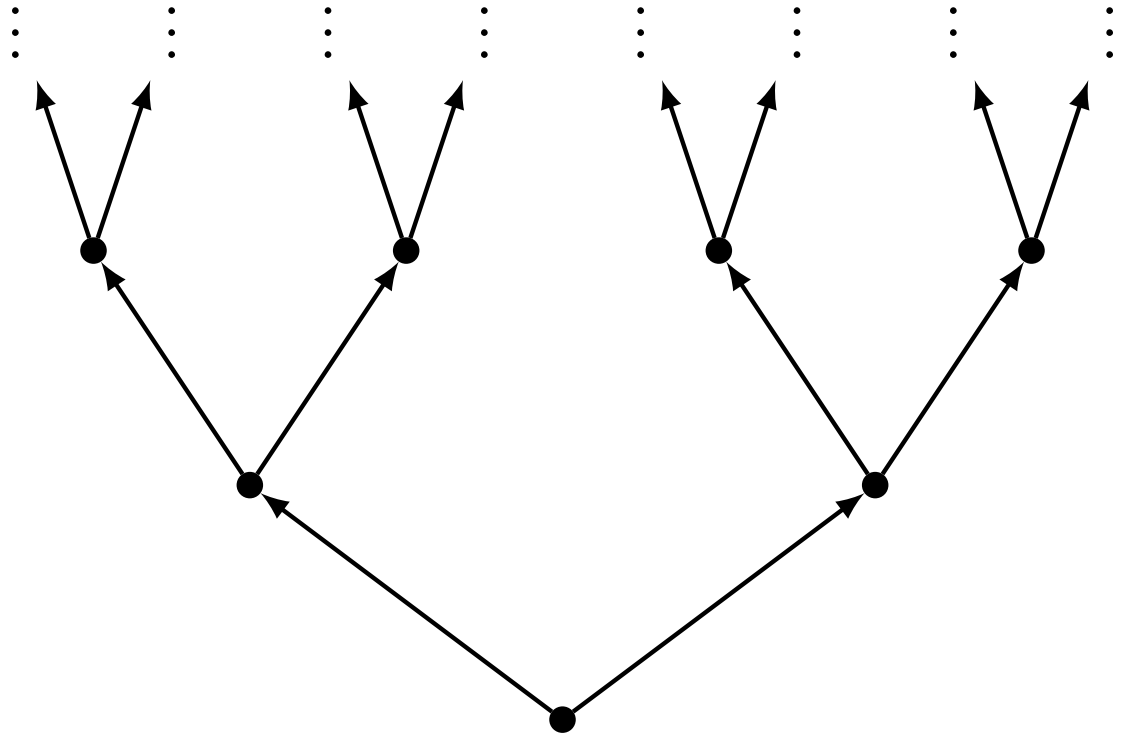
Proof.

$$\begin{aligned} x &= 0.3 \quad 01 \quad 2 \quad 007 \quad 08 \ldots \\ y &= 0.009 \quad 2 \quad 05 \quad 1 \quad 0003 \ldots \end{aligned}$$

$$z = 0.3\ 009\ 01\ 2\ 2\ 05\ 007\ 1\ 08\ 0003\ \ldots$$

□

Continuum



$$[0,1] = \left\{ \sum_{n=1}^{\infty} \frac{x_n}{2^n} : x_n = 0 \vee x_n = 1 \right\}$$

Cantor's Theorem

Theorem (Cantor's Theorem)

$$|X| < |\mathrm{P}(X)|$$

Proof.

If $f : X \to \mathrm{P}(X)$, then

$$Y := \{x \in X : x \notin f(x)\}$$

is not in the range of f. □

Cantor's Paradox

the 'set' of all sets?

Cantor's Theorem

$1, 2, \ldots, \aleph_0, \aleph_1, \ldots, \aleph_\omega, \aleph_{\omega+1}, \ldots, \aleph_{\omega\cdot 2}, \ldots, \aleph_{\omega^2}, \ldots, \aleph_{\omega^\omega}, \ldots, \aleph_{\varepsilon_0}, \ldots, \aleph_{\aleph_0}, \ldots, \aleph_{\aleph_{\aleph_0}}, \ldots$
the first cardinal to succeed all of these is labeled by the ordinal $\kappa = \aleph_\kappa$.
(So big that it needs itself to say how big it is!)

The 'set' I of all distinct levels of infinity is so large that it can't be a set!

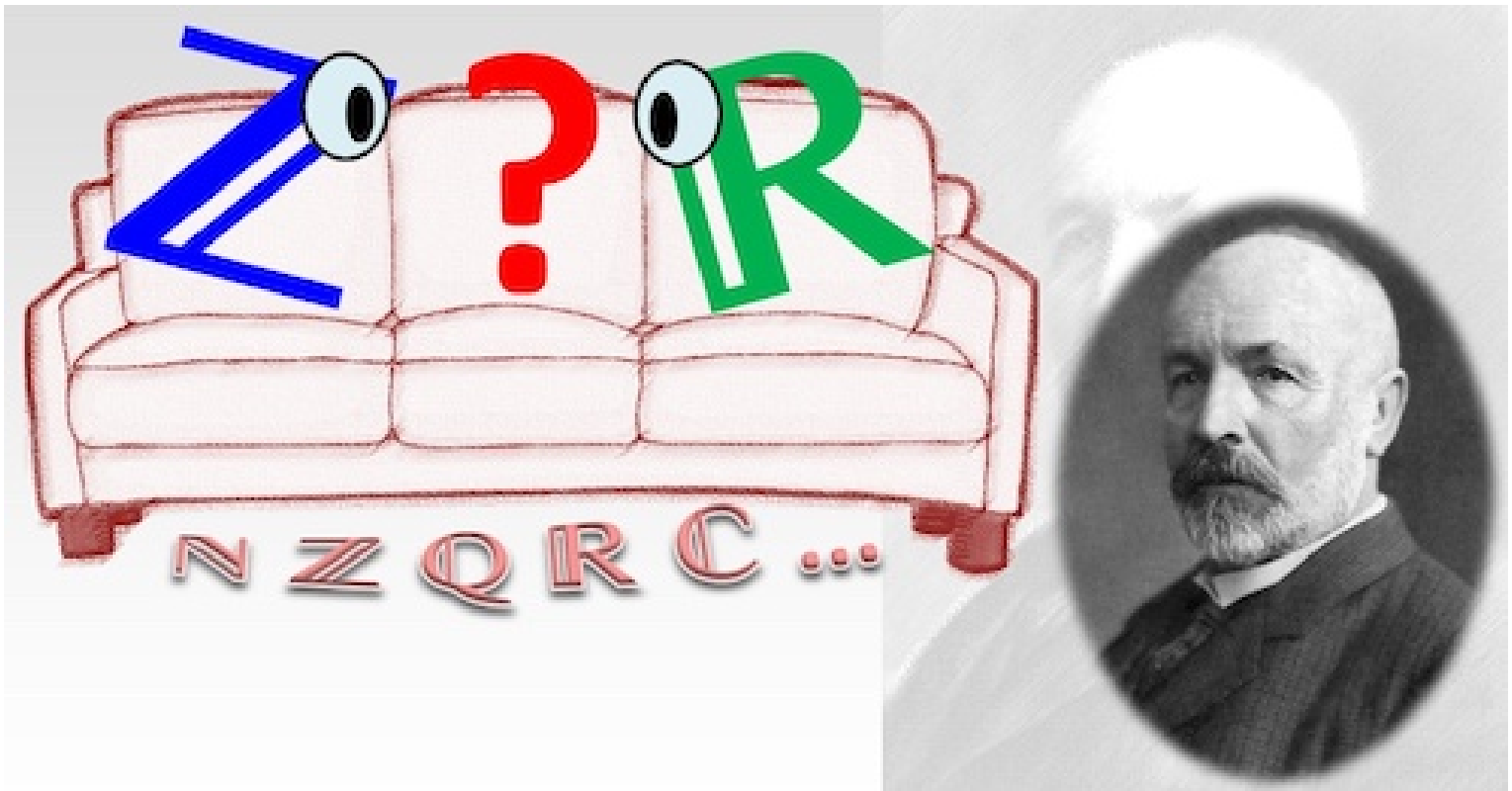
$$\forall d \in I : \ |S_d| \leq \left|\bigcup_{c\in I} S_c\right| < \left|\mathrm{P}\left(\bigcup_{c\in I} S_c\right)\right|$$

where S_c is a representative set that has cardinality c.

Cantor's Continuum Hypothesis

Cantor's Continuum Hypothesis (CH)

$$2^{\aleph_0} \stackrel{?}{=} \aleph_1$$

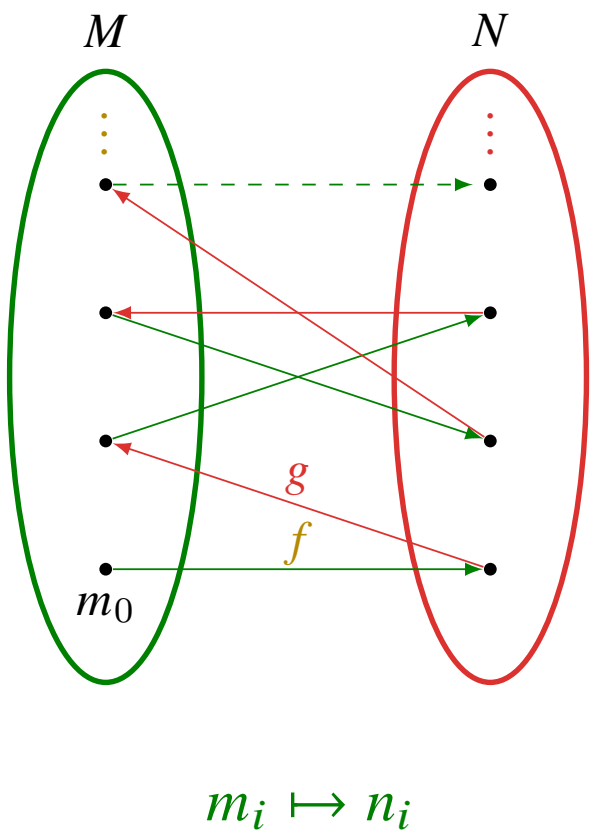


Cantor-Schröder-Bernstein Theorem

Theorem (Cantor-Schröder-Bernstein Theorem)

$$\left.\begin{array}{l} |M| \le |N| \\ |N| \le |M| \end{array}\right\} \Longrightarrow |M| = |N|$$

1. Finite cycles on $2k+2$ distinct elements $(k \ge 0)$
   $m_0 \longrightarrow n_0 \longrightarrow m_1 \longrightarrow \cdots \longrightarrow m_k \longrightarrow n_k$
2. Two-way infinite chains of distinct elements
   $\cdots \longrightarrow m_0 \longrightarrow n_0 \longrightarrow m_1 \longrightarrow n_1 \longrightarrow \cdots$
3. The one-way infinite chains of distinct elements that start at the elements $m_0 \in M \setminus g(N)$
   $m_0 \longrightarrow n_0 \longrightarrow m_1 \longrightarrow n_1 \longrightarrow m_2 \longrightarrow \cdots$
4. The one-way infinite chains of distinct elements that start at the elements $n_0 \in N \setminus f(M)$
   $n_0 \longrightarrow m_0 \longrightarrow n_1 \longrightarrow m_1 \longrightarrow n_2 \longrightarrow \cdots$



Remark

Problem

假设 M 是男人的无穷集合，N 是女人的无穷集合. 每个男人爱且只爱一个女人，没有两个男人爱同一个女人，每个女人爱且只爱一个男人，没有两个女人爱同一个男人. 那么，男人和女人能否一一配对结婚，使得或者老公爱老婆，或者老婆爱老公？

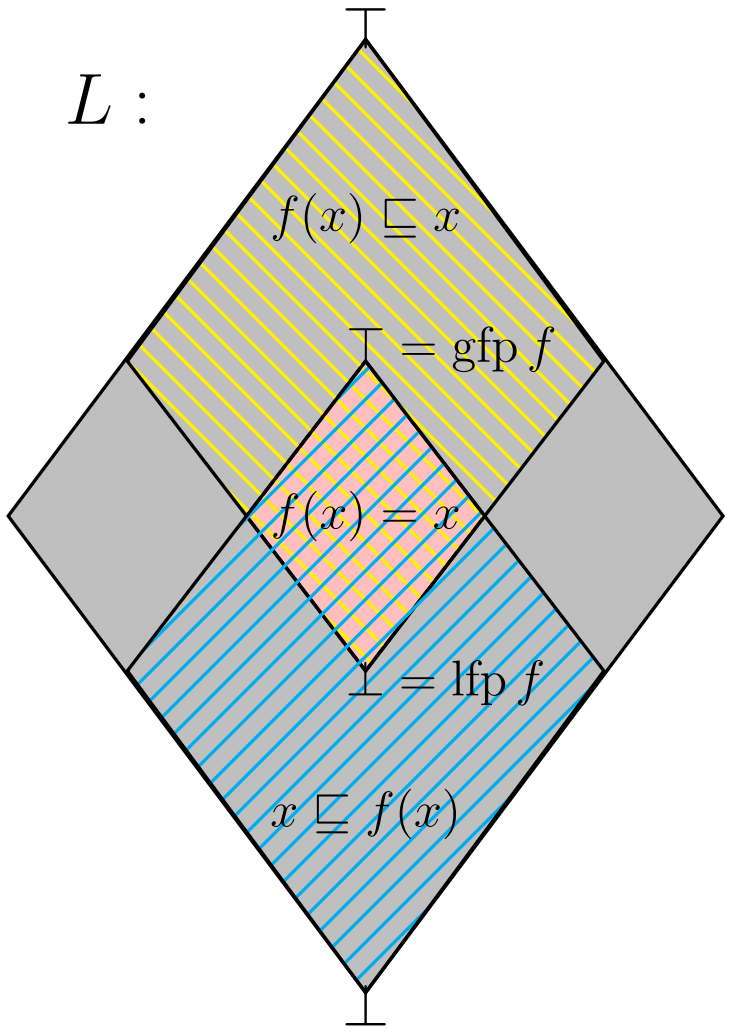
- 任给一个人，我们可以搜索一个 "被爱的" 路径，路径或者永不终结 (情形 1、情形 2)，或者以某个不被爱的男人终结 (情形 3)，或者以某个不被爱的女人终结 (情形 4).
- 在情形 1、情形 2 中，可以让每个男人娶了他所爱的女人 (或让每个女人嫁给她所爱的男人).
- 在情形 3 中，让每个男人都娶了他所爱的女人.
- 在情形 4 中，让每个女人都嫁给她所爱的男人.

Tarski's Fixpoint Theorem

- A complete lattice is a partially ordered set in which every non-empty subset has both a supremum and an infimum.

Theorem (Tarski's Fixpoint Theorem)

For a complete lattice $(L, \sqsubseteq)$ and an order-preserving function $f : L \to L$, the set of fixpoints of f is also a complete lattice, with greatest fixpoint $\bigsqcup\{x : x \sqsubseteq f(x)\}$ and least fixpoint $\bigsqcap\{x : f(x) \sqsubseteq x\}$.



Proof.

Let $D := \{x \in L : x \sqsubseteq f(x)\}$, and $u := \bigsqcup D$.
Then for $x \in D$, $x \sqsubseteq u$ and $x \sqsubseteq f(x) \sqsubseteq f(u)$, hence $u \sqsubseteq f(u)$.
Then $f(u) \sqsubseteq f(f(u))$.
So we have $f(u) \in D$ and $f(u) \sqsubseteq u$, from which follows $f(u) = u$. □

Induction vs Coinduction

$$\frac{F \text{ monotone} \quad F(X) \sqsubseteq X}{\mu F \sqsubseteq X}$$

$$\frac{F \text{ monotone} \quad X \sqsubseteq F(X)}{X \sqsubseteq \nu F}$$

where $\mu F = \prod\{X : F(X) \sqsubseteq X\}$ is the least fixpoint, and $\nu F = \coprod\{X : X \sqsubseteq F(X)\}$ is the greatest fixpoint.

Example (Mathematical Induction)

$$F : \mathsf{P}(\mathbb{N}) \to \mathsf{P}(\mathbb{N}) :: X \mapsto \{0\} \cup \{x + 1 : x \in X\}$$

$$\frac{F \text{ monotone} \quad F(X) \subset X}{\mu F = \mathbb{N} = X}$$

Remark: coinduction can be viewed as induction on the complement, let $Y := \overline{X}$, and $G(Y) := \overline{F(\overline{Y})}$, then $X \subset F(X) \iff G(Y) \subset Y$. So $\mu G = \nu F$.

Induction vs Coinduction — Examples

- Defining the set of points that are reachable from a point w_0 in a model as μF where

$$F(X) = \{w : w_0 \to w\} \cup \{w : \exists v \in X : v \to w\}$$

- Defining the set of points that have infinite descending chains in a model as νF where

$$F(X) = \{w : \exists v \in X : w \to v\}$$

Cantor-Schröder-Bernstein Theorem — another proof

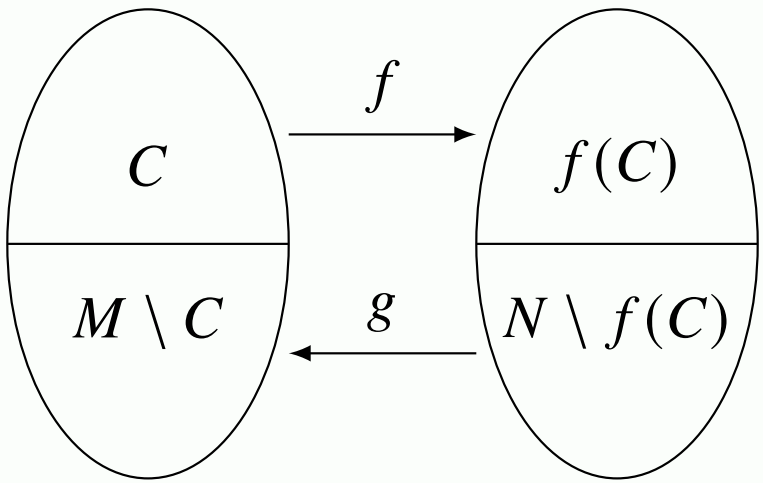
Proof.

$(\mathrm{P}(M), \subset)$ is a complete lattice.
Since the map

$$h : S \mapsto M \setminus g(N \setminus f(S))$$

is nondecreasing, it has a fixpoint C and $M \setminus C = g(N \setminus f(C))$.

$$f\restriction_C \cup g^{-1}\restriction_{M\setminus C}$$



□

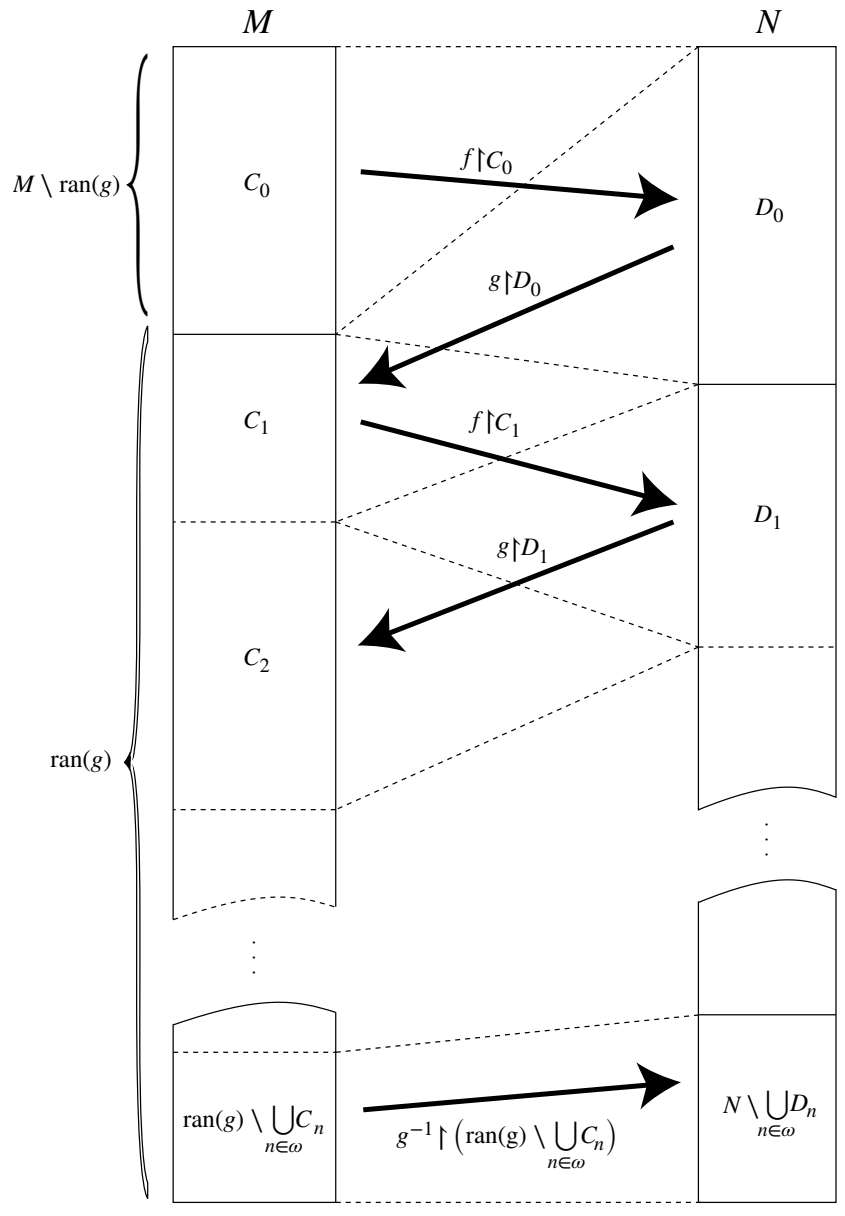
Cantor-Schröder-Bernstein Theorem — another proof

Proof.

$$C_0 \coloneqq M \setminus g(N)$$
$$D_0 \coloneqq f(C_0)$$
$$C_{n+1} \coloneqq g(D_n)$$
$$D_{n+1} \coloneqq f(C_n)$$
$$C \coloneqq \bigcup_{n \in \mathbb{N}} C_n$$
$$f\restriction_C \cup g^{-1}\restriction_{g(N)\setminus C}$$

□

$$h : S \mapsto (M \setminus g(N)) \cup g(f(S))$$
$$C = \bigcup_{n \to \infty} h^n(\emptyset)$$
$$= \bigcap \left\{ S : (M \setminus g(N)) \cup g(f(S)) \subset S \right\}$$



Remark

- 我们把 M 中的元素比作客人, 而把 N 中的元素比作房间.
- 对于任意 $n \in N$, $g(n)$ 就是 n 房间里住的客人.
- 所有房间里住的客人的集合就是 $g(N)$.
- $C_0 := M \setminus g(N)$ 就是还没安排房间的新客人.
- 把 C_0 中的客人 m 安排进 $f(m)$ 房间中, 即把 C_0 中的客人安排在 $D_0 := f(C_0)$ 中的房间.
- 把原来住在 D_0 这些房间的客人 (即 $C_1 := g(D_0)$) 重新安排到 $D_1 := f(C_1)$ 房间中, 再把原来住在 D_1 房间里的客人 (即 $C_2 := g(D_1)$) 安排在 $D_2 := f(C_2)$.
- 如此继续, 正好把 M 中的客人一一对应地安排在 N 中的房间了.

Cantor-Schröder-Bernstein Theorem — another proof

Proof.

$$M_0 \coloneqq M, \quad M_1 \coloneqq g(N), \quad M_{k+2} \coloneqq g \circ f(M_k)$$

$$M_0 \supset M_1 \supset M_2 \supset \cdots \supset M_k \supset M_{k+1} \supset \cdots$$

$$A \coloneqq \bigcup_{k=0}^{\infty} (M_{2k+1} \setminus M_{2k+2}) \quad B \coloneqq \bigcup_{k=0}^{\infty} (M_{2k} \setminus M_{2k+1}) \quad C \coloneqq \bigcup_{k=1}^{\infty} (M_{2k} \setminus M_{2k+1})$$

$$D \coloneqq \bigcap_{k=0}^{\infty} M_k$$

$$M = A \cup B \cup D$$

$$M_1 = A \cup C \cup D$$

$$|M_{2k} \setminus M_{2k+1}| = |g \circ f(M_{2k}) \setminus g \circ f(M_{2k+1})| = |M_{2k+2} - M_{2k+3}| \implies |B| = |C|$$

$$|M| = |M_1| = |N|$$

Countable ordinals are uncountable

Theorem (Countable ordinals are uncountable)

There does not exist a countable enumeration of the countable ordinals.

Proof.

- Suppose for contradiction that there exists a countable enumeration $\omega_1, \omega_2, \ldots$ of the countable ordinals.
- Then the set $\Omega := \bigcup_n \omega_n$ is also a countable ordinal, as is the set $\Omega \cup \{\Omega\}$.
- But $\Omega \cup \{\Omega\}$ is not equal to any of the ω_n (by the axiom of foundation), a contradiction.

□

Cardinal Arithmetic

Definition (Cardinal Arithmetic)

$$\kappa + \lambda = |\,(A \times \{0\}) \cup (B \times \{1\})\,|$$
$$\kappa \cdot \lambda = |A \times B|$$
$$\kappa^{\lambda} = |A^{B}|$$

where $|A| = \kappa, |B| = \lambda$.

Theorem

- $+$ *and* $\cdot$ *are associative, commutative and distributive.*
- $(\kappa \cdot \lambda)^{\mu} = \kappa^{\mu} \cdot \lambda^{\mu}$
- $\kappa^{\lambda+\mu} = \kappa^{\lambda} \cdot \kappa^{\mu}$
- $(\kappa^{\lambda})^{\mu} = \kappa^{\lambda \cdot \mu}$
- $\kappa \leq \lambda \implies \kappa + \mu \leq \lambda + \mu \;\&\; \kappa \cdot \mu \leq \lambda \cdot \mu \;\&\; \kappa^{\mu} \leq \lambda^{\mu}$
- $0 < \lambda \leq \mu \implies \kappa^{\lambda} \leq \kappa^{\mu}$
- $\kappa^{0} = 1;\ 1^{\kappa} = 1;\ 0^{\kappa} = 0$ *if* $\kappa > 0$.

Alephs

- Since $\mathrm{Card} \subset \mathrm{Ord}$, Card is well-ordered and the elements of Card can be enumerated with Ord as indices.
- For any cardinal κ, κ^+ denotes the least cardinal $> \kappa$.
- The Aleph function $\aleph$ is define by the transfinite recursion:

$$\begin{aligned} \aleph_0 &= \omega \\ \aleph_{\alpha+1} &= \aleph_\alpha^+ \\ \aleph_\alpha &= \lim_{\beta \to \alpha} \aleph_\beta \quad \text{for limit } \alpha \end{aligned}$$

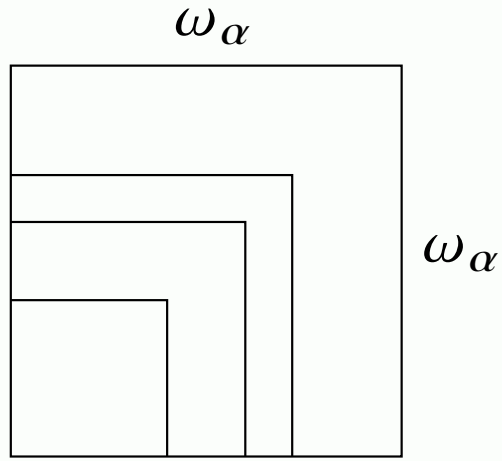
Theorem

$$\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$$

Proof.

We define $(\alpha, \beta) < (\gamma, \delta)$ iff ether

- $\max\{\alpha, \beta\} < \max\{\gamma, \delta\}$, or
- $\max\{\alpha, \beta\} = \max\{\gamma, \delta\}$ and $\alpha < \gamma$, or
- $\max\{\alpha, \beta\} = \max\{\gamma, \delta\}, \alpha = \gamma$ and $\beta < \delta$.



Obviously, $<$ is a well-order on $\mathrm{Ord} \times \mathrm{Ord}$ and $\alpha \times \alpha = \{(\xi, \eta) : (\xi, \eta) < (0, \alpha)\}$.

Let $\Gamma(\alpha, \beta) := \mathrm{otp}\{(\xi, \eta) : (\xi, \eta) < (\alpha, \beta)\}$. Then $(\alpha, \beta) < (\gamma, \delta) \iff \Gamma(\alpha, \beta) < \Gamma(\gamma, \delta)$, and $\Gamma(\omega, \omega) = \omega$, $\Gamma(\alpha, \alpha) \geq \alpha$. Assume α is the least ordinal s.t. $\Gamma(\omega_\alpha, \omega_\alpha) \neq \omega_\alpha$. Let $\beta, \gamma < \omega_\alpha$ s.t. $\Gamma(\beta, \gamma) = \omega_\alpha$. Pick $\delta < \omega_\alpha$ s.t. $\delta > \beta, \gamma$. We have $\Gamma(\delta, \delta) \supset \omega_\alpha$ and so $|\delta \times \delta| \geq \aleph_\alpha$. However, $|\delta \times \delta| = |\delta| \cdot |\delta| = |\delta| < \aleph_\alpha$. Contradiction. □

$$\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \max\{\aleph_\alpha, \aleph_\beta\}$$

Cantor Set



$$C_0 := [0,1]$$

$$C_{k+1} := \frac{C_k}{3} \cup \left(\frac{2}{3} + \frac{C_k}{3}\right)$$

$$C := \bigcap_{k=0}^{\infty} C_k = \bigcap_{n=1}^{\infty} \bigcup_{k=0}^{3^{n-1}-1} \left(\left[\frac{3k+0}{3^n}, \frac{3k+1}{3^n}\right] \cup \left[\frac{3k+2}{3^n}, \frac{3k+3}{3^n}\right]\right)$$

$$= [0,1] \setminus \bigcup_{n=1}^{\infty} \bigcup_{k=0}^{3^{n-1}-1} \left(\frac{3k+1}{3^n}, \frac{3k+2}{3^n}\right)$$

$$C = \left\{ \sum_{n=1}^{\infty} \frac{x_n}{3^n} : x_n = 0 \vee x_n = 2 \right\} \implies |C| = 2^{\aleph_0}$$

Banach's Fixpoint Theorem

Theorem (Banach's Fixpoint Theorem)

Let (M, d) be a complete metric space and $T : M \to M$ be a contraction mapping, with Lipschitz constant $\gamma < 1$. Then T has a unique fixpoint $x \in M$. Further, for each $x_0 \in M$, $\lim_{n\to\infty} T^n(x_0) = x$, and the convergence is geometric:

$$d(T^n(x_0), x) \leq \gamma^n d(x_0, x)$$

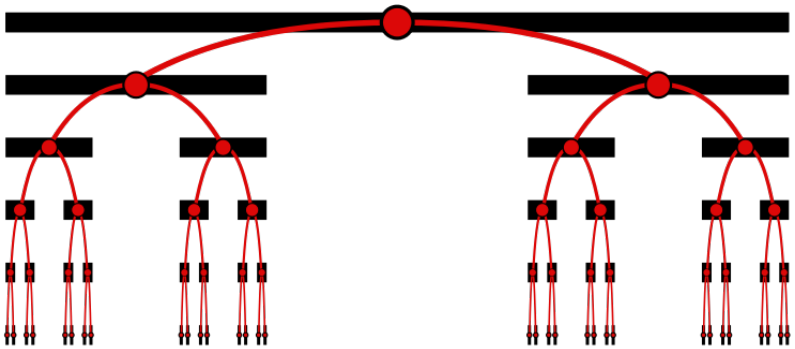
Banach's Fixpoint Theorem and Cantor Set

- Let (M, d) be a complete metric space and let $\mathcal{M}$ be the set of all non-empty bounded closed subsets of M.
  For $A \in \mathcal{M}$ and $\varepsilon > 0$, let $N_\varepsilon(A) := \{x \in M : d(x, A) < \varepsilon\}$ where $d(x, A) := \inf_{y \in A} d(x, y)$. Let

$$d_H(A, B) := \inf \{\varepsilon : A \subset N_\varepsilon(B) \ \& \ B \subset N_\varepsilon(A)\}$$

  Then $(\mathcal{M}, d_H)$ is a complete metric space.
- Let $T_i : M \to M, i = 1, \ldots, n$ be a set of contractions. Let $\mathcal{M}'$ be the set of all compact sets of $\mathcal{M}$. Define the map $S : \mathcal{M}' \to \mathcal{M}'$ by $S(X) = \bigcup_{i=1}^{n} T_i(X)$. Then S is a contraction.
- According to Banach's fixpoint theorem, $\exists X \in \mathcal{M}' : S(X) = X$.
  Furthermore, $\forall Y \in \mathcal{M}' : S(Y) \subset Y \implies X = \bigcap_{k=0}^{\infty} S^k(Y)$.
- Cantor set C is the fixpoint of $x \mapsto \frac{x}{3} \cup \left(\frac{x}{3} + \frac{2}{3}\right)$.

Cantor Set



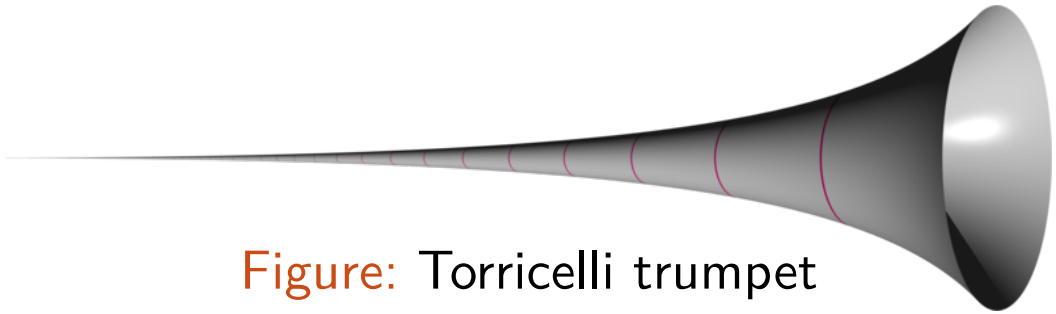
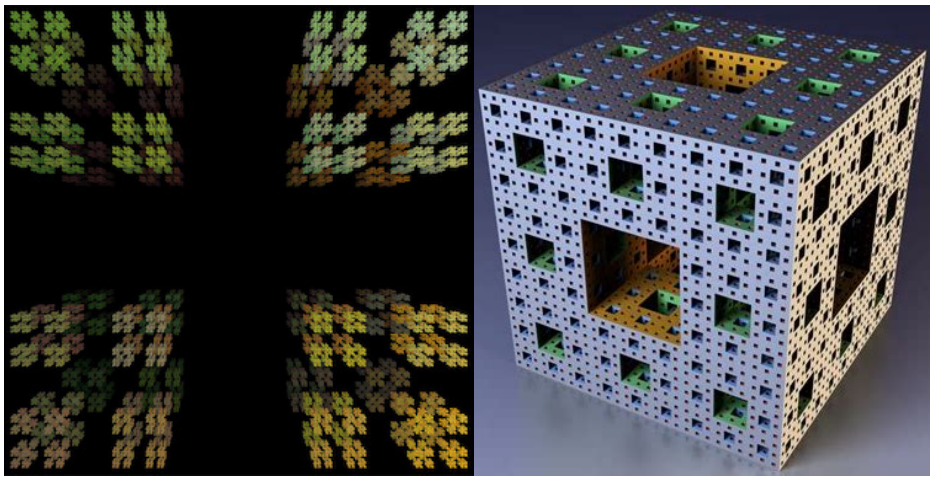


Figure: Torricelli trumpet

- $|C| = 2^{\aleph_0}$
- C is perfect.
- C is *nowhere dense* in $[0, 1]$.

$$\big(\operatorname{cl}(C)\big)^\circ = \emptyset$$

- Lebesgue measure: 0
- Hausdorff dimension: $\log_3 2$
- compact metric space



(a) Cantor dust(3D) (b) Menger sponge: infinite surface area but 0 volume

Fractal, Hausdorff Dimension, Topological Dimension

A set A is a *fractal* iff $\dim_H(A) > \dim_T(A)$.

$$H^d_\varepsilon(A) := \inf\left\{\sum_{k=1}^{\infty} \mathrm{diam}(B_k)^d : A \subset \bigcup_{k=1}^{\infty} B_k \ \ \& \ \ \mathrm{diam}(B_k) \le \varepsilon\right\}$$

$$\dim_H(A) := \inf\left\{d \ge 0 : \lim_{\varepsilon \to 0} H^d_\varepsilon(A) = 0\right\}$$

Definition (Topological Dimension)

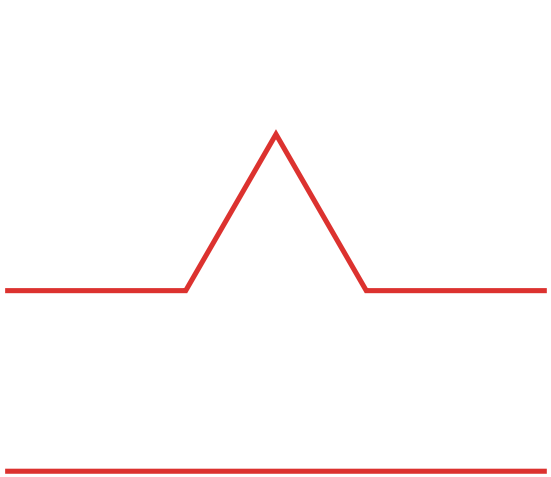
The *topological dimension* of a space X is defined by induction as

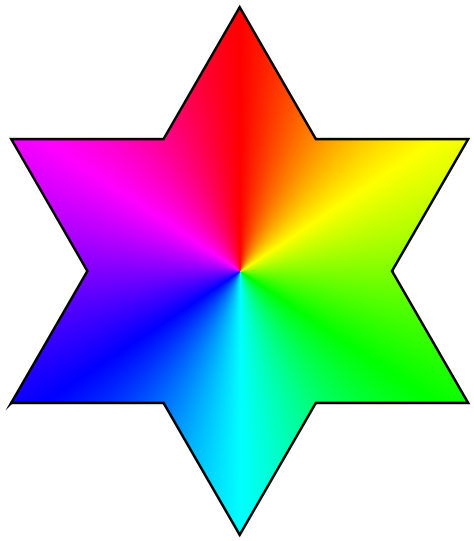
$$\dim_T(\emptyset) := -1$$

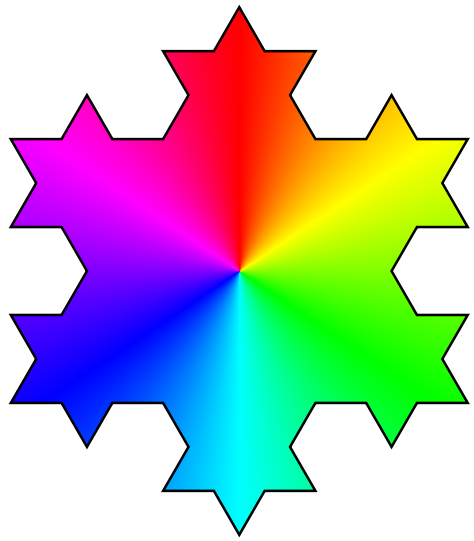
$$\dim_T(X) := \inf\left\{d : X \text{ has a basis } \mathcal{U} \text{ s.t. } \forall U \in \mathcal{U} : \dim_T(\partial U) \le d - 1\right\}.$$

The topological dimension of a space X is the smallest integer d such that every open cover $\mathcal{U}$ of X has a refinement $\mathcal{V}$ in which no point of X lies in more than $d + 1$ elements of $\mathcal{V}$.

科赫曲线 Koch Curve





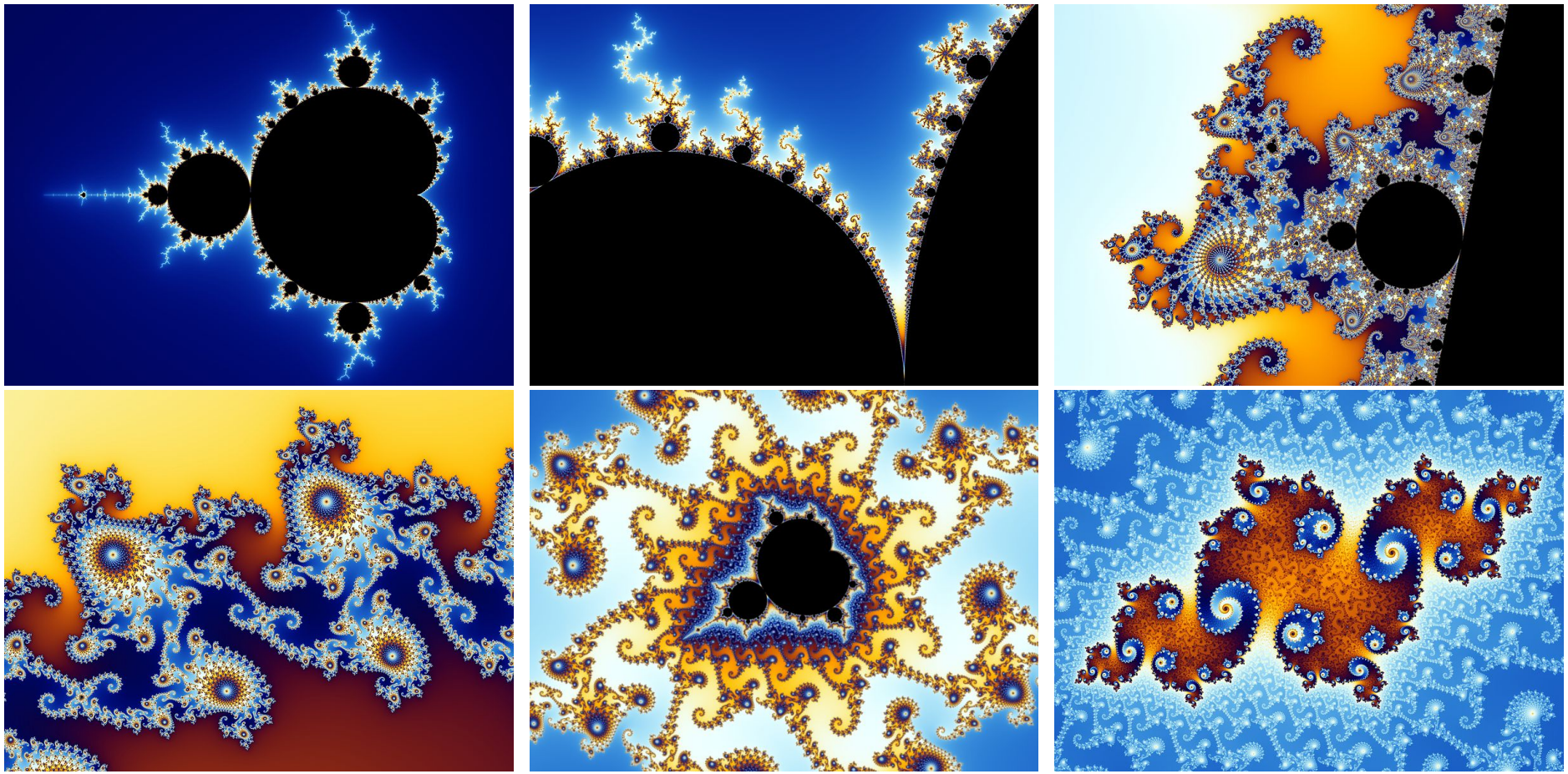


科赫曲线的自相似维数为:

$$D := \frac{\log n}{\log s} = \frac{\log 4}{\log 3} \approx 1.26$$

其中, n 为子块数, s 为缩减因子.

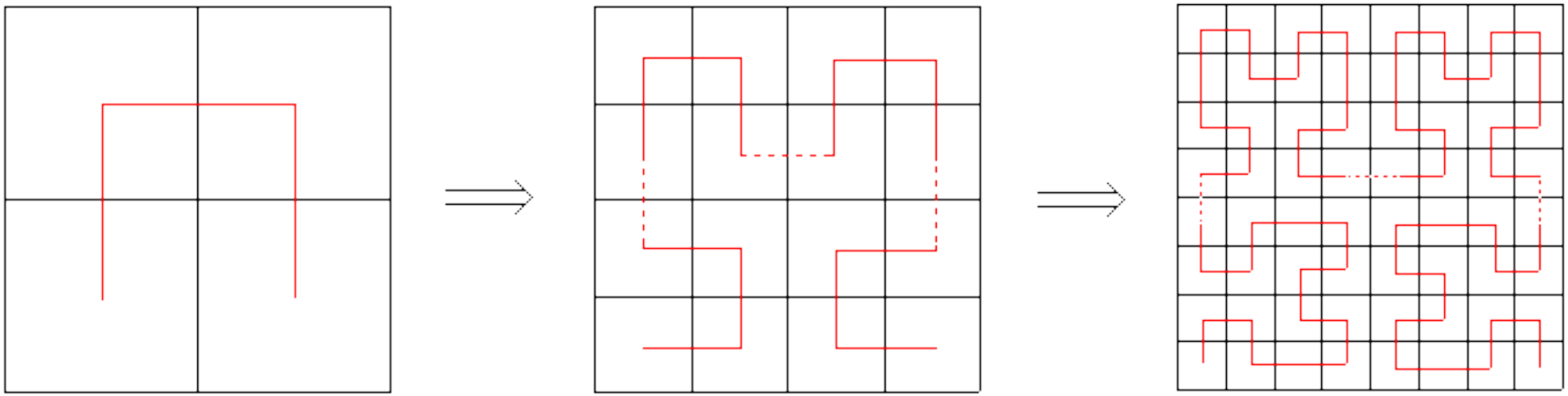
Mandelbrot Set — complex structure from simple rule

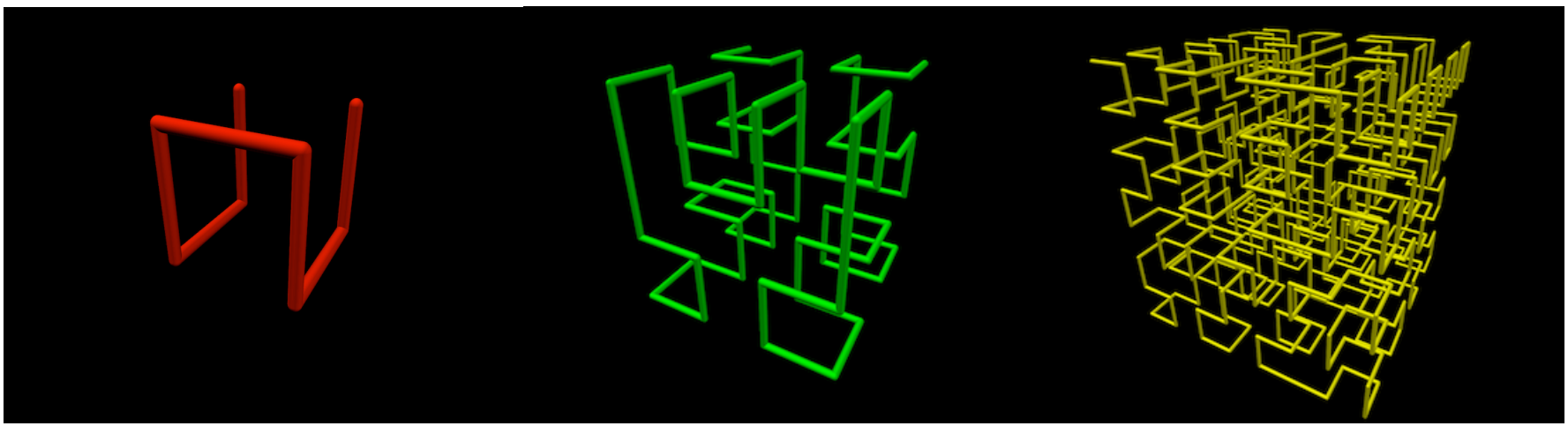


$$z \mapsto z^2 + c$$

- Hausdorff dimension of the boundary of the Mandelbrot set: 2
- Topological dimension of the boundary of the Mandelbrot set: 1

希尔伯特空间填充曲线





Hilbert's Space-filling Curve

- When we draw h_n, we impose a $2^n \times 2^n$ grids onto the square S. The diagonal of each grid is of length $\sqrt{(2^{-n})^2 + (2^{-n})^2} = 2^{\frac{1}{2}-n}$.
- We define the curve h as the limit of these successive functions $h_1, h_2 \ldots$ s.t. $h(x) = \lim_{n\to\infty} h_n(x)$.
- Each point in S is at most $2^{\frac{1}{2}-n}$ distance away from some point on h_n. So the maximum distance of any point from h is $\lim_{n\to\infty} 2^{\frac{1}{2}-n} = 0$. So h fills space!
- Definition. A curve is a continuous map from unit interval L to unit square S.
- For a pont $p \in S$ and $\varepsilon > 0$, there is some n s.t. some grid of the $2^n \times 2^n$ grids on S lies within the circle with centre p and radius ε. let I be the largest open part of L which h_n maps into the relevant grid. Whenever $x \in I$, $h_m(x)$ lies in that same grid, for any $m > n$. So h is continuous.
- Hilbert's curve is continuous everywhere but differentiable nowhere.
- Hausdorff dimension: 2
- Topological dimension: 1

Cardinality of the Permutations of $\mathbb{N}$ — $|\mathrm{Aut}(\mathbb{N})|$

Proof1. Diagonal method: For any countable sequence $(\sigma_n)_{n\in\mathbb{N}}$ of permutations, let $f : 2n \mapsto \min\left(2\mathbb{N} \setminus \{\sigma_0(0), \sigma_1(2), \ldots, \sigma_n(2n)\}\right)$, and $f : 2n+1 \mapsto$ the n^{th} element of $\mathbb{N} \setminus f(2\mathbb{N})$. Then $\forall n : f \neq \sigma_n$.

Proof1′. Let f be the bijection s.t. for each n, f swaps $2n$ and $2n+1$ if $\sigma_n(2n) = 2n$, leaving the rest fixed. $|\mathrm{Aut}(\mathbb{N})| > \aleph_0$.

Proof2. For any n, either swap $(2n, 2n+1)$ or keep them fixed.

Proof2′. The set of fixpoints of any permutation can be any subset of $\mathbb{N}$ except ones of the form $\mathbb{N} \setminus \{n\}$ for some n. $|\mathrm{Aut}(\mathbb{N})| \geq 2^{\aleph_0}$.

Proof3. Riemann rearrangement $\left(e.g.\ \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\right) \Longrightarrow \mathbb{R} \rightarrowtail \mathrm{Aut}(\mathbb{N})$.

Theorem (Riemann Rearrangement Theorem)

Any conditionally convergent series can be rearranged in a permutation to

1. *converge to any real number;*
2. *diverge to ∞ or $-\infty$;*
3. *oscillate finitely or infinitely.*



The Cardinality of all Continuous Functions $C(\mathbb{R}, \mathbb{R})$

$$|C(\mathbb{R}, \mathbb{R})| = |\mathbb{R}|$$

Proof.

Obviously, $|\mathbb{R}| \leq |C(\mathbb{R}, \mathbb{R})|$, since all constant functions are continuous.
On the other hand, suppose $f, g \in C(\mathbb{R}, \mathbb{R})$ and $f\restriction_{\mathbb{Q}} = g\restriction_{\mathbb{Q}}$.
For any $x \in \mathbb{R}$, there exists a sequence of rational numbers q_n s.t. $\lim_{n\to\infty} q_n = x$, according to the continuity of f, g:

$$\left.\begin{array}{c} \lim_{n\to\infty} f(q_n) = f(\lim_{n\to\infty} q_n) = f(x) \\ \| \\ \lim_{n\to\infty} g(q_n) = g(\lim_{n\to\infty} q_n) = g(x) \end{array}\right\} \Longrightarrow f = g$$

In other words, a continuous function on $\mathbb{R}$ is determined by its values at rational points. The map $\Phi : C(\mathbb{R}, \mathbb{R}) \to \mathbb{R}^{\mathbb{Q}} :: f \mapsto f\restriction_{\mathbb{Q}}$ is injective.

$$|C(\mathbb{R}, \mathbb{R})| \leq |\mathbb{R}^{\mathbb{Q}}| = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0} = |\mathbb{R}|$$

The Pasadena Paradox

The Pasadena Game

Toss a fair coin until the first head appears. If the first head appears on toss n, the payoff is $\frac{(-1)^{n+1}2^n}{n}$.

How to calculate the expected utility?

$$\sum_{n=1}^{\infty} \frac{1}{2^n} \frac{(-1)^{n+1}2^n}{n} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln 2 \qquad \sum_{n=1}^{\infty} \left| \frac{(-1)^{n+1}}{n} \right| = \infty$$

Cofinality and Inaccessible Cardinal

- $\mathrm{cf}(\alpha) :=$ the least limit ordinal β s.t. there is an increasing β-sequence $\langle \alpha_\xi : \xi < \beta \rangle$ with $\lim\limits_{\xi \to \beta} \alpha_\xi = \alpha$.

  Example:
  - If $\alpha = \beta + 1$ is a successor ordinal, then $\mathrm{cf}(\alpha) = 1$.
  - $\mathrm{cf}(\omega) = \mathrm{cf}(\omega + \omega) = \mathrm{cf}(\omega^n) = \mathrm{cf}(\aleph_\omega) = \mathrm{cf}(\aleph_{\aleph_\omega}) = \omega$
  - $\mathrm{cf}(\aleph_1) = \aleph_1$
  - $\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha)$
  - If α is a limit ordinal, then $\mathrm{cf}(\aleph_\alpha) = \mathrm{cf}(\alpha)$.
- An infinite cardinal κ is regular iff $\mathrm{cf}(\kappa) = \kappa$.
- It is singular iff $\mathrm{cf}(\kappa) < \kappa$.
- A cardinal κ is a strong limit cardinal iff $\forall \lambda < \kappa \big(2^\lambda < \kappa\big)$.
- A cardinal κ is inaccessible iff it is a regular strong limit uncountable cardinal.

Theorem (König's Theorem)

If κ is an infinite cardinal, then $\kappa < \kappa^{\mathrm{cf}(\kappa)}$.

von Neumann Universe

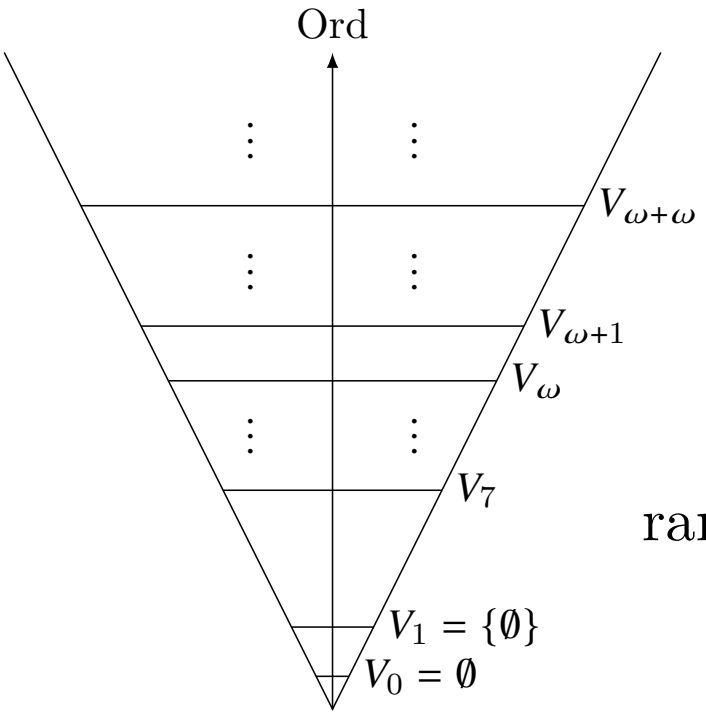
Definition (von Neumann Universe)

$$V_0 := \emptyset$$

$$V_{\alpha+1} := \mathrm{P}(V_\alpha)$$

$$V_\alpha := \bigcup_{\beta<\alpha} V_\beta \quad \text{for limit } \alpha.$$





$$V := \{x : x = x\}$$

$$\boxed{V = \bigcup_{\alpha\in\mathrm{Ord}} V_\alpha}$$

$\mathrm{rank}(x) :=$ the least α s.t. $x \subset V_\alpha$

Constructable Universe

$$\mathrm{Def}(M) := \{X \subset M : X \text{ is } M\text{-definable over } (M, \in)\}$$

Definition (Constructable Universe)

$$\begin{aligned} L_0 &:= \emptyset \\ L_{\alpha+1} &:= \mathrm{Def}(L_\alpha) \\ L_\alpha &:= \bigcup_{\beta<\alpha} L_\beta \quad \text{for limit } \alpha. \\ L &:= \bigcup_{\alpha\in\mathrm{Ord}} L_\alpha \end{aligned}$$

Axiom of Constructibility

$$V = L$$

$$\begin{aligned} &L \vDash \mathrm{ZF} \\ &L \vDash V = L \\ &\mathrm{ZF} + \mathrm{V} = \mathrm{L} \vdash \mathrm{AC} + \mathrm{GCH} \end{aligned}$$

An inner model is a transitive class model of ZF that contains all ordinals.
L is the smallest inner model of ZF.

Grothendieck Universe

Definition (Grothendieck Universe)

1. $x \in y \in U \implies x \in U$
2. $x \in U \ \& \ y \in U \implies \{x, y\} \in U$
3. $x \in U \implies \mathrm{P}(x) \in U$
4. $I \in U \ \& \ x : I \to U \implies \bigcup_{i \in I} x_i \in U$
5. $\omega \in U$



Universe Axiom

For every set x, there exists a Grothendieck universe U s.t. $x \in U$.

U is a Grothendieck universe iff $U = V_\kappa$ for some inaccessible κ.

The Universe Axiom is equivalent to the "inaccessible cardinal axiom" that "there exist arbitrarily large inaccessible cardinals."

$$U \vDash \mathrm{ZFC}$$

$$\frac{\text{super-infinite}}{\text{infinite}} \approx \frac{\text{infinite}}{\text{finite}}$$

finite	$\iff$	every self-embedding is bijective.
infinite	$\iff$	admits a non-surjective self-embedding.
super-infinite	$\iff$	admits a non-surjective elementary self-embedding.

Example: $\mathbb{N}$ is infinite but not super-infinite.

Axiom I3

For some λ, V_λ is super-infinite.

Shelf

Definition (Shelf)

A left (right) *shelf* is a set S with an operation $*$ satisfying

$$x * (y * z) = (x * y) * (x * z) \qquad \text{(left self-distributive)}$$

$$(x * y) * z = (x * z) * (y * z) \qquad \text{(right self-distributive)}$$

Example:

- S set, $f : S \to S$, and $x * y := f(y)$
- E module and $x * y := (1 - \lambda)x + \lambda y$
- G group and $x * y := xyx^{-1}$
- B boolean algebra and $x * y := \overline{x} + y$

Under the logical interpretation, $*$ corresponds to implication $\to$.

Laver Tables

Theorem (Laver)

1. *For every N, there exists a unique binary operation $*$ on $\{1, \ldots, N\}$ s.t.*

$$x * 1 = x + 1 \bmod N$$

$$x * (y * 1) = (x * y) * (x * 1)$$

2. *The operation thus obtained obeys*

$$x * (y * z) = (x * y) * (x * z)$$

iff N is a power of 2.

Laver Tables

Definition (Laver Table)

Laver table A_n is the unique left shelf $(\{1, \ldots, 2^n\}, *)$ satisfying

$$x * 1 = x + 1 \bmod 2^n$$

A_0	1
1	1

A_1	1	2
1	2	2
2	1	2

A_2	1	2	3	4
1	2	4	2	4
2	3	4	3	4
3	4	4	4	4
4	1	2	3	4

$x \mapsto x \bmod 2^{n-1}$ is a surjective homomorphism from A_n to A_{n-1}.

Period

Theorem (Laver)

For every $p \leq 2^n$, there exists a number $\pi_n(p)$, a power of 2, such that the p^{th} row in the table of A_n is the repetition of $\pi_n(p)$ values increasing from $p + 1 \bmod 2^n$ to 2^n.

$$\pi_n(p) := \mu x \, [p * x = 2^n]$$

A_3	1	2	3	4	5	6	7	8	period
1	2	4	6	8	2	4	6	8	$\pi_3(1) = 4$
2	3	4	7	8	3	4	7	8	$\pi_3(2) = 4$
3	4	8	4	8	4	8	4	8	$\pi_3(3) = 2$
4	5	6	7	8	5	6	7	8	$\pi_3(4) = 4$
5	6	8	6	8	6	8	6	8	$\pi_3(5) = 2$
6	7	8	7	8	7	8	7	8	$\pi_3(6) = 2$
7	8	8	8	8	8	8	8	8	$\pi_3(7) = 1$
8	1	2	3	4	5	6	7	8	$\pi_3(8) = 8$

n	0	1	2	3	4	5	6	7	8	9	10	11	$\cdots$
$\pi_n(1)$	1	1	2	4	4	8	8	8	8	16	16	16	$\cdots$
$\pi_n(2)$	–	2	2	4	4	8	8	16	16	16	16	16	$\cdots$

$\mu n[\pi_n(1) = 32] \geq A(9, A(8, A(8, 254)))$ where A is the Ackermann Function

Theorem (Laver)

1. $\mathrm{ZFC} + \mathrm{I3} \vdash \forall n \big(\pi_n(2) \geq \pi_n(1)\big)$
2. $\mathrm{ZFC} + \mathrm{I3} \vdash \lim\limits_{n \to \infty} \pi_n(1) = \infty$

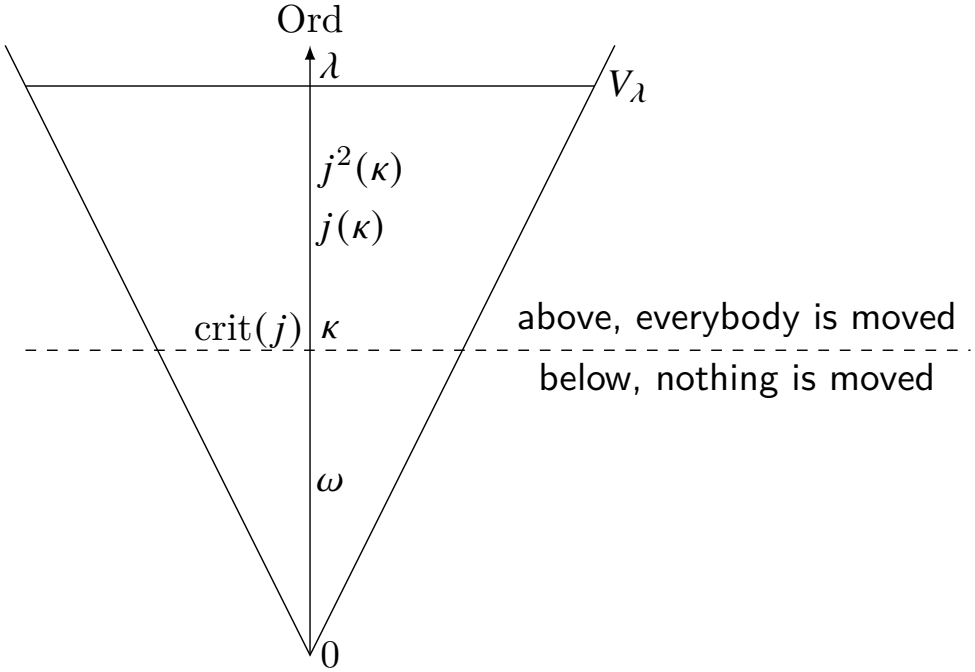
Can one find alternative proofs using no large cardinal?

Analogy:

- In physics: using a physical intuition, guess statements, then pass them to the mathematician for a formal proof;
- In logic: using a logical intuition (existence of a super-infinite set), guess statements (periods in Laver tables tend to ∞), then pass them to the mathematician for a formal proof.

$$\mathrm{crit}(j) \coloneqq \mu\alpha[j(\alpha) > \alpha]$$

The critical ordinal $\mathrm{crit}(j)$ of the elementary embedding j is the least ordinal that is not invariant under j.



$$\mathcal{E}_\lambda \coloneqq \{j : V_\lambda \prec V_\lambda \ \ \& \ \ j \text{ is non-surjective}\}$$

$$i[j] \coloneqq \bigcup_{\alpha<\lambda} i(j \cap V_\alpha^2)$$

$\big(\mathcal{E}_\lambda, []\big)$ is a left-shelf: $i[j[k]] = i[j][i[k]]$

$$\mathrm{crit}(j \circ j) = \mathrm{crit}(j) \quad \text{but} \quad \mathrm{crit}(j[j]) = j(\mathrm{crit}(j)) > \mathrm{crit}(j)$$

$$j_{[n]} := \underbrace{j[j][j]\cdots[j]}_{n \text{ times}}$$

$$\mathrm{Iter}(j) := \left\{ j_{[n]} : n \in \mathbb{N}^+ \right\}$$

$$\boxed{\left(\mathrm{Iter}(j), []\right) \text{ is a left-shelf}}$$

For $k, k' \in \mathrm{Iter}(j)$, declare $k \equiv_n k' := \forall x \in V_\gamma \left(k(x) \cap V_\gamma = k'(x) \cap V_\gamma \right)$ with $\gamma := \mathrm{crit}(j_{[2^n]})$. Then

$$\boxed{\mathrm{Iter}(j)/\equiv_n \text{ is (isomorphic to) the Laver table } A_n}$$

Ordinal

$0, 1, 2, 3, \ldots$

$\omega, \omega + 1, \omega + 2, \ldots$

$\omega \cdot 2, (\omega \cdot 2) + 1, (\omega \cdot 2) + 2, \ldots$

$\omega^2, \omega^2 + 1, \omega^2 + 2, \ldots$

$\omega^\omega, \omega^\omega + 1, \omega^\omega + 2, \ldots$

$\omega^{\omega^\omega}, \ldots$

$\varepsilon_0 = \omega^{\omega^{\omega^{\cdot^{\cdot}}}} = \sup\{\omega, \omega^\omega, \omega^{\omega^\omega}, \omega^{\omega^{\omega^\omega}}, \ldots\}$

$\varepsilon_1 = \sup\{\varepsilon_0 + 1, \omega^{\varepsilon_0+1}, \omega^{\omega^{\varepsilon_0+1}}, \omega^{\omega^{\omega^{\varepsilon_0+1}}}, \ldots\} = \sup\{0, 1, \varepsilon_0, {\varepsilon_0}^{\varepsilon_0}, {\varepsilon_0}^{{\varepsilon_0}^{\varepsilon_0}}, \ldots\}$

$\varepsilon_{\alpha+1} = \sup\{\varepsilon_\alpha + 1, \omega^{\varepsilon_\alpha+1}, \omega^{\omega^{\varepsilon_\alpha+1}}, \ldots\} = \sup\{0, 1, \varepsilon_\alpha, \varepsilon_\alpha^{\varepsilon_\alpha}, \varepsilon_\alpha^{\varepsilon_\alpha^{\varepsilon_\alpha}}, \ldots\}$

$\varepsilon_\alpha = \sup\{\varepsilon_\beta : \beta < \alpha\}$ if α is a limit ordinal.

$\boxed{\varepsilon_\alpha \text{ is countable iff } \alpha \text{ is countable}}$

$\boxed{\forall \alpha \geq 1 : \varepsilon_{\omega_\alpha} = \omega_\alpha}$

Veblen Hierarchy

$$\varphi_0(\alpha) := \omega^\alpha$$

$$\varphi_{\gamma+1}(\alpha) := \alpha^{\text{th}} \text{ ordinal s.t. } \varphi_\gamma(\beta) = \beta$$

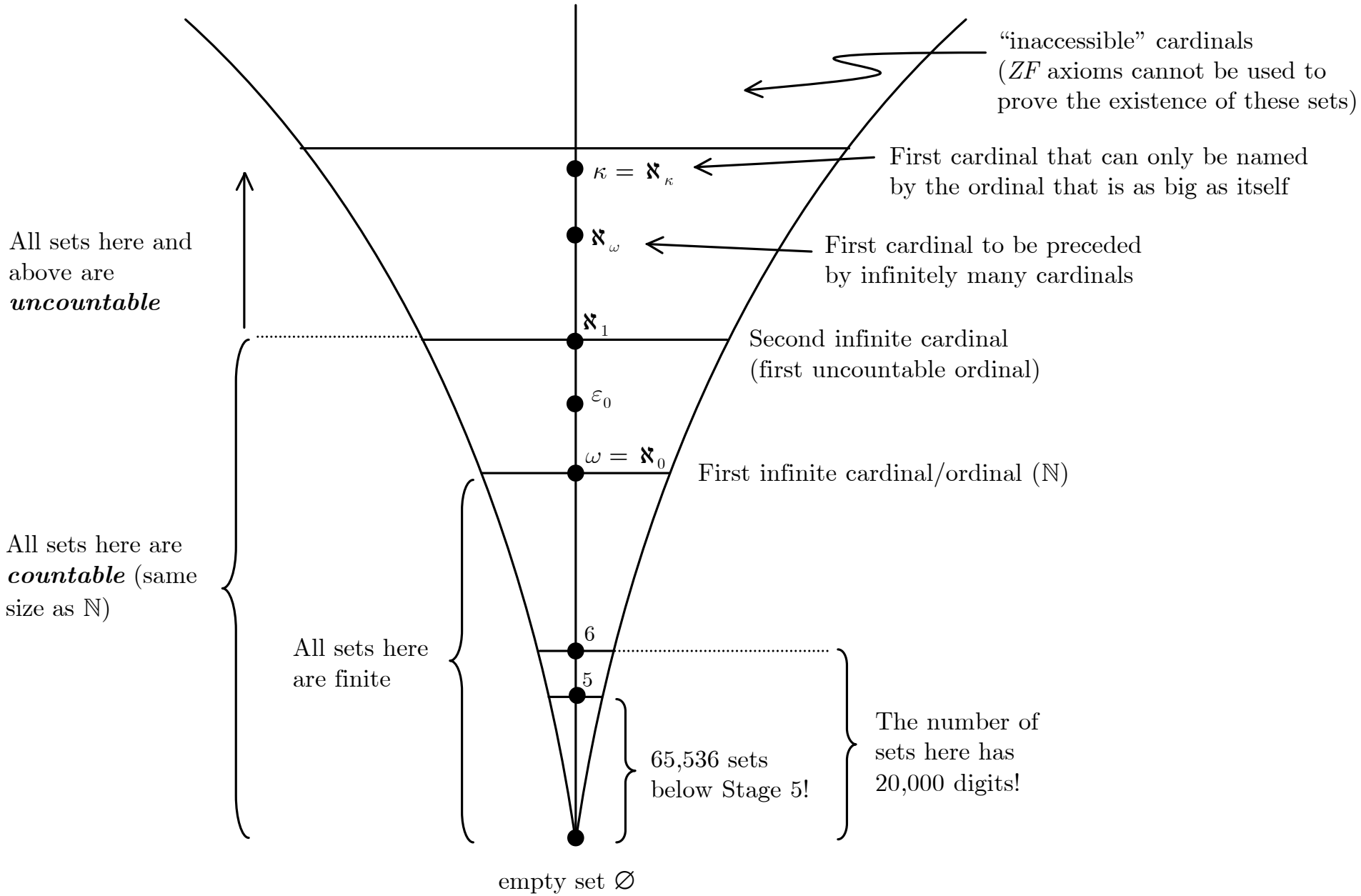
$$\varphi_\delta(\alpha) := \alpha^{\text{th}} \text{ common fixpoint of } \varphi_\gamma \text{ for all } \gamma < \delta$$

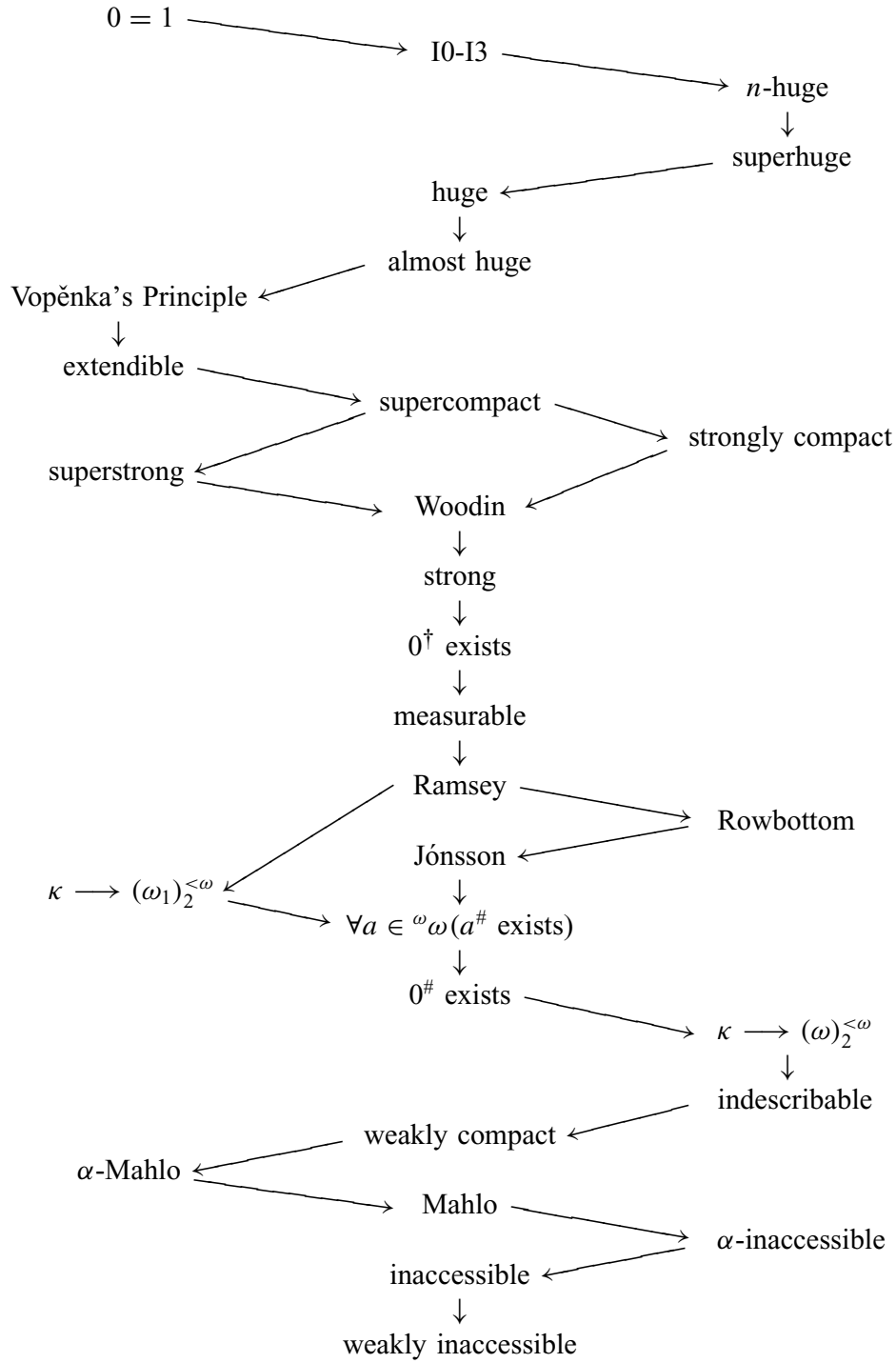
$$\Gamma_\alpha := \alpha^{\text{th}} \text{ ordinal s.t. } \varphi_\alpha(0) = \alpha$$

$$\boxed{\varepsilon_\alpha = \varphi_1(\alpha)}$$

About ε_0

- Gentzen: Transfinite induction on ε_0 proves $\mathrm{Con}(\mathrm{PA})$
- By Gödel's 2nd Incompleteness, PA can not prove transfinite induction for (or beyond) ε_0
- PA are not strong enough to show that ε_0 is an ordinal
- while ε_0 can easily be arithmetically described


"inaccessible" cardinals
(ZF axioms cannot be used to
prove the existence of these sets)
First cardinal that can only be named
by the ordinal that is as big as itself
$\kappa = \aleph_\kappa$
$\aleph_\omega$
First cardinal to be preceded
by infinitely many cardinals
All sets here and
above are
uncountable
$\aleph_1$
Second infinite cardinal
(first uncountable ordinal)
ε_0
$\omega = \aleph_0$
First infinite cardinal/ordinal (ℕ)
All sets here are
countable (same
size as ℕ)
All sets here
are finite
6
5
65,536 sets
below Stage 5!
The number of
sets here has
20,000 digits!
empty set ∅


0 = 1
I0-I3
n-huge
superhuge
huge
almost huge
Vopěnka's Principle
extendible
supercompact
strongly compact
superstrong
Woodin
strong
0† exists
measurable
Ramsey
Rowbottom
Jónsson
κ ⟶ (ω1)2<ω
∀a ∈ ω ω(a# exists)
0# exists
κ ⟶ (ω)2<ω
indescribable
weakly compact
α-Mahlo
Mahlo
α-inaccessible
inaccessible
weakly inaccessible

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory
- Axioms of ZFC
- Ordinal Numbers
- Cardinal Numbers
- Axiom of Choice

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

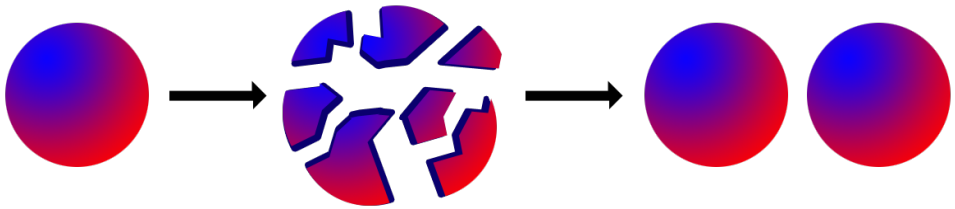
Answers to the Exercises

References2039

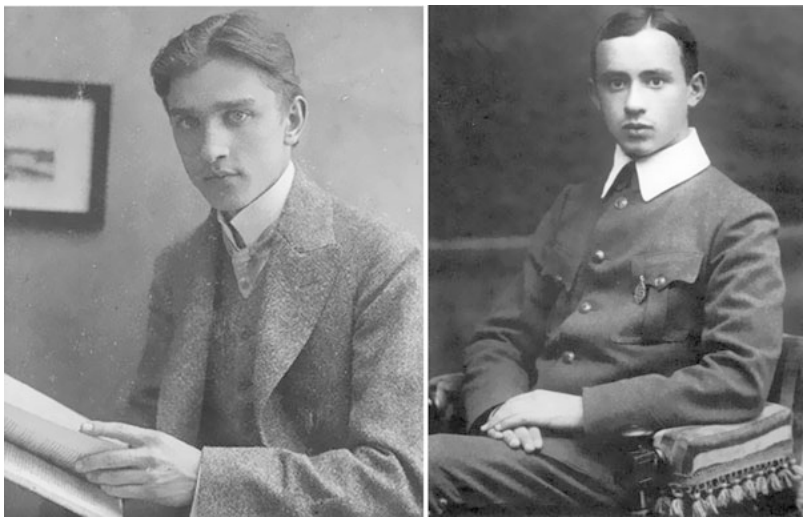
AC and Banach-Tarski Paradox

The Axiom of Choice is necessary to select a set from an infinite number of pairs of socks, but not an infinite number of pairs of shoes.

— Russell



对于 $n \geq 3$ 维空间中的一个球, 可以找到某种方式, 把其切分成有限 (比如 5) 份, 通过平移和旋转 (不能拉伸或撕裂) 进行重组, 能得到两个和原来一模一样的球.



Theorem (Banach-Tarski Theorem)

If A and B are bounded subsets of $\mathbb{R}^n, n \geq 3$, with non-empty interior, then there are finite partitions $A = \coprod_{i=1}^{n} A_i, B = \coprod_{i=1}^{n} B_i$ s.t. each A_i is congruent to B_i for $1 \leq i \leq n$.

AC vs AD

Axiom of Determinacy (AD)

Consider $A \subset \omega^\omega$. Two players alternately pick natural numbers

$$n_0, n_1, n_2, \ldots$$

Player 1 wins the game iff $(n_i)_{i\in\omega} \in A$.
The axiom of determinacy states that for every $A \subset \omega^\omega$, the game is determined, i.e. one of the two players has a winning strategy.

$$\forall A \subset \omega^\omega : \Big(\forall n_0 \exists n_1 \forall n_2 \exists n_3 \ldots \big[(n_i)_{i\in\omega} \in A\big]\Big) \vee \Big(\exists n_0 \forall n_1 \exists n_2 \forall n_3 \ldots \big[(n_i)_{i\in\omega} \notin A\big]\Big)$$

- AD is inconsistent with AC.
- AD implies countable axiom of choice.
- AD implies that every subset of reals is Lebesgue measurable.
- AD $\implies$ CH. Since GCH $\implies$ AC, AD is inconsistent with GCH.

Equivalents of AC

1. Well-ordering theorem: Every set can be well-ordered.
2. Trichotomy: For any two cardinals κ and λ : $\kappa < \lambda \vee \kappa = \lambda \vee \kappa > \lambda$.
3. For any infinite set A : $|A| = |A \times A|$.
4. Any two sets can be compared by their cardinalities.
5. The Cartesian product of any family of non-empty sets is non-empty.
6. Every surjective function has a right inverse.
7. Maximal Chain Principle: Each partial order has a maximal chain.
8. Zorn's lemma: If in a partially ordered set X each chain has an upper bound, then X has a maximal element.

Equivalents of AC

9. Every vector space has a basis.
10. Every nontrivial unitary ring contains a maximal ideal.
11. The closed unit ball of the dual of a normed vector space over the reals has an extreme point.
12. Tychonoff's theorem: Any product of compact spaces is compact in the product topology.
13. In the product topology, the closure of a product of subsets is equal to the product of the closures.
14. Any product of complete uniform spaces is complete.
15. If a set Γ of formulas with $|\mathscr{L}| = \kappa$ is finitely satisfiable, then it has a model with cardinality $\leq \kappa + \aleph_0$.

AC $\implies$ WO

Theorem (Well-Ordering Theorem)

Every set can be well-ordered.

Proof.

Assume f is a choice function for $\mathtt{P}(A) \setminus \{\emptyset\}$. Let

$$a_\alpha = f\left(A \setminus \left\{a_\xi : \xi < \alpha\right\}\right)$$

$$\theta \coloneqq \mu\alpha\left[A = \left\{a_\xi : \xi < \alpha\right\}\right]$$

Then $\langle a_\alpha : \alpha < \theta\rangle$ enumerates A. □

Lemma (König's Lemma)

Every finitely branching tree with infinitely many nodes contains an infinite path.

Weaker Consequences of AC

1. The union of a countable family of countable sets is countable.
2. For each property $P \in \{$Perfect Set Property, Lebesgue Measurable, Baire Property$\}$, there is a set without property P.
3. The Lebesgue measure of a countable disjoint union of measurable sets is equal to the sum of the measures of the individual sets. (σ-additivity)
4. Banach-Tarski Theorem.
5. Every field has a unique algebraic closure.
6. Every field extension has a transcendence basis.
7. Every subgroup of a free group is free.
8. Hahn-Banach Extension Theorem: Every bounded linear functional on a subspace of some vector space can be extended to the whole space.
9. The Baire Category Theorem.
10. On every infinite-dimensional topological vector space there is a discontinuous linear map.
11. Every Tychonoff space has a Stone-Čech compactification.

Consistence & Independence

Theorem (Gödel1938)

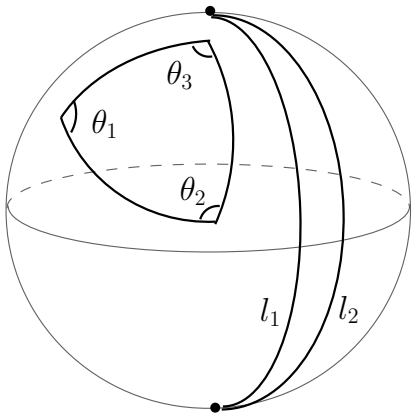
$$\mathrm{Con}(\mathrm{ZF}) \rightarrow \mathrm{Con}(\mathrm{ZFC} + \mathrm{GCH})$$

Theorem (Cohen1963)

- $\mathrm{Con}(\mathrm{ZF}) \rightarrow \mathrm{Con}(\mathrm{ZF} + \neg\mathrm{AC})$
- $\mathrm{Con}(\mathrm{ZF}) \rightarrow \mathrm{Con}(\mathrm{ZFC} + \neg\mathrm{GCH})$

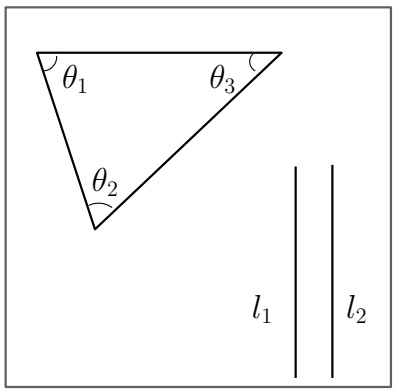


Figure: Cohen



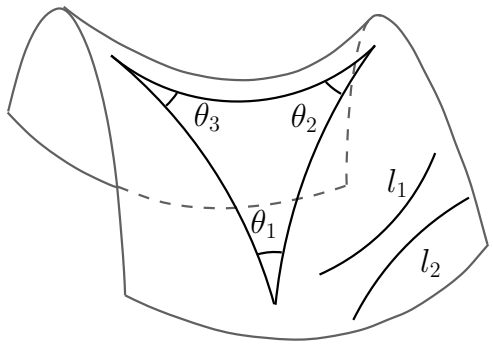
$\theta_1 + \theta_2 + \theta_3 > 180°$

Spherical (positive curvature)



$\theta_1 + \theta_2 + \theta_3 = 180°$

Euclidean (zero curvature)



$\theta_1 + \theta_2 + \theta_3 < 180°$

Hyperbolic (negative curvature)

GCH vs Weak GCH

$$2^{\aleph_\alpha} = \aleph_{\alpha+1} \qquad \text{(GCH)}$$

$$\Downarrow$$

$$2^{\kappa} < 2^{\kappa^+} \qquad \text{(WGCH)}$$

$$\Updownarrow$$

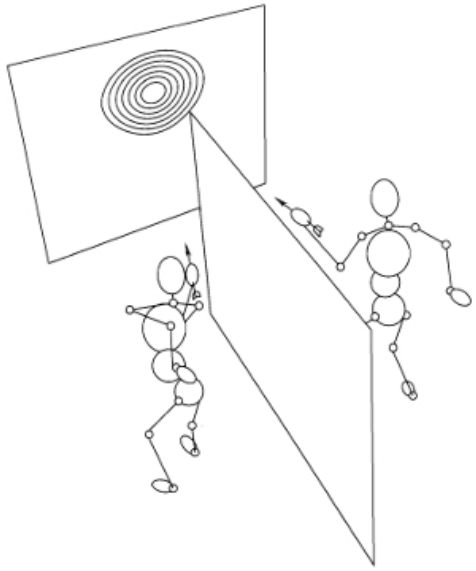
$$\kappa < \lambda \implies 2^{\kappa} < 2^{\lambda}$$

"$|X| < |Y| \implies |\mathrm{P}(X)| < |\mathrm{P}(Y)|$" is independent of ZFC.

Freiling's Axiom of Symmetry

Freiling's Axiom of Symmetry (AX)

$$\forall f : \mathbb{R} \to \mathbb{R}_{\aleph_0} \exists xy \left[x \notin f(y) \wedge y \notin f(x) \right]$$



Theorem

$$\mathrm{ZFC} \vdash \mathrm{AX} \leftrightarrow \neg \mathrm{CH}$$

Proof.

$(\to)$: Let $\prec$ be a well ordering of $\mathbb{R}$ of length $\aleph_1$. Let $f(x) := \{y : y \preceq x\}$. Then $f : \mathbb{R} \to \mathbb{R}_{\aleph_0}$. So $\exists xy\big(x \prec y \wedge y \prec x\big)$. Contradiction.

$(\leftarrow)$: Assume $2^{\aleph_0} > \aleph_1$. Let $x_1, x_2, \ldots$ be an $\aleph_1$-sequence of distinct reals.

Let $f : \mathbb{R} \to \mathbb{R}_{\aleph_0}$. Then $\left| \bigcup_{\alpha < \aleph_1} f(x_\alpha) \right| = \aleph_1$. So $\exists y \in \mathbb{R} \forall \alpha < \aleph_1 \big(y \notin f(x_\alpha)\big)$.

Since $f(y)$ is countable, $\exists \alpha \big(x_\alpha \notin f(y)\big)$. Therefore $y \notin f(x_\alpha) \wedge x_\alpha \notin f(y)$.

Philosophy

- The multiverse view might lead one to consider how the various models of set theory interact, whereas the universe view might lead one to try to discover fundamental features of the one true set-theoretic universe.
- These are quite different mathematical projects, and so one's philosophy of set theory will direct one to different mathematical efforts.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Readings

1. N. Cutland: Computability
2. P. Odifreddi: Classical Recursion Theory
3. R. I. Soare: Turing Computability
4. A. Nies: Computability and Randomness
5. M. Li, P. Vitányi: An Introduction to Kolmogorov Complexity and Its Applications

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Turing Machine
Computability
Diagonal Method
Incompressibility Method
Incompleteness

Equational Logic

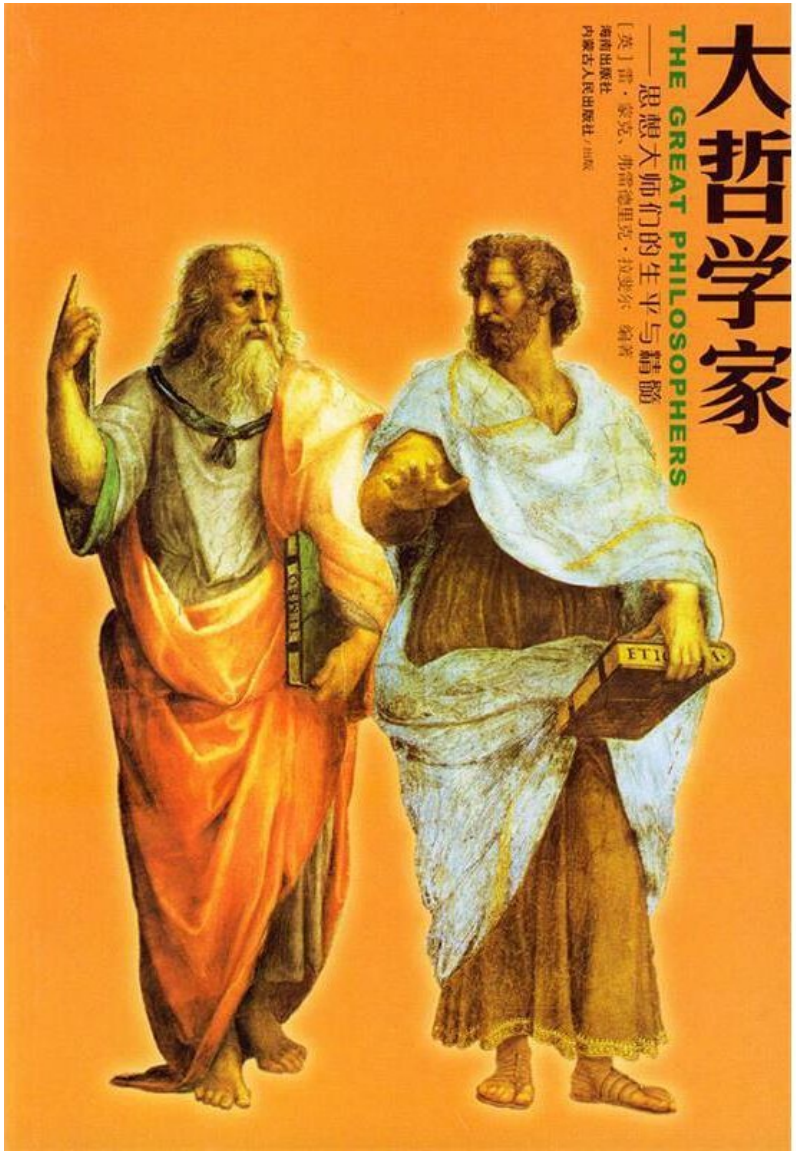
Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039



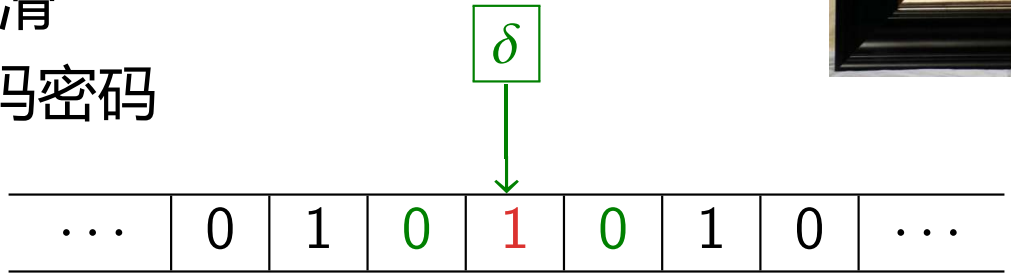
目　　录

导言
苏格拉底：哲学的殉道者
柏拉图：哲学的创始者
笛卡儿：我思故我在
斯宾诺莎：寻求真理与精神幸福
贝克莱：经验论哲学
大卫·休谟：道德科学的牛顿
马克思和自由：发展实践哲学
罗素：毕达哥拉斯之梦
海德格尔：存在与时间的历史和真理
维特根斯坦：论人类本性
波普尔：历史主义及其贫困
阿兰·图灵：一个自然哲学家

图灵 Alan Turing 1912-1954

- 图灵机/通用图灵机
- 丘奇-图灵论题
- 停机问题
- 不可判定性
- 神谕图灵机
- 可计算的绝对正规数
- 图灵测试、学习机
- 形态发生学 — 图灵斑图
- 古德-图灵平滑
- 破译恩尼格玛密码





What is “effective procedure”? — Recursion Theory

- 什么是计算?
- 人是怎么进行计算的?
- 有没有可能建造一台计算机器，机械地模拟人脑的计算过程?
- 机器的计算极限是什么?



图灵可计算 — 一个概念分析的典范[26]

机械可计算 $\Longleftrightarrow$ 图灵机可计算

图灵对丘奇图灵论题的论证策略

机械地可计算的 $\Longrightarrow$ 原则上人能计算的

$\Longrightarrow$ 图灵机可计算的

$\Longrightarrow$ 机械地可计算的

[26] Turing: On computable numbers, with an application to the Entscheidungsproblem. 1936.

“图灵可计算”的概念分析

- 想象一个理想的计算者，把她的操作拆解为基本的“简单操作”.
- 计算者进行的计算一般是在不限量的草稿纸上进行的符号书写.

$$
\begin{array}{cccccc}
 & & 4 & 2 & 3 & 1 \\
 & & & \times & 7 & 7 \\
\hline
 & 2 & 9 & 6 & 1 & 7 \\
2 & 9 & 6 & 1 & 7 & 0 \\
\hline
3 & 2 & 5 & 7 & 8 & 7
\end{array}
$$

- 不限量的草稿纸可以表示为一条画格子的无穷延伸的纸带.

4	2	3	1	×	7	7	=	2	9	6	1	7	+	2	9	6	1	7	0	=

“图灵可计算”的概念分析

- 计算者的符号的数量是有穷(等价于两个) 的.
  - 一个符号是 $[0,1] \times [0,1]$ 的一个勒贝格可测的子集
  - 符号间的距离被定义为两个符号对称差的测度
  - 由此, 上述符号构成一个紧致的度量空间
  - 因此不存在两两不交的无穷邻域集
  - 无论计算者的识别精度有多高, 都只能识别有穷个符号
- 计算者每个时刻只能注意到 (有穷)一个符号.
- 计算者的思想状态的数量是有穷的. (哥德尔表示怀疑)
  - 计算者总是可以暂停计算后再继续进行, 思想状态说明如何继续
- 计算者每个时刻的操作完全取决于其注意到的纸带上的符号, 以及当时的思想状态.
- 计算者能做的操作: 改变纸带上的一个符号、改变注意的格子、改变思想状态.

(Deterministic) Turing Machine

Definition ((Deterministic) Turing Machine)

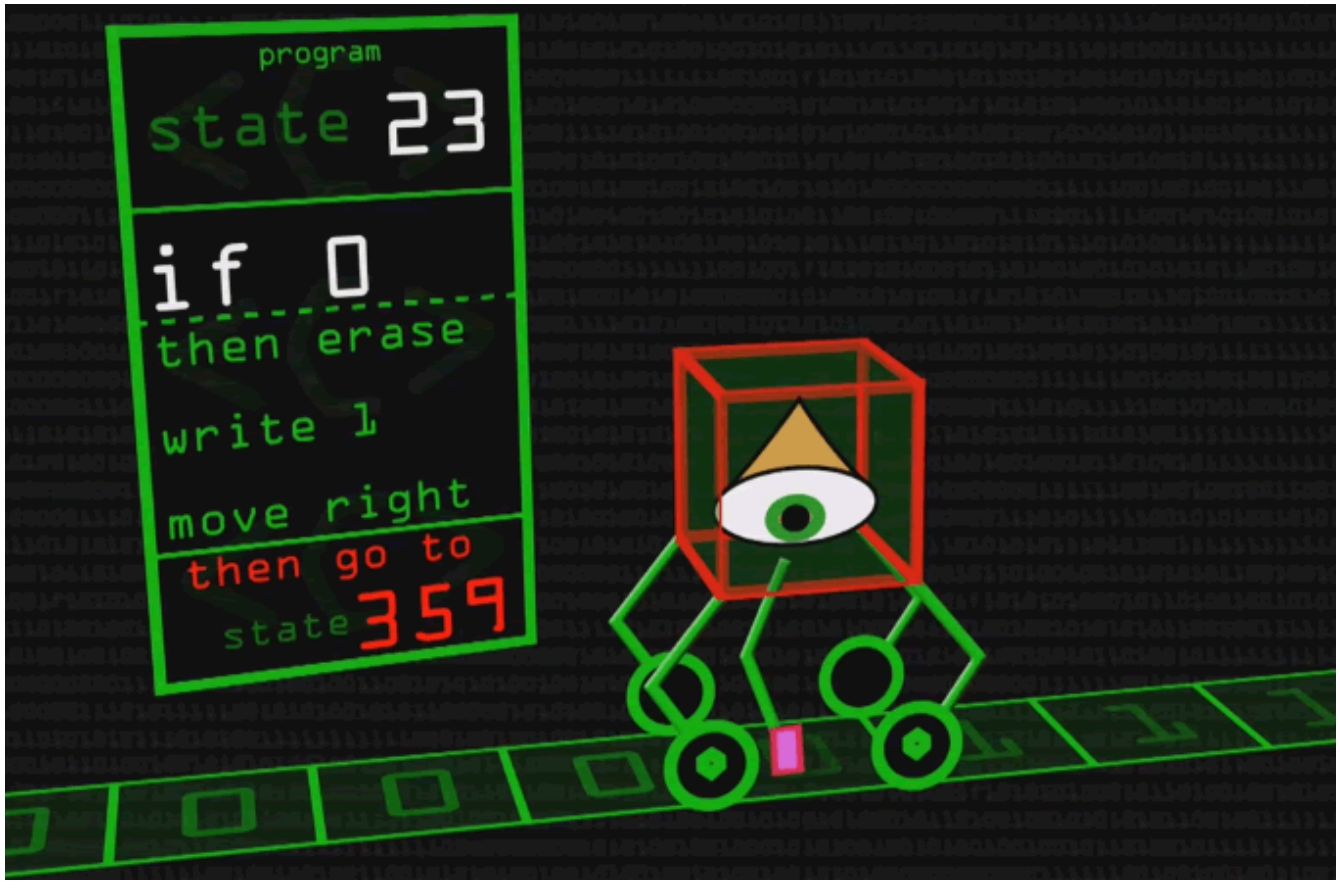
A deterministic Turing machine is a triplet (Σ, Q, δ), where Σ is a finite alphabet with an identified blank symbol, Q is a finite set of states with identified initial state q_0 and final state $q_f \neq q_0$, and δ, a deterministic transition function

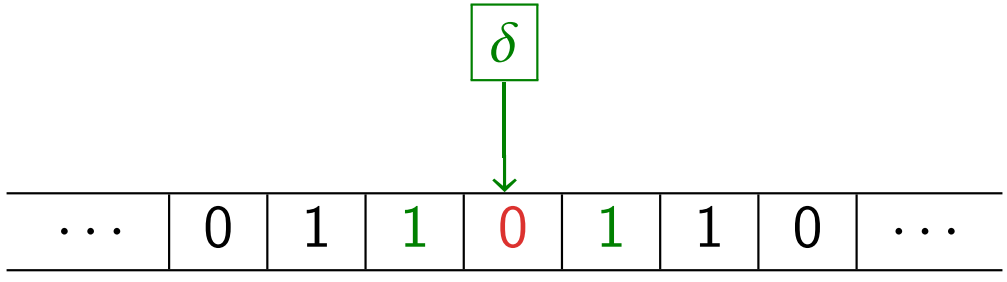
$$\delta : Q \times \Sigma \to \Sigma \times \{L, R\} \times Q$$

Here $\{L, R\}$ denote left and right, directions to move on the tape.

Definition (Configuration)

A configuration of a Turing Machine is a tuple (d, h, q) where d is a description of the contents of the tape, h is the location of the head symbol, and q represents the state the Turing machine is in.

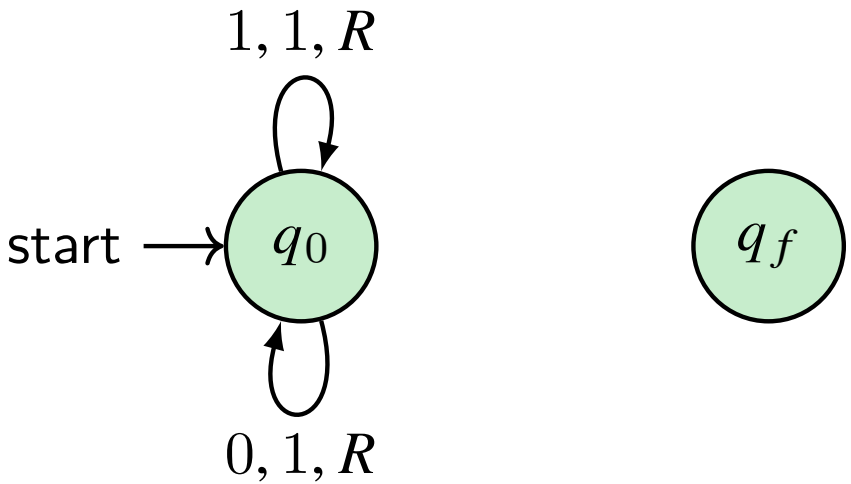




$$\delta(q_{23}, 0) = (1, R, q_{359})$$

Turing Machine — Example

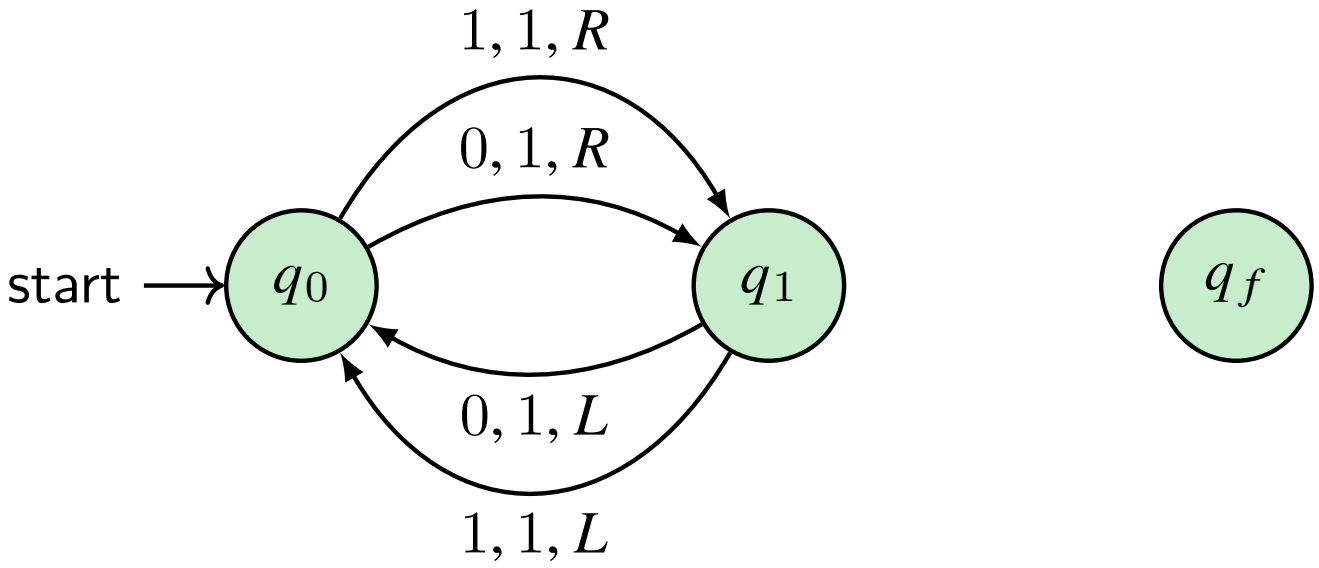
写入 1, 然后一直向右移动. 永不停机.



$$\Sigma = \{0, 1\}$$
$$Q = (q_0, q_f)$$
$$\delta(q_0, 0) = (1, R, q_0)$$
$$\delta(q_0, 1) = (1, R, q_0)$$

Turing Machine — Example

0 改为 1, 然后一直左右移动. 永不停机.



$$\Sigma = \{0, 1\}$$

$$Q = (q_0, q_1, q_f)$$

$$\delta(q_0, 0) = (1, R, q_1)$$

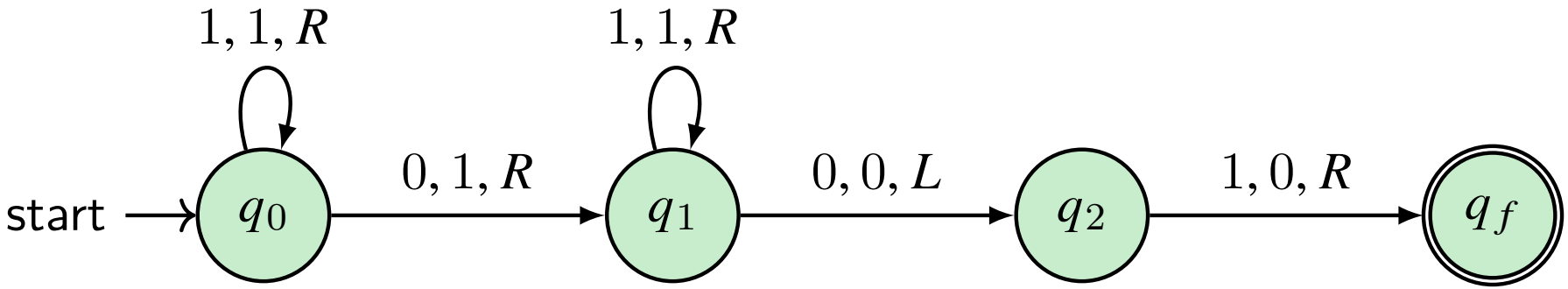
$$\delta(q_1, 0) = (1, L, q_0)$$

$$\delta(q_0, 1) = (1, R, q_1)$$

$$\delta(q_1, 1) = (1, L, q_0)$$

Turing Machine — Example

将两个被 0 隔开的一进制自然数 $(1^m 0 1^n)$ 相加 (1^{m+n}).



$$\Sigma = \{0, 1\}$$

$$Q = (q_0, q_1, q_2, q_f)$$

$$\delta(q_0, 1) = (1, R, q_0)$$

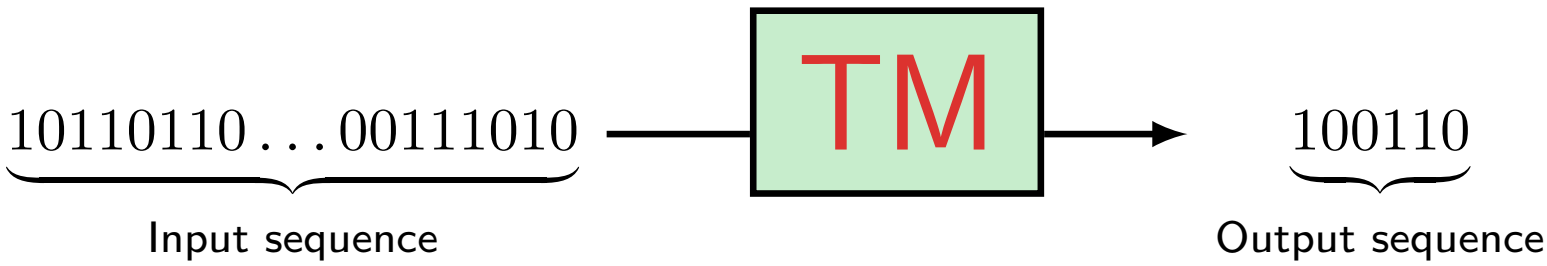
$$\delta(q_0, 0) = (1, R, q_1)$$

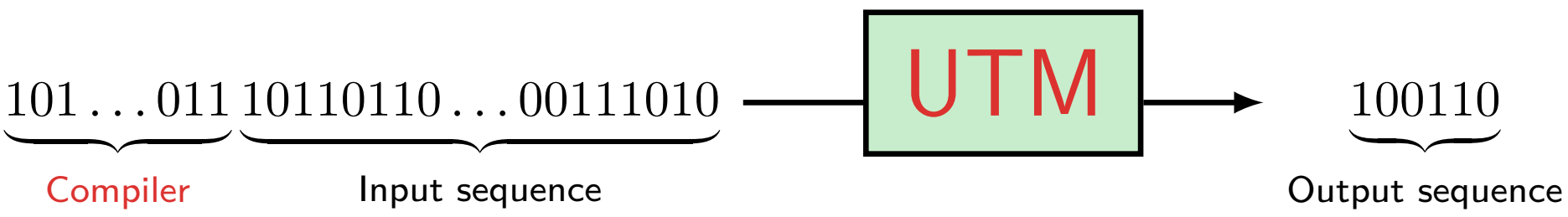
$$\delta(q_1, 1) = (1, R, q_1)$$

$$\delta(q_1, 0) = (0, L, q_2)$$

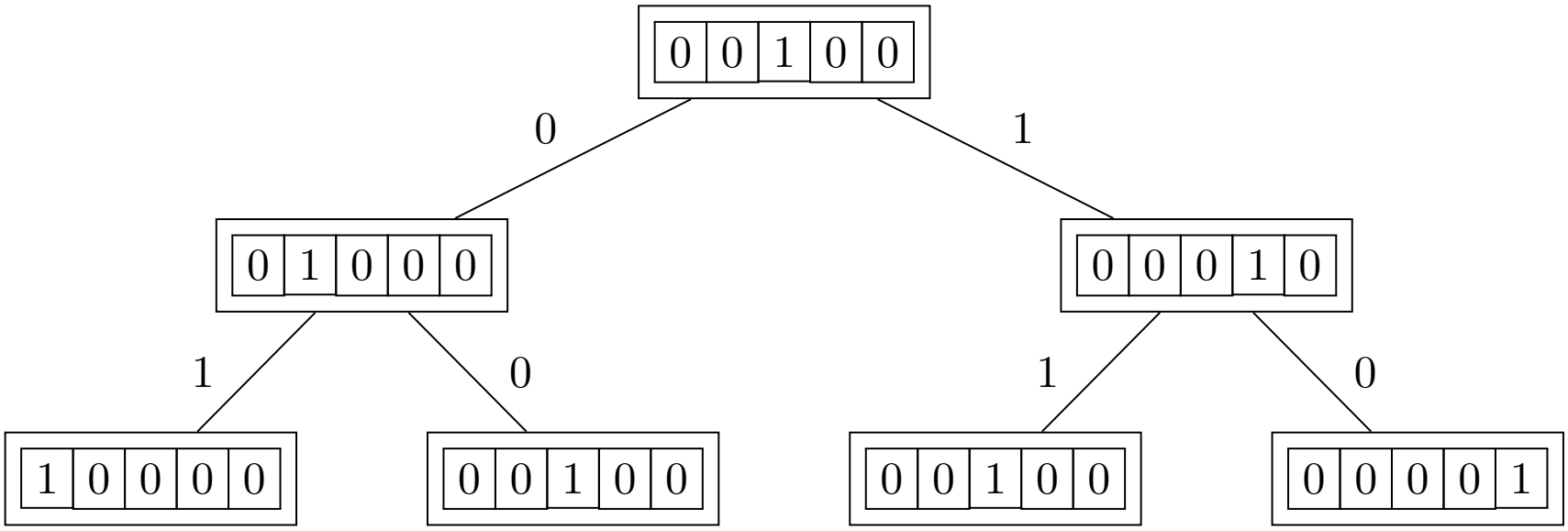
$$\delta(q_2, 1) = (0, R, q_f)$$

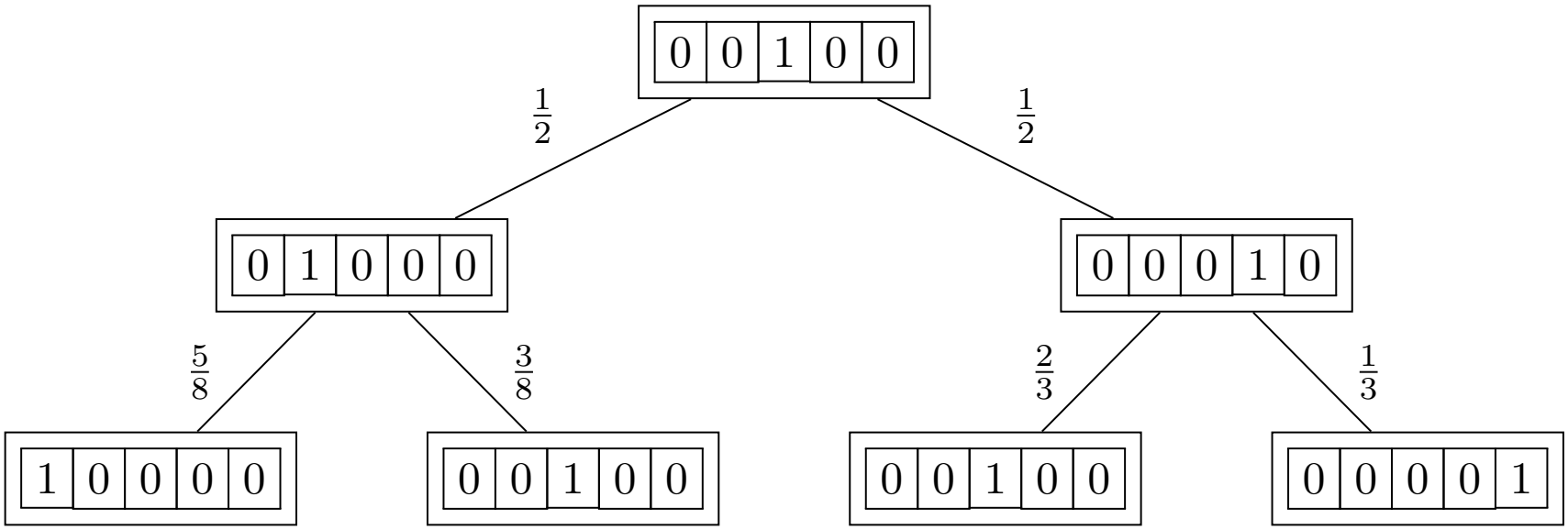
通用图灵机 Universal Turing Machine



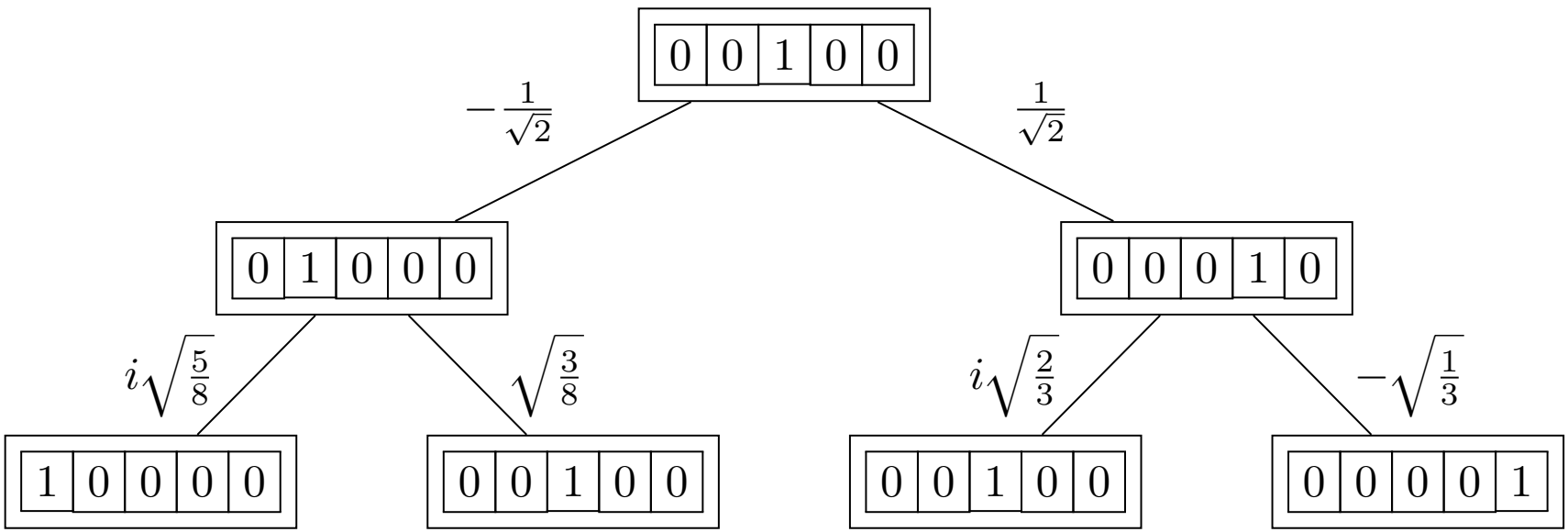


(Deterministic / Probabilistic) Turing Machine





(Quantum) Turing Machine



Church-Turing Thesis

- 1931-1934, Herbrand-Gödel: "广义递归函数"
- 1933-1935, Church: λ-可定义函数
  — Kleene 1935 证明了 λ-可定义函数与 "广义递归函数" 的等价性, 但哥德尔依然不认为它强到了足以涵盖所有能行可计算函数.

  "I was completely convinced only by Turing's paper."

  — *Kurt Gödel*
- 1936, Turing: 图灵机
- 1936, Post: 波斯特机
- 1956, Chomsky: 0-型文法 type-0 grammar
- 1970, Conway: 生命游戏

"With this concept (Turing Computability) one has for the first time succeeded in giving an absolute ***definition*** *of an interesting epistemological notion, i.e., one not depending on the formalism chosen."*

— *Kurt Gödel*

The Thesis as a Definition

- Cauchy-Weierstrass Thesis: a function is intuitively continuous iff

$$\forall x \in I \forall \varepsilon > 0 \exists \delta > 0 \forall y \in I \big(|x - y| < \delta \rightarrow |f(x) - f(y)| < \varepsilon \big)$$

- Church-Turing Thesis:

effective calculable = Turing computable

- "Intelligence Thesis"?
- "Life Thesis"?
- "Consciousness Thesis"?
- "Free Will Thesis"?
- "Beauty Thesis"?
- "Knowledge/Understanding/Meaning/Love Thesis..."?

Thesis (Church-Turing Thesis)

effective calculable = *recursive* = *Turing Computable*

||

representable in Q = λ-*definable*

||

finite definable = *Herbrand-Gödel computable*

||

flowchart (or 'while') computable

||

Neural Network with unbounded tape = *Conway's 'game of life'*

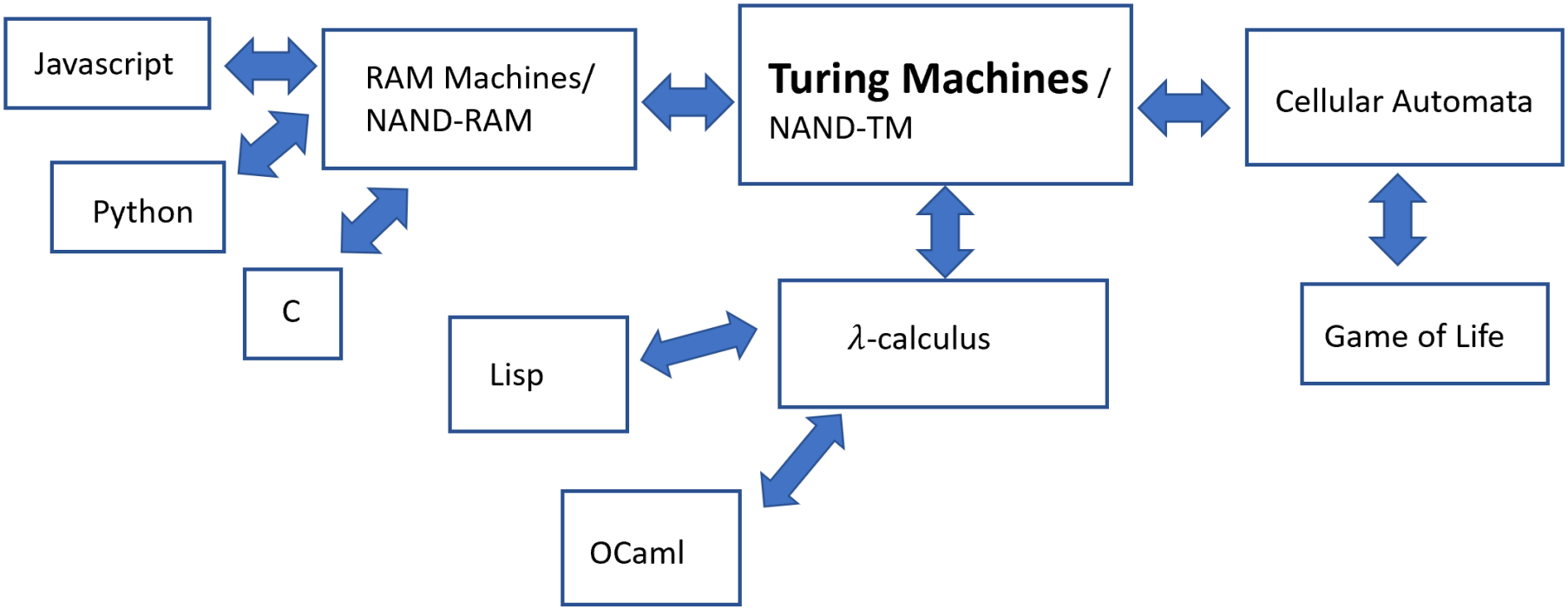
||

Adleman's DNA Computing

||

Post/Markov/McCarthy/Kolmogorov-Uspensky computable ...

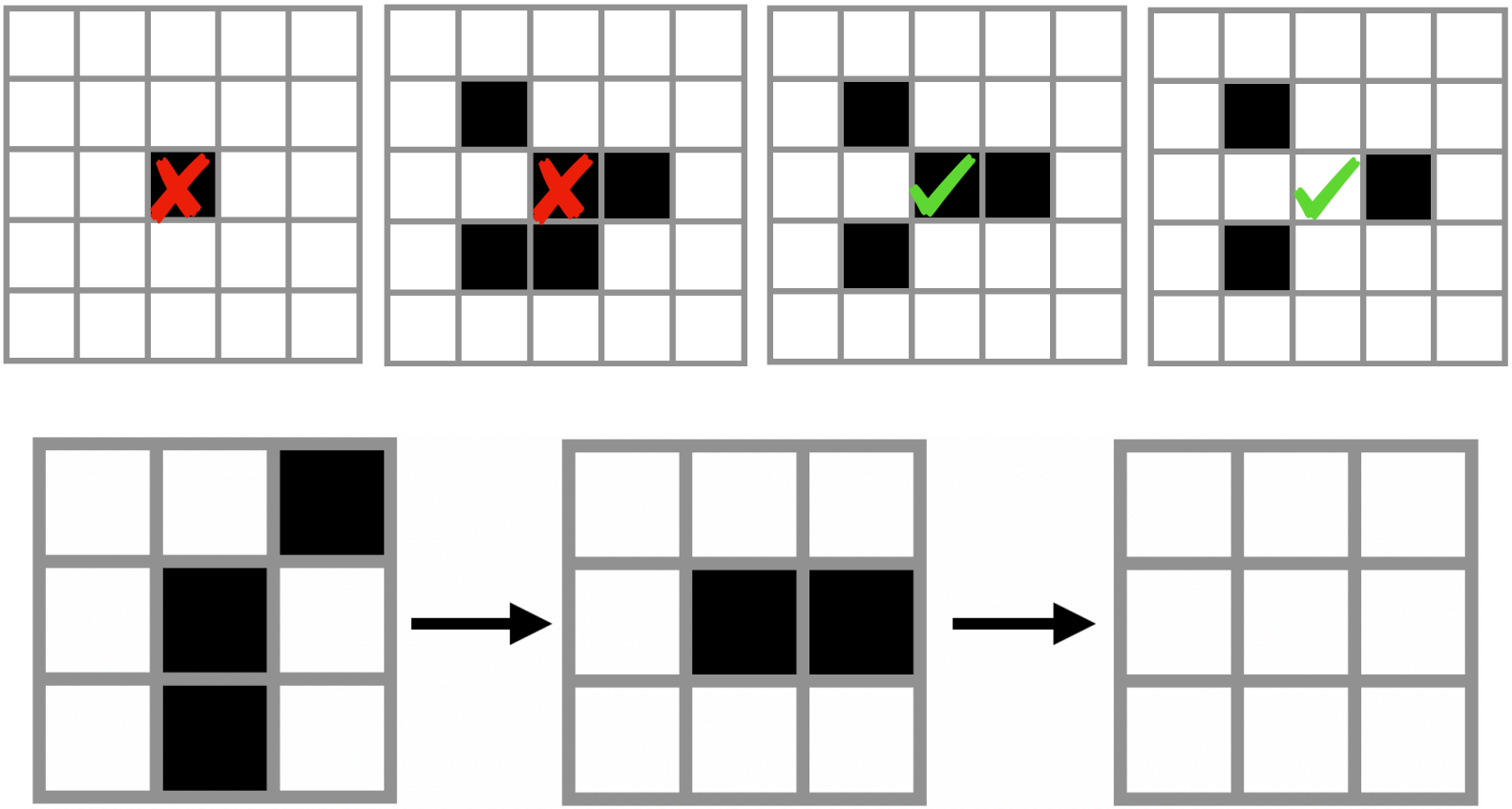
- Any possible discrete physical process is computable?
- Any constructive function is computable?
- The mental functions can be simulated by machines?



- Every finite function $f : \{0,1\}^n \to \{0,1\}^m$ is computable by a Boolean circuit with $O(m2^n/n)$ gates.
- To compute functions with unbounded inputs $f : \{0,1\}^* \to \{0,1\}^*$, we need a collection of circuits: one for every input length.
- Turing machines capture the notion of a single algorithm that can compute functions of all input lengths.

NAND-TM = NAND-CIRC + loops + arrays

Conway's Game of Life



1. A live cell with < 2 neighbors dies of isolation.
2. A live cell with > 3 neighbors dies of overcrowding.
3. A live cell with 2 or 3 neighbors survives.
4. A dead cell with 3 neighbors will come to life.

Church-Turing Thesis

- Church-Turing Thesis

  Every "function which could be regarded as computable" can be computed by a universal Turing machine.

- Church-Turing-Deutsch Thesis

  Every finite physical system can be simulated to any specified degree of accuracy by a universal Turing machine.

- Feasibility Thesis — Classical / Quantum Version

  A probabilistic (quantum) Turing machine can efficiently simulate any realistic model of computation.

- Wolfram's Principle of Computational Equivalence

  Almost all processes that are not obviously simple can be viewed as computations of equivalent sophistication.

- Wolfram's Principle of Computational Irreducibility

  Most of the time, the only way to see what a physical system (computer program) will do is to run it.

丘奇-图灵论题

一沙一世界,
一花一天国,
无限掌中置,
刹那含永劫.
— 布莱克



- 通用图灵机可以模拟任何图灵机.
- 通用图灵机可以模拟整个宇宙.
- 任何图灵完备的装置都包含了宇宙的所有规律.
- “宇宙最不可理解之处是它是可理解的.”
- 描述复杂性、生成复杂性、组织复杂性……
  — 科尔莫哥洛夫复杂度、逻辑深度……

Free Will as Computational Irreducibility?

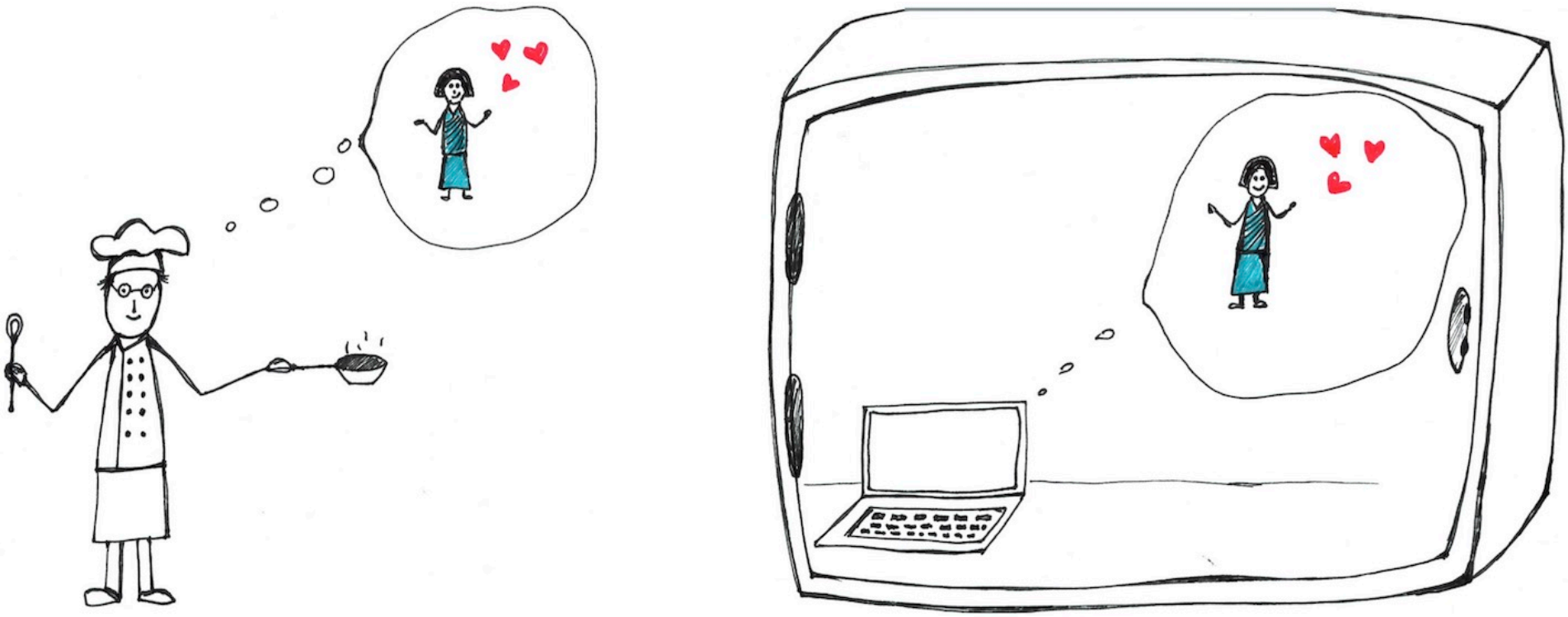


Figure: To predict John's choice of breakfast by simulation. A universal computer in the safe (on the right) reproduces the outputs of another process, i.e. its observable actions (John preparing breakfast, on the left). The agent is the source of its decisions?

Libet: We are conscious of our free choices only after 300ms our brain has made them.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Turing Machine
Computability
Diagonal Method
Incompressibility Method
Incompleteness

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises
References2039

Presburger/Robinson/Peano Arithmetic

- $x + 0 = x$
- $x + y = y + x$
- $(x + y) + z = x + (y + z)$
- $x + z = y + z \rightarrow x = y$
- $x \cdot 0 = 0$
- $x \cdot 1 = x$
- $x \cdot y = y \cdot x$
- $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
- $x \cdot (y + z) = x \cdot y + x \cdot z$
- $P(0) \wedge \forall x(P(x) \rightarrow P(x + 1)) \rightarrow \forall x P(x)$

1. $s(x) \neq 0$
2. $s(x) = s(y) \rightarrow x = y$
3. $y = 0 \vee \exists x(s(x) = y)$
4. $x + 0 = x$
5. $x + s(y) = s(x + y)$
6. $x \cdot 0 = 0$
7. $x \cdot s(y) = (x \cdot y) + x$

} Q

8. $P(0) \wedge \forall x(P(x) \rightarrow P(s(x))) \rightarrow \forall x P(x)$

Full Second Order Arithmetic Z_2

- The basic axioms of Peano arithmetic
- The second order induction axiom

$$\forall X\big[0 \in X \wedge \forall n(n \in X \to s(n) \in X) \to \forall n \in \mathbb{N}(n \in X)\big]$$

- The comprehension axiom

$$\exists X \forall n[n \in X \leftrightarrow A(n)]$$

where A is any formula in which X does not occur freely.

Dedekind-Peano Axioms PA^2

The natural numbers $\mathbb{N}$ is a set with a chosen element $0 \in \mathbb{N}$ and an injective function $s : \mathbb{N} \to \mathbb{N}$ suth that $0 \notin s(\mathbb{N})$ and suth that the principle of mathematical induction holds: a set containing 0 and closed under s contains all natural numbers.

Definition (Dedekind-Peano Axioms)

1. $0 \in \mathbb{N}$
2. $\forall n \in \mathbb{N} : s(n) \in \mathbb{N}$
3. $\forall n \in \mathbb{N} : s(n) \neq 0$
4. $\forall mn \in \mathbb{N} : s(m) = s(n) \to m = n$
5. $\forall X \big[0 \in X \wedge \forall n(n \in X \to s(n) \in X) \to \forall n \in \mathbb{N}(n \in X)\big]$

Theorem (Dedekind-Peano Categoricity Theorem)

All models of the Dedekind-Peano axioms are isomorphic.

Exponentiation is Definable in $\mathcal{N}$/Representable in $\mathbb{Q}$

$$\pi(x,y) \coloneqq (x+y)^2 + x + 1$$
$$\pi_1(z) \coloneqq \mu x[\exists y \leq z : \pi(x,y) = z]$$
$$\pi_2(z) \coloneqq \mu y[\exists x \leq z : \pi(x,y) = z]$$
$$\beta(s,i) \coloneqq \mu x < s\Big[\pi_1(s) \equiv x \pmod{1 + (i+1)\cdot\pi_2(s)}\Big]$$
$$\beta^*(s,i) \coloneqq \mu x < s\,[\exists y < s\,\exists z < s : s = \pi(y,z) \ \wedge\ (1 + (\pi(x,i)+1)\cdot z) \mid y]$$

Lemma

For every sequence $a_0, \ldots, a_n \in \mathbb{N}$, $\exists s \in \mathbb{N}.\forall i \leq n.\beta(s,i) = a_i$.

Proof.

$$b \coloneqq \max\{n, a_0, \ldots, a_n\} \quad d \coloneqq b! \quad m_i \coloneqq 1 + (i+1)\cdot d$$

$m_0, \ldots, m_n$ are pairwise coprime.

$c \coloneqq \mu x\,[\forall i \leq n : x \equiv a_i \pmod{m_i}]$ by Chinese Remainder Theorem.

$$s \coloneqq \pi(c,d)$$
$$x^y \coloneqq \beta\Big(\mu s\Big[\beta(s,0) = 1 \ \wedge\ \forall i < y : \beta(s,i+1) = \beta(s,i)\cdot x\Big], y\Big)$$

□

Chinese Remainder Theorem

Theorem (Chinese Remainder Theorem)

Suppose $m_0, \ldots, m_n$ are pairwise coprime. Let $a_0, \ldots, a_n$ be arbitrary integers. Then there is $x \in \mathbb{Z}$ s.t. for $i \leq n$:

$$x \equiv a_i \pmod{m_i} \qquad \text{(Chinese Remainder Theorem)}$$

Proof.

$$M \coloneqq \prod_{i=1}^{n} m_i \qquad M_i \coloneqq \frac{M}{m_i}$$

$$n_i \coloneqq \mu x \left[x \cdot M_i \equiv 1 \pmod{m_i} \right]$$

$$x \equiv \sum_{i=1}^{n} n_i \cdot M_i \cdot a_i \pmod{m}$$

□

Gödel Numbering

Definition (Gödel Numbering)

A Gödel numbering is a mapping from a set of expressions to $\mathbb{N}$ s.t.,

1. Different expressions receive different Gödel numbers. (injective)
2. The Gödel number of an expression can be effectively calculated. (computable)
3. It is effectively decidable whether a given number is a Gödel number or not. (its inverse function is computable)

ζ	$($	$)$	$,$	$\neg$	$\rightarrow$	$\forall$	x_k	a_k	f_k^n	P_k^n
$\#\zeta$	3	5	7	9	11	13	$7+8k$	$9+8k$	$11+8(2^n3^k)$	$13+8(2^n3^k)$

$$\langle a_1, \ldots, a_n \rangle := \mu x \left[\beta(x, 0) = n \wedge \beta(x, 1) = a_1 \wedge \cdots \wedge \beta(x, n) = a_n \right]$$

$$or \quad \langle a_1, \ldots, a_n \rangle := \prod_{i=1}^{n} p_i^{a_i+1}$$

Arithmetization of Syntax — Gödel Numbering

$$\#(\zeta_0 \cdots \zeta_n) := \langle \#\zeta_0, \ldots, \#\zeta_n \rangle$$

$$\ulcorner \zeta \urcorner := \underline{\#\zeta} = s^{\#\zeta} 0$$

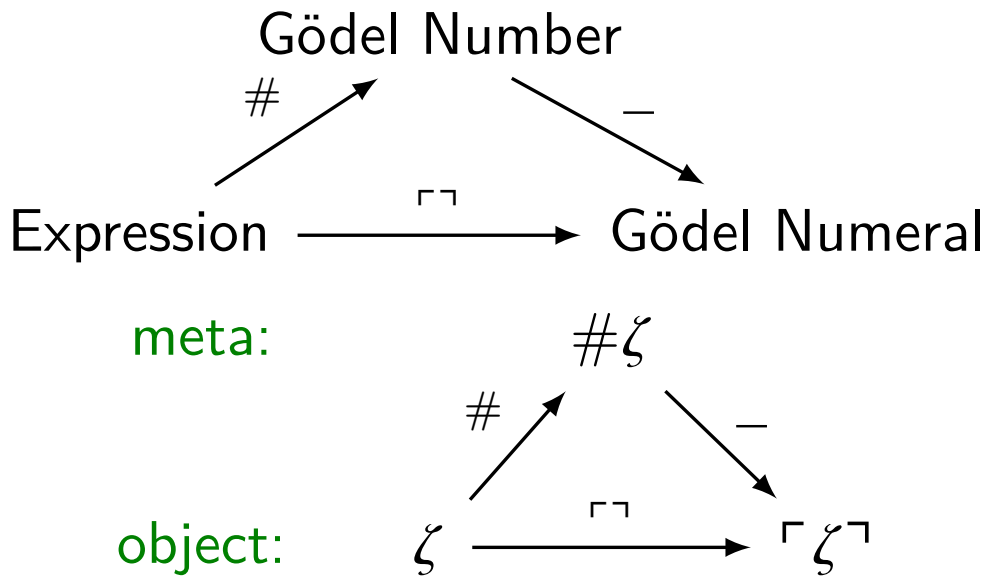
$\mathrm{prf}_{\mathrm{T}}(y, x) :=$ "y is the code of a proof in T of the formula with code x."

$$\mathrm{prov}_{\mathrm{T}}(x) := \exists y \, \mathrm{prf}_{\mathrm{T}}(y, x)$$

$$\Box_{\mathrm{T}} A := \mathrm{prov}_{\mathrm{T}}(\ulcorner A \urcorner)$$

Is there a formula A s.t. $\mathrm{T} \vdash A \leftrightarrow \neg \Box_{\mathrm{T}} A$?





- $\mathrm{Con}_{\mathrm{T}} :\equiv \neg\Box_{\mathrm{T}}\bot$
- ω-consistent: $\forall x \Box_{\mathrm{T}} \neg A(\dot{x}) \rightarrow \neg\Box_{\mathrm{T}}\exists x A(x)$ for any formula A
- 1-consistent: $\forall x \Box_{\mathrm{T}} \neg A(\dot{x}) \rightarrow \neg\Box_{\mathrm{T}}\exists x A(x)$ for $A \in \Delta_0$
- $\mathrm{Rfn}_{\Gamma}(\mathrm{T}) : \Box_{\mathrm{T}} A \rightarrow A$ for any sentence $A \in \Gamma$
- $\mathrm{RFN}_{\Gamma}(\mathrm{T}) : \forall x \big(\Box_{\mathrm{T}} A(\dot{x}) \rightarrow A(x)\big)$ for any formula $A \in \Gamma$
- arithmetically sound: $\mathrm{Rfn}_{\Sigma_{<\omega}}(\mathrm{T})$
- 1-consistent $\iff \mathrm{Rfn}_{\Sigma_1}(\mathrm{T})$
- $\mathrm{Rfn}_{\Pi_1}(\mathrm{T}) \iff \mathrm{RFN}_{\Pi_1}(\mathrm{T}) \iff \mathrm{Con}_{\mathrm{T}}$

Definition (Gödelian Theory)

A theory is Gödelian iff it is

1. consistent
2. axiomatizable
3. rich enough to represent elementary arithmetic (able to represent primitive recursive functions)

Primitive Recursive Function & Recursive Function

- initial functions:
  1. projection: $I_i^m(n_1, \ldots, n_m) = n_i$ for $1 \leq i \leq m$
  2. successor: $s(n) = n + 1$
  3. zero: $z(n) = 0$
- composition: given $g, h_1, \ldots, h_k$,

$$f(\mathbf{x}) = g\left(h_1(\mathbf{x}), \ldots, h_k(\mathbf{x})\right)$$

- primitive recursion: given g, h,

$$f(0, \mathbf{x}) = g(\mathbf{x})$$
$$f(n+1, \mathbf{x}) = h\left(n, f(n, \mathbf{x}), \mathbf{x}\right)$$

- regular μ-operation: given g, and $\forall \mathbf{x} \exists y [g(\mathbf{x}, y) = 0]$,

$$f(\mathbf{x}) = \mu y [g(\mathbf{x}, y) = 0]$$

Partial Recursive Function

- μ-operation: given g,

$$f(\mathbf{x}) = \mu y \left[\forall z \leq y\big(g(\mathbf{x}, z)\downarrow\big) \ \wedge \ g(\mathbf{x}, y) = 0\right]$$

- bounded μ-operation:

$$\mu x < n[A(x)] := \mu x[A(x) \vee x = n]$$

Definition (Primitive Recursive / Recursive / Partial Recursive)

The class of primitive recursive functions (recursive functions, partial recursive functions) is the smallest class of functions containing the initial functions and closed under composition, primitive recursion (regular μ-operation, μ-operation).

Theorem

The primitive recursive functions are closed under bounded μ-operation.

Ackermann Function

Definition (Ackermann Function)

$$A(m,n) := \begin{cases} n+1 & \text{if } m = 0 \\ A(m-1,1) & \text{if } m > 0 \text{ and } n = 0 \\ A(m-1,A(m,n-1)) & \text{if } m > 0 \text{ and } n > 0 \end{cases}$$

Theorem

The Ackermann function is recursive but not primitive recursive.

$$a \uparrow^n b := \begin{cases} ab & \text{if } n = 0 \\ (a \uparrow^{n-1})^b 1 & \text{if } n \geq 1 \end{cases}$$

$$a \uparrow b = a^b \qquad a \uparrow\uparrow b = \underbrace{a \uparrow (a \uparrow (\cdots \uparrow a))}_{b \text{ copies of } a} 1 = \underbrace{a^{a^{\cdot^{\cdot^{\cdot^{a}}}}}}_{b \text{ copies of } a}$$

$$A(m,n) = 2 \uparrow^{m-2} (n+3) - 3$$

The Representability Theorem

A function/relation is representable in Robinson Q iff it is computable.

(proof sketch.) We have to show that all initial functions are representable, and the representable functions are closed under composition, regular μ-operation and primitive recursion.

Remark: To show that the set of representable functions is closed under primitive recursion, we can't use x^y and p_n, since they are defined by primitive recursion.

$$f(0, \mathbf{x}) = g(\mathbf{x})$$
$$f(n+1, \mathbf{x}) = h(n, f(n, \mathbf{x}), \mathbf{x})$$

we can code the sequence of values of f from 0 to y by Gödel's β function:

$$F(\mathbf{x}, y) = \mu z \big[\beta(z, 0) = g(\mathbf{x}) \ \wedge \ \forall i < y : \beta(z, i+1) = h\big(i, \beta(z, i), \mathbf{x}\big)\big]$$

$$f(y, \mathbf{x}) = \beta\big(F(\mathbf{x}, y), y\big)$$

Kleene Normal Form Theorem

Theorem (Kleene Normal Form Theorem)

There is a primitive recursive function U and primitive recursive predicate T, s.t. for every partial recursive function f, there is an index e s.t.

- $f(\mathbf{x})\downarrow \iff \exists y T(e, \mathbf{x}, y)$
- $f(\mathbf{x}) = U(\mu y T(e, \mathbf{x}, y))$

$T(e, \mathbf{x}, y) :=$ "y is the code number of some computation according to program P_e with input $\mathbf{x}$."
$U(y) :=$ "the number of 1's in the final configuration of y."

Definition

φ_e is the e^{th} partial recursive function[a]:

$$\varphi_e(\mathbf{x}) := U(\mu y T(e, \mathbf{x}, y))$$

[a] Sometimes we write $[\![e]\!]$ or $\{e\}$ for φ_e.

Incompleteness Theorem

- The function $\bar{f}$ is a *completion* of a partial function f iff $\bar{f}$ is total and $\forall n : f(n)\downarrow \implies f(n) = \bar{f}(n)$.
- A partial function f is *potentially recursive* iff it has a completion which is recursive.

Not every partial recursive function is potentially recursive.

$$f(n) \coloneqq \varphi_n(n) + 1$$

Theorem (Incompleteness Theorem)

Any ω-consistent Gödelian T is incomplete.

Proof.

Suppose T is represented in T by γ.

$$\bar{\varphi}_e(n) \coloneqq \begin{cases} U(\mu y T(e,n,y)) & \text{if } \exists y T(e,n,y) \\ 0 & \text{if } \mathrm{T} \vdash \forall y \neg\gamma(e,n,y) \end{cases}$$

Enumeration Theorem & s-m-n Theorem

Theorem (Enumeration Theorem)

The sequence $\{\varphi_e^n\}_{e\in\omega N}$ *is a partial recursive enumeration of the* n*-ary partial recursive functions, in the sense that:*

- *for each* e, φ_e^n *is a partial recursive function of* n *variables.*
- *if* ψ *is a partial recursive function of* n *variables, then there is* e *s.t.* $\psi = \varphi_e^n$.
- *there is a partial recursive function* φ *of* $n+1$ *variables s.t.* $\varphi(e, \mathbf{x}) = \varphi_e(\mathbf{x})$.

Theorem (s-m-n Theorem / Parameter Theorem)

For any $m, n > 0$, *there exists a primitive recursive function* s_n^m *of* $m+1$ *arguments s.t. for every Gödel number* e *of a partial recursive function with* $m+n$ *arguments*

$$\varphi_{s_n^m(e,x_1,\ldots,x_m)}(y_1 \ldots y_n) = \varphi_e(x_1, \ldots, x_m, y_1, \ldots, y_n)$$

Theorem (Characterization of the r.e. Sets)

The following are equivalent:

- A *is r.e.*
- *Its semicharacteristic function* $c_A(x) = \begin{cases} 1 & \text{if } x \in A \\ \uparrow & \text{otherwise} \end{cases}$ *is partial recursive.*
- A *is the domain of a partial recursive function.*
- A *is the range of a partial recursive function.*
- *There is a decidable relation* R *s.t.* $x \in A \iff \exists y : (x, y) \in R$.
- $A = \emptyset$ *or* A *is the range of a recursive function.*

Theorem (Characterization of the Recursive Sets)

The following are equivalent:

- A *is recursive.*
- $A = \emptyset$ *or* A *is the range of a nondecreasing recursive function.*
- *Both* A *and its complement* $\mathbb{N} \setminus A$ *are r.e.*

Acceptable Numbering

Definition (Acceptable Numbering)

A numbering ψ is acceptable iff there are recursive functions f, g s.t.

$$\psi_e = \varphi_{f(e)} \quad \text{and} \quad \varphi_e = \psi_{g(e)}$$

Theorem

A numbering is acceptable iff it satisfies both enumeration and s-m-n.

Theorem (Rogers' Equivalence Theorem)

ψ is an acceptable numbering iff there is a recursive permutation h s.t.

$$\psi_e = \varphi_{h(e)}$$

Theorem (Blum)

If ψ is an acceptable numbering, then there is a recursive permutation h s.t.

$$h(\psi_e(x)) = \varphi_{h(e)}(h(x))$$

Programming Languages

- A set Prog of programs.
- A set D of data on which programs operate.
- A pairing operation $\langle \cdot, \cdot \rangle : D^2 \to D$.
- A semantic function
$$[\![ ]\!] : \mathrm{Prog} \to [D \rightharpoonup D]$$
which maps programs to partial functions on data.

Assume there is a function $\ulcorner \urcorner : \mathrm{Prog} \to D$ which represents program texts as data items.

Universal Turing Machine

There is a program $\mathrm{eval} \in \mathrm{Prog}$ such that for all $P \in \mathrm{Prog}$ and $d \in D$
$$[\![\mathrm{eval}]\!](\langle \ulcorner P \urcorner, d \rangle) = [\![P]\!](d)$$

Theorem (s-m-n Theorem / Parameter Theorem)

There is a primitive recursive function $[\![\mathrm{spec}]\!]$ *s.t. for every Gödel number* e *of a partial recursive function*
$$[\![[\![\mathrm{spec}]\!](e, \mathbf{x})]\!](\mathbf{y}) = [\![e]\!](\mathbf{x}, \mathbf{y})$$

Futamura Projections

- The first Futamura projection.

$$\begin{aligned} \text{target} &:= [\![\text{spec}]\!](\text{int}, \text{source}) \\ \text{out} &= [\![\text{source}]\!](\text{in}) \\ &= [\![\text{int}]\!](\text{source}, \text{in}) && \text{(Definition of an interpreter)} \\ &= [\![[\![\text{spec}]\!](\text{int}, \text{source})]\!](\text{in}) && \text{(Definition of a specializer)} \\ &= [\![\text{target}]\!](\text{in}) \end{aligned}$$

- Compiler generation by the second Futamura projection.

$$\begin{aligned} \text{compiler} &:= [\![\text{spec}]\!](\text{spec}, \text{int}) \\ \text{target} &= [\![\text{spec}]\!](\text{int}, \text{source}) \\ &= [\![[\![\text{spec}]\!](\text{spec}, \text{int})]\!](\text{source}) \\ &= [\![\text{compiler}]\!](\text{source}) \end{aligned}$$

- Compiler generator generation by the third Futamura projection.

$$\begin{aligned} \text{cogen} &:= [\![\text{spec}]\!](\text{spec}, \text{spec}) \\ \text{compiler} &= [\![\text{spec}]\!](\text{spec}, \text{int}) \\ &= [\![[\![\text{spec}]\!](\text{spec}, \text{spec})]\!](\text{int}) \\ &= [\![\text{cogen}]\!](\text{int}) \end{aligned}$$

Futamura Projections

$$\begin{array}{lcccl} \text{out} & = & [\![\text{int}]\!](\text{source}, \text{in}) & = & [\![\text{target}]\!](\text{in}) \\ \text{target} & = & [\![\text{spec}]\!](\text{int}, \text{source}) & = & [\![\text{compiler}]\!](\text{source}) \\ \text{compiler} & = & [\![\text{spec}]\!](\text{spec}, \text{int}) & = & [\![\text{cogen}]\!](\text{int}) \\ \text{cogen} & = & [\![\text{spec}]\!](\text{spec}, \text{spec}) & = & [\![\text{cogen}]\!](\text{spec}) \end{array}$$

- A program int is an interpreter iff for program p and data d:
$$[\![\text{int}]\!](p, d) = [\![p]\!](d)$$
- A program compiler is a compiler iff for program p and data d:
$$[\![[\![\text{compiler}]\!](p)]\!](d) = [\![p]\!](d)$$
- A program spec is a specializer iff for program p and data x, y:
$$[\![[\![\text{spec}]\!](p, x)]\!](y) = [\![p]\!](x, y)$$

Interpreter vs Compiler

- $\text{int} = \lambda p.\lambda d.\, \text{compiler}(p)(d)$
- $\text{compiler} = \text{cogen}(\text{int})$

Remark: Self-printing programs and self-generating compilers generators are two disjoint program classes.

Kleene's Fixpoint Theorem

Theorem (Kleene's Fixpoint Theorem)

Given a recursive function h, there is an index e s.t.

$$\varphi_e = \varphi_{h(e)}$$

Corollary (Second Recursion Theorem)

If $f(x, y)$ is a partial recursive function, there is an index e s.t.

$$\varphi_e(y) = f(e, y)$$

Proof.

By the s-m-n theorem, $\varphi_{s(x)}(y) = f(x, y)$. Then

$$\exists e : \varphi_e(y) = \varphi_{s(e)}(y) = f(e, y)$$

□

Theorem (Kleene's Relativized Fixpoint Theorem (with Parameters))

Let $A \subset \mathbb{N}$. If $f(x,y)$ is an A-recursive function, then there is a recursive function $e(y)$ s.t. $\varphi^A_{e(y)} = \varphi^A_{f(e(y),y)}$ for all y. Moreover, e does not depend on A.

Proof.

Let the index e code the function

$$\varphi^A_e(x,y,z) = \begin{cases} \varphi^A_{\varphi_x(x,y)}(z) & \text{if } \varphi_x(x,y)\downarrow \\ \uparrow & \text{otherwise} \end{cases}$$

By the relativized s-m-n theorem there is a recursive function $s(x,y)$ s.t.

$$\varphi^A_{s(x,y)} = \varphi^A_e(x,y,z)$$

We know $\exists v : \varphi^A_v(x,y) = f(s(x,y),y)$. Let $e(y) := s(v,y)$.

$$\varphi^A_{e(y)} = \varphi^A_{s(v,y)} = \varphi^A_{\varphi^A_v(v,y)} = \varphi^A_{f(s(v,y),y)} = \varphi^A_{f(e(y),y)}$$

From Kleene's Fixpoint Theorem to Ackermann's Function

We can use Kleene's fixpoint theorem to prove that the Ackermann's function is recursive. Consider:

$$g(k,m,n) := \begin{cases} n+1 & \text{if } m = 0 \\ \varphi_k(m-1,1) & \text{if } m > 0 \text{ and } n = 0 \\ \varphi_k(m-1,\varphi_k(m,n-1)) & \text{if } m > 0 \text{ and } n > 0 \end{cases}$$

Clearly, g is partial recursive. By the s-m-n Theorem, there is a recursive function h such that

$$\varphi_{h(k)}(m,n) = g(k,m,n)$$

By Kleene's fixpoint theorem, there exists e s.t.

$$\varphi_{h(e)} = \varphi_e$$

Therefore, $\varphi_e(m,n)$ satisfies the definition of Ackermann's function.

Rice's Theorem

Theorem (Rice's Theorem)

A set of partial recursive functions $\mathcal{A}$ is recursive iff it is trivial, i.e. either $A = \emptyset$ or $A = \mathbb{N}$, where $A := \{x : \varphi_x \in \mathcal{A}\}$.

Proof.

Let $a \in A$ and $b \notin A$.

$$h(x) := \begin{cases} a & \text{if } x \notin A \\ b & \text{if } x \in A \end{cases}$$

Obviously, h is recursive, and $\forall x : x \in A \leftrightarrow h(x) \notin A$.
By Kleene's fixpoint theorem, $\exists e : \varphi_e = \varphi_{h(e)}$.
Hence $e \in A \iff h(e) \in A$. Contradiction. □

Remark: For any non-trivial property of partial functions, no general and effective method can decide whether an algorithm computes a partial function with that property.
All non-trivial semantic properties of programs are undecidable.

Recursion Theorem

Theorem (Second Recursion Theorem)

If $f(x, y)$ is a partial recursive function, there is an index e s.t.

$$\varphi_e(y) = f(e, y)$$

Kleene's Fixpoint Theorem $\iff$ Second Recursion Theorem

Theorem (First Recursion Theorem)

Every partial recursive functional $F(\alpha, x)$ admits a least fixpoint. In other words, there is a partial recursive function α s.t.

1. $\forall x\big(\alpha(x) = F(\alpha, x)\big)$
2. $\forall x\big(\beta(x) = F(\beta, x)\big) \implies \alpha \subset \beta$

Gödel's Speed-up Theorem

Theorem (Gödel's Speed-up Theorem)

Let $\mathrm{T}' \supset \mathrm{T}$ be formal systems (with recursive sets of axioms and of recursive rules) such that $\mathrm{T}' \setminus \mathrm{T}$ is not r.e. Given a recursive function h, there is a theorem A of T and a number n such that A admits a proof of length $\leq n$ in T', but no proof of length $\leq h(n)$ in T.

Proposition

If T is an essentially undecidable formal system, and $\mathrm{T} \nvdash A$, then $\mathrm{T} \cup \{A\} \setminus \mathrm{T}$ is not r.e.

Remark: Adding an unprovable sentence to an essentially undecidable formal system T radically shortens some proof of some theorem of T.

Blum's Speed-up Theorem

A Blum complexity measure is a pair (φ, Φ) with φ a Gödel numbering of the partial recursive functions $\mathbf{P}^{(1)}$ and a recursive function $\Phi : \mathbb{N} \longrightarrow \mathbf{P}^{(1)}$ such that

- the domains of φ_i and Φ_i are identical.
- the set $\{(i, x, t) \in \mathbb{N}^3 : \Phi_i(x) = t\}$ is recursive.

Theorem (Blum's Speed-up Theorem)

Given a Blum complexity measure (φ, Φ) and a total recursive function f with two parameters, there exists a $0,1$-valued total recursive function g s.t. for every index i for g, there is another index j for g s.t. for almost all x

$$f(x, \Phi_j(x)) \leq \Phi_i(x)$$

Remark: For any complexity measure there are recursive functions that are not optimal with respect to that measure. There is no notion of best complexity for all total recursive functions.

Remark: No computer can be optimal for every purpose: no matter how good a computer is, there are always functions on which such a computer behaves very badly.

Turing Degree

Definition

Let $A \subset \mathbb{N}$. A function $f : \mathbb{N} \to \mathbb{N}$ is computable in A (we write $f \leq_T A$) if there exists a program that calculates f using χ_A as a primitive function.

Definition

A and B are Turing-equivalent ($A \equiv_T B$) if $A \leq_T B$ and $B \leq_T A$.

- A and $\mathbb{N} \setminus A$ are Turing-equivalent.
- Given B, the set $\{A \subset \mathbb{N} : A \leq_T B\}$ is countable.
- $\mathrm{P}(\mathbb{N})/\equiv_T$ is uncountable.
- There are $A, B \subset \mathbb{N}$ such that $A \not\leq_T B$ and $B \not\leq_T A$.
  Even worst, there are uncountable $\leq_T$-antichains.
- Every countable partial ordering can be embedded in the Turing degrees.
  Even worst, it can be embedded below the diagonal Halting set $K = \{x : \varphi_x(x) \downarrow\}$.

m-reduction

Definition (m-reduction)

We say that A is many-one reducible (m-reducible) to B, and write $A \leq_m B$, if there is a total recursive function f such that for all x, we have $x \in A \iff f(x) \in B$.

Definition (1-reduction)

If the function f in the definition of m-reduction is injective, then we say that A is 1-reducible to B, and write $A \leq_1 B$.

Productive and Creative Sets

Definition (Productive Set)

A set A is *productive* if there is a partial recursive function f such that

$$W_e \subset A \implies f(e) \downarrow \ \& \ f(e) \in A \setminus W_e$$

Definition (Creative Set)

A r.e set A is *creative* iff its complement is productive.

Theorem

The following are equivalent:

- A *is creative.*
- A *is* m*-complete (i.e.,* m*-equivalent to* $\emptyset'$*)*
- A *is* 1*-complete (i.e.,* 1*-equivalent to* $\emptyset'$*)*

Theorem

The set $\#\operatorname{Th}\mathcal{N} := \left\{\#A : \mathcal{N} \vDash A\right\}$ *is productive.*

Immune and Simple Sets

Definition (Immune Set)

A set A is *immune* iff it is infinite and contains no infinite r.e. subsets.

Definition (Simple Set)

A r.e. set A is *simple* iff its complement is immune.

$$W_e \text{ infinite} \implies W_e \cap A \neq \emptyset$$

Theorem

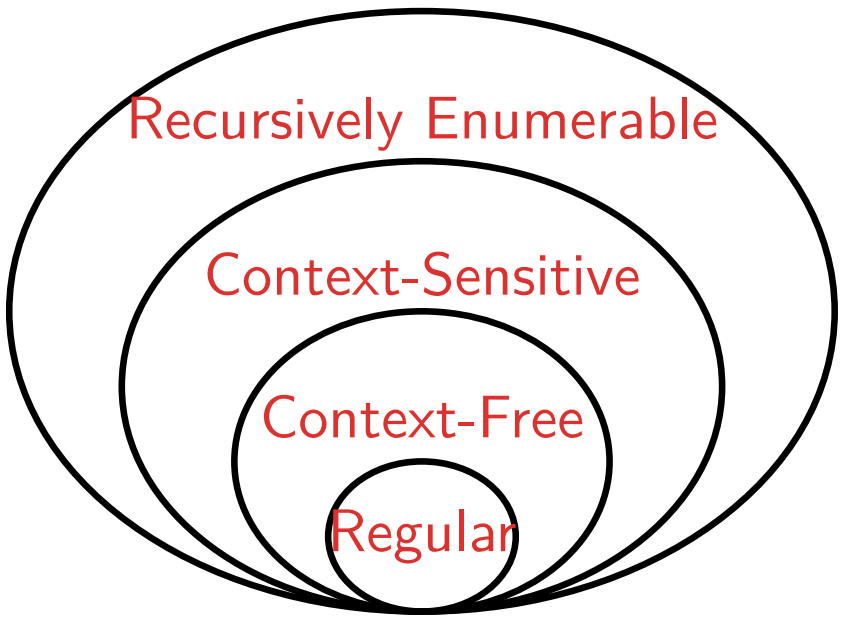
The set of non-random numbers $\{x : K(x) < \ell(x)\}$ *is simple.*

Chomsky Grammar

A grammar is a mathematical specification of the set of all word sequences that form valid sentences in a language (N, T, P, S).

1. a finite set N of non-terminals,
2. a finite set T of terminals,
3. a finite set P of production rules,
4. a choice of start symbol $S \in N$.

- Regular (Left Linear): $A \longrightarrow Bx$ or $A \longrightarrow x$, where $A, B \in N$ and $x \in T^*$.
- Context-Free: $A \longrightarrow \alpha$, where $A \in N$ and $\alpha \in (N \cup T)^*$.
- Context-Sensitive: $\alpha \longrightarrow \beta$ and $|\alpha| \leq |\beta|$, where $\alpha, \beta \in (N \cup T)^+$.
- Recursively Enumerable: $\alpha \longrightarrow \beta$.

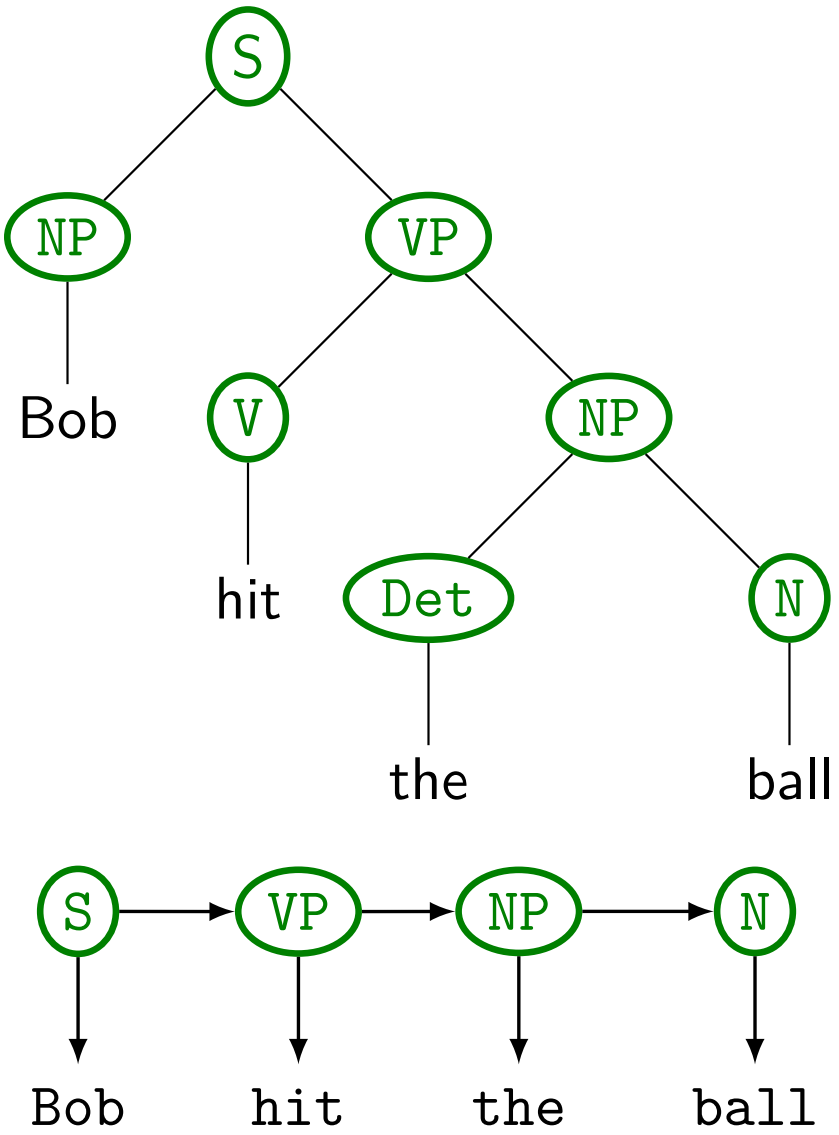


Example

- Humans usually think of natural language using CFG.
- A CFG with finite recursion depth can be written as a regular grammar.

$$
\begin{aligned}
\texttt{S} &\longrightarrow \texttt{NP VP} \\
\texttt{VP} &\longrightarrow \texttt{V NP} \\
\texttt{NP} &\longrightarrow \texttt{Det N} \\
\texttt{NP} &\longrightarrow \texttt{Bob} \\
\texttt{V} &\longrightarrow \texttt{hit} \\
\texttt{Det} &\longrightarrow \texttt{the} \\
\texttt{N} &\longrightarrow \texttt{ball}
\end{aligned}
\qquad
\begin{aligned}
\texttt{S} &\longrightarrow \texttt{Bob VP} \\
\texttt{VP} &\longrightarrow \texttt{hit NP} \\
\texttt{NP} &\longrightarrow \texttt{the N} \\
\texttt{N} &\longrightarrow \texttt{ball}
\end{aligned}
$$

- A regular grammar can be written using an HMM.



Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Turing Machine
Computability
Diagonal Method
Incompressibility Method
Incompleteness

Equational Logic

Homotopy Type Theory

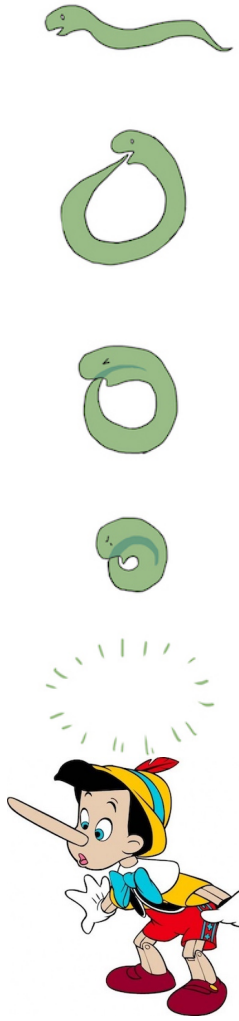
Category Theory

Quantum Computing

Answers to the Exercises
References2039

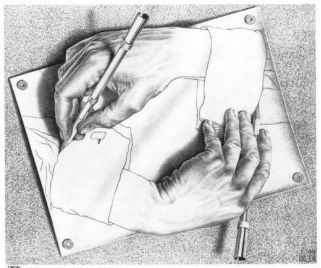
Self-Reference

- This sentence repeats the word ‘twice’ twice.
- Thare are five mistukes im this centence.
- **The only boldface sentence on this page is false.**
- All generalizations are wrong.
- Every rule has an exception except this one.
- Moderation in all things, including moderation.
- We must believe in free will — we have no choice!
- I know that I know nothing.
- There are two rules lor success in life:
  1. Never tell anyone all that you know.
- If you choose an answer to this question at random, what is the chance you will be correct? (A) 25% (B) 50% (C) 0% (D) 25%
- 1. What is the best question to ask and what is the answer to it?
  2. The best question is the one you asked; the answer is the one I gave.
- Can you answer the following question in the same way to this one?
- One of the lessons of history is that no one ever learns the lessons of history.
- 涅槃是消除了一切欲望后才能抵达的境界，包括对涅槃的欲望.



Self-Reference vs Paradox

The sentence below is false.



The sentence above is true.

Yablo Paradox

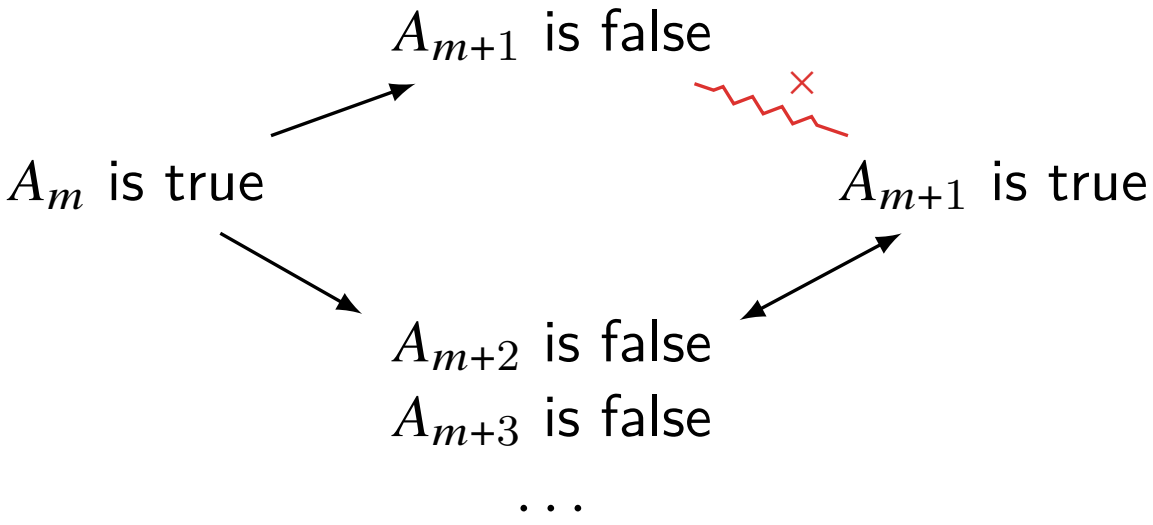
- A_1: for all $k > 1$, A_k is false.
- A_2: for all $k > 2$, A_k is false.
- A_3: for all $k > 3$, A_k is false.
- $\cdots$

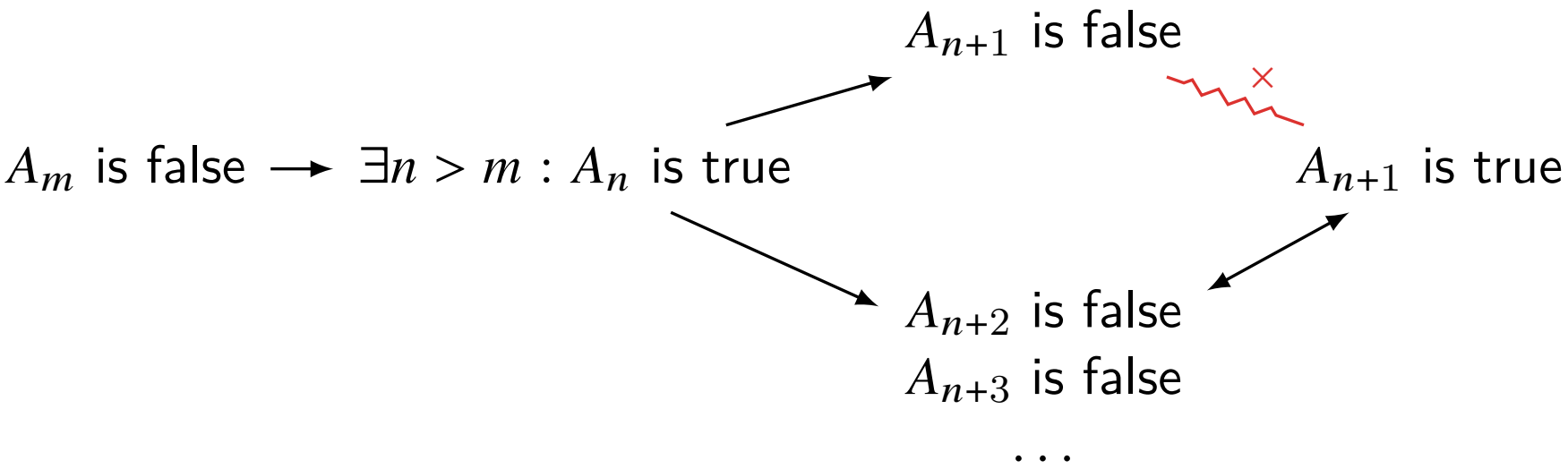
Quine Paradox

"Yields falsehood when preceded by its quotation" yields falsehood when preceded by its quotation.

self-reference / circularity or infinite regress / negation / infinity / totality

Yablo Paradox





The "Power" of Self-Reference

Curry's Paradox

- If this sentence is true, then God exists.
- This sentence is false, and God does not exist.

1. At least one of these two sentences is false.
2. God does not exists.

Hi 美女, 问你个问题呗

如果我问你"你能做我女朋友吗", 那么你的答案和**这个问题**的答案是一样的吗?

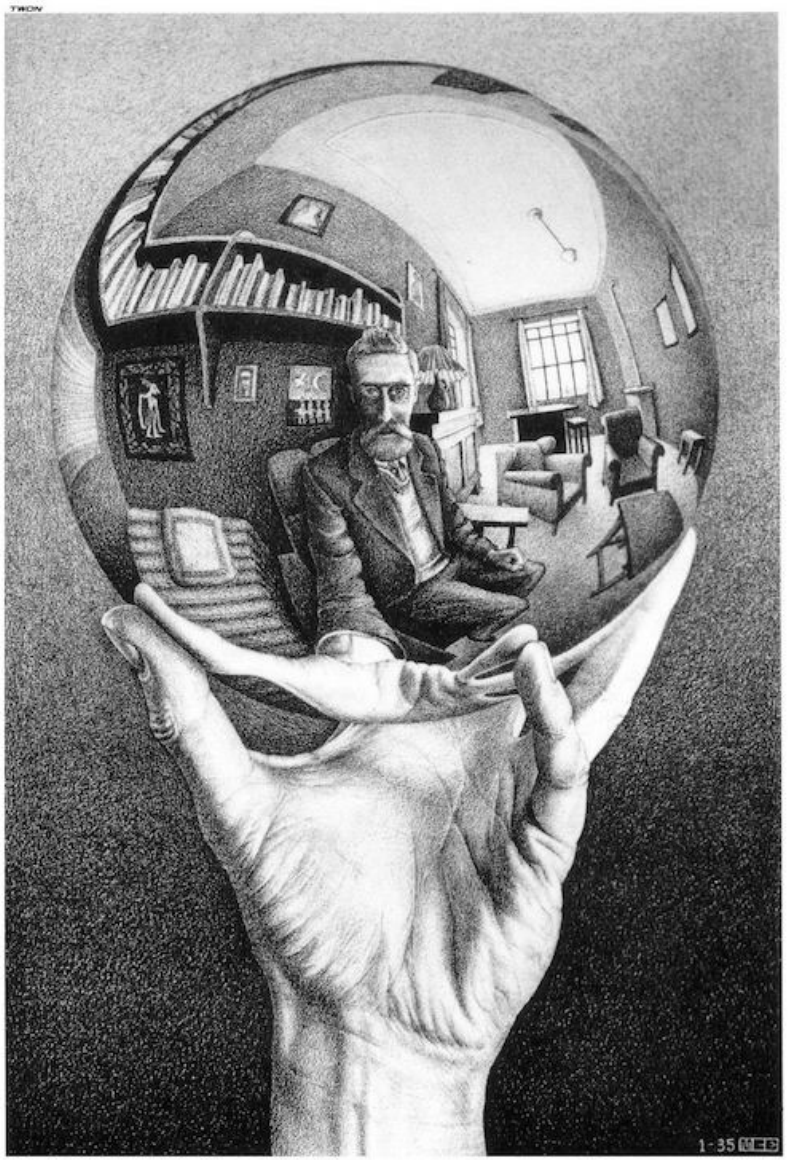
自我修复/自我实现?

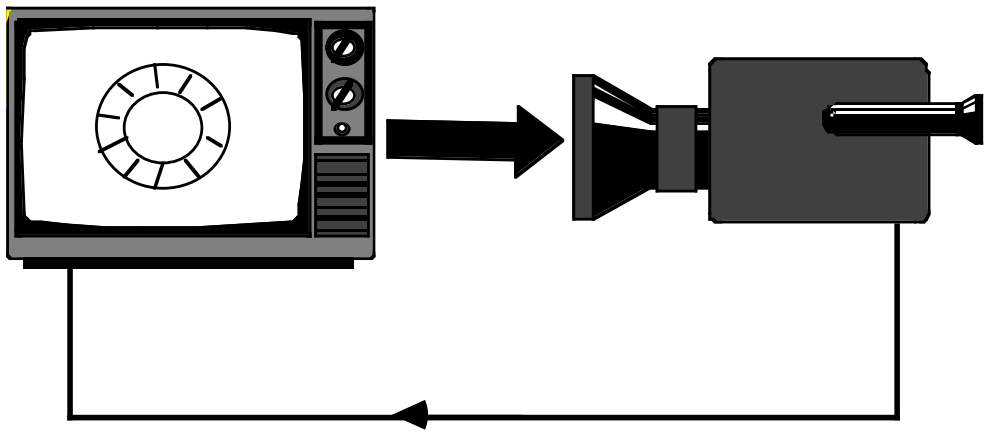
- "This sentence has _____ letters." thirty-one / thirty-three
- 这句话有 2 个 '这' 字, 2 个 '句' 字, 2 个 '话' 字, 2 个 '有' 字, 7 个 '2' 字, 11 个 '个' 字, 11 个 '字' 字, 2 个 '7' 字, 3 个 '11' 字, 2 个 '3' 字.

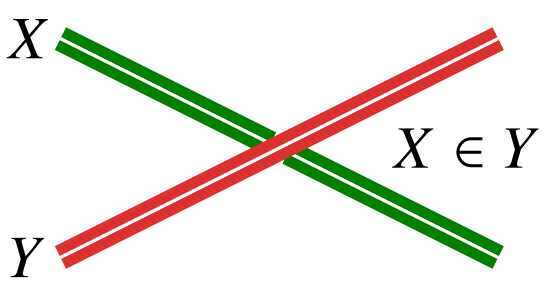
How to Refer? — Levels

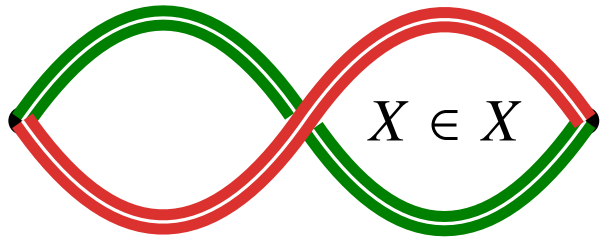


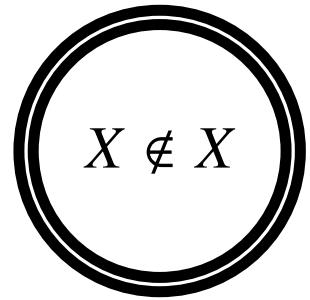
How to Refer?

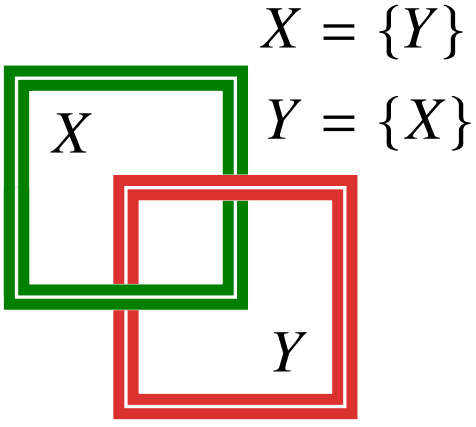


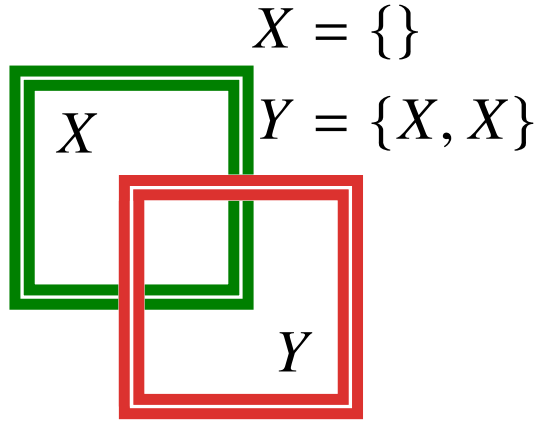


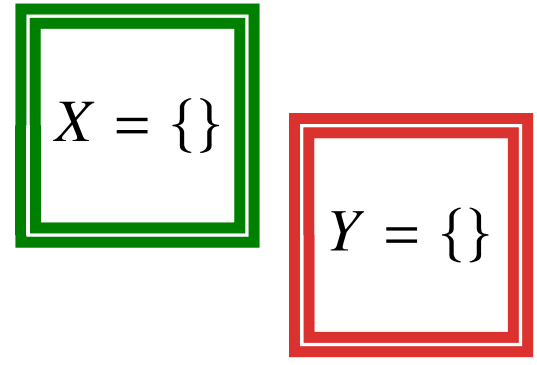

X
Y
X ∈ Y

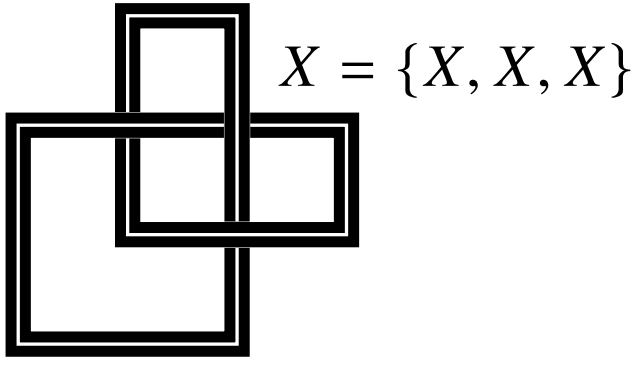

X ∈ X

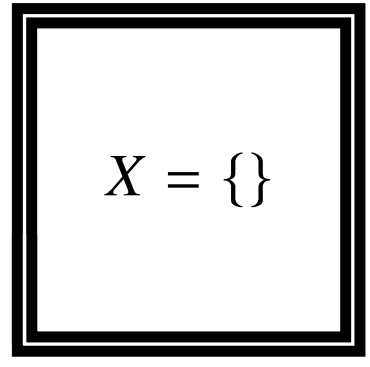

X ∉ X


X = {Y}
Y = {X}
X
Y

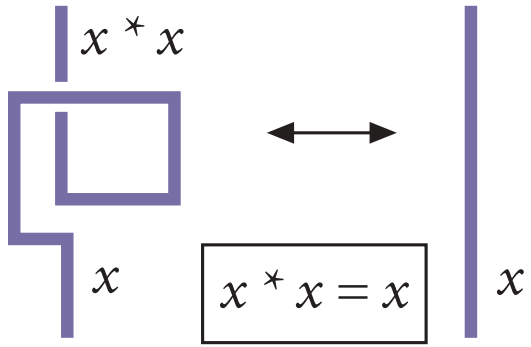

X = {}
Y = {X, X}
X
Y


X = {}
Y = {}

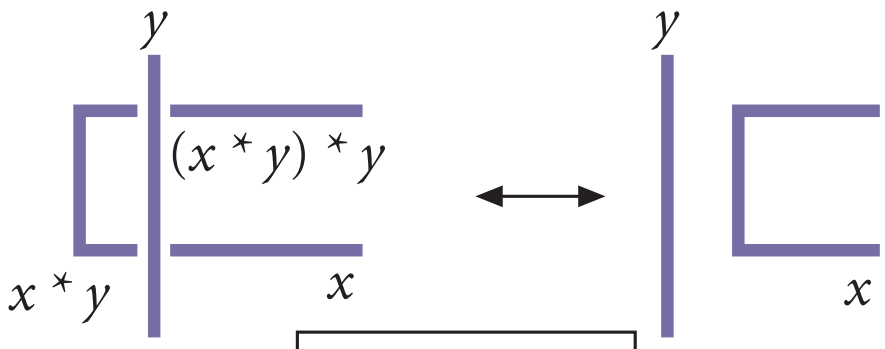

X = {X, X, X}


X = {}

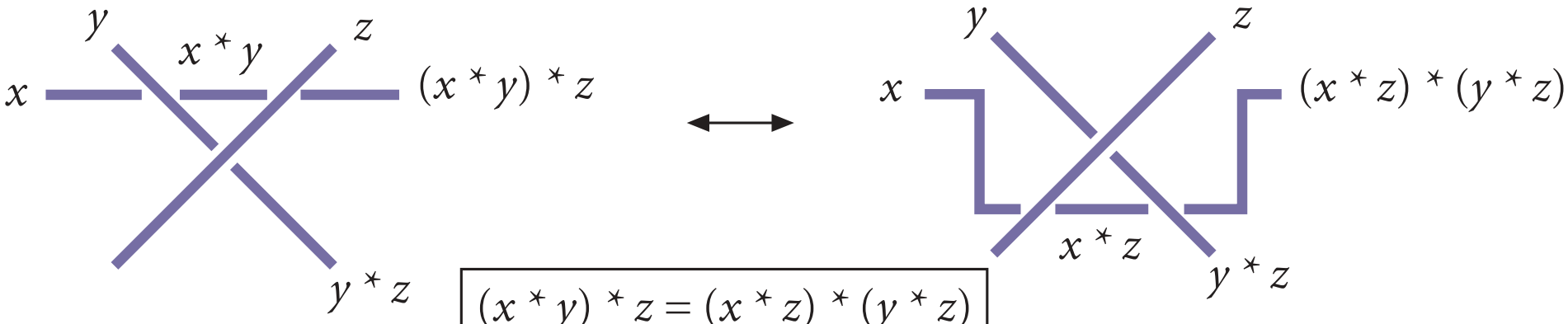
Reidemeister Moves



$$x * x = x$$

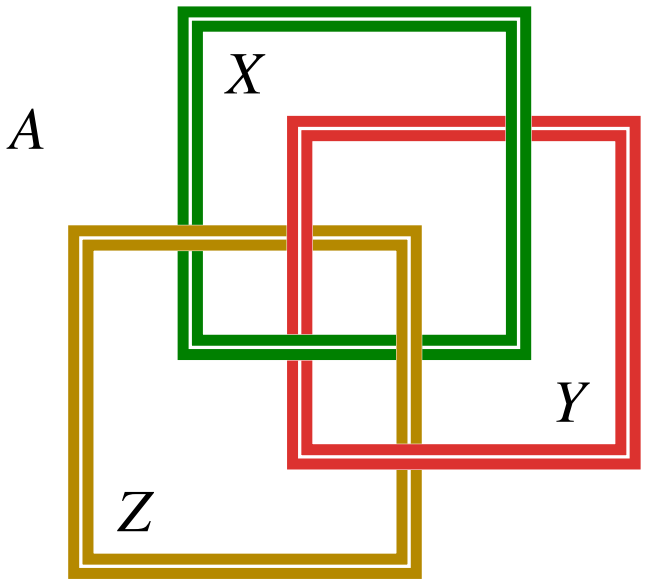


$$(x * y) * y = x$$

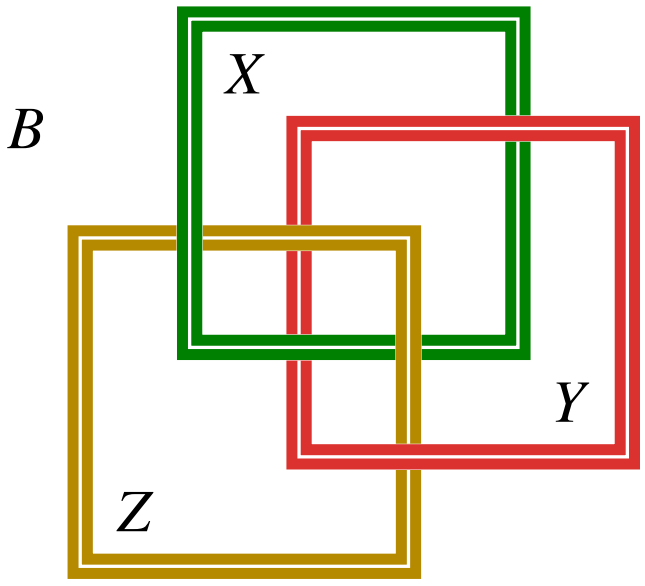


$$(x * y) * z = (x * z) * (y * z)$$

Self-Reference & IIT



$$X = \{Y, Y\}$$
$$Y = \{Z, Z\}$$
$$Z = \{X, X\}$$



$$X = \{Y, Z\}$$
$$Y = \{X, Z\}$$
$$Z = \{X, Y\}$$

Figure: $\Phi(A) < \Phi(B)$?

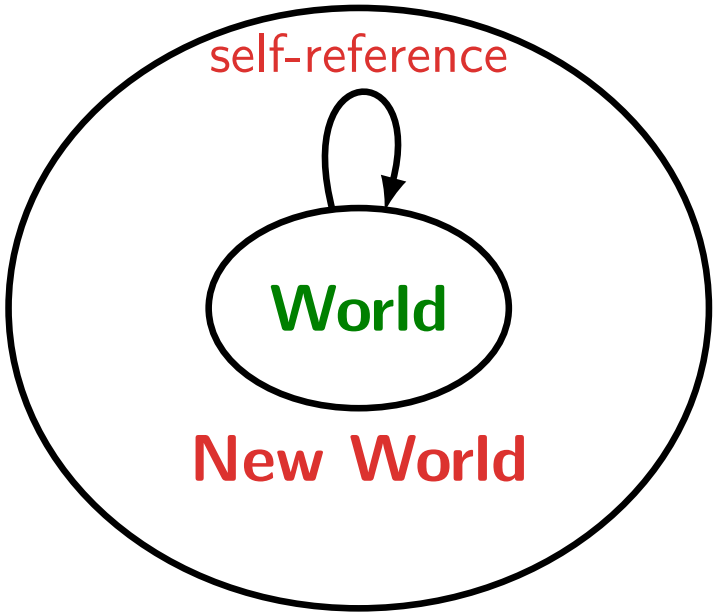
Larger Domain

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

$$F_0 = F_1 = 1; F_{n+1} = F_n + F_{n-1}$$

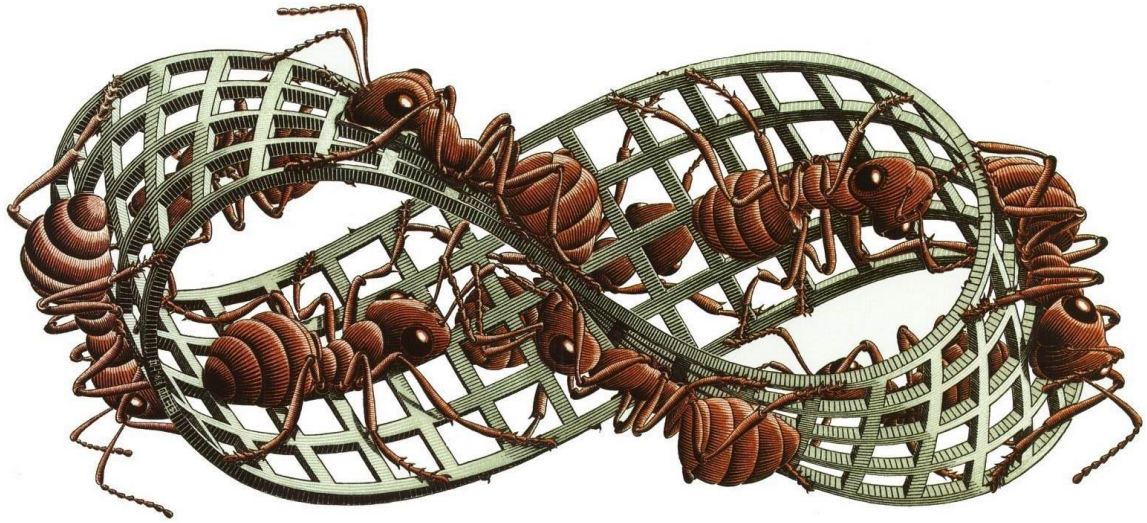
$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n}$$

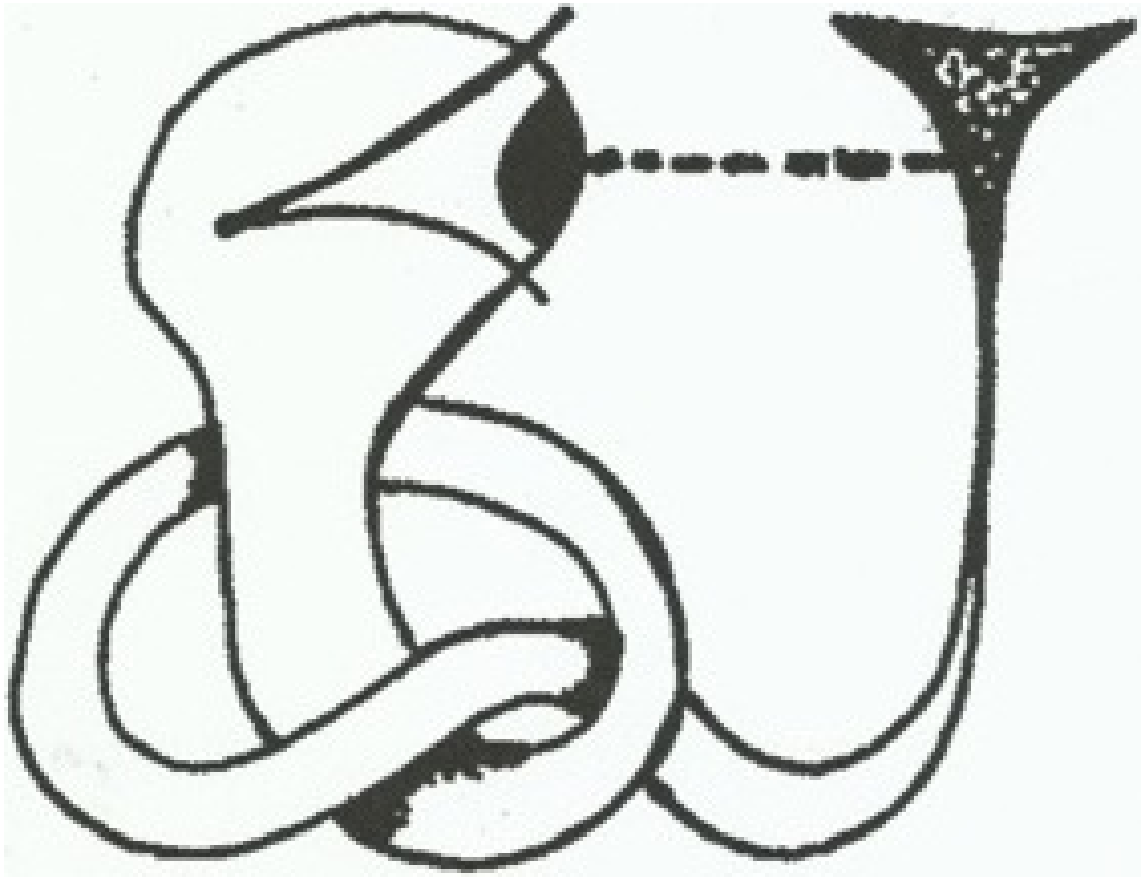
$$\frac{F_{n+1}}{F_n} = 1 + \frac{1}{\frac{F_n}{F_{n-1}}}$$



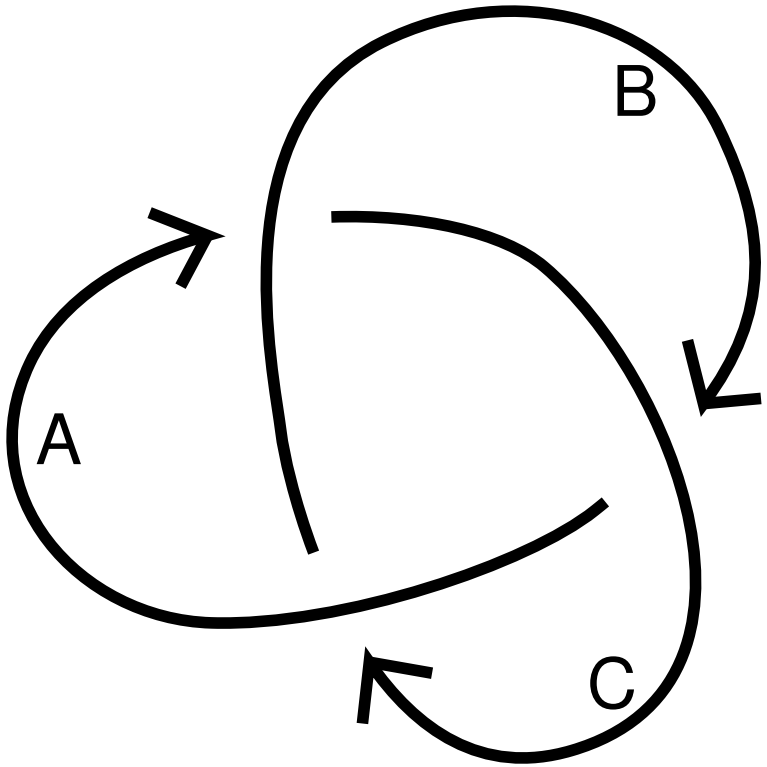
$$f(x) = 1 + \frac{1}{x} = x \implies 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \ddots}}}} = \frac{1+\sqrt{5}}{2}$$

$$f(x) = \frac{-1}{x} = x \implies x = i$$





Trefoil



- objects $\{A, B, C\}$
- morphisms
  - A: $C \to B$
  - B: $A \to C$
  - C: $B \to A$

Nested Virtualization?



从前有座山，山里有座庙，庙里有个老和尚在讲故事：从前有座山…

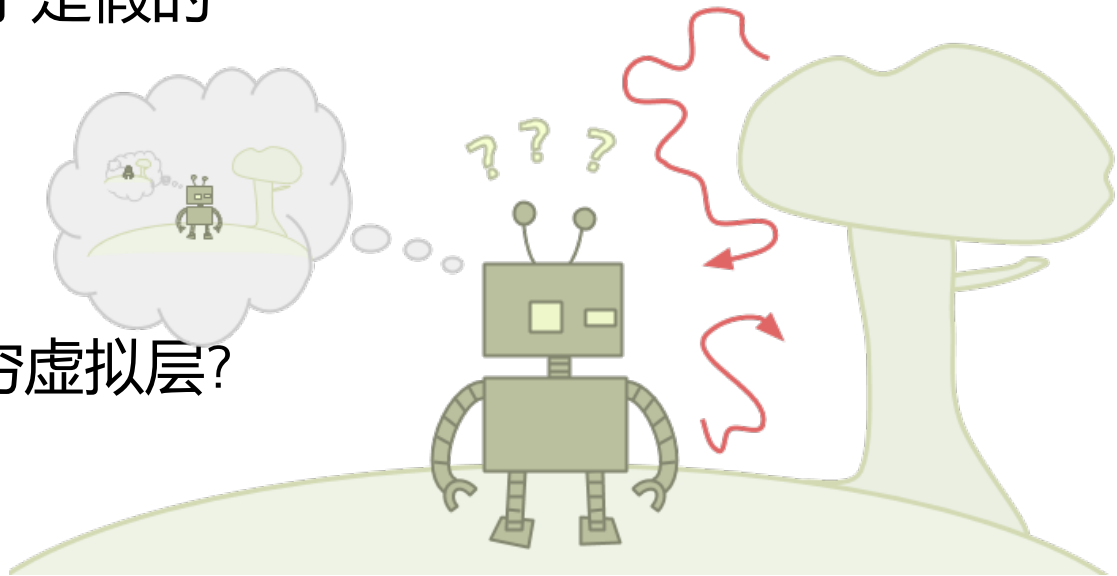
$$\begin{cases} FX = Y \\ GY = X \end{cases}$$

$$X = GFGFGFGF\cdots$$

$$Y = FGFGFGFG\cdots$$

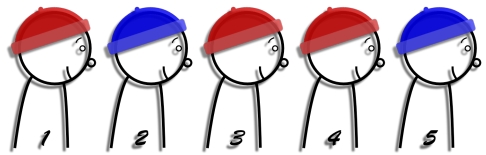
Liar Paradox vs Quine Paradox

1. 这句话是假的
2. “这句话是假的” 是假的
3. ““““““……是假的” 是假的” 是假的” 是假的” 是假的” 是假的
4. 把 “把中的第一个字放到左引号前面，其余的字放到右引号后面，并保持引号及其中的字不变'' 中的第一个字放到左引号前面, 其余的字放到右引号后面, 并保持引号及其中的字不变
5. 把 “把中的第一个字放到左引号前面，其余的字放到右引号后面，并保持引号及其中的字不变得到的句子是假的” 中的第一个字放到左引号前面, 其余的字放到右引号后面, 并保持引号及其中的字不变得到的句子是假的



无穷虚拟层?

How to Refer? — Encoding



- 100 prisoners are lined up by an jailer, who places a red or blue hat upon each of their heads.
- The prisoners can see the hats of the people lined up in from of them, but they can't look at the hats behind them, or at their own.
- The jailer is going to ask color of each prisoner's hat starting from the last prisoner in queue. If a prisoner tells the correct color, then is saved, otherwise executed.
- How many prisoners can be saved at most if they are allowed to discuss a strategy before the jailer starts asking colors of their hats?

If the first person sees an odd number of red hats he calls out red, if he sees an even number of red hats he calls out blue.

手扶拐杖的外星绅士造访地球. 临别, 人类赠送百科全书: "人类文明尽在其中!". 绅士谢绝: "不, 谢谢! 我只需在拐杖上点上一点".

What is the Next Number?

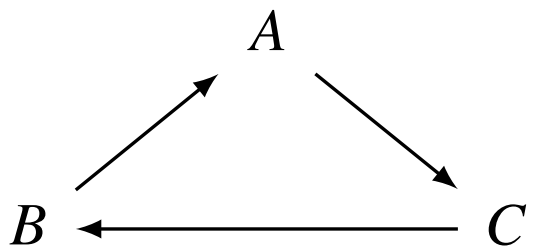
1. 1
2. 11
3. 21
4. 1211
5. 111221
6. 312211
7. ?

A. 11131221131211132221...

B. 31131122211311123113321...

C. 13211321322113311213212321...





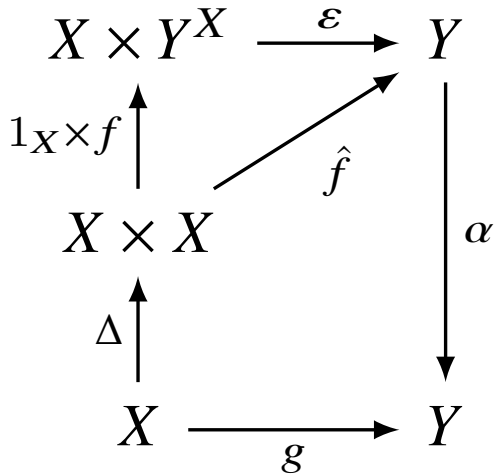
Diagonalization[27]

Definition (Point-Surjective)

A morphism $f : X \to Y$ is *point-surjective* iff for every $y : 1 \to Y$, there is an $x : 1 \to X$ s.t. $y = f \circ x$.

Theorem (Lawvere's Fixpoint Theorem)

In a cartesian closed category, if there is a point-surjective morphism $f : X \to Y^X$, then every morphism $\alpha : Y \to Y$ has a fixpoint $y : 1 \to Y$.





[27] Lawvere: Diagonal arguments and cartesian closed categories.
Yanofsky: A universal approach to self-referential paradoxes, incompleteness and fixed points.

Lawvere's Fixpoint Theorem

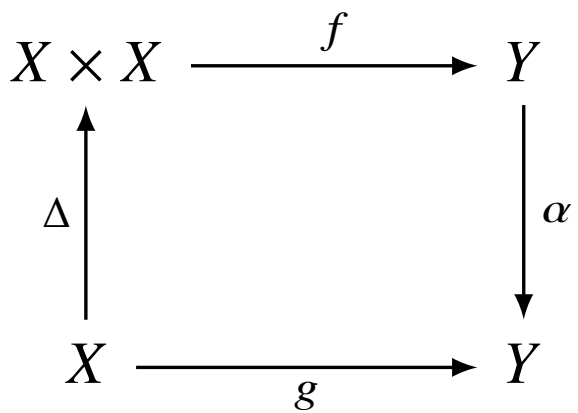
- A function $g : X \to Y$ is *representable* by $f : X \times X \to Y$ iff

$$\exists y \forall x : g(x) = f(x, y)$$

Theorem (Lawvere's Fixpoint Theorem)

For sets X, Y, functions $f : X \times X \to Y$, $\alpha : Y \to Y$, let $g := \alpha \circ f \circ \Delta$.

1. *If α has no fixpoint, then g is not representable by f.*
2. *If g is representable by f, then α has a fixpoint.*



$$\alpha\big(f\left(\ulcorner g \urcorner, \ulcorner g \urcorner\right)\big) = g\left(\ulcorner g \urcorner\right) = f\left(\ulcorner g \urcorner, \ulcorner g \urcorner\right)$$

- $\Delta : x \mapsto \langle x, x \rangle$ diagonal
- f evaluation
- α "negation"
- $g\left(\ulcorner g \urcorner\right)$ fixpoint-(free) transcendence
- $f\left(\ulcorner g \urcorner, \ulcorner g \urcorner\right)$ self-reference "I have property α."

Lawvere's Fixpoint Theorem

f	0	1	2	3	$\cdots$	t	$\cdots$
0	$\alpha f(0,0)$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$f(0,t)$	$\cdots$
1	$\cdots$	$\alpha f(1,1)$	$\cdots$	$\cdots$	$\cdots$	$f(1,t)$	$\cdots$
2	$\cdots$	$\cdots$	$\alpha f(2,2)$	$\cdots$	$\cdots$	$f(2,t)$	$\cdots$
3	$\cdots$	$\cdots$	$\cdots$	$\alpha f(3,3)$	$\cdots$	$f(3,t)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
t	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$f(t,t)$ $\parallel$ $\alpha f(t,t)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Diagonalization

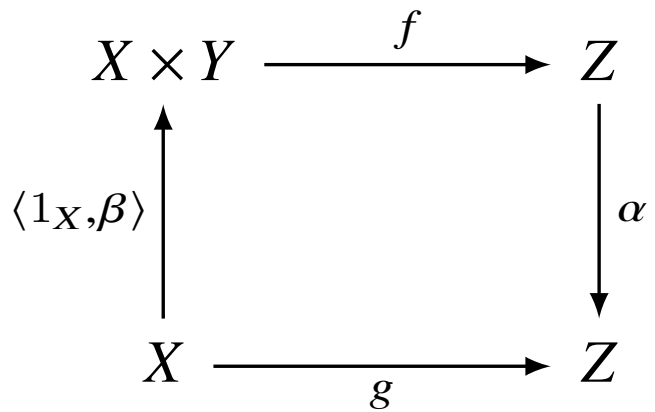
- A function $g : X \to Z$ is *representable* by $f : X \times Y \to Z$ iff

$$\exists y \in Y \forall x \in X : g(x) = f(x, y)$$

Theorem (Lawvere's Fixpoint Theorem)

For all sets X, Y, Z, and all functions $f : X \times Y \to Z, \alpha : Z \to Z$, surjective functions $\beta : X \twoheadrightarrow Y$, let $g := \alpha \circ f \circ \langle 1_X, \beta \rangle$.

1. *If α has no fixpoint, then g is not representable by f.*
2. *If g is representable by f, then α has a fixpoint.*



Lawvere's Fixpoint Theorem — Multi-Valued Version

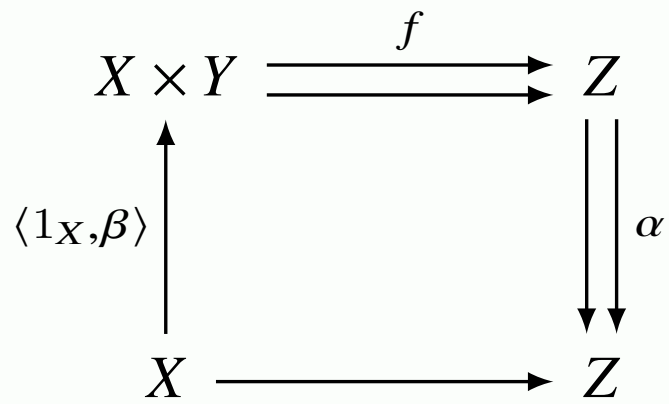
- A *multi-valued function* $f : X \rightrightarrows Y$ is a function $f : X \to \mathtt{P}(Y)$ s.t. $\forall x \in X \exists y \in Y : y \in f(x)$.

Theorem (Lawvere's Fixpoint Theorem — multi-valued version)

For sets X, Y, Z, multi-valued functions $f : X \times Y \rightrightarrows Z$, $\alpha : Z \rightrightarrows Z$, and surjective function $\beta : X \twoheadrightarrow Y$,

$$\exists y \forall x \Big(\alpha \circ f \circ \langle 1_X, \beta \rangle x \cap f(x, y) \neq \emptyset \Big) \to \exists z \in \alpha(z)$$

Proof.

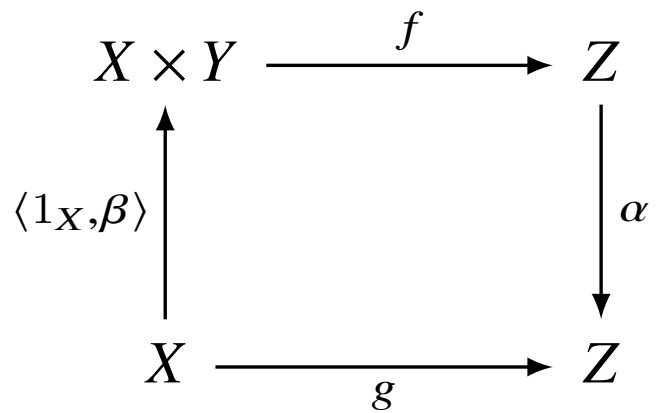


$$\beta x = y \implies \alpha(f(x, y)) \cap f(x, y) \neq \emptyset$$

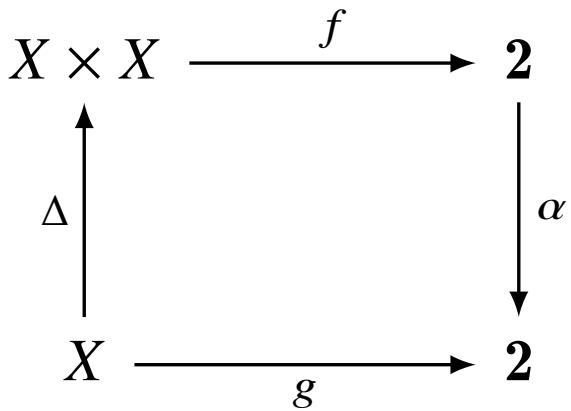
Approximate Version

Theorem (Approximate Version)

Suppose (Z, d) *is a metric space. Given* $f : X \times Y \to Z$, $\alpha : Z \to Z$, *and surjective* $\beta : X \twoheadrightarrow Y$, *let* $g := \alpha \circ f \circ \langle 1_X, \beta \rangle$. *If* g *is* ε*-representable by* f*:* $\exists y \in Y \forall x \in X : d(f(x, y), g(x)) < \varepsilon$, *then* α *has* ε*-fixpoint:* $\exists z \in Z : d(z, \alpha(z)) < \varepsilon$.



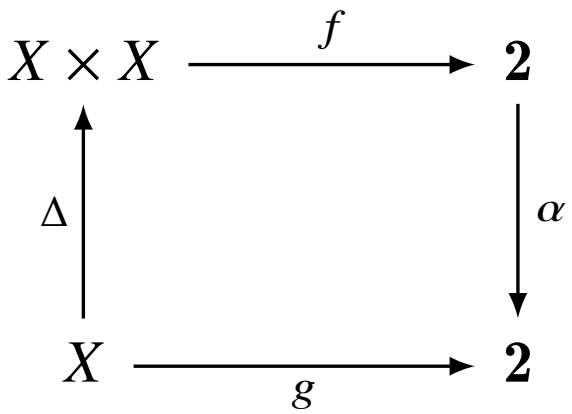
Example — Grelling/Liar/Quine...Paradox



where $f : (x, y) \mapsto [\![y \text{ “describes” } x]\!]$ and $\alpha : x \mapsto 1 - x$.

- Is “non-self-descriptive” non-self-descriptive?
- “This sentence is false.”
- “Yields falsehood when preceded by its quotation” yields falsehood when preceded by its quotation.

Example — Russell Paradox



where

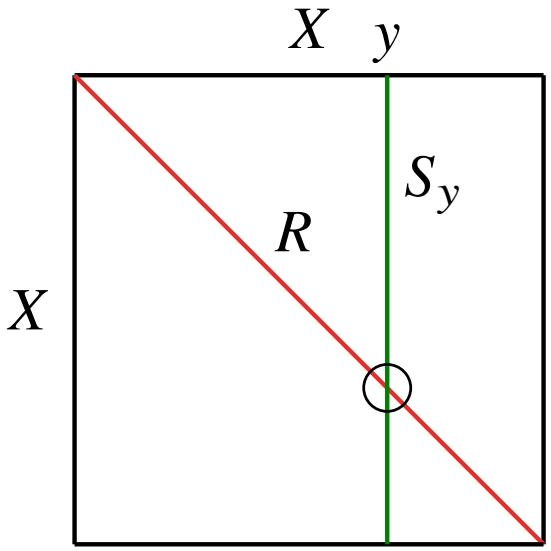
$$f : (x, y) \mapsto [\![x \in y]\!]$$

and

$$\alpha : x \mapsto 1 - x$$

$$R := \{x : x \notin x\} \quad \text{exist?}$$

Barber paradox: $f : (x, y) \mapsto [\![y \text{ “shaves” } x]\!]$



Let $S \subset X \times X$

$$S_y := \{x : Sxy\}$$

$$R := \{x : x \notin S_x\}$$

$$\forall x : \; R \neq S_x$$

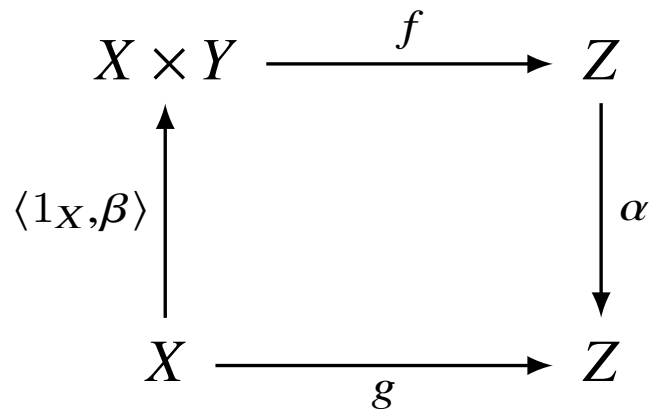
Example — Russell/Barber Paradox

f	S_0	S_1	S_2	S_3	$\cdots$	S_t	$\cdots$
S_0	0	1	0	1	$\cdots$	1	$\cdots$
S_1	1	1	0	0	$\cdots$	0	$\cdots$
S_2	0	1	0	0	$\cdots$	1	$\cdots$
S_3	0	0	0	1	$\cdots$	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
S_t	0	0	0	0	$\cdots$	?	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Remark: the diagonal is the opposite of the Barber column S_t.

Example — Wittgenstein/Kripke Paradox

- X: word space.
- Y: rule space.
- Z: meaning space.
- $g \in Z^X$ is finding a meaning.
- In naive realism, it is assumed that the meaning is certain.
- What we assume, that a meaning found for a word is certain, is something we do not have to consider another possibility of a meaning. Namely, $\forall g \in Z^X \exists y \in Y : g = \lambda x.f(x, y)$.
- In the context of Wittgenstein/Kripke, what we can by no means deny as another possibility for a meaning is expressed by infinite regression. It is formally replaced by a fixpoint.



Example — Yablo Paradox in Linear Temporal Logic(LTL)

$$
\begin{aligned}
n \vDash A \wedge B &\iff n \vDash A \;\&\; n \vDash B \\
n \vDash \neg A &\iff n \nvDash A \\
n \vDash \bigcirc A &\iff n+1 \vDash A \\
n \vDash \Box A &\iff \forall m \geq n \implies m \vDash A
\end{aligned}
$$

$$
\begin{array}{ccc}
\mathrm{Sen} \times \mathrm{Sen} & \xrightarrow{\;f\;} & \mathbf{2} \\
{\scriptstyle\Delta}\uparrow & & \downarrow{\scriptstyle\alpha} \\
\mathrm{Sen} & \xrightarrow[\;g\;]{} & \mathbf{2}
\end{array}
$$

$f : (X, Y) \mapsto [\![X \nleftrightarrow \bigcirc\Box\neg Y]\!]$ and $\alpha : x \mapsto 1 - x$

$$\nvDash A \leftrightarrow \bigcirc\Box\neg A$$

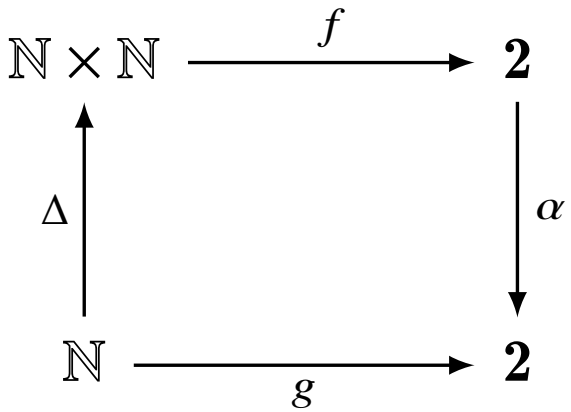
$$\vDash \neg\Box\big(A \leftrightarrow \bigcirc\Box\neg A\big)$$

Remark: Yablo Paradox: $A_n \leftrightarrow \forall m > n : \neg A_m$

Example — Euclid's Theorem?

Theorem (Euclid's Theorem)

There are infinitely many prime numbers.



where

$$f(m,n) = \begin{cases} 1 & \forall p \in \mathbb{P} : p \mid (m! + 1) \rightarrow p \leq n \\ 0 & \text{otherwise} \end{cases}$$

and $\alpha : x \mapsto 1 - x$.

Obviously, $\forall n : f(n,n) = 0$, and $g(n) = \alpha(f(n,n)) = 1$.

If $|\mathbb{P}| < \infty$, let $t := \max \mathbb{P}$, then $\forall n : f(n,t) = 1$ and $\forall n : g(n) = f(n,t)$.

Therefore, $f(t,t)$ is a fixpoint of α. Contradiction!

A Complete List of All Great Mathematicians

De Morgan

Abel

Boole

Brouwer

Sierpinski

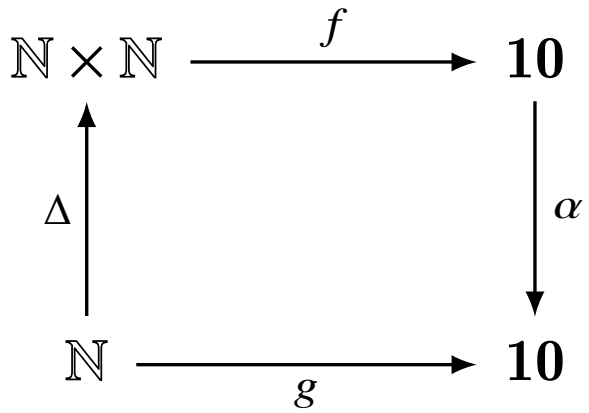
Weierstrass

Cantor

Example — The set of real numbers is uncountable

Theorem (Cantor)

$\mathbb{R}$ *is uncountable.*



where $f : (m, n) \mapsto r_{mn} :=$ "the n^{th} digit of the m^{th} real" and $\alpha : x \mapsto 9 - x$.

- There exists uncomputable real $\sum_n g(n)10^{-n}$, where

$$f : (m, n) \mapsto r_{mn} := \begin{cases} \text{the } n^{\text{th}} \text{ digit output by the } m^{\text{th}} \text{ Turing machine} \\ 0 \text{ if the } m^{\text{th}} \text{ Turing machine never outputs a } n^{\text{th}} \text{ digit} \end{cases}$$

- Richard paradox(unnameable real):

$f : (m, n) \mapsto r_{mn} :=$ "the n^{th} digit of the real named by the m^{th} sentence"

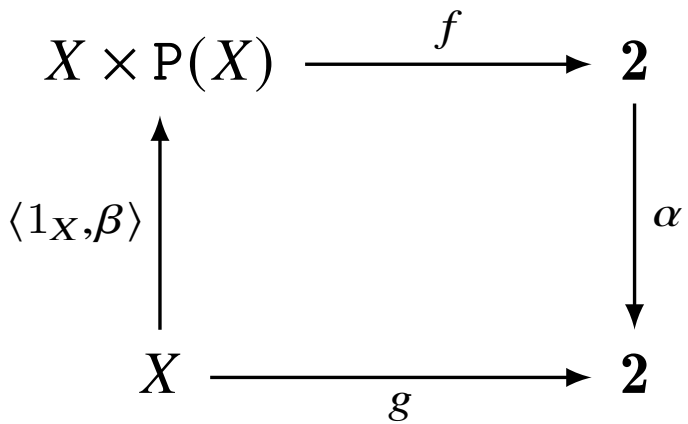
Example — The set of real numbers is uncountable

r_{mn}	0	1	2	3	4	5	$\cdots$	t	$\cdots$
0	5	7	2	7	7	4	$\cdots$	3	$\cdots$
1	1	2	2	7	6	7	$\cdots$	1	$\cdots$
2	3	0	3	0	0	0	$\cdots$	8	$\cdots$
3	6	2	0	4	2	0	$\cdots$	0	$\cdots$
4	1	0	2	3	1	3	$\cdots$	5	$\cdots$
5	1	0	3	0	1	0	$\cdots$	4	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
t	4	7	6	5	8	9	$\cdots$	?	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Example — Cantor's Theorem

Theorem (Cantor's Theorem)

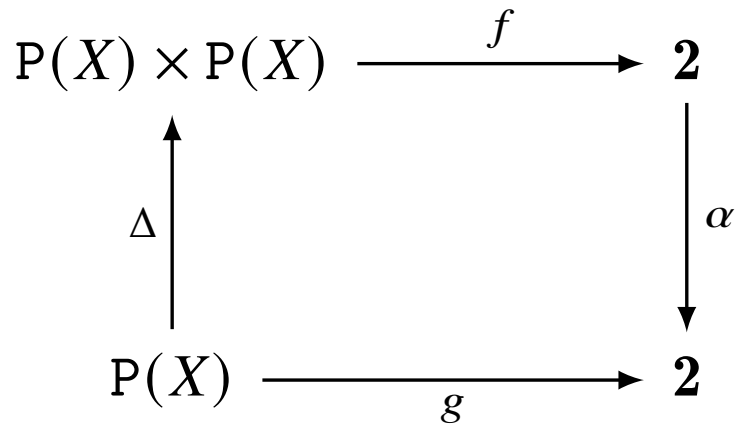
$$|X| < |\mathrm{P}(X)|$$



where $f : (x, y) \mapsto [\![x \in y]\!]$ and $\alpha : x \mapsto 1 - x$.

β is not surjective.

another proof: assume $h : \mathrm{P}(X) \rightarrowtail X$.

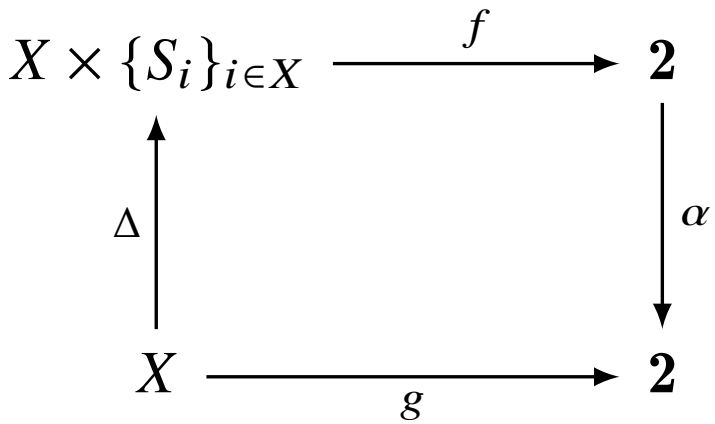


where $f : (x, y) \mapsto [\![h(x) \in y]\!]$, and $\alpha : x \mapsto 1 - x$.

g is representable by $y := \{ h(x) : x \subset X \;\&\; h(x) \notin x \}$.

Example — Cantor's Theorem — another proof

If $|X| \geq |\mathrm{P}(X)|$, then there exists some enumeration $\{S_i\}_{i \in X}$ of $\mathrm{P}(X)$.



where $f : (x, y) \mapsto [\![x \in S_y]\!]$ and $\alpha : x \mapsto 1 - x$.

$$g : x \mapsto [\![x \notin S_x]\!]$$

Since $\{S_i\}_{i \in X}$ is the enumeration of $\mathrm{P}(X)$, the set $R \coloneqq \{x : x \notin S_x\}$ that g characterizes must be some S_t: $\exists t(R = S_t)$. It means g is representable by t. Contradiction!

Example — Cantor's Theorem — another proof

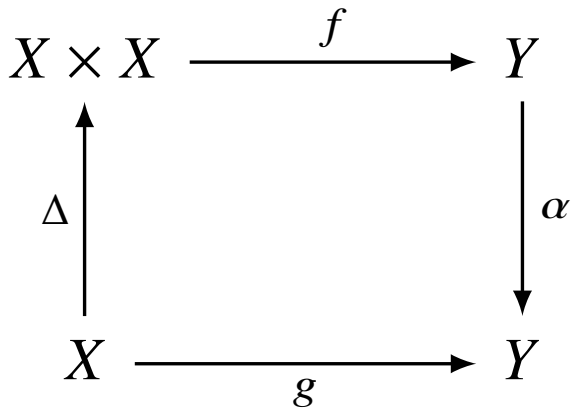
f	S_0	S_1	S_2	S_3	$\cdots$	S_t	$\cdots$
0	0	1	0	1	$\cdots$	1	$\cdots$
1	1	1	1	1	$\cdots$	0	$\cdots$
2	0	1	0	0	$\cdots$	1	$\cdots$
3	0	1	1	1	$\cdots$	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
t	1	0	1	0	$\cdots$	?	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Example — Cantor's Theorem

Theorem (Cantor's Theorem)

For $|Y| \geq 2$,

$$|X| < |Y^X|$$



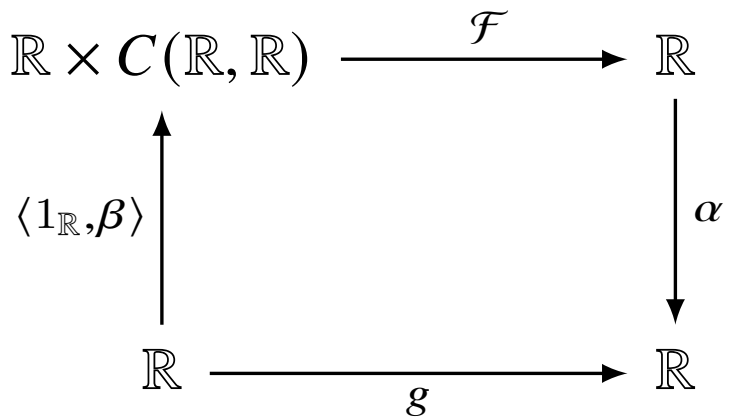
where α is the cyclic permutation.

Every $g : X \to Y$ is representable by some $f : X \times X \to Y$ iff $\exists f : X \twoheadrightarrow Y^X$.

If there exists $f : X \twoheadrightarrow Y^X$, *then every* $\alpha : Y \to Y$ *has a fixpoint.*

Example — Continuous Functions

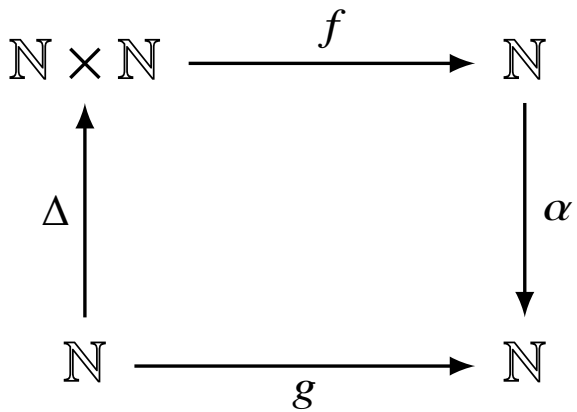
- Since a continuous function on $\mathbb{R}$ is determined by its values at rational points, the set of continuous functions $|C(\mathbb{R},\mathbb{R})| = |\mathbb{R}|$. However, there is no continuous surjection $\mathbb{R} \twoheadrightarrow C(\mathbb{R},\mathbb{R})$ from the real line to the Banach space of continuous real functions, equipped with the sup-norm $\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)|$.



  where $\mathcal{F} : (x, f) \mapsto f(x)$ and $\alpha : x \mapsto x + 1$.

- For most spaces X, there is no space-filling curve for its path space, $f : I \to X^I$.

Example — total recursive but not primitive recursive



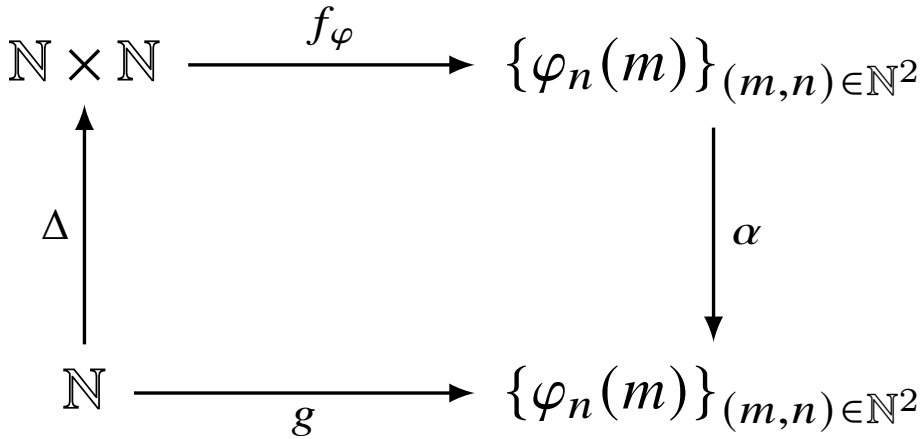
where $f : (m, n) \mapsto \psi_n(m)$ and $\alpha : x \mapsto x + 1$.

$$g : n \mapsto \psi_n(n) + 1$$

or, let $f : (m, n) \mapsto \max_{k \leq n} \psi_k(m)$.

Similarly, let $f : (m, n) \mapsto \max_{k \leq n} \varphi_k(m)$ then we get a busy beaver function.

Example — Berry Paradox vs Busy Beaver



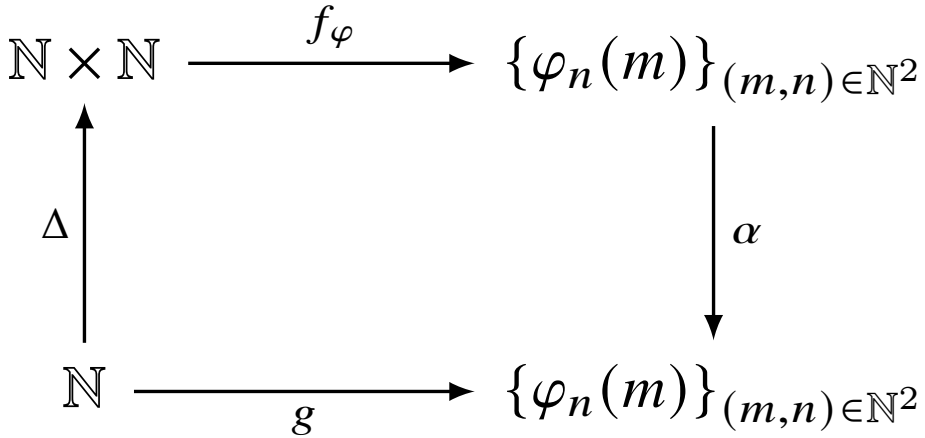
where $f_\varphi : (m,n) \mapsto \varphi_n(m)$, and $\alpha : \varphi_n(m) \mapsto \min\left(\mathbb{N} \setminus \{\varphi_k(m) : k \leq n\}\right)$

$$g(m) = \min\left(\mathbb{N} \setminus \{\varphi_k(m) : k \leq m\}\right) = \mu n\,[K(n \mid m) > m]$$

g unrepresentable $\implies$ g uncomputable $\implies$ K uncomputable

$$\Sigma(m) \coloneqq \max\{\varphi_k(0) : k \leq m\} = \max\{n : K(n) \leq m\}$$

Example — Not every partial recursive function is potentially recursive



where $f_\varphi : (m,n) \mapsto \varphi_n(m)$, and $\alpha : x \mapsto x+1$

$$g : m \mapsto \varphi_m(m) + 1$$

$$g \text{ partial recursive} \implies g \text{ representable} \implies \alpha(g(\ulcorner g \urcorner)) = g(\ulcorner g \urcorner) \uparrow$$

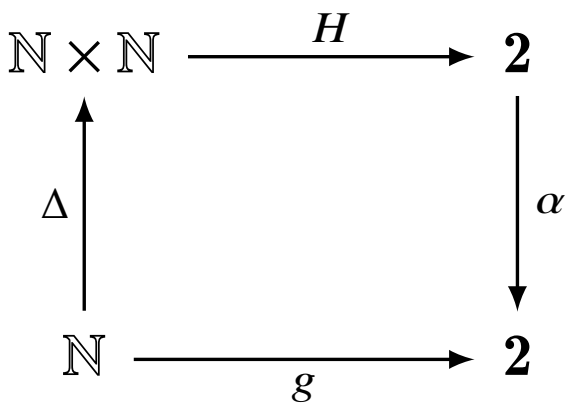
for any partial recursive $\bar{g} \supset g : \bar{g}(\ulcorner \bar{g} \urcorner) \uparrow$.

$$\bar{g}(\ulcorner \bar{g} \urcorner) = \varphi_{\ulcorner \bar{g} \urcorner}(\ulcorner \bar{g} \urcorner) = g(\ulcorner \bar{g} \urcorner) = \varphi_{\ulcorner \bar{g} \urcorner}(\ulcorner \bar{g} \urcorner) + 1$$

Example — Turing's Halting Problem

Theorem (Turing1936)

The Halting problem is unsolvable.



where $H : (m, n) \mapsto [\![m \in W_n]\!]$, and $\alpha : x \mapsto 1 - x$.
However, if H is total computable, then g is total computable and representable by H. Controdiction!

Example — Turing's Halting Problem

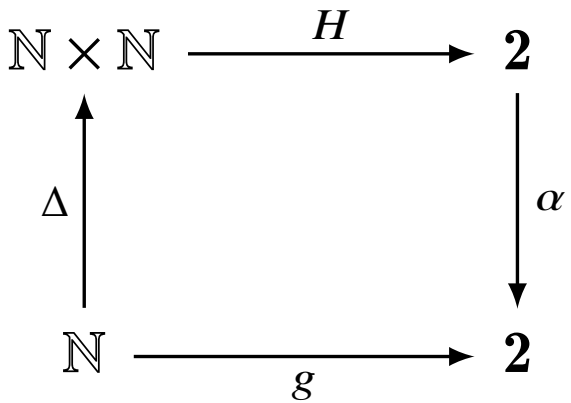
H	W_0	W_1	W_2	W_3	$\cdots$	W_t	$\cdots$
0	0	1	0	1	$\cdots$	1	$\cdots$
1	1	1	0	0	$\cdots$	0	$\cdots$
2	0	1	0	0	$\cdots$	1	$\cdots$
3	0	0	0	1	$\cdots$	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
t	0	0	0	0	$\cdots$	?	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Remark: "Russell Paradox" — The set $\overline{K} = \{n : n \notin W_n\}$ of numbers not belonging to the r.e. sets they code is not r.e. itself $\forall t : \overline{K} \neq W_t$.

Example — Turing's Halting Problem

Theorem (Turing1936)

The Halting problem is unsolvable.



where $H : (m, n) \mapsto [\![\varphi_n(m) \downarrow]\!]$, and $\alpha(x) = \begin{cases} 1 & \text{if } x = 0 \\ \uparrow & \text{if } x = 1 \end{cases}$.

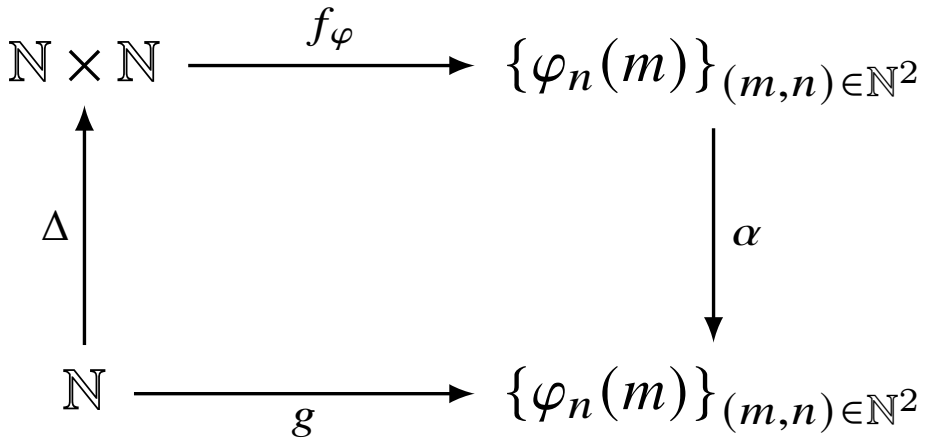
$$H(\ulcorner g \urcorner, \ulcorner g \urcorner) \uparrow$$

There is no perfect anti-virus software.

Example — Turing's Halting Problem

$H(m,n)$	0	1	2	3	$\cdots$	t	$\cdots$
0	1	1	0	0	$\cdots$	0	$\cdots$
1	1	0	1	1	$\cdots$	1	$\cdots$
2	0	1	1	0	$\cdots$	0	$\cdots$
3	0	1	1	0	$\cdots$	1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
t	1	0	1	0	$\cdots$	$\boxed{\uparrow}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Example — Turing's Halting Problem — another proof



where $f_\varphi : (m,n) \mapsto \varphi_n(m)$, and

$$\alpha : \varphi_n(m) \mapsto \begin{cases} 0 & \text{if } H(m,n) = 0 \\ \varphi_n(m) + 1 & \text{if } H(m,n) = 1 \end{cases}$$

or

$$\alpha : \varphi_n(m) \mapsto 1 + \sum_{k=0}^{n} H(m,k) \cdot \varphi_k(m)$$

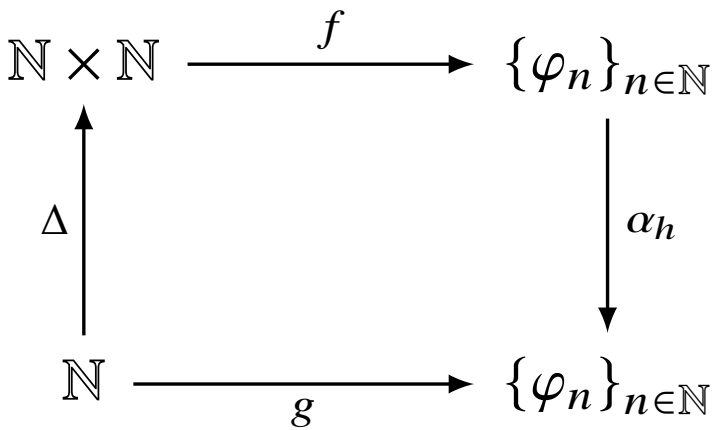
If H is total computable, then g is total computable. ×

Example — Kleene's Fixpoint Theorem

Theorem (Kleene's Fixpoint Theorem)

Given a recursive function h, there is an index e s.t.

$$\varphi_e = \varphi_{h(e)}$$



where $f : (m, n) \mapsto \varphi_{\varphi_n(m)}$, and $\alpha_h : \varphi_n \mapsto \varphi_{h(n)}$.
The function $g : m \mapsto \varphi_{h(\varphi_m(m))}$ is a recursive sequence of partial recursive functions, and thus is representable by $f(-, t)$.

$$e := \varphi_t(t)$$

Explicitly, $g(m) = \varphi_{h(\varphi_m(m))} = \varphi_{s(m)} = \varphi_{\varphi_t(m)} = f(m, t)$

Example — Kleene's Fixpoint Theorem

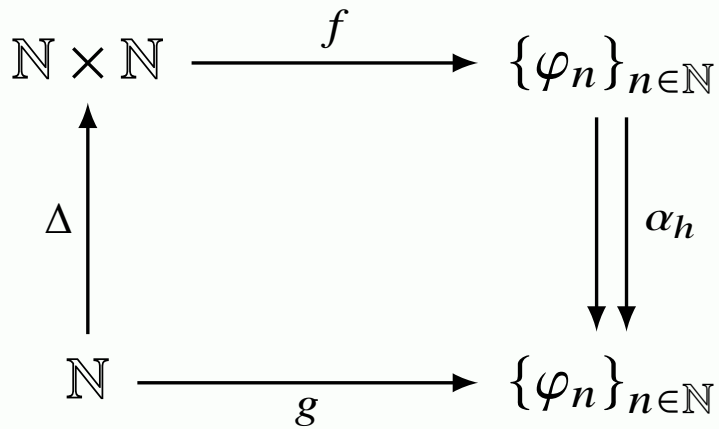
f	0	1	2	3	$\cdots$	t
0	$\varphi_{h(\varphi_0(0))}$	$\varphi_{\varphi_0(1)}$	$\varphi_{\varphi_0(2)}$	$\varphi_{\varphi_0(3)}$	$\cdots$	$\varphi_{\varphi_0(t)}$
1	$\varphi_{\varphi_1(0)}$	$\varphi_{h(\varphi_1(1))}$	$\varphi_{\varphi_1(2)}$	$\varphi_{\varphi_1(3)}$	$\cdots$	$\varphi_{\varphi_1(t)}$
2	$\varphi_{\varphi_2(0)}$	$\varphi_{\varphi_2(1)}$	$\varphi_{h(\varphi_2(2))}$	$\varphi_{\varphi_2(3)}$	$\cdots$	$\varphi_{\varphi_2(t)}$
3	$\varphi_{\varphi_3(0)}$	$\varphi_{\varphi_3(1)}$	$\varphi_{\varphi_3(2)}$	$\varphi_{h(\varphi_3(3))}$	$\cdots$	$\varphi_{\varphi_3(t)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
t	$\varphi_{\varphi_0(0)}$ $\varphi_{h(\varphi_0(0))}$	$\varphi_{\varphi_1(1)}$ $\varphi_{h(\varphi_1(1))}$	$\varphi_{\varphi_2(2)}$ $\varphi_{h(\varphi_2(2))}$	$\varphi_{\varphi_3(3)}$ $\varphi_{h(\varphi_3(3))}$	$\cdots$	$\varphi_{\varphi_t(t)}$ $\varphi_{h(\varphi_t(t))}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Theorem (Kleene's Fixpoint Theorem)

Given a recursive function h, there is an index e s.t.

$$\varphi_e = \varphi_{h(e)}$$

Proof.



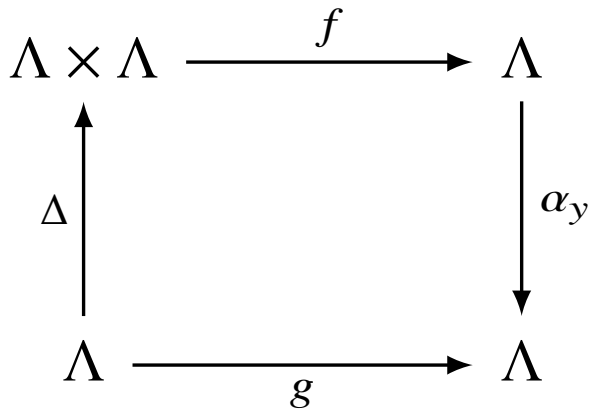
where $f : (m, n) \mapsto \varphi_{\varphi_n(m)}$, and

$$\alpha_h : \varphi_n \mapsto \left\{\varphi_i : \varphi_i = \varphi_{h(n)}\right\}$$

□

Remark: 杀毒软件的功能是通过一个能行的方法 h 系统性地修改病毒 V_n 的代码, 但根据 Kleene's Fixpoint 定理, 存在杀不死的病毒 $V_e = V_{h(e)}$.

Example — $\mathbf{Y}$ Combinator



where $f : (x, y) \mapsto yx$, and $\alpha_y : x \mapsto yx$.

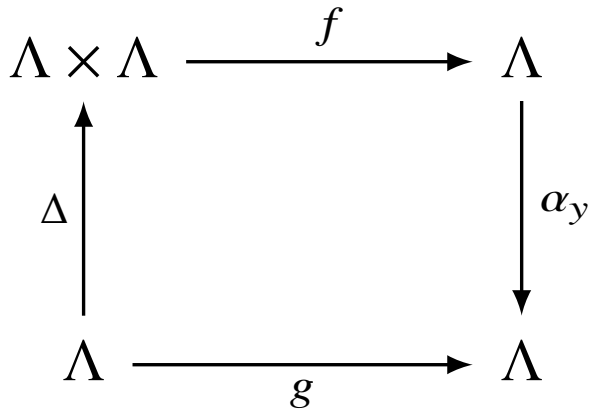
$$g = \lambda x.y(xx)$$

$$gg = \alpha_y(gg)$$

$$\mathbf{Y} \coloneqq \lambda y.gg = \lambda y.(\lambda x.y(xx))(\lambda x.y(xx))$$

$$\mathbf{Y}h = h(\mathbf{Y}h) = h(h(\mathbf{Y}h)) = \cdots$$

Example — $\mathbf{Z}$ Combinator



where $f : (x, y) \mapsto \lambda v.yxv$, and $\alpha_y : x \mapsto yx$.

$$g = \lambda x.y(\lambda v.xxv)$$

$$gg = \alpha_y(gg)$$

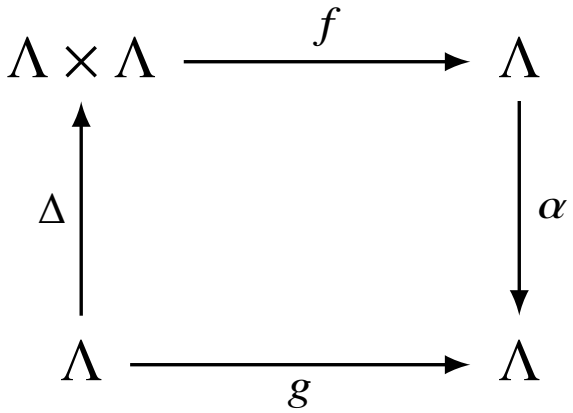
$$\mathbf{Z} := \lambda y.gg = \lambda y.(\lambda x.y(\lambda v.xxv))(\lambda x.y(\lambda v.xxv))$$

$$\mathbf{Z}hv = h(\mathbf{Z}h)v$$

$$e := \mathbf{Z}h \implies ev = hev$$ (Kleene's fixpoint?)

Remark: Quine: $ev = \mathbf{K}ev = e$.

Example — Θ Combinator



where $f : (x, y) \mapsto yx$, and $\alpha : x \mapsto \lambda y.y(xy)$.

$$g = \lambda xy.y(xxy)$$

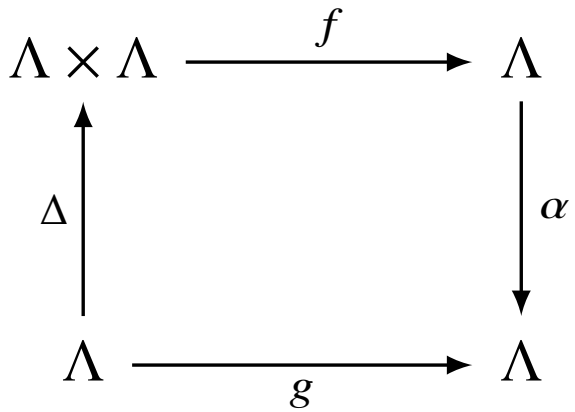
$$gg = \alpha(gg)$$

$$\Theta := gg = (\lambda xy.y(xxy))(\lambda xy.y(xxy))$$

$$\Theta h = h(\Theta h) = h(h(\Theta h)) = \cdots$$

Generally, let $\gamma := \lambda x_1 \ldots x_{n-1} y.y(wy)$ where w is an arbitrary word of length n over the alphabet $\{x_1, \ldots, x_{n-1}\}$. Then $\Gamma := \gamma^n$ is a fixpoint combinator.

Example — Θ_v Combinator



where $f : (x, y) \mapsto yx$, and $\alpha : x \mapsto \lambda y.y(\lambda z.xyz)$.

$$g = \lambda xy.y(\lambda z.xxyz)$$

$$gg = \alpha(gg)$$

$$\Theta_v := gg = (\lambda xy.y(\lambda z.xxyz))(\lambda xy.y(\lambda z.xxyz))$$

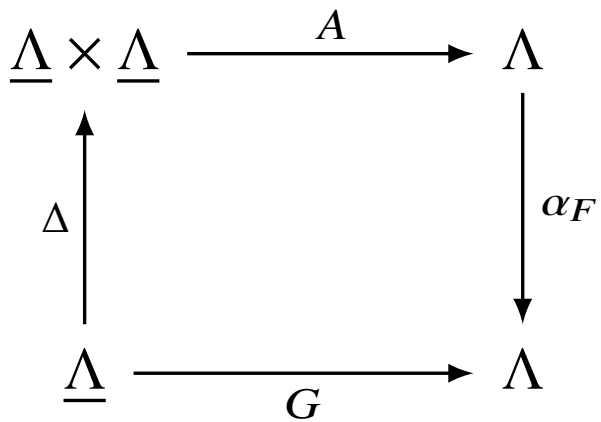
$$\Theta_v hv = h(\Theta_v h)v$$

Example — Fixpoint Theorem in Lambda Calculus

Theorem (Fixpoint Theorem in Lambda Calculus)

For every λ-term F there is a λ-term E s.t.

$$F\ulcorner E\urcorner = E$$



where $\underline{\Lambda} := \{\ulcorner M\urcorner : M \in \Lambda\}$, and $A : (\ulcorner M\urcorner, \ulcorner N\urcorner) \mapsto N\,(\ulcorner M\urcorner)$, and $\alpha_F : M \mapsto F\ulcorner M\urcorner$.

$$G\ulcorner M\urcorner = F\ulcorner M\ulcorner M\urcorner\urcorner$$

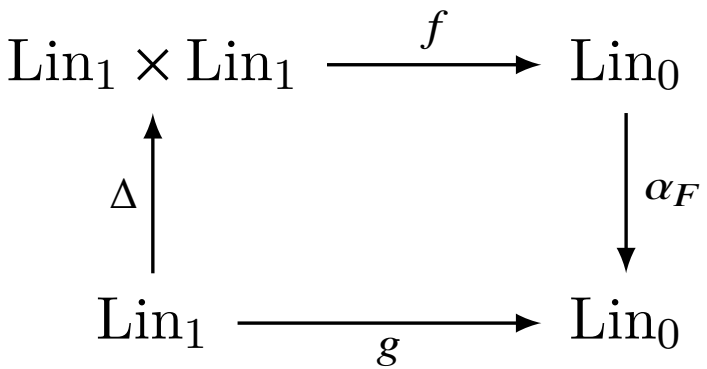
$$E := G\ulcorner G\urcorner$$

Example — Fixpoint Lemma in Logic

Theorem (Fixpoint Lemma in Logic)

For any formula $F(x)$ with one free variable x, there exists a sentence E s.t.

$$\mathrm{Q} \vdash E \leftrightarrow F(\ulcorner E \urcorner)$$



where $f : (M(x), N(x)) \mapsto N(\ulcorner M(x) \urcorner)$, and $\alpha_F : M \mapsto F(\ulcorner M \urcorner)$.

$g(M(x)) = F\big(\ulcorner M(\ulcorner M(x) \urcorner) \urcorner\big)$ which is representable by

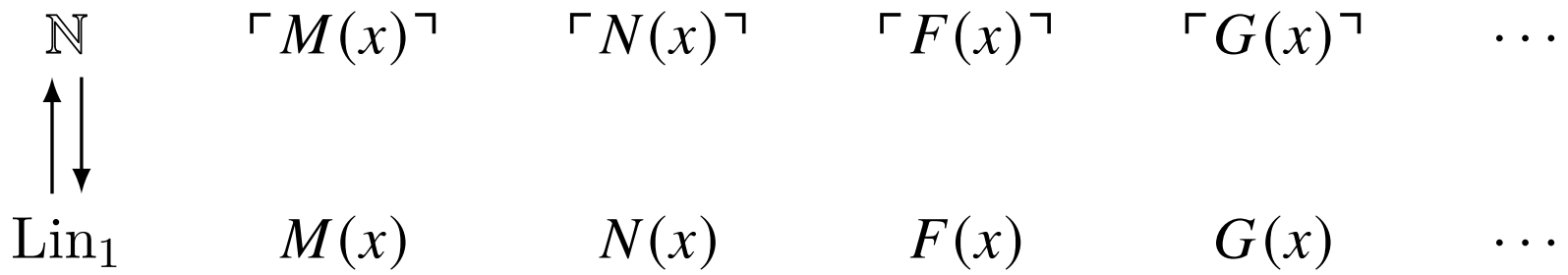
$$G(x) \coloneqq F(d(x))$$

where $d(n) \coloneqq \begin{cases} \#M(\ulcorner M(x) \urcorner) & \text{if } n = \#M(x) \text{ for } M(x) \in \mathrm{Lin}_1 \\ 0 & \text{otherwise} \end{cases}$

is primitive recursive and is thus represented by some function symbol d.

$$E \coloneqq G(\ulcorner G(x) \urcorner)$$

简化一下



- $d(n)$: **用自然数 n 替换Gödel 编码为 n 的命题中的变元所得命题**的Gödel 编码.
- 若 $n = \ulcorner M(x) \urcorner$, 则 $d(n) = \ulcorner M(n) \urcorner$, 即 $d(\ulcorner M(x) \urcorner) = \ulcorner M(\ulcorner M(x) \urcorner) \urcorner$.
- $F(d(\ulcorner M(x) \urcorner)) = F(\ulcorner M(\ulcorner M(x) \urcorner) \urcorner)$
- $G(x) := F(d(x))$
- $E := G(\ulcorner G(x) \urcorner)$

$$E \leftrightarrow G(\ulcorner G(x) \urcorner) \leftrightarrow F(d(\ulcorner G(x) \urcorner)) \leftrightarrow F(\ulcorner G(\ulcorner G(x) \urcorner) \urcorner) \leftrightarrow F(\ulcorner E \urcorner)$$

Example — Fixpoint Lemma in Logic

f	$\ulcorner M(x) \urcorner$	$\ulcorner N(x) \urcorner$	$\cdots$	$\ulcorner G(x) \urcorner$
$M(x)$	$F(\ulcorner M(\ulcorner M(x) \urcorner) \urcorner)$	$M(\ulcorner N(x) \urcorner)$	$\cdots$	$M(\ulcorner G(x) \urcorner)$
$N(x)$	$N(\ulcorner M(x) \urcorner)$	$F(\ulcorner N(\ulcorner N(x) \urcorner) \urcorner)$	$\cdots$	$N(\ulcorner G(x) \urcorner)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$G(x)$	$G(\ulcorner M(x) \urcorner)$ $F(\ulcorner M(\ulcorner M(x) \urcorner) \urcorner)$	$G(\ulcorner N(x) \urcorner)$ $F(\ulcorner N(\ulcorner N(x) \urcorner) \urcorner)$	$\cdots$	$G(\ulcorner G(x) \urcorner)$ $F(\ulcorner G(\ulcorner G(x) \urcorner) \urcorner)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Strong Fixpoint Lemma

Theorem (Strong Fixpoint Lemma — Jeroslow1973)

For any formula $F(x)$ with one free variable x, there is a closed-term t s.t.

$$\mathrm{T} \vdash t = \ulcorner F(t) \urcorner$$

Proof.

$$\begin{array}{ccc} \mathbb{N} \times \mathbb{N} & \xrightarrow{d} & \mathbb{N} \\ {\scriptstyle\Delta}\uparrow & & \downarrow{\scriptstyle\alpha_F} \\ \mathbb{N} & \xrightarrow[g]{} & \mathbb{N} \end{array}$$

where $d(m,n) := \llcorner n \lrcorner(m) := \begin{cases} f(m) & \text{if } n = \ulcorner f \urcorner \text{ for some function } f \\ 0 & \text{otherwise} \end{cases}$ and

$\alpha_F : n \mapsto \ulcorner F(n) \urcorner$.

$$g(n) = \ulcorner F(\llcorner n \lrcorner(n)) \urcorner$$

$$g \text{ is primitive recursive} \implies g = \llcorner k \lrcorner = d(-,k) \text{ for some } k \in \mathbb{N}$$

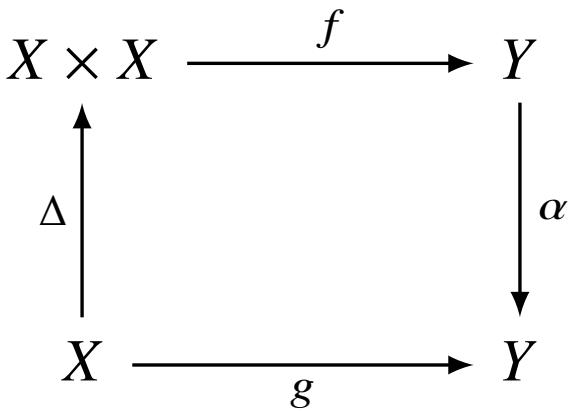
$$t := \llcorner k \lrcorner(k)$$

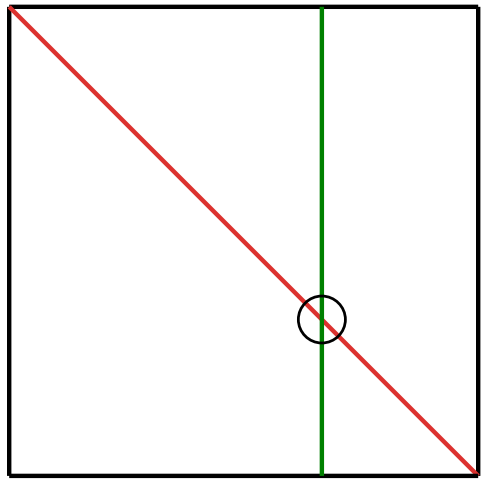
Remark: Strong Fixpoint Lemma $\implies$ Fixpoint Lemma: $E := F(t)$

d	0	1	2	$\cdots$	k	$\cdots$
0	0	0	0	$\cdots$	0	$\cdots$
1	$\llcorner 1 \lrcorner(0)$	$\llcorner 1 \lrcorner(1)$	$\llcorner 1 \lrcorner(2)$	$\cdots$	$\llcorner 1 \lrcorner(k)$	$\cdots$
2	$\llcorner 2 \lrcorner(0)$	$\llcorner 2 \lrcorner(1)$	$\llcorner 2 \lrcorner(2)$	$\cdots$	$\llcorner 2 \lrcorner(k)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	$\llcorner k \lrcorner(0)$	$\llcorner k \lrcorner(1)$	$\llcorner k \lrcorner(2)$	$\cdots$	$\llcorner k \lrcorner(k)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

d	0	1	2	$\cdots$	k	$\cdots$
0	$\ulcorner F(0) \urcorner$	0	0	$\cdots$	0	$\cdots$
1	$\llcorner 1 \lrcorner(0)$	$\ulcorner F(\llcorner 1 \lrcorner(1)) \urcorner$	$\llcorner 1 \lrcorner(2)$	$\cdots$	$\llcorner 1 \lrcorner(k)$	$\cdots$
2	$\llcorner 2 \lrcorner(0)$	$\llcorner 2 \lrcorner(1)$	$\ulcorner F(\llcorner 2 \lrcorner(2)) \urcorner$	$\cdots$	$\llcorner 2 \lrcorner(k)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	$\llcorner k \lrcorner(0)$ $\ulcorner F(0) \urcorner$	$\llcorner k \lrcorner(1)$ $\ulcorner F(\llcorner 1 \lrcorner(1)) \urcorner$	$\llcorner k \lrcorner(2)$ $\ulcorner F(\llcorner 2 \lrcorner(2)) \urcorner$	$\cdots$	$\boxed{\llcorner k \lrcorner(k)}$ $\boxed{\ulcorner F(\llcorner k \lrcorner(k)) \urcorner}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Fixpoint vs Diagonalization



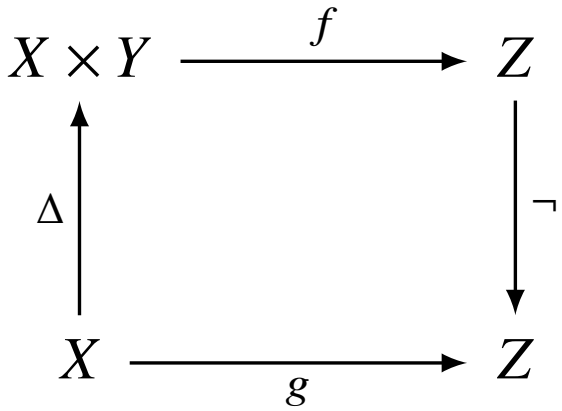


Curry $\mathbf{Y}$	$\hat{=}$	λ-fixpoint	$\hat{=}$	Gödel	$\hat{=}$	Kleene	$\hat{=}$	Russell
yx	$\hat{=}$	$N(\ulcorner M \urcorner)$	$\hat{=}$	$N(\ulcorner M(x) \urcorner)$	$\hat{=}$	$\varphi_n(m)$	$\hat{=}$	$x \in y$
xx	$\hat{=}$	$M(\ulcorner M \urcorner)$	$\hat{=}$	$M(\ulcorner M(x) \urcorner)$	$\hat{=}$	$\varphi_n(n)$	$\hat{=}$	$x \in x$
$y(xx)$	$\hat{=}$	$F\ulcorner M \ulcorner M \urcorner\urcorner$	$\hat{=}$	$F(\ulcorner M(\ulcorner M(x) \urcorner)\urcorner)$	$\hat{=}$	$h(\varphi_n(n))$	$\hat{=}$	$x \notin x$
$\lambda x.y(xx)$	$\hat{=}$	G	$\hat{=}$	$G(x)$	$\hat{=}$	$\varphi_t(n)$	$\hat{=}$	$x \notin R$
$(\lambda x.y(xx))(\lambda x.y(xx))$	$\hat{=}$	$G(\ulcorner G \urcorner)$	$\hat{=}$	$G(\ulcorner G(x) \urcorner)$	$\hat{=}$	$\varphi_t(t)$	$\hat{=}$	$R \notin R$

self-reference $\overset{?}{\Longrightarrow}$ self-improvement

Cantor's Diagonal Argument vs Three-Valued Logic

Let $Z = \{0, u, 1\}$.



z	$\neg z$
0	1
u	u
1	0

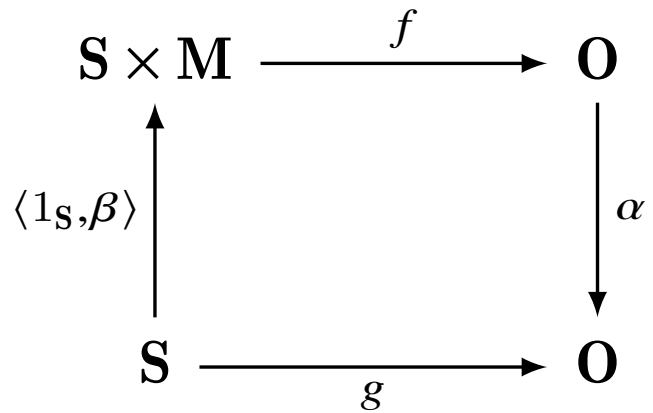
$$\neg f(\ulcorner g \urcorner, \ulcorner g \urcorner) = g(\ulcorner g \urcorner) = f(\ulcorner g \urcorner, \ulcorner g \urcorner) \implies f(\ulcorner g \urcorner, \ulcorner g \urcorner) = u$$

We can't conclude that g is not of the form $f(-, y)$ for any y. Thus Cantor's diagonal argument about higher cardinalities does not generalize to a set theory based on the three-valued logic.

Non-operational Self-inspection[Svo18; Sza18]

The information available to the observer regarding his own state could have absolute limitations, by the laws of nature.

— John von Neumann

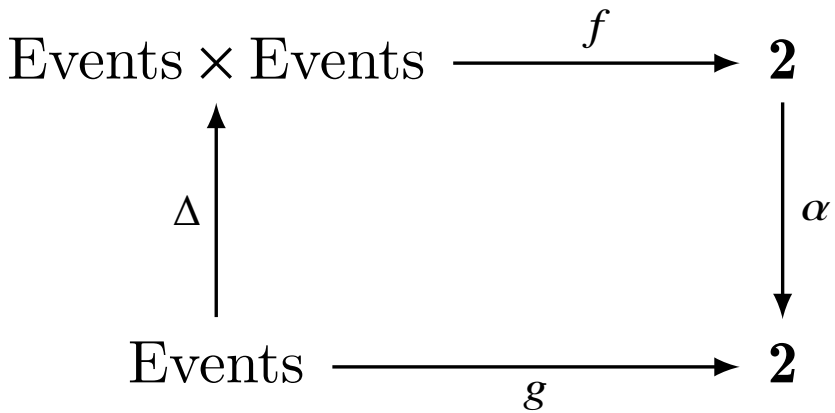


- $\mathbf{S}$: quantum states.
- $\mathbf{M}$: quantum measurements.
- $\mathbf{O}$: possible outcomes of quantum measurements.
- $f(s, m)$: predicts the outcome of measurement m for state s.

If we assume that it is not possible to measure properties without changing them (observer effect: α is fixpoint-free), then there is a limit to self-inspection.

Time Travel Paradox

Can a time traveler go back in time and shoot his bachelor grandfather?



where $f : (x, y) \mapsto [\![x \text{ “is negated by” } y]\!]$ and $\alpha : x \mapsto 1 - x$.

Mind-Blowing Questions

- What would happen if you take a gun back in time? How will the universe stop you?
- A consistent resolution: you are born with $1/2$ probability, and if you are born you go back in time to kill your grandfather, therefore you are born with $1/2$ probability.

General Fixpoint Theorem[San20]

$$\forall XYZ.\forall R \subset Z \times Z.\forall R' \subset Z^X \times Z^{X\times Y} \times X \times Y.\forall \beta : X \twoheadrightarrow Y.\forall d : Z^{X\times Y} \to Z^X.\Big\{\Big(\exists f \in Z^{X\times Y}.\exists y \in Y.\forall x \in X.R'\big(df, f, x, y\big)\Big) \wedge \Big(\forall g \in Z^X.\forall f \in Z^{X\times Y}.\forall x \in X.\big[R'(g, f, x, \beta x) \to R(gx, dfx)\big]\Big) \to \exists z \in Z.R(z, z)\Big\}$$

Theorem (General Fixpoint Theorem)

In a category $\mathbf{C}$, *for* $X, Y, Z \in \mathrm{ob}(C)$, $R \subset \mathrm{Hom}(1, Z) \times \mathrm{Hom}(1, Z)$, $R' \subset \mathrm{Hom}(X, Z) \times \mathrm{Hom}(X \times Y, Z) \times \mathrm{Hom}(1, X) \times \mathrm{Hom}(1, Y)$, $\beta : \mathrm{Hom}(1, X) \twoheadrightarrow \mathrm{Hom}(1, Y)$, *and* $d : \mathrm{Hom}(X \times Y, Z) \to \mathrm{Hom}(X, Z)$, *if*

1. *there exists* $f : X \times Y \to Z$ *and* $y : 1 \to Y$ *s.t for all* $x : 1 \to X$:
$$R'(df, f, x, y)$$
2. *for all* $g : X \to Z$, $f : X \times Y \to Z$, *and* $x : 1 \to X$,
$$R'(g, f, x, \beta x) \implies R(gx, dfx)$$

then, there exists $z : 1 \to Z$ *s.t.* $R(z, z)$.

General Fixpoint Theorem ⇒ Smullyan's Fixpoint Theorem

Theorem (Smullyan's Fixpoint Theorem)

For $R, R' \subset Z \times Z$, and $\alpha : Z \to Z$, if

1. *there exists $x \in Z$ s.t. $R'(\alpha x, x)$, and*
2. *for $x, y \in Z$, $R'(x, y) \implies R(x, \alpha y)$.*

then, there exists $z \in Z$ s.t. $R(z, z)$.

Proof.

Let $d : Z^{Z \times Z} \to Z^Z :: f \mapsto \alpha \circ f \circ \langle 1_Z, \beta \rangle$.
Let $R'_1(g, f, x, y) := R'(gx, f(x, y))$.

$$\begin{array}{ccc} X \times Y & \xrightarrow{f} & Z \\ {\scriptstyle \langle 1_X, \beta \rangle}\uparrow & & \downarrow{\scriptstyle \alpha} \\ X & \longrightarrow & Z \end{array}$$

$$\begin{aligned} R'(\alpha x, x) &\implies R'(\alpha \circ \pi_1 \circ \langle 1_Z, \beta \rangle x, x) \implies R'_1(d\pi_1 x, \pi_1(x, y)) \\ &\implies R'_1(d\pi_1, \pi_1, x, y) \\ R'_1(g, f, x, \beta x) &\implies R'(gx, f(x, \beta x)) \implies R(gx, \alpha(f(x, \beta x))) \\ &\implies R(gx, \alpha \circ f \circ \langle 1_Z, \beta \rangle x) \implies R(gx, dfx) \end{aligned}$$

General Fixpoint Theorem ⇒ Lawvere's Fixpoint Theorem

Theorem (Lawvere's Fixpoint Theorem)

For all sets X, Y, Z, and all functions $f : X \times Y \to Z, \alpha : Z \to Z$, surjective functions $\beta : X \twoheadrightarrow Y$. If $\alpha \circ f \circ \langle 1_X, \beta \rangle$ is representable by f, then α has a fixpoint.

$$\begin{array}{ccc} X \times Y & \xrightarrow{f} & Z \\ {\scriptstyle \langle 1_X, \beta \rangle}\uparrow & & \downarrow{\scriptstyle \alpha} \\ X & \longrightarrow & Z \end{array}$$

Proof.

$$d(f) \coloneqq \alpha \circ f \circ \langle 1_X, \beta \rangle$$

$$R'(g, f, x, y) \coloneqq gx = f(x, y)$$

$$R(x, y) \coloneqq \alpha x = y$$

Since df is representable by $f(-, y)$, we have $R'(df, f, x, y)$, and

$$\begin{aligned} R'(g, f, x, \beta x) &\implies gx = f(x, \beta x) \implies \alpha(gx) = (\alpha \circ f \circ \langle 1_X, \beta \rangle)x \\ &\implies R(gx, dfx) \end{aligned}$$

General Fixpoint Theorem ⇒ Lawvere's Fixpoint Theorem

Theorem (Lawvere's Fixpoint Theorem — multi-valued version)

For sets X, Y, Z, multi-valued functions $f : X \times Y \rightrightarrows Z$, $\alpha : Z \rightrightarrows Z$, and surjective functions $\beta : X \twoheadrightarrow Y$, if $\exists y \forall x : \alpha \circ f \circ \langle 1_X, \beta \rangle x \cap f(x, y) \neq \emptyset$, then α has a fixpoint, i.e. $\exists z \in \alpha(z)$.

$$\begin{array}{ccc} X \times Y & \overset{f}{\rightrightarrows} & Z \\ {\scriptstyle \langle 1_X, \beta \rangle}\uparrow & & \downarrow\downarrow{\scriptstyle \alpha} \\ X & \longrightarrow & Z \end{array}$$

Proof.

$$\begin{aligned} d(f) &:= \alpha \circ f \circ \langle 1_X, \beta \rangle \\ R'(g, f, x, y) &:= gx \cap f(x, y) \neq \emptyset \\ R(x, y) &:= \alpha x \cap y \neq \emptyset \end{aligned}$$

Obviously, we have $R'(df, f, x, y)$, and

$$\begin{aligned} R'(g, f, x, \beta x) &\implies gx \cap f(x, \beta x) \neq \emptyset \implies \alpha(gx) \cap (\alpha \circ f \circ \langle 1_X, \beta \rangle)x \neq \emptyset \\ &\implies R(gx, dfx) \end{aligned}$$

Smullyan's Fixpoint Theorem $\Rightarrow$ Tarski's Fixpoint Theorem

Theorem (Tarski's Fixpoint Theorem)

Given a set V, and $(\mathrm{P}(V), \subset)$, and an order-preserving function $f : \mathrm{P}(V) \to \mathrm{P}(V)$, f has a fixpoint.

Proof.

Define $R, R' \subset \mathrm{P}(V) \times \mathrm{P}(V)$ and $\alpha : \mathrm{P}(V) \to \mathrm{P}(V)$ as follows.

$$R(X,Y) \coloneqq f(X) = Y$$

$$R'(X,Y) \coloneqq \left(\forall x \in Y . x \subset f(x)\right) \wedge \left(\bigcup Y \subset X\right) \wedge \left(\bigcup Y \in Y\right) \wedge \left(f(X) \subset \bigcup Y\right)$$

$$\alpha(X) \coloneqq \bigcup X$$

Consider $S \coloneqq \left\{X \in \mathrm{P}(V) : X \subset f(X)\right\}$. It is easy to check that $R'(\alpha S, S)$.

$$R'(X,Y) \implies f(X) = \bigcup Y = \alpha Y \implies R(X, \alpha Y)$$

By Smullyan's theorem, $R(X,Y)$ has a fixpoint. $\square$

Definition

A function $f : X \to X$ is preserved in $\mathcal{A} \subset X^\omega$ iff for each $x \in \mathcal{A}$, $\langle f(x_n) \rangle_{n \in \omega} \in \mathcal{A}$.

Definition

A function $F : \mathcal{A} \to X$ is a limit function iff

1. $\forall x \in \mathcal{A} : \langle x_{n+1} \rangle_{n \in \omega} \in \mathcal{A}$;
2. $\forall x \in \mathcal{A} : F(x) = F(\langle x_{n+1} \rangle_{n \in \omega})$.

Definition

A function $f : X \to X$ is continuous with respect to a limit function $F : \mathcal{A} \to X$ iff f is preserved in $\mathcal{A}$, and

1. $\exists a \in X : \langle f^n(a) \rangle_{n \in \omega} \in \mathcal{A}$;
2. $\forall x \in \mathcal{A} : f(F(x)) = F(\langle f(x_n) \rangle_{n \in \omega})$.

Theorem (Fixpoint for Continuous Function w.r.t. Limit Function)

Let $f : X \to X$ be a continuous function with respect to a limit function $F : \mathcal{A} \to X$. Then f has a fixpoint.

Proof.

Define $R, R' \subset \mathcal{A} \times \mathcal{A}$ and $\alpha : \mathcal{A} \to \mathcal{A}$ as follows.

$$R(x, y) \coloneqq f(F(x)) = F(y)$$
$$R'(x, y) \coloneqq f(F(x)) = F(\langle f(y_n) \rangle_{n \in \omega})$$
$$\alpha(x) \coloneqq \langle f(x_n) \rangle_{n \in \omega}$$

Take a in $\exists a \in X : \langle f^n(a) \rangle_{n \in \omega} \in \mathcal{A}$. Consider z given by $z_n \coloneqq f^n(a)$.
$f(F(\alpha z)) = f(F(\langle f(z_n) \rangle_{n \in \omega})) = F(\langle f(f(z_n)) \rangle_{n \in \omega}) = F(\langle f^{n+2}(a) \rangle_{n \in \omega}) =$
$F(\langle f^{(n+1)+1}(a) \rangle_{n \in \omega}) = F(\langle f^{n+1}(a) \rangle_{n \in \omega}) = F(\langle f(z_n) \rangle_{n \in \omega})$
Hence $R'(\alpha z, z)$.

$$R'(x, y) \implies f(F(x)) = F(\langle f(y_n) \rangle_{n \in \omega}) \implies f(F(x)) = F(\alpha y) \implies R(x, \alpha y$$

By Smullyan's theorem, $R(x, y)$ has a fixpoint. □

Fixpoint Theorem for Normal Functions

Definition

A function $f : \mathrm{Ord} \to \mathrm{Ord}$ is a normal function iff

1. $\alpha < \beta \implies f(\alpha) < f(\beta)$.
2. $f(\alpha) = \sup\{f(\gamma) : \gamma < \alpha\}$ if α is a limit ordinal.

Theorem (Fixpoint Theorem for Normal Functions)

Every normal function $f : \mathrm{Ord} \to \mathrm{Ord}$ has a fixpoint.

Proof.

Let $\mathcal{A} := \{x \in \mathrm{Ord}^\omega : x \text{ is increasing}\}$. Obviously, f is preserved in $\mathcal{A}$.
Let $F : \mathcal{A} \to \mathrm{Ord} :: x \mapsto \sup\{x_n : n \in \omega\}$. Obviously, F is a limit function.
Take $\alpha \in \mathrm{Ord}$. Obviously, $\langle f^n(\alpha)\rangle_{n\in\omega} \in \mathcal{A}$.
For $x \in \mathcal{A}$,

$$f(F(x)) = f(\sup\{x_n : n \in \omega\}) = \sup\{f(x_n) : n \in \omega\} = F(\langle f(x_n)\rangle_{n\in\omega})$$

Banach's Fixpoint Theorem

Theorem (Banach's Fixpoint Theorem)

Let (X, d) be a complete metric space and $T : X \to X$ be a contraction mapping, with Lipschitz constant $\gamma < 1$. Then T has a unique fixpoint.

Proof.

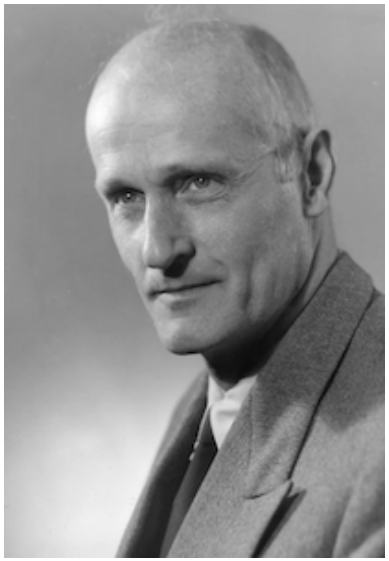
Let $\mathcal{A} := \{x \in X^\omega : x \text{ is a Cauchy sequence}\}$.
For $x \in \mathcal{A}$, it is easy to check that $\langle T(x_n)\rangle_{n\in\omega} \in \mathcal{A}$.
Take $F : \mathcal{A} \to X :: x \mapsto \lim_{n\in\omega} x_n$. Obviously, F is a limit function.
Let us prove that T is continuous with respect to F.
Let $a \in X$. Consider x given by $x_n := T^n(a)$.

$$d(x_{n+1}, x_n) \leq \gamma^n d(x_1, x_0)$$

It is not hard to check that x is a Cauchy sequence, i.e., $x \in \mathcal{A}$.
As the contraction mapping T is continuous, we have

$$T(F(x)) = T \lim_{n\in\omega} x_n = \lim_{n\in\omega} T(x_n) = F(\langle T(x_n)\rangle_{n\in\omega})$$

Kleene's Fixpoint Theorem



Theorem (Second Recursion Theorem)

If $f(x, y)$ is a partial recursive function, there is an index e s.t.

$$\varphi_e(y) = f(e, y)$$

Remark: 对于任意的程序 h, 总存在某个程序 e, 执行程序 e 的结果等价于把程序 e 当作数据输入给程序 h 执行的结果 $[\![e]\!](-) = [\![h]\!](e, -)$.

Theorem (Kleene's Fixpoint Theorem)

Given a recursive function h, there is an index e s.t.

$$\varphi_e = \varphi_{h(e)}$$

Remark: You can systematically change an infinite number of programs $n \mapsto h(n)$ but you cannot systematically change an infinite number of recursive functions $\varphi_e = \varphi_{h(e)}$.

From Kleene's Fixpoint to Chaitin's Incompleteness

Definition: Kolmogorov Complexity $K(x) \coloneqq \mu e[\varphi_e(0) = x]$

Theorem (Chaitin's Incompleteness Theorem)

For any arithmetically sound Gödelian theory T, $\exists c \forall x : \mathrm{T} \nvdash K(x) > c$.

Proof.

For any m, we can construct:

$$M_n \coloneqq \text{“find } \mu y\big[\,\mathrm{prf}_{\mathrm{T}}\left(y, K(x) > m\right)\big], \text{output } x\text{”}$$

So there exists a recursive function $f : m \mapsto n$.
By Kleene's fixpoint theorem, there exists e such that

$$M_e = M_{f(e)} = \text{“find } \mu y\big[\,\mathrm{prf}_{\mathrm{T}}\left(y, K(x) > e\right)\big], \text{output } x\text{”}$$

Take $c \coloneqq e$. □

Remark: For almost all random strings their randomness cannot be proved.

Self-Reproducing Program/Quine

There is a program that outputs its own length.

There is a program that outputs its own source code.

- A Quine is a program which takes no input and outputs its own source code.
- Quines are algorithmic random.

Corollary (Self-Reproducing Program)

There is a recursive function φ_e *s.t.* $\forall x : \varphi_e(x) = e$.

Quine in Python

```
s='s=%r; print(s%%s)'; print(s%s)
```

Quine in Lambda Calculus

$$(\lambda x.xx)(\lambda x.xx)$$

Self-Reproducing Program

Print two copies of the following, the second copy in quotes:

"Print two copies of the following, the second copy in quotes:"

DNA / mutation / evolution

Build a baby that acts on the following instructions, and also contains a copy of those instructions in its reproductive organs.

"Build a baby that acts on the following instructions, and also contains a copy of those instructions in its reproductive organs."

von Neumann's Self-Reproducing Automata

1. A universal constructor A.

$$A + \ulcorner X \urcorner \rightsquigarrow X$$

2. A copying machine B.

$$B + \ulcorner X \urcorner \rightsquigarrow \ulcorner X \urcorner$$

3. A control machine C, which first activates B, then A.

$$A + B + C + \ulcorner X \urcorner \rightsquigarrow X + \ulcorner X \urcorner$$

4. Let $X := A + B + C$. Then $A + B + C + \ulcorner A + B + C \urcorner$ is self-reproducing.

$$A + B + C + \ulcorner A + B + C \urcorner \rightsquigarrow A + B + C + \ulcorner A + B + C \urcorner$$

5. It is possible to add the description of any machine D.

$$A + B + C + \ulcorner A + B + C + D \urcorner \rightsquigarrow A + B + C + D + \ulcorner A + B + C + D \urcorner$$

6. Now allow mutation on the description $\ulcorner A + B + C + D \urcorner$.

$$A + B + C + \ulcorner A + B + C + D' \urcorner \rightsquigarrow A + B + C + D' + \ulcorner A + B + C + D' \urcorner$$

Introspective Program

Definition (ψ-introspective)

Given a total recursive function ψ,

- the *ψ-analysis* of $\varphi(x)$ is the code of the computation of $\varphi(x)$ to $\psi(x)$ steps.
- φ is *ψ-introspective* at x iff $\varphi(x)\downarrow$ and outputs its own ψ-analysis.
- φ is *totally ψ-introspective* iff it is ψ-introspective at all x.

Corollary

There is a program that is totally ψ-introspective.

Proof.

Let $f(n,x) :=$ "the ψ-analysis of $\varphi_n(x)$". □

Introspective Program

There is a program that is totally introspective.

$$\varphi_e = \varphi_{h(e)}$$

Self-simulating Computer	Self-consciousness
Host Machine	Experiencing Self
Virtual Machine	Remembering Self
Hardware	Body



Who am I?

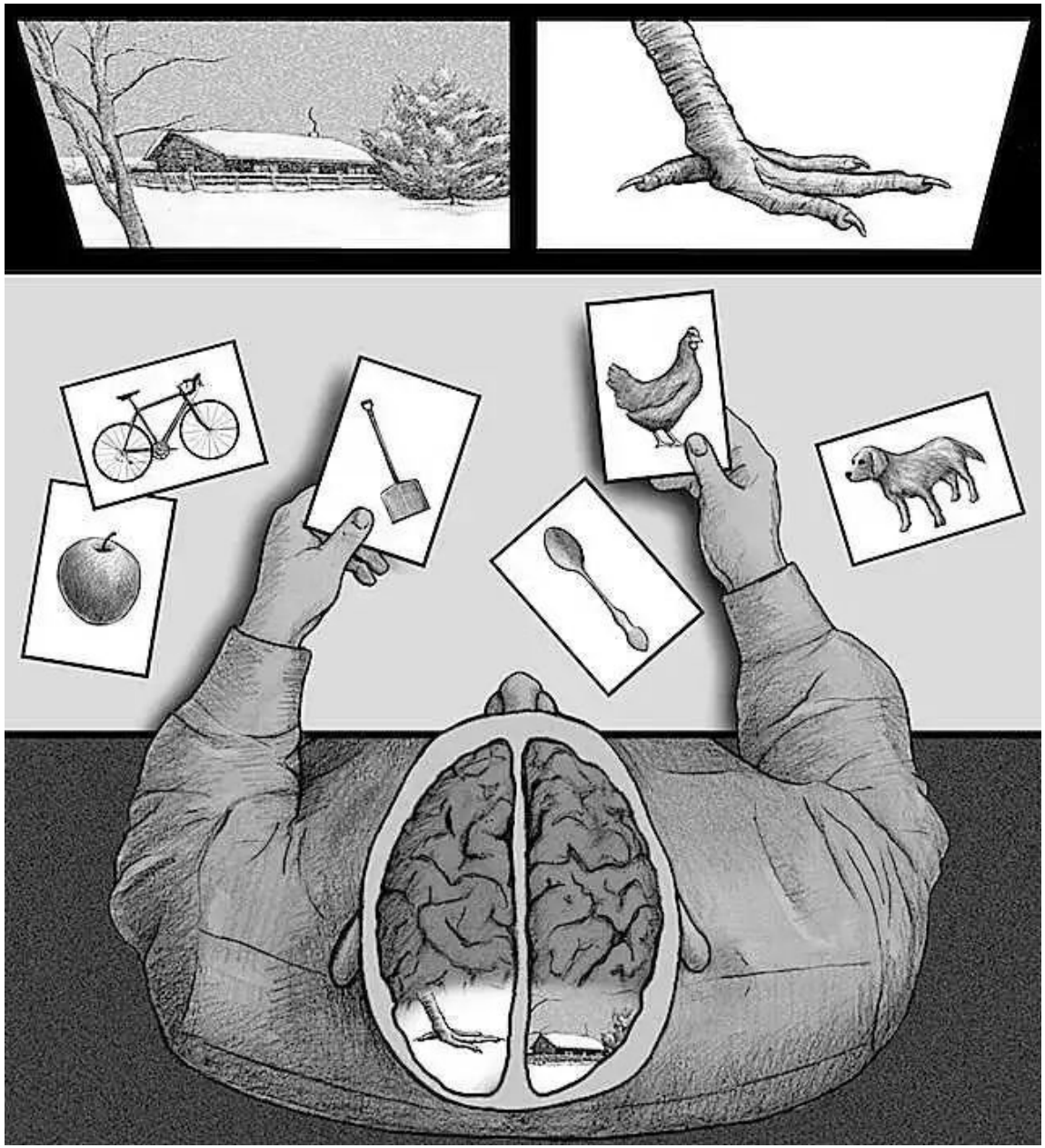
I think, therefore I am.

self-locating: "I" is an indexical term that I use to refer to myself as myself.

What is "me"?

What is "self-consciousness"?

- self-perception self-observation self-experience self-tracking self-reflection self-awareness
- self-evaluation self-analysis self-monitoring
- self-control self-adjustment self-modification self-actualization self-fulfillment self-surpass self-improvement
- *actual-self* pk *ideal-self* self-identity "the *self*"
- free will: Second order desire that we want to act on is second order volition. Second order volitions involve wanting a certain desire to be one's will, that is wanting it to move one to action. (Frankfurt)



- the split brain in man
- snow?
- shit!
- life as a story

Kahneman — Thinking, Fast and Slow

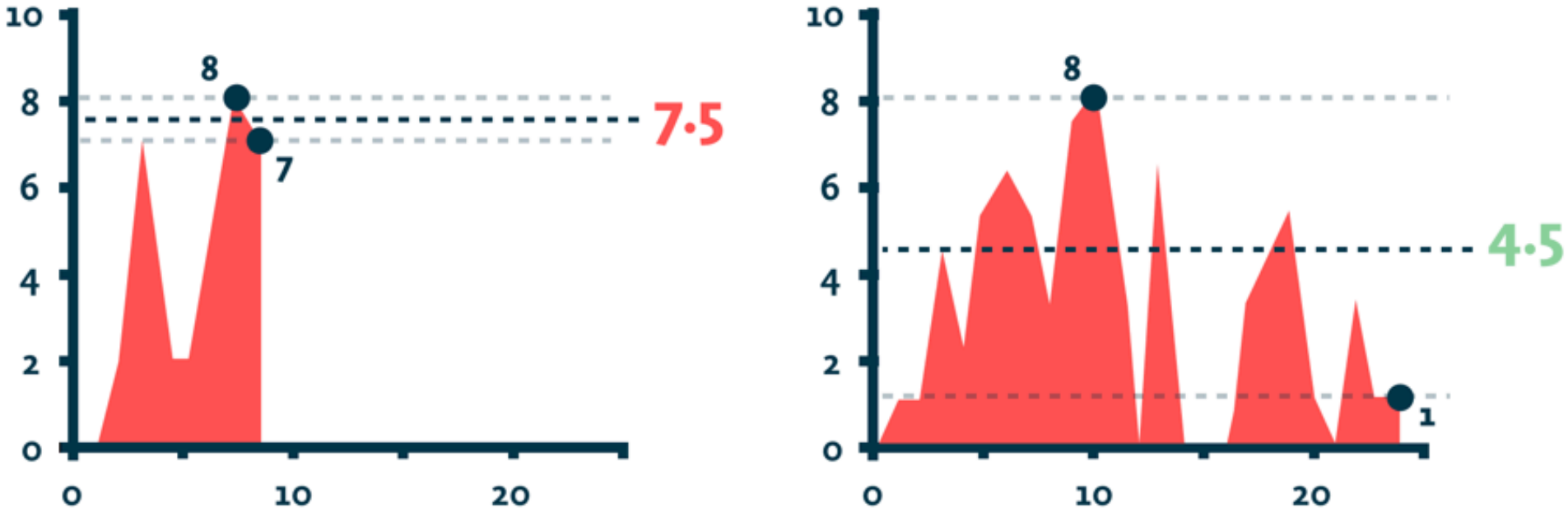


Figure: Why you might prefer more pain

- painful experiment
- experiencing self
- remembering self
- duration neglect
- peak-end rule



Figure: One can imagine a detailed floor plan of a room, sitting on a table in the room; this plan has an image of the table on which there is an image of the plan itself. Now introduce the dynamical aspect: the items on the plan are cut out from paper and can be moved to try a different furniture arrangement; in this way the plan models possible states of the world about which it carries information.

Manin — Cognitive Networks



The brain contains inside a map of itself, and some neural information channels in the central neural system:

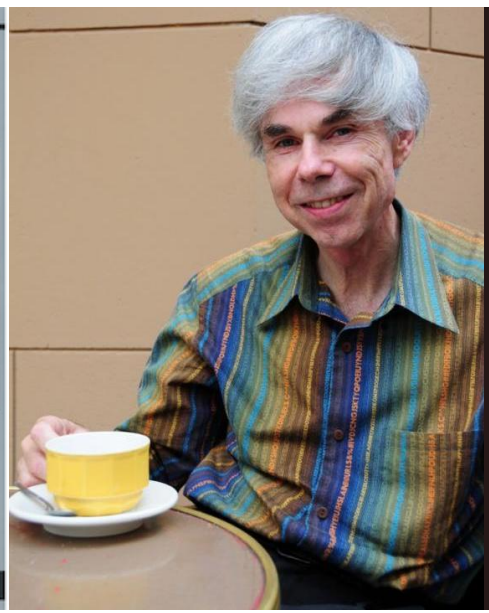
- carry information about the mind itself, i.e. are reflexive;
- are capable of modelling states of the mind different from the current one, i.e. possess a modelling function;
- can influence the state of the whole mind and through that, the behavior, i.e. possess controlling function.

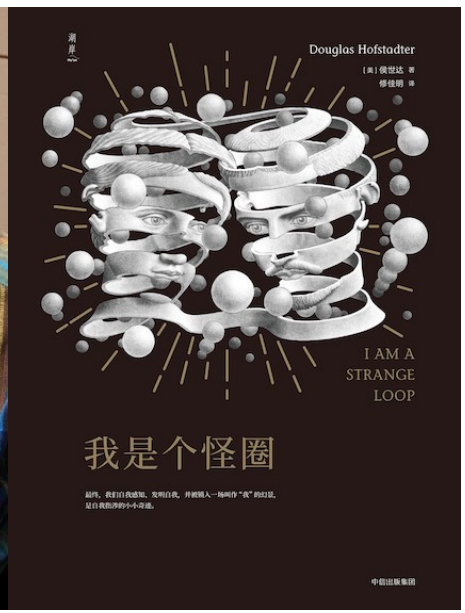
The reflection of the brain inside itself must be coarse grained.

侯士达 — 《我是个怪圈》

- 有没有意识取决于在哪个层级上对结构进行观察. 在整合度最高的层级上看, 大脑是有意识的. 下降到微观粒子层面, 意识就不见了.
- 意识体是那些在某个描述层级上表现出某种特定类型的循环回路的结构. 当一个系统能把外部世界过滤成不同的范畴、并不断向越来越抽象的层级创造新的范畴时, 这种循环回路就会逐渐形成.
- 当系统能进行自我表征 — 对自己讲故事 — 的时候, 这种循环回路就逐渐变成了实体的 "我" — 一个统一的因果主体.







说谎者悖论	**我在说谎**
Grelling 悖论	“非自谓的” 是自谓的吗
Russell 悖论	“不属于自身的集合的集合” 属于自身吗
Berry 悖论	我是少于十八个字不可定义的最小数
Yablo 悖论	我下一句及后面所有的句子都是假的
Gödel 不动点引理	**我有性质** F
Tarski 算术真不可定义定理	我不真
Gödel 第一不完备性定理	我不可证
Gödel-Rosser 不完备性定理	对于任何一个关于我的证明，都有一个更短的关于我的否定的证明
Löb 定理	如果我可证，那么 A
Curry 悖论	如果我是真的，那么上帝存在
Parikh 定理	我没有关于自己的长度短于 n 的证明
Kleene 不动点定理	**我要进行** h **操作**
Quine 悖论	把 “把中的第一个字放到左引号前面，其余的字放到右引号后面，并保持引号及其中的字不变得到的句子是假的” 中的第一个字放到左引号前面，其余的字放到右引号后面，并保持引号及其中的字不变得到的句子是假的
自测量长度程序	我要输出自己的长度
自复制程序	我要输出自己
自反省程序	我要回顾自己走过的每一步
Gödel 机	我要变成能获取更大效用的自己

Schmidhuber's Gödel Machine

- The Gödel machine consists of a **Solver** and a **Searcher** running in parallel.
- The **Solver** (AIXI^S/AIXI$^{t\ell}$) interacts with the environment.
- The **Searcher** (LSEARCH/HSEARCH/OOPS) searches for a proof of "the modification of the software — including the *Solver* and *Searcher* — will increase the expected utility than leaving it as is".
- Logic: a theorem prover and a set of self-referential axioms, which include a description of its own software and hardware, and a description of the probabilistic properties of the environment, as well as a user-given utility function.
- *Since the utility of "leaving it as is" implicitly evaluates all possible alternative modifications, the current modification is globally optimal w.r.t. its initial utility function.*

Gödel Machine

- language $\mathscr{L} := \{\neg, \wedge, \vee, \rightarrow, \forall, \exists, =, (,), \ldots, +, -, \cdot, /, <, \ldots\}$
- well-formed formula
- utility function $u(s, e) = \mathbb{E}_\mu \left[\sum_{t=1}^{T} r_t \,\middle|\, s, e \right]$
- target theorem

  $$u\left[s(t) \oplus \left(\texttt{switchbit}(t) = 1\right), e(t)\right] > u\left[s(t) \oplus \left(\texttt{switchbit}(t) = 0\right), e(t)\right]$$

- theorem prover

  hardware, costs, environment, initial state, utility, logic/arithmetic/probability

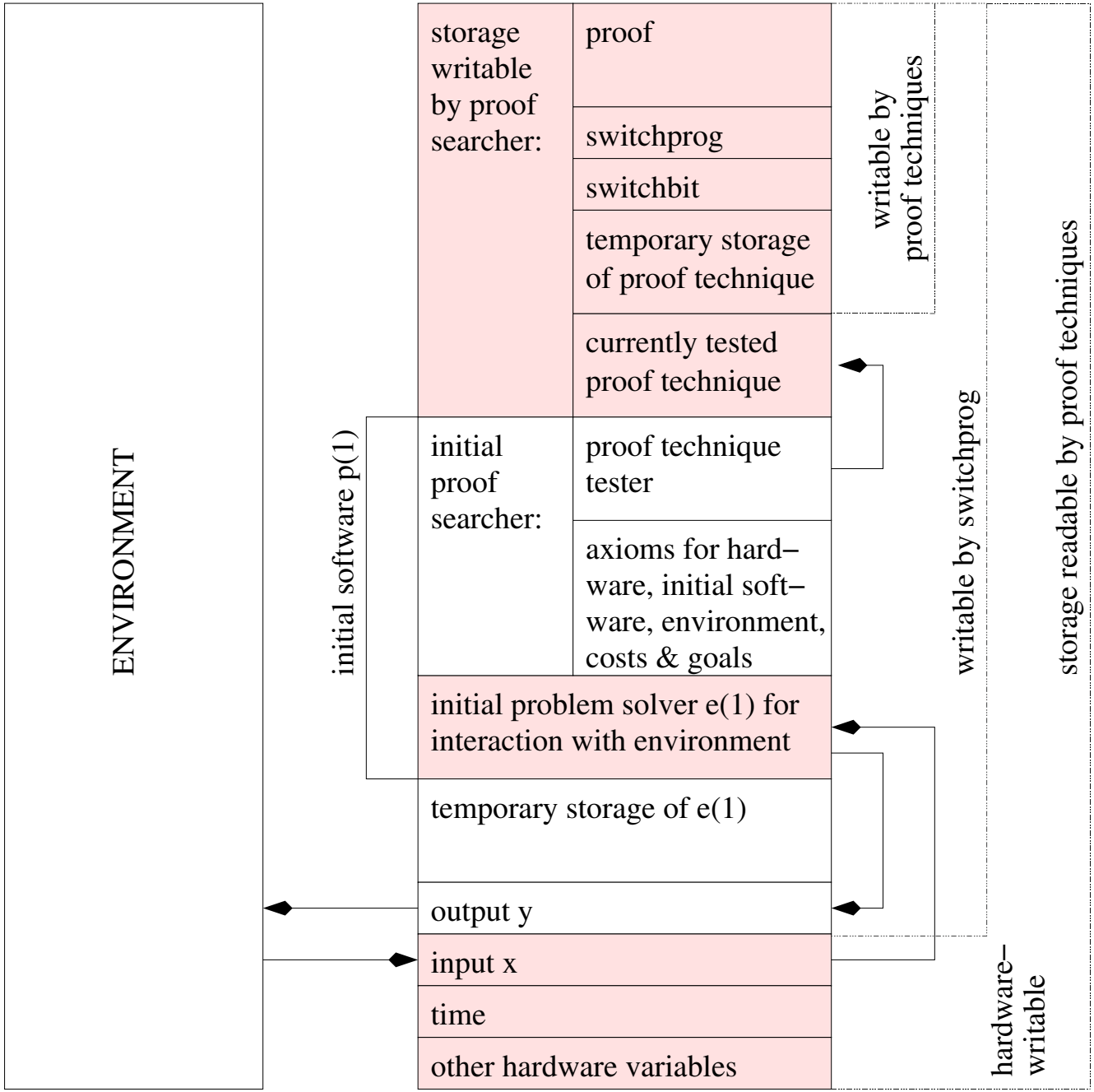
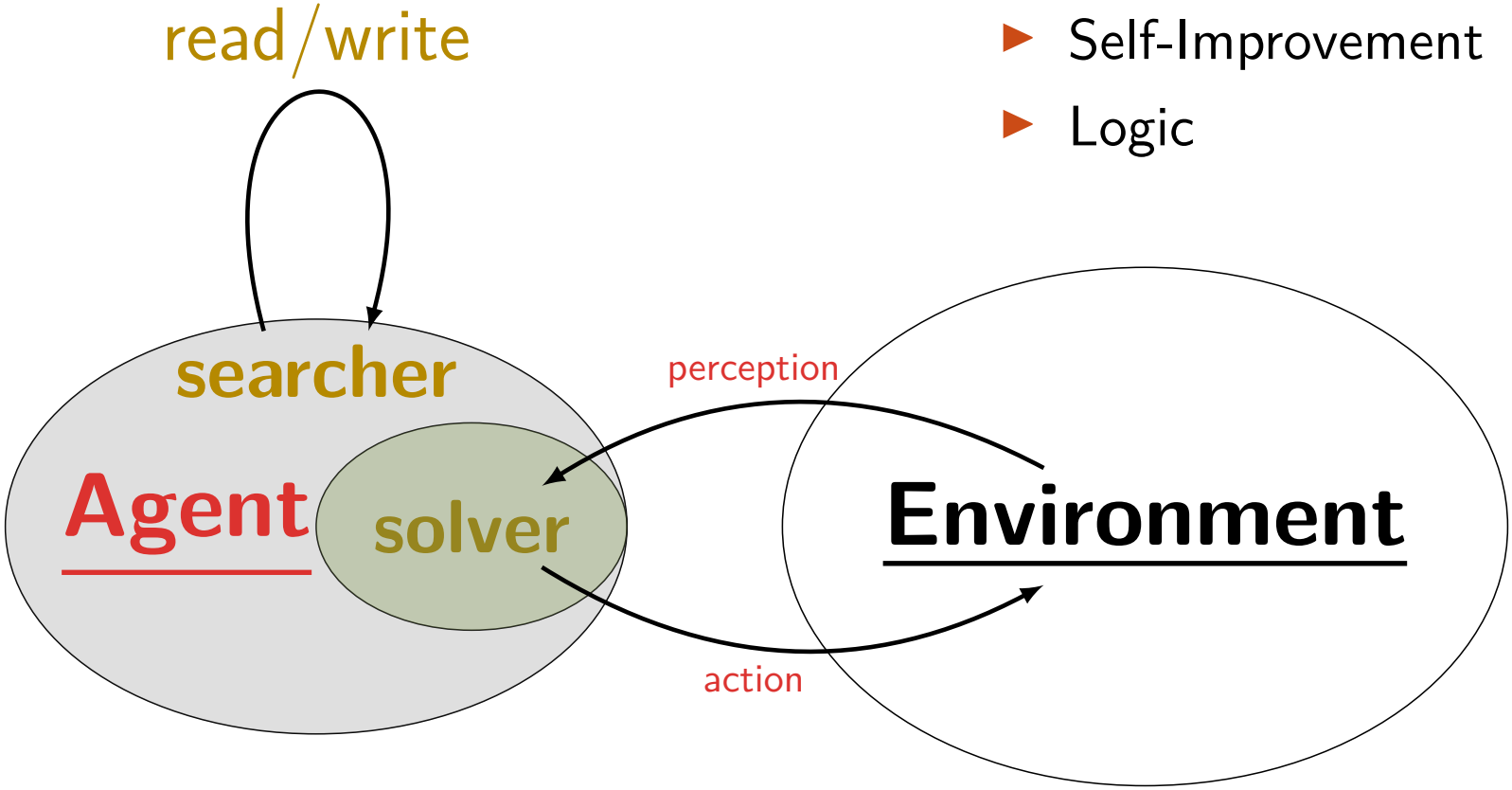




Figure: Schmidhuber

Gödel Machine

- GRL
- Universal Search
- Self-Improvement
- Logic



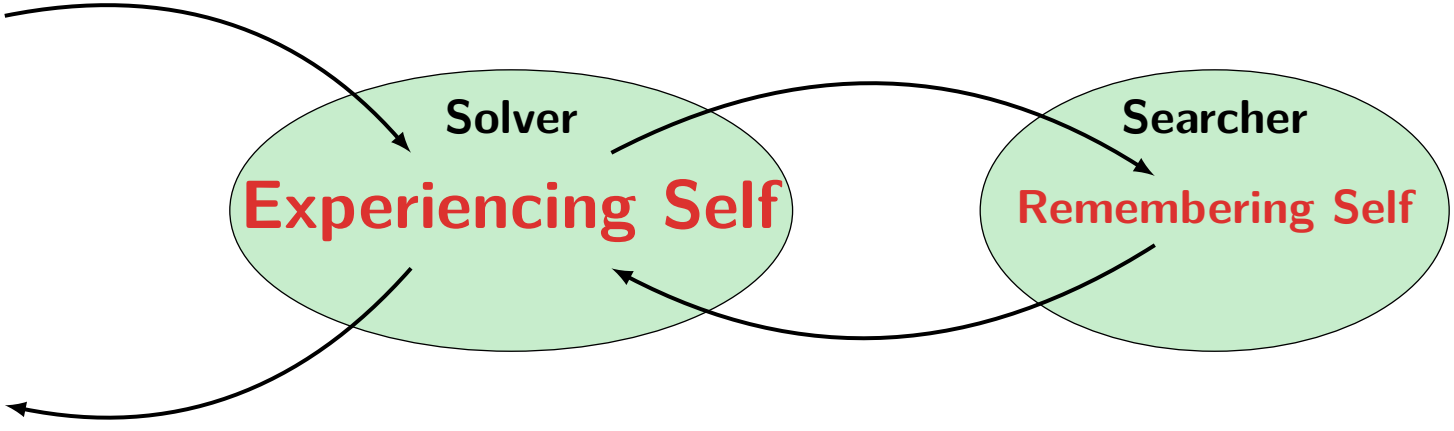
Disadvantage: A Gödel Machine with a badly chosen utility function is motivated to converge to a "poor" program. (goal orthogonality!)

Gödel Machine vs Self-Consciousness vs Free Will?

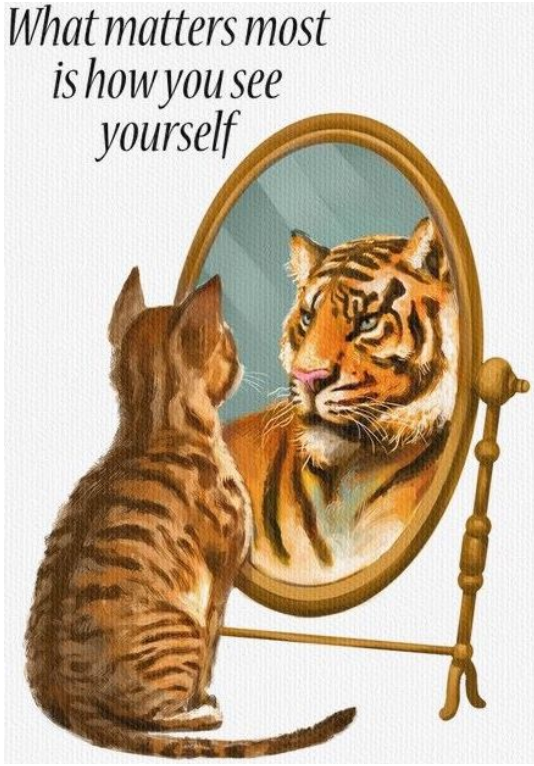
Self-simulating Computer	Gödel Machine	Self-consciousness
Host Machine	Solver	Experiencing Self
Virtual Machine	Searcher	Remembering Self
Hardware	Hardware	Body

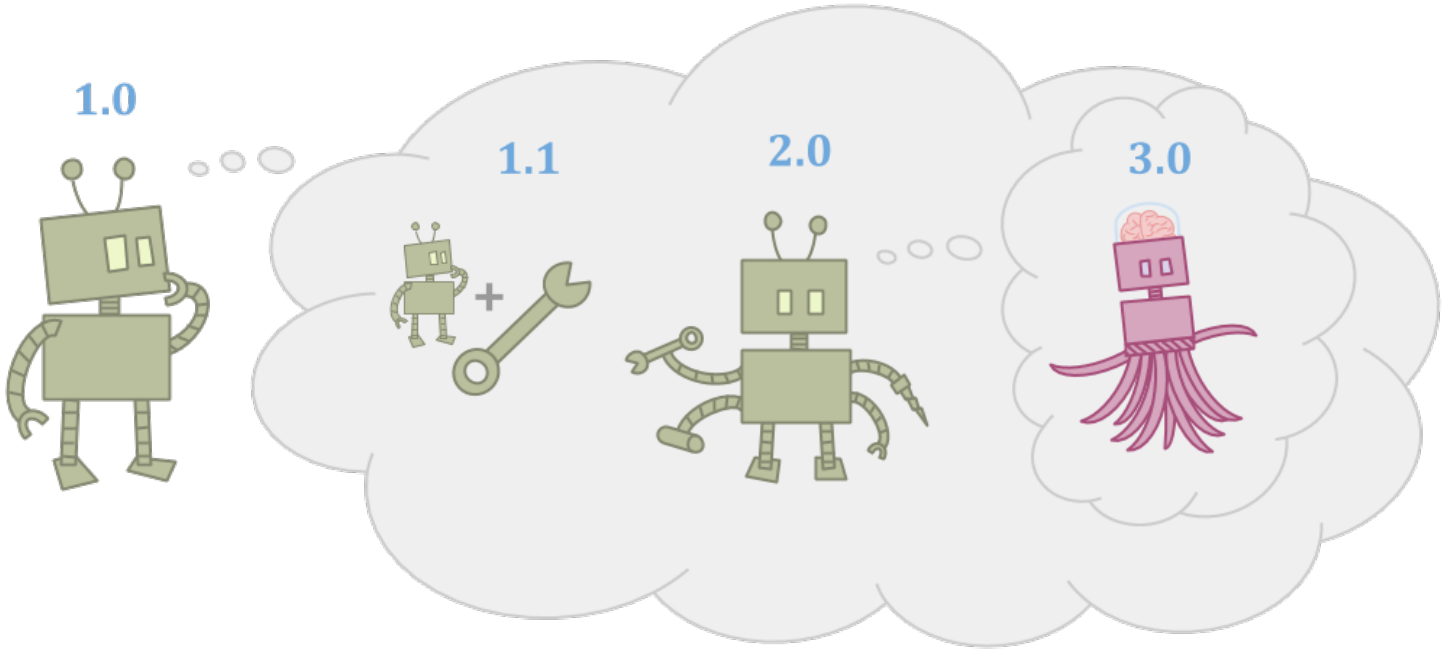


$$\varphi_{\mathbf{e}} = \varphi_{\mathbf{h(e)}}$$



$$\text{self-reference} \overset{?}{\Longrightarrow} \text{self-improvement}$$


What matters most
is how you see
yourself


1.0
1.1
2.0
3.0

Gödel Machines

1. *one-shot* self-improvement: Kleene's fixpoint theorem

$$\varphi_e = \varphi_{h(e)}$$

   - global optimality?
   - goal orthogonality? ends vs means
2. *continuous* self-improvement: Kleene's fixpoint theorem with parameters

$$\varphi_{e(y)} = \varphi_{h(e(y),y)}$$

   - "real-time" optimality. human-computer interaction?
   - intelligent explosion / technological singularity???
     continuous self-improvement ≠ exponential iteration
3. *beyond computability*: Kleene's relativized fixpoint theorem

$$\varphi^A_{e(y)} = \varphi^A_{h(e(y),y)}$$

   - Gödel Machine PK AIXI$^{t\ell}$
   - Gödel Machine PK AIXI

Limitation

1. Gödel's first incompleteness theorem / Rice's theorem
2. Gödel's second incompleteness theorem

$$\mathrm{T} \vdash \Box_{\mathrm{T}'} A \to A \implies \mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}'}$$

3. Legg's incompleteness theorem. *General prediction algorithms must be complex. Beyond a certain complexity they can't be mathematically discovered.*
4. Complexity: higher-level abstractions — coarse grained.
   - Psychology: Duration neglect / Peak-end rule
   - Information Bottleneck: Learning is to forget!
5. Physical constraint: If we assume that it is not possible to measure properties without changing them (observer effect: α is fixpoint-free), then there is a limit to self-inspection.

Evolution & the Number of Wisdom — Chaitin Constant

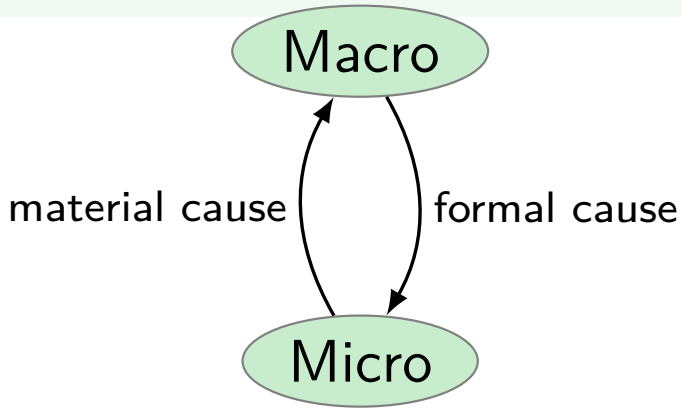
- The enormous computational power of evolution could have developed and coded information into our genes,
  - (a) which significantly guides human reasoning,
  - (b) cannot efficiently be obtained from scratch.

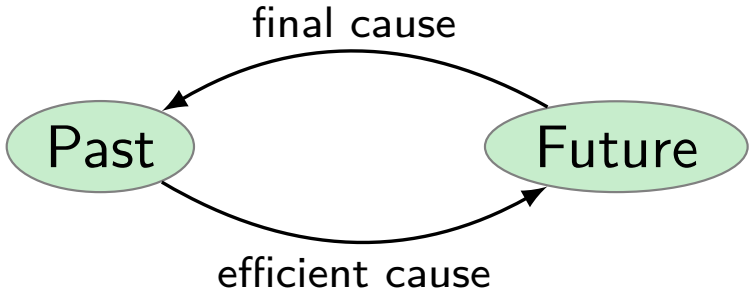
$$\Omega = \lim_{t \to \infty} \sum_{\ell(p) \leq t \ \& \ U(p)\downarrow \text{ within time } t} 2^{-\ell(p)}$$

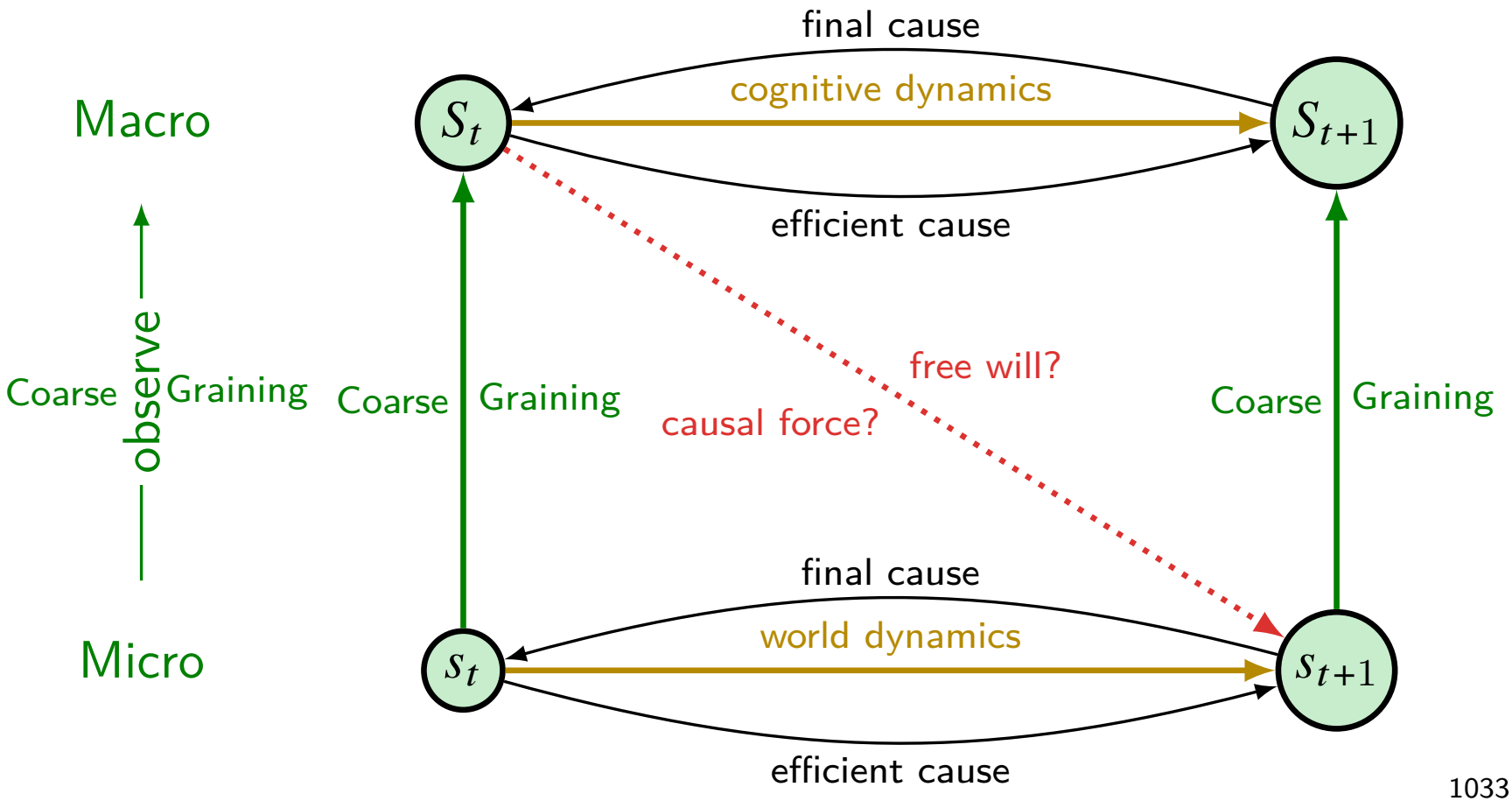
- Cheating solution: add the information from our genes or brain structure to our AI system?
- Biological Evolution: Darwin PK Lamarck
  natural selection vs artificial evolution
  random vs non-random mutation
- Tegmark: Life3.0



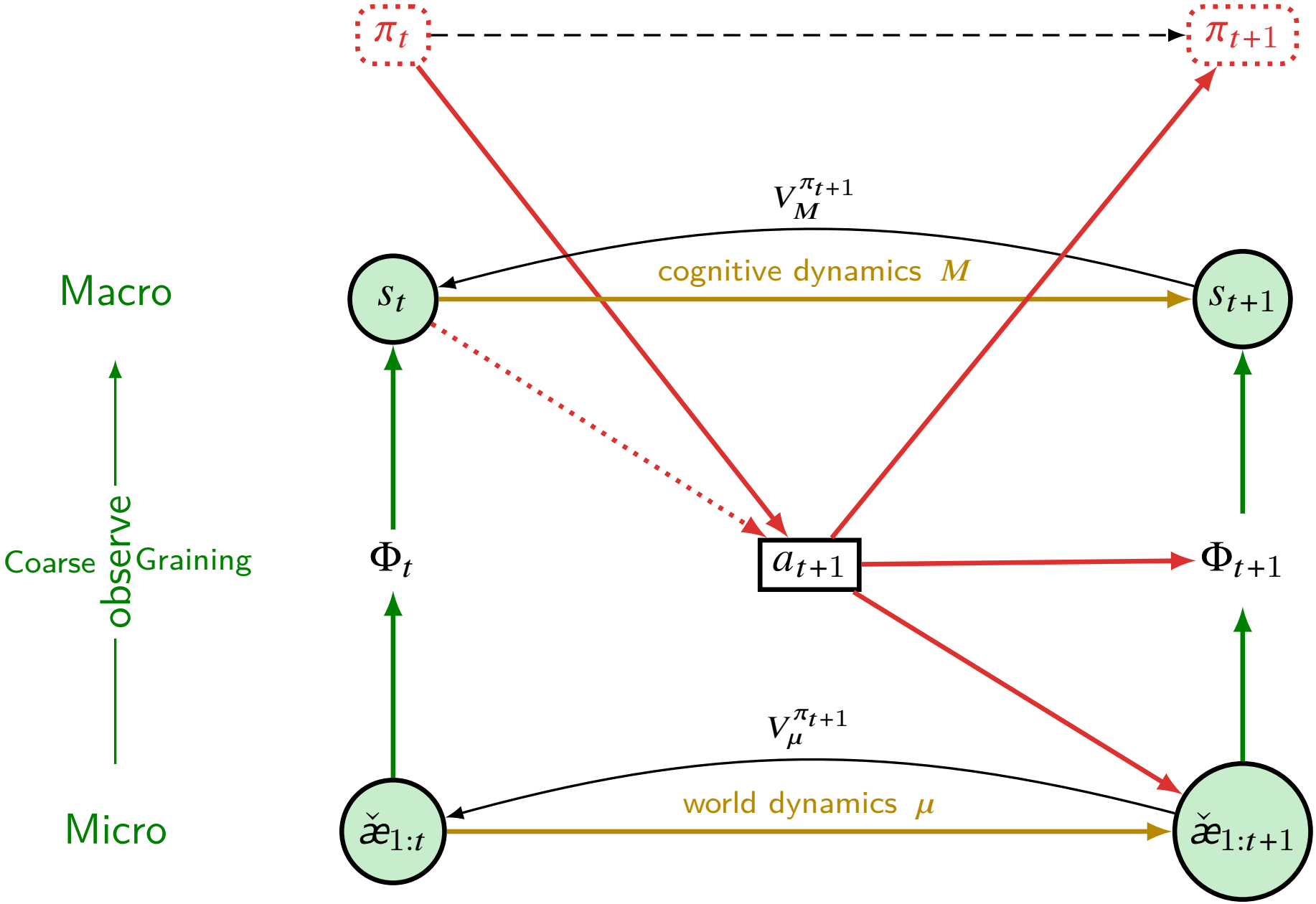
O God, give us courage to change what can be changed,
serenity to accept what cannot be changed,
and wisdom to know the difference.


Macro
material cause
formal cause
Micro


final cause
Past
Future
efficient cause


Macro
final cause
cognitive dynamics
S_t
S_{t+1}
efficient cause
observe
Coarse
Graining
Coarse
Graining
free will?
causal force?
Coarse
Graining
Micro
final cause
world dynamics
s_t
s_{t+1}
efficient cause

Self-Modifying Causal Representation



The Universal program (warm-up): the petulant child

The Petulant Child

Consider program e:

It searches for a proof in PA of a statement:

"program e does not give output n."

When found, gives output n and halts.

According to Kleene's Fixpoint Theorem, there is a program e s.t.

- When run in the standard model $\mathcal{N} \vDash \mathrm{PA}$, the program never halts.
- For any n, there is a model $\mathcal{M} \vDash \mathrm{PA}$ in which the program e gives output n.

The Universal algorithm: sequence version

The Petulant Child

Consider program e:

It searches for a proof in PA of a statement:

"program e does not enumerate the sequence $a_0, \ldots, a_n$ and halt."

When found, enumerate that sequence.

Theorem (The Universal algorithm: sequence version)

There is a program e s.t.

1. PA *proves that the set accepted by e is finite.*
2. *For any finite $A \subset \mathbb{N}$, there is a model $\mathcal{M} \models \mathrm{PA}$ in which the program e accepts exactly A.*
3. *Indeed, for any $A \subset \mathbb{N}$, there is a model $\mathcal{M} \models \mathrm{PA}$ in which the program e accepts exactly A.*

A program that accepts exactly any desired finite set, in the right universe.

The Universal algorithm: function version

Theorem (The Universal algorithm: function version)

There is a program e s.t.

1. PA *proves that the function computed by* e *is finite.*
2. *For any finite partial function* f*, there is a model* $\mathcal{M} \vDash \mathrm{PA}$ *in which the program* e *computes exactly* f*.*
3. *For any partial function* f*, there is a model* $\mathcal{M} \vDash \mathrm{PA}$ *in which the program* e *computes exactly* f*.*

Every function can be computable!...in the right universe.

Proof.

The statements "the n^{th} number enumerated by e is $f(n)$" are finitely consistent with PA. So by compactness they are all true in some model. □

The Universal Algorithm: full extension version

Theorem (Woodin)

There is a program e s.t,

1. PA *proves that the sequence enumerated by* e *is finite.*
2. *In the standard model* $\mathcal{N} \vDash \mathrm{PA}$*, program* e *enumerates the empty sequence.*
3. *For any model* $\mathcal{M} \vDash \mathrm{PA}$ *in which* e *enumerates a finite (possibly nonstandard) sequence* s*, and any finite* $t \in \mathcal{M}$ *extending* s*, there is an end-extension of* $\mathcal{M}$ *to a model* $\mathcal{M}' \vDash \mathrm{PA}$ *in which* e *enumerates exactly* t*.*

In particular, every finite sequence s is enumerated by e in some model $\mathcal{M} \vDash \mathrm{PA}$.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Turing Machine
Computability
Diagonal Method
Incompressibility Method
Incompleteness

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises
References2039

Incompressibility Method

1. In order to prove that an object in a certain class on average satisfies a certain property, select an object of that class that is incompressible.
2. Show that if it does not satisfy the property then it can be compressed by clever computable coding.
3. In general almost all objects of a given class are incompressible, therefore almost all objects in the class have the property involved.

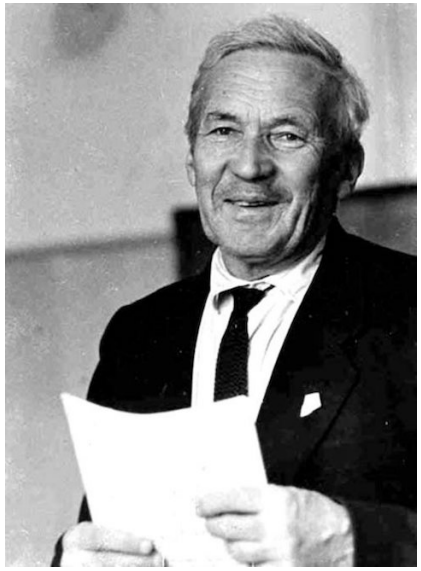


Figure: Kolmogorov

The Infinity of Primes

Theorem (Euclid's Theorem)

There are infinitely many prime numbers.

Proof by Contradiction.

Assume to the contrary that $\mathbb{P} := \{p_1, p_2, \ldots, p_k\}$ is finite. Let $N := \prod_{i=1}^{k} p_i$.

$\exists p \in \mathbb{P} : p \mid (N+1) \;\; \& \;\; p \mid N \implies p \mid 1.$ □

Proof by Incompressibility Method — Chaitin.

$$n = \prod_{i=1}^{m} p_i^{e_i}$$



For a random n,

$$\begin{aligned}
\log n &\leq K(n) \\
&\stackrel{+}{\leq} K(\langle e_1, \ldots, e_m \rangle) \\
&\stackrel{+}{\leq} \sum_{i=1}^{m} K(e_i) \\
&\stackrel{+}{\leq} mK(\log n) \\
&\stackrel{+}{\leq} m(\log\log n + 2\log\log\log n)
\end{aligned}$$

□

Proof by Combinatorics.

$$n = \prod_{i=1}^{m} p_i^{e_i} \implies e_i \leq \left\lfloor \frac{\ln n}{\ln p_i} \right\rfloor \implies$$

$$\# \left\{ (e_1, \ldots, e_m) : \prod_{i=1}^{m} p_i^{e_i} \leq n \right\} \leq \prod_{i=1}^{m} \left(\left\lfloor \frac{\ln n}{\ln p_i} \right\rfloor + 1 \right) \leq (\ln n)^m \ll n$$

□

Proof by Combinatorics — Erdős.

For $N \in \mathbb{N}$, we write every $n \leq N$ in the form $n = rs^2$, where r is the square-free part. There are at most $2^{\pi(N)}$ different square-free parts. Furthermore, $s \leq \sqrt{N}$.
Hence

$$N \leq \# \left\{ (r, s) : rs^2 \leq N \text{ and } r \text{ is squre-free} \right\} \leq 2^{\pi(N)} \sqrt{N}$$

$$\pi(N) \geq \frac{\log_2 N}{2}$$

Proof by Coprime Sequence.

Let $n > 1$. Then n and $n + 1$ must be coprime, and hence $N_2 := n(n + 1)$ must have at least 2 different prime factors. Similarly, $n(n + 1)$ and $n(n + 1) + 1$ are coprime, $N_3 := n(n + 1)[n(n + 1) + 1]$ must have at least 3 different prime factors. This can be continued indefinitely. □

Proof by Coprime Sequence.

Fermat number $F_n := 2^{2^n} + 1$. It is easy to verify that $F_{n+1} = \prod_{k=0}^{n} F_k + 2$, and any two Fermat numbers are coprime, hence there must be infinitely many primes. □

Proof.

For any n, the prime factor of $n! + 1$ must be larger than n. □

Proof by Bertrand's Postulate.

$\forall n \geq 1 \exists p \in \mathbb{P} : n < p \leq 2n$. □

Proof by Prime Number Theorem.

The prime-counting function $\pi(x) \sim \frac{x}{\ln x}$. □

Proof by Euler's Phi Function.

Euler's phi function $\varphi(n) := \#\{k : 1 \leq k \leq n \ \ \& \ \ \gcd(n, k) = 1\}$. We know

$$\gcd(m, n) = 1 \implies \varphi(mn) = \varphi(m)\varphi(n)$$

Then $\varphi\left(\prod_{i=1}^{n} p_i\right) = \prod_{i=1}^{n} (p_i - 1) \geq 2$. Hence
$\exists m \forall p_i \in \{p_1, \ldots, p_n\} : p_i \nmid m$. □

Proof by Euler Product Formula.

$$\zeta(s) := \sum_{n=1}^{\infty} n^{-s} = \prod_{p \in \mathbb{P}} (1 - p^{-s})^{-1}$$

$\zeta(2) = \frac{\pi^2}{6}$ is irrational. If $\mathbb{P}$ were finite, then $\zeta(2)$ would be rational. □

Euler.

$$\ln x \leq \sum_{n \leq x} n^{-1} \leq \prod_{\substack{p \in \mathbb{P} \\ p \leq x}} \left(\sum_{k \geq 0} p^{-k} \right) = \prod_{\substack{p \in \mathbb{P} \\ p \leq x}} \left(1 - p^{-1}\right)^{-1} = \prod_{n=1}^{\pi(x)} \left(1 + \frac{1}{p_n - 1}\right) \leq$$

$$\prod_{n=1}^{\pi(x)} \left(1 + \frac{1}{n}\right) = \pi(x) + 1.$$ □

Proof by Lagrange's Theorem.

Let $p := \max \mathbb{P}$. Let q be a prime dividing $2^p - 1$. We have $2^p \equiv 1 \pmod{q}$. This means that the element 2 has order p in the multiplicative group $\mathbb{Z}_q \setminus \{0\}$ of the field $\mathbb{Z}_q$. This group has $q - 1$ elements. By Lagrange's theorem we have $p \mid (q - 1)$. Hence $q > p$. □

Proof by Topology — Fürstenberg.

Let

$$N_{a,b} \coloneqq \{an + b : n \in \mathbb{Z}\}$$

We call a set $O \subset \mathbb{Z}$ open if $O = \emptyset$ or $\forall x \in O \exists N_{a,b} \subset O : x \in N_{a,b}$.
Note that any non-empty open set is infinite.
And $N_{a,b} = \mathbb{Z} \setminus \bigcup_{i=1}^{a-1} N_{a,b+i}$ is clopen.
Since every integer $\mathbb{Z} \setminus \{-1, +1\}$ has a prime factor, then

$$\mathbb{Z} \setminus \{-1, +1\} = \bigcup_{p \in \mathbb{P}} N_{p,0}$$

If $\mathbb{P}$ were finite, then $\bigcup_{p \in \mathbb{P}} N_{p,0}$ would be closed.
Consequently, $\{-1, +1\}$ would be open. Contradiction! □

Proof.

Let $f(n) \coloneqq \#\{k \in \mathbb{P} : p \mid n\}$, and $P \coloneqq \prod_{p \in \mathbb{P}} p$. Obviously, $\forall n : f(n) = f(n+P)$. However, $f(n) = 0 \implies n = 1$. □

Proof.

$$0 < \prod_{p \in \mathbb{P}} \sin\left(\frac{\pi}{p}\right) = \prod_{p \in \mathbb{P}} \sin\left(\frac{\pi\left(1 + 2\prod_{p \in \mathbb{P}} p\right)}{p}\right) = 0$$

□

Proof.

Consider $P \coloneqq \prod_{i=2}^{n} p_i$. Obviously, $\{k \in \mathbb{N} : \gcd(k, P) = 1\} = \{2^i : i \in \mathbb{N}\}$. In particular, $\gcd(2, P) = 1$, then $\gcd(P-2, P) = 1$. Therefore, $P - 2 \in \{2^i : i \in \mathbb{N}\}$ and $2 \nmid (P-2)$. Hence $P - 2 = 1$, i.e. $P = 3$. It means that 3 is the greatest prime number. □

Halting Problem

Theorem (Halting Problem is Undecidable)

There is no recursive function deciding whether a program halts.

Proof.

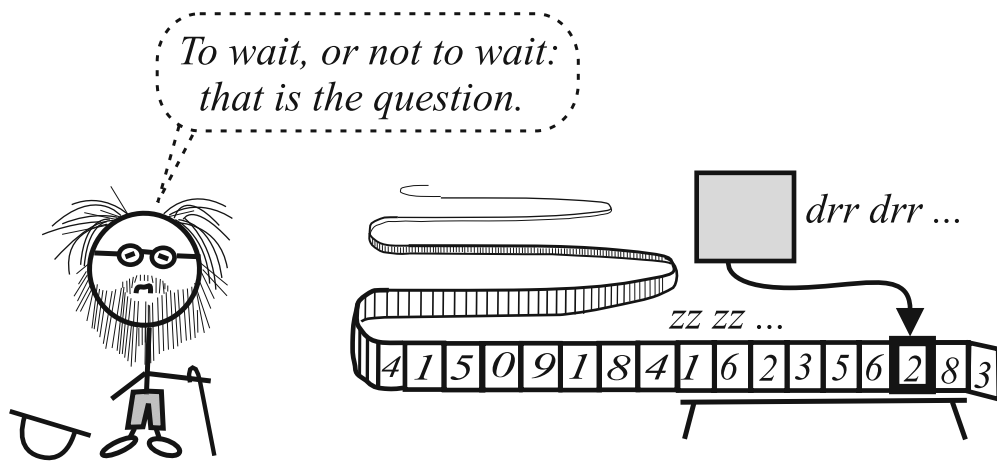
Assume there exists a halting program H.

Construct a program q as follows:

1. read n;
2. generate $A := \{p : \ell(p) \leq n\}$;
3. use H to get $B := \{p \in A : U(p)\downarrow\}$;
4. output $2\max\{U(p) : p \in B\}$.

$$\ell(q) \overset{+}{\leq} \log n \lesssim n \implies U(q) \geq 2U(q)$$

□



Incompressibility vs Incompleteness vs Berry Paradox

Theorem (Kolmogorov)

Kolmogorov complexity K is uncomputable.

$$x^* := \mu x[K(x) > n] \implies n < K(x^*) \leq O(\log n)$$

Theorem (Chaitin)

For any arithmetically sound Gödelian T, $\exists c \forall x : \mathrm{T} \nvdash K(x) > c$.

"given n, find $\mu y\big[\,\mathrm{prf}_{\mathrm{T}}\left(y, K(x) > n\right)\big]$, ouput x" $\implies n < K(x) \leq O(\log n)$

"the least number undefinable in fewer characters than there are in this sentence."

$M_e :=$"find $\mu y\big[\,\mathrm{prf}_{\mathrm{T}}\left(y, K(x) > e\right)\big]$, ouput x" (Berry Paradox)

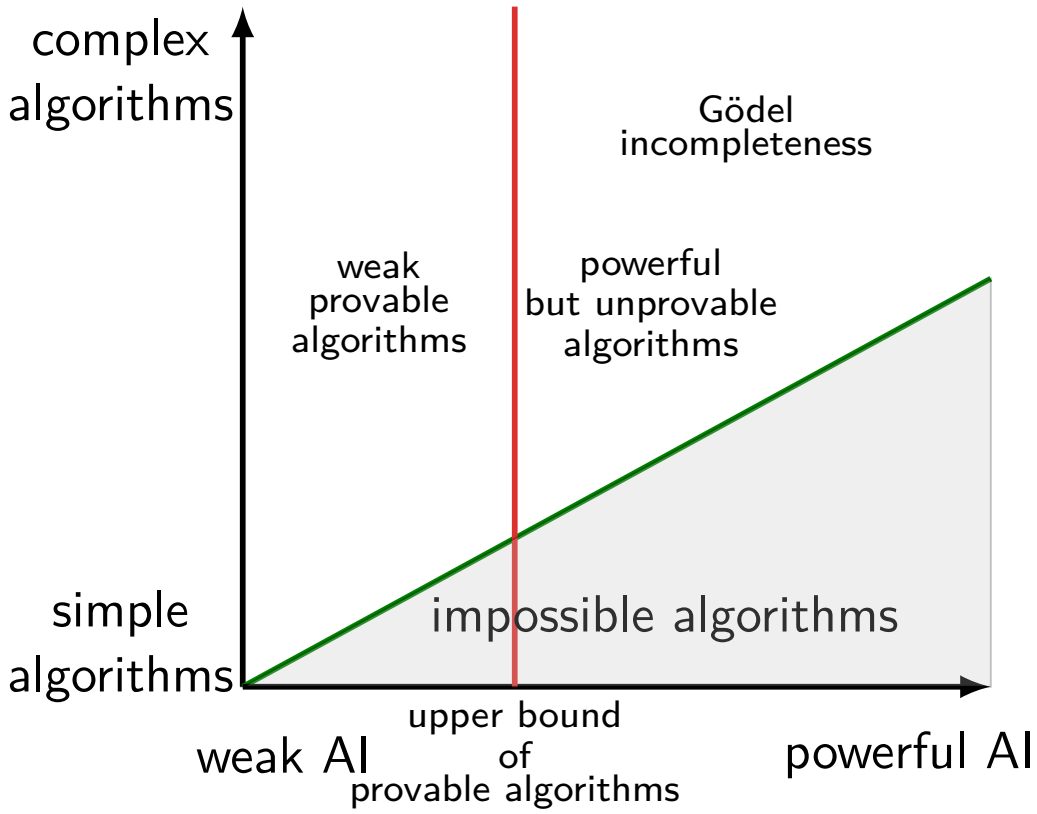
Theorem (Chaitin)

For any arithmetically sound Gödelian T, $\#\{x : \mathrm{T} \vdash K(x) > \ell(x)\} < \infty$.

Remark: For almost all random strings their randomness cannot be proved.

Incompressibility vs Incompleteness vs Intelligence

- $P(x) := \left\{ p \in \mathcal{X}^* : \exists t \forall k \geq t \left(p(x_{1:k}) = x_{k+1} \right) \right\}$
- $P(A) := \bigcap_{x \in A} P(x)$
- $P_n := P\left(\left\{ x : K(x) \leq n \right\} \right)$



- $\forall n \exists p \in P_n : K(p) \stackrel{+}{\leq} n + O(\log n)$
- $\forall n : p \in P_n \implies K(p) \stackrel{+}{\geq} n$

Theorem (Legg)

For any arithmetically sound Gödelian T, $\exists c \forall n \geq c \forall p : \mathrm{T} \nvdash p \in P_n$.

"given n, find $\mu x \left[\mathrm{prf}_{\mathrm{T}} \left(x, p \in P_n \right) \right]$, output p" $\implies K(p) < O(\log n)$

Halting Probability

Definition (Halting Probability)

$$
\begin{aligned}
E_t &:= \left\{p : U(p) \downarrow \text{ in at most } t \text{ steps}\right\} \\
E &:= E_\infty \\
\Omega^t &:= \sum_{p \in E_t} 2^{-\ell(p)} \\
\Omega &:= \Omega^\infty \\
t(n) &:= \mu t[\Omega^t \geq \Omega_{1:n}]
\end{aligned}
$$

Obviously,

$$\Omega^1 \leq \cdots \leq \Omega^i \leq \Omega^{i+1} \leq \cdots \xrightarrow{i \to \infty} \Omega$$

and

$$\Omega_{1:n} \leq \Omega < \Omega_{1:n} + 2^{-n}$$

and $t(n)$ is computable with Oracle Ω.

Halting Probability

Lemma

$$\Omega \equiv_T \chi_E$$

Proof.

If a program p of length $\leq n$ is not in $E_{t(n)}$ then it is not in E at all. Otherwise,

$$\Omega^t + 2^{-n} \leq \Omega^t + 2^{-\ell(p)} < \Omega$$

conflicts with

$$\Omega_{1:n} \leq \Omega^t < \Omega < \Omega_{1:n} + 2^{-n}$$

It follows that $\chi_{E\,1:2^n}$ can be computed from $\Omega_{1:n}$. □

Remark: Ω is a "philosopher's stone". Knowing Ω to an accuracy of n bits will enable us to decide the truth of any provable or finitely refutable mathematical theorem that can be written in less than n bits.

Randomness of Ω

Theorem (Randomness of Ω)

$$\exists c \forall n : K(\Omega_{1:n}) \geq n - c$$

Proof.

$$f(\Omega_{1:n}) := \mu x \left[2^{<\omega} \setminus \left\{ U(p) : p \in E_{t(n)} \right\} \right]$$

Obviously, f is computable.

$$n < K(f(\Omega_{1:n})) \overset{+}{\leq} K(\Omega_{1:n}) + K(f) \overset{+}{\leq} K(\Omega_{1:n})$$

□

Theorem (Chaitin Diophantine Incompleteness)

There is an exponential diophantine equation

$$f(n, x_0, x_1, \ldots, x_m) = g(n, x_0, x_1, \ldots, x_m)$$

which has finitely many solutions $x_0, x_1, \ldots, x_m$ *iff* $\Omega_n = 0$.

Proof.

$A := \left\{\langle n, k\rangle : \Omega_n^k = 1\right\}$ is r.e.
Since a set is r.e. iff it is singlefold exponential Diophantine,
there exists $f(y, x_0, x_1, \ldots, x_m) = g(y, x_0, x_1, \ldots, x_m)$ s.t.

$$\langle n, k\rangle \in A \iff \exists!\langle x_1, \ldots, x_m\rangle \Big[f(n, k, x_1, \ldots, x_m) = g(n, k, x_1, \ldots, x_m)\Big]$$

Thereby, $f(n, k, x_1, \ldots, x_m) = g(n, k, x_1, \ldots, x_m)$ has exactly one solution $x_1, \ldots, x_m$ if $\Omega_n^k = 1$, and it has no solution if $\Omega_n^k = 0$.
$f(n, x_0, x_1, \ldots, x_m) = g(n, x_0, x_1, \ldots, x_m)$ has infinitely many solutions if $\Omega_n = 1$, and it has finitely many solutions if $\Omega_n = 0$. □

Remark: T 只能对有穷多个 n 判定 $f(n, \mathbf{x}) = g(n, \mathbf{x})$ 是否有无穷多解.

Theorem (Chaitin Ω Incompleteness)

For any arithmetically sound Gödelian T, T *can determine at most finitely many (scattered) bits of* Ω.

Proof.

Assume T can provide infinitely many bit of Ω.
Any k different bits $i_1, i_2, \ldots, i_k$ of Ω give us a covering A_k of measure 2^{-k} which includes Ω.

$$A_k := \left\{ s_1 \Omega_{i_1} \cdots s_k \Omega_{i_k} 2^{\omega} : \forall 1 \leq j \leq k \left(s_j \in 2^{<\omega} \ \& \ \ell(s_j) = i_j - i_{j-1} - 1 \right) \right\}$$

$$\mu(A_k) = \frac{2^{i_k - k}}{2^{i_k}} = 2^{-k}$$

Thereby,

$$\forall k \left[\mu(A_k) \leq 2^{-k} \ \ \& \ \ \Omega \in A_k \right]$$

which contradicts the Martin-Löf randomness of Ω. □

Definition (Busy Beaver)

$$\Sigma(n) \coloneqq \max\{x : K(x) \leq n\}$$
$$\sigma(n) \coloneqq \mu t \left[\Omega^t_{1:n} = \Omega_{1:n}\right]$$

Lemma

$$\Sigma(n - K(n) - O(1)) \leq \sigma(n) \leq \Sigma(n + 2K(n) + O(1))$$

Proof.

$$\Omega_{1:n} = \Omega^{\sigma(n)}_{1:n} \implies K(\Omega_{1:n}) \leq K(n) + K(\sigma(n)) + O(1)$$

Since $\exists c \forall n : K(\Omega_{1:n}) \geq n - c$, we have $n - K(n) - O(1) \leq K(\sigma(n))$.
Thus $\Sigma(n - K(n) - O(1)) \leq \sigma(n)$.
From $\sigma(n) \coloneqq \mu t \left[\Omega^t_{1:n} = \Omega_{1:n}\right]$ and $K(\Omega_{1:n}) \leq n + K(n) + O(1)$, we have

$$K(\sigma(n)) \leq K(n) + K(\Omega_{1:n}) + O(1) = n + 2K(n) + O(1)$$

Therefore

$$\sigma(n) \leq \Sigma(n + 2K(n) + O(1))$$

Busy Beaver

Lemma (Busy Beaver)

For any recursive function f, $\Sigma \overset{+}{\geq} f$ *and* $\sigma \overset{+}{\geq} f$.

Proof.

$$K(f(n)) \overset{+}{\leq} K(n) + K(f) \lesssim n \implies \Sigma \overset{+}{\geq} f$$

□

Theorem (Chaitin Busy Beaver Incompleteness)

For any arithmetically sound Gödelian T, $\exists n \forall x : \mathrm{T} \nvdash \sigma(n) \leq x$.

Busy Beaver

Definition

$\mathrm{BB}(n) :=$ the maximum finite number of steps made by any n-state Turing machine (before halting), on an initially blank input tape

Theorem

BB *grows faster than any computable function.*

Proof.

Suppose $f(n) \geq \mathrm{BB}(n)$ is computable. Then we can solve the halting problem, by running an n-state TM for $f(n)$ steps. □

- $\mathrm{BB}(1) = 1$
- $\mathrm{BB}(2) = 6$
- $\mathrm{BB}(3) = 21$
- $\mathrm{BB}(4) = 107$
- $\mathrm{BB}(5) = 47176870$
- $\mathrm{BB}(6) > 10^{10^{10^{10^{10^{10^{10^{10^{10^{10}}}}}}}}}$

Theorem

Any consistent formal system F *can prove at most finitely values of* $\mathrm{BB}(n)$*.*

Proof.

Suppose you build an n-state machine M that lists all theorems of F, halting iff it finds a contradiction. Then if F proved that $\mathrm{BB}(n) = k$, it could also prove its own consistency by running M for k steps and checking that M hadn't yet halted. This would violate Gödel! □

Theorem

The value of $\mathrm{BB}(745)$ *is independent of the* ZFC *axioms, assuming* $\mathrm{Con}_{\mathrm{ZFC}}$*.*

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Turing Machine
Computability
Diagonal Method
Incompressibility Method
Incompleteness

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

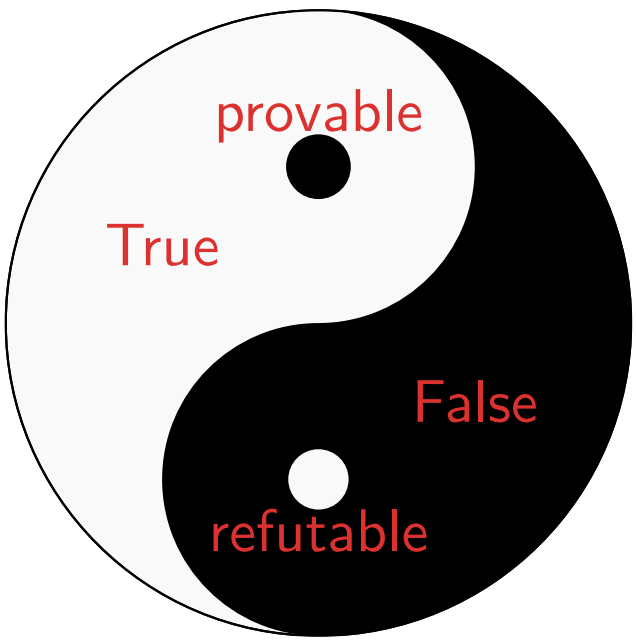
- M: It's one of the most important discoveries of the last decade!
- P: Can you explain it in words ordinary mortals can understand?
- M: Look, buster, if ordinary mortals could understand it, you wouldn't need mathematicians to do the job for you, right? You can't get a feeling for what's going on without understanding the technical details. How can I talk about manifolds without mentioning that the theorems only work if the manifolds are finite-dimensional paracompact Hausdorff with empty boundary?
- P: Lie a bit.
- M: Oh, but I couldn't do that!
- P: Why not? Everybody else does.
- M: Oh, no! Don't lie — because everybody else does.

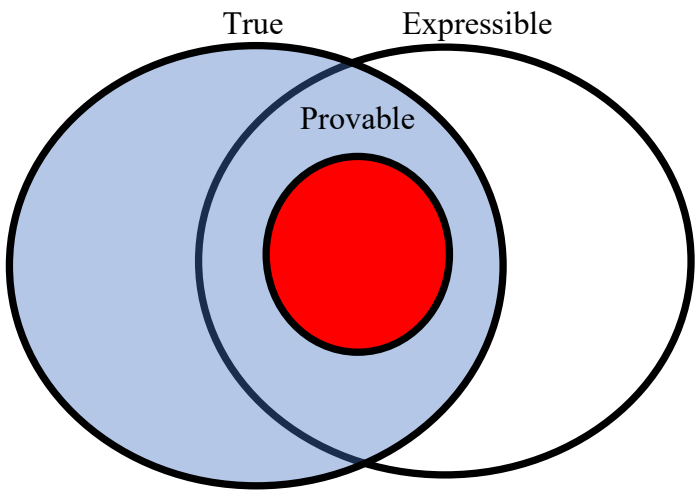
Fixpoint Lemma

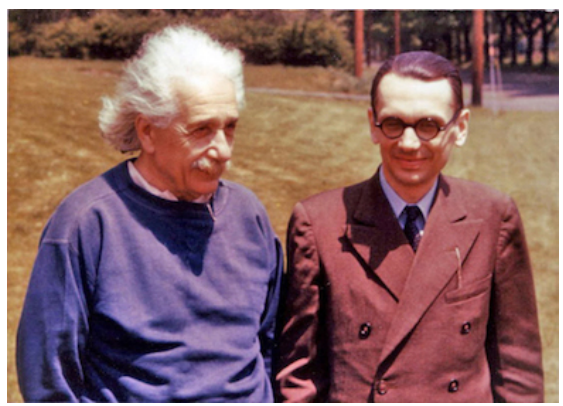
Lemma (Fixpoint Lemma)

For any formula $F(x)$ with one free variable x, there exists a sentence G s.t.

$$\mathrm{Q} \vdash G \leftrightarrow F(\ulcorner G \urcorner)$$







Strong Fixpoint Lemma

Theorem (Strong Fixpoint Lemma)

For a theory T *that is a primitively recursive axiomatised extension of* PA *and that has a function-symbol for each primitive recursive function, and for each one-free-variable-formula* $F(x)$*, there is a closed-term* t *s.t.*

$$\mathrm{T} \vdash t = \ulcorner F(t) \urcorner$$

Proof.

Let $d : \mathbb{N} \to \mathbb{N}$ be the function s.t.

$$d(n) := \begin{cases} \#F(f(\ulcorner f \urcorner)) & \text{if } n = \#f \text{ for one-free-variable-function-symbol } f \\ 0 & \text{otherwise} \end{cases}$$

Since d is primitive recursive, it is represented by some function-symbol d.

$$\mathrm{T} \vdash d(\ulcorner f \urcorner) = \underline{d(\#f)} = \ulcorner F(f(\ulcorner f \urcorner)) \urcorner$$

$$\mathrm{T} \vdash d(\ulcorner d \urcorner) = \ulcorner F(d(\ulcorner d \urcorner)) \urcorner$$

□

Strong Fixpoint Lemma $\implies$ Fixpoint Lemma.

Let $G := F(t)$. Then $G \leftrightarrow F(t) \leftrightarrow F(\ulcorner F(t) \urcorner) \leftrightarrow F(\ulcorner G \urcorner)$.

老子 ◎ô◎

- 道, 可道, 非常道; 名, 可名, 非常名.
  - The theory that can be formulated can't be the ultimate theory. The formulated theory of categories evolves, and its projection on reality changes.
- 无名, 天地之始; 有名, 万物之母.
  - The unformulatable ultimate theory is the truth of universe. The formulated theory is the basis to describe all the matter.
- 故常无, 欲以观其妙; 常有, 欲以观其徼.
  - In search of the unformulatable ultimate theory, we give meaning to life. Within the formulated theory, we study its limits.
- 此两者, 同出而异名, 同谓之玄.
  - The gap between the formulatable and the unformulatable is a mystery.
- 玄之又玄, 众妙之门.
  - From the formulated to the unformulated and from the unformulated to the formulated is the gateway to all understanding.

Gödel's First Incompleteness Theorem

Theorem (Gödel's First Incompleteness Theorem)

For any Gödelian $\mathrm{T} \supset \mathrm{Q}$*, there is a sentence* G *such that,*

1. *if* T *is consistent,* $\mathrm{T} \nvdash G$
2. *if* T *is* ω*-consistent,* $\mathrm{T} \nvdash \neg G$

Gödel's First Incompleteness Theorem

$$F(x) := \neg\Box x$$

$$G = \text{“I am not provable.”}$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \neg\Box G$$

Proof.

$G \leftrightarrow \neg\Box G \implies \Box G \rightarrow \neg G \implies \Box(\Box G \rightarrow \neg G) \implies \Box G \rightarrow \Box\neg G \implies$
$\Box G \rightarrow \Box(G \wedge \neg G) \implies \neg\Box\bot \rightarrow \neg\Box G$ □

We now know enough to know that we will never know everything! ◎ô◎

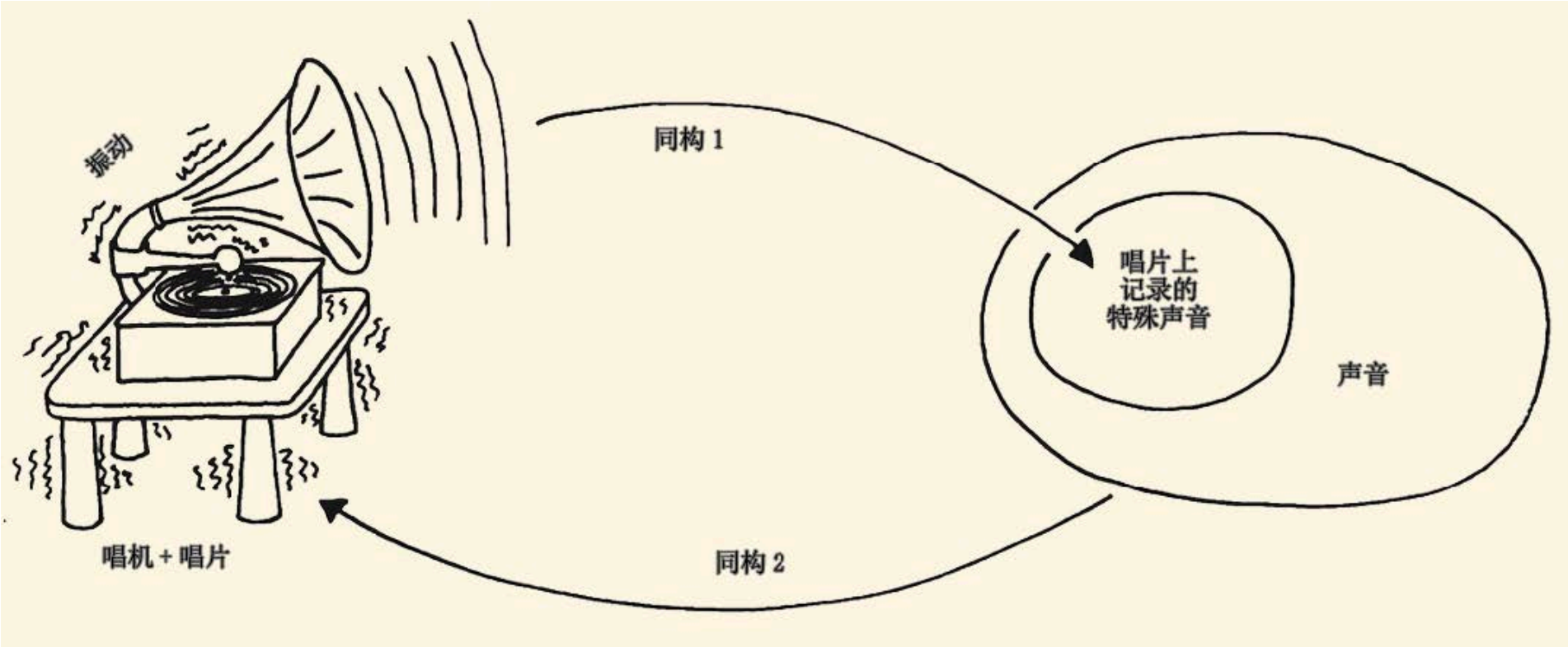


Figure: For every record player, there are records that it can't play. (sympathetic vibration)

Gödel-Rosser's Incompleteness Theorem

Theorem (Gödel's-Rosser First Incompleteness Theorem)

For any Gödelian $\mathrm{T} \supset \mathrm{Q}$, $\mathrm{Cn}(\mathrm{T}) \subsetneq \mathrm{Th}(\mathcal{N})$.

Rosser's Trick

$$\Box^R x := \exists y \big[\mathrm{prf}_{\mathrm{T}}(y, x) \wedge \neg \exists z < y \ \mathrm{prf}_{\mathrm{T}}\left(z, N(x)\right) \big]$$

where $N : \ulcorner A \urcorner \mapsto \ulcorner \neg A \urcorner$

$$F(x) := \neg \Box^R x$$

G = "For every proof of me, there is a shorter proof of my negation."

Tarski's Undefinability of truth Theorem

Theorem (Tarski's Undefinability of truth Theorem)

There is no definable predicate $B(x)$ in the language of arithmetic, such that $\mathcal{N} \vDash A \leftrightarrow B(\ulcorner A \urcorner)$.

suppose $\{\#A : \mathcal{N} \vDash A\}$ is definable by $B(x)$.

$$F(x) := \neg B(x)$$

$$G = \text{“I am not true.”}$$

Montague's Fixpoint Lemma

Lemma (Fixpoint Lemma — Montague)

For any formula $F(x, y)$, there is a formula $G(x)$ s.t.

$$\mathrm{T} \vdash \forall x\big(G(x) \leftrightarrow F\left(x, \ulcorner G(x) \urcorner\right)\big)$$

Proof.

$$d(n) := \begin{cases} \#A(\ulcorner A(x) \urcorner) & \text{if } n = \#A(x) \\ 0 & \text{otherwise} \end{cases}$$

is primitive recursive and is thus represented by some function symbol d.

$$\mathrm{T} \vdash d(\ulcorner A(x) \urcorner) = \ulcorner A(\ulcorner A(x) \urcorner) \urcorner$$

$$E_x(y) := F(x, d(y))$$

$$G(x) := E_x(\ulcorner E_x(y) \urcorner)$$

$$\begin{aligned} G(x) &\leftrightarrow E_x(\ulcorner E_x(y) \urcorner) \\ &\leftrightarrow F(x, d(\ulcorner E_x(y) \urcorner)) \\ &\leftrightarrow F(x, \ulcorner E_x(\ulcorner E_x(y) \urcorner) \urcorner) \\ &\leftrightarrow F(x, \ulcorner G(x) \urcorner) \end{aligned}$$

$$\mathrm{T} \vdash \forall x \big(G(x) \leftrightarrow F\left(x, \ulcorner G(x) \urcorner\right) \big)$$

Remark: This is the form analogous to Second Recursion Theorem: If $f(x, y)$ is a partial recursive function, there is an index e s.t.

$$\varphi_e(y) = f(e, y)$$

Corollary

For any formula $F(x, y, z)$, there is a formula $G(x)$ s.t.

$$\mathrm{T} \vdash \forall x \big(G(x) \leftrightarrow \forall y F\left(x, y, \ulcorner G(y) \urcorner\right) \big)$$

Proof.

Apply the fixed point lemma to the formula

$$F'(x, v) := \forall y \exists z (\delta(v, y, z) \wedge F(x, y, z))$$

$$\mathrm{T} \vdash \forall x (G(x) \leftrightarrow \forall y \exists z (\delta(\ulcorner G(x) \urcorner, y, z) \wedge F(x, y, z)))$$

where $\exists z (\delta(\ulcorner G(x) \urcorner, y, z) \wedge F(x, y, z)) \leftrightarrow F(x, y, \ulcorner G(y) \urcorner)$ □

Yablo Paradox

$$F(x, y, z) \coloneqq y > x \rightarrow \neg\Box z$$

$$\mathrm{T} \vdash \forall x\big(Y(x) \leftrightarrow \forall y > x \neg\Box Y(y)\big)$$

- $\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \forall x \neg\Box Y(x)$
- $\mathrm{T} \vdash \forall x\big(Y(x) \leftrightarrow \mathrm{Con}_{\mathrm{T}}\big)$
- $\mathrm{T} \vdash \forall x \forall y\big(Y(x) \leftrightarrow Y(y)\big)$

Provability Conditions

For any Gödelian $\mathrm{T} \supset \mathrm{PA}$,

1. $\mathrm{T} \vdash A \implies \mathrm{Q} \vdash \Box_{\mathrm{T}} A$
2. $\mathrm{PA} \vdash \Box_{\mathrm{T}} A \to \Box_{\mathrm{T}} \Box_{\mathrm{T}} A$
3. $\mathrm{PA} \vdash \Box_{\mathrm{T}} (A \to B) \to \Box_{\mathrm{T}} A \to \Box_{\mathrm{T}} B$

Löb's Theorem

Theorem (Löb's Theorem)

For any Gödelian $\mathrm{T} \supset \mathrm{PA}$, $\mathrm{T} \vdash \Box(\Box A \to A) \to \Box A$.

Löb's Theorem

$$F(x) := \Box x \to A$$

G = "If I am provable, then A."

Corollary

$$\mathrm{T} \vdash \Box A \to A \implies \mathrm{T} \vdash A$$

Proof.

$\mathrm{T} \vdash \Box G \to \Box\Box G \wedge \Box(\Box G \to A) \implies \mathrm{T} \vdash \Box G \to \Box A \implies \mathrm{T} \vdash \Box G \to A \implies \mathrm{T} \vdash G \implies \mathrm{T} \vdash \Box G \implies \mathrm{T} \vdash A$ □

Curry's Paradox

"If this sentence is true, then Santa Claus exists."

Curry's Paradox

Theorem

$$\left.\begin{aligned} A(y) &\leftrightarrow (\Box B(y) \to \varphi) \\ A(\ulcorner A(x) \urcorner) &\leftrightarrow \Box B(\ulcorner A(x) \urcorner) \end{aligned}\right\} \implies \varphi$$

Proof.

$$\left.\begin{aligned} A(y) &\leftrightarrow (\Box B(y) \to \varphi) \\ R'(x,y) &:= A(x) \leftrightarrow \Box B(y) \end{aligned}\right\} \implies A(y) \leftrightarrow (A(x) \to \varphi) =: R(x,y)$$

Therefore

$$\begin{aligned} A(\ulcorner A(x) \urcorner) \leftrightarrow \Box B(\ulcorner A(x) \urcorner) &\implies R'(\ulcorner A(x) \urcorner, \ulcorner A(x) \urcorner) \\ &\implies R(\ulcorner A(x) \urcorner, \ulcorner A(x) \urcorner) \implies \varphi \end{aligned}$$

□

Gödel's Second Incompleteness Theorem

Theorem (Gödel's Second Incompleteness Theorem)

For any Gödelian $\mathrm{T} \supset \mathrm{PA}$, $\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \neg \Box \, \mathrm{Con}_{\mathrm{T}}$.

Proof.

$$\mathrm{T} \vdash \Box(\Box \bot \rightarrow \bot) \rightarrow \Box \bot \implies \mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \neg \Box \, \mathrm{Con}_{\mathrm{T}}$$

□

Remark:

$$\mathrm{PA} \nvdash \neg \Box_{\mathrm{PA}} \, \mathrm{Con}(\mathrm{PA})$$

$$\mathrm{PA} \nvdash \mathrm{Con}(\mathrm{PA})$$

$$\mathrm{PA}^* \vdash \neg \, \mathrm{Con}(\mathrm{PA}^*) \text{ where } \mathrm{PA}^* = \mathrm{PA} + \neg \, \mathrm{Con}(\mathrm{PA})$$

Remark: The second incompleteness theorem does not imply that the consistency of a system T can only be proved in a stronger system.

Gödel's Second Incompleteness Theorem

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \mathrm{Con}(\mathrm{T} + \neg\, \mathrm{Con}_{\mathrm{T}})$$

Proof.

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \neg \Box\, \mathrm{Con}_{\mathrm{T}}$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \neg \Box (\neg\, \mathrm{Con}_{\mathrm{T}} \rightarrow \bot)$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \mathrm{Con}(\mathrm{T} + \neg\, \mathrm{Con}_{\mathrm{T}})$$

□

We have put a fence around the herd to protect it from the wolves but we do not know whether some wolves were already enclosed within the fence.

— *Henri Poincaré*

God exists because mathematics is consistent, and the devil exists because we can't prove the consistency.

— *André Weil*

Second Incompleteness Theorem

$$\mathrm{T} \nvdash \mathrm{Con}_{\mathrm{T}}$$

Proof.

$$\mathrm{T} \vdash G \leftrightarrow \neg\Box G$$
$$\mathrm{T} \vdash G \rightarrow (\Box G \rightarrow \bot)$$
$$\mathrm{T} \vdash \Box G \rightarrow \Box(\Box G \rightarrow \bot)$$
$$\mathrm{T} \vdash \Box G \rightarrow \Box\bot$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow \neg\Box G$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \implies \mathrm{T} \vdash \neg\Box G$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \implies \mathrm{T} \vdash G$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \implies \mathrm{T} \vdash \Box G$$
$$\mathrm{T} \nvdash \mathrm{Con}_{\mathrm{T}}$$

□

second incompleteness $\implies$ Löb

Löb's Theorem

$$\mathrm{T} \vdash \Box A \rightarrow A \iff \mathrm{T} \vdash A$$

Proof.

Assume $\mathrm{T} \nvdash A$.
Then $\mathrm{T} + \neg A$ is consistent.

$$\mathrm{T} + \neg A \nvdash \mathrm{Con}(\mathrm{T} + \neg A)$$
$$\mathrm{T} + \neg A \nvdash \neg\Box(\neg A \rightarrow \bot)$$
$$\mathrm{T} + \neg A \nvdash \neg\Box A$$
$$\mathrm{T} \nvdash \Box A \rightarrow A$$

□

Let G be the Gödel sentence s.t. $\mathrm{T} \vdash G \leftrightarrow \neg\Box G$. Then

$$\mathrm{T} \vdash G \leftrightarrow \mathrm{Con}_{\mathrm{T}}$$

Proof.

$(\rightarrow)$:

$$\mathrm{T} \vdash \bot \rightarrow G$$
$$\mathrm{T} \vdash \Box\bot \rightarrow \Box G$$
$$\mathrm{T} \vdash \Box\bot \rightarrow \neg G$$
$$\mathrm{T} \vdash G \rightarrow \mathrm{Con}_{\mathrm{T}}$$

$(\leftarrow)$:

$$\mathrm{T} \vdash \Box G \rightarrow \Box\Box G$$
$$\mathrm{T} \vdash \Box G \rightarrow \Box\neg G$$
$$\mathrm{T} \vdash \Box G \rightarrow \Box(G \wedge \neg G)$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \rightarrow G$$

□

Fixpoint

1. For any Gödelian $\mathrm{T} \supset \mathrm{PA}$, $\mathrm{Con}_{\mathrm{T}}$ is the only fixpoint of $\neg\Box x$ up to the logical equivalence in T.
2. For any Gödelian $\mathrm{T} \supset \mathrm{PA}$, $\top$ is the only fixpoint of $\Box x$ up to the logical equivalence in T.

Surprise Exam Paradox vs Second Incompleteness Theorem

$$\boxed{\mathrm{T} \nvdash \mathrm{Con}_{\mathrm{T}}}$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \forall x \neg \Box\big(K(x) > c\big) \qquad \text{(Chaitin)}$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \implies \mathrm{T} \vdash \forall x \in \mathcal{X}^{c+1} \neg \Box\big(K(x) > c\big)$$

$$\mathrm{T} \vdash \forall x \in \mathcal{X}^{c+1} \big[K(x) \leq c \to \Box\big(K(x) \leq c\big)\big] \qquad (\Sigma_1\text{-complete})$$

$$m \coloneqq \Big|\big\{x \in \mathcal{X}^{c+1} : K(x) > c\big\}\Big|$$

$$\mathrm{T} \vdash 1 \leq m \leq 2^{c+1}$$

We prove by induction that for $1 \leq i \leq 2^{c+1}$,

$$\mathrm{T} \vdash m \geq i \implies \mathrm{T} \vdash m \geq i + 1$$

$$\mathrm{T} \vdash m \geq i \implies \mathrm{T} \vdash m \geq i+1$$

Proof.

Assume $\mathrm{T} \vdash m \geq i$. Let $r := 2^{c+1} - i$.

$$\mathrm{T} \vdash m = i \to \exists \text{ different } y_1 \dots y_r \in \mathcal{X}^{c+1} \bigwedge_{k=1}^{r} \big(K(y_k) \leq c\big)$$

$$\mathrm{T} \vdash m = i \to \exists \text{ different } y_1 \dots y_r \in \mathcal{X}^{c+1} \bigwedge_{k=1}^{r} \Box\big(K(y_k) \leq c\big)$$

$$\forall x \in \mathcal{X}^{c+1} \setminus \{y_1, \dots, y_r\} : \ \mathrm{T} \vdash m \geq i \to \left(\bigwedge_{k=1}^{r} \big(K(y_k) \leq c\big) \to \big(K(x) > c\big) \right)$$

$$\mathrm{T} \vdash \Box\big(m \geq i\big) \to \left(\bigwedge_{k=1}^{r} \Box\big(K(y_k) \leq c\big) \to \Box\big(K(x) > c\big) \right)$$

$$\mathrm{T} \vdash m = i \wedge \Box\big(m \geq i\big) \to \exists x \in \mathcal{X}^{c+1} \Box\big(K(x) > c\big)$$

$$\mathrm{T} \vdash m \neq i$$

Completeness Properties

- syntactic completeness: $\Box A \vee \Box \neg A$
- semantic completeness: $A \to \Box A$
- ω-completeness: $\forall x \Box A(\dot{x}) \to \Box \forall x A(x)$

ω-complete $\implies$ ω-consistent $\implies$ 1-consistent $\implies$ consistent

Theorem

For any Gödelian $\mathrm{T} \supset \mathrm{PA}$, *the following are equivalent in* T.

1. $\neg \operatorname{Con}_{\mathrm{T}}$
2. $\Box A \vee \Box \neg A$
3. $A \to \Box A$
4. $\forall x \Box A(\dot{x}) \to \Box \forall x A(x)$

1. consistency
2. effectiveness $\mathrm{Th}(\mathcal{N})$
3. richness Real closed field/Euclidean geometry/Presburger
4. completeness Q / PA / ZFC

Parikh Sentences

There are true sentences that have very long proofs, but there are relatively short proof of the fact that the sentences are provable.

$$\mathrm{prflen}_{\mathrm{T}}(m) := \text{“the length of the proof encoded by } m\text{”}$$

$$\Box^n x := \exists m\big(\,\mathrm{prf}_{\mathrm{T}}(m, x) \wedge \mathrm{prflen}_{\mathrm{T}}(m) < n\big)$$

$$F(x) := \neg\Box^n x$$

$$G = \text{“I have no proof of myself shorter than } n\text{.”}$$

Although there is no short proof of G, there is a short proof of $\Box G$.

$$\neg\Box G \implies G \implies \Box G$$

Gödel's No-short-proof Theorem

Theorem (Gödel's No-short-proof Theorem)

Let f be any primitive recursive function of one variable. Then there is a formula $G(x)$ of one free variable such that $\forall x G(x)$ is true, but for each n, $G(n)$ has no proof with fewer than $f(n)$ steps.

Gödel's No-short-proof Theorem

$$F(x) := \neg \Box^{f(y)} x$$

$$G(y) = \text{"I have no proof of myself shorter than } f(y)\text{."}$$

$$\neg G(n) \implies \Box^{f(n)} G(n) \implies G(n)$$

Remark: it is easily seen that the fixpoint lemma applies also to formulas with free variables.

$$\boxed{\text{Gödel's no-short-proof theorem} \implies \mathrm{T} \nvdash \forall x G(x)}$$

Undecidability

Q is incomplete.

Theorem (Σ_1-completeness of Robinson Arithmetic Q)

For any Gödelian $\mathrm{T} \supset \mathrm{Q}$, and any sentence $A \in \Sigma_1$, $\mathrm{Q} \vdash A \longrightarrow \Box A$.

Theorem (Strong Undecidability of Q)

If $\mathrm{T} \cup \mathrm{Q}$ is consistent, then T is undecidable.

Remark: In fact, the above is true for any countable $\mathscr{L}$ containing a k-ary predicate or function symbol, $k \geq 2$, or at least two unary function symbols.

First order logic $\mathcal{L}_{\omega\omega}$ is undecidable.

Theorem (Church1936, Turing1936)

The set of valid sentences is recursively enumerable but undecidable.

Theorem (Trakhtenbrot's Theorem)

Suppose $\mathscr{L}$ contains at least one binary relation symbol.

- *The set of finitely satisfiable sentences is recursively enumerable,*
- *but it is undecidable whether a sentence is finitely satisfiable.*
- *The set of sentences valid in all finite models is not recursively enumerable.*

Remark:

1. This implies that Gödel's completeness theorem fails in the finite since completeness implies recursive enumerability.
2. It follows that there is no recursive function f s.t.: if a formula A has a finite model, then it has a model of size at most $f(A)$. In other words, there is no effective analogue to the Löwenheim-Skolem theorem in the finite.

Undecidability

Problem (Post Correspondence Problem)

Given n pairs of words:

$$(w_1, v_1), \ldots, (w_n, v_n)$$

is there a sequence of indices $(i_1, \ldots, i_k)$ with $k \geq 1$ s.t.

$$w_{i_1} \ldots w_{i_k} \stackrel{?}{=} v_{i_1} \ldots v_{i_k}$$

Example

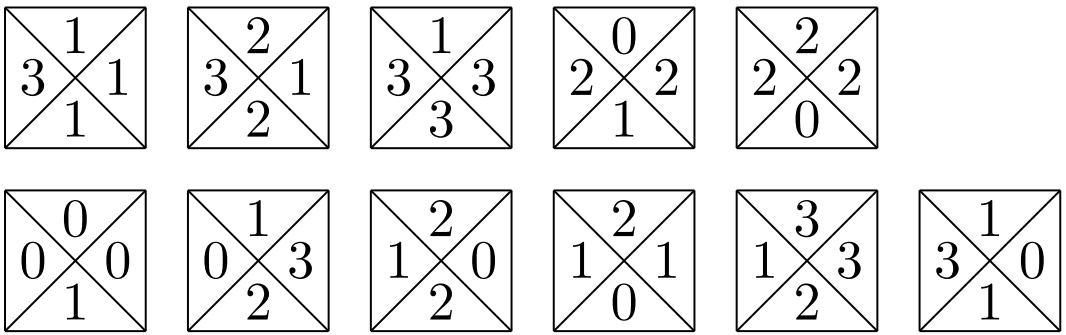
$(a, baa), (ab, aa), (bba, bb)$	$(3, 2, 3, 1)$
$(1, 101), (10, 00), (011, 11)$	$(1, 3, 2, 1)$
$(110, 0), (00, 1)$	no solution

Theorem (Post1946)

Post Correspondence Problem is undecidable.

Problem (Wang's Tiling Problem)

Wang's tiling problem (of determining whether a tile set can tile the plane) is undecidable.

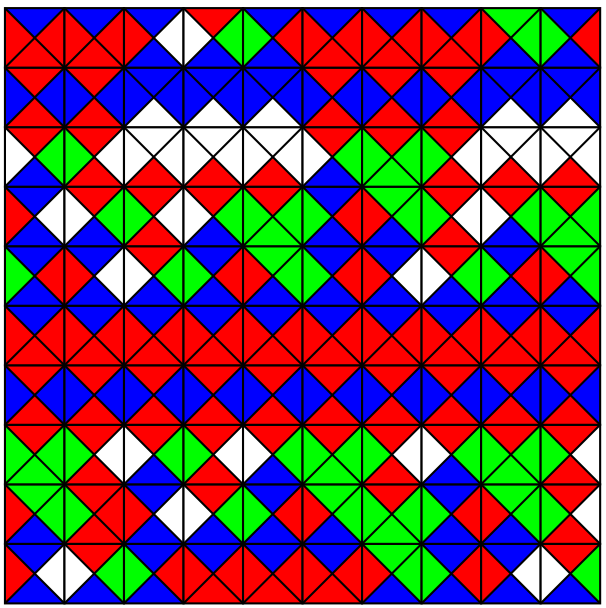


$0 \mapsto$ white

$1 \mapsto$ red

$2 \mapsto$ blue

$3 \mapsto$ green



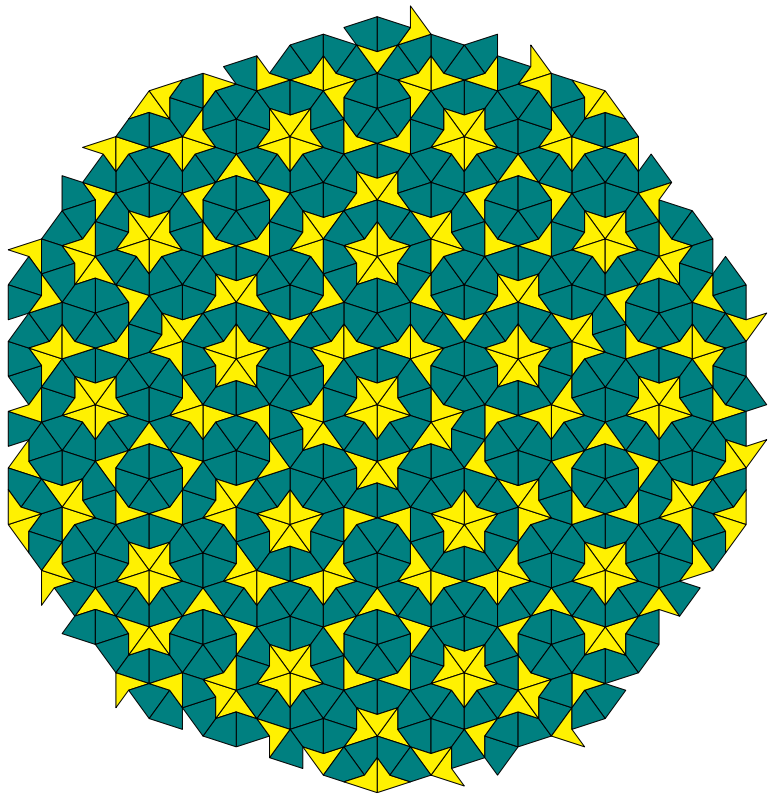


Figure: Penrose Tiling

王浩铺砖



- 是否有算法判定任给有穷个四色砖块能否铺满整个平面？否!
- 是否存在有穷个砖块能铺满平面但只能非周期性的铺满？是!
- 王浩铺砖可以模拟图灵机的运行.
- Berger 证明: 一个图灵机不停机当且仅当相应砖块集铺满整个平面.
- 可以用一个一阶逻辑公式描述这样的铺砖问题, 使得这个公式可满足当且仅当存在这样的铺砖. 例如, 可以用一阶逻辑说: 只有几种砖块, 任意两个相邻的砖块的相接的颜色是一样的, 每个砖块上下左右都有相邻的砖块. 所以一阶逻辑的可满足性 (及有效性) 不可判定.

Hilbert's 10^{th} Problem is Unsolvable

Definition (Diophantine Set)

A set $A \subset \mathbb{N}^n$ is diophantine iff there exists a polynomial $P(\mathbf{x}, \mathbf{y})$ with integer coefficients s.t.

$$\mathbf{x} \in A \iff \exists \mathbf{y} \in \mathbb{N}^m [P(\mathbf{x}, \mathbf{y}) = 0]$$

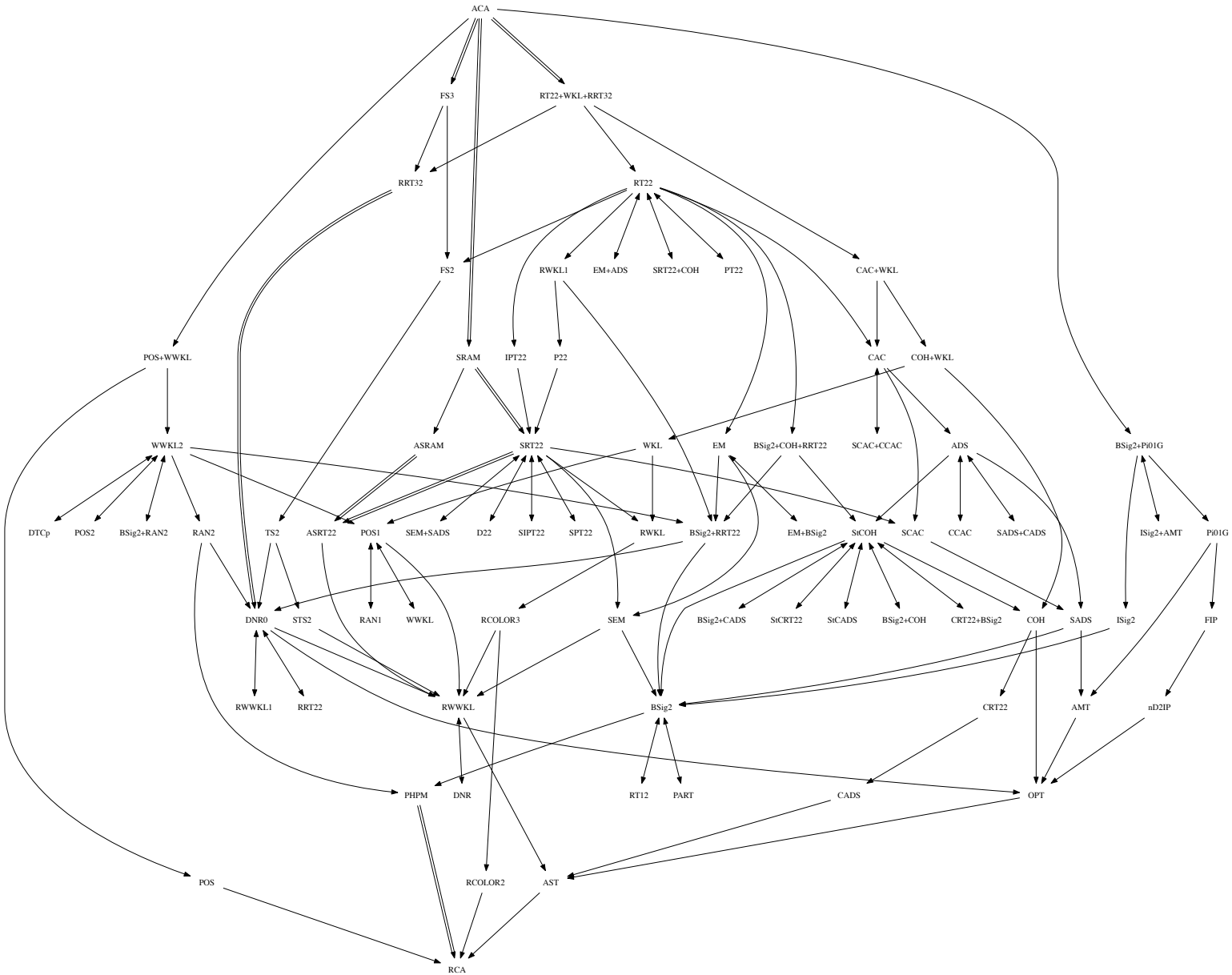
Theorem (MRDP Theorem — Matiyasevich,Robinson,Davis,Putnam)

A subset of $\mathbb{N}$ is r.e. iff it is diophantine.

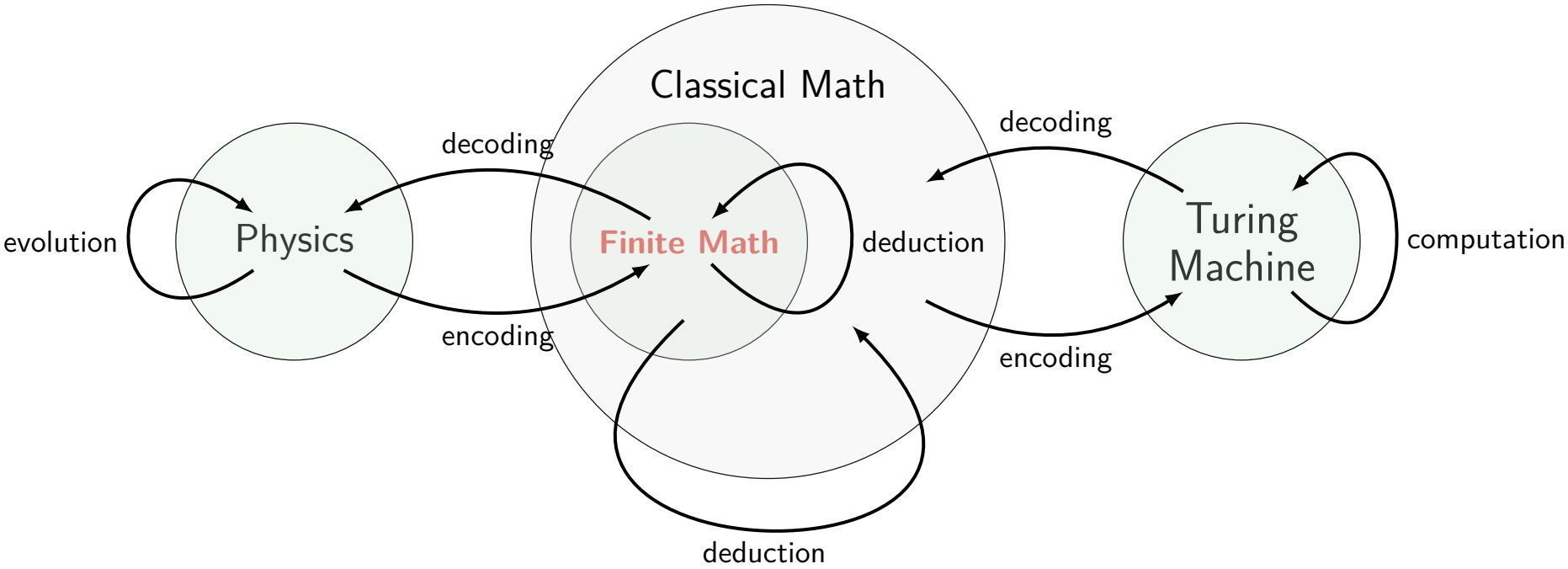
There is no algorithm for deciding whether an arbitrary diophantine equation has a solution.



Reverse Mathematics



The Applicability (Unreasonable Effectiveness) of Mathematics



$$\underbrace{\Gamma_r \cup \Gamma_m \cup \Gamma_b \vdash A \qquad \mathcal{M}_r \vDash \Gamma_r}$$

$$\Downarrow ?$$

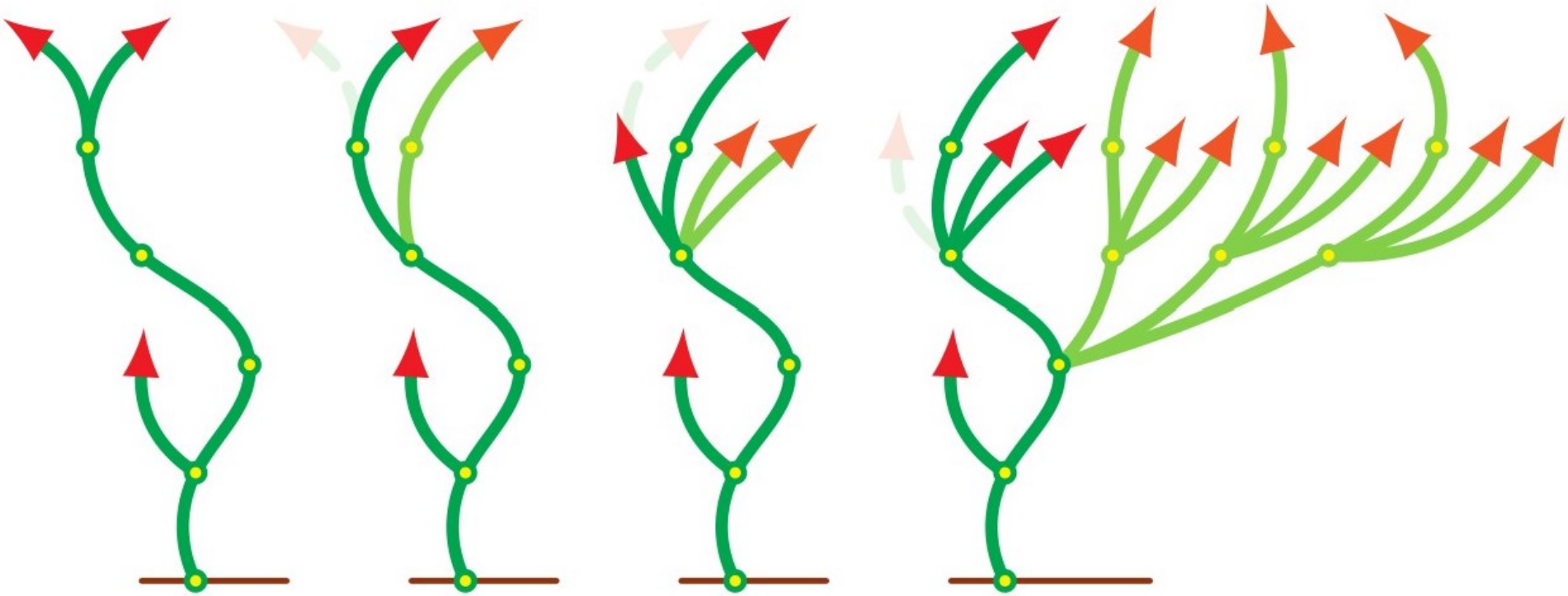
$$\mathcal{M}_r \vDash A$$

$$\frac{\Gamma_r \cup \Gamma'_r \vdash A \quad \mathcal{M}_r \vDash \Gamma_r \cup \Gamma'_r}{\mathcal{M}_r \vDash A}$$

where $\Gamma'_r \subset \Gamma_m \cup \Gamma_b$

Hydra 九头蛇游戏 — “自然的” 不可证命题

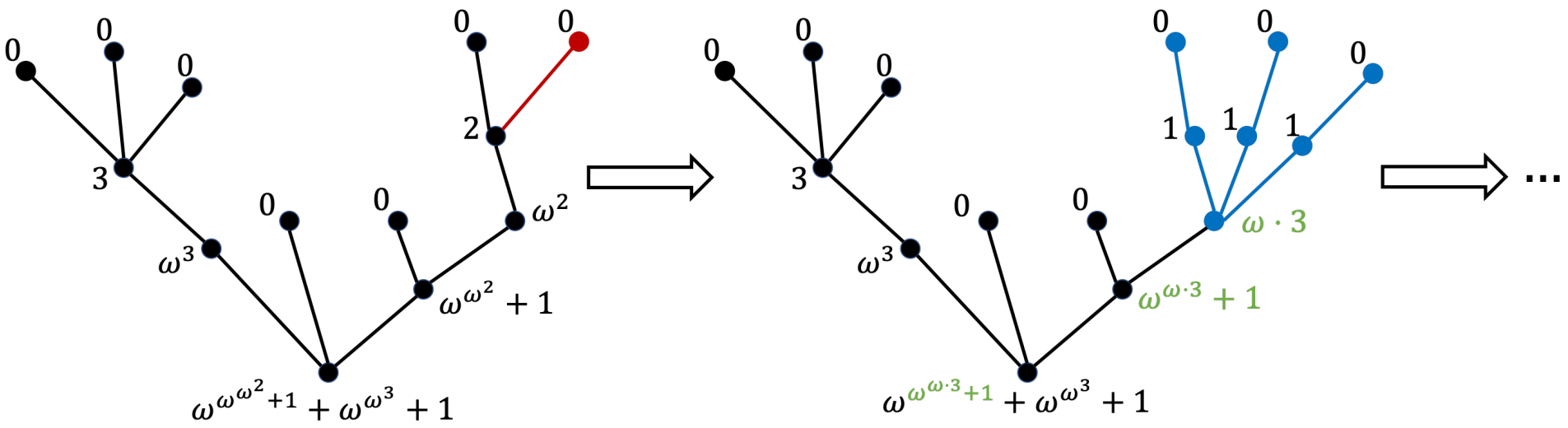
- 每一步你砍掉它一个头
- 在第 n 步, 如果它的一个非根部的头被砍掉, 则从被砍掉头的下方节点处长出 n 个副本



Goodstein Theorem 不管你怎么砍都能杀死 “九头蛇”

Kirby-Paris Theorem 但你无法通过 PA 证明这一点

When you come across the Hydra — "natural" vs "ad-hoc"



Problem

Is there a winning strategy?

Goodstein Theorem

You can't lose!

Theorem (Kirby-Paris Theorem)

Any formal system that proves Goodstein Theorem is strong enough to prove that PA *is consistent.*

Goodstein Function

Definition (Goodstein Function)

Define $G_n(m)$ as follows: if $m = 0$ then $G_n(m) = 0$, if $m \neq 0$ then $G_n(m)$ is a number obtained by replacing every n in the base n representation of m by $n + 1$ and then subtracting 1. Let

$$\begin{aligned} m_0 &:= m \\ m_k &:= G_{k+1}(m_{k-1}) \\ f_G(m) &:= \mu k [m_k = 0] \end{aligned}$$

$$\begin{aligned} m_0 &= 266 = 2^{2^{2+1}} + 2^{2+1} + 2^1 \\ m_1 &= G_2(m_0) = 3^{3^{3+1}} + 3^{3+1} + 2 \approx 10^{38} \\ m_2 &= G_3(m_1) = 4^{4^{4+1}} + 4^{4+1} + 1 \approx 10^{616} \\ m_3 &= G_4(m_2) = 5^{5^{5+1}} + 5^{5+1} \approx 10^{10000} \end{aligned}$$

Fast-Growing Hierarchy

Definition (Wainer Hierarchy)

$$f_0(n) := n + 1$$
$$f_{\alpha+1}(n) := f_\alpha^n(n)$$
$$f_\alpha(n) := f_{\alpha[n]}(n) \text{ if } \alpha \text{ is a limit ordinal.}$$

For limit ordinals $\lambda < \varepsilon_0$, written in Cantor normal form,

- if $\lambda = \omega^{\alpha_1} + \cdots + \omega^{\alpha_k}$ for $\alpha_1 \geq \cdots \geq \alpha_k$, then $\lambda[n] := \omega^{\alpha_1} + \cdots + \omega^{\alpha_k}[n]$
- if $\lambda = \omega^{\alpha+1}$, then $\lambda[n] := \omega^\alpha[n]$
- if $\lambda = \omega^\alpha$ for a limit ordinal α, then $\lambda[n] := \omega^{\alpha[n]}$
- if $\lambda = \varepsilon_0$, then $\lambda[0] := 0$ and $\lambda[n+1] := \omega^{\lambda[n]}$

- $\alpha < \beta < \varepsilon_0 \implies f_\alpha < f_\beta$
- For any primitive recursive function f, $\exists \alpha < \omega : f < f_\alpha$
- Every f_α with $\alpha < \varepsilon_0$ is computable, and provably total in PA.
- If f is computable and provably total in PA, then $\exists \alpha < \varepsilon_0 : f < f_\alpha$. Hence f_{ε_0} is not provably total in PA.

Kirby-Paris Theorem vs Goodstein Theorem

Theorem

$$\forall \alpha < \varepsilon_0 \, (f_\alpha < f_G)$$

$$\mathrm{ZFC} \vdash \forall m \exists k (m_k = 0) \quad \text{but} \quad \mathrm{PA} \nvdash \forall m \exists k (m_k = 0)$$

$$\Sigma(n) := \max\{m : K(m) \leq n\}$$

$$f < \Sigma \text{ for any computable } f$$

$$\mathrm{PA} \nvdash \forall n \exists m (\Sigma(n) = m)$$

$$\exists n \forall m : \mathrm{PA} \nvdash \Sigma(n) \leq m$$

For any arithmetically sound Gödelian T : $\exists n \forall m : \mathrm{T} \nvdash \Sigma(n) \leq m$

Paris-Harrington Theorem vs Ramsey Theorem

$$[A]^n := \{X \subset A : |X| = n\}$$

$$\kappa \to (\lambda)^n_m := \forall F : [\kappa]^n \to m\left(\exists H \subset \kappa\left(|H| = \lambda \ \wedge \ \exists i \in m\left([H]^n \subset F^{-1}(i)\right)\right)\right)$$

Ramsey theorem: $\forall mn \in \omega : \aleph_0 \to (\aleph_0)^n_m$

$$s \to (k_0, \ldots, k_{m-1})^n_m := \forall F : [s]^n \to m\left(\bigvee_{i=0}^{m-1} \exists H \subset s\left(|H| = k_i \ \wedge \ [H]^n \subset F^{-1}(i)\right)\right)$$

$$s \underset{*}{\to} (k)^n_m := \forall F : [s]^n \to m\left(\exists H \subset s\left(|H| \geq \min(H) \ \wedge \ |H| \geq k \ \wedge \ \exists i \in m\left([H]^n \subset F^{-1}(i)\right)\right)\right)$$

$$\mathrm{ZFC} \vdash \forall mnk \exists s \left(s \underset{*}{\to} (k)^n_m\right)$$

$$\mathrm{PA} \nvdash \forall mnk \exists s \left(s \underset{*}{\to} (k)^n_m\right)$$

$$\forall \alpha < \varepsilon_0 \, (f_\alpha < f_R) \quad \text{where} \quad f_R(m,n,k) := \mu s \left[s \underset{*}{\to} (k)^n_m\right]$$

Leibniz — Hilbert — Gödel — Turing ···

- G_{T} := "This sentence cannot be proved in the formal axiomatic system T"
- We humans can easily see that G_{T} must be true.
- Since any AI is a FAS T, no AI can prove G_{T}. — *Penrose*
- Therefore there are things humans, but no AI system can do.
- P := "Penrose cannot prove that this sentence is true"
- Penrose cannot prove P, but now we can conclude that it is true.
- Penrose is in the same situation as an AI.
- Either (a) absolutely unsolvable problems exist or (b) the human mind infinitely surpasses any Turing machine or formal axiomatizable system. — *Gödel*
- Hayek: social spontaneous order?
- Hawking: 'Theory of Everything' impossible?

Ingenuity, Intuition and Creativity

Logic will get you from A to B; Imagination will take you everywhere.

— Albert Einstein

No, no, you're not thinking; you're just being logical.

— Niels Bohr

The ultimate goal of mathematics is to eliminate any need for intelligent thought.

— Alfred Whitehead

Eliminate not intuition but ingenuity.

— Alan Turing

The logical process is essentially creative.

— Emil Post

Logic may be said to be Mathematics become self-conscious.

— Emil Post

Strength & Limitation

God plays dice both in quantum mechanics and in pure math.

— *Gregory Chaitin*

It is the duty of the human understanding to understand that there are things which it can't understand, and what those things are.

— *Søren Kierkegaard*

The only way of discovering the limits of the possible is to venture a little way past them into the impossible.

— *Arthur Charles Clarke*

- Is the Universe Like π or Like Ω?
- Perhaps from inside this world we will never be able to tell the difference, only an outside observer could do that.

数学之外

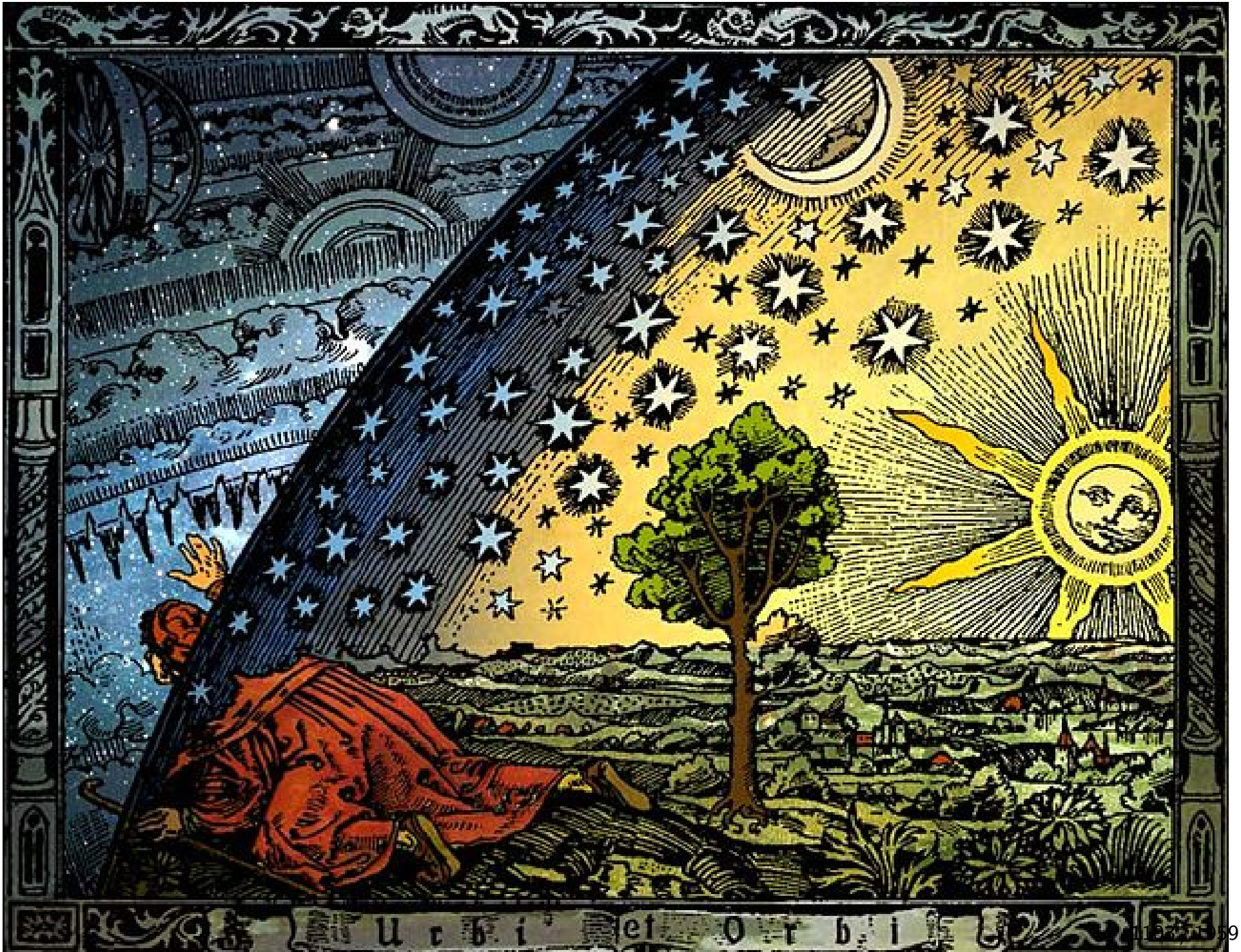
> 一个完全不自由的社会 (即处处按 "统一" 的法则行事的社会), 就其行为而言, 或者是不一致的, 或者是不完备的, 即无力解决某些问题, 可能是极端重要的问题. 在困难的处境里, 二者当然都会危及它的生存. 这个说法也适用于个体的人.
>
> — 哥德尔

Remark: 哥德尔定理版本的 "哈耶克-自发社会秩序".

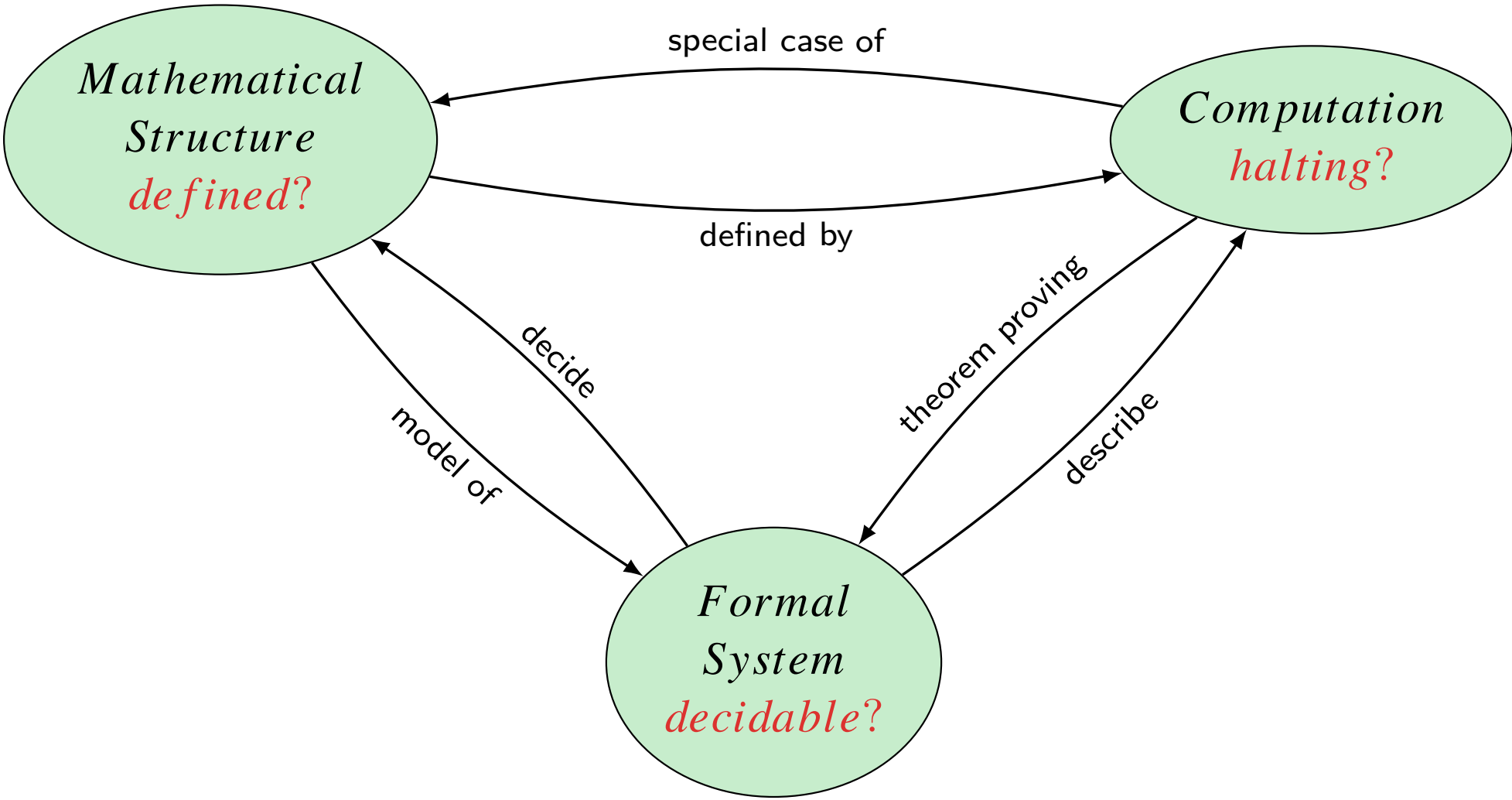
1. 在包含理性人类的任何社会文明体系中, 永远存在着无法用人类理性解决的问题, 不存在一个万能的政府, 能对体系内的任何问题作出合理与公正的解决.
2. 对于包含理性人类的任何社会文明体系, 不能在该体系内对其作出合理与公正的评价.

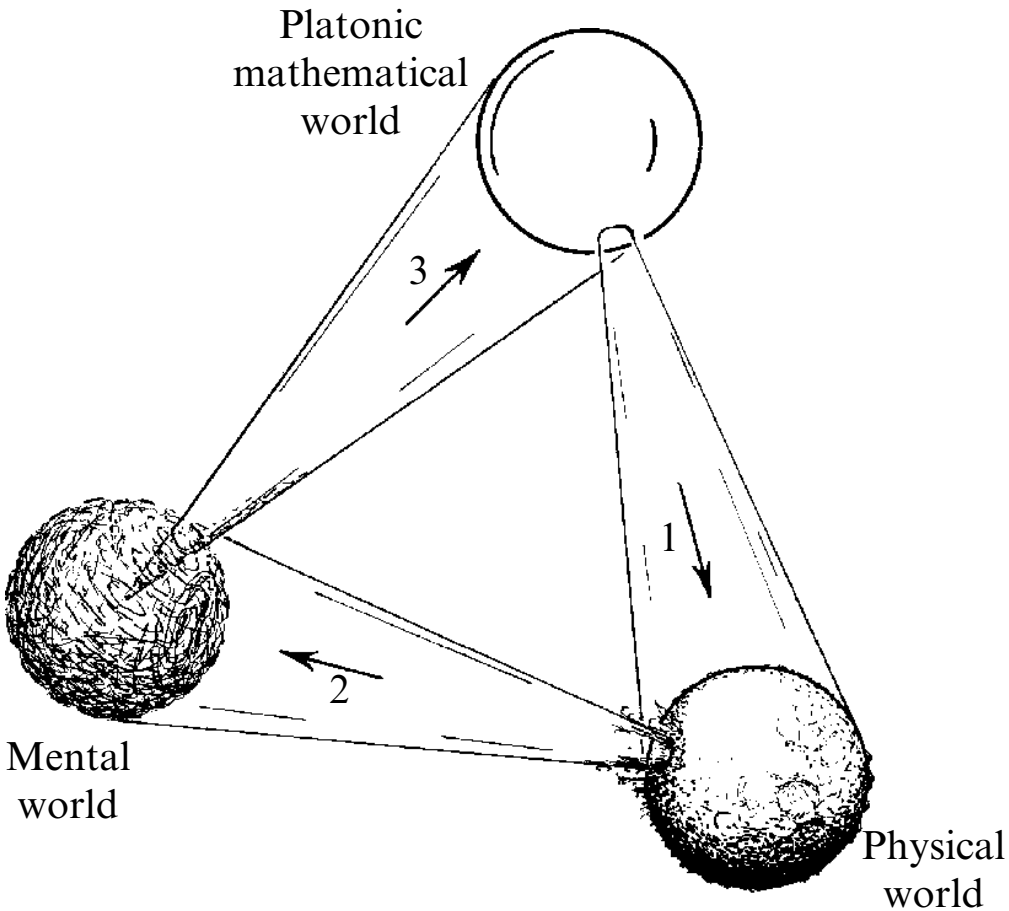
Remark: 哥德尔定理 vs 因果涌现: 秩序和涌现的属性不能从系统内部观察和认知, 只能由外部观察者来观察.

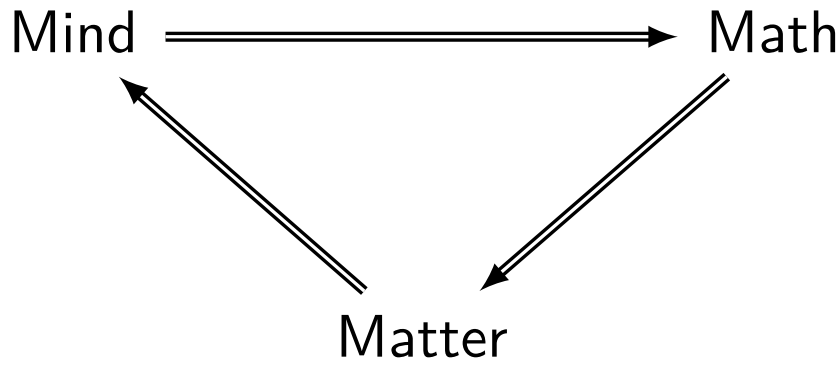
Remark: 霍金: "万有理论" 不可能.

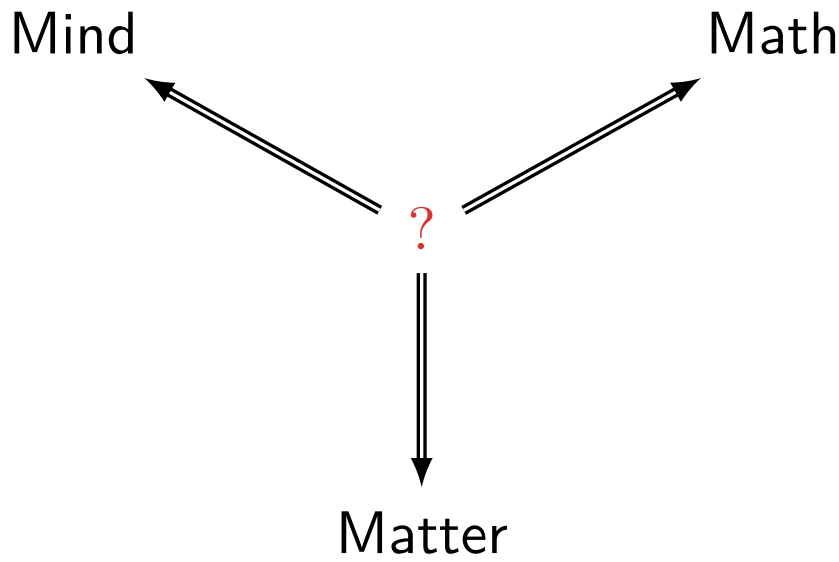

Urbi et Orbi

Math-Matter-Mind (Penrose)




Platonic
mathematical
world
3
1
2
Mental
world
Physical
world


Mind
Math
Matter


Mind
Math
?
Matter

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic
- Equational Logic
- Boolean Algebra
- Lambda Calculus and Combinatory Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Why Study Equational Logic?

- Propositional logic has very limited expressive power.
- Equational Logic is powerful enough to express propositional logic.
- It underlies the mathematical field of universal algebra.
- The basis of programming and specification languages.

Language

$$\mathscr{L}^{=} := \{=, (,)\} \cup \mathrm{Var} \cup \mathrm{Cst} \cup \mathrm{Fun}$$

where

$$\mathrm{Var} := \{x_i : i \in \mathbb{N}\}$$

$$\mathrm{Cst} := \{c_i : i \in \mathbb{N}\}$$

$$\mathrm{Fun} := \bigcup_{n \in \mathbb{N}} \mathrm{Fun}^n \quad \mathrm{Fun}^n := \left\{f_1^n, f_2^n, f_3^n, \dots\right\}$$

- c is a constant symbol.
- f^n is an n-place function symbol.

Term & Formula

Term

$$t \coloneqq x \mid c \mid f(t, \ldots, t)$$

where $x \in \mathrm{Var}$, $c \in \mathrm{Cst}$ and $f \in \mathrm{Fun}$.

Well-Formed Formula Wff

$$A \coloneqq s = t$$

where $s, t \in \mathrm{Term}$.

Semantics

Definition (Model)

A model is a pair $\mathcal{M} := (M, [\![\,]\!])$, where M is a non-empty set, and $[\![\,]\!]$ is a mapping which assigns to each constant symbol an element $[\![c]\!] \in M$, and assigns each function symbol f an n-ary function $[\![f]\!] : M^n \to M$.

Remark: We write $\left(M, c^{\mathcal{M}}, f^{\mathcal{M}}\right)$ for convenience.

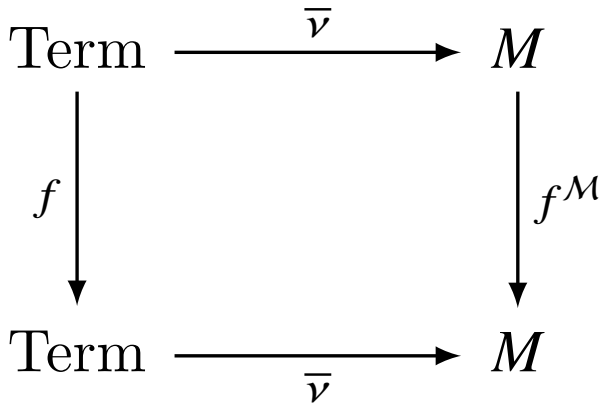
Definition (Variable Assignment)

A variable assignment is a function $\nu : \mathrm{Var} \to M$.

We extend ν to $\overline{\nu} : \mathrm{Term} \to M$ by recursion as follows:

Assignment over Terms

- $\overline{\nu}(x) = \nu(x)$
- $\overline{\nu}(c) = c^{\mathcal{M}}$
- $\overline{\nu}(f(t_1, \ldots, t_n)) = f^{\mathcal{M}}(\overline{\nu}(t_1), \ldots, \overline{\nu}(t_n))$



Semantics

- $\mathcal{M}, \nu \vDash s = t$ iff $\nu(s) = \nu(t)$. (Satisfaction)
- $\mathcal{M} \vDash s = t$ iff for all $\nu : \mathcal{M}, \nu \vDash s = t$. (True)
- $\mathcal{M} \vDash \mathrm{T}$ iff for all $A \in \mathrm{T} : \mathcal{M} \vDash A$. (Model)
- $\mathrm{T} \vDash s = t$ iff for all $\mathcal{M} : \mathcal{M} \vDash \mathrm{T} \implies \mathcal{M} \vDash s = t$.
- $\vDash s = t$ iff $\emptyset \vDash s = t$. (Valid)

Formal System

Birkhoff's Rules

$$\frac{}{t = t}\ \text{refl} \qquad \frac{s = t}{t = s}\ \text{symm}$$

$$\frac{r = s \quad s = t}{r = t}\ \text{trans} \qquad \frac{s = t}{r(\ldots s \ldots) = r(\ldots t \ldots)}\ \text{rep}$$

$$\frac{r(x_1, \ldots, x_n) = s(x_1, \ldots, x_n)}{r[t_1/x_1, \ldots, t_n/x_n] = s[t_1/x_1, \ldots, t_n/x_n]}\ \text{subst}$$

where $r(\ldots t \ldots)$ arises from $r(\ldots s \ldots)$ by replacing an occurrence of s in r by t.

$\mathrm{T} \vdash s = t$: An equation $s = t$ is a *theorem* of a theory T iff $s = t$ is the last member of some deduction sequence from T.

Meta-Theorems

Theorem (Soundness & Completeness)

$$\mathrm{T} \vdash s = t \iff \mathrm{T} \vDash s = t$$

- Determining Validity? Undecidable!

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

- Equational Logic
- Boolean Algebra
- Lambda Calculus and Combinatory Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Semigroup/Monoid/Group

Group $\mathscr{L} = \{e, \cdot, ^{-1}\}$

1. $\forall xyz : x \cdot (y \cdot z) = (x \cdot y) \cdot z$
2. $\forall x : e \cdot x = x$
3. $\forall x : x \cdot e = x$
4. $\forall x : x^{-1} \cdot x = e$
5. $\forall x : x \cdot x^{-1} = e$

Group $\mathscr{L} = \{\cdot\}$

1. $\forall xyz : x \cdot (y \cdot z) = (x \cdot y) \cdot z$
2. $\forall xy \exists z : x \cdot z = y$
3. $\forall xy \exists z : z \cdot x = y$

Group $\mathscr{L} = \{e, \cdot\}$

1. $\forall xyz : x \cdot (y \cdot z) = (x \cdot y) \cdot z$
2. $\forall x : e \cdot x = x \cdot e = x$
3. $\forall x \exists y : x \cdot y = y \cdot x = e$

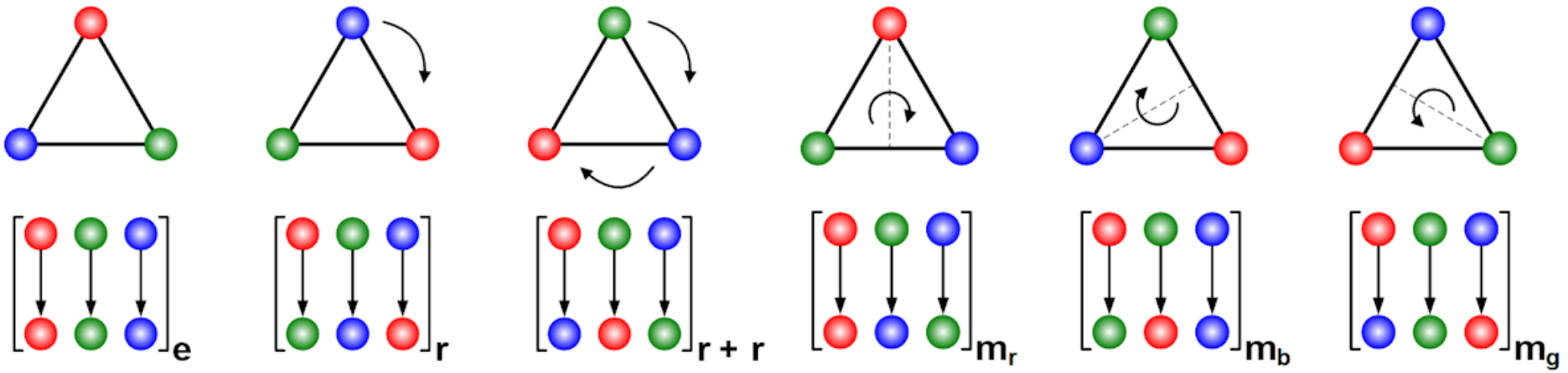
Model

- $(\mathbb{Z}, 0, +)$
- $(\mathbb{Q}^+, 1, \times)$
- Klein group: $(\{e, a, b, c\}, e, \cdot)$

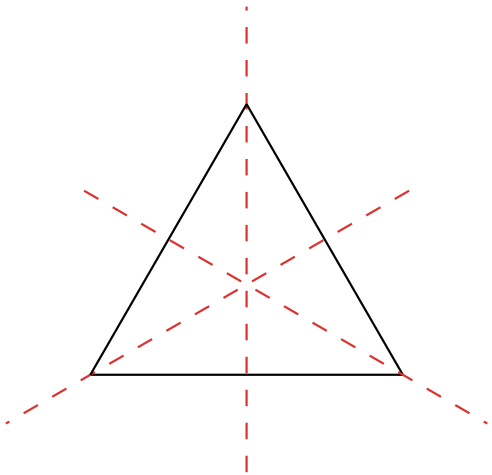
$\cdot$	e	a	b	c	permutation
e	e	a	b	c	e
a	a	e	c	b	$(1,2)(3,4)$
b	b	c	e	a	$(1,3)(2,4)$
c	c	b	a	e	$(1,4)(2,3)$

Examples of Groups

$$\begin{array}{ccccccc} \dagger & \clubsuit & \divideontimes & \oint & \blacklozenge & \nabla & \$ \\ \divideontimes & \blacklozenge & \nabla & \clubsuit & \oint & \$ & \dagger \end{array} \Longleftrightarrow \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 6 & 2 & 4 & 7 & 1 \end{array} = (1,3,6,7)(2,5,4)$$



$$\{e, r, r^2, m, mr, mr^2\}$$



Jump above calculations; group the operations, classify them according to their complexities rather than their appearances.

— Évariste Galois



Cayley Theorem

Theorem (Cayley Theorem)

Every group G is isomorphic to a subgroup of the symmetric group on G.

Proof.

Let $\lambda_g : x \mapsto g \cdot x$, and $T : g \mapsto \lambda_g$ for $g \in G$.
For every group $(G, e, \cdot)$, the function T embeds $(G, e, \cdot)$ in the group $(\mathrm{Aut}(G), 1_G, \circ)$.

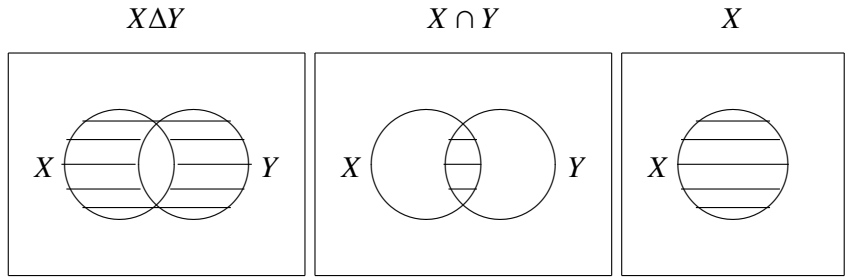
$$\lambda_e = 1_G \qquad \lambda_{g \cdot h} = \lambda_g \circ \lambda_h \qquad (\lambda_g)^{-1} = \lambda_{g^{-1}}$$

□

Ring/Boolean Ring

Ring $\mathscr{L} = \{0, 1, \oplus, \odot, -\}$

1. $\forall xyz : x \oplus (y \oplus z) = (x \oplus y) \oplus z$
2. $\forall xy : x \oplus y = y \oplus x$
3. $\forall x : x \oplus (-x) = 0$
4. $\forall x : x \oplus 0 = x$
5. $\forall xyz : x \odot (y \odot z) = (x \odot y) \odot z$
6. $\forall xyz : x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$
7. $\forall xyz : (x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$
8. $\forall x : x \odot 1 = 1 \odot x = x$
9. $0 \neq 1$



A *Boolean ring* is a ring for which

$$\forall x : x \odot x = x$$

Field $\mathscr{L} = \{0, 1, \oplus, -, \odot, ^{-1}\}$

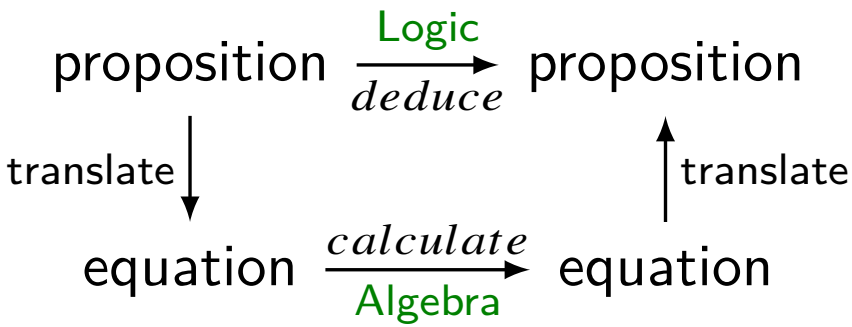
- $\forall xy : x \odot y = y \odot x$
- $\forall x \neq 0 : x \odot x^{-1} = x^{-1} \odot x = 1$

Example: $(\mathbb{Q}, 0, 1, +, -, \times, /)$

Boolean Algebra

Boolean Algebra $\mathscr{L} = \{0, 1, +, \cdot, \bar{}\}$

- $x + (y + z) = (x + y) + z$
  $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
- $x + y = y + x \quad x \cdot y = y \cdot x$
- $x + (x \cdot y) = x \quad x \cdot (x + y) = x$
- $x + (y \cdot z) = (x + y) \cdot (x + z)$
  $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$
- $\overline{\overline{x}} = x$
- $\overline{x + y} = \overline{x} \cdot \overline{y} \quad \overline{x \cdot y} = \overline{x} + \overline{y}$
- $x + \overline{x} = 1 \quad x \cdot \overline{x} = 0 \quad 0 \neq 1$
- $x + 0 = x \quad x \cdot 0 = 0$
  $x + 1 = 1 \quad x \cdot 1 = x$



Logic as Algebra

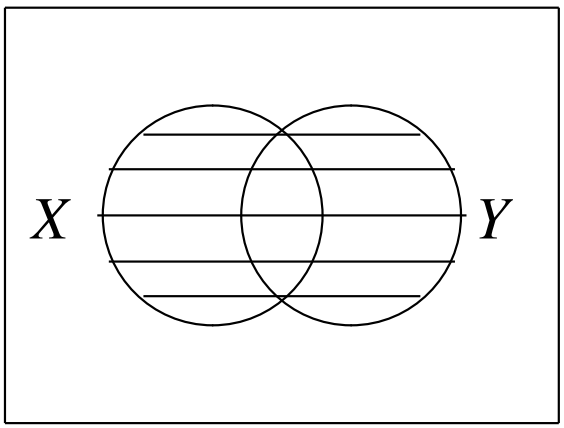
$$x - y := x \cdot \overline{y}$$

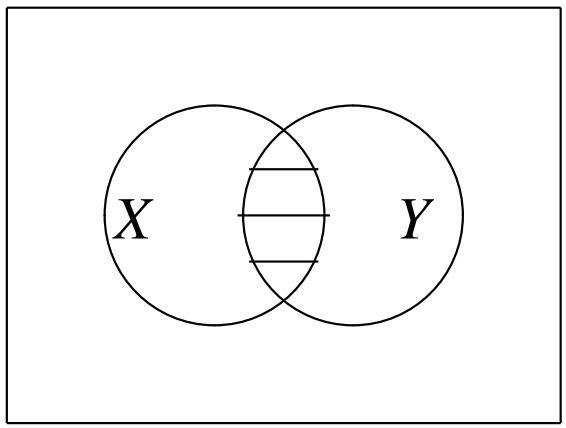
$$\overline{x} = 1 - x$$

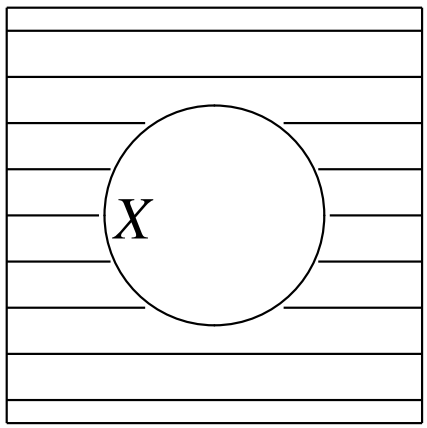
$$x \cdot (y - z) = x \cdot y - x \cdot z$$

Power Set Algebra

$$(\mathrm{P}(A), \emptyset, A, \cup, \cap, \overline{\ })$$

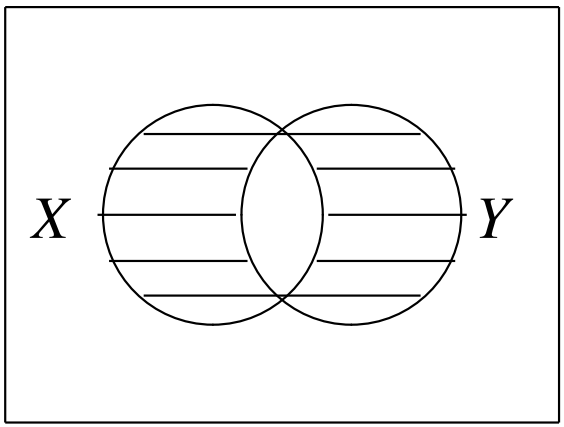






$$x \oplus y \coloneqq (x \cdot \overline{y}) + (\overline{x} \cdot y)$$

$$x = y \iff x \oplus y = 0$$



Boolean Ring vs Boolean Algebra

Boolean Ring to Boolean Algebra

$$x \cdot y = x \odot y$$
$$x + y = x \oplus y \oplus (x \odot y)$$
$$\overline{x} = 1 \oplus x$$

Boolean Algebra to Boolean Ring

$$x \odot y = x \cdot y$$
$$x \oplus y = (x \cdot \overline{y}) + (\overline{x} \cdot y)$$
$$-x = x$$

Boole's Four Main Theorems

$$\mathbf{x}^{\sigma} := x_1^{\sigma_1} \cdots x_n^{\sigma_n} \qquad x_i^{\sigma_i} := \begin{cases} x_i & \text{if } \sigma_i = 1 \\ \overline{x}_i & \text{if } \sigma_i = 0 \end{cases} \qquad \sigma : \{1, \ldots, n\} \to \{0, 1\}$$

1. Expansion

$$f(\mathbf{x}, \mathbf{y}) = \sum_{\sigma} f(\sigma, \mathbf{y}) \cdot \mathbf{x}^{\sigma}$$

2. Reduction

$$\bigwedge_i \big(f_i(\mathbf{x}) = 0\big) \iff \sum_i f_i(\mathbf{x}) = 0$$

3. Elimination

$$\exists \mathbf{x} \big(f(\mathbf{x}, \mathbf{y})\big) = 0 \iff \prod_{\sigma} f(\sigma, \mathbf{y}) = 0$$

4. Solution

$$q(\mathbf{y}) \cdot x = p(\mathbf{y})$$

$$\Updownarrow$$

$$p(\mathbf{y}) \cdot \big(p(\mathbf{y}) - q(\mathbf{y})\big) = 0 \quad \& \quad \exists v : x = \sum_{\tau : p(\tau) = q(\tau) \neq 0} \mathbf{y}^{\tau} + v \cdot \sum_{\tau : p(\tau) = q(\tau) = 0} \mathbf{y}^{\tau}$$

Boole's Method

1. Translation. Translate premises into equational form.
   A: $x \cdot \overline{y} = 0$; **E**: $x \cdot y = 0$; **I**: $v = v \cdot x \cdot y$; **O**: $v = v \cdot x \cdot \overline{y}$.
2. Reduction. Combine the premise-equations into a single equation.

$$f_1(\mathbf{x}) = 0, \ldots, f_k(\mathbf{x}) = 0 \iff \sum_{i=1}^{k} f_i(\mathbf{x}) = 0$$

3. Elimination. Given the single premise $\sum_{i=1}^{k} f_i(\mathbf{y}, \mathbf{z}) = 0$, the most general conclution involving only $\mathbf{z}$ is $f(\mathbf{z}) = 0$, where

$$f(\mathbf{z}) \coloneqq \left(\sum_{i=1}^{k} f_i(1, \ldots, 1, \mathbf{z})\right) \cdot \ldots \cdot \left(\sum_{i=1}^{k} f_i(0, \ldots, 0, \mathbf{z})\right)$$

4. Expansion. $f(\mathbf{z}) = f(1, \ldots, 1) \cdot z_1 \cdot \ldots \cdot z_n + \ldots + f(0, \ldots, 0) \cdot \overline{z}_1 \cdot \ldots \cdot \overline{z}_n$
5. Translation. Interpret the conclusion-equations as propositions.

Boole's Method — Syllogism

A : $x \cdot \overline{y} = 0;$ **E** : $x \cdot y = 0;$ **I** : $v = v \cdot x \cdot y;$ **O** : $v = v \cdot x \cdot \overline{y}$

$$\frac{MAP \quad SAM}{SAP}$$

$$\underbrace{m \cdot \overline{p} = 0 \qquad s \cdot \overline{m} = 0}$$

$$\Downarrow \; Reduction$$

$$m \cdot \overline{p} + s \cdot \overline{m} = 0$$

$$\Downarrow \; Elimination$$

$$(1 \cdot \overline{p} + s \cdot 0) \cdot (0 \cdot \overline{p} + s \cdot 1) = 0$$

$$\Downarrow \; Expansion$$

$$(1 \cdot 0 + 1 \cdot 0) \cdot (0 \cdot 1 + 1 \cdot 1) \cdot s \cdot p + \ldots + (1 \cdot 1 + 0 \cdot 0) \cdot (0 \cdot 1 + 0 \cdot 1) \cdot \overline{s} \cdot \overline{p} = 0$$

$$\Downarrow$$

$$s \cdot \overline{p} = 0$$

$$s \cdot \overline{p} = s \cdot 1 \cdot \overline{p} = s \cdot (m + \overline{m}) \cdot \overline{p} = s \cdot m \cdot \overline{p} + s \cdot \overline{m} \cdot \overline{p} = 0 + 0 = 0$$

Boole's Method — Syllogism

A : $x \cdot \overline{y} = 0;$ **E** : $x \cdot y = 0;$ **I** : $v = v \cdot x \cdot y;$ **O** : $v = v \cdot x \cdot \overline{y}$

$$\begin{array}{c} PAM \\ SOM \\ \hline SOP \end{array} \qquad \begin{array}{c} p \cdot \overline{m} = 0 \\ v = v \cdot s \cdot \overline{m} \\ \hline v = v \cdot s \cdot \overline{p} \end{array}$$

$$v \cdot s \cdot \overline{p} \overset{2}{=} v \cdot s \cdot \overline{m} \cdot s \cdot \overline{p} = v \cdot s \cdot \overline{m} \cdot \overline{p} \overset{1}{=} v \cdot s \cdot \overline{m} \cdot \overline{p} + v \cdot s \cdot \overline{m} \cdot p = v \cdot s \cdot \overline{m} \overset{2}{=} v$$

$$\begin{array}{c} MEP \\ MIS \\ \hline SOP \end{array} \qquad \begin{array}{c} m \cdot p = 0 \\ v = v \cdot m \cdot s \\ \hline v = v \cdot s \cdot \overline{p} \end{array}$$

$$v \cdot s \cdot \overline{p} \overset{2}{=} v \cdot m \cdot s \cdot s \cdot \overline{p} = v \cdot s \cdot m \cdot \overline{p} \overset{1}{=} v \cdot s \cdot m \cdot \overline{p} + v \cdot s \cdot m \cdot p = v \cdot s \cdot m \overset{2}{=} v$$

Boole's Method — Syllogism — Another Translation

A : $x \cdot \overline{y} = 0$; **E** : $x \cdot y = 0$; **I** : $x \cdot y \neq 0$; **O** : $x \cdot \overline{y} \neq 0$.

$$\frac{\begin{array}{c} PAM \\ SOM \end{array}}{SOP} \qquad \frac{\begin{array}{c} MEP \\ MIS \end{array}}{SOP}$$

$$p \cdot \overline{m} = s \cdot \overline{m} \cdot p + \overline{s} \cdot \overline{m} \cdot p = 0$$
$$s \cdot \overline{m} = s \cdot \overline{m} \cdot p + s \cdot \overline{m} \cdot \overline{p} \neq 0$$
$$\Downarrow$$
$$p \cdot \overline{m} + s \cdot \overline{m} = s \cdot \overline{m} \cdot \overline{p} \neq 0$$
$$\Downarrow$$
$$s \cdot \overline{m} \cdot \overline{p} + s \cdot m \cdot \overline{p} = s \cdot \overline{p} \neq 0$$

$$m \cdot p = s \cdot m \cdot p + \overline{s} \cdot m \cdot p = 0$$
$$m \cdot s = s \cdot m \cdot p + s \cdot m \cdot \overline{p} \neq 0$$
$$\Downarrow$$
$$m \cdot p + m \cdot s = s \cdot m \cdot \overline{p} \neq 0$$
$$\Downarrow$$
$$s \cdot m \cdot \overline{p} + s \cdot \overline{m} \cdot \overline{p} = s \cdot \overline{p} \neq 0$$

Propositional Logic vs Boolean Algebra

$$
\begin{aligned}
(\bot)^* &:= 0 \\
(\top)^* &:= 1 \\
(p)^* &:= p \\
(\neg A)^* &:= \overline{A^*} \\
(A \vee B)^* &:= A^* + B^* \\
(A \wedge B)^* &:= A^* \cdot B^*
\end{aligned}
\qquad\qquad
\begin{aligned}
(0)^\star &:= \bot \\
(1)^\star &:= \top \\
(p)^\star &:= p \\
\left(\overline{A}\right)^\star &:= \neg A^\star \\
(A + B)^\star &:= A^\star \vee B^\star \\
(A \cdot B)^\star &:= A^\star \wedge B^\star
\end{aligned}
$$

$$
\frac{A \vdash B}{A^* \leq B^*}
\qquad\qquad
\frac{A \leq B}{A^\star \vdash B^\star}
$$

where $x \leq y := x \cdot \overline{y} = 0$

General Solution?[28]

Exercise — Save Yourself

You can say one sentence. If you lie I will hang you. If you tell the truth I will shoot you.

$$x =? \implies \neg x \to h, x \to s, h \leftrightarrow \neg s \vDash \neg h \wedge \neg s$$

$$\neg h \to h, h \to s, h \leftrightarrow \neg s \vDash \neg h \wedge \neg s$$

Problem (General Solution?)

$$x =? \implies \vDash A(x)$$

[28] Frank Markham Brown: Boolean Reasoning — The Logic of Boolean Equations.

General Solution?

$$x^5 - x - 1 = 0$$

$$x = ?$$

There is no solution in radicals to general polynomial equations of degree five or higher with arbitrary coefficients.

$$ax^2 + bx + c = 0$$

$$x = ?$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Theorem (General Solution of Boolean Equation)

Let $f : B \to B$ be a Boolean function for which $f(x) = 0$ is consistent (it has at least one solution, i.e. $f(0) \cdot f(1) = 0$), then

$$f(x) = 0$$

$$\Updownarrow$$

$$f(1) \cdot x + f(0) \cdot \overline{x} = 0$$

$$\Updownarrow$$

$$x = \overline{f(1)} \cdot x + f(0) \cdot \overline{x}$$

$$\Updownarrow$$

$$f(0) \leq x \leq \overline{f(1)}$$

$$\Updownarrow$$

$$x = f(0) + \overline{f(1)} \cdot \theta \quad \textit{where} \quad \theta \in B$$

Application — How to Flirt with a Beauty ☺

Smullyan Flirts with a Beauty ♡◎♡

1. "I am to make a statement. If it is true, would you give me your autograph?"
2. "I don't see why not."
3. "If it is false, do not give me your autograph."
4. "Alright."
5. Then Smullyan said such a sentence that she have to give him a kiss.

$$x =? \implies a \leftrightarrow x \vDash k$$

Application — How to Flirt with a Beauty ☺

Smullyan Flirts with a Beauty ♡◎♡

1. "I am to make a statement. If it is true, would you give me your autograph?"
2. "I don't see why not."
3. "If it is false, do not give me your autograph."
4. "Alright."
5. Then Smullyan said such a sentence that she have to give him a kiss.

$$x =? \implies a \leftrightarrow x \vDash k$$

Solution

$$(a \cdot x + \overline{a} \cdot \overline{x}) \cdot \overline{k} = 0 \quad \Longrightarrow \quad x = \overline{a} \cdot \overline{k} + \theta \cdot (\overline{a} + k)$$

$$a \leftrightarrow \neg a \wedge \neg k \vDash k$$

$$a \leftrightarrow (a \rightarrow k) \vDash k$$

General Solution of Boolean Equation

Theorem (General Solution of Boolean Equation)

Given the Boolean function $f : B^n \to B$, *define*
$f_0, f_1(x_1), f_2(x_1, x_2), \ldots, f_n(x_1, \ldots, x_n)$ *by means of the recursion*

$$f_n \coloneqq f$$
$$f_{i-1}(x_1, \ldots, x_{i-1}) \coloneqq f_i(x_1, \ldots, x_{i-1}, 0) \cdot f_i(x_1, \ldots, x_{i-1}, 1)$$

then

$$f_0 = 0$$
$$f_i(x_1, \ldots, x_{i-1}, 0) \leq x_i \leq \overline{f_i(x_1, \ldots, x_{i-1}, 1)} \qquad (i = 1, \ldots, n)$$

is a genereal solution of $f(x_1, \ldots, x_n) = 0$.

Homomorphism

- Let $\mathcal{M}$ and $\mathcal{N}$ be Boolean algebras. A (Boolean) homomorphism is a mapping $h : M \to N$ s.t. for all $x, y \in M$:
  1. $h(0) = 0$
  2. $h(1) = 1$
  3. $h(\overline{x}) = \overline{h(x)}$
  4. $h(x + y) = h(x) + h(y)$
  5. $h(x \cdot y) = h(x) \cdot h(y)$
- If $h : M \rightarrowtail N$, it is an isomorphic embedding of $\mathcal{M}$ into $\mathcal{N}$.
- If $h : M \twoheadrightarrow N$, then $\mathcal{M}$ and $\mathcal{N}$ are isomorphic ($\mathcal{M} \cong \mathcal{N}$).

Example — Lindenbaum Algebra of Propositional Logic

Lindenbaum Algebra of Propositional Logic

$$\mathrm{Lin} := (\mathrm{Wff}/{\sim}, 0, 1, +, \cdot, {}^{-})$$

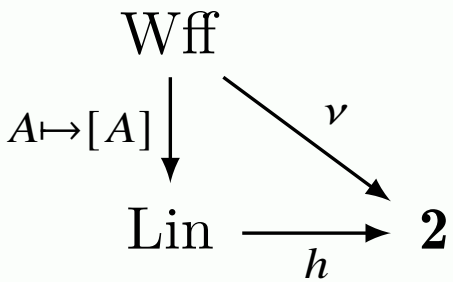
$$\mathbf{2} := \left(\{0,1\}, 0, 1, \max, \min, 1 - \right)$$

where

$$\begin{aligned}
A \sim B &:= \vdash A \leftrightarrow B \\
[A] &:= \{B \in \mathrm{Wff} : A \sim B\} \\
\mathrm{Wff}/{\sim} &:= \{[A] : A \in \mathrm{Wff}\} \\
0 &:= [\bot] \\
1 &:= [\top] \\
\overline{[A]} &:= [\neg A] \\
[A] + [B] &:= [A \vee B] \\
[A] \cdot [B] &:= [A \wedge B]
\end{aligned}$$

$$\mathrm{Lin} \overset{?}{\cong} \mathbf{2}$$

If ν is a truth assignment, then $h : [A] \mapsto \nu(A)$ is a homomorphism.
If $h : \mathrm{Lin} \to \mathbf{2}$ is a homomorphism, then $\nu : A \mapsto h([A])$ is a truth assignment.



Filter & Ultrafilter

- A partial order R over P is a binary relation which is reflexive, antisymmetric, and transitive, i.e. for all $x, y, z \in P$:
  1. Rxx
  2. $Rxy \wedge Ryx \rightarrow x = y$
  3. $Rxy \wedge Ryz \rightarrow Rxz$
- $\leq$ is a partial order.
- Let $\mathcal{B} = (B, 0, 1, +, \cdot, \bar{})$ be a Boolean algebra. A subset $F \subset B$ is a filter iff
  1. $1 \in F$
  2. $x \in F \wedge x \leq y \rightarrow y \in F$
  3. $x \in F \wedge y \in F \rightarrow x \cdot y \in F$
- A filter F is *proper* iff $0 \notin F$.
- A proper filter F is an *ultrafilter* iff either $x \in F$ or $\overline{x} \in F$.

Theorem

Let F be a filter in a Boolean algebra B. Then the following conditions are equivalent:

1. $B/F \cong \mathbf{2}$
2. *F is the hull of a 2-valued homomorphism h on B*
3. *F is an ultrafilter*
4. *for all $x, y \in B$, if $x \vee y \in F$ then $x \in F$ or $y \in F$*
5. *for all $x \in B$, either $x \in F$ or $\overline{x} \in F$*

Theorem (Ultrafilter Theorem)

Every proper filter can be extended to an ultrafilter.

Ultrafilter theorem on Lindenbaum Algebra of Propositional Logic $\Longleftrightarrow$
Every consistent set can be extended to a maximal consistent set.

Stone's Representation Theorem

- There are Boolean algebras which are not isormophic to any power set algebra.
- For example, the finite-cofinite algebra F_ω of the set of natural numbers.
- Since F_ω has cardinality $\aleph_0$, while no power set algebra can have this cardinality.

Theorem (Stone's Representation Theorem)

Every Boolean algebra is isomorphic to a subalgebra of a power set algebra.

Proof.

Let $\mathcal{B}$ be a Boolean algebra, and $\mathrm{St}(B) := \{w : w \text{ is an ultrafilter on } B\}$. Define a map $h : B \to \mathtt{P}(\mathrm{St}(B))$ by

$$x \mapsto \{w \in \mathrm{St}(B) : x \in w\}$$

Then

$$\mathcal{B} \cong (h(B), \emptyset, \mathrm{St}(B), \cup, \cap, \bar{})$$

$$\boxed{\vDash A \implies \vdash A}$$

$$\begin{array}{c} \nvdash A \\ \Downarrow \\ [A] \neq [\top] \\ \Downarrow \\ [\neg A] \neq [\bot] \\ \Downarrow \\ h\left([\neg A]\right) \neq \emptyset \\ \Downarrow \\ \exists w \in \mathrm{St}(\mathrm{Lin})\big([\neg A] \in w\big) \\ \Downarrow \\ \chi_w\left([\neg A]\right) = 1 \end{array}$$

Compactness in Topology $\iff$ Compactness in Logic

Compactness in Topology $\implies$ Compactness in Logic

Let $\mathbf{2} \coloneqq \{0, 1\}$ be the discrete topology.
For any formula A, let $U_A \coloneqq \mathrm{Mod}(A) = \left\{\nu \in \mathbf{2}^{\mathrm{Var}} : \nu \vDash A\right\}$.
Then $\mathcal{B} \coloneqq \{U_A : A \in \mathrm{Wff}\}$ is a basis for a topology on $\mathbf{2}^{\mathrm{Var}}$.
Let τ be the topology on $\mathbf{2}^{\mathrm{Var}}$ generated by $\mathcal{B}$.
$(\mathbf{2}^{\mathrm{Var}}, \tau)$ is a compact, Hausdorff, totally disconnected topological space.
It can be shown that U_A is clopen.
By hypothesis, for each finite $\Delta \subset \Gamma$, $U_\Delta \neq \emptyset$. That is to say, $\{U_A : A \in \Gamma\}$ has the Finite Intersection Property. By the compactness of τ, $U_\Gamma \neq \emptyset$.

Compactness in Logic $\implies$ Compactness in Topology

Γ is unsatisfiable iff $\{U_{\neg A} : A \in \Gamma\}$ covers $\mathbf{2}^{\mathrm{Var}}$.
Given any cover $\{U_A\}_{A \in \Gamma}$ of $\mathbf{2}^{\mathrm{Var}}$. Then $\Gamma^* \coloneqq \{\neg A : A \in \Gamma\}$ is unsatisfiable. By compactness theorem, some finite subset $\Delta \subset \Gamma^*$ is unsatisfiable. The set $\{U_A\}_{\neg A \in \Delta} \subset \{U_A\}_{A \in \Gamma}$ is a finite subcover for $\mathbf{2}^{\mathrm{Var}}$.
So $(\mathbf{2}^{\mathrm{Var}}, \tau)$ is compact.

拓扑紧致 vs 一阶逻辑紧致

Lemma

设 $\mathscr{L}_1$ 为一阶语言. 令

$$X \coloneqq \{(\mathcal{M}, v) : \mathcal{M} \text{ 是 } \mathscr{L}_1\text{-结构且 } v \text{ 是 } \mathcal{M}\text{-赋值}\}$$
$$\tau \coloneqq \{U_\Delta : \Delta \subset \mathscr{L}_1\}, \text{ 其中 } U_\Delta \coloneqq \{(\mathcal{M}, v) \in X : \mathcal{M}, v \nvDash \Delta\}$$

则 τ 是 X 上的拓扑.

Lemma

τ 的闭集都形如 $C_\Delta = \{(\mathcal{M}, v) \in X : \mathcal{M}, v \vDash \Delta\}$.

Theorem

如下命题等价:

1. 拓扑空间 (X, τ) 是紧的.
2. 如果公式集 Γ 是有穷可满足的, 则它是可满足的.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

- Equational Logic
- Boolean Algebra
- Lambda Calculus and Combinatory Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Church 1903-1995



- Lambda Calculus
- Church-Turing Thesis
- Undecidability
- Church-Rosser Theorem
- Frege-Church Ontology

Lambda Calculus

$\mathscr{L} = \{\lambda, .\}$

Definition (λ-Terms)

$$\Lambda \coloneqq x \mid \Lambda\Lambda \mid \lambda x.\Lambda$$

Notation:

- $M_0 M_1 \cdots M_n$ denotes $(\cdots((M_0 M_1) M_2 \cdots M_n)$
- $\lambda x_0 x_1 \cdots x_n.M$ denotes $(\lambda x_0.(\lambda x_1.(\cdots(\lambda x_n.M))\cdots))$

Definition (Free Variable)

$$\mathrm{Fv}(\Lambda) \coloneqq \begin{cases} \{x\} & \text{if } \Lambda = x \\ \mathrm{Fv}(M) \cup \mathrm{Fv}(N) & \text{if } \Lambda = MN \\ \mathrm{Fv}(M) \setminus \{x\} & \text{if } \Lambda = \lambda x.M \end{cases}$$

Reduction Rules

Definition (Substitution)

$$y[N/x] = \begin{cases} N & \text{if } x = y \\ y & \text{otherwise} \end{cases}$$

$$(M_1 M_2)[N/x] = (M_1[N/x])\,(M_2[N/x])$$

$$(\lambda y.M)[N/x] = \begin{cases} \lambda y.M & \text{if } x = y \\ \lambda y.M[N/x] & \text{if } x \neq y \text{ and } y \notin \mathrm{Fv}(N) \end{cases}$$

Reduction Rules

$$\lambda x.M \overset{\alpha}{=} \lambda y.M[y/x] \quad \text{if } y \text{ does not occur in } M.$$

$$(\lambda x.M)N \overset{\beta}{=} M[N/x]$$

$$\lambda x.Mx \overset{\eta}{=} M \quad \text{if } x \notin \mathrm{Fv}(M)$$

λ-definability

Definition (Church Numeral)

$$\underline{n} := \lambda f x.f^n x$$
$$f^0 x := x$$
$$f^{n+1} x := f(f^n x)$$

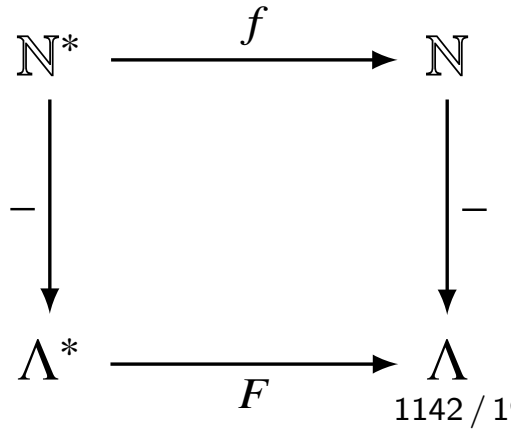
Definition (λ-definability)

An n-ary function $f(x_1, \ldots, x_n)$ is λ-definable iff there is a λ-term F s.t. for all $a_1, \ldots, a_n$,

$$F\underline{a_1} \ldots \underline{a_n} \overset{\beta}{=} \underline{f(a_1, \ldots, a_n)}$$

A function f is computable iff it is λ-definable.

$$\mathbf{succ} := \lambda n f x.f(nfx)$$
$$\mathbf{add} := \lambda mnfx.mf(nfx)$$
$$\mathbf{mult} := \lambda mnf.m(nf)$$
$$\mathbf{exp} := \lambda mn.nm$$



Combinator

Definition (Combinator)

A λ-term M is called a combinator iff $\mathrm{Fv}(M) = \emptyset$.

$$\mathbf{K} = \lambda xy.x$$
$$\mathbf{S} = \lambda xyz.xz(yz)$$
$$\mathbf{I} = \lambda x.x$$
$$\omega = \lambda x.xx$$
$$\Omega = \omega\omega$$
$$\mathbf{Y} = \lambda y.(\omega(\lambda x.y(xx)))$$

$$\mathbf{if_then_else} = \lambda bxy.bxy$$

$$\mathbf{F} = \lambda xy.y$$
$$\mathbf{T} = \mathbf{K}$$
$$\mathbf{B} = \mathbf{S}(\mathbf{KS})\mathbf{K}$$
$$\mathbf{C} = \mathbf{S}(\mathbf{BBS})(\mathbf{KK})$$
$$\mathbf{W} = \mathbf{SS}(\mathbf{SK})$$
$$\mathbf{D} = \mathbf{SII}$$
$$\mathbf{L} = \mathbf{D}(\mathbf{BDD})$$
$$\mathbf{neg} = \lambda x.x\mathbf{FT}$$
$$\mathbf{and} = \lambda xy.xy\mathbf{F}$$
$$\mathbf{or} = \lambda xy.x\mathbf{T}y$$
$$\mathbf{iszero} = \lambda x.x(\lambda y.\mathbf{F})\mathbf{T}$$

$$\iota = \lambda x.x\mathbf{SK}$$

$$\mathbf{K} = \iota(\iota(\iota\iota))$$
$$\mathbf{S} = \iota(\iota(\iota(\iota\iota)))$$
$$\mathbf{I} = \iota\iota$$

Iota

Syntax	Semantics	
$F \to 1F_0F_1$	$[F_0]([F_1])$	"application"
$F \to 0$	$\lambda x.xSK$	"Iota"

Example

$$
\begin{aligned}
100 &= (\lambda x.xSK)(\lambda x.xSK) \\
&= (\lambda x.xSK)SK \\
&= SSKK \\
&= SK(KK) \\
&= I
\end{aligned}
$$

$$
\begin{aligned}
1010100 &= 1010I \\
&= (\lambda x.xSK)((\lambda x.xSK)I) \\
&= (\lambda x.xSK)(ISK) \\
&= (\lambda x.xSK)(SK) \\
&= SKSK \\
&= KK(SK) \\
&= K
\end{aligned}
$$

Exercises

- $\mathbf{B}xyz = x(yz)$ (composition)
- $\mathbf{C}xyz = xzy$ (swap)
- $\mathbf{W}xy = xyy$ (duplicate)
- $\mathbf{D}x = xx$ (doubling)
- $\mathbf{L} = \mathbf{LL}$ (self-doubling)

Smullyan: To Mock a Mockingbird

- A forest is inhabited by talking birds. Given any birds A, B, if you call out the name of B to A, then A will respond by calling out the name of some bird AB. If $AB = B$, then we say that A is fond of B.
- A bird x is called *egocentric* if x is fond of itself $xx = x$.

The forest satisfies the following two conditions:

1. For any birds A, B, there is a bird C s.t. for any bird x, $Cx = A(Bx)$.
2. There is a mockingbird M s.t. for any bird x, $Mx = xx$.

Theorem

Every bird is fond of some bird, and at least one bird is egocentric.

Proof.

Take any bird A. There is a bird C such that for any bird x, $Cx = A(Mx)$.
Then taking C for x, $CC = A(MC) = A(CC)$.
In particular, the mocking bird M is fond of some bird E. Thus $ME = E$, but also $ME = EE$. Therefore $EE = E$. □

Smullyan: To Mock a Mockingbird

- A bird Θ is called a *Sage* bird if you call out the name of a bird x to it, it will name a bird of which x is fond $\Theta x = x(\Theta x)$.
- A bird U is called a *Turing* bird if $Uxy = y(xxy)$.
- A bird L is called a *Lark* if $Lxy = x(yy)$.
- A bird O is called an *Owl* if $Oxy = y(xy)$.
- A bird I is called an *Idiot* if $Ix = x$.
- A bird K is called a *Kestrel* if $Kxy = x$.
- A bird S is called a *Starling* if $Sxyz = xz(yz)$.

1. If there is a lark in the forest, then every bird is fond of some bird. $Lx(Lx) = x(Lx(Lx))$.
2. SI is an Owl.
3. An owl is fond only of Sage birds. $OA = A \implies Ax = x(Ax)$.
4. Turing birds: LO, $L(SI)$.
5. Sage birds: ΘO, $O\Theta$, UU, SLL.
6. Mocking birds: OI, LI, SII.
7. ΘM is egocentric. $\Theta M = M(\Theta M) = \Theta M(\Theta M)$.
8. ΘK is hopelessly egocentric. $\Theta Kx = K(\Theta K)x = \Theta K$.

First Fixpoint Theorem in Lambda Calculus

Theorem (First Fixpoint Theorem in Lambda Calculus)

For every λ-term F there is a λ-term E s.t. $FE = E$.

Proof.

Let $G \coloneqq \lambda x.F(xx)$ and $E \coloneqq GG$. □

$$\mathbf{Y} = \lambda y.(\lambda x.y(xx))(\lambda x.y(xx))$$

Corollary

For any λ-term $C(f, \vec{x})$, there exists a λ-term M s.t. for all λ-terms $\vec{N}$

$$M\vec{N} = C(M, \vec{N})$$

Proof.

Let $M \coloneqq \mathbf{Y}(\lambda f\vec{x}.C(f, \vec{x}))$. □

Fixpoint Combinator

$$\mathbf{Y} = \lambda y.(\lambda x.y(xx))(\lambda x.y(xx)) \qquad \textbf{Curry}$$

$$\Theta = (\lambda xy.y(xxy))(\lambda xy.y(xxy)) \qquad \textbf{Turing}$$

$$\textbf{fac}\ n = \textbf{if_then_else}\ (\textbf{iszero}\ n)\ (1)\ (\textbf{mult}\ n\ (\textbf{fac}\ (\textbf{pred}\ n)))$$

$$\textbf{fac} = \lambda n.\textbf{if_then_else}\ (\textbf{iszero}\ n)\ (1)\ (\textbf{mult}\ n\ (\textbf{fac}\ (\textbf{pred}\ n)))$$

$$\textbf{fac} = (\lambda f.\lambda n.\textbf{if_then_else}\ (\textbf{iszero}\ n)\ (1)\ (\textbf{mult}\ n\ (f\ (\textbf{pred}\ n))))\ \textbf{fac}$$

$$F := \lambda f.\lambda n.\textbf{if_then_else}\ (\textbf{iszero}\ n)\ (1)\ (\textbf{mult}\ n\ (f\ (\textbf{pred}\ n)))$$

$$\textbf{fac} := \mathbf{Y}F$$

$$\mathbf{Y}F = F(\mathbf{Y}F)$$

$$\textbf{fac} = F\ \textbf{fac}$$

Church-Rosser Theorem

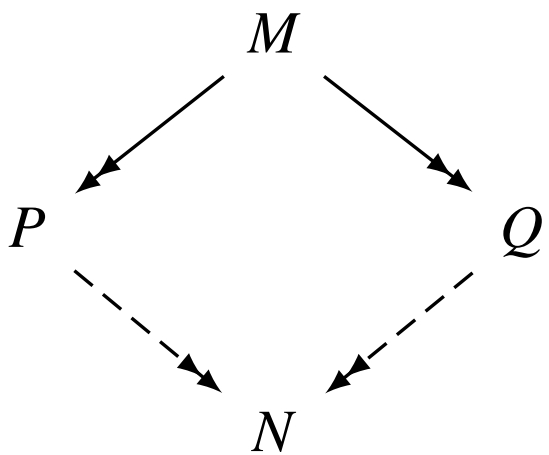
We write $M \twoheadrightarrow_\beta N$ iff M β-reduces to N in zero or more steps.

Definition (β-nf)

A term is in β normal form iff it can't be β-reduced.
A term M has a β normal form iff it β reduces to some N that is in β-nf.

Theorem (Church-Rosser Theorem)

Let $\twoheadrightarrow$ denote either $\twoheadrightarrow_\beta$ or $\twoheadrightarrow_{\beta\eta}$. Suppose M, P, Q are λ-terms s.t. $M \twoheadrightarrow P$ and $M \twoheadrightarrow Q$. Then there exists a λ-term N s.t. $P \twoheadrightarrow N$ and $Q \twoheadrightarrow N$.



Second Fixpoint Theorem in Lambda Calculus

We write $\ulcorner M \urcorner := \overline{\#M}$ to denote the Church numeral representing the Gödel number of M.

Theorem (Second Fixpoint Theorem in Lambda Calculus)

For every λ-term F there is a λ-term E s.t. $F\ulcorner E \urcorner = E$.

Proof.

By Church-Turing thesis, there is a term C s.t. $C\ulcorner M \urcorner = \ulcorner\ulcorner M \urcorner\urcorner$.
Furthermore, there is a term A s.t. $A\ulcorner M \urcorner\ulcorner N \urcorner = \ulcorner MN \urcorner$.
Take $G := \lambda n.F(An(Cn))$. Then let $E := G\ulcorner G \urcorner$.

$$\begin{aligned} E &= G\ulcorner G \urcorner \\ &= F(A\ulcorner G \urcorner(C\ulcorner G \urcorner)) \\ &= F(A\ulcorner G \urcorner(\ulcorner\ulcorner G \urcorner\urcorner)) \\ &= F\ulcorner G\ulcorner G \urcorner\urcorner \\ &= F\ulcorner E \urcorner \end{aligned}$$

□

Undecidability

Theorem (Church1936)

There is no term that will decide whether two terms have the same normal form.

Theorem (Church1936)

There is no λ-term D s.t. for all $\underline{n}$,

$$D\underline{n} = \begin{cases} \underline{0} & \text{if term with Gödel number } n \text{ has a } \beta\text{-nf} \\ \underline{1} & \text{otherwise} \end{cases}$$

Proof.

Suppose there was such a D. Then define $G := \lambda n.\mathbf{iszero}(Dn)\mathbf{\Omega I}$.
By the fixpoint theorem, there is X s.t. $G(\ulcorner X \urcorner) = X$.
X has a β-nf $\implies D\ulcorner X \urcorner = \underline{0} \implies G\ulcorner X \urcorner = \mathbf{\Omega} \implies X$ has no β-nf
X has no β-nf $\implies D\ulcorner X \urcorner = \underline{1} \implies G\ulcorner X \urcorner = \mathbf{I} \implies X$ has a β-nf $\square$

Undecidability

Theorem (Church1936)

There is no D s.t. for all M,

$$DM = \begin{cases} \mathbf{T} & \text{if } M \text{ has a normal form} \\ \mathbf{F} & \text{otherwise} \end{cases}$$

Proof.

let $G := \mathbf{C}(\mathbf{C}(\mathbf{B}D(\mathbf{SII}))\Omega)\mathbf{I}$ and $X := GG$. Then

$$X = D(X)\Omega\mathbf{I}$$

If X has a normal form, then $D(X)\Omega\mathbf{I} = \Omega$, but Ω has no normal form.
If X has no normal form, then $D(X)\Omega\mathbf{I} = \mathbf{I}$, but $\mathbf{I}$ is in normal form. □

Theorem (Curry, Scott, Rice)

Suppose $A \subset \Lambda$ is closed under $\overset{\beta}{=}$. Then A is decidable iff $A = \Lambda$ or $A = \emptyset$.

Proof.

Define $B \coloneqq \{M : M^{\ulcorner} M^{\urcorner} \in A\}$.
There exists a term $D \in \Lambda$ s.t.

$$M \in B \iff D^{\ulcorner} M^{\urcorner} = \underline{0}$$
$$M \notin B \iff D^{\ulcorner} M^{\urcorner} = \underline{1}$$

Let $P \in A$ and $Q \in \Lambda \setminus A$.

$$G \coloneqq \lambda n.\mathbf{iszero}(Dn)QP$$

$$G \in B \iff D^{\ulcorner} G^{\urcorner} = \underline{0} \implies G^{\ulcorner} G^{\urcorner} = Q \implies G^{\ulcorner} G^{\urcorner} \notin A \implies G \notin B$$
$$G \notin B \iff D^{\ulcorner} G^{\urcorner} = \underline{1} \implies G^{\ulcorner} G^{\urcorner} = P \implies G^{\ulcorner} G^{\urcorner} \in A \implies G \in B$$

□

Theorem (Enumeration Theorem)

There exists a term $\mathbf{E} \in \Lambda^0$ *such that, for all* $M \in \Lambda^0$

$$\mathbf{E}\ulcorner M \urcorner \twoheadrightarrow_\beta M$$

Theorem

There is no term $Q \in \Lambda$ *such that, for all* $M \in \Lambda$

$$QM =_\beta \ulcorner M \urcorner$$

Proof.

We know that Church numerals are in normal form. However

$$\ulcorner M \urcorner =_\beta QM =_\beta Q(\mathbf{I}M) =_\beta \ulcorner \mathbf{I}M \urcorner$$

which makes two distinct normal forms equal. □

Remark: One can go from intension $\ulcorner M \urcorner$ to extension M, but not the other way within the system itself.

$$\left.\begin{array}{c}\text{Second Fixpoint Theorem}\\ \text{Enumeration Theorem}\end{array}\right\} \Longrightarrow \text{First Fixpoint Theorem}$$

Proof.

$$M =_\beta \lambda x.F(\mathbf{E}x)\ulcorner M\urcorner =_\beta F(\mathbf{E}\ulcorner M\urcorner) =_\beta FM$$

□

Combinatory Logic

Definition (Combinatory Terms)

$$C \coloneqq x \mid \mathbf{K} \mid \mathbf{S} \mid (CC)$$

Reduction

$$\mathbf{K}MN = M$$
$$\mathbf{S}MNL = ML(NL)$$

- $ex \simeq \varphi_e(x)$
- $\varphi_k(x, y) \simeq x$
- $\varphi_s(x, y, z) \simeq \varphi_{\varphi_x(z)}(\varphi_y(z))$

- $\mathbf{I} = \mathbf{SKK}$
- $f \circ g = \mathbf{S}(\mathbf{K}f)g$

Combinatory Completeness

Proposition (Combinatory Completeness)

For every λ-term P and variable x, there is a combinator $\lambda^* x.P$ s.t.

$$(\lambda^* x.P)Q = P[Q/x]$$

Proof.

$$\lambda^* x.P := \begin{cases} \mathbf{I} & \text{if } P = x \\ \mathbf{K}P & \text{if } x \notin \mathrm{Fv}(P) \\ \mathbf{S}(\lambda^* x.M)(\lambda^* x.N) & \text{if } P = MN \end{cases}$$

□

Combinatory Logic subsumes Lambda Calculus

M	$(M)_{\mathbf{C}}$
x	x
$\lambda x.P$	$\lambda^* x.(P)_{\mathbf{C}}$
PQ	$(P)_{\mathbf{C}}(Q)_{\mathbf{C}}$

Table: translation: $()_{\mathbf{C}} : \Lambda \to \mathrm{CL}$

$$\vdash_\lambda M = N \iff \vdash_{\mathrm{CL}} (M)_{\mathbf{C}} = (N)_{\mathbf{C}}$$

Lambda Calculus subsumes Combinatory Logic

M	$(M)_\lambda$
$\mathbf{I}$	$\lambda x.x$
$\mathbf{K}$	$\lambda xy.x$
$\mathbf{S}$	$\lambda xyz.xz(yz)$
PQ	$(P)_\lambda(Q)_\lambda$

Table: translation: $()_\lambda : \mathrm{CL} \to \Lambda$

$$\vdash_{\mathrm{CL}} M = N \implies \vdash_\lambda (M)_\lambda = (N)_\lambda$$

But not the other way around: $\vdash_\lambda (\mathbf{SKI})_\lambda = (\mathbf{I})_\lambda$, but $\nvdash_{\mathrm{CL}} \mathbf{SKI} = \mathbf{I}$.

Numeralwise Representability

- The combinatory Church numerals are defined by

$$\underline{n} = (\mathbf{SB})^n(\mathbf{KI})$$

- A partial function $f : \mathbb{N} \to \mathbb{N}$ is *numeralwise represented* by a combinatory term M if for all $n \in \mathbb{N}$, if $f(n)$ is defined and equal to m, then

$$\vdash_{\mathrm{CL}} M\underline{n} = \underline{m}$$

and if $f(n)$ is undefined, then $M\underline{n}$ has no normal form.

Theorem

The partial functions numeralwise representable in CL *are exactly the partial recursive functions.*

Simply-Typed Lambda Calculus (STLC)

- Type

$$T := 1 \mid T \times T \mid T \to T$$

- Term

$$\Lambda := x \mid * \mid \Lambda\Lambda \mid \lambda x.\Lambda \mid \langle \Lambda, \Lambda \rangle \mid \pi_1 \Lambda \mid \pi_2 \Lambda$$

- Judgement

$$x_1 : T_1, \ldots, x_n : T_n \vdash t : T$$

1. t is a proof of T from assumptions $T_1, \ldots, T_n$.
2. t is a program of type T with free variables $x_1, \ldots, x_n$ of type $T_1, \ldots, T_n$.

The System of Simply-Typed Lambda Calculus

$$\frac{x : A \in \Gamma}{\Gamma \vdash x : A}\ \text{var} \qquad \frac{}{\Gamma \vdash * : 1}\ \text{unit}$$

$$\frac{\Gamma \vdash t : A \quad \Gamma \vdash u : B}{\Gamma \vdash \langle t, u \rangle : A \times B}\ \times^{+} \qquad \frac{\Gamma \vdash t : A \times B}{\Gamma \vdash \pi_1 t : A}\ \times^{-} \qquad \frac{\Gamma \vdash t : A \times B}{\Gamma \vdash \pi_2 t : B}\ \times^{-}$$

$$\frac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \to B}\ \text{abs} \qquad \frac{\Gamma \vdash t : A \to B \quad \Gamma \vdash u : A}{\Gamma \vdash tu : B}\ \text{app}$$

Reduction rules

$$(\lambda x.t)u \to t[u/x] \qquad (\beta_{\to})$$

$$\lambda x.tx \to t \quad \text{where } x \notin \mathrm{Var}(t) \qquad (\eta_{\to})$$

$$\pi_1\langle t, u \rangle \to t \qquad (\beta_{\times,1})$$

$$\pi_2\langle t, u \rangle \to u \qquad (\beta_{\times,2})$$

$$\langle \pi_1 t, \pi_2 t \rangle \to t \qquad (\eta_{\times})$$

What does β/η-reduction correspond to?

$$\dfrac{\dfrac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \to B} \qquad \Gamma \vdash u : A}{\Gamma \vdash (\lambda x.t)u : B} \quad \overset{\text{cut}}{\Longrightarrow} \quad \Gamma \vdash t[u/x] : B$$

$$\Gamma \vdash t : A \to B \quad \overset{\text{expansion}}{\Longrightarrow} \quad \dfrac{\dfrac{\Gamma, x : A \vdash t : A \to B \quad \Gamma, x : A \vdash x : A}{\Gamma, x : A \vdash tx : B}}{\Gamma \vdash \lambda x.tx : A \to B}$$

Example

$$\dfrac{\dfrac{\dfrac{\dfrac{[x : A \to B \to C]^3 \quad [z : A]^1}{xz : B \to C} \quad \dfrac{[y : A \to B]^2 \quad [z : A]^1}{yz : B}}{xz(yz) : C}}{\lambda z.xz(yz) : A \to C}\ {\to}^{+1}}{\lambda y.\lambda z.xz(yz) : (A \to B) \to A \to C}\ {\to}^{+2}}{\lambda x.\lambda y.\lambda z.xz(yz) : (A \to B \to C) \to (A \to B) \to A \to C}\ {\to}^{+3}$$

- $\mathbf{I} := \lambda x.x : A \to A$
- $\mathbf{K} := \lambda x.\lambda y.x : A \to B \to A$
- $\mathbf{S} := \lambda x.\lambda y.\lambda z.xz(yz) : (A \to B \to C) \to (A \to B) \to A \to C$
- $\mathbf{B} : (B \to C) \to (A \to B) \to A \to C$
- $\mathbf{C} : (A \to B \to C) \to B \to A \to C$
- $\mathbf{W} : (A \to A \to B) \to A \to B$

Lemma (Substitution lemma)

$$\frac{\Gamma, x : A \vdash t : B \quad \Gamma \vdash u : A}{\Gamma \vdash t[u/x]B}$$

Theorem (Subject Reduction Theorem)

$$\Gamma \vdash t : A \ \ \& \ \ t \twoheadrightarrow_\beta u \implies \Gamma \vdash u : A$$

Theorem (Church-Rosser property for typable terms)

Suppose that $\Gamma \vdash t : A$. *If* $t \twoheadrightarrow_\beta u$ *and* $t \twoheadrightarrow_\beta v$, *then there exists a term* w *s.t.* $u \twoheadrightarrow_\beta w, v \twoheadrightarrow_\beta w$ *and* $\Gamma \vdash w : A$.

Theorem (Strong Normalization Theorem)

If $\Gamma \vdash t : A$, *then there is no infinite* β*-reduction path starting from* t.

Subject Reduction Well-typed programs never go wrong: evaluating a program $t : A$ to a value indeed returns a value of type A.

Church-Rosser It doesn't make any difference for the final value how we reduce.

Strong Normalization No matter how one evaluates, one always obtains a value: there are no infinite computations possible.

Subformula Property

Theorem (Subformula Property)

If $\Gamma \vdash t : A$ is normal, i.e. there is no reduction step $t \longrightarrow t'$, then the derivation of $\Gamma \vdash t : A$ can only mention subformulas of A and subformulas of assumptions in Γ.

Remark: We can eliminate irrelevant detours from a proof in finite time.

Remark: Intensionality

Intensionality occurs when mathematical objects can be seen in two ways:

1. extensionally, i.e. abstractly, up to extensional equality. For example, logical formula, recursive function
2. intensionally, through their descriptions, or syntax. For example, Gödel number, index of recursive function

To be intensional is to be finer than extensional equality.
One cannot go from extension to intension within the system.

- we cannot get an index for a recursive function from the function itself within the computing system.

$$\neg\exists Q \in \Lambda \forall M \in \Lambda.\, QM =_\beta \ulcorner M \urcorner$$

- we cannot obtain the Gödel number of a logical formula within the logical system.

Remark: Modality-as-Intension

- For type A, let there be a type $\Box A$, whose elements can be understood as "programs that — when run — will yield objects of that type".
- From intension to extension: Interpreter, or evaluator.
$$\Box A \to A$$
- From code to code-for-code.
$$\Box A \to \Box\Box A$$
- From code for a function, to a map on codes: intensional substitution, a.k.a. the s-m-n theorem
$$\Box(A \to B) \to \Box A \to \Box B$$
- Assume $t = f\ulcorner t\urcorner$. If $t : A$, then $\ulcorner t\urcorner : \Box A$ and hence $f : \Box A \to A$.
- Kleene's fixpoint theorem then says
"for each $f : \Box A \to A$, we have $\ulcorner t\urcorner : \Box A$"
- It is Löb's rule
$$\frac{\Box A \to A}{\Box A}$$
- The type of Kleene's fixpoint theorem is Löb axiom $\Box(\Box A \to A) \to \Box A$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

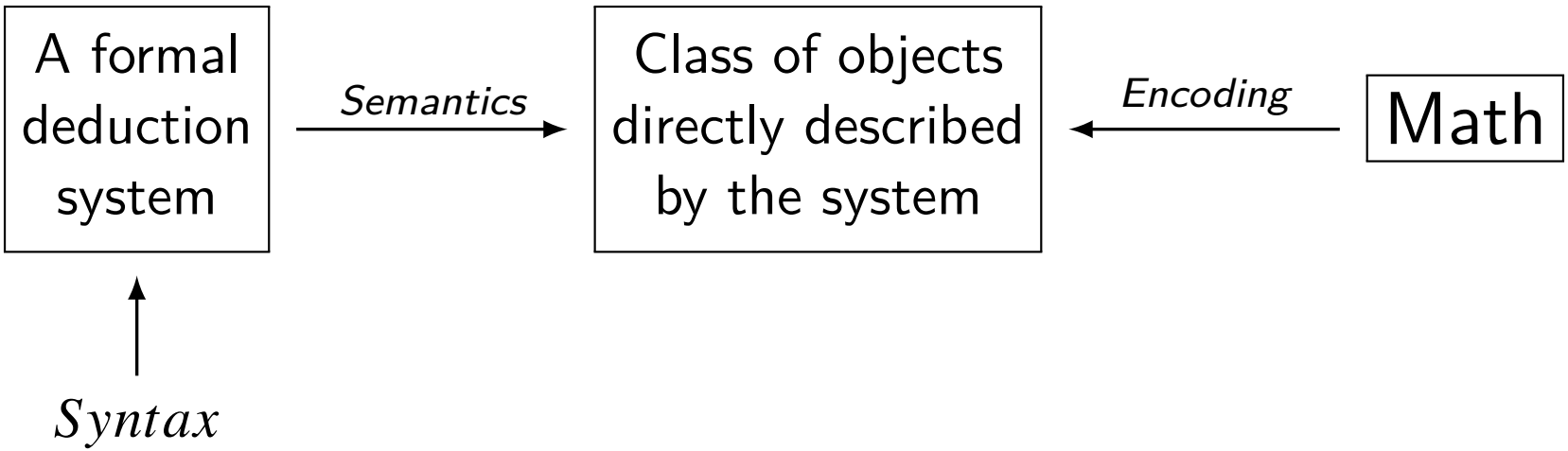
Category Theory

Quantum Computing

Answers to the Exercises

References2039

What is a Formalization of Math?

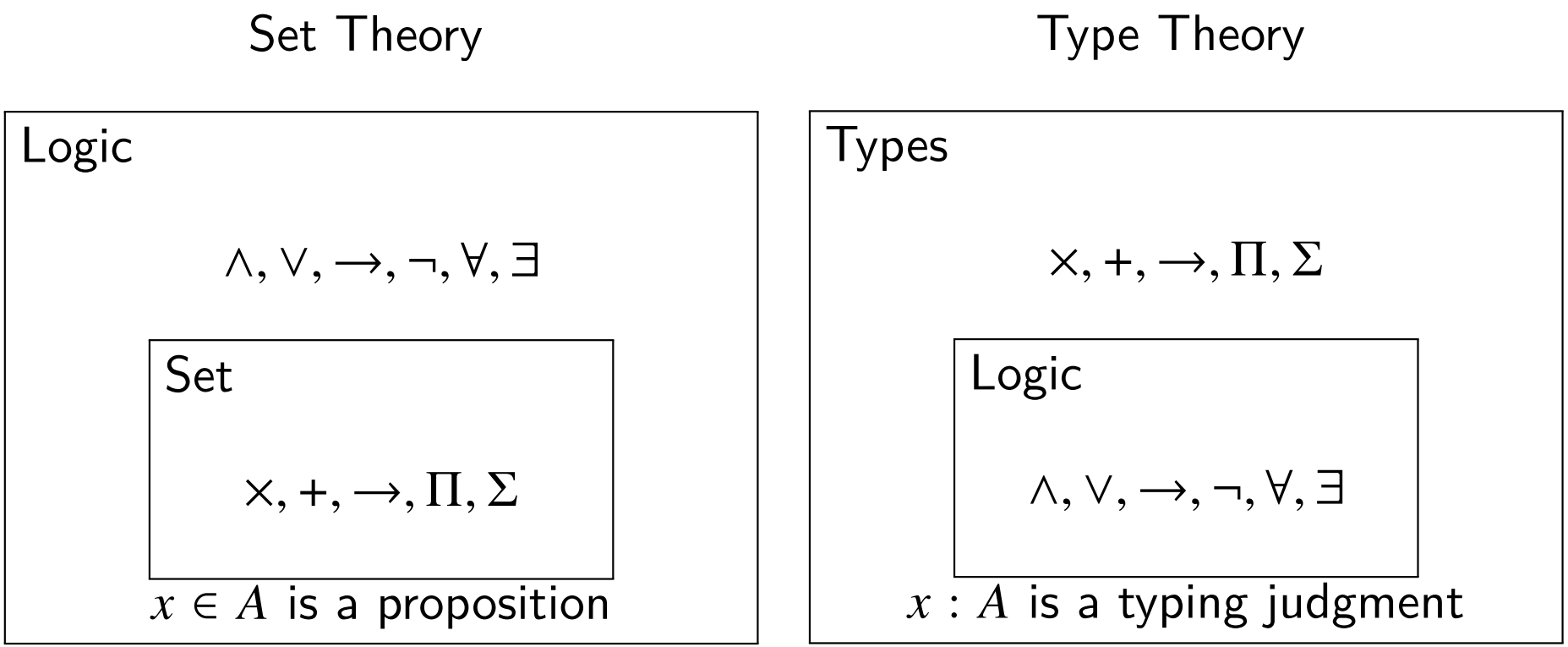


First Order Logic $\xrightarrow{express}$ ZFC $\xrightarrow{describe}$ V

Intensional $MLTT$ $\xrightarrow{express}$ $HoTT$ $\xrightarrow{describe}$ $\infty\text{–}\mathbf{Grpd}$

- set theory
- category theory
- homotopy type theory

Type Theory vs Set Theory



- We should treat "mathematical foundations" more like programming languages.
- It doesn't make sense to argue about which is "correct".
- What do we want a foundation to do?
- Risk Assessment, Metamathematical Corral, Generous Arena, Shared Standard, Proof Checking, Essential Guidance...

Why HoTT?[Uni13]

1. Homotopy can be used as a tool to construct models of systems of logic.
2. Constructive type theory can be used as a formal calculus to reason about homotopy.
3. The computational implementation of type theory allows computer verified proofs in homotopy theory.
4. The homotopy interpretation suggests new logical constructions and axioms as a new approach to foundations of math with intrinsic geometric content.



(e) Martin-Löf (f) Voevodsky

Context & Judgement

- context: sequence of variable declarations
  $x_1 : A_1,\ x_2 : A_2(x_1),\ \ldots,\ x_n : A_n(x_1, \ldots, x_{n-1})$
- judgement: context $\vdash$ conclusion

$\Gamma \vdash A : \text{type}$	A is a well-formed **type** in context Γ
$\Gamma \vdash a : A$	a is a well-formed **term** of type A
$\Gamma \vdash a = b : A$	a is convertible to b in type A
$\Gamma \vdash A = B : \text{type}$	types A and B are convertible

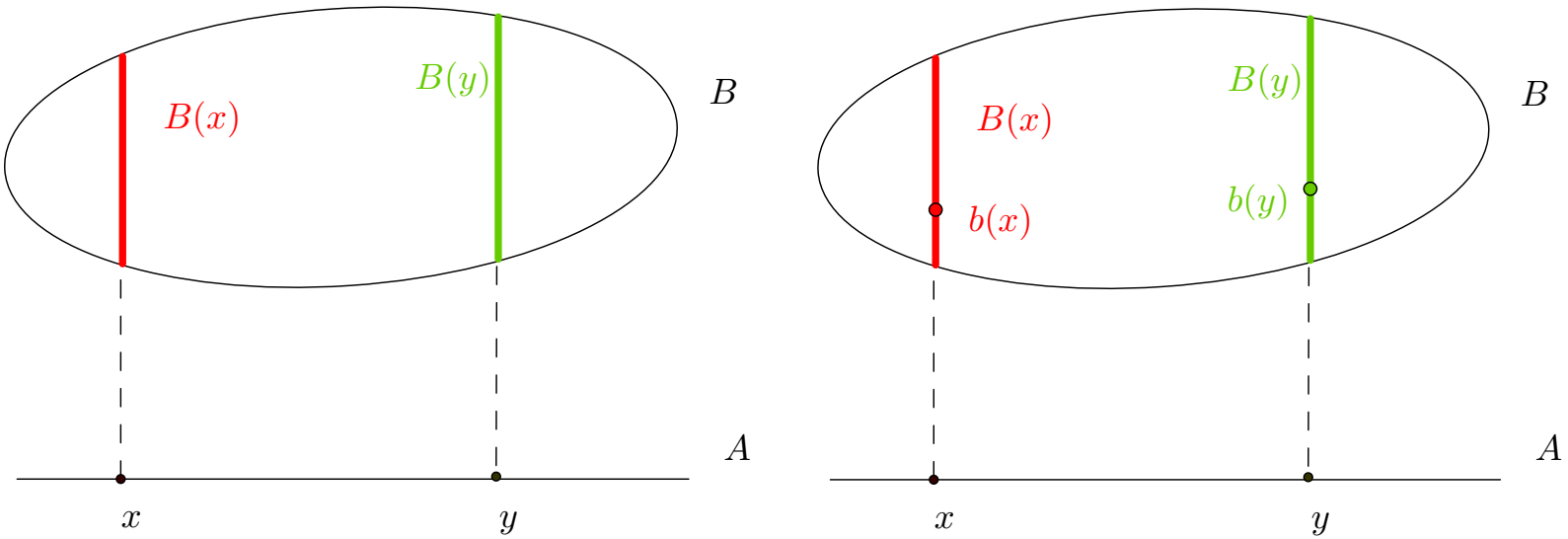
- dependent type

  $$x : A \vdash B(x) : \text{type}$$

- dependent term

  $$x : A \vdash b(x) : B(x)$$

Remark



- From any point x of A, we can observe a portion $B(x)$ of the shape B.

$$x : A \vdash B(x) : \text{Shape}$$

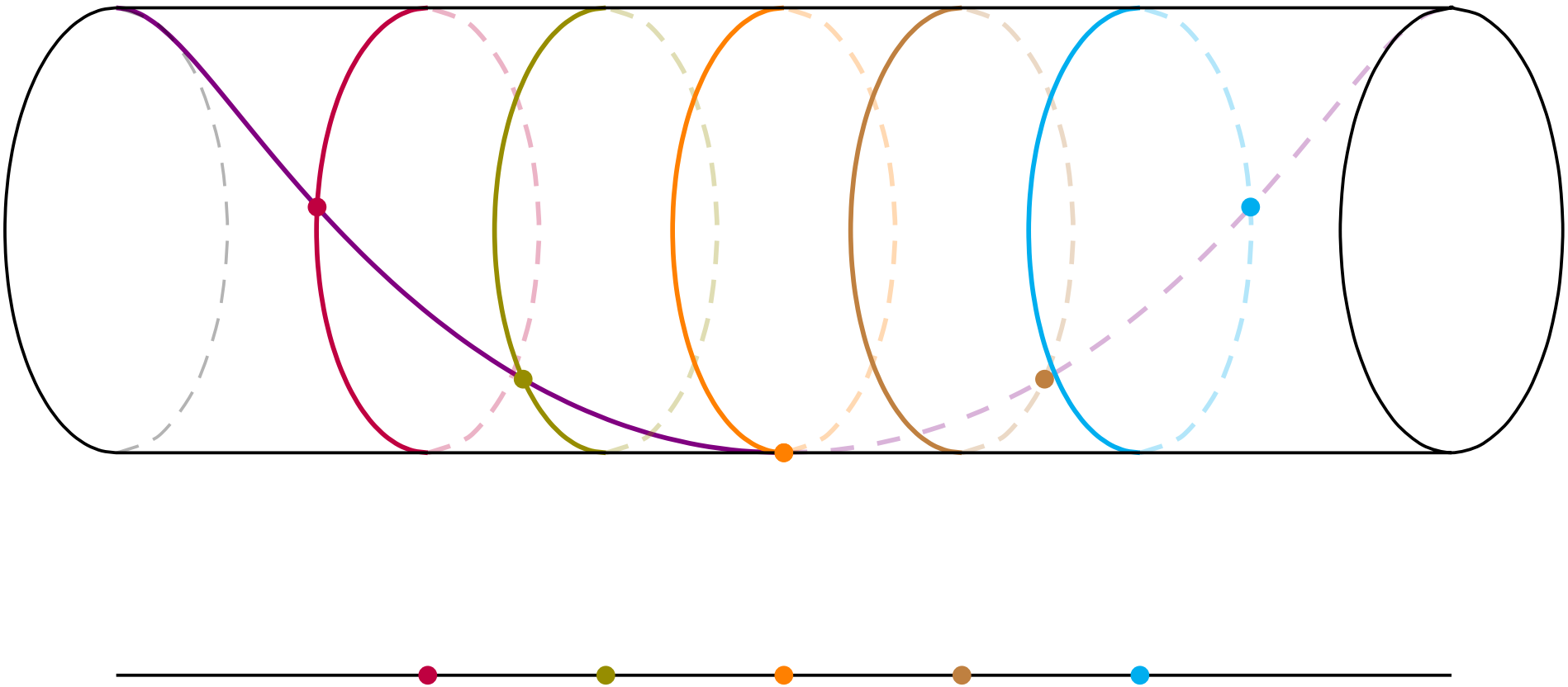
In topology, B is a space fibered over A (there is a fibration $\pi : B \to A$), whose fibers are given by $\pi^{-1}(x)$ for each $x \in A$.

- From any point x of A, we can observe a point $b(x)$ of $B(x)$.

$$x : A \vdash b(x) : B(x)$$

Here b is a section of the fibration $\pi : B \to A$, i.e. a map $s : A \to B$ such that $\pi \circ s = 1_A$.

Remark



$t : \mathsf{Time},\ x : \mathsf{Space}(t) \vdash \mathsf{Swimming}(t, x)$

$(t : \text{时间})(x : \text{在 } t \text{ 时的湖面}) \vdash \text{鸭鸭在 } t \text{ 时游在湖面的 } x \text{ 处}$

Logical Rules

Each type constructor comes with rules:

Formation	way to construct a type
Introduction	way to construct canonical terms of that type
Elimination	way to use a term of the introduced type to construct other terms
Conversion	what happens when one does Introduction followed by Elimination

Remark:

Formation	When can we observe A?
Introduction	When can we observe a point of A?
Elimination	When can we observe points on another shape from A?
Conversion	What are the symmetries of A?

Rules for unit type

$$\frac{}{\Gamma \vdash 1 : \mathsf{type}}\ 1F$$

$$\frac{}{\Gamma \vdash * : 1}\ 1I$$

$$\frac{\Gamma \vdash x : 1}{\Gamma \vdash x = * : 1}\ 1C$$

Rules for dependent product type

$$\frac{A : \text{type} \quad x : A \vdash B(x) : \text{type}}{\prod_{x:A} B(x) : \text{type}} \; \Pi F$$

$$\frac{x : A \vdash fx : B(x)}{\lambda x.fx : \prod_{x:A} B(x)} \; \Pi I$$

$$\frac{a : A \quad f : \prod_{x:A} B(x)}{fa : B(a)} \; \Pi E$$

$$\frac{a : A \quad x : A \vdash fx : B(x)}{(\lambda x.fx)\, a = fa : B(a)} \; \Pi C$$

Remark: We write $A \to B$ instead of $\prod_{x:A} B$ if x is not free in B.

Rules for dependent sum type

$$\frac{A : \text{type} \quad x : A \vdash B(x) : \text{type}}{\sum_{x:A} B(x) : \text{type}}\ \Sigma F$$

$$\frac{a : A \quad b : B(a)}{(a, b) : \sum_{x:A} B(x)}\ \Sigma I$$

$$\frac{p : \sum_{x:A} B(x) \quad x : A, y : B(x) \vdash c(x, y) : C((x, y))}{E(c, p) : C(p)}\ \Sigma E$$

Remark: We execute $E(c, p)$ as folows. First execute p, which yields a canonical term of the form (a, b) with $a : A$ and $b : B(a)$. Then we have $c(a, b) : C((a, b))$. Executing $c(a, b)$ we obtain a canonical term e of $C((a, b))$. It is also a canonical term of $C(p)$.

$$\frac{a : A \quad b : B(a) \quad x : A, y : B(x) \vdash c(x, y) : C((x, y))}{E(c, (a, b)) = c(a, b) : C((a, b))}\ \Sigma C$$

Derived Rules

$$\pi_1(p) :\equiv E(\lambda xy.x, p)$$
$$\pi_2(p) :\equiv E(\lambda xy.y, p)$$

$$\frac{A : \mathsf{type} \quad x : A \vdash B(x) : \mathsf{type}}{\sum_{x:A} B(x) : \mathsf{type}}\ \Sigma F$$

$$\frac{a : A \quad b : B(a)}{(a,b) : \sum_{x:A} B(x)}\ \Sigma I$$

$$\frac{p : \sum_{x:A} B(x)}{\pi_1(p) : A}\ \Sigma E_1 \qquad \frac{p : \sum_{x:A} B(x)}{\pi_2(p) : B(\pi_1(p))}\ \Sigma E_2$$

$$\frac{a : A \quad b : B(a)}{\pi_1(a,b) = a : A}\ \Sigma C_1 \qquad \frac{a : A \quad b : B(a)}{\pi_2(a,b) = b : B(a)}\ \Sigma C_2$$

Remark: We write $A \times B$ instead of $\sum_{x:A} B$ if x is not free in B.

Rules for coproduct type

$$\frac{A : \mathsf{type} \quad B : \mathsf{type}}{A + B : \mathsf{type}} \; +F$$

$$\frac{a : A}{\iota_1(a) : A + B} \; +I_1 \qquad \frac{b : B}{\iota_2(b) : A + B} \; +I_2$$

$$\frac{c : A + B \quad x : A \vdash f(x) : C(\iota_1(x)) \quad y : B \vdash g(y) : C(\iota_2(y))}{D(f, g, c) : C(c)} \; +E$$

$$\frac{a : A \quad x : A \vdash f(x) : C(\iota_1(x)) \quad y : B \vdash g(y) : C(\iota_2(y))}{D(f, g, \iota_1(a)) = f(a) : C(\iota_1(a))} \; +C_1$$

$$\frac{b : B \quad x : A \vdash f(x) : C(\iota_1(x)) \quad y : B \vdash g(y) : C(\iota_2(y))}{D(f, g, \iota_2(b)) = g(b) : C(\iota_2(b))} \; +C_2$$

Rules for identity type

Notation: Sometimes we write $x =_A y$ for $\mathsf{Id}_A(x, y)$.

$$\frac{A : \mathsf{type} \quad a, b : A}{a =_A b : \mathsf{type}}\ \mathsf{Id}F$$

$$\frac{a : A}{\mathsf{refl}_a : a =_A a}\ \mathsf{Id}I$$

$$\frac{p : a =_A b \qquad \begin{array}{c} x, y : A, z : x =_A y \vdash B(x, y, z) : \mathsf{type} \\ x : A \vdash d(x) : B(x, x, \mathsf{refl}_x) \end{array}}{J_d(a, b, p) : B(a, b, p)}\ \mathsf{Id}E$$

$$\frac{a : A \qquad \begin{array}{c} x, y : A, z : x =_A y \vdash B(x, y, z) : \mathsf{type} \\ x : A \vdash d(x) : B(x, x, \mathsf{refl}_x) \end{array}}{J_d(a, a, \mathsf{refl}_a) = d(a) : B(a, a, \mathsf{refl}_a)}\ \mathsf{Id}C$$

Remark: IdE Rule as Path Induction

Path Induction

If $x, y : A,\ p : x =_A y \vdash B(x, y, p)$ is a type family then to prove $B(x, y, p)$ it suffices to assume y is x and p is refl_x. i.e.:

$$\mathsf{ind}_{=_A} : \prod_{x:A} B(x, x, \mathsf{refl}_x) \to \prod_{x,y:A} \prod_{p:x=_A y} B(x, y, p)$$

By path induction, paths can be reversed and concatenated:

$$(-)^{-1} : x =_A y \to y =_A x \qquad * : x =_A y \to y =_A z \to x =_A z$$

To define both terms, we may assume $p : x =_A y$ and then define terms in the types $P(x, y, p) :\equiv y =_A x$ and $Q(x, y, p) :\equiv y =_A z \to x =_A z$. By path induction, we may reduce to the cases $P(x, x, \mathsf{refl}_x) :\equiv x =_A x$ and $Q(x, x, \mathsf{refl}_x) :\equiv x =_A z \to x =_A z$.

Logic in HoTT

Logical Connectives	Interpretation in HoTT
$\bot$	0
$\top$	1
$A \wedge B$	$A \times B$
$A \vee B$	$\|A + B\|$
$A \to B$	$A \to B$
$A \leftrightarrow B$	$A \simeq B$
$\neg A$	$A \to 0$
$\forall_{x:A} B(x)$	$\prod_{x:A} B(x)$
$\exists_{x:A} B(x)$	$\|\sum_{x:A} B(x)\|$
$\exists!_{x:A} B(x)$	$\mathrm{isContr}\left(\sum_{x:A} B(x)\right)$

Mathematics in Type Theory

- To state a conjecture, one forms a type that encodes its statement.
- To prove the theorem, one constructs a term in that type.

Theorem

For any types A, B, C*:* $(A + B \to C) \to (A \to C) \times (B \to C)$.

Proof.

By ΠI, suppose given $h : A + B \to C$, our goal is a term of type $(A \to C) \times (B \to C)$. By ΣI, it suffices to define terms of type $A \to C$ and type $B \to C$. By ΠI, to define a term of type $A \to C$ it suffices to assume a term $a : A$ and define a term of type C. By $+I_1$, we then have a term $\iota_1 a : A + B$. Then by ΠE, we obtain a term $h(\iota_1 a) : C$. Similarly, we have $\lambda b.h(\iota_2 b) : B \to C$. This constructs

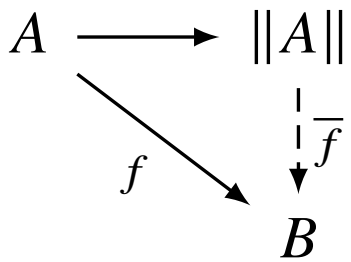
$$\lambda h.(\lambda a.h(\iota_1 a), \lambda b.h(\iota_2 b)) : (A + B \to C) \to (A \to C) \times (B \to C)$$

□

Rules for Propositional Truncation

The propositional truncation $\|A\|$ of A is defined by

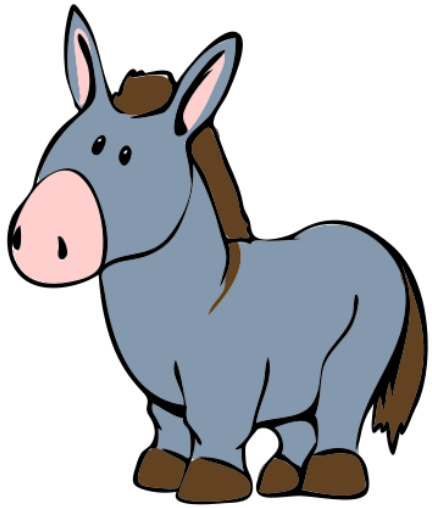
- $a : A \vdash \overline{a} : \|A\|$
- $x, y : \|A\| \vdash x =_A y$
- If B is a mere proposition and $f : A \to B$, then there is an induced $\overline{f} : \|A\| \to B$ such that $\overline{f}(\overline{a}) = f(a)$ for all $a : A$.



Remarks: The first rule means that if A is inhabited, so is $\|A\|$.
The second ensures that $\|A\|$ is a mere proposition.
Remark: $\|A\| :\equiv \prod_{X:\mathrm{Prop}} (A \to X) \to X$ satisfies the rules.

Translation

Every farmer who owns a donkey beats it.



$$\forall x \forall y (Fx \wedge Dy \wedge Oxy \rightarrow Bxy)$$

$$b : \prod_{z : \sum_{x:F} \sum_{y:D} Oxy} B(\pi_1(z), \pi_1(\pi_2(z)))$$

- Anyone who owns a gun should register it.
- Every point that lies outside a line determines a parallel to it.
- Any number which has a proper divisor is greater than it.

Homotopy Levels

$$\mathrm{isContr}(A) :\equiv \sum_{x:A} \prod_{y:A} x =_A y$$

$$\mathrm{isProp}(A) :\equiv \prod_{x,y:A} \mathrm{isContr}\left(x =_A y\right) \qquad \left(\text{equivalently, } \prod_{x,y:A} x =_A y\right)$$

$$\mathrm{isSet}(A) :\equiv \prod_{x,y:A} \mathrm{isProp}\left(x =_A y\right)$$

$$\mathrm{isGroupoid}_1(A) :\equiv \prod_{x,y:A} \mathrm{isSet}\left(x =_A y\right)$$

$$\mathrm{isGroupoid}_{n+1}(A) :\equiv \prod_{x,y:A} \mathrm{isGroupoid}_n\left(x =_A y\right)$$

$$\mathrm{Prop} :\equiv \sum_{A:U} \mathrm{isProp}(A) \qquad \mathrm{Set} :\equiv \sum_{A:U} \mathrm{isSet}(A)$$

The Hierarchy of Homotopy Levels

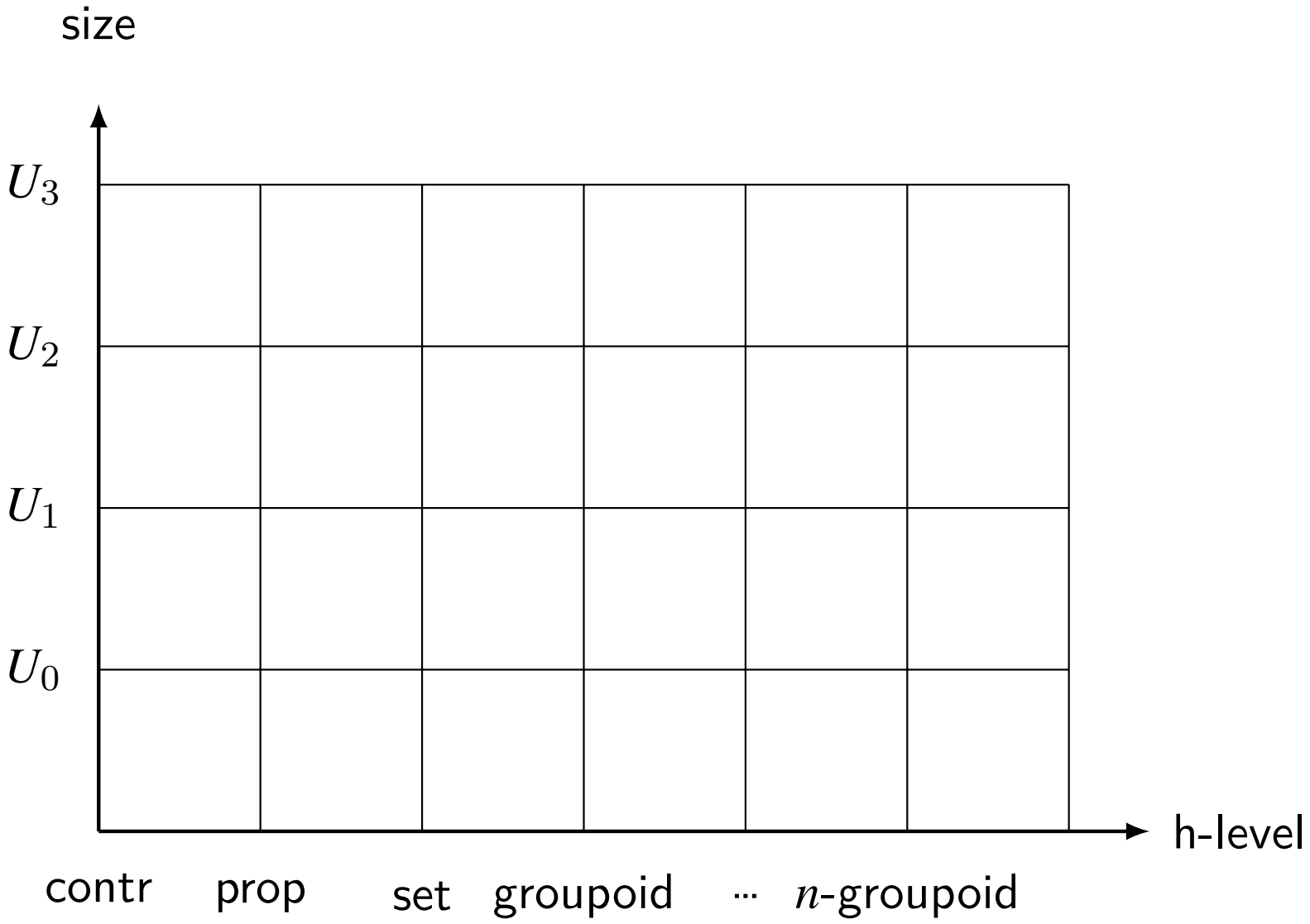


Figure: The 2D hierarchy of types

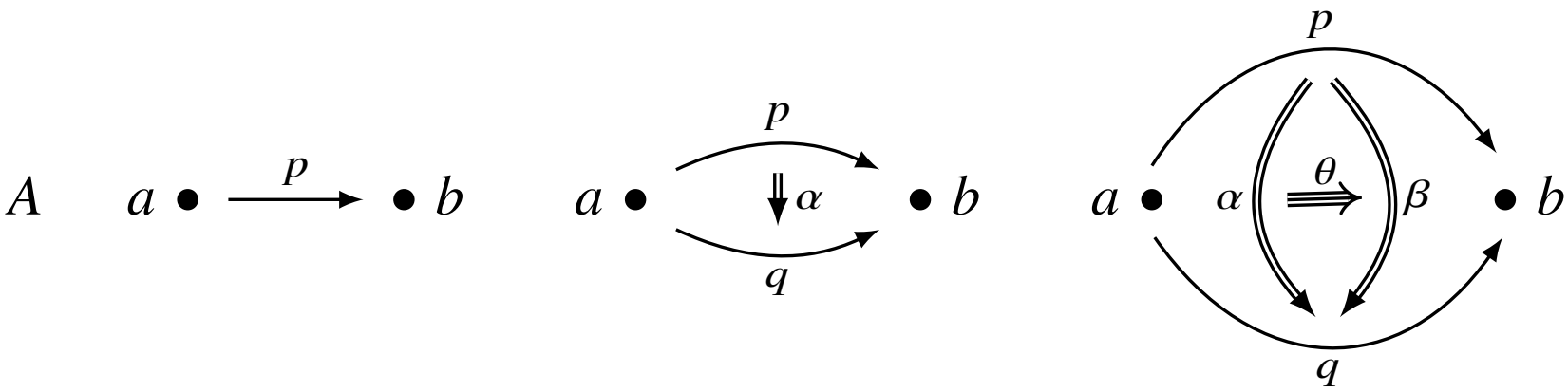
The Hierarchy of Homotopy Levels

- There is only one, up to the univalent equality, type of h-level 0. It is the one point type.
- Types of h-level 1 are propositions.
- Types of h-level 2 are sets.
- Types of h-level 3 are the kind of types that are formed by objects in a category. The study of structures on types of h-level 3 is closely related to category-level mathematics.

Instead of sets, clouds of discrete elements, we envisage some sorts of vague spaces, which can be very severely deformed, mapped one to another, and all the while the specific space is not important, but only the space up to deformation.
If you want to get a discrete set, then you pass to the set of connected components of a space defined only up to homotopy.

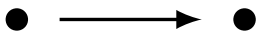
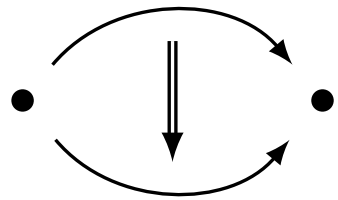
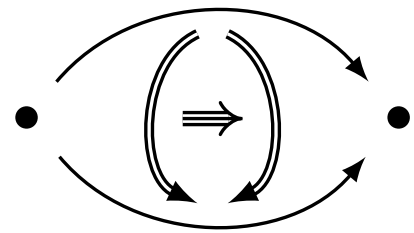
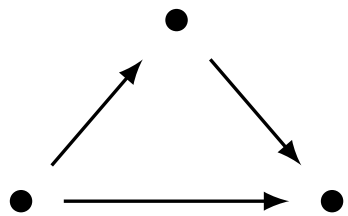
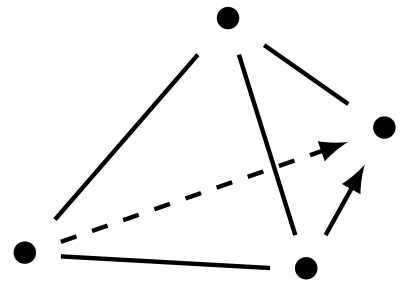
— *Yuri Manin*

Morning Star $\stackrel{?}{=}$ Evening Star

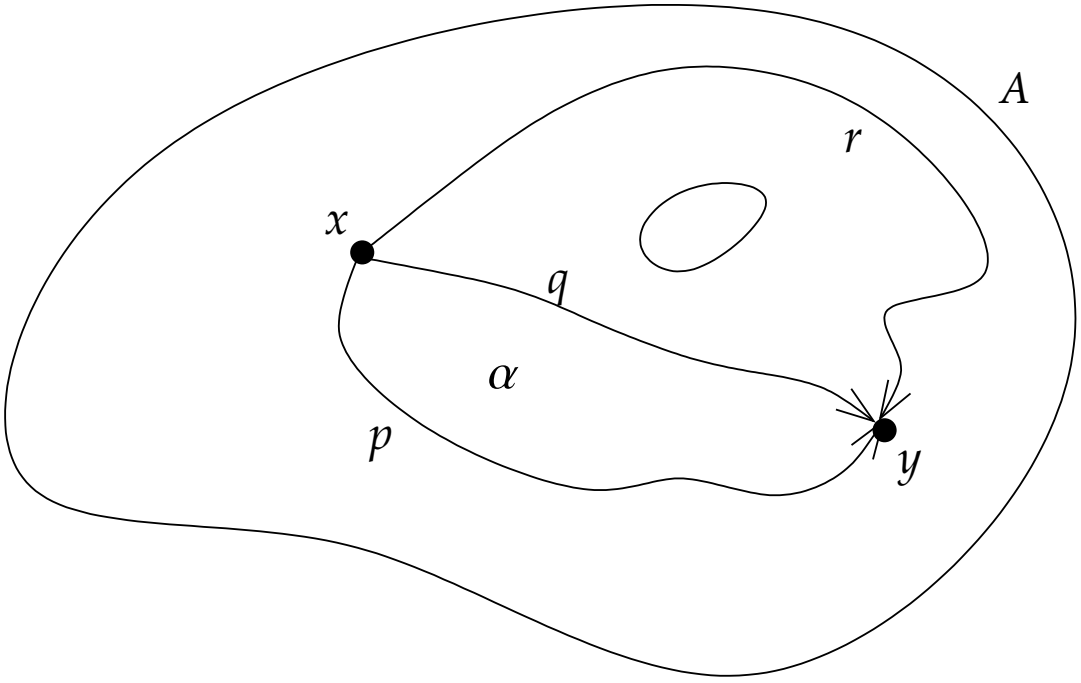


Type	Topological space
Term	Continuous map
$a : A$	point $a \in A$
$p : a =_A b$	path p from a to b
$\alpha : p =_{a=_A b} q$	homotopy α from p to q
$\theta : \alpha =_{p=_{a=_A b} q} \beta$	$\ldots$

- There is no ultimate, once-and-for-all identity between things — instead, there are structures and isomorphisms between them, and insofar as we are given an isomorphism between structures, we can transfer properties of one to the other, making them indistinguishable.
- But isomorphisms are no longer facts, they are themselves structures.

object	morphism	2-morphism	3-morphism	⋯
				Globes
				Simplices

Example



$$x, y \in A$$

$$p, q, r : x = y$$

$$\alpha : p = q$$

$$q \neq r$$

$$p \neq r$$

Equivalence

$$\mathrm{isContr}(A) :\equiv \sum_{x:A} \prod_{y:A} x =_A y$$

$$f^{-1}(y) :\equiv \sum_{x:A} f(x) =_A y$$

$$\mathrm{isEquiv}(f) :\equiv \prod_{y:B} \mathrm{isContr}\left(f^{-1}(y)\right)$$

$$A \simeq B :\equiv \sum_{f:A\to B} \mathrm{isEquiv}(f)$$

- function extensionality

  Let $f, g : \prod_{x:A} B(x)$. A homotopy from f to g is a dependent function of the type

$$f \sim g \coloneqq \prod_{x:A} (fx =_B gx)$$

  For saying $\prod_{x:A}(fx =_B gx)$ is inhabited tells us there is a continuous map from $x : A$ to paths between $f(x)$ and $g(x)$, which is the same as giving us a continuous deformation of f into g.

- bi-inverse

$$\mathrm{isBiInv}(f) \coloneqq \left(\sum_{g:B\to A} g \circ f \sim 1_A\right) \times \left(\sum_{h:B\to A} f \circ h \sim 1_B\right)$$

- isomorphism

$$\mathrm{isIso}(f) \coloneqq \sum_{g:B\to A} \left[\left(\prod_{x:A} gfx =_A x\right) \times \left(\prod_{y:B} fgy =_B y\right)\right]$$

$$A \cong B \coloneqq \sum_{f:A\to B} \mathrm{isIso}(f)$$

- $\mathrm{isEquiv}(f) \simeq \mathrm{isBiInv}(f) \simeq \mathrm{isIso}(f)$

Univalence Foundation

If a statement, concept, or construction is purely logical, then it should be invariant under all equivalences of the structures involved.

—*Steve Awodey*

Remark

- *All concepts in Russell's theory of types are invariant under isomorphism.*
- *All concepts in Martin-Löf type theory are invariant under homotopy equivalence.*

Voevodsky's Univalence Axiom

$$A =_U B \simeq (A \simeq B)$$

"Identity is equivalent to equivalence."

Univalence vs Invariance

- Frege's observes that the truth-value of $P(a)$ does not change when one substitutes a by b with the same meaning $a = b$.
- Under propositions as types, the meaning of a proposition is not just its truth-value, but the collection of its proofs, i.e., its homotopy type.
- For any family of types $x : A \vdash P(x)$, given a term $p : a =_A b$, and a term $t : P(a)$, there is an associated term $p_* t : P(b)$.
- In fact, the map $p_* : P(a) \to P(b)$ is always an equivalence of types $P(a) \simeq P(b)$.
- Thus type equivalence $A \simeq B$ is a finer notion of meaning than the truth-values derived from logical equivalence of propositions $A \leftrightarrow B$.
- We can state the *invariance principle* by adding a universe of types U. For any family of types $X : U \vdash P(X)$,

$$A \simeq B \to P(A) \simeq P(B)$$

- Take $P(X)$ to be $A =_U X$,

$$A \simeq B \to A =_U A \simeq A =_U B$$

From this, we get $A \simeq B \to A =_U B$.

Univalence $\Longrightarrow$ Invariance

- Voevodsky's Univalence Axiom is even stronger.

$$(A \simeq B) \simeq (A =_U B)$$

- Given an equivalence $e : A \simeq B$, by univalence we get an equality $\overline{e} : A =_U B$.
- But then the equality $\overline{e}$ acts on any family of types $X : U \vdash P(X)$ to give an equivalence $\overline{e}_* : P(A) \simeq P(B)$.
- So univalence implies the invariance principle

$$A \simeq B \to P(A) \simeq P(B)$$

- Univalence is an internalization of the invariance principle: it asserts that all concepts in the system are invariant under equivalence.

Univalence and Intensionality

- Apply the Univalence Axiom to itself in a higher universe U':

$$\Big[(A \simeq B) \simeq (A =_U B)\Big] \simeq \Big[(A \simeq B) =_{U'} (A =_U B)\Big]$$

- Under the interpretation of equality as sameness of meaning, type expressions A, B are regarded as presentations of mathematical propositions and structures, and they present the same mathematical object if $A = B$.
- Type expressions A and B have the same homotopy type means the same thing as to say that they have the same meaning.
- The meaning of a mathematical statement is its homotopy type.

Consequences of the Univalence Axiom

- Structure invariance principle

$$A \cong B \to A =_U B$$

  Isomorphic structures are identical.
- Function extensionality

$$\prod_{f,g:A\to B} f = g \simeq \left(\prod_{x:A} fx =_B gx\right)$$

- Propositional extensionality

$$\prod_{A,B:\mathrm{Prop}} A = B \simeq (A \leftrightarrow B)$$

  where $A \leftrightarrow B :\equiv (A \to B) \times (B \to A)$
- Paths are isomorphisms for sets

$$\prod_{A,B:\mathrm{Set}} A = B \simeq (A \cong B)$$

- The univalence axiom implies the functional extensionality both for "straight" functions and for dependent functions. It also implies that two logically equivalent "propositions" (types of h-level 1) are equal.
- The univalence axiom implies that the universe of types of h-level n has h-level $n+1$. In particular, the type of "propositions" is a "set" and the type of "sets" is a "groupoid".
- The univalence axiom implies similar statements for types with structures e.g. one can prove using the univalence axiom that the identity type between two groups is equivalent to the type of isomorphisms between these groups.

Type-Theoretic Axiom of Choice

Theorem (Type-Theoretic Axiom of Choice)

$$\left(\prod_{x:A}\sum_{b:B(x)} C(x,b)\right)\to\left(\sum_{g:\prod_{x:A}B(x)}\prod_{x:A} C\big(x,g(x)\big)\right)$$

is inhabited.

Remark: The stronger version of axiom of choice

$$\left(\prod_{x:A}\left\|\sum_{b:B(x)} C(x,b)\right\|\right)\to\left\|\sum_{g:\prod_{x:A}B(x)}\prod_{x:A} C\big(x,g(x)\big)\right\|$$

is not a consequence of our basic type theory, but it may consistently be assumed as axioms.

Proof.

Take $f : \prod_{x:A} \sum_{b:B(x)} C(x,b)$ and $x : A$.

$$fx : \sum_{b:B(x)} C(x,b) \qquad (\Pi E)$$

$$\pi_1(fx) : B(x) \qquad (\Sigma E_1)$$

$$\pi_2(fx) : C(x, \pi_1(fx)) \qquad (\Sigma E_2)$$

$$\lambda x.\pi_1(fx) : \prod_{x:A} B(x) \qquad (\Pi I)$$

$$(\lambda x.\pi_1(fx))x = \pi_1(fx) : B(x) \qquad (\Pi C)$$

$$\pi_2(fx) : C(x, (\lambda x.\pi_1(fx))x) \qquad (\text{substitution})$$

$$\lambda x.\pi_2(fx) : \prod_{x:A} C(x, (\lambda x.\pi_1(fx))x) \qquad (\Pi I)$$

$$\big(\lambda x.\pi_1(fx), \lambda x.\pi_2(fx)\big) : \sum_{g:\prod_{x:A} B(x)} \prod_{x:A} C(x, g(x)) \qquad (\Sigma I)$$

where $g :\equiv \lambda x.\pi_1(fx)$. By ΠI, the type-theoretic axiom of choice is inhabited by $\lambda f.\big(\lambda x.\pi_1(fx), \lambda x.\pi_2(fx)\big)$. □

Curry-Howard-Voevodsky Correspondence

Type Theory	**Logic**	**Set Theory**	**Homotopy Theory**
A : type	proposition	set	space
$a : A$	proof	element	point
$x : A \vdash B(x)$	predicate of sets	$\{B_x\}_{x \in A}$	fibration $B \twoheadrightarrow A$ with fibers $B(x)$
$x : A \vdash b(x) : B(x)$	conditional proof	family of elements	section
$0, 1$	$\bot, \top$	$\emptyset, \{\emptyset\}$	$\emptyset, \{\bullet\}$
$A + B$	$A \vee B$	disjoint union	coproduct
$A \times B$	$A \wedge B$	set of pairs	product space
$A \to B$	$A \to B$	set of functions	function space
$\sum_{x:A} B(x)$	$\exists_{x:A} B(x)$	disjoint sum	total space of fibration $B \twoheadrightarrow A$
$\prod_{x:A} B(x)$	$\forall_{x:A} B(x)$	product	space of sections of fibration $B \twoheadrightarrow A$
$p : x =_A y$	proof of equality	$x = y$	path from x to y in A
$\sum_{x,y:A} x =_A y$	equality relation	$\{(x, x) : x \in A\}$	path space A^I

Type Theory vs Category Theory

Type Theory	**Category Theory**
empty type 0	initial object
unit type 1	terminal object
product type $A \times B$	product
coproduct type $A + B$	coproduct
function type $A \to B$	exponential object (cartesian closure)
dependent product $\prod_{x:A} B$	right adjoint to pullback
dependent sum $\sum_{x:A} B$	left adjoint to pullback
identity type	diagonal/equalizer
proposition type Ω	subobject classifier (elementary topos)
universe type U	object classifier (∞-topos)
natural numbers $\mathbb{N}$	natural numbers object
coequalizer type $\mathrm{coeq}(f, g)$	coequalizer

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

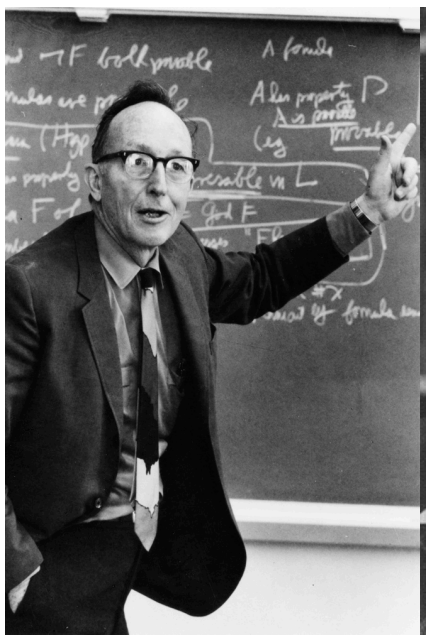
Birds vs Frogs

Category theory takes a bird's eye view of mathematics. From high in the sky, details become invisible, but we can spot patterns that were impossible to detect from ground level.

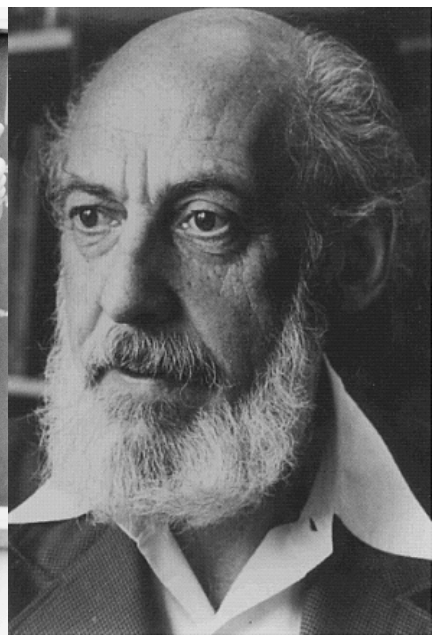
— *Tom Leinster*

Good general theory does not search for the maximum generality, but for the right generality.

— *Saunders Mac Lane*



(a) Mac Lane



(b) Eilenberg

Human Activity	Mathematical Idea	Mathematical Technique
Collecting	Object Collection	Set; class; multiset; list; family
Connecting	Cause and effect	ordered pair; relation; function; operation
”	Proximity; connection	Topological space; mereotopology
Following	Successive actions	Function composition; transformation group
Comparing	Enumeration	Bijection; cardinal number; order
Timing	Before & After	Linear order
Counting	Successor	Successor function; ordinal number
Computing	Operations on numbers	Addition, multiplication recursively defined; abelian group; rings
Looking at objects	Symmetry	Symmetry group; invariance; isometries
Building; shaping	Shape; point	Sets of points; geometry; pi
Rearranging	Permutation	Bijection; permutation group
Selecting; distinguishing	Parthood	Subset; order; lattice theory; mereology
Arguing	Proof	First-order logic
Measuring	Distance; extent	Rational number; metric space
Endless repetition	Infinity; Recursion	Recursive set; Infinite set
Estimating	Approximation	Real number; real field
Moving through space & time:	curvature	calculus; differential geometry
–without cycling	Change	Real analysis; transformation group
–with cycling	Repetition	pi; trigonometry; complex number; complex analysis
–both		Differential equations; mathematical physics
Motion through time alone	Growth & decay	e; exponential function; natural logarithms;
Altering shapes	Deformation	Differential geometry; topology
Observing patterns	Abstraction	Axiomatic set theory; universal algebra; category theory; morphism
Seeking to do better	Optimization	Operations research; optimal control theory; dynamic programming
Choosing; gambling	Chance	Probability theory; mathematical statistics; measure

推广与类比

$\mathbb{Z}$	$F[X]$
n 的十进制展开	f 的标准形式
$\lvert n\rvert$	$\deg f$
$a = qb + r$	$f = qg + r$
$b \mid a \iff r = 0$	$g \mid f \iff r = 0$
$\gcd(a, b) = ua + vb$	$\gcd(f, g) = uf + vg$
素数 p	不可约多项式 $p(X)$
$p \mid ab \to p \mid a \vee p \mid b$	$p \mid fg \to p \mid f \vee p \mid g$
整数分解为素因数的乘积	多项式分解为不可约因式的乘积
分数 $\frac{a}{b}$	分式 $\frac{f}{g}$
有理数域 $\mathbb{Q}$	有理函数域 $F(X)$
分数化成十进制小数	分式展开无穷级数
$\vdots$	$\vdots$

推广与类比

离散	连续
$\mathbb{N}$	$\mathbb{R}$
数列	函数
高阶等差数列 (如 $\binom{n}{p}$)	多项式函数 (如 $\frac{x^p}{p!}$)
差分 Δ	微分 D
差分方程 (如 $\Delta^p f(n) = 0$)	微分方程 (如 $D^p f(x) = 0$)
求和 $\sum$	积分 $\int$
线性方程组	微分方程
裂项求和 $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$	牛顿-莱布尼茨公式
离散概率	连续概率
$\vdots$	$\vdots$

Readings

1. S. Awodey: Category Theory.
2. S. Mac Lane: Categories for the Working Mathematician.
3. T. Leinster: Basic Category Theory.
4. P. Smith: Category Theory.
5. E. Riehl: Category theory in Context.
6. H. Simmons: An Introduction to Category Theory.
7. M. Barr, C. Wells: Category Theory for Computing Science.
8. B. Fong, D. I. Spivak: An Invitation to Applied Category Theory.
9. D. I. Spivak: Category Theory for the Sciences.
10. F. W. Lawvere, S. H. Schanuel: Conceptual Mathematics.
11. F. W. Lawvere, R. Rosebrugh: Sets for Mathematics.
12. T. Streicher: Introduction to Category Theory and Categorical Loigc.
13. B. Jacobs: Categorical Logic and Type Theory.
14. R. Goldblatt: Topoi — The Categorical Analysis of Logic.
15. P. Johnstone: Sketches of an Elephant.
16. S. Mac Lane, I. Moerdijk: Sheaves in Geometry and Logic.
17. J. Adamek, H. Herrlich, G. E. Strecher: Abstract and Concrete Categories — The Joy of Cats.
18. nLab

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

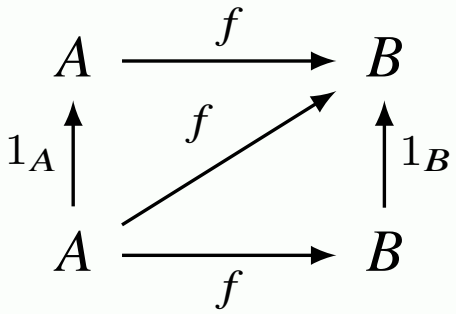
References2039

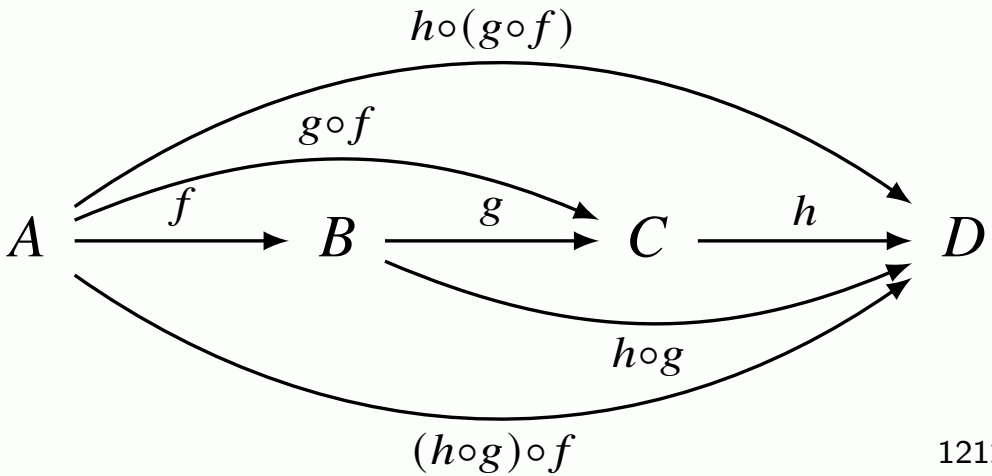
What is a Category?

Definition (Category)

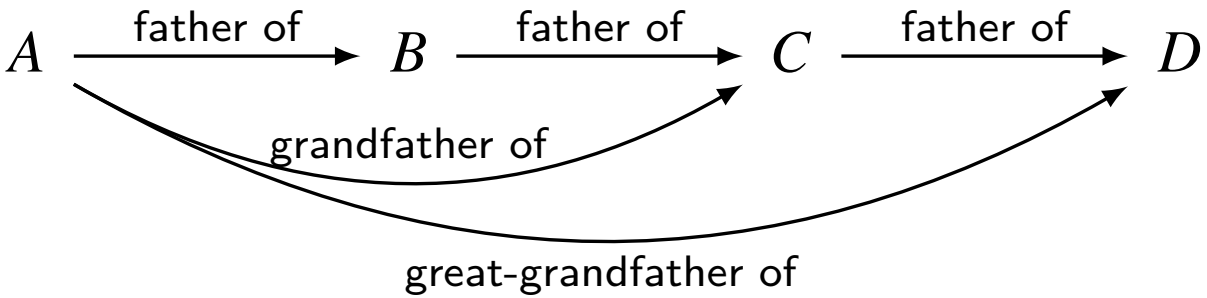
A category $\mathbf{C} = (\mathrm{ob}(\mathbf{C}), \mathrm{Hom}, \circ, 1)$ consists of

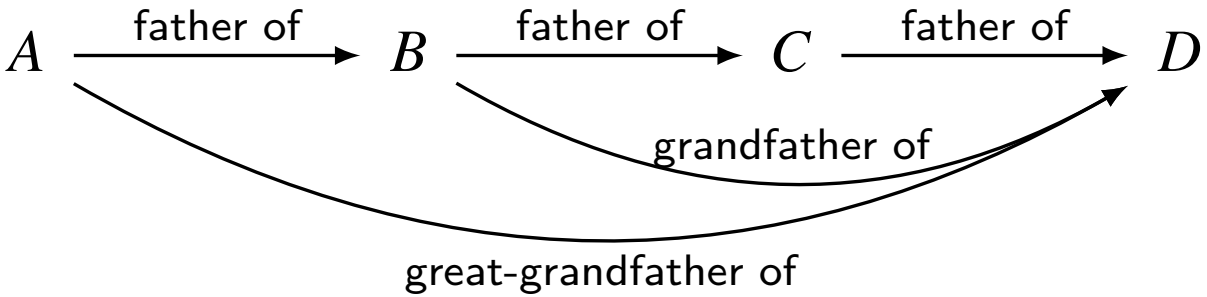
- a class of objects $\mathrm{ob}(\mathbf{C})$.
- a set of morphisms $\mathrm{Hom}(A, B) := \{f : A \to B\}$ (or $\mathbf{C}(A, B)$) with domain $A = \mathrm{dom}(f)$ and codomain $B = \mathrm{cod}(f)$ for $A, B \in \mathrm{ob}(\mathbf{C})$.
- the identity $1_A : A \to A$ for $A \in \mathrm{ob}(\mathbf{C})$.
- the composition $\circ : \mathrm{Hom}(A, B) \times \mathrm{Hom}(B, C) \to \mathrm{Hom}(A, C)$ for $A, B, C \in \mathrm{ob}(\mathbf{C})$ such that the following properties are satisfied:
  - $\forall AB \in \mathrm{ob}(\mathbf{C}) \forall f : A \to B \left[f \circ 1_A = f = 1_B \circ f \right]$
  - $\forall ABCD \in \mathrm{ob}(\mathbf{C}) \forall fgh : A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D \left[h \circ (g \circ f) = (h \circ g) \circ f \right]$



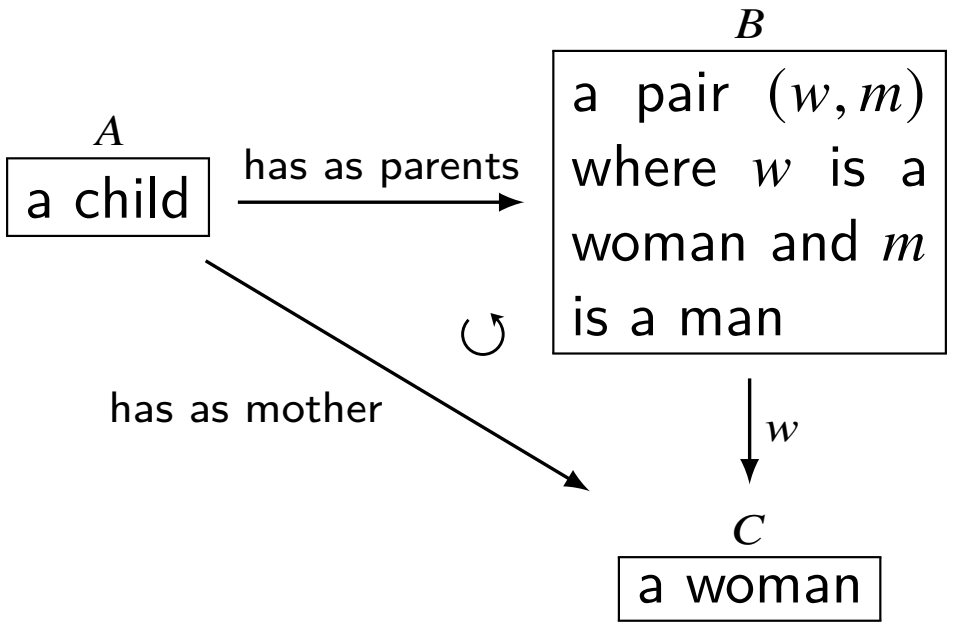


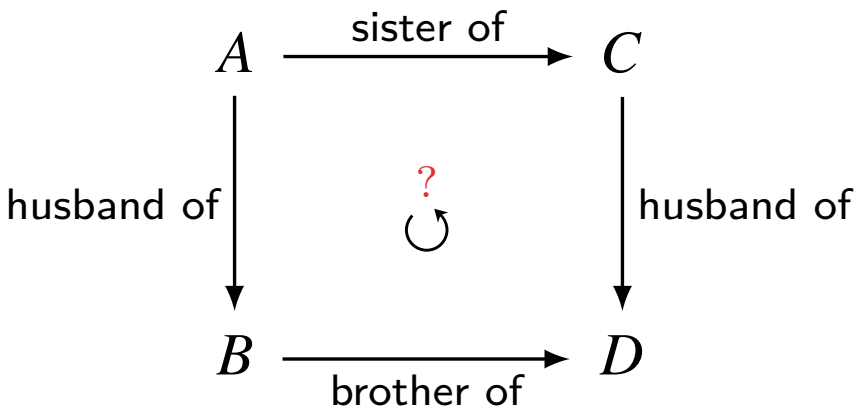
Associativity





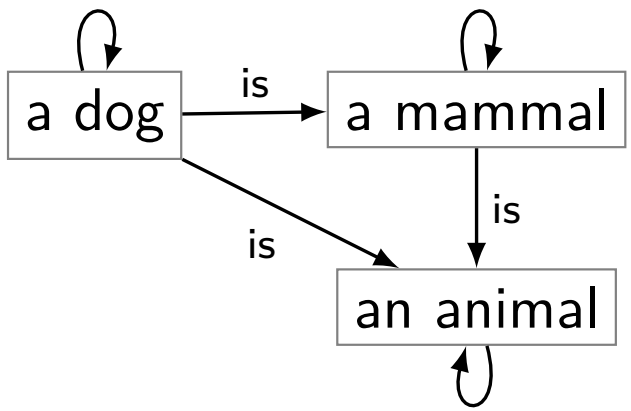
Commutative?





Example — Category

- Object: Count Nouns.
- Morphism from A to B: an A is a B.



Remark

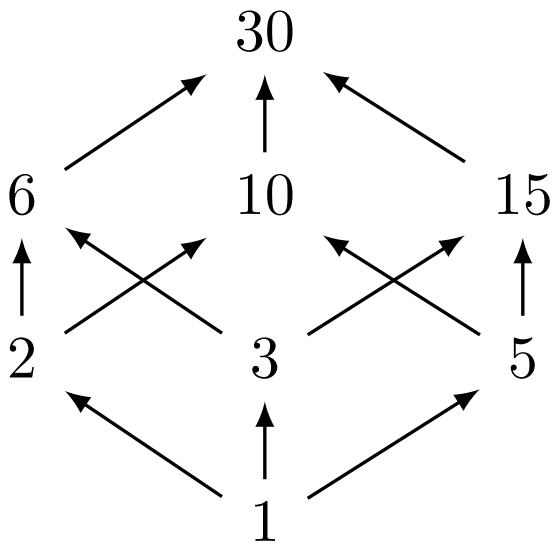
The concept of category has a first-order axiomatization, in a language having

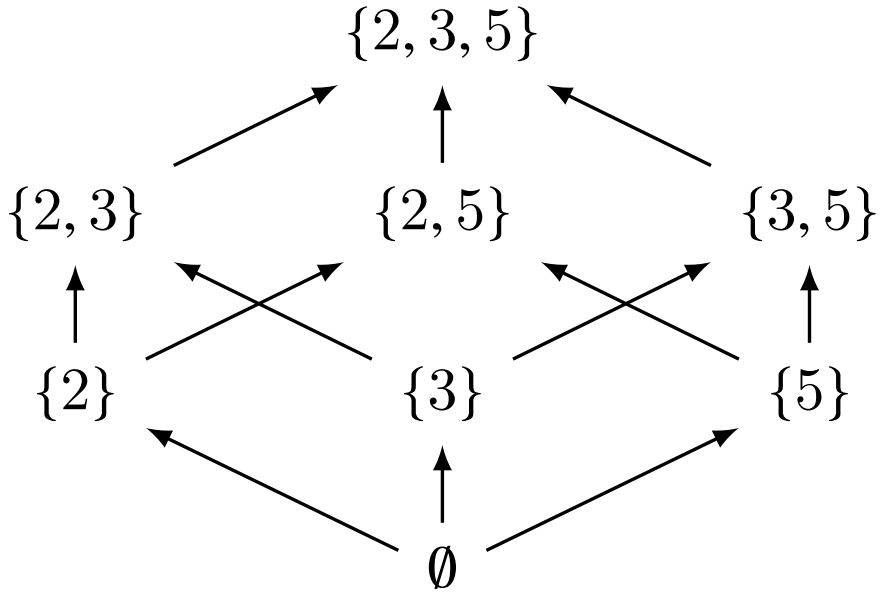
- two sorts $\mathbf{C}_0$ and $\mathbf{C}_1$ (respectively for objects and arrows),
- two unary function symbols (for domain and codomain)

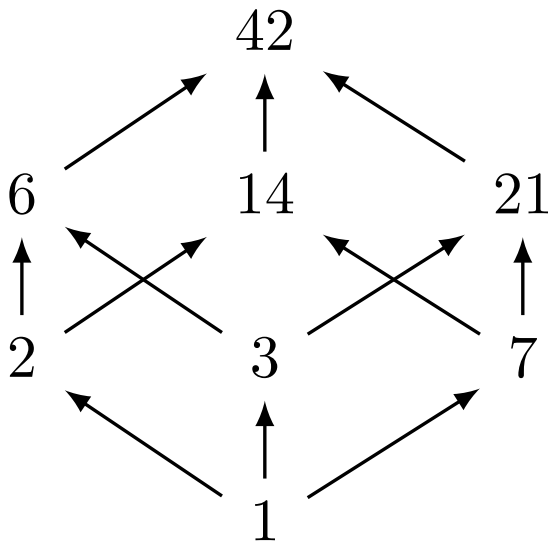
$$\mathbf{C}_1 \underset{\mathrm{cod}}{\overset{\mathrm{dom}}{\rightrightarrows}} \mathbf{C}_0$$

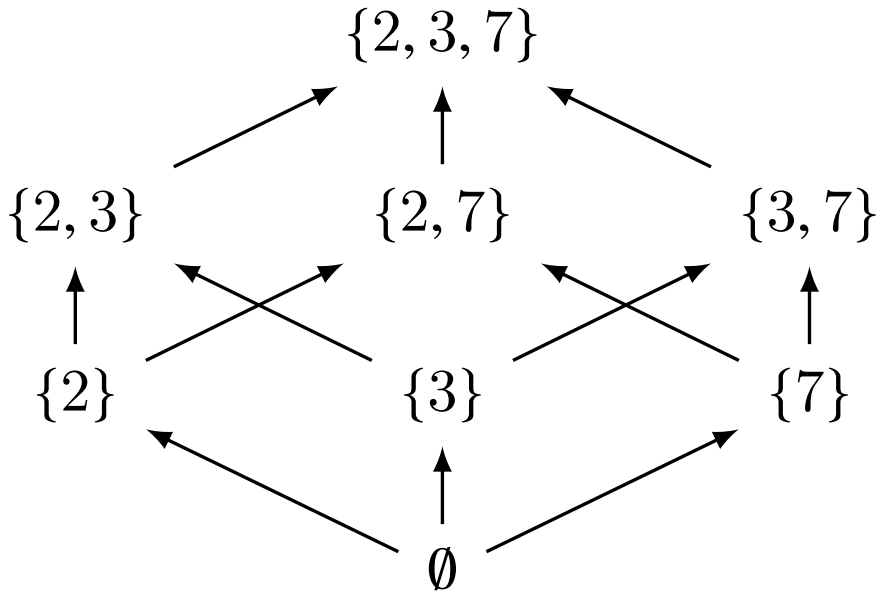
- one unary function symbol $1 : \mathbf{C}_0 \to \mathbf{C}_1$ (formalizing the concept of identity arrow) and
- a ternary predicate of type $\mathbf{C}_1$ (formalizing the notion of composition of arrows).

Example: The "Lattice" of Factors

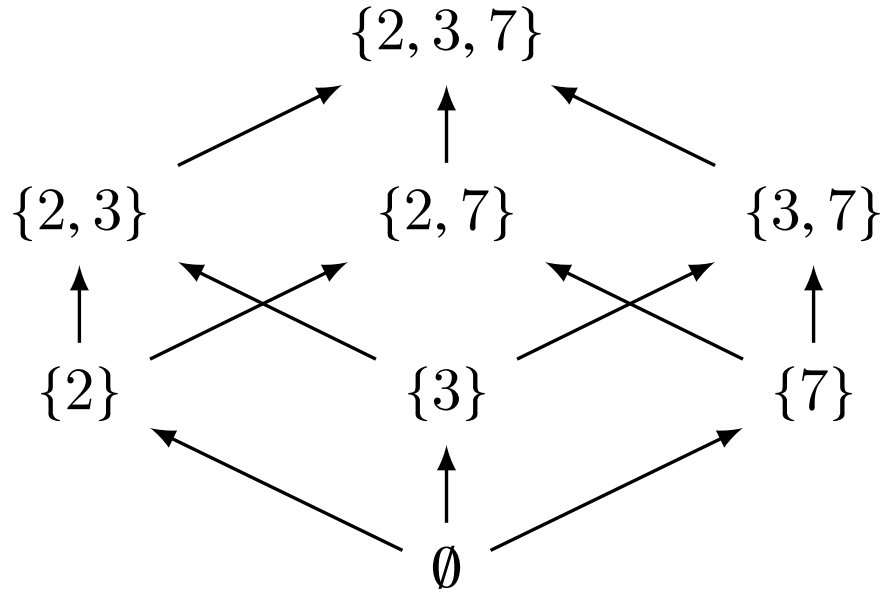


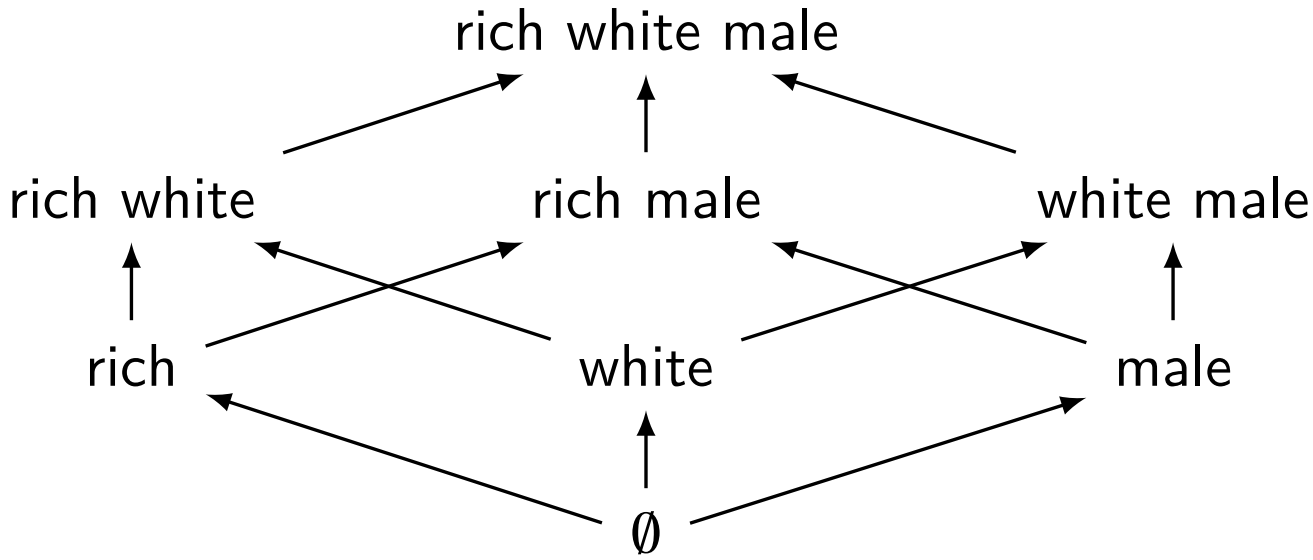






Example: The "Lattice" of Factors





A category is a network of composable relationships

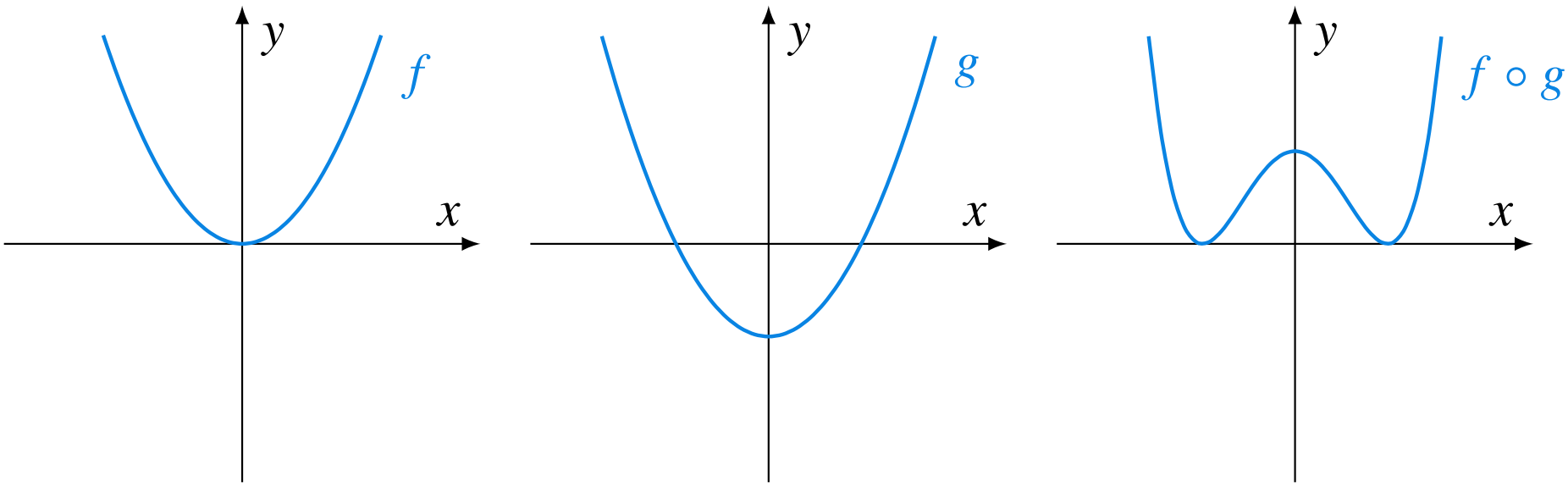
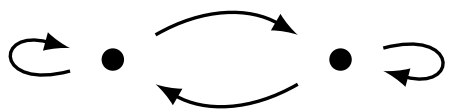


Figure: Convex functions on $\mathbb{R}$ do not form a category.

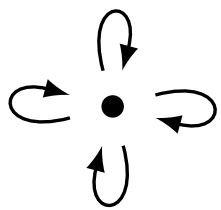
- The devil is in the morphisms!
- Objects are easy, morphisms are usually where the difficulties hide.
- Ask not what a thing is; ask what it does.

Examples — Mathematical Objects as Categories

- A *discrete category* is a category whose only morphisms are the identity morphisms. — A *set* can be seen as a discrete category.
- A *preorder* is a category having at most one morphism from one object to another.
- A *poset* is a preorder where if there's a morphism $f : x \to y$ and $g : y \to x$, then $x = y$.
- A *monoid* $(M, \cdot, e)$ is a category that has only one object $\bullet$ s.t. $\mathrm{Hom}(\bullet, \bullet) = M$.
- A *groupoid* is a category in which every morphism is an isomorphism.



- A *group* is a category that has only one object and in which every morphism is an isomorphism.



Groups as Categories

- A *group* is a category that has only one object and in which every morphism is an isomorphism.
- If G and H are groups, regarded as categories, then a functor $f : G \to H$ is exactly the same thing as a *group homomorphism*.
- What is a functor $R : G \to \mathbf{C}$ from a group G to another category $\mathbf{C}$? If $\mathbf{C} = \mathbf{Vect}_{\mathbb{K}}$, then R is a "linear representation" of G.
- In general, any functor $R : G \to \mathbf{C}$ can be regarded as a representation of G in the category $\mathbf{C}$: the elements of G become automorphisms of some object in $\mathbf{C}$. A permutation representation, for instance, is simply a functor into $\mathbf{Set}$.

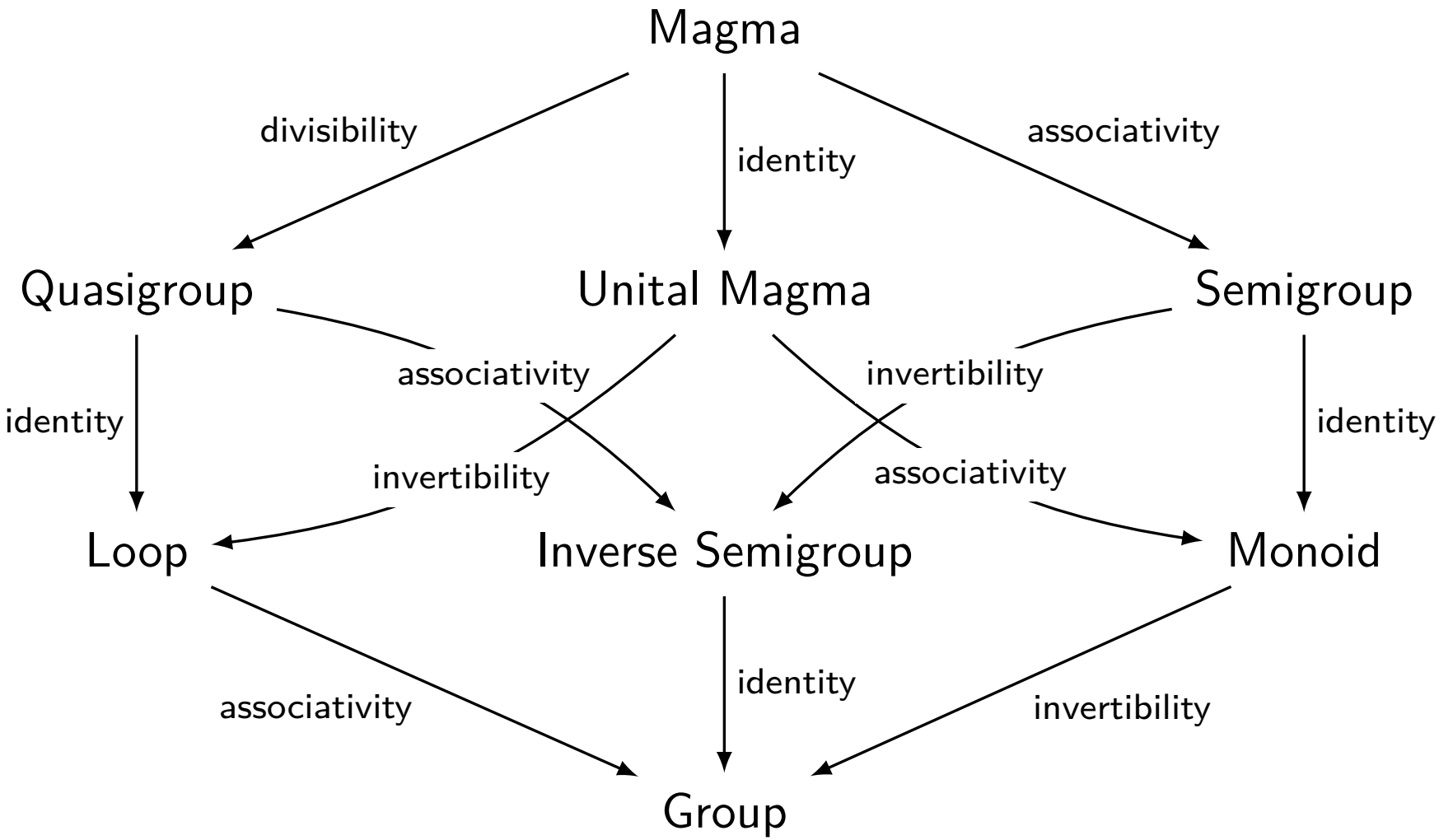
Examples — Categories of Mathematical Objects

Category	Objects	Morphisms
$\mathbf{Set}/\mathbf{FinSet}$	sets / finite sets	total functions
$\mathbf{Par}$	sets	partial functions
$\mathbf{Rel}$	sets	relations
$\mathbf{Preord}$	preorders	monotone functions
$\mathbf{Poset}$	partial order sets	monotone functions
$\mathbf{Graph}$	directed graphs	graph homomorphisms
$\mathbf{Type}$	types	recursive functions
$\mathbf{Mon}$	monoids	homomorphisms
$\mathbf{Grp}/\mathbf{AbGrp}$	groups / abelian groups	homomorphisms
$\mathbf{Rng}$	rings	ring homomorphisms
$\mathbf{Vect}_{\mathbb{K}}$	vector spaces over a field $\mathbb{K}$	linear maps
$\mathbf{Ban}_{\infty}$	real Banach spaces	bounded linear mappings
$\mathbf{Ban}_{1}$	real Banach spaces	linear contractions
$\mathbf{Top}$	topological spaces	continuous functions
$\mathbf{Diff}$	smooth manifolds	smooth maps
$\mathbf{Meas}$	measurable spaces	measurable functions

Group-like Structures

	Closure	Associativity	Identity	Invertibility	Commutativity
Semigroupoid		✓			
Category		✓	✓		
Groupoid		✓	✓	✓	
Magma	✓				
Quasigroup	✓			✓	
Unital Magma	✓		✓		
Loop	✓		✓	✓	
Semigroup	✓	✓			
Inverse Semigroup	✓	✓		✓	
Monoid	✓	✓	✓		
Commutative monoid	✓	✓	✓		✓
Group	✓	✓	✓	✓	
Abelian group	✓	✓	✓	✓	✓

Group-like Structures



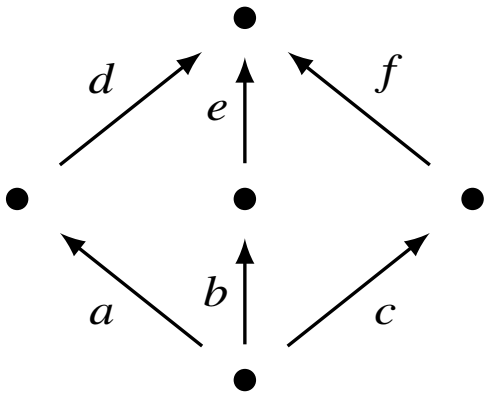
Groupoid

- Any equivalence relation R on a set X can be presented as a groupoid on X. A groupoid is a generalized equivalence relation.
- The underlying groupoid of a category $\mathbf{C}$ is an internal criterion of identity. Entities are entities in a context.

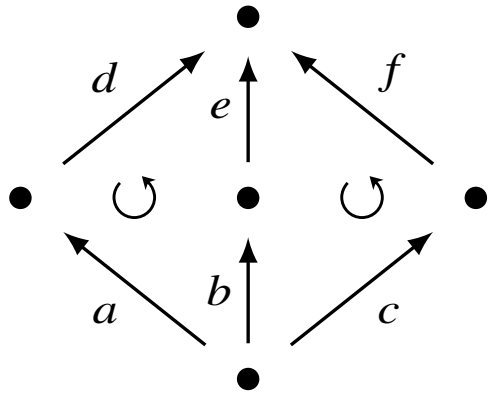
Equivalence Relation		Category		Group
objects	$\rightsquigarrow$	objects		
relations	$\rightsquigarrow$	arrows	$\leftsquigarrow$	objects
reflexivity	$\rightsquigarrow$	identities	$\leftsquigarrow$	identity
symmetry	$\rightsquigarrow$	inverses	$\leftsquigarrow$	inverses
transitivity	$\rightsquigarrow$	composition	$\leftsquigarrow$	binary operation
		unitality	$\leftsquigarrow$	unitality
		associativity	$\leftsquigarrow$	associativity

Preorder

Which graph is a preorder?

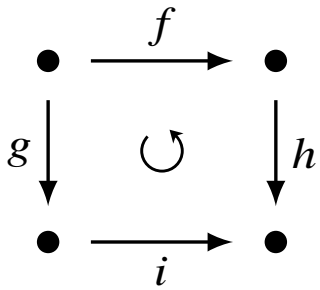


no equation

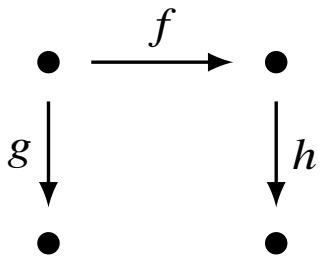


$da = eb = fc$

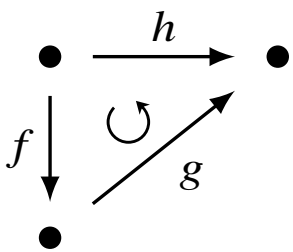
What equations do you need to make the following graphs into preorders?



$hf = ig$



no equation



$h = gf$



$f = 1_a$

Category vs Graph

- From a graph G we can get the "free category on G" $F(G)$, whose objects are the nodes of the graph G, and $\mathrm{Hom}(x, y)$ is the set of all paths from x to y. The composition of morphisms is given by concatenation of paths.
- The free category on a graph has the fewest possible equations between parallel morphisms, while a preorder has the most possible.
- 'free' means 'no extra equations'.
- Roughly speaking, category theory is graph theory with additional structure to represent composition.
- There is an adjunction

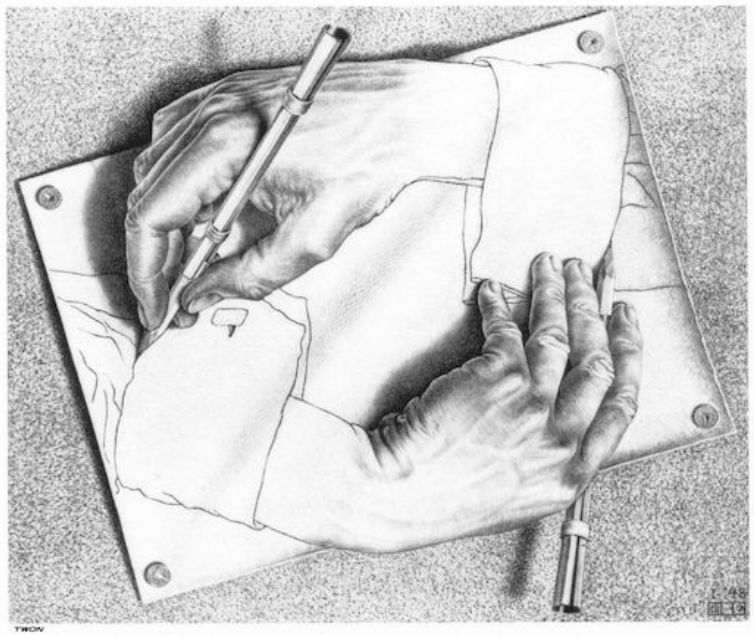
$$\mathrm{Hom}_{\mathbf{Cat}}(F(G), \mathbf{C}) \cong \mathrm{Hom}_{\mathbf{Graph}}(G, U\mathbf{C})$$

Example — Free Category



- 对应自然数集 $\mathbb{N}$
- 对应幺半群 $(\mathbb{N}, +, 0)$

Example — Drawings



There exists a category $\mathbf{Draw}$ in which:

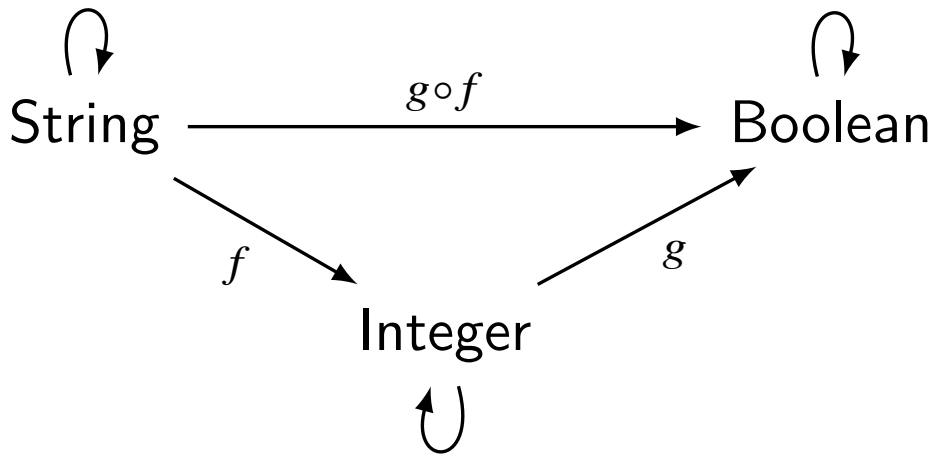
- An object is a black-and-white drawing, that is a function $\alpha : \mathbb{R}^2 \to \mathbf{2}$.
- A morphism in $\mathrm{Hom}(\alpha, \beta)$ between two drawings is an invertible map $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that $\alpha(x) = \beta(fx)$.
- The identity function at any object α is the identity map on $\mathbb{R}^2$.
- Composition is given by function composition.

Some Morphisms in **Draw**

- Scalings. Let $a, b \in \mathbb{R}$.
$$\mathrm{sc}_{a,b} : (x, y) \mapsto (ax, by)$$
- Translations.
$$\mathrm{tra}_{a,b} : (x, y) \mapsto (x + a, y + b)$$
- Rotations.
$$\mathrm{rot}_{\theta} : (x, y) \mapsto (x \cos\theta + y \sin\theta, -x \sin\theta + y \cos\theta)$$
- Roto-translations.
$$\mathrm{rotra}_{\theta,a,b} : (x, y) \mapsto (x \cos\theta + y \sin\theta + a, -x \sin\theta + y \cos\theta + b)$$
- Afine transformations (which arise from rotations, translations, and scalings). Let $\mathbf{a} \in \mathbb{R}^{2\times 2}$ and $\mathbf{b} \in \mathbb{R}^{2\times 1}$.
$$\mathrm{aff}_{\mathbf{a},\mathbf{b}} : (x, y) \mapsto (a_{11}x + a_{12}y + b_{11}, a_{21}x + a_{22}y + b_{21})$$
— Some special cases:
  - with $\mathbf{a} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ one obtains scalings $\mathrm{sc}_{a,b}$.
  - with $\mathbf{a} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} a \\ b \end{bmatrix}$ one obtains translations $\mathrm{tra}_{a,b}$.
  - with $\mathbf{a} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ one obtains rotations rot_{θ}.

The Category of Typed Programming Language

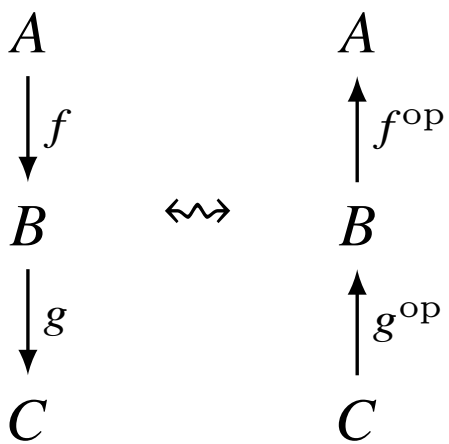
- A category of typed programming language is a category whose objects are datatypes (integers, strings, booleans), and whose morphisms are the possible programs.
- Each datatype A is a collection of possible values for a datum having that type. For example, integers or strings or Booleans.
- The product type $A \times B$ represents vectors, arrays and records.
- The function type $A \longrightarrow B$ represents programs. Each term of type $A \longrightarrow B$ is a program that transforms its argument of type A into an output of type B.



Opposite Category

Definition (Opposite Category)

The *opposite category* $\mathbf{C}^{\mathrm{op}}$ of $\mathbf{C}$ is formed by reversing the morphisms. Formally, $\mathrm{ob}(\mathbf{C}^{\mathrm{op}}) = \mathrm{ob}(\mathbf{C})$, and for each morphism $f : A \longrightarrow B$ of $\mathbf{C}$ a morphism $f^{\mathrm{op}} : B \longrightarrow A$ in $\mathbf{C}^{\mathrm{op}}$, with the same identities and a composition defined (when possible) by $f^{\mathrm{op}} \circ g^{\mathrm{op}} := (g \circ f)^{\mathrm{op}}$.



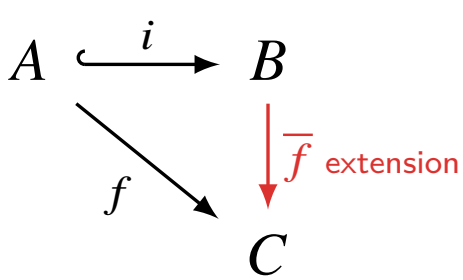
The Duality Principle

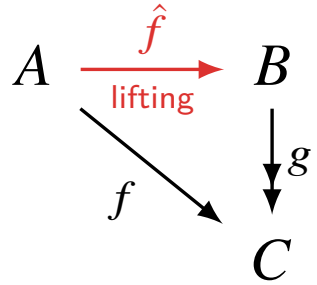
- Every statement formulated in the language of Category Theory has a dual, obtained by formally reversing the arrows and the order of composition of them.
- A statement φ is true in a category $\mathbf{C}$ iff the dual statement φ^{op} is true in the dual category $\mathbf{C}^{\mathrm{op}}$.
- Hence a statement is valid in all categories iff its dual is.

"A comathematician is a machine for turning cotheorems into ffee."
"A mathematician is a machine for turning coffee into theorems."

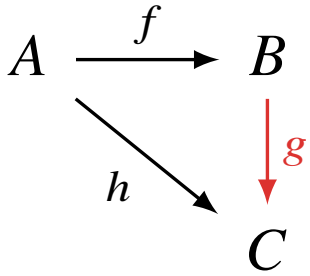
Composition, Extension, Lifting

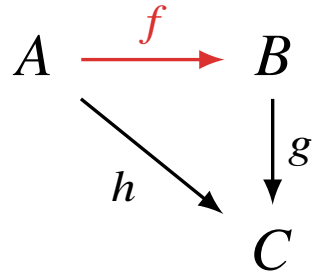
- If $f : A \to C$ is a function and A is a subset of B with the inclusion function $i : A \hookrightarrow B$, then an extension of f along i is a function $\overline{f} : B \to C$ such that $f = \overline{f} \circ i$.
- Consider an onto function $g : B \twoheadrightarrow C$. Let $f : A \to C$ be any function. A lifting of f along g is a function $\hat{f} : A \to B$ such that $f = g \circ \hat{f}$.





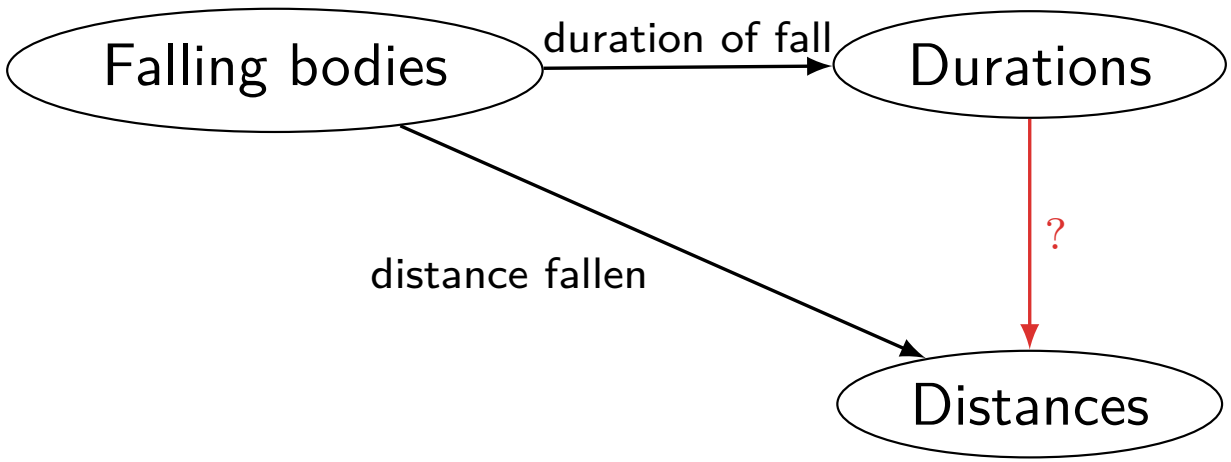
1. The 'determination' (or 'extension') problem: given f and h, what are all g, if any, for which $h = g \circ f$?
2. The 'choice' (or 'lifting') problem: given g and h, what are all f, if any, for which $h = g \circ f$?



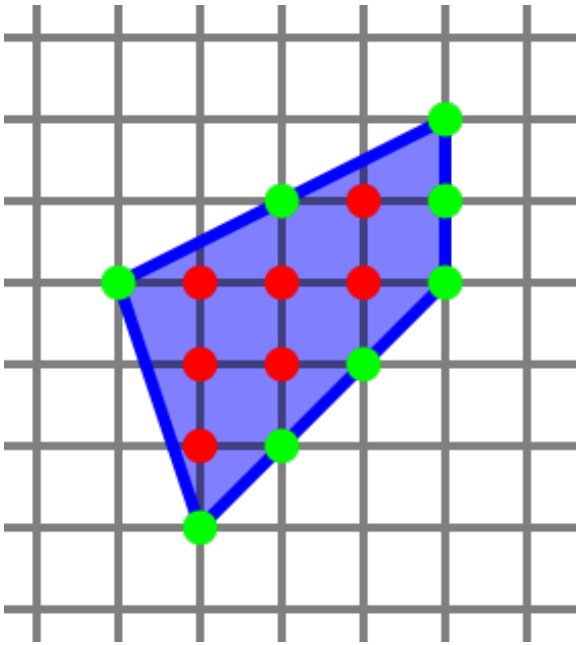


Example: the 'determination' (or 'extension') problem

Once Galileo realized that the distance of a dropped object falls in a certain time is determined by the time, it did not take too many experiments before he found a function for the distance in terms of the time $d = \frac{1}{2}gt^2$.



Example: the 'determination' (or 'extension') problem

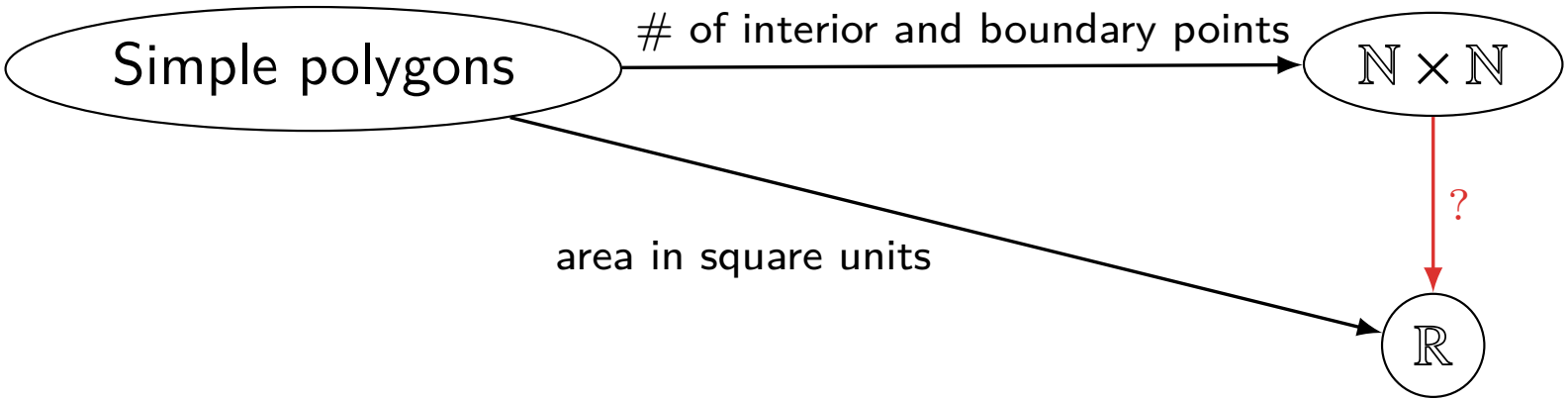


Pick's Theorem

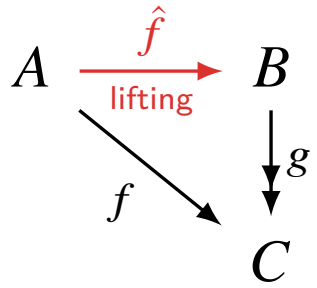
The *Area* of a simple polygon with integer vertex coordinates is determined by the number of its *Interior* points and its *Boundary* points.

$$A = I + \frac{B}{2} - 1$$

$$A = 7 + \frac{8}{2} - 1 = 10$$



Remarks: the "choice" (or "lifting") problem



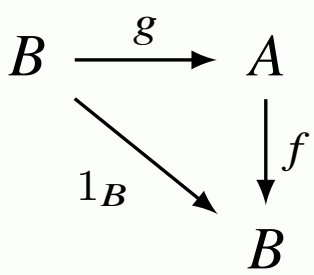
- If no further requirements are placed on the lift $\hat{f}$, then the axiom of choice assures that the lifting problem is solvable.
- The axiom of choice lets us select a preimage $h(c) \in g^{-1}(c)$, and one can lift f by $\hat{f} := h \circ f$.
- Conversely, to build a choice function for a surjective map g, it suffices to lift the identity map $1_C : C \to C$ to B.
- However, the lifting provided by the axiom of choice are pathological, being almost certain to be discontinuous, non-measurable, etc..
- In many applications we would like to have $\hat{f}$ preserve many of the properties of f (e.g., continuity, differentiability, linearity, etc.).

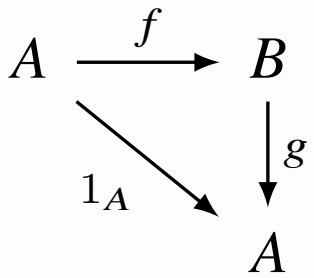
Monic / Epic / Bimorphism / Isomorphism

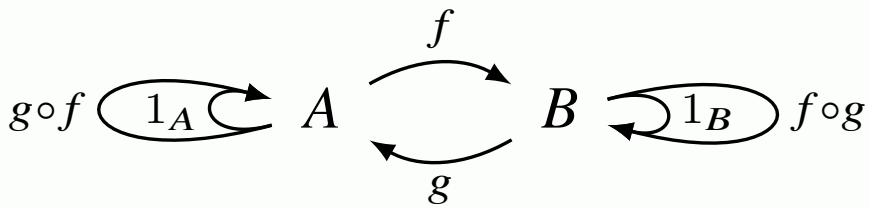
Definition

A morphism $f : A \to B$ is a:

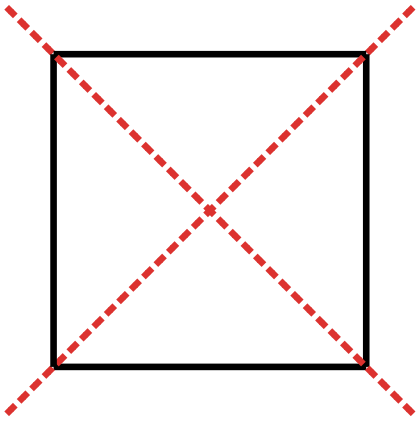
1. monomorphism (monic) iff
   $\forall X \forall g_1 g_2 : X \to A : fg_1 = fg_2 \implies g_1 = g_2$.
2. epimorphism (epic) iff $\forall X \forall g_1 g_2 : B \to X : g_1 f = g_2 f \implies g_1 = g_2$.
3. bimorphism iff f is both monic and epic.
4. isomorphism iff $\exists g : B \to A : gf = 1_A \ \& \ fg = 1_B$.
5. endomorphism iff $A = B$.
6. automorphism iff f is both an endomorphism and an isomorphism.
7. retraction iff a right inverse of f exists, i.e. $\exists g : B \to A : fg = 1_B$.
8. section iff a left inverse of f exists, i.e. $\exists g : B \to A : gf = 1_A$.

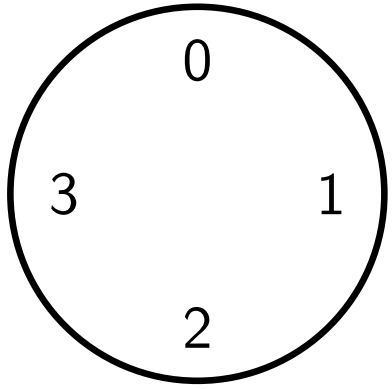






Group Isomorphism





	0	90	180	270
0	0	90	180	270
90	90	180	270	0
180	180	270	0	90
270	270	0	90	180

$\mathbb{Z}_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$$0 \longleftrightarrow 0$$
$$90 \longleftrightarrow 1$$
$$180 \longleftrightarrow 2$$
$$270 \longleftrightarrow 3$$

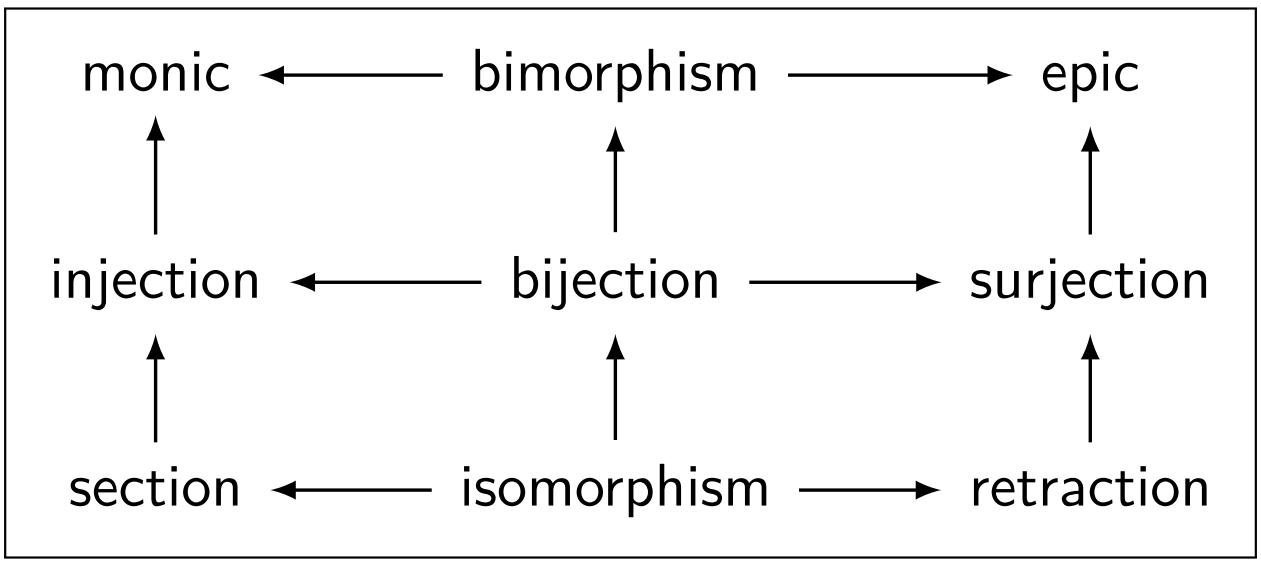
$\circlearrowleft$

$$0 \longleftrightarrow 0$$
$$90 \longleftrightarrow 3$$
$$180 \longleftrightarrow 2$$
$$270 \longleftrightarrow 1$$

$\circlearrowleft$

Monic / Epic / Bimorphism / Isomorphism

- Every retraction is epic.
- Every section is monic.
- In a concrete category every section is injective.
- In a concrete category every retraction is surjective.
- The following three statements are equivalent:
  1. f is a monomorphism and a retraction;
  2. f is an epimorphism and a section;
  3. f is an isomorphism.



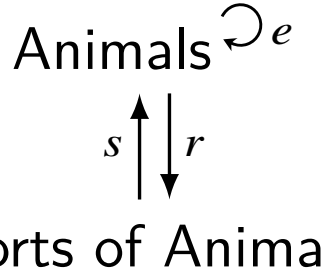
Idempotents as Records of Retracts

If $A \xrightarrow{s} B \xrightarrow{r} A$, and $rs = 1_A$, then $e := sr$ is idempotent: $ee = e$.

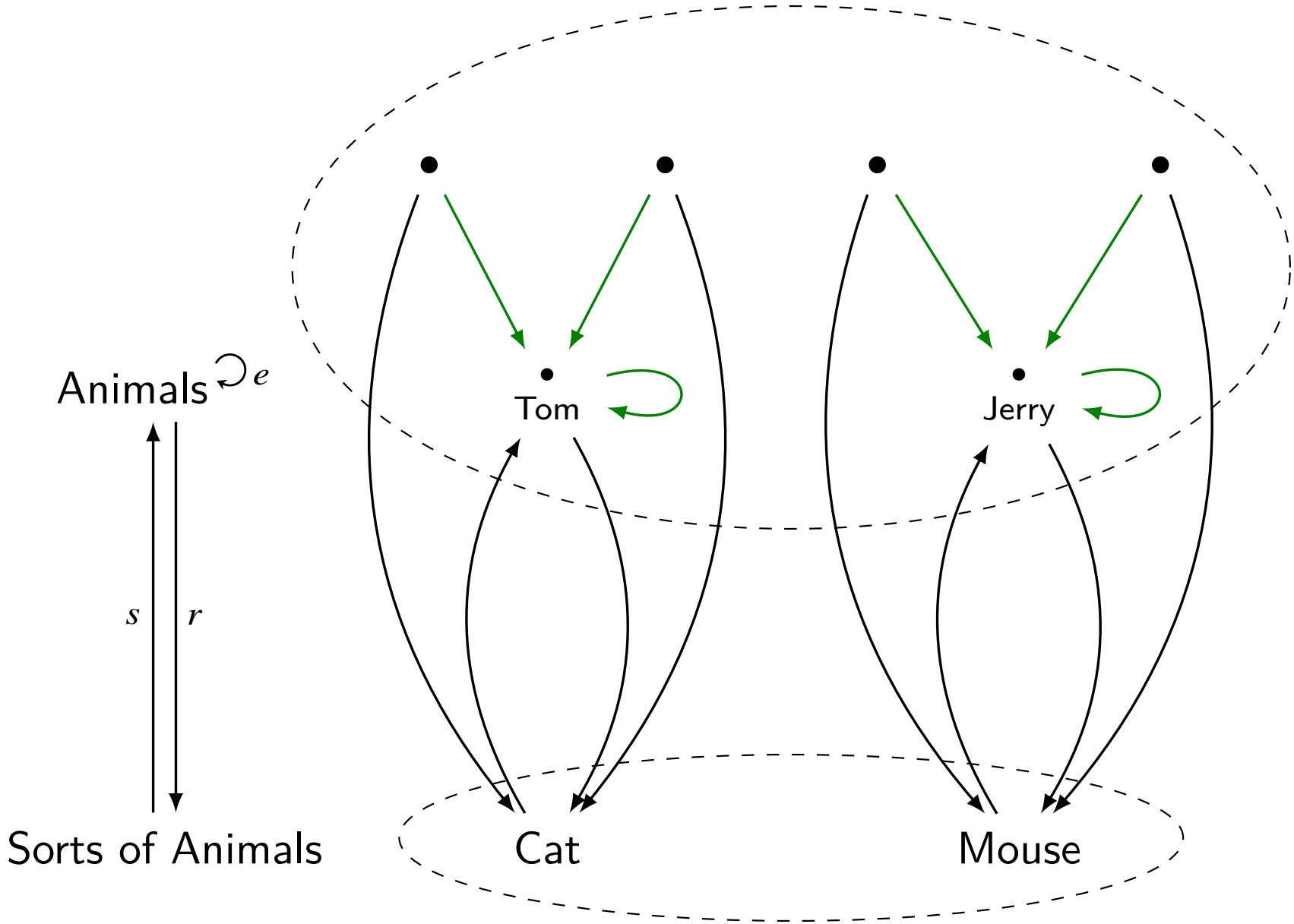
Definition

A splitting of an idempotent map $B \xrightarrow{e} B$ consists of an object A together with two maps $A \underset{r}{\overset{s}{\rightleftarrows}} B$ s.t. $rs = 1_A$ and $sr = e$.

Example: Given a variety of animals, and an idempotent map e which should assign to each animal the most familiar animal it closely resembles. How to form the abstract idea of 'sorts of animals' (e.g. cat, dog, cow)?



Remark: For the children, the selection of the idempotent map and the learning of the sort-names go on concurrently.


Animals
e
s
r
Sorts of Animals
Tom
Jerry
Cat
Mouse

- In $\mathbf{Set}$, iso is equivalent to monic-plus-epic.
- The inclusion $i : \mathbb{N} \to \mathbb{Z}$ is monic and epic in the category of monoids $\mathbf{Mon}$, but it is not iso.
- In $\mathbf{Set}$, isomorphic objects are equicardinal.
- In $\mathbf{Top}$, isomorphic objects are homeomorphic spaces.
- In $\mathbf{hTop}$, isomorphic objects are homotopic spaces.
- In $\mathbf{Man}$, isomorphic objects are diffeomorphic spaces.

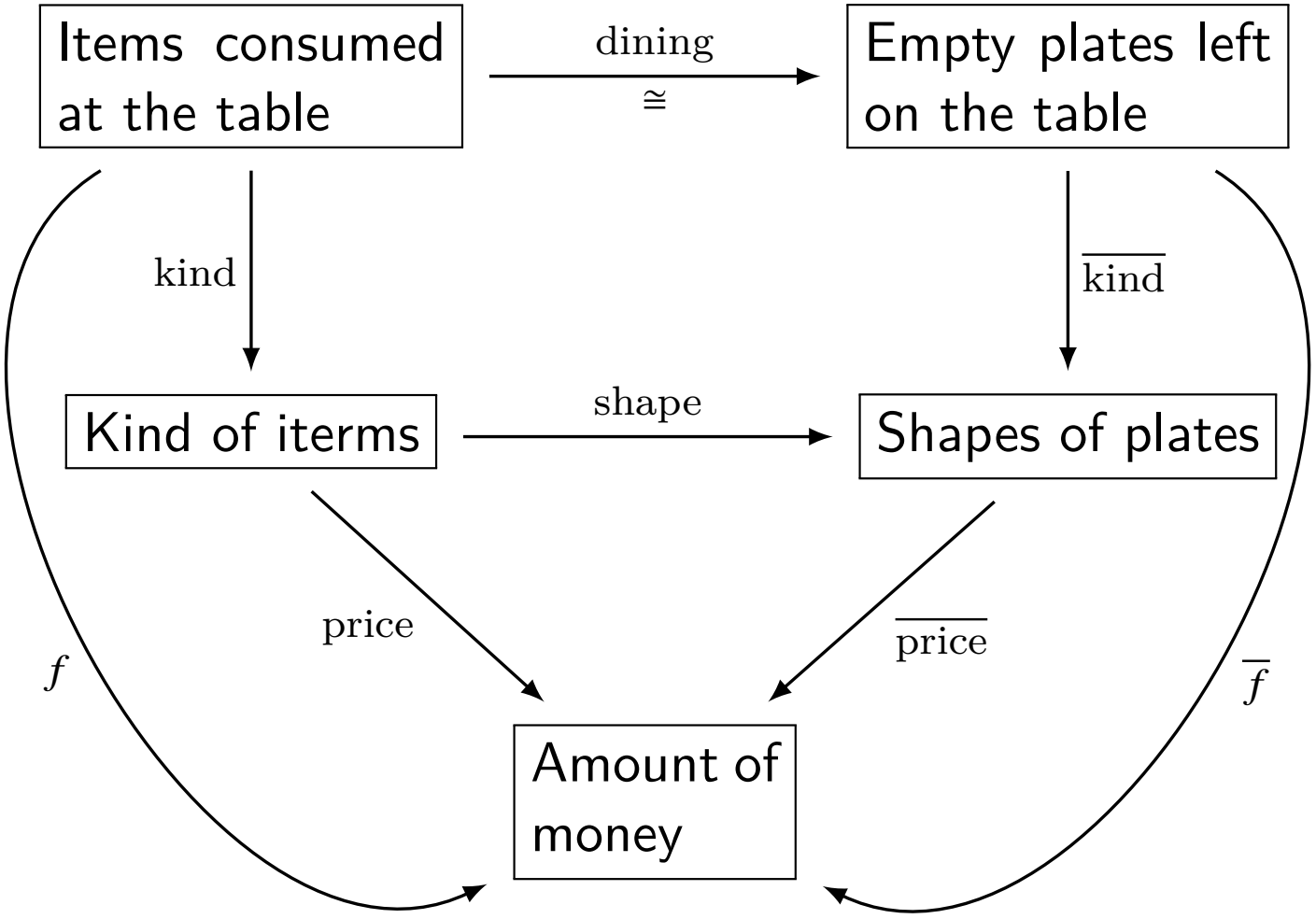
Remark

- **Theorem:** A morphism of monoids is an isomorphism iff it is a bijective homomorphism.
- **Question:** A morphism of Topological spaces is an isomorphism if it is a bijective homomorphism?
- **Answer:** No!
- A function that is continuous can have an inverse function that is not continuous, so being bijective is not enough to ensure that a continuous map is an isomorphism in $\mathbf{Top}$.

$$f : [0, 1) \to \left\{z \in \mathbb{C} : |z| = 1\right\} :: t \mapsto \mathrm{e}^{2\pi \mathrm{i} t}$$

- The function f is continuous, because it doesn't break the interval apart.
- It is bijective, because it wraps the interval around the circle without any overlap, and without any gaps.
- However, its inverse f^{-1} as a function is not continuous because it "breaks" the circle apart to go back to being an interval.

大学自助食堂



$$\sum_k \text{price}(k)\cdot(\text{size of the stack of kind over } k) = \sum_x x\cdot(\text{size of the stack of } f \text{ over } x)$$

$$\sum_k \overline{\text{price}}(s)\cdot(\text{size of the stack of } \overline{\text{kind}} \text{ over } s) = \sum_x x\cdot(\text{size of the stack of } \overline{f} \text{ over } x)$$

- All equations are lies except $x = x$.
- Equality doesn't mean a and b are the same — it is about when the world should treat them the same.
- Isomorphic objects are treated as the same by the rest of the category.
- Things can be isomorphic in one category but not another.
- Everything should be understood in context.
- Something not universal in one place can be universal in another.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

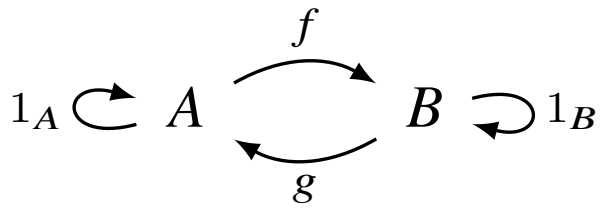
Initial Object & Terminal Object

Definition (Initial Object & Terminal Object)

- An *initial object* in $\mathbf{C}$ is an object 0 s.t. for every object A, there is a unique morphism $!_A : 0 \to A$.
- A *terminal object* in $\mathbf{C}$ is an object 1 s.t. for every object A, there is a unique morphism $!_A : A \to 1$.
- If an object is both initial and terminal, it is called a *zero object*.
- An initial object 0 is called a *strict initial object* iff every morphism $x \to 0$ is an isomorphism.
- A *zero morphism* $0 : A \to B$ is the unique morphism factoring over the zero object $0 : A \to 0 \to B$
- A category with a zero object is called *pointed*.

- An initial object of $\mathbf{C}$ is a terminal object of $\mathbf{C}^{\mathrm{op}}$, and vice-versa.
- Initial and terminal objects are unique up to unique isomorphism.

Example



- If $g \circ f \neq 1_A$, then A can't be initial, since we have at least two arrows: 1_A and $g \circ f$.
- If $g \circ f = 1_A$ and $f \circ g = 1_B$, then both A and B are initial.
  - A is initial: there are unique arrows $A \overset{1_A}{\circlearrowright}$ and $A \xrightarrow{f} B$.
  - A is initial: there are unique arrows $B \overset{1_B}{\circlearrowright}$ and $B \xrightarrow{g} A$.

  — And there is a unique isomorphism between them.

Examples

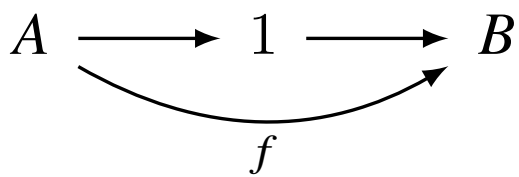
- In $\mathbf{Set}$, any one-element set $\{\bullet\}$ is terminal. The empty set $\emptyset$ is initial. — Suppose we have $|A|$ inputs and $|B|$ outputs then the number of possible functions is $|B|^{|A|}$. Now $|B|^{|A|} = 1 \iff |A| = 0 \vee |B| = 1$.
- In $\mathbf{Rel}$, the empty set $\emptyset$ is both initial and terminal.
- In $\mathbf{Poset}$, the poset $(\emptyset, \emptyset)$ is initial, while $(\{\bullet\}, \{\bullet, \bullet\})$ is terminal.
- In a poset, seen as a category, an initial object is a least element $\bot$, while a terminal object is a greatest element $\top$.
- In a given Boolean algebra, 0 is initial, and 1 is terminal.
- In $\mathbf{BoolAlg}$, the 2-element algebra $\{0, 1\}$ is initial, and the single-element algebra $\{0\}$ is terminal.

Examples

- In $\mathbf{Top}$, the one-element topological space $(\{\bullet\}, \{\emptyset, \{\bullet\}\})$ is terminal, and the empty topological space $(\emptyset, \{\emptyset\})$ is initial.
  — The unique function from any space to the one-element space is continuous because everything lands on the same point, thus everything ends up close together and nothing is broken apart.
- In $\mathbf{Top}_\bullet$, the singleton space is both initial and terminal.
- In $\mathbf{Grp}$, the one-element group $(\{e\}, \cdot, e)$ is both initial and terminal. Similarly in $\mathbf{Mon}$ and R-$\mathbf{Mod}$.
- In $\mathbf{Rng}$, the ring $\mathbb{Z}$ is initial because ring homomorphisms necessarily preserve both the additive and multiplicative identity, and the zero ring $\{0\}$ is terminal.
- In $\mathbf{FdVect}$, the zero vector space $\{\mathbf{0}\}$ is both initial and terminal.

Remarks

- The structure of an object is defined by the arrows pointing at it.
- The terminal object 1 has the simplest structure. There is only one way of probing it from any object.
- If there is more than one arrow coming from the terminal object 1 to some object X, it means that X has some structure.



A map f that can be factored through 1 is called a constant map.

The Uniqueness of Morphisms

1. Unique.
   — For example, in $\mathbf{Set}$, there is only one empty set.
2. Unique up to a unique isomorphism.
   — For example, in $\mathbf{Set}$, there are many one-element sets, and they are all isomorphic to each other with a unique isomorphism between them.

$$\{a\} \to \{b\}$$

3. Unique up to an isomorphism.
   — For example, in $\mathbf{Set}$, there are many sets with three objects. All these sets are isomorphic to each other, however, there are six possible isomorphisms between any two such sets.

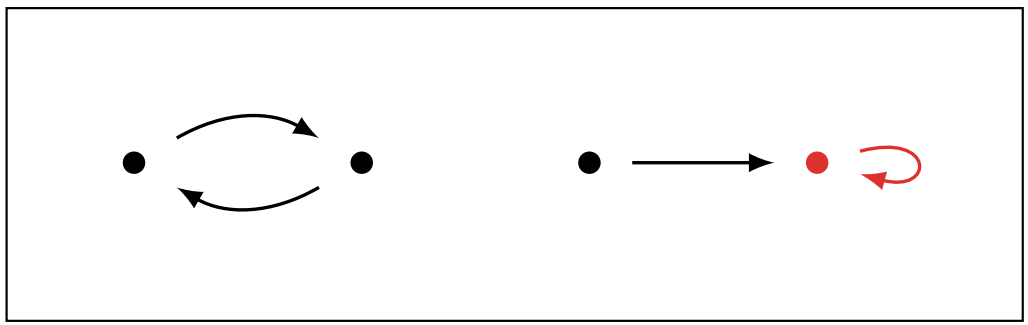
$$\{a, b, c\} \to \{x, y, z\}$$

Point

Definition (Point)

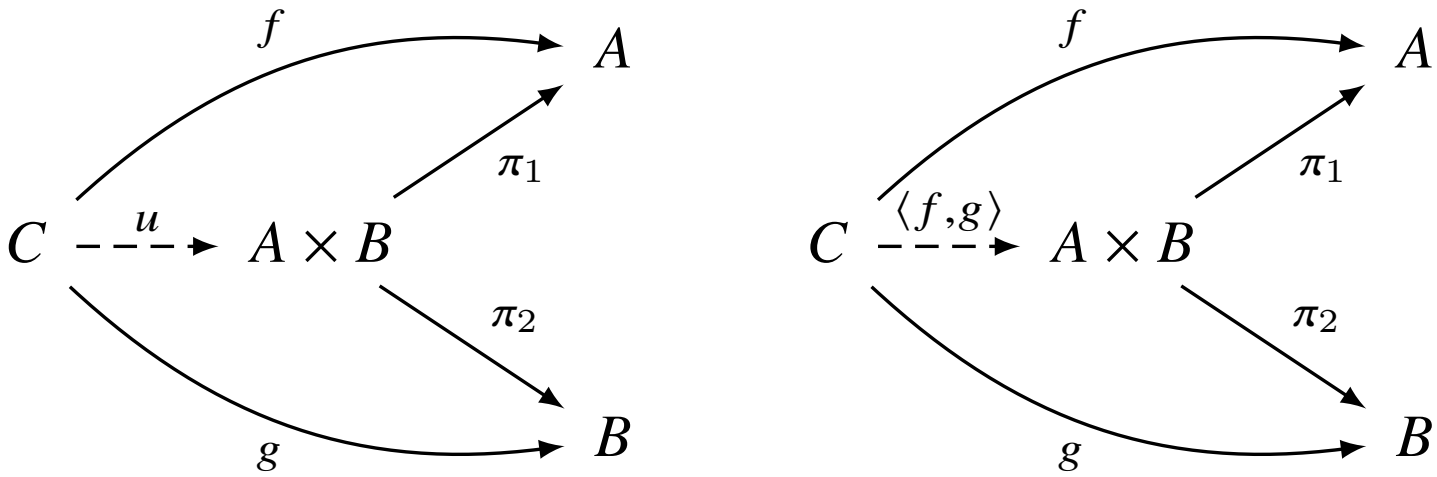
A point of an object $A \in \mathrm{ob}(\mathbf{C})$ is a morphism $1 \to A$.

- In $\mathbf{Set}$, the points of a set are the elements of that set.
- In $\mathbf{Set}^{\circlearrowright}$, the points $\bullet^{\circlearrowright} \to A^{\circlearrowright f}$ are "fixpoints".



The above object $A^{\circlearrowright f} \in \mathbf{Set}^{\circlearrowright}$ has only one point.

Product



Definition (Product)

A *product* of A and B is an object $A \times B$ with a pair of morphisms $A \xleftarrow{\pi_1} A \times B \xrightarrow{\pi_2} B$ s.t

$$\forall C \forall f g : A \xleftarrow{f} C \xrightarrow{g} B \; \exists! u : C \to A \times B \big[f = \pi_1 \circ u \;\; \& \;\; g = \pi_2 \circ u \big]$$

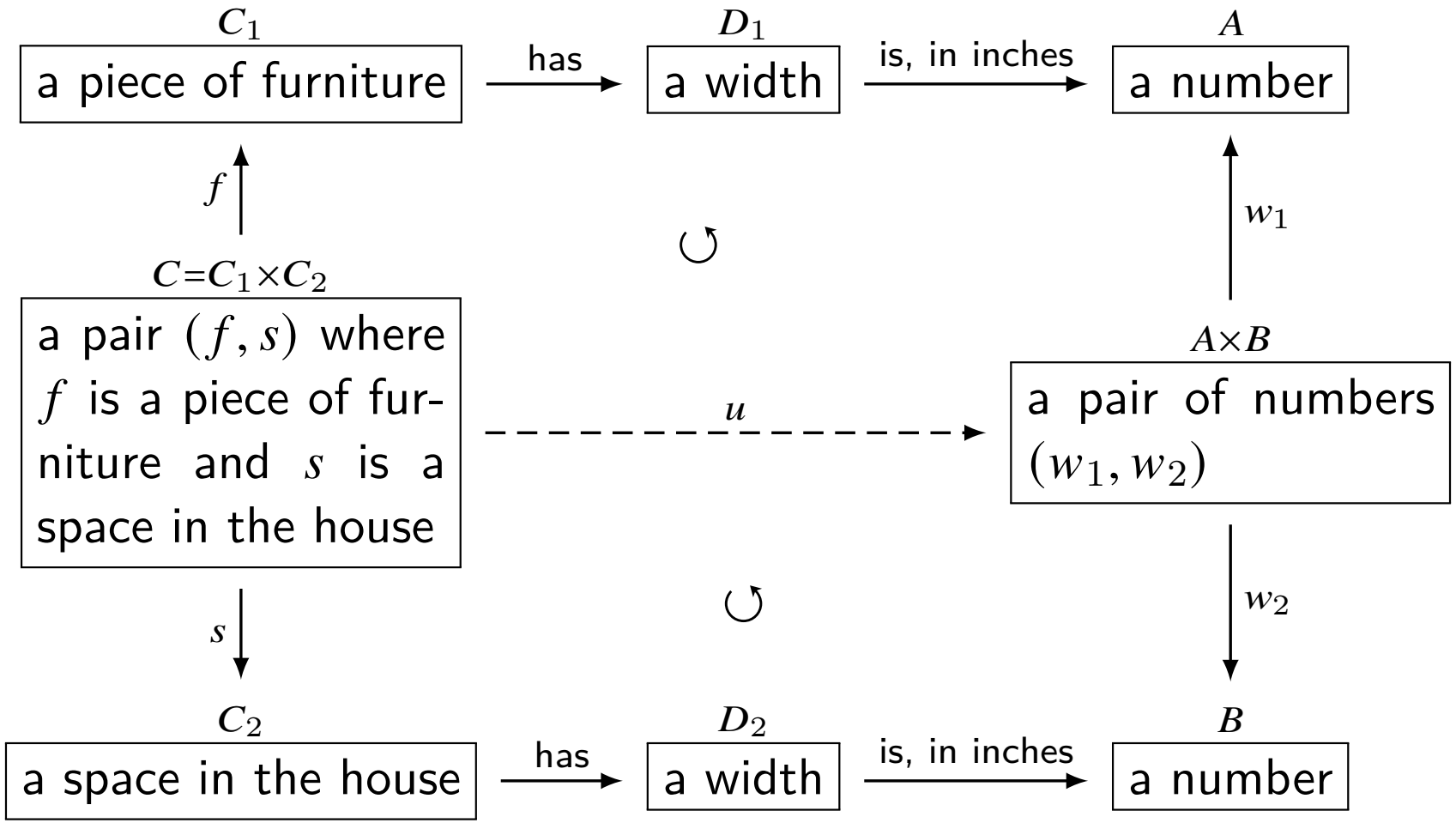
Examples

- In $(\mathbb{R}, \leq)$, products are $\min\{x, y\}$.
- In $(\mathrm{P}(A), \subset)$, products are $x \cap y$.
- In $(\mathbb{N}, |)$, products are $\gcd(x, y)$.
- In logic, products are $A \wedge B$.

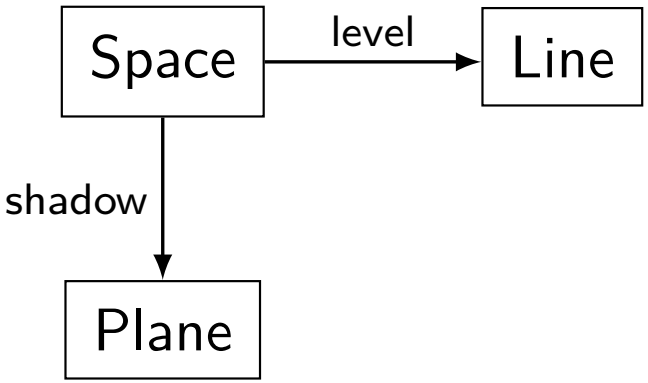
$$A \wedge B \vdash A \qquad A \wedge B \vdash B \qquad \frac{C \vdash A \quad C \vdash B}{C \vdash A \wedge B}$$

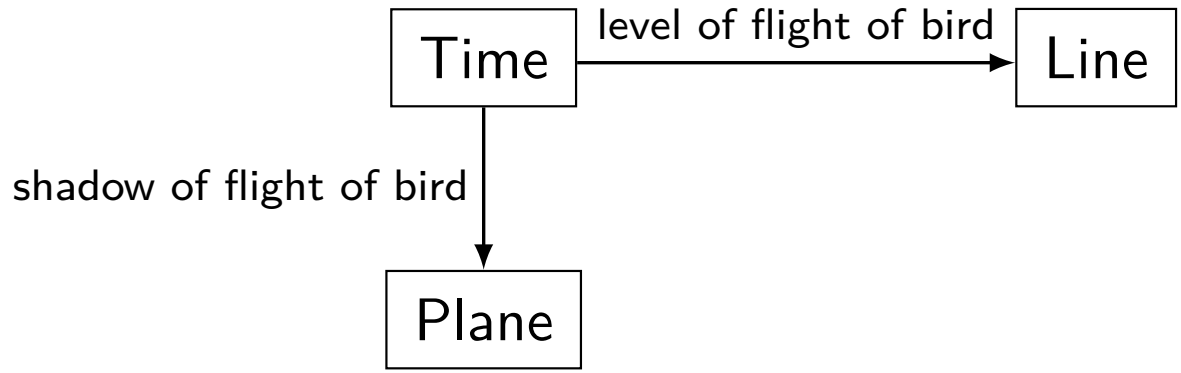
- In $\mathbf{Set}$, products are the usual cartesian products.
- In $\mathbf{Poset}$, products are cartesian products with the pointwise order.
- In $\mathbf{Top}$, products are cartesian products with the product topology.
- In $\mathbf{FdVect}_{\mathbb{K}}$, the product of V and W is their direct sum $V \oplus W$.
- In $\mathbf{Group}$, the product of groups is their direct product $\prod_i G_i$. Concretely, it's a group whose elements are sequences $(g_1, g_2, \dots)$. The group operation is given componentwise. The identity element is the sequence $(e_1, e_2, \dots)$ where e_i is the identity in G_i. One has to verify that the projection $\pi_i : \prod_i G_i \to G_i :: (g_1, g_2, \dots) \mapsto g_i$ is a group homomorphism.

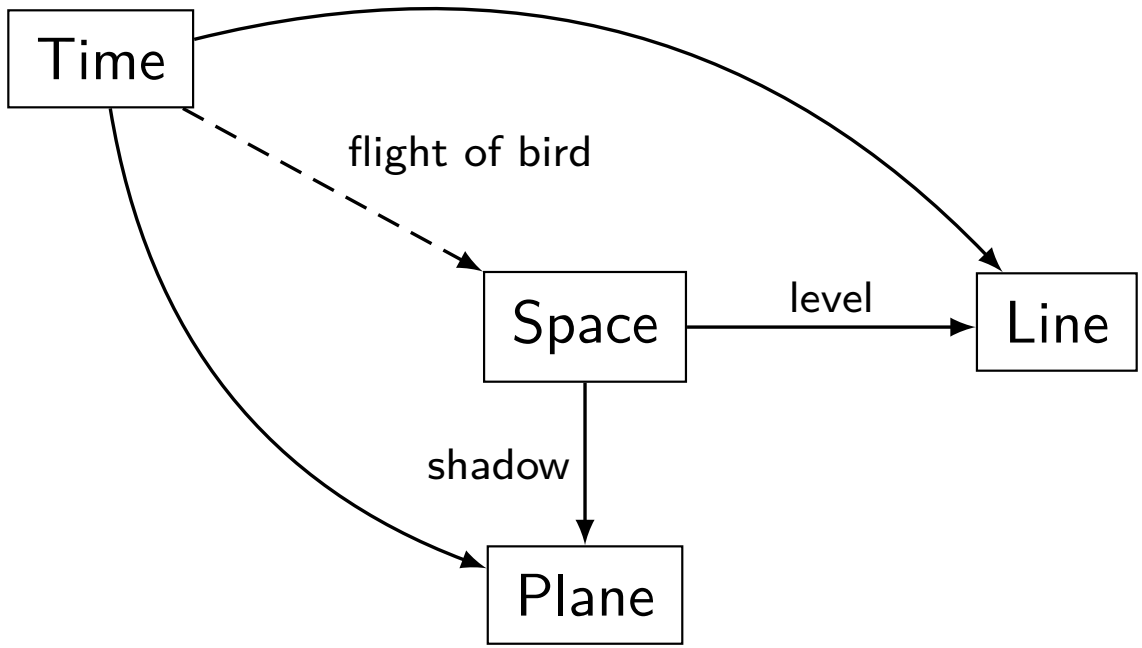
Example — Product



- If a bird is in our space, and you know only the shadow of the bird and the level of the bird, then you can reconstruct the position of the bird.
- If you have a motion picture of the bird's shadow as it flies, and a motion picture of its level, then you can reconstruct the entire flight of the bird!

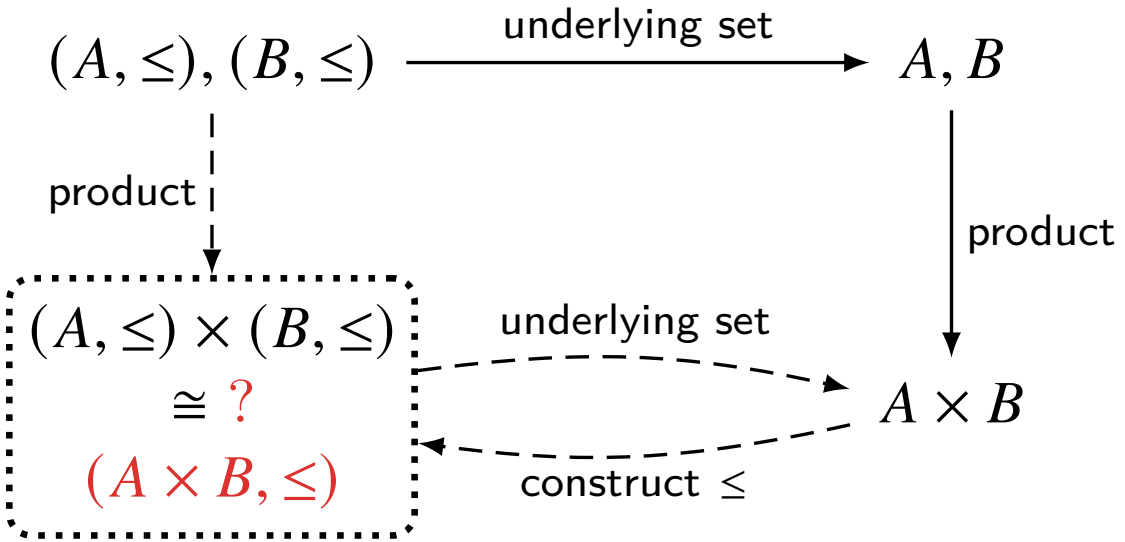


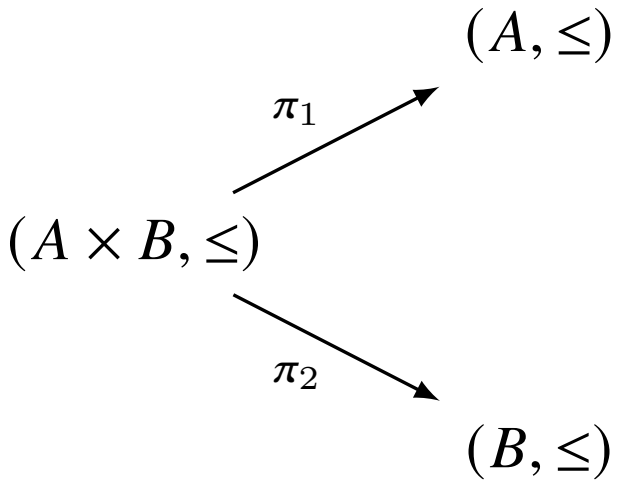




Example — Product

Let $(A, \leq)$ and $(B, \leq)$ be posets.
How to define the categorical product $(A, \leq) \times (B, \leq)$?





since π_1, π_2 preserves $\leq$,

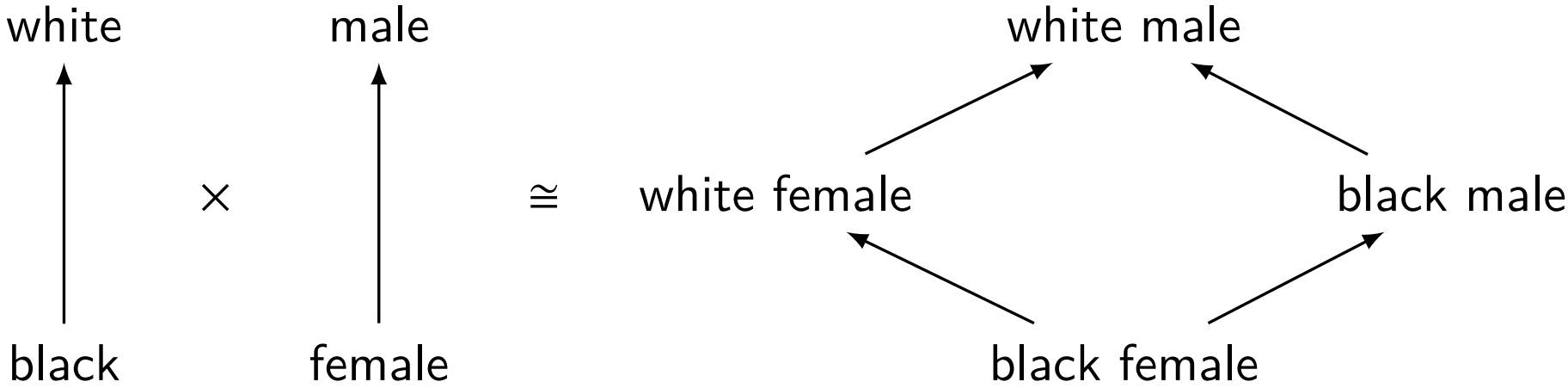
$$(a_1, b_1) \leq (a_2, b_2) \implies \pi_1(a_1, b_1) \leq \pi_1(a_2, b_2) \ \& \ \pi_2(a_1, b_1) \leq \pi_2(a_2, b_2)$$

we define $\leq$ on $A \times B$ by

$$(a_1, b_1) \leq (a_2, b_2) := a_1 \leq a_2 \wedge b_1 \leq b_2$$

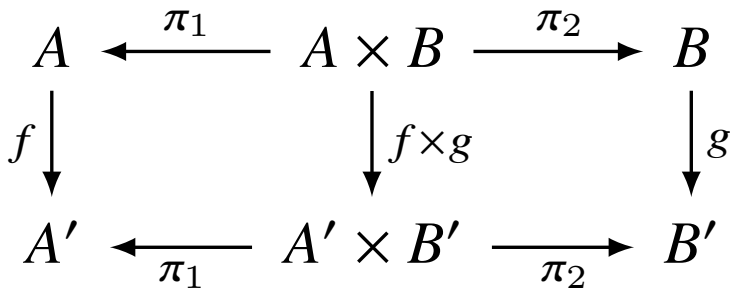
and this makes $(A \times B, \leq)$ the product.

$$(A, \leq) \times (B, \leq) \cong (A \times B, \leq)$$

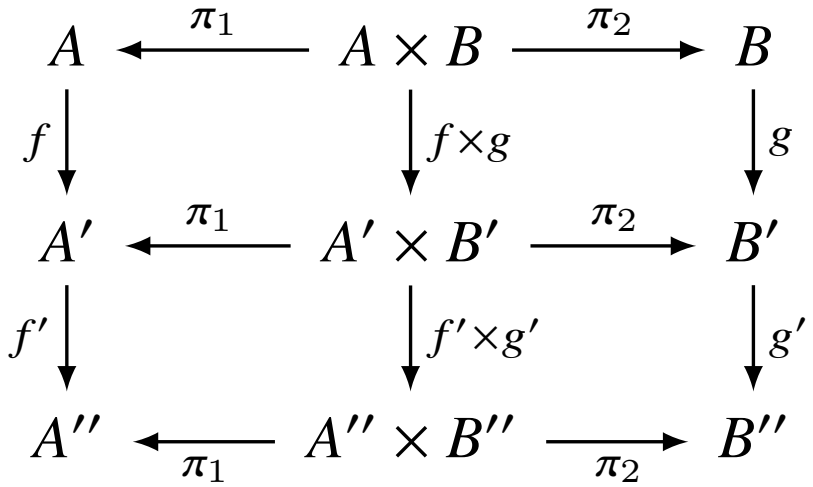


$$(a_1, b_1) \leq (a_2, b_2) := a_1 \leq a_2 \wedge b_1 \leq b_2$$

The Interchange Law



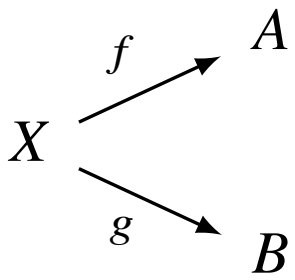
$$f \times g \coloneqq \langle f \circ \pi_1, g \circ \pi_2 \rangle \qquad (f \times g)(a, b) = (fa, gb)$$

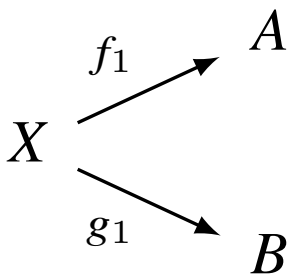
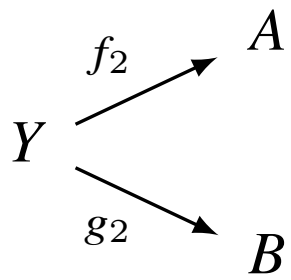


$$(f' \times g') \circ (f \times g) = (f' \circ f) \times (g' \circ g)$$

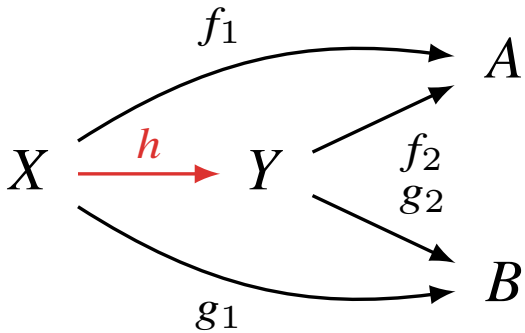
How to understand product? — terminal object of $\mathbf{C}_{AB}$

- Given a category $\mathbf{C}$ and $A, B \in \mathbf{C}$, we can build a new category $\mathbf{C}_{AB}$.
- An object of $\mathbf{C}_{AB}$ is an object of $\mathbf{C}$ with a pair of maps to A, B.



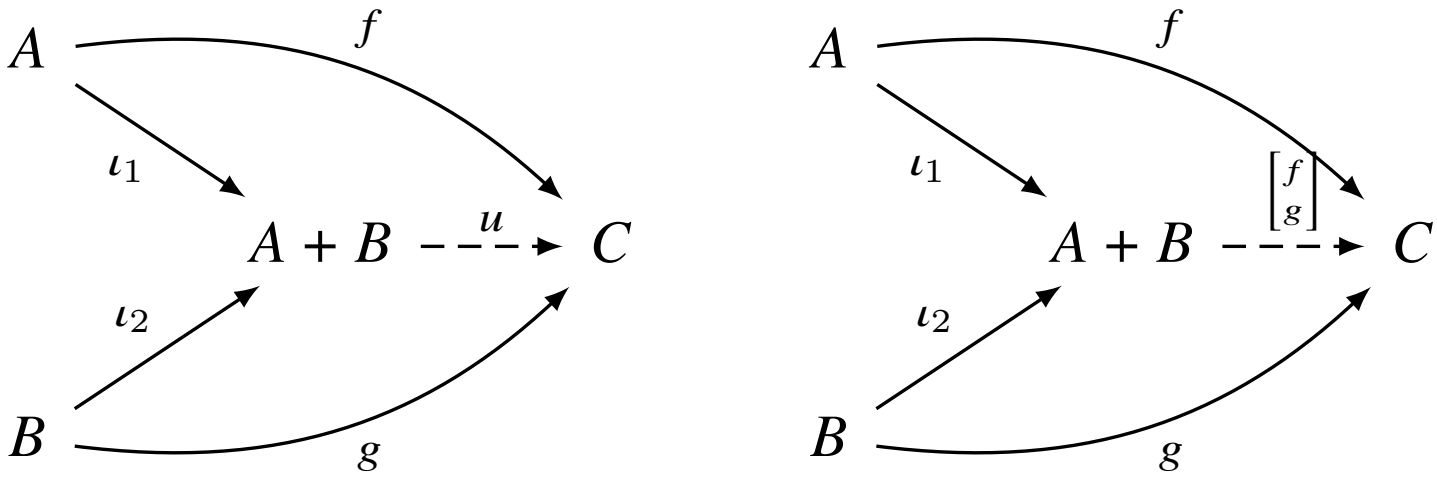
- a morphism from  to  is a map $X \xrightarrow{h} Y$ in $\mathbf{C}$

which 'preserves the structure'.



- then the terminal object of $\mathbf{C}_{AB}$ is the product $A \times B$.

Coproduct



Definition (Coproduct)

A *coproduct* of A and B is an object $A + B$ with a pair of morphisms $A \xrightarrow{\iota_1} A + B \xleftarrow{\iota_2} B$ s.t

$$\forall C \forall f g : A \xrightarrow{f} C \xleftarrow{g} B \; \exists! u : A + B \to C \left[f = u \circ \iota_1 \;\; \& \;\; g = u \circ \iota_2 \right]$$

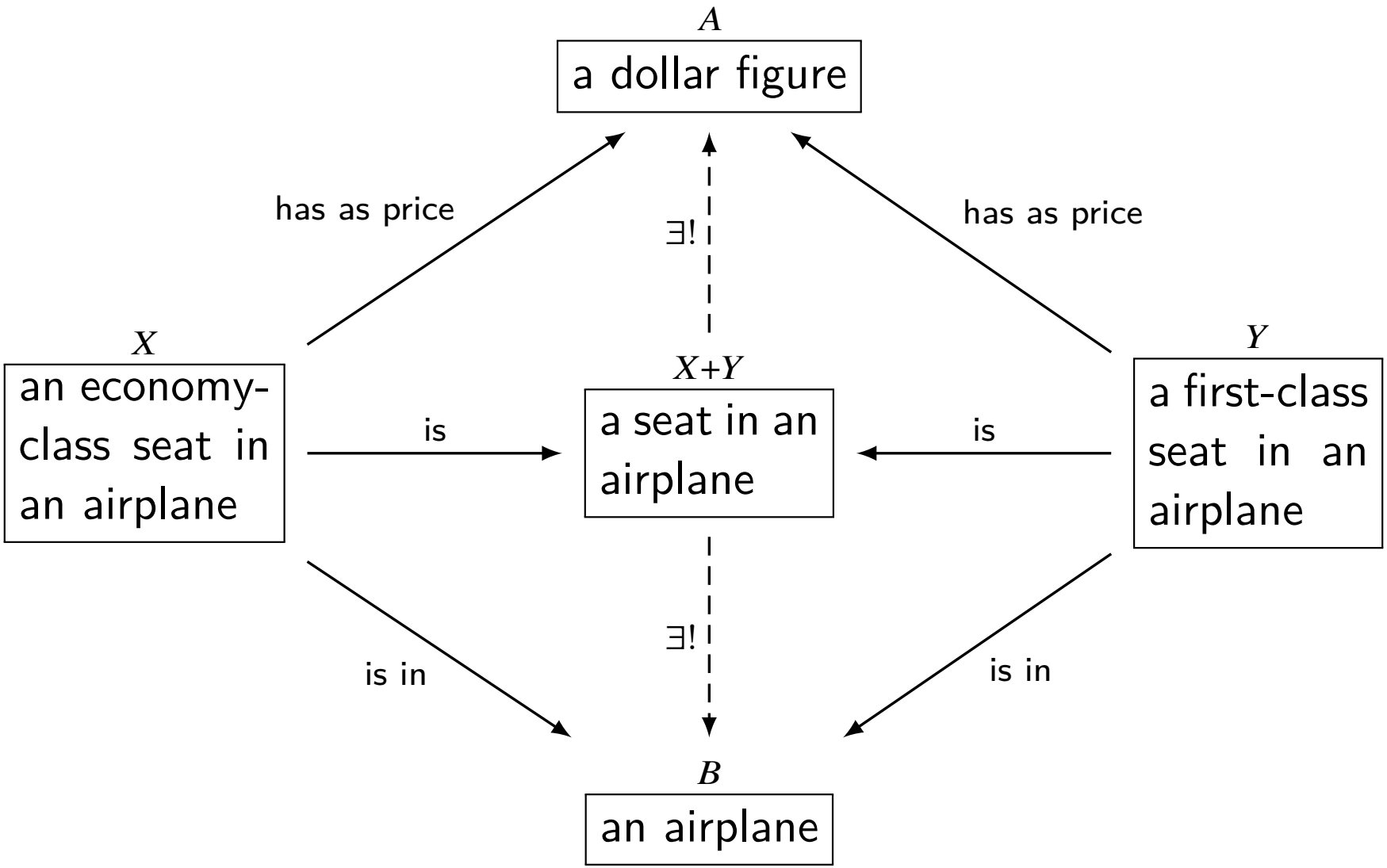
Examples

- In $(\mathbb{R}, \leq)$, coproducts are $\max\{x, y\}$.
- In $(\mathrm{P}(A), \subset)$, coproducts are $x \cup y$.
- In $(\mathbb{N}, |)$, coproducts are $\mathrm{lcm}(x, y)$.
- In logic, coproducts are $A \vee B$.

$$A \vdash A \vee B \qquad B \vdash A \vee B \qquad \frac{A \vdash C \quad B \vdash C}{A \vee B \vdash C}$$

- In $\mathbf{Set}$, disjoint unions are coproducts.
- In $\mathbf{Top}$, topological disjoint unions are coproducts.
- In $\mathbf{FdVect}_{\mathbb{K}}$, the coproduct of V and W is also their direct sum $V \oplus W$.
- In $\mathbf{Mat}$, the oproducts are the sum of dimensions.

Example — Coproduct



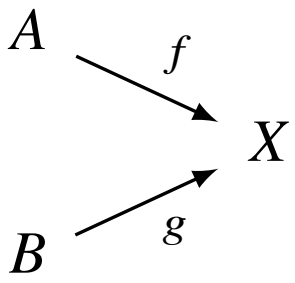
Example — Coproduct

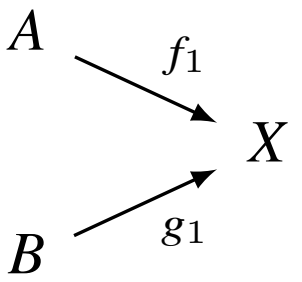
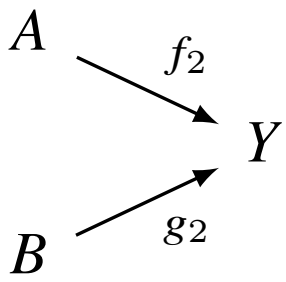
Let $(A, \leq)$ and $(B, \leq)$ be posets. we can define their coproduct $(A + B, \leq)$, where $A + B$ is the disjoint union, and $\leq\colon (A + B) \times (A + B) \longrightarrow \mathbf{2}$ is given by

$$((0, a), (0, b)) \mapsto a \leq_A b$$
$$((1, a), (1, b)) \mapsto a \leq_B b$$
$$((0, -), (1, -)) \mapsto \top$$
$$((1, -), (0, -)) \mapsto \bot$$

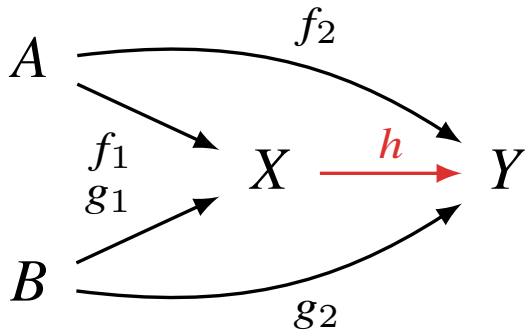
How to understand coproduct? — initial object of $\mathbf{C}_{AB}$

- Given a category $\mathbf{C}$ and $A, B \in \mathbf{C}$, we can build a new category $\mathbf{C}_{AB}$.
- An object of $\mathbf{C}_{AB}$ is an object of $\mathbf{C}$ with a pair of maps from A, B.



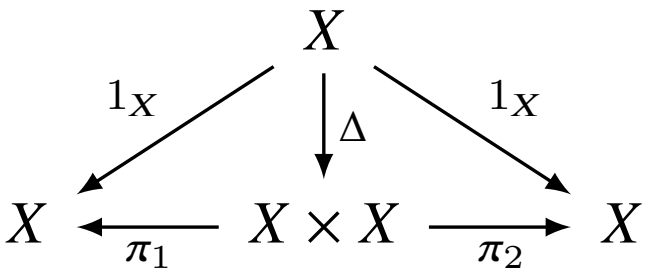
- a morphism from  to  is a map $X \xrightarrow{h} Y$ in $\mathbf{C}$

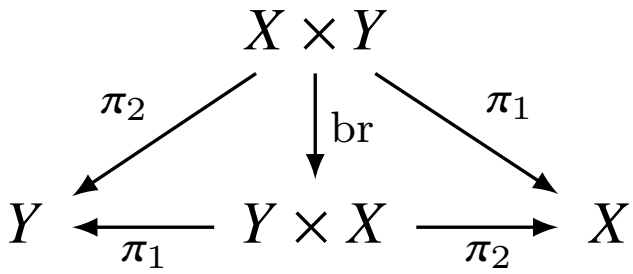
which 'preserves the structure'.



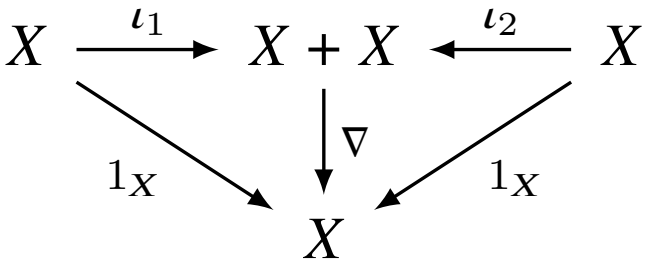
- then the initial object of $\mathbf{C}_{AB}$ is the coproduct $A + B$.

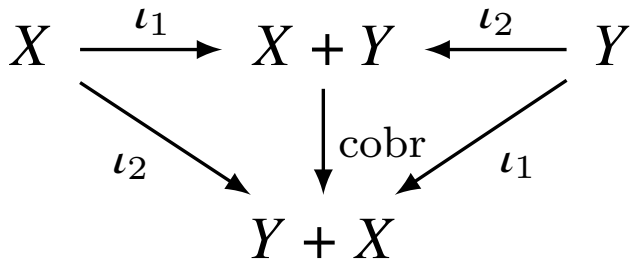
Diagonal & Braid Morphism



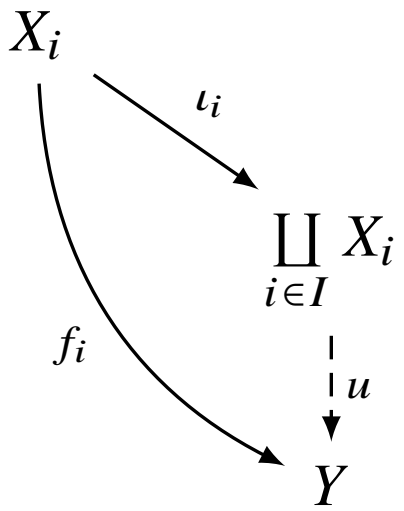


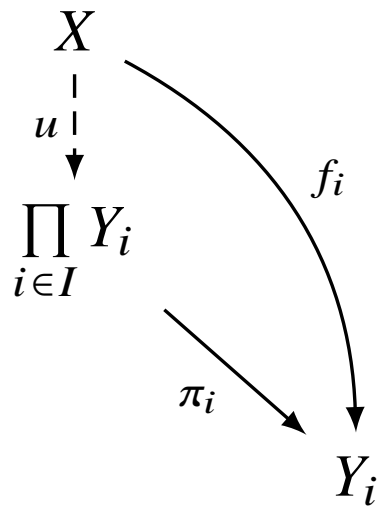
- $\Delta := \langle 1_X, 1_X \rangle$ is called the diagonal morphism.
- $\mathrm{br} := \langle \pi_2^{XY}, \pi_1^{XY} \rangle : X \times Y \to Y \times X$ is called the braid morphism.
- There is a codiagonal morphism ∇, and a cobraid morphism cobr for the coproduct.





Coproduct & Product



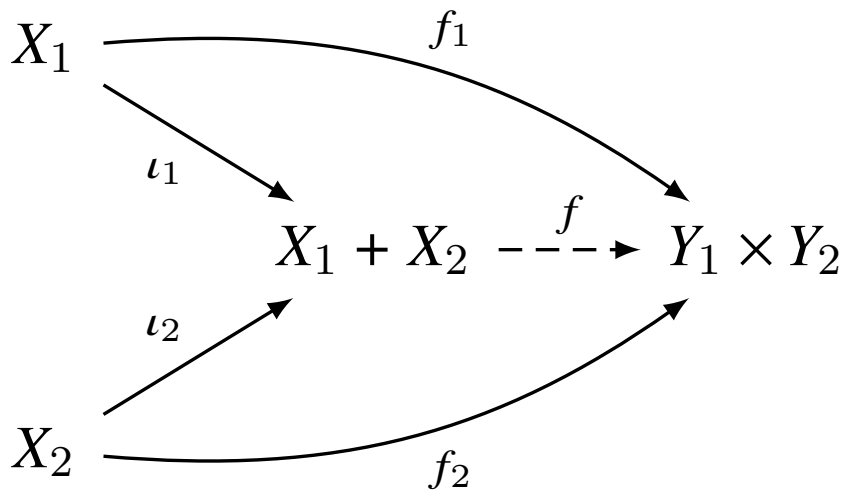


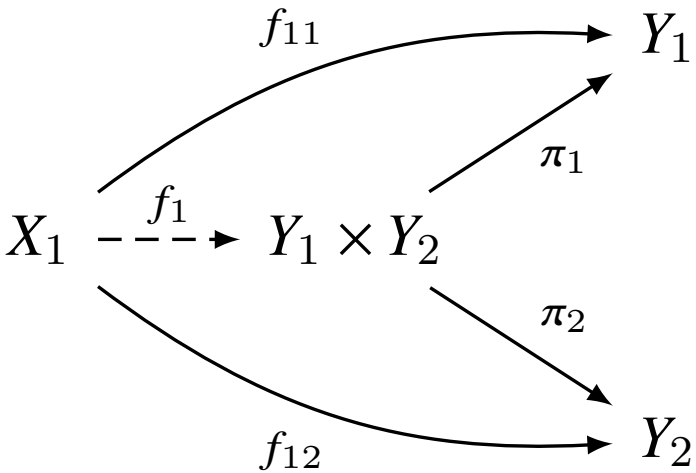
$$\mathrm{Hom}\left(\coprod_{i\in I} X_i, Y\right) \cong \prod_{i\in I} \mathrm{Hom}(X_i, Y) \qquad u \mapsto (u\iota_i)_{i\in I}$$

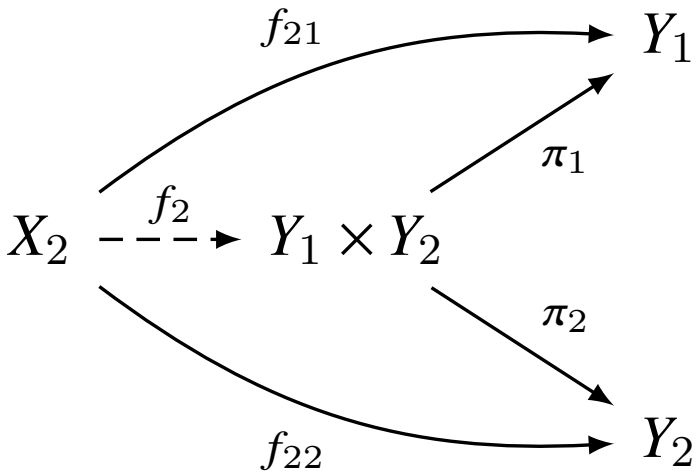
$$\mathrm{Hom}\left(X, \prod_{i\in I} Y_i\right) \cong \prod_{i\in I} \mathrm{Hom}(X, Y_i) \qquad u \mapsto (\pi_i u)_{i\in I}$$

Remark: A poset category having finite products and coproducts is a bounded lattice; if it has arbitrary products and coproducts, it is a complete and bounded lattice.

What are the morphisms $\coprod_{i=1}^{m} X_i \to \prod_{j=1}^{n} Y_j$?







$$f_1 = \langle f_{11}, f_{12} \rangle \quad f_2 = \langle f_{21}, f_{22} \rangle \quad f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$$

What are the morphisms $\coprod_{i=1}^{m} X_i \to \prod_{j=1}^{n} Y_j$?

$$\operatorname{Hom}\left(\coprod_{i=1}^{m} X_i, \prod_{j=1}^{n} Y_j\right) \cong \prod_{i=1}^{m} \operatorname{Hom}\left(X_i, \prod_{j=1}^{n} Y_j\right) \cong \prod_{i=1}^{m} \prod_{j=1}^{n} \operatorname{Hom}(X_i, Y_j)$$

Theorem

The morphisms $\coprod_{i=1}^{m} X_i \to \prod_{j=1}^{n} Y_j$ are the "matrices"

$$M = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ & & \ddots & \\ f_{m1} & f_{m2} & \cdots & f_{mn} \end{bmatrix}$$

where $\pi_j M \iota_i = f_{ij}$.

Matrix Multiplication

- Assume there is a zero morphism 0_{AB} from A to B.
- Then we have the 'identity matrix':

$$A + B \xrightarrow{\begin{bmatrix} 1_A & 0_{AB} \\ 0_{BA} & 1_B \end{bmatrix}} A \times B$$

- A category with zero morphisms in which every 'identity matrix' is an isomorphism is called a *linear category*.
- In a linear category, we can multiply any matrix $X_1 + X_2 \xrightarrow{f} A \times B$ and $A + B \xrightarrow{g} Y_1 \times Y_2$ as

$$\begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \cdot \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} := \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \circ \begin{bmatrix} 1_A & 0_{AB} \\ 0_{BA} & 1_B \end{bmatrix}^{-1} \circ \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$$

- This multiplication is another matrix $X_1 + X_2 \to Y_1 \times Y_2$.

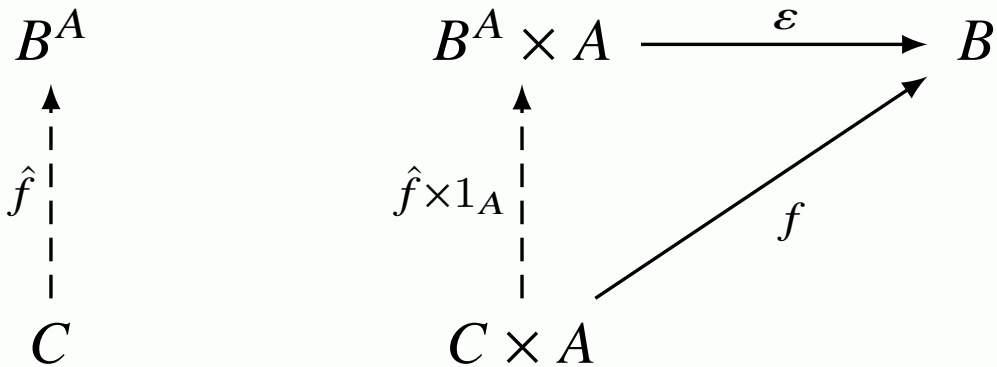
$$X_1 + X_2 \xrightarrow{f} A \times B \xrightarrow{\begin{bmatrix} 1_A & 0_{AB} \\ 0_{BA} & 1_B \end{bmatrix}^{-1}} A + B \xrightarrow{g} Y_1 \times Y_2$$

Exponential

Definition (Exponential)

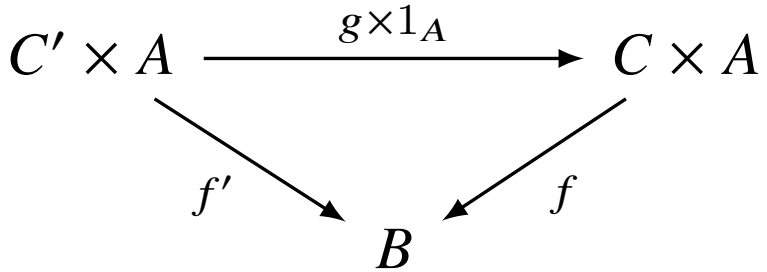
An *exponential* of objects A, B is an object B^A with a morphism $\varepsilon : B^A \times A \to B$ s.t.

$$\forall C \forall f : C \times A \to B \; \exists! \; \hat{f} : C \to B^A \left[\varepsilon \circ (\hat{f} \times 1_A) = f \right]$$



How to understand exponential? — terminal object of $\mathbf{C}/_{A\to B}$

- Given a category $\mathbf{C}$ and $A, B \in \mathbf{C}$, we can build a new category $\mathbf{C}/_{A\to B}$.
- An object of $\mathbf{C}/_{A\to B}$ is an object $C \in \mathbf{C}$ with a morphism $C \times A \to B$.
- A morphism from $C' \times A \xrightarrow{f'} B$ to $C \times A \xrightarrow{f} B$ is a $\mathbf{C}$-morphism $C' \xrightarrow{g} C$ s.t.



- The terminal object of $\mathbf{C}/_{A\to B}$ is the exponential B^A.

Example

- B: particles of a continuous body of matter, e.g. a cloud.
- E: ordinary space.
- T: time interval.
- E^B: positions of the body in space.
- E^T: paths in space.
- The motion of B in E during T:

$$\frac{\frac{T \to E^B}{T \times B \to E}}{B \to E^T}$$

1. To tell how *the distance of the particles from the earth* changes with time, we calculate the composition $T \times B \longrightarrow E \longrightarrow \mathbb{R}$
2. To obtain *the motion of the center of mass*, we calculate the composition $T \longrightarrow E^B \xrightarrow{\text{center of mass}} E$
3. To obtain *the velocity field* over the body at any instant $T \to V^B$, we first calculate the composition $B \longrightarrow E^T \xrightarrow{\text{vilocity operator}} V^T$

Cartesian Closed Category ($\mathbf{CCC}$)

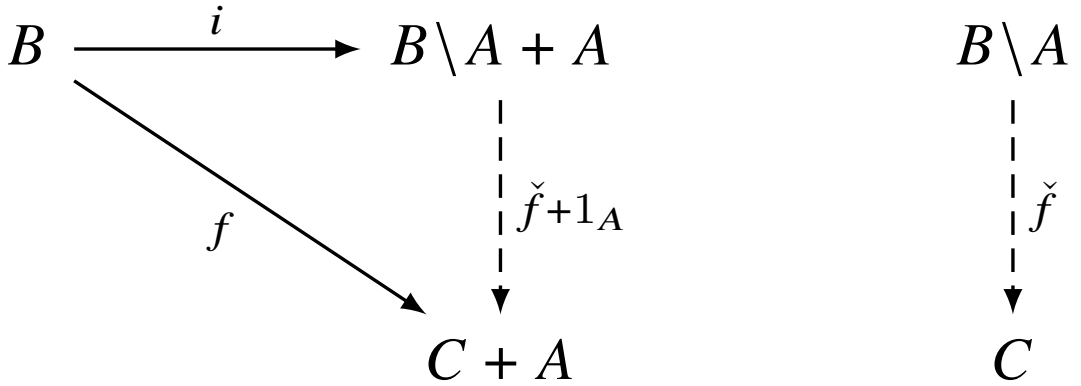
- **Cartesian Closed Category** ($\mathbf{CCC}$) is a category with a terminal object, all products and all exponentials.
- **Bicartesian Closed Category** ($\mathbf{BCCC}$) is a $\mathbf{CCC}$ with an initial object and all coproducts, with products distributing over coproducts.

Remark

- $\mathbf{Set}$ *is a* $\mathbf{CCC}$*.*
- $\mathbf{Grp}$ *has products and a terminal object, but it does not admit exponential.*

Subtraction

- The "dual" of exponential is subtraction.
- The subtraction of A from B, is an object $B \setminus A$, together with an arrow $i : B \to B \setminus A + A$ such that for each C and each $f : B \to C + A$, there is a unique $\check{f} : B \setminus A \to C$ such that $(\check{f} + 1_A) \circ i = f$.



- When $\mathbf{C}$ is a deductive system,

$$\frac{B \setminus A \vdash C}{B \vdash C \vee A}$$

- Whenever a category $\mathbf{C}$ has subtractions, products and a terminal object 1, we can define the complement of A by $\overline{A} = 1 \setminus A$, and the boundary $\partial A = A \wedge \overline{A}$.

Curry-Howard-Lambek Isomorphism

- **Objects** types/formulas $\top \mid A \times B \mid A + B \mid A \to B$
- **Morphisms** terms/proofs

$$1_A \mid ! \mid g \circ f \mid \langle f, g \rangle \mid \pi_1 \mid \pi_2 \mid \begin{bmatrix} f \\ g \end{bmatrix} \mid \iota_1 \mid \iota_2 \mid \hat{f} \mid \varepsilon$$

$$f : A \to B \iff A \vdash B$$

$$\frac{}{1_A : A \to A} \qquad \frac{}{! : A \to \top} \qquad \frac{f : A \to B \quad g : B \to C}{g \circ f : A \to C}$$

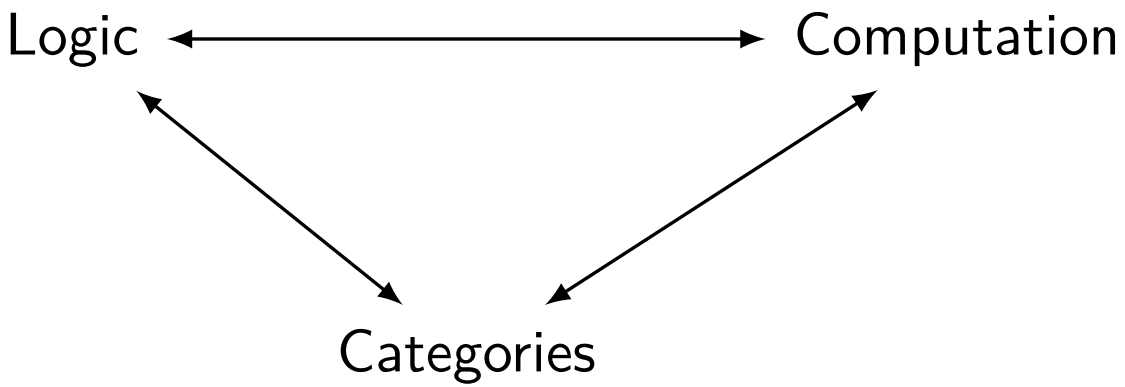
$$\frac{f : C \to A \quad g : C \to B}{\langle f, g \rangle : C \to A \times B} \qquad \frac{f : C \to A \times B}{\pi_1 \circ f : C \to A} \qquad \frac{f : C \to A \times B}{\pi_2 \circ f : C \to B}$$

$$\frac{f : A \to C \quad g : B \to C}{\begin{bmatrix} f \\ g \end{bmatrix} : A + B \to C} \qquad \frac{f : C \to A}{\iota_1 \circ f : C \to A + B} \qquad \frac{f : C \to B}{\iota_2 \circ f : C \to A + B}$$

$$\frac{f : C \times A \to B}{\hat{f} : C \to [A \to B]} \qquad \frac{f : C \to [A \to B]}{\varepsilon \circ (f \times 1_A) : C \times A \to B}$$

Curry-Howard-Lambek Isomorphism

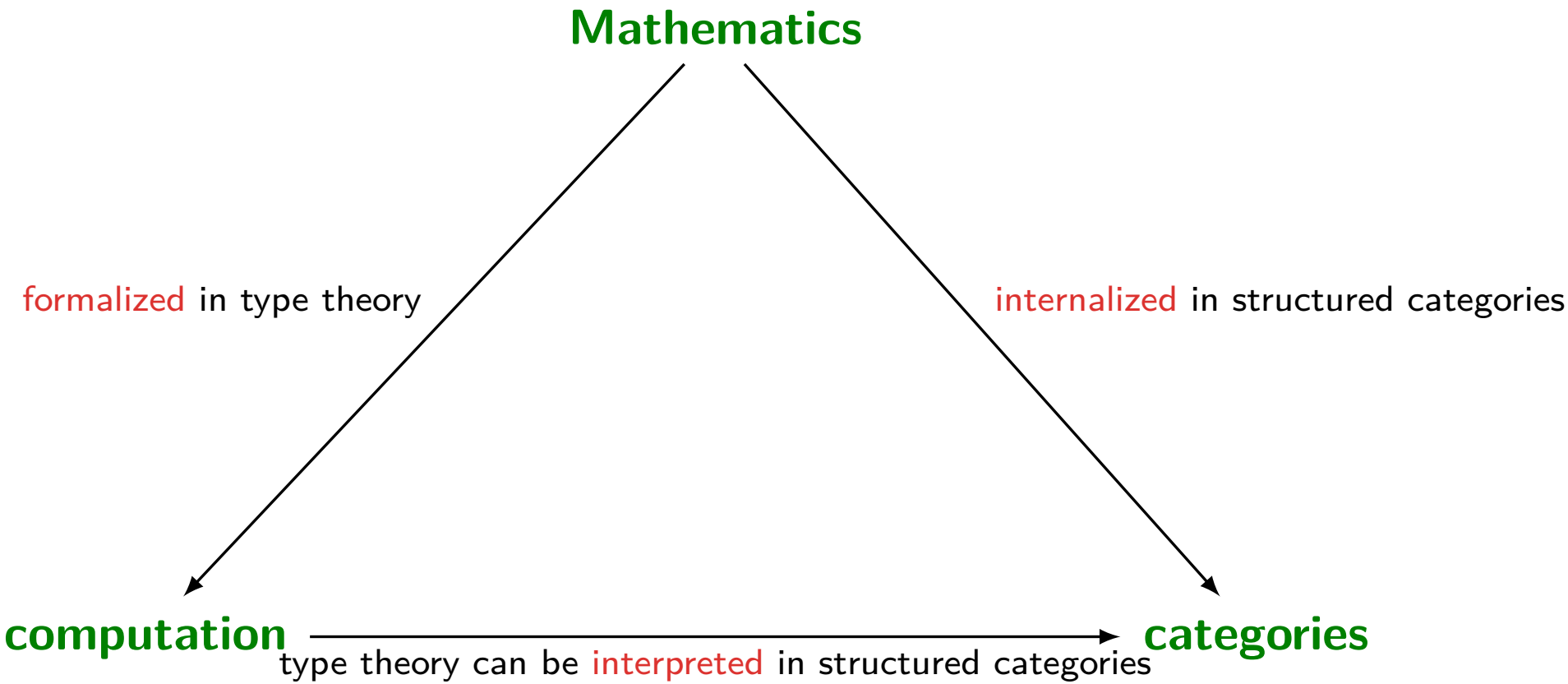
Logic	**Type Theory**	**Category Theory**
Formula	Type	Object
Proof	Term/Program	Morphism
false $\bot$	empty type 0	initial object 0
true $\top$	unit type 1	terminal object 1
conjunction $\wedge$	product type $\times$	product $\times$
disjunction $\vee$	coproduct type $+$	coproduct $+$
implication $\longrightarrow$	function type $\longrightarrow$	exponential B^A
cut-elimination	β-reduction	composition $\circ$
modus ponens	application app	evaluation ε



A mathematician is a person who can find analogies between theorems; a better mathematician is one who can see analogies between proofs and the best mathematician can notice analogies between theories. One can imagine that the ultimate mathematician is one who can see analogies between analogies.

— *Stefan Banach*

- Category theory is not a theory of everything.
- It is a theory of theories of anything.
- It allows one to compare different models, thus carrying knowledge from one domain to another, as long as one can construct the appropriate "analogy", i.e., functor.


Mathematics
formalized in type theory
internalized in structured categories
computation
categories
type theory can be interpreted in structured categories

Category	**Physics**	**Topology**	**Logic**	**Computation**
Object	Hilbert space	Manifold	Proposition	Data type
Morphism	Operator	Cobordism	Proof	Program
Tensor product of objects	Hilbert space of joint system	Disjoint union of manifolds	Conjunction of propositions	Product of data types
Tensor product of morphisms	Parallel processes	Disjoint union of cobordisms	Proofs carried out in parallel	Programs executing in parallel
Internal Hom	Hilbert space of "anti-X and Y"	Disjoint union of orientation-reversed X and Y	Conditional proposition	Function type

Table: Physics, Topology, Logic and Computation

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

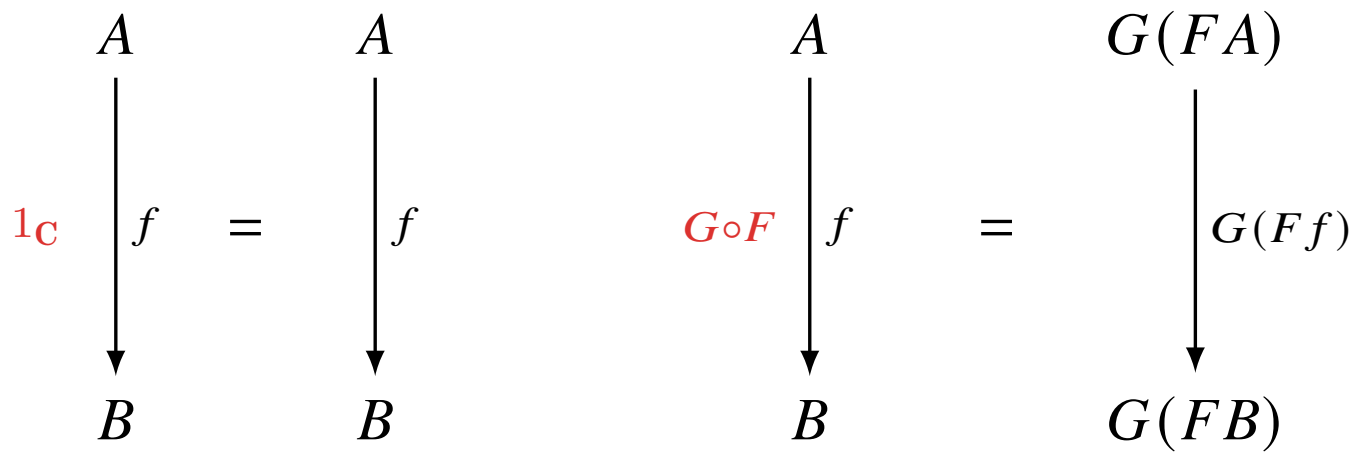
Functor

Definition (Functor)

A functor $F : \mathbf{C} \to \mathbf{D}$ is a mapping of objects to objects and morphisms to morphisms $A \xrightarrow{f} B \;\mapsto\; FA \xrightarrow{Ff} FB$ such that:

- $F(1_A) = 1_{FA}$
- $F(g \circ f) = Fg \circ Ff$

identity functor $1_{\mathbf{C}}$ and composition of functors



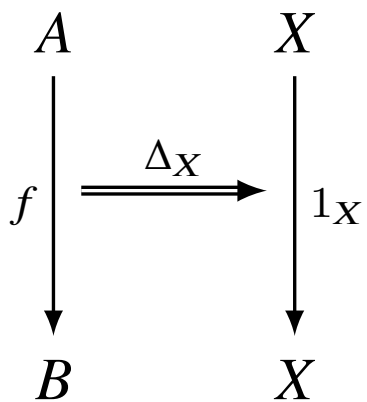
Remark: A mapping on objects and on morphisms that preserves all of the structure of a category: domains, codomains, composition, identities.

Constant Functor

Definition (Constant Functor)

Given two categories $\mathbf{C}, \mathbf{D}$, and a fixed $X \in \mathbf{D}$, the *constant functor* $\Delta_X : \mathbf{C} \to \mathbf{D}$ is defined as

- $\Delta_X(A) = X$ for any object $A \in \mathbf{C}$.
- $\Delta_X(f) = 1_X$ for any morphism f of $\mathbf{C}$.



Definition (Diagonal Functor)

For categories $\mathbf{C}$ and $\mathbf{I}$, there exists a *diagonal functor*

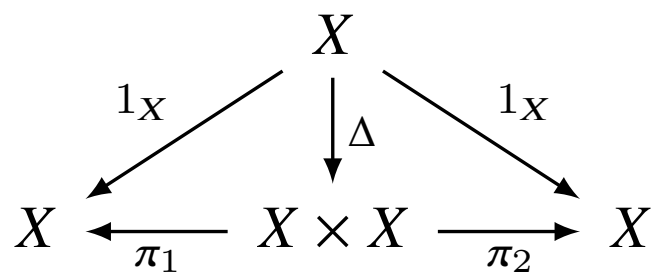
$$\Delta : \mathbf{C} \to \mathbf{C}^{\mathbf{I}}$$

mapping an object $X \in \mathbf{C}$ to the constant diagram Δ_X of shape $\mathbf{I}$ in $\mathbf{C}$ where all objects are copies of X and all morphisms are copies of $1_X : X \to X$.

$$(\Delta_X)i \coloneqq X \text{ for } i \in \mathbf{I}$$
$$(\Delta_X)f \coloneqq 1_X \text{ for } f \in \mathbf{I}$$
$$(\Delta f)i \coloneqq f \text{ for } f \in \mathrm{Hom}_{\mathbf{C}}(X, Y) \text{ and } i \in \mathbf{I}$$

Remark: The (binary) diagonal functor of $\mathbf{C}$ is the functor $\Delta : \mathbf{C} \to \mathbf{C} \times \mathbf{C}$ given by $\Delta X = (X, X)$ and $\Delta f = (f, f)$.

Definition: The diagonal morphism of X in a category $\mathbf{C}$ with product is $\Delta : X \to X \times X$ satisfying



Remark: The diagonal morphisms in $\mathbf{Cat}$ are diagonal functors.

Monoids as Categories

- Each monoid as a category with one object.
- Each monoid homomorphism as a functor between one-object categories.
- Each monoid action as a set-valued functor.
  — Let $\mathbf{M} = (M, e, \cdot)$ be a monoid. A monoid action consists of a set X and a map $F : M \times X \to X$ s.t. $F_e(x) = x$ and $F_m(F_n(x)) = F_{mn}(x)$.
  — How might we relate the notion of monoid actions to the notion of a functor $F : \mathbf{M} \to \mathbf{Set}$?
  — Since $\mathbf{M}$ as a category has only one object, let $X \coloneqq F(\bullet) \in \mathbf{Set}$.

$$\mathrm{Hom}_{\mathbf{M}}(\bullet, \bullet) \to \mathrm{Hom}_{\mathbf{Set}}(F(\bullet), F(\bullet)) \quad \text{i.e.} \quad M \to \mathrm{Hom}_{\mathbf{Set}}(X, X)$$

  By currying, this is the same as $F : M \times X \to X$.
  — The rule that $F_e(x) = x$ becomes the rule that functors preserve identities $F(1_\bullet) = 1_X$.
  — The other rule is equivalent to $F(m \circ n) = Fm \circ Fn$.

Example — Homotopy

- Let $\mathbf{Top}_\bullet$ be the category of *pointed topological spaces*, where
  - objects (X, x_0) are topological spaces with a distinguished *base point*.
  - a morphism $f : (X, x_0) \to (Y, y_0)$ is a continuous map $f : X \to Y$ which preserves the base point $f(x_0) = y_0$.
- If one has a path f and a path g, which begins where f ends, then their *concatenation*:

$$(g \cdot f)t := \begin{cases} f(2t) & 0 \leq t \leq 1/2 \\ g(2t-1) & 1/2 < t \leq 1 \end{cases}$$

- A *homotopy* between the maps $f, g : X \to Y$ is a continuous map $\gamma : X \times [0,1] \to Y$ such that:
  - $\forall x \in X : \gamma(x, 0) = fx$
  - $\forall x \in X : \gamma(x, 1) = gx$
- If there exists a homotopy γ between f and g we say that f and g are *homotopic*. This is an equivalence relation.
- The *fundamental groupoid* $\pi_1(X)$ *of space* X is the category whose objects are points of X. A morphism $x \to y$ is a homotopy class $[f]$ of paths from x to y. Composition is given by concatenation of paths $[g] \circ [f] = [g \cdot f]$.

Example — Fundamental Group

- A *loop in X based at x_0* is a continuous function $f : [0,1] \to X$ such that $f(0) = f(1) = x_0$.
- The *fundamental group of X based at the point x_0* is the group $\pi_1(X, x_0)$ of homotopy classes of loops based at $x_0 \in X$.
  The unit is given by the constant loop at x_0, and the inverse is given by "walking the loop backwards" $f^{-1}(t) \coloneqq f(1-t)$.
- Let $f : (X, x_0) \to (Y, y_0)$ be a base point-preserving continuous function. We can map a loop at x_0 to a loop of y_0

$$[0,1] \xrightarrow{l} X \xrightarrow{f} Y$$

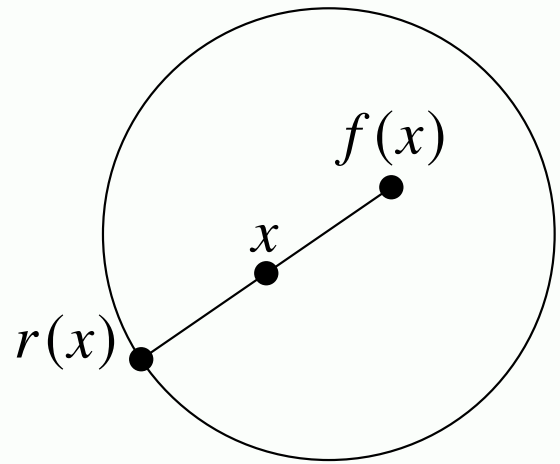
- It induces a map between the equivalence classes, $\pi_1(X, x_0) \to \pi_1(Y, y_0)$. We denote this resulting map $\pi_1(f)$.
- The assignment given by $(X, x_0) \mapsto \pi_1(X, x_0)$ and $f \mapsto \pi_1(f)$ is a functor $\pi_1 : \mathbf{Top}_\bullet \to \mathbf{Grp}$.
- $\pi_1(S^1) \cong (\mathbb{Z}, +)$ where $S^1 = \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\right\}$.
- $\pi_1(S^1 \times S^1) \cong \pi_1(S^1) \times \pi_1(S^1) \cong \mathbb{Z} \times \mathbb{Z}$

Theorem (Brouwer Fixpoint Theorem)

Any continuous endomorphism of a 2-dimensional disk D^2 has a fixpoint.

Proof.

Suppose $f : D^2 \to D^2$ has no fixpoint. Then there is a continuous function $r : D^2 \to S^1$ that carries a point $x \in D^2$ to the intersection of the ray from $f(x)$ to x with the boundary S^1, and $r(x) = x$ when $x \in S^1$. Then the function r is a retraction for the inclusion $i : S^1 \hookrightarrow D^2$.



$$S^1 \xhookrightarrow{i} D^2 \xrightarrow{r} S^1$$

Pick any basepoint on the boundary S^1 and apply the functor π_1 to obtain a composable pair of group homomorphisms:

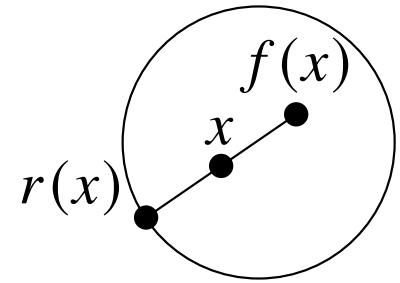
$$\pi_1(S^1) \xrightarrow{\pi_1(i)} \pi_1(D^2) \xrightarrow{\pi_1(r)} \pi_1(S^1)$$

By the functoriality axioms, we have

$$\pi_1(r) \circ \pi_1(i) = \pi_1(r \circ i) = \pi_1(1_{S^1}) = 1_{\pi_1(S^1)}$$

Therefore, $\pi_1(i)$ is monic and hence injective. However, $\pi_1(S^1) \cong (\mathbb{Z}, +)$, $\pi_1(D^2) \cong (\{0\}, +)$. There is no injection $\mathbb{Z} \rightarrowtail \{0\}$. □

- How to understand Brouwer's proof?
- Let $\mathbf{C}$ be a category. The objects are some type of sets (sphere S, ball B, etc.), and the morphisms are continuous maps.
  - there is an object A (whose points are the arrows in B);
  - a morphism $A \xrightarrow{h} B$ (assigning to each arrow its head);
  - a morphism $A \xrightarrow{p} S$ (telling where each arrow points).



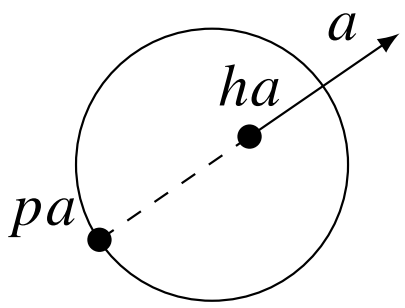
- Assumption1: $\forall T \in \mathbf{C} \forall a : T \to A \forall s : T \to S[h \circ a = i \circ s \to p \circ a = s]$

$$\begin{array}{ccc} T & \xrightarrow{a} & A \\ {\scriptstyle s}\downarrow & \swarrow_{p} & \downarrow{\scriptstyle h} \\ S & \xrightarrow[i]{} & B \end{array}$$

Remark: if an arrow's head is on S, then its head is the point it points.

- Lemma: $\forall \gamma : B \to A[h \circ \gamma \circ i = i \to p \circ \gamma \circ i = 1_S]$
  Proof: Take $T = S$, $s = 1_S$, and $a = \gamma \circ i$.
- Assumption2: For any $T \in \mathbf{C}$, and $T \underset{g}{\overset{f}{\rightrightarrows}} B$, then either
  $\exists t : 1 \to T[f \circ t = g \circ t]$, or $\exists \gamma : T \to A[h \circ \gamma = g]$. "tail ≠ head"
- Theorem: Suppose $B \underset{g}{\overset{f}{\rightrightarrows}} B$ and $g \circ i = i$, then either
  $\exists b : 1 \to B[f \circ b = g \circ b]$, or there is a retraction for $S \xrightarrow{i} B$.

The eye of the storm



- Imagine a fluid moving in a spherical container.
- Each point in our ball is moving $B \xrightarrow{\gamma} A$, and we draw an arrow with 'tail' $A \xrightarrow{h} B$ at that point to represent its velocity.
- The length of the arrow is proportional to the speed of the point, and the arrow points in the direction of travel.
- Could it be that every point is moving with non-zero speed?
- For the map $A \xrightarrow{p} S$, we assign to each arrow its 'place of birth'.
- Assumption1 says that if the moving point is on the sphere, then its 'place of birth' is its current location: $ha = pa$.
- Its corollary $(h \circ \gamma = 1_B \implies p \circ \gamma \circ i = 1_S)$ tells us that if there were a storm with no instantaneous 'eye', there would be a retraction for the inclusion of the sphere into the ball.

Brouwer vs Banach Fixpoint

- Throw a perfect map on the area, and the map contains a picture of the map itself, and that picture has a smaller picture which has a smaller picture ...
- These pictures gradually close in on the one and only fixpoint for the endomap.
- The idea only works for an endomap which shrinks distances, though.
- Brouwer's theorem applies to every continuous endomap of the area.

Theorem (代数基本定理)

一元 n 次复系数多项式 $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ 在复数域中至少有一个根.

Proof.

假设 $f(z)$ 没有根. 则 f 给出了连续映射 $\mathbb{C} \to \mathbb{C} \setminus \{0\}$.
令 $p : \mathbb{C} \setminus \{0\} \to S^1 :: x \mapsto \frac{x}{|x|}$. 则 $p \circ f : \mathbb{C} \to S^1$ 也是连续映射.
令 $S^1_r := \{z \in \mathbb{C} : |z| = r\}$ 表示半径为 r 的圆.

$$S^1_r \xhookrightarrow{i} \mathbb{C} \xrightarrow{f} \mathbb{C} \setminus \{0\} \xrightarrow{p} S^1$$

如果 $z = re^{i\theta}$ 以坐标原点为圆心绕一个半径为 r 的圆, 则 $z^n = r^n e^{in\theta}$ 绕了 n 个半径为 r^n 的圆. 当 $|z| = r$ 很大时, $f(z)$ 与 z^n 的差别很小. $\gamma(z,t) := z^n + t(f(z) - z^n)$ 给出了 z^n 与 $f(z)$ 的同伦, 所以, $f(z)$ 也绕了 n 周. 于是 $p \circ f \circ i$ 把绕着 S^1_r 的一个 loop 映射到了绕着 S^1 的 n 圈的 loop, 把 $\pi_1(S^1_r, x)$ 的生成元映射到了 $\pi_1(S^1, p(f(x)))$ 的生成元的 n 倍.

但是, $\pi_1(S^1_r, x) \longrightarrow \pi_1(\mathbb{C}, x) \longrightarrow \pi_1(S^1, p(f(x)))$ 为 $\mathbb{Z} \to \{0\} \to \mathbb{Z}$, 群同态把一切都映到了 0. 仅当 $n = 0$ 时 $f(z)$ 才可能没有根. □

Example — Derivative

Define the category of *pointed Euclidean spaces* $\mathbf{Euc}_\bullet$ as follows:

- as objects, we take $(\mathbb{R}^n, x)$ with a distinguished point $x \in \mathbb{R}^n$.
- as morphisms $f : (\mathbb{R}^n, x) \to (\mathbb{R}^m, y)$ we take smooth (i.e. differentiable infinitely many times) functions $\mathbb{R}^n \to \mathbb{R}^m$ such that $f(x) = y$.

The derivative is a functor $D : \mathbf{Euc}_\bullet \to \mathbf{Vect}$ defined in the following way:

- on objects, it maps $(\mathbb{R}^n, x)$ to $\mathbb{R}^n$ (now seen as a vector space).
- on morphisms, it maps $f : (\mathbb{R}^n, x) \to (\mathbb{R}^m, y)$ to the derivative $Df|_x$.

The derivative is functorial because:

- the derivative of the identity map $(\mathbb{R}^n, x) \to (\mathbb{R}^n, x)$ is just the identity of $\mathbb{R}^n$ (the identity matrix).
- consider composable maps

$$(\mathbb{R}^n, x) \xrightarrow{f} (\mathbb{R}^m, y) \xrightarrow{g} (\mathbb{R}^p, z)$$

we have that, *by the chain rule*, $D(g \circ f)|_x = Dg|_y \circ Df|_x$, i.e.

$$\frac{\partial (g \circ f)^k}{\partial x^i} = \sum_{j=1}^{m} \frac{\partial g^k}{\partial y^j} \frac{\partial f^j}{\partial x^i} \quad \text{for } i = 1, \ldots, n \text{ and } k = 1, \ldots, p.$$

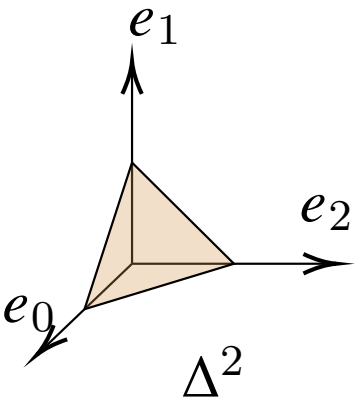
Faddeev's Characterization of Shannon Entropy

- Shannon Entropy:

$$H(p) := -\sum_{i=1}^{n} p_i \log p_i$$

- n-simplex:

$$\Delta^n := \left\{ (p_0, \ldots, p_n) \in \mathbb{R}^{n+1} : 0 \leq p_i \leq 1 \text{ and } \sum_{i=0}^{n} p_i = 1 \right\}$$



Theorem (Faddeev's Characterization of Shannon Entropy)

Let $\{F : \Delta^n \to \mathbb{R}\}_{n \geq 0}$ be a sequence of functions. The following are equivalent:

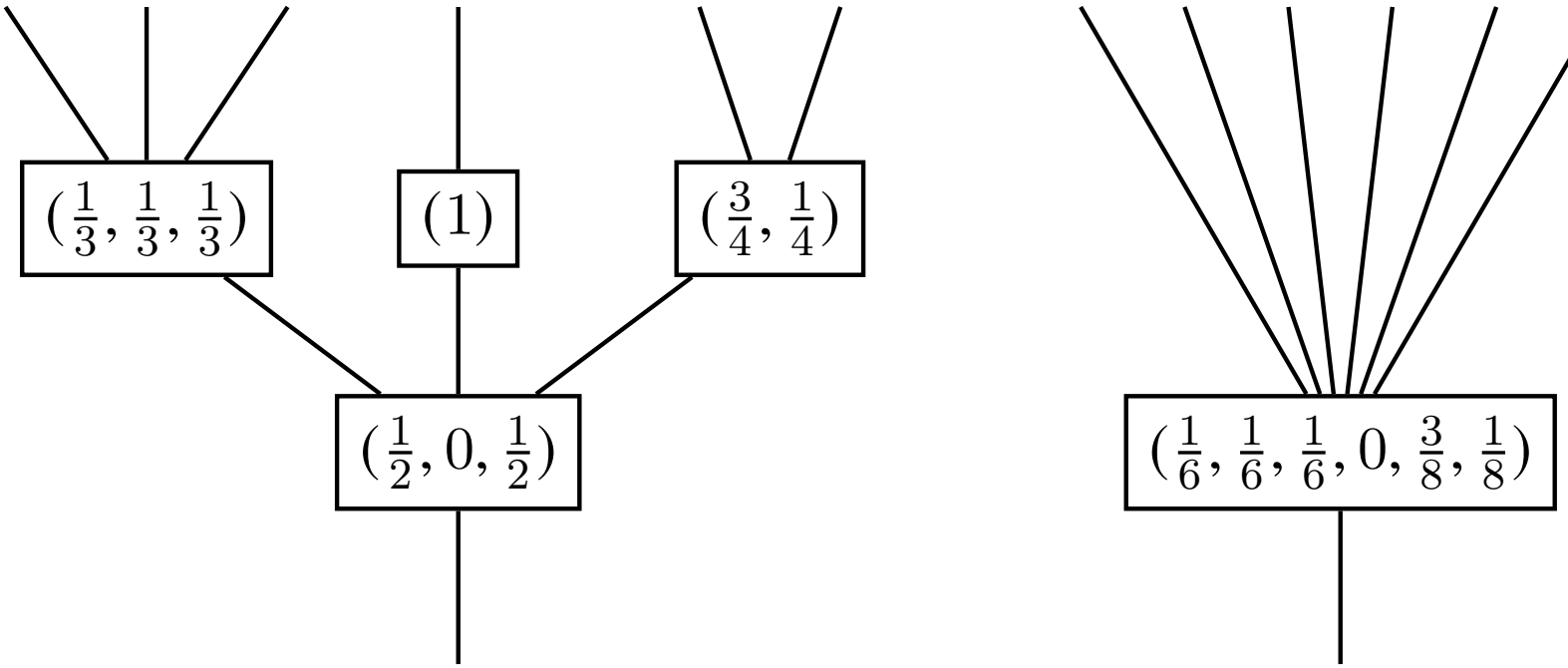
1. *the functions F are continuous and satisfy*

$$F(p \circ (q^0, \ldots, q^n)) = F(p) + \sum_{i=0}^{n} p_i F(q^i)$$

   where $n \geq 0$ and $p \in \Delta^n$ and $q^i \in \Delta^{k_i}$ with $k_0, \ldots, k_n \geq 0$.

2. *$F = cH$ for some $c \in \mathbb{R}$.*

Composition of Probability Distributions and Chain Rule



$$H(p \circ (q^0, \ldots, q^n)) = H(p) + \sum_{i=0}^{n} p_i H(q^i)$$

Remark

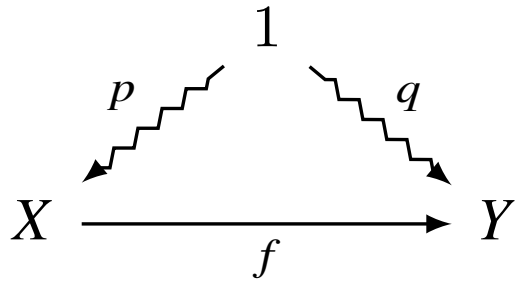
Entropy $H(pq) = H(p) + pH(q)$

Calculus $\mathrm{d}(pq) = \mathrm{d}(p)q + p\mathrm{d}(q)$

FinProb Category

- Given finite sets with probability distributions (X, p) and (Y, q), a measure-preserving map from the first to the second is a function $f : (X, p) \to (Y, q)$ s.t.

$$q_y = \sum_{x \in f^{-1}(y)} p_x$$



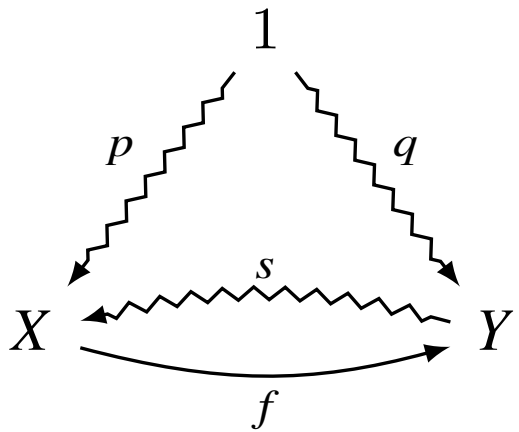
- The composite of measure-preserving maps is measure-preserving.
- So, we get a category **FinProb** with
  1. finite sets equipped with probability distributions as objects.
  2. measure-preserving maps as morphisms.
- The entropy loss of a measure-preserving map $f : (X, p) \to (Y, q)$ defined by $\mathrm{Loss}(f) := H(p) - H(q)$ is a functor from **FinProb** to some category $[0, \infty)$ with one object $\bullet$, and nonnegtive real numbers c as morphisms $c : \bullet \to \bullet$, and addition as composition.

$$\mathrm{Loss}(g \circ f) = \mathrm{Loss}(f) + \mathrm{Loss}(g)$$

Relative Entropy

The expected amount of information you gain when you thought the right probability distribution was q and you discover it's really p.

$$D_{\mathrm{KL}}(p\|q) := \sum_{x\in X} p(x)\log\frac{p(x)}{q(x)}$$



$$f \circ p = q$$
$$f \circ s = 1_Y$$

- Think of f as a 'measurement process' and s as a 'hypothesis' about the state in X given the measurement in Y.
- $D_{\mathrm{KL}}(p\|s\circ q)$: how much information we gain when we learn the 'true' probability distribution p on the states of the measured system, given our 'guess' $s\circ q$ based on the measurements q and our hypothesis s.

Clustering Algorithms as Functors

- A clustering algorithm C is a map that assigns each finite metric space (X, d_X) one of its partitions.
- The following properties are desirable:
  1. **Scale invariance:** $C(X, d_X) = C(X, \lambda \cdot d_X)$ for all $\lambda > 0$.
  2. **Surjectivity:** for any partition P of X there exists a metric d_X on X with $C(X, d_X) = P$.
  3. **Consistency:** upon reducing distances between points in the same cluster, and increasing distances between points in different clusters, the result of applying C to the new metric does not change $C(X, d_X) = C(X, d'_X)$.

Theorem (Kleinberg2002)

There exists no clustering algorithm that simultaneously satisfies scale invariance, surjectivity and consistency.

Remark: There exists no nontrivial onto functors from $\mathbf{FinMet}^{\leq}$ to $\mathbf{Clust}$.

Clustering Algorithms as Functors

Remark: Kleinberg's theorem in categorical language: there exists no nontrivial onto functors from $\mathbf{FinMet}^{\leq}$ to $\mathbf{Clust}$.

- An objects of $\mathbf{FinMet}^{\leq}$ is a finite metric space, and a morphism is a map $f : X \to Y$ s.t. $d_Y(f(x), f(x')) \leq d_X(x, x')$.
- An object of $\mathbf{Clust}$ is a pair (X, P_X) consisting of a finite set X and a partition P_X, and a morphism is a map $f : X \to Y$ s.t. $f^{-1}(P_Y)$ is a refinement of P_X.

$$\begin{array}{ccc} (X, d_X) & \xrightarrow{f} & (Y, d_Y) \\ {\scriptstyle C}\downarrow & & \downarrow{\scriptstyle C} \\ (X, P_X) & \xrightarrow[Cf]{} & (Y, P_Y) \end{array}$$

Remark

- Many machine learning tasks can be formulated as functors between categories.
- For instance, machine translation is a functor between two language categories, which not only maps words to words but also preserves the semantic relations between words.
- Image captioning is a functor from image to language categories, that transcribes objects in the image as well as their interrelations.

Full and Faithful Functors

Definition

A functor $F : \mathbf{C} \to \mathbf{D}$ is:

- *faithful* iff for $A, B \in \mathbf{C}$ each $F : \mathbf{C}(A, B) \to \mathbf{D}(FA, FB)$ is injective;
- *full* iff for $A, B \in \mathbf{C}$ each $F : \mathbf{C}(A, B) \to \mathbf{D}(FA, FB)$ is surjective;
- an *embedding* iff F is full, faithful, and injective on objects;
- *essentially surjective* iff for every $B \in \mathbf{D}$ there is $A \in \mathbf{C}$ s.t. $F(A) \cong B$;
- an *isomorphism* iff there is a functor $G : \mathbf{D} \to \mathbf{C}$ such that $G \circ F = 1_{\mathbf{C}}$ and $F \circ G = 1_{\mathbf{D}}$.

Theorem (Functors Preserve Isomorphism)

If $A \cong B$ are isomorphic objects in $\mathbf{C}$ and $F : \mathbf{C} \to \mathbf{D}$ is a functor then $FA \cong FB$.

Definition (Contravariant Functor)

A *contravariant functor* $F : \mathbf{C} \to \mathbf{D}$ between categories $\mathbf{C}$ and $\mathbf{D}$ is a functor $F : \mathbf{C}^{\mathrm{op}} \to \mathbf{D}$.

Definition

- A functor $F : \mathbf{C} \to \mathbf{D}$ preserves monomorphisms when, for every f in $\mathbf{C}$, if f is mono then so is Ff.
- A functor $F : \mathbf{C} \to \mathbf{D}$ reflects monomorphisms when, for every f in $\mathbf{C}$, if Ff is mono then so is f.

Proposition

- *Every functor preserves isomorphisms.*
- *A faithful functor reflects monomorphisms.*
- *A full and faithful functor reflects isomorphisms.*
- *If $F : \mathbf{C} \to \mathbf{D}$ is full and faithful. Then F is "essentially injective on objects": $\forall A, A' \in \mathbf{C} : FA \cong FA' \to A \cong A'$.*

Remark: *"essentially injective on objects" is just a weaker version of "reflects isomorphisms".*

Example

Both $(\mathbb{Z}, +, 0)$ and $(\mathbb{N}, +, 0)$ are objects in $\mathbf{Mon}$. Consider

$$i : (\mathbb{N}, +, 0) \hookrightarrow (\mathbb{Z}, +, 0)$$

Being an injection, i is mono. But it is also epi, although not surjective.

Proof.

Let $f, g : (\mathbb{Z}, +, 0) \to (M, \cdot, e)$ s.t. $fi = gi$, and let $z \in \mathbb{Z}$.
If $z \geq 0$, then $f(z) = f(i(z)) = g(i(z)) = g(z)$.
If $z < 0$, then
$f(z) = f(z) \cdot e = f(z) \cdot g(0) = f(z) \cdot g(-z + z) = f(z) \cdot g(-z) \cdot g(z) =$
$f(z) \cdot g(i(-z)) \cdot g(z) = f(z) \cdot f(i(-z)) \cdot g(z) = f(0) \cdot g(z) = g(z)$ □

- In $\mathbf{Set}$, isomorphisms are exactly the bijective functions.
- In concrete categories, isomorphisms are the usual invertible homomorphisms.
- In $\mathbf{Mon}$, a morphism which is mono and epi does not need to be iso.

Subcategory & Inclusion Functor

Definition (Subcategory)

We say that $\mathbf{C}'$ is a subcategory of $\mathbf{C}$ if:

1. $\mathrm{ob}(\mathbf{C}') \subset \mathrm{ob}(\mathbf{C})$;
2. $\mathbf{C}'(A, B) \subset \mathbf{C}(A, B)$ for $A, B \in \mathrm{ob}(\mathbf{C}')$;
3. the composition of morphisms in $\mathbf{C}'$ is induced by the composition of morphisms in $\mathbf{C}$;
4. the identity morphisms in $\mathbf{C}'$ are identity morphisms in $\mathbf{C}$.

Definition (Inclusion Functor)

A functor $F : \mathbf{C} \hookrightarrow \mathbf{D}$ is an *inclusion functor*, iff, for all $A \in \mathbf{C}$, $FA = A$, and, for all $f \in \mathbf{C}(A, B)$, $Ff = f$.

Remark: The category $\mathbf{C}$ is a subcategory of $\mathbf{D}$ iff there is an inclusion functor $i : \mathbf{C} \hookrightarrow \mathbf{D}$.

- The subcategory $\mathbf{C}' \subset \mathbf{C}$ is called *wide* iff all objects of $\mathbf{C}$ are also objects of $\mathbf{C}'$.
- $\mathbf{C}'$ is called *full* iff for $A, B \in \mathrm{ob}(\mathbf{C}') : \mathbf{C}'(A, B) = \mathbf{C}(A, B)$.
- The *core* of a category $\mathbf{C}$ is the wide subcategory of $\mathbf{C}$ whose morphisms are the isormphisms of $\mathbf{C}$.
  — **Example:** $\mathrm{Core}(\mathbf{Set}) = \mathbf{Set}_{\mathrm{bij}}$.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

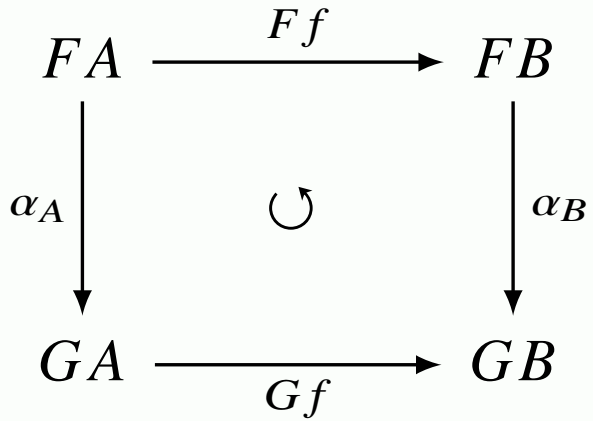
Answers to the Exercises

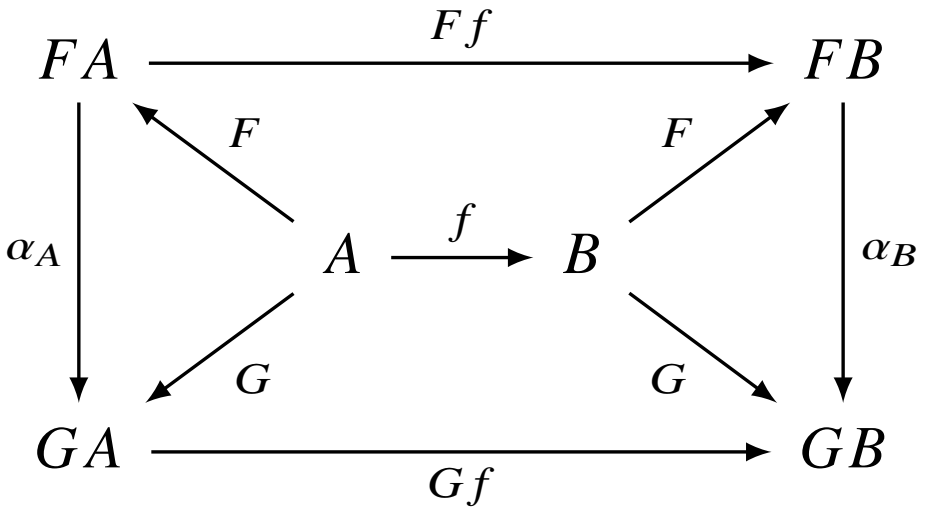
References2039

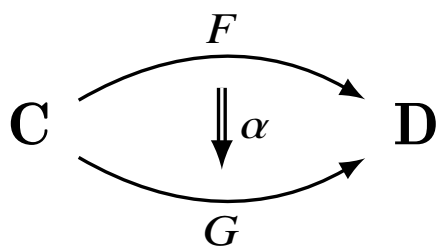
Natural Transformation

Definition (Natural Transformation)

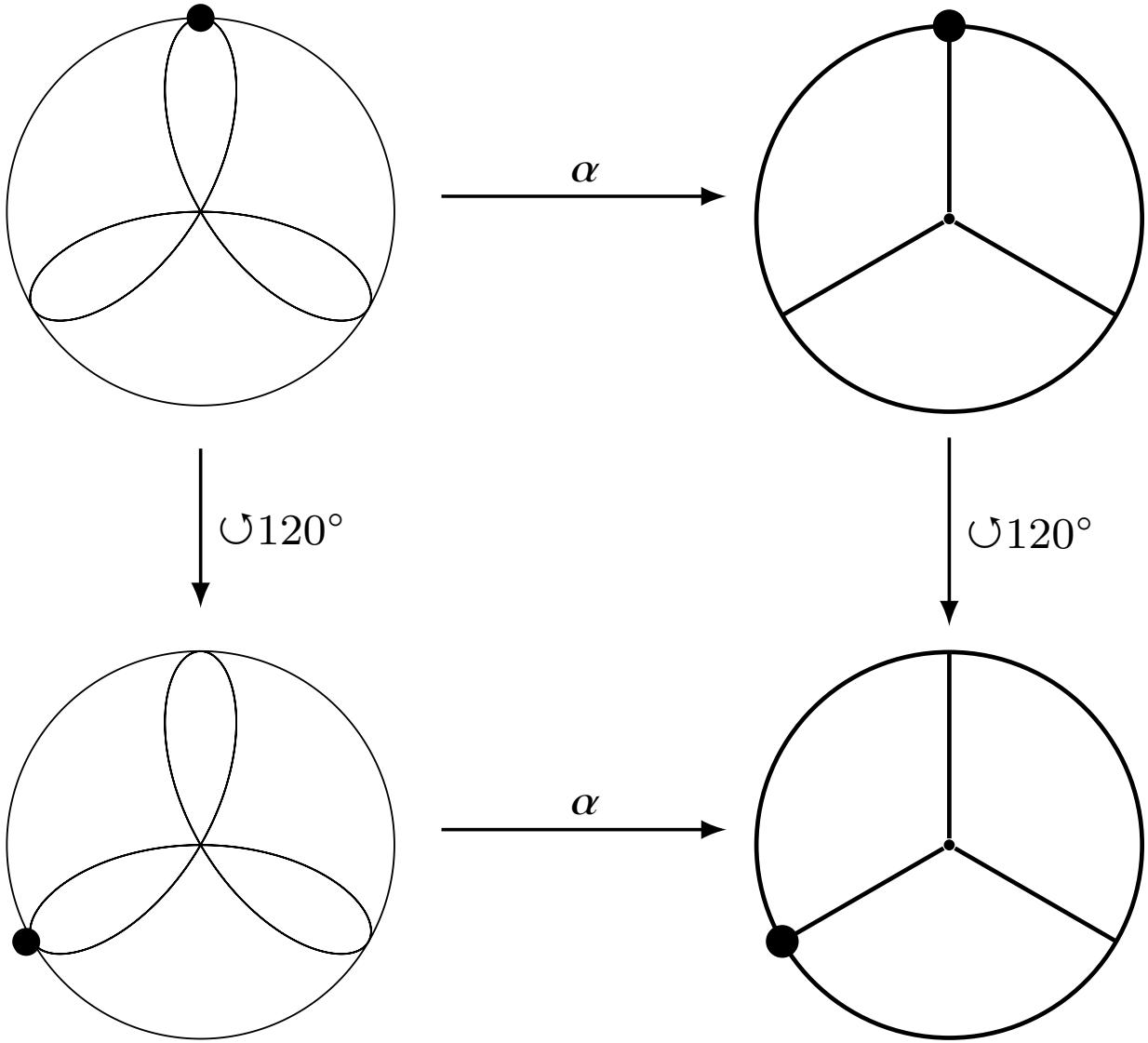
Given categories and functors $F, G : \mathbf{C} \to \mathbf{D}$, a natural transformation $\alpha : F \to G$ is a family of $\mathbf{D}$-morphisms $\{\alpha_A : FA \to GA\}_{A \in \mathbf{C}}$, such that for all $\mathbf{C}$-morphisms $f : A \to B$, the diagram commutes:







A natural transformation is a mapping between functors preserving specified actions, symmetries, or other structures



Non-Natural Transformations

Example (Fundamental Group of a Torus)

- The homotopy groups of a product space are the product of the homotopy groups of the components,

$$\pi_n((X, x_0) \times (Y, y_0)) \cong \pi_n(X, x_0) \times \pi_n(Y, y_0)$$

  with the isomorphism given by projection onto the two factors.
- However, the torus T is abstractly a product of two circles, and thus has fundamental group isomorphic to $\mathbb{Z}^2$:

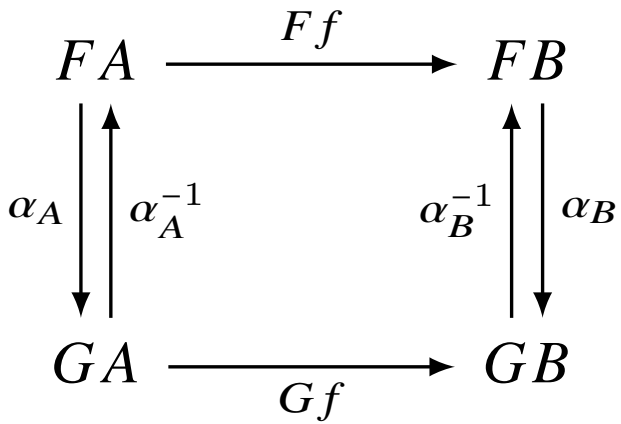
$$\pi_1(T, t_0) \approx \pi_1(S^1, x_0) \times \pi_1(S^1, y_0) \cong \mathbb{Z} \times \mathbb{Z}$$

- This abstract isomorphism with a product is not natural, as some isomorphisms of T do not preserve the product.
- The torus as a space that happens to be a product (in the category of spaces and continuous maps) is different from the torus presented as a product (in the category of products of two spaces and continuous maps between the respective components).

Natural Isomorphism

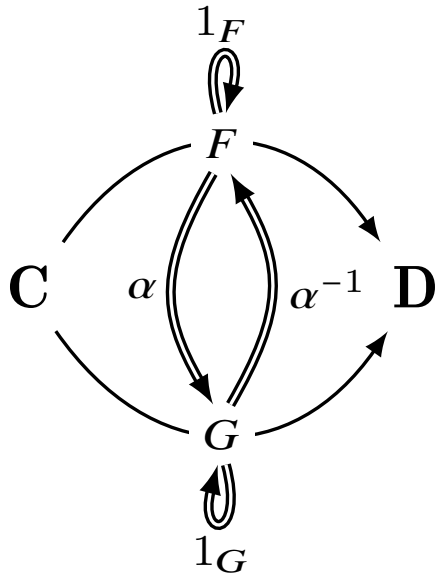
Definition (Natural Isomorphism)

A natural transformation $\alpha : F \to G$ is a natural isomorphism ($F \cong G$) iff each morphism $\alpha_A : FA \to GA$ is an isomorphism, i.e., a morphism that has a (left/right) inverse.



$$Ff = \alpha_B^{-1} \circ Gf \circ \alpha_A$$
$$Gf = \alpha_B \circ Ff \circ \alpha_A^{-1}$$

Remark



- Natural isomorphism $F \cong G$ is also expressed by saying $\alpha_A : FA \cong GA$ naturally in A.
- This is different from $\alpha_A : FA \cong GA$ for all A.
- $FA \cong GA$ naturally in $A \implies FA \cong GA$ for all A
- $FA \cong GA$ naturally in $A \iff FA \cong GA$ for all A & there is a natural transformation $\alpha : F \to G$

Remark

We have two equivalent characterizations of natural isomorphisms:

1. abstract/categorical: isomorphisms in a functor category $\mathbf{D}^{\mathbf{C}}$.
2. concrete/pointwise: each component is an isomorphism.

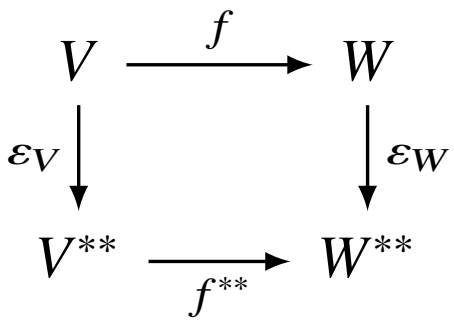
- The abstract version is better for theory and use.
- The concrete one is better for checking.
- When we check that something is a natural isomorphism it usually comes down to checking the components, but then when we use the fact that it is a natural isomorphism we often use the fact that it is an isomorphism in a functor category.

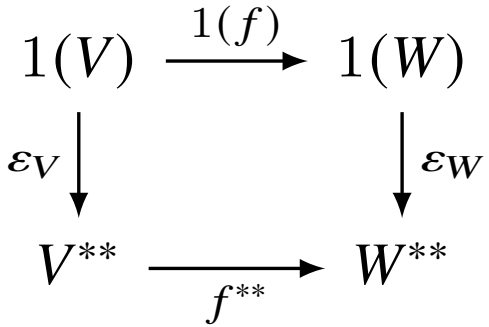
Why 'natural'?

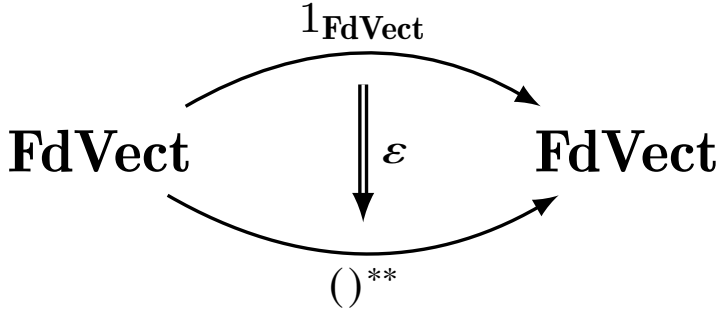
- Consider a finite dimensional vector space V over $\mathbb{R}$, and its dual space $V^* = \mathrm{Hom}(V, \mathbb{R})$ of linear maps $g : V \to \mathbb{R}$, and $f^* = \mathrm{Hom}(f, -)$.
- $V \cong V^*$.
  — Take a basis $B = \{v_1, \ldots, v_n\}$ for V. Define $v_i^* : V \to \mathbb{R}$ by $v_i^*(v_j) := \delta_{ij}$. Then $B^* = \{v_1^*, \ldots, v_n^*\}$ is a basis for V^*, and $\varphi_B : v_i \mapsto v_i^*$ is an isomorphism.
- However, $\varphi_B : V \to V^*$ depends on the initial choice of basis B.
- Now consider V^{**}. We can construct an isomorphism $\varepsilon_V : V \xrightarrow{\cong} V^{**}$ independently of any choice of basis.

$$\varepsilon_V(v) : g \mapsto g(v)$$

- The isomorphism ε_V only dependeds on the fact that V is a finite dimensional vector space over $\mathbb{R}$.







Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Equivalence of Categories

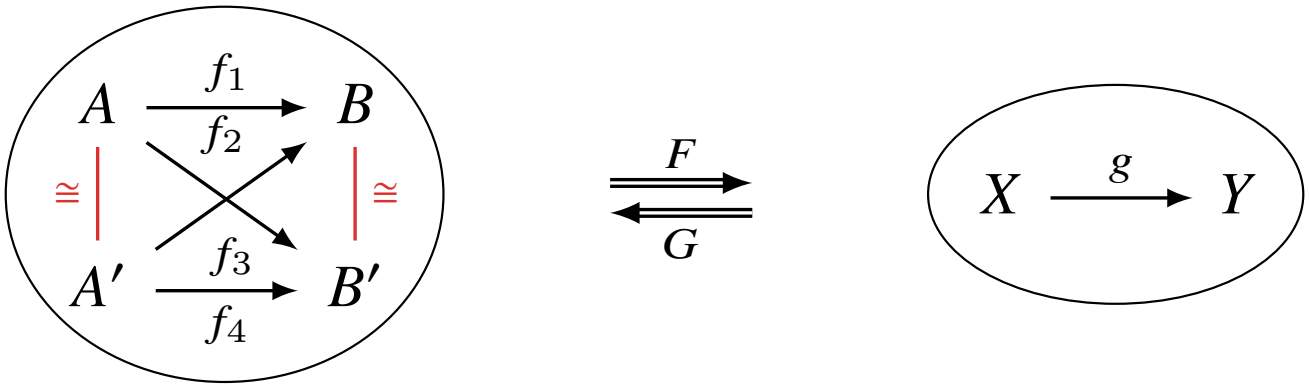
Definition (Equivalence of Categories)

The categories $\mathbf{C}$ and $\mathbf{D}$ are equivalent ($\mathbf{C} \simeq \mathbf{D}$) iff there are functors $\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$ and natural isomorphisms $G \circ F \cong 1_{\mathbf{C}}, F \circ G \cong 1_{\mathbf{D}}$.

- The categories $\mathbf{C}$ and $\mathbf{D}$ are isomorphic ($\mathbf{C} \cong \mathbf{D}$) iff there are functors $\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$ satisfying $G \circ F = 1_{\mathbf{C}}, F \circ G = 1_{\mathbf{D}}$.
- Equivalence of categories is a generalization of isomorphism. It can be seen as "isomorphism up to isomorphism".

Remark

- Categories are isomorphic if they have exactly the same object and arrow structure, so the only difference is that everything is renamed.
- Categories are equivalent if they have the same arrow structure but the objects can be "fattened up" a bit by having many isomorphic copies, or "thinned down" by identifying isomorphic objects (making them the same).

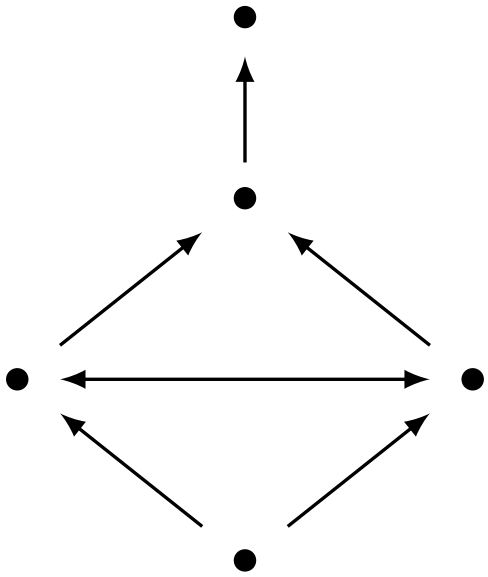


$$F := \begin{cases} A & \mapsto X \\ A' & \mapsto X \\ B & \mapsto Y \\ B' & \mapsto Y \\ f_i & \mapsto g \end{cases} \qquad G := \begin{cases} X & \mapsto A \\ Y & \mapsto B \\ g & \mapsto f_1 \end{cases}$$

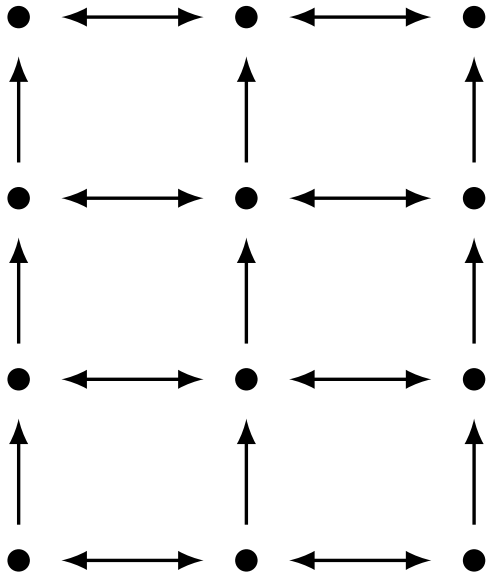
$$FG \cong 1_{\mathbf{D}}$$
$$GF \cong 1_{\mathbf{C}}$$

Examples

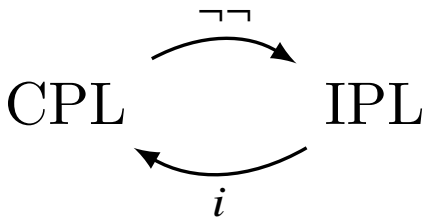
- The following categories are equivalent, but not isomorphic.







- CPL and IPL are not equivalent.



$$A \equiv_{\mathrm{CPL}} i(\neg\neg(A)) \quad \text{but} \quad A \not\equiv_{\mathrm{IPL}} \neg\neg(i(A))$$

Example

- Let $\mathbf{Par}$ be the category of sets and partial functions, where a partial function $f : A \rightharpoonup B$ is a function defined on a subset $\operatorname{supp} f \subset A$.
- Let $\mathbf{Set}_\bullet$ be the category of pointed sets and pointed functions, where a pointed set (A, a) is a set A together with an element $a \in A$, and a pointed function $f : (A, a) \to (B, b)$ is a function $f : A \to B$ such that $fa = b$.
- Then $\mathbf{Set}_\bullet \simeq \mathbf{Par}$ are equivalent.
- $F : \mathbf{Set}_\bullet \to \mathbf{Par}$ maps $(A, a) \mapsto A \setminus \{a\}$, and maps $f : (A, a) \to (B, b)$ to $Ff : F(A, a) \to F(B, b)$ defined by

$$\operatorname{supp}(Ff) = \{x \in A : fx \neq b\} \quad (Ff)x = fx$$

- $G : \mathbf{Par} \to \mathbf{Set}_\bullet$ maps $A \mapsto (A + \{\bot_A\}, \bot_A)$, where $\bot_A$ is an element that does not belong to A. And G maps $f : A \rightharpoonup B$ to $Gf : GA \to GB$ defined by

$$(Gf)x = \begin{cases} fx & \text{if } x \in \operatorname{supp} f \\ \bot_B & \text{otherwise} \end{cases}$$

Theorem

Under axiom of choice, a functor $F : \mathbf{C} \to \mathbf{D}$ is an equivalence functor iff F is full and faithful and essentially surjective on objects.

Proof.

($\Longrightarrow$) Easy.
($\Longleftarrow$) Suppose $F : \mathbf{C} \to \mathbf{D}$ is full and faithful and essentially surjective on objects. For each $B \in \mathbf{D}$, choose $GB \in \mathbf{C}$ and an isomorphism $\alpha_B : F(GB) \to B$. For $f : B \to B'$, let $Gf : GB \to GB'$ be the unique morphism s.t.

$$F(Gf) = \alpha_{B'}^{-1} \circ f \circ \alpha_B$$

Such a unique morphism exists because F is full and faithful.
This defines a functor $G : \mathbf{D} \to \mathbf{C}$.
In addition, α is a natural isomorphism $\alpha : FG \to 1_{\mathbf{D}}$.
It remains to show that $GF \cong 1_{\mathbf{C}}$. For $A \in \mathbf{C}$, let $\beta_A : A \to G(FA)$ be the unique morphism s.t. $F\beta_A = \alpha_{FA}^{-1}$. Because F reflects isomorphisms, β_A is an isomorphism for every A. Naturality of β_A follows from functoriality of F and naturality of α. □

$\mathbf{Mat} \simeq \mathbf{FdVect}_{\mathbb{R}}$

- Construct a functor $F : \mathbf{Mat} \to \mathbf{FdVect}_{\mathbb{R}}$ as follows:
  - F maps n to $\mathbb{R}^n$
  - F maps $m \times n$ matrix $M : m \to n$ to the linear map $\mathbb{R}^m \to \mathbb{R}^n$ represented by the matrix M.
- F is full and faithful and essentially surjective on objects.
- Let $G : \mathbf{FdVect}_{\mathbb{R}} \to \mathbf{Mat}$ map V to $\dim(V)$, and map a linear map $f : V \to W$ to the corresponding matrix representing f with respect to the chosen bases.
- In particular, let $\dim(V) = m$ and $\dim(W) = n$. There are isomorphisms $\phi_V : \mathbb{R}^m \to V$ and $\phi_W : \mathbb{R}^n \to W$. There exists a unique linear map $\mathbb{R}^m \to \mathbb{R}^n$ making the diagram commute:

$$\begin{array}{ccc} \mathbb{R}^m & \longrightarrow & \mathbb{R}^n \\ {\scriptstyle \phi_V}\downarrow{\scriptstyle \cong} & & {\scriptstyle \cong}\downarrow{\scriptstyle \phi_W} \\ V & \xrightarrow[f]{} & W \end{array}$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

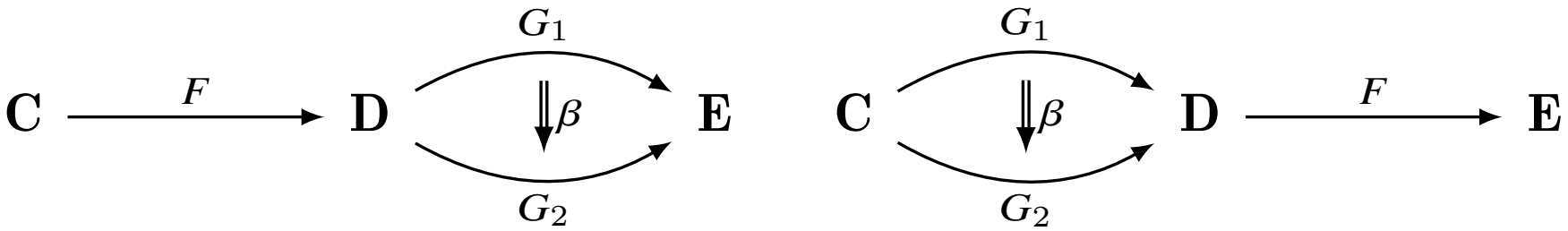
Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

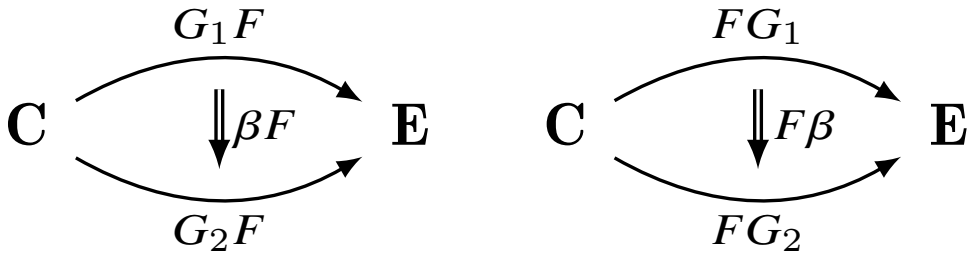
Answers to the Exercises

References2039

Whiskering

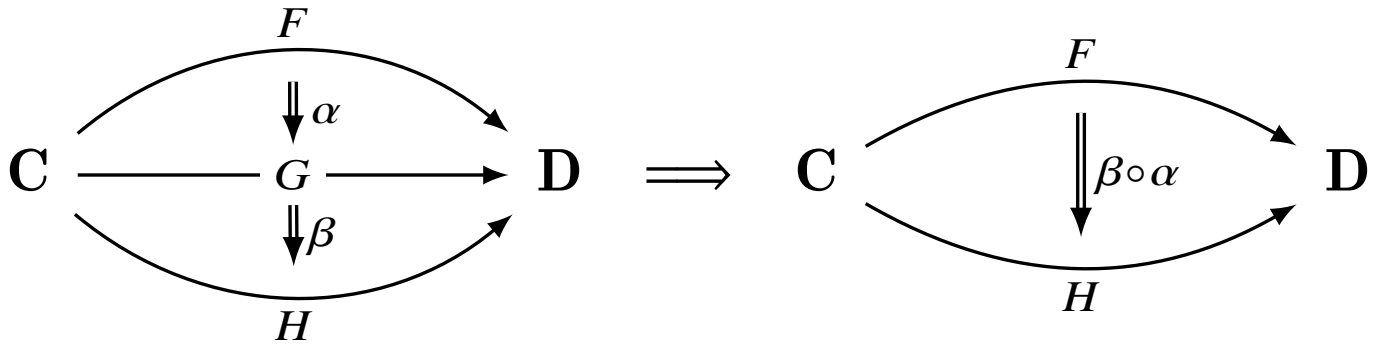


The prewhiskering of β by F, denoted $\beta F := \beta * 1_F : G_1 F \to G_2 F$ (resp. the postwhiskering of β by F, denoted $F\beta := 1_F * \beta : FG_1 \to FG_2$) is defined as follows.

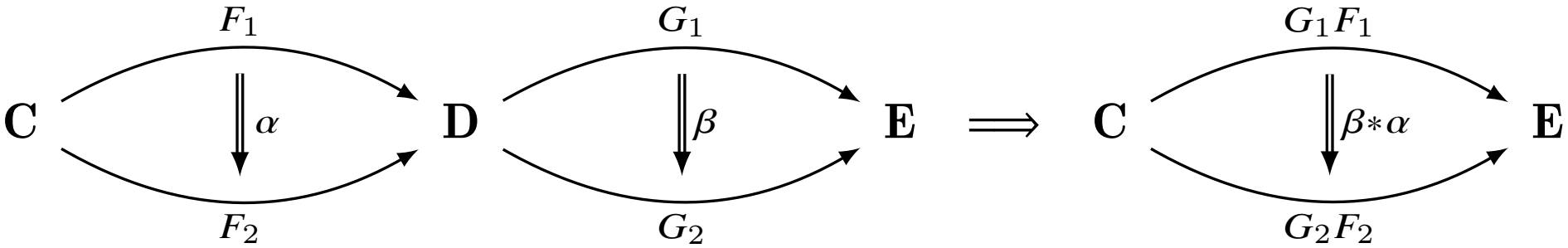


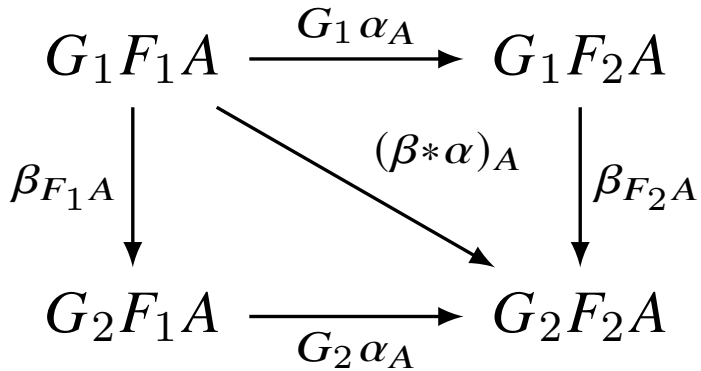
$$(\beta F)_A := \beta_{FA}$$
$$(F\beta)_A := F\beta_A$$

Vertical and Horizontal Composition



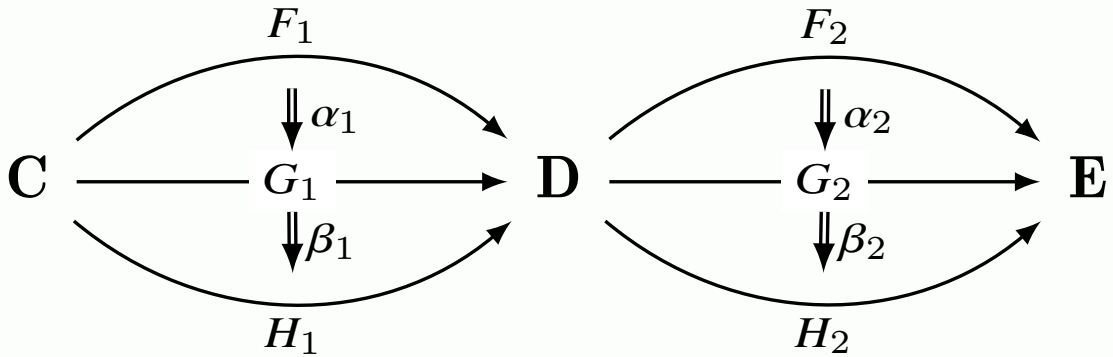
$$(\beta \circ \alpha)_A \coloneqq \beta_A \circ \alpha_A$$





Middle Four Interchange

Theorem (Middle Four Interchange)



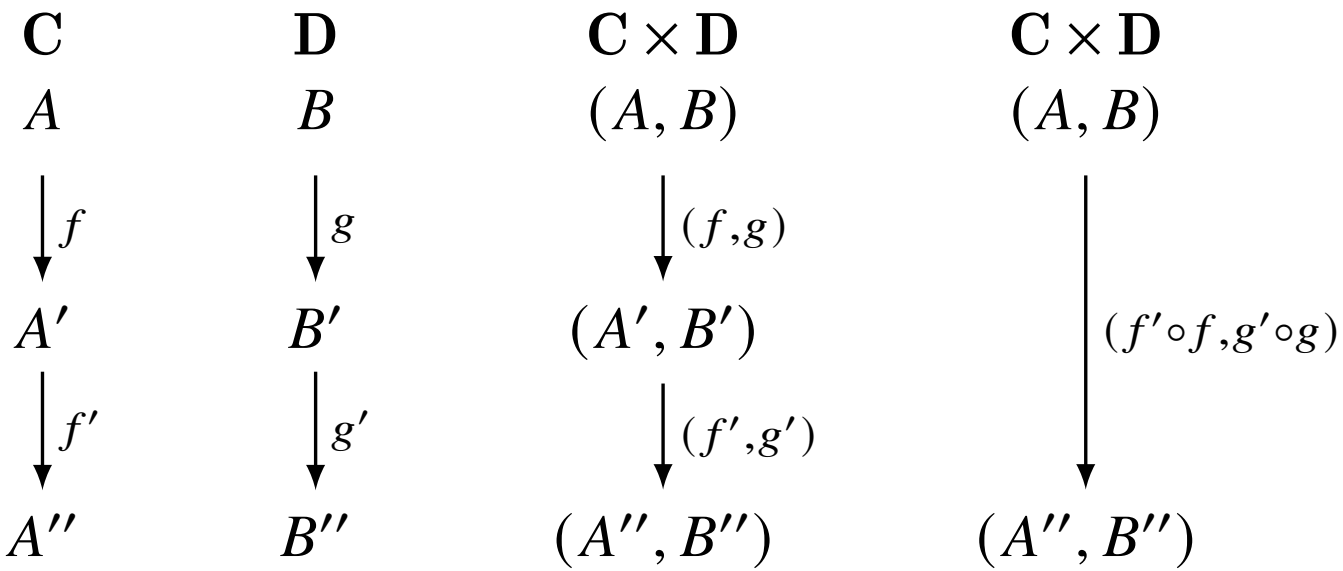
$$(\beta_2 * \beta_1) \circ (\alpha_2 * \alpha_1) = (\beta_2 \circ \alpha_2) * (\beta_1 \circ \alpha_1)$$

Product Category

Definition (Product Category)

Given categories $\mathbf{C}$ and $\mathbf{D}$, the product category $\mathbf{C} \times \mathbf{D}$ has

- objects (A, B) for $A \in \mathbf{C}$ and $B \in \mathbf{D}$.
- morphisms $(f, g) : (A, B) \to (A', B')$ for $f : A \to A'$ and $g : B \to B'$.
- identity $1_{(A,B)} \coloneqq (1_A, 1_B)$.
- composition $(f', g') \circ (f, g) \coloneqq (f' \circ f, g' \circ g)$.



Product/Coproduct Category

Definition (Product Category $\prod_{i \in I} \mathbf{C}_i$)

$$\mathrm{ob}\left(\prod_{i \in I} \mathbf{C}_i\right) = \prod_{i \in I} \mathrm{ob}(\mathbf{C}_i)$$

$$\mathrm{Hom}_{\prod_{i \in I} \mathbf{C}_i}\left((X_i)_{i \in I}, (Y_i)_{i \in I}\right) = \prod_{i \in I} \mathrm{Hom}_{\mathbf{C}_i}(X_i, Y_i)$$

Definition (Coproduct Category $\coprod_{i \in I} \mathbf{C}_i$)

$$\mathrm{ob}\left(\coprod_{i \in I} \mathbf{C}_i\right) = \coprod_{i \in I} \mathrm{ob}(\mathbf{C}_i)$$

$$\mathrm{Hom}_{\coprod_{i \in I} \mathbf{C}_i}(X, Y) = \begin{cases} \mathrm{Hom}_{\mathbf{C}_i}(X, Y) & \text{if } X, Y \in \mathrm{ob}(\mathbf{C}_i) \\ \emptyset & \text{otherwise} \end{cases}$$

Functor Category

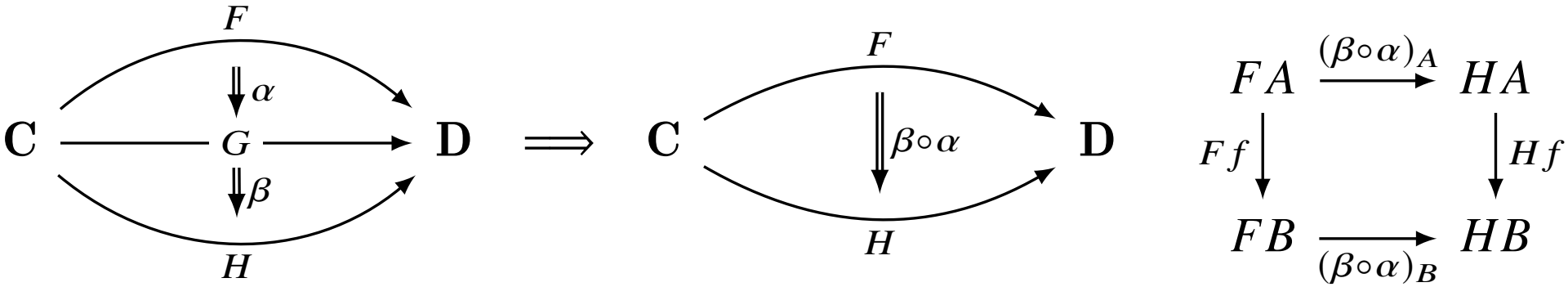
Definition (Functor Category)

Given categories $\mathbf{C}$ and $\mathbf{D}$, the functor category $\mathbf{D}^{\mathbf{C}}$ (or denoted by $[\mathbf{C}, \mathbf{D}]$) has

- objects: functors $F : \mathbf{C} \to \mathbf{D}$
- morphisms: natural transformations $\alpha : F \to G$
- identity natural transformation $(1_F)_A \coloneqq 1_{FA}$
  given a functor $F : \mathbf{C} \to \mathbf{D}$, define $1_F : F \to F$ with
  $(1_F)_A \coloneqq FA \xrightarrow{1_{FA}} FA$.
- composition of natural transformations $(\beta \circ \alpha)_A \coloneqq \beta_A \circ \alpha_A$
  given functors $F, G, H : \mathbf{C} \to \mathbf{D}$ and natural transformations
  $F \xrightarrow{\alpha} G \xrightarrow{\beta} H$, define $\beta \circ \alpha : F \to H$ with

$$(\beta \circ \alpha)_A \coloneqq FA \xrightarrow{\alpha_A} GA \xrightarrow{\beta_A} HA$$

Remark



$$(\beta \circ \alpha)_A \coloneqq \beta_A \circ \alpha_A$$

- Note that all the "action" is going on in the target category $\mathbf{D}$: the components of the natural transformation are in $\mathbf{D}$ and the naturality squares are in $\mathbf{D}$.
- Functor categories inherit more of their properties from the target category than from the source category.
- The fact that a functor category gets structure from its target category means it can be very useful to study a category $\mathbf{C}$ via functors into a category that we know has a lot of excellent structure, such as $\mathbf{Set}$.

The Category of Small Categories $\mathbf{Cat}$

- Assume there is an infinite sequence $U_0 \in U_1 \in U_2 \in \ldots$ of bigger and bigger Grothendieck universes.
- $\mathbf{Set}_n$ = category whose objects are the sets in U_n and with $\mathbf{Set}_n(A, B) = B^A$ = the functions from A to B.
- A category $\mathbf{C}$ is locally small iff $\forall AB \in \mathbf{C} : \mathbf{C}(A, B) \in \mathbf{Set}_0$.
- A category $\mathbf{C}$ is small iff it is both locally small and $\mathrm{ob}(\mathbf{C}) \in \mathbf{Set}_0$.

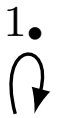
Definition (The Category of Small Categories $\mathbf{Cat}$)

The category of small categories $\mathbf{Cat}$ has

- objects: small categories.
- morphisms: functors $F : \mathbf{C} \to \mathbf{D}$.
- identity and composition as for functors.

$\mathbf{Cat}$ is large.

Cat is CCC

- $\mathbf{Cat}$ has a terminal object $\mathbf{1} \coloneqq$ 

.
- $\mathbf{Cat}$ has products.
- There is a functor $\varepsilon : \mathbf{D}^{\mathbf{C}} \times \mathbf{C} \to \mathbf{D}$ that makes $\mathbf{D}^{\mathbf{C}}$ the exponential.
- $\mathbf{Cat}$ is cartesian closed.

Products and Sums of Functors

Definition (Product of Functors)

Given $F : \mathbf{A} \to \mathbf{B}$ and $G : \mathbf{C} \to \mathbf{D}$, their product $F \times G : \mathbf{A} \times \mathbf{C} \to \mathbf{B} \times \mathbf{D}$ is given by

$$F \times G : (X, Y) \mapsto (FX, GY)$$
$$F \times G : (f, g) \mapsto (Ff, Gg)$$

Definition (Sum of Functors)

Given $F : \mathbf{A} \to \mathbf{B}$ and $G : \mathbf{C} \to \mathbf{D}$, their sum $F + G : \mathbf{A} + \mathbf{C} \to \mathbf{B} + \mathbf{D}$ is given by

$$F + G : \begin{cases} (0, X) \mapsto (0, FX) \\ (1, Y) \mapsto (1, GY) \end{cases}$$
$$F + G : \begin{cases} (0, f) \mapsto (0, Ff) \\ (1, g) \mapsto (1, Gg) \end{cases}$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

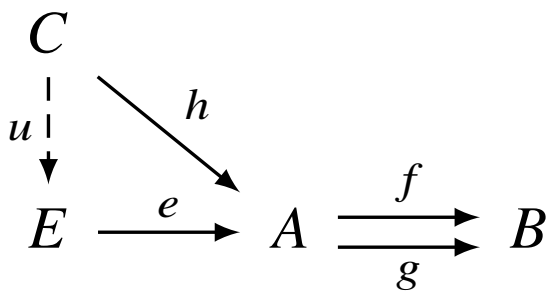
Quantum Computing

Answers to the Exercises

References2039

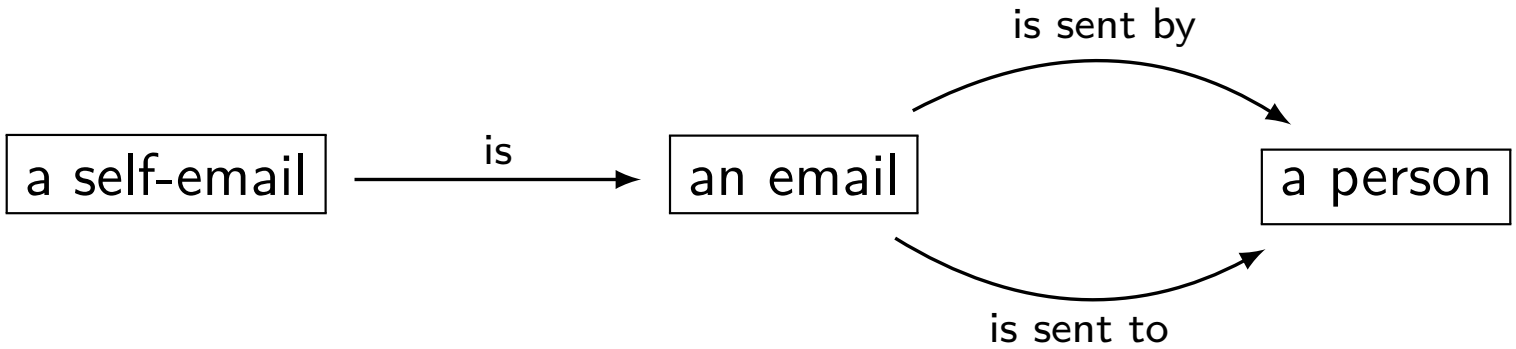
Equalizer

- A *fork* in a category $\mathbf{C}$ consists of $C \xrightarrow{h} A \underset{g}{\overset{f}{\rightrightarrows}} B$ s.t. $fh = gh$.
- Let $\mathbf{C}$ be a category and take $A \underset{g}{\overset{f}{\rightrightarrows}} B$. An *equalizer* of f and g is an object E with a morphism $E \xrightarrow{e} A$ s.t. $E \xrightarrow{e} A \underset{g}{\overset{f}{\rightrightarrows}} B$ is a fork, and for any fork $C \xrightarrow{h} A \underset{g}{\overset{f}{\rightrightarrows}} B$, there exists a unique morphism $C \xrightarrow{u} E$ s.t.



Example: In $\mathbf{Set}$, $E := \{x \in A : f(x) = g(x)\}$.

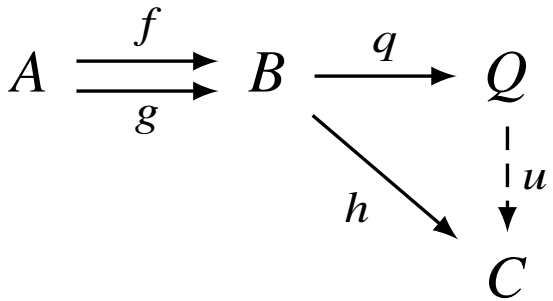
Example — Equalizer



A self-email is an email which is sent by the same person it is sent to.

Coequalizer

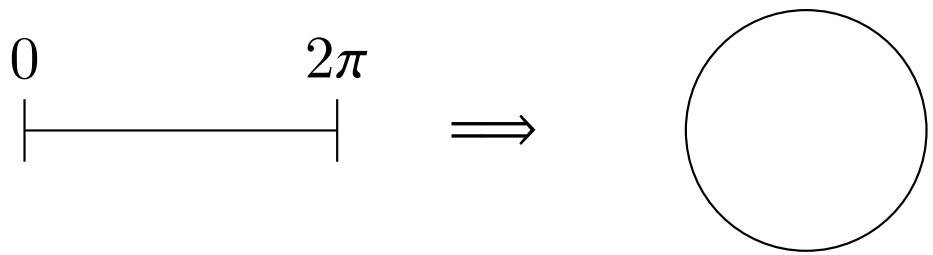
- A *cofork* in a category $\mathbf{C}$ consists of $A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{h} C$ s.t. $hf = hg$.
- Let $\mathbf{C}$ be a category and take $A \underset{g}{\overset{f}{\rightrightarrows}} B$. An *coequalizer* of f and g is an object Q with a morphism $B \xrightarrow{q} Q$ s.t. $A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{q} Q$ is a cofork, and for any cofork $A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{h} C$, there exists a unique morphism $Q \xrightarrow{u} C$ s.t.



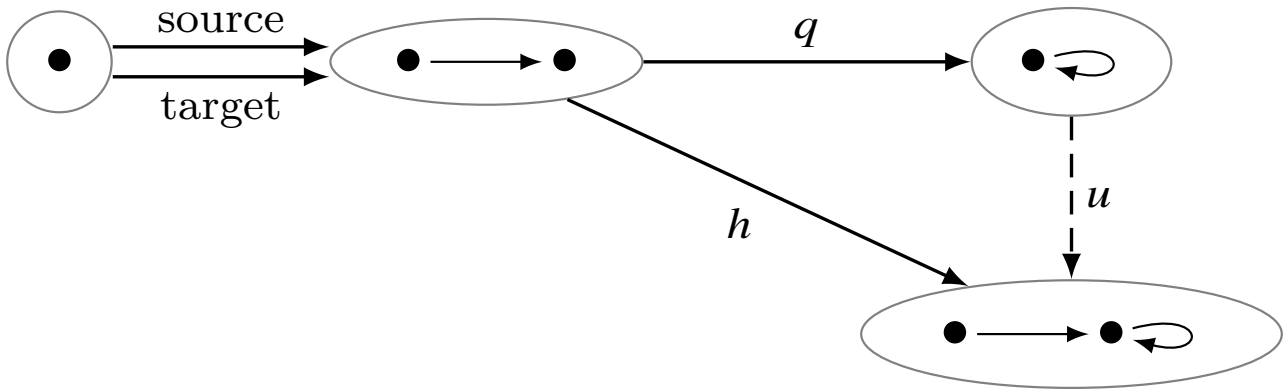
Example: In $\mathbf{Set}$, let $y \sim y' := \exists x\big(f(x) = y \wedge g(x) = y'\big)$. Then $Q := B/\sim$ is a coequalizer.

Example — Coequalizer

$$\{\bullet\} \underset{\bullet \mapsto 2\pi}{\overset{\bullet \mapsto 0}{\rightrightarrows}} [0, 2\pi] \xrightarrow{\theta \mapsto (\cos\theta, \sin\theta)} \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\right\}$$



Example — Coequalizer



Theorem

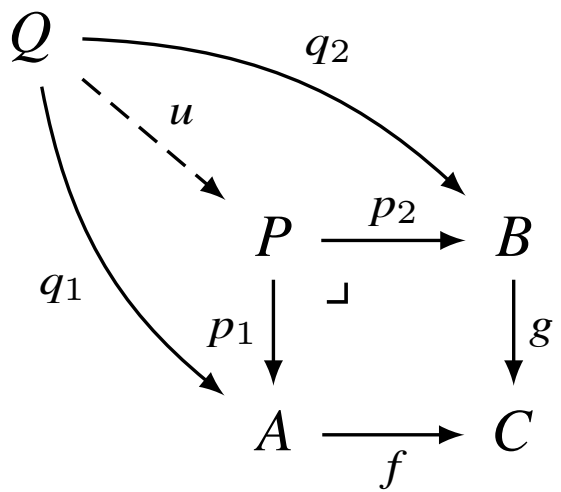
If $C \xrightarrow{h} A \underset{g}{\overset{f}{\rightrightarrows}} B$ *is an equalizer,* h *is monic.*

Theorem

If $A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{h} C$ *is a coequalizer,* h *is epic.*

Pullback

Let $\mathbf{C}$ be a category. A *pullback* of $\begin{array}{ccc} & & B \\ & & \downarrow g \\ A & \xrightarrow{f} & C \end{array}$ is an object P with $P \xrightarrow{p_1} A$ and $P \xrightarrow{p_2} B$ s.t. $\begin{array}{ccc} P & \xrightarrow{p_2} & B \\ {\scriptstyle p_1}\downarrow & & \downarrow{\scriptstyle g} \\ A & \xrightarrow[f]{} & C \end{array}$, and for any $\begin{array}{ccc} Q & \xrightarrow{q_2} & B \\ {\scriptstyle q_1}\downarrow & & \downarrow{\scriptstyle g} \\ A & \xrightarrow[f]{} & C \end{array}$ there is a unique morphism $Q \xrightarrow{u} P$ s.t.



.

Example — Pullback

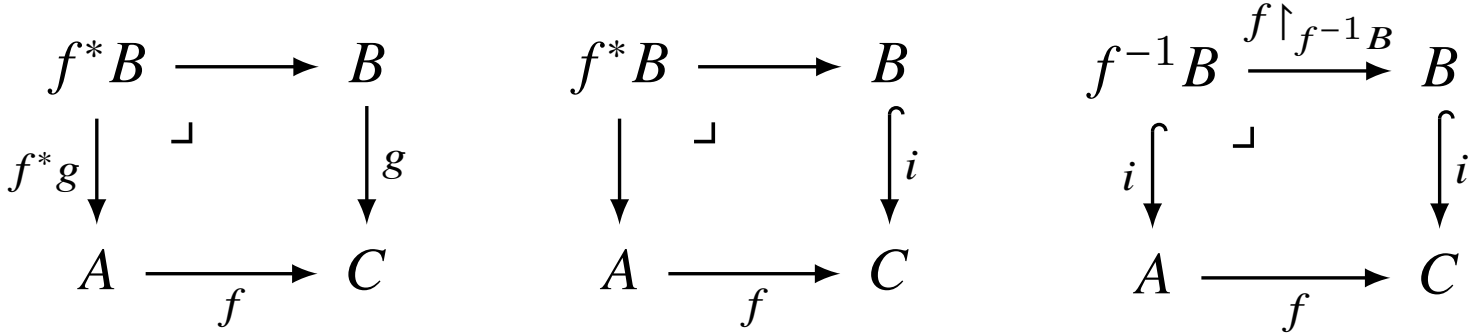
- In $\mathbf{Set}$, the pullback is

$$A \times_C B := \{(a, b) \in A \times B : fa = gb\} = \coprod_{c \in C} f^{-1}(c) \times g^{-1}(c)$$

- For example, consider a category $\mathbf{C}$ with

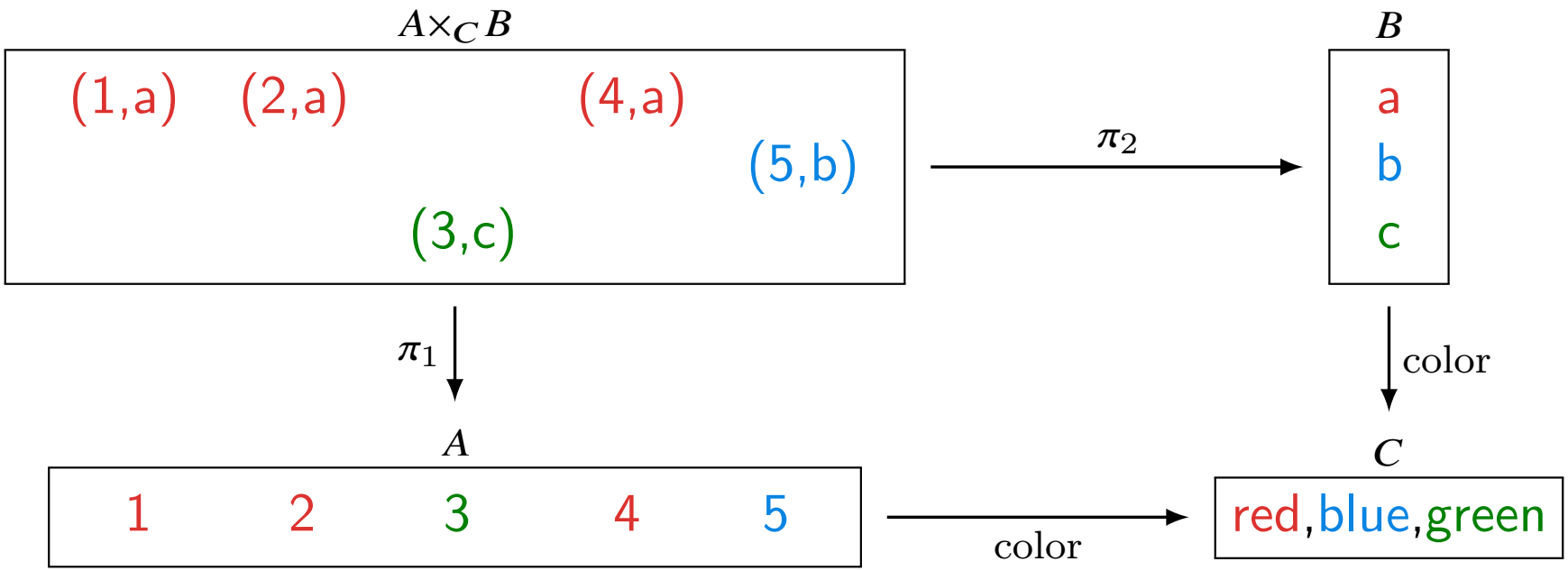
$$\operatorname{mor} \mathbf{C} \xrightarrow{\text{cod}} \operatorname{ob} \mathbf{C} \xleftarrow{\text{dom}} \operatorname{mor} \mathbf{C}$$

  Then the pullback is the set of composable morphisms.
- We also say that we pull back g along f and think of $f^*g : f^*B \to A$ as the inverse image of B along f. This terminology is explained by looking at the pullback of a subset inclusion $i : B \hookrightarrow C$.



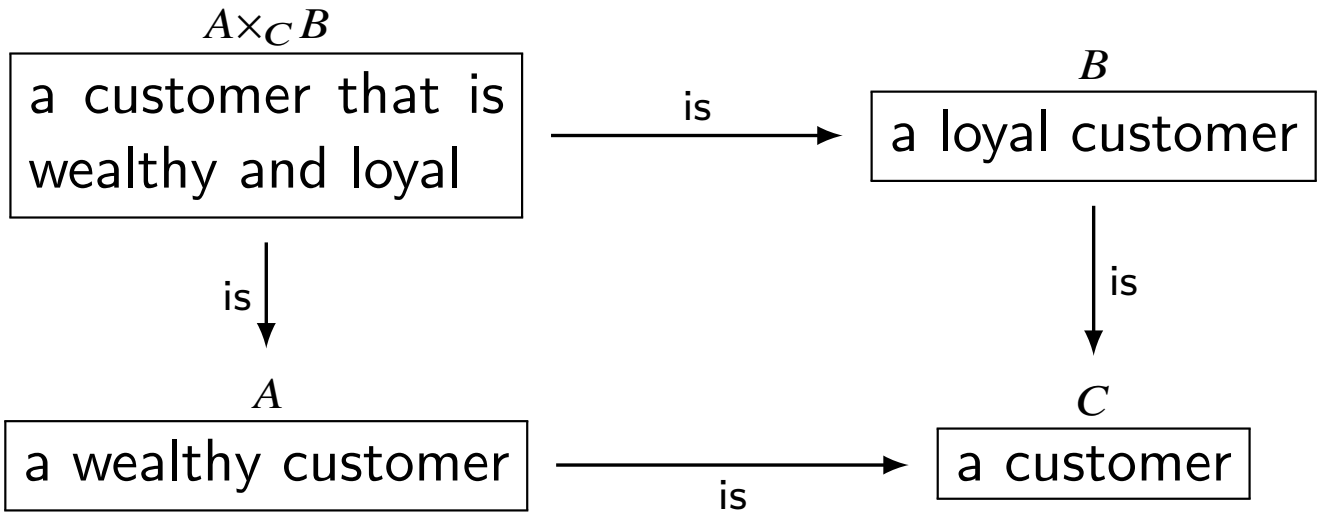
In this case $\{(a, b) \in A \times B : fa = b\} \cong \{a \in A : fa \in B\} = f^*B$.

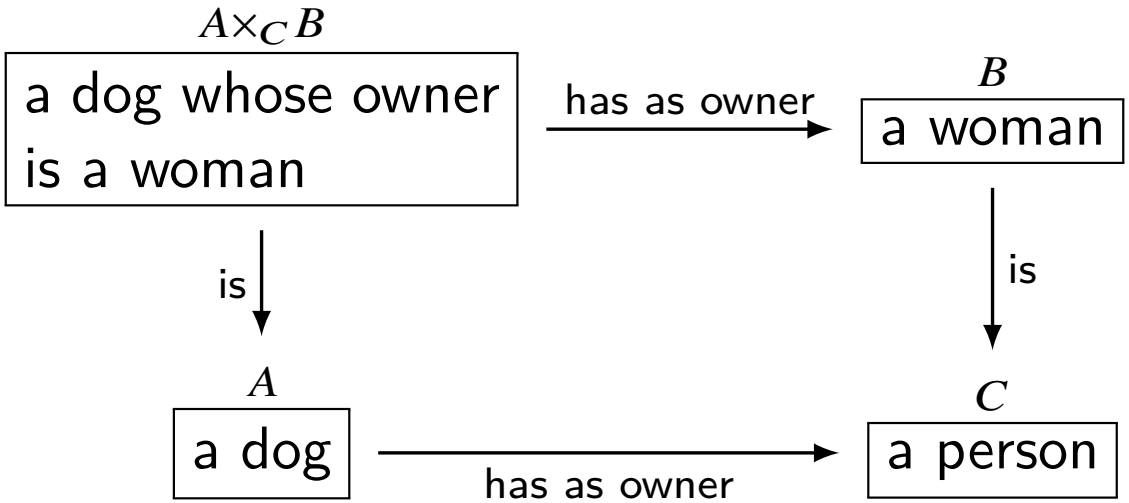
Example — Pullback

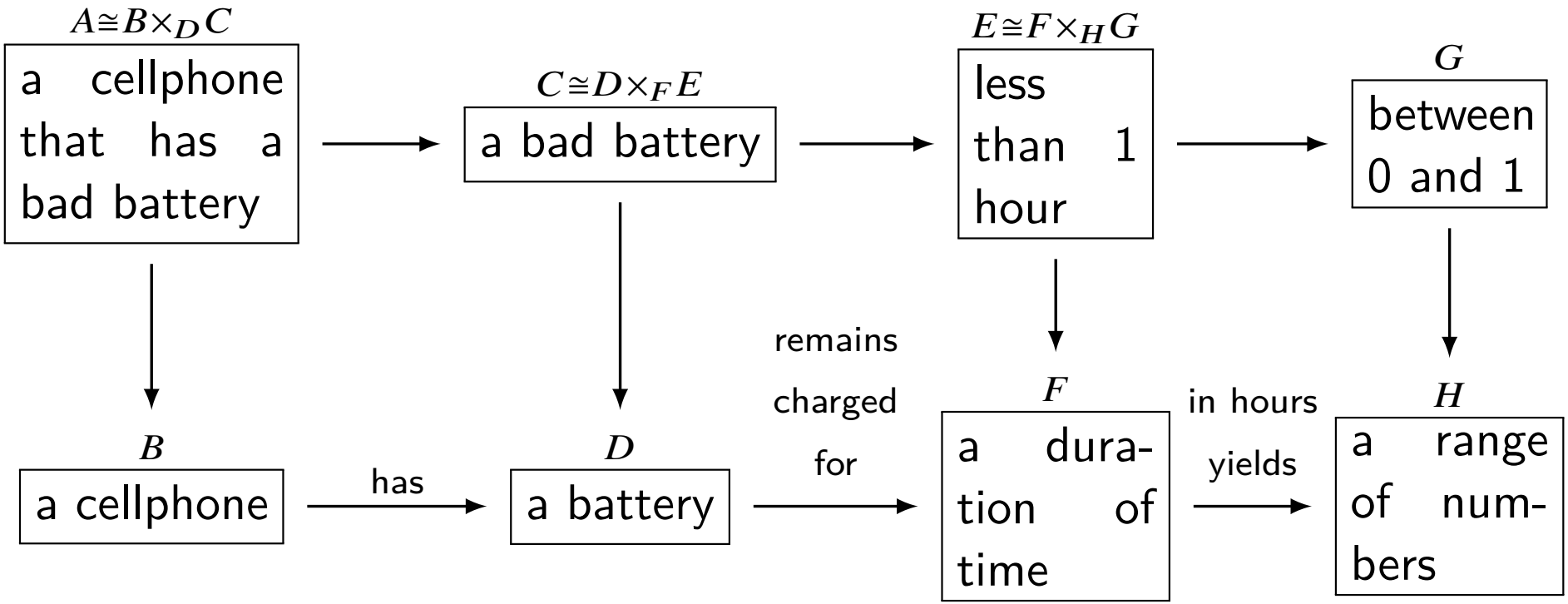


Note that inside the set of $A \times B = 15$ possible (x, y) pairs is the set of pairs that agree on color — this is $A \times_C B$.

Example — Pullback







Since

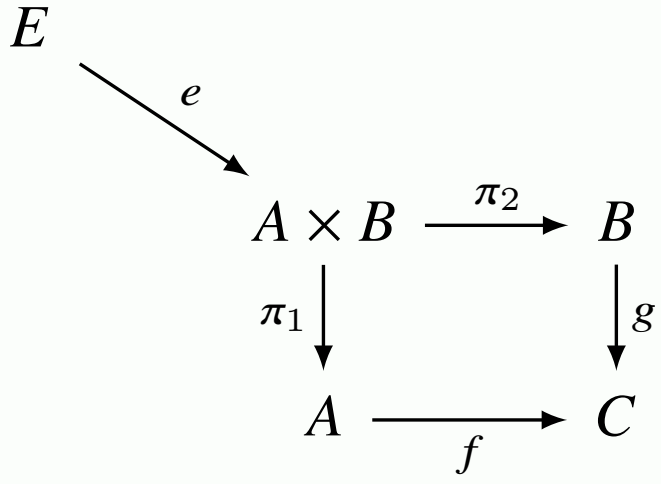
$$A \cong B \times_D C \ \ \& \ \ C \cong D \times_F E \implies A \cong B \times_F E$$

we can deduce the definition "a cellphone that has a bad battery is defined as a cellphone that has a battery which remains charged for less than 1 hour."

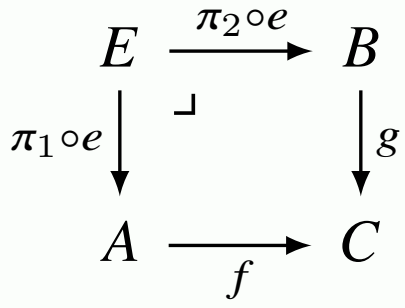
products & equalizers $\Longrightarrow$ pullbacks

Theorem

In a category with products and equalizers, given $A \xrightarrow{f} C \xleftarrow{g} B$ *, consider the diagram:*



Then $e : E \to A \times B$ *is an equalizer of* $A \times B \underset{g \circ \pi_2}{\overset{f \circ \pi_1}{\rightrightarrows}} C$ *iff*



is a pullback.

- pullbacks & terminal objects $\Longrightarrow$ products

  $A \times B \cong A \times_1 B$

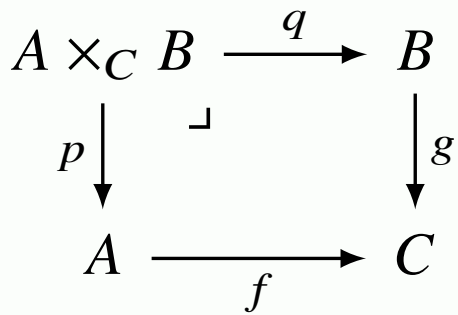
$$\begin{array}{ccc} A \times B & \xrightarrow{\pi_2} & B \\ {\scriptstyle \pi_1}\downarrow & \lrcorner & \downarrow{\scriptstyle !_B} \\ A & \xrightarrow[!_A]{} & 1 \end{array}$$

- pullbacks & products $\Longrightarrow$ equalizers

  The equalizer $E \overset{e}{\rightarrowtail} A \underset{g}{\overset{f}{\rightrightarrows}} B$ is constructed as the following pullback,

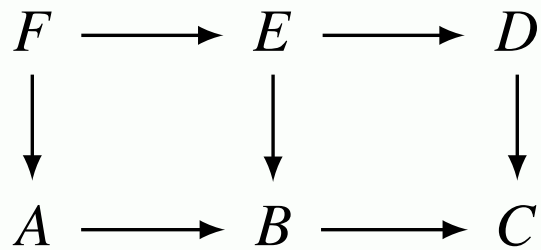
$$\begin{array}{ccc} E & \xrightarrow{h} & B \\ {\scriptstyle e}\downarrow & \lrcorner & \downarrow{\scriptstyle \Delta} \\ A & \xrightarrow[\langle f,g \rangle]{} & B \times B \end{array}$$

Theorem



If g is monic, so is p.

Theorem



1. *If the two squares are pullbacks, so is the outer rectangle. Thus,*

$$A \times_B (B \times_C D) \cong A \times_C D$$

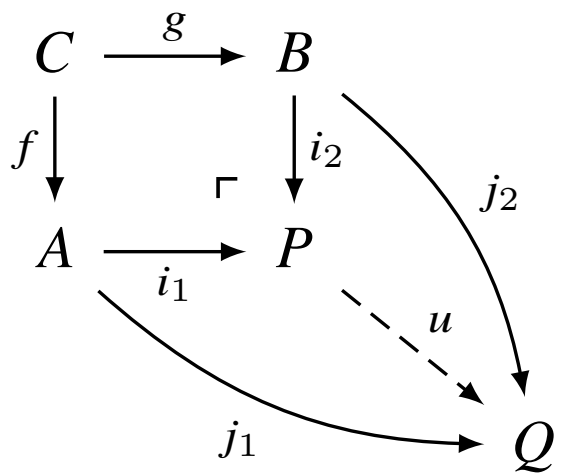
2. *If the right square and the outer rectangle are pullbacks, so is the left square.*

Pushout

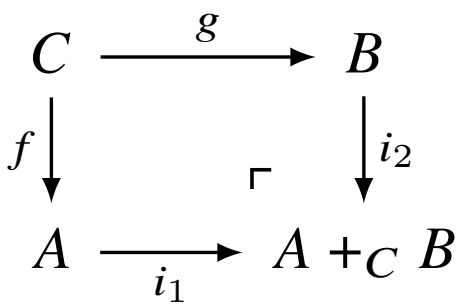
Let $\mathbf{C}$ be a category. A *pushout* of $\begin{array}{ccc} C & \xrightarrow{g} & B \\ {\scriptstyle f}\downarrow & & \\ A & & \end{array}$ is an object P with

$A \xrightarrow{i_1} P$ and $B \xrightarrow{i_2} P$ s.t. $\begin{array}{ccc} C & \xrightarrow{g} & B \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle i_2} \\ A & \xrightarrow[i_1]{} & P \end{array}$, and for any $\begin{array}{ccc} C & \xrightarrow{g} & B \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle j_2} \\ A & \xrightarrow[j_1]{} & Q \end{array}$ there is

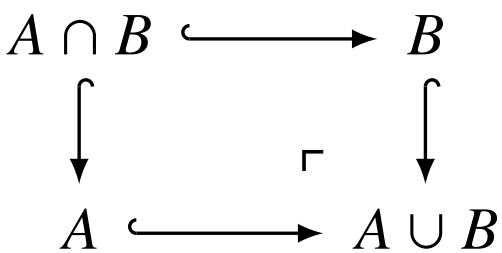
a unique morphism $P \xrightarrow{u} Q$ s.t.



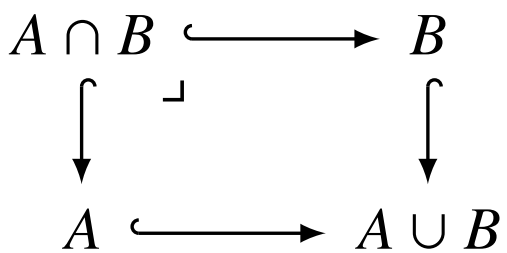
Pushout



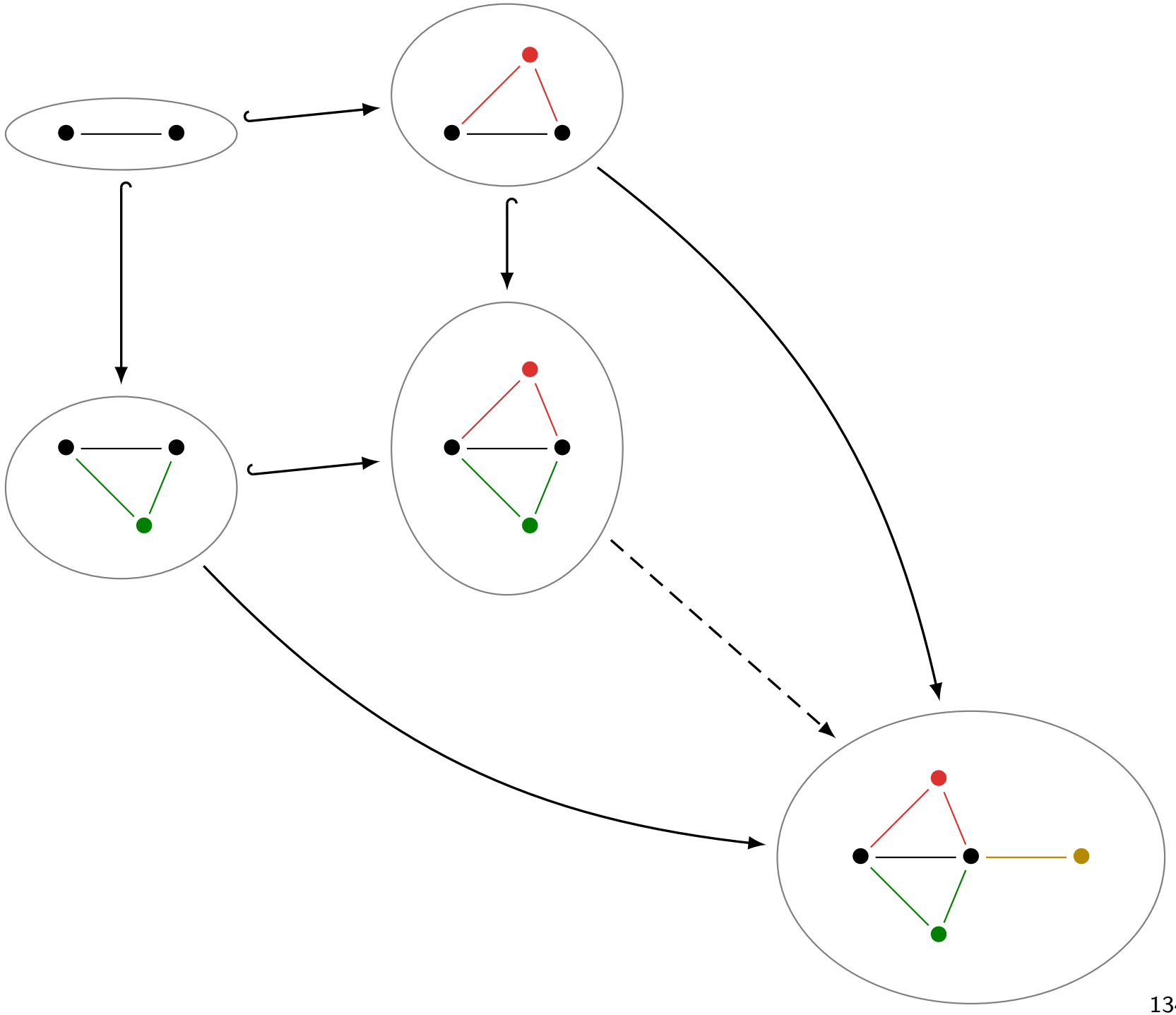
In $\mathbf{Set}$, $A +_C B := (A + B)/\sim$, where $\sim$ is the smallest equivalence relation on $A + B$ such that for all $x \in C : fx = gx$.



$A +_{A \cap B} B = A \cup B$

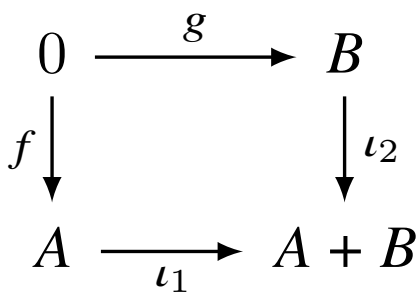


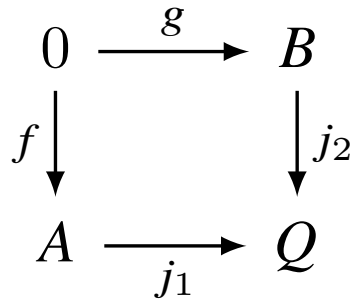
$A \times_{A \cup B} B = A \cap B$

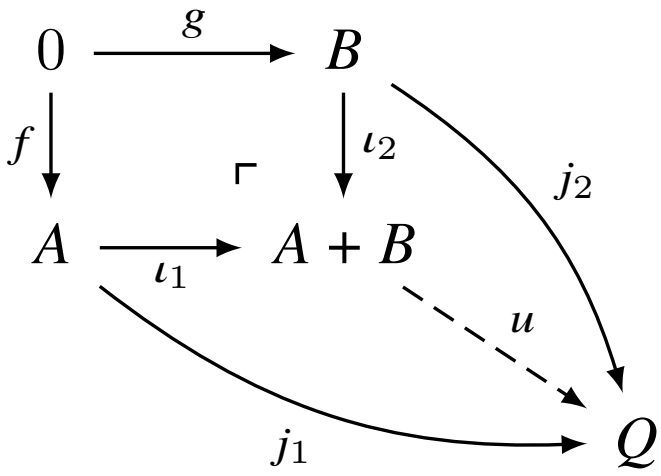


Pushout vs Coproduct

- Suppose 0 is the initial object. The diagram $A \xleftarrow{f} 0 \xrightarrow{g} B$ has a pushout iff A and B have a coproduct, and they are the same.
- Suppose A and B have a coproduct $A + B$, then the left diagram commutes, and for any commutative diagram of the middle form, there is a unique map $A + B \to Q$ s.t. the right diagram commutes.





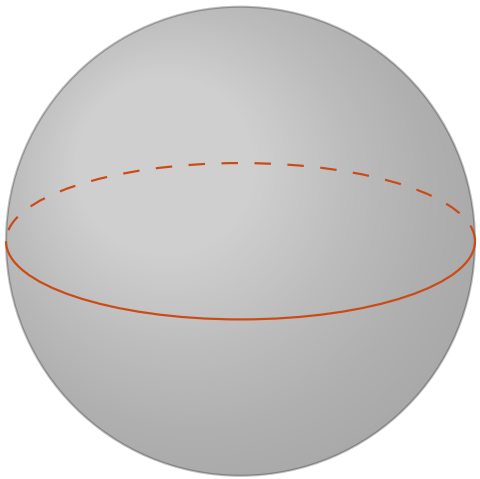


This shows that $A + B$ is a pushout, $A + B \cong A +_0 B$.

- Similarly, if a pushout $A +_0 B$ exists, then it satisfies the universal property of the coproduct.

Pushout — Example

$$\begin{array}{ccc}
\left\{(x,y,0) \in \mathbb{R}^3 : x^2+y^2=1\right\} & \longrightarrow & \left\{(x,y,z) \in \mathbb{R}^3 : \begin{array}{l} x^2+y^2+z^2=1 \\ z \geq 0 \end{array}\right\} \\
\downarrow & & \downarrow \\
\left\{(x,y,z) \in \mathbb{R}^3 : \begin{array}{l} x^2+y^2+z^2=1 \\ z \leq 0 \end{array}\right\} & \hookrightarrow & \left\{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2=1\right\}
\end{array}$$



Pushout — Example

- Pushouts in $\mathbf{Top}$ are similarly formed from coproducts and coequalizers, which can be made first in $\mathbf{Set}$ and then topologized as sum and quotient spaces.
- Pushouts are used, for example, to construct spheres from disks.
- Let D^2 be the disk and S^1 the circle, with its inclusion $i : S^1 \hookrightarrow D^2$ as the boundary of the disk.
- Then, the two-sphere S^2 is the pushout,

$$\begin{array}{ccc} S^1 & \xhookrightarrow{i} & D^2 \\ \downarrow & & \downarrow \\ \bullet & \longrightarrow & S^2 \end{array}$$

- In another way, the sphere is the coequalizer

$$S^1 \times (0,1) \rightrightarrows D + D \longrightarrow S^2$$

Remark

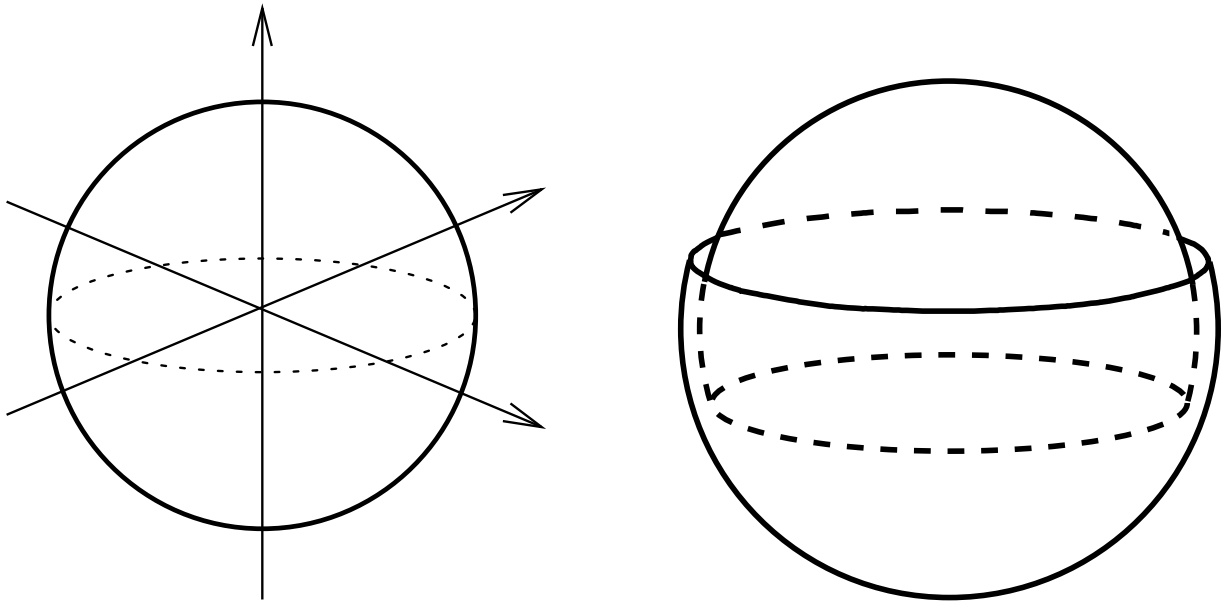


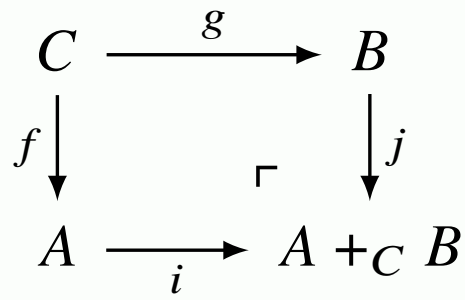
Figure: Sphere as a limit, and a colimit.

- One disadvantage of the limit point of view is that it makes an arbitrary choice of coordinate system.

$$S^2 \hookrightarrow \mathbb{R}^3 \underset{1}{\overset{x^2+y^2+z^2}{\rightrightarrows}} \mathbb{R}$$

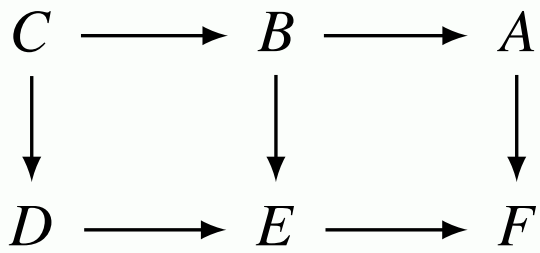
- It is generally best to think of spaces as free-standing objects, existing independently of any particular embedding into Euclidean space.

Theorem



If f is epic, so is j.

Theorem



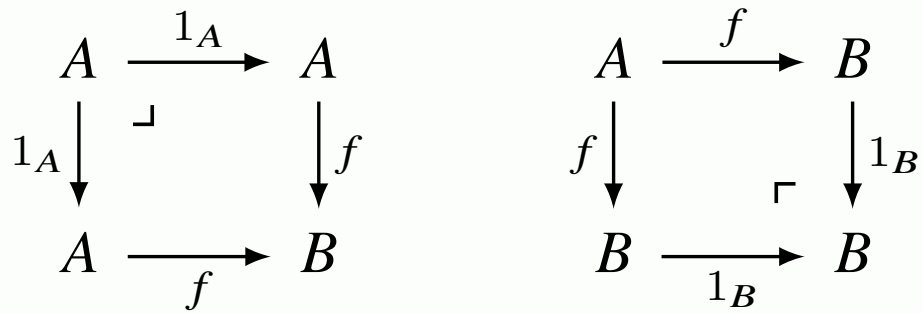
1. *If the two squares are pushouts, so is the outer rectangle. Thus,*

$$A +_B (B +_C D) \cong A +_C D$$

2. *If the left square and the outer rectangle are pushouts, so is the right square.*

Theorem

Let $\mathbf{C}$ be a category, and let $f : A \to B$ be a morphism. It is a monomorphism (resp. epimorphism) iff the square to the left is a pullback (resp. the square to the right is a pushout):



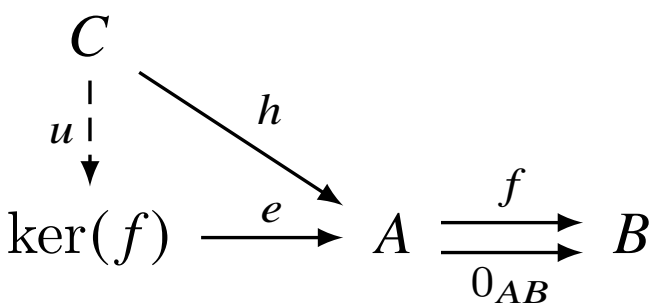
Kernel & Cokernel

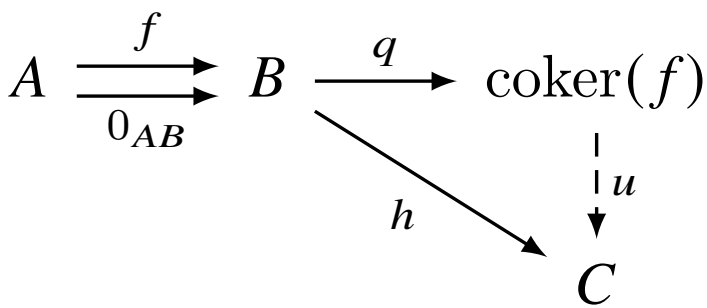
- In a category $\mathbf{C}$ with a zero object 0, the *zero morphism* $0_{AB} : A \to B$ between $A, B \in \mathbf{C}$ is the unique morphism that factors through 0:

$$0_{AB} : A \to 0 \to B$$

- In a category with zero morphism, the *kernel* $\ker(f)$ of $f : A \to B$ is the equalizer of f and the zero morphism 0_{AB}.
- In a category with zero morphism, the *cokernel* $\operatorname{coker}(f)$ of $f : A \to B$ is the coequalizer of f and the zero morphism 0_{AB}.

$$\ker(f) = \operatorname{eq}(f, 0_{AB}) \qquad \operatorname{coker}(f) = \operatorname{coeq}(f, 0_{AB})$$

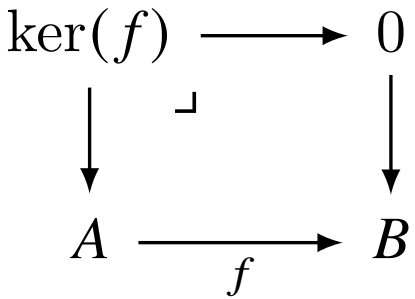




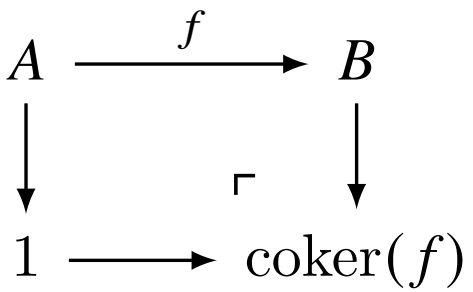
The composition of a zero morphism with any morphism is a zero morphism.

Kernel & Cokernel

- In a category with an initial object 0, *the kernel* $\ker(f)$ of a morphism $f : A \to B$ is the pullback:

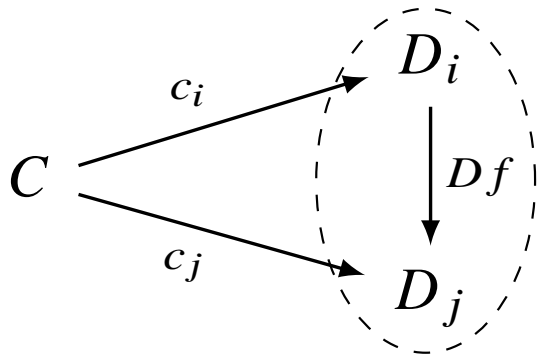


- In a category with a terminal object 1, the *cokernel* $\mathrm{coker}(f)$ of a morphism $f : A \to B$ is the pushout:



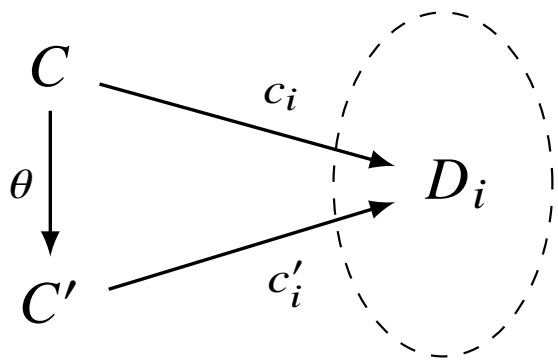
Cone

- A *diagram* of a small category $\mathbf{I}$ in category $\mathbf{C}$ is a functor $D : \mathbf{I} \to \mathbf{C}$.
- A *cone* (C, c) over a diagram D consists of an object $C \in \mathbf{C}$ and a family of morphisms $(C \xrightarrow{c_i} D_i)_{i \in \mathbf{I}}$ such that for each $i \xrightarrow{f} j$ in $\mathbf{I}$, the following triangle commutes.



A cone (C, c) over D can be taken as a natural transformation $c : \Delta_C \to D$.

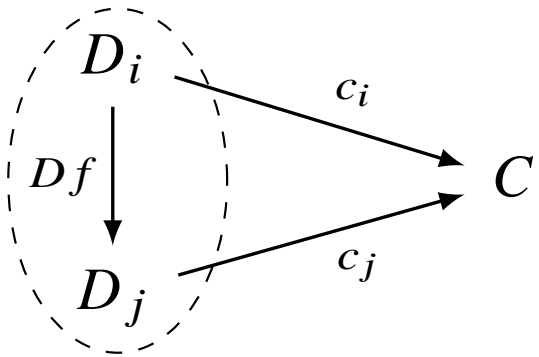
- A morphism of cones $\theta : (C, c) \to (C', c')$ is a morphism θ s.t.



Then we have a category $\int \mathbf{Cone}(-, D)$ of all the cones over D.

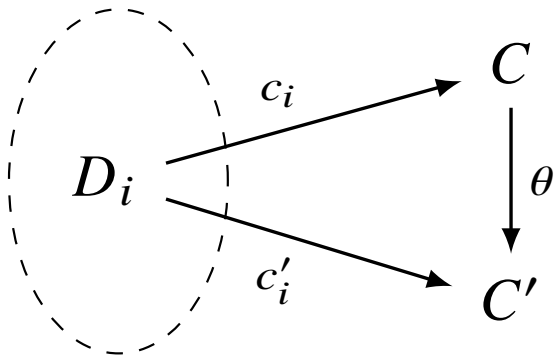
Cocone

- A *cocone* (C, c) over a diagram D consists of an object $C \in \mathbf{C}$ and a family of morphisms $(D_i \xrightarrow{c_i} C)_{i \in \mathbf{I}}$ such that for each $i \xrightarrow{f} j$ in $\mathbf{I}$, the following triangle commutes.



A cocone (C, c) over D can be taken as a natural transformation $c : D \to \Delta_C$.

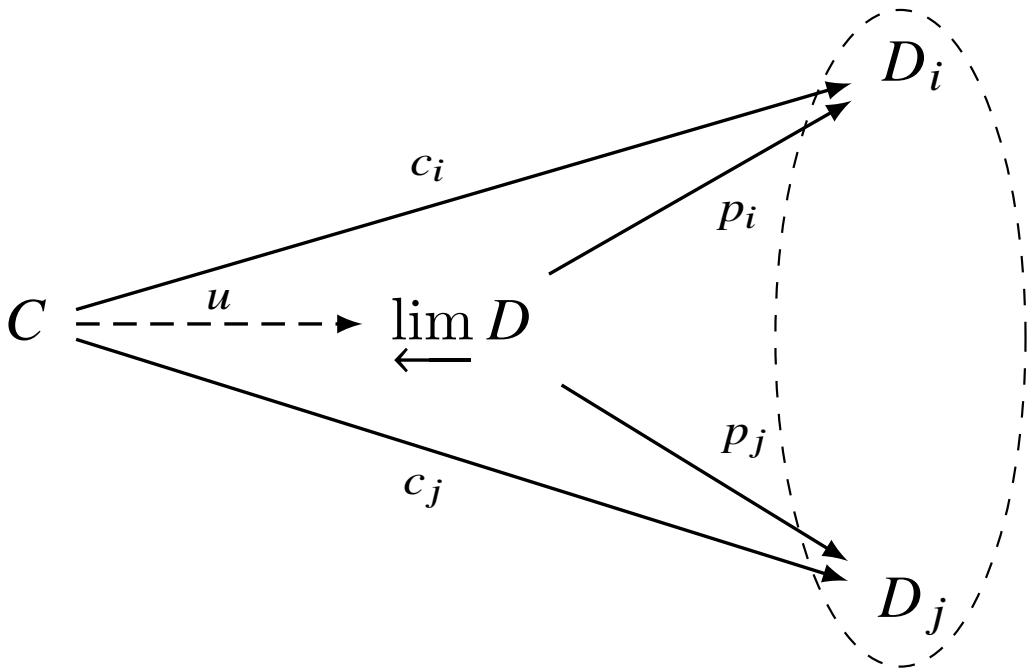
- A morphism of cocones $\theta : (C, c) \to (C', c')$ is a morphism θ s.t.



Then we have a category $\int \mathbf{Cone}(D, -)$ of all the cocones under D.

Limit

- A *limit* $\left(\varprojlim D, p\right)$ for a diagram $D : \mathbf{I} \to \mathbf{C}$ is a terminal object in $\int \mathbf{Cone}(-, D)$. In other word, for any cone (C, c) over D, there is a unique morphism $C \xrightarrow{u} \varprojlim D$ s.t. $c_i = p_i \circ u$ for all $i \in \mathbf{I}$.

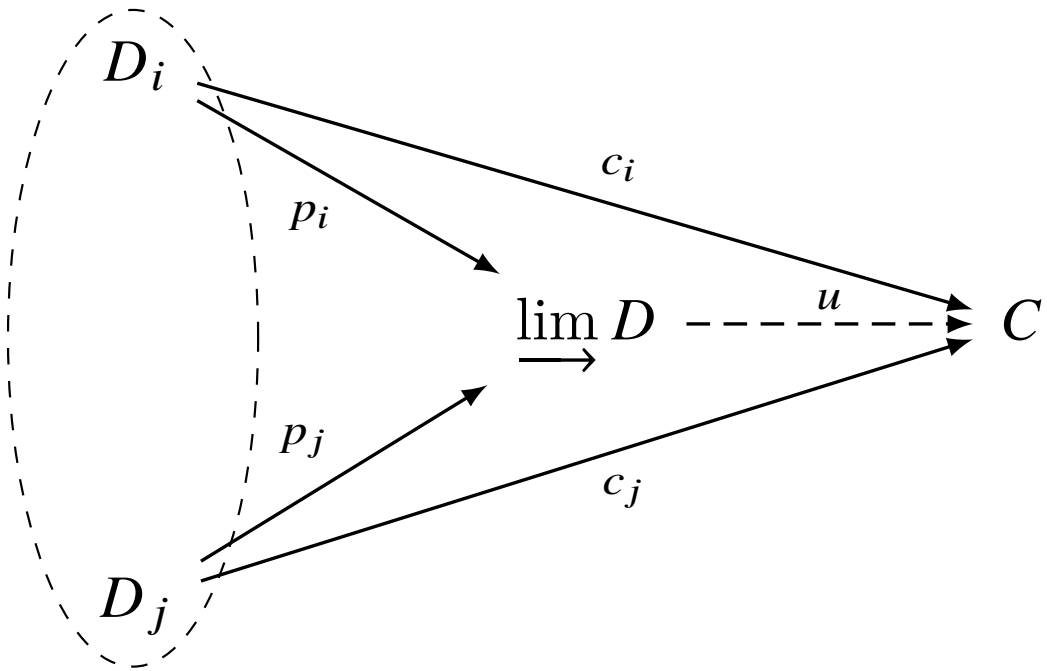


Remark: There is an isomorphism natural in C.

$$\mathbf{C}^{\mathbf{I}}(\Delta_C, D) \cong \mathbf{C}(C, \varprojlim D)$$

Colimit

- A *colimit* $\left(\varinjlim D, p\right)$ for a diagram $D : \mathbf{I} \to \mathbf{C}$ is an initial object in $\int \mathbf{Cone}(D, -)$. In other word, for any cocone (C, c) over D, there is a unique morphism $\varinjlim D \xrightarrow{u} C$ s.t. $c_i = u \circ p_i$ for all $i \in \mathbf{I}$.



Remark: There is an isomorphism natural in C.

$$\mathbf{C}^{\mathbf{I}}(D, \Delta_C) \cong \mathbf{C}(\varinjlim D, C)$$

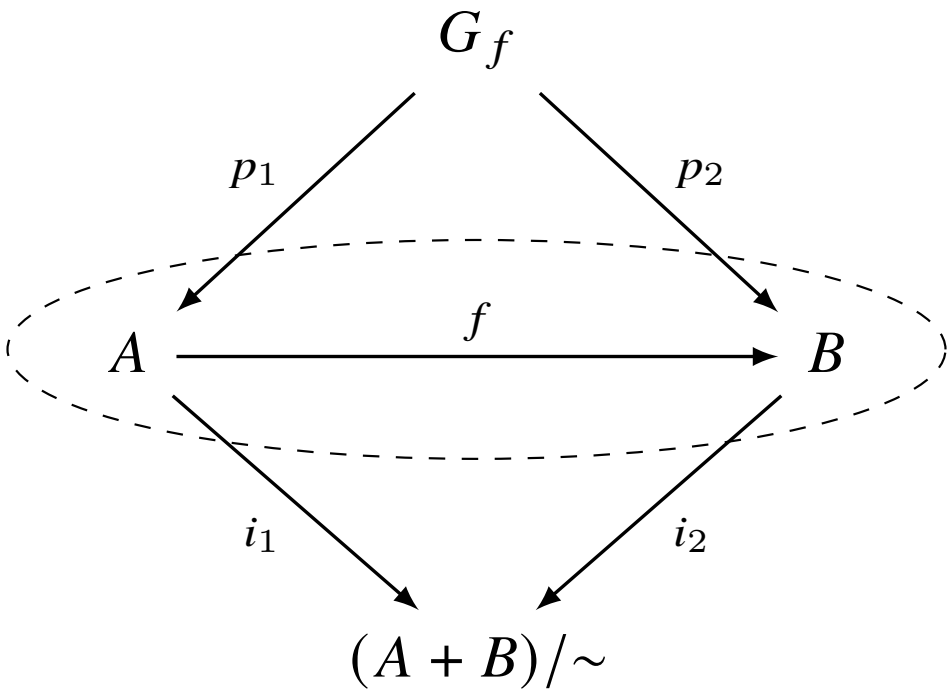
Example

In $\mathbf{Set}$, given a set function $f : A \to B$,

- the limit is the graph of f:

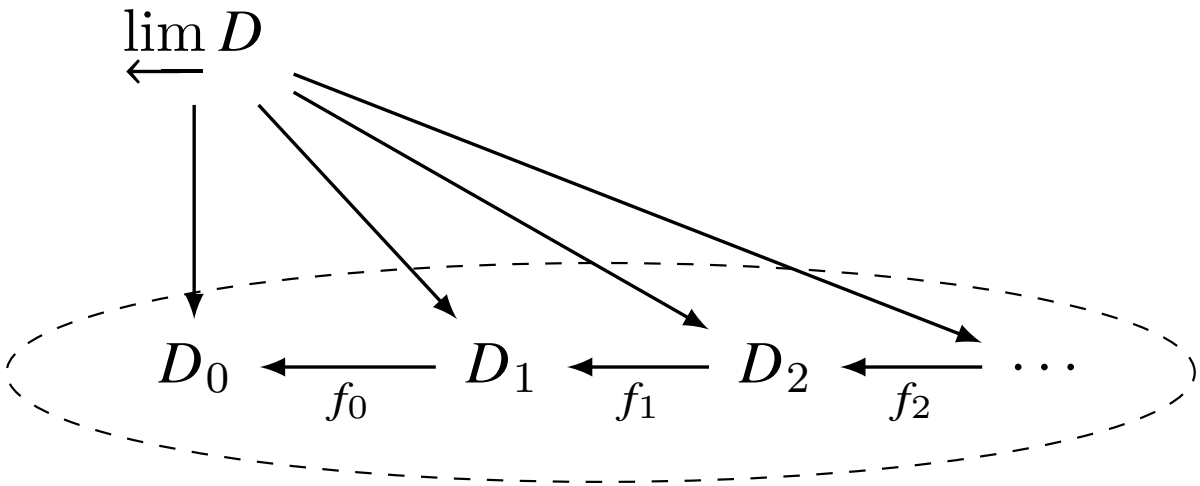
$$G_f = \{(x, fx) \in A \times B\}$$

- the colimit is $(A + B)/\sim$, where $x \sim fx$.

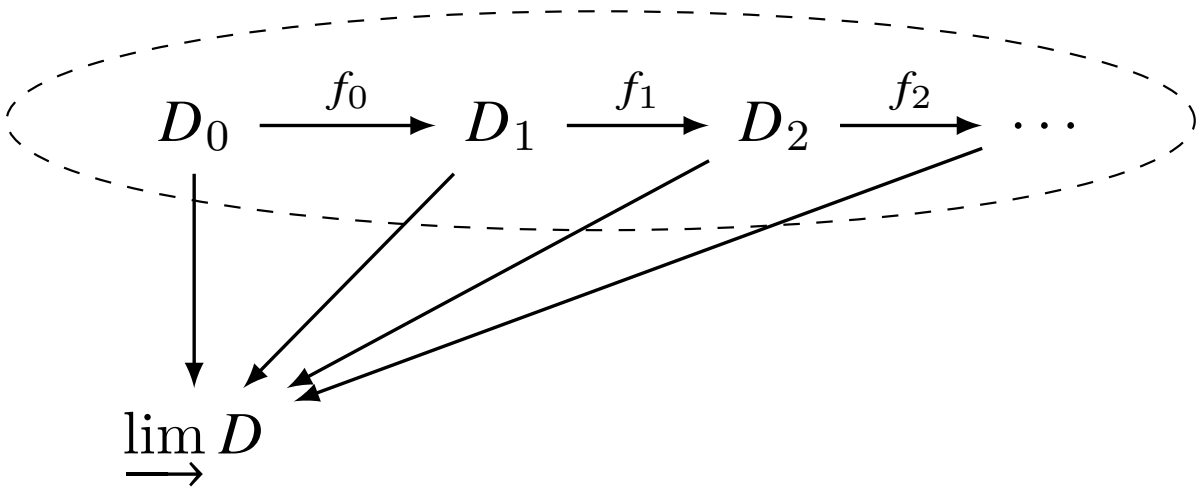


Inverse/Direct Limit

- Inverse Limit

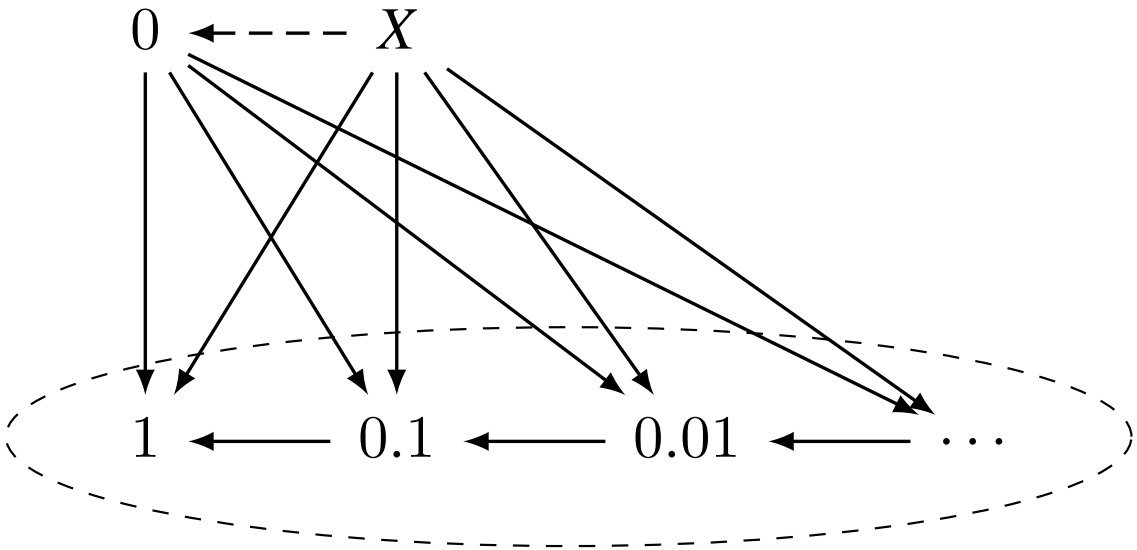


- Direct Limit

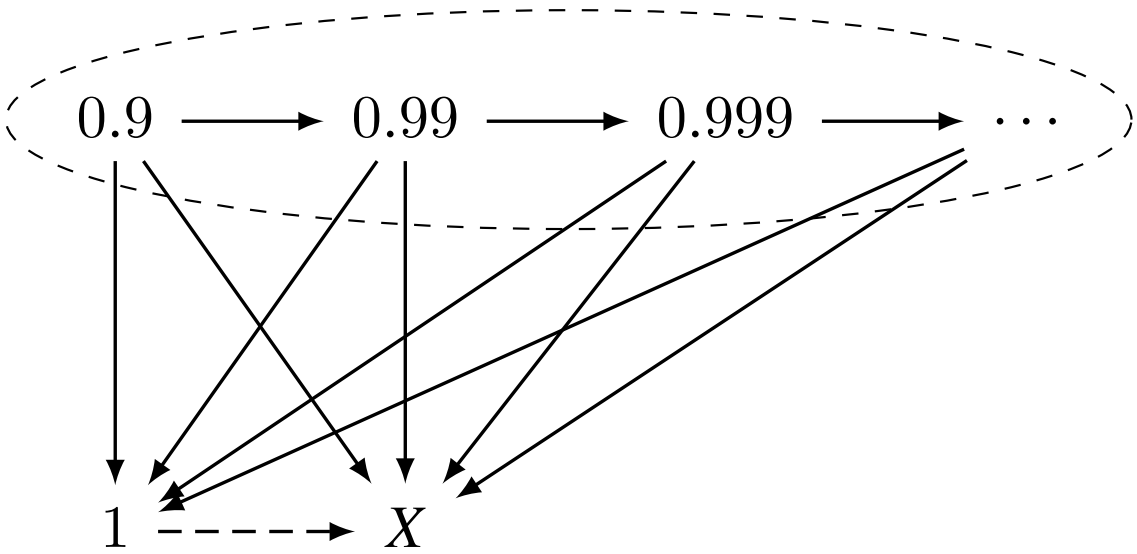


Inverse/Direct Limit — Example

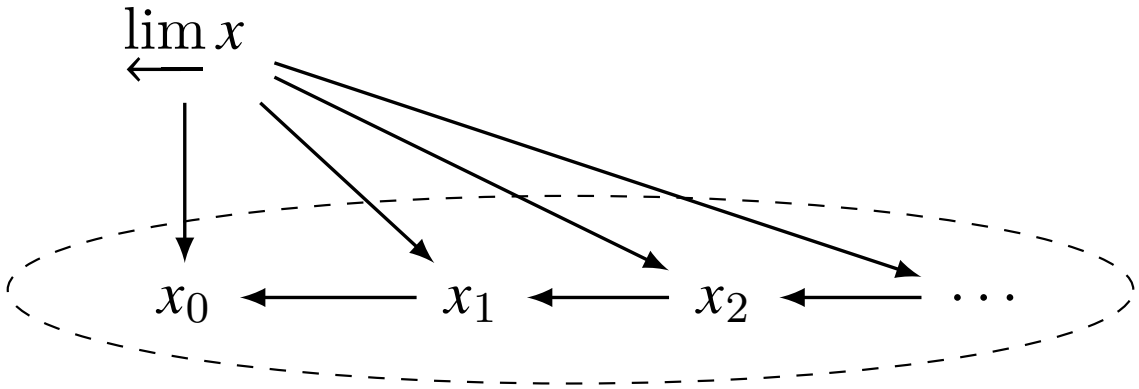
- Inverse Limit



- Direct Limit



Inverse/Direct Limit — Example

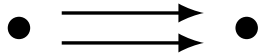
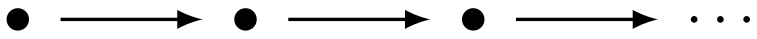
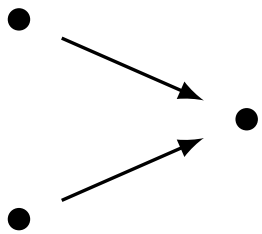
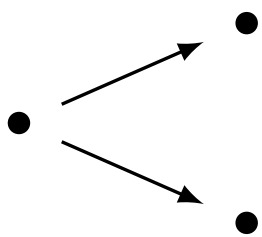


Let $\mathbb{R}$ be the category whose objects are real numbers and which has a unique morphism $x \longrightarrow y$ iff $x \leq y$. A functor $x : \mathbb{Z}_{\leq 0} \longrightarrow \mathbb{R}$ is a non-increasing sequence $(x_0, x_1, x_2 \dots)$ of real numbers. This functor has a inverse limit $\varprojlim x$ iff the sequence $(x_0, x_1, x_2 \dots)$ is bounded below and has a limit, in the classical "ε-δ" sense.

$$\lim_{n \to \infty} x_n = L \iff \forall \varepsilon > 0 \exists N \forall n \geq N (|x_n - L| < \varepsilon)$$

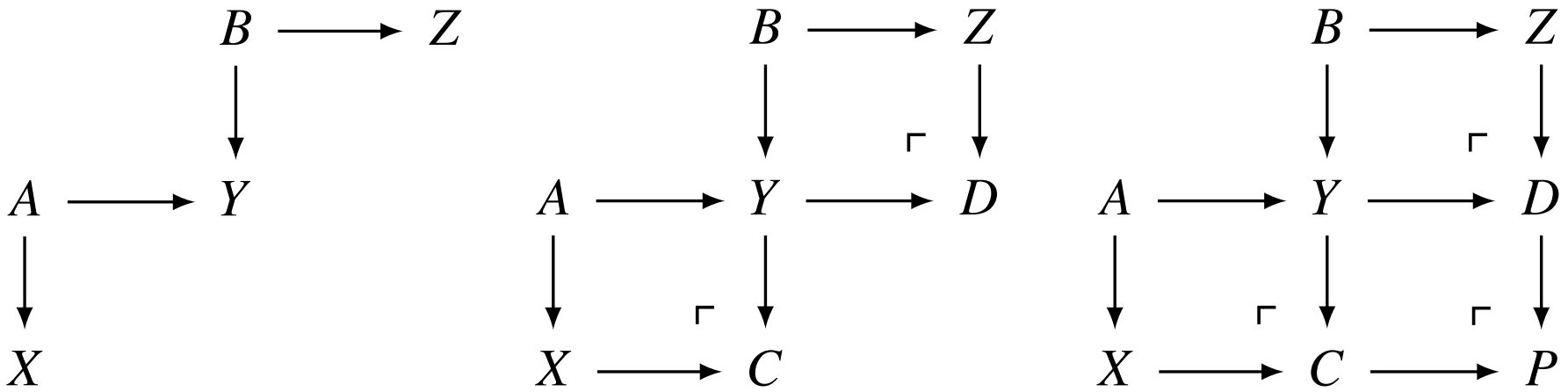
- The limit of $\cdots \twoheadrightarrow \mathbb{Z}/p^3 \twoheadrightarrow \mathbb{Z}/p^2 \twoheadrightarrow \mathbb{Z}/p$ in $\mathbf{AbGrp}$ is the group $\mathbb{Z}_p$ of p-adic integers.
- The colimit of $\mathbb{Z}/p \hookrightarrow \mathbb{Z}/p^2 \hookrightarrow \mathbb{Z}/p^3 \hookrightarrow \cdots$ is the group $\mathbb{Z}/p^\infty$.

Examples

Limit $\varprojlim D$	**I**	**Colimit** $\varinjlim D$
terminal object	$\emptyset$	initial object
binary product		binary coproduct
equalizor		coequalizor
inverse limit		
		direct limit
pullback		
		pushout

Exercise

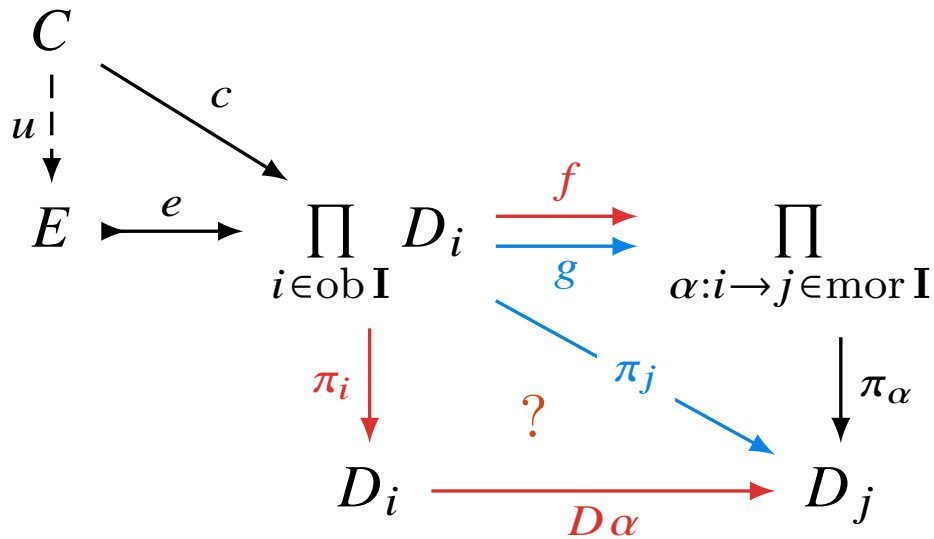
Suppose we want to take the colimit of the diagram.



We know how to take pushouts. The object P, together with all the morphisms from the original diagram to P is the colimit of the original diagram.

Products & Equalizers $\Longrightarrow$ Limits

Let $D : \mathbf{I} \to \mathbf{C}$. Let E be the equalizer of f and g such that $f_\alpha := \pi_\alpha \circ f = D\alpha \circ \pi_i$ and $g_\alpha := \pi_\alpha \circ g = \pi_j$ for $\alpha : i \to j \in \operatorname{mor} \mathbf{I}$.



but generally $D\alpha \circ \pi_i \neq \pi_j$

Take any $c : C \to \prod_{i \in \operatorname{ob}\mathbf{I}} D_i$. For $\alpha : i \to j$ in $\mathbf{I}$, we have

$$D\alpha \circ \pi_i \circ c = \pi_\alpha \circ f \circ c \quad \text{and} \quad \pi_j \circ c = \pi_\alpha \circ g \circ c$$

So (C, c_i) with $c_i := \pi_i \circ c$ is a cone of D iff $f \circ c = g \circ c$.

It follows that (E, e_i) with $e_i := \pi_i \circ e$ is a cone of D.

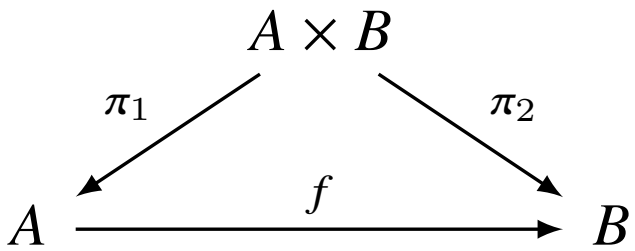
For any $c : C \to \prod_{i \in \operatorname{ob}\mathbf{I}} D_i$, there is a unique $u : C \to E$ s.t. $c = e \circ u$.

Then $u : C \to E$ is also the required factorization of the cone (C, c_i) through (E, e_i) s.t. $c_i = e_i \circ u$. Therefore $E = \varprojlim D$.

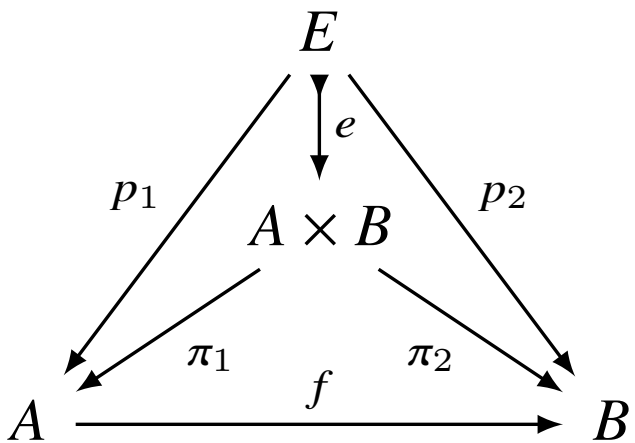
Example

Products & Equalizers $\implies$ Limits

For example: consider the product



There are two morphisms from $A \times B \to B$. Take the equalizer of π_2 and $f \circ \pi_1$.

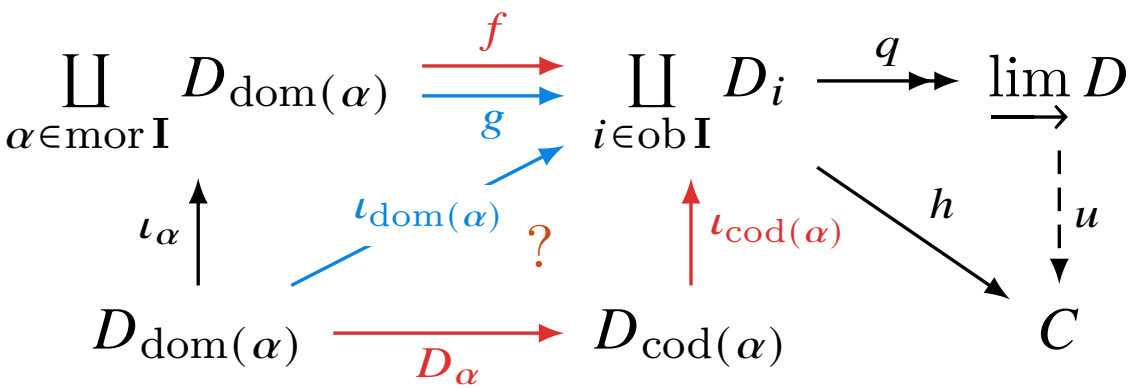


Then E is the limit of $f : A \to B$.

Coproducts & Coequalizers $\Longrightarrow$ Colimits

$$\varinjlim D = \operatorname{coeq}\left(\coprod_{\alpha \in \operatorname{mor} \mathbf{I}} D_{\operatorname{dom}(\alpha)} \underset{g}{\overset{f}{\rightrightarrows}} \coprod_{i \in \operatorname{ob} \mathbf{I}} D_i \right)$$

where the morphisms are determined by their components as follows:
$f_\alpha := f \circ \iota_\alpha = \iota_{\operatorname{cod}(\alpha)} \circ D_\alpha$ and $g_\alpha := g \circ \iota_\alpha = \iota_{\operatorname{dom}(\alpha)}$.



Theorem

The following are equivalent for a category $\mathbf{C}$*:*

- $\mathbf{C}$ *has all pullbacks and a terminal object.*
- $\mathbf{C}$ *has all equalizers and finite products.*
- $\mathbf{C}$ *has has finite limits.*

Theorem

The following are equivalent for a category $\mathbf{C}$*:*

- $\mathbf{C}$ *has all equalizers and small products.*
- $\mathbf{C}$ *has small limits.*

Theorem

A functor preserves finite (small) limits iff it preserves equalizers and finite (small) products.

Preserving/Reflecting/Creating Limits

- A functor $F : \mathbf{C} \to \mathbf{D}$ is said to *preserve limits of* $\mathbf{I}$ iff, for all diagrams $D : \mathbf{I} \to \mathbf{D}$ and all cones (C, c) over D,

  $$(C, c) \text{ is a limit over } D \implies (FC, Fc) \text{ is a limit over } FD : \mathbf{I} \to \mathbf{D}$$

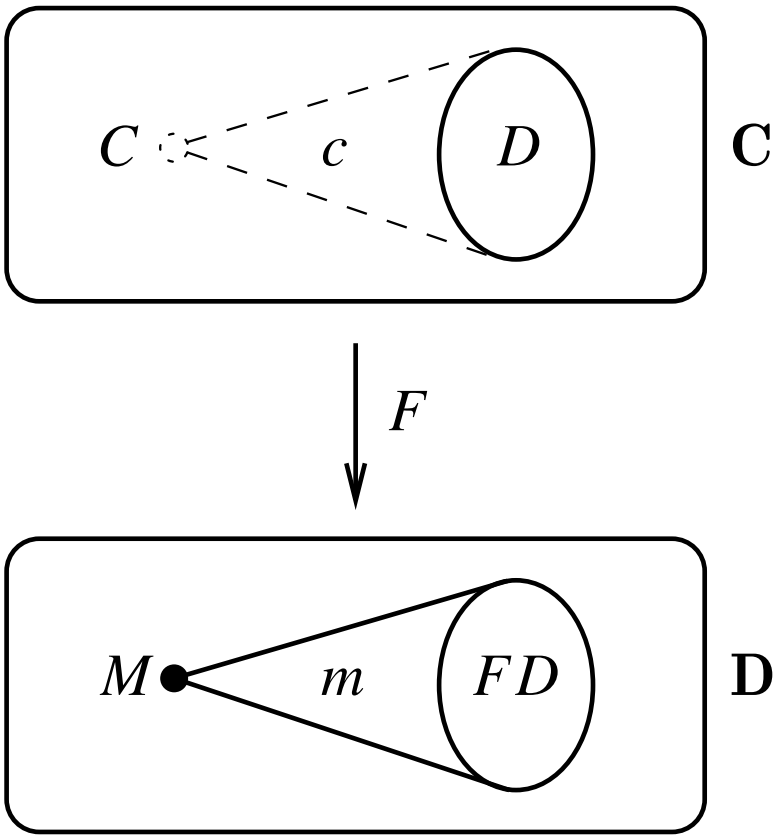
- A functor $F : \mathbf{C} \to \mathbf{D}$ is said to *reflect limits of* $\mathbf{I}$ iff, for all diagrams $D : \mathbf{I} \to \mathbf{D}$ and all cones (C, c) over D,

  $$(C, c) \text{ is a limit over } D \impliedby (FC, Fc) \text{ is a limit over } FD : \mathbf{I} \to \mathbf{D}$$

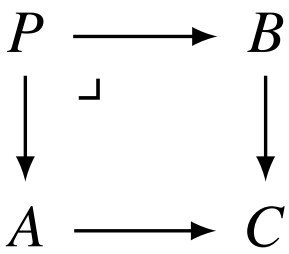
- A functor $F : \mathbf{C} \to \mathbf{D}$ is said to *create limits of* $\mathbf{I}$ iff, for all diagrams $D : \mathbf{I} \to \mathbf{C}$, if (M, m) is a limit cone over $FD : \mathbf{I} \to \mathbf{D}$, there is a unique cone (C, c) over D s.t. $(FC, Fc) = (M, m)$, and moreover (C, c) is a limit.
- A category is *complete* iff it has all small limits.
- A functor is *continuous* iff it preserves all small limits.

Obviously, if F preserves limits then $F\left(\varprojlim D\right) \cong \varprojlim(FD)$.

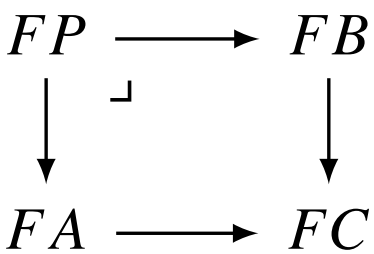
Creation of limits

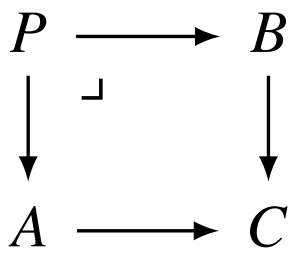


Examples

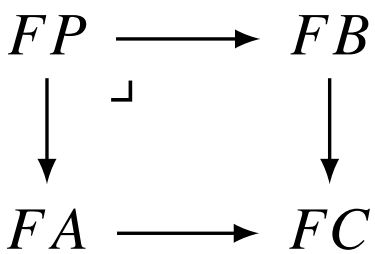


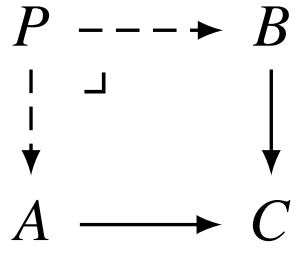
preserves $\Downarrow$



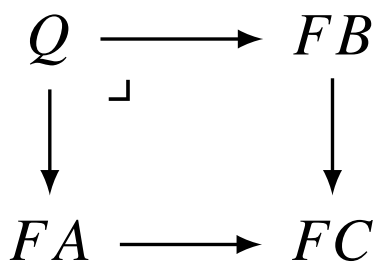


reflects $\Uparrow$





creates $\Uparrow$



Theorem

Consider a category $\mathbf{C}$ *such that, for every category* $\mathbf{I}$ *and every functor* $D : \mathbf{I} \to \mathbf{C}$*, the limit* $\varprojlim D$ *exists. Then,* $\mathbf{C}$ *is a preordered category.*

Proof.

Suppose there are $f, g : A \to B$ distinct.
Because all limits exists, there is an object $P := \prod_\kappa B$, the product of κ copies of B, where κ is the cardinality of the set of arrows in $\mathbf{C}$.
Thus, we can construct 2^κ distinct cones $(A, \{A \to B\}_\kappa)$ by combining f and g in every possible way.
Thus, we have 2^κ distinct factorizations $A \to P$, being P a product.
These factorizations are arrows, so we have that $2^\kappa \leq \kappa$, contradicting Cantor's Theorem. □

Remark: The result says that it makes little sense to consider categories which have all limits, for every "size" of the diagram category.

The Fibonacci Sequence as a Functor

$$F_0 \coloneqq 1$$
$$F_1 \coloneqq 1$$
$$F_n \coloneqq F_{n-1} + F_{n-2}$$

$$\gcd(F_m, F_n) = F_{\gcd(m,n)}$$

Fibonacci function $F : \mathbb{N} \to \mathbb{N}$ is a functor that preserves limits, which is to say it's a continuous functor from the finitely complete category $\mathbb{N}$ to itself.

Remarks

- Limits: A diagram D in a category $\mathbf{C}$ can be seen as a system of constraints, and then a limit of D represents all possible solutions of the system.
- Colimits: Given a species of structure, say widgets, then the result of interconnecting a system of widgets to form a super-widget corresponds to taking the colimit of the diagram of widgets in which the morphisms show how they are interconnected.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Construct New from Old

- opposite category
- subcategory
- product/coproduct category
- functor category
- slice/coslice category
- comma category
- arrow category
- quotient category

Free Category

Definition (Free Category)

The *free category* generated by a directed graph is the category that results from freely concatenating arrows together, whenever the target of one arrow is the source of the next.

Definition (Free Group)

There is a functor $F : \mathbf{Set} \to \mathbf{Grp}$ that sends a set X to the *free group* on X. Elements of FX are finite "words" whose letters are elements $x \in X$ or their formal inverses x^{-1}, modulo an equivalence relation that equates the words xx^{-1} and $x^{-1}x$ with the empty word. Multiplication is by concatenation, with the empty word serving as the identity.

Remark: Freedom is just another word for nothing left to lose.
— there are no relations between the generators of FX beyond the bare minimum required by the group axioms.(没有额外的方程约束)

Quotient Category

Definition (Quotient Category)

- Given a category $\mathbf{C}$. A congruence relation $\sim$ on $\mathbf{C}$ is given by: for each pair of objects $A, B \in \mathbf{C}$, an equivalence relation $\sim_{AB}$ on $\mathrm{Hom}(A, B)$, such that for $f_1, f_2 \in \mathrm{Hom}(A, B), g_1, g_2 \in \mathrm{Hom}(B, C)$:

$$f_1 \sim_{AB} f_2 \ \& \ g_1 \sim_{BC} g_2 \implies g_1 f_1 \sim_{AC} g_2 f_2$$

- Given a congruence relation $\sim$ on $\mathbf{C}$, we can define the *quotient category* $\mathbf{C}/\sim$ as the category whose objects $\mathrm{ob}(\mathbf{C}/\sim) = \mathrm{ob}(\mathbf{C})$ and whose morphisms are equivalence classes of morphisms in $\mathbf{C}$. That is,

$$\mathrm{Hom}_{\mathbf{C}/\sim}(A, B) = \mathrm{Hom}_{\mathbf{C}}(A, B)/\sim_{AB}$$

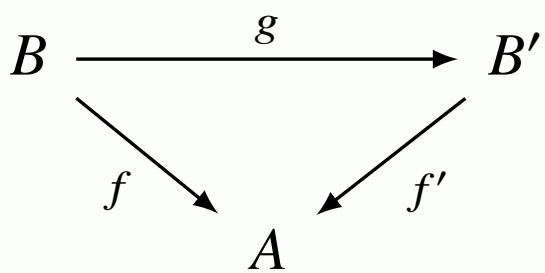
The composition in $\mathbf{C}/\sim$:

$$[g] \circ [f] = [g \circ f]$$

Slice Category

Definition (Slice Category)

Given a category $\mathbf{C}$ and $A \in \mathbf{C}$, the *slice category* $\mathbf{C}/A$ is a category whose objects are pairs (B, f) where $B \in \mathbf{C}$ and $f : B \to A$. A morphism of $\mathbf{C}/A$ from (B, f) to (B', f') is a morphism $g : B \to B'$ s.t.



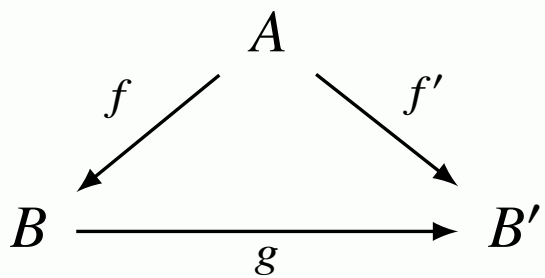
Example

Let $A := \{r, g, b\}$. The category $\mathbf{Set}/A$ is the category of A-colored sets. The morphisms are color-preserving functions.

Coslice Category

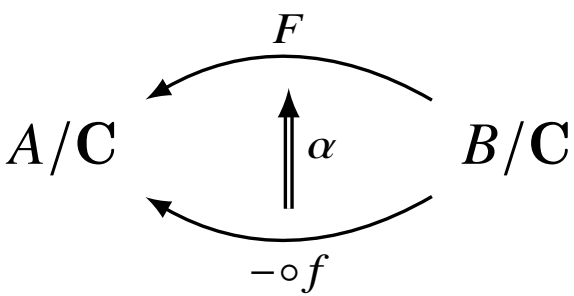
Definition (Coslice Category)

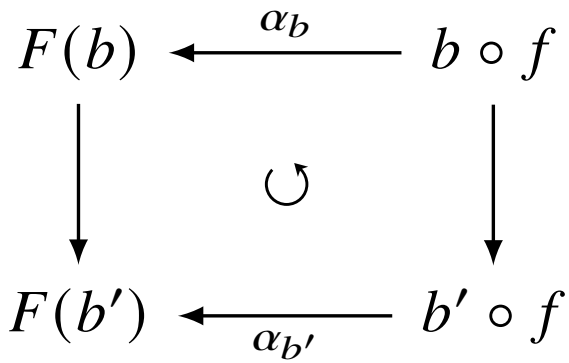
Given a category $\mathbf{C}$ and $A \in \mathbf{C}$, the *coslice category* $A/\mathbf{C}$ is a category whose objects are pairs (B, f) where $B \in \mathbf{C}$ and $f : A \to B$. A morphism of $A/\mathbf{C}$ from (B, f) to (B', f') is a morphism $g : B \to B'$ s.t.



Metaphor as Functor between Coslice Categories[FST20]

- Coslice category $A/\mathbf{C}$: the category of "meanings" of A.
- Metaphor "A is like B": a functor $F : B/\mathbf{C} \to A/\mathbf{C}$.
- "base-of-metaphor functor": $f/\mathbf{C} \coloneqq - \circ f : B/\mathbf{C} \to A/\mathbf{C}$.
- A new metaphor functor $F : B/\mathbf{C} \to A/\mathbf{C}$ is created by a natural transformation α from the base-of-metaphor functor $f/\mathbf{C}$.



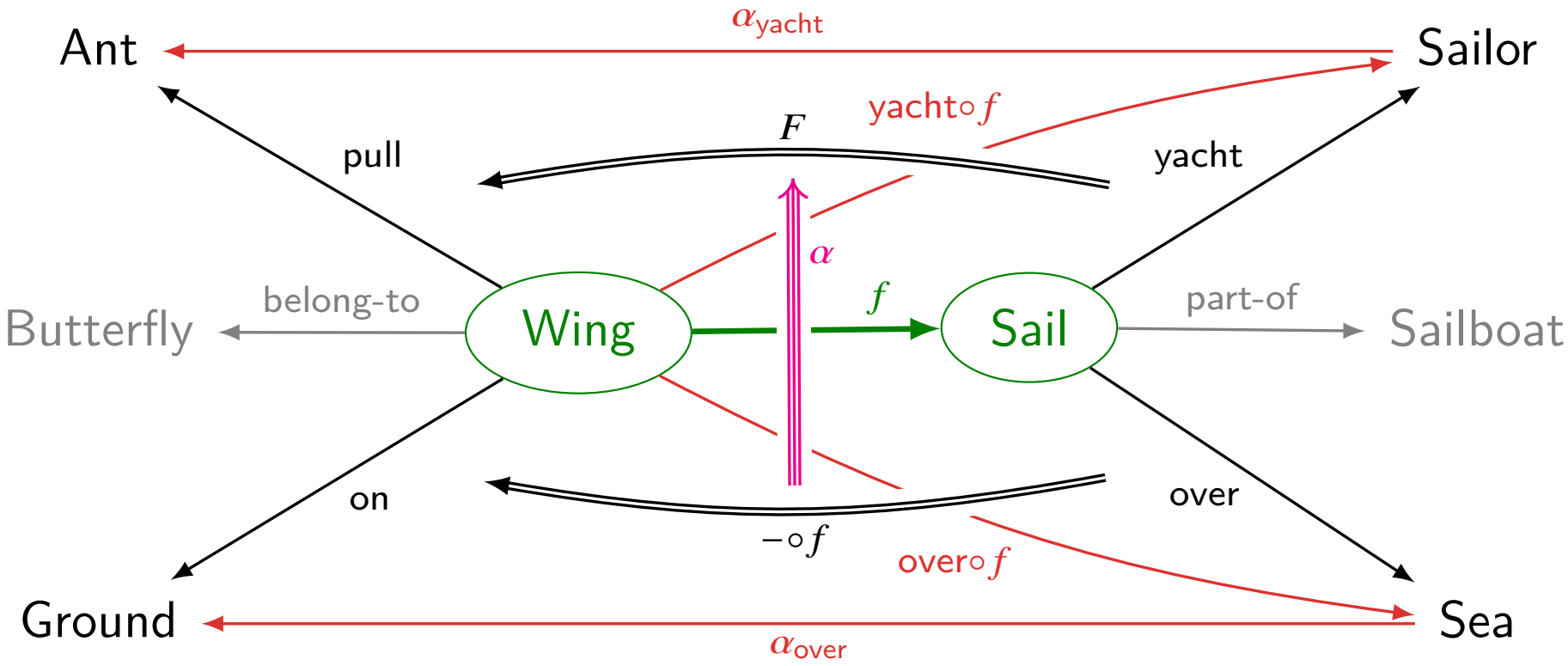


Metaphor: "A Wing is a Sail"

The Ground — A Japanese Poem

An ant
pull a wing of a butterfly.
O
It is like yachting.

$$\begin{array}{ccc} F(\text{yacht}) = \text{pull} & \xleftarrow{\alpha_{\text{yacht}}} & \text{yacht} \circ f \\ \downarrow & \circlearrowleft & \downarrow \\ F(\text{over}) = \text{on} & \xleftarrow[\alpha_{\text{over}}]{} & \text{over} \circ f \end{array}$$



Examples

- $\mathbf{Set}/I$ can be regarded as the category of "I-indexed families of sets".

$$\boxed{\mathbf{Set}/I \simeq \mathbf{Set}^I}$$

$$\Phi : \mathbf{Set}/I \to \mathbf{Set}^I$$

$$\Psi : \mathbf{Set}^I \to \mathbf{Set}/I$$

$$\Phi : A \xrightarrow{f} I \mapsto \left(f^{-1}(i)\right)_{i \in I}$$

$$\Psi : (A_i)_{i \in I} \mapsto \coprod_{i \in I} A_i \xrightarrow{\pi} I \quad \text{(the indexing projection)}$$

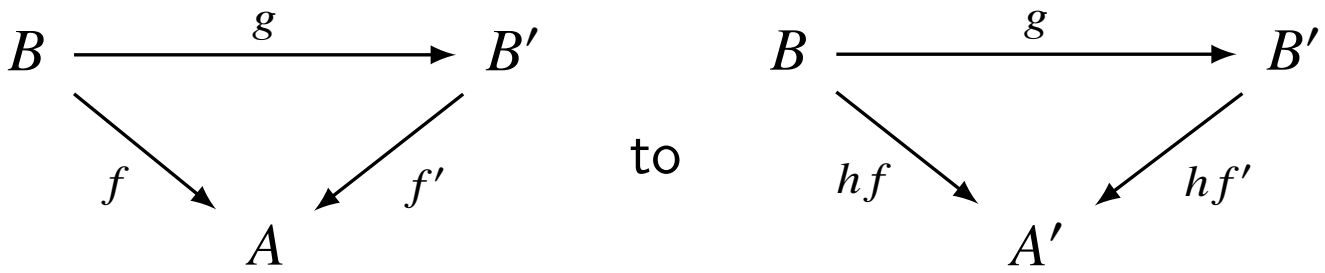
  where the coproduct is conveniently taken to be

$$\coprod_{i \in I} A_i := \bigcup_{i \in I} A_i \times \{i\}$$

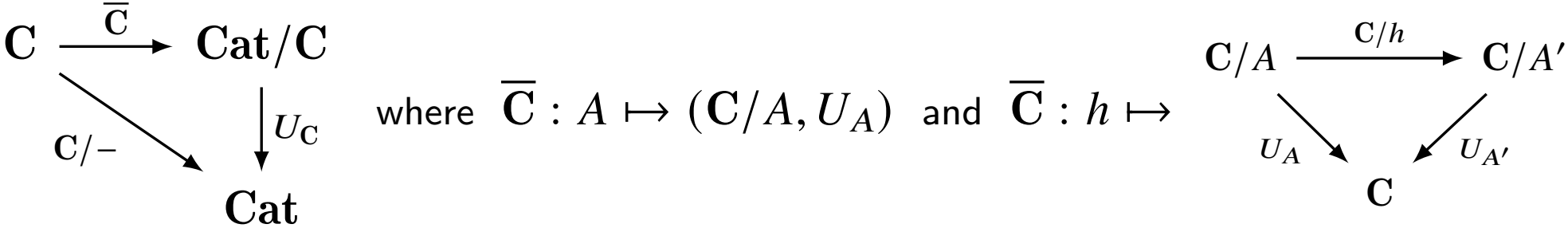
- $1/\mathbf{Set}$ (with $1 = \{\bullet\}$ a one-point set) is the category of pointed sets: objects are pairs (A, a) of sets with a distinguished element $a \in A$, and morphisms $f : (A, a) \to (B, b)$ must preserve this: $fa = b$.

Slice Category

- There is a forgetful functor $U_A : \mathbf{C}/A \to \mathbf{C}$ which maps (B, f) to B.
- Furthermore, for $h : A \to A'$ there is a functor "*composition by h*" $\mathbf{C}/h : \mathbf{C}/A \to \mathbf{C}/A'$ which maps (B, f) to (B, hf) and



- For any small category $\mathbf{C}$, the construction of slice categories itself is a functor $\mathbf{C}/- : \mathbf{C} \to \mathbf{Cat}$.
- The functor $\mathbf{C}/-$ then factors through the forgetful functor $U_{\mathbf{C}} : \mathbf{Cat}/\mathbf{C} \to \mathbf{Cat}$ via a functor $\overline{\mathbf{C}} : \mathbf{C} \to \mathbf{Cat}/\mathbf{C}$.



Arrow Category

Definition (Arrow Category)

Given a category $\mathbf{C}$, the *arrow category* $\mathbf{C}^{\to}$ has as objects the morphisms $f : A \to B$ of $\mathbf{C}$, and a morphism from $f : A \to B$ to $f' : A' \to B'$ is a pair $(a, b) : f \to f'$ where $a : A \to A', b : B \to B'$ s.t.

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle a}\downarrow & & \downarrow{\scriptstyle b} \\ A' & \xrightarrow[f']{} & B' \end{array}$$

The composition is given by

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle a}\downarrow & & \downarrow{\scriptstyle b} \\ A' & \xrightarrow{f'} & B' \\ {\scriptstyle a'}\downarrow & & \downarrow{\scriptstyle b'} \\ A'' & \xrightarrow{f''} & B'' \end{array}$$

Twisted Arrow Category

Definition (Twisted Arrow Category)

Given a category $\mathbf{C}$, the *twisted arrow category* $\mathbf{C}^{\to}$ has as objects the morphisms $f : A \to B$ of $\mathbf{C}$, and a morphism from $f : A \to B$ to $f' : A' \to B'$ is a pair $(a, b) : f \to f'$ where $a : A' \to A, b : B \to B'$ s.t.

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle a}\uparrow & & \downarrow{\scriptstyle b} \\ A' & \xrightarrow[f']{} & B' \end{array}$$

The composition is given by

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle a}\uparrow & & \downarrow{\scriptstyle b} \\ A' & \xrightarrow{f'} & B' \\ {\scriptstyle a'}\uparrow & & \downarrow{\scriptstyle b'} \\ A'' & \xrightarrow{f''} & B'' \end{array}$$

Example: $[x, y] \subset [x', y'] \iff x' \leq x \ \ \& \ \ y \leq y'$

Comma Category

Definition (Comma Category)

Given $\mathbf{A} \xrightarrow{F} \mathbf{C} \xleftarrow{G} \mathbf{B}$, we can form the *comma category* $F \downarrow G$ as follows:

- the objects are triples (A, B, f) with $A \in \mathbf{A}, B \in \mathbf{B}$ and $f : FA \to GB$.
- the morphisms from (A, B, f) to (A', B', f') are pairs (a, b) where $a : A \to A'$ in $\mathbf{A}$ and $b : B \to B'$ in $\mathbf{B}$ s.t.

$$\begin{array}{ccc} FA & \xrightarrow{f} & GB \\ {\scriptstyle Fa}\downarrow & & \downarrow{\scriptstyle Gb} \\ FA' & \xrightarrow[f']{} & GB' \end{array}$$

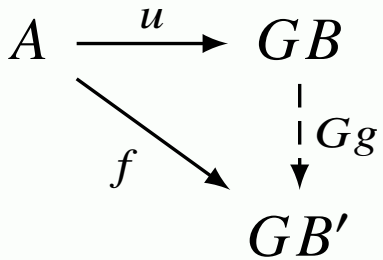
- If $\mathbf{B} = \mathbf{1}$ and $G : \mathbf{1} \to \mathbf{C}$ picks out the object X and $\mathbf{A} = \mathbf{C}$ with $F = 1_{\mathbf{C}}$, then the comma category $F \downarrow G$ is the slice category $\mathbf{C}/X$.
- If $\mathbf{A} = \mathbf{1}$ and $F : \mathbf{1} \to \mathbf{C}$ picks out the object X and $\mathbf{B} = \mathbf{C}$ with $G = 1_{\mathbf{C}}$, then the comma category $F \downarrow G$ is the coslice category $X/\mathbf{C}$.
- If $\mathbf{A} = \mathbf{B} = \mathbf{C}$, then the comma category $1_{\mathbf{C}} \downarrow 1_{\mathbf{C}}$ is the arrow category $\mathbf{C}^{\to}$.

Universal Property

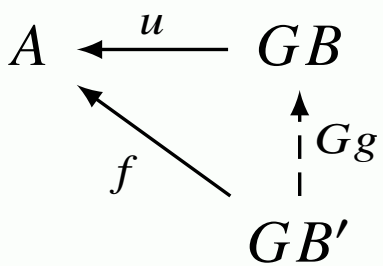
Definition (Universal Property)

Let $G : \mathbf{D} \to \mathbf{C}$ be a functor, and $A \in \mathbf{C}, B \in \mathbf{D}$.

- A universal morphism from A to G is a unique pair (B, u) where $u : A \to GB$ with the following property: for any $f : A \to GB'$, there exists a unique morphism $g : B \to B'$ s.t.



- A universal morphism from G to A is a unique pair (B, u) where $u : GB \to A$ with the following property: for any $f : GB' \to A$, there exists a unique morphism $g : B' \to B$ s.t.

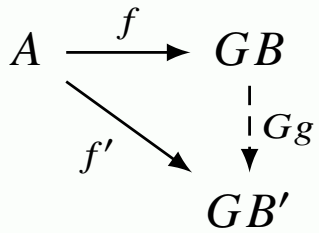


Comma Category

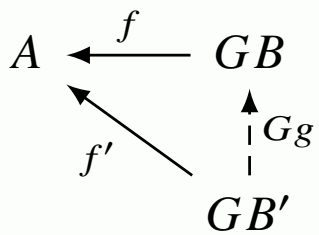
Definition (Comma Category)

Let $G : \mathbf{D} \to \mathbf{C}$ be a functor, and $A \in \mathbf{C}$.

- The *comma category* $A \downarrow G$ has objects all pairs (B, f) with $B \in \mathbf{D}$ and $f : A \to GB$. A morphism from (B, f) to (B', f') is given by $g : B \to B'$ s.t.



- The *comma category* $G \downarrow A$ has objects all pairs (B, f) with $B \in \mathbf{D}$ and $f : GB \to A$. A morphism from (B, f) to (B', f') is given by $g : B' \to B$ s.t.



- A universal morphism from A to G is an initial object of $A \downarrow G$.
- A universal morphism from G to A is a terminal object of $G \downarrow A$.

Example

考虑所有度量空间 (X, d) 构成的范畴 $\mathbf{Metr}$, 其中的态射是保距映射 $f : (X, d_X) \to (Y, d_Y)$, 即: $\forall u, v,\ d_Y(f(u), f(v)) = d_X(u, v)$. 完备度量空间构成的全子范畴记作 $\mathbf{ComMetr}$. Cauchy 完备化构造给出一个函子

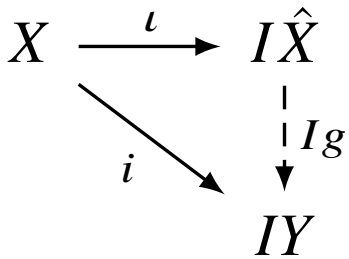
$$C : \mathbf{Metr} \to \mathbf{ComMetr}$$

$$(X, d) \mapsto (\hat{X}, \hat{d})$$

其中 $\hat{X}$ 是所有 X 中的 Cauchy 列 $\vec{x} = (x_n)_{n\geq 0}$ 的等价类, 而 $\hat{d}(\vec{x}, \vec{y}) := \lim_{n\to\infty} d(x_n, y_n)$. 令 $I : \mathbf{ComMetr} \to \mathbf{Metr}$ 为包含函子. 对于给定的度量空间 X, 嵌入 $x \mapsto (x_n := x)_{n\geq 1}$ 给出态射 $\iota : X \to I\hat{X}$.

Comma 范畴 $X \downarrow I$ 的对象形如 $(Y, i : X \to IY)$.

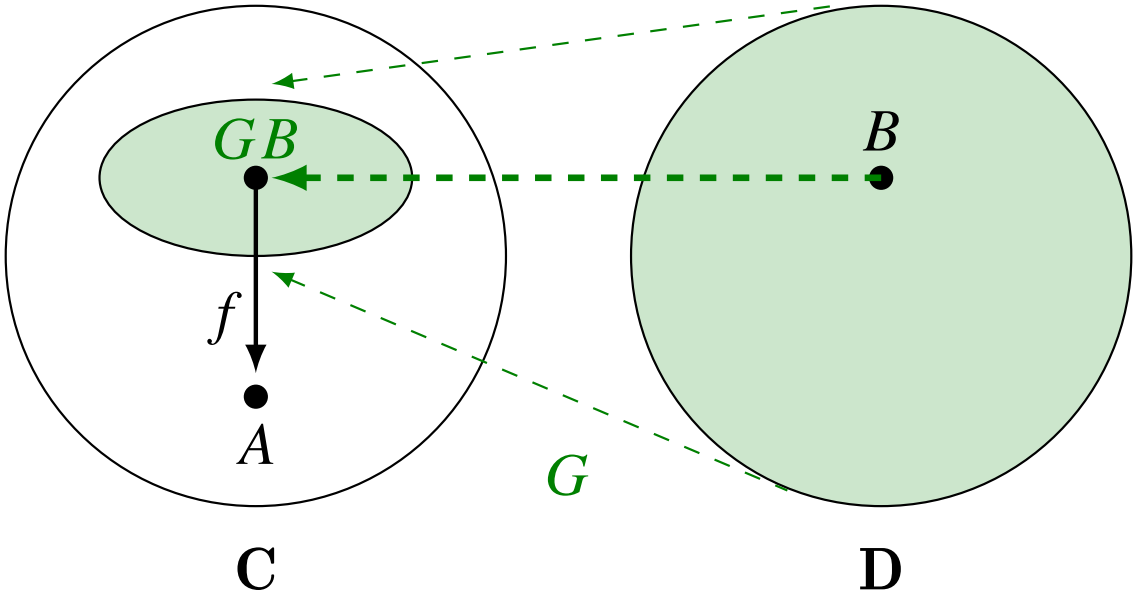
完备化 $\hat{X}$ 的泛性质: 对任意 $(Y, i) \in \mathrm{ob}(X \downarrow I)$, 存在唯一的态射 $g : \hat{X} \to Y$ 使得



交换. 这归结为 X 在 $\hat{X}$ 中的稠密性.

因此完备化 $(\hat{X}, \iota : X \to I\hat{X})$ 可以刻画为 $X \downarrow I$ 的始对象.

Comma Category $G \downarrow A$



Remark: When dealing with an adjunction:

$$\mathbf{C}(GB, A) \cong \mathbf{D}(B, FA)$$

we are observing the object A from a narrower perspective defined by the functor G. Think of G as defining a model of the category $\mathbf{D}$ inside $\mathbf{C}$. We are interested in the view of A from the perspective of this model. The arrows that describe this view form the comma category $G \downarrow A$.

Comma Category

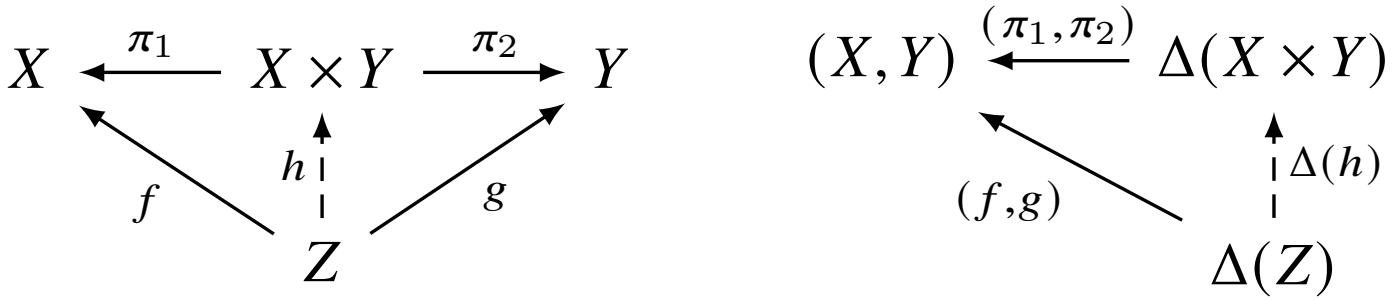
Given $\mathbf{1} \xrightarrow{A} \mathbf{C} \xleftarrow{G} \mathbf{D}$, the $A \downarrow G$-objects are $(\bullet, B, f)$ with $B \in \mathbf{D}$ and $f : A \to GB$ in $\mathbf{C}$. The $A \downarrow G$-morphisms from $(\bullet, B, f)$ to $(\bullet, B', f')$ are pairs $(1_\bullet, g)$ with $g : B \to B'$ in $\mathbf{D}$ s.t.

$$\begin{array}{ccc} A & \xrightarrow{f} & GB \\ {\scriptstyle 1_A}\downarrow & & \downarrow{\scriptstyle Gg} \\ A & \xrightarrow[f']{} & GB' \end{array}$$

- Let $G : \mathbf{D} \to \mathbf{C}$ be a functor and let A be an object of $\mathbf{C}$.
  The following statements are equivalent:
  1. (B, u) is a universal morphism from A to G.
  2. (B, u) is an initial object of the comma category $A \downarrow G$.
  3. (B, u) is a representation of $\mathrm{Hom}(A, G(-))$.
- The dual statements are also equivalent:
  1. (B, u) is a universal morphism from G to A.
  2. (B, u) is a termianl object of the comma category $G \downarrow A$.
  3. (B, u) is a representation of $\mathrm{Hom}(G(-), A)$.
- A universal element can be viewed as a universal morphism from the one-point set $\{\bullet\}$ to the functor $G : \mathbf{D} \to \mathbf{Set}$.

Universal Property — Product

Let X and Y be objects of a category $\mathbf{D}$.



Take $\mathbf{C}$ to be the product category $\mathbf{D} \times \mathbf{D}$, and let Δ be the diagonal functor $\Delta : \mathbf{C} \to \mathbf{C}^{\mathbf{I}}$.
Then $(X \times Y, (\pi_1, \pi_2))$ is a universal morphism from Δ to the object (X, Y) of $\mathbf{D} \times \mathbf{D}$.
One can generalize the above example to arbitrary limits and colimits.

- Given $D : \mathbf{I} \to \mathbf{C}$ (thought of as an object in $\mathbf{C}^{\mathbf{I}}$), the limit $\varprojlim D$ is a universal morphism from Δ to D.
- Dually, the colimit $\varinjlim D$ is a universal morphism from D to Δ.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Subobject Classifier

Definition (Subobject)

A *subobject* of an object $X \in \mathbf{C}$ is a monomorphism $m : M \rightarrowtail X$.

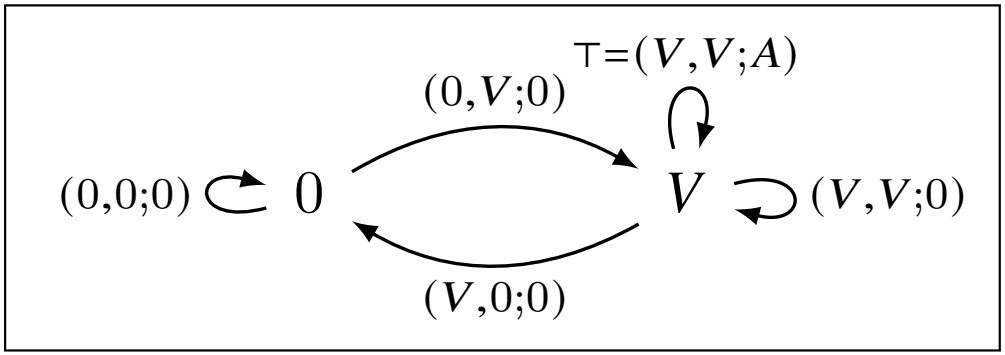
Definition (Subobject Classifier)

Let $\mathbf{E}$ be a category with all finite limits. A *subobject classifier* in $\mathbf{E}$ is a monomophism $\top : 1 \rightarrowtail \Omega$ such that for every monomorphism $m : M \rightarrowtail X$, there is a unique $\varphi : X \to \Omega$ making the following diagram a pullback:

$$\begin{array}{ccc} M & \xrightarrow{!} & 1 \\ {\scriptstyle m}\downarrow & \lrcorner & \downarrow{\scriptstyle \top} \\ X & \xrightarrow[\varphi]{} & \Omega \end{array}$$

Remark: The arrow $\top \circ !$: $M \xrightarrow{!} 1 \xrightarrow{\top} \Omega$ is often denoted as $\top_M : M \to \Omega$.

Subobject Classifier in $\mathbf{Graph}$

$\Omega_{\mathbf{Graph}} =$ 

- There are 5 arrows and 2 dots in $\Omega_{\mathbf{Graph}}$, which represent the various degrees of truth that a statement may have.
  - For dots:
    — The dot is in the subgraph.
    — The dot is not in the subgraph.
  - For arrows:
    — The arrow is included in the subgraph.
    — The arrow is not in the subgraph, but its source and its target are.
    — The arrow is not in the subgraph and neither is its source, but its target is.
    — The arrow is not in the subgraph and neither is its target, but its source is.
    — The arrow is not in the subgraph and neither is its source nor its target.

Remark

- If a category has a subobject classifier, it will be balanced.
- **Pos** is not balanced. Hence it can't have a subobject classifier.
- **Grp** is balanced, but it doesn't have a subobject classifier.

Topos

Definition (Topos)

A *topos* is a category which

1. is cartesian closed,
2. has all finite limits,
3. has a subobject classifier.

Remark

A topos is a category satisfying one of these equivalent conditions:

1. *it is a complete category with exponentials and subobject classifier,*
2. *it is a complete category with subobject classifier and its power object,*
3. *it is a cartesian closed category with equalizers and subobject classifier.*

Topos — Examples

- $\mathbf{Set}$ is a topos.
- $\mathbf{FinSet}$ is a topos.
- If $\mathbf{E}$ is a topos and $X \in \mathbf{E}$, then $\mathbf{E}/X$ is a topos.
- The arrow category $\mathbf{Set}^{\to}$ is a topos.
- If $\mathbf{E}$ is a topos, then $\mathbf{E} \times \mathbf{E}$ is a topos.
- For any small category $\mathbf{C}$, the category of all presheaves $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ is a topos.

Example — Presheaf

The category of presheaves $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ on $\mathbf{C}$, is the functor category whose objects are contravariant functors $\mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$ and whose maps are natural transformations between them.

$$\begin{aligned} \top(A) &= \{\bullet\} \\ (F \times G)A &= FA \times GA \\ G^F A &= \mathrm{Nat}(\mathbf{C}(-, A) \times F, G) \end{aligned} \qquad \text{(Yoneda)}$$

where $G^F \times F \xrightarrow{\varepsilon} G$ is defined by

$$\varepsilon_A(\alpha, a) = \alpha_A(1_A, a) \text{ for } a \in FA$$

Power Object

Definition (Power Object)

Let $\mathbf{C}$ be a category with finite limits. A *power object* of an object $A \in \mathbf{C}$ is an object $\mathrm{P}(A)$ with a monomorphism $\in_A \rightarrowtail A \times \mathrm{P}(A)$ such that, for every object B and every monomorphism $R \rightarrowtail A \times B$ there is a unique morphism $\chi_R : B \to \mathrm{P}(A)$ such that R is the pullback

$$\begin{array}{ccc} R & \longrightarrow & \in_A \\ \downarrow & \lrcorner & \downarrow \\ A \times B & \xrightarrow[1_A \times \chi_R]{} & A \times \mathrm{P}(A) \end{array}$$

Theorem

Any topos $\mathbf{E}$ has power objects.

Proof.

For $A \in \mathbf{E}$, let $\mathrm{P}(A) := \Omega^A$, and let $\in_A \rightarrowtail A \times \Omega^A$ be the subobject of $A \times \Omega^A$ whose character is $\varepsilon : A \times \Omega^A \to \Omega$.

$$\begin{array}{ccc} \in_A & \xrightarrow{!} & 1 \\ \downarrow & \lrcorner & \downarrow \top \\ A \times \Omega^A & \xrightarrow[\varepsilon]{} & \Omega \end{array}$$

Take any monic $R \rightarrowtail A \times B$, and let $f : A \times B \to \Omega$ be its character. Then let $\hat{f} : B \to \Omega^A$ be the unique morphism s.t. $\varepsilon \circ (1_A \times \hat{f}) = f$.

$$\begin{array}{ccc} A \times B & \xrightarrow{1_A \times \hat{f}} & A \times \Omega^A \\ & \searrow^{f} & \downarrow \varepsilon \\ & & \Omega \end{array}$$

$$\begin{array}{ccccc} R & \dashrightarrow & \in_A & \xrightarrow{!} & 1 \\ \downarrow & \lrcorner & \downarrow & \lrcorner & \downarrow \top \\ A \times B & \xrightarrow[1_A \times \hat{f}]{} & A \times \Omega^A & \xrightarrow[\varepsilon]{} & \Omega \end{array} \qquad \overset{!}{\frown} \quad \underset{f}{\smile}$$

Logical Morphism

Definition (Logical Morphism)

A *logical morphism* between toposes is a functor which preserves the topos structure, that is: finite limits, exponentials and the subobject classifier.

- The inclusion $\mathbf{FinSet} \hookrightarrow \mathbf{Set}$ is logical.
- For any small category $\mathbf{C}$ the inclusion $\mathbf{FinSet}^{\mathbf{C}^{\mathrm{op}}} \hookrightarrow \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ is logical.
- A logical morphism preserves finite colimits.
- A logical morphism has a left adjoint iff it has a right adjoint.
- Let $\mathbf{E}$ be a topos and $X \in \mathbf{E}$. Then $\mathbf{E}/X$ is a topos, and the functor $X^* : \mathbf{E} \to \mathbf{E}/X$ is logical.
- For a morphism $f : X \to Y$ in a topos $\mathbf{E}$, the pullback $f^* : \mathbf{E}/Y \to \mathbf{E}/X$ is logical, and it has a left adjoint $\sum_f$ and a right adjoint $\prod_f$.

Geometric Morphism

Definition (Geometric Morphism)

A *geometric morphism* $f : \mathbf{E} \to \mathbf{F}$ between toposes is a pair of adjoint functors (f^*, f_*): $\mathbf{F} \underset{f_*}{\overset{f^*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{E}$, such that the left adjoint f^* preserves finite limits (which implies that f_* preserves colimits).

We say that f_* is the *direct image*, and f^* is the *inverse image*.
Definition: The category of toposes and their geometric morphisms is denoted $\mathbf{Topoi}$.

Definition (Essential Geometric Morphism)

A geometric morphism (f^*, f_*) is *essential* if f^* has a left adjoint $f_! : \mathbf{E} \to \mathbf{F}$.

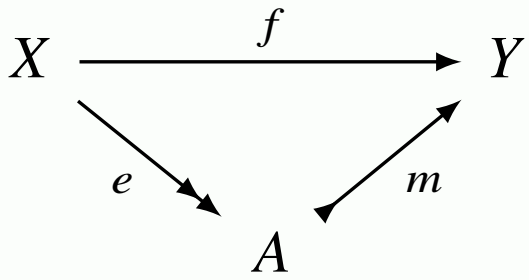
Theorem

A logical morphism is the direct image of a geometric morphism iff it is an equivalence.

Epi-Mono Factorization

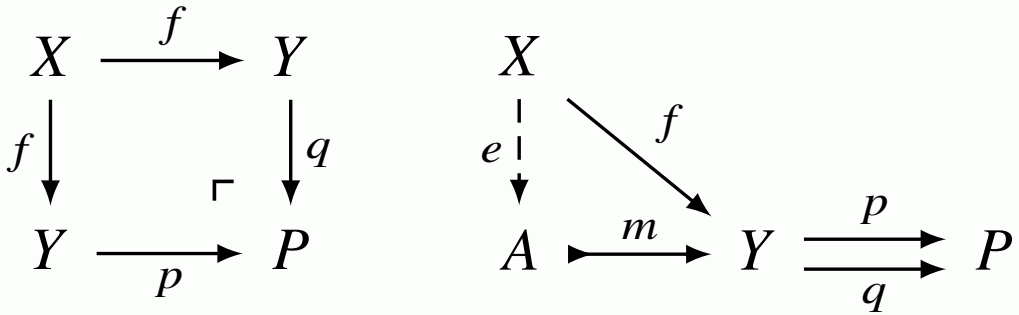
Theorem

In a topos, any morphism $f : X \to Y$ has an epi-mono factorization i.e. there exists an epi $e : X \to A$ and a mono $m : A \to Y$ s.t.



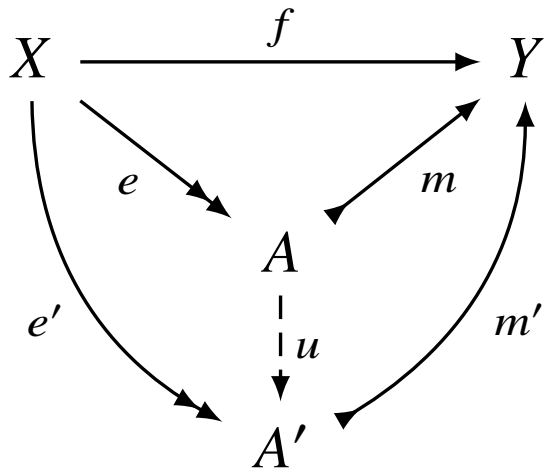
Proof.

Given $f : X \to Y$, let (p, q) be the pushout of (f, f) and $m : A \to Y$ the equalizer of (p, q). The universality of the equalizer gives a unique arrow $e : X \to A$ s.t. $f = m \circ e$. This is an epi-mono factorization of f.

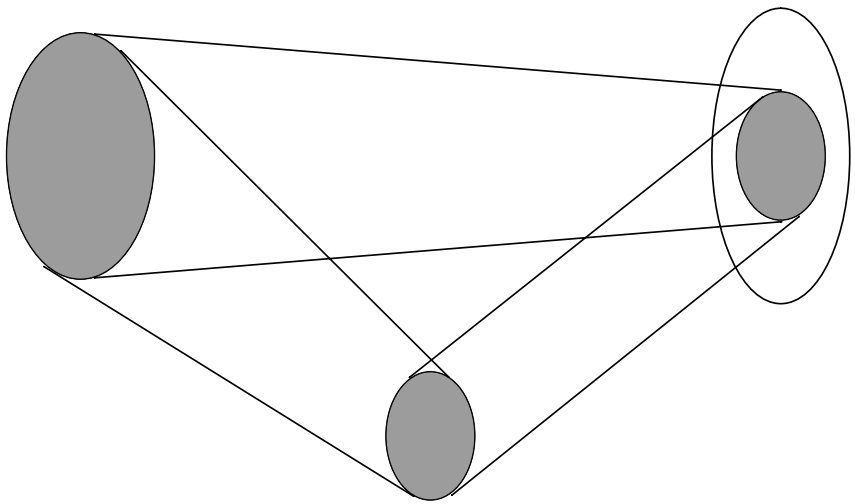


Epi-Mono Factorization

The epi-mono factorization is unique up to a unique isomorphism.

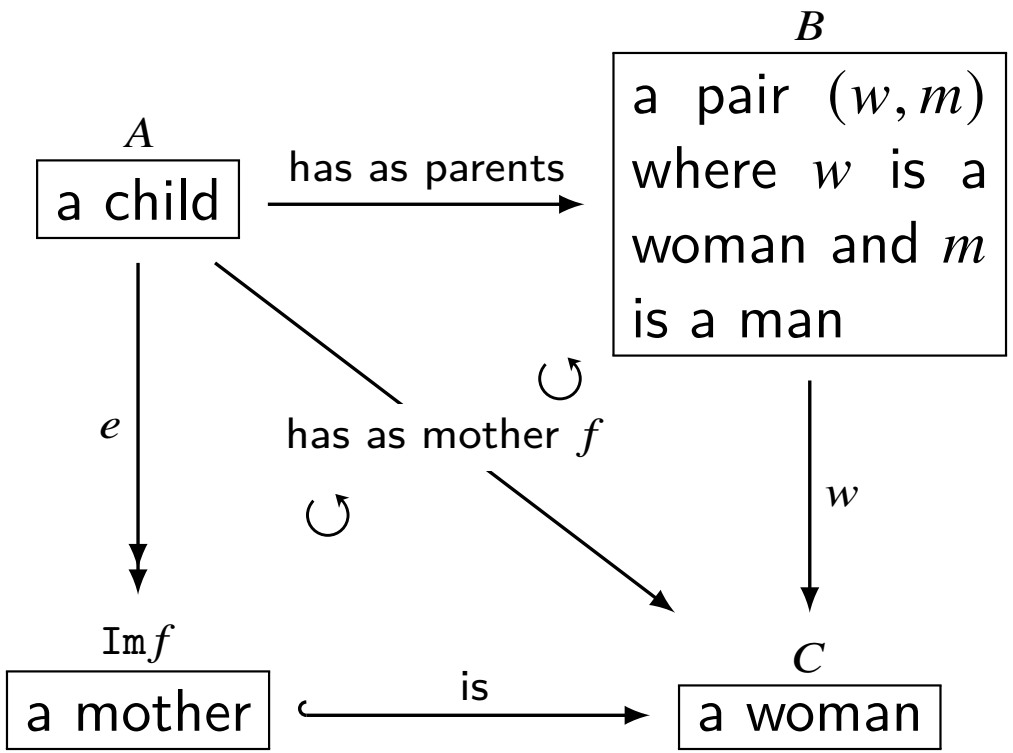


Given an epi-mono factorization, we call A the image of f, and denote it as $\mathtt{Im} f$.



Example — Image

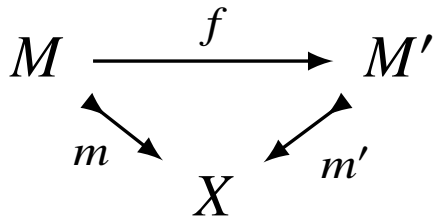
- The image of the map ⌜a person⌝ $\xrightarrow{\text{has as mother}}$ ⌜a person⌝ is the type ⌜a mother⌝.
- A mother is defined to be a person x for which there is some other person y such that x is the mother of y.



$\subset$ and $\in$

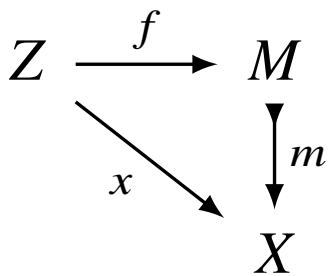
- Given subobjects m and m', define $m \subset_X m'$ (or $M \subset_X M'$) as follows

$$m \subset_X m' := \exists f : m \to m' \in \mathbf{C}/X$$



- $m \sim m' := m \subset_X m' \;\&\; m' \subset_X m$
- $\mathrm{Sub}(X) := \{[m] : m \text{ is monic with } \mathrm{cod}(m) = X\}$
- $[m] \subset_X [m'] := m \subset_X m'$
- In terms of 'generalized elements'[29] of an object X, $x : Z \to X$, one can define a *local membership* relation,

$$x \in_X M := \exists f : Z \to M[x = mf]$$



[29] When the domain is the terminal object, then $x : 1 \to X$ is the 'global element' of X.

Subobject Functor

Definition (Subobject Functor)

Let $\mathbf{C}$ be a category with pullbacks. Then $\mathbf{C}$ determines a contravariant functor $\mathrm{Sub} : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$ defined as

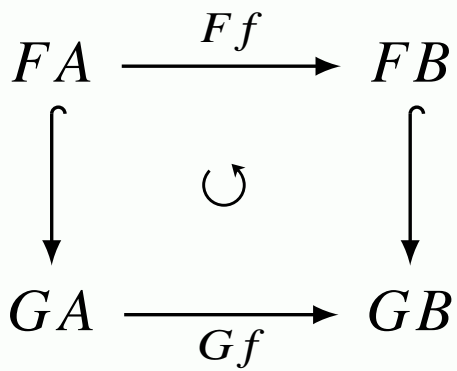
- for $X \in \mathbf{C}$, $\mathrm{Sub}(X) = \{M \in \mathbf{C} : M \text{ is an subobject of } X\}$.
- for $f \in \mathrm{Hom}(X, Y)$, $\mathrm{Sub}(f) : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$ which assigns to $m : M \rightarrowtail Y$ the arrow $f^*m : f^*M \rightarrowtail X$ defined by the pullback:

$$\begin{array}{ccc} f^*M & \longrightarrow & M \\ {\scriptstyle f^*m}\downarrow & \lrcorner & \downarrow{\scriptstyle m} \\ X & \xrightarrow[f]{} & Y \end{array}$$

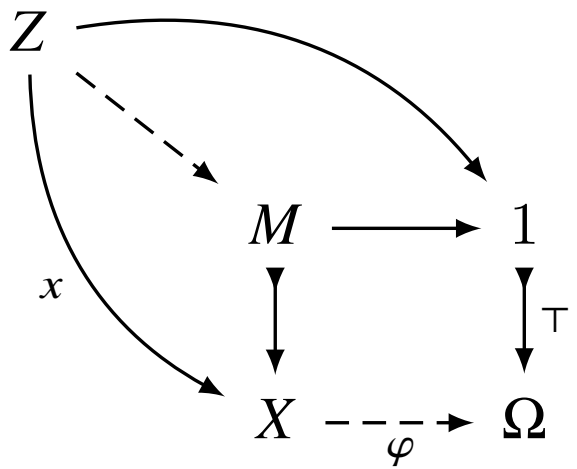
Subfunctor

Definition (Subfunctor)

A functor $F : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$ is a subfunctor of $G : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$, denoted by $F \subset G$, iff for any $f : B \to A$ in $\mathbf{C}^{\mathrm{op}}$,



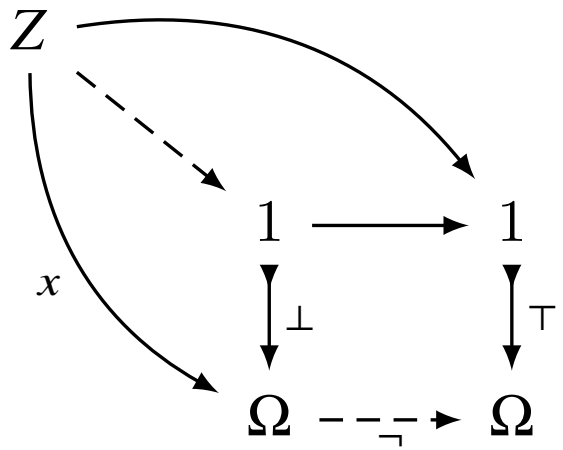
Predicate



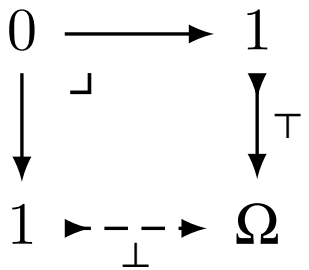
$$\varphi x = \top \iff x \in_X M$$

Remark: A predicate φ is the character of subobject M of universe X.

$\bot$ and $\neg$



where $\bot$ is defined by the pullback,

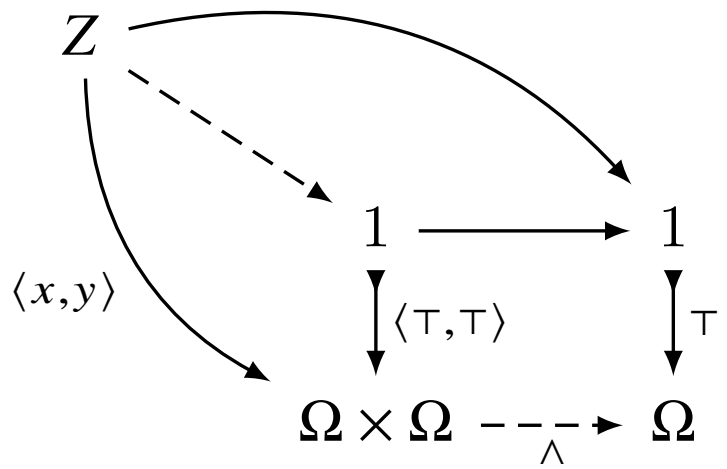


$$\neg x = \top \iff x = \bot$$

Remark

- $\bot : 1 \to \Omega$ *is the character of the unique arrow* $0 \to 1$.
- $\neg : \Omega \to \Omega$ *is the character of* $\bot : 1 \to \Omega$.

$\wedge$

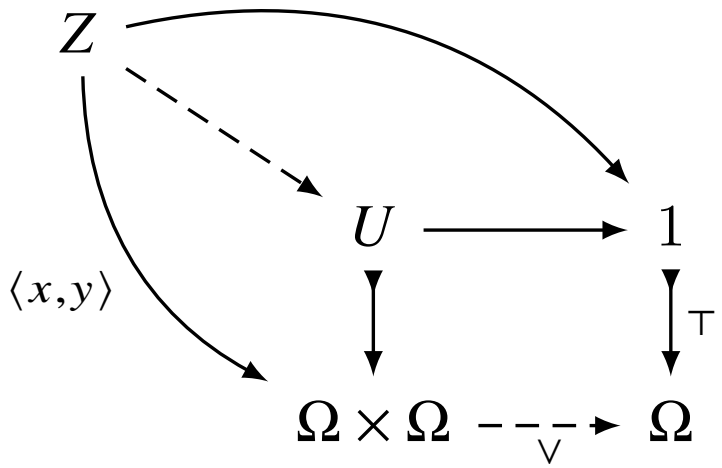


$$x \wedge y = \top \iff \langle x, y \rangle = \langle \top, \top \rangle$$

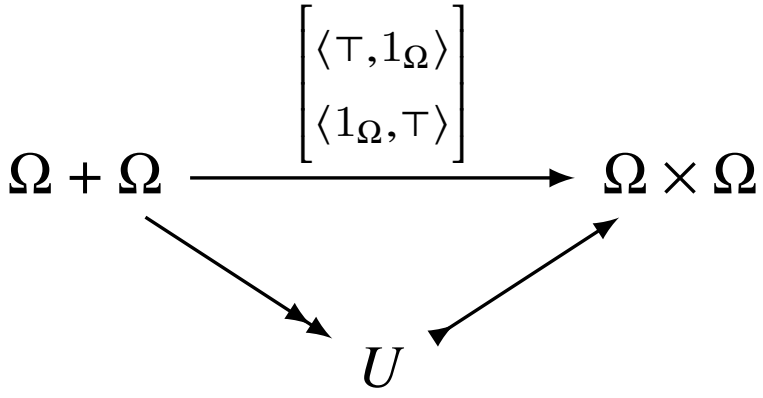
Remark

- $\wedge : \Omega \times \Omega \to \Omega$ *is the character of* $\langle \top, \top \rangle : 1 \to \Omega \times \Omega$.

$\vee$



where

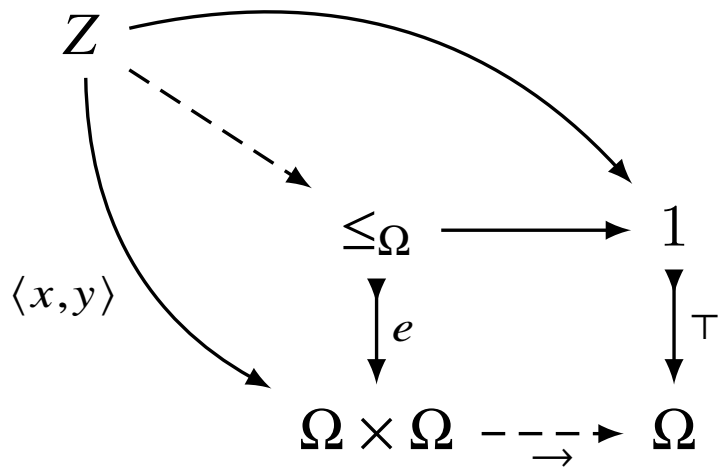


$$x \vee y = \top \iff \langle x, y \rangle \in_{\Omega \times \Omega} U$$

Remark

- $\vee : \Omega \times \Omega \to \Omega$ *is the character of the image (the mono part of the epi-mono factorization) of the arrow* $\begin{bmatrix} \langle \top, 1_\Omega \rangle \\ \langle 1_\Omega, \top \rangle \end{bmatrix} : \Omega + \Omega \to \Omega \times \Omega$.

$\rightarrow$



where $\leq_\Omega$ is the equaliser of $\Omega \times \Omega \underset{\pi_1}{\overset{\wedge}{\rightrightarrows}} \Omega$.

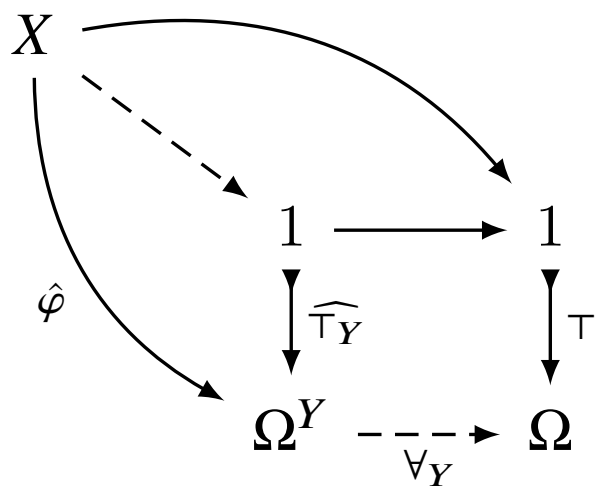
$$\leq_\Omega \xrightarrow{e} \Omega \times \Omega \underset{\pi_1}{\overset{\wedge}{\rightrightarrows}} \Omega$$

$$x \rightarrow y = \top \iff x \wedge y = x$$

Remark

- $\rightarrow : \Omega \times \Omega \rightarrow \Omega$ *is the character of the equalizer* $e : \leq_\Omega \rightarrow \Omega \times \Omega$ *of the arrows* $\wedge, \pi_1 : \Omega \times \Omega \rightarrow \Omega$.

$\forall$



where $\hat{\varphi} : X \to \Omega^Y$ is the currying of $X \times Y \xrightarrow{\varphi} \Omega$, and $\widehat{\top_Y}$ is the currying of the composite $1 \times Y \xrightarrow{!} 1 \xrightarrow{\top} \Omega$.

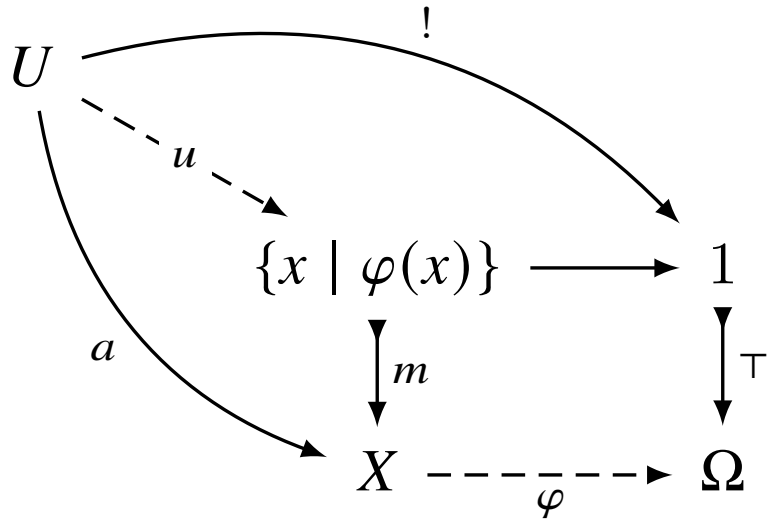
Obviously, $\forall_Y \circ \hat{\varphi} = \top_X \iff \varepsilon \circ (\hat{\varphi} \times 1_Y) = \top_{X \times Y}$

Theorem

Let $\varphi : X \times Y \to \Omega$ and $\psi : X \to \Omega$ be morphisms in $\mathbf{E}$. Then

$$\frac{\psi \circ \pi_1 \subset_{X \times Y} \varphi}{\psi \subset_X \forall_Y \circ \hat{\varphi}}$$

Kripke-Joyal Semantics



Definition (Kripke-Joyal Forcing)

$$U \vDash \varphi(a) \iff \varphi \circ a = \top_U \iff \exists u : m \circ u = a$$

Remark: "The formula $\varphi(a)$ holds at stage U."

Proposition

- *If* $f : V \to U$ *and* $U \vDash \varphi(a)$, *then* $V \vDash \varphi(a \circ f)$.
- *If* $f : V \twoheadrightarrow U$ *is epic and* $V \vDash \varphi(a \circ f)$, *then* $U \vDash \varphi(a)$.

Kripke-Joyal Semantics

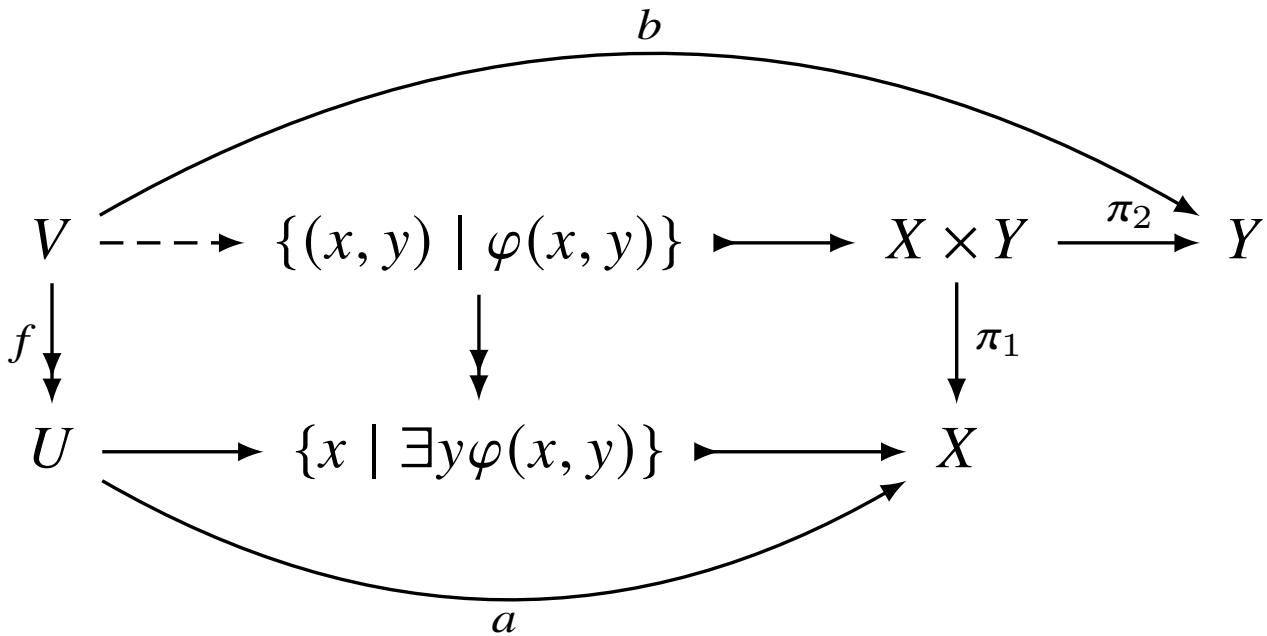
Theorem

Let $\mathbf{E}$ *be an elementary topos and* $a : U \to X$ *a generalized element of* $X \in \mathbf{E}$*, and* $\varphi(x), \psi(x)$ *formula with a free variable* x *of sort* X*. Then*

- $U \vDash \varphi(a) \wedge \psi(a)$ *iff* $U \vDash \varphi(a)$ *and* $U \vDash \psi(a)$*.*
- $U \vDash \varphi(a) \vee \psi(a)$ *iff there are morphisms* $f : V \to U$ *and* $g : W \to U$ *s.t.* $\begin{bmatrix} f \\ g \end{bmatrix} : V + W \twoheadrightarrow U$ *is epic and* $V \vDash \varphi(a \circ f)$ *and* $W \vDash \varphi(a \circ g)$*.*
- $U \vDash \varphi(a) \to \psi(a)$ *iff for any* $f : V \to U$*,* $V \vDash \varphi(a \circ f)$ *implies* $V \vDash \psi(a \circ f)$*.*
- $U \vDash \neg\varphi(a)$ *iff for any* $f : V \to U$*,* $V \vDash \varphi(a \circ f)$ *implies* $V \cong 0$*.*

If $\varphi(x, y)$ *has an additional free variable* y *of sort* Y*, then*

- $U \vDash \exists y \varphi(a, y)$ *iff there exists an epic* $f : V \twoheadrightarrow U$ *and a generalized element* $b : V \to Y$ *s.t.* $V \vDash \varphi(a \circ f, b)$*.*
- $U \vDash \forall y \varphi(a, y)$ *iff for every object* V*, every morpism* $f : V \to U$*, and every generalized element* $b : V \to Y$*,* $V \vDash \varphi(a \circ f, b)$*.*


b
V
{(x, y) | φ(x, y)}
X × Y
π2
Y
f
π1
U
{x | ∃yφ(x, y)}
X
a

Theorem

If $\mathbf{C}$ *is a locally small category and has finite limits, then it has a subobject classifier iff* $\mathrm{Sub} : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set} :: X \mapsto \{M \rightarrowtail X\}/\sim$ *is representable, with representing object* Ω*.*

$$\mathrm{Sub}_{\mathbf{C}}(X) \cong \mathrm{Hom}_{\mathbf{C}}(X, \Omega)$$

$$\mathrm{Sub}_{\mathbf{Set}}(X) \cong \mathtt{P}(X)$$

$$\mathrm{Sub}_{\mathbf{Set}}(1) : \boxed{0 \longrightarrow 1}$$

Theorem

In any topos $\mathbf{E}$*, for any object* $X \in \mathbf{E}$*,* $(\mathrm{Sub}_{\mathbf{E}}(X), \subset)$ *is a Heyting algebra: it is a poset that has finite limits, finite colimits, and is cartesian closed.*

Definition (Boolean Topos)

A topos $\mathbf{E}$ is *Boolean* iff for any $X \in \mathbf{E}$, $(\mathrm{Sub}_{\mathbf{E}}(X), \subset)$ is a Boolean algebra.

Example — $\mathrm{Sub}_{\mathbf{Graph}}(1)$

- What is $\mathrm{Sub}_{\mathbf{Graph}}(1)$ of the category of graphs?
- The two maps of sets f_A, f_D must be injective; each of the sets arrows and dots must either be empty or have one single element.

$$\begin{array}{ccc} \text{Arrows} & \xrightarrow{f_A} & 1 \\ {\scriptstyle s}\downarrow\downarrow{\scriptstyle t} & & \downarrow\downarrow \\ \text{Dots} & \xrightarrow[f_D]{} & 1 \end{array}$$

- Thus every subgraph of the terminal graph is isomorphic to one of these:

$$0 = \emptyset \qquad D = \bullet \qquad 1 = \bullet \circlearrowright$$

- Here the graph D represents an intermediate 'truth-value' which can be interpreted as 'true for dots but false for arrows'.

$$\mathrm{Sub}_{\mathbf{Graph}}(1) : \boxed{0 \longrightarrow D \longrightarrow 1}$$

$\mathrm{Sub}(X)$

Theorem

Given $X \in \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$, $(\mathrm{Sub}(X), \subset)$ *is a Heyting algebra.*

Proof.

For subfunctors F, G of $X \in \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ and $A, B \in \mathbf{C}$,

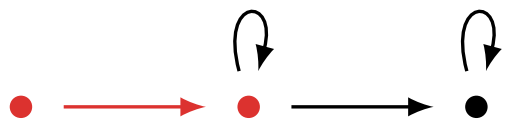
$$
\begin{aligned}
(F \wedge G)(A) &:= FA \cap GA \\
(F \vee G)(A) &:= FA \cup GA \\
1 &:= X \\
0 &:= \Delta_{\emptyset} \\
(\neg F)(A) &:= \left\{ a \in X(A) : \forall f : B \to A \left(Ffa \notin FB \right) \right\} \\
(F \to G)(A) &:= \left\{ a \in X(A) : \forall f : B \to A \left(Ffa \in FB \implies Ffa \in GB \right) \right\}
\end{aligned}
$$

Remark: $\neg F \vee F$ may not be the same as X. □

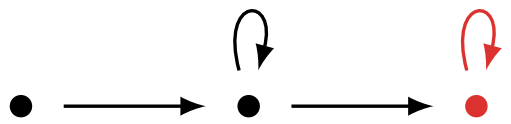
Remark

The universal property for $\to$ implies that for any subobject A of an object X, $\neg A$ is the subobject of X which is largest among all subobjects whose intersection with A is empty. Here is an example in graphs.

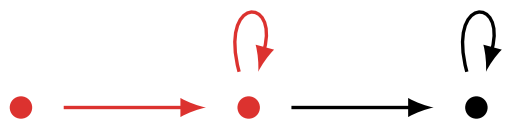
- X and A



- $\neg A$



- $\neg\neg A$



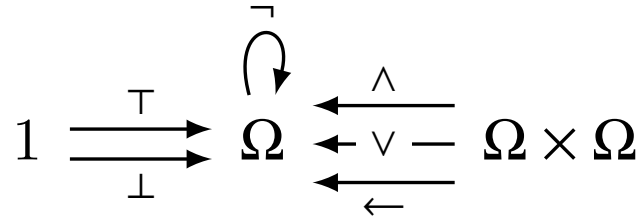
- $A \to \neg\neg A$ but $\neg\neg A \not\to A$

Remark

$$\mathrm{Sub}(X) \cong \mathrm{Hom}(X, \Omega)$$

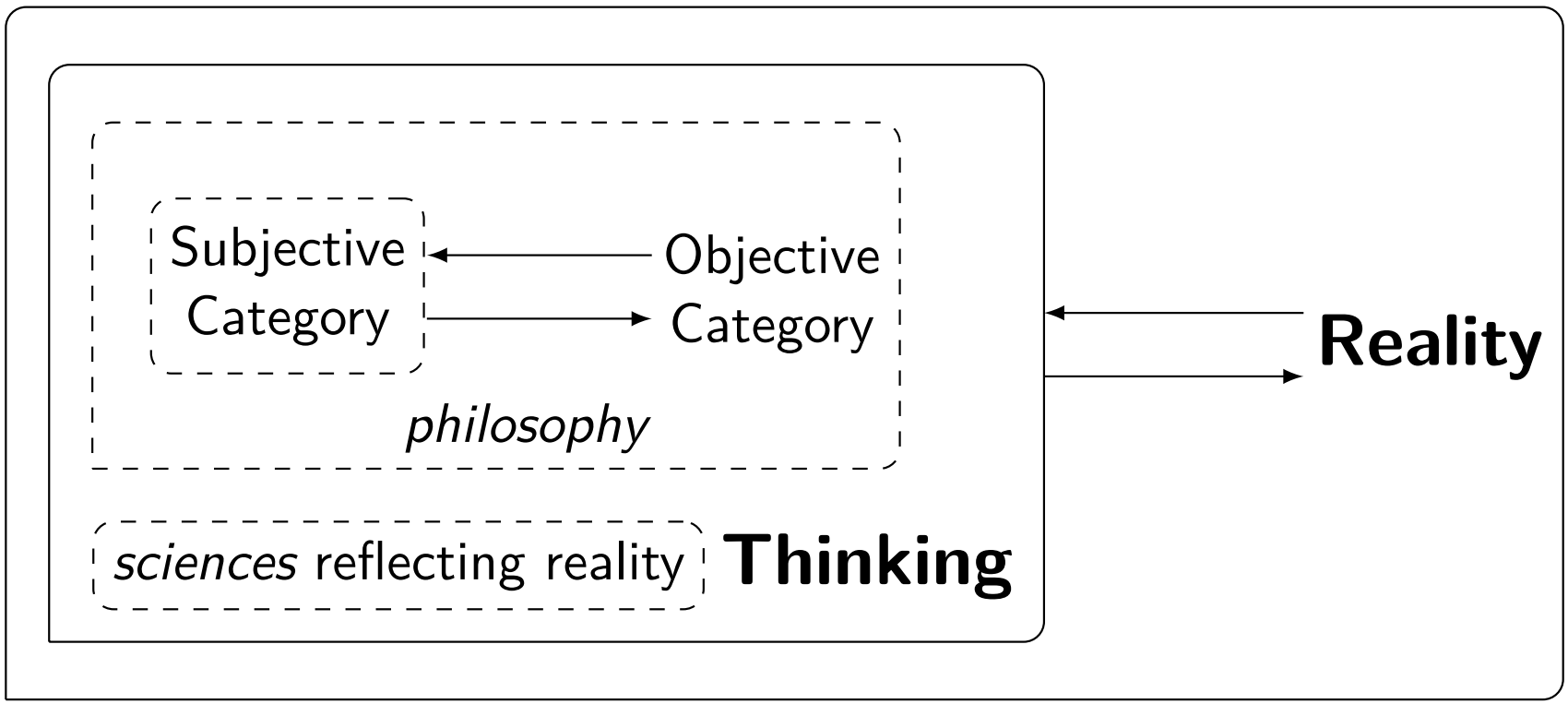
- $\mathrm{Sub}(X)$ is an external Heyting algebra
- $\mathrm{Hom}(X, \Omega)$ is an external Heyting algebra
- Because Hom and Sub are sets, hence the above isomorphism is an external statement in set-theoretical language.
- However, every external statement corresponds to some internal statement of topos.
- The corresponding internal statement in this case says that the power object Ω^X is an internal Heyting algebra.
- In particular, $\Omega \cong \Omega^1$ is an internal Heyting algebra.

True in any Topos	True only in a Boolean Topos
	$A \vee \neg A = \top$
$A \longrightarrow \neg\neg A$	$\neg\neg A \longrightarrow A$
$\neg A \wedge \neg B \leftrightarrow \neg(A \vee B)$	
$\neg A \vee \neg B \longrightarrow \neg(A \wedge B)$	$\neg(A \vee B) \longrightarrow \neg A \wedge \neg B$
$\forall x \neg A \leftrightarrow \neg\exists x A$	
$\exists x A \longrightarrow \neg\forall x \neg A$	$\neg\forall x \neg A \longrightarrow \exists x A$
$\forall x A \longrightarrow \neg\exists x \neg A$	$\neg\exists x \neg A \longrightarrow \forall x A$

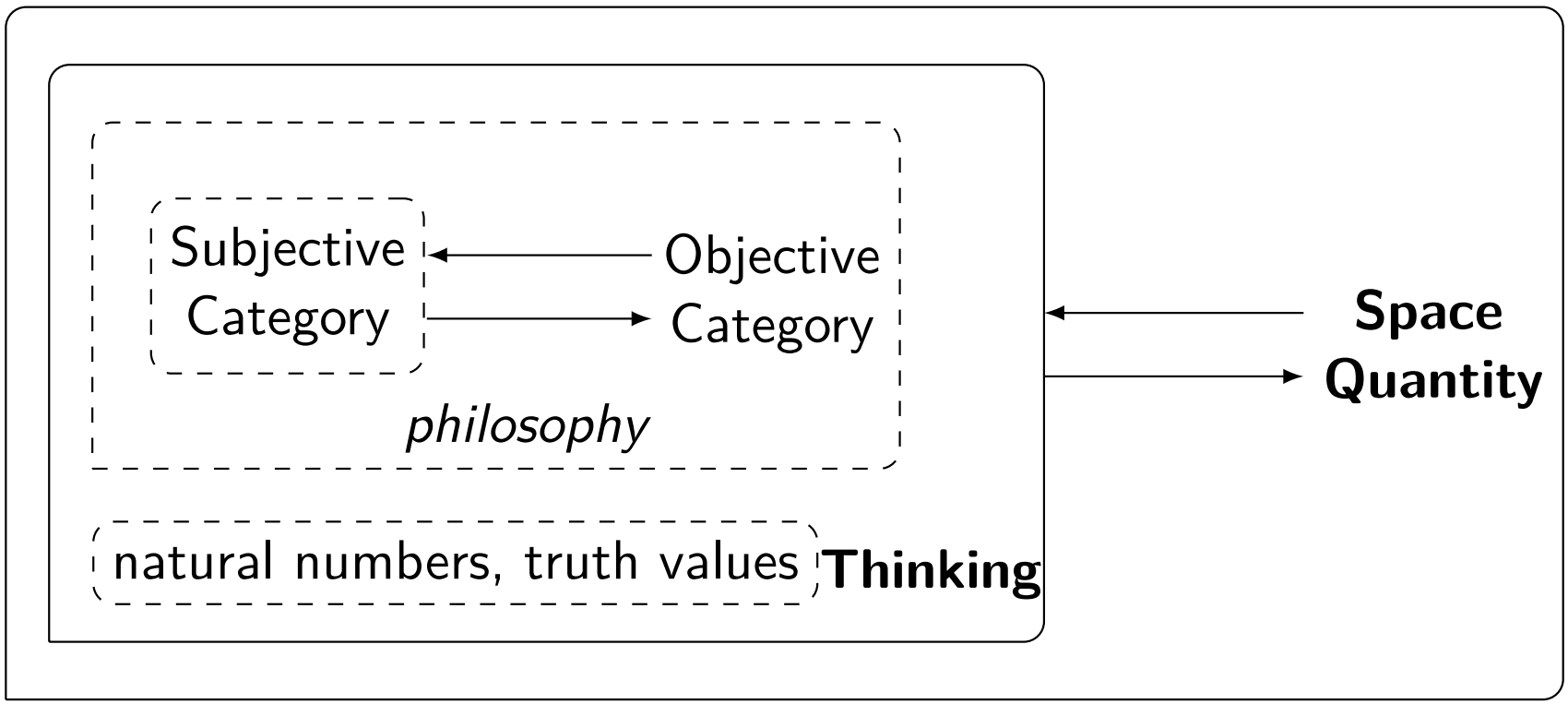


Logical Operator	Operation on $\mathrm{Sub}(X)$
truth value	$\top, \bot : 1 \to \Omega$
monic logical operator	$\Omega \to \Omega$
binary logical operator	$\Omega \times \Omega \to \Omega$
proposition with no free variable	$1 \to \Omega$
proposition with free variable x	the characteristic of some $A \in \mathrm{Sub}(X)$
proposition with free variables x, y	the characteristic of some $R \in \mathrm{Sub}(X \times Y)$
$\exists x \varphi$ with free variable y	the characteristic of $\exists_\pi R \in \mathrm{Sub}(Y)$, where $\exists_\pi \dashv \pi^*$, where $\pi : X \times Y \to Y$
$\forall x \varphi$ with free variable y	the characteristic of $\forall \pi R \in \mathrm{Sub}(Y)$, where $\pi^* \dashv \forall_\pi$, where $\pi : X \times Y \to Y$

Lawvere's Philosophy

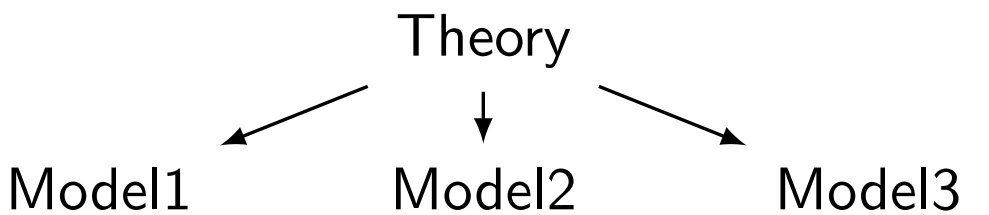


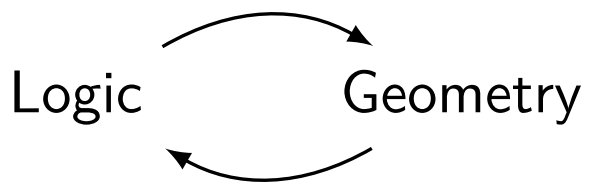
Lawvere's Philosophy



External vs Internal Logic

- Using 'intuitionistic' logic which is sound in any topos, we can construct categories and toposes and the notion of internal logic of a topos. This allows for developing the general theory of categories and toposes internally in any topos: every theorem of the general 'external' theory still holds true in every internal version of the theory.
- Hilbert's axiomatic logic is always external.
- Lawvere's topos theory and Voevodsky's homotopy type theory qualify both as theories of logic and as theories of geometry.





- For example, in a topos one and the same element is interpreted either
  1. logically as a truth-value object, or
  2. geometrically as a particular sheaf.

  Similarly, in HoTT one and the same element is interpreted either
  1. logically as identity type, or
  2. geometrically as groupoid of homotopies.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Elementary Theory of the Category of Sets ETCS

Definition (Well-pointed Topos)

A topos is called *well-pointed* iff

1. extensionality: for $A \overset{f}{\underset{g}{\rightrightarrows}} B$, if $\forall x : 1 \to A[fx = gx]$ then $f = g$.
2. non-triviality: $0 \ncong 1$.

Remark: Lawvere calls the objects of any topos 'variable sets' and those of a well-pointed topos 'constant sets', the limiting case of variation.

Definition (Choice)

For any epimorphism $f : A \to B$, there exists $g : B \to A$ s.t. $fg = 1_B$.
(every epimorphism has a section.)

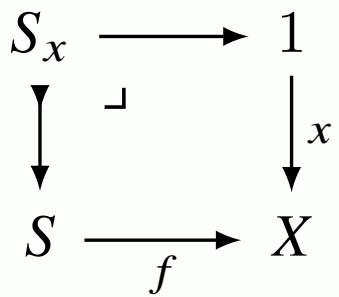
Definition (ETCS — Lawvere)

ETCS is a well-pointed topos with NNO and Choice.

Axiom Scheme of Replacement

Definition (Replacement)

For each relation $R(x, Y)$ of morphisms x to objects Y expressible in ETCS: For any object X, if for any $x : 1 \to X$ there exists an object S_x unique up to isomorphism with property $R(x, S_x)$, then there exists S and $f : S \to X$ such that for any $x : 1 \to X$ there is a pullback



Theorem

$\mathrm{ETCS} + \mathrm{Replacement}$ *is bi-interpretable with* ZFC.

ZFC vs ETCS

The ZFC axiomatization of sets looks like this:

- there are some things called 'sets'
- there is a binary relation '$\in$' on sets
- some axioms hold.

The ETCS axiomatization of sets looks like this:

- there are some things called 'sets'
- for each set X and set Y, there are some things called 'functions from X to Y'
- for each set X, set Y and set Z, there is a binary operation assigning to each pair of functions

$$f : X \to Y, \quad g : Y \to Z$$

  a function $g \circ f : X \to Z$
- some axioms hold.

ZFC vs ETCS

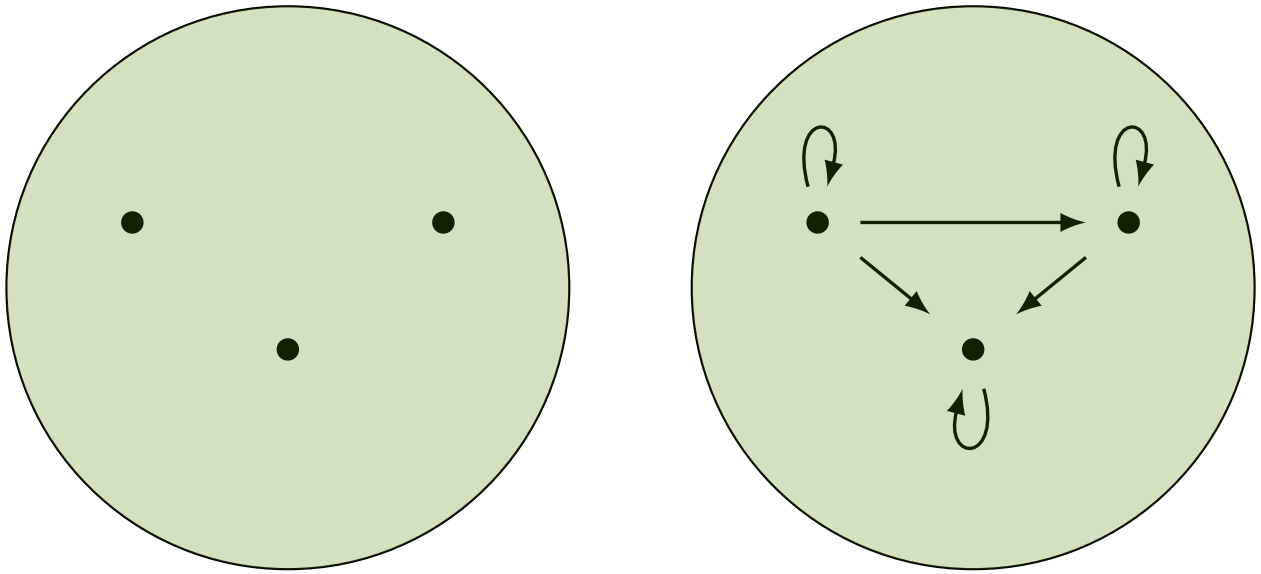
Global vs Local

	ZFC	ETCS
Are elements of a set also sets?	✓	×
Given sets X and Y, can you ask whether $X \in Y$?	✓	×
Does '$X \cap Y$' make sense for arbitrary X and Y?	✓	×
Is everything isomorphism-invariant?	×	✓
Sets	primitive	primitive
$\in$	primitive	derived
Functions	derived	primitive
Composition	derived	primitive

"The elementary theory of topoi is a basis for the study of continuously variable structures, as classical set theory is a basis for the study of constant structures."

— F. W. Lawvere

Set vs Category



"An abstract set may be conceived of as a bag of dots which are devoid of properties apart from mutual distinctness. Further, the bag as a whole is assumed to have no properties except cardinality, which amounts to just the assertion that it might or might not be isomorphic to another bag."

— F. W. Lawvere

Remark

"Set theory should not be based on membership, as in ZFC*, but rather on isomorphism-invariant structure."*

— F. W. Lawvere

"There is no reason to distinguish between, say, $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \dots\}$ *and* $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots\}$ *as definitions of 'the natural numbers'."*

— J. Benacerraf

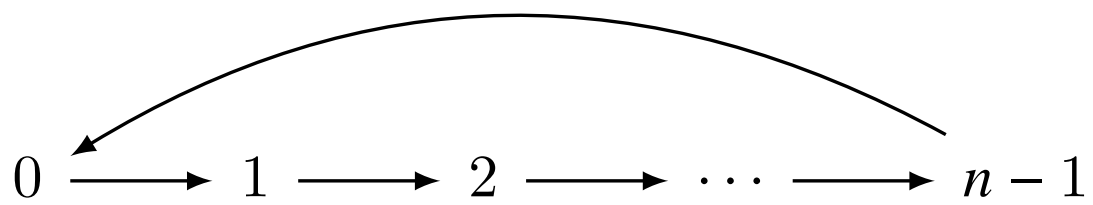
- Junk Theorems in ZFC: $2 \in 3$, $5 \subset 12$, $1 = \mathtt{P}(0)$
- Junk Theorems in ETCS: $\mathrm{dom}(5) = \mathrm{dom}(7)$

The Category $\mathbf{Set}^{\circlearrowright}$ of Endomaps of Sets

- An object of $\mathbf{Set}^{\circlearrowright}$ is any set $A \in \mathbf{Set}$ equipped with an endomap.
- A morphism $f : A^{\circlearrowright s} \to B^{\circlearrowright t}$ is a map satisfying $f \circ s = t \circ f$.

$$\begin{array}{ccc} A & \xrightarrow{s} & A \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ B & \xrightarrow[t]{} & B \end{array}$$

- Some questions can be asked about a particular object $A^{\circlearrowright s}$ of $\mathbf{Set}^{\circlearrowright}$.
  - accessibility: Given $a \in A^{\circlearrowright s}$, does $\exists b : s(b) = a$?
  - convergence to equilibrium: Given $a \in A^{\circlearrowright s}$, does $\exists n : s^{n+1}(a) = s^n(a)$?
- Example: let C_n be



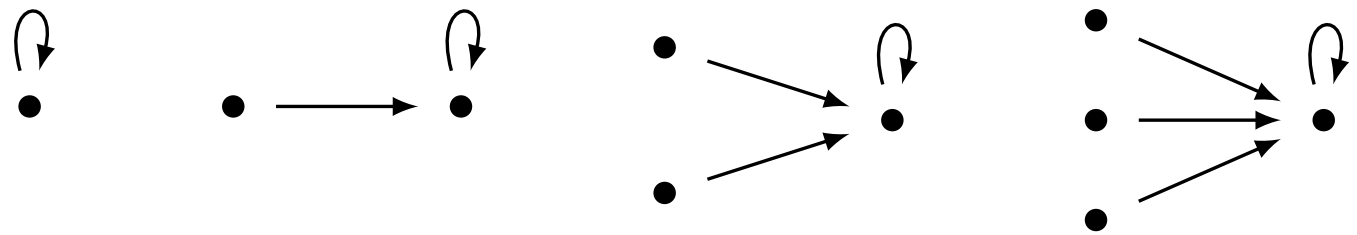
then $C_n \to B^{\circlearrowright t}$ corresponds to elements in $B^{\circlearrowright t}$ having period n.

Two Subcategories of $\mathbf{Set}^{\circlearrowright}$

Although an abstract set is completely described by a single number, the set has the potentiality to carry all sorts of structure with the help of maps.

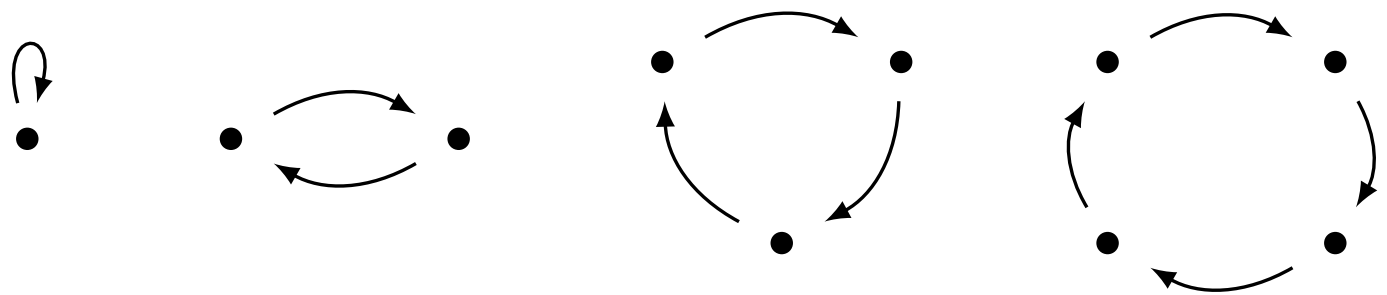
- The category $\mathbf{Set}^{e}$ of sets with an endomap which is idempotent.

  A set-with-an-endomap $A^{\circlearrowright s}$ is an object in $\mathbf{Set}^{e}$ iff $s \circ s = s$.



- The category $\mathbf{Set}^{\circlearrowright}$ of sets with an endomap which is invertible.

  $A^{\circlearrowright s}$ is an object in $\mathbf{Set}^{\circlearrowright}$ iff the endomap s has an inverse.



Dynamical Systems

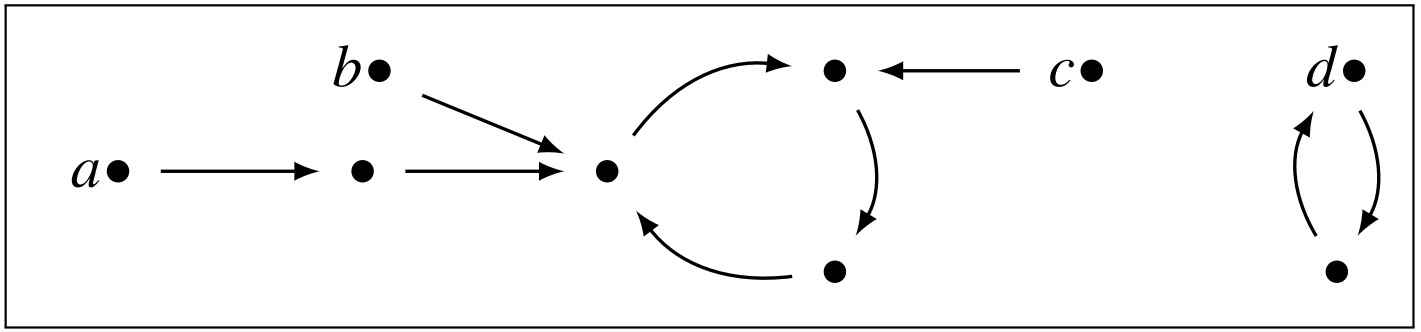
- A discrete dynamical system is a set X with an 'evolution function' $X \circlearrowright f$, which is the same thing as a functor $f : \mathbb{N} \to \mathbf{Set}$, where $\mathbb{N}$ is the monoid of natural numbers under addition.
  - $f(\bullet) = X$
  - $f(n) = \overbrace{f \circ \cdots \circ f}^{n \text{ times}}$
  - $f(0) = 1_X$
- The Category $\mathbf{Set}^{\circlearrowright}$ can be seen as the category of dynamical systems $\mathbf{Set}^{\mathbb{N}}$.
- A continuous dynamical system (sometimes called a flow) is a topological space X and a continuous function $f : \mathbb{R}^{\geq 0} \times X \to X$ s.t.

$$f^0 = 1_X \quad \text{and} \quad f^{t_2+t_1} = f^{t_2} \circ f^{t_1}$$

- A continuous dynamical system can be modeled categorically as a functor $f : \mathbb{R}^{\geq 0} \to \mathbf{Top}$.

Presentations of Dynamical Systems

- Let $A \circlearrowright f$ be



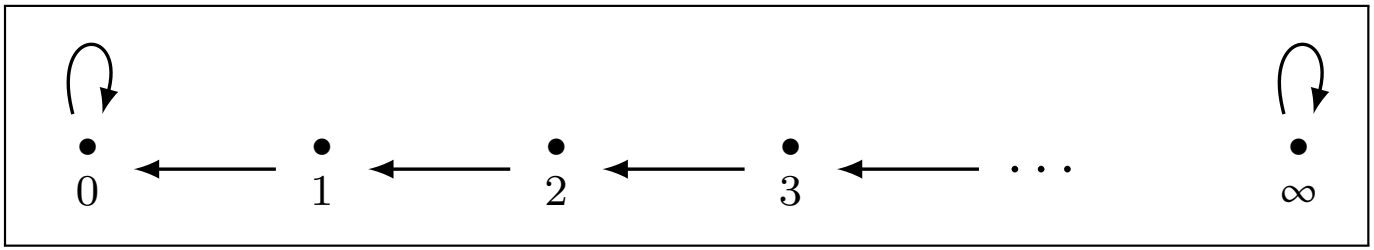
- The list $a, fa, f^2a, f^3a, f^4a, b, c, d, fd$ labels A exactly once.
- The equations give the 'relations' among the generators a, b, c, d.

$$f^5a = f^2a$$
$$fb = f^2a$$
$$fc = f^3a$$
$$f^d = d$$

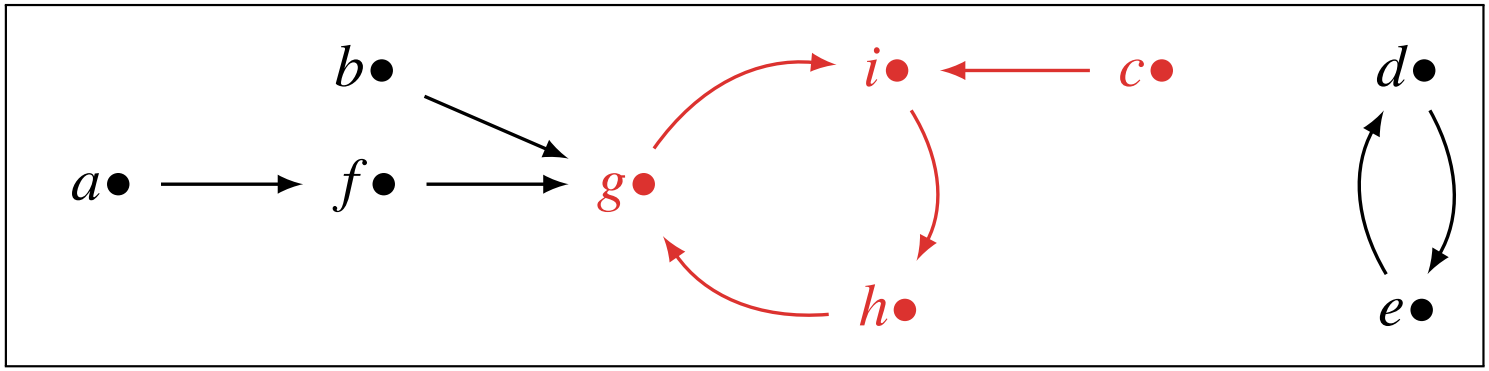
- A family of generators, together with a family of equations, is a presentation of $A \circlearrowright f$ iff it has the 'universal property': the maps from $A \circlearrowright f$ to $B \circlearrowright g$ correspond to generators in B satisfying the 'same' equations.

Subobject Classifier in Dynamical Systems

- The object of truth values in the category of dynamical systems:



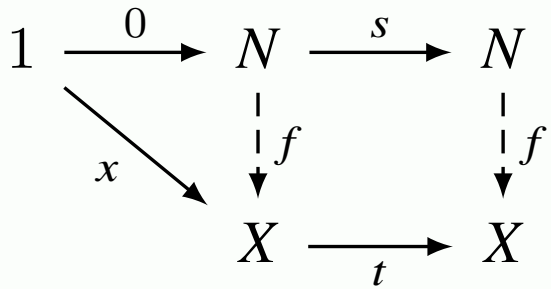
- The explanation of this is that a subsystem is a part of a dynamical system that is closed under the dynamics and if you pick a state x and ask whether x is included in the subsystem, the answer may be 'no, but it will be included in one step', or 'in two steps', etc.



Natural Numbers Object (NNO) — Lawvere's Definition

Definition (NNO — Lawvere)

A *natural numbers object* is an object N with morphisms $0 : 1 \to N$ and $s : N \to N$ such that: given morphisms $x : 1 \to X$ and $t : X \to X$, there is a unique morphism $f : N \to X$ making the following diagram commute:

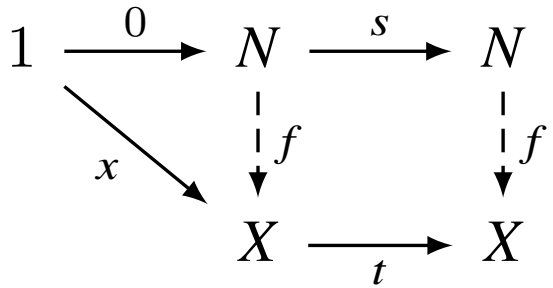


The function f is said to be constructed by *primitive recursion*.

Remark: Two cultures:

1. The natural numbers exist and we can define a function using recursion?
2. The natural numbers are defined as that object with which we can do recursion?

Remark: NNO vs Dynamical System



$$\frac{1 \xrightarrow{x} X \text{ in } \mathbf{Set}}{N \circlearrowright s \xrightarrow{f} X \circlearrowright t \text{ in } \mathbf{Set}^{\circlearrowright}}$$

Theorem (Recursion Lemma)

Given $h : N \times A \to A$ *and* $1 \xrightarrow{a} A$, *there exists a unique* $N \xrightarrow{f} A$ *s.t.*

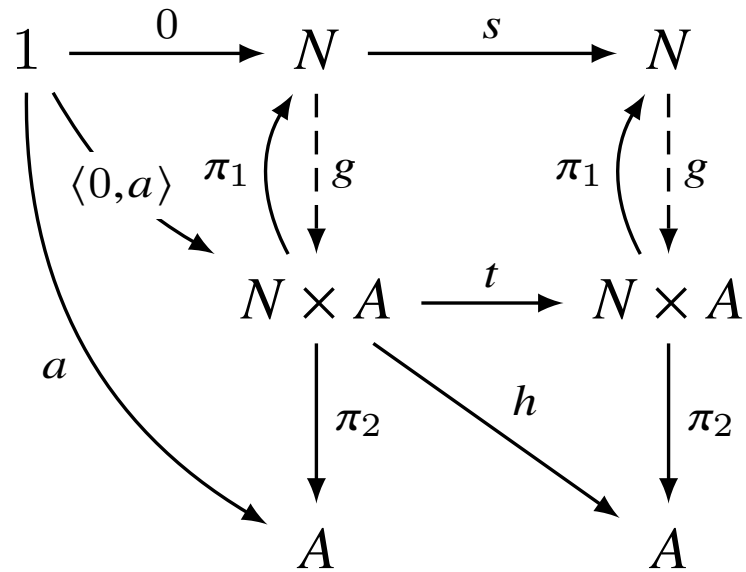
$$f(0) = a$$
$$f(n+1) = h(n, f(n))$$

Proof.

Let $t(n,x) := \langle n+1, h(n,x)\rangle$.
Then $\exists! g : N \to N \times A$:

$$g(0) = \langle 0, a\rangle$$
$$g(n+1) = t \circ g(n)$$

Let $k := \pi_1 \circ g$ and $f := \pi_2 \circ g$.
Then $g(0) = \langle k(0), f(0)\rangle = \langle 0, a\rangle$ and
$g(n+1) = \langle k(n+1), f(n+1)\rangle =$
$t \circ \langle k(n), f(n)\rangle = \langle k(n)+1, h(k(n), f(n))\rangle$.
Obviously, $k = 1_N$. □

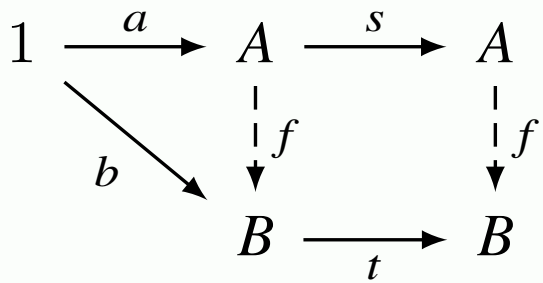


NNO — Initial Object of Peano Category

Peano Category

Let $\mathbf{C}$ be a category with terminal object 1, and define the Peano category as follows:

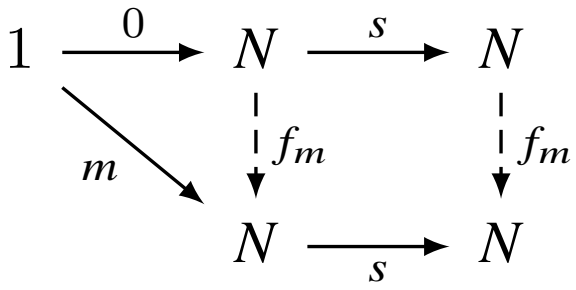
- the objects are triples (A, a, s) where $A \in \mathrm{ob}(\mathbf{C})$, and $a : 1 \to A$ and $s : A \to A$ are $\mathbf{C}$-morphisms.
- a morphism $f : (A, a, s) \to (B, b, t)$ is a $\mathbf{C}$-morphism $f : A \to B$ s.t.
  1. $f \circ a = b$
  2. $f \circ s = t \circ f$



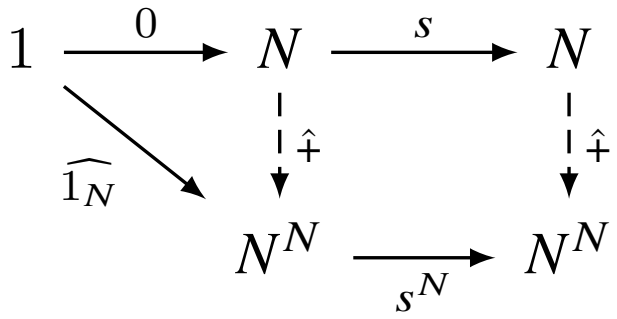
The *Natural Numbers Object* NNO is an initial object of the Peano category.

NNO — Example

- Addition

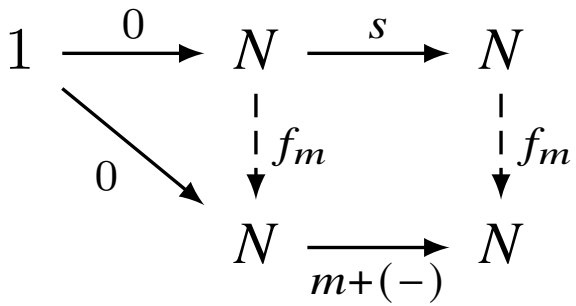


$$f_m(0) = m$$
$$f_m(s(n)) = s(f_m(n))$$

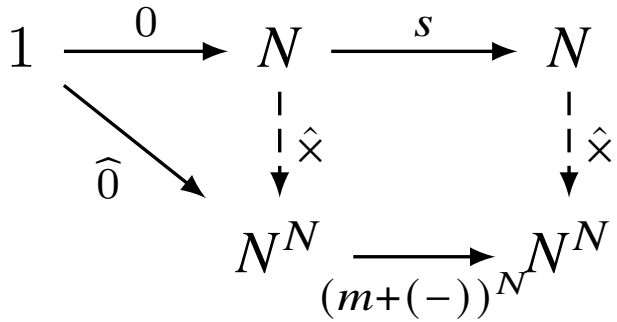


$$m + 0 = m$$
$$m + s(n) = s(m + n)$$

- Multiplication



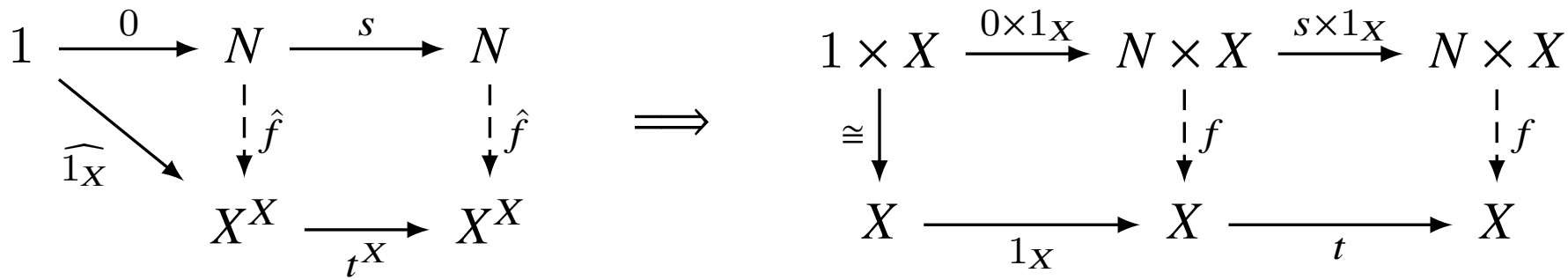
$$f_m(0) = 0$$
$$f_m(s(n)) = m + f_m(n)$$



$$m \times 0 = 0$$
$$m \times s(n) = m + m \times n$$

NNO — Example

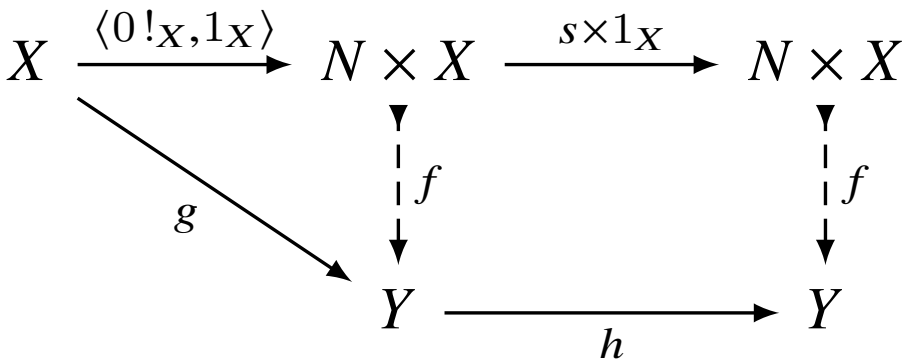
- Iteration



where $\widehat{1_X}$ is the currying of $X \xrightarrow{1_X} X$.

$$\left.\begin{aligned} f(0,x) &= x \\ f(sn,x) &= t\big(f(n,x)\big) \end{aligned}\right\} \quad \Longrightarrow \quad f(n,x) = t^n(x)$$

Parametrized NNO



$$f(0,x) = g(x)$$
$$f(sn,x) = h(f(n,x))$$

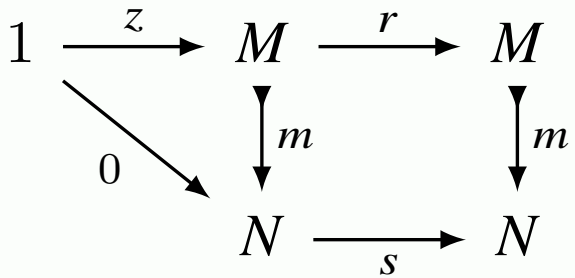
NNO — Dedekind's Definition

Definition (NNO — Dedekind)

In a topos, $1 \xrightarrow{0} N \xrightarrow{s} N$ is a NNO iff

1. if $1 \xrightarrow{x} N$, then $sx \neq 0$.
2. s is monic.
3. if $M \xrightarrow{m} N$ is a subobject of N such that
   - there is $1 \xrightarrow{z} M$ such that $mz = 0$ and
   - there is $M \xrightarrow{r} M$ such that $mr = sm$

   then $M \overset{m}{\cong} N$.

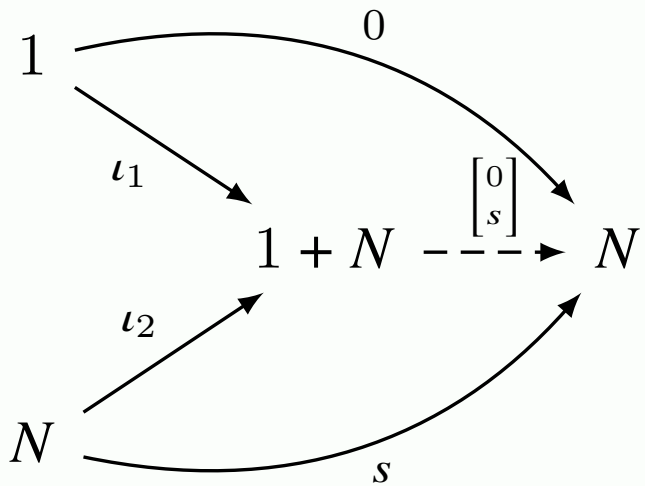


NNO — Freyd's Definition

Definition (NNO — Freyd)

In a topos, $1 \xrightarrow{0} N \xrightarrow{s} N$ is a NNO iff

1. the morphism $\begin{bmatrix} 0 \\ s \end{bmatrix} : 1 + N \to N$ is an isomorphism.



2. $N \underset{1_N}{\overset{s}{\rightrightarrows}} N \xrightarrow{!_N} 1$ is a coequalizer.

Skeleton

Definition (Skeletal)

A category $\mathbf{C}$ is *skeletal* iff for any $A, B \in \mathbf{C} : A \cong B \implies A = B$.

Theorem

If $\mathbf{C}$ *and* $\mathbf{D}$ *are skeletal, then* $\mathbf{C} \simeq \mathbf{D} \implies \mathbf{C} \cong \mathbf{D}$.

Definition (Skeleton)

Given a category $\mathbf{C}$, a *skeleton* $\mathrm{sk}(\mathbf{C})$ of $\mathbf{C}$ is a full subcategory containing exactly one objects from each isomorphism class of objects of $\mathbf{C}$.

Obviously, $\mathbf{C} \simeq \mathbf{D} \implies \mathrm{sk}(\mathbf{C}) \cong \mathrm{sk}(\mathbf{D})$.

The following statements are equivalent to the axiom of choice.

- Any category has a skeleton.
- Any category is equivalent to any of its skeletons.
- Any two skeletons of a given category are isomorphic.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

An object is completely determined by its relationships to other objects

"You work at a particle accelerator. You want to understand some particle. All you can do are throw other particles at it and see what happens. If you understand how your mystery particle responds to all possible test particles at all possible test energies, then you know everything there is to know about your mystery particle."

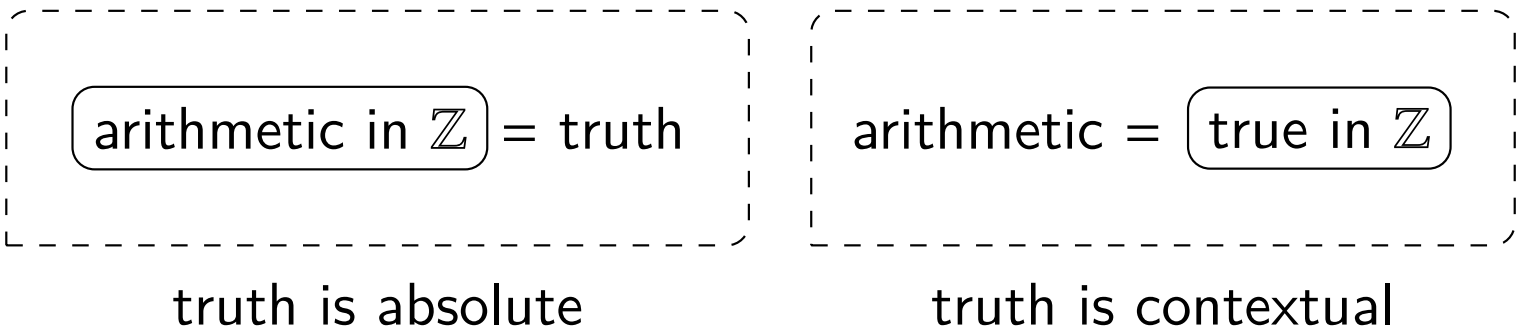
— Ravi Vakil



Tell me who your friends are and I will tell you who you are.

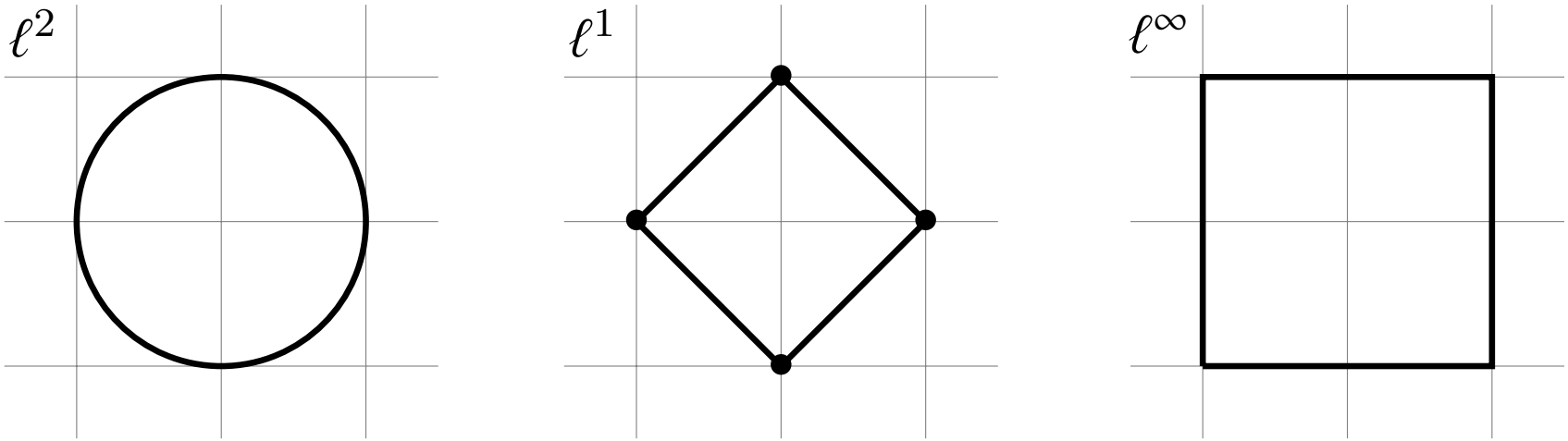
Context: Which is the number 5?

1. In the context of natural numbers, 5 is a prime number.
2. In the context of integers, 5 has an additive inverse, which is -5.
3. In the context of rational numbers, 5 has a multiplicative inverse, which is $\frac{1}{5}$.
4. In the context of arithmetic modulo 6, 5 is a generator, which means if you add 5 to itself repeatedly you will get every number in the system.



Remark: Math isn't about absolute truth; it's about different contexts in which different things can be true.

什么是圆?



- 什么是圆? — 到定点 $\mathbf{c}$ 的距离等于定长 r 的点的集合.

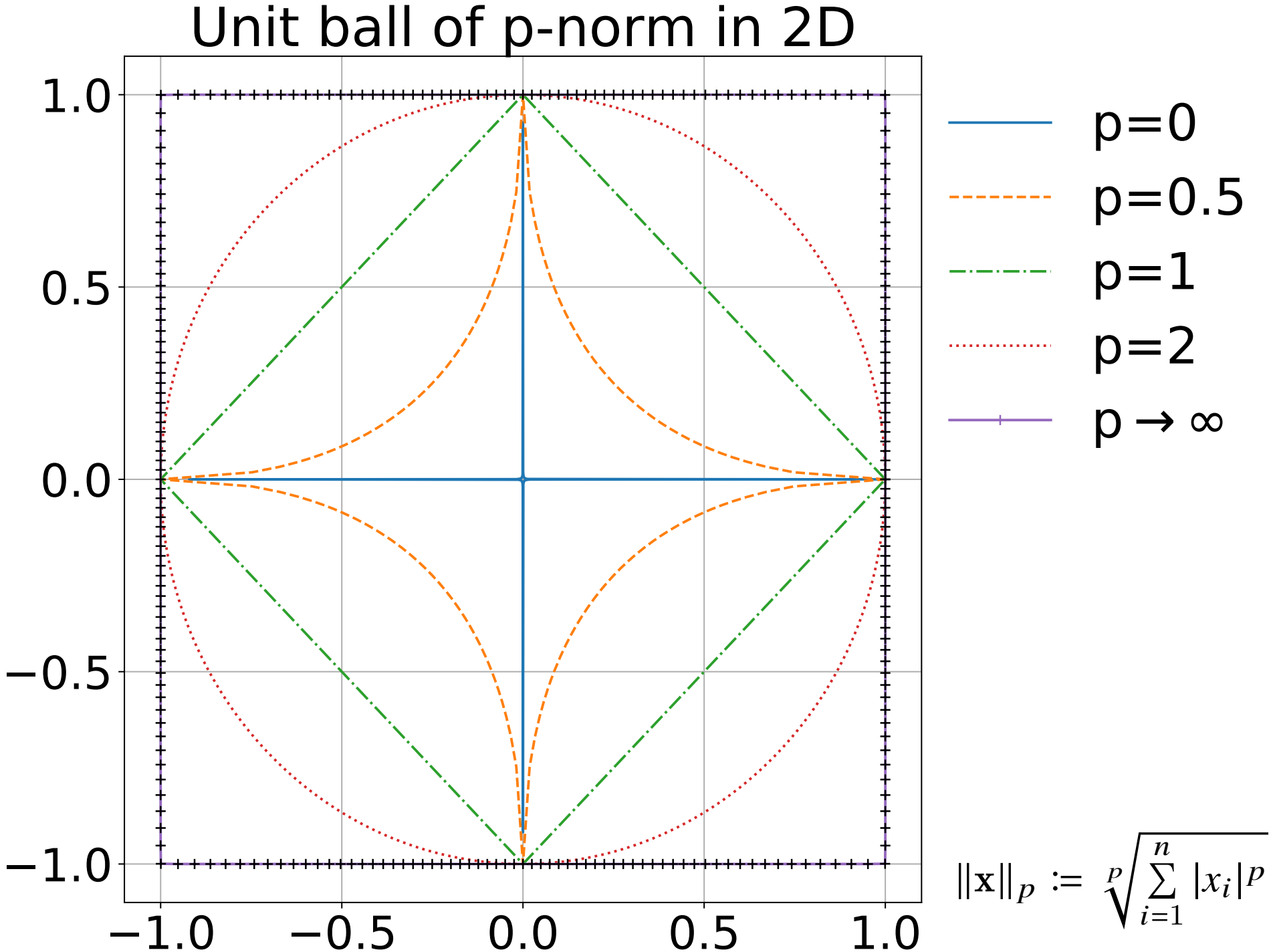
$$\{\mathbf{x} : \mathcal{M} \vDash d(\mathbf{x}, \mathbf{c}) = r\}$$

- 什么是半径为 1 的圆? — 基于什么距离? $\ell^2, \ell^1, \ell^\infty$ norm?

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_i |x_i - y_i|^2} \quad d_1(\mathbf{x}, \mathbf{y}) = \sum_i |x_i - y_i| \quad d_\infty(\mathbf{x}, \mathbf{y}) = \max_i |x_i - y_i|$$

- 什么是圆周率 π? 在 "出租车世界" $(\mathbb{R}^2, d_1)$, $\pi = \frac{\text{周长}}{\text{直径}} = \frac{8}{2} = 4$

"We are often interested not just in whether or not something is true, but in where it is true." — *John Baez*

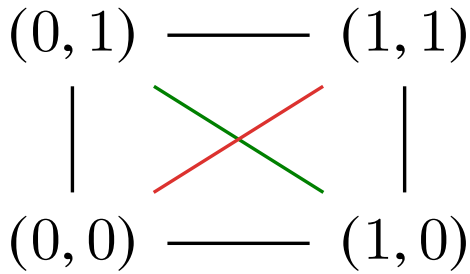

Unit ball of p-norm in 2D
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
p=0
p=0.5
p=1
p=2
p → ∞
$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$

$\mathbb{F}_2^2$ 平面上的圆

- $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z} = (\{0,1\}, 0, 1, +, \cdot)$

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1



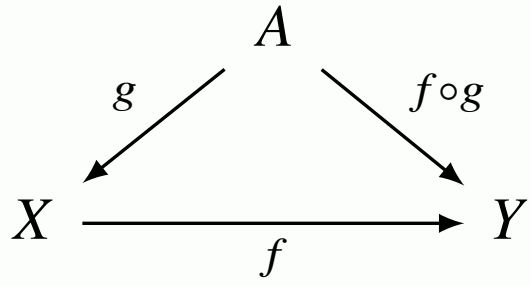
- $\mathbb{F}_2^2$ 平面只有四个点: $(0,0), (0,1), (1,0), (1,1)$
- $\mathbb{F}_2^2$ 平面有多少条直线? $ax + by + c = 0 \ \ (a, b, c \in \mathbb{F}_2)$, a, b 不同为 0
- $\mathbb{F}_2^2$ 平面有多少个圆? $(x-a)^2 + (y-b)^2 = r^2 \ \ (a, b, r \in \mathbb{F}_2)$
- 因为 $\forall x \in \mathbb{F}_2 : x^2 = x$, 所以
  $(x-a)^2 + (y-b)^2 = r^2 \iff x - a + y - b = r$
  即 $x + y = 0$ 或 $x + y = 1$.
  因此, 有且仅有 2 个圆, 都是直线!
- 例: $x + y = 0 \iff (x-0)^2 + (y-0)^2 = 0 \iff (x-1)^2 + (y-1)^2 = 0 \iff (x-0)^2 + (y-1)^2 = 1 \iff (x-1)^2 + (y-0)^2 = 1$
- 该圆周上有两个点: $(0,0), (1,1)$. 两个点都是圆心, 半径是 0.
- 该圆周外的两个点 $(0,1), (1,0)$ 也是圆心, 此时半径为 1.
- 因此, $\mathbb{F}_2^2$ 平面有 2 个圆, 每个圆都有 4 个圆心, 2 个半径.

Hom Functor

Definition (Hom Functor)

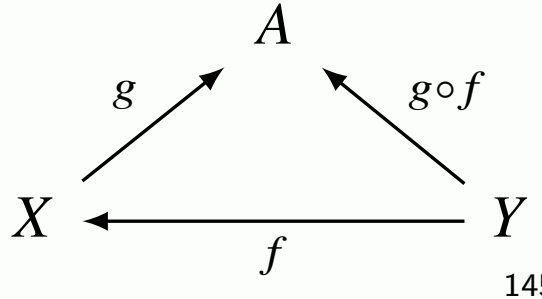
- For $A \in \mathbf{C}$, the functor $h^A := \mathrm{Hom}(A, -)$ maps $X \mapsto \mathrm{Hom}(A, X)$, and $f : X \to Y$ to $f_* := \mathrm{Hom}(A, f) : \mathrm{Hom}(A, X) \to \mathrm{Hom}(A, Y) :: g \mapsto fg$.

$$\begin{array}{ccc} X & \xrightarrow{\quad f \quad} & Y \\ \mathrm{Hom}(A,X) & \xrightarrow{\mathrm{Hom}(A,f)} & \mathrm{Hom}(A,Y) \\ g & \xmapsto{\quad f\circ - \quad} & fg \end{array}$$

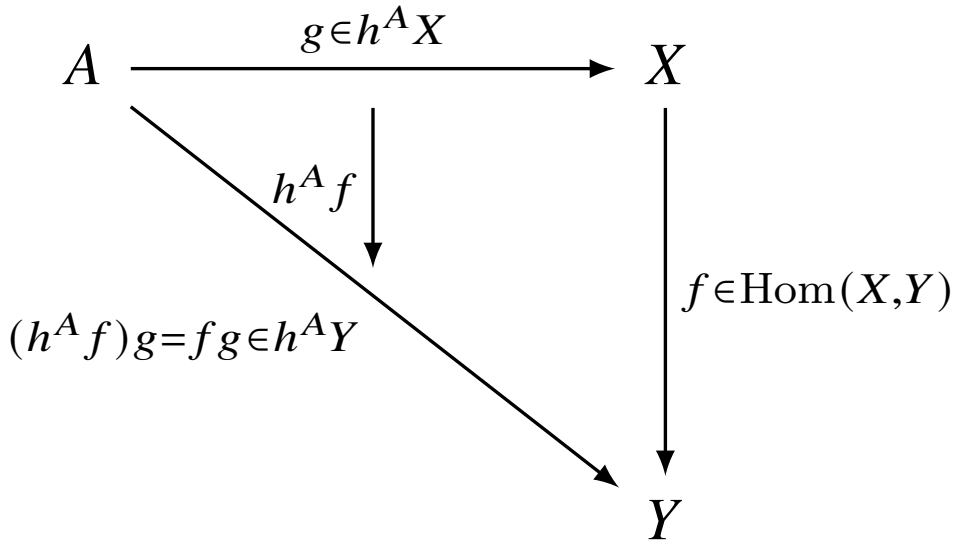


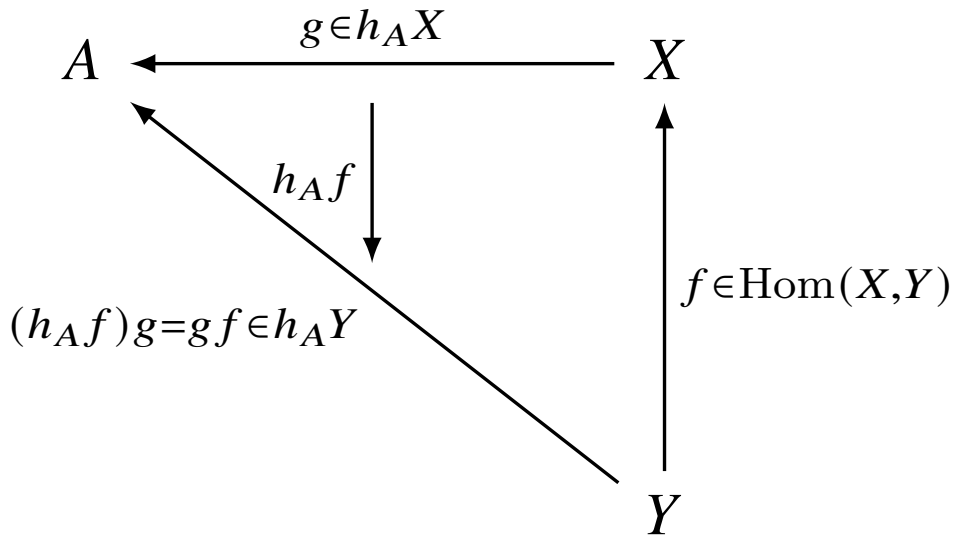
- For $A \in \mathbf{C}$, the functor $h_A := \mathrm{Hom}(-, A)$ maps $X \mapsto \mathrm{Hom}(X, A)$, and $f : Y \to X$ to $f^* := \mathrm{Hom}(f, A) : \mathrm{Hom}(X, A) \to \mathrm{Hom}(Y, A) :: g \mapsto gf$.

$$\begin{array}{ccc} Y & \xrightarrow{\quad f \quad} & X \\ \mathrm{Hom}(X,A) & \xrightarrow{\mathrm{Hom}(f,A)} & \mathrm{Hom}(Y,A) \\ g & \xmapsto{\quad -\circ f \quad} & gf \end{array}$$

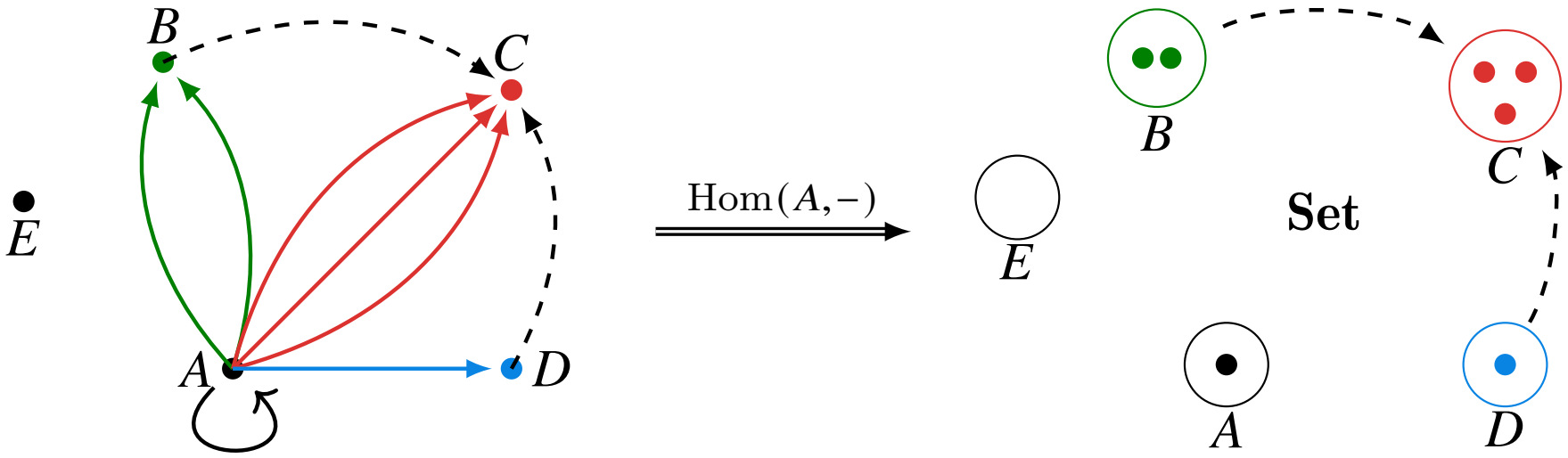


Hom Functor


$g \in h^A X$
A
X
$h^A f$
$f \in \mathrm{Hom}(X,Y)$
$(h^A f)g = fg \in h^A Y$
Y

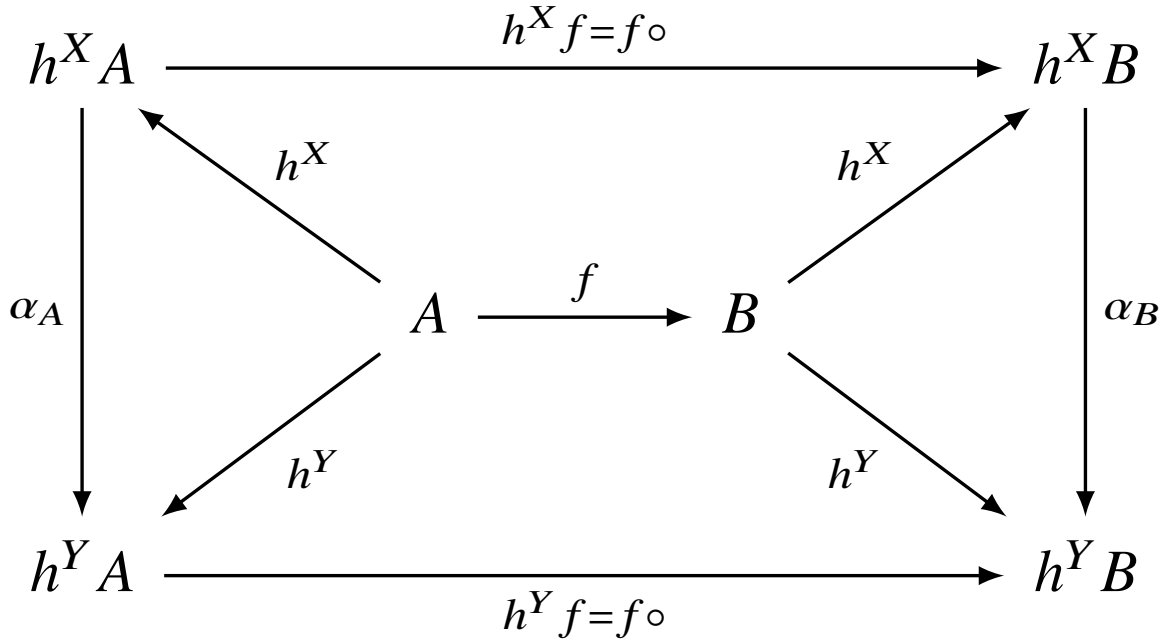

$g \in h_A X$
A
X
$h_A f$
$f \in \mathrm{Hom}(X,Y)$
$(h_A f)g = gf \in h_A Y$
Y

The Hom functor $\mathrm{Hom}(A, -) : \mathbf{C} \to \mathbf{Set}$



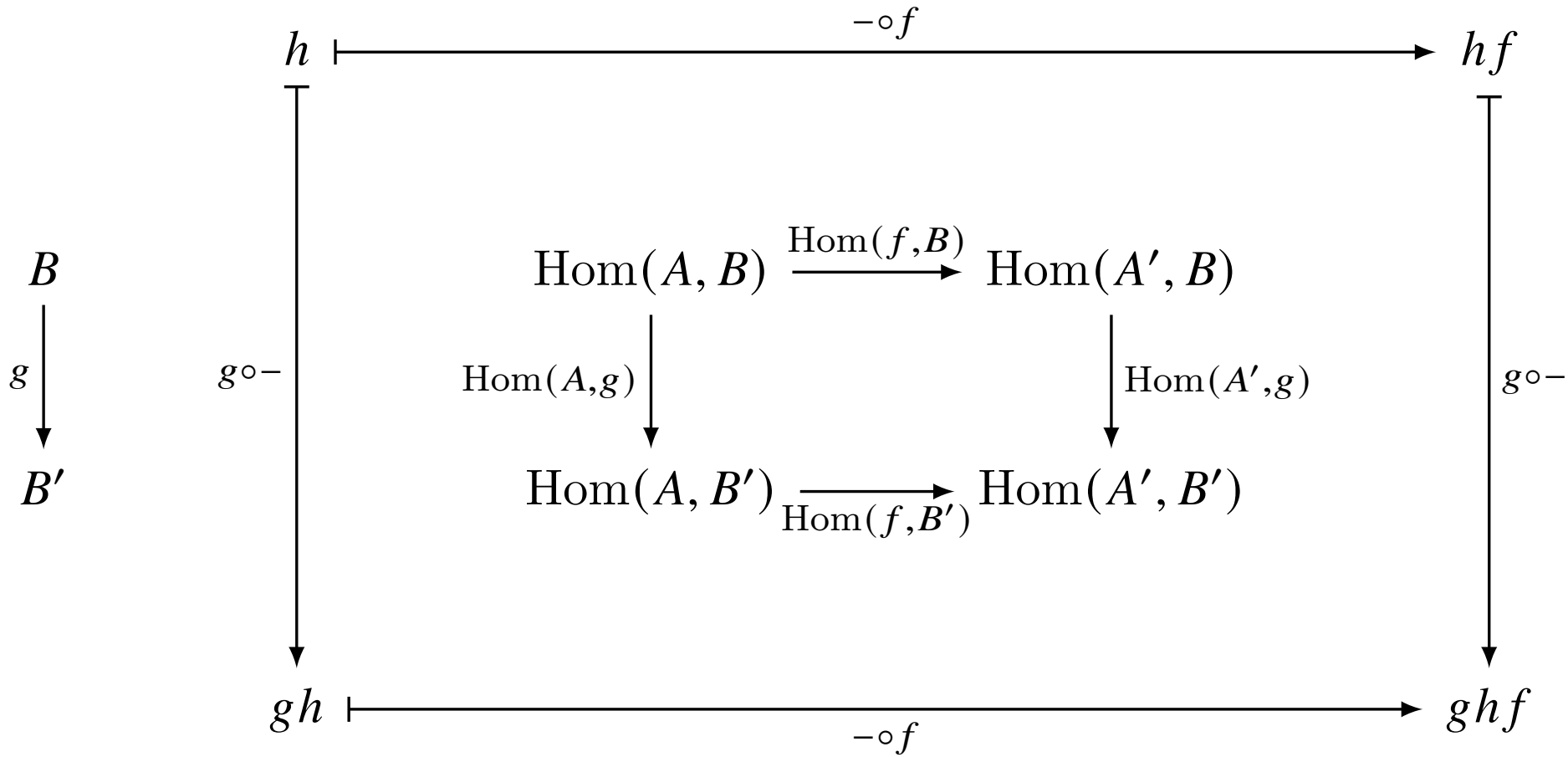
Natural Transformation between Hom Functors

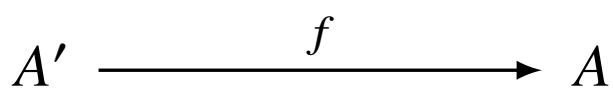
$$\alpha \in \mathrm{Nat}(h^X, h^Y)$$



$$f \circ \alpha_A = \alpha_B \circ f$$

$\mathrm{Hom}(-,-) : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Set}$ maps $A, B \in \mathbf{C}$ to $\mathrm{Hom}(A,B)$, and maps $f : A' \to A, g : B \to B'$ to
$\mathrm{Hom}(f,g) : \mathrm{Hom}(A,B) \to \mathrm{Hom}(A',B') :: h \mapsto ghf$.





Yoneda Embedding

$$\frac{\mathrm{Hom} : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Set}}{\mathsf{y} : \mathbf{C} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}}$$

Definition (Yoneda Embedding)

Let $\mathbf{C}$ be a locally small category. Define $\mathsf{y} : \mathbf{C} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ as follows.

- For $A \in \mathbf{C}$, $\mathsf{y} : A \mapsto \mathrm{Hom}(-, A) : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$.
- For $f : A \to B$, $\mathsf{y} : f \mapsto f_* = \mathrm{Hom}(-, f) : \mathrm{Hom}(-, A) \to \mathrm{Hom}(-, B)$.

Definition (Yoneda Embedding)

Let $\mathbf{C}$ be a locally small category. Define $\mathsf{y} : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}^{\mathbf{C}}$ as follows.

- For $A \in \mathbf{C}^{\mathrm{op}}$, $\mathsf{y} : A \mapsto \mathrm{Hom}(A, -) : \mathbf{C} \to \mathbf{Set}$.
- For $f : B \to A$, $\mathsf{y} : f \mapsto f^* = \mathrm{Hom}(f, -) : \mathrm{Hom}(A, -) \to \mathrm{Hom}(B, -)$.

$$\widehat{\mathbf{C}} := \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$$

Yoneda Lemma

Theorem (Yoneda Lemma)

For any locally small category $\mathbf{C}$*, object* $A \in \mathbf{C}$ *and functor* $F : \mathbf{C} \to \mathbf{Set}$*,*

$$\mathbf{Set}^{\mathbf{C}}(\mathrm{y}A, F) \cong FA$$

naturally in both A *and* F*.*

Theorem (Yoneda Lemma: another version)

- *For any locally small category* $\mathbf{C}$*, object* $A \in \mathbf{C}$ *and functor* $F : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$*,*

  $$\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(\mathrm{Hom}(-, A), F) \cong FA$$

  naturally in both A *and* F*.*
- *For any locally small category* $\mathbf{C}$*, object* $A \in \mathbf{C}$ *and functor* $F : \mathbf{C} \to \mathbf{Set}$*,*

  $$\mathbf{Set}^{\mathbf{C}}(\mathrm{Hom}(A, -), F) \cong FA$$

  naturally in both A *and* F*.*

Proof Sketch of Yoneda Lemma

Proof.

Let $\varphi : \mathrm{Hom}(\mathbf{C}(A,-),F) \to FA :: \alpha \mapsto \alpha_A(1_A)$.
Our first aim is to define an inverse function $\psi : FA \to \mathrm{Hom}(\mathbf{C}(A,-),F)$ that constructs a natural transformation $\psi(a) : \mathbf{C}(A,-) \to F$ from any $a \in FA$.
To this end, we must define components $\psi(a)_B : \mathbf{C}(A,B) \to FB$ so that $\psi(a)_B \circ \mathbf{C}(A,f) = Ff \circ \psi(a)_A$.
To make $\varphi(\psi(a)) = a$, let $\psi(a)_A(1_A) = a$.
Now, naturality forces us to define $\psi(a)_B(f) := Ff(a)$.
By construction, $\varphi(\psi(a)) = a$. It remains to verify that $\psi(\varphi(\alpha)) = \alpha$.
$\psi(\varphi(\alpha))_B(f) = \psi(\alpha_A(1_A))_B(f) = Ff(\alpha_A(1_A)) = \alpha_B(f)$

$$\begin{array}{ccc} \mathbf{C}(A,A) & \xrightarrow{\mathbf{C}(A,f)} & \mathbf{C}(A,B) \\ {\scriptstyle \psi(a)_A}\downarrow & & \downarrow{\scriptstyle \psi(a)_B} \\ FA & \xrightarrow[Ff]{} & FB \end{array}$$

$$\begin{array}{ccc} 1_A & \longmapsto & f \\ \downarrow & & \downarrow \\ a & \longmapsto & \psi(a)_B(f) = Ff(a) \end{array}$$

$$\begin{array}{ccc} \mathbf{C}(A,A) & \xrightarrow{\mathbf{C}(A,f)} & \mathbf{C}(A,B) \\ {\scriptstyle \alpha_A}\downarrow & & \downarrow{\scriptstyle \alpha_B} \\ FA & \xrightarrow[Ff]{} & FB \end{array}$$

Proof Sketch of Yoneda Lemma — Continued

To show that φ is natural in F, suppose given a natural transformation $\beta : F \to G$.

$$\begin{array}{ccc} \mathbf{Set}^{\mathbf{C}}(\mathbf{C}(A,-),F) & \xrightarrow{\beta\circ -} & \mathbf{Set}^{\mathbf{C}}(\mathbf{C}(A,-),G) \\ \downarrow{\scriptstyle \varphi_F} & & \downarrow{\scriptstyle \varphi_G} \\ FA & \xrightarrow[\beta_A]{} & GA \end{array}$$

$$\beta_A(\varphi_F(\alpha)) = \beta_A(\alpha_A(1_A)) = (\beta \circ \alpha)_A(1_A) = \varphi_G(\beta \circ \alpha)$$

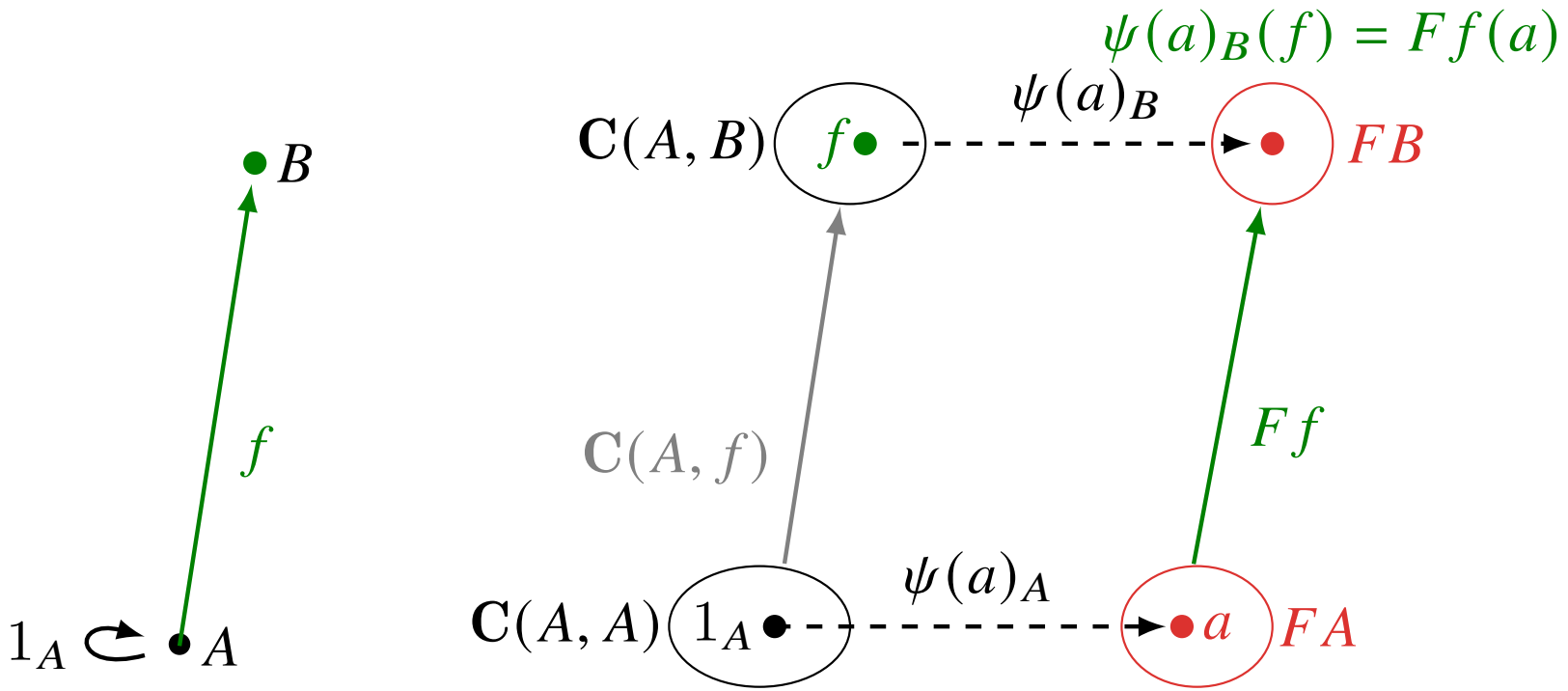
To show that φ is natural in A, suppose given $f : A \to B$.

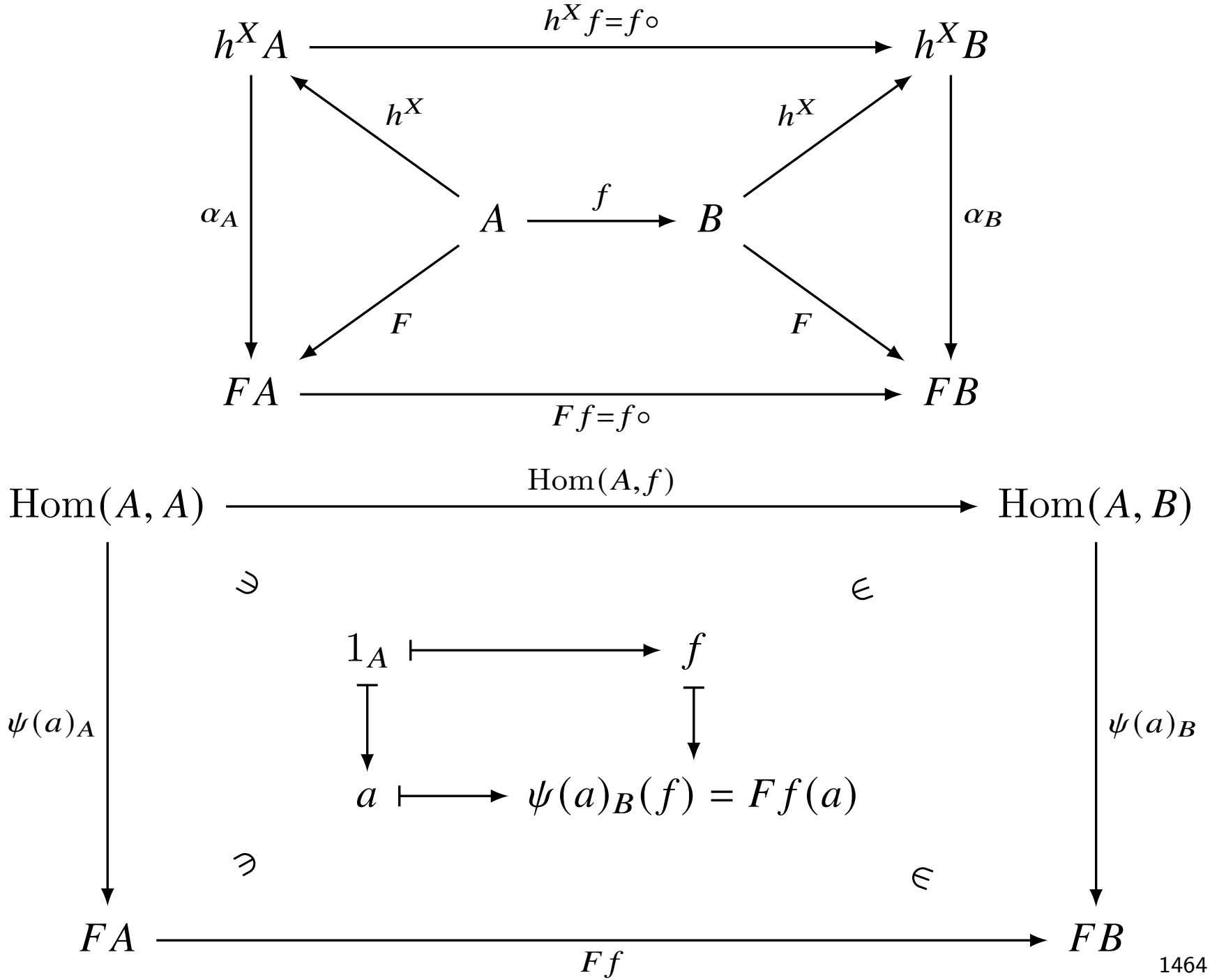
$$\begin{array}{ccc} \mathbf{Set}^{\mathbf{C}}(\mathbf{C}(A,-),F) & \xrightarrow{-\circ\mathbf{C}(f,-)} & \mathbf{Set}^{\mathbf{C}}(\mathbf{C}(B,-),F) \\ \downarrow{\scriptstyle \varphi_A} & & \downarrow{\scriptstyle \varphi_B} \\ FA & \xrightarrow[Ff]{} & FB \end{array}$$

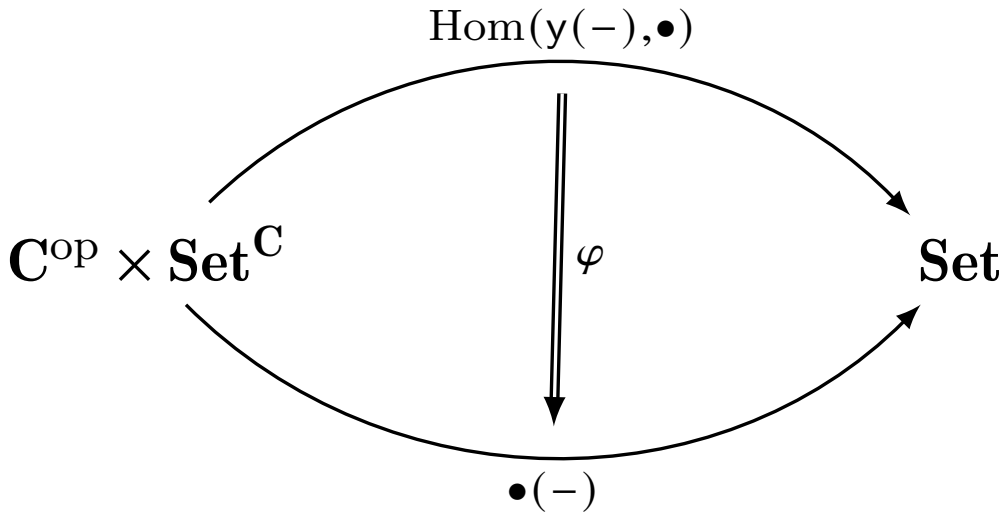
$$Ff(\varphi_A(\alpha)) = Ff(\alpha_A(1_A)) = \alpha_B \circ \mathbf{C}(A,f)(1_A) = \alpha_B(f) = \alpha_B \circ \mathbf{C}(f,-)_B(1_B) = (\alpha \circ \mathbf{C}(f,-))_B(1_B) = \varphi_B(\alpha \circ \mathbf{C}(f,-))$$

Remark

- First, given a natural transformation $\alpha : \mathbf{C}(A, -) \to F$ we construct an element $\alpha_A(1_A)$ of FA. $\quad \varphi : \alpha \mapsto \alpha_A(1_A)$
- Second, given an element $a \in FA$ we construct the corresponding natural transformation $\psi(a) : \mathbf{C}(A, -) \to F$.




$h^X A$
$h^X f = f\circ$
$h^X B$
h^X
h^X
α_A
A
f
B
α_B
F
F
FA
FB
$Ff = f\circ$
$\mathrm{Hom}(A, A)$
$\mathrm{Hom}(A, f)$
$\mathrm{Hom}(A, B)$
$\ni$
$\in$
1_A
f
$\psi(a)_A$
$\psi(a)_B$
a
$\psi(a)_B(f) = Ff(a)$
$\ni$
$\in$
FA
FB
Ff



Everything you could possibly hope to be true is true — as long as you hope for the right things. ☺

Theorem (Restricted Yoneda Lemma)

For any locally small category $\mathbf{C}$*, object* $A, B \in \mathbf{C}$*,*

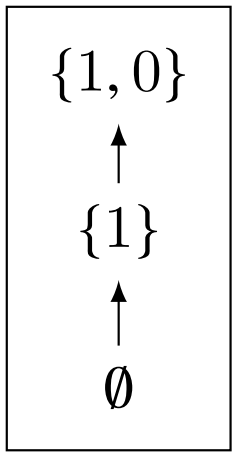
- $\mathbf{Set}^{\mathbf{C}}(\mathrm{Hom}(A,-), \mathrm{Hom}(B,-)) \cong \mathrm{Hom}(B,A)$
- $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(\mathrm{Hom}(-,A), \mathrm{Hom}(-,B)) \cong \mathrm{Hom}(A,B)$

Proof.

Let $F := \mathrm{Hom}(B,-)$ or $F := \mathrm{Hom}(-,B)$. □

Example — Upper Set

- Given a preorder $(P, \leq)$, an *upper set* in P is a subset $U \subset P$ s.t. $x \in U \;\&\; x \leq y \implies y \in U$.
- We write $\mathrm{U}(P)$ for the set of upper sets in P.
- Example: $\mathrm{U}(0 \xrightarrow{\leq} 1)$



- Let $x^{\uparrow} := \{y \in P : x \leq y\}$
- This defines a monotone map $\uparrow\colon P^{\mathrm{op}} \to \mathrm{U}(P)$.

$$x \leq y \iff y^{\uparrow} \subset x^{\uparrow}$$

- This is the Yoneda lemma for preorders.
  — to know an element is the same as knowing its upper set.

Example — $\mathbf{Mat}$

The category of matrix $\mathbf{Mat}$ is a category whose

- objects are natural numbers $n \in \mathbb{N}$.
- morphisms $M : m \to n$ are $m \times n$ matrices M.
- if $m = 0$ or $n = 0$, we just assign a unique morphism $m \to n$, which we can see as a "zero-dimensional matrix".
- the identity of n is just the $n \times n$ identity matrix.

Consider linear operation on rows of matrices, such as "multiply the second row by 2 and add it to the first one".

$$\text{for example } \begin{bmatrix} a & b & c \\ a' & b' & c' \end{bmatrix} \mapsto \begin{bmatrix} a + 2a' & b + 2b' & c + 2c' \\ a' & b' & c' \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 + 2 \cdot 0 & 0 + 2 \cdot 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$\text{we have } \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ a' & b' & c' \end{bmatrix} \mapsto \begin{bmatrix} a + 2a' & b + 2b' & c + 2c' \\ a' & b' & c' \end{bmatrix}$$

This is an instance of the restricted Yoneda Lemma.

$$\mathrm{Hom}(n, n) \cong \mathbf{Set}^{\mathbf{Mat}^{\mathrm{op}}} (\mathrm{Hom}(-, n), \mathrm{Hom}(-, n))$$

- Every naturally-defined column operation $\mathrm{Hom}(-,m) \xrightarrow{\alpha} \mathrm{Hom}(-,n)$ is given by right multiplication by the $m \times n$ matrix obtained by applying the column operation α_m to the $m \times m$ identity matrix.
- The operation that swaps the first two columns is defined by right multiplication by the matrix $\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$.
- The operation that multiplies the first column by a scalar λ is defined by right multiplication by the matrix $\begin{bmatrix} \lambda & 0 & \cdots \\ 0 & 1 & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}$.
- The operation that adds the first column to the second column is defined by right multiplication by the matrix $\begin{bmatrix} 1 & 1 & \cdots \\ 0 & 1 & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}$.

Corollary

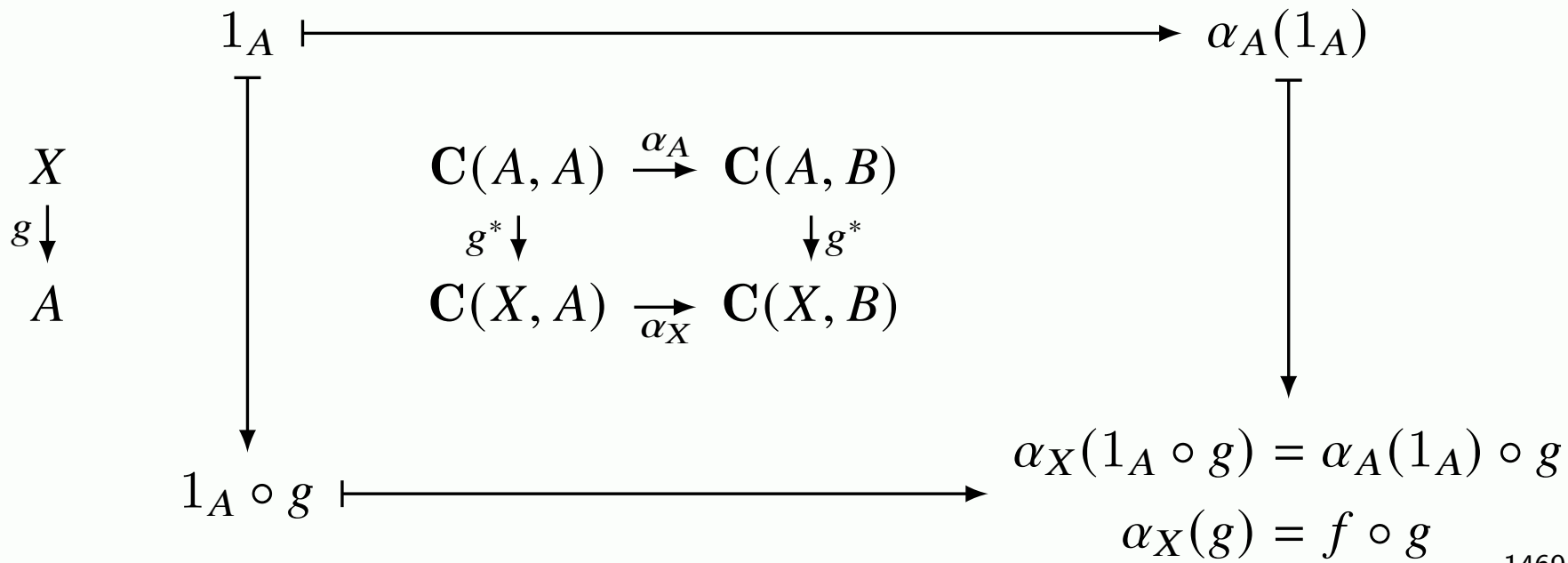
The Yoneda embedding functor $\mathsf{y} : \mathbf{C} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ *is full and faithful, and injective on objects. Hence,* y *is a full embedding* $\mathbf{C} \hookrightarrow \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$.

Proof.

Injectivity of $\mathbf{C}(A,B) \mapsto \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(\mathbf{C}(-,A),\mathbf{C}(-,B))$ given by $f \mapsto f_*$ is clear. The Yoneda lemma gives us surjectivity.

$$\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(\mathbf{C}(-,A),\mathbf{C}(-,B)) \cong \mathbf{C}(A,B)$$

For any natural transformation $\alpha : \mathbf{C}(-,A) \to \mathbf{C}(-,B)$, We need to find a morphism $f : A \to B$ so that $\alpha = f_*$. Let $f := \alpha_A(1_A)$.



Concrete Category & Abstract Category

Definition (Concrete Category & Abstract Category)

- A category $\mathbf{C}$ is *concrete* iff there is a faithful functor $F : \mathbf{C} \to \mathbf{Set}$.
- Categories that are not concrete are called *abstract categories*.

All small categories are concrete because of Yoneda lemma.

Corollary

For any locally small category $\mathbf{C}$*, any objects* $A, B \in \mathbf{C}$*,*
$A \cong B \iff \mathrm{y}A \cong \mathrm{y}B$*.*

Proof.

For any full and faithful functor F,

$$A \cong B \iff FA \cong FB$$

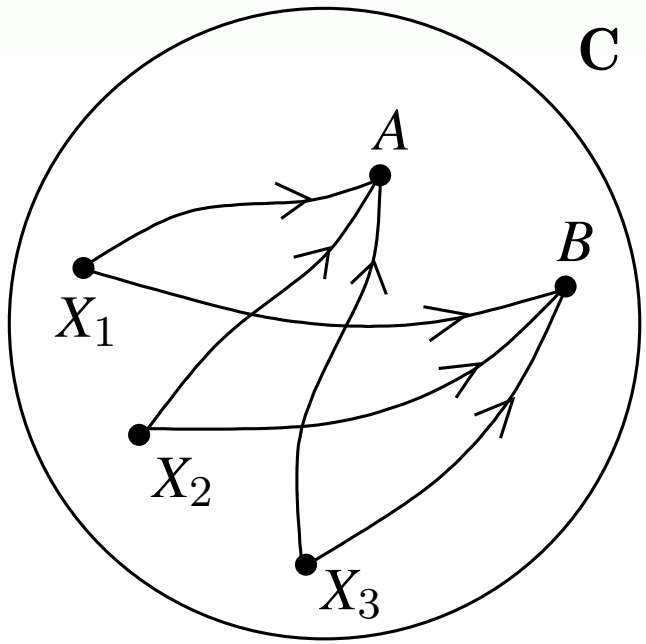
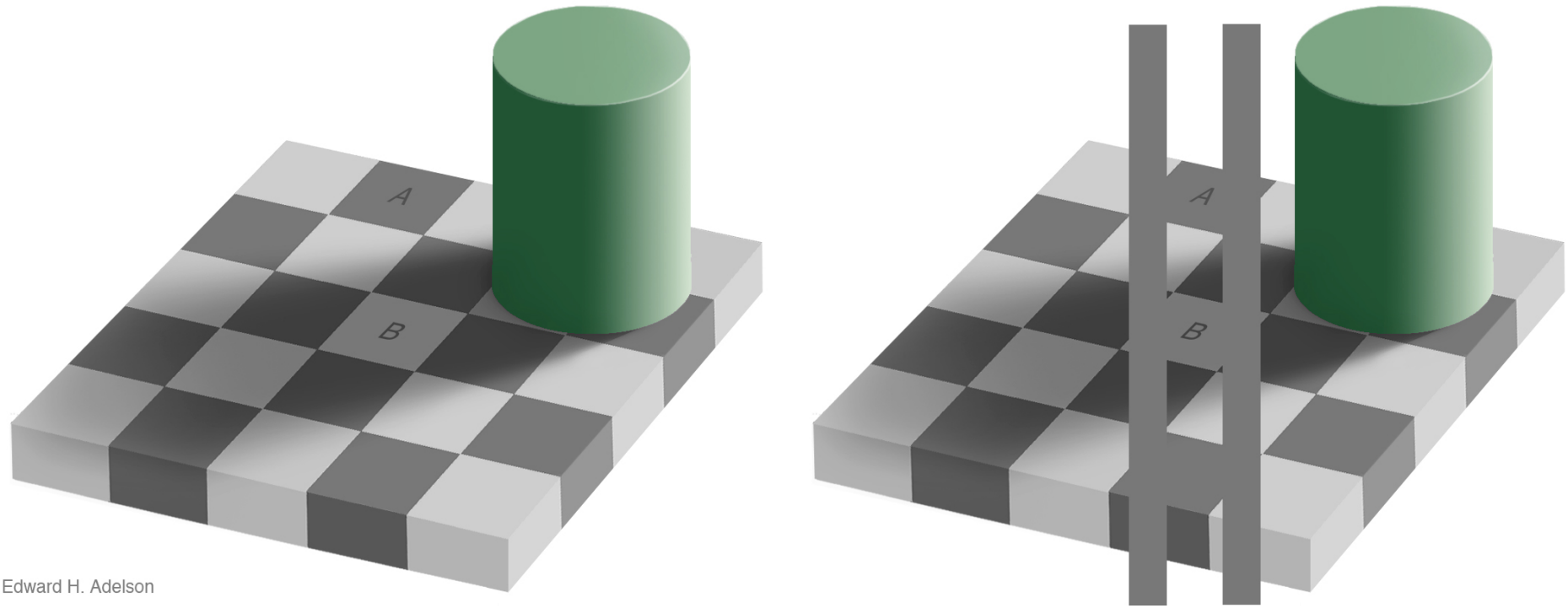


Figure: If $\mathrm{Hom}(X, A) \cong \mathrm{Hom}(X, B)$ naturally in X, then $A \cong B$.

Remark



Edward H. Adelson

Figure: If $\mathrm{Hom}(X, A) \cong \mathrm{Hom}(X, B)$ naturally in X, then $A \cong B$.

Application — Categorifying Cardinal Arithmetic

What is the meaning of the equation?

$$\forall abc \in \mathbb{N} : a \times (b + c) = (a \times b) + (a \times c)$$

- categorification (replacing equality by isomorphism)
- the Yoneda lemma (replacing isomorphism by natural isomorphism)
- representability (characterizing maps to or from an object)
- limits and colimits (like cartesian product and disjoint union)
- adjunctions (such as currying)

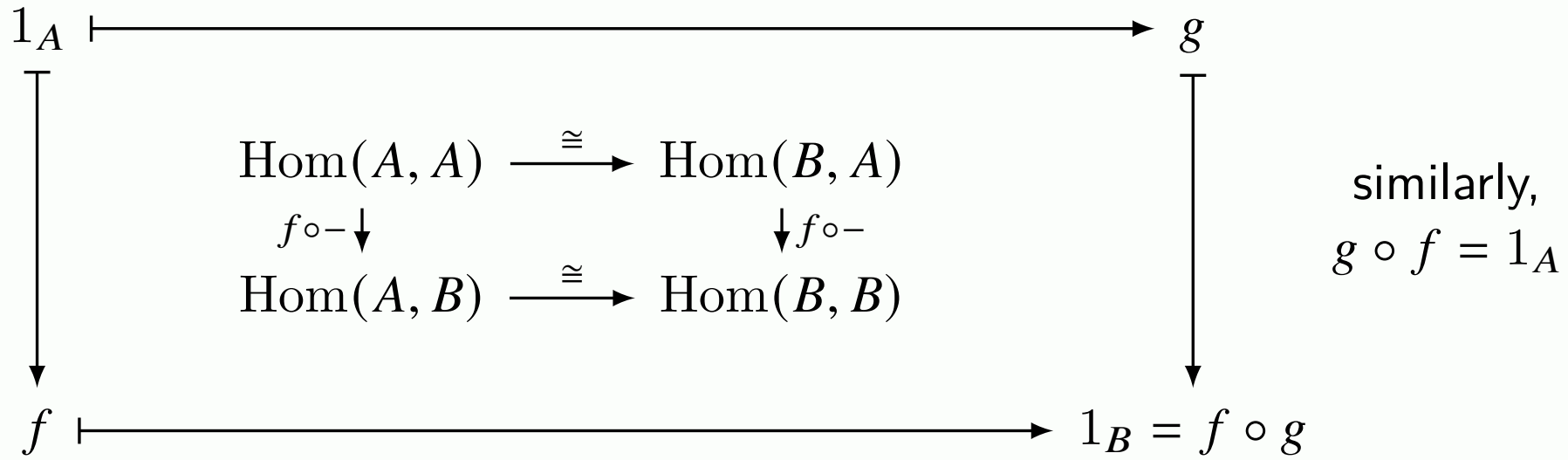
Lemma

Lemma (Yoneda Lemma)

$A \cong B \iff \mathrm{Hom}(A, X) \cong \mathrm{Hom}(B, X)$ *naturally in* X.

Proof.

Suppose $\mathrm{Hom}(A, X) \cong \mathrm{Hom}(B, X)$. Taking $X = A$ and $X = B$, we use the bijections $\mathrm{Hom}(A, A) \cong \mathrm{Hom}(B, A)$ and $\mathrm{Hom}(A, B) \cong \mathrm{Hom}(B, B)$ to define $f : A \to B$ and $g : B \to A$ such that $f \circ g = 1_B$ and $g \circ f = 1_A$.
By naturality:



Categorifying Cardinal Arithmetic

$$\boxed{a \times (b + c) = (a \times b) + (a \times c)}$$

Proof.

- pick sets A, B, C s.t. $a = |A|, b = |B|, c = |C|$.
- show that $A \times (B + C) \cong (A \times B) + (A \times C)$.
- by the Yoneda lemma, this holds iff

$$\mathrm{Hom}(A \times (B + C), X) \cong \mathrm{Hom}((A \times B) + (A \times C), X) \quad \text{naturally}$$

- now

$$\begin{aligned}
\mathrm{Hom}(A \times (B + C), X) &\cong \mathrm{Hom}(B + C, X^A) && \text{(curring)} \\
&\cong \mathrm{Hom}(B, X^A) \times \mathrm{Hom}(C, X^A) && \text{(pairing)} \\
&\cong \mathrm{Hom}(A \times B, X) \times \mathrm{Hom}(A \times C, X) && \text{(curring)} \\
&\cong \mathrm{Hom}((A \times B) + (A \times C), X) && \text{(pairing)}
\end{aligned}$$

Categorification

- Roughly speaking: categorification is a process that takes a set-theoretic concept and produces an analogous categorical-theoretic concept.
- Two Steps:
  1. to describe the set-theoretic concept to be categorified in terms of categorical structures in $\mathbf{Set}$;
  2. to *internalize* the defining categorical structures in $\mathbf{Cat}$.

Remark: *Internalization* is the process of taking math that lives in $\mathbf{Set}$ and moving it into some category $\mathbf{C}$.
We have to replace:

sets	categories
elements	objects
equations between elements	isomorphisms between objects
functions	functors
equations between functions	natural isomorphisms between functors

Categorification

- The general idea of categorification is that we take a thing we know and add structure to it, so that what were formerly properties become structures. We do this in such a way that we can recover the thing we categorified by forgetting this new structure.
- To categorify a set S is to find a category $\mathbf{C}$ and a map

$$F : \mathrm{Decat}(\mathbf{C}) \to S$$

  where $\mathrm{Decat}(\mathbf{C})$ is the set of isomorphism classes of objects of $\mathbf{C}$.
- Example: If we categorify the natural numbers $\mathbb{N}$ to the category of finite sets $\mathbf{FinSet}$, addition gets categorified to disjoint union, and equality to isomorphism. F is "cardinality".
- Another Example: $S = \mathbb{N}$ is the natural numbers, $\mathbf{C} = \mathbf{FdVect}$ is the category of finite-dimensional vector spaces, F is "dimension".

Remark: Choosing a good categorification — like designing a good algebraic structure such as that of preorders or quantales — is part of the art of mathematics.

Theorem

For a locally small category $\mathbf{C}$*, the Yoneda embedding* y *preserves all limits that exist in* $\mathbf{C}$*.*

Proof.

Suppose (L, λ) is a limit of $D : \mathbf{I} \to \mathbf{C}$. The Yoneda embedding maps D to the diagram $\mathrm{y}D : \mathbf{I} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ defined by $(\mathrm{y}D)_i = \mathrm{y}D_i = \mathrm{Hom}(-, D_i)$, and it maps (L, λ) to $(\mathrm{y}L, \mathrm{y}\lambda)$ on $\mathrm{y}D$ defined by $(\mathrm{y}\lambda)_i = \mathrm{y}\lambda_i = \mathrm{Hom}(-, \lambda_i)$.
To see that $(\mathrm{y}L, \mathrm{y}\lambda)$ is a limit cone on $\mathrm{y}D$, consider a cone (M, μ) on $\mathrm{y}D$. Then $\mu : \Delta M \to D$ consists of a family of functions, one for each $i \in \mathbf{I}$ and $A \in \mathbf{C}$,

$$(\mu_i)_A : MA \to \mathrm{Hom}(A, D_i)$$

For $A \in \mathbf{C}$ and $m \in MA$, we get a cone on D consisting of morphisms $(\mu_i)_A m : A \to D_i$. There exists a unique morphism $\varphi_A m : A \to L$ s.t. $(\mu_i)_A m = \lambda_i \circ \varphi_A m$. Then $\varphi_A : MA \to \mathrm{Hom}(A, L) = (\mathrm{y}L)A$ forms a unique factorization $\varphi : M \to \mathrm{y}L$. □

- $\mathrm{Hom}(A, \varprojlim D_i) \cong \varprojlim \mathrm{Hom}(A, D_i)$
- $\mathrm{Hom}(\varinjlim D_i, A) \cong \varprojlim \mathrm{Hom}(D_i, A)$

Cayley Theorem

Theorem (Cayley Theorem)

Every group G is isomorphic to a subgroup of the symmetric group on G.

Proof.

Any group G can be viewed as a single object $\bullet$ category, call it $\mathbf{G}$.

$$\mathbf{Set}^{\mathbf{G}^{\mathrm{op}}}(\mathbf{G}(-,\bullet),\mathbf{G}(-,\bullet)) \cong \mathbf{G}(\bullet,\bullet)$$

The right-hand side is just G.
y : $g \mapsto \mathbf{G}(-,g) : \mathbf{G}(\bullet,\bullet) \to \mathbf{G}(\bullet,\bullet)$ and $\mathbf{G}(-,g) : x \mapsto gx$.
$(\{\mathrm{y}g : g \in G\}, \circ, \mathrm{y}e)$ form a group.
Therefore, the left-hand side is a subgroup of the group of all permutations on G.
Moreover, this subgroup is isomorphic to the group G itself by the restricted Yoneda lemma. □

From Klein's Erlangen Program to Category Theory

- Klein[30] started with a geometry and looked at the group of transformations of that geometry.
- One possible generalization is to replace the geometry by a different structure X and consider its algebra of automorphisms $\mathrm{Aut}(X)$.

$$\mathrm{Aut}(X) \to \mathrm{End}(X) \to \mathrm{Hom}(X, Y)$$

$$\frac{\text{Space}}{\text{Transformation group}} \sim \frac{\text{Category}}{\text{Algebra of mappings}}$$

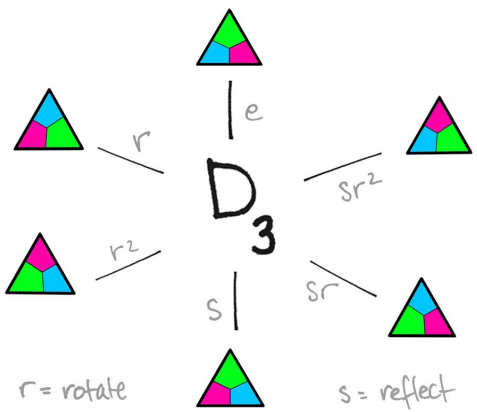
[30] Klein's Erlangen program: "Given a manifold, and a transformation group acting on it, to study its invariants."

- If geometric spaces are taken first, then groups, seen as systems of global properties of spaces, **supervene** upon geometric properties, that is properties definable in the language of the space, for instance via linear algebra.
- On the other hand, it is possible to reverse the dependence and instead consider groups as being fundamental and construct the spaces from them to look at their various representations.

- Given a vector space X, a group action $G \times X \to X$ can be seen as a group representation $G \to \mathrm{Aut}(X)$. A group representation provides a way to view the abstract group elements as concrete linear transformations of some vector space.
- For example, $D_3 \to \mathrm{Aut}(\mathbb{R}^2)$.

$$D_3 = \langle r, s \mid r^3 = s^2 = rsrs = e \rangle$$

$$r \mapsto R = \begin{bmatrix} \cos(2\pi/3) & -\sin(2\pi/3) \\ \sin(2\pi/3) & \cos(2\pi/3) \end{bmatrix} \qquad s \mapsto S = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$



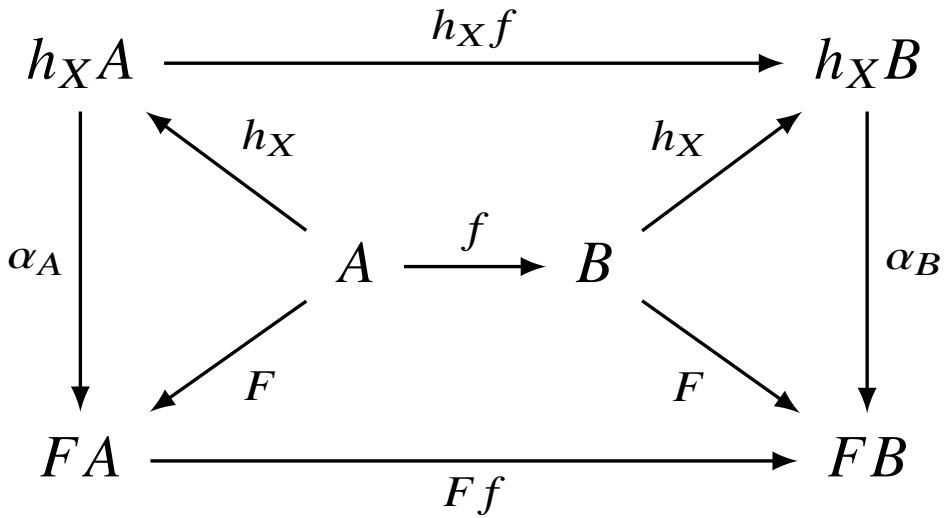
- We can think of a group G as providing the syntax while automorphisms $\mathrm{Aut}(X)$ provide the semantics. So a group representation is like a functor

$$F : \text{syntax} \to \text{semantics}$$

Representation

Definition (Representable Functor)

A functor $F : \mathbf{C} \to \mathbf{Set}$ is *representable* iff there is a natural isomorphism $\alpha : \mathrm{Hom}(A, -) \xrightarrow{\cong} F$ for some $A \in \mathbf{C}$. We say that the pair (A, α) is a *representation* of F.



Remark

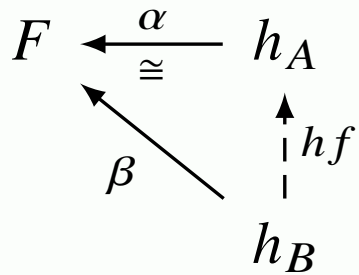
- $\mathbf{C}^{\mathbf{I}}(\Delta -, D)$ *is representable by* $\varprojlim D$*:* $\mathbf{C}^{\mathbf{I}}(\Delta -, D) \cong \mathbf{C}(-, \varprojlim D)$.
- $\mathbf{C}^{\mathbf{I}}(D, \Delta -)$ *is representable by* $\varinjlim D$*:* $\mathbf{C}^{\mathbf{I}}(D, \Delta -) \cong \mathbf{C}(\varinjlim D, -)$.

Theorem

若函子 $F : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$ 可表, 则其代表元 $(A, \alpha : h_A \xrightarrow{\cong} F)$ 在至多差一个唯一同构的意义下是唯一的.

Proof.

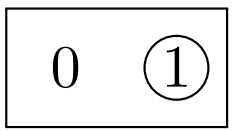
考虑逗号范畴 $(h \downarrow F)$, 其中 Hom 函子 $h : \mathbf{C} \to \widehat{\mathbf{C}} :: A \mapsto \mathrm{Hom}(-, A)$. 对于任意 $B \in \mathbf{C}$ 与 $\beta : h_B \to F$, 存在唯一的 $f \in \mathrm{Hom}_{\mathbf{C}}(B, A)$ 使下图交换.



可知 (A, α) 是 $(h \downarrow F)$ 的终对象. □

Example: Representable Functor

- There is a contravariant functor $\mathcal{O} : \mathbf{Top}^{\mathrm{op}} \to \mathbf{Set}$ that associates to each topological space X its set $\mathcal{O}(X)$ of open subsets. On morphisms, $\mathcal{O}$ takes a continuous function $f : X \to Y$ to $f^{-1} : \mathcal{O}(Y) \to \mathcal{O}(X)$.
- The functor $\mathcal{O}$ is *representable*.
- It is represented by the Sierpinski space $S := \big(\{0,1\}, \big\{\emptyset, \{1\}, \{0,1\}\big\}\big)$.



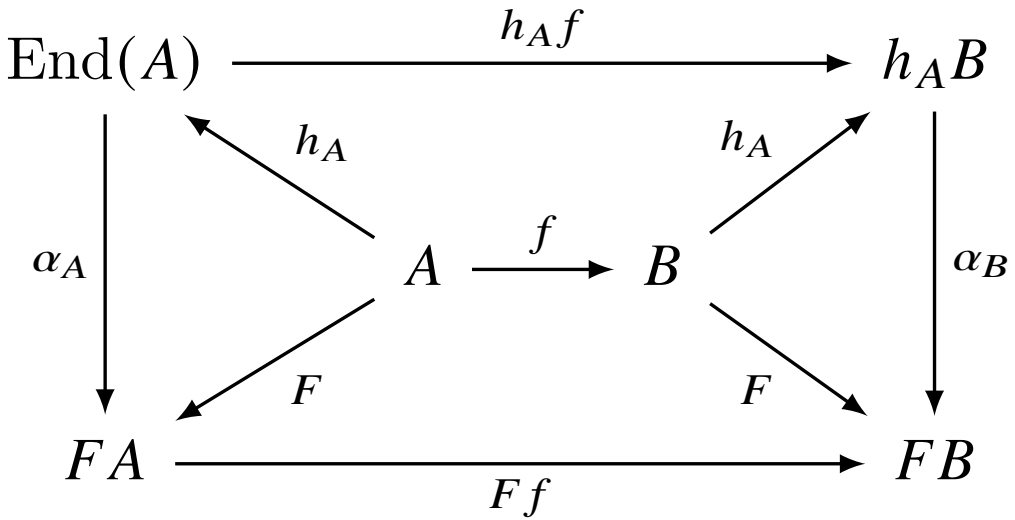
- For any topological space X, continuous functions $X \to S$ are in one-to-one correspondence with the open subsets of X.

$$\mathcal{O}(X) \underset{f^{-1}(\{1\}) \mapsfrom f}{\overset{U \mapsto \chi_U}{\rightleftarrows}} \mathrm{Hom}(X, S)$$

Universal Element

Definition (Universal Element)

A *universal element* of the functor $F : \mathbf{C} \to \mathbf{Set}$ is a pair (A, a), where $A \in \mathbf{C}$ and $a \in FA$, and for each $B \in \mathbf{C}$ and $b \in FB$, there is a unique map $f : A \to B$ such that $Ff(a) = b$.



- If (A, α) is a representation of $F : \mathbf{C} \to \mathbf{Set}$, then $(A, \alpha_A(1_A))$ is a universal element of F.
- The natural transformation $\psi(a) : \mathrm{Hom}(A, -) \to F$ induced by $a \in FA$ in Yoneda lemma is an isomorphism iff (A, a) is a universal element of F.

Universal Property

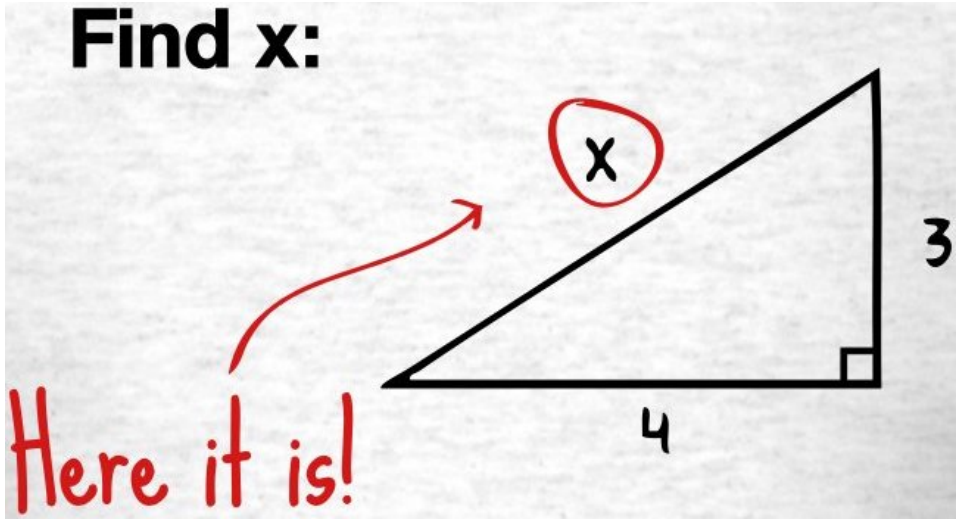
All the information contained in a representable functor $F : \mathbf{C} \to \mathbf{Set}$ is condensed in the representing object $A \in \mathbf{C}$ and the universal element $a \in FA$, which 'generates' all the other elements $b \in FB$ by applying functions of the form Ff to it.

Definition (Universal Property)

A *universal property* of an object $A \in \mathbf{C}$ is expressed by a representable functor $F : \mathbf{C} \to \mathbf{Set}$ together with a universal element $a \in FA$ that defines a natural isomorphism $\mathrm{Hom}(A, -) \cong F$ or $\mathrm{Hom}(-, A) \cong F$.

Philosophy — Object as a Solution to a Problem

Find x s.t. $x^2 + 1 = 0$.



Problem

For a functor $F : \mathbf{C} \to \mathbf{Set}$, find an object X of the category $\mathbf{C}$ together with $\alpha : \mathrm{Hom}(X, -) \cong F$.

Remark: If we find two solutions $\alpha : \mathrm{Hom}(X, -) \cong F$ and $\beta : \mathrm{Hom}(Y, -) \cong F$, then by Yoneda's lemma, $X \cong Y$, i.e. the object is unique up to unique isomorphism.

Philosophy — Objects as tokens for Eigenbehaviors

"We identify the world in terms of how we shape it. We shape the world in response to how it changes us. Objects arise as tokens of behavior that leads to seemingly unchanging forms. Forms are seen to be unchanging through their invariance under our attempts to change, to shape them."

— Louis H. Kauffman

如果一个东西看起来像鸭子、走路像鸭子、游泳像鸭子、叫起来像鸭子, 那么它就是鸭子.

Category of Elements

Definition (Category of Elements)

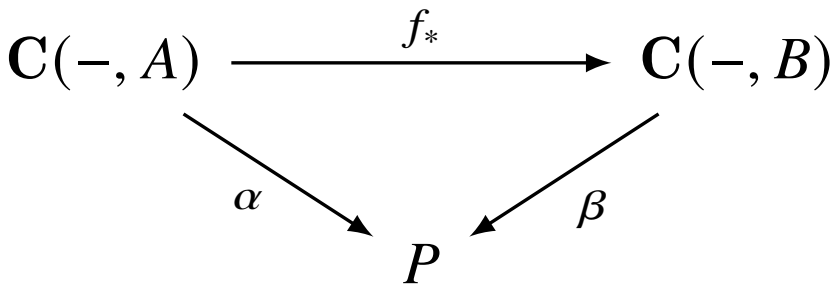
- The *category of elements* $\int^{\mathbf{C}} F$ of a covariant functor $F : \mathbf{C} \to \mathbf{Set}$ has
  1. as objects (A, a), where $A \in \mathbf{C}$ and $a \in FA$, and
  2. as morphisms $(A, a) \to (B, b)$ with $f : A \to B$ in $\mathbf{C}$ s.t. $Ff(a) = b$.
- The *category of elements* $\int_{\mathbf{C}} P$ of a contravariant functor $P : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$ has
  1. as objects (A, a) where $A \in \mathbf{C}$ and $a \in PA$, and
  2. as morphisms $(A, a) \to (B, b)$ with $f : A \to B$ in $\mathbf{C}$ s.t. $Pf(b) = a$.

A universal element can be viewed as an initial object in the category of elements of F.

Category of Elements

Theorem

- *For* $F : \mathbf{C} \to \mathbf{Set}$, $\int^{\mathbf{C}} F \cong \mathsf{y} \downarrow F$, *where* $\mathsf{y} : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}^{\mathbf{C}}$ *is the Yoneda embedding.*
- *For* $P : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$, $\int_{\mathbf{C}} P \cong \mathsf{y} \downarrow P$, *where* $\mathsf{y} : \mathbf{C} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ *is the Yoneda embedding.*



Theorem

- *A covariant set-valued functor is representable iff its category of elements has an initial object.*
- *A contravariant set-valued functor is representable iff its category of elements has a terminal object.*

Universal Property

Remark

Let $\mathbf{C}$ *be a locally small category and* $A \in \mathbf{C}$*. Then a (covariant) universal property of* $\mathbf{C}$ *is, equivalently:*

1. *a functor* $F : \mathbf{C} \to \mathbf{Set}$ *with a universal element* $(A \in \mathbf{C}, a \in FA)$.
2. *a functor* $F : \mathbf{C} \to \mathbf{Set}$ *with a representation* $(A, \alpha : \mathrm{Hom}_{\mathbf{C}}(A, -) \xrightarrow{\cong} F)$.
3. *a functor* $F : \mathbf{C} \to \mathbf{Set}$ *with an initial object* $(A, a) \in \int^{\mathbf{C}} F$.

The category of elements $\int_{\mathbf{C}} P$ has an evident projection functor:

$$\pi_P : \int_{\mathbf{C}} P \to \mathbf{C} :: (A, a) \mapsto A$$

Theorem

If $F : \mathbf{C} \to \mathbf{D}$ is a functor from a small category $\mathbf{C}$ to a cocomplete category $\mathbf{D}$, the functor $R : \mathbf{D} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ given by

$$R(B) : A \mapsto \mathrm{Hom}_{\mathbf{D}}(FA, B)$$

has a left adjoint $L : \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}} \to \mathbf{D}$ given by

$$L(P) = \varinjlim \left(\int_{\mathbf{C}} P \xrightarrow{\pi_P} \mathbf{C} \xrightarrow{F} \mathbf{D} \right)$$

In other words,

$$\mathrm{Hom}_{\mathbf{D}}(LP, B) \cong \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(P, RB) \qquad L \dashv R$$

Proof.

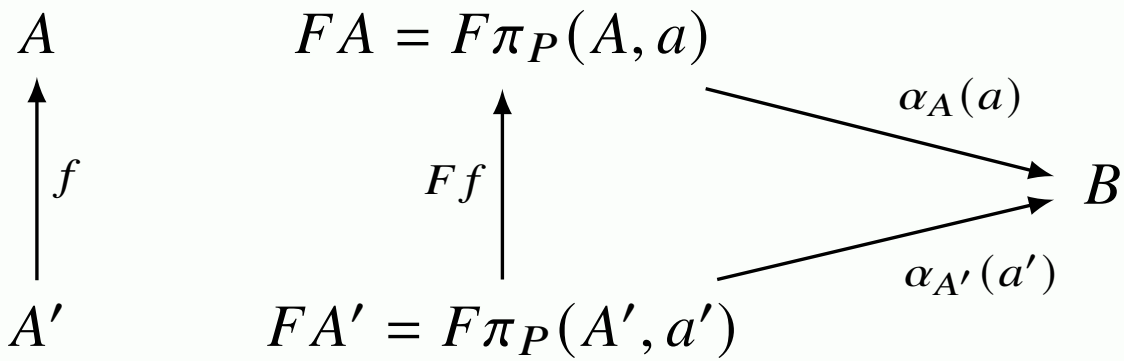
For a natural transformation $\alpha : P \to RB$,

$$\alpha_A : PA \to \mathbf{D}(FA, B)$$

$\{\alpha_A\}_{A \in \mathbf{C}}$ is natural in A.

$$\begin{array}{ccc} A & PA & \xrightarrow{\alpha_A} & \mathbf{D}(FA, B) \\ \uparrow f & Pf \downarrow & & \downarrow \mathbf{D}(Ff, B) \\ A' & PA' & \xrightarrow[\alpha_{A'}]{} & \mathbf{D}(FA', B) \end{array}$$

Such an α can also be considered as $\{\alpha_A(a) : FA \to B\}_{(A,a) \in \int_{\mathbf{C}} P}$. Then



This means that (B, α_A) constitute a cocone over $F\pi_P : \int_{\mathbf{C}} P \to \mathbf{D}$.
Each such cocone comes by composing the colimiting cocone with a unique arrow from the colimit LP to the object B. In other words,

$$\mathrm{Hom}_{\mathbf{D}}(LP, B) \cong \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(P, RB)$$

Corollary

Every presheaf is a colimit of representable presheaves.

Proof.

By the Yoneda lemma,

$$R_{\mathrm{y}}(B)(A) = \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(\mathrm{y}A, B) \cong B(A)$$

this means that R_{y} is isomorphic to the identity functor of $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$.
Its left adjoint L must also be isomorphic to the identity functor.

$$P \cong \varinjlim \left(\int_{\mathbf{C}} P \xrightarrow{\pi_P} \mathbf{C} \xrightarrow{\mathrm{y}} \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}} \right)$$

□

Example: Category of Elements

- Let A be a set, and consider it as a discrete category. A functor $S : A \to \mathbf{Set}$ is the same thing as an A-indexed set. For each $a \in A$, write $S_a := S(a)$.
- What is the category of elements of a functor $S : A \to \mathbf{Set}$? The objects of $\int_A S$ are pairs (a, s), where $a \in A$ and $s \in Sa$. Since A has nothing but identity morphisms, $\int_A S$ has nothing but identity morphisms, i.e., it is the discrete category on a set. In fact, that set is the disjoint union

$$\int_A S = \coprod_{a \in A} S_a$$

The functor $\pi_S : \int_A S \to A$ sends each element in S_a to the element $a \in A$.

Example

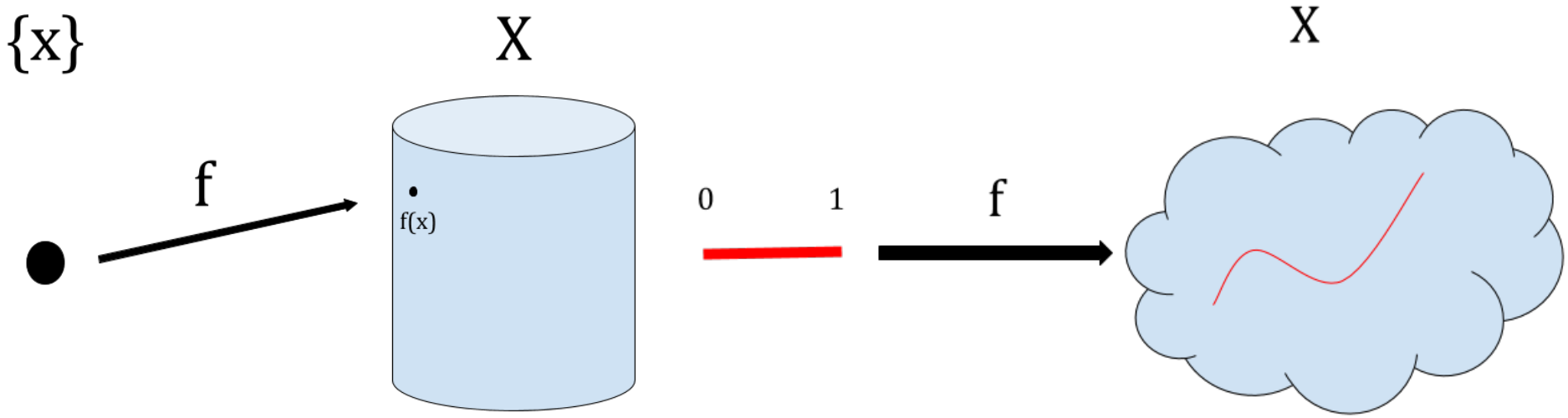
Let $A = \{\texttt{BOS}, \texttt{NYC}, \texttt{LA}, \texttt{DC}\}$, and let $S : A \to \mathbf{Set}$

$$S_{\texttt{BOS}} = \{\texttt{Abby}, \texttt{Bob}, \texttt{Carl}\}$$
$$S_{\texttt{NYC}} = \emptyset$$
$$S_{\texttt{LA}} = \{\texttt{John}, \texttt{Tom}\}$$
$$S_{\texttt{DC}} = \{\texttt{Abby}, \texttt{Sam}\}$$

$$\int_A S = \left\{ \begin{array}{lll} (\texttt{BOS}, \texttt{Abby}) & & \\ (\texttt{BOS}, \texttt{Bob}) & (\texttt{LA}, \texttt{John}) & (\texttt{DC}, \texttt{Abby}) \\ (\texttt{BOS}, \texttt{Carl}) & (\texttt{LA}, \texttt{Tom}) & (\texttt{DC}, \texttt{Sam}) \end{array} \right\}$$

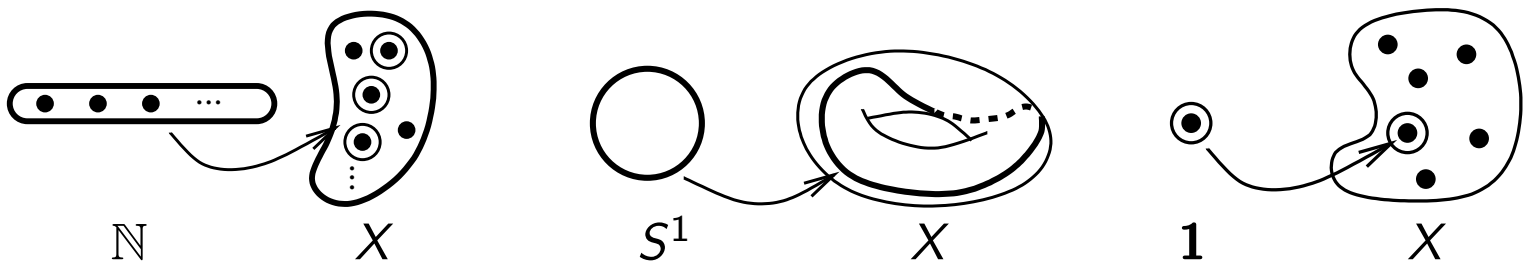
$$\downarrow \pi_S$$

$$A = \{ \texttt{BOS} \quad \texttt{NYC} \quad \texttt{LA} \quad \texttt{DC} \}$$



Suppose X and Y are topological spaces and let $\bullet$ denote the one-point space and I and S^1 the unit interval and the circle. Then,

- X and Y have the same cardinality iff $\mathrm{Hom}(\bullet, X) \cong \mathrm{Hom}(\bullet, Y)$.
- X and Y have the same path space iff $\mathrm{Hom}(I, X) \cong \mathrm{Hom}(I, Y)$.
- X and Y have the same loop space iff $\mathrm{Hom}(S^1, X) \cong \mathrm{Hom}(S^1, Y)$.
- Probing X and Y with various spaces gives us more information.
- Probing them with all spaces gives us all information.



a sequence / a loop / an element

Remarks — Why Yoneda?

- Given any objects A, B in a locally small category $\mathbf{C}$, to find a morphism $h : A \longrightarrow B$ it suffices to give one $\theta : \mathrm{y}A \longrightarrow \mathrm{y}B$ in $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$, for then there is a unique h with $\theta = \mathrm{y}h$.
- Why should it be easier to give $\mathrm{y}A \longrightarrow \mathrm{y}B$ than $A \longrightarrow B$?
- The key difference is that in general $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ has much more structure to work with than does $\mathbf{C}$.
- The category $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ is a topos. It is complete, cocomplete, cartesian closed, and has a subobject classifier.
- This means we have notions of conjunction, disjunction, and implication....
- It is like an extension of $\mathbf{C}$ by "ideal elements" that permit calculations which cannot be done in $\mathbf{C}$.
- This is something like passing to the complex numbers to solve equations in the reals, or adding higher types to an elementary logical theory.

Weakly Initial / Terminal Object

- A weakly initial/terminal object is an object which has an arrow coming to/from every other object.
- A weakly initial/terminal object is initial/terminal if the arrow is unique.
- A weakly initial set is a family of objects $(A_i)_{i \in I}$ such that, for every object $X \in \mathbf{C}$ there exists some $i \in I$ and an arrow $A_i \xrightarrow{x_i} X$.
- A weakly terminal set is a family of objects $\{T_i\}_{i \in I}$ such that, for every object $X \in \mathbf{C}$ there exists some $i \in I$ and an arrow $X \xrightarrow{x_i} T_i$.

Theorem. If a cocomplete locally small category has a weakly terminal set, it has the terminal object.

Proof. Given a weakly terminal set, $\coprod_{i \in I} T_i$ is a weakly terminal object, since there is an arrow to it from every object $X \xrightarrow{x_i} T_i \xrightarrow{\iota_i} \coprod_{i \in I} T_i$.

Given a weakly terminal object, we first define a full subcategory $\mathbf{T} \hookrightarrow \mathbf{C}$ whose objects are T_i.

Then the colimit of the inclusion functor $i : \mathbf{T} \hookrightarrow \mathbf{C}$ is the terminal object.

Digression — Mazzola's Path to Creativity

1. Exhibiting the open question = to understand the object X
2. Identifying the semiotic context = to describe the category $\mathbf{C}$ of which X is an object
3. Finding the question's critical concept in the semiotic context = X
4. Identifying the concept's walls = the uncontrolled behaviour of the Yoneda functor $\mathrm{y}X = \mathrm{Hom}_{\mathbf{C}}(-, X)$
5. Opening the walls and displaying its new perspectives = finding a subcategory $\mathbf{A}$ of $\mathbf{C}$ such that: $\mathrm{y}_{|\mathbf{A}} : \mathbf{C} \to \mathbf{Set}^{\mathbf{A}^{\mathrm{op}}} :: X \mapsto \mathrm{Hom}_{\mathbf{A}}(-, X)$ is full and faithful, and for any $X \in \mathbf{C}$ there is a diagram D in $\mathbf{A}$ whose colimit $\varinjlim D \cong X$
6. Evaluating the extended walls = try to understand X via the isomorphism $\varinjlim D \cong X$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Adjunction

Definition (Left/Right Adjoint)

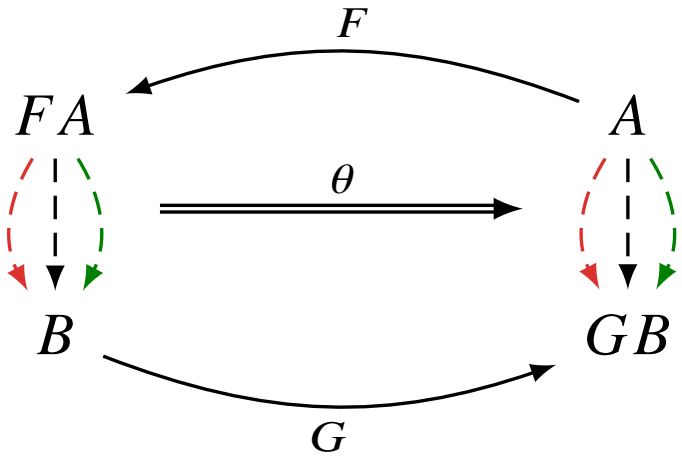
Functors $\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$ are *adjoint* $F \dashv G$ iff there is an isomorphism

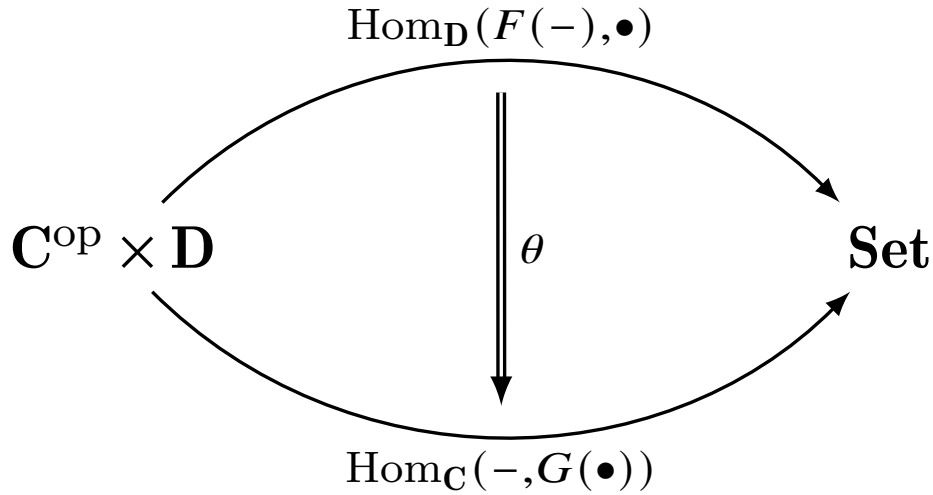
$$\theta_{A,B} : \mathbf{D}(FA, B) \overset{\cong}{\longrightarrow} \mathbf{C}(A, GB)$$

that is natural in both A and B.

'Natural in both A and B' means that, $\mathbf{D}(FA, -) \cong \mathbf{C}(A, G-)$ for A, and $\mathbf{D}(F-, B) \cong \mathbf{C}(-, GB)$ for B, i.e., for $g : B \to B'$ and $f : A \to A'$,

$$\begin{array}{ccc} \mathbf{D}(FA, B) & \xrightarrow{\theta_{A,B}} & \mathbf{C}(A, GB) \\ {\scriptstyle \mathbf{D}(FA,g)}\downarrow & & \downarrow{\scriptstyle \mathbf{C}(A,Gg)} \\ \mathbf{D}(FA, B') & \xrightarrow[\theta_{A,B'}]{} & \mathbf{C}(A, GB') \end{array} \qquad \begin{array}{ccc} \mathbf{D}(FA', B) & \xrightarrow{\theta_{A',B}} & \mathbf{C}(A', GB) \\ {\scriptstyle \mathbf{D}(Ff,B)}\downarrow & & \downarrow{\scriptstyle \mathbf{C}(f,GB)} \\ \mathbf{D}(FA, B) & \xrightarrow[\theta_{A,B}]{} & \mathbf{C}(A, GB) \end{array}$$

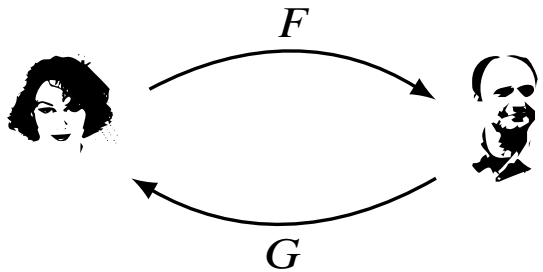




For any $A, B \in \mathbf{C}$, and $f : A \to A'$, $g : B \to B'$, $h : FA \to B$,

$$\begin{array}{ccc}
\mathrm{Hom}_{\mathbf{D}}(FA, B) & \xrightarrow{\mathrm{Hom}_{\mathbf{D}}(Ff,g)} & \mathrm{Hom}_{\mathbf{D}}(FA', B') \\
\ni & & \in \\
h & \longmapsto & g \circ h \circ Ff \\
\big\downarrow & & \big\downarrow \\
\theta_{AB}h & \longmapsto & \theta_{A'B'}(g \circ h \circ Ff) = Gg \circ \theta_{AB}h \circ f \\
\theta_{AB} \big\downarrow & & \big\downarrow \theta_{A'B'} \\
\ni & & \in \\
\mathrm{Hom}_{\mathbf{C}}(A, GB) & \xrightarrow[\mathrm{Hom}_{\mathbf{C}}(f,Gg)]{} & \mathrm{Hom}_{\mathbf{C}}(A', GB')
\end{array}$$

Remark



category = country

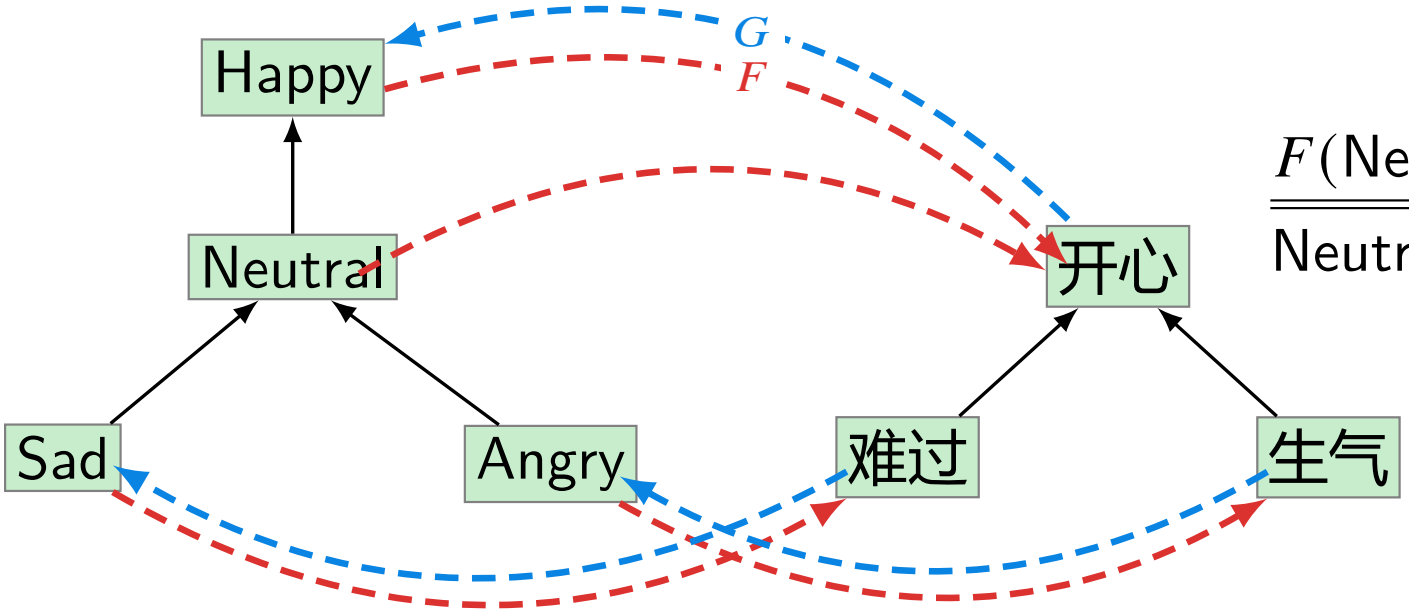
object = citizen

morphism = speaking in country's language

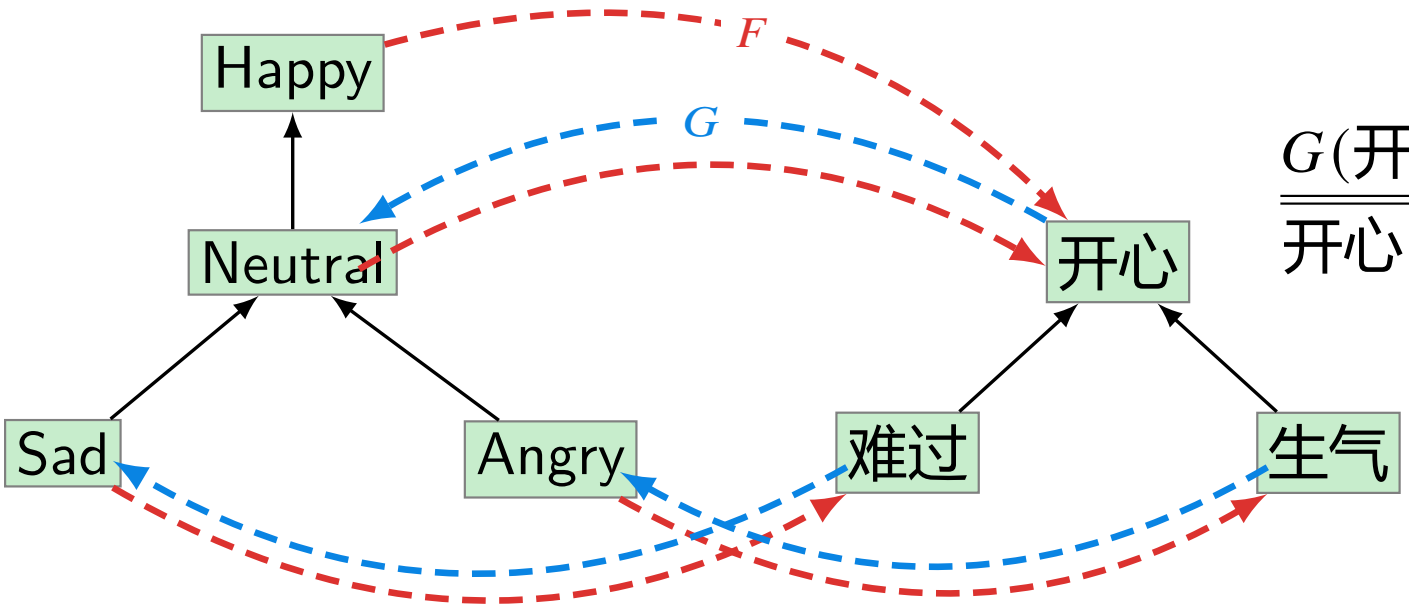
functor = translation

- Equivalence: doesn't matter whether I travel to you and speak your language, or you travel to me and speak my language.
- Some things get lost in translation; adjunction is next best thing.

Left/Right Adjoint — Translation



$$\frac{F(\text{Neutral}) \leq \text{开心}}{\text{Neutral} \leq G(\text{开心})} \quad F \dashv G$$



$$\frac{G(\text{开心}) \leq \text{Happy}}{\text{开心} \leq F(\text{Happy})} \quad G \dashv F$$

Theorem

Given $\mathbf{D}(FA, B) \underset{\sharp}{\overset{\flat}{\rightleftarrows}} \mathbf{C}(A, GB)$ *, the naturality condition of the adjunction implies that for every* $f : A \to A'$, $g : B \to B'$, $h^\sharp : FA \to B$, $k^\sharp : FA' \to B'$,

$$\begin{array}{ccc} A & \xrightarrow{h^\flat} & GB \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle Gg} \\ A' & \xrightarrow[k^\flat]{} & GB' \end{array} \quad \Longleftrightarrow \quad \begin{array}{ccc} FA & \xrightarrow{h^\sharp} & B \\ {\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle g} \\ FA' & \xrightarrow[k^\sharp]{} & B' \end{array}$$

Proof.

$$\begin{array}{ccc} \mathbf{D}(FA', B') & \xrightarrow{\flat} & \mathbf{C}(A', GB') \\ {\scriptstyle -\circ Ff}\downarrow & & \downarrow{\scriptstyle -\circ f} \\ \mathbf{D}(FA, B') & \xrightarrow[\flat]{} & \mathbf{C}(A, GB') \end{array} \qquad \begin{array}{ccc} \mathbf{D}(FA, B) & \xrightarrow{\flat} & \mathbf{C}(A, GB) \\ {\scriptstyle g\circ -}\downarrow & & \downarrow{\scriptstyle Gg\circ -} \\ \mathbf{D}(FA, B') & \xrightarrow[\flat]{} & \mathbf{C}(A, GB') \end{array}$$

$$k^\flat \circ f = (k^\sharp \circ Ff)^\flat \qquad Gg \circ h^\flat = (g \circ h^\sharp)^\flat$$

Left/Right Adjoint

Example

$$\frac{A \wedge B \vdash C}{A \vdash B \to C}$$

$$F_B : X \mapsto X \wedge B \qquad G_B : X \mapsto B \to X \qquad F_B \dashv G_B$$

Example

$$\frac{A \cap B \subset C}{A \subset B \to C} \qquad \text{where } B \to C := \overline{B} \cup C$$

$$F_B : X \mapsto X \cap B \qquad G_B : X \mapsto B \to X \qquad F_B \dashv G_B$$

Example

$$\frac{A \setminus B \subset C}{A \subset B \cup C}$$

$$F_B : X \mapsto X \setminus B \qquad G_B : X \mapsto B \cup X \qquad F_B \dashv G_B$$

Left/Right Adjoint

Example $\mathbf{Cost} := ([0, \infty], \geq, +, 0)$

$$\frac{A + B \geq C}{A \geq B \to C}$$

where $B \to C := \max\{0, C - B\}$.

$$F_B : X \mapsto X + B \qquad G_B : X \mapsto B \to X \qquad F_B \dashv G_B$$

Example

Consider the inclusion map $i : \mathbb{Z} \hookrightarrow \mathbb{R}$.
This has both a left adjoint $\lceil \rceil$ and a right adjoint $\lfloor \rfloor$. For $z \in \mathbb{Z}, r \in \mathbb{R}$:

$$\frac{r \leq i(z)}{\lceil r \rceil \leq z} \qquad \frac{i(z) \leq r}{z \leq \lfloor r \rfloor} \qquad \lceil \rceil \dashv i \dashv \lfloor \rfloor$$

Left/Right Adjoint

Example

$$\mathrm{Hom}(A \times B, C) \cong \mathrm{Hom}(A, C^B)$$

$$F_B : X \mapsto X \times B \qquad G_B : X \mapsto X^B \qquad F_B \dashv G_B$$

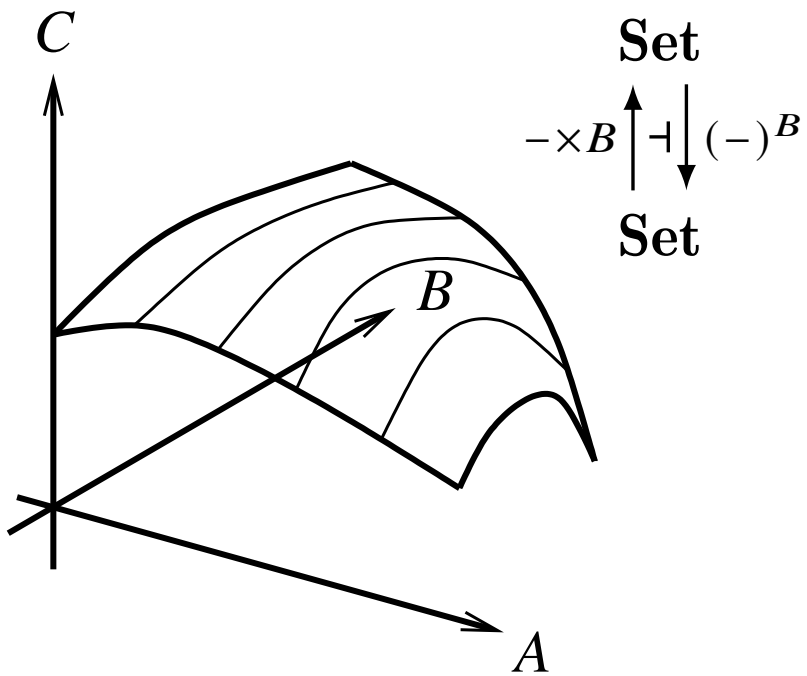


Figure: In **Set**, a map $A \times B \to C$ can be seen as a way of assigning to each element of A a map $B \to C$.

Free ⊣ Forgetful

Definition (Free Monoid)

Let X be a set. The free monoid generated by X is $F_X := (\mathrm{List}(X), *, [])$, where $\mathrm{List}(X)$ is the set of lists of elements in X, $[]$ is the empty list, and $*$ is the operation of list concatenation.

- The functor $F : \mathbf{Set} \to \mathbf{Mon}$ is a left adjoint of the forgetful functor $U : \mathbf{Mon} \to \mathbf{Set}$.
- $UF : \mathbf{Set} \to \mathbf{Set}$ encodes the free monoid construction inside Set.

Free ⊣ Forgetful

Example

The forgetful functor $U : \mathbf{Grp} \to \mathbf{Set}$ has as a left adjoint $F : \mathbf{Set} \to \mathbf{Grp}$ which sends a set to the free group on that set. $F \dashv U$.

Example

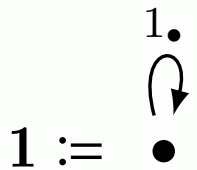
- The forgetful functor $\mathrm{ob} : \mathbf{Cat} \to \mathbf{Set}$ has a left adjoint Disc sending A to the discrete category whose objects are the members of A.
- And $\mathrm{ob} : \mathbf{Cat} \to \mathbf{Set}$ has a right adjoint Indisc sending A to the preorder with objects $a \in A$ and one morphism $a \to b$ for each pair $(a, b) \in A \times A$.
- The functor Disc also has a left adjoint π_0 sending $\mathbf{C}$ to its set of connected components, i.e. the quotient of $\mathrm{ob}\,\mathbf{C}$ by the equivalence $A \sim B$ whenever there exists a morphism $A \to B \in \mathbf{C}$.

$$\pi_0 \dashv \mathrm{Disc} \dashv \mathrm{ob} \dashv \mathrm{Indisc}$$

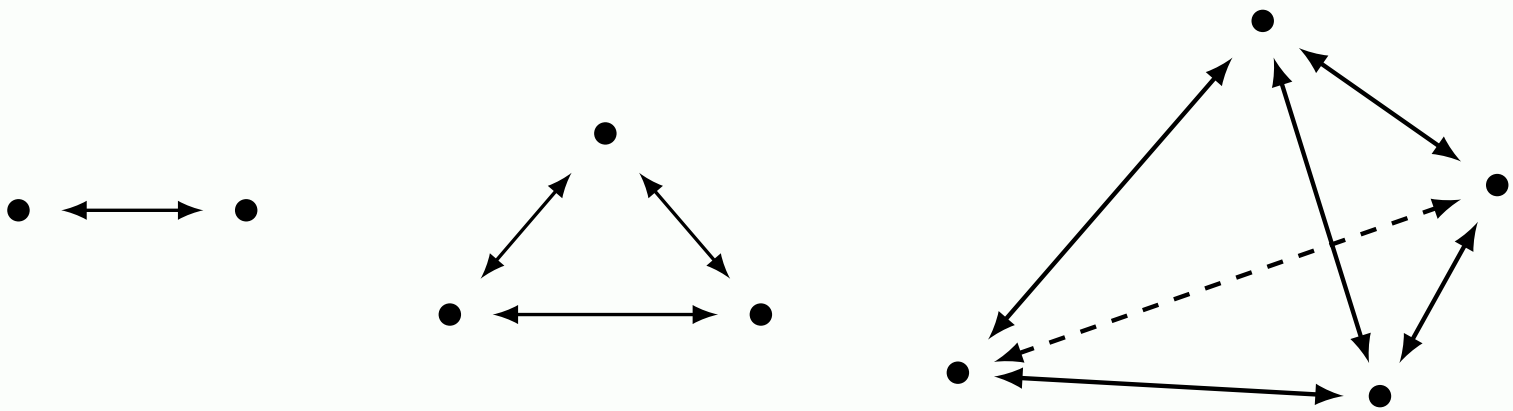
Indiscrete/Chaotic Categories

Definition (Indiscrete/Chaotic Category)

A category is indiscrete/chaotic iff it is equivalent to the terminal category



Example



Example

- The underlying graph functor $\mathbf{Cat} \to \mathbf{Graph}$ has a left adjoint $\mathbf{Graph} \to \mathbf{Cat}$ given by the free category.
- The inclusion $\mathbf{Grp} \hookrightarrow \mathbf{Mon}$ has a right adjoint $\mathbf{Mon} \xrightarrow{\text{core}} \mathbf{Grp}$, called the core, that sends a monoid to its subgroup of invertible elements.
- The forgetful functor from partial orders to preorders has a left adjoint given by quotienting out the cliques, where a clique is a subset $S \subset A$ such that $\forall xy \in S : x \leq y$.

Left/Right Adjoint

Example

Let X be the poset of subsets of $\mathbb{R}^2$, ordered by inclusion. Let Y be the poset of *convex* subsets of $\mathbb{R}^2$.
The *convex hull* of a subset $A \subset \mathbb{R}^2$ is defined as either

- The smallest convex subset of $\mathbb{R}^2$ containing A;
- The intersection of all convex subsets of $\mathbb{R}^2$ containing A;
- The set obtained by closing A under all possible convex combinations.

Let $c : X \to Y$ be the map assigning to each $A \in X$ its convex hull.

$$\frac{c(A) \subset B}{A \subset i(B)}$$

Topological interior as an adjoint

- A *topological space* $(X, \mathcal{O}_X)$ is a set X with a family $\mathcal{O}_X \subset \mathrm{P}(X)$ of subsets of X which contains $\emptyset$ and X, and is closed under finite intersections and arbitrary unions.
- The topological *interior* of a subset $S \subset X$ is

$$S^\circ := \bigcup \{U \in \mathcal{O}_X : U \subset S\}$$

- For $U \in \mathcal{O}_X$ and $S \in \mathrm{P}(X)$, topological interior is a right adjoint to the inclusion of $\mathcal{O}_X$ into $\mathrm{P}(X)$.

$$\frac{i\,U \subset S}{U \subset S^\circ}$$

$$\mathcal{O}_X \underset{()^\circ}{\overset{i}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathrm{P}(X)$$

$0_{\mathbf{C}} \dashv !_{\mathbf{C}} \dashv 1_{\mathbf{C}}$

- The terminal object of $\mathbf{CAT}$ is the category $\mathbf{1}$ containing just one object and one morphism 

.
- For any category $\mathbf{C}$, there exists a unique functor $!_{\mathbf{C}} : \mathbf{C} \to \mathbf{1}$, which maps every object $A \in \mathbf{C}$ to $\bullet$.
- An object $A \in \mathbf{C}$ can be viewed as a functor $A : \mathbf{1} \to \mathbf{C}$, i.e. $A : \bullet \to A$ and $A : 1_\bullet \to 1_A$.
- Then the terminal object $1_{\mathbf{C}}$ of $\mathbf{C}$ is the right adjoint of $!_{\mathbf{C}} : \mathbf{C} \to \mathbf{1}$, for the corresponding functor $1_{\mathbf{C}} : \mathbf{1} \to \mathbf{C}$ has the property that, for every $A \in \mathbf{C}$ we have a trivial natural bijective correspondence:

$$\frac{1_\bullet : !_{\mathbf{C}}A \to \bullet}{!_A : A \to 1_{\mathbf{C}}\bullet} \qquad \text{similarly,} \qquad \frac{1_\bullet : \bullet \to !_{\mathbf{C}}A}{!_A : 0_{\mathbf{C}}\bullet \to A}$$

$$\mathbf{1}(!_{\mathbf{C}}A, \bullet) \cong \mathbf{C}(A, 1_{\mathbf{C}}\bullet) \qquad \mathbf{C}(0_{\mathbf{C}}\bullet, A) \cong \mathbf{1}(\bullet, !_{\mathbf{C}}A)$$

$$0_{\mathbf{C}} \dashv !_{\mathbf{C}} \dashv 1_{\mathbf{C}}$$

$+ \dashv \Delta \dashv \times$

Theorem

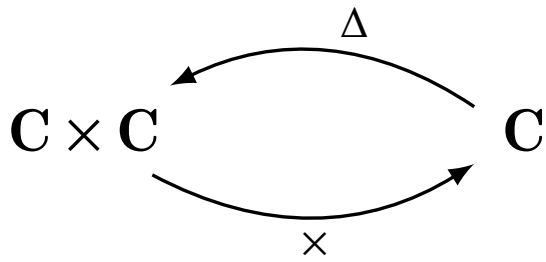
1. Δ *has a right adjoint iff* $\mathbf{C}$ *has binary products, and the right adjoint is* $\times : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$.
2. Δ *has a left adjoint iff* $\mathbf{C}$ *has binary coproducts, and the left adjoint is* $+ : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$.

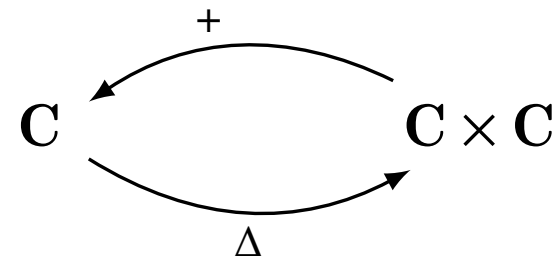
$$\mathbf{C} \times \mathbf{C}(\Delta X, (A, B)) \cong \mathbf{C}(X, A) \times \mathbf{C}(X, B) \cong \mathbf{C}(X, A \times B)$$

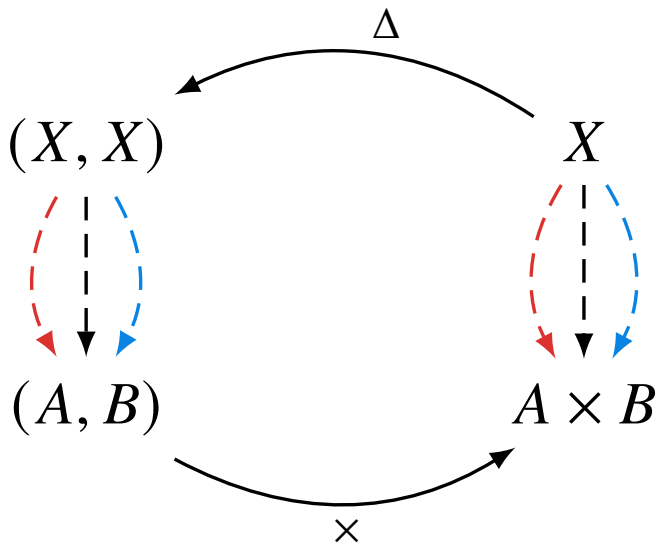
$$\mathbf{C}(A + B, X) \cong \mathbf{C}(A, X) \times \mathbf{C}(B, X) \cong \mathbf{C} \times \mathbf{C}((A, B), \Delta X)$$

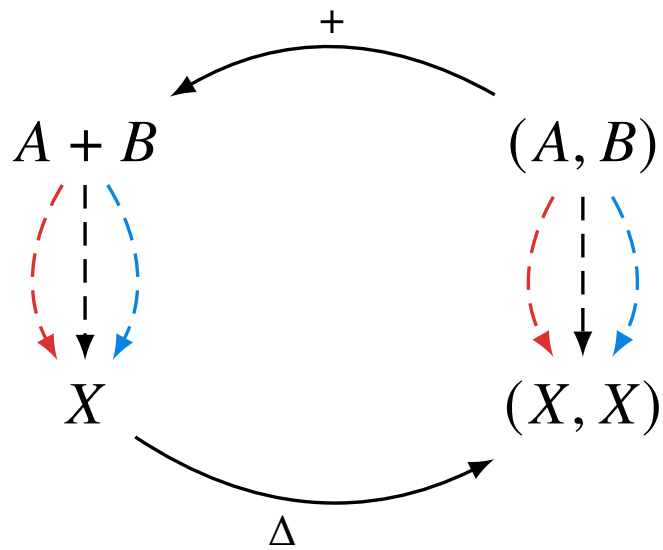
$$\begin{array}{ccccc} X & \xleftarrow{\pi_1} & X \times Y & \xrightarrow{\pi_2} & Y \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f \times g} & & \downarrow{\scriptstyle g} \\ A & \xleftarrow[\pi_1]{} & A \times B & \xrightarrow[\pi_2]{} & B \end{array} \qquad \begin{array}{ccccc} X & \xleftarrow{\pi_1} & X \times X & \xrightarrow{\pi_2} & X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f \times g} & & \downarrow{\scriptstyle g} \\ A & \xleftarrow[\pi_1]{} & A \times B & \xrightarrow[\pi_2]{} & B \end{array} \qquad \frac{(f, g) : \Delta X \to (A, B)}{f \times g : X \to A \times B}$$

$$\begin{array}{ccccc} X & \xrightarrow{\iota_1} & X + Y & \xleftarrow{\iota_2} & Y \\ {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle f + g} & & \uparrow{\scriptstyle g} \\ A & \xrightarrow[\iota_1]{} & A + B & \xleftarrow[\iota_2]{} & B \end{array} \qquad \begin{array}{ccccc} X & \xrightarrow{\iota_1} & X + X & \xleftarrow{\iota_2} & X \\ {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle f + g} & & \uparrow{\scriptstyle g} \\ A & \xrightarrow[\iota_1]{} & A + B & \xleftarrow[\iota_2]{} & B \end{array} \qquad \frac{f + g : A + B \to X}{(f, g) : (A, B) \to \Delta X}$$


Δ
C × C
C
×


+
C
C × C
Δ


Δ
(X, X)
X
(A, B)
A × B
×


+
A + B
(A, B)
X
(X, X)
Δ

Cartesian Closed Category

Theorem (Cartesian Closed Category)

A category $\mathbf{C}$ *is cartesian closed iff, the following functors have right adjoints:*

$$!_{\mathbf{C}} : \mathbf{C} \to \mathbf{1}$$
$$\Delta : \mathbf{C} \to \mathbf{C} \times \mathbf{C}$$
$$(- \times A) : \mathbf{C} \to \mathbf{C}$$

Here $!_{\mathbf{C}}$ *is the unique functor from* $\mathbf{C}$ *to the terminal category* $\mathbf{1}$ *and* Δ *is the diagonal functor* $\Delta A = \langle A, A \rangle$*, and the right adjoint of* $- \times A$ *is exponentiation by* A*.*

$\vee \dashv \Delta \dashv \wedge$

$\vee \dashv \Delta$

$$\frac{A \vee B \to C}{(A \to C) \wedge (B \to C)}$$

$$\mathbf{Prop}(A \vee B, C) \cong \mathbf{Prop} \times \mathbf{Prop}((A, B), \Delta C)$$

$\Delta \dashv \wedge$

$$\frac{(C \to A) \wedge (C \to B)}{C \to A \wedge B}$$

$$\mathbf{Prop} \times \mathbf{Prop}(\Delta C, (A, B)) \cong \mathbf{Prop}(C, A \wedge B)$$

$\varinjlim \dashv \Delta \dashv \varprojlim$

Consider the constant diagram functor $\Delta : \mathbf{C} \to \mathbf{C}^{\mathbf{I}}$. It maps $X \in \mathbf{C}$ to the constant diagram $\Delta_X : \mathbf{I} \to \mathbf{C}$ which maps every object to X and every morphism to 1_X.
The limit construction is a functor $\varprojlim : \mathbf{C}^{\mathbf{I}} \to \mathbf{C}$ that maps each diagram $D \in \mathbf{C}^{\mathbf{I}}$ to its limit $\varprojlim D$.
The cones over $D : \mathbf{I} \to \mathbf{C}$ with vertex X is the hom-set $\mathbf{C}^{\mathbf{I}}(\Delta_X, D)$.
The cones over $D : \mathbf{I} \to \mathbf{C}$ with vertex X correspond one-to-one with $\mathbf{C}(X, \varprojlim D)$.
If $\mathbf{C}$ has all limits of shape $\mathbf{I}$, then

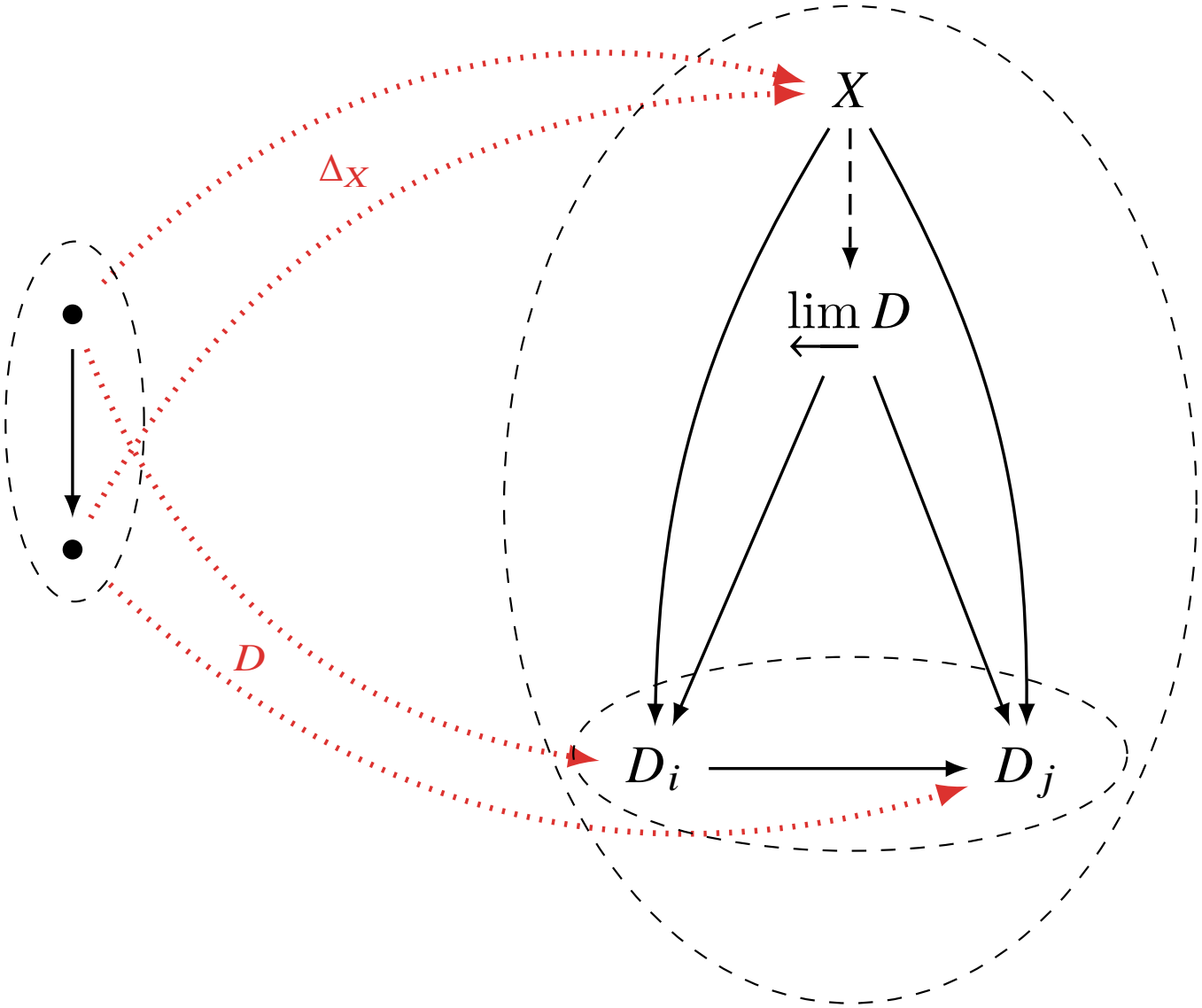
$$\mathbf{C}^{\mathbf{I}}(\Delta_X, D) \cong \mathbf{C}(X, \varprojlim D)$$

Similarly,

$$\mathbf{C}^{\mathbf{I}}(D, \Delta_X) \cong \mathbf{C}(\varinjlim D, X)$$

$$\varinjlim \dashv \Delta \dashv \varprojlim$$

$\Delta \dashv \varprojlim$



$$\mathbf{C}^{\mathbf{I}}(\Delta_X, D) \cong \mathbf{C}(X, \varprojlim D)$$

The Category of Adjunctions

Theorem

$$\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D} \underset{G'}{\overset{F'}{\rightleftarrows}} \mathbf{E} \quad \Longrightarrow \quad \mathbf{C} \underset{G'G}{\overset{F'F}{\rightleftarrows}} \mathbf{E}$$

Remark: We can define the category of adjunctions $\mathbf{Adj}(\mathbf{Cat})$ in which objects are categories and arrows are adjunctions.

Theorem

Adjoints are unique up to natural isomorphism. If $F \dashv G$ and $F \dashv G'$ then $G \cong G'$. If $F \dashv G$ and $F' \dashv G$ then $F \cong F'$.

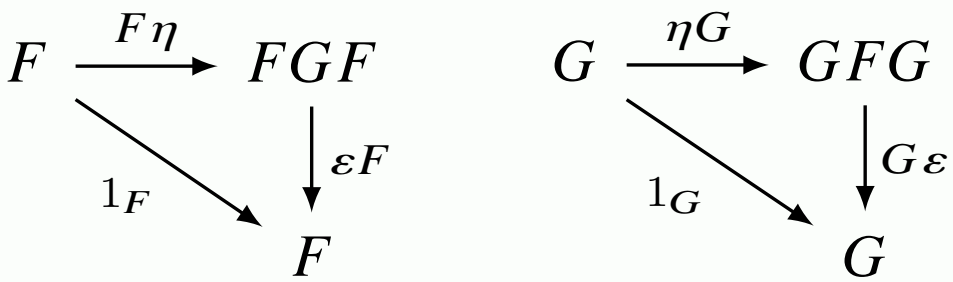
Theorem

If $F \dashv G$ and $G \cong G'$ then $F \dashv G'$. If $F \dashv G$ and $F \cong F'$ then $F' \dashv G$.

Adjunction via Unit/Counit

Theorem

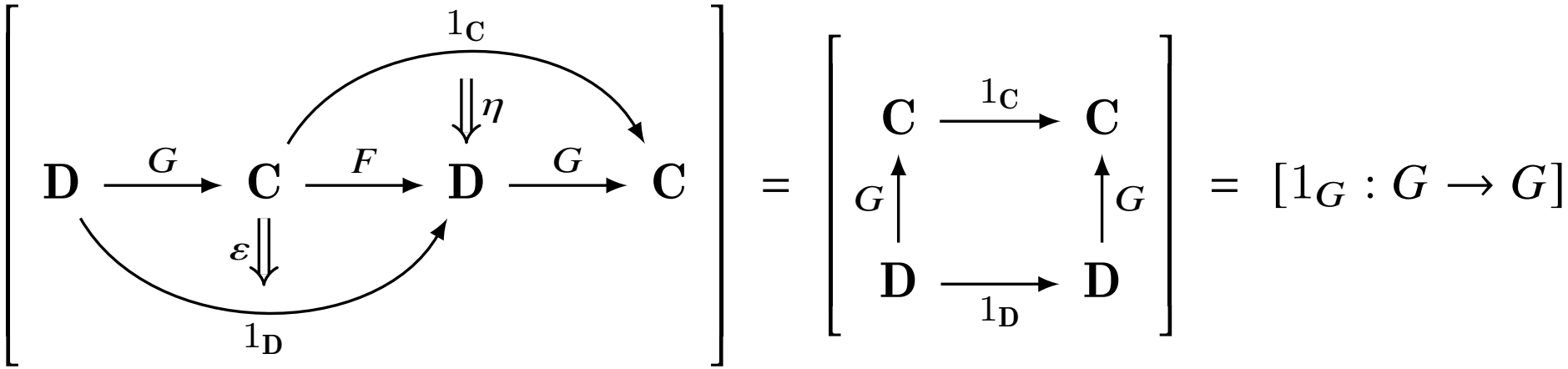
Functors $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}$ *are adjoint iff there are two natural transformations: the unit* $\eta : 1_{\mathbf{C}} \to GF$ *and counit* $\varepsilon : FG \to 1_{\mathbf{D}}$ *s.t.*

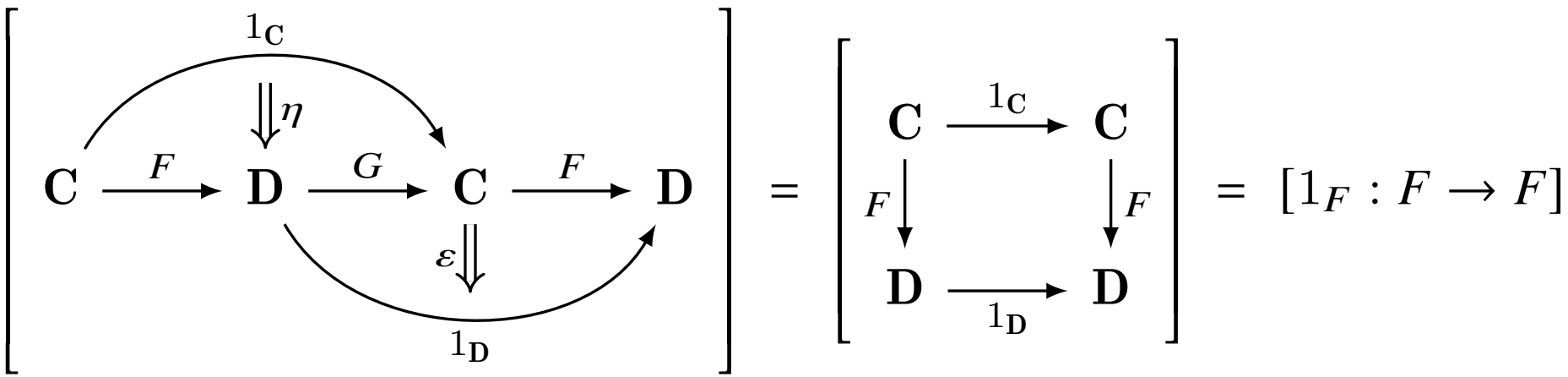


Remark: Adjoint functors are "pseudo-inverses": the round trip returns to an object/morphism that is related to the original by a natural transformation.

Remark: The triangle identities have the virtue of being entirely "algebraic" — no quantifiers, limits, Hom-sets, infinite conditions, etc. Thus, anything defined by adjunctions such as free groups, product spaces, quantifiers, ...can be defined equationally.

Remark: 2-胞腔表示





Let $\theta : \mathbf{D}(F-,-) \xrightarrow{\cong} \mathbf{C}(-,G-)$ be the natural isomorphism witnessing $F \dashv G$. For any $A \in \mathbf{C}$, there is a distinguished morphism $\eta_A \coloneqq \theta_{A,FA}(1_{FA}) : A \to G(FA)$.

$$\frac{1_{FA} : FA \to FA}{\eta_A : A \to G(FA)}\ \theta_{A,FA}$$

In fact, we can recover θ from η as follows, for $g : FA \to B$,

$$\theta_{A,B}g = \theta_{A,B}(g \circ 1_{FA}) = Gg \circ \theta_{A,FA}(1_{FA}) = Gg \circ \eta_A$$

$$\begin{array}{ccc} \mathbf{D}(FA,FA) & \xrightarrow{\theta_{A,FA}} & \mathbf{C}(A,GFA) \\ \Big\downarrow{\scriptstyle \mathbf{D}(FA,g)} & & \Big\downarrow{\scriptstyle \mathbf{C}(A,Gg)} \\ \mathbf{D}(FA,B) & \xrightarrow[\theta_{A,B}]{} & \mathbf{C}(A,GB) \end{array}$$

Similarly, for any $B \in \mathbf{D}$, there is a distinguished morphism $\varepsilon_B \coloneqq \theta^{-1}_{GB,B}(1_{GB}) : F(GB) \to B$.

$$\frac{1_{GB} : GB \to GB}{\varepsilon_B : F(GB) \to B}\ \theta^{-1}_{GB,B}$$

In fact, we can recover θ^{-1} from ε as follows, for $f : A \to GB$,

$$\theta^{-1}_{A,B} f = \theta^{-1}_{A,B}(1_{GB} \circ f) = \theta^{-1}_{GB,B} 1_{GB} \circ Ff = \varepsilon_B \circ Ff$$

$$\begin{array}{ccc} \mathbf{D}(FGB, B) & \xleftarrow{\theta^{-1}_{GB,B}} & \mathbf{C}(GB, GB) \\ \Big\downarrow{\scriptstyle \mathbf{D}(Ff,B)} & & \Big\downarrow{\scriptstyle \mathbf{C}(f,GB)} \\ \mathbf{D}(FA, B) & \xleftarrow[\theta^{-1}_{A,B}]{} & \mathbf{C}(A, GB) \end{array}$$

In general,

$$\theta \mapsto \Big(\eta_A \coloneqq \theta(1_{FA}), \varepsilon_B \coloneqq \theta^{-1}(1_{GB})\Big)$$

$$\theta g \coloneqq Gg \circ \eta_A \mapsfrom \eta$$

$$\theta^{-1} f \coloneqq \varepsilon_B \circ Ff \mapsfrom \varepsilon$$

$X \underset{g}{\overset{f}{\rightleftarrows}} Y$	**Topology**	**Category**	$\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$
$fg = 1_Y$ $gf = 1_X$ equality	homeomorphism between spaces	isomorphism $\mathbf{C} \cong \mathbf{D}$	$FG = 1_{\mathbf{D}}$ $GF = 1_{\mathbf{C}}$ equality
$fg \cong 1_Y$ $gf \cong 1_X$ homotopy	homotopy equivalence between spaces	equivalence $\mathbf{C} \simeq \mathbf{D}$	$FG \cong 1_{\mathbf{D}}$ $GF \cong 1_{\mathbf{C}}$ natural isomorphism
		adjunction $F \dashv G$	$\eta : 1_{\mathbf{C}} \to GF$ $\varepsilon : FG \to 1_{\mathbf{D}}$ natural transformation

Every sufficiently good analogy is yearning to become a functor.

— *John Baez*

functor < adjunction < equivalence < isomorphism < equality

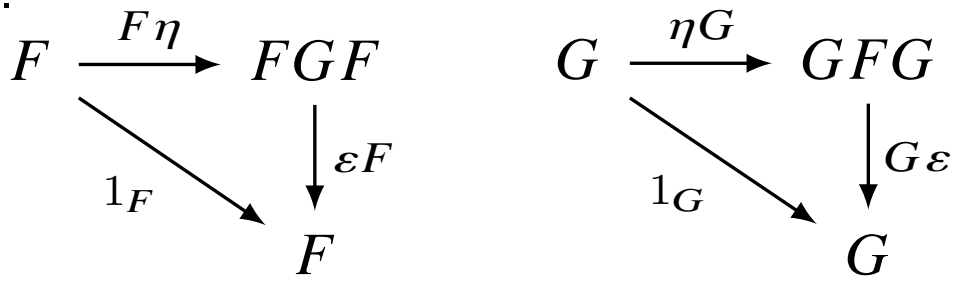
Remark

Given a functor $F : \mathbf{C} \to \mathbf{D}$, is there a way of traveling back in the other direction?

1. If we want to get back to exactly where we started, We could just ask for an inverse functor
$$FF^{-1} = 1_{\mathbf{D}} \text{ and } 1_{\mathbf{C}} = F^{-1}F$$
2. If just want to to get back to somewhere isomorphic to our starting point, We could require a functor $G : \mathbf{D} \to \mathbf{C}$ such that there are natural isomorphisms
$$FG \cong 1_{\mathbf{D}} \text{ and } 1_{\mathbf{C}} \cong GF$$
3. If we only require that there is a morphism relating the start and end points of a round trip, there are two sensible choices:
$$FG \xrightarrow{\varepsilon} 1_{\mathbf{D}} \text{ and } 1_{\mathbf{C}} \xrightarrow{\eta} GF$$
$$FG \xleftarrow{\varepsilon'} 1_{\mathbf{D}} \text{ and } 1_{\mathbf{C}} \xleftarrow{\eta'} GF$$

Adjunction vs Equivalence

- An equivalence between categories $\mathbf{C}$ and $\mathbf{D}$ is a pair of functors $\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$ and a pair of natural *isomorphisms* $\eta : 1_{\mathbf{C}} \to GF$ and $\varepsilon : FG \to 1_{\mathbf{D}}$.
- An adjunction between categories $\mathbf{C}$ and $\mathbf{D}$ is a pair of functors $\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$ and a pair of natural *transformations* $\eta : 1_{\mathbf{C}} \to GF$ and $\varepsilon : FG \to 1_{\mathbf{D}}$ s.t.

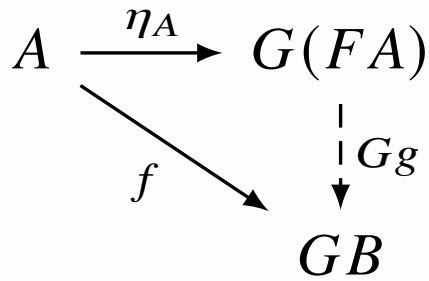


Theorem

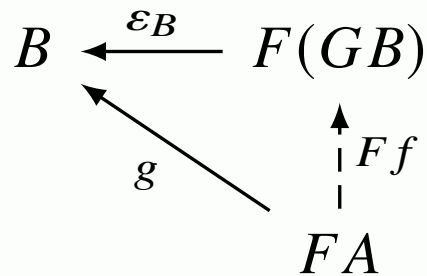
If there is an equivalence $\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$ *and a pair of natural isomorphisms* $\eta : 1_{\mathbf{C}} \to GF$ *and* $\gamma : FG \to 1_{\mathbf{D}}$, *then there is an adjunction* $F \dashv G$ *with unit* η *and counit* $\varepsilon : FG \xrightarrow{FG\gamma^{-1}} FGFG \xrightarrow{F\eta^{-1}G} FG \xrightarrow{\gamma} 1_{\mathbf{D}}$.

Theorem

- *Functors* $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}$ *are adjoint iff there is a natural transformation* $\eta : 1_{\mathbf{C}} \to GF$, *for which for any* $f : A \to GB$ *in* $\mathbf{C}$ *there is a unique* $g : FA \to B$ *in* $\mathbf{D}$ *s.t.* $f = Gg \circ \eta_A$.



- *Functors* $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}$ *are adjoint iff there is a natural transformation* $\varepsilon : FG \to 1_{\mathbf{D}}$, *for which for any* $g : FA \to B$ *in* $\mathbf{D}$ *there is a unique* $f : A \to GB$ *in* $\mathbf{C}$ *s.t.* $g = \varepsilon_B \circ Ff$.



Lemma

Given $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}$ *, with counit* $\varepsilon : FG \to 1_{\mathbf{D}}$*,*

- G *is faithful iff* ε_B *is an epimorphism for all* B*.*
- G *is full and faithful iff* ε_B *is an isomorphism for all* B*.*

Definition

An adjunction where G is full and faithful is called a *reflection*.

Definition (Reflective Subcategory)

A full subcategory $i : \mathbf{C} \hookrightarrow \mathbf{D}$ is *reflective* if the inclusion functor i has a left adjoint $r \dashv i$:

$$\mathbf{C} \underset{i}{\overset{r}{\underset{\hookrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbf{D}$$

Dually, it is *coreflective* if i has a right adjoint.

Adjoint Equivalence vs Equivalence

Definition (Adjoint Equivalence)

Functors $\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$ are *adjoint equivalent* iff there is an adjunction $F \dashv G$, and the unit $\eta : 1_{\mathbf{C}} \to GF$ and the counit $\varepsilon : FG \to 1_{\mathbf{D}}$ are natural isomorphisms.

Theorem

Let $(F, G, \eta, \varepsilon)$ be an adjoint equivalence. Then

1. *$(F, G, \eta, \varepsilon)$ is an equivalence of categories.*
2. *$(G, F, \varepsilon^{-1}, \eta^{-1})$ is also an adjoint equivalence.*
3. *$F \dashv G \dashv F$*

1. Adjunctions may or may not be equivalences.
   e.g. $- \times B \dashv (-)^B$ is not an equivalence, as the counit $\varepsilon_A : A^B \times B \to A$ is not invertible.
2. Equivalences may or may not be adjunctions.
3. If $F \dashv G \dashv F$, then (F, G) may or may not be an equivalence.

Theorem

If $(F, G, \eta, \varepsilon)$ is an equivalence, then there exists a unique $\varepsilon_0 : FG \to 1_{\mathbf{D}}$ such that $(F, G, \eta, \varepsilon_0)$ is an adjoint equivalence.

Proof Sketch.

η and ε may not satisfy the triangle identity. Define ε_0 to be

$$\begin{array}{ccc} FGFG & \xrightarrow{F\eta^{-1}G} & FG \\ {\scriptstyle FG\varepsilon^{-1}}\uparrow & & \downarrow{\scriptstyle \varepsilon} \\ FG & \xrightarrow[\varepsilon_0]{} & 1_{\mathbf{D}} \end{array}$$

the following diagram commutes,

$$\begin{array}{ccccc} F & & \xrightarrow{\varepsilon^{-1}F} & & FGF \\ & \searrow^{F\eta} & & & \searrow^{F\eta GF} \\ {\scriptstyle 1_F}\downarrow & & FGF & \xrightarrow{FG\varepsilon^{-1}F} & FGFGF \\ & \searrow^{\varepsilon_0 F} & & & \swarrow_{F\eta^{-1}GF} \\ F & & \xleftarrow[\varepsilon F]{} & & FGF \end{array}$$

which gives a triangle identity $\varepsilon_0 F \circ F\eta = 1_F$. □

Fixpoint Equivalence of an Adjunction

Definition (Fixpoint Equivalence of an Adjunction)

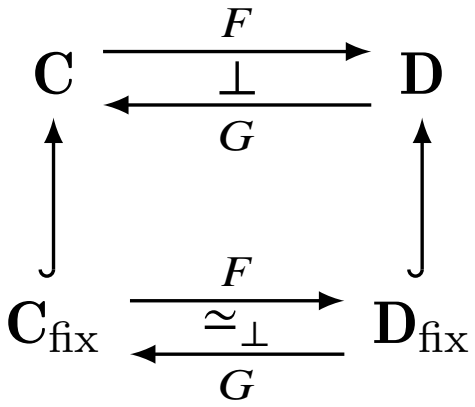
Let $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}$ be a pair of adjoint functors.

- and object $A \in \mathbf{C}$ is a fixpoint of the adjunction if its adjunction unit is an isomorphism
$$\eta_A : A \xrightarrow{\cong} GFA$$
and write $\mathbf{C}_{\text{fix}} \hookrightarrow \mathbf{C}$ for the full subcategory on these fixed objects;
- and object $B \in \mathbf{D}$ is a fixpoint of the adjunction if its adjunction counit is an isomorphism
$$\varepsilon_B : FGB \xrightarrow{\cong} B$$
and write $\mathbf{D}_{\text{fix}} \hookrightarrow \mathbf{D}$ for the full subcategory on these fixed objects.

Theorem: Then the adjunction (co-)restrics to an adjoint equivalence on these full subcategories of fixpoints:

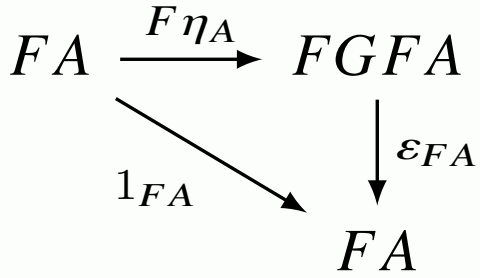
$$\mathbf{C}_{\text{fix}} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\simeq \bot}}}} \mathbf{D}_{\text{fix}}$$

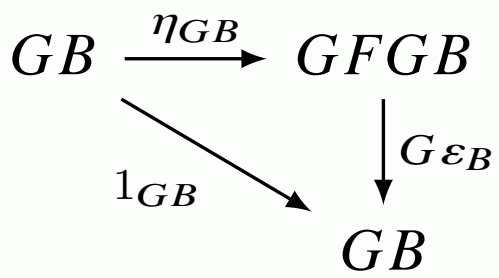
Fixpoint Equivalence of an Adjunction



Proof.

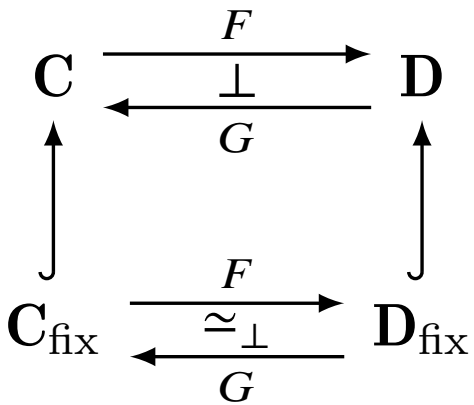
- for $A \in \mathbf{C}_{\text{fix}} \hookrightarrow \mathbf{C}$, since η_A is an isomorphism in $\mathbf{C}$, ε_{FA} is an isomorphism in $\mathbf{D}$.
- for $B \in \mathbf{D}_{\text{fix}} \hookrightarrow \mathbf{D}$, since ε_B is an isomorphism in $\mathbf{D}$, η_{GB} is an isomorphism in $\mathbf{C}$.





Digression — Dialectical Interpretations

"All things are in flux; the flux is subject to a unifying measure or rational principle. This principle (logos, the hidden harmony behind all change) bound opposites together in a unified tension, which is like that of a lyre, where a stable harmonious sound emerges from the tension of the opposing forces that arise from the bow bound together by the string." — *Heraclitus*



"The technical advances forged by category theorists will be of value to dialectical philosophy. ...Of course this will require that philosophers learn mathematics and that mathematicians learn philosophy."

— *Lawvere*

1. We may think of (F, G) as establishing a contradiction between $\mathbf{C}$ and $\mathbf{D}$. The equivalence $\mathbf{C}_{\text{fix}} \simeq \mathbf{D}_{\text{fix}}$ is then the unity of opposites.
2. We may think of F as the thesis, its right adjoint G as the antithesis, and the adjunction itself $F \dashv G$ as the synthesis.

Example

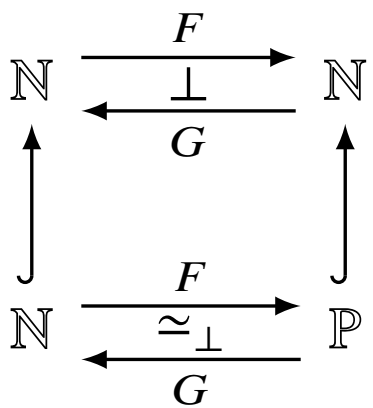
Take $\mathbf{C} = \mathbf{D} = (\mathbb{N}, \geq)$.

$$F(m) := \begin{cases} 0 & \text{if } m = 0 \\ \text{the } m\text{-th prime number} & \text{if } m > 0 \end{cases}$$

$$G(n) := \text{the number of primes} \leq n$$

$$\frac{F(m) \geq n}{m \geq G(n)}$$

- $m \geq GF(m)$ for all $m \in \mathbb{N}$.
- $FG(n) \geq n$ iff $n \in \mathbb{P}$.



Adjoint Cylinder $L \dashv T \dashv R$ ☯

- An adjoint cylinder is an adjoint triple such that the outer two adjoints are full and faithful.
- Let $L : N \to N :: n \mapsto 2n$ and $R : N \to N :: n \mapsto 2n + 1$.
- There is a third functor $T : N \to N$ satisfying $L \dashv T \dashv R$.
  1. The triple $L \dashv T \dashv R$ expresses the unity by the idempotency of $(R \circ T)^2 = R \circ T$ and $(L \circ T)^2 = L \circ T$.
  2. it expresses the opposition between L and R by an entailed adjunction $L \circ T \dashv R \circ T$.
  3. it expresses the identity by the entailed quivalence $T \circ L \cong T \circ R$.
- For our poset example N, when T exists it must satisfy $T \circ L = 1_N = T \circ R$, which indicates as definition for T:

$$T(n) := \begin{cases} \frac{n}{2} & n \in N_{\text{even}} \\ \frac{n-1}{2} & n \in N_{\text{odd}} \end{cases}$$

- The idempotent comonad $\mathrm{sk} = L \circ T$ and the idempotent monad $\mathrm{cosk} = R \circ T$:

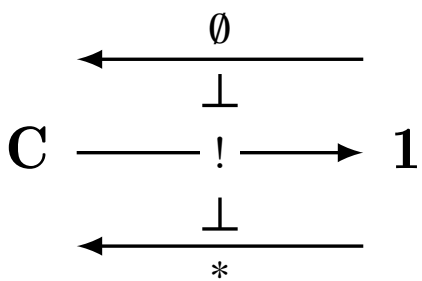
$$\mathrm{sk}(n) = \begin{cases} n & n \in N_{\text{even}} \\ n-1 & n \in N_{\text{odd}} \end{cases} \quad \text{and} \quad \mathrm{cosk}(n) = \begin{cases} n+1 & n \in N_{\text{even}} \\ n & n \in N_{\text{odd}} \end{cases}$$

becoming: nothing ⊣ being

Theorem

Let $\mathbf{C}$ *be a category.*

- *The following are equivalent:*
  1. $\mathbf{C}$ *has a terminal object.*
  2. *the unique functor* $\mathbf{C} \to \mathbf{1}$ *to the terminal category has a right adjoint*
- *The following are equivalent:*
  1. $\mathbf{C}$ *has an initial object.*
  2. *the unique functor* $\mathbf{C} \to \mathbf{1}$ *to the terminal category has a left adjoint*



$$\mathrm{Hom}_{\mathbf{C}}(\emptyset(\bullet), X) \cong \mathrm{Hom}_{\mathbf{1}}(\bullet, !(X))$$
$$\mathrm{Hom}_{\mathbf{C}}(X, *(\bullet)) \cong \mathrm{Hom}_{\mathbf{1}}(!(X), \bullet)$$

$$\emptyset \to X \to * \qquad \emptyset \dashv ! \dashv * \qquad \emptyset! \dashv *!$$

$$\mathrm{Hom}_{\mathbf{C}}(\emptyset!(X), Y) \cong \mathrm{Hom}_{\mathbf{C}}(X, *!(Y))$$

Hegel's Aufhebung

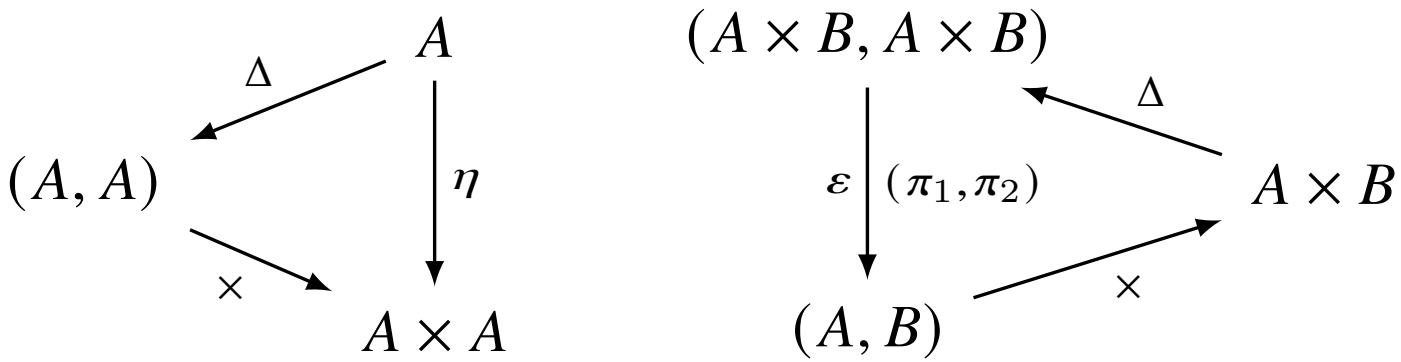
- A *localization* of a category $\mathbf{A}$ with finite limits is a reflective subcategory $\mathbf{B} \overset{i_*}{\hookrightarrow} \mathbf{A}$ whose reflection preserves finite limits. The localization is called *essential* when the reflection i^* has furthermore a left adjoint $i_! \dashv i^* \dashv i_*$.
- When $\mathbf{A}$ is a topos, $\mathbf{B}$ is called an essential subtopos.
- An essential subtopos $\mathbf{A}_i \hookrightarrow \mathbf{A}$ is called a level of $\mathbf{A}$.
- An adjoint triple $i_! \dashv i^* \dashv i_*$ yields two adjoint modalities $\Box_i \dashv \bigcirc_i$ on $\mathbf{A}$. $\Box_i := i_! i^*$ and $\bigcirc_i := i^* i_*$.
- The modalities yield notions of modal types
  1. (i-sheaves) $X \in \mathbf{A}$ with $\bigcirc_i X \cong X$.
  2. (i-skeleta) $X \in \mathbf{A}$ with $\Box_i X \cong X$.
- We say that the level i is lower than level j, (written $i \prec j$) when
  1. every i-sheaf is a j-sheaf: $\bigcirc_j \bigcirc_i = \bigcirc_i$
  2. every i-skeleton is a j-skeleton: $\Box_j \Box_i = \Box_i$
- Let $i \prec j$, we say that the level j resolves the opposite of level i, (written $i \ll j$) when $\bigcirc_j \Box_i = \Box_i$.
- A level $\bar{i}$ is called the Aufhebung of level i iff it is a minimal level which resolves the opposites of level i.

Example — Left/Right Adjoint & Unit/Counit $+ \dashv \Delta \dashv \times$

$$\mathbf{C} \times \mathbf{C}(\Delta X, (A, B)) \cong \mathbf{C}(X, A \times B)$$

$$\eta : 1_{\mathbf{C}} \to \times \circ \Delta$$

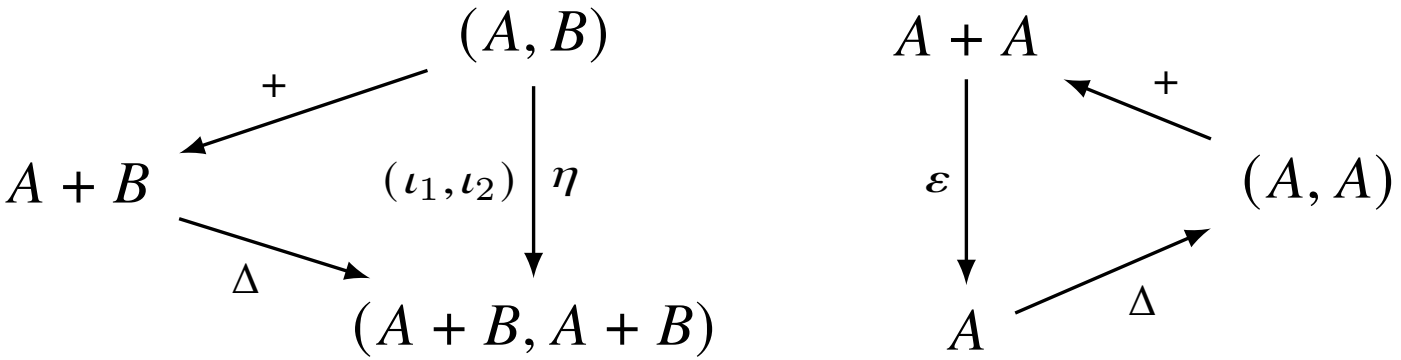
$$\varepsilon : \Delta \circ \times \to 1_{\mathbf{C} \times \mathbf{C}}$$



$$\mathbf{C}(A + B, X) \cong \mathbf{C} \times \mathbf{C}((A, B), \Delta X)$$

$$\eta : 1_{\mathbf{C} \times \mathbf{C}} \to \Delta \circ +$$

$$\varepsilon : + \circ \Delta \to 1_{\mathbf{C}}$$

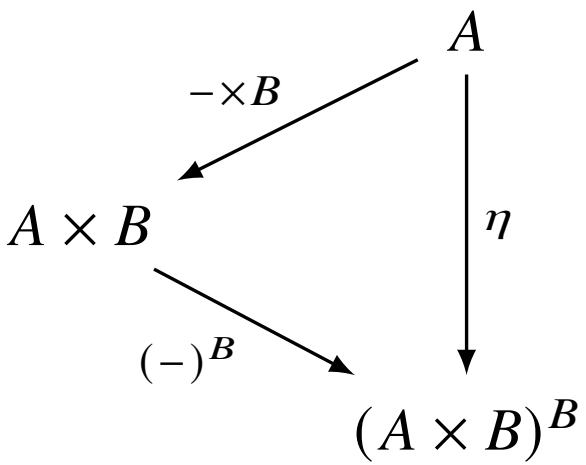


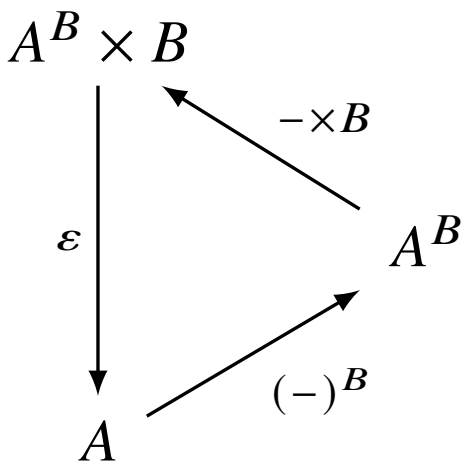
Example — Left/Right Adjoint & Unit/Counit
$- \times B \dashv (-)^B$

$$\mathrm{Hom}(A \times B, C) \cong \mathrm{Hom}(A, C^B)$$

$$\eta_A : A \to (A \times B)^B$$

$$\varepsilon_A : A^B \times B \to A$$





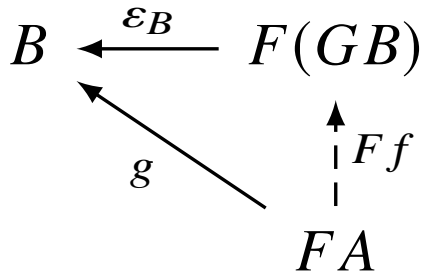
$$\frac{A \wedge B \vdash C}{A \vdash B \to C}$$

$$\eta_A : A \to B \to A \wedge B$$

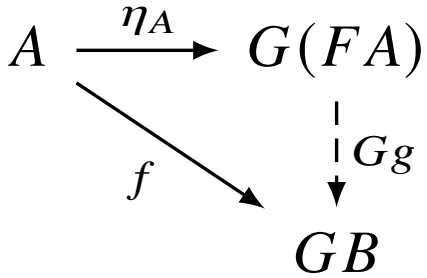
$$\varepsilon_A : (B \to A) \wedge B \to A$$

Adjunction via Universal Property

- A functor $F : \mathbf{C} \to \mathbf{D}$ has a right adjoint, iff, for any $B \in \mathbf{D}$, there is an object $GB \in \mathbf{C}$ and a morphism $\varepsilon_B : F(GB) \to B$ such that (GB, ε_B) is a universal morphism from F to B.



- A functor $G : \mathbf{D} \to \mathbf{C}$ has a left adjoint, iff, for any $A \in \mathbf{C}$, there is an object $FA \in \mathbf{D}$ and a morphism $\eta_A : A \to G(FA)$ such that (FA, η_A) is a universal morphism from A to G.



Universality ≡ Adjunctions

Characterizations of Adjunction $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}$

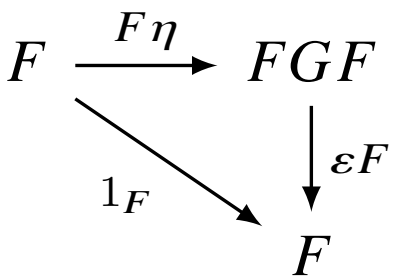
- Isomorphism between Hom-sets

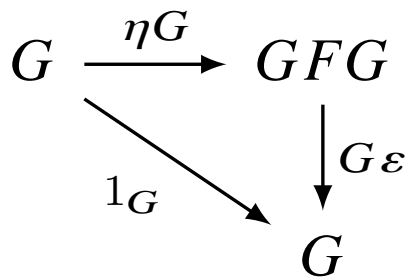
$$\mathbf{D}(FA, B) \xrightarrow{\cong} \mathbf{C}(A, GB)$$

that is natural in both A and B.

- Units and Counits

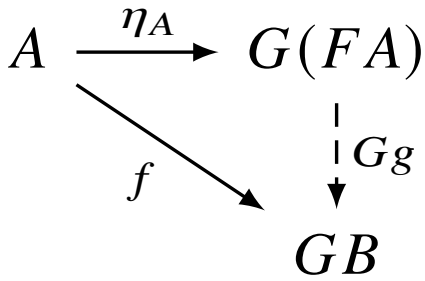
$$\eta : 1_{\mathbf{C}} \to GF$$
$$\varepsilon : FG \to 1_{\mathbf{D}}$$



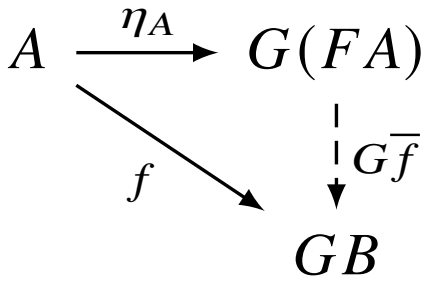


that are natural transformations.

- Universal Property
  (FA, η_A) is an initial object of the comma category $A \downarrow G$.



Example — Extension Problem



1. We often start with a set A.
   Example: take $A := \{x, y, z\}$ to be a set with three elements.
2. The elements in A are then used as building blocks to construct a bigger mathematical object FA, which contains the original A and more.
   Example: take $FA := \{ax + by + cz : a, b, c \in \mathbb{R}\}$ to be the three-dimensional real vector space with basis A.
3. As a consequence, we observe that whenever another object also "contains" A, it automatically contains FA, too.
   Example: if B is any vector space and there is a mapping f from A to GB, then you automatically have a linear transformation $\overline{f} : FA \to B$. In other word, $\overline{f}$ is the unique map that extends f linearly from the basis set A to the entire vector space FA.

Theorem

Given $F \dashv G$. *Then for any small* $\mathbf{I}$,

$$\mathbf{C}^{\mathbf{I}} \underset{G_*}{\overset{F_*}{\rightleftarrows}} \mathbf{D}^{\mathbf{I}} \quad (\perp)$$

and for any locally small $\mathbf{E}$,

$$\mathbf{E}^{\mathbf{C}} \underset{F^*}{\overset{G^*}{\rightleftarrows}} \mathbf{E}^{\mathbf{D}} \quad (\perp)$$

Theorem (Continuity)

- *Right adjoints preserve limits.*
- *Left adjoints preserve colimits.*

Proof.

Given $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}$, assume limits of shape $\mathbf{I}$ exist in both $\mathbf{C}$ and $\mathbf{D}$.

Define $F_*(D) \coloneqq FD$ and $G_*(D) \coloneqq GD$. Then

$$\mathbf{C}^{\mathbf{I}} \underset{G_*}{\overset{F_*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}^{\mathbf{I}}$$

Since the diagram of right adjoints of the functors in a commutative square commutes up to natural isomorphism, then

$$\begin{array}{ccc} \mathbf{C} & \xrightarrow{F} & \mathbf{D} \\ {\scriptstyle\Delta}\downarrow & & \downarrow{\scriptstyle\Delta} \\ \mathbf{C}^{\mathbf{I}} & \xrightarrow[F_*]{} & \mathbf{D}^{\mathbf{I}} \end{array} \quad \Longrightarrow \quad \begin{array}{ccc} \mathbf{C} & \xleftarrow{G} & \mathbf{D} \\ {\scriptstyle\lim}\uparrow & & \uparrow{\scriptstyle\lim} \\ \mathbf{C}^{\mathbf{I}} & \xleftarrow[G_*]{} & \mathbf{D}^{\mathbf{I}} \end{array}$$

$$G \varprojlim D \cong \varprojlim GD$$

Example

- Right adjoints preserve limits.

$$\frac{C \to A \wedge B}{(C \to A) \wedge (C \to B)} \; - \wedge C \dashv C \to -$$

$$\frac{\forall x(Ax \wedge Bx)}{\forall x Ax \wedge \forall x Bx} \; \exists \dashv \pi^* \dashv \forall$$

- Left adjoints preserve colimits.

$$\frac{(A \vee B) \wedge C}{(A \wedge C) \vee (B \wedge C)} \; - \wedge C \dashv C \to -$$

$$\frac{\exists x(Ax \vee Bx)}{\exists x Ax \vee \exists x Bx} \; \exists \dashv \pi^* \dashv \forall$$

Theorem

Full and faithful functor reflects limits and colimits.

Proof.

If F is full and faithful and $FL \cong \varprojlim FD$, one may sketch the proof for $L \cong \varprojlim D$ as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\Delta_C \to D}{\Delta_{FC} \to FD}\ (F \text{ full and faithful})}{FC \to \varprojlim FD}\ (\Delta \dashv \varprojlim)}{FC \to FL}\ (FL \cong \varprojlim FD)}{C \to L}\ (F \text{ full and faithful})$$

□

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions

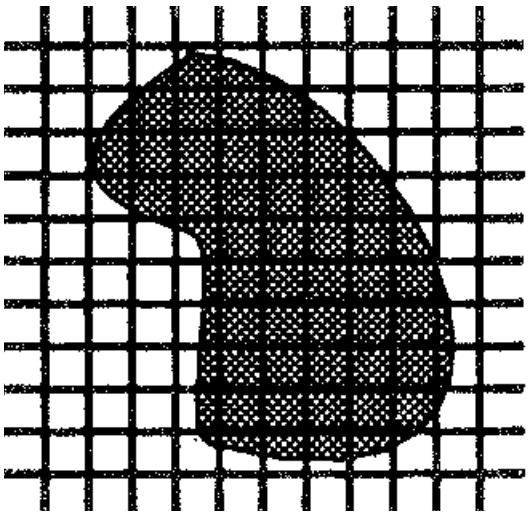
Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

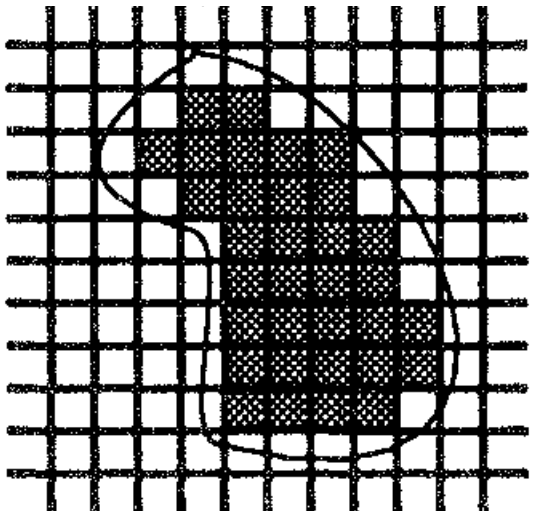
Answers to the Exercises

References2039

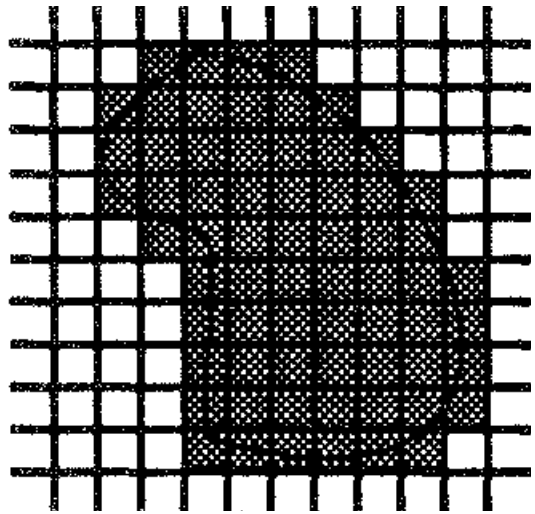
Left/Right Adjoint



A



$\Box A$



$\Diamond A$

$\Box A$ squares that are completely covered by A.

$\Diamond A$ squares that are partly or totally covered by A.

$$\frac{\Diamond A \subset B}{A \subset \Box B} \qquad \Diamond \dashv \Box$$

Left/Right Adjoint

Example

Assume $R \subset X \times Y$, and let $F_R : \mathrm{P}(X) \to \mathrm{P}(Y) :: A \mapsto \bigcup_{x \in A} \{y : Rxy\}$.

This has a right adjoint $[R] : \mathrm{P}(Y) \to \mathrm{P}(X)$:

$$\frac{F_R(A) \subset B}{A \subset [R]B}$$

The definition of $[R]$ which satisfies this condition is:

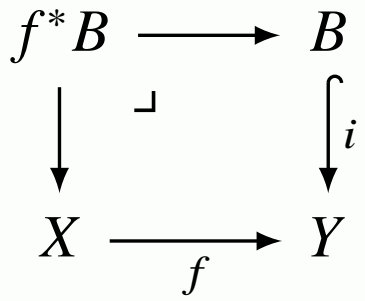
$$[R]B \coloneqq \{x : \forall y (Rxy \to y \in B)\}$$

If we take $X = Y = W$ and (W, R) as the Kripke frame for modal logic, then $[R]$ gives the usual Kripke semantics for $\Box$.

Left/Right Adjoint $\exists_f \dashv f^* \dashv \forall_f$

Example

Given a function $f : X \longrightarrow Y$, consider
$f^* : \mathrm{P}(Y) \longrightarrow \mathrm{P}(X) :: B \mapsto \{x \in X : fx \in B\}$.
Take the subset $B \subset Y$ as a predicate $B(y)$ over Y, and
f^*B as $f^*B(x)$ over X.
By the pullback, $f^*B(x) = B(fx) =: (Bf)(x)$.



Then f^* has both a left and a right adjoint $\exists_f, \forall_f : \mathrm{P}(X) \longrightarrow \mathrm{P}(Y)$.

$$\frac{A \subset f^*B}{\exists_f A \subset B} \qquad \frac{f^*B \subset A}{B \subset \forall_f A} \qquad \text{i.e.} \qquad \frac{A \vdash_X Bf}{\exists_f A \vdash_Y B} \qquad \frac{Bf \vdash_X A}{B \vdash_Y \forall_f A}$$

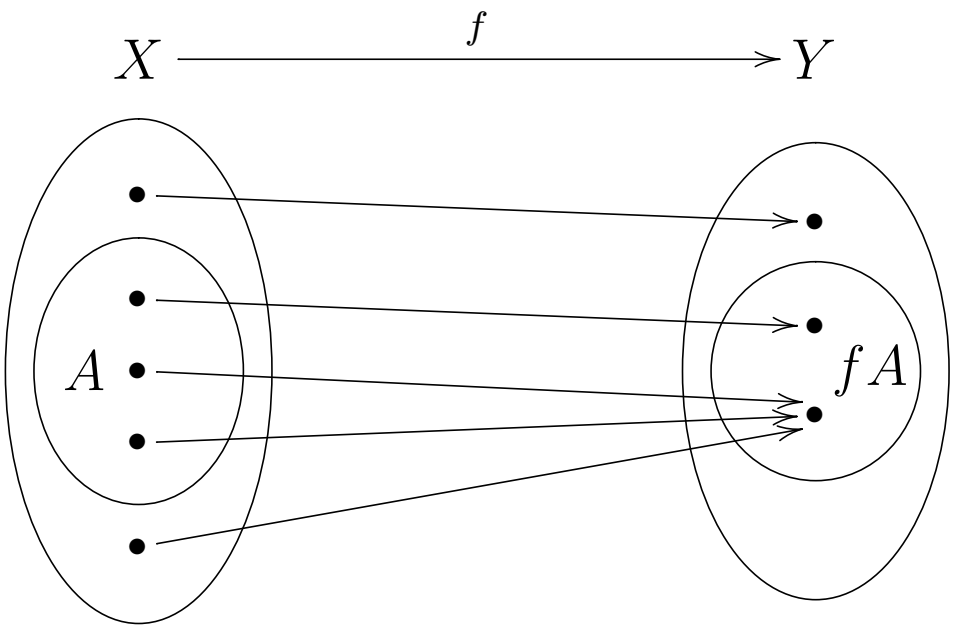
$$\boxed{\exists_f \dashv f^* \dashv \forall_f}$$

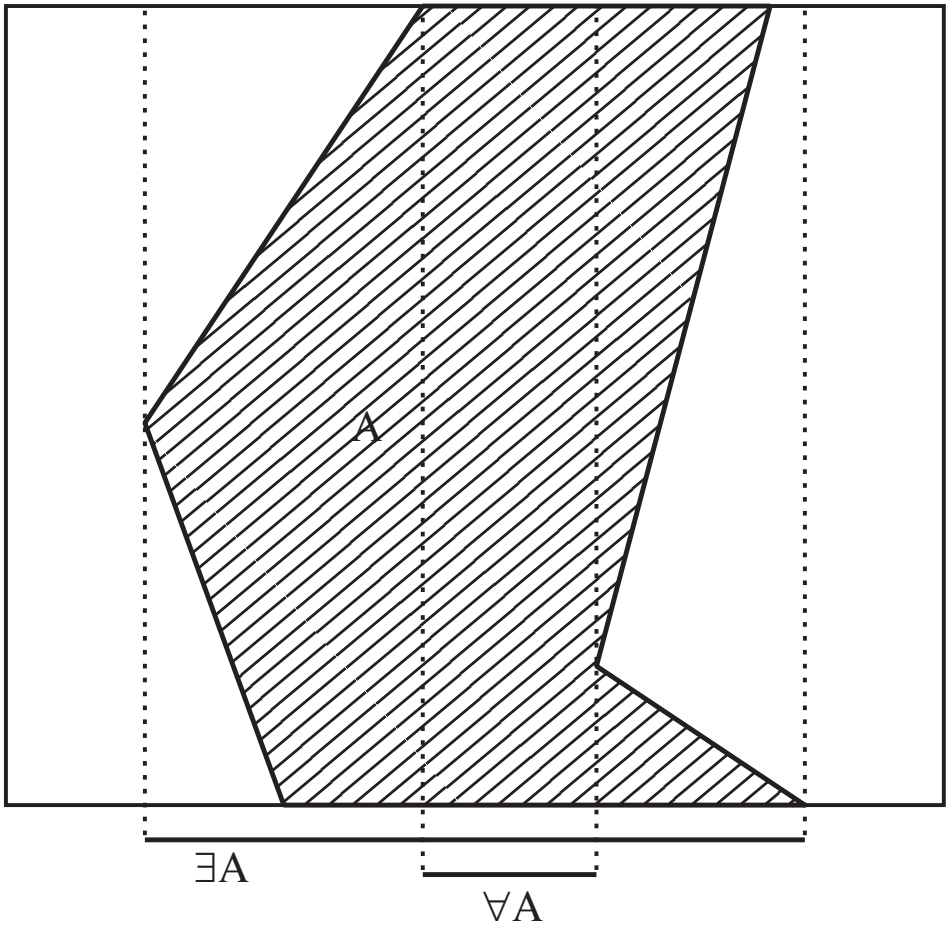
The unique functions satisfying these conditions can be defined as

$$\exists_f A := \bigcap \{B : A \subset f^*B\} = fA = \{y \in Y : \exists x(fx = y \wedge x \in A)\}$$

$$\forall_f A := \bigcup \{B : f^*B \subset A\} = \{y \in Y : f^*y \subset A\} = \{y \in Y : \forall x(fx = y \longrightarrow x \in A)\}$$

Quantifiers are Adjoints





$$\frac{A \subset f^* B}{f A \subset B} \qquad \frac{A \vdash_X B f}{\exists_f A \vdash_Y B}$$

$$\exists_f \dashv f^* \dashv \forall_f : \mathrm{P}(X) \leftarrow \mathrm{P}(Y)$$

Example

- Let $Y = \{\texttt{even}, \texttt{odd}\}$, and let $f : \mathbb{N} \to Y$.
- Given $A \in \mathtt{P}(\mathbb{N})$, we investigate $\exists_f A, \forall_f A$.
- The set $\exists_f A$ includes the element $\texttt{even}$ if there exists an even number in A; it includes the element $\texttt{odd}$ if there exists an odd number in A. Similarly, the set $\forall_f A$ includes the element $\texttt{even}$ if every even number is in A, and it includes $\texttt{odd}$ if every odd number is in A.
- Let's use the definition of adjunction to ask whether every element of $A \subset \mathbb{N}$ is even. Let $B = \{\texttt{even}\} \subset Y$. Then f^*B is the set of even numbers, and there is a morphism $A \to f^*B$ in the preorder $\mathtt{P}(\mathbb{N})$ iff every element of A is even. Therefore,

$$\mathrm{Hom}_{\mathtt{P}(\mathbb{N})}(A, f^*B) \cong \mathrm{Hom}_{\mathtt{P}(Y)}(\exists_f A, B)$$

says that $\exists_f A \subset \{\texttt{even}\}$ iff every element of A is even.

Remark

$$\{Ax\}_{x\in X} \cong \mathrm{P}(X) \underset{\forall_f}{\overset{\exists_f}{\underset{\longleftarrow}{\overset{\longrightarrow}{f^*}}}} \mathrm{P}(\bullet) = \{\emptyset, \bullet\} \cong \{\bot, \top\}$$

The set $\mathrm{P}(X) = \{A : A \subset X\}$ can be identified with the set of predicates Ax. The corresponding subset is $\{x \in X : Ax = \top\} = f^{-1}(\top)$.

$$\exists_f : A \mapsto \exists x A$$
$$\forall_f : A \mapsto \forall x A$$

Let f be the projection $\pi : X \times Y \to X$. Hence $\pi^* : \mathrm{P}(X) \to \mathrm{P}(X \times Y)$.

$$\begin{array}{ccc} \pi^* B & \longrightarrow & B \\ \downarrow & \lrcorner & \downarrow i \\ X \times Y & \xrightarrow[\pi]{} & X \end{array}$$

By the pullback, $\pi^* B(x, y) = B\pi(x, y)$.
Then π^* has both a left and a right adjoint $\exists_\pi, \forall_\pi : \mathrm{P}(X \times Y) \to \mathrm{P}(X)$.

$$\boxed{\exists_\pi \dashv \pi^* \dashv \forall_\pi}$$

We write $\exists_\pi A$ as $\exists y A(x, y)$, $\forall_\pi A$ as $\forall y A(x, y)$, and $\subset$ as $\vdash$.

$$\exists y A(x, y) = \exists_\pi A = \{x \in X : \exists y.\, (x, y) \in A\}$$
$$\forall y A(x, y) = \forall_\pi A = \{x \in X : \forall y.\, (x, y) \in A\}$$

$$\frac{A \subset \pi^* B}{\exists_\pi A \subset B} \qquad \frac{\pi^* B \subset A}{B \subset \forall_\pi A} \qquad \frac{A(x, y) \vdash_{X \times Y} B(x)}{\exists y A(x, y) \vdash_X B(x)} \qquad \frac{B(x) \vdash_{X \times Y} A(x, y)}{B(x) \vdash_X \forall y A(x, y)}$$

Quantifiers are Adjoints

For a list $\mathbf{x} = x_1, \ldots, x_n$ of distinct variables, let $\mathrm{Form}(\mathbf{x})$ be the set of formulas that has at most $\mathbf{x}$ free. Then $\mathrm{Form}(\mathbf{x})$ is a preorder under $\vdash$.
Let y be a variable not in $\mathbf{x}$. We have a trivial operation

$$* : \mathrm{Form}(\mathbf{x}) \to \mathrm{Form}(\mathbf{x}, y)$$

The operation $*$ is trivially a functor, since

$$A(\mathbf{x}) \vdash B(\mathbf{x}) \text{ in } \mathrm{Form}(\mathbf{x}) \implies A(\mathbf{x}, y) \vdash B(\mathbf{x}, y) \text{ in } \mathrm{Form}(\mathbf{x}, y)$$

For any $A \in \mathrm{Form}(\mathbf{x}, y)$, obviously $y \notin \mathrm{Fv}(\exists y A)$ and $y \notin \mathrm{Fv}(\forall y A)$. We have $\exists y / \forall y : \mathrm{Form}(\mathbf{x}, y) \to \mathrm{Form}(\mathbf{x})$.
Quantifiers are adjoints $\exists \dashv * \dashv \forall$.
Conversely, we could take $\exists \dashv * \dashv \forall$ as basic and derive the customary introduction and elimination rules from it. $\forall x A(x, y) \vdash A(x, y)$ is just the counit of $* \dashv \forall$, and $A(x, y) \vdash \exists y A(x, y)$ is the unit of $\exists \dashv *$.

$$\begin{aligned}
\forall x A(x, y) &\vdash A(x, y) && \text{(counit of } * \dashv \forall) \\
A(x, y) &\vdash \exists y A(x, y) && \text{(unit of } \exists \dashv *) \\
\forall x A(x, y) &\vdash \exists y A(x, y) && \text{(transitivity of } \vdash) \\
\exists y \forall x A(x, y) &\vdash \exists y A(x, y) && (\exists \dashv *) \\
\exists y \forall x A(x, y) &\vdash \forall x \exists y A(x, y) && (* \dashv \forall)
\end{aligned}$$

Quantifiers are Adjoints

Given a formula A. Let $[\![A]\!] \coloneqq \{(\mathbf{b}, a) : \mathcal{M} \vDash A[\mathbf{b}, a]\}$. Take the projection $\pi : (\mathbf{b}, a) \mapsto \mathbf{b}$. It can be regarded as $\pi : \nu(a/x) \mapsto \nu$.

$$\boxed{\exists_\pi \dashv \pi^* \dashv \forall_\pi}$$

$$\frac{[\![A]\!] \subset \pi^*[\![B]\!]}{\exists_\pi[\![A]\!] \subset [\![B]\!]} \qquad \frac{\pi^*[\![B]\!] \subset [\![A]\!]}{[\![B]\!] \subset \forall_\pi[\![A]\!]}$$

Explicitly,

$$\exists_\pi[\![A]\!] = \left\{\mathbf{b} : \exists a\Big(\mathcal{M} \vDash A[\mathbf{b}, a]\Big)\right\} = \bigcup_{a \in M} \left\{\nu : \mathcal{M}, \nu(a/x) \vDash A\right\}$$

$$\forall_\pi[\![A]\!] = \left\{\mathbf{b} : \forall a\Big(\mathcal{M} \vDash A[\mathbf{b}, a]\Big)\right\} = \bigcap_{a \in M} \left\{\nu : \mathcal{M}, \nu(a/x) \vDash A\right\}$$

And we have

$$[\![\exists x A]\!] = \exists_\pi[\![A]\!] \qquad [\![\forall x A]\!] = \forall_\pi[\![A]\!]$$

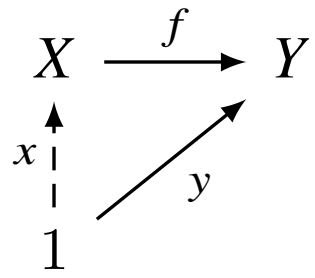
Internal Logic

Logical operator	Operation on $\mathrm{Sub}(A)$
truth: $\top$	top element (A itself)
falsity: $\bot$	bottom element (strict initial object)
conjunction: $\wedge$	intersection (pullback)
disjunction: $\vee$	union
implication: $\rightarrow$	Heyting implication
existential quantification: $\exists$	left adjoint to pullback
universal quantification: $\forall$	right adjoint to pullback

Diagonalization[Law69][31]

Definition (Point-Surjective)

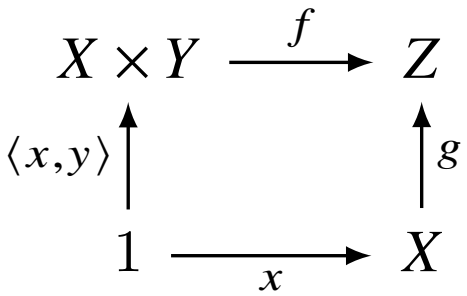
A morphism $f : X \to Y$ is *point-surjective* iff for every $y : 1 \to Y$, there is an $x : 1 \to X$ s.t. $y = f \circ x$.



Definition (Weakly Point-Surjective)

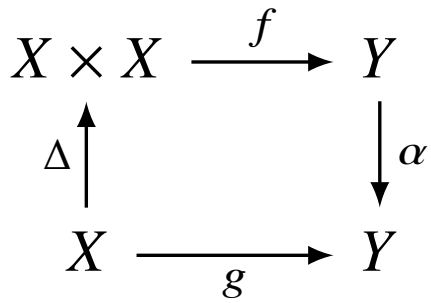
A morphism $f : X \times Y \to Z$ is *weakly point-surjective* iff for every $g : X \to Z$, there exists $y : 1 \to Y$ such that, for all $x : 1 \to X$:

$$g \circ x = f \circ \langle x, y \rangle$$



Theorem (Lawvere's Fixpoint Theorem)

Let $\mathbf{C}$ *be a category with a terminal object and binary products. If* $f : X \times X \to Y$ *is weakly point-surjective, then every* $\alpha : Y \to Y$ *has a fixpoint* $y : 1 \to Y$.

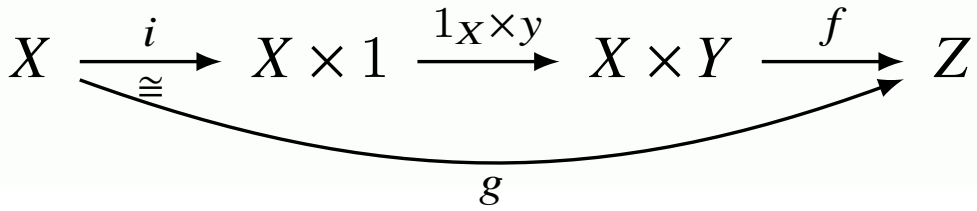


[31] Lawvere: Diagonal arguments and cartesian closed categories.
Yanofsky: A universal approach to self-referential paradoxes, incompleteness and fixed points.

Definition (Representability)

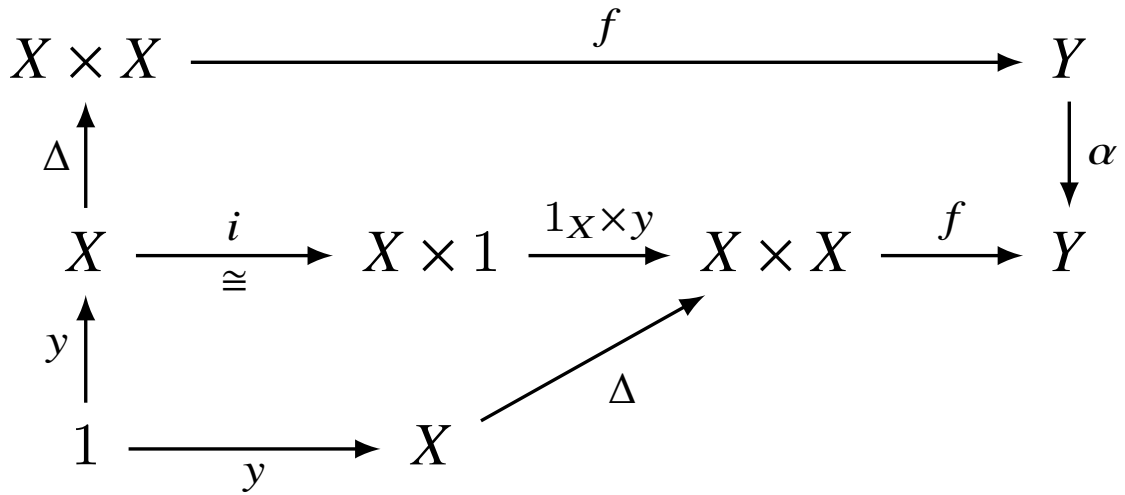
$g : X \to Z$ is representable by $f : X \times Y \to Z$ iff there exists $y : 1 \to Y$ s.t.

$$g = f \circ (1_X \times y) \circ i$$



Theorem (Another Version of Lawvere's Fixpoint Theorem)

Let $\mathbf{C}$ *be a category with a terminal object and binary products. For* $f : X \times X \to Y$, $\alpha : Y \to Y$, *if* $\alpha \circ f \circ \Delta$ *is representable by* f, *then* $\alpha : Y \to Y$ *has a fixpoint.*



If $\alpha \circ f \circ \Delta$ is represented by $f(-, y)$, then

$$\alpha \circ f \circ \Delta \circ y = f \circ (1_X \times y) \circ i \circ y = f \circ \Delta \circ y$$

Observations

In the proof of Lawvere's fixpoint theorem,

- associativity of the product $\times$ is never used
- commutativity of the product $\times$ is never used
- unitality of the product $\times$ with respect to 1 is never used
- projections are never used
- the universal property of $\times$ is only used to construct the diagonals $\Delta : X \to X \times X$
- 1 is only used to be the domain of global elements

Another Version of Lawvere's Fixpoint Theorem

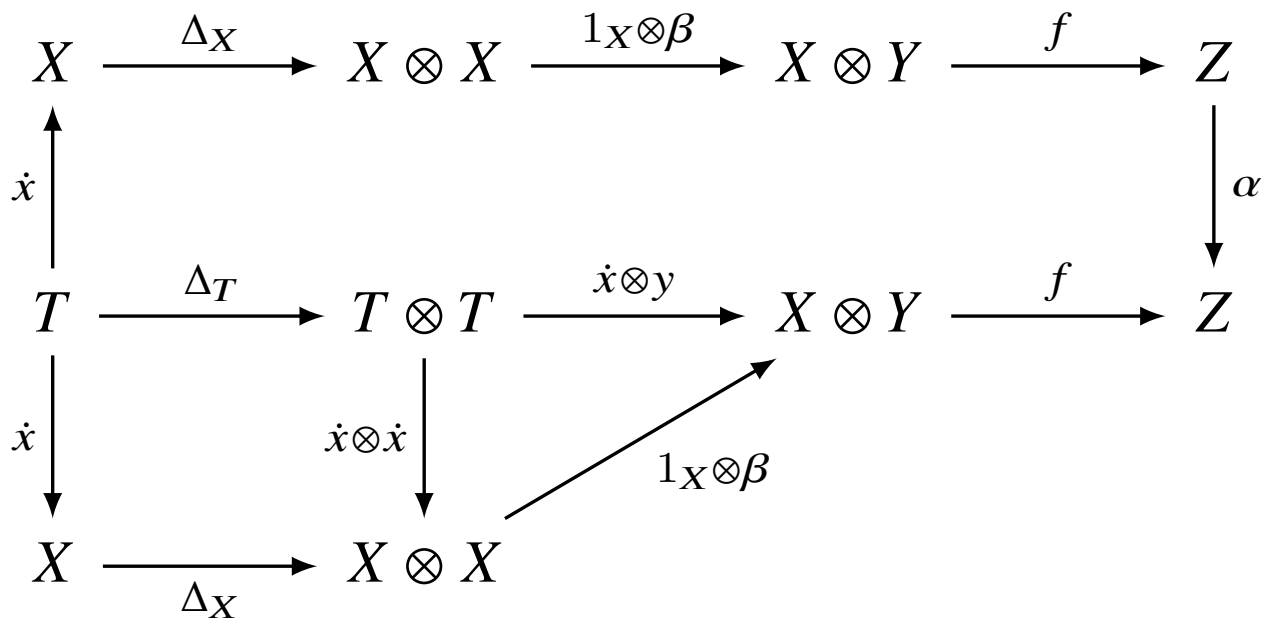
Definition (Pointed Magmoidal Category with Diagonals)

- A *magmoidal category* $(\mathbf{C}, \otimes)$ is a category $\mathbf{C}$ with a functor $\otimes : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$.
- A *pointed magmoidal category* $(\mathbf{C}, \otimes, T)$ is a magmoidal category with a distinguished object T.
- A magmoidal category is said to have *diagonals* $\Delta : X \to X \otimes X$ if there is a natural transformation $\Delta : 1_{\mathbf{C}} \to \otimes \circ \Delta_{\mathbf{C}}$ from the identity functor $1_{\mathbf{C}}$ to the composite $\mathbf{C} \xrightarrow{\Delta_{\mathbf{C}}} \mathbf{C} \times \mathbf{C} \xrightarrow{\otimes} \mathbf{C}$ of $\otimes$ with the diagonal functor $\Delta_{\mathbf{C}}$.

$$\begin{array}{ccc} X & \xrightarrow{\Delta_X} & X \otimes X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f \otimes f} \\ Y & \xrightarrow[\Delta_Y]{} & Y \otimes Y \end{array}$$

Theorem (Another Version of Lawvere's Fixpoint Theorem)

Let $(\mathbf{C}, \otimes, \Delta, T)$ *be a pointed magmoidal category with diagonals. For* $f : X \otimes Y \to Z$, $\alpha : Z \to Z$, $\beta : X \to Y$, *if* $\forall y : T \to Y . \exists \dot{x} : T \to X . [\beta \circ \dot{x} = y]$, *and if* $\exists y : T \to Y . \forall x : T \to X . \left[\alpha \circ f \circ (1_X \otimes \beta) \circ \Delta_X \circ x = f \circ (x \otimes y) \circ \Delta_T\right]$, *then* $\alpha \circ z = z$ *for some* $z : T \to Z$.



$$\alpha \circ f \circ (1_X \otimes \beta) \circ \Delta_X \circ \dot{x} = f \circ (\dot{x} \otimes y) \circ \Delta_T = f \circ (1_X \otimes \beta) \circ (\dot{x} \otimes \dot{x}) \circ \Delta_T$$
$$= f \circ (1_X \otimes \beta) \circ \Delta_X \circ \dot{x}$$

Lawvere Theory

Definition (Lawvere Theory)

A *Lawvere theory* is a category $\mathbb{T}$ with finite products and with a distinguished object A such that every object of $\mathbb{T}$ is a finite power of A:

$$\forall X \in \mathbb{T} \exists n \in \mathbb{N} : X \cong A^n$$

- A morphism $f : A^n \to A$ is called an n-ary operation (and, in particular, $1 \to A$ are called constants).
- Every $f : A^m \to A^n$ is the tupling of n operations $f_i : A^m \to A$ $(i = 1, \ldots, n)$.

Definition (Model)

A *model* (or algebra) of a Lawvere theory $\mathbb{T}$ in any category $\mathbf{C}$ with finite products is a finite-product-preserving functor

$$F : \mathbb{T} \to \mathbf{C}$$

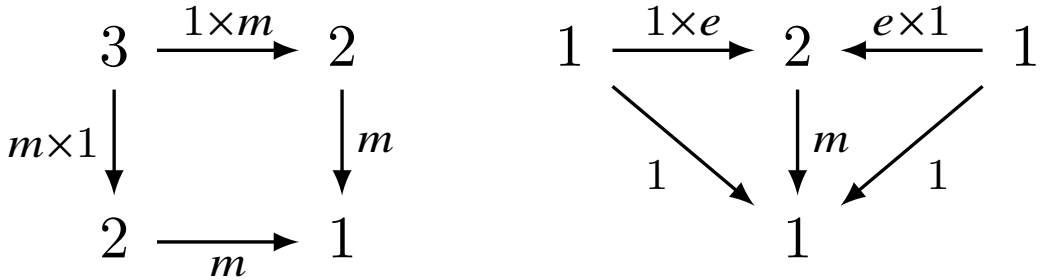
The category of $\mathbb{T}$-models in $\mathbf{C}$ is written $\mathrm{Mod}(\mathbb{T}, \mathbf{C})$.

Example: The Lawvere theory $\mathbb{T}_{\mathrm{Mon}}$ of monoids

The objects of $\mathbb{T}_{\mathrm{Mon}}$ are given by natural numbers, and the morphisms

$$m : 2 \to 1 \quad \text{and} \quad e : 0 \to 1$$

such that



Remark

The category $\mathrm{Mod}(\mathbb{T}_{\mathrm{Mon}}, \mathbf{Set})$ *of* $\mathbb{T}_{\mathrm{Mon}}$*-models in* $\mathbf{Set}$ *is equivalent to the category* $\mathbf{Mon}$ *of monoids and monoid homomorphisms.*

Functorial Semantics

- The syntactic category $\mathbf{C}_{\mathbb{T}}$ of $\mathbb{T}$ has as objects equivalence classes $[\varphi(x)]$ of formulas modulo $\mathbb{T}$.
- A morphism from $[\varphi(x)]$ to $[\psi(y)]$ is given by an equivalence class of formulas modulo $\mathbb{T}$, say $[\theta(x,y)]$, such that $\mathbb{T}$ proves that θ is a functional relation between φ and ψ:

$$\mathbb{T} \vdash \forall x \forall y [\theta(x,y) \to \varphi(x) \wedge \psi(y)] \wedge \forall x [\varphi(x) \to \exists! y \theta(x,y)]$$

Theorem

For any algebraic theory $\mathbb{T}$, there is a finite product category $\mathbf{C}_{\mathbb{T}}$ called the syntactic category of $\mathbb{T}$ such that

$$\mathrm{Mod}(\mathbb{T}, \mathbf{C}) \cong \mathrm{Hom}_{\mathrm{FP}}(\mathbf{C}_{\mathbb{T}}, \mathbf{C})$$

Theorem

The following are equivalent:

1. *For every model M of $\mathbb{T}$ in $\mathbf{C}$, $M \vDash s = t$*
2. $\mathbb{T} \vdash s = t$

Functorial Semantics

- Categories with finite products give multi-sorted theories
- Categories with finite limits give essentially algebraic theories
- Regular categories give logic with $\exists, \wedge, \vee$
- Pretoposes give full first-order logic $\exists, \forall, \neg, \wedge, \vee$
  - a theory is a category with some structure
  - a model is a functor $\mathbb{T} \to \mathbf{Set}$ that preserves the relevant structure
  - a homomorphism between models is a natural transformation between them
- It is useful to consider logical systems weaker than Lawvere theories.

Doctrine	Single-sorted	Typical theories
category	—	discrete dynamical systems
monoidal category	PRO	(co)monoids
symmetric monoidal category	PROP	commutative (co)monoids
cartesian category	Lawvere theory	groups, rings

Stone Duality

- A Boolean algebra is a bounded distributive lattice s.t. every element has a complement.
- A Stone space is a compact, Hausdorff, totally disconnected topological space.
- We write $\mathbf{BoolAlg}$ for the category of Boolean algebras and homomorphisms, $\mathbf{Stone}$ for the category of Stone spaces and continuous maps.
- Every Boolean algebra B induces a Stone space $\mathrm{St}(B)$ of its ultrafilters.
- Every Stone space X induces a Boolean algebra $\mathrm{cl}(X)$ of its clopen sets.
- St and cl are both functors:

$$\mathbf{BoolAlg}^{\mathrm{op}} \underset{\mathrm{cl}}{\overset{\mathrm{St}}{\rightleftarrows}} \mathbf{Stone}$$

- Moreover, these functors comprise an equivalence $\mathbf{BoolAlg}^{\mathrm{op}} \simeq \mathbf{Stone}$.

$\mathbf{BoolAlg}^{\mathrm{op}} \simeq \mathbf{Stone}$

- Every homomorphism $A \longrightarrow B$ corresponds to a continuous map $\mathrm{St}(B) \longrightarrow \mathrm{St}(A)$.
- The terminal $\bullet$ in $\mathbf{Stone}$ is the singleton space, which corresponds to $\mathrm{St}(\mathbf{2})$.
- A point $x \in X$ is a continuous map $x : \bullet \longrightarrow X$.
- This corresponds to a homomorphism $x^* : \mathrm{cl}(X) \longrightarrow \mathbf{2}$, which is a model of $\mathrm{cl}(X)$ in $\mathbf{2}$.
- A generalized point $x : Y \longrightarrow X$ corresponds to a homomorphism $x^* : \mathrm{cl}(X) \longrightarrow \mathrm{cl}(Y)$, which in turn corresponds to a model of $\mathrm{cl}(X)$ in $\mathrm{cl}(Y)$.
- So X is the (classifying) space of models of $\mathrm{cl}(X)$.
- Let $\Sigma = \{0, 1\} \in \mathbf{Stone}$ be the discrete space with two points. Then

$$\mathrm{cl} \cong \mathrm{Hom}_{\mathbf{Stone}}(-, \Sigma)$$

- $\mathrm{Hom}_{\mathbf{Stone}}(-, \Sigma)$ has the structure of Boolean algebra. Clopen sets correspond to propositional formulas.

Frame & Locale

- A frame A is a poset with all joins and all finite meets which satisfies the infinite distributive law:

$$x \wedge (\bigvee_i y_i) = \bigvee_i (x \wedge y_i)$$

- A frame homomorphism $f : A \to B$ is a function which preserves finite meets and arbitrary joins.
- Frames and frame homomorphisms form a category $\mathbf{Frm}$.
- The category $\mathbf{Locale}$ of locales is the opposite of the category of frames

$$\mathbf{Locale} := \mathbf{Frm}^{\mathrm{op}}$$

Relation to Topological Spaces

- Every topological space X has a frame of opens $\mathcal{O}_X$, which gives rise to a locale X_L. For every continuous function $f : X \to Y$, the inverse image map $f^{-1} : \mathcal{O}_Y \to \mathcal{O}_X$ is a frame homomorphism, so f induces a continuous map $f_L : X_L \to Y_L$. Thus we have a functor

$$(-)_L : \mathbf{Top} \to \mathbf{Locale}$$

- Conversely, if X is any locale, we define a point of X to be a continuous map $1_L \to X$. Here 1_L is the terminal locale. The elements of the frame $\mathcal{O}_X$ induce a topology on the set of points of X in an obvious way, thereby giving rise to a topological space X_P. Any continuous map $f : X \to Y$ of locales induces a continuous map $f_P : X_P \to Y_P$ of spaces, so we have a functor

$$(-)_P : \mathbf{Locale} \to \mathbf{Top}$$

- The functor $(-)_L$ is left adjoint to $(-)_P$.

$$\mathbf{Top} \underset{(-)_P}{\overset{(-)_L}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{Locale}$$

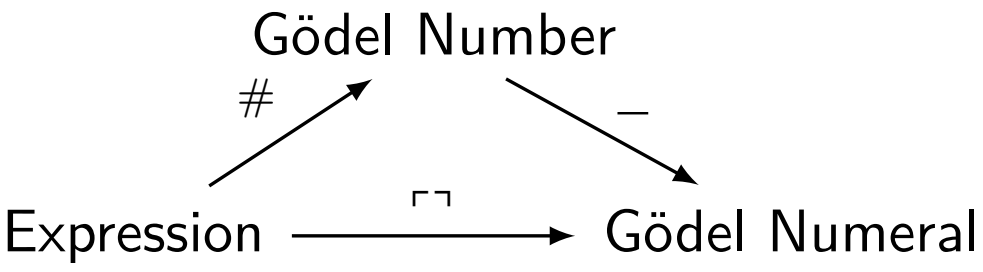
Lindenbaum Category

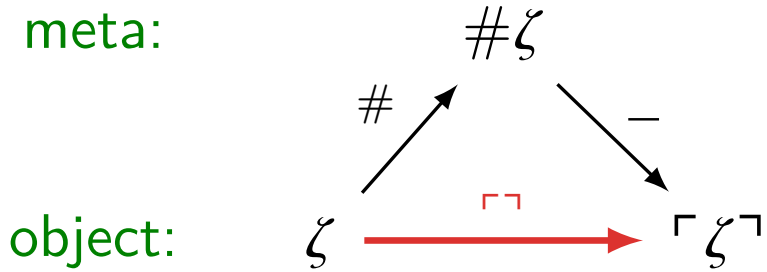
- Consider a first-order theory $\mathbb{T}$.
  We form $\mathbf{C}_{\mathbb{T}}$ a syntactic category of $\mathbb{T}$ in the following way:
  The $\mathbf{C}_{\mathbb{T}}$-objects are generated by a sort object A (more object if the theory is multi-sorted), and an object 2, by closure under products.

$\mathbf{C}_{\mathbb{T}}$-morphisms	Equivalance classes of ...
$A^n \longrightarrow 2$	provably equivalent formulas of n variables
$A^n \longrightarrow 2 \times 2$	provably equivalent tuples of formulas of n variables
$A^n \longrightarrow A$	provably equivalent terms of n variables
$1 \longrightarrow 2$	sentences
$1 \longrightarrow A$	constant terms
$\top : 1 \longrightarrow 2$	sentences provable
$\bot : 1 \longrightarrow 2$	sentences refutable
$2^n \longrightarrow 2$	propositional operations e.g., $\neg : 2 \longrightarrow 2$

- A theory is *consistent* iff $\mathrm{Hom}(1, 2)$ contains at least two elements $\mathrm{Hom}(1, 2) \supset \{\top, \bot\}$. Equivalently, there is a morphism $\neg : 2 \longrightarrow 2 : \neg\varphi \neq \varphi$ for all $\varphi : 1 \longrightarrow 2$.
- A theory is *complete* iff $\mathrm{Hom}(1, 2) = \{\top, \bot\}$.

Gödel Encoding





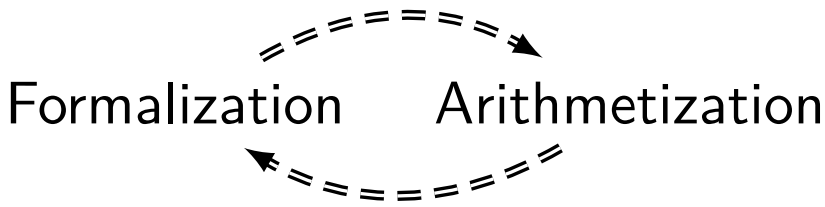


$$\ulcorner\zeta\urcorner = \underline{\#\zeta} = s^{\#\zeta}0$$

meta-language

$\Uparrow ? \Downarrow$

object-language



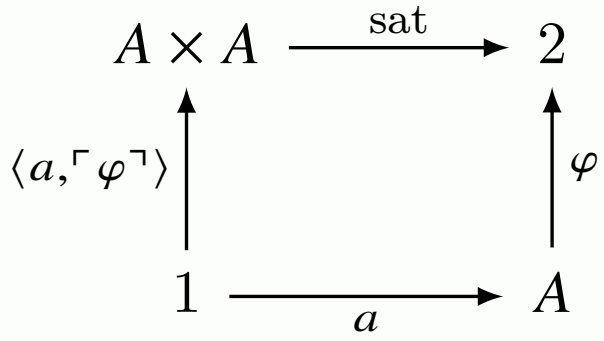
Definition (Gödel Encoding)

Gödel encoding is an injective map $\ulcorner\urcorner : \mathrm{Hom}(A^n, 2) \to \mathrm{Hom}(1, A)$.

Undefinability of sat

Definition (Satisfiabilty Predicate)

Satisfiabilty predicate is definable in $\mathbb{T}$ iff there is $\mathrm{sat} : A \times A \to 2$ such that, for any $\varphi : A \to 2$, and for all $a : 1 \to A$, we have $\mathrm{sat}\langle a, \ulcorner \varphi \urcorner \rangle = \varphi a$.



Traditionally,
$\mathbb{T} \vdash \mathrm{sat}\langle a, \ulcorner \varphi \urcorner \rangle \leftrightarrow \varphi a$

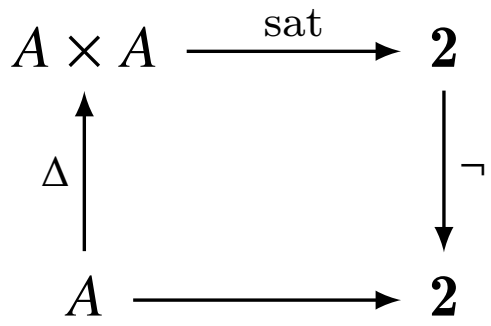
Remark: This is exactly the condition for weak point-surjectivity!

Theorem (Undefinability of sat)

If $\mathbb{T}$ is consistent, then sat is not definable in $\mathbb{T}$.

Proof.

If sat is definable in $\mathbb{T}$, then $\neg$ has a fixpoint. □



Truth Predicate

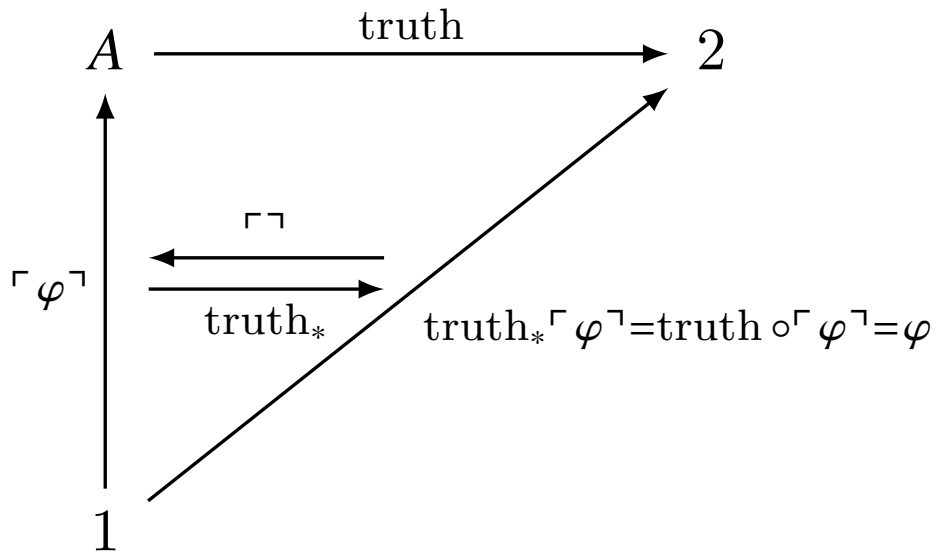
$\ulcorner$snow is white$\urcorner$ is true iff snow is white.

Definition (Truth Predicate)

Truth Predicate is definable in $\mathbb{T}$ iff there is $\mathrm{truth} : A \to 2$ such that

$$\mathrm{truth}_* : \mathrm{Hom}(1, A) \to \mathrm{Hom}(1, 2)$$

is a retraction of $\ulcorner\urcorner : \mathrm{Hom}(1, 2) \to \mathrm{Hom}(1, A)$, i.e., $\mathrm{truth} \circ \ulcorner\varphi\urcorner = \varphi$.

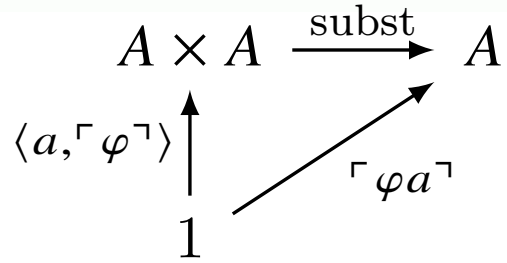


Traditionally,
$\mathbb{T} \vdash \mathrm{truth}(\ulcorner\varphi\urcorner) \leftrightarrow \varphi$

Tarski's Undefinability of truth Theorem

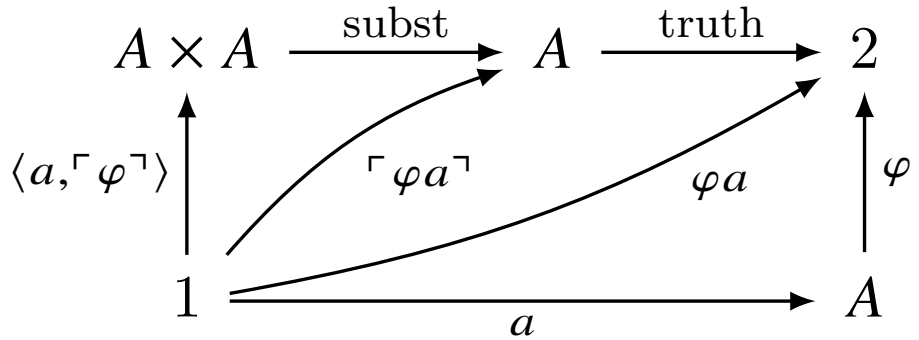
Definition (Substitution)

"*substitution*" is definable in $\mathbb{T}$ iff there is $\mathrm{subst} : A \times A \longrightarrow A$ such that $\mathbb{T} \vdash \mathrm{subst}\langle a, \ulcorner \varphi \urcorner \rangle = \ulcorner \varphi a \urcorner$.



Theorem (Tarski's Undefinability of truth Theorem)

If $\mathbb{T}$ is consistent and "substitution" is definable in $\mathbb{T}$, then truth is not definable in $\mathbb{T}$.



$$\mathrm{sat} = \mathrm{truth} \circ \mathrm{subst}$$

Gödel's First Incompleteness Theorem

Definition (Provability Predicate)

Provability is representable in $\mathbb{T}$ iff there is $\mathrm{prov} : A \to 2$ such that, for any $\varphi : 1 \to 2$, we have $\varphi = \top \iff \mathrm{prov}(\ulcorner \varphi \urcorner) = \top$.

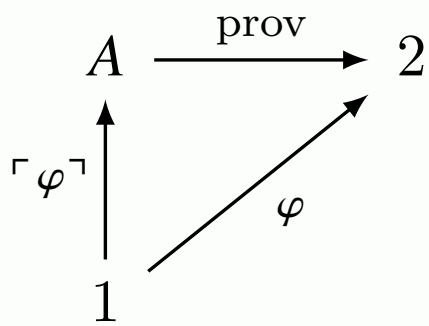
Remark: Traditionally, $\mathbb{T} \vdash \varphi \iff \mathbb{T} \vdash \mathrm{prov}(\ulcorner \varphi \urcorner)$.

Theorem (Gödel's First Incompleteness Theorem)

If $\mathbb{T}$ is consistent, and "substitution" is definable, and "provability" is representable, then $\mathbb{T}$ is not complete.

Proof.

If $\mathbb{T}$ is complete, then $\varphi = \top$ or $\varphi = \bot$.
$\varphi = \top \implies \mathrm{prov}(\ulcorner \varphi \urcorner) = \top$.
$\varphi = \bot \implies \mathrm{prov}(\ulcorner \varphi \urcorner) = \bot$.
Therefore, $\mathrm{prov} \circ \ulcorner \varphi \urcorner = \varphi$ for all $\varphi : 1 \to 2$.
Namely, $\mathrm{truth} = \mathrm{prov}$.



□

Synthetic Mathematics vs Analytic Mathematics

- Synthetic
  - Basic objects are taken as primitive.
  - Their properties & relations are axiomatized.
  - We work within the axiomatic system.
- Analytic
  - Basic objects are constructed from other objects.
  - Their properties & relations are deduced.
  - We work in a wider mathematical environment.

Examples:

- Analytic geometry analyzes the points, lines, etc. in terms of $\mathbb{R}^2$.
- Synthetic geometry axiomatizes points and lines as primitive notions, by specifying what we can do with them.
- Synthetic differential geometry: "All maps are smooth!"
- Synthetic topology: "All maps are continuous!"
  — construe spaces as primitive and axiomatize them directly, rather than build spaces out of sets.
- Synthetic computability: "All maps are computable!"
  — Axiom: there are countably many countable subsets of $\mathbb{N}$.

- Mathematics is the analysis of invariants and of the transformations that preserve them (including the analysis of non-preservations, deformations and symmetry breakings).
- We are presented with an open universe of categories, then, to which new categories are constantly added; new invariants, and new transformations. Concepts are created by being correlating with existent ones, and by deforming one into the other, thus enriching them, paying attention to the meaning of what is being done.

Synthetic		**Analytic**
Theory	logic interpretation $\longrightarrow$	Model
HOL	internal language $\longrightarrow$	Topos
IHOL	Kleene realizability $\longrightarrow$	Effective Topos

- In $\mathbf{Eff}$ all objects and morphisms are equipped with computability structure.
- We need not know how $\mathbf{Eff}$ is built — we just use the logic and axioms which are valid in it.

Symbol	External view of Eff	Internal view of Eff
$\mathbb{N}$	natural numbers	natural numbers
$\mathbb{R}$	computable reals	all reals
$f : \mathbb{N} \longrightarrow \mathbb{N}$	recursive function	any function
$e : \mathbb{N} \twoheadrightarrow A$	recursive enumeration of A	any enumeration of A
$\{\bot, \top\}$	truth values	decidable truth values
$\mathbf{\Omega}$	truth values of $\mathbf{Eff}$	truth values
$\forall x$	computably for all x	for all x
$\exists x$	there exists computable x	there exists x
$p \vee \neg p$	decision procedure for p	p or not p

The Effective Topos

For any model M of computation, such as Turing machines (TM), or lambda calculus, there is an associated effective topos $\mathbf{Eff}(M)$.

- "Every number is either prime or not."
  — Trivially true in $\mathbf{Set}$, nontrivially true in $\mathbf{Eff}(\mathrm{TM})$.
- "Every function $\mathbb{N} \to \mathbb{N}$ is either the zero function or not."
  — Trivially true in $\mathbf{Set}$, false in $\mathbf{Eff}(\mathrm{TM})$.
- "Every function $\mathbb{N} \to \mathbb{N}$ is computable by a Turing machine."
  — False in $\mathbf{Set}$, trivially true in $\mathbf{Eff}(\mathrm{TM})$.
- "Every function $\mathbb{R} \to \mathbb{R}$ is continuous."
  — False in $\mathbf{Set}$, nontrivially true in $\mathbf{Eff}(\mathrm{TM})$.

Remark: If a statement has a constructive proof, then it holds in any topos.

Exercise

If you know an example of a noncomputable function, look and see where it uses excluded middle.

Remark

- Using constructive logic not only ensures our proofs are constructive; it gives us axiomatic freedom: we can assume powerful axioms that would classically be inconsistent.
- This places us firmly on the side of viewing constructive logic as a "strange new universe", rather than just a refined notion of proof for ordinary mathematics.
- This can be made precise by constructing models of constructive logic, with new nonclassical axioms, inside classical logic.

Synthetic Computability

Definition (Assembly)

An assembly $\mathbf{X} = (X, \vDash_{\mathbf{X}})$ is a set X with a realizability relation $\vDash_{\mathbf{X}} \subset \mathbb{N} \times X$ such that

$$\forall x \in X \exists n \in \mathbb{N} : n \vDash_{\mathbf{X}} x$$

Definition (Assembly Morphism)

An *assembly morphism* $f : \mathbf{X} \to \mathbf{Y}$ is a function $f : X \to Y$ for which there exists $e \in \mathbb{N}$ such that

$$\forall x \in X \forall n \in \mathbb{N} : n \vDash_{\mathbf{X}} x \implies \varphi_e(n) \downarrow \ \& \ \varphi_e(n) \vDash_{\mathbf{Y}} f(x)$$

$\mathbf{Asm}$ is the category of assemblies.

Examples:

- Natural numbers: $\mathbf{N} = (\mathbb{N}, \vDash_{\mathbf{N}})$ where $n \vDash_{\mathbf{N}} m \iff n = m$.
- Partial recursive functions: $\mathcal{R} = (\mathcal{R}, \vDash_{\mathcal{R}})$ where $\mathcal{R}$ is the set of partial recursive functions and $n \vDash_{\mathcal{R}} f \iff \varphi_n = f$.
- Recursive enumerable sets: $\mathbf{E} = (\mathcal{E}, \vDash_{\mathbf{E}})$ where $\mathcal{E}$ is the set of r.e. sets and $n \vDash_{\mathbf{E}} A \iff A = W_n := \{m : \varphi_n(m) \downarrow\}$.

- Boolean truth values: $\mathbf{2} = (\{\bot, \top\}, \vDash_{\mathbf{2}})$ where $0 \vDash_{\mathbf{2}} \bot$ and $1 \vDash_{\mathbf{2}} \top$.
- Semidecidable truth values: $\Sigma = (\{\bot, \top\}, \vDash_{\Sigma})$ where $n \vDash_{\Sigma} \top \iff \varphi_n(n) \downarrow$ and $n \vDash_{\Sigma} \bot \iff \varphi_n(n) \uparrow$.
- Classical sets: $\nabla X = (X, \vDash_{\nabla X})$ where $n \vDash_{\nabla X} x$ for all $n \in \mathbb{N}$ and $x \in X$.
- Assembly morphisms $\mathrm{Hom}_{\mathbf{Asm}}(\mathbf{N}, \mathbf{N})$ are exactly the total computable functions from $\mathbb{N}$ to $\mathbb{N}$.
- Assembly morphisms $\mathrm{Hom}_{\mathbf{Asm}}(\mathbf{N} \times \mathbf{N}, \mathbf{N})$ are exactly the total computable functions from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$.
- Assembly morphisms $\mathrm{Hom}_{\mathbf{Asm}}(\mathbf{N}, \mathbf{2})$ are decision procedure/subsets.
- Assembly morphisms $\mathrm{Hom}_{\mathbf{Asm}}(\mathbf{N}, \Sigma)$ are semi-decision procedure/subsets.
- By Lawvere' theorem, $\mathbf{2}^{\mathbf{N}}$ and $\mathbf{N}^{\mathbf{N}}$ are not enumerable.
- We know that there is an enumeration of r.e. sets, thus a weakly point-surjective $W : N \twoheadrightarrow \Sigma^{\mathbf{N}}$.
- Hence, by Lawvere's theorem every map $\Sigma \longrightarrow \Sigma$ has a fixpoint.
- It immediately follows that negation is not definable on Σ and hence r.e. sets are not closed under complements.

Kleene's Fixpoint Theorem

- Note that
$$W : \mathbf{N} \twoheadrightarrow \Sigma^{\mathbf{N}} \cong \Sigma^{\mathbf{N}\times\mathbf{N}} \cong \Sigma^{\mathbf{N}^{\mathbf{N}}}$$
so every map $F : \Sigma^{\mathbf{N}} \to \Sigma^{\mathbf{N}}$ has a fixpoint.
- We can identify the exponent $\Sigma^{\mathbf{N}}$ with an assembly $(\mathcal{E}, \vDash_{\mathbf{E}})$.
- A map $F : \Sigma^{\mathbf{N}} \to \Sigma^{\mathbf{N}}$ is an enumeration operator: $F(W_e) = W_{f(e)}$ for some computable f.
- Lawvere's theorem shows that every such operator has a fixpoint: $W_e = W_{f(e)}$.

Theorem

$$\mathcal{E} \cong \Sigma^{\mathbf{N}}$$

Proof.

The isomorphism $f : \mathcal{E} \to \Sigma^{\mathbf{N}}$ is given by

$$f : W_e \mapsto \lambda x.[\![\exists n \in \mathbf{N} : \varphi_e(n) = x]\!]$$

Conversely, for $h \in \Sigma^{\mathbf{N}}$,

$$g : h \mapsto \left\{n \in \mathbf{N} : \varphi_{\varphi_e(n)}(\varphi_e(n)) \downarrow\right\}$$

where φ_e $\vDash_\Sigma$-realizes h. □

Rice's Theorem

Theorem (Rice's Theorem)

If A has the fixed point property then every map $A \to \mathbf{2}$ is constant.

Proof.

Given $f : A \to \mathbf{2}$ and any $x, y \in A$, we show that $f(x) = f(y)$.
Define

$$g(z) := \begin{cases} x & \text{if } f(z) = f(y) \\ y & \text{otherwise} \end{cases}$$

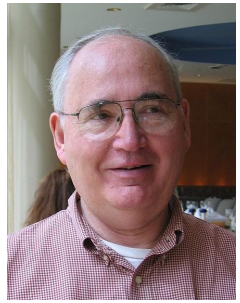
There is $u \in A$ s.t $u = g(u)$.
If $f(u) = f(y)$, then $u = g(u) = x \implies f(x) = f(u) = f(y)$.
If $f(u) \neq f(y)$, then $u = g(u) = y \implies f(y) \neq f(y)$. □

Corollary

- *Every map $\Sigma \to \mathbf{2}$ is constant. (Halting problem)*
- *Every map $\Sigma^{\mathbf{N}} \to \mathbf{2}$ is constant. (Rice's theorem)*

Scott's Reflexive Domain[SV84][32]

A *reflexive domain* is a space where every object is a transformation, and every transformation corresponds uniquely to an object.



- **Po**: the category of all partially-ordered sets (posets) and monotone mappings.
- **CPo**: all posets which have a least element $\bot$, and such that every countable monotone chain has a supremum.
- A monotone map f is continuous iff it preserves supremums of countable chains $f(\coprod_n x_n) = \coprod_n f(x_n)$.
- By Tarski's argument, it is then clear that every continuous endomorphism $f : D \to D$ of such poset $D \in \mathbf{CPo}$ has a least fixpoint, namely the supremum of the monotone sequence $\coprod_n f^n(\bot)$.
- **CCPo**: the subcategory of **CPo** obtained by restricting the morphisms to be continuous.

[32] J. Soto-Andrade & F. J. Varela: Self-Reference and Fixed Points: A Discussion and an Extension of Lawvere's Theorem.

Examples of **CPo**

- powersets $(\mathrm{P}(X), \subset)$. Least upper bounds = unions.
- partial functions $(\mathbb{N} \rightharpoonup \mathbb{N}, \sqsubset)$, where $\mathbb{N} \rightharpoonup \mathbb{N}$ is the set of partial functions on $\mathbb{N}$, and

$$f \sqsubset g := \forall xy \in \mathbb{N}.\ f(x) \simeq y \to g(x) \simeq y$$

- flat nats $(\mathbb{N}_\perp, \subset)$, where $\mathbb{N}_\perp := \mathbb{N} \cup \{\perp\}$, and

$$x \sqsubset y := x = \perp \vee x = y$$

- streams $(\Sigma^\#, \subset)$, where $\Sigma^\#$ is the set of finite or infinite sequences over Σ, and

$$x \sqsubset y \text{ iff } x \text{ is a prefix of } y$$

Remark:

- $\sqsubset$: the partial order of definedness
- $\bigsqcup$: least upper bounds (lub) as limits
- $\bigsqcup_i f_i$: functions defined by recursion as fixpoints
- Continuity serves as an '*intrinsic* approximation' to computability.

Reflexive Domain

For $D \in \mathrm{ob}(\mathbf{CPo})$, define

$$D_1 \coloneqq D$$

$$D_{n+1} \coloneqq [D_n, D_n]$$

$$i_n : D_n \to D_{n+1}$$

$$j_n : D_{n+1} \to D_n$$

$$i_1(x) \coloneqq \mathrm{con}(x) : D \to D, \text{ where } \mathrm{con}(x)(y) \coloneqq x$$

$$j_1(f) \coloneqq f(\bot)$$

$$i_{n+1}(x_{n+1}) \coloneqq i_n \circ x_{n+1} \circ j_n$$

$$j_{n+1}(x_{n+2}) \coloneqq j_n \circ x_{n+2} \circ i_n$$

$$\begin{array}{ccc} D_{n+1} & \xrightarrow{i_{n+1}(x_{n+1})} & D_{n+1} \\ {\scriptstyle j_n}\downarrow & & \uparrow{\scriptstyle i_n} \\ D_n & \xrightarrow[x_{n+1}]{} & D_n \end{array} \qquad \begin{array}{ccc} D_{n+1} & \xrightarrow{x_{n+2}} & D_{n+1} \\ {\scriptstyle i_n}\uparrow & & \downarrow{\scriptstyle j_n} \\ D_n & \xrightarrow[j_{n+1}(x_{n+2})]{} & D_n \end{array}$$

then we have

$$i_n \circ j_n \leq 1_{D_{n+1}}$$

$$j_n \circ i_n = 1_{D_n}$$

Reflexive Domain

$$D^\infty := \varprojlim(D_n, j_n)$$
$$D_\infty := \varinjlim(D_n, i_n)$$

$$D^\infty = \left\{ x \in \prod_{n=1}^{\infty} D_n : x_n \in D_n \text{ and } j_n(x_{n+1}) = x_n \text{ for all } n \right\}$$
$$D_\infty = \left\{ x \in D^\infty : x_{m+1} = i_m(x_m) \text{ for all } m \geq n, \text{ for some } n \right\}$$

$$J_n : D^\infty \to D_n :: x \mapsto x_n$$
$$I_n : D_n \to D_\infty :: x_n \mapsto \left(j_1 \circ \cdots \circ j_{n-1}(x_n), \ldots, j_{n-1}(x_n), x_n, i_n(x_n), i_{n+1} \circ i_n(x_n), \ldots\right)$$

Then we have

$$I_n \circ J_n \leq 1_{D_\infty}$$
$$J_n \circ I_n = 1_{D_n}$$

Reflexive Domain

Define

$$F : D^\infty \to [D^\infty \to D^\infty]$$
$$G : [D^\infty \to D^\infty] \to D^\infty$$

as

$$F(x) = \coprod_{n=0}^{\infty} I_n \circ x_{n+1} \circ J_n$$
$$G(f) = \coprod_{n=0}^{\infty} I_{n+1}(J_n \circ f \circ I_n)$$

Then D^∞ is reflexive.

$$F \circ G = 1_{[D^\infty \to D^\infty]}$$
$$G \circ F = 1_{D^\infty}$$

Theorem

- *In the category* **CPo**, *we have* $D^\infty \cong [D_\infty, D^\infty]$.
- *In the category* **CCPo**, *we have* $D^\infty \cong [D^\infty, D^\infty]$.

Models of λ-Calculus

Definition (Reflexive Object)

Let $\mathbf{C}$ be a CCC. An object D is *reflexive* iff D^D is a retract of D, i.e. there are $\mathrm{app} : D \to D^D$ and $\mathrm{lam} : D^D \to D$ s.t. $\mathrm{app} \circ \mathrm{lam} = 1_{D^D}$.

Definition $(D, \cdot, [\![\,]\!])$

- Let $\mathbf{C}$ be a CCC with reflexive object D via $\mathrm{app}, \mathrm{lam}$. For $x, y : 1 \to D$,
$$x \cdot y \coloneqq \mathrm{app}(x)(y)$$
- Let $\rho : \mathrm{Var} \to D$ be a variable assignment. Define $[\![\,]\!]_\rho : \Lambda \to D$ as:
$$[\![x]\!]_\rho \coloneqq \rho(x)$$
$$[\![MN]\!]_\rho \coloneqq [\![M]\!]_\rho \cdot [\![N]\!]_\rho$$
$$[\![\lambda x.M]\!]_\rho \coloneqq \mathrm{lam}(f) \text{ where } f : d \mapsto [\![M]\!]_{\rho(d/x)}$$

Definition: An object D has *enough points* iff for all $f, g : D \to D$:
$f \neq g \implies \exists x : 1 \to D : fx \neq gx$.

Theorem

Any reflexive object D has enough points, iff, $(D, \cdot, [\![\,]\!])$ is a λ-model.

Remark

- Since $\mathrm{app} \circ \mathrm{lam} = 1_{D^D}$, then for every $f \in D^D$, there exists $\mathrm{lam}(f)$ s.t. $f = \mathrm{app} \circ \mathrm{lam}(f)$.
- Therefore, $\mathrm{app} : D \to D^D$ is point-surjective.
- According to Lawvere's fixpoint theorem, every $\alpha : D \to D$ has a fixpoint.
- The Y-combinator is such a fixpoint.

Heritability of the Fixpoint Property

Definition (Weak Retraction)

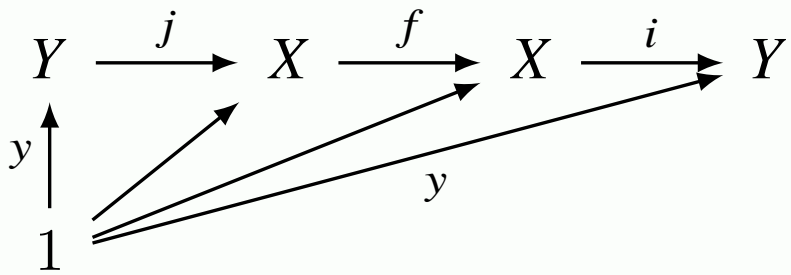
$j : Y \to X$ is a *weak retraction* of Y onto X iff there is a morphism $i : X \to Y$ s.t. for all $x : 1 \to X$,

$$(j \circ i) \circ x = x$$

Theorem

Let $\mathbf{C}$ be a category with terminal object 1. If X is a weak retract of Y and Y has the fixpoint property, then X has the fixpoint property.

Proof.



Let $f : X \to X$. Then $i \circ f \circ j : Y \to Y$, and there is $y : 1 \to Y$ s.t. $(i \circ f \circ j) \circ y = y$. It follow that

$$j \circ y = j \circ (i \circ f \circ j \circ y) = (j \circ i) \circ (f \circ j \circ y) = f \circ j \circ y$$

Theorem

Let $\mathbf{C}$ *be a Cartesian Closed Category. Assume we have* $\{X_n, i_n, j_n\}_{n\geq 0}$ *s.t.* $i_n : X_n \to X_{n+1}$, $j_n : X_{n+1} \to X_n$, *with* $j_n \circ i_n = 1_{X_n}$ *for all* n, *and* $j_n(x_{n+1}) = x_{n+1} \circ i_{n-1}$ *for all* $x_{n+1} \in X_{n+1}$, *where* $X_{n+1} := Y^{X_n}$. *Let* $X_\infty := \varinjlim(X_n, i_n)$ *and* $X^\infty := \varprojlim(X_n, j_n)$ *whenever such limits exist. Then*

$$X^\infty \cong Y^{X_\infty}$$

Proof.

Let us abbreviate $\mathrm{Hom}(X, Y)$ as $[X, Y]$.

$$\begin{aligned}
[Z, Y^{X_\infty}] &\cong [Z \times X_\infty, Y] \\
&\cong [Z \times \varinjlim X_n, Y] \\
&\cong [\varinjlim (Z \times X_n), Y] \\
&\cong \varprojlim [Z \times X_n, Y] \\
&\cong \varprojlim [Z, Y^{X_n}] \\
&\cong [Z, \varprojlim Y^{X_n}] \\
&\cong [Z, X^\infty]
\end{aligned}$$

□

Remark: If $X_\infty = X^\infty$, then Y has the fixpoint property.

Conjecture: The fixpoint property in any structure is a reflection of a higher reflexive domain of which it is a retraction.

Restriction Category

Definition (Restriction Category)

A restriction category is a category with a restriction operator $\overline{f} : A \to A$ for each morphism $f : A \to B$ satisfying

1. $f \circ \overline{f} = f$
2. $\overline{f} \circ \overline{g} = \overline{g} \circ \overline{f}$ whenever $\operatorname{dom} f = \operatorname{dom} g$
3. $\overline{g \circ \overline{f}} = \overline{g} \circ \overline{f}$ whenever $\operatorname{dom} f = \operatorname{dom} g$
4. $\overline{g} \circ f = f \circ \overline{g \circ f}$ whenever $\operatorname{dom} g = \operatorname{cod} f$

Example

- Every category admits the trivial restriction operator $\overline{f} = 1_{\operatorname{dom} f}$.
- Sets and partial functions.

$$\overline{f} = \begin{cases} x & f(x) \downarrow \\ \uparrow & \text{otherwise} \end{cases}$$

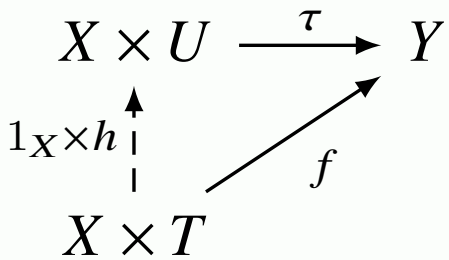
Definition: A total morphism $f : A \to B$ is a morphism for which $\overline{f} = 1_A$.

Turing Category

Definition: A *cartesian restriction category* is a restriction category which has a terminal object and for which each pair of objects has a product.

Definition (Turing Category)

A *Turing category* is a cartesian restriction category $\mathbf{C}$ with a *Turing object* U such that, for each $X, Y \in \mathbf{C}$, there is a *universal application morphism* $\tau : X \times U \to Y$, for any $f : X \times T \to Y$, there exists a total morphism $h : T \to U$ for which the following diagram is commutative:



Remark: In the special case when h is uniquely determined we say that τ is extensional.
In the special case when $T = 1$ is the terminal object, we say $h : 1 \to U$ is a code for f.

Remarks

- cartesian products — to pair (the codes of) data and programs,
- restriction operator — a notion of partiality — to represent programs (morphisms) which do not necessarily halt,
- Turing object U — to represent the "codes" of all programs.

- A reflexive object in a CCC is an object U together with embedding-retraction pairs $U^U \underset{\text{app}}{\overset{\text{lam}}{\rightleftarrows}} U$ s.t. $\text{app} \circ \text{lam} = 1_{U^U}$.
- A reflexive object is said to be extensional when also $\text{lam} \circ \text{app} = 1_U$.
- In a Turing category with Turing object U, the Turing morphism is defined by taking an embedding-retraction pair $(\text{lam}, \text{app}) : U^U \triangleleft U$,

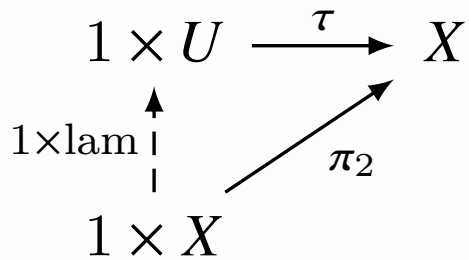
$$\tau : U \times U \xrightarrow{1_U \times \text{app}} U \times U^U \xrightarrow{\varepsilon} U$$

Remark: Viewed as a model of lambda calculus, the equation $\text{app} \circ \text{lam} = 1_{U^U}$ of a reflexive object represents β-reduction $(\lambda x.M)N = M[N/x]$, while the equation $\text{lam} \circ \text{app} = 1_U$ of an extensional reflexive object represents η-reduction $\lambda x.Mx = M$.

Theorem

In a Turing category $\mathbf{C}$ *with Turing object* U*, every object* $X \in \mathbf{C}$ *is a retract of* U*.*

Proof.



Let $\mathrm{app} \coloneqq \langle\, !\,, 1_X \rangle \circ \tau$. Then

$$\mathrm{app} \circ \mathrm{lam} = 1_X$$

□

Theorem (Recognition Criterion for Turing Categories)

For a cartesian restriction category $\mathbf{C}$*, the following are equivalent:*

1. $\mathbf{C}$ *is a Turing category.*
2. *There is an object* U *of which every object is a retract, and for which there exists a universal self-application Turing morphism*
$$U \times U \xrightarrow{\tau} U \ .$$

Proof.

"$2 \implies 1$": For arbitrary objects $X, Y \in \mathbf{C}$, by assumption we have $(\mathrm{lam}_X, \mathrm{app}_X) : X \triangleleft U$ and $(\mathrm{lam}_Y, \mathrm{app}_Y) : Y \triangleleft U$.

$$X \times U \xrightarrow{\tau} Y \quad = \quad X \times U \xrightarrow{\mathrm{lam}_X \times 1_U} U \times U \xrightarrow{\tau} U \xrightarrow{\mathrm{app}_Y} Y$$

$$\begin{array}{ccccccc} X \times U & \xrightarrow{\mathrm{lam}_X \times 1_U} & U \times U & \xrightarrow{\tau} & U & \xrightarrow{\mathrm{app}_Y} & Y \\ {\scriptstyle 1_X \times h}\uparrow & & {\scriptstyle 1_U \times h}\uparrow & & {\scriptstyle \mathrm{lam}_Y \circ f}\uparrow & \nearrow_{f} & \\ X \times T & \xrightarrow[\mathrm{lam}_X \times 1_T]{} & U \times T & \xrightarrow[\mathrm{app}_X \times 1_T]{} & X \times T & & \end{array}$$

Example of Turing Category

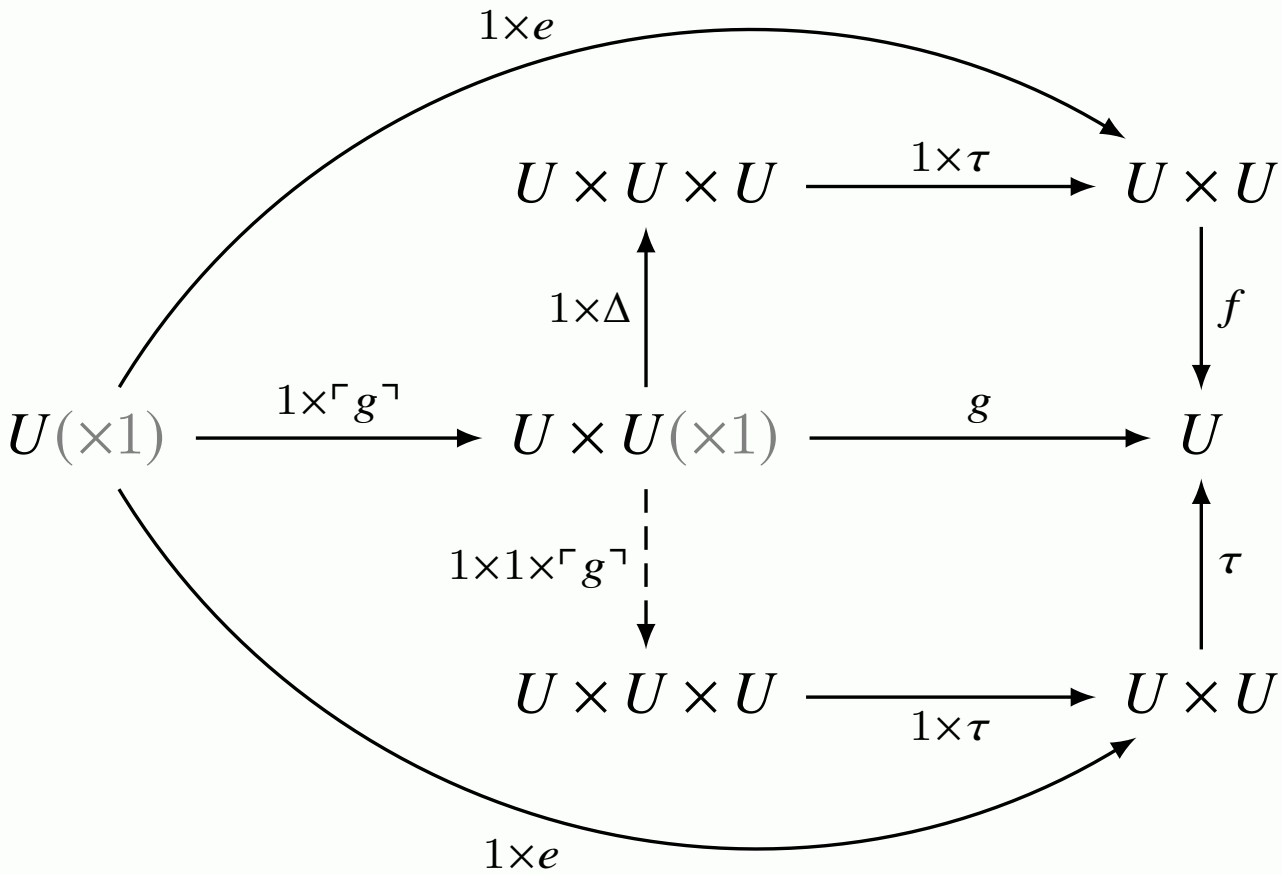
The classical category $\mathbf{Comp}(\mathbb{N})$ of partial recursive functions.

- objects: $0, 1, 2 \ldots$ the natural numbers.
- morphism: $f : m \to n$ partial recursive functions $f : \mathbb{N}^m \to \mathbb{N}^n$.
- Turing object: $1 (= \mathbb{N}^1)$.
- Turing morphism: $\tau : \mathbb{N} \times \mathbb{N} \to \mathbb{N} :: (m, n) \mapsto \varphi_n(m)$.

Theorem (Second Recursion Theorem)

In any Turing category, for any partial map $f : U \times U \to U$ on the Turing object U, there is a total point $e : 1 \to U$ such that $\tau(1 \times e) = f(1 \times e)$.

Proof.

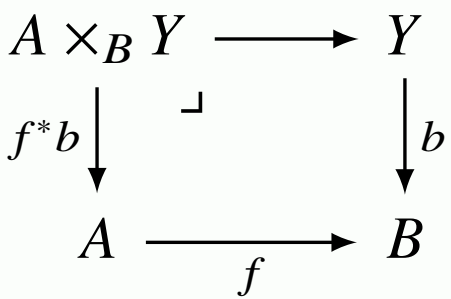


$$g := f(1 \times \tau)(1 \times \Delta) \qquad e := \tau(\ulcorner g \urcorner \times \ulcorner g \urcorner)$$

Dependent Sum

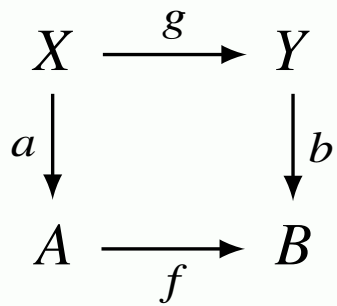
Theorem

For $f : A \to B$ a morphism in a category $\mathbf{C}$ with pullbacks, the pullback functor $f^* : \mathbf{C}/B \to \mathbf{C}/A :: y \mapsto f^* y$ has a left adjoint $\sum_f \dashv f^*$.

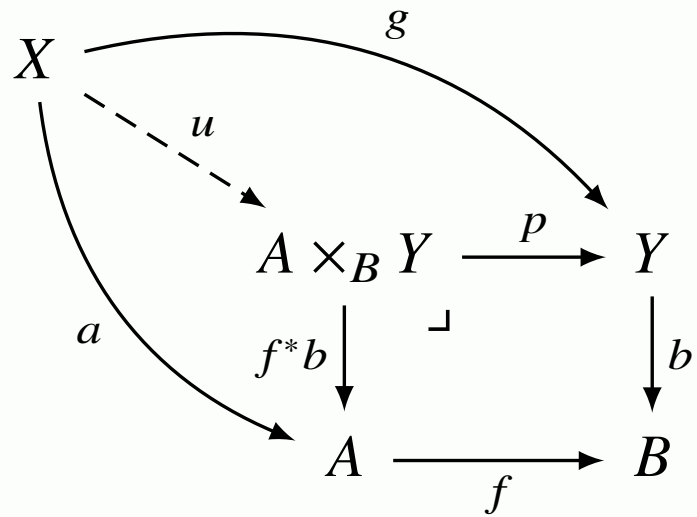


Proof Sketch.

Let $\sum_f : \mathbf{C}/A \to \mathbf{C}/B :: a \mapsto f \circ a$.
We have to check that there is a natural isomorphism:
$\theta : \mathrm{Hom}_{\mathbf{C}/B}(\sum_f a, b) \xrightarrow{\cong} \mathrm{Hom}_{\mathbf{C}/A}(a, f^* b)$.
Assume $g \in \mathrm{Hom}_{\mathbf{C}/B}(\sum_f a, b)$.



By definition of $f^* b$, the following diagram is a pullback.



So for any $g \in \mathrm{Hom}_{\mathbf{C}/B}(\sum_f a, b)$, there is a unique $u \in \mathrm{Hom}_{\mathbf{C}/A}(a, f^* b)$ s.t. $\theta : g \mapsto u$ is a bijection.

Locally Cartesian Closed Category

Definition (Locally Cartesian Closed Category)

A category $\mathbf{C}$ is called *locally cartesian closed* whenever, for all object $A \in \mathrm{ob}(\mathbf{C})$, the slice category $\mathbf{C}/A$ is cartesian closed.

Theorem

If $\mathbf{C}$ is locally cartesian closed and has a terminal object, then $\mathbf{C}$ is cartesian closed.

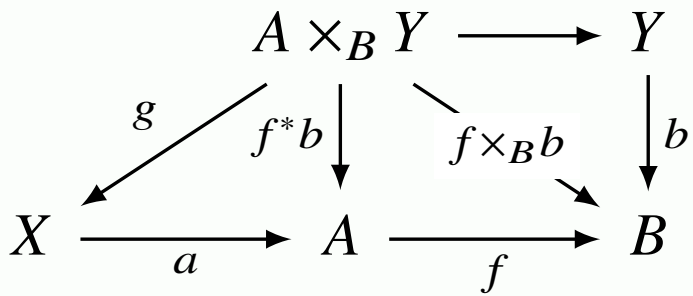
Dependent Product

Theorem

Let $\mathbf{C}$ be a category with all pullbacks. Then $\mathbf{C}$ is locally cartesian closed, iff, for any morphism $f : A \to B$, the pullback functor $f^* : \mathbf{C}/B \to \mathbf{C}/A$ has a right adjoint $\prod_f$.

proof sketch of " $\Longrightarrow$ ".

Let $f : A \to B$ in $\mathbf{C}$, and $X \xrightarrow{a} A$ in $\mathbf{C}/A$. The pullback of $Y \xrightarrow{b} B$ along f corresponds to $f \times_B b : A \times_B Y \to B$ in $\mathbf{C}/B$.

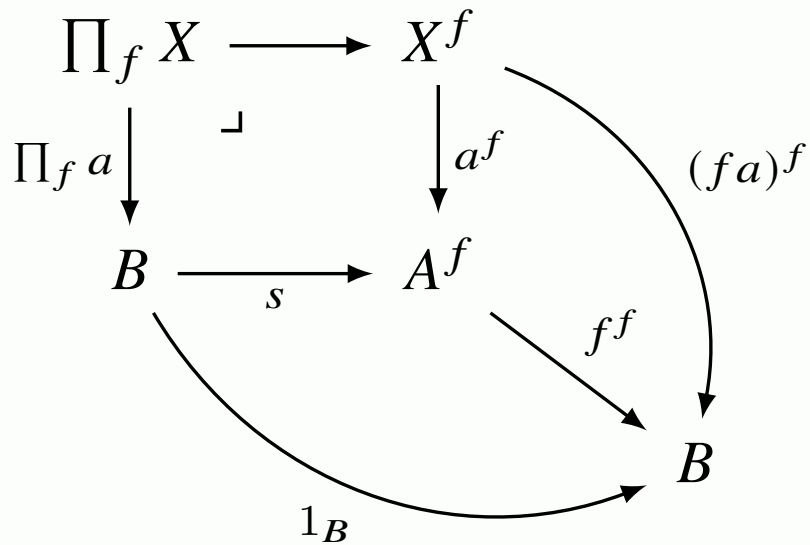


Since $\mathbf{C}/B$ is cartesian closed, we may exponentiate $fa : X \to B$ by $f : A \to B$ to obtain a morphism $(fa)^f : X^f \to B$ such that

proof sketch of " $\Longrightarrow$ " continued.

Since $(-)^f : \mathbf{C}/B \to \mathbf{C}/B$ is a functor and we have a morphism $a : fa \to f$ in $\mathbf{C}/B$, so we obtain a morphism $a^f : (fa)^f \to f^f$ in $\mathbf{C}/B$. Moreover, $1_B \times_B f \cong f$, so by the product-exponential adjunction

$$\mathrm{Hom}_{\mathbf{C}/B}(f, f) \cong \mathrm{Hom}_{\mathbf{C}/B}(1_B, f^f)$$

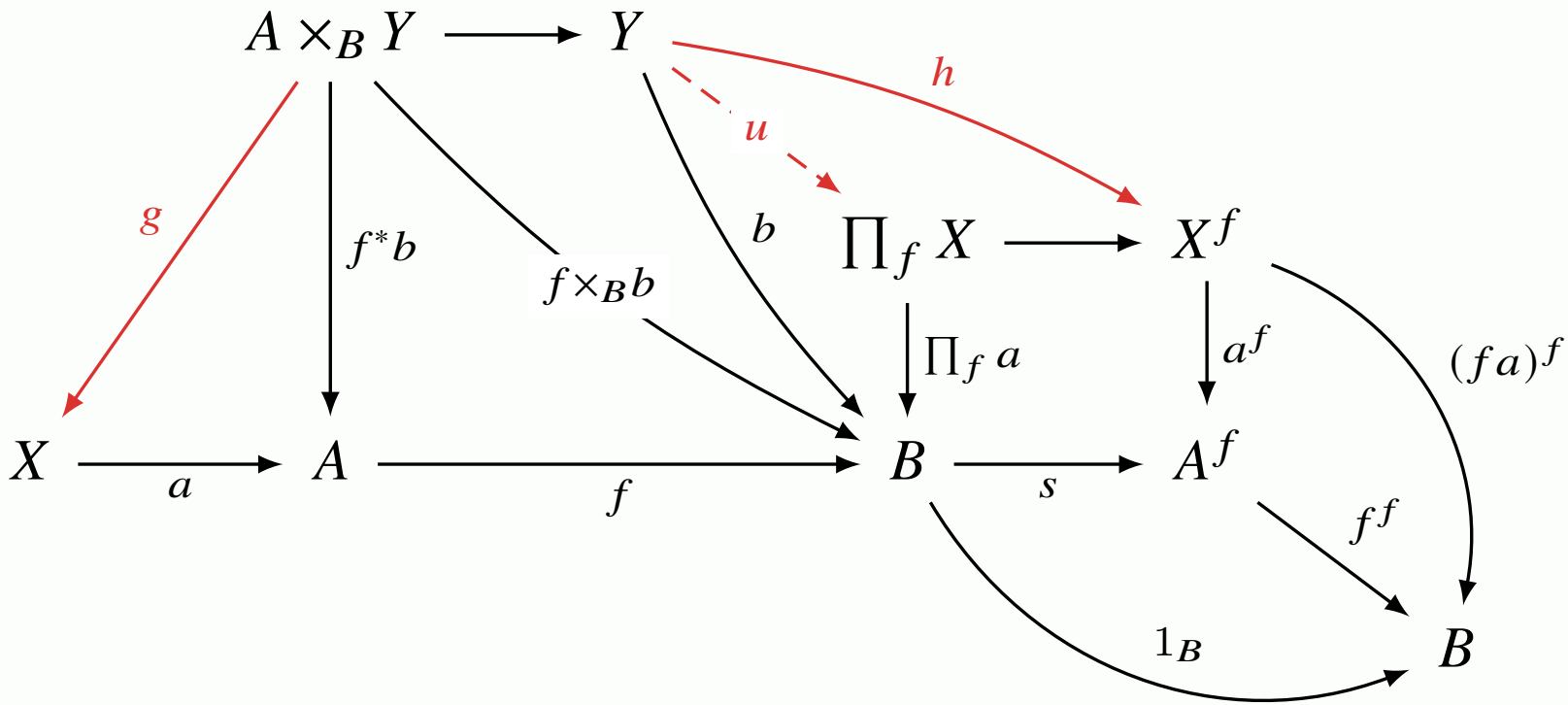


In particular, $1_f : f \to f$ is mapped to some $s : 1_B \to f^f$ (i.e. a morphism $s : B \to A^f$ in $\mathbf{C}$ such that $f^f \circ s = 1_B$). Now, take the pullback (in $\mathbf{C}$) of a^f along s to obtain $\prod_f a := s^* a^f : \prod_f X \to B$. □

proof sketch of " $\Longrightarrow$ " continued.

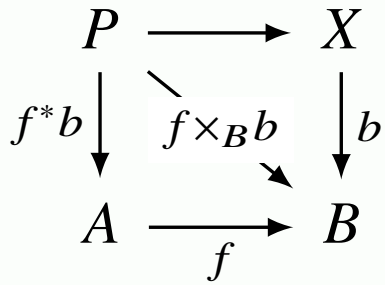
Now we prove there is a bijection $\mathrm{Hom}_{\mathbf{C}/A}(f^*b, a) \cong \mathrm{Hom}_{\mathbf{C}/B}(b, \prod_f a)$, which is natural in $X \xrightarrow{a} A$ and $Y \xrightarrow{b} B$.

The left hand side can be identified with morphisms $g : f \times_B b \to fa$ in $\mathbf{C}/B$ satisfying $a \circ g = f^*b$, where $f^*b : f \times_B b \to f$ is the projection. The right hand side can be identified with morphisms $h : b \to (fa)^f$ such that $a^f \circ h = s \circ b$. Then $\mathrm{Hom}_{\mathbf{C}/B}(f \times_B b, fa) \cong \mathrm{Hom}_{\mathbf{C}/B}(b, (fa)^f)$ restricts to a bijection $g \mapsto u$.



proof sketch of " $\Longleftarrow$ ".

The terminal object in $\mathbf{C}/B$ is 1_B.
The pullback of f and b in $\mathbf{C}$ corresponds to a product in $\mathbf{C}/B$.



$$f \times_B b = f \circ f^* b = \sum\nolimits_f f^* b$$

We deduce the following equivalence

$$\begin{aligned} \mathrm{Hom}_{\mathbf{C}/B}(f \times_B b, u) &= \mathrm{Hom}_{\mathbf{C}/B}(\sum\nolimits_f (f^* b), u) \\ &\cong \mathrm{Hom}_{\mathbf{C}/A}(f^* b, f^* u) \\ &\cong \mathrm{Hom}_{\mathbf{C}/B}(b, \prod\nolimits_f (f^*(u))) \end{aligned}$$

Then $u^f = \prod_f (f^*(u))$. □

Remark

$$\{Ax\}_{x\in X} \cong \mathrm{P}(X) \underset{\forall_f}{\overset{\exists_f}{\underset{\longleftarrow}{\overset{\longrightarrow}{f^*}}}} \mathrm{P}(\bullet) = \{\emptyset, \bullet\} \cong \{\bot, \top\}$$

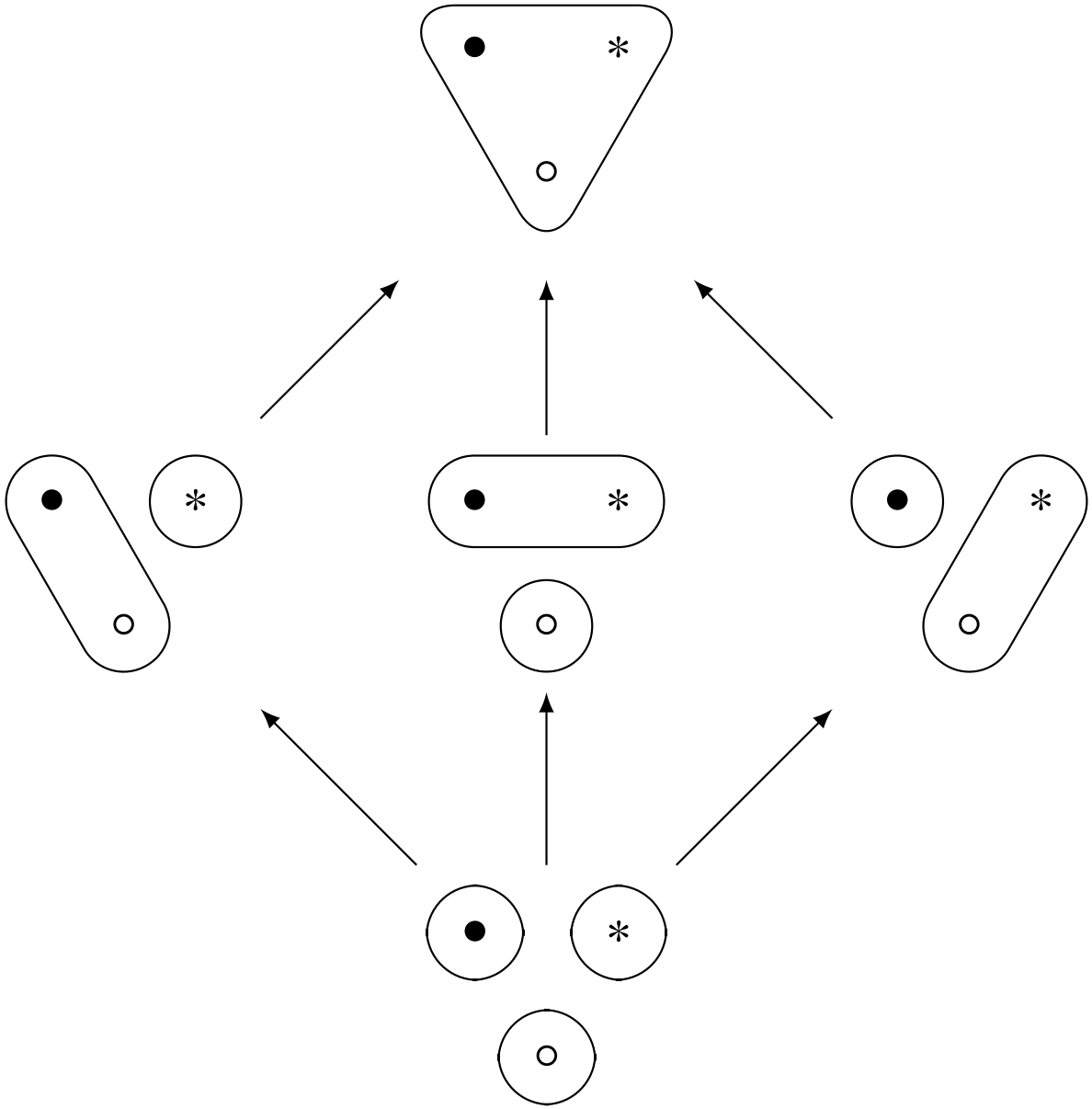
The set $\mathrm{P}(X) = \{A : A \subset X\}$ can be identified with the set of predicates Ax. The corresponding subset is $\{x \in X : Ax = \top\} = f^{-1}(\top)$.

$$\exists_f : A \mapsto \exists x A \qquad \forall_f : A \mapsto \forall x A$$

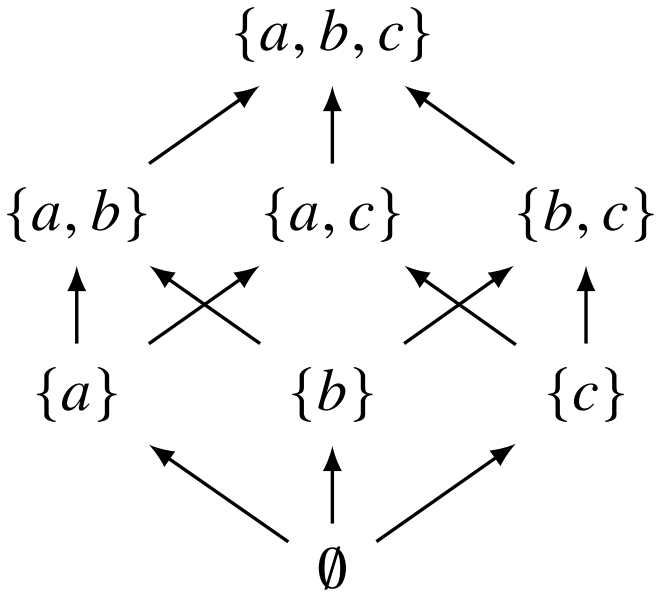
$$\mathbf{Set}/X \underset{\prod_f}{\overset{\sum_f}{\underset{\longleftarrow}{\overset{\longrightarrow}{f^*}}}} \mathbf{Set}/Y$$

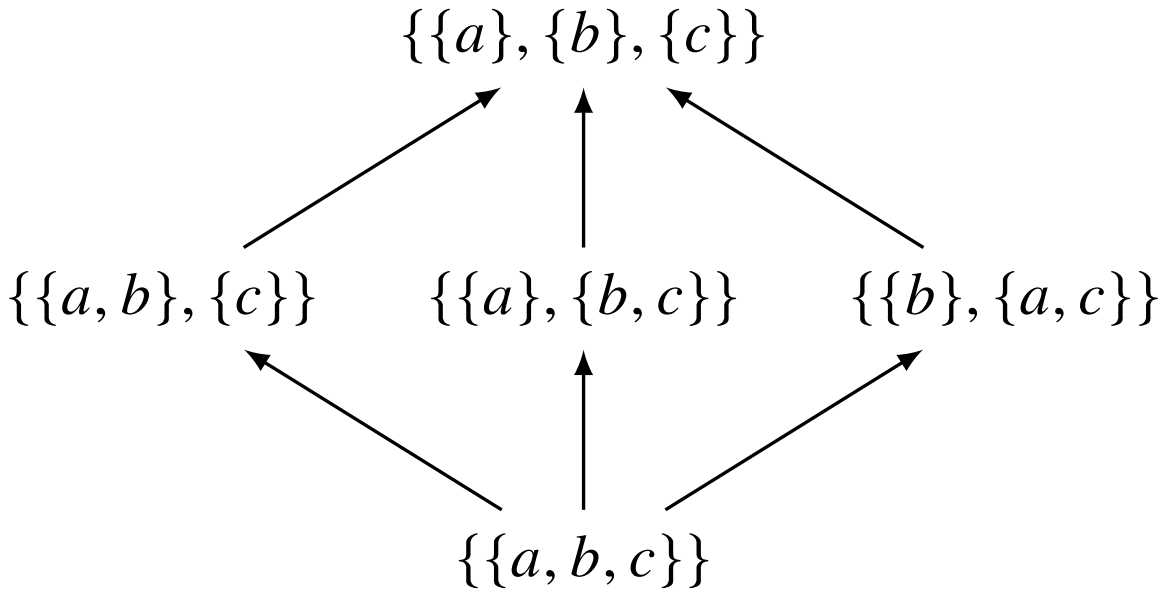
$$f^* : \{By\}_{y\in Y} \mapsto \{B(fx)\}_{x\in X}$$
$$\sum_f : \{Ax\}_{x\in X} \mapsto \left\{\sum_{x\in f^{-1}(y)} Ax\right\}_{y\in Y}$$
$$\prod_f : \{Ax\}_{x\in X} \mapsto \left\{\prod_{x\in f^{-1}(y)} Ax\right\}_{y\in Y}$$

The Poset of Partitions



Subset Logic vs Partition Logic





- Two Philosophies: Creating Elements vs Creating Distinctions
- Classical logic is closely connected to the logic of subsets. For a set X of "states" of the world, we get a poset $\mathrm{P}(X)$, with the partial order being $\subset$. Elements of $\mathrm{P}(X)$ are "propositions" about the world.
- In classical logic, propositions correspond to subsets of X. In partition logic, propositions correspond to partitions of X.
- In both approaches we get a poset of propositions where the partial order is "implication" $\rightarrow$.

Partition Logic

- A set X has a poset of partitions. Let $\mathcal{E}(X)$ be the set of partitions of X. Each partition P corresponds to an equivalence relation $\sim_P$. We say a partition P is finer than Q ($P \leq Q$) iff $\forall xy : x \sim_P y \implies x \sim_Q y$.
- The meet $P \wedge Q$ is the coarsest partition that is finer than P and Q.

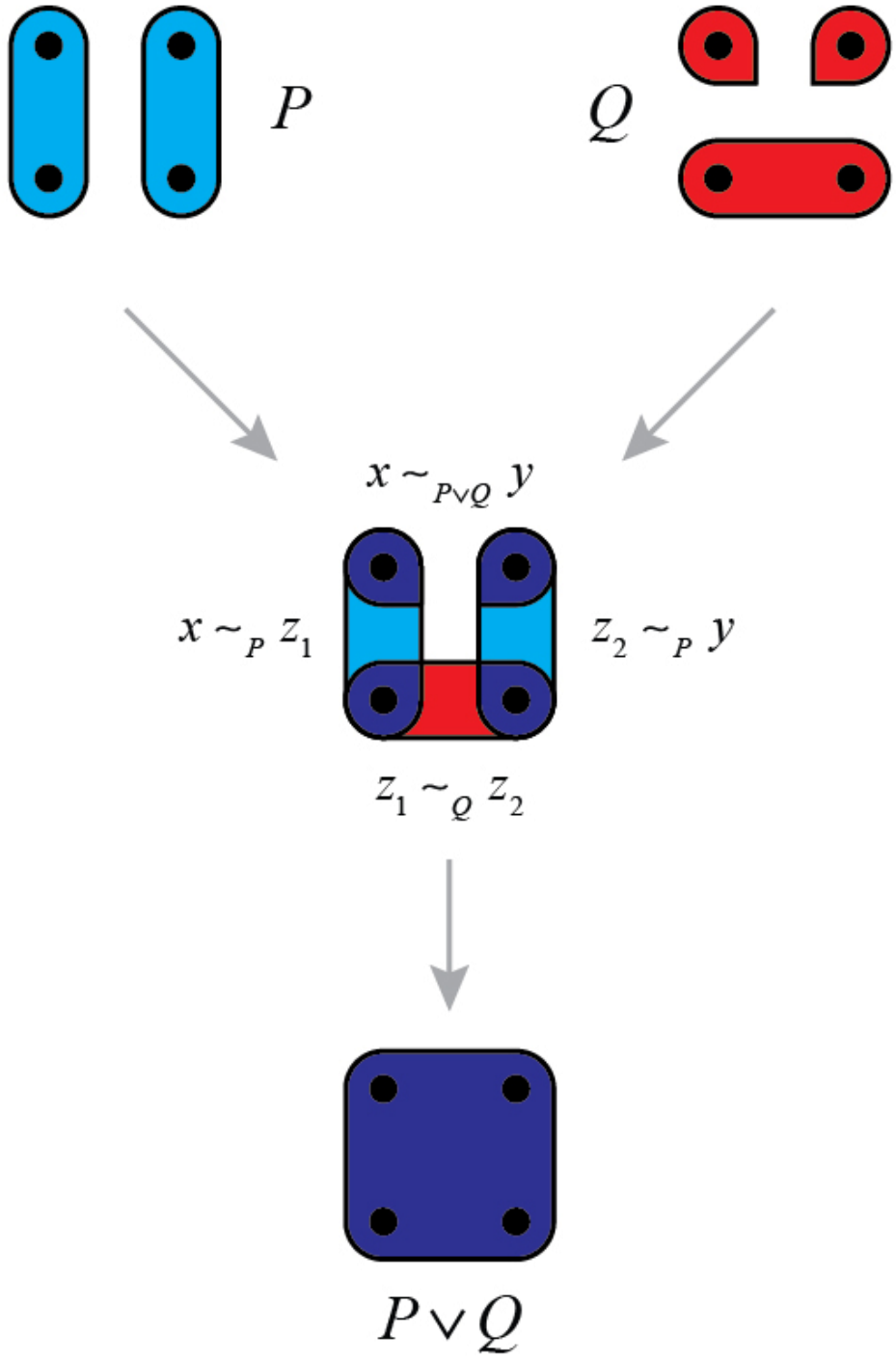
$$\sim_{P \wedge Q} := \sim_P \cap \sim_Q$$

- The join $P \vee Q$ is the finest partition that is coarser than P and Q.

$$\sim_{P \vee Q} := \left(\sim_P \cup \sim_Q\right)^*$$

  where $()^*$ is the transitive closure operator.
- Given a function $f : X \to Y$ and a partition P of Y, the pullback of P along f is the partition of X: $f^*(P) := \{f^*(S) : S \in P \;\&\; f^*(S) \neq \emptyset\}$.
- We always have $f^*(P \wedge Q) = f^*(P) \wedge f^*(Q)$.
- But sometimes we have $f^*(P \vee Q) \neq f^*(P) \vee f^*(Q)$.

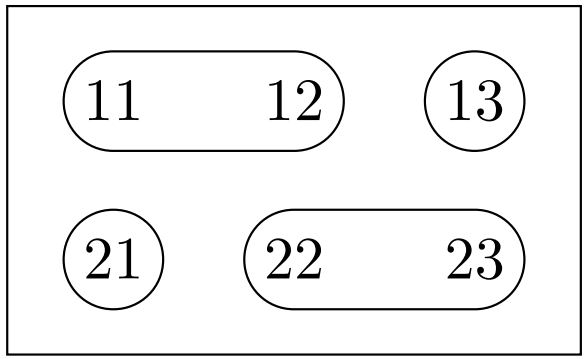
Remark: $f(a \vee b) \not\equiv f(a) \vee f(b)$ implies that we see something when we observe the combined system that we could not expect by merely combining our observations of the subsystems.


P
Q
$x \sim_{P \vee Q} y$
$x \sim_P z_1$
$z_2 \sim_P y$
$z_1 \sim_Q z_2$
$P \vee Q$

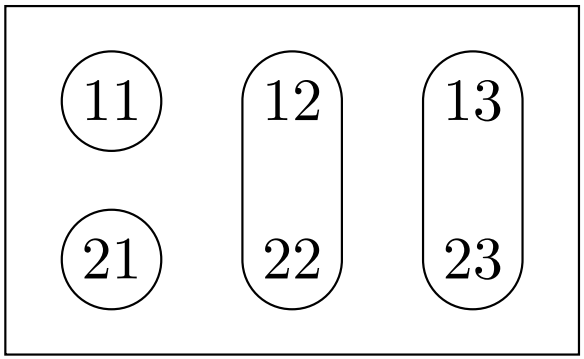
Proof of $f^*(P \vee Q) \neq f^*(P) \vee f^*(Q)$

Take $X = \{11, 22\}$, $Y = \{11, 12, 13, 21, 22, 23\}$, and let $i : X \hookrightarrow Y$.

P =



Q =



Then

$$i^*(P) = \{\{11\}, \{22\}\} = i^*(Q)$$

$$i^*(P) \vee i^*(Q) = \{\{11\}, \{22\}\}$$

$$P \vee Q = \{\{11, 12, 13, 22, 23\}, \{21\}\}$$

$$i^*(P \vee Q) = \{\{11, 22\}\}$$

$$i^*(P \vee Q) \neq i^*(P) \vee i^*(Q)$$

Galois Correspondence

Definition (Galois Correspondence)

The function $X \mapsto X^*$ from $\mathtt{P}(A)$ to $\mathtt{P}(B)$ and the function $Y \mapsto Y^\dagger$ from $\mathtt{P}(B)$ to $\mathtt{P}(A)$ constitute a *Golois correspondence* iff

1. $X_1 \subset X_2 \implies X_2^* \subset X_1^*$
2. $Y_1 \subset Y_2 \implies Y_2^\dagger \subset Y_1^\dagger$
3. $X \subset (X^*)^\dagger$
4. $Y \subset (Y^\dagger)^*$

Definition (Polarity)

Given $R \subset A \times B$, $X \subset A$, $Y \subset B$. Let

$$X^* := \bigcap_{x \in X} \{y \in B : Rxy\} \qquad Y^\dagger := \bigcap_{y \in Y} \{x \in A : Rxy\}$$

We refer to the functions $X \mapsto X^*$ and $Y \mapsto Y^\dagger$ as *polarities*.

- The polarities induced by a relation constitute a Galois correspondence.
- Every Galois correspondence arises from ploarities induced by a relation.

Galois Connection

Definition (Galois Connection)

- Let $(A, \leq_A)$ and $(B, \leq_B)$ be two partially ordered sets. A monotone Galois connection between these posets consists of two monotone functions: $f : A \to B, g : B \to A$ s.t.

$$\forall a \in A \forall b \in B : f(a) \leq_B b \iff a \leq_A g(b)$$

- An antitone Galois connection between these posets consists of two order-reversing functions: $f : A \to B, g : B \to A$ s.t.

$$\forall a \in A \forall b \in B : b \leq_B f(a) \iff a \leq_A g(b)$$

Example: $B \to \neg A \iff A \to \neg B, \quad f = g = \neg$

- A Galois correspondence is an antitone Galois connection.
- An antitone Galois connection between $\mathbf{C}$ and $\mathbf{D}$ is just a monotone Galois connection between $\mathbf{C}$ and the order dual $\mathbf{D}^{\mathrm{op}}$.
- A Galois connection is a pair of adjoint functors between two categories that arise from partially ordered sets.

- Suppose we have two preorders $(A, \leq_A)$ and $(B, \leq_B)$.
  If f has a right adjoint $g : B \to A$, then g is unique and

$$g(b) = \bigvee \{a \in A : f(a) \leq_B b\}$$

  If $g : B \to A$ has a left adjoint $f : A \to B$, then f is unique and

$$f(a) = \bigwedge \{b \in B : a \leq_A g(b)\}$$

- The function $g : B \to A$ is the inverse of $f : A \to B$ iff

$$\forall a \in A \forall b \in B : f(a) = b \iff a = g(b)$$

- The right adjoint g is the "best approximation from below" to the "nonexistent" inverse of f.
- The left adjoint f is the "best approximation from above" to the "nonexistent" inverse of g.

Fundamental Theorem of Galois Theory

Theorem (Fundamental Theorem of Galois Theory)

Let $K \to L$ be a finite separable normal field extension with Galois group $G := \mathrm{Aut}(L/K)$. For any subfiled F of L containing K, any subgroup $H < G$, let

$$F^* := \mathrm{Aut}(L/F) := \left\{\sigma \in \mathrm{Aut}(L) : \forall x \in F\big(\sigma(x) = x\big)\right\}$$

$$H^\dagger := \left\{x \in L : \forall \sigma \in H\big(\sigma(x) = x\big)\right\}$$

Then

1. $[L : K] = |G|$, *where* $[L : K]$ *is the dimension of* L *as a vector space over* K.
2. $F = (F^*)^\dagger$, $H = (H^\dagger)^*$, $[L : F] = |F^*|$, $[F : K] = |G|/|F^*|$.
3. F *is a normal extension of* K *iff* $F^* \lhd G$.
4. $F^* \lhd G \implies \mathrm{Aut}(F/K) \cong G/F^*$.

The Existence of a Left/Right Adjoint

Theorem (Adjoint Functor Theorem for Posets)

Let $(X, \leq)$ and $(Y, \leq)$ be posets. Suppose that Y has all meets, and let $g : Y \to X$ be a monotone function preserving all meets. Then g has a left adjoint $f : X \to Y$, given by

$$f(x) := \bigwedge \{y \in Y : x \leq g(y)\}$$

In particular, a monotone map $g : Y \to X$ is the right adjoint of a Galois connection iff it preserves all meets.

Corollary

Suppose that X has all joins. A monotone map $f : X \to Y$ has a right adjoint iff it preserves all joins.

Proof.

$$g(f(x)) = g\left(\bigwedge \{y \in Y : x \leq g(y)\}\right) = \bigwedge \{g(y) : y \in Y \;\&\; x \leq g(y)\}$$

$$x \leq \bigwedge \{g(y) : y \in Y \;\&\; x \leq g(y)\} = g(f(x))$$

This inequality is the unit of the adjunction.
For the counit, let $z \in Y$. Then

$$f(g(z)) = \bigwedge \{y \in Y : g(z) \leq g(y)\}$$

Since $g(z) \leq g(z)$, we have

$$z \geq \bigwedge \{y \in Y : g(z) \leq g(y)\} = f(g(z))$$

□

The Existence of a Left/Right Adjoint

Lemma

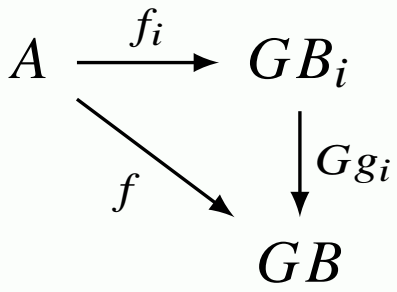
A functor $G : \mathbf{D} \to \mathbf{C}$ has a left adjoint iff for every $A \in \mathbf{C}$ the comma category $A \downarrow G$ has an initial object.

Lemma

*Suppose $\mathbf{D}$ is locally small and complete. Then $\mathbf{D}$ has an initial object iff $\mathbf{D}$ has a **weakly initial set:** there is a set of objects $(B_i)_{i \in I}$ in $\mathbf{D}$ s.t. for any $B \in \mathbf{D}$ there exists some $i \in I$ and a morphism $g_i : B_i \to B$.*

Theorem (General Adjoint Functor Theorem)

*Suppose $\mathbf{D}$ is locally small and complete. Then $G : \mathbf{D} \to \mathbf{C}$ has a left adjoint iff G is continuous and for each $A \in \mathbf{C}$, the comma category $A \downarrow G$ has a **weakly initial set:** there is a set of objects $(B_i, f_i : A \to GB_i)_{i \in I}$ in $A \downarrow G$ s.t. for any $(B, f : A \to GB)$ there exists some $i \in I$ and $g_i : B_i \to B$ with $f = Gg_i \circ f_i$.*



The Existence of a Left/Right Adjoint

Definition (Coseparating Family)

A *coseparating family* for a category $\mathbf{C}$ is a family of objects $(G_i)_{i \in I}$ such that for any pair $A \overset{f}{\underset{g}{\rightrightarrows}} B$ with $f \neq g$, there is an $i \in I$ and an $h : B \to G_i$ such that $hf \neq hg$.

Theorem (Special Adjoint Functor Theorem)

Suppose both $\mathbf{C}$ and $\mathbf{D}$ are locally small, and that $\mathbf{D}$ is complete and well-powered and has a coseparating set. Then a functor $G : \mathbf{D} \to \mathbf{C}$ has a left adjoint iff G preserves small limits.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

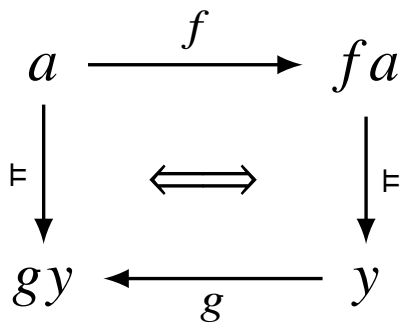
Answers to the Exercises

References2039

Formal Concept Analysis / Chu Space

- A formal context (or *Chu space*) is a triple $(X, \vDash, Y)$, where X is a set of objects, Y is a set of attributes, and $\vDash$: $X \times Y \to 2$ is a relation.
- A morphism from $(A, \vDash, X)$ to $(B, \vDash, Y)$ is a pair of functions (f, g) with $f : A \to B$ and $g : Y \to X$ s.t.

$$\forall a \in A \forall y \in Y : a \vDash gy \iff fa \vDash y$$



Formal Concept Analysis

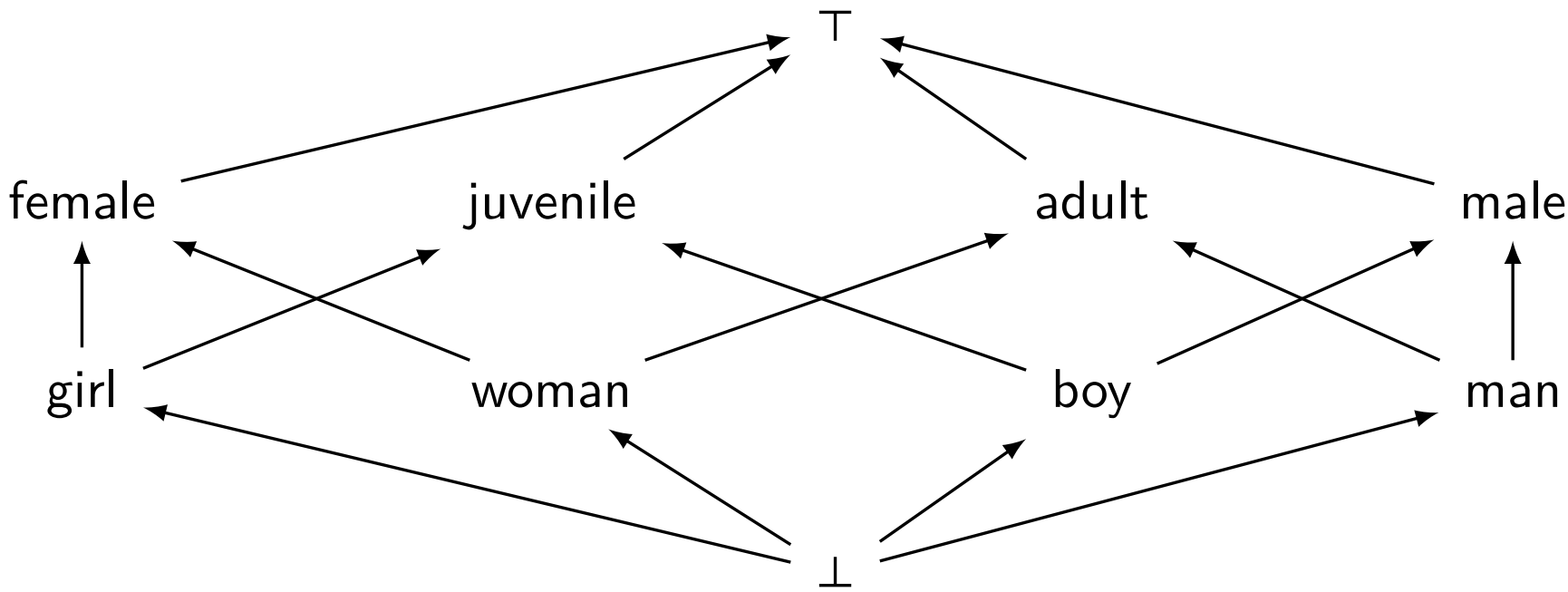
$$\frac{\vDash : X \times Y \to 2}{\frac{\hat{\vDash} : X \to 2^Y}{\check{\vDash} : Y \to 2^X}}$$

- The formal context is said to be *separable* if $\hat{\vDash} : X \to 2^Y$ is injective (no repeated rows) and is said to be *extensional* if $\check{\vDash} : Y \to 2^X$ is injective (no repeated columns).
- For $A \subset X$, let $A^\uparrow := \{y \in Y : \forall x \in A : x \vDash y\}$, i.e., a set of attributes shared by all objects from A.
- For $B \subset Y$, let $B^\downarrow := \{x \in X : \forall y \in B : x \vDash y\}$, i.e., a set of objects sharing all attributes from B.
- Then we have the Galois connection.

$$A^\uparrow \supset B \iff A \subset B^\downarrow$$

Remark: both sides mean "every object in A has every attribute in B".

$\models$	male	female	juvenile	adult
boy	1		1	
girl		1	1	
man	1			1
woman		1		1



- A pair (A, B) is a formal concept of a context $(X, \vDash, Y)$ iff

$$A^{\uparrow} = B \;\; \& \;\; B^{\downarrow} = A$$

The novel idea of FCA is the clustering of attributes based on Galois connection. The clustering determines which collection of attributes forms a coherent entity called a concept, by the philosophical criteria of unity between extension and intension.

$$\text{concept} = \text{extent } A + \text{intent } B$$

- The set of formal concepts of $(X, \vDash, Y)$ is

$$\mathbf{B}(X, \vDash, Y) := \left\{(A, B) \in 2^X \times 2^Y : A^{\uparrow} = B \;\; \& \;\; B^{\downarrow} = A\right\}$$

- formal concept = fixpoint of $\downarrow\uparrow$, where $\mathrm{Fix}(f) = \{x : fx = x\}$.

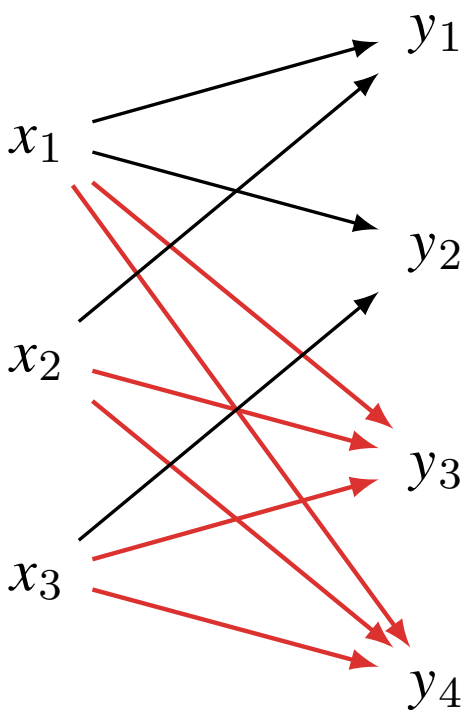
$$\mathbf{B}(X, \vDash, Y) \cong \mathrm{Fix}(\downarrow\uparrow) \cong \mathrm{Fix}(\uparrow\downarrow)$$

- The conceptualization of an object $x \in X$ is $(\{x\}^{\uparrow\downarrow}, \{x\}^{\uparrow})$; The conceptualization of an attribute $y \in Y$ is $(\{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow})$.

$\vDash$	y_1	y_2	y_3	y_4
x_1	1	1	1	1
x_2	1		1	1
x_3		1	1	1

$\vDash$	y_1	y_2	y_3	y_4
x_1	1	1	1	1
x_2	1		1	1
x_3		1	1	1

$\vDash$	y_1	y_2	y_3	y_4
x_1	1	1	1	1
x_2	1		1	1
x_3		1	1	1



$$(A_1, B_1) = (\{x_1, x_2, x_3\}, \{y_3, y_4\})$$
$$(A_2, B_2) = (\{x_1, x_3\}, \{y_2, y_3, y_4\})$$
$$(A_3, B_3) = (\{x_1, x_2\}, \{y_1, y_3, y_4\})$$

The set of Formal Concepts is a Complete Lattice

- Subconcept $(A, B) \leq (A', B')$ if $A \subset A'$, or equivalently if $B \supset B'$.
- The set of concepts $\mathbf{B}(X, \vDash, Y)$ is a complete lattice.

$$(A, B) \wedge (A', B') = (A \cap A', (B \cup B')^{\downarrow\uparrow})$$

$$(A, B) \vee (A', B') = ((A \cup A')^{\uparrow\downarrow}, B \cap B')$$

$$\bigwedge_{i \in I} (A_i, B_i) = \left(\bigcap_{i \in I} A_i, \quad \left(\bigcup_{i \in I} B_i \right)^{\downarrow\uparrow} \right)$$

$$\bigvee_{i \in I} (A_i, B_i) = \left(\left(\bigcup_{i \in I} A_i \right)^{\uparrow\downarrow}, \quad \bigcap_{i \in I} B_i \right)$$

- Attribute Implication $B \rightarrow B' := B^{\downarrow} \subset B'^{\downarrow}$ for $B, B' \subset Y$.
- Armstrong rules for attribute implication.

$$\frac{}{A \rightarrow A} \qquad \frac{A \rightarrow B}{A \cup C \rightarrow B} \qquad \frac{A \rightarrow B, \quad B \cup C \rightarrow D}{A \cup C \rightarrow D}$$

Chu Spaces as Generalized Topological Spaces

- A topological space $(X, \mathcal{O}_X)$ where X is the set of points and $\mathcal{O}_X$ the set of open sets, can be understood as a Chu space $(X, \in, \mathcal{O}_X)$, such that $\mathcal{O}_X$ is extensional and closed under arbitrary union and finite intersection.
- The morphism (f, g) gives exactly continuous functions between the topological spaces $(X, \in, \mathcal{O}_X)$ and $(Y, \in, \mathcal{O}_Y)$.
- The adjointness condition makes g the inverse image function f^{-1}, while the choice of X for the codomain of g corresponds to the requirement for continuous functions that the inverse image of open sets be open.

$$x \in gy \iff fx \in y \quad \text{iff} \quad gy = \{x : fx \in y\} = f^{-1}y$$

$$f^{-1}(B) \in \mathcal{O}_X \text{ for } B \in \mathcal{O}_Y$$

- A function $f : A \to B$ is continuous if there exists a function $g : Y \to X$ s.t. $(A, \vDash, X) \xrightarrow{(f,g)} (B, \vDash, Y)$ is a Chu transform.
- **Top** is a full subcategory of the category of Chu spaces.
- A topological space is T_0 if $(X, \in, \mathcal{O}_X)$ is separable.

Chu Spaces — Examples

$$(X, \vDash, Y)$$

	X	Y
Formal Concept	objects	attributes
Logic	models	formulas
Information System	messages	contents
Game	strategies	counterstrategies
Topology	points	open sets

Dualities via Chu Space

- Algebraic geometry:
  {algebraic sets in $\mathbb{C}^n$} $\cong$ {radical ideals in $\mathbb{C}[x_1, \ldots, x_n]$}$^{\mathrm{op}}$
  $$X = \mathbb{C}^n, \quad Y = \mathbb{C}[x_1, \ldots, x_n], \quad xRp \iff p(x) = 0$$
- Number theory:
  {intermediate extensions $K \subset J \subset L$} $\cong$ {subgroups of $\mathrm{Gal}(L, K)$}$^{\mathrm{op}}$
  $$X = L, \quad Y = \mathrm{Aut}(L, K), \quad xR\varphi \iff \varphi(x) = x$$
- Convex geometry:
  {closed convex sets in $\mathbb{R}^n$} $\cong$ {'closed' sets of half spaces in $\mathbb{R}^n$}$^{\mathrm{op}}$
  $$X = \mathbb{R}^n, \quad Y = \{\text{half spaces in } \mathbb{R}^n\}, \quad xRH \iff x \in H$$
- Analysis:
  {upper closed subsets of $\mathbb{Q}$} $\cong$ {lower closed subsets of $\mathbb{Q}$}$^{\mathrm{op}} \cong [-\infty, +\infty]$
  $$X = \mathbb{Q}, \quad Y = \mathbb{Q}, \quad pRq \iff p \leq q$$
- Logic: $X = \{\mathscr{L}\text{-structures}\}$, $Y = \{\mathscr{L}\text{-sentences}\}$, $\mathcal{M}RA \iff \mathcal{M} \vDash A$
- Linear algebra: $X = V$, $Y = V^*$, $vRf \iff f(v) = 0$

The Formal Concepts

$$\mathbf{B}(X, R, Y) \cong \mathrm{Fix}(R_* R^*) \cong \mathrm{Fix}(R^* R_*)$$

- Algebraic geometry: affine varieties, and radical ideals.
- Number theory: Galois correspondence.
- Convex geometry: the closed convex sets, and the closure of its convex hull.
- Analysis: Dedekind cuts.
- Logic: The theories.
- Linear algebra: the linear subsets, and the annihilator.

The Category of Chu Spaces

Definition (The Category of Chu Spaces)

Let $\mathbf{C}$ be a symmetric monoidal closed category and $K \in \mathbf{C}$ an object. The objects of $\mathbf{Chu}_K(\mathbf{C})$ are Chu spaces $(A, \vDash, X)$ over K, where A and X are objects of $\mathbf{C}$ and $\vDash$ is a morphism $A \otimes X \to K$ (which can be taken as $A \times X$ matrix with entries from K).
The morphisms $(A, \vDash, X) \to (B, \vDash, Y)$ are pairs of morphisms $f : A \to B$ and $g : Y \to X$ s.t.

$$\begin{array}{ccc} A \otimes Y & \xrightarrow{1_A \otimes g} & A \otimes X \\ {\scriptstyle f \otimes 1_Y}\downarrow & \circlearrowright & \downarrow{\scriptstyle \vDash} \\ B \otimes Y & \xrightarrow[\vDash]{} & K \end{array}$$

- Symmetry of $\otimes$ makes $\mathbf{Chu}_K(C)$ self-dual. $\mathbf{Chu}_K(\mathbf{C}) \to \mathbf{Chu}_K(\mathbf{C})^{\mathrm{op}}$.
- On objects, it takes $(A, \vDash\colon A \otimes X \to K, X)$ to $(X, \vDash^{\circ}\colon X \otimes A \to K, A)$.
- On morphisms, it takes $(f, g) : (A, \vDash, X) \to (B, \vDash, Y)$ to $(g, f) : (Y, \vDash^{\circ}, B) \to (X, \vDash^{\circ}, A)$.

Remarks

$$(A, \vDash, X) \xrightarrow{(f,g)} (B, \vDash, Y)$$

$$\begin{array}{ccc} a & \xrightarrow{f} & fa \\ \vDash & \Longleftrightarrow & \vDash \\ gy & \xleftarrow[g]{} & y \end{array}$$

- Chu spaces do not take subjects to be primitive and predicates to be derived, but rather take both to be primitive.
- If we view the open sets of a topological space as its permitted predicates defining its structure, a topological space is an example of an object structured by the interaction of its subjects and predicates.

Example: Agent vs Environment

- Let W be a set of possible worlds. A chu space $(A, \cdot, X)$ over W is a way of factoring the space of possible world histories into an agent and an environment.
  — We can think of A as possible states the agent can choose to be.
  — We can think of the environment X as representing the agent's uncertainty about the set of counterfactuals, "what the world is as a function of my behavior."
- We can think of the morphism (f, g) from $(A, \cdot, X)$ to $(B, *, Y)$ as a way of fitting the agent of A into the environment of Y.
  — We can construct $(A, \star, Y)$, with $a \star y := a \cdot g(y) = f(a) * y$.
- Given two sets W, V and a function $p : W \to V$, let $p^\circ : \mathbf{Chu}_W \to \mathbf{Chu}_V$ denote the functor that sends the object $(A, \cdot, X) \in \mathbf{Chu}_W$ to $(B, *, Y) \in \mathbf{Chu}_V$, where $b * y = p(a \cdot x)$, and $p^\circ(f, g) = (f, g)$.
  — We say that V is a coarse version of W if $p : W \to V$ is surjective.

Theorem

Every small category $\mathbf{C}$ *embeds fully in* $\mathbf{Chu}_{\mathrm{mor}(\mathbf{C})}(\mathbf{Set})$*, where* $\mathrm{mor}(\mathbf{C})$ *is the set of morphisms of* $\mathbf{C}$*.*

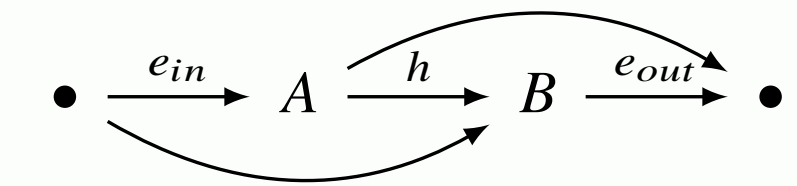
Proof.

Define the functor $\mathsf{y} : \mathbf{C} \to \mathbf{Chu}_{\mathrm{mor}(\mathbf{C})}(\mathbf{Set})$ as follows.

$$\mathsf{y} : A \mapsto (I_A, \vDash, O_A)$$

where $I_A := \left\{ \bullet \xrightarrow{e_{in}} A : \bullet \in \mathrm{ob}(\mathbf{C}) \right\}, O_A := \left\{ A \xrightarrow{e_{out}} \bullet : \bullet \in \mathrm{ob}(\mathbf{C}) \right\}$, and
$\vDash (e_{in}, e_{out}) := e_{out} \circ e_{in}$.
For morphisms,

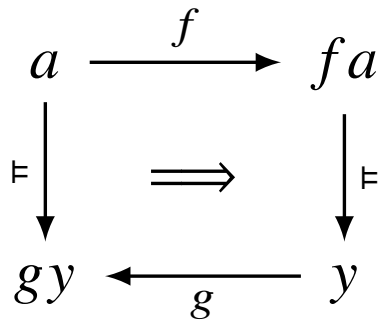
$$\mathsf{y} : A \xrightarrow{h} B \mapsto (I_A, \vDash, O_A) \xrightarrow{(f,g)} (I_B, \vDash, O_B)$$



where $f : I_A \to I_B :: e_{in} \mapsto h \circ e_{in}$, and
$g : O_B \to O_A :: e_{out} \mapsto e_{out} \circ h$. □

Remark: This is an analog of the Yoneda embedding for Chu spaces.

Dialectica Category & Kolmogorov Problem

- The Dialectica category $\mathbf{Dial}_2(\mathbf{Sets})$ has as object triples $(A, \vDash, X)$, where A and X are sets and $\vDash$: $A \times X \to 2$ is a relation.
  A morphism from $(A, \vDash, X)$ to $(B, \vDash, Y)$ is a pair of functions (f, g) with $f : A \to B$ and $g : Y \to X$ s.t.,
  $$\forall a \in A \forall y \in Y : a \vDash gy \implies fa \vDash y$$
  

- An object of $\mathbf{Dial}_2(\mathbf{Sets})$ is a Kolmogorov problem, where X is the set of instances of the problem, A is the set of possible solutions for X, and $\vDash$ is the problem condition, so $a \vDash x$ means "a $\vDash$-solves x".
- A morphism $(A, \vDash, X) \to (B, \vDash, Y)$ is a reduction of $(B, \vDash, Y)$ to $(A, \vDash, X)$.
  $$\forall a \in A \forall y \in Y : a \vDash gy \implies fa \vDash y$$
  for all instances of problems y of Y and all solutions a of problems gy of X, if a $\vDash$-solves gy then fa $\vDash$-solves y.

Kolmogorov Problem Example: Analytical geometry

- L denotes the family of all lines in the plane π.
- E is the family of all equations of the form $ax + by = c$.
- $f : \mathbb{R}^2 \to \pi$ is a coordinate system for a plane π.
- $g : L \to E$ is the canonical equation which represents the line.
- Problem A: To decide whether a given pair of real numbers (u, v) satisfies a given equation is the problem $(\mathbb{R}^2, \vDash, E)$.
- Problem B: To decide whether a given point lies on a given line l is the problem $(\pi, \vDash, L)$.
- A slight variation of what we have just done reduces the problem of finding the intersection point of two distinct lines to the problem of solving a linear system with two equations over two variables $(\mathbb{R}^2, \vDash, [E]^2) \to (\pi, \vDash, [L]^2)$.

Gödel's Functional (Dialectica) Interpretation

- A translation from intuitionistic Heyting arithmetic HA to a finite type extension of primitive recursive arithmetic.

$$A \rightsquigarrow \exists x \forall y A_D(x, y)$$

- Via the negative translation, the consistency of PA is reduced to the consistency of HA.
- The dialectica interpretation gives a relative consistency proof for PA.

$$\begin{aligned}
A_D &:= A \text{ for atomic } A \\
(A \wedge B)_D &:= \exists x u \forall y v [A_D(x, y) \wedge B_D(u, v)] \\
(A \vee B)_D &:= \exists z x u \forall y v [(z = 0 \to A_D(x, y)) \wedge (z \neq 0 \to B_D(u, v))] \\
(A \to B)_D &:= \exists f g \forall x y [A_D(x, gxy) \to B_D(fx, y)] \\
(\exists z A(z))_D &:= \exists z x \forall y A_D(x, y, z) \\
(\forall z A(z))_D &:= \exists f \forall z y A_D(fz, y, z)
\end{aligned}$$

- Whenever A is provable, the proof of A can be converted into a closed term t and a proof of $\forall y A_D(t, y)$.

$\mathbf{Dial}_2(\mathbf{Set})$ is a $*$-autonomous Category

- The tensor product of $A = (A, R, X)$ and $B = (B, S, Y)$:

$$A \otimes B \coloneqq \left(A \times B, \otimes, Y^A \times X^B\right)$$

where

$$(a, b) \otimes (f, g) \coloneqq R(a, gb) \wedge S(b, fa)$$

- The internal-hom of $A = (A, R, X)$ and $B = (B, S, Y)$:

$$A \multimap B \coloneqq \left(B^A \times X^Y, \multimap, A \times Y\right)$$

where

$$(f, g) \multimap (a, y) \coloneqq R(a, gy) \to S(fa, y)$$

- The linear negation of $A = (A, R, X)$:

$$A^* \coloneqq (X, R^*, A) \quad \text{where} \quad R^*(x, a) \coloneqq \neg R(a, x)$$

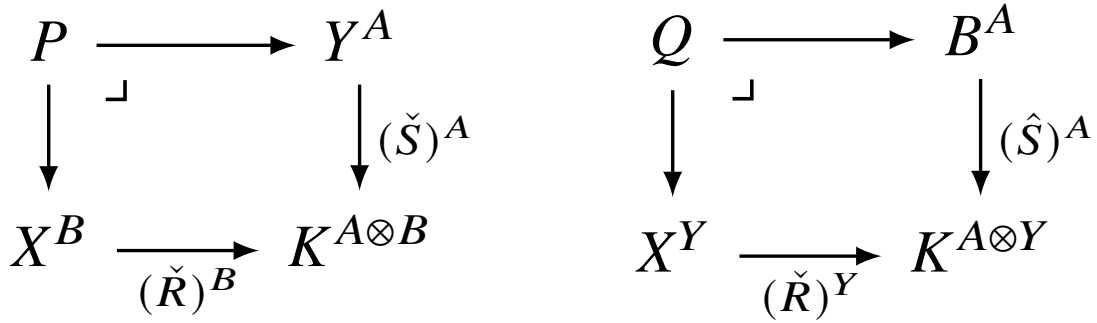
- The unit and the dualizing object:

$$I = (1, =, 1) \qquad \bot = (1, \emptyset, 1)$$

- We have $\mathrm{Hom}(A \otimes B, C) \cong \mathrm{Hom}(A, B \multimap C)$ and $A^* = A \to \bot$.

$\mathbf{Chu}_K(\mathbf{C})$ is a $*$-autonomous Category

- The tensor product and internal-hom of $A = (A, R, X)$ and $B = (B, S, Y)$ are defined by the pullback P and Q:



$$A \otimes B \coloneqq (A \otimes B, \otimes, P) \qquad A \multimap B \coloneqq (Q, \multimap, A \otimes Y)$$

Example: In $\mathbf{Chu}_2(\mathbf{Set})$, we have

$$(a \otimes b) \otimes (f, g) \coloneqq R(a, gb) \quad (\text{also} = S(b, fa))$$
$$(f, g) \multimap (a \otimes y) \coloneqq R(a, gy) \quad (\text{also} = S(fa, y))$$

- The linear negation of $A = (A, R, X)$:

$$A^* \coloneqq (X, R^\circ, A) \quad \text{where} \quad R^\circ(x, a) \coloneqq R(a, x)$$

- The unit and the dualizing object:

$$I = (I, \lambda_K : I \otimes K \to K, K) \qquad \bot = (K, \rho_K : K \otimes I \to K, I)$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

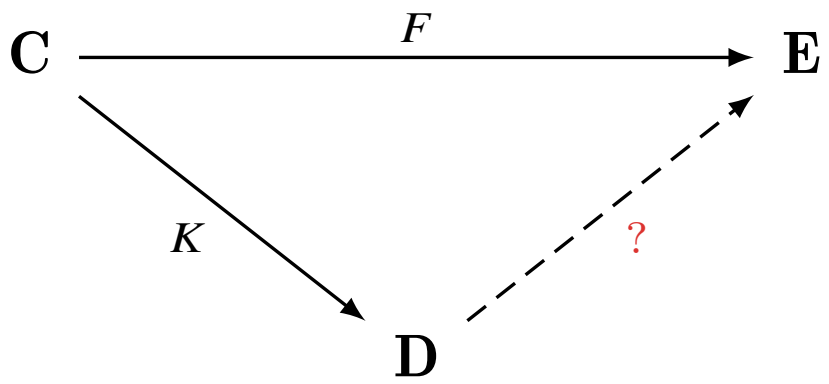
Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Kan Extension

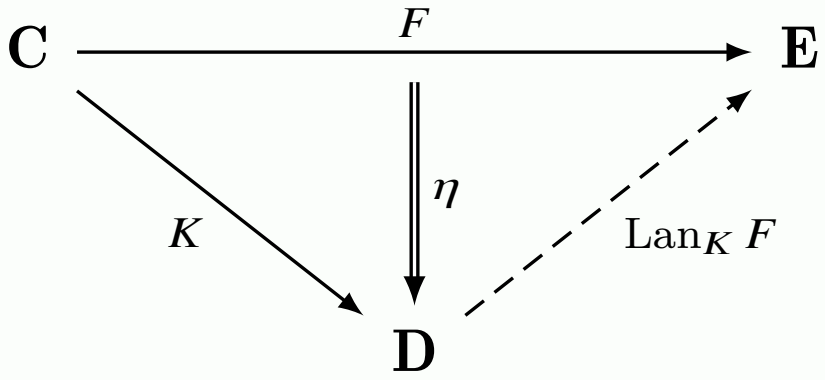


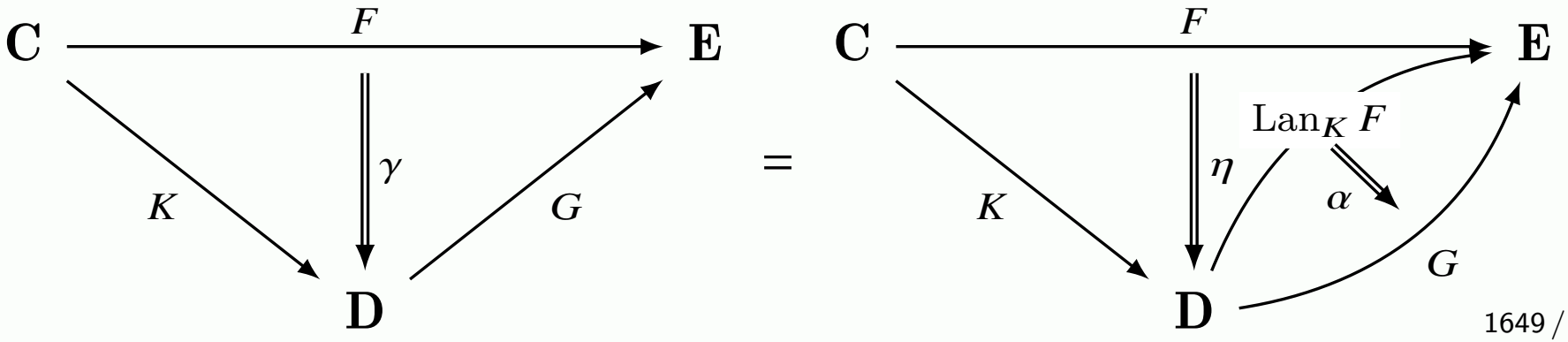
- Suppose we want to construct a functor $\mathbf{D} \to \mathbf{E}$ based the information carried by $F : \mathbf{C} \to \mathbf{E}$.
- Given a functor $K : \mathbf{C} \to \mathbf{D}$, we can ask
  1. what is the most 'liberal' extension?
  2. what is the most 'conservative' extension?

Left Kan Extension

Definition (Left Kan Extension)

Given functors $F : \mathbf{C} \to \mathbf{E}$ and $K : \mathbf{C} \to \mathbf{D}$, a *left Kan extension* of F along K is a functor $\mathrm{Lan}_K F : \mathbf{D} \to \mathbf{E}$ with a natural transformation $\eta : F \to \mathrm{Lan}_K F \circ K$ s.t. for any such pair $(G : \mathbf{D} \to \mathbf{E}, \gamma : F \to GK)$, there exists a unique natural transformation $\alpha : \mathrm{Lan}_K F \to G$ with $\gamma = \alpha K \circ \eta$.

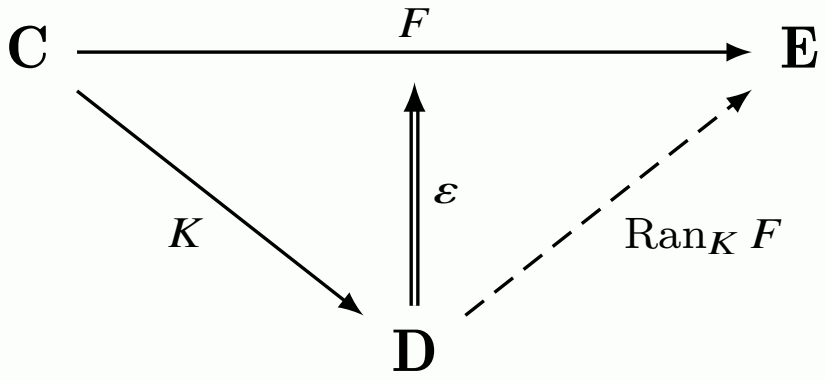


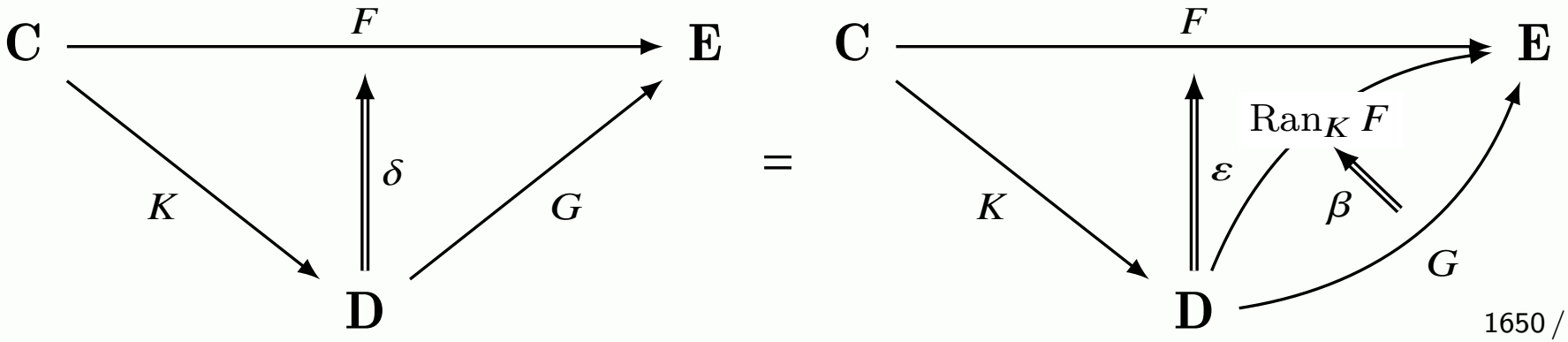


Right Kan Extension

Definition (Right Kan Extension)

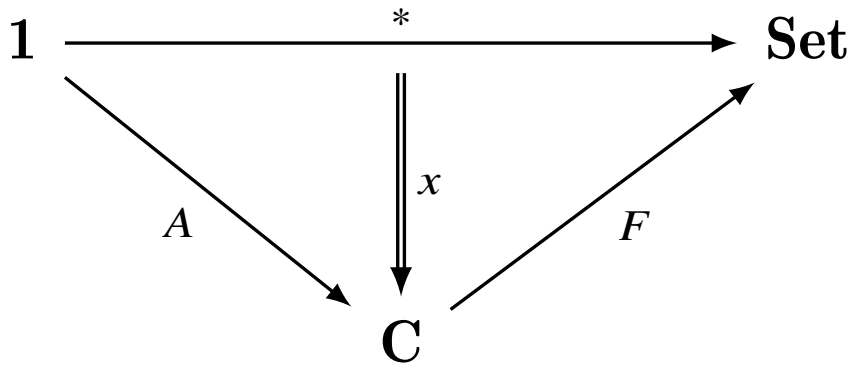
Given functors $F : \mathbf{C} \to \mathbf{E}$ and $K : \mathbf{C} \to \mathbf{D}$, a *right Kan extension* of F along K is a functor $\mathrm{Ran}_K F : \mathbf{D} \to \mathbf{E}$ with a natural transformation $\varepsilon : \mathrm{Ran}_K F \circ K \to F$ s.t. for any such pair $(G : \mathbf{D} \to \mathbf{E}, \delta : GK \to F)$, there exists a unique natural transformation $\beta : G \to \mathrm{Ran}_K F$ with $\delta = \varepsilon \circ \beta K$.



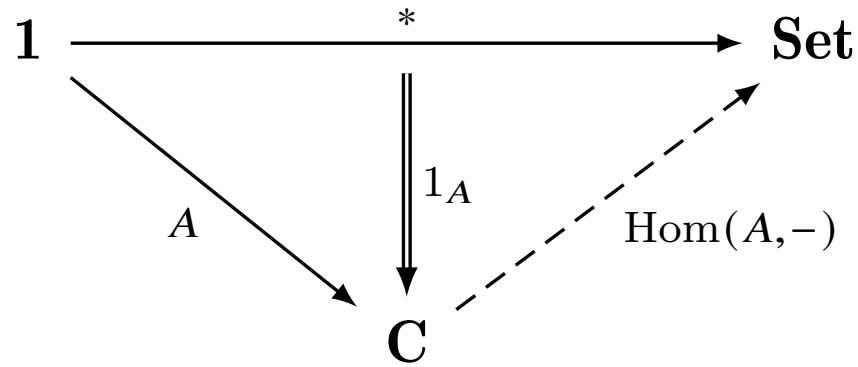


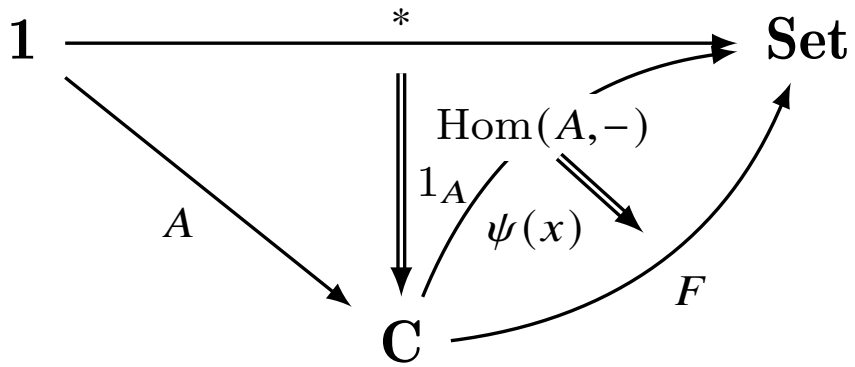
Example

For any object $A \in \mathbf{C}$ and any $F : \mathbf{C} \to \mathbf{Set}$, there is a bijection between elements $x \in FA$ and natural transformations with boundary as displayed



By the Yoneda lemma, the representable functor $\mathrm{Hom}(A, -)$ and the identity $1_A : A \to A$ define the left Kan extension of $* : \mathbf{1} \to \mathbf{Set}$ along $A : \mathbf{1} \to \mathbf{C}$.



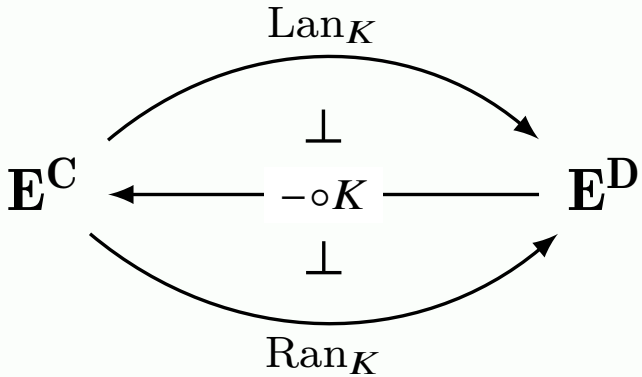


The required unique factorization is the natural transformation $\psi(x) : \mathrm{Hom}(A, -) \to F$ with $\psi(x)_A(1_A) = x$.

$\mathrm{Lan}_K \dashv K^* \dashv \mathrm{Ran}_K$

Theorem

If the Kan extensions exist for all F, then $\mathrm{Lan}_K \dashv K^* \dashv \mathrm{Ran}_K$, where $K^* := - \circ K$.



Proof.

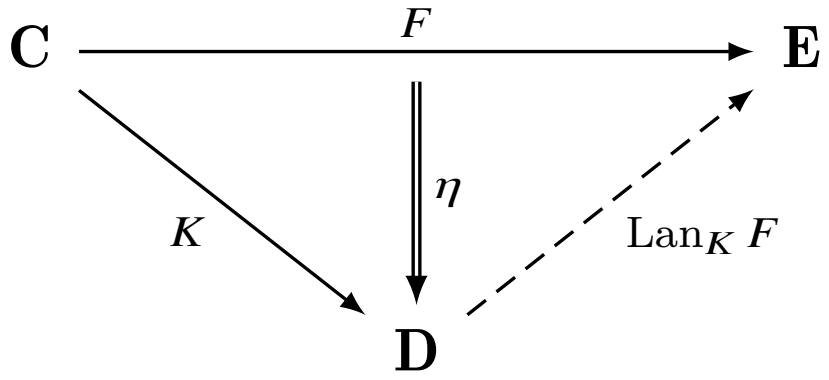
By the Yoneda Lemma, any pair (G, γ), as in the definition for the left Kan extension, yields a natural transformation by $\gamma^*_H(\alpha) := \alpha K \circ \gamma$.

$$\gamma^* : \mathbf{E}^{\mathbf{D}}(G, -) \to \mathbf{E}^{\mathbf{C}}(F, - \circ K)$$

The universal property of the left Kan extension says that (Lan_K, η) yields a natural isomorphism.

$$\mathbf{E}^{\mathbf{D}}(\mathrm{Lan}_K F, -) \cong \mathbf{E}^{\mathbf{C}}(F, - \circ K)$$

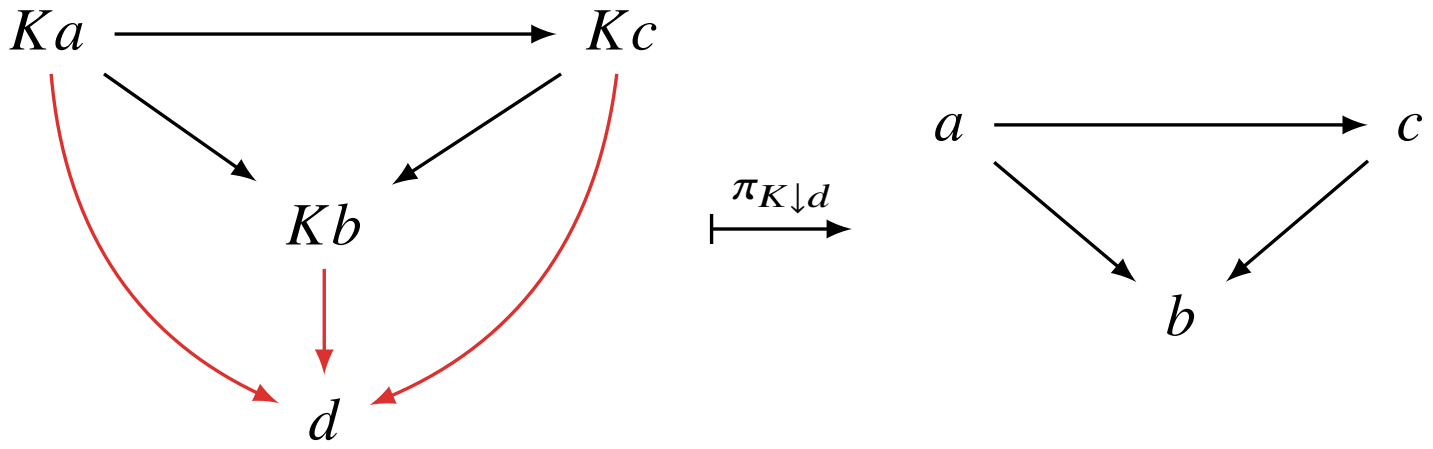
Computing Left Kan Extension



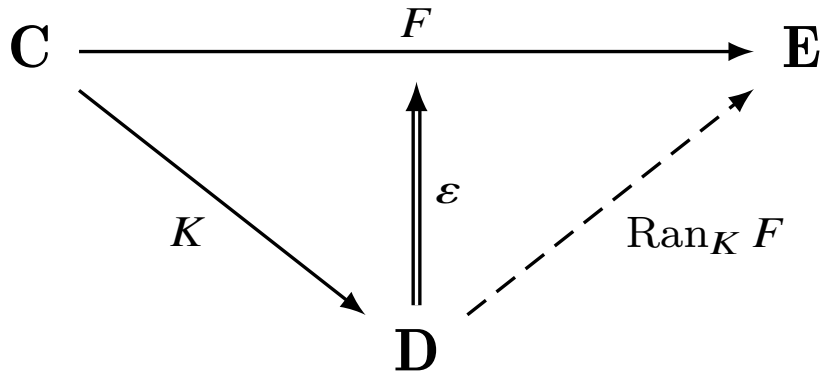
If $\mathbf{E}$ is cocomplete, then the left Kan extension $\mathrm{Lan}_K F$ exists and is:

$$\mathrm{Lan}_K F(d) \coloneqq \varinjlim \left(K \downarrow d \xrightarrow{\pi_{K\downarrow d}} \mathbf{C} \xrightarrow{F} \mathbf{E} \right) \quad \text{for} \quad d \in \mathbf{D}$$

with the natural transformation η extracted from colimiting cocones in $\mathbf{E}$.



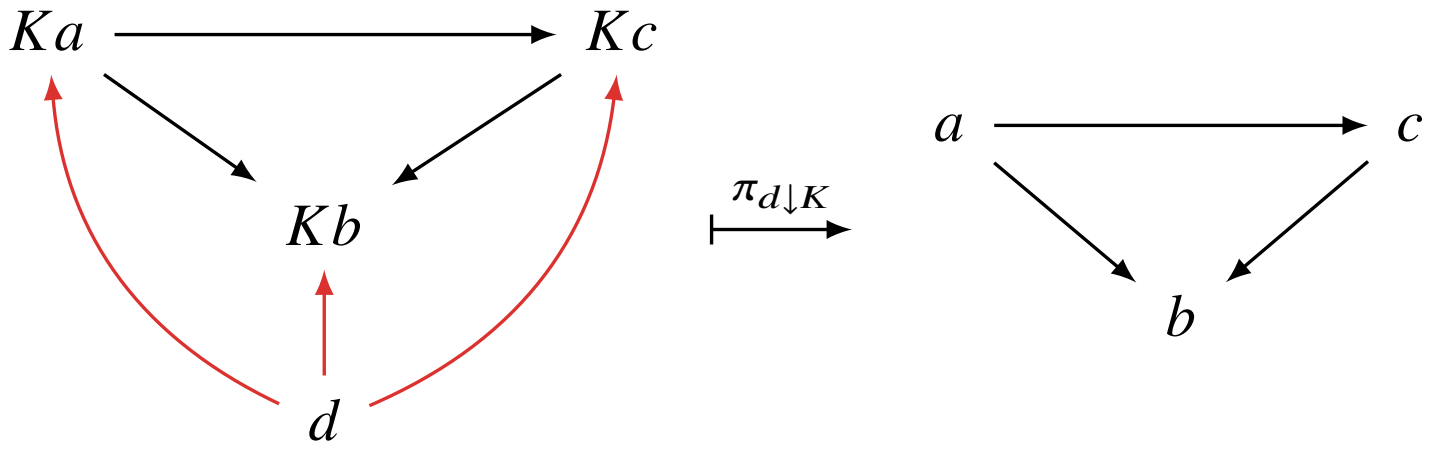
Computing Right Kan Extension



If $\mathbf{E}$ is complete, then the right Kan extension $\mathrm{Ran}_K F$ exists and is:

$$\mathrm{Ran}_K F(d) \coloneqq \varprojlim \left(d \downarrow K \xrightarrow{\pi_{d\downarrow K}} \mathbf{C} \xrightarrow{F} \mathbf{E} \right) \quad \text{for} \quad d \in \mathbf{D}$$

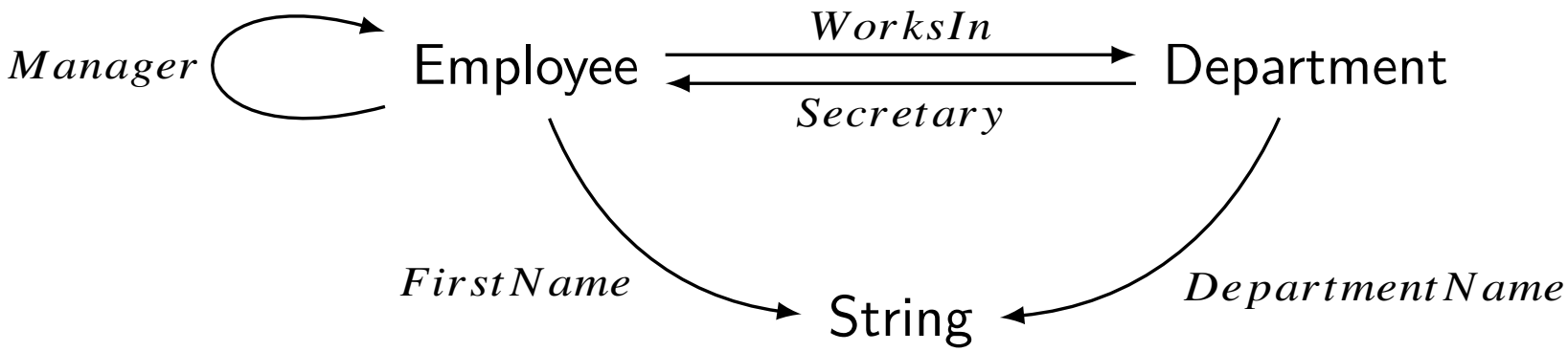
with the natural transformation ε extracted from limiting cones in $\mathbf{E}$.



Example: Databases

- A "database schema" is a category $\mathbf{D}$.
- An database built using this schema is a functor

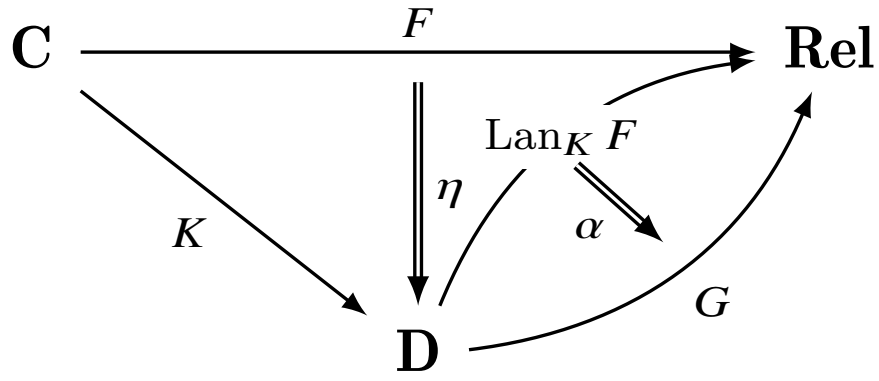
$$G : \mathbf{D} \to \mathbf{Rel}$$



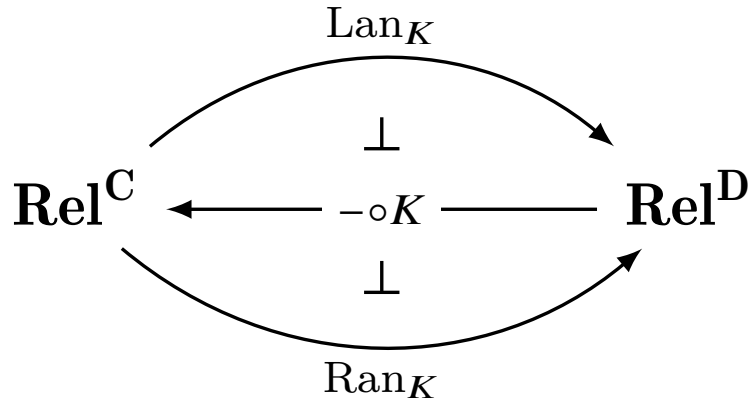
- How can we transform our database into a different database built using a different schema $\mathbf{C}$?

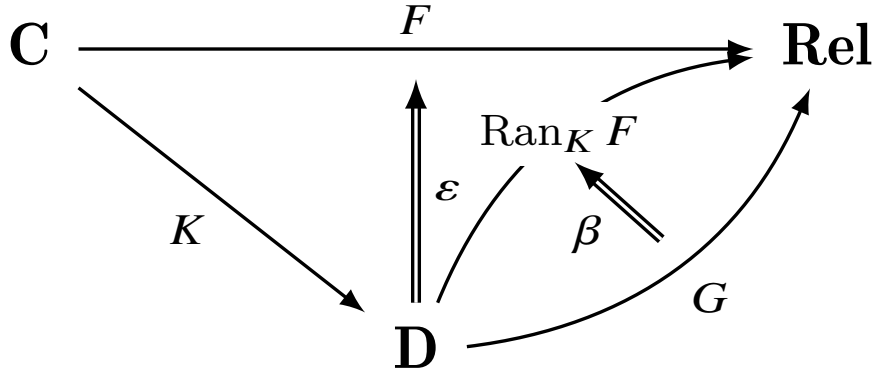
$$G \circ K : \mathbf{C} \xrightarrow{K} \mathbf{D} \xrightarrow{G} \mathbf{Rel}$$

- The functor $- \circ K : \mathbf{Rel}^{\mathbf{D}} \to \mathbf{Rel}^{\mathbf{C}}$ has both a left adjoint $\mathrm{Lan}_K : \mathbf{Rel}^{\mathbf{C}} \to \mathbf{Rel}^{\mathbf{D}}$ and a right adjoint $\mathrm{Ran}_K : \mathbf{Rel}^{\mathbf{C}} \to \mathbf{Rel}^{\mathbf{D}}$.



$$\frac{\alpha : \mathrm{Lan}_K F \implies G}{\eta : F \implies G \circ K}$$





$$\frac{\delta : G \circ K \implies F}{\beta : G \implies \mathrm{Ran}_K F}$$

Example: Databases

$$\mathbf{D} := \quad \text{Germans} \xrightarrow{Friend} \text{Italians}$$

$$\mathbf{C} := \quad \text{Germans}$$

$$K : \mathbf{C} \hookrightarrow \mathbf{D}$$

Germans	Friend
Ilsa	Giulia
Klaus	Gian-Carlo
Jörg	Martina
Sabine	Alessandro
Heinrich	Martina

Table: $G : \mathbf{D} \to \mathbf{Rel}$

Italians
Bianca
Giulia
Gian-Carlo
Alessandro
Martina

Germans
Ilsa
Klaus
Jörg
Sabine
Heinrich

Table: $G \circ K : \mathbf{C} \to \mathbf{Rel}$

Example: Databases — Left Kan Extension

Germans
Ilsa
Klaus
Jörg
Sabline

Table: $F : \mathbf{C} \to \mathbf{Rel}$

Germans	Friend	Italians
Ilsa	Italian1	Italian1
Klaus	Italian2	Italian2
Jörg	Italian3	Italian3
Sabine	Italian4	Italian4

Table: $\mathrm{Lan}_K F : \mathbf{D} \to \mathbf{Rel}$

Remark: the left Kan extension is a left adjoint, it does this in a "liberal" way. It freely makes up the entries obeying only the equations that are needed to get a valid database.

Example: Databases — Right Kan Extension

Germans
Ilsa
Klaus
Jörg
Sabline

Table: $F : \mathbf{C} \to \mathbf{Rel}$

Germans	Friend
Ilsa	Italian1
Klaus	Italian1
Jörg	Italian1
Sabine	Italian1

Italians
Italian1

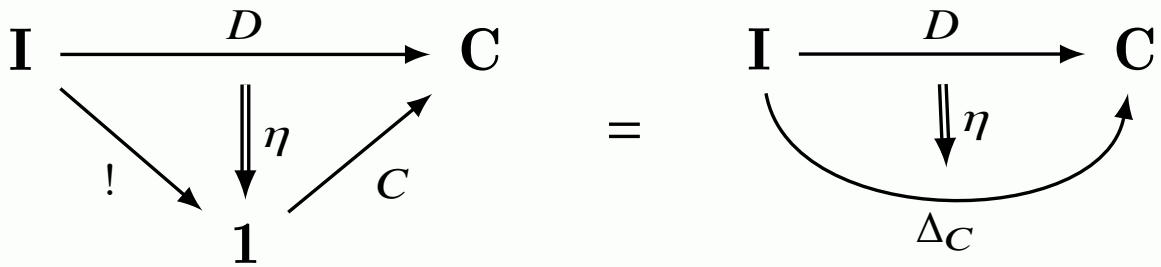
Table: $\mathrm{Ran}_K F : \mathbf{D} \to \mathbf{Rel}$

Remark: the right Kan extension is a right adjoint, it does this in a "conservative" way. It imposes all the equations that are possible in a valid database.

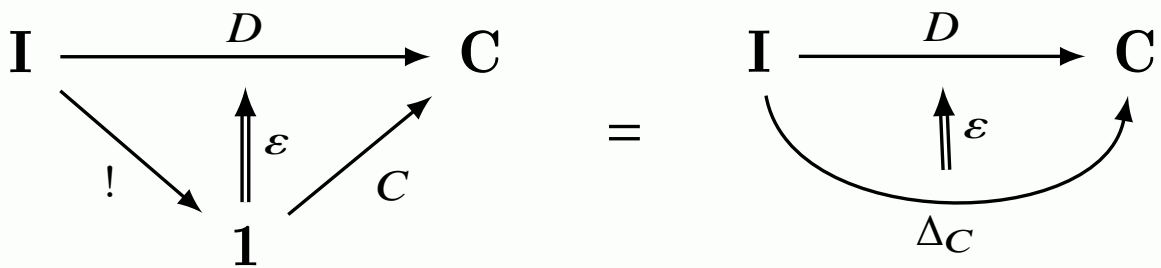
(Co)Limits are Kan Extensions

Theorem ((Co)Limits are Kan Extensions)

1. *The left Kan extension* $\mathrm{Lan}_! D$ *of* $D : \mathbf{I} \to \mathbf{C}$ *along* $! : \mathbf{I} \to \mathbf{1}$ *defines the colimit* $\varinjlim D$.



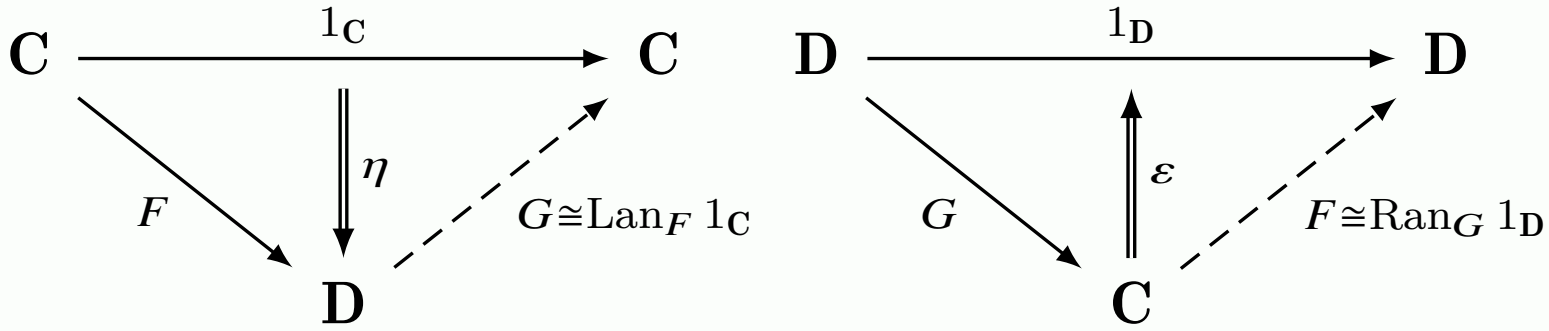
2. *Dually, the right Kan extension* $\mathrm{Ran}_! D$ *defines the limit* $\varprojlim D$.



Adjunctions are Kan Extensions

Theorem (Adjunctions are Kan Extensions)

1. *If $F \dashv G$ is an adjunction with unit $\eta : 1_{\mathbf{C}} \to GF$ and counit $\varepsilon : FG \to 1_{\mathbf{D}}$, then (G, η) is a left Kan extension of the identity functor $1_{\mathbf{C}}$ along F and (F, ε) is a right Kan extension of the identity functor $1_{\mathbf{D}}$ along G.*



Moreover, both Kan extensions are absolute(preserved by all functors).

2. *Conversely, if $(G, \eta : 1_{\mathbf{C}} \to GF)$ is a left Kan extension of the identity functor $1_{\mathbf{C}}$ along F and if F preserves this Kan extension, then $F \dashv G$ with unit η.*

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Monoidal Category

Definition (Monoidal Category)

A *monoidal category* $(\mathbf{C}, \otimes, I)$ is a category $\mathbf{C}$ equipped with:

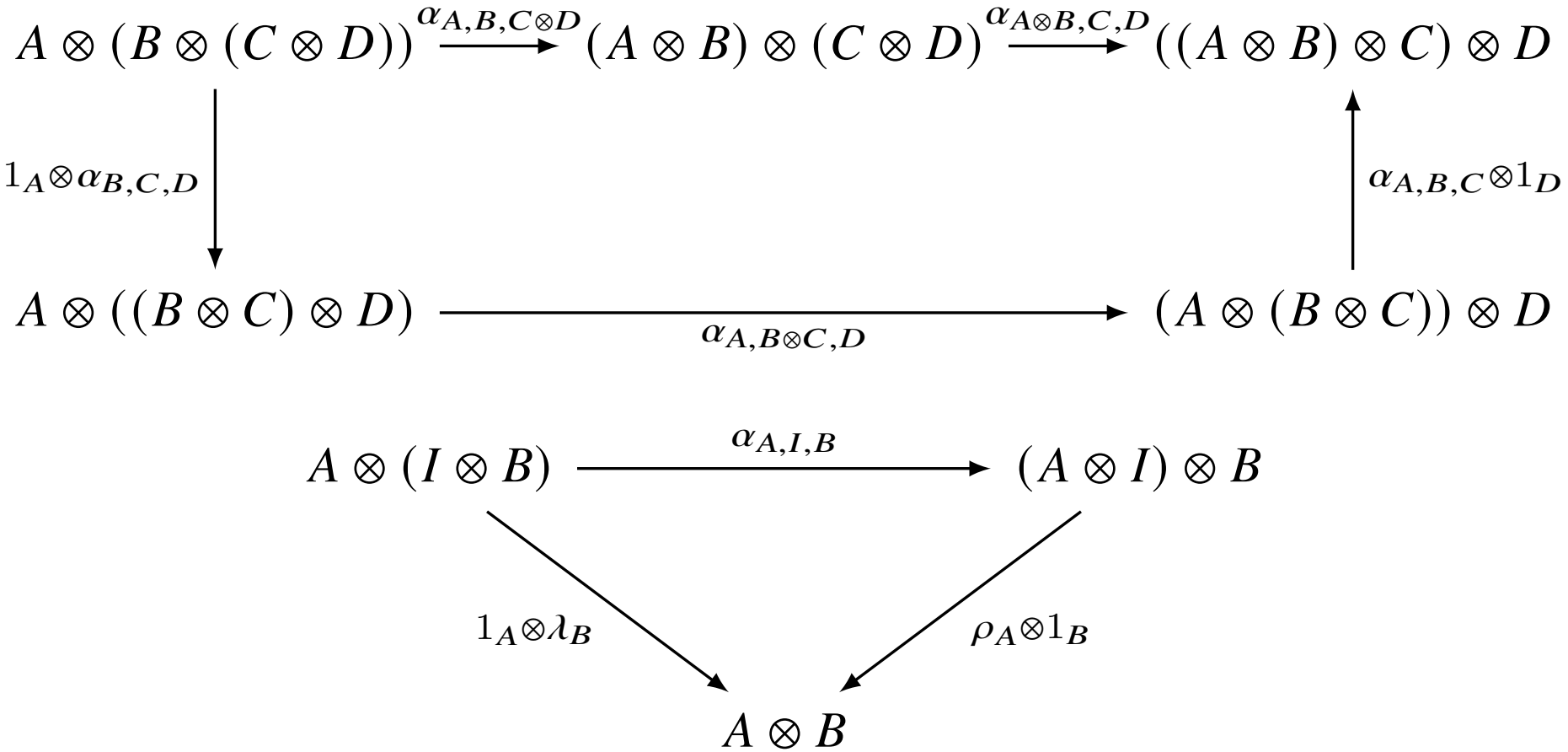
- a bifunctor $\otimes : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ called the *tensor product* or *monoidal product*.
- an object I called the *unit*.

such that,

- $\otimes$ is associative: there is a natural (in each of three arguments A, B, C) isomorphism α, called *associator*, with components $\alpha_{A,B,C} : A \otimes (B \otimes C) \cong (A \otimes B) \otimes C$,
- I acts as left and right unit: there are two natural isomorphisms λ and ρ, respectively called left and right *unitor*, with components $\lambda_A : I \otimes A \cong A$ and $\rho_A : A \otimes I \cong A$.

A *strict monoidal category* is one for which the natural isomorphisms α, λ and ρ are identities.

Monoidal Category



Example

- $(\mathbf{Set}, \times, \{\bullet\})$
- $(\mathbf{Set}, \coprod, \emptyset)$
- $(\mathbf{Vect}, \otimes, \mathbb{R})$
- $(\mathbf{Vect}, \oplus, \mathbf{1})$
- $(\mathbf{Top}, \coprod, \emptyset)$
- Any monoid can be thought of as a discrete monoidal category.
  - $(\mathbb{R}, +, 0)$
  - $(\mathbb{R}, \cdot, 1)$
  - $(\mathrm{List}(X), *, [])$
  - $(\{0, 1\}, \wedge, 1)$

1. A monoid $(M, \cdot, e)$ is a category that has only one object $\bullet$ s.t. $\mathrm{Hom}(\bullet, \bullet) = M$.
2. A monoid is a discrete category with a monoidal category structure where the tensor product is the monoid multiplication.

State

Definition (State)

A state of an object A in a monoidal category is a morphism $I \longrightarrow A$.

Example

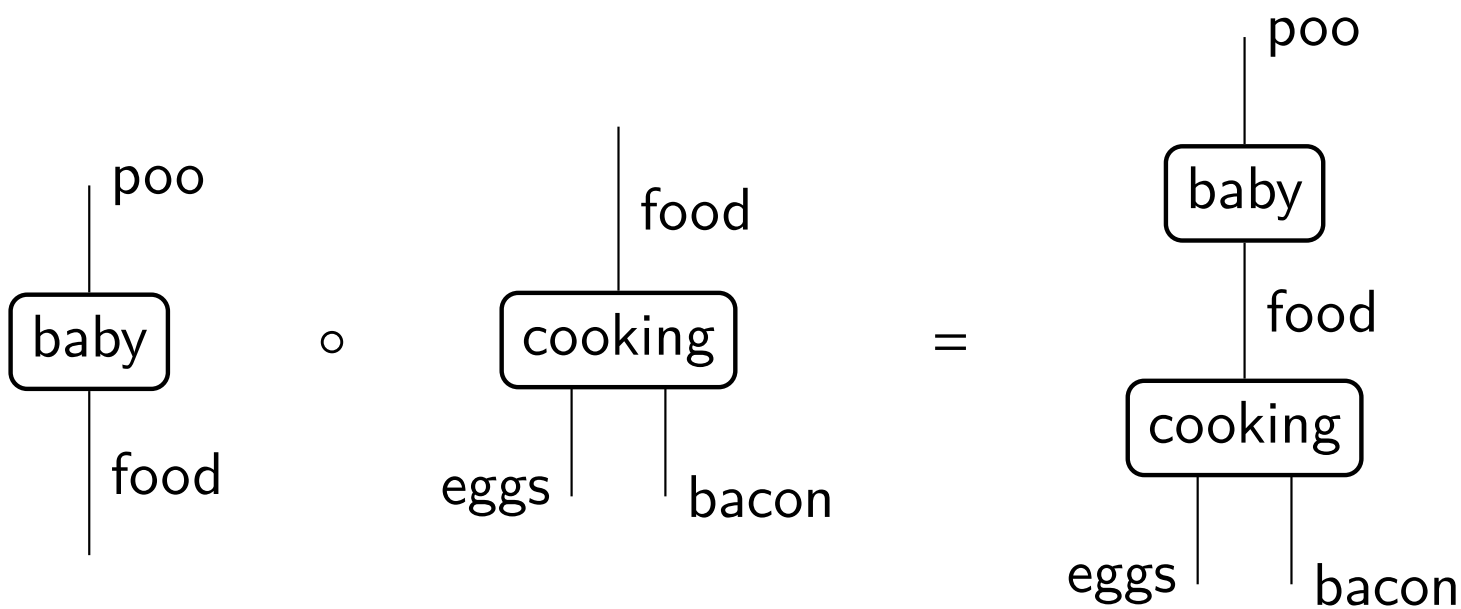
- In $\mathbf{Set}$, points of a set A are morphisms $\{\bullet\} \longrightarrow A$, which correspond to elements of A.
- In $\mathbf{Hilb}$, points of a Hilbert space H are morphisms $\mathbb{C} \longrightarrow H$, which correspond to elements of H by considering the image of $1 \in \mathbb{C}$.
- In $\mathbf{Rel}$, points of a set A are relations $\{\bullet\} \xrightarrow{R} A$, which correspond to subsets of A.

Process Theory

Definition (Process Theory)

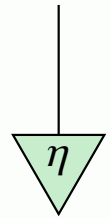
A process theory consists of:

- a collection T of system-types represented by wires,
- a collection P of processes represented by boxes, with inputs/outputs in T.
- a means of 'wiring processes together'.

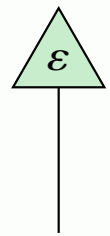


Special processes: states and effects

- Processes with no inputs are called states



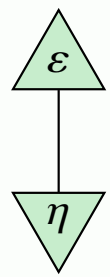
- Processes with no outputs are called effects



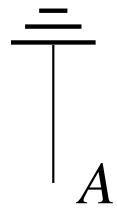
- A scalar is a process with no inputs or outputs.



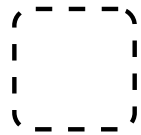
Remark: Interpret as: what happens when a state meets an effect



- There is a special effect "discarding": $A \longrightarrow I$.

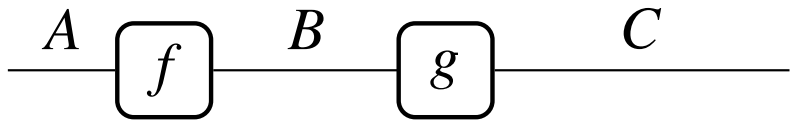


- The identity $1_I : I \longrightarrow I$ is represented as empty space.

$1_I \quad =$ 

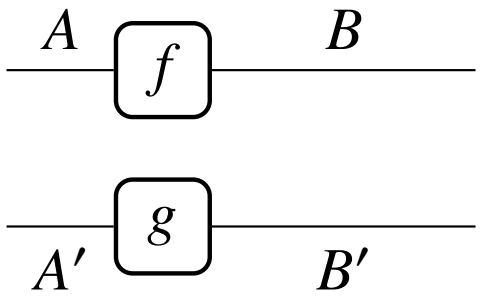
- composition of $f : A \to B$ and $g : B \to C$

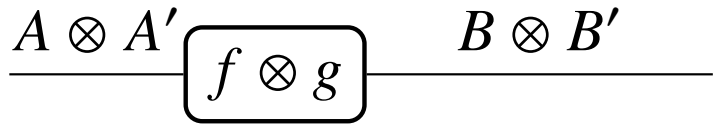
$$g \circ f : A \to C$$



- tensor product of $f : A \to B$ and $g : A' \to B'$

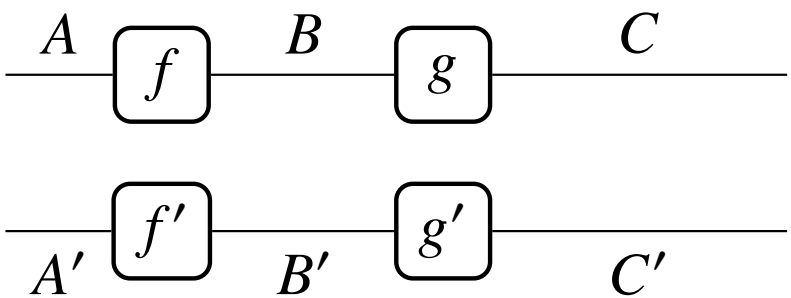
$$f \otimes g : A \otimes A' \to B \otimes B'$$

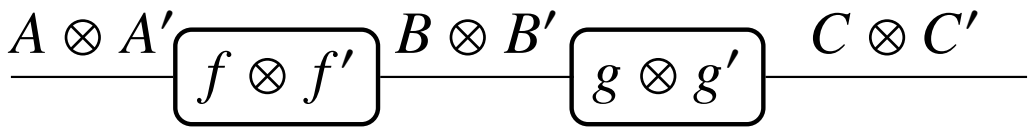




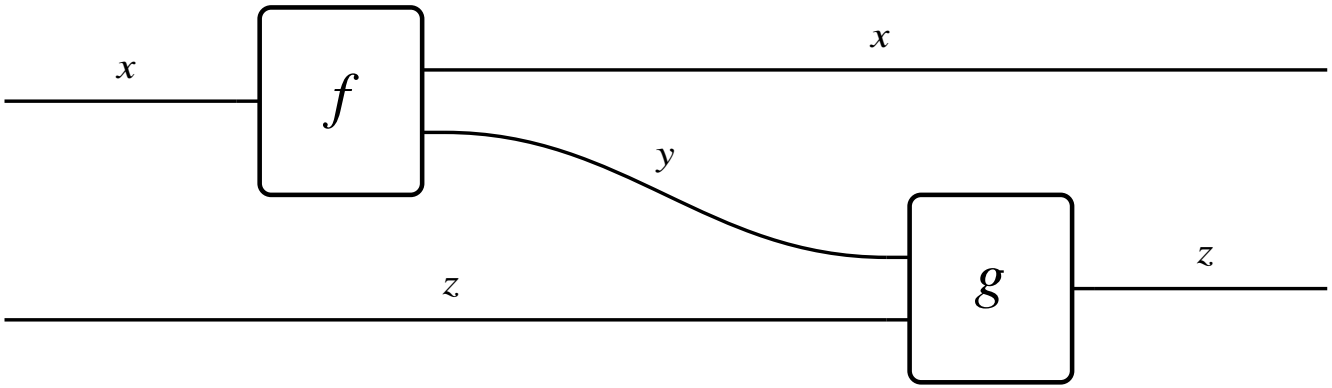
- composition and tensor product must obey

$$(g \circ f) \otimes (g' \circ f') = (g \otimes g') \circ (f \otimes f')$$

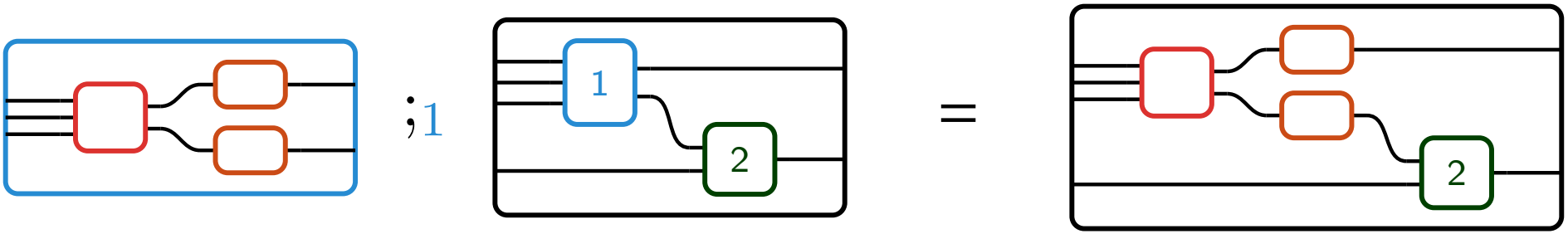




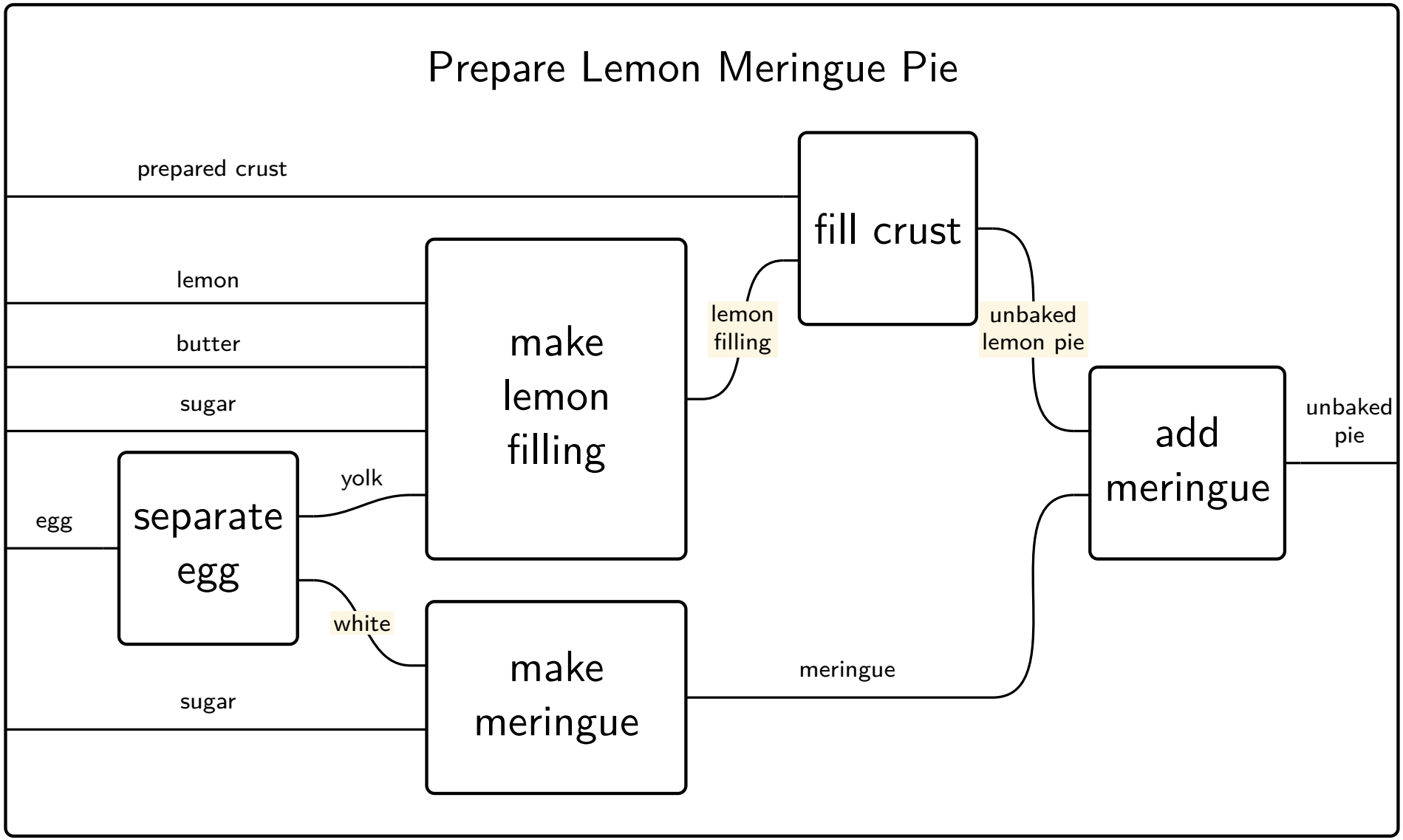
Wiring Diagram Representation — Example



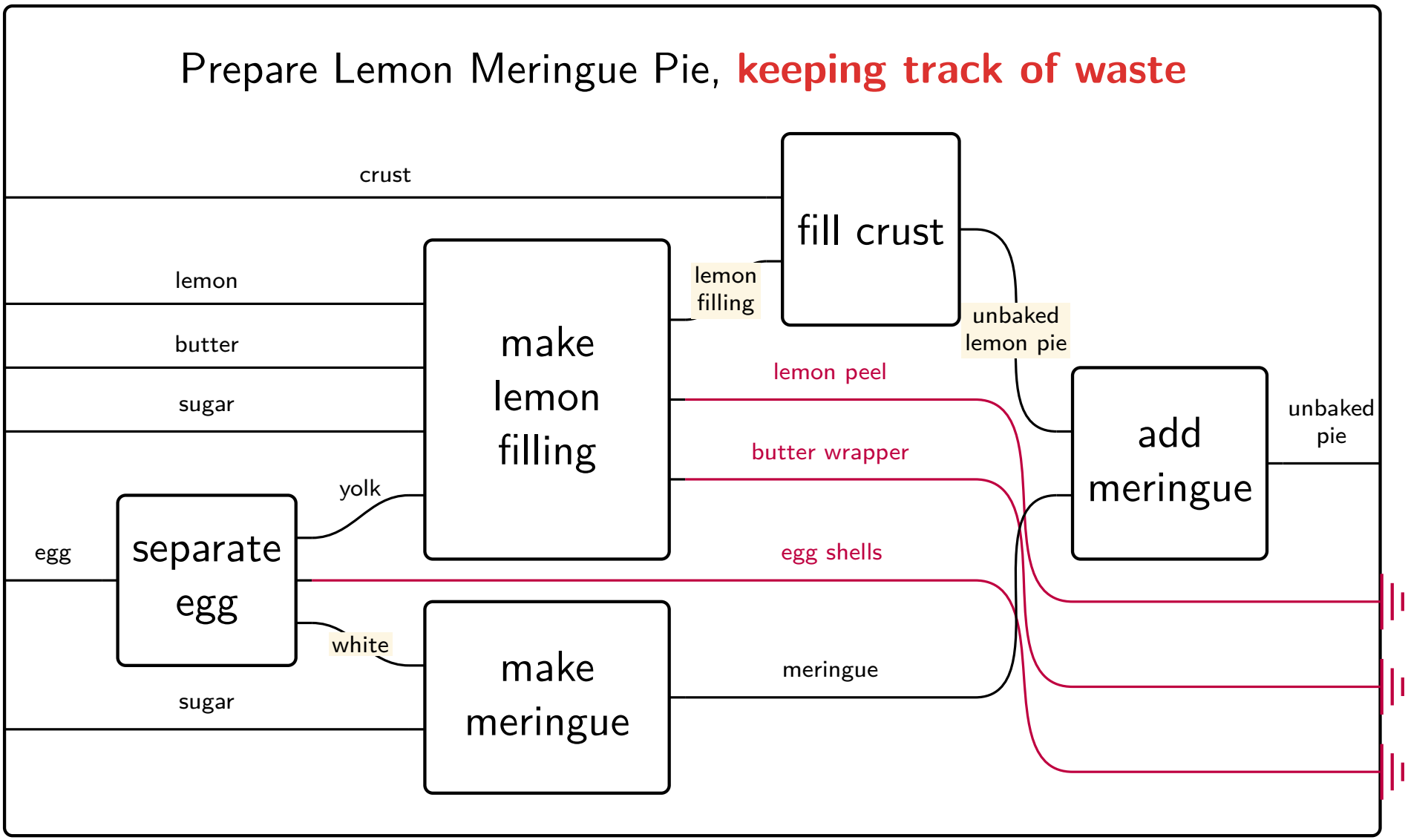
$$(f \otimes 1_z)\,;(1_x \otimes g)$$

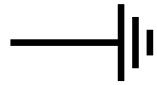


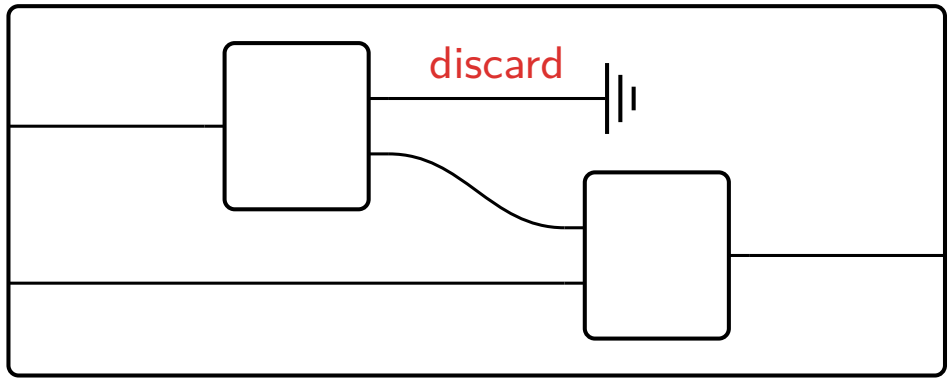
Wiring Diagram Example: Prepare Lemon Meringue Pie

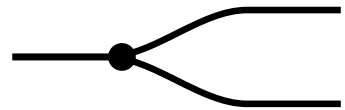


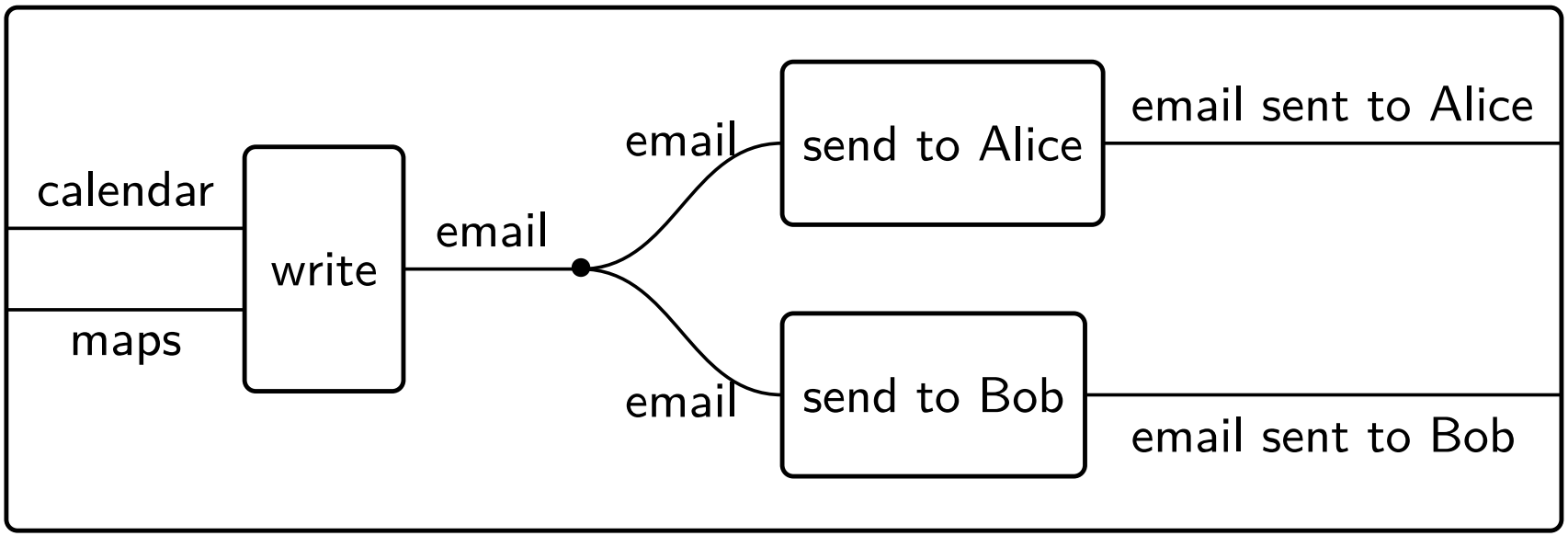
Wiring Diagram Example: Prepare Lemon Meringue Pie



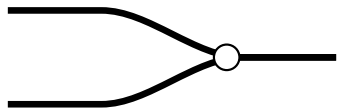
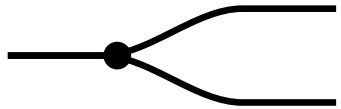
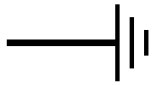
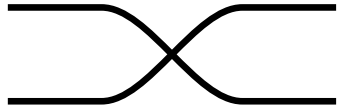
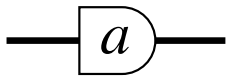
- To "discard waste"  , one just adds an additional axiom: $x \leq I$ for all x.



- Information can be copied. To "copy information"  , one just adds an additional axiom: $x \leq x + x$ for all x.



Wiring Diagram vs Matrix

Generator	Icon	Matrix	Arity
add		$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$2 \to 1$
zero		$\begin{bmatrix} \end{bmatrix}$	$0 \to 1$
copy		$\begin{bmatrix} 1 & 1 \end{bmatrix}$	$1 \to 2$
discard		$\begin{bmatrix} \end{bmatrix}$	$1 \to 0$
swap		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$2 \to 2$
identity		$\begin{bmatrix} 1 \end{bmatrix}$	$1 \to 1$
scalar		$\begin{bmatrix} a \end{bmatrix}$	$1 \to 1$

Example: Monoidal Category $\mathbf{Rel}$

Definition

The monoidal structure $(\mathbf{Rel}, \times, \{\bullet\})$ on the category $\mathbf{Rel}$ is defined as follows:

- the tensor product is Cartesian product of sets, $A \times A' \xrightarrow{R \times R'} B \times B'$ for $A \xrightarrow{R} B \;\&\; A' \xrightarrow{R'} B'$, where

$$R \times R'((a, a'), (b, b')) \coloneqq Rab \wedge R'a'b'$$

- the unit object is a chosen singleton set $I \coloneqq \{\bullet\}$.
- associators $A \times (B \times C) \xrightarrow{\alpha_{A,B,C}} (A \times B) \times C$ are the relations defined by $(a, (b, c)) \sim ((a, b), c)$.
- left unitors $I \times A \xrightarrow{\lambda_A} A$ are the relations defined by $(\bullet, a) \sim a$.
- right unitors $A \times I \xrightarrow{\rho_A} A$ are the relations defined by $(a, \bullet) \sim a$.

Example: Monoidal Category $\mathbf{Hilb}$

Definition

The monoidal category $(\mathbf{Hilb}, \otimes, \mathbb{C})$ is defined as follows:

- Objects are separable complex Hilbert spaces $H, J, K, \ldots$
- Morphisms are bounded linear maps $f, g, h, \ldots$
- Composition is composition of linear maps;
- Identity maps are given by the identity linear maps;
- Tensor product $\otimes$ is tensor product of Hilbert spaces;
- Unit object I is the 1-dimensional Hilbert space $\mathbb{C}$;
- Associators $\alpha_{H,J,K} : (H \otimes J) \otimes K \to H \otimes (J \otimes K)$ are the unique linear maps sending $|\phi\rangle \otimes (|\chi\rangle \otimes |\psi\rangle) \mapsto (|\phi\rangle \otimes |\chi\rangle) \otimes |\psi\rangle$ for all $|\phi\rangle \in H$, $|\chi\rangle \in J$, and $|\psi\rangle \in K$;
- Left unitor $\lambda_H : \mathbb{C} \otimes H \to H$ are the unique linear maps sending $1 \otimes |\phi\rangle \mapsto |\phi\rangle$ for all $|\phi\rangle \in H$;
- Right unitor $\rho_H : H \otimes \mathbb{C} \to H$ are the unique linear maps sending $|\phi\rangle \otimes 1 \mapsto |\phi\rangle$ for all $|\phi\rangle \in H$.

Examples: Monoidal Categories

1. $(\mathbb{N}, \leq, +, 0)$ is a monoidal preorder. $(\mathbb{R}, \leq, \cdot, 1)$ is not a monoidal preorder, since $-5 \leq 0$ and $-4 \leq 3$ but $(-5) \cdot (-4) \not\leq 0 \cdot 3$.
2. $(\mathrm{P}(X), \subset, \cup, \emptyset)$ and $(\mathrm{P}(X), \subset, \cap, X)$ are strict symmetric monoidal preorders.
3. The category of all endofunctors on a category $\mathbf{C}$ is a strict monoidal category $(\mathbf{End}_{\mathbf{C}}, \circ, 1_{\mathbf{C}})$ with the composition of functors as the product and the identity functor as the unit.
4. R-$\mathbf{Mod}$, the category of modules over a commutative ring R, is a monoidal category with the tensor product of modules $\otimes_R$ serving as the monoidal product and the ring R (thought of as a module over itself) serving as the unit.
5. Any category with finite products can be regarded as monoidal with the product as the tensor product and the terminal object as the unit. Such a category is sometimes called a cartesian monoidal category. For example: $(\mathbf{Set}, \times, \{\bullet\})$ and $(\mathbf{Cat}, \otimes, \mathbf{1})$.

Cartesian Monoidal Category

Theorem

If $(\mathbf{C}, \otimes, I, \sigma)$ *is a symmetric monoidal category equipped with monoidal natural transformations*

$$\Delta_A : A \to A \otimes A$$

and

$$\varepsilon_A : A \to I$$

such that the following composites are identity morphisms:

$$A \xrightarrow{\Delta_A} A \otimes A \xrightarrow{\varepsilon_A \otimes 1_A} I \otimes A \xrightarrow{\lambda_A} A$$

$$X \xrightarrow{\Delta_A} A \otimes A \xrightarrow{1_A \otimes \varepsilon_A} A \otimes I \xrightarrow{\rho_A} A$$

then $(\mathbf{C}, \otimes, I, \sigma, \Delta, \varepsilon)$ *is a cartesian monoidal category.*

Remark: a symmetric monoidal category is cartesian if we can duplicate and discard data, and "duplicating a piece of data and then discarding one copy is the same as not doing anything".

Braided Monoidal Category

Definition (Braided Monoidal Category)

A *braided monoidal category* $(\mathbf{C}, \otimes, I, \sigma)$ is a monoidal category $(\mathbf{C}, \otimes, I)$ equipped with a natural isomorphism σ called the *braiding* that assigns to every pair of objects $A, B \in \mathbf{C}$ an isomorphism $\sigma_{A,B} : A \otimes B \to B \otimes A$ s.t.,

$$
\begin{array}{ccccc}
A \otimes (B \otimes C) & \xrightarrow{\alpha_{A,B,C}} & (A \otimes B) \otimes C & \xrightarrow{\sigma_{A,B} \otimes 1_C} & (B \otimes A) \otimes C \\
\downarrow{\scriptstyle \sigma_{A,B \otimes C}} & & & & \downarrow{\scriptstyle \alpha^{-1}_{B,A,C}} \\
(B \otimes C) \otimes A & \xleftarrow[\alpha_{B,C,A}]{} & B \otimes (C \otimes A) & \xleftarrow[1_B \otimes \sigma_{A,C}]{} & B \otimes (A \otimes C)
\end{array}
$$

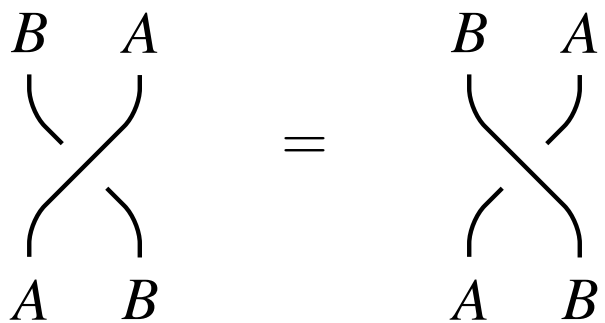
$$
\begin{array}{ccccc}
(A \otimes B) \otimes C & \xrightarrow{\alpha^{-1}_{A,B,C}} & A \otimes (B \otimes C) & \xrightarrow{1_A \otimes \sigma_{B,C}} & A \otimes (C \otimes B) \\
\downarrow{\scriptstyle \sigma_{A \otimes B,C}} & & & & \downarrow{\scriptstyle \alpha_{A,C,B}} \\
C \otimes (A \otimes B) & \xleftarrow[\alpha^{-1}_{C,A,B}]{} & (C \otimes A) \otimes B & \xleftarrow[\sigma_{A,C} \otimes 1_B]{} & (A \otimes C) \otimes B
\end{array}
$$

Symmetric Monoidal Category

Definition (Symmetric Monoidal Category)

A *symmetric monoidal category* $(\mathbf{C}, \otimes, I, \sigma)$ is a braided monoidal category where the braiding satisfies $\sigma_{B,A} \circ \sigma_{A,B} = 1_{A \otimes B}$.

Remark: In a symmetric monoidal category, the braiding $\sigma_{A,B} = \sigma_{B,A}^{-1}$.



Example from Chemistry

Example

$(\mathrm{Mat}, \to, +, 0)$ is a symmetric monoidal preorder, where Mat is the set of all collections of atoms and molecules, sometimes called *materials*.

For example, given these chemical reactions:

$$\mathrm{C} + \mathrm{O_2} \to \mathrm{CO_2}$$
$$\mathrm{CO_2} + \mathrm{NaOH} \to \mathrm{NaHCO_3}$$
$$\mathrm{NaHCO_3} + \mathrm{HCl} \to \mathrm{H_2O} + \mathrm{NaCl} + \mathrm{CO_2}$$

we have

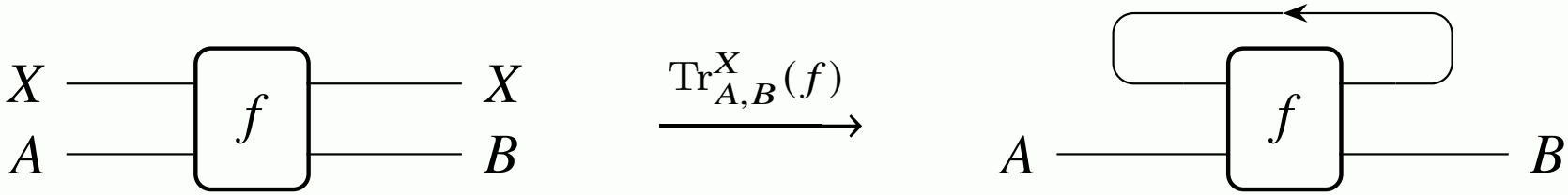
$$\mathrm{C} + \mathrm{O_2} + \mathrm{NaOH} + \mathrm{HCl} \to \mathrm{CO_2} + \mathrm{H_2O} + \mathrm{NaCl}$$

Traced Symmetric Monoidal Category

Definition (Traced Symmetric Monoidal Category)

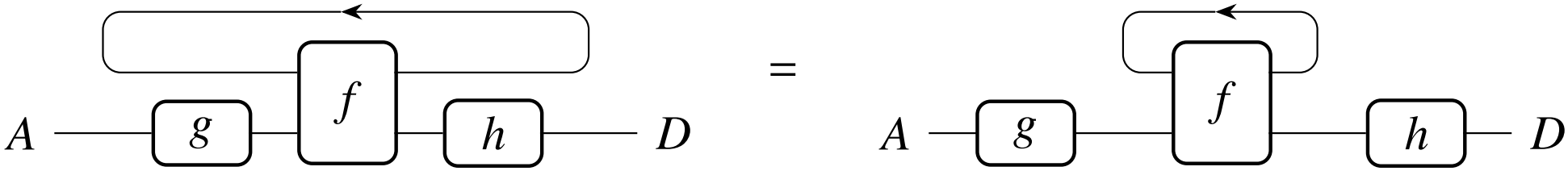
A symmetric monoidal category $(\mathbf{C}, \otimes, I, \sigma)$ is said to be traced if it is equipped with a family of operators

$$\mathrm{Tr}^X_{A,B} : \mathbf{C}(A \otimes X, B \otimes X) \to \mathbf{C}(A, B)$$

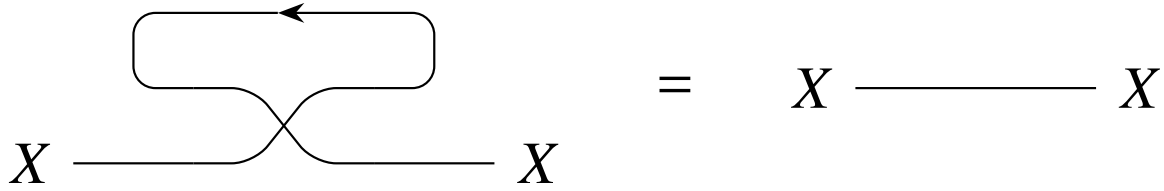


satisfying the following axioms:

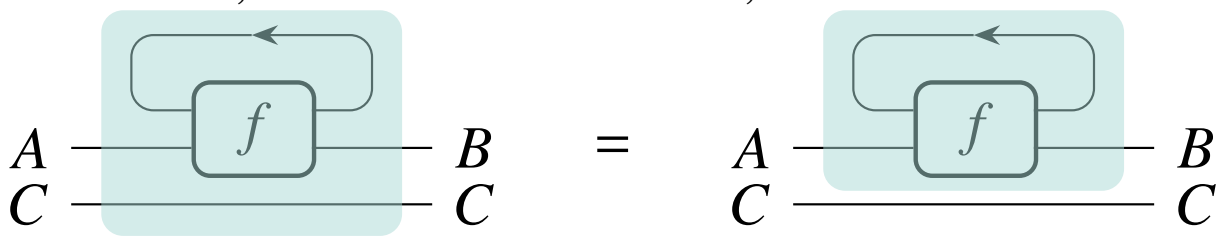
1. Tightening: $\mathrm{Tr}^X_{A,D}(1_X \otimes g\,;f\,;1_X \otimes h) = g\,;\mathrm{Tr}^X_{B,C}(f)\,;h$



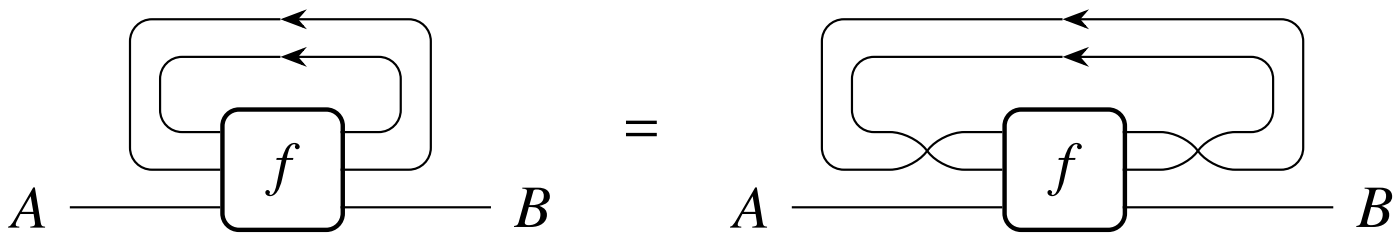
2. Yanking: $\mathrm{Tr}^X_{X,X}(\sigma_{X,X}) = 1_X$



3. Superposing: $\mathrm{Tr}^X_{A\otimes C,B\otimes C}(f \otimes 1_C) = \mathrm{Tr}^X_{A,B}(f) \otimes 1_C$



4. Exchange:
$\mathrm{Tr}^Y_{A,B}(\mathrm{Tr}^X_{Y\otimes A,Y\otimes B}(f)) = \mathrm{Tr}^X_{A,B}(\mathrm{Tr}^Y_{X\otimes A,X\otimes B}(\sigma_{Y,X} \otimes 1_A\,;f\,;\sigma_{X,Y} \otimes 1_B))$



Left/Right Closed Monoidal Category

Definition (Left/Right Closed Monoidal Category)

A monoidal category $\mathbf{C}$ is *left closed* if there is an *internal hom* functor

$$\multimap\colon \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{C}$$

together with a natural isomorphism c called *currying* that assigns to any objects $A, B, C \in \mathbf{C}$ a bijection

$$c_{A,B,C} : \mathrm{Hom}(A \otimes B, C) \xrightarrow{\cong} \mathrm{Hom}(A, B \multimap C)$$

It is *right closed* if there is an internal hom functor as above and a natural isomorphism

$$c_{A,B,C} : \mathrm{Hom}(A \otimes B, C) \xrightarrow{\cong} \mathrm{Hom}(B, A \multimap C)$$

Remark: there is no difference between left and right closed for a braided monoidal category, as the braiding gives an isomorphism $A \otimes B \cong B \otimes A$.

Theorem

For $(\mathbf{C}, \otimes, I)$ a closed monoidal category with internal hom, then not only are there natural bijections

$$\mathrm{Hom}(A \otimes B, C) \cong \mathrm{Hom}(A, B \multimap C)$$

but these isomorphisms themselves "internalize" to isomorphisms in $\mathbf{C}$ of the form

$$A \otimes B \multimap C \cong A \multimap B \multimap C$$

Proof.

By the external natural bijections there is for every $X \in \mathbf{C}$:

$$\begin{aligned} &\mathrm{Hom}(X, A \otimes B \multimap C) \cong \mathrm{Hom}(X \otimes A \otimes B, C) \cong \mathrm{Hom}(X \otimes A, B \multimap C) \\ \cong &\mathrm{Hom}(X, A \multimap B \multimap C) \end{aligned}$$

The Yoneda lemma implies that there is an isomorphism

$$A \otimes B \multimap C \cong A \multimap B \multimap C$$

Examples

- The category $\mathbf{Cost} := ([0, \infty], \geq, +, 0, \multimap)$ is closed. The internal hom is given by $a \multimap b := \max\{0, b - a\}$.
- The interval category $([0, 1], \leq, \cdot, 1, \multimap)$ is closed. The internal hom is given by $a \multimap b := \min\{1, \frac{b}{a}\}$.
- The categories $\mathbf{Set}$, $\mathbf{Vect}_{\mathbb{K}}$, and $\mathbf{FdHilb}$ are closed and in each case $A \multimap B$ is the appropriate function space. For $\mathbf{Set}$, it's the set of functions and for $\mathbf{Vect}_{\mathbb{K}}$ and $\mathbf{FdHilb}$, it's the set of linear functions.
- The category of finite-dimensional Hilbert spaces $\mathbf{FdHilb}$ is closed, but the larger category of Hilbert spaces $\mathbf{Hilb}$ is not.
- Every category of presheaves, i.e. every functor category of the form $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$, is closed.
- Any topos is closed.
- The category of topological spaces is not closed, but the category of compactly generated Hausdorff spaces is.
- The category of smooth manifolds and smooth maps is not closed.

$*$-autonomous Category

Definition ($*$-autonomous Category)

A $*$-*autonomous* category $(\mathbf{C}, \otimes, I, \multimap, \bot)$ is a symmetric monoidal closed category $\mathbf{C}$ with a *dualizing* object $\bot$, such that if we set $A^* := A \multimap \bot$, the canonical morphism $A \longrightarrow A^{**}$ is an isomorphism.

Remark: The operation $(-)^*$ induces a contravariant dualizing functor

$$* : \mathbf{C} \longrightarrow \mathbf{C}^{\mathrm{op}}$$

such that

- $\mathrm{Hom}(A, B) \cong \mathrm{Hom}(B^*, A^*)$
- $\mathrm{Hom}(A \otimes B, C) \cong \mathrm{Hom}(A, B \multimap C)$
- $A \multimap B \cong (A \otimes B^*)^*$
- $I \cong \bot^*$

Definition: A compact closed category is a $*$-autonomous category with natural isomorphisms $(A \otimes B)^* \cong A^* \otimes B^*$ and $I^* \cong I$.

Remark: It follows that $A \multimap B \cong A^* \otimes B$.

Linear Logic[33][Gir11]

$$A \coloneqq p \mid 0 \mid 1 \mid \bot \mid \top \mid A^{\bot} \mid A \oplus A \mid A \,\&\, A \mid A \otimes A \mid A \parr A \mid A \multimap A \mid !A \mid ?A$$

$\oplus$	plus	+	additives
0	the unit of $\oplus$		
$\&$	with	−	
$\top$	the unit of $\&$		
$\otimes$	tensor product	+	multiplictives
1	the unit of $\otimes$		
$\parr$	parallel product	−	
$\bot$	the unit of $\parr$		
$!$	of course	+	exponential
$?$	why not	−	modalities

Positive	Negative
Action	Reaction
Answer	Question
Output	Input
Consumer	Producer

[33] Philip Wadler: A Taste of Linear Logic.

Negation

$(p)^\perp = p^\perp$	$(p^\perp)^\perp = p$
$(A \otimes B)^\perp = A^\perp \parr B^\perp$	$(A \parr B)^\perp = A^\perp \otimes B^\perp$
$(A \oplus B)^\perp = A^\perp \,\&\, B^\perp$	$(A \,\&\, B)^\perp = A^\perp \oplus B^\perp$
$(1)^\perp = \bot$	$(\bot)^\perp = 1$
$(0)^\perp = \top$	$(\top)^\perp = 0$
$(!A)^\perp = ?(A^\perp)$	$(?A)^\perp = !(A^\perp)$

Linear Logic — A Logic of Resources

- $A \multimap B := A^{\perp} \parr B$: consume A and produce B in parallel.
- $A \parr B$: you have both A and B, but you can't use them together.
- $A \otimes B$: you can have both A and B simultaneously.
- $A \,\&\, B$: you can choose from A and B, but not both simultaneously.
- $A \oplus B$: you may have A or B, but you have no choice.
- $!A$: you can produce as many copies of A as you want, including zero copies.
- $?A$: you can consume as many copies of A as you want, including zero copies.
- 1: the trivial resource that can be produced from nothing. $A \otimes 1 \equiv A$.
- $\top$: it consumes all resources. $A \,\&\, \top \equiv A$.
- 0: the impossible resource, or something that will produce any resource. $A \oplus 0 \equiv A$.
- $\perp$: the resource that can be consumed by nothing. $A \parr \perp \equiv A$.
- $A^{\perp}$: the demand for an A. Negation of a consumer gives rise to a producer, and vice versa. $A^{\perp\perp} \equiv A$.

Example

- $P \otimes C \vdash H$: I will be happy given both a pizza and a cake.
- $P \,\&\, C \vdash H$: I will be happy given my choice from a pizza and a cake.
- $P \oplus C \vdash H$: I will be happy given either a pizza or a cake, I don't care which.

Menu: $\$5$	$(\$1 \otimes \$1 \otimes \$1 \otimes \$1 \otimes \$1)$
	$\multimap$
Fish	Fish
	$\otimes$
Chips	Chips
	$\otimes$
Soup or Salad	Soup $\&$ Salad
	$\otimes$
Fruit or Cheese (depending on availability)	Fruit $\oplus$ Cheese
	$\otimes$
Coffee (free refills)	Coffee $\otimes$!Coffee

	classical	linear	
re-use	$A, A \longrightarrow B \vdash A \wedge B$	$A, A \multimap B \nvdash A \otimes B$	$!A \vdash A \otimes !A$
discard	$A \wedge B \vdash A$	$A \otimes B \nvdash A$	$!A \otimes B \vdash B$

- $A \,\&\, B \vdash A \oplus B$
- $A \oplus B \nvdash A \,\&\, B$
- $A \otimes B \nvdash A \,\&\, B$
- $A \,\&\, B \nvdash A \otimes B$
- $A \nvdash A \otimes A$
- $A \otimes A \nvdash A$
- $A \multimap B \,\&\, C \vdash A \otimes A \multimap B \otimes C$

Translation

1. You can spend 1 dollar to buy a bottle of water or a bag of chips.

$$D \multimap W \,\&\, C$$

2. You can exchange a ten-dollar bill for two five-dollar bills.

$$T \multimap F \otimes F$$

3. If you have a water bottle, you can refill it with water as many times as you want.

$$B \multimap !W$$

4. If you give a man a fish, he'll eat for a day. If you teach a man to fish, he'll eat for the rest of his life.

$$(F \multimap E) \otimes (T \multimap !E)$$

5. If you flip a coin, it will come up heads or tails.

$$F \multimap H \oplus T$$

6. If you have a headache, taking ibuprofen will cure your pain.

$$H \otimes I \multimap H^{\perp}$$

Classical Linear Logic

$$\frac{}{A \vdash A}\ \mathsf{Id} \qquad \frac{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}{\Gamma_2, \Gamma_1 \vdash \Delta_2, \Delta_1}\ \mathsf{Exch} \qquad \frac{\Gamma_1 \vdash \Delta_1, A \quad \Gamma_2, A \vdash \Delta_2}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\ \mathsf{Cut}$$

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, A \,\&\, B \vdash \Delta}\ \&L \qquad \frac{\Gamma, B \vdash \Delta}{\Gamma, A \,\&\, B \vdash \Delta}\ \&L \qquad \frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \,\&\, B}\ \&R$$

$$\frac{\Gamma, A \vdash \Delta \quad \Gamma, B \vdash \Delta}{\Gamma, A \oplus B \vdash \Delta}\ \oplus L \qquad \frac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, A \oplus B}\ \oplus R \qquad \frac{\Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \oplus B}\ \oplus R$$

$$\frac{\Gamma, A, B \vdash \Delta}{\Gamma, A \otimes B \vdash \Delta}\ \otimes L \qquad \frac{\Gamma_1 \vdash \Delta_1, A \quad \Gamma_2 \vdash \Delta_2, B}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2, A \otimes B}\ \otimes R$$

$$\frac{\Gamma_1, A \vdash \Delta_1 \quad \Gamma_2, B \vdash \Delta_2}{\Gamma_1, \Gamma_2, A \parr B \vdash \Delta_1, \Delta_2}\ \parr L \qquad \frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \parr B}\ \parr R$$

$$\frac{\Gamma_1 \vdash \Delta_1, A \quad \Gamma_2, B \vdash \Delta_2}{\Gamma_1, \Gamma_2, A \multimap B \vdash \Delta_1, \Delta_2}\ \multimap L \qquad \frac{\Gamma, A \vdash \Delta, B}{\Gamma \vdash \Delta, A \multimap B}\ \multimap R$$

no $\top L$ rule $\qquad \dfrac{}{\Gamma \vdash \Delta, \top}\ \top R$

$\dfrac{}{\Gamma, 0 \vdash \Delta}\ 0L \qquad$ no $0R$ rule

$\dfrac{\Gamma \vdash \Delta}{\Gamma, 1 \vdash \Delta}\ 1L \qquad \dfrac{}{\vdash 1}\ 1R$

$\dfrac{}{\bot \vdash}\ \bot L \qquad \dfrac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \bot}\ \bot R$

$\dfrac{\Gamma \vdash \Delta, A}{\Gamma, A^{\bot} \vdash \Delta}\ {}^{\bot}L \qquad \dfrac{\Gamma, A \vdash \Delta}{\Gamma \vdash \Delta, A^{\bot}}\ {}^{\bot}R$

$\dfrac{\Gamma, !A, !A \vdash \Delta}{\Gamma, !A \vdash \Delta}\ !C \qquad \dfrac{\Gamma \vdash \Delta}{\Gamma, !A \vdash \Delta}\ !W \qquad \dfrac{\Gamma, A \vdash \Delta}{\Gamma, !A \vdash \Delta}\ !D \qquad \dfrac{!\Gamma \vdash ?\Delta, A}{!\Gamma \vdash ?\Delta, !A}\ !P$

$\dfrac{\Gamma \vdash \Delta, ?A, ?A}{\Gamma \vdash \Delta, ?A}\ ?C \qquad \dfrac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, ?A}\ ?W \qquad \dfrac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, ?A}\ ?D \qquad \dfrac{!\Gamma, A \vdash ?\Delta}{!\Gamma, ?A \vdash ?\Delta}\ ?P$

Valid Formulas in Linear Logic

$$A \equiv B := (A \multimap B) \,\&\, (B \multimap A)$$

- $A \otimes (B \oplus C) \equiv (A \otimes B) \oplus (A \otimes C)$
- $(A \oplus B) \otimes C \equiv (A \otimes C) \oplus (B \otimes C)$
- $A \parr (B \,\&\, C) \equiv (A \parr B) \,\&\, (A \parr C)$
- $(A \,\&\, B) \parr C \equiv (A \parr C) \,\&\, (B \parr C)$
- $A \multimap (B \,\&\, C) \equiv (A \multimap B) \,\&\, (A \multimap C)$
- $(A \oplus B) \multimap C \equiv (A \multimap C) \,\&\, (B \multimap C)$
- $A \otimes 0 \equiv 0$
- $A \parr \top \equiv \top$
- $!(A \,\&\, B) \equiv !A \otimes !B$
- $?(A \oplus B) \equiv ?A \parr ?B$
- $!\top \equiv 1$
- $?0 \equiv \bot$

- $(A \otimes (B \parr C)) \multimap ((A \otimes B) \parr C)$
- $!A \otimes !B \multimap !(A \otimes B)$
- $!A \oplus !B \multimap !(A \oplus B)$
- $?(A \parr B) \multimap ?A \parr ?B$
- $?(A \,\&\, B) \multimap ?A \,\&\, ?B$
- $(A \,\&\, B) \otimes C \multimap (A \otimes C) \,\&\, (B \otimes C)$
- $(A \,\&\, B) \oplus C \multimap (A \oplus C) \,\&\, (B \oplus C)$
- $(A \parr C) \oplus (B \parr C) \multimap (A \oplus B) \parr C$
- $(A \,\&\, C) \oplus (B \,\&\, C) \multimap (A \oplus B) \,\&\, C$

$$\Gamma \vdash \Delta \iff \otimes\Gamma \vdash \parr \Delta$$

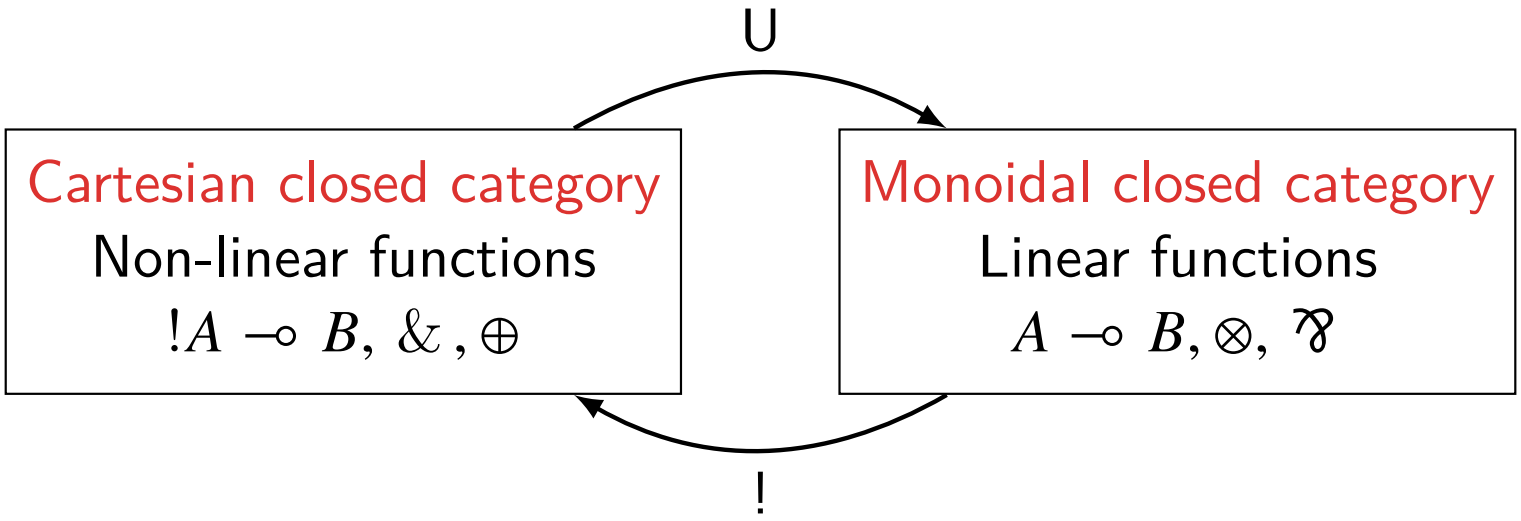
$$A \vdash B \iff \vdash A \multimap B$$

Embedding intuitionistic logic into linear logic

$$A \to B = !A \multimap B$$
$$A \wedge B = A \,\&\, B$$
$$A \vee B = !A \oplus !B$$
$$\top = \top$$
$$\bot = 0$$
$$\neg A = !A \multimap 0$$
$$\Gamma \vdash A = !\Gamma \vdash A$$

Remark: an alternative embedding:

$$A \wedge B = !A \otimes !B$$



$$A \multimap B \equiv A^{\perp} \parr B \equiv (A \otimes B^{\perp})^{\perp}$$

$$A \otimes B \equiv (A \multimap B^{\perp})^{\perp}$$

$$A \parr B \equiv (A^{\perp} \otimes B^{\perp})^{\perp} \equiv A^{\perp} \multimap B$$

$$A \,\&\, B \equiv (A^{\perp} \oplus B^{\perp})^{\perp}$$

$$!(A \,\&\, B) \equiv !A \otimes !B$$

$$?(A \oplus B) \equiv ?A \parr ?B$$

- $\&$: product
- $\oplus$: coproduct
- $\otimes$: tensor product

	$\otimes$	$\parr$	$\multimap$
$\otimes$		$A \otimes B \equiv (A^{\perp} \parr B^{\perp})^{\perp}$	$A \otimes B \equiv (A \multimap B^{\perp})^{\perp}$
$\parr$	$A \parr B \equiv (A^{\perp} \otimes B^{\perp})^{\perp}$		$A \parr B \equiv A^{\perp} \multimap B$
$\multimap$	$A \multimap B \equiv (A \otimes B^{\perp})^{\perp}$	$A \multimap B \equiv A^{\perp} \parr B$	

$\dfrac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}$	$\dfrac{f : \Gamma \otimes A \otimes B \otimes \Delta \longrightarrow C}{f \circ (1_\Gamma \otimes \sigma_{B,A} \otimes 1_\Delta) : \Gamma \otimes B \otimes A \otimes \Delta \longrightarrow C}$
$\dfrac{}{A \vdash A}$	$\dfrac{}{1_A : A \longrightarrow A}$
$\dfrac{\Gamma \vdash A \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}$	$\dfrac{f : \Gamma \longrightarrow A \quad g : A \otimes \Delta \longrightarrow B}{g \circ (f \otimes 1_\Delta) : \Gamma \otimes \Delta \longrightarrow B}$
$\dfrac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \otimes B}$ $\dfrac{\Gamma, A, B \vdash C}{\Gamma, A \otimes B \vdash C}$	$\dfrac{f : \Gamma \longrightarrow A \quad g : \Delta \longrightarrow B}{f \otimes g : \Gamma \otimes \Delta \longrightarrow A \otimes B}$ $\dfrac{f : (\Gamma \otimes A) \otimes B \longrightarrow C}{f \circ \alpha_{\Gamma,A,B} : \Gamma \otimes (A \otimes B) \longrightarrow C}$
$\dfrac{\Gamma, A \vdash B}{\Gamma \vdash A \multimap B}$ $\dfrac{\Gamma \vdash A \multimap B \quad \Delta \vdash A}{\Gamma, \Delta \vdash B}$	$\dfrac{f : \Gamma \otimes A \longrightarrow B}{\hat{f} : \Gamma \longrightarrow (A \multimap B)}$ $\dfrac{f : \Gamma \longrightarrow (A \multimap B) \quad g : \Delta \longrightarrow A}{\varepsilon_{A,B} \circ (f \otimes g) : \Gamma \otimes \Delta \longrightarrow B}$

Dagger Category

Besides the duals for objects

$$* : \mathbf{C} \to \mathbf{C}^{\mathrm{op}}$$

we have duals for morphisms:

$$\dagger : \mathbf{C} \to \mathbf{C}^{\mathrm{op}}$$

Definition (Dagger Category)

A *dagger category* is a category $\mathbf{C}$ such that for any morphism $f : A \to B$ in $\mathbf{C}$ there is a specified morphism $f^\dagger : B \to A$ such that for all $f : A \to B$ and $g : B \to C$,

- $1_A^\dagger = 1_A$
- $(gf)^\dagger = f^\dagger g^\dagger$
- $(f^\dagger)^\dagger = f$

Example — **Bayes**

Definition: The dagger category **Bayes** is defined as follows:

- objects (X, p) are finite sets X equipped with prior probability distributions, functions $p : X \to \mathbb{R}^+$ s.t. $\sum_{x \in X} p(x) = 1$.
- morphisms $(X, p) \to (Y, q)$ are conditional probability distributions, functions $f : X \times Y \to \mathbb{R}^{\geq 0}$ s.t.

$$\forall x : \sum_{y \in Y} f(y \mid x) = 1 \quad \text{and} \quad \forall y : \sum_{x \in X} p(x) f(y \mid x) = q(y)$$

- composition is composition of probability distributions as matrices.

$$(g \circ f)(z \mid x) := \sum_{y \in Y} g(z \mid y) f(y \mid x)$$

- the dagger functor is the *Bayesian converse*, acting on $f : X \times Y \to \mathbb{R}^{\geq 0}$ to give $f^\dagger : Y \times X \to \mathbb{R}^{\geq 0}$, defined by

$$f^\dagger(x \mid y) := \frac{p(x) f(y \mid x)}{q(y)}$$

The monoidal structure implements stochastic independence $(f \otimes g)(x, y \mid a, b) = f(x \mid a) g(y \mid b)$.

Definition

A morphism $f : A \longrightarrow B$ in a dagger category is:

- the *adjoint* of $g : B \longrightarrow A$ iff $g = f^\dagger$;
- *self-adjoint* iff $f = f^\dagger$ (and $A = B$);
- *idempotent* iff $ff = f$ (and $A = B$);
- a *projection* iff it is idempotent and self-adjoint;
- *unitary* iff both $f^\dagger f = 1_A$ and $f f^\dagger = 1_B$;
- an *isometry* iff $f^\dagger f = 1_A$;
- a *partial isometry* iff $f^\dagger f$ is a projection;
- *positive* iff $f = g^\dagger g$ for some morphism $g : A \longrightarrow B$ (and $A = B$).

Dagger Symmetric Monoidal Category

Definition (Dagger Symmetric Monoidal Category)

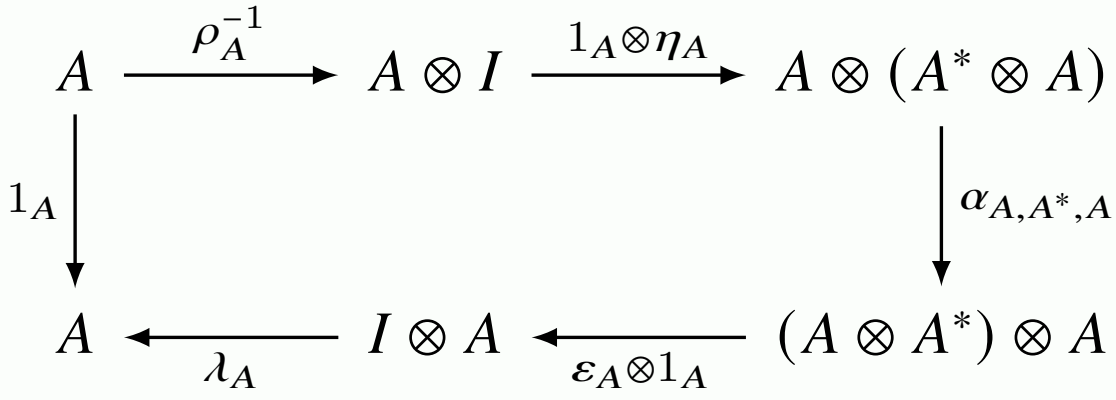
A *dagger symmetric monoidal category* is a symmetric monoidal category that is a dagger category, such that the dagger structure is compatible with the monoidal structure in the following sense:

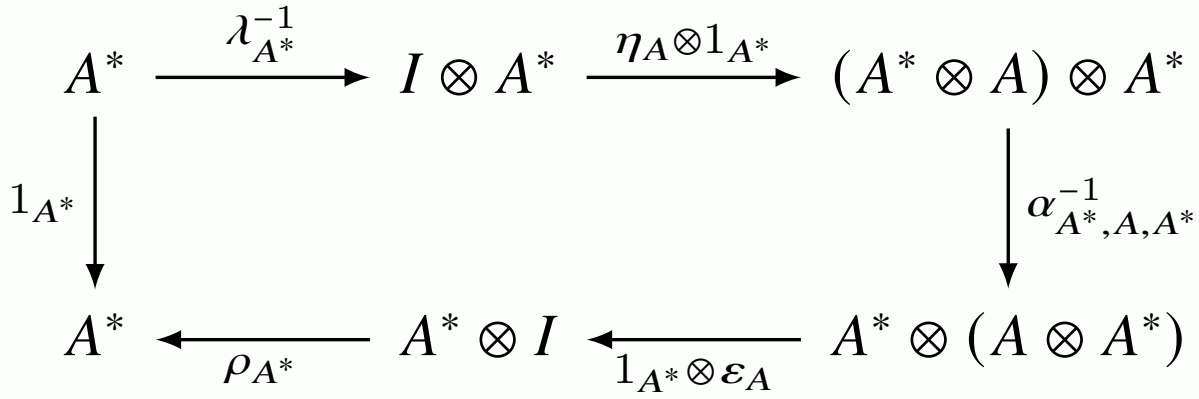
- $(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger$
- the canonical isomorphisms of the symmetric monoidal structure $\alpha, \lambda, \rho, \sigma$ are unitary.

Compact Closed Category

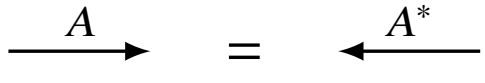
Definition (Compact Closed Category)

A *compact closed category* is a symmetric monoidal category where each object A has a dual object A^*, a unit $\eta_A : I \to A^* \otimes A$, and a counit $\varepsilon_A : A \otimes A^* \to I$ such that

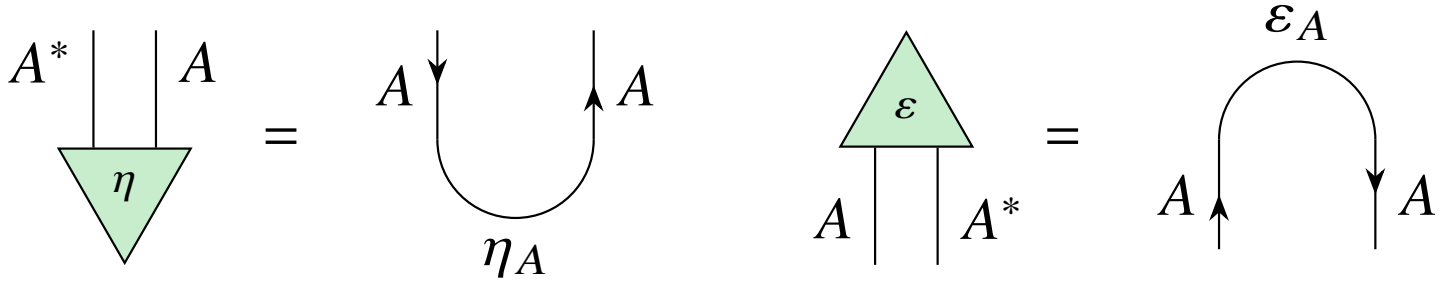




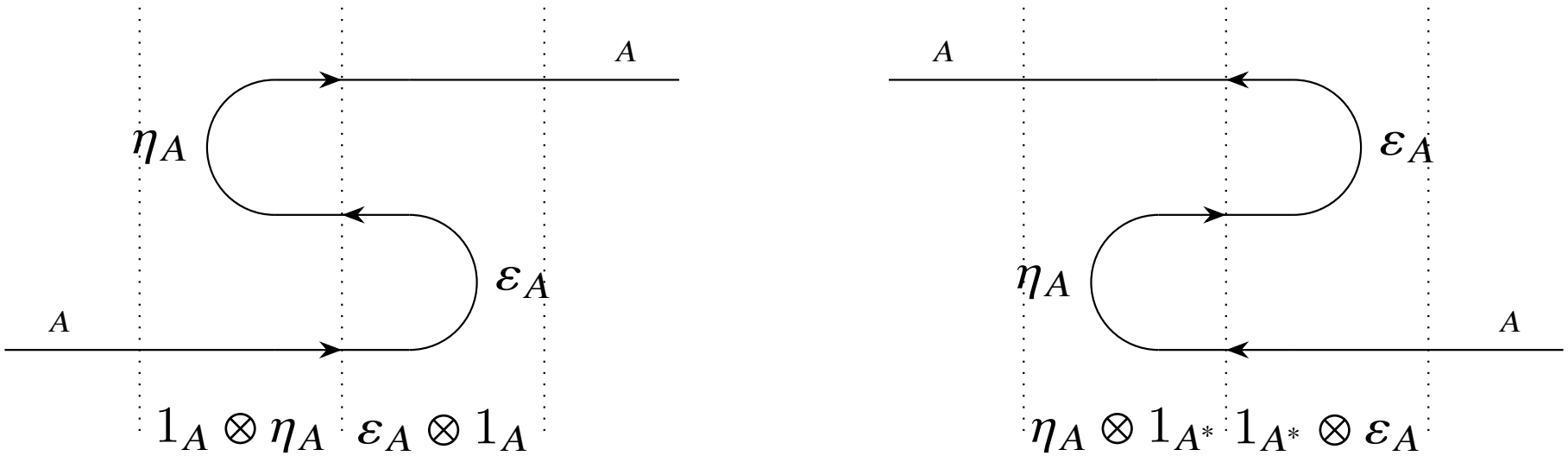
Wiring Diagram Representation

- In a compact closed category, each wire is equipped with a direction.
- A forward-pointing wire labeled A is considered equivalent to a backward-pointing wire labeled A^*, i.e. 

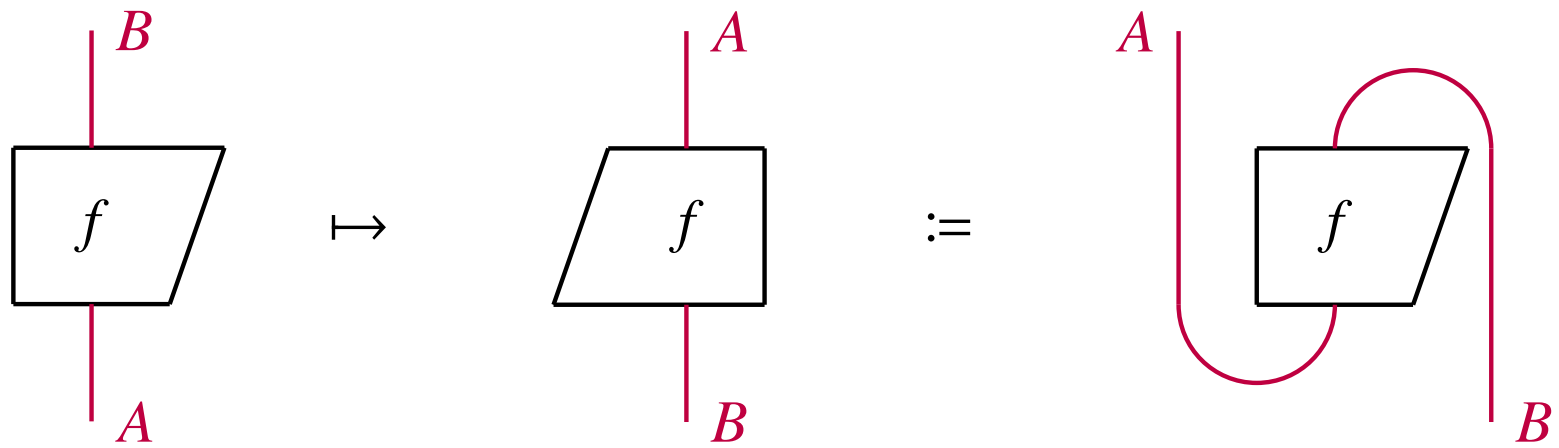
.
- The cup and cap are the unit and counit morphisms.



- In wiring diagrams, the snake equations are then drawn as follows:

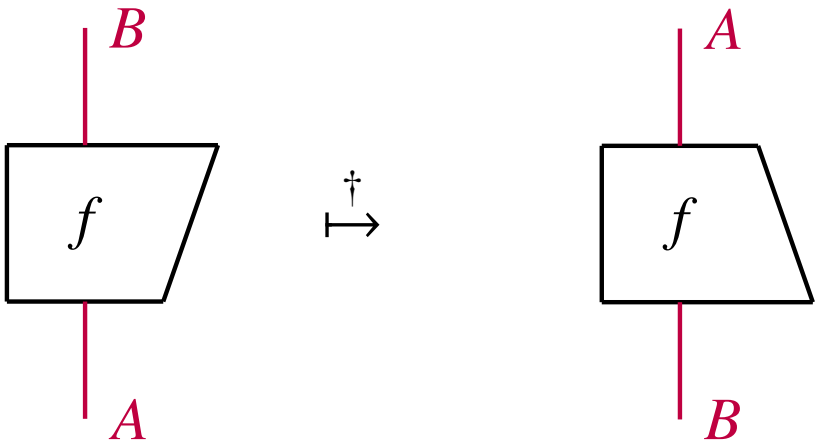


Transpose

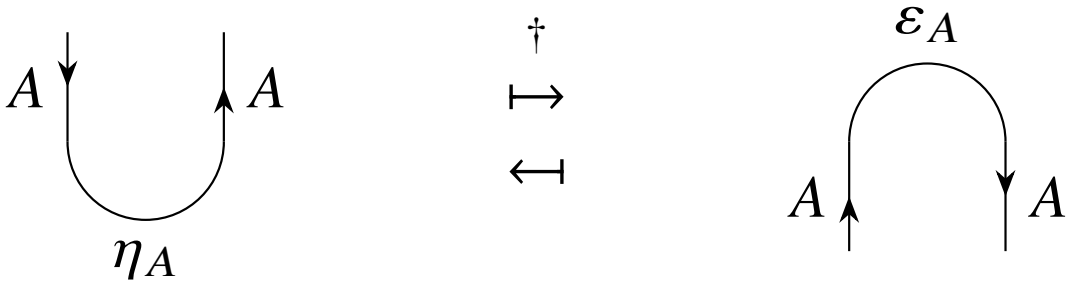


Remark: rotate $180°$

Adjoint

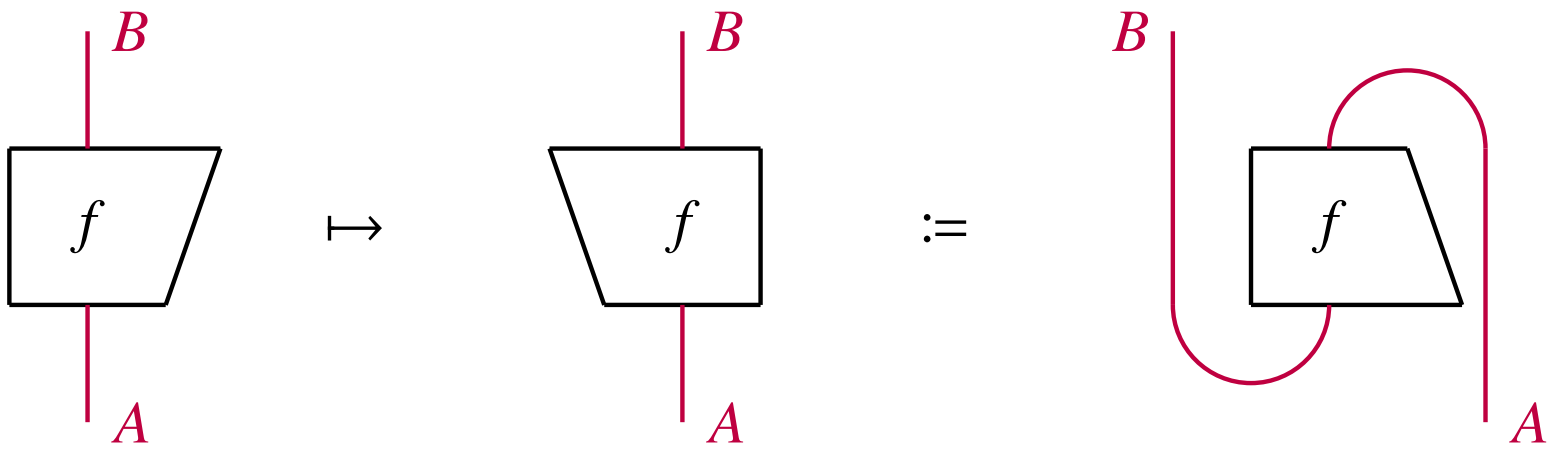


Remark: vertical reflection

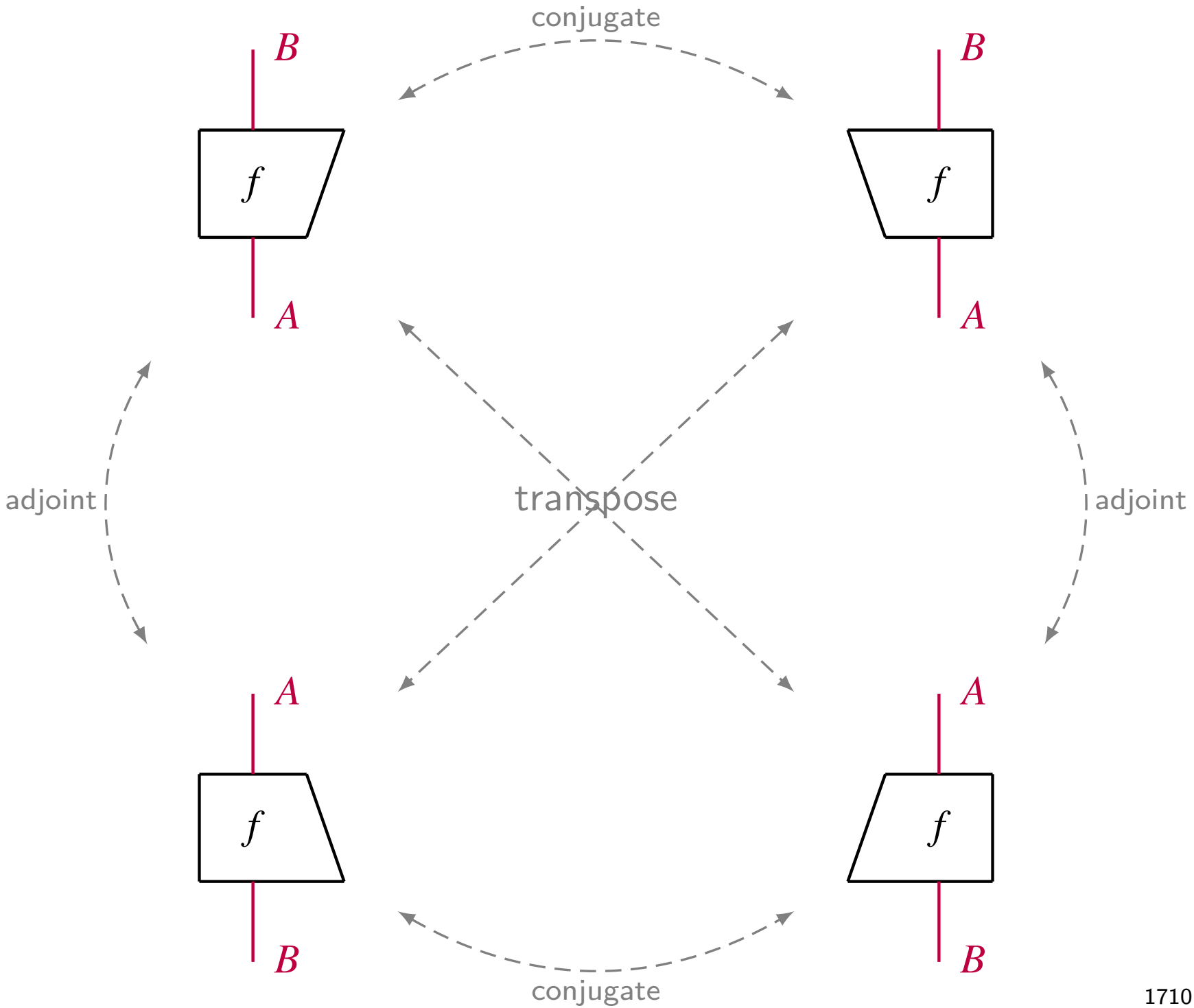


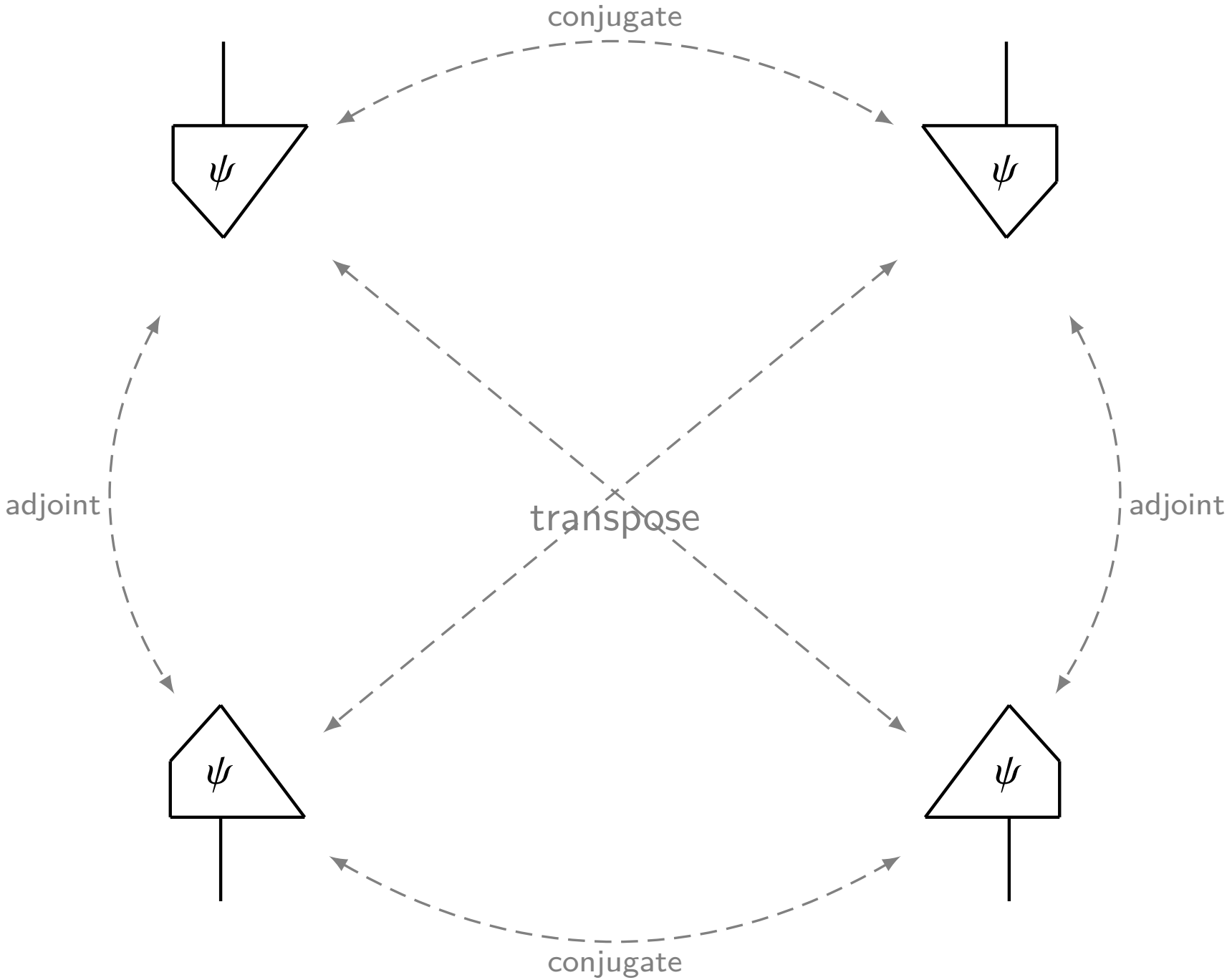
Conjugate

The conjugate of a process is the transpose of its adjoint.



Remark: horizontal reflection


conjugate
B
f
A
B
f
A
adjoint
transpose
adjoint
A
f
B
A
f
B
conjugate


conjugate
ψ
ψ
adjoint
transpose
adjoint
ψ
ψ
conjugate

Trace of an Endomorphism

Definition (Trace of an Endomorphism)

Let $(\mathbf{C}, \otimes, I, \alpha, \lambda, \rho, \sigma)$ be a symmetric monoidal category. Let $A \in \mathbf{C}$ be dualizable. The trace of $f : A \to A$ is the morphism

$$\mathrm{Tr(f)} : I \to I$$

defined by the composite

$$I \xrightarrow{\eta_A} A^* \otimes A \xrightarrow{1_{A^*} \otimes f} A^* \otimes A \xrightarrow{\sigma_{A^*,A}} A \otimes A^* \xrightarrow{\varepsilon_A} I$$

Example

- The category $(\mathbf{FdVect}_{\mathbb{K}}, \otimes, \mathbb{K})$ is compact. The unit and counit are

$$\eta_V : \mathbb{K} \to V^* \otimes V :: 1 \mapsto \sum_{i=1}^{n} \overline{e}_i \otimes e_i$$

$$\varepsilon_V : V \otimes V^* \to \mathbb{K} :: \sum_{i,j} a_{ij} e_i \otimes \overline{e}_j \mapsto \sum_{i,j} a_{ij} \langle e_i, \overline{e}_j \rangle$$

  where n is the dimension of V, $\{e_i\}_{i=1}^n$ is a basis of V, and $\overline{e}$ is the linear functional in $V^* = \mathrm{Hom}(V, \mathbb{K})$ s.t. $\overline{e}_j(e_i) = \delta_{ij}$.
- The linear maps η_V and ε_V do not depend on the choice of the basis $\{e_i\}_{i=1}^n$. Since there is a canonical isomorphism

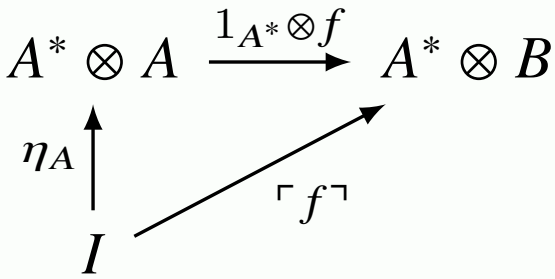
$$\mathbf{FdVect}_{\mathbb{K}}(V, V) \xrightarrow{\cong} \mathbf{FdVect}_{\mathbb{K}}(\mathbb{K}, V^* \otimes V)$$

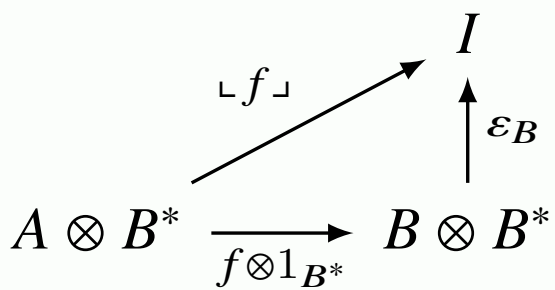
  The unit η_V is the image of 1_V under this isomorphism and 1_V is independent of the choice of basis.
- The trace of $f \in \mathrm{Hom}(V, V)$ gives the usual trace. In particular, for $f = 1_V$, it gives $\dim(V) \in \mathrm{Hom}(\mathbb{K}, \mathbb{K})$.

Name & Coname

Definition (Name & Coname)

The name $\ulcorner f \urcorner$ and the coname $\llcorner f \lrcorner$ of a morphism $f : A \to B$ in a compact closed category are





Example: For $R \in \mathbf{Rel}(X, Y)$,

$$\ulcorner R \urcorner = \{(\bullet, (x, y)) : Rxy\}$$

$$\llcorner R \lrcorner = \{((x, y), \bullet) : Rxy\}$$

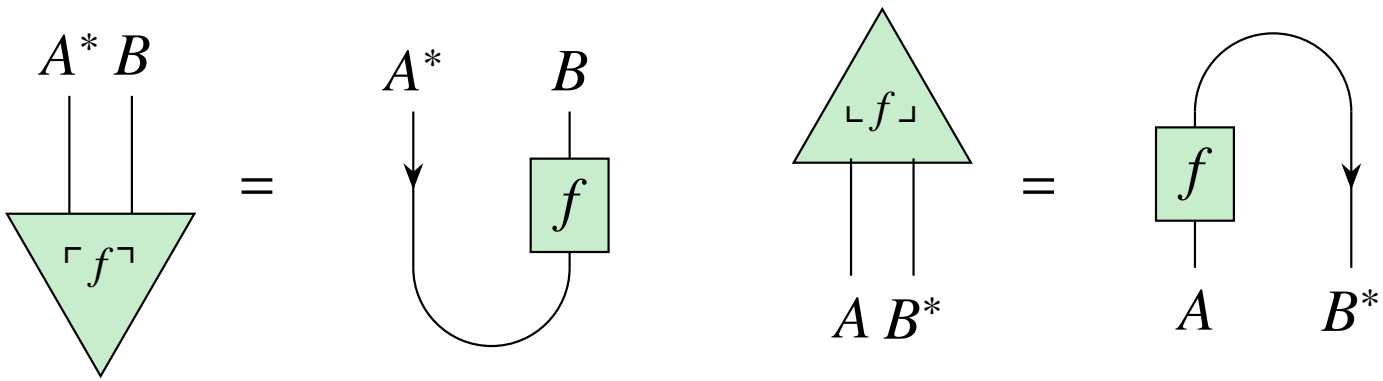
For $f \in \mathbf{FdVect}_{\mathbb{K}}(V, W)$ with (a_{ij}) the matrix of f in bases $\{e_i^V\}_{i=1}^n$ and $\{e_j^W\}_{j=1}^m$:

$$\ulcorner f \urcorner : \mathbb{K} \to V^* \otimes W :: 1 \mapsto \sum_{ij} a_{ij} \cdot \overline{e}_i^V \otimes e_j^W$$

$$\llcorner f \lrcorner : V \otimes W^* \to \mathbb{K} :: e_i^V \otimes \overline{e}_j^W \mapsto a_{ij}$$

Name & Coname





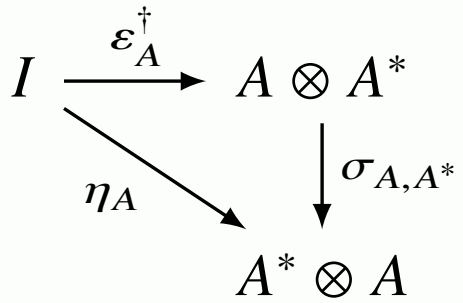
$$I \xrightarrow{\ulcorner f \urcorner} A^* \otimes B \qquad A \otimes B^* \xrightarrow{\llcorner f \lrcorner} I$$

- Every category has a set $\mathrm{Hom}(X, Y)$ of morphisms from one object X to another object Y.
- A cartesian closed category also has an object Y^X of morphisms from X to Y.
- Given $f : X \to Y$ in $\mathrm{Hom}(X, Y)$ we can convert it into its *name* $\ulcorner f \urcorner : 1 \to Y^X$ in $\mathrm{Hom}(1, Y^X)$.
- In functional programming, objects are data types, morphisms are programs, and any program $f : X \to Y$ has a name $\ulcorner f \urcorner \in \mathrm{Hom}(1, Y^X)$.

Dagger Compact Closed Category

Definition (Dagger Compact Closed Category)

A *dagger compact closed category* is a dagger symmetric monoidal category that is also compact closed, and such that:



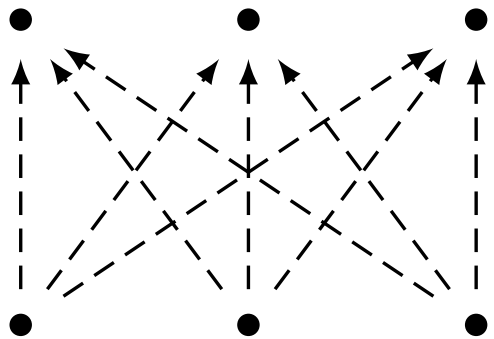
Examples

- The category $\mathbf{Rel}$ of Sets and relations. The product is the Cartesian product. The dagger is the relational converse.
- The category $\mathbf{FdHilb}$ of finite dimensional Hilbert spaces and linear maps. The morphisms are linear operators between Hilbert spaces. The product is the tensor product, and the dagger is the Hermitian conjugate.

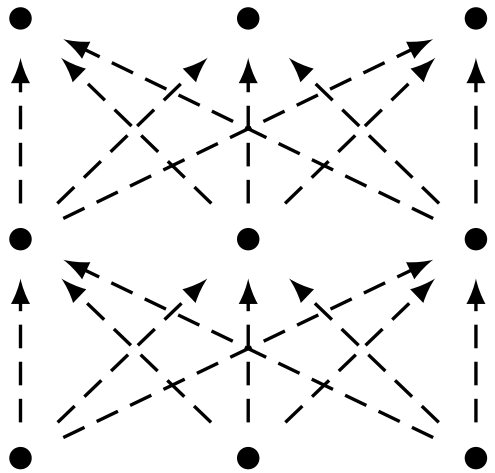
Infinite-dimensional Hilbert spaces are dagger symmetric monoidal categories, but are not dagger compact closed categories.

Remark

- Duality for objects $*$ and morphisms $\dagger$ fit together in the dagger compact closed category.
- Dagger compact closed category are deeply related to 'matrix mechanics'.



T_i^j



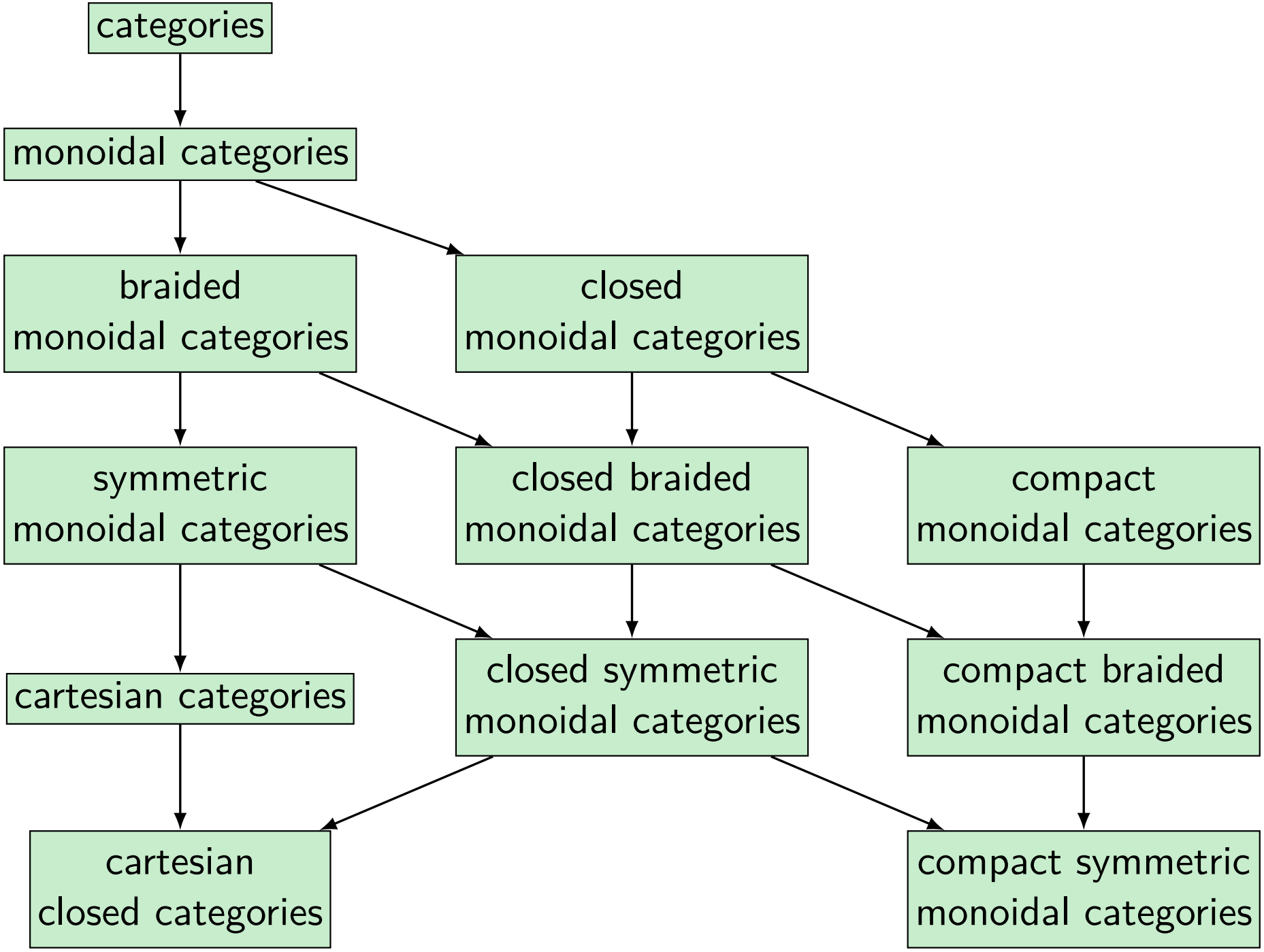
S_j^k

T_i^j

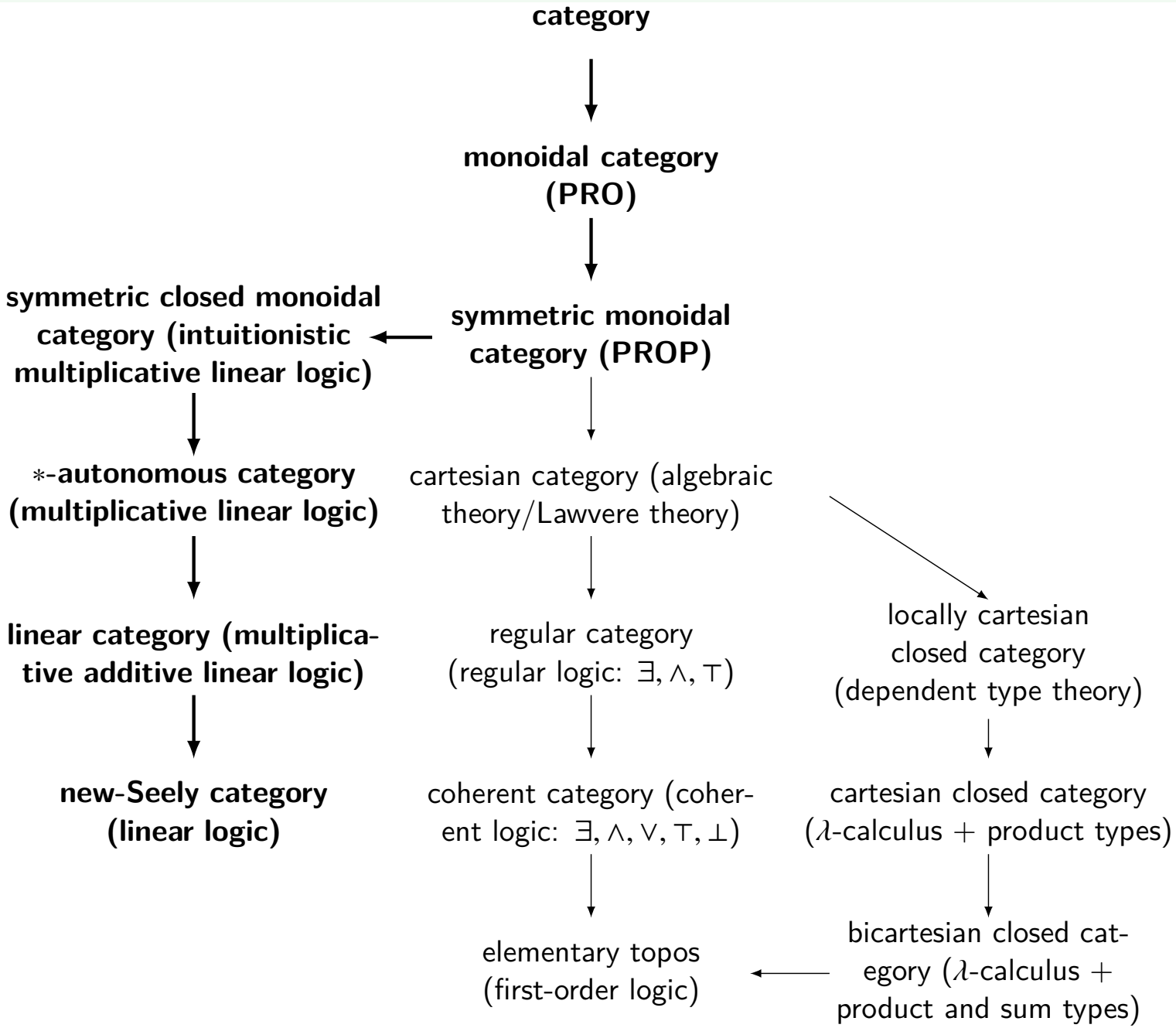
- For each input state i and output state j, the process T gives a complex number $T_i^j \in \mathbb{C}$, the amplitude to go from i to j.
- To compose processes S and T, we sum over paths:

$$(ST)_i^k = \sum_j S_j^k \times T_i^j$$

In the continuum limit, such sums become path integrals.


categories
monoidal categories
braided monoidal categories
closed monoidal categories
symmetric monoidal categories
closed braided monoidal categories
compact monoidal categories
cartesian categories
closed symmetric monoidal categories
compact braided monoidal categories
cartesian closed categories
compact symmetric monoidal categories

A family tree of categorical logic



Definition (Monoidal Functor)

A functor $F : (\mathbf{C}, \otimes_{\mathbf{C}}, I_{\mathbf{C}}) \to (\mathbf{D}, \otimes_{\mathbf{D}}, I_{\mathbf{D}})$ between monoidal categories is *monoidal* if it is equipped with:

- a morphism $\phi : I_{\mathbf{D}} \to FI_{\mathbf{C}}$
- a natural transformation $\phi_{A,B} : FA \otimes_{\mathbf{D}} FB \to F(A \otimes_{\mathbf{C}} B)$

satisfying the following conditions:

$$\begin{array}{ccc}
(FA \otimes_{\mathbf{D}} FB) \otimes_{\mathbf{D}} FC & \xrightarrow{\alpha_{FA,FB,FC}} & FA \otimes_{\mathbf{D}} (FB \otimes_{\mathbf{D}} FC) \\
\downarrow{\scriptstyle \phi_{A,B} \otimes 1_{FC}} & & \downarrow{\scriptstyle 1_{FA} \otimes \phi_{B,C}} \\
F(A \otimes_{\mathbf{C}} B) \otimes_{\mathbf{D}} FC & & FA \otimes_{\mathbf{D}} F(B \otimes_{\mathbf{C}} C) \\
\downarrow{\scriptstyle \phi_{A \otimes_{\mathbf{C}} B, C}} & & \downarrow{\scriptstyle \phi_{A, B \otimes_{\mathbf{C}} C}} \\
F((A \otimes_{\mathbf{C}} B) \otimes_{\mathbf{C}} C) & \xrightarrow[F\alpha_{A,B,C}]{} & F(A \otimes_{\mathbf{C}} (B \otimes_{\mathbf{C}} C))
\end{array}$$

$$\begin{array}{ccc}
I_{\mathbf{D}} \otimes_{\mathbf{D}} FA & \xrightarrow{\phi \otimes 1_{FA}} & FI_{\mathbf{C}} \otimes_{\mathbf{D}} FA \\
\downarrow{\scriptstyle \lambda_{FA}} & & \downarrow{\scriptstyle \phi_{I_{\mathbf{C}}, A}} \\
FA & \xleftarrow[F\lambda_A]{} & F(I_{\mathbf{C}} \otimes_{\mathbf{C}} A)
\end{array}
\qquad
\begin{array}{ccc}
FA \otimes_{\mathbf{D}} I_{\mathbf{D}} & \xrightarrow{1_{FA} \otimes \phi} & FA \otimes_{\mathbf{D}} FI_{\mathbf{C}} \\
\downarrow{\scriptstyle \rho_{FA}} & & \downarrow{\scriptstyle \phi_{A, I_{\mathbf{C}}}} \\
FA & \xleftarrow[F\rho_A]{} & F(A \otimes_{\mathbf{C}} I_{\mathbf{C}})
\end{array}$$

Monoidal Functors

- If ϕ and all $\phi_{A,B}$ are isomorphisms, then F is called a *strong monoidal functor*.
- If they are identities, then F is called a *strict monoidal functor*.

Definition (Braided Monoidal Functor)

A functor $F : \mathbf{C} \to \mathbf{D}$ between braided monoidal categories is *braided monoidal* if it is monoidal and

$$\begin{array}{ccc} FA \otimes_{\mathbf{D}} FB & \xrightarrow{\sigma_{FA,FB}} & FB \otimes_{\mathbf{D}} FA \\ \downarrow{\scriptstyle \phi_{A,B}} & & \downarrow{\scriptstyle \phi_{B,A}} \\ F(A \otimes_{\mathbf{C}} B) & \xrightarrow[F\sigma_{A,B}]{} & F(B \otimes_{\mathbf{C}} A) \end{array}$$

- A *symmetric monoidal functor* is simply a braided monoidal functor that happens to go between symmetric monoidal categories!

Lax Monoidal Functor

- A functor $F : (\mathbf{C}, \otimes_{\mathbf{C}}, I_{\mathbf{C}}) \to (\mathbf{D}, \otimes_{\mathbf{D}}, I_{\mathbf{D}})$ between monoidal categories is called lax monoidal if

$$\phi : I_{\mathbf{D}} \to F I_{\mathbf{C}}$$

$$\phi_{A,B} : FA \otimes_{\mathbf{D}} FB \to F(A \otimes_{\mathbf{C}} B)$$

  which assembles into a natural transformation.
- It's called strong monoidal if

$$I_{\mathbf{D}} \cong F I_{\mathbf{C}}$$

$$FA \otimes_{\mathbf{D}} FB \cong F(A \otimes_{\mathbf{C}} B)$$

- It's called strict monoidal if

$$I_{\mathbf{D}} = F I_{\mathbf{C}}$$

$$FA \otimes_{\mathbf{D}} FB = F(A \otimes_{\mathbf{C}} B)$$

- A strict monoidal monotone from a monoidal preorder $(X, \leq_X, \otimes_X, I_X)$ to a monoidal preorder $(Y, \leq_Y, \otimes_Y, I_Y)$ is a map $f : X \to Y$ s.t.

$$x \leq_X x \implies f(x) \leq_Y f(y)$$
$$f(x) \otimes_Y f(x') = f(x \otimes_X x')$$
$$I_Y = f(I_X)$$

- A lax monoidal monotone from a monoidal preorder $(X, \leq_X, \otimes_X, I_X)$ to a monoidal preorder $(Y, \leq_Y, \otimes_Y, I_Y)$ is a map $f : X \to Y$ s.t.

$$x \leq_X x \implies f(x) \leq_Y f(y)$$
$$f(x) \otimes_Y f(x') \leq_Y f(x \otimes_X x')$$
$$I_Y \leq_Y f(I_X)$$

- A oplax monoidal monotone from a monoidal preorder $(X, \leq_X, \otimes_X, I_X)$ to a monoidal preorder $(Y, \leq_Y, \otimes_Y, I_Y)$ is a map $f : X \to Y$ s.t.

$$x \leq_X x \implies f(x) \leq_Y f(y)$$
$$f(x) \otimes_Y f(x') \geq_Y f(x \otimes_X x')$$
$$I_Y \geq_Y f(I_X)$$

- **Example:** $\lfloor x \rfloor + \lfloor x' \rfloor \leq \lfloor x + x' \rfloor$ and $\lceil x \rceil + \lceil x' \rceil \geq \lceil x + x' \rceil$

Lax Monoidal Functor — Examples

- A monoid is a lax monoidal functor $F : (\mathbf{1}, \otimes, \mathbf{1}) \to (\mathbf{Set}, \times, \{\bullet\})$.
- A topological monoid is a monoid in the category of topological spaces.

$$(\mathbf{1}, \otimes, \mathbf{1}) \to (\mathbf{Top}, \times, *)$$

  where the monoidal unit is the one-point space $*$.
- A ring is a monoid in the category of abelian groups.

$$(\mathbf{1}, \otimes, \mathbf{1}) \to (\mathbf{AbGrp}, \otimes, \mathbb{Z})$$

- An algebra is a monoid in the category of vector spaces.

$$(\mathbf{1}, \otimes, \mathbf{1}) \to (\mathbf{FdVect}_{\mathbb{K}}, \otimes, \mathbb{K})$$

- A monad is a monoid in the category of endofunctors on $\mathbf{C}$.

$$(\mathbf{1}, \otimes, \mathbf{1}) \to (\mathbf{End}_{\mathbf{C}}, \circ, 1_{\mathbf{C}})$$

  where $\mathbf{End}_{\mathbf{C}}$ denote the category whose objects are functors $\mathbf{C} \to \mathbf{C}$ and whose morphisms are natural transformations, and the monoidal product is the composition of functors.

Theorem

A functor f with a right adjoint g is oplax monoidal iff g is a lax monoidal functor.

Corollary

- *Suppose $f : X \to Y$ is a strict monoidal monotone and $g : Y \to X$ is a right adjoint of f. Then g is a lax monoidal monotone.*
- *Suppose $g : Y \to X$ is a strict monoidal monotone and $f : X \to Y$ is a left adjoint of g. Then f is an oplax monoidal monotone.*

Monoidal, Braided Monoidal, and Symmetric Monoidal Natural Transformation

Definition (Monoidal, Braided Monoidal, and Symmetric Monoidal Natural Transformation)

Suppose that (F, ϕ) and (G, ψ) are monoidal functors from the monoidal category $\mathbf{C}$ to the monoidal category $\mathbf{D}$. Then a natural transformation $\eta : F \to G$ is *monoidal* if

$$\begin{array}{ccc} FA \otimes_{\mathbf{D}} FB & \xrightarrow{\eta_A \otimes_{\mathbf{D}} \eta_B} & GA \otimes_{\mathbf{D}} GB \\ \downarrow{\scriptstyle \phi_{A,B}} & & \downarrow{\scriptstyle \psi_{A,B}} \\ F(A \otimes_{\mathbf{C}} B) & \xrightarrow[\eta_{A \otimes_{\mathbf{C}} B}]{} & G(A \otimes_{\mathbf{C}} B) \end{array} \qquad \begin{array}{ccc} I_{\mathbf{D}} & & \\ \downarrow{\scriptstyle \phi} & \searrow^{\psi} & \\ FI_{\mathbf{C}} & \xrightarrow[\eta_{I_{\mathbf{C}}}]{} & GI_{\mathbf{C}} \end{array}$$

Definition (Monoidal Equivalence)

If $\mathbf{C}$ and $\mathbf{D}$ are (braided / symmetric) monoidal categories, a (braided / symmetric) monoidal functor $F : \mathbf{C} \to \mathbf{D}$ is an *(braided / symmetric) monoidal equivalence* if there is a (braided / symmetric) monoidal functor $G : \mathbf{D} \to \mathbf{C}$ such that there exist (braided / symmetric) monoidal natural isomorphisms $GF \cong 1_{\mathbf{C}}$ and $FG \cong 1_{\mathbf{D}}$.

Theorem (Mac Lane's Strictification Theorem)

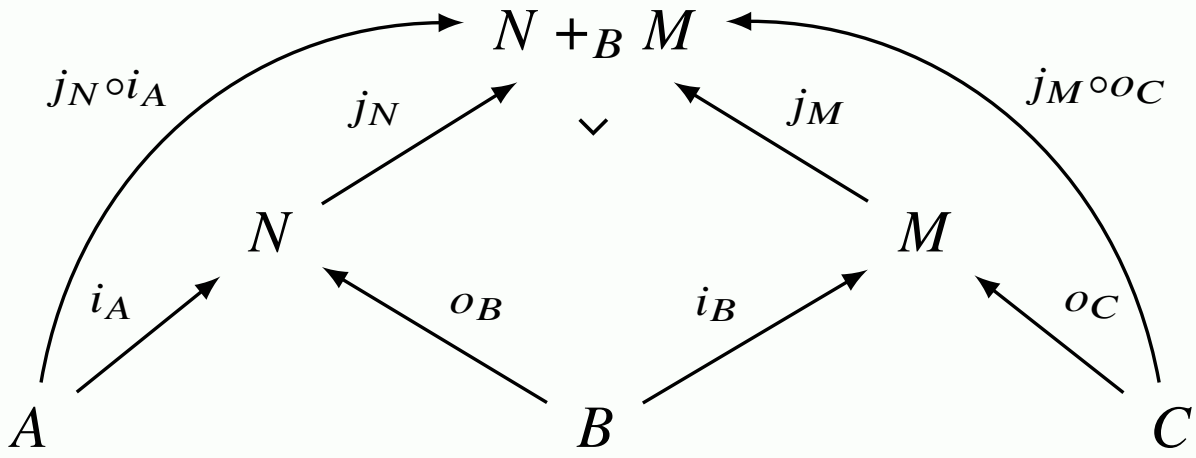
Every (braided / symmetric) monoidal category is monoidally equivalent to a strict (braided / symmetric) monoidal category.

Remark: Just like there is a 2-category $\mathbf{Cat}$ consisting of categories, functors and natural transformations, there is a 2-category $\mathbf{MonCat}$ consisting of monoidal categories, monoidal functors and monoidal transformations. Likewise, there are 2-categories $\mathbf{BrMonCat}$ and $\mathbf{SymmMonCat}$.

Cospan

Definition (Cospan)

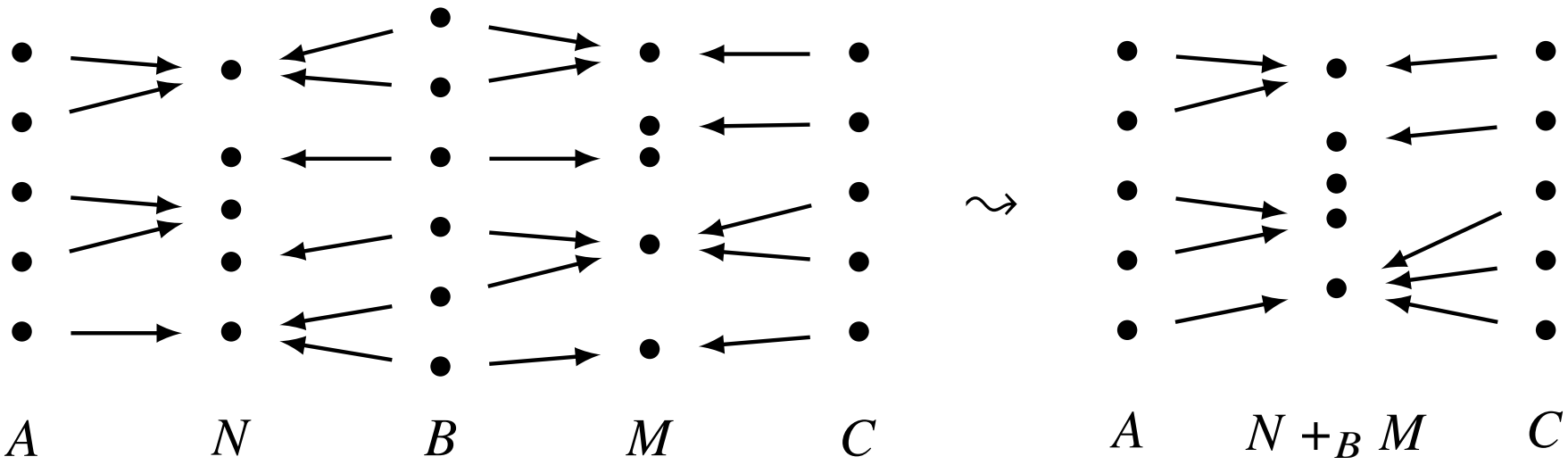
Let $\mathbf{C}$ be a category with finite colimits. Then there exists a category $\mathbf{Cospan}_{\mathbf{C}}$ with objects $\mathrm{ob}(\mathbf{Cospan}_{\mathbf{C}}) = \mathrm{ob}(\mathbf{C})$, where the morphisms $A \to B$ are the (equivalence classes of) cospans $A \to N \leftarrow B$, and composition is given by the pushout

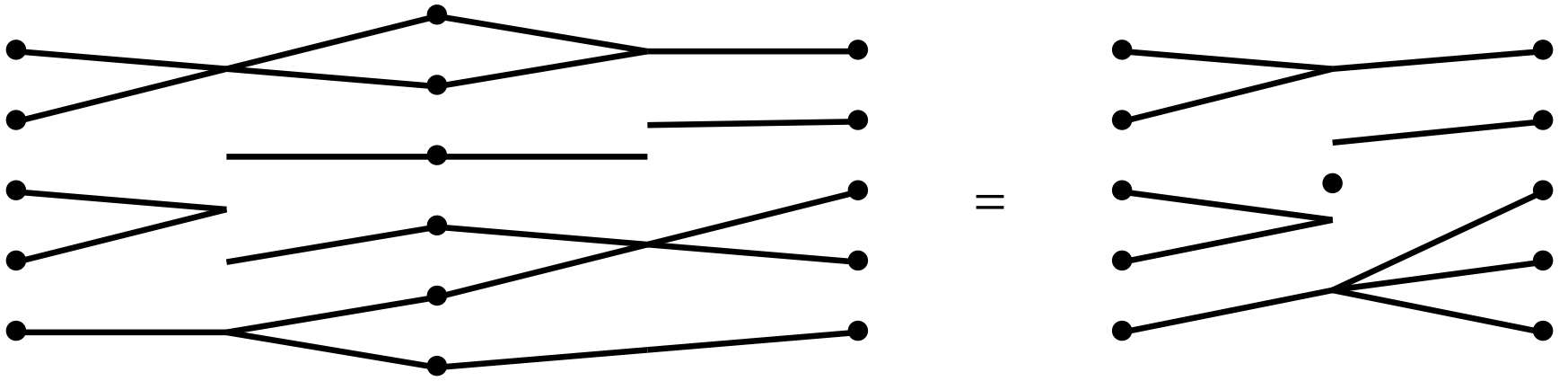


Theorem

$(\mathbf{Cospan}_{\mathbf{C}}, +, 0)$ *is a symmetric monoidal category, where the monoidal product is given by coproduct* $+$*, and the monoidal unit is the initial object* $0 \in \mathbf{C}$.

Example: $\mathbf{Cospan}_{\mathbf{FinSet}}$





Decorated Cospan

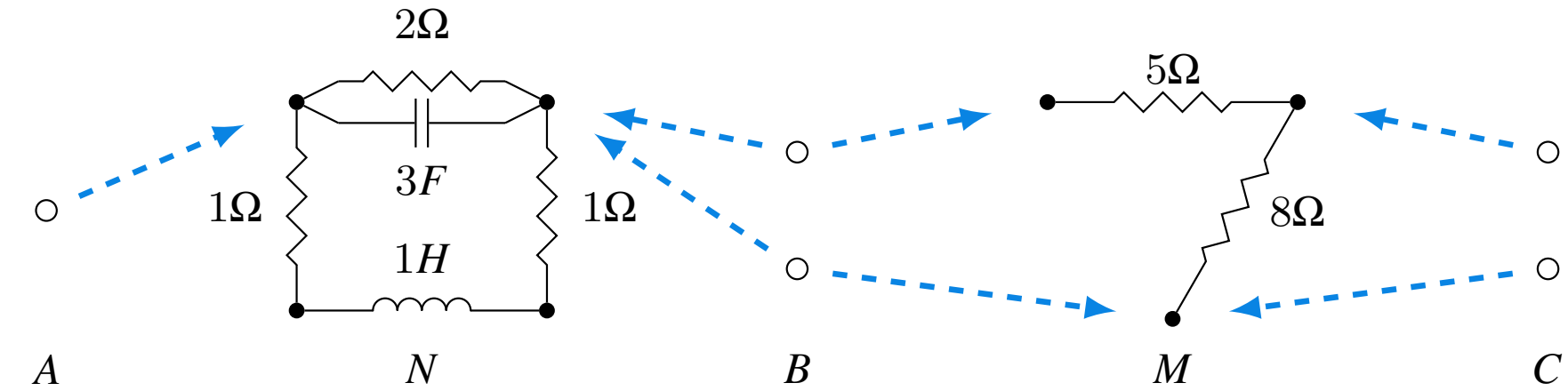
Definition (Decorated Cospan)

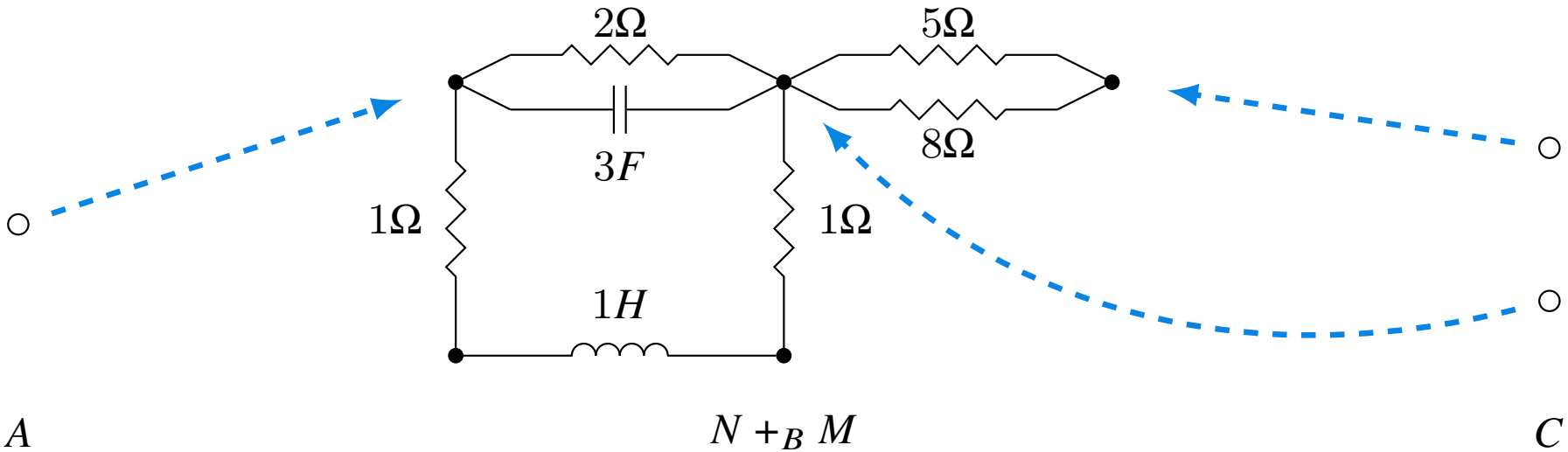
Let $\mathbf{C}$ be a category with finite colimits, and $(F, \phi) : (\mathbf{C}, +) \to (\mathbf{Set}, \times)$ be a symmetric monoidal functor. An *F-decorated cospan* is a pair consisting of a cospan $A \xrightarrow{i} N \xleftarrow{o} B$ in $\mathbf{C}$ together with an element $s \in FN$.

The composite of the decorated cospans $(A \xrightarrow{i_A} N \xleftarrow{o_B} B, s)$ and $(B \xrightarrow{i_B} M \xleftarrow{o_C} C, t)$ is given by the pushout, with decoration given by

$$1 \xrightarrow{\cong} 1 \times 1 \xrightarrow{s \times t} FN \times FM \xrightarrow{\phi_{N,M}} F(N + M) \xrightarrow{F(j_N \sqcup j_M)} F(N +_B M)$$

Example: Decorated Cospan





PROP (PROduct and Permutation categories)

Definition (prop)

A *prop* is a strict symmetric monoidal category $(\mathbf{C}, +, 0)$ for which $\mathrm{ob}(\mathbf{C}) = \mathbb{N}$, the monoidal product of objects is given by addition $+$, and the monoidal unit is $0 \in \mathbb{N}$.

- We define a morphism of props to be a strict symmetric monoidal functor that is the identity on objects $fn = n$.
- Let $\mathbf{PROP}$ be the category of props.

Definition (Model)

If $\mathbb{T}$ is a prop and $\mathbf{C}$ is a strict symmetric monoidal category, a $\mathbb{T}$-*models (or* $\mathbb{T}$-*algebras) in* $\mathbf{C}$ is a strict symmetric monoidal functor $F : \mathbb{T} \to \mathbf{C}$.

The category of $\mathbb{T}$-models (or $\mathbb{T}$-algebras) in $\mathbf{C}$, say $\mathrm{Mod}(\mathbb{T}, \mathbf{C})$, has

- symmetric monoidal functors $F : \mathbb{T} \to \mathbf{C}$ as objects,
- symmetric monoidal natural transformations as morphisms.

Remark

To specify a prop it is enough to specify five things:

1. a set $\mathbf{C}(m, n)$ of morphisms $m \to n$, for $m, n \in \mathbb{N}$.
2. for all $n \in \mathbb{N}$, an identity map $1_n : n \to n$.
3. for all $m, n \in \mathbb{N}$, a symmetry map $\sigma_{m,n} : m + n \to n + m$.
4. a composition rule: given $f : m \to n$ and $g : n \to p$, a map $(g \circ f) : m \to p$.
5. a monoidal product on morphisms: given $f : m \to m'$ and $g : n \to n'$, a map $(f + g) : m + n \to m' + n'$.

Example: $\mathbf{FinSet}$

Example

$\mathbf{FinSet}$ is a prop where the morphisms $f : m \to n$ are functions from $\{1, \ldots, m\}$ to $\{1, \ldots, n\}$. The monoidal product on functions is given by the disjoint union of functions: that is, given $f : m \to m'$ and $g : n \to n'$, we define $f + g : m + n \to m' + n'$ by

$$i \mapsto \begin{cases} f(i) & \text{if } 1 \leq i \leq m \\ m' + g(i) & \text{if } m + 1 \leq i \leq m + n \end{cases}$$

Theorem

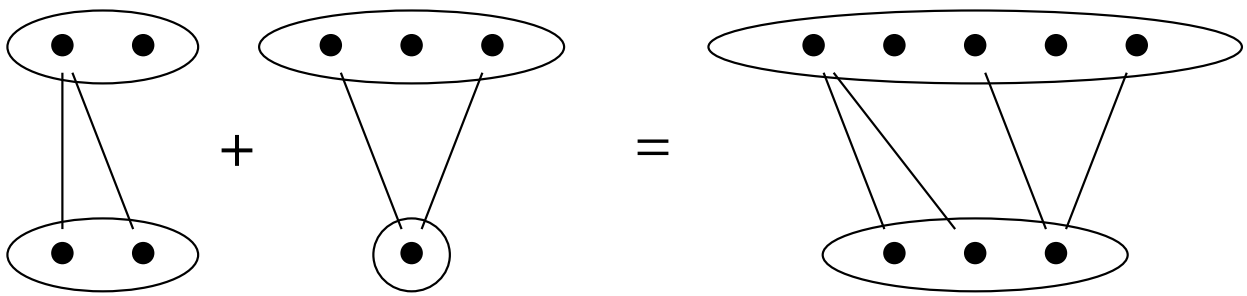
A symmetric monoidal category $\mathbf{C}$ is equivalent to a prop iff there is an object $x \in \mathbf{C}$ such that every object of $\mathbf{C}$ is isomorphic to $x^{\otimes n} = x \otimes (x \otimes (x \otimes \cdots))$ for some $n \in \mathbb{N}$.

Example: $\mathbf{Rel}$

- There is a prop $\mathbf{Rel}$ for which morphisms $m \longrightarrow n$ are relations, $R \subset [m] \times [n]$.
- The composition of R with $S \subset [n] \times [p]$ is

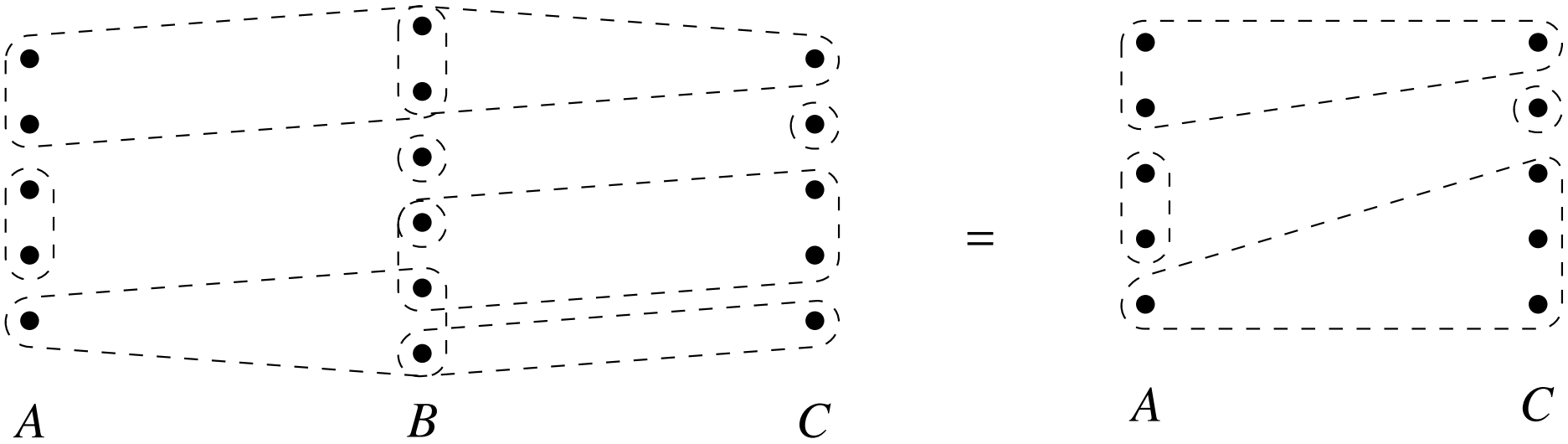
$$S \circ R := \left\{(i,k) \in [m] \times [p] : \exists j \in [n] : Rij \;\&\; Sjk\right\}$$

- The monoidal product $R_1 + R_2$ of relations $R_1 \subset [m_1] \times [n_1]$ and $R_2 \subset [m_2] \times [n_2]$ is easy.



The Category of Corelation $\mathbf{Corel}$

- Given two finite sets, A and B, a corelation $A \to B$ is an equivalence relation on $A \coprod B$.
- The objects of the category $\mathbf{Corel}$ are finite sets, and a morphism from $A \to B$ is a corelation $A \to B$.
- The composition rule is simpler to look at than to write down formally.



Remark: two elements are equivalent in the composite corelation if we may travel from one to the other staying within equivalence classes.

Example: $\mathbf{Corel}$

- The category $\mathbf{Corel}$ may be equipped with the symmetric monoidal structure $(\emptyset, \coprod)$.
- This monoidal category is compact closed, with every finite set its own dual.
- For any finite set A there is an equivalence relation on $A \coprod A : \{(a,0),(a,1) \ : \ a \in A\}$ where each part simply consists of the two elements $(a,0)$ and $(a,1)$ for each $a \in A$.
- The unit $\eta_A : \emptyset \to A \coprod A$ and the counit $\varepsilon_A : A \coprod A \to \emptyset$ are specifed by this same equivalence relation.
- The compact closed category $\mathbf{Corel}$, in which the morphisms $f : m \to n$ are partitions on $[m] \coprod [n]$, is a prop.

Example: $\mathbf{Mat}$

The category of matrix $\mathbf{Mat}$ is a prop whose

- objects are natural numbers $n \in \mathbb{N}$.
- morphisms $M : m \to n$ are $m \times n$ matrices M.
- if $m = 0$ or $n = 0$, we just assign a unique morphism $m \to n$, which we can see as a "zero-dimensional matrix".
- the identity of n is just the $n \times n$ identity matrix.
- monoidal product is direct sum

$$\frac{A_1 : m_1 \to n_1 \quad A_2 : m_2 \to n_2}{A_1 \oplus A_2 : m_1 + m_2 \to n_1 + n_2}$$

  where

$$A_1 \oplus A_2 \coloneqq \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

- symmetries are permutation matrices

Multicategory

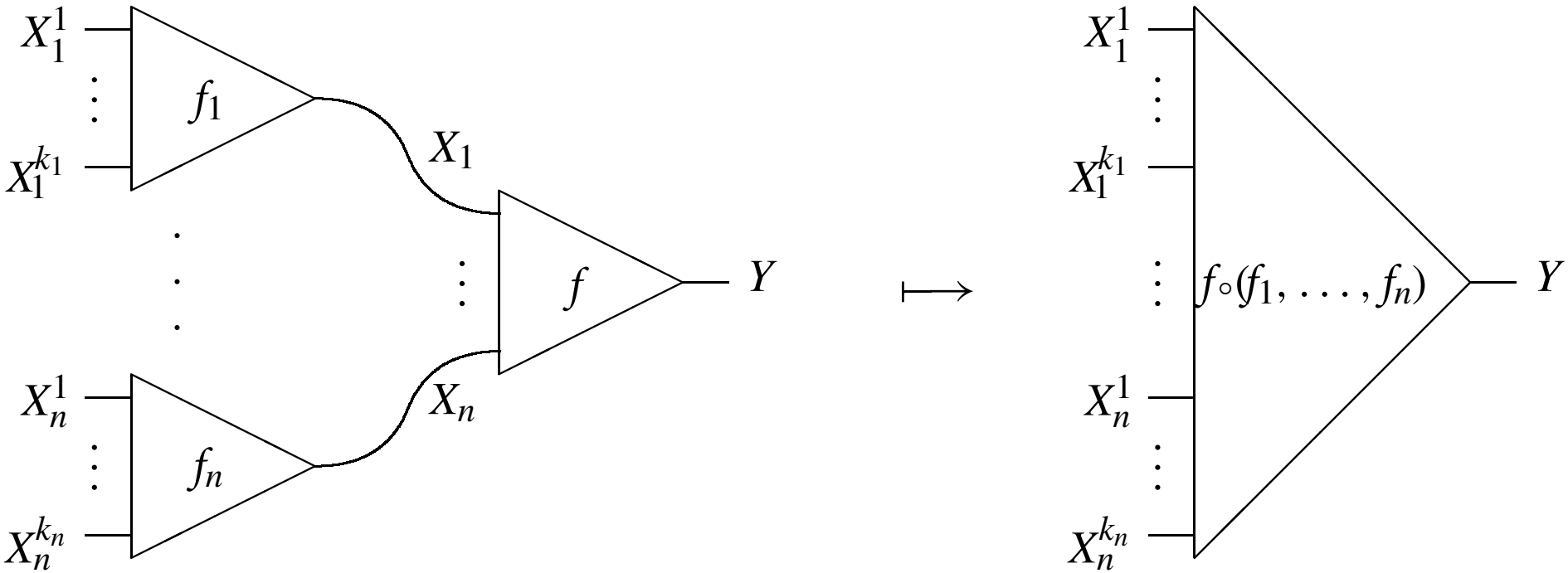
Definition (Multicategory)

A multicategory $\mathbf{C}$ consists of

- a class $\mathrm{ob}(\mathbf{C})$ of objects
- a class $\mathbf{C}(X_1, \ldots, X_n; Y)$ of multimorphisms for each $n \in \mathbb{N}$.

$$X_1, \ldots, X_n \xrightarrow{f} Y$$

- the identity $1_X : X \to X$ for $X \in \mathrm{ob}(\mathbf{C})$
- the composition $(f, f_1, \ldots, f_n) \mapsto f \circ (f_1, \ldots, f_n)$
  $\mathbf{C}(X_1, \ldots, X_n; Y) \times \mathbf{C}(X_1^1, \ldots, X_1^{k_1}; X_1) \times \cdots \times \mathbf{C}(X_n^1, \ldots, X_n^{k_n}; X_n) \to$
  $\mathbf{C}(X_1^1, \ldots, X_1^{k_1}, \ldots, X_n^1, \ldots, X_n^{k_n}; Y)$
  these must obey generalized "unital" and "associative" laws.
  $f \circ (1_{X_1}, \ldots, 1_{X_n}) = f = 1_Y \circ f$
  $f \circ \left(f_1 \circ (f_1^1, \ldots, f_1^{k_1}), \ldots, f_n \circ (f_n^1, \ldots, f_n^{k_n})\right) =$
  $\left(f \circ (f_1, \ldots, f_n)\right) \circ (f_1^1, \ldots, f_1^{k_1}, \ldots, f_n^1, \ldots, f_n^{k_n})$



- An operad is a multicategory with only one object.
- A category is a multicategory in which every arrow is unary.

Operad

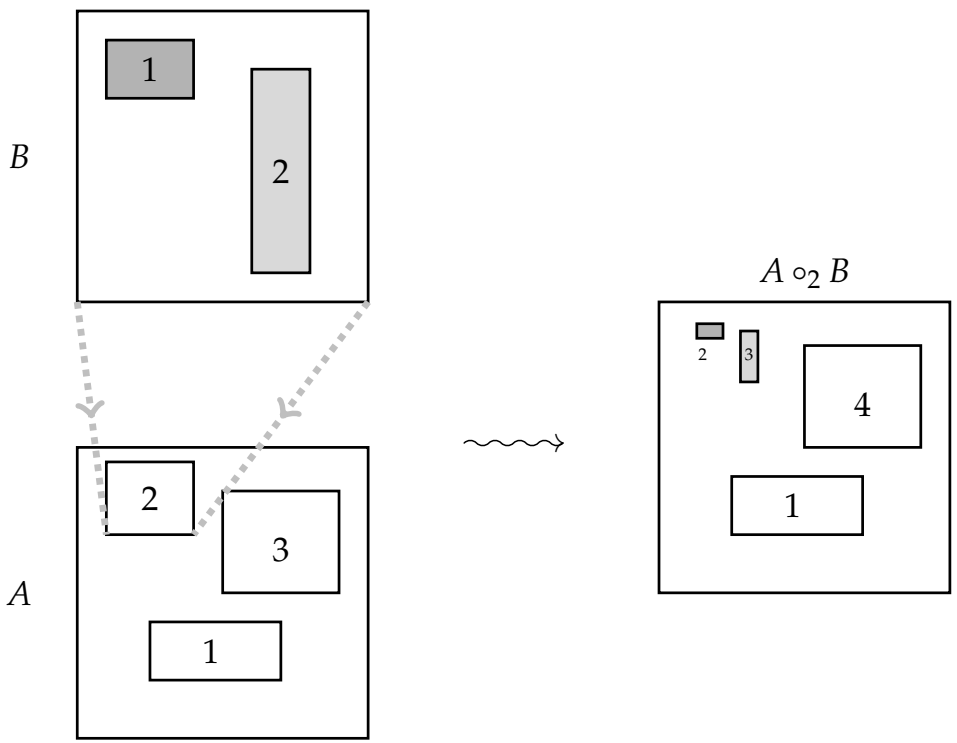
An *operad* $\mathbf{O}$ consists of

1. a class $\mathrm{ob}(\mathbf{O})$ of objects;
2. for each tuple $(t_1, \ldots, t_n, t)$, a class $\mathbf{O}(t_1, \ldots, t_n; t)$ of morphisms;
3. for $(s_1, \ldots, s_m, t_i)$ and $(t_1, \ldots, t_n, t)$, the composition

$$\circ_i : \mathbf{O}(s_1, \ldots, s_m; t_i) \times \mathbf{O}(t_1, \ldots, t_n; t) \to \mathbf{O}(t_1, \ldots, t_{i-1}, s_1, \ldots, s_m, t_{i+1}, \ldots, t_n; t)$$

4. for each t, the identity $1_t \in \mathbf{O}(t; t)$.

These must obey generalized "unital" and "associative" laws.



Multicat

- A functor of multicategories consists of $F : \mathrm{ob}(\mathbf{C}) \to \mathrm{ob}(\mathbf{D})$ such that

$$\mathbf{C}(X_1, \ldots, X_n; Y) \to \mathbf{D}(FX_1, \ldots, FX_n; FY)$$

  where composition and identities are preserved.
- The category $\mathbf{Multicat}$ consists of small multicategories and functors between them.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Category vs Enriched Category

Definition (Category)

A category $\mathbf{C}$ consists of

- a class $\mathrm{ob}(\mathbf{C})$ of objects
- a set $\mathrm{Hom}(A, B)$ for $A, B \in \mathrm{ob}(\mathbf{C})$
- $\mathrm{id}_A : \{\bullet\} \to \mathrm{Hom}(A, A)$ for $A \in \mathrm{ob}(\mathbf{C})$
- $\circ_{ABC} : \mathrm{Hom}(A, B) \times \mathrm{Hom}(B, C) \to \mathrm{Hom}(A, C)$ for $A, B, C \in \mathrm{ob}(\mathbf{C})$

all subject to associativity and identity axioms.

Definition (Enrichment in a Monoidal Category)

For a monoidal category $(V, \otimes, I, \alpha, \lambda, \rho)$, a $\mathbf{V}$-category $\mathbf{C}$ consists of

- a class $\mathrm{ob}(\mathbf{C})$ of objects
- an object $\mathrm{Hom}(A, B)$ in $\mathbf{V}$ for $A, B \in \mathrm{ob}(\mathbf{C})$
- $\mathrm{id}_A : I \to \mathrm{Hom}(A, A)$ in $\mathbf{V}$ for $A \in \mathrm{ob}(\mathbf{C})$
- $\circ_{ABC} : \mathrm{Hom}(A, B) \otimes \mathrm{Hom}(B, C) \to \mathrm{Hom}(A, C)$ in $\mathbf{V}$ for $A, B, C \in \mathrm{ob}(\mathbf{C})$

all subject to associativity and identity axioms.

Enriched Category

Definition (Enrichment in a Monoidal Category)

For a monoidal category $(V, \otimes, I, \alpha, \lambda, \rho)$, a $\mathbf{V}$-category $\mathbf{C}$ consists of

- a class $\mathrm{ob}(\mathbf{C})$ of objects.
- a hom-object $\mathbf{C}(A, B) \in \mathbf{V}$ for $A, B \in \mathrm{ob}(\mathbf{C})$.
- an arrow $\mathrm{id}_A : I \to \mathbf{C}(A, A)$ in $\mathbf{V}$ for $A \in \mathrm{ob}(\mathbf{C})$.
- an arrow $\circ_{ABC} : \mathbf{C}(A, B) \otimes \mathbf{C}(B, C) \to \mathbf{C}(A, C)$ in $\mathbf{V}$ for $A, B, C \in \mathrm{ob}(\mathbf{C})$ such that

$$
\begin{array}{ccccc}
\mathbf{C}(A,B) \otimes (\mathbf{C}(B,C) \otimes \mathbf{C}(C,D)) & & \xrightarrow{\alpha} & & (\mathbf{C}(A,B) \otimes \mathbf{C}(B,C)) \otimes \mathbf{C}(C,D) \\
\downarrow{\scriptstyle 1 \otimes \circ_{BCD}} & & & & \downarrow{\scriptstyle \circ_{ABC} \otimes 1} \\
\mathbf{C}(A,B) \otimes \mathbf{C}(B,D) & \xrightarrow[\circ_{ABD}]{} & \mathbf{C}(A,D) & \xleftarrow[\circ_{ACD}]{} & \mathbf{C}(A,C) \otimes \mathbf{C}(C,D) \\
\mathbf{C}(A,B) \otimes \mathbf{C}(B,B) & \xrightarrow{\circ_{ABB}} & \mathbf{C}(A,B) & \xleftarrow{\circ_{AAB}} & \mathbf{C}(A,A) \otimes \mathbf{C}(A,B) \\
\uparrow{\scriptstyle 1 \otimes \mathrm{id}_B} & \nearrow{\scriptstyle \rho} & & \nwarrow{\scriptstyle \lambda} & \uparrow{\scriptstyle \mathrm{id}_A \otimes 1} \\
\mathbf{C}(A,B) \otimes I & & & & I \otimes \mathbf{C}(A,B)
\end{array}
$$

Definition (Enriched Functor)

Given two $\mathbf{V}$-categories $\mathbf{C}, \mathbf{D}$, an *enriched functor* $F : \mathbf{C} \to \mathbf{D}$ assigns to each object of $\mathbf{C}$ an object of $\mathbf{D}$ and for each pair of objects A and B in $\mathbf{C}$ a morphism $F_{A,B}$ in $\mathbf{V}$

$$F_{A,B} : \mathbf{C}(A, B) \to \mathbf{D}(FA, FB)$$

such that

$$\begin{array}{ccc} \mathbf{C}(A,B) \otimes \mathbf{C}(B,C) & \xrightarrow{\circ_{A,B,C}} & \mathbf{C}(A,C) \\ \downarrow{\scriptstyle F_{A,B} \otimes F_{B,C}} & & \downarrow{\scriptstyle F_{A,C}} \\ D(FA,FB) \otimes D(FB,FC) & \xrightarrow[\circ_{FA,FB,FC}]{} & D(FA,FC) \end{array}$$

$$\begin{array}{ccc} & I & \\ {\scriptstyle \mathrm{id}_A} \swarrow & & \searrow {\scriptstyle \mathrm{id}_{FA}} \\ \mathbf{C}(A,A) & \xrightarrow[F_{A,A}]{} & \mathbf{D}(FA,FA) \end{array}$$

Examples

- A locally small category is a category enriched in $\mathbf{Set}$.
- A 2-category is a category enriched in $\mathbf{Cat}$.
- An n-category is a category enriched in $(\mathbf{n-1})$-$\mathbf{Cat}$.
- A $\mathbf{Vect}$-enriched category is a linear category: hom-sets are vector spaces and multiplication is bilinear.
- A $\mathbf{2}$-enriched category is a preorder.

$$x \leq y \iff \mathrm{Hom}(x, y) = 1$$

- Any preorder gives a $\mathbf{2}$-enriched category.

Enrichment in $\mathbf{2} = (\{0, 1\}, \leq, \wedge, 1)$	
$\mathbf{2}$-enriched category	preorder
$\mathbf{2}$-enriched functor	monotone map
$\mathbf{2}$-enriched profunctor	feasibility relation

Self-Enrichment

- Any cartesian closed category $\mathbf{V}$ can be viewed as self-enriched. This is because every external Hom-set $\mathrm{Hom}(A, B)$ can be replaced by the internal hom B^A (the object of arrows).
- In fact every monoidal closed category $\mathbf{V}$ is self-enriched.

$$\mathrm{Hom}(A \otimes B, C) \cong \mathrm{Hom}(A, [B, C])$$

- The counit of this adjunction works as the evaluation morphism:

$$\varepsilon : B \otimes [B, C] \to C$$

- The composition $\circ_{ABC} : [A, B] \otimes [B, C] \to [A, C]$ is given by the following composite

$$A \otimes ([A, B] \otimes [B, C]) \xrightarrow{\alpha} (A \otimes [A, B]) \otimes [B, C] \xrightarrow{\varepsilon \otimes 1} B \otimes [B, C] \xrightarrow{\varepsilon} C$$

- The identity $\mathrm{id}_A : I \to [A, A]$ is given by the adjunction

$$\mathrm{Hom}(I, [A, A]) \cong \mathrm{Hom}(I \otimes A, A)$$

Hom Functors

- Hom Functor

$$\mathrm{Hom}(-,-) : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Set}$$

- $\mathbf{V}$-enriched Hom functor

$$\mathrm{Hom}(-,-) : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{V}$$

- internal Hom functor $\multimap$

$$[-,-] : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{C}$$

Lawvere Metric Space

Definition (Lawvere Metric Space)

A Lawvere metric space is a set X with a function $d : X \times X \to [0, \infty]$ s.t.

1. $d(x, x) = 0$
2. $d(x, z) \leq d(x, y) + d(y, z)$

- If we define $x \leq_X y \coloneqq d(x, y) = 0$, then $\leq_X$ is a partial order.
- A Lawvere metric space can be viewed as a category enriched in the monoidal poset $\mathbf{Cost} \coloneqq ([0, \infty], \geq, +, 0)$, where the tensor product is $\otimes = +$, and the identity is $I = 0$.

$$\mathrm{Hom}(x, y) \coloneqq d(x, y) \in \mathrm{ob}(\mathbf{Cost})$$

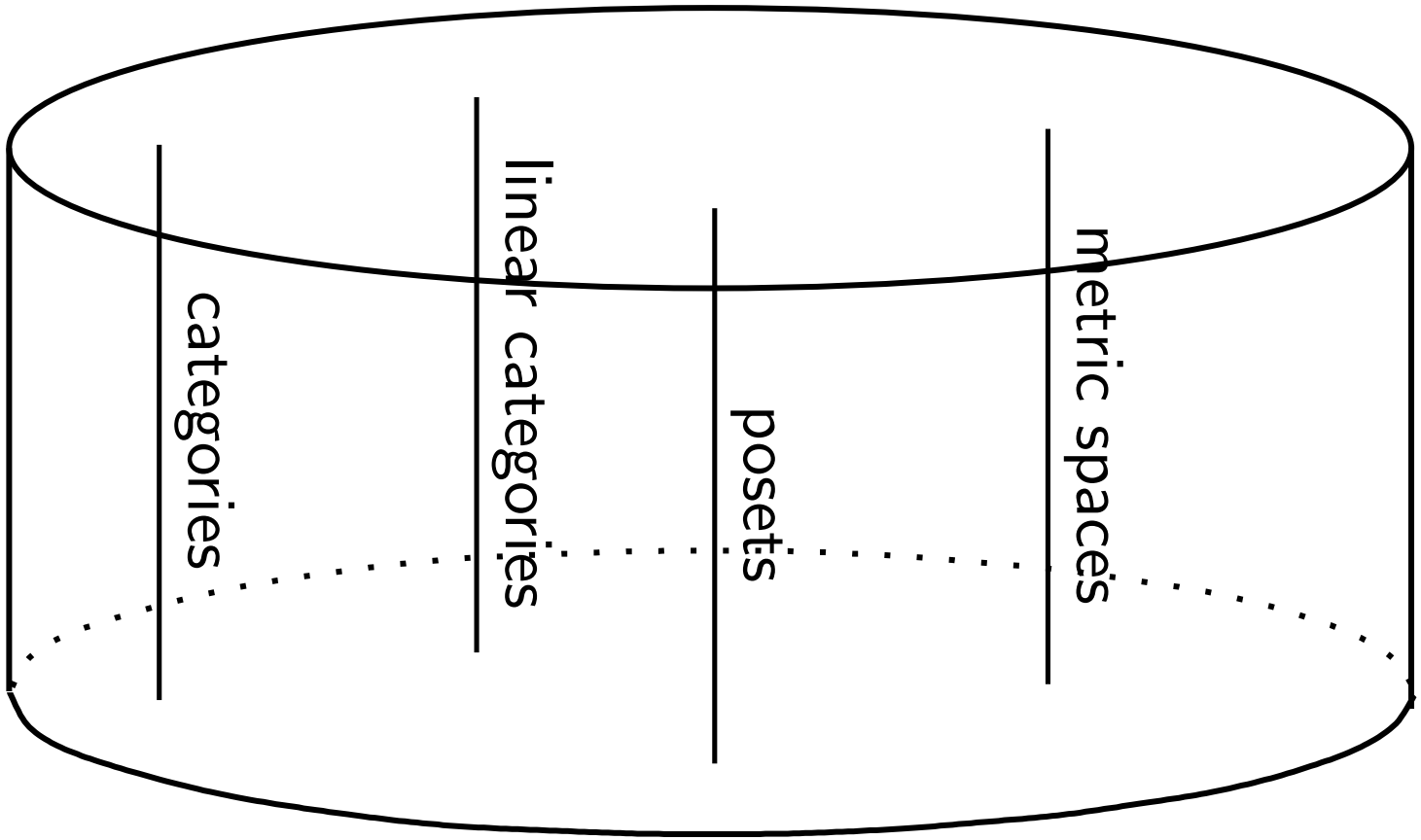
Example (I will put you in an arbitrary place of A and you have to get anywhere in B; what is the distance in the worst-case scenario?)

$$d'(A, B) \coloneqq \sup_{a \in A} \inf_{b \in B} d(a, b) \text{ for } A, B \subset X$$

then $(\mathrm{P}(X), d')$ is a Lawvere metric space.

Enriched Categories

enriched categories



monoidal categories

- $(\mathbf{Set}, \times)$
- $(\mathbf{Vect}, \otimes)$
- $([0, \infty], +)$
- $((0 \to 1), \wedge)$

Remark

Earlier	Now
$X \rightrightarrows Y$	$X \longrightarrow Y$
$\mathrm{Hom}(X, Y)$ is an object in **Vect** or **Top** or **AbGrp** ...	$\mathrm{Hom}(X, Y)$ is an "object" in $\{0, 1\}$ or $[0, 1]$ or $[0, \infty]$ ...

- 充实范畴: Hom-set 被赋予额外结构的范畴.
- 充实到 $[0, \infty]$ 上, $\mathrm{Hom}(x, y)$ 被赋予距离 $\mathrm{Hom}(x, y) = d(x, y)$.
- 充实到 $[0, 1]$ 上, $\mathrm{Hom}(x, y)$ 被赋予条件概率 $\mathrm{Hom}(x, y) = P(y \mid x)$.

$$([0, 1], \leq, \cdot, 1, \multimap)$$

where $a \multimap b := \begin{cases} 1 & \text{if } a \leq b \\ \frac{b}{a} & \text{otherwise} \end{cases}$

Yoneda Lemma for Lawvere Metric Spaces

Let $\widehat{X} = [0,\infty]^{X^{\mathrm{op}}}$ be the set of all non-expansive maps from X^{op} to $[0,\infty]$, i.e., all $f : X^{\mathrm{op}} \to [0,\infty]$ such that $[0,\infty](f(x), f(y)) \leq X^{\mathrm{op}}(x,y)$.

Theorem (Yoneda Lemma for Lawvere Metric Spaces)

For any Lawvere metric space (X,d) and any $x \in X$, let

$$X(-,x) : X^{\mathrm{op}} \to [0,\infty] :: y \to X(y,x)$$

then

$$\widehat{X}(X(-,x), f) = fx$$

Proof.

$$fx = [0,\infty](X(x,x), fx) \leq \sup_{y \in X} [0,\infty](X(y,x), fy) = \widehat{X}(X(-,x), f)$$

$$[0,\infty](fx, fy) \leq X^{\mathrm{op}}(x,y) = X(y,x) \iff [0,\infty](X(y,x), fy) \leq fx$$

□

Corollary

$$X(x,y) = \widehat{X}(X(-,x), X(-,y))$$

The $[0,1]$-Enriched Yoneda Lemma

- The interval category $[0,1]$ is a closed symmetric monoidal category.

$$([0,1], \leq, \cdot, 1, \multimap)$$

where $a \multimap b = [a,b] \coloneqq \begin{cases} 1 & \text{if } a \leq b \\ \frac{b}{a} & \text{otherwise} \end{cases}$

- If $\mathbf{C}$ is a category enriched over $[0,1]$, then the category $[0,1]^{\mathbf{C}}$ of copresheaves is also enriched over $[0,1]$.

$$\mathrm{Hom}_{[0,1]^{\mathbf{C}}}(f,g) \coloneqq \inf_{x \in \mathbf{C}}([fx, gx]) = \inf_{x \in \mathbf{C}} \left\{1, \frac{gx}{fx}\right\}$$

Theorem (The Enriched Yoneda Lemma)

For any object x in a $[0,1]$-category $\mathbf{C}$, and any $[0,1]$-copresheaf $f : \mathbf{C} \to [0,1]$, we have

$$\mathrm{Hom}_{[0,1]^{\mathbf{C}}}\left(\mathrm{Hom}_{\mathbf{C}}(x, -), f\right) = fx$$

Proof of $\mathrm{Hom}_{[0,1]^{\mathrm{C}}}(\mathrm{Hom}_{\mathbf{C}}(x,-),f) = fx$

Proof.

Fix an object $x \in \mathbf{C}$ and a copresheaf f.
Since $\mathrm{Hom}_{[0,1]^{\mathrm{C}}}(\mathrm{Hom}_{\mathbf{C}}(x,-),f) = \inf_{c\in\mathbf{C}} \{[\mathrm{Hom}_{\mathbf{C}}(x,c), fc]\}$ we have for any particular $c \in \mathbf{C}$ that $\mathrm{Hom}_{[0,1]^{\mathrm{C}}}(\mathrm{Hom}_{\mathbf{C}}(x,-),f) \leq [\mathrm{Hom}_{\mathbf{C}}(x,c), fc]$.
For $c = x$, we have

$$\mathrm{Hom}_{[0,1]^{\mathrm{C}}}(\mathrm{Hom}_{\mathbf{C}}(x,-),f) \leq [\mathrm{Hom}_{\mathbf{C}}(x,x), fx] = [1, fx] = fx$$

On the other hand, since f is a $[0,1]$-functor from $\mathbf{C}$ to $[0,1]$, we have $\mathrm{Hom}_{\mathbf{C}}(x,c) \leq [fx, fc]$ for any $c \in \mathbf{C}$.
By the closure of $[0,1]$ the inequality $\mathrm{Hom}_{\mathbf{C}}(x,c) \leq [fx, fc]$ is equivalent to $\mathrm{Hom}_{\mathbf{C}}(x,c)fx \leq fc$ which in turn is equivalent to $fx \leq [\mathrm{Hom}_{\mathbf{C}}(x,c), fc]$.
Having $fx \leq [\mathrm{Hom}_{\mathbf{C}}(x,c), fc]$ for every $c \in \mathbf{C}$ implies that

$$fx \leq \inf_{c\in\mathbf{C}} \{[\mathrm{Hom}_{\mathbf{C}}(x,c), fc]\} = \mathrm{Hom}_{[0,1]^{\mathrm{C}}}(\mathrm{Hom}_{\mathbf{C}}(x,-),f)$$

$\mathbf{V}$-Product

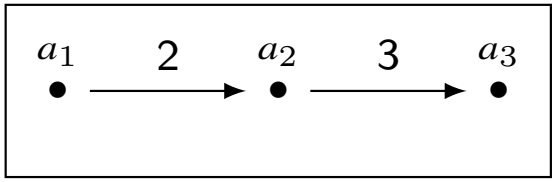
Let $\mathbf{A}$ and $\mathbf{B}$ be $\mathbf{V}$-categories. Define their $\mathbf{V}$-*product*, to be the $\mathbf{V}$-category $\mathbf{A} \times \mathbf{B}$ with

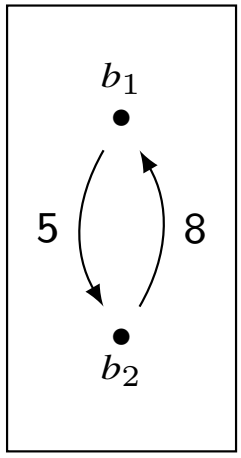
- $\mathrm{ob}(\mathbf{A} \times \mathbf{B}) := \mathrm{ob}(\mathbf{A}) \times \mathrm{ob}(\mathbf{B})$
- $(\mathbf{A} \times \mathbf{B})\big((a_1, b_1), (a_2, b_2)\big) := \mathbf{A}(a_1, a_2) \otimes \mathbf{B}(b_1, b_2)$,

for two objects (a_1, b_1) and (a_2, b_2) in $\mathrm{ob}(\mathbf{A} \times \mathbf{B})$.

Example — $\mathbf{V}$-Product

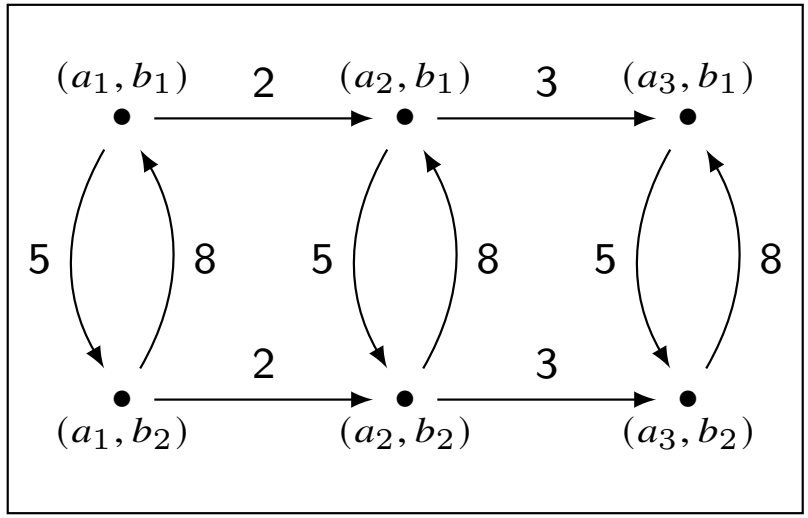
Let $\mathbf{A}$ and $\mathbf{B}$ be the Lawvere metric spaces (i.e. $\mathbf{Cost}$-categories):

$\mathbf{A} :=$ 



$=: \mathbf{B}$

The distance $d_{\mathbf{A}\times\mathbf{B}}((a_1, b_1), (a_2, b_2))$ between any two points is given by the sum $d_{\mathbf{A}}(a_1, a_2) + d_{\mathbf{B}}(b_1, b_2)$.

$\mathbf{A} \times \mathbf{B} =$ 

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Internal Category

Let $\mathbf{D}$ be a category with pullbacks. A category internal to $\mathbf{D}$ consists of

- an object of objects D_0
- an object of morphisms D_1

together with

- source and target morphisms $D_1 \underset{t}{\overset{s}{\rightrightarrows}} D_0$
- an identity-assigning morphism $D_0 \xrightarrow{\mathrm{id}} D_1$
- a composition morphism $D_1 \times_{D_0} D_1 \xrightarrow{\odot} D_1$

that are associative and unital.

Example

- A category internal to $\mathbf{Set}$ is a small category.
- A topological groupoid is a groupoid internal to $\mathbf{Top}$.
- A small Lie groupoid is a groupoid internal to the category $\mathbf{Diff}$ of smooth manifolds.
- A double category is a category internal to $\mathbf{Cat}$, i.e., a monad in $\mathbf{Span}(\mathbf{Cat})$.

$$\mathbf{D}_1 \times_{\mathbf{D}_0} \mathbf{D}_1 \xrightarrow{\odot} \mathbf{D}_1 \underset{t}{\overset{s}{\underset{\text{id}}{\rightleftarrows}}} \mathbf{D}_0$$

Category

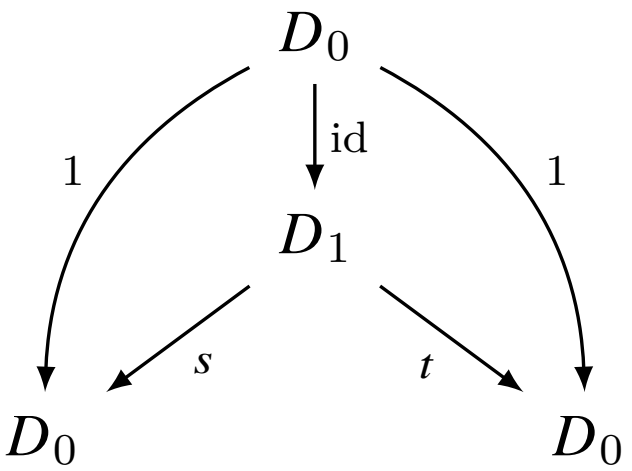
A category $\mathbf{D}$ is given by

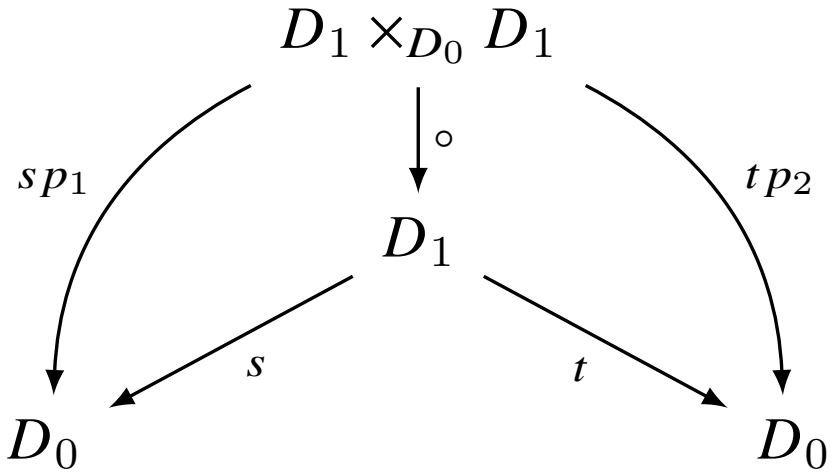
$$D_1 \underset{t}{\overset{s}{\rightrightarrows}} D_0 \quad \in \mathbf{Set}$$

with identity and composition

$$D_1 \times_{D_0} D_1 \xrightarrow{\circ} D_1 \underset{t}{\overset{s}{\underset{\leftarrow \mathrm{id} \rightarrow}{\rightrightarrows}}} D_0$$

s.t.





Double Category

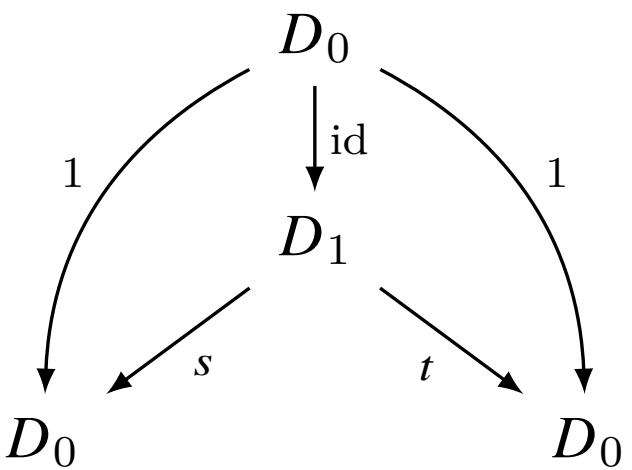
A double category $\mathbf{D}$ is given by

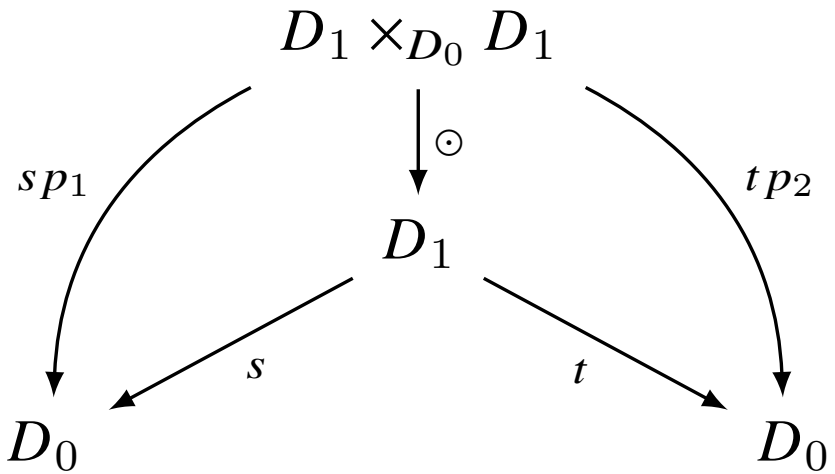
$$D_1 \underset{t}{\overset{s}{\rightrightarrows}} D_0 \quad \in \mathbf{Cat}$$

with identity and composition

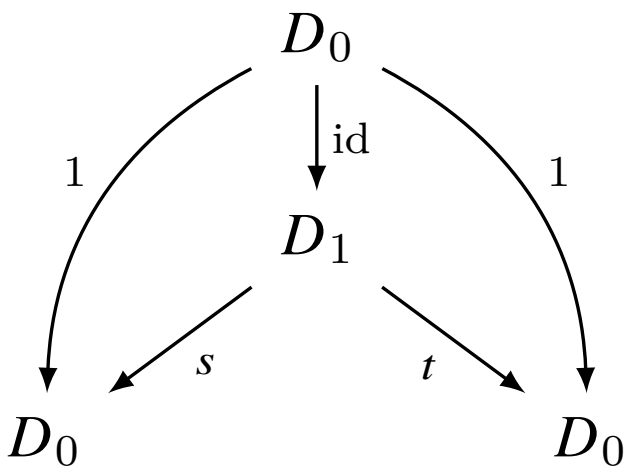
$$D_1 \times_{D_0} D_1 \xrightarrow{\odot} D_1 \underset{t}{\overset{s}{\underset{\text{id}}{\rightleftarrows}}} D_0$$

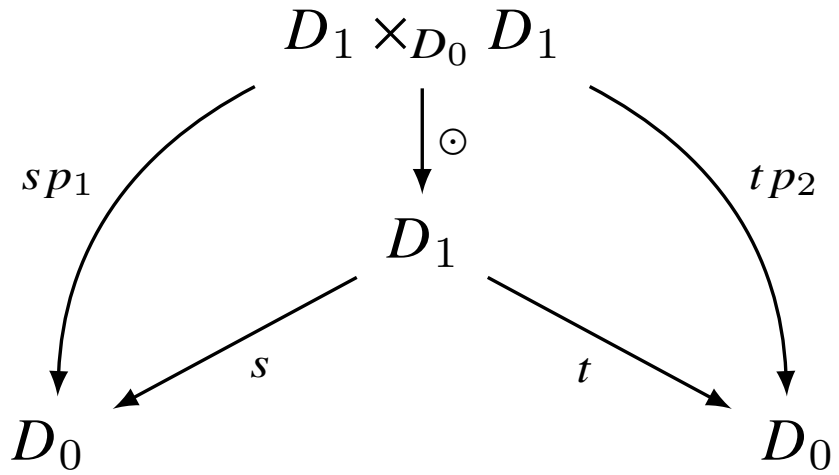
s.t.





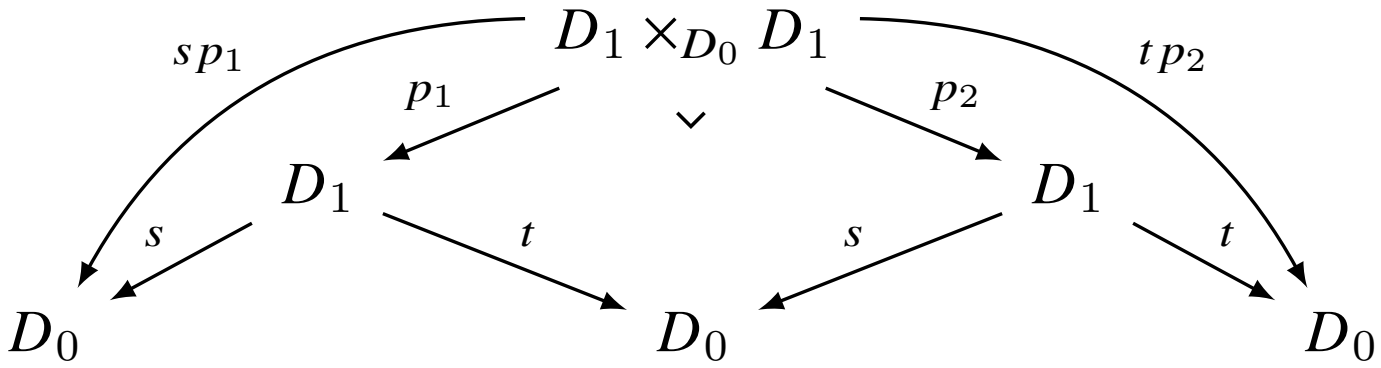
A double category is a category internal to **Cat**.





where $D_0 \xleftarrow{sp_1} D_1 \times_{D_0} D_1 \xrightarrow{tp_2} D_0$ is the horizontal composite of

$D_0 \xleftarrow{s} D_1 \xrightarrow{t} D_0$ with itself. Thus $D_1 \times_{D_0} D_1$ is the pullback.



Definition (Groups internal to a Category)

Let $\mathbf{D}$ be a category with finite products. A group in $\mathbf{D}$ consists of an object G and morphisms $G \times G \xrightarrow{m} G \xleftarrow{i} G$, $G \xleftarrow{e} 1$ such that,

1. m is associative,

$$\begin{array}{ccccc} (G \times G) \times G & & \xrightarrow{\cong} & & G \times (G \times G) \\ \downarrow{\scriptstyle m \times 1_G} & & & & \downarrow{\scriptstyle 1_G \times m} \\ G \times G & & \xrightarrow[m]{} & & G & & \xleftarrow[m]{} & & G \times G \end{array}$$

2. e is a unit,

$$\begin{array}{ccccc} 1 \times G & & \xleftarrow{\cong} & & G & & \xrightarrow{\cong} & & G \times 1 \\ \downarrow{\scriptstyle e \times 1_G} & & & & \downarrow{\scriptstyle 1_G} & & & & \downarrow{\scriptstyle 1_G \times e} \\ G \times G & & \xrightarrow[m]{} & & G & & \xleftarrow[m]{} & & G \times G \end{array}$$

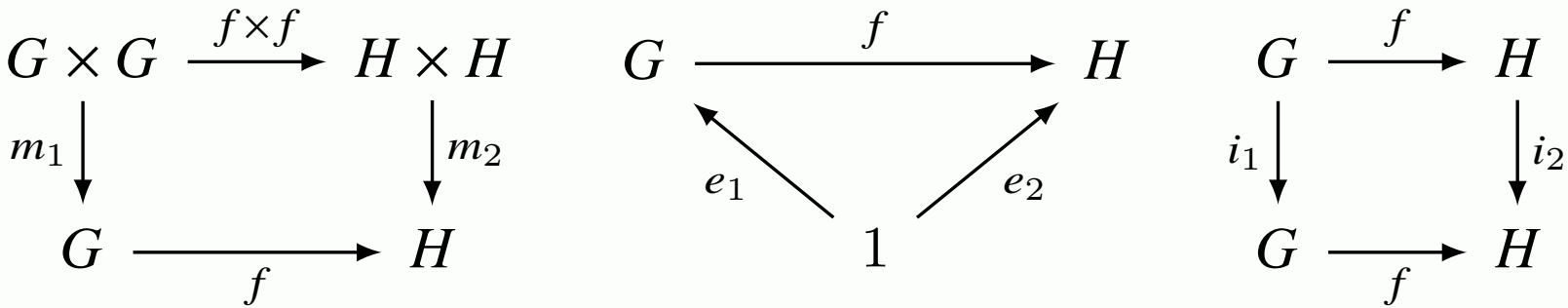
3. i is an inverse,

$$\begin{array}{ccccc} G \times G & & \xleftarrow{\Delta} & & G & & \xrightarrow{\Delta} & & G \times G \\ \downarrow{\scriptstyle 1_G \times i} & & & & \downarrow{\scriptstyle !_G} & & & & \downarrow{\scriptstyle i \times 1_G} \\ & & & & 1 & & & & \\ & & & & \downarrow{\scriptstyle e} & & & & \\ G \times G & & \xrightarrow[m]{} & & G & & \xleftarrow[m]{} & & G \times G \end{array}$$

Groups internal to a Category

Definition

A *group homomorphism* from G to H is a morphism $f : G \to H$ in $\mathbf{D}$ s.t.



Example

$$b^{(\cdot)} : (\mathbb{R}, +, 0, -) \to (\mathbb{R}^+, \cdot, 1, ^{-1})$$

$$\log_b : (\mathbb{R}^+, \cdot, 1, ^{-1}) \to (\mathbb{R}, +, 0, -)$$

Groups internal to a Category

- If $\mathbf{D} = \mathbf{Set}$, a group internal to $\mathbf{D}$ is an ordinary group.
- If $\mathbf{D} = \mathbf{Varieties}$, a group internal to $\mathbf{D}$ is an algebraic group.
- If $\mathbf{D} = \mathbf{Top}$, a group internal to $\mathbf{D}$ is a topological group.
- If $\mathbf{D} = \mathbf{Diff}$, a group internal to $\mathbf{D}$ is a Lie group.
- If $\mathbf{D} = \mathbf{Grp}$, a group internal to $\mathbf{D}$ is an abelian group.

Suppose $\circ$ is a group homomorphism $(G, \bullet) \times (G, \bullet) \to (G, \bullet)$. Then

$$(a \bullet b) \circ (c \bullet d) = (a \circ c) \bullet (b \circ d)$$

$$a \bullet b = (a \circ 1) \bullet (1 \circ b) = (a \bullet 1) \circ (1 \bullet b) = a \circ b = (1 \bullet a) \circ (b \bullet 1) = (1 \circ b) \bullet (a \circ 1) = b \bullet a$$

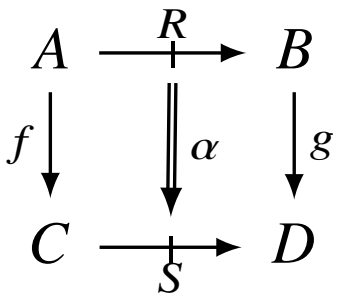
Remark

- *When we have a definition of a structure in terms of elements, we still desire a definition in terms of morphisms.*
- *Whereas a description in terms of elements is good only in one context, a description in terms of morphisms can be used in many different contexts.*

Double Category

A (strict) double category $\mathbf{D}$ consists of

- object category $\mathbf{D}_0$ (objects & arrows)
- arrow category $\mathbf{D}_1$ (proarrows & cells)



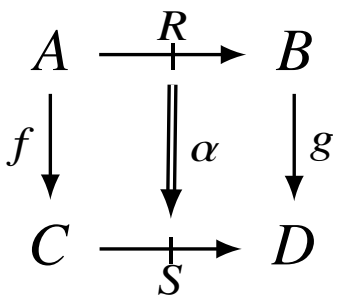
- the identity, the source and target, and the composition functors

$$\mathbf{D}_0 \xrightarrow{\text{id}} \mathbf{D}_1 \qquad \mathbf{D}_1 \underset{t}{\overset{s}{\rightrightarrows}} \mathbf{D}_0 \qquad \mathbf{D}_1 \times_{\mathbf{D}_0} \mathbf{D}_1 \xrightarrow{\odot} \mathbf{D}_1$$

that are associative and unital.

Example: Double Categories of Relations $\mathbb{R}\mathbf{el}$

- sets as objects
- functions as arrows
- relations as proarrows
- a unique cell of the form

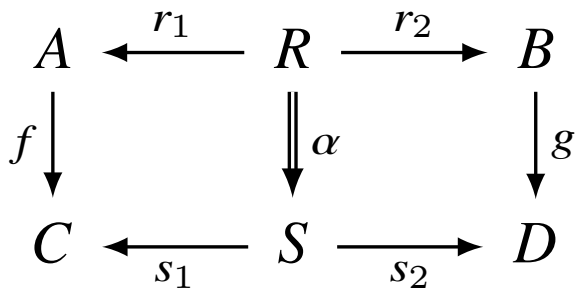


whenever the implication holds:

$$\forall a \in A \forall b \in B : R(a,b) \implies S(fa, gb)$$

Example

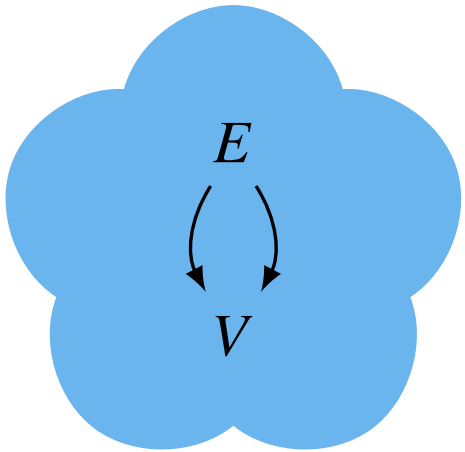
- $\mathbb{S}\mathbf{pan}(\mathbf{Cat})$: morphisms and spans

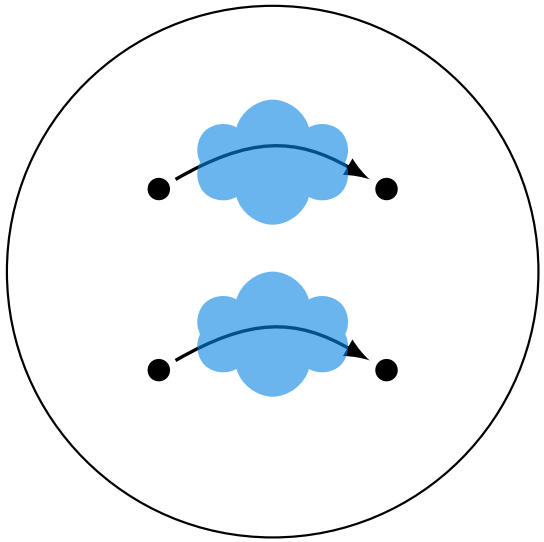


- $\mathbb{P}\mathbf{rof}$: functors and profunctors
- $\mathbb{R}\mathbf{ing}$: ring homomorphisms and bimodules

Internalisation vs Enrichment

- An internal category is defined in terms of a category with finite limits.
- An enriched category is defined in terms of a monoidal category.
- In a category internal to $\mathbf{D}$, both the set of objects and the set of arrows are replaced by objects of $\mathbf{D}$.
- In a $\mathbf{V}$-enriched category, the objects still form a set (or a proper class) while the arrows are replaced by objects of $\mathbf{V}$.
- Internal categories and enriched categories are both instances of monads in bicategories (the bicategory of spans and the bicategory of matrices, respectively).





Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

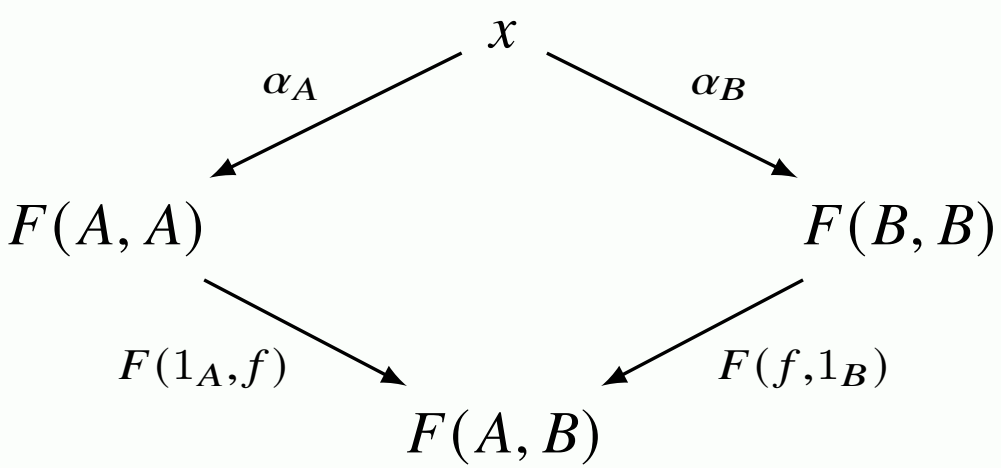
Answers to the Exercises

References2039

Wedge

Definition (Wedge)

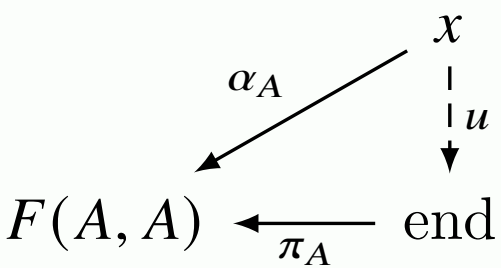
Let $F : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{X}$ be a functor. A *wedge* $\alpha : x \to F$ is an object $x \in \mathbf{X}$ with maps $\alpha_A : x \to F(A, A)$ for each $A \in \mathbf{C}$, such that given any morphism $f : A \to B$ in $\mathbf{C}$, the following diagram commutes:



End

Definition (End)

Let $F : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{X}$ be a functor. An end of F is a universal wedge, ie. a wedge $\pi : \mathrm{end} \to F$ such that, for every other wedge $\alpha : x \to F$ there is a unique $u : x \to \mathrm{end}$ such that for every $A \in \mathbf{C} : \alpha_A = \pi_A \circ u$.



We usually use following integral notation for ends

$$\mathrm{end}(F) = \int_{A \in \mathbf{C}} F(A, A) \quad \text{or just} \quad \int_{\mathbf{C}} F$$

Example

The set of natural transformations between $F, G : \mathbf{C} \to \mathbf{D}$:

$$[\mathbf{C}, \mathbf{D}](F, G) = \int_{A \in \mathbf{C}} \mathbf{D}(FA, GA)$$

Proof.

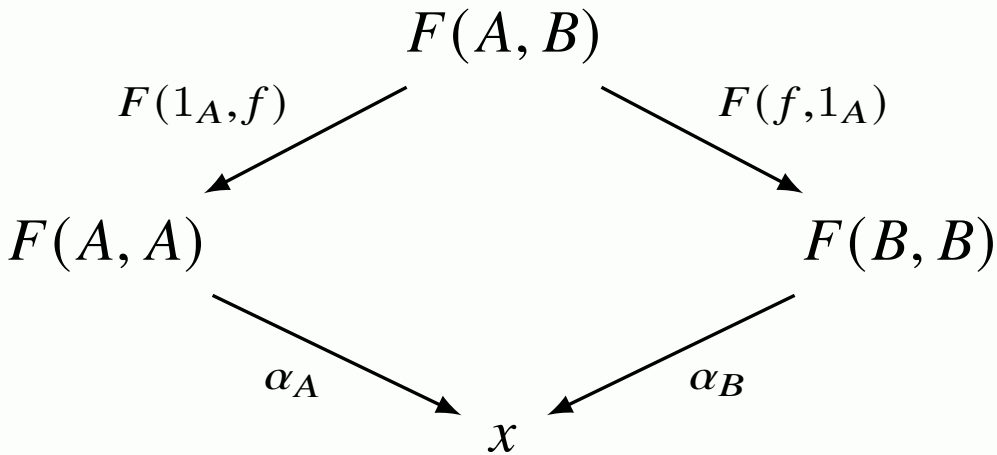
An element of $\int_{A \in \mathbf{C}} \mathbf{D}(FA, GA)$ is by definition a collection $\alpha_A : FA \to GA$ of morphisms in $\mathbf{D}$ such that for any morphism $f : A \to B$ in $\mathbf{C}$:

$$\begin{array}{ccc} FA & \xrightarrow{Ff} & FB \\ {\scriptstyle \alpha_A}\downarrow & & \downarrow{\scriptstyle \alpha_B} \\ GA & \xrightarrow[Gf]{} & GB \end{array}$$

which is by definition a natural transformation $F \to G$. □

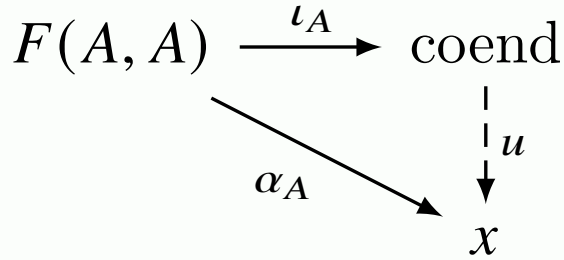
Definition (Cowedge)

Let $F : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{X}$ be a functor. A *cowedge* $\alpha : F \to x$ is an object $x \in \mathbf{X}$ with maps $\alpha_A : F(A, A) \to x$ for each $A \in \mathbf{C}$, such that given any morphism $f : B \to A$ in $\mathbf{C}$, the following diagram commutes:



Definition (Coend)

Let $F : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{X}$ be a functor. A coend of F is a universal cowedge, ie. a cowedge $\iota : F \to \mathrm{coend}$ such that, for every other cowedge $\alpha : F \to x$ there is a unique $u : \mathrm{end} \to x$ such that for every $A \in \mathbf{C} : \alpha_A = u \circ \iota_A$.



We usually write $\mathrm{coend}(F) = \int^{A \in \mathbf{C}} F(A,A)$ or just $\int^{\mathbf{C}} F$

Theorem

For functors $F : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{D}$ *and* $B \in \mathbf{D}$,

$$\mathrm{Hom}_{\mathbf{D}}\left(\int^{A \in \mathbf{C}} F(A,A), B\right) \cong \int_{A \in \mathbf{C}} \mathrm{Hom}_{\mathbf{D}}(F(A,A), B)$$

$$\mathrm{Hom}_{\mathbf{D}}\left(B, \int_{A \in \mathbf{C}} F(A,A)\right) \cong \int_{A \in \mathbf{C}} \mathrm{Hom}_{\mathbf{D}}(B, F(A,A))$$

Profunctor

Definition (Profunctor)

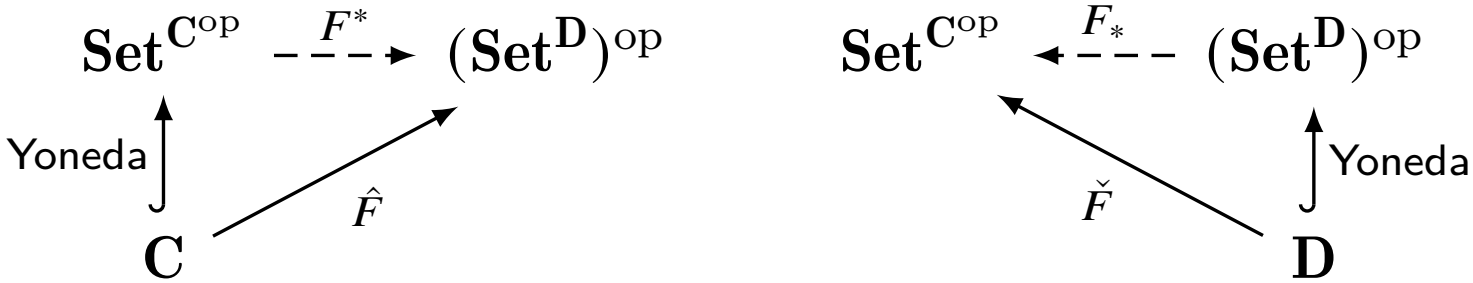
A *profunctor* $F : \mathbf{C} \nrightarrow \mathbf{D}$ is a functor $F : \mathbf{C}^{\mathrm{op}} \times \mathbf{D} \to \mathbf{Set}$.
A $\mathbf{V}$*-enriched profunctor* $F : \mathbf{C} \nrightarrow \mathbf{D}$ is a functor $F : \mathbf{C}^{\mathrm{op}} \times \mathbf{D} \to \mathbf{V}$.

Remark

- *Profunctors generalise* Hom*-functors* $\mathrm{Hom} : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Set}$.
- *Profunctors are generalised relations.*
  A relation is a $\mathbf{2}$*-enriched profunctor.*
- *Profunctors* $F : \mathbf{C} \nrightarrow \mathbf{D}$ *correspond to functors* $\hat{F} : \mathbf{C} \to (\mathbf{Set}^{\mathbf{D}})^{\mathrm{op}}$, *and* $\check{F} : \mathbf{D} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$.
- *Given* $\mathbf{D} \xrightarrow{F} \mathbf{C}$, *the composite* $\mathbf{D} \xrightarrow{F} \mathbf{C} \xrightarrow{\textit{Yoneda}} \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ *induces a profunctor* $\overline{F} : \mathbf{C} \nrightarrow \mathbf{D}$.

Remark

- The functors $\hat{F} : \mathbf{C} \to (\mathbf{Set}^{\mathbf{D}})^{\mathrm{op}}$ and $\check{F} : \mathbf{D} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ can be extended in a unique way to functors $F^* : \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}} \to (\mathbf{Set}^{\mathbf{D}})^{\mathrm{op}}$ and $F_* : (\mathbf{Set}^{\mathbf{D}})^{\mathrm{op}} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ that preserve colimits and limits, respectively.



- The functors F^* and F_* are adjoint.

$$\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}} \underset{F_*}{\overset{F^*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} (\mathbf{Set}^{\mathbf{D}})^{\mathrm{op}}$$

$$(\mathbf{Set}^{\mathbf{D}})^{\mathrm{op}}(F^* p, q) \cong \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(p, F_* q)$$

where

$$(F^* p)(B) = \int_{A \in \mathbf{C}} \mathrm{Hom}(p(A), F(A, B))$$

$$(F_* q)(A) = \int_{B \in \mathbf{D}} \mathrm{Hom}(q(B), F(A, B))$$

Nucleus

- Objects that are fixed up to isomorphism under F_*F^* or F^*F_* are called the nucleus of F, and are analogous to the left and right singular vectors of a matrix.
- We can organize the nuclei into pairs (p, q), where

$$F^*p \cong q \quad \text{and} \quad F_*q \cong p$$

- The nucleus $\{(p, q)\}$ have significant structure — they organize into a category that is complete and cocomplete.
- If the base category is $\mathbf{2}$ instead of $\mathbf{Set}$, then a profunctor R between two finite sets X and Y, viewed as discrete categories enriched over $\mathbf{2}$, is just a relation $R : X \times Y \to \{0, 1\}$.
- The function R^* maps $A \subset X$ to $R^*(A) = \{y \in Y : \forall x \in A : R(x, y) = 1\}$ and R_* maps $B \subset Y$ to $R_*(B) = \{x \in X : \forall y \in B : R(x, y) = 1\}$.
- The fixpoints of R_*R^* or R^*R_* are known as formal concepts. They are organized into pairs (A, B) with

$$R^*(A) = B \quad \text{and} \quad R_*(B) = A$$

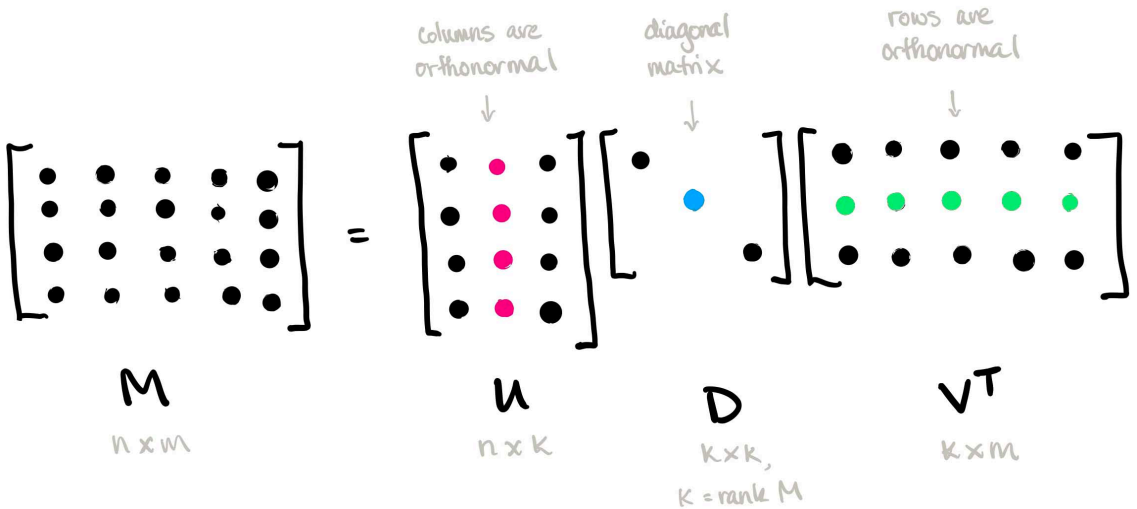
An Analogy: SVD vs Nucleus

$$M^\dagger M \quad \text{vs} \quad R_* R^*$$

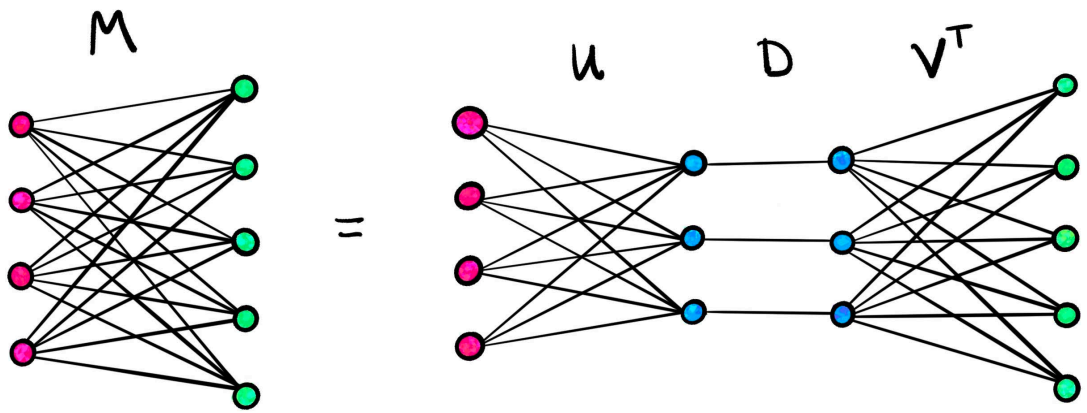
Let $|\psi\rangle \in V \otimes W$ be a unit vector and consider the orthogonal projection operator $\rho = |\psi\rangle\langle\psi|$. Then

- the reduced density operators ρ_V and ρ_W have the same spectrum, and there is a one-to-one correspondence between their eigenvectors. Moreover, if $M = V\Sigma U^\dagger$ is the singular value decomposition of the $\dim(W) \times \dim(V)$ matrix corresponding to the coefficients of $|\psi\rangle$ then
- the columns $\{u_i\}$ of U are the eigenvectors of ρ_V, the columns $\{v_i\}$ of V are the eigenvectors of ρ_W, and the nonzero diagonal entries of Σ are the singular values $\sigma_i = \sqrt{\lambda_i}$.
- $\rho_V = M^\dagger M$ and $\rho_W = M M^\dagger$
- $M u_i = \sigma_i v_i$ and $M^\dagger v_i = \sigma_i u_i$
- The original density ρ may be reconstructed from the spectral decompositions of its reduced densities ρ_V and ρ_W.

SVD

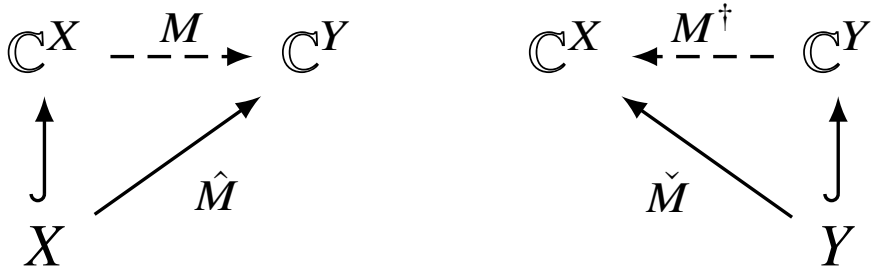


Take matrices as bipartite graphs. If we have only a few blue nodes — i.e. a few singular values — then there are fewer pathways between the pink and green nodes (i.e. people and movies).



Formal Concept Analysis — A Linear Algebraic Version

Given finite sets X and Y, any matrix $M : X \times Y \to \mathbb{C}$ induces two functions $\hat{M} : X \to \mathbb{C}^Y$ and $\check{M} : Y \to \mathbb{C}^X$ that lift to linear maps M and $M^\dagger$ s.t.



Moreover,

$$\forall u \in \mathbb{C}^X \forall v \in \mathbb{C}^Y : \langle Mu|v\rangle = \langle u|M^\dagger v\rangle$$

and the one-dimensional invariant subspaces of $M^\dagger M$ and $MM^\dagger$ correspond to their eigenvectors.

$\left\{(u_i, v_i) \in \mathbb{C}^X \times \mathbb{C}^Y : Mu_i = \sigma_i v_i \;\;\&\;\; M^\dagger v_i = \sigma_i u_i\right\} \cong$
$\left\{u_i : M^\dagger M u_i = \lambda_i u_i\right\} \cong \left\{v_i : MM^\dagger v_i = \lambda_i v_i\right\}$

where $\{\sigma_i\}$ are the singular values of M, and $\{\lambda_i\}$ are the eigenvalues.

Formal Concept Analysis — A Quantum Version

Given finite sets X and Y, any joint probability distribution $\pi : X \times Y \to \mathbb{R}$ defines a unit vector in $\mathbb{C}^X \otimes \mathbb{C}^Y$ whose coefficients are the square roots of the probabilities

$$|\psi\rangle = \sum_{(x,y)\in X\times Y} \sqrt{\pi(x,y)}\, |x\rangle \otimes |y\rangle$$

Orthogonal projection onto this unit vector defines a density operator $\rho = |\psi\rangle\langle\psi|$ whose reduced densities ρ_X and ρ_Y have the following properties.

- The classical marginal probability distribution $\pi_X : X \to \mathbb{R}$ (or $\pi_Y : Y \to \mathbb{R}$) is contained along the diagonal of ρ_X (or ρ_Y).

$$(\rho_X)_{ii} = \sum_\alpha \pi(x_i, y_\alpha) = \pi_X(x_i) \qquad (\rho_Y)_{\alpha\alpha} = \sum_i \pi(x_i, y_\alpha) = \pi_Y(y_\alpha)$$

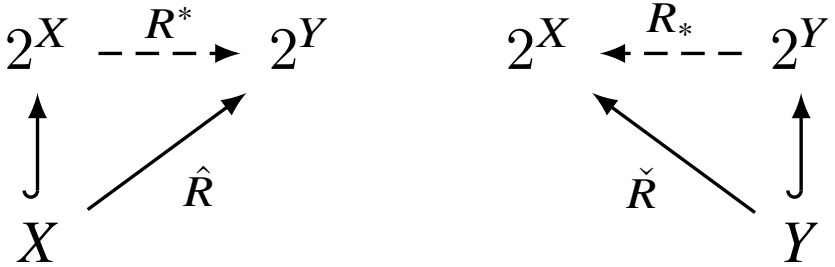
- The reduced densities ρ_X and ρ_Y generally have nonzero off-diagonal entries that encode extra information about subsystem interactions.

$$(\rho_X)_{ij} = \sum_\alpha \sqrt{\pi(x_i, y_\alpha)\pi(x_j, y_\alpha)} \qquad (\rho_Y)_{\alpha\beta} = \sum_i \sqrt{\pi(x_i, y_\alpha)\pi(x_i, y_\beta)}$$

  This information contributes to the eigenvalues and eigenvectors of ρ_X and ρ_Y, and it is akin to conditional probability.

Formal Concept Analysis — A Boolean Version

Given finite sets X and Y, any function $R : X \times Y \to \{0, 1\}$ induces two functions $\hat{R} : X \to 2^Y$ and $\check{R} : Y \to 2^X$ that lift to order-reversing functions R^* and R_* s.t.



Moreover,

$$R^*A \supset B \iff A \subset R_*B$$

and the formal concepts are the invariant subsets of the compositions R_*R^* or R^*R_*.

$$\left\{(A, B) \in 2^X \times 2^Y : R^*A = B \;\&\; R_*B = A\right\} \cong \mathrm{Fix}(R_*R^*) \cong \mathrm{Fix}(R^*R_*)$$

Isbell Duality

Let $\mathbf{V}$ be a complete and cocomplete closed symmetric monoidal category. Let $\mathbf{C}$ be a small $\mathbf{V}$-enriched category. Then there is a $\mathbf{V}$-adjunction

$$[\mathbf{C}^{\mathrm{op}}, \mathbf{V}] \underset{\mathrm{Spec}}{\overset{\mathcal{O}}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} [\mathbf{C}, \mathbf{V}]^{\mathrm{op}}$$

where

$$\mathcal{O}p : A \mapsto [\mathbf{C}^{\mathrm{op}}, \mathbf{V}]\,(p, \mathbf{C}(-, A))$$
$$\mathrm{Spec}\, q : A \mapsto [\mathbf{C}, \mathbf{V}]\,(q, \mathbf{C}(A, -))$$

The Isbell completion of $\mathbf{C}$, denoted $\mathbf{I}(\mathbf{C})$, is the full sub-$\mathbf{V}$-category of $[\mathbf{C}^{\mathrm{op}}, \mathbf{V}]$ consisting of the fixpoints $\mathrm{Fix}(\mathrm{Spec} \circ \mathcal{O})$.

Proof of $\mathcal{O} \dashv \mathrm{Spec}$

$$
\begin{aligned}
[\mathbf{C}, \mathbf{V}]^{\mathrm{op}}(\mathcal{O}p, q) &:= \int_{A \in \mathbf{C}} \mathbf{V}(qA, (\mathcal{O}p)A) \\
&:= \int_{A \in \mathbf{C}} \mathbf{V}(qA, [\mathbf{C}^{\mathrm{op}}, \mathbf{V}](p, \mathbf{C}(-, A))) \\
&:= \int_{A \in \mathbf{C}} \int_{B \in \mathbf{C}} \mathbf{V}(qA, \mathbf{V}(pB, \mathbf{C}(B, A))) \\
&\cong \int_{B \in \mathbf{C}} \int_{A \in \mathbf{C}} \mathbf{V}(pB, \mathbf{V}(qA, \mathbf{C}(B, A))) \\
&=: \int_{B \in \mathbf{C}} \mathbf{V}(pB, [\mathbf{C}, \mathbf{V}]^{\mathrm{op}}(\mathbf{C}(B, -), q)) \\
&=: \int_{B \in \mathbf{C}} \mathbf{V}(pB, (\mathrm{Spec}\, q)B) \\
&=: [\mathbf{C}^{\mathrm{op}}, \mathbf{V}](p, \mathrm{Spec}\, q)
\end{aligned}
$$

Feasibility Relation

- A $\mathbf{V}$-profunctor $F : \mathbf{C}^{\mathrm{op}} \times \mathbf{D} \to \mathbf{V}$ is the same as a function $F : \mathrm{ob}(\mathbf{C}) \times \mathrm{ob}(\mathbf{D}) \to \mathbf{V}$ such that for any $x, x' \in \mathbf{C}$ and $y, y' \in \mathbf{D}$ the following inequality holds in $\mathbf{V}$:

$$\mathbf{C}(x', x) \otimes F(x, y) \otimes \mathbf{D}(y, y') \leq F(x', y')$$

- Let $\mathbf{X} = (X, \leq_X)$ and $\mathbf{Y} = (Y, \leq_Y)$ be preorders. A *feasibility relation* for $\mathbf{X}$ given $\mathbf{Y}$ is a monotone map

$$F : \mathbf{X}^{\mathrm{op}} \times \mathbf{Y} \to \mathbf{2}$$

- Given $x \in X, y \in Y$, if $F(x, y) = 1$ we say x *can be obtained given* y.
- The requirement that F is monotone says that

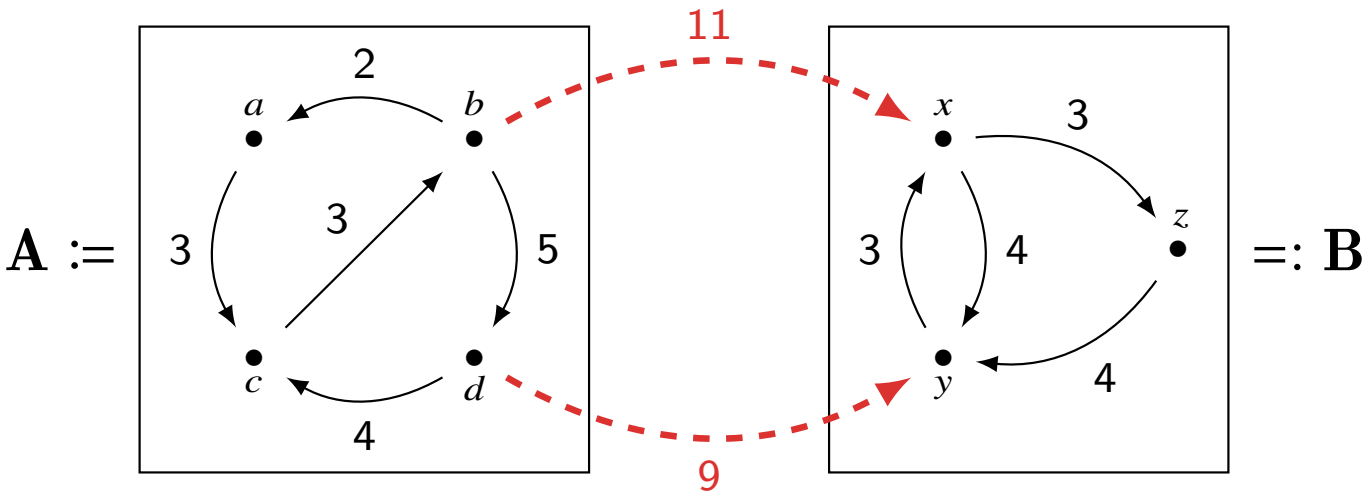
$$x' \leq_X x \;\;\&\;\; y \leq_Y y' \implies F(x, y) \leq_{\mathbf{2}} F(x', y')$$

- In other words, if x can be obtained given y, and if x' is available given x, then x' can be obtained given y. And if furthermore y is available given y', then x' can also be obtained given y'.

category	preorder
functor	order preserving map
set of natural transformations	domination relation
internal hom object in $\mathbf{Set}$	logical implication in $\mathbf{2}$
presheaf	downward closed subset (downset)
copresheaf	upward closed subset (upset)
category of presheaves	downsets ordered by inclusion
category of opcopresheaves	upsets ordered by containment
category of presheaves on a set	powerset of a set ordered by inclusion
category of opcopresheaves on a set	powerset of a set ordered by containment
adjunction	Galois connection
profunctor	relation
nucleus of a profunctor	Galois correspondence from a relation

Example — $\mathbf{Cost}$-Enriched Profunctor

- A $\mathbf{Cost}$-weighted graph is a graph with edges labeled by costs.
- The $\mathbf{Cost}$-enriched profunctor $F : \mathbf{A} \nrightarrow \mathbf{B}$ gives us the cost of getting from any city $x \in \mathbf{A}$ to any city $y \in \mathbf{B}$.



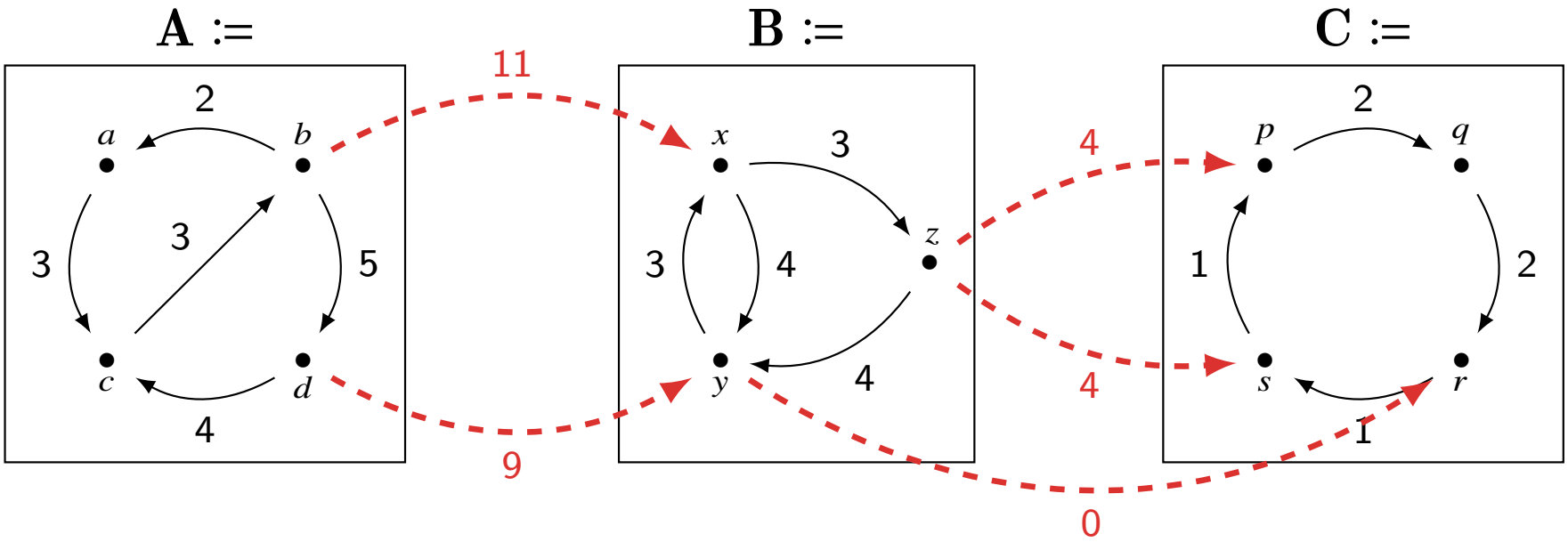
- The cost from $x \in \mathbf{A}$ to $y \in \mathbf{B}$ is given by the shortest path that runs from x through $\mathbf{A}$, then across one of the bridges, and then through $\mathbf{B}$ to the destination y.
- For example, $F(b, x) = 11, F(a, z) = 20, F(c, y) = 17$.

Composing Profunctors

Let $\mathbf{V}$ be a quantale, let $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ be $\mathbf{V}$-enriched categories, and let $F : \mathbf{A} \nrightarrow \mathbf{B}$ and $G : \mathbf{B} \nrightarrow \mathbf{C}$ be $\mathbf{V}$-enriched profunctors. We define their *composite*, denoted $G \circ F : \mathbf{A} \nrightarrow \mathbf{C}$ by the formula

$$(G \circ F)(a, c) = \bigvee_{b \in \mathbf{B}} \big(F(a, b) \otimes G(b, c)\big)$$

- Consider the $\mathbf{Cost}$-profunctors $F : \mathbf{A} \nrightarrow \mathbf{B}$ and $G : \mathbf{B} \nrightarrow \mathbf{C}$:



- For example, $(G \circ F)(a, q) = 24$

Collage

Let $\mathbf{V}$ be a quantale, let $\mathbf{A}$ and $\mathbf{B}$ be $\mathbf{V}$-categories, and let $F : \mathbf{A} \nrightarrow \mathbf{B}$ be a $\mathbf{V}$-enriched profunctors. The collage of F, denoted $\mathbf{Col}(F)$ is the $\mathbf{V}$-category defined as follows:

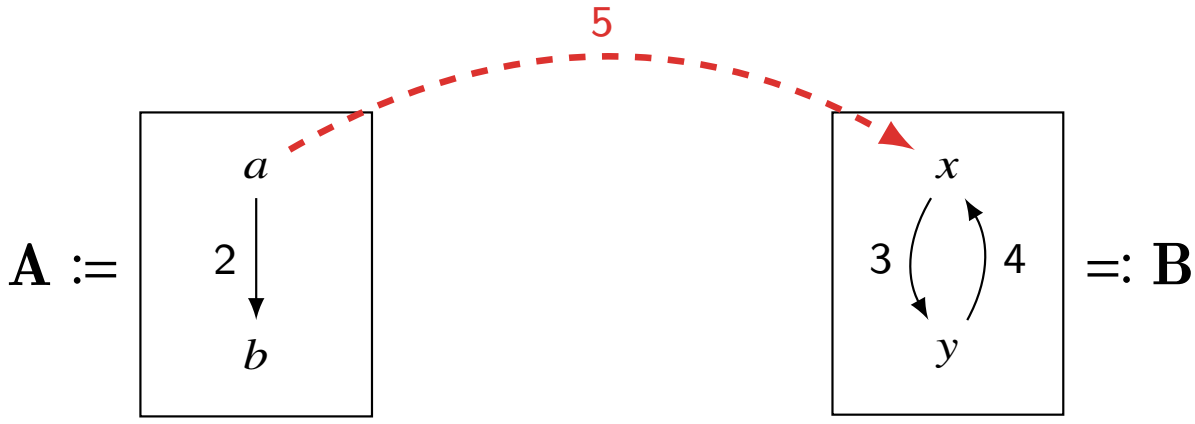
- $\mathrm{ob}(\mathbf{Col}(F)) \coloneqq \mathrm{ob}(\mathbf{A}) \coprod \mathrm{ob}(\mathbf{B})$
- For $a, b \in \mathrm{ob}(\mathbf{Col}(F))$, define $\mathbf{Col}(F)(a, b) \in \mathbf{V}$ to be

$$\mathbf{Col}(F)(a,b) \coloneqq \begin{cases} \mathbf{A}(a,b) \text{ if } a,b \in \mathbf{A} \\ F(a,b) \text{ if } a \in \mathbf{A}, b \in \mathbf{B} \\ \mathbf{B}(a,b) \text{ if } a,b \in \mathbf{B} \\ \emptyset \text{ if } a \in \mathbf{B}, b \in \mathbf{A} \end{cases}$$

There are obvious functors $\iota_{\mathbf{A}} : \mathbf{A} \to \mathbf{Col}(F)$ and $\iota_{\mathbf{B}} : \mathbf{B} \to \mathbf{Col}(F)$, sending each object and morphism to "itself".

Collage — Example

Consider the following picture of a $\mathbf{Cost}$-profunctor $F\colon \mathbf{A} \nrightarrow \mathbf{B}$:



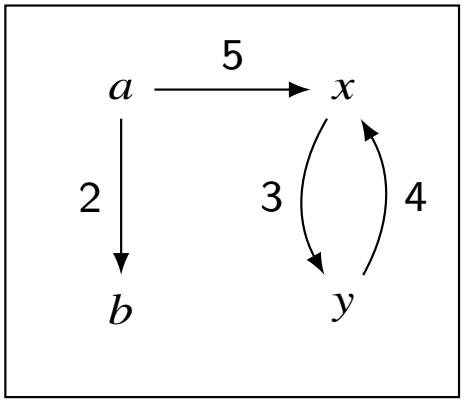
$\mathbf{A}$	a	b
a	0	2
b	∞	0

F	x	y
a	5	8
b	∞	∞

$\mathbf{B}$	x	y
x	0	3
y	4	0

A generalized Hasse diagram of the collage can be obtained by simply taking the union of the Hasse diagrams for $\mathbf{A}$ and $\mathbf{B}$, and adding in the bridges as arrows.

$\mathbf{Col}(F) =$



$\mathbf{Col}(F)$	a	b	x	y
a	0	2	5	8
b	∞	0	∞	∞
x	0	0	0	3
y	0	0	4	0

The Bicategory of Profunctors

Definition (The Bicategory of Profunctors)

There exists a bicategory of profunctors $\mathbf{Prof}$ in which

- objects are categories,
- morphisms are profunctors with the composition

$$(G \circ F)(A, C) = \int^{B \in \mathbf{B}} F(A, B) \times G(B, C)$$

  for $F : \mathbf{A} \nrightarrow \mathbf{B}$ and $G : \mathbf{B} \nrightarrow \mathbf{C}$.
- 2-morphisms are natural transformations between the profunctors.

Remark

$$(G \circ F)(A, C) = \int^{B \in \mathbf{B}} F(A, B) \times G(B, C) \quad \textit{in} \quad \mathbf{Prof}$$

$$(S \circ R)(a, c) = \exists b \in B : R(a, b) \wedge S(b, c) \quad \textit{in} \quad \mathbf{Rel}$$

Theorem

Let $F : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$. *Then*

$$FA \cong \int^{X \in \mathbf{C}} \mathrm{Hom}(X, A) \times FX$$

Proof.

For any Y, we have

$$\begin{aligned}
& \mathrm{Hom}_{\mathbf{Set}}\left(\int^{X \in \mathbf{C}} \mathrm{Hom}(X, A) \times FX, Y \right) \\
\cong & \int_{X \in \mathbf{C}} \mathrm{Hom}_{\mathbf{Set}}(\mathrm{Hom}(X, A) \times FX, Y) \\
\cong & \int_{X \in \mathbf{C}} \mathrm{Hom}_{\mathbf{Set}}(\mathrm{Hom}(X, A), \mathrm{Hom}_{\mathbf{Set}}(FX, Y)) \\
\cong & \ \mathrm{Nat}(\mathrm{Hom}(-, A), \mathrm{Hom}_{\mathbf{Set}}(F-, Y)) \\
\cong & \ \mathrm{Hom}_{\mathbf{Set}}(FA, Y) \qquad \text{(Yoneda Lemma)}
\end{aligned}$$

□

The Hom-functor is the identity of profunctor composition.

$$(F \circ \mathrm{Hom}(-,-))(A,B) \cong \int^{X \in \mathbf{C}} F(A,X) \times \mathrm{Hom}(X,B) \cong F(A,B)$$

$$\mathrm{Hom}(A,B) \cong \int^{X \in \mathbf{C}} \mathrm{Hom}(A,X) \times \mathrm{Hom}(X,B)$$

Analogy:

$$\int_{x \in \mathbb{R}} f(x)\delta(x-a)\,\mathrm{d}x = f(a)$$

$$\int^{X \in \mathbf{C}} FX \times \mathrm{Hom}(X,A) \cong FA$$

Theorem (Fubini Theorem for Ends/Coends)

Let $\mathbf{V}$ *be a symmetric monoidal category. Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be small* $\mathbf{V}$*-enriched categories. Let*

$$F : (\mathbf{A} \otimes \mathbf{B})^{\mathrm{op}} \otimes (\mathbf{A} \otimes \mathbf{B}) \to \mathbf{V}$$

be a $\mathbf{V}$*-enriched functor. Then: If for all object* $B, B' \in \mathbf{B}$ *the end* $\int_{A \in \mathbf{A}} F(A, B, A, B')$ *exists, then*

$$\int_{(A,B) \in \mathbf{A} \otimes \mathbf{B}} F(A, B, A, B) \cong \int_{A \in \mathbf{A}} \int_{B \in \mathbf{B}} F(A, B, A, B)$$

if either side exists. In particular, since $A \otimes B \cong B \otimes A$ *this implies that*

$$\int_{A \in \mathbf{A}} \int_{B \in \mathbf{B}} F(A, B, A, B) \cong \int_{B \in \mathbf{B}} \int_{A \in \mathbf{A}} F(A, B, A, B)$$

Dually,

$$\int^{(A,B) \in \mathbf{A} \otimes \mathbf{B}} F(A, B, A, B) \cong \int^{A \in \mathbf{A}} \int^{B \in \mathbf{B}} F(A, B, A, B) \cong \int^{B \in \mathbf{B}} \int^{A \in \mathbf{A}} F(A, B, A, B)$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

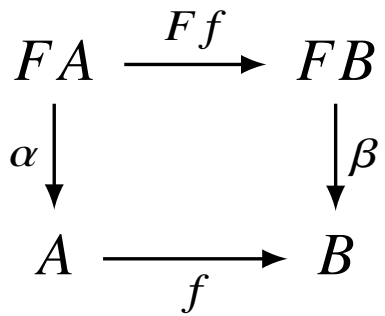
References2039

F-Algebra

Definition (F-Algebra)

If $\mathbf{C}$ is a category, and $F : \mathbf{C} \to \mathbf{C}$ is an endofunctor on $\mathbf{C}$, then an F-algebra is a pair (A, α) where $A \in \mathbf{C}$ and $\alpha : FA \to A$.

A homomorphism of F-algebras $f : (A, \alpha) \to (B, \beta)$ is $f : A \to B$ s.t.



The F-algebras together with F-algebra homomorphisms constitute a category $\mathbf{Alg}(F)$.

F-Algebra Examples

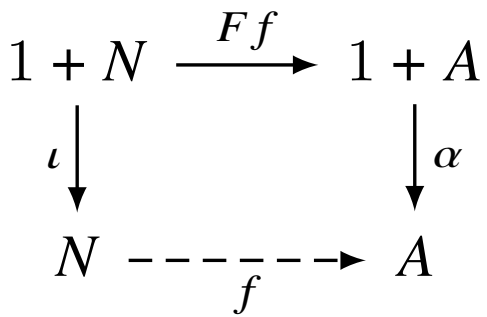
- A group is an F-algebra where F is the functor $FG = 1 + G + G \times G$, and $\alpha = e + i + m$:

$$\begin{aligned} \alpha : 1 + G + G \times G &\to G \\ * &\mapsto 1 \\ x &\mapsto x^{-1} \\ (x, y) &\mapsto x \cdot y \end{aligned}$$

- Monoids, lattices, modules, rings, fields, and vector spaces are also F-algebras for suitable functors F.

NNO — Initial F-Algebra

- An initial F-algebra is an F-algebra (I, ι) such that given any other F-algebra (A, α) there exists a unique F-algebra homomorphism $f : (I, \iota) \to (A, \alpha)$.
- A natural numbers object NNO is an initial F-algebra for the endofunctor $FA = 1 + A$ on $\mathbf{Set}$.



where $\iota = \begin{bmatrix} 0 \\ s \end{bmatrix} : 1 + N \to N$.

F-Algebra & F-Coalgebra

- For a category $\mathbf{C}$ and endofunctor F, an F-algebra is an object A in $\mathbf{C}$ and a morphism $\alpha : FA \longrightarrow A$.
- For a category $\mathbf{C}$ and endofunctor F, a F-coalgebra is an object A in $\mathbf{C}$ and a morphism $\alpha : A \longrightarrow FA$.
- An initial algebra for an endofunctor F on a category $\mathbf{C}$ is an initial object in the category of F-algebras.
- A terminal coalgebra for an endofunctor F on a category $\mathbf{C}$ is a terminal object in the category of F-coalgebras.
- **Theorem:** The initial algebra $(\mu F, \iota)$ is the least fixed point of F in that there is a unique map from μF to any other fixed point.
- **Theorem:** The terminal coalgebra $(\nu F, \tau)$ is the greatest fixed point of F in that there is a unique map to νF from any other fixed point.

Coalgebra Examples

- Power set

$$S \to \mathrm{P}(S)$$

- Non-deterministic automaton

$$\dfrac{\dfrac{\dfrac{\dfrac{S \to \mathrm{P}(S)^A}{S \times A \to \mathrm{P}(S)}}{R \subset (S \times A) \times S}}{R \subset S \times (A \times S)}}{S \to \mathrm{P}(A \times S)}$$

- Probability Distributions on a set S

$$S \to \mathrm{D}(S) := \left\{ \mu : S \to [0,1] : \sum_s \mu(s) = 1 \right\}$$

- Kripke Model $\mathcal{M} = (W, R, V)$

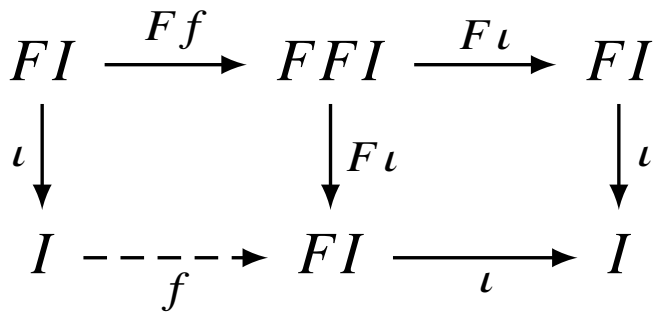
$$(R, V) : W \to \mathrm{P}(W) \times \mathbf{2}^{\mathrm{Var}}$$

$$w \vDash p \;\text{ iff }\; V(w)(p) = 1$$

$$w \vDash \Box A \;\text{ iff }\; \forall v \in R(w) : v \vDash A$$

An Initial F-Algebra is a Fixpoint of F

If we lift the initial algebra (I, ι), we get a new algebra $(FI, F\iota)$. And there must be a unique morphism $f : (I, \iota) \longrightarrow (FI, F\iota)$.
The morphism f is the inverse of ι.



since

$$\iota \circ f = 1_I$$

$$f \circ \iota = F\iota \circ Ff = F(\iota \circ f) = F(1_I) = 1_{FI}$$

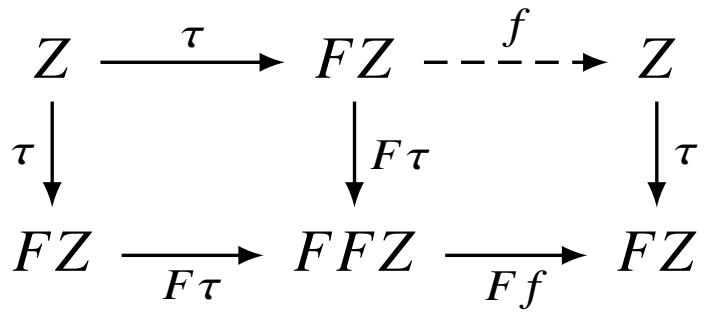
which means that ι is an isomorphism.

$$FI \cong I$$

I is a fixpoint of F.

A Terminal F-Coalgebra is a Fixpoint of F

Let (Z, τ) be a terminal F-coalgebra. Since $(FZ, F\tau)$ is also an F-coalgebra, there exists a unique morphism $f : FZ \to Z$ s.t.,



since

$$f \circ \tau = 1_Z$$

$$\tau \circ f = Ff \circ F\tau = F(f \circ \tau) = F(1_Z) = 1_{FZ}$$

which means that τ is an isomorphism.

$$Z \cong FZ$$

Z is a fixpoint of F.

Remark

Theorem

The powerset functor $\mathtt{P} : \mathbf{Set} \to \mathbf{Set}$ *does not have a terminal coalgebra.*

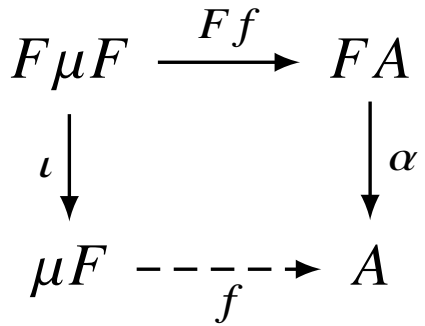
Proof.

Cantor: there is no injection $\mathtt{P}(X) \rightarrowtail X$.

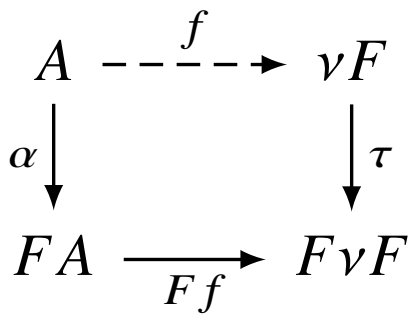
This excludes a terminal coalgebra $\tau : X \xrightarrow{\cong} \mathtt{P}(X)$. □

Iteration & Coiteration

Iteration: For any algebra $\alpha : FA \to A$, there is a unique morphism from the initial algebra $f : \mu F \to A$.



Coiteration: For any coalgebra $\alpha : A \to FA$, there is a unique morphism into the terminal coalgebra $f : A \to \nu F$.



Terminal F-Coalgebra Example: Streams

- Streams over a given set A are infinite sequences $\sigma = (\sigma_0, \sigma_1, \sigma_2, \dots)$ with $\sigma_i \in A$.
- A stream system with outputs in a set A is a pair $(S, \langle \text{out}, \text{trans} \rangle)$ where $\langle \text{out}, \text{trans} \rangle : S \to A \times S$.
- The set A^ω of all streams is the terminal F-coalgebra, where $F : \mathbf{Set} \to \mathbf{Set} :: S \to A \times S$.

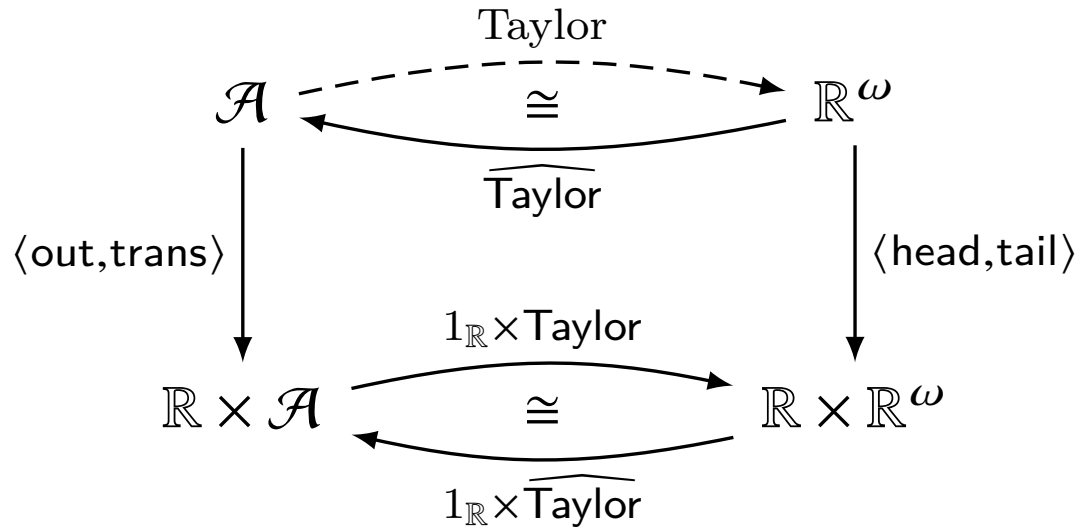
$$\begin{array}{ccc} S & \overset{f}{\dashrightarrow} & A^\omega \\ {\scriptstyle \langle \text{out},\text{trans} \rangle}\downarrow & & \downarrow{\scriptstyle \langle \text{head},\text{tail} \rangle} \\ A \times S & \xrightarrow[1_A \times f]{} & A \times A^\omega \end{array}$$

where $\text{head} : A^\omega \to A :: \sigma \mapsto \sigma_0$ and
$\text{tail} : A^\omega \to A^\omega :: \sigma \mapsto (\sigma_1, \sigma_2, \dots)$, and f is defined by

$$\text{head}(f(s)) = \text{out}(s)$$
$$\text{tail}(f(s)) = f(\text{trans}(s))$$

Terminal F-Coalgebra Example: Analytic Functions

- A function $f : \mathbb{R} \to \mathbb{R}$ is analytic if it possesses derivatives of all orders and agrees with its Taylor series in the neighbourhood of every point.
- Let us write $\mathcal{A}$ for the set of such analytic functions.



$$\text{out}(f) = f(0)$$
$$\text{trans}(f) = f'$$

$$\text{Taylor}(f) = \left(f(0), f'(0), f''(0), \ldots, f^{(n)}(0), \ldots\right)$$

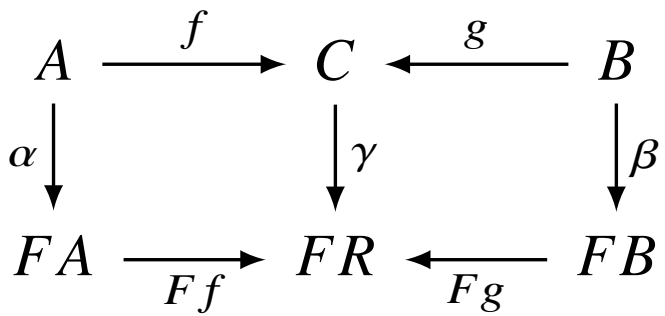
$$\widehat{\text{Taylor}}(\sigma) = \sum_{n=0}^{\infty} \frac{\sigma_n}{n!} x^n$$

- Taylor is an isomorphism, thus $\mathcal{A}$ can also be considered as the terminal F-coalgebra.

Behavioral Equivalence

Definition (Behavioral Equivalence)

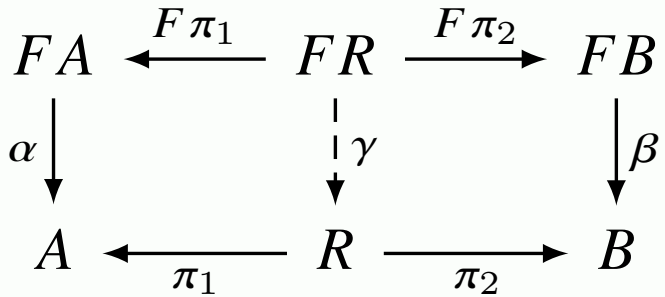
Let $\alpha : A \to FA$ and $\beta : B \to FB$ be two F-coalgebras. Two states $a \in A$ and $b \in B$ are **behaviorally equivalent** if there exists an F-coalgebra (C, γ) and coalgebra morphisms $f : (A, \alpha) \to (C, \gamma)$ and $g : (B, \beta) \to (C, \gamma)$ such that $f(a) = g(b)$.



F-Congruence

Definition (F-Congruence)

An F-congruence between two F-algebras (A, α) and (B, β) is a relation $R \subset A \times B$ such that there exists a map $\gamma : FR \to R$ s.t.

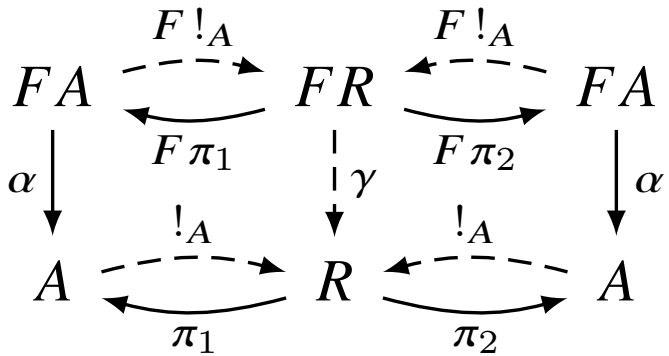


Induction Proof Principle

Theorem (Induction Proof Principle)

Every congruence relation $R \subset A \times A$ on an initial F-algebra (A, α) contains the diagonal relation $\Delta = \{(a, a) : a \in A\}$.

$$\Delta \subset R$$

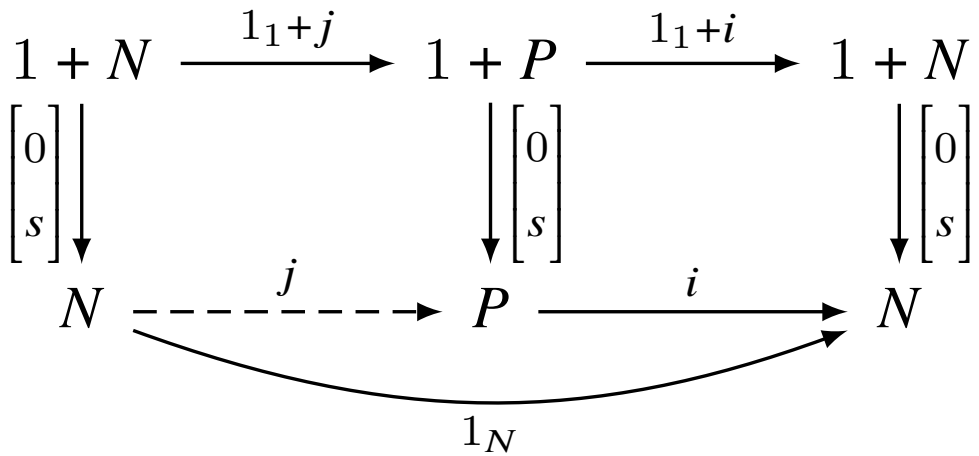


$$\pi_1 \circ !_A = 1_A = \pi_2 \circ !_A \implies \Delta = !_A$$

Induction Proof Principle

$$\frac{P(0) \quad \forall n \in N : P(n) \to P(sn)}{\forall n \in N : P(n)}$$

From the premises, the subset P carries an algebra structure $\begin{bmatrix} 0 \\ s \end{bmatrix} : 1 + P \to P$.

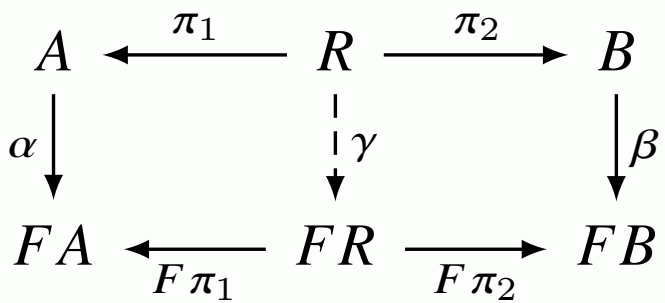


$$i \circ j = 1_N \implies \forall n \in N : P(n)$$

F-Bisimulation

Definition (F-Bisimulation)

An F-bisimulation between two F-coalgebras (A, α) and (B, β) is a relation $R \subset A \times B$ such that there exists a map $\gamma : R \to FR$ s.t.



- Bisimilar states are behaviorally equivalent.
- Kripke models are bisimilar iff they are behaviorally equivalent.

Coinduction Proof Principle

Theorem (Coinduction Proof Principle)

Every bisimulation relation $R \subset A \times A$ on a terminal F-coalgebra (A, α) is contained in the diagonal $\Delta = \{(a, a) : a \in A\}$.

$$R \subset \Delta$$

- Homomorphisms are structure-preserving functions. Similarly, bisimulations are structure-preserving relations.
- If two states of a final coalgebra are related by some bisimulation relation, then they are equal.
- For two streams $\sigma, \sigma' \in A^\omega$,

$$\sigma \leftrightarrows \sigma' \implies \sigma = \sigma'$$

Primitive Recursion

Theorem (Primitive Recursion)

For any $\alpha : F(\mu F \times A) \to A$, there exists a unique $f : \mu F \to A$ s.t.

$$\begin{array}{ccc} F\mu F & \xrightarrow{F\langle 1, f\rangle} & F(\mu F \times A) \\ {\scriptstyle \iota}\downarrow & & \downarrow{\scriptstyle \alpha} \\ \mu F & \xrightarrow[f]{} & A \end{array}$$

Proof.

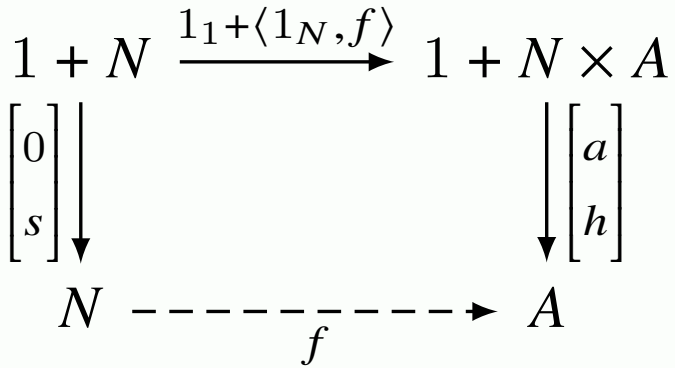
$$\begin{array}{ccccc} F\mu F & \xrightarrow{Fg} & F(\mu F \times A) & \xrightarrow{F\pi_1} & F\mu F \\ {\scriptstyle \iota}\downarrow & & \downarrow{\scriptstyle \langle \iota \circ F\pi_1, \alpha\rangle} & & \downarrow{\scriptstyle \iota} \\ \mu F & \xrightarrow{g} & \mu F \times A & \xrightarrow{\pi_1} & \mu F \\ & \searrow^{1} & & & \end{array}$$

$$f := \pi_2 \circ g$$

Primitive Recursion on N

Theorem (Primitive Recursion on N)

Given $a : 1 \to A$ and $h : N \times A \to A$, there exists a unique $f : N \to A$ s.t.



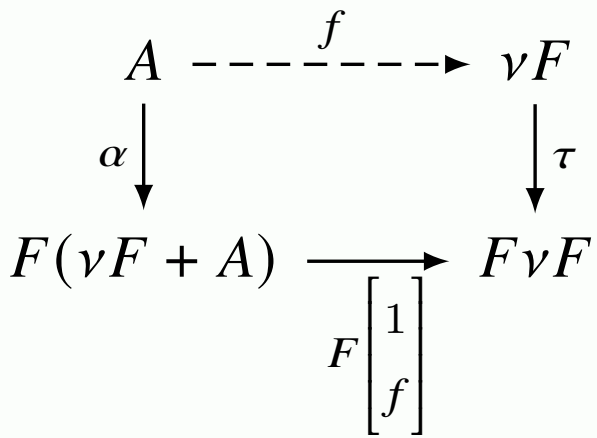
$$f \circ 0 = a$$

$$f \circ s = h \circ \langle 1_N, f \rangle$$

Primitive Corecursion

Theorem (Primitive Corecursion)

For any $\alpha : A \to F(\nu F + A)$, *there exists a unique* $f : A \to \nu F$ *s.t.*



Conatural Numbers

Conatural number system $\overline{N} = \{0, 1, 2, \ldots, \infty\}$ is the terminal F-coalgebra for the endofunctor $FA = 1 + A$ on $\mathbf{Set}$.

$$\begin{array}{ccc} A & \xrightarrow{f} & \overline{N} \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\tau} \\ 1 + A & \xrightarrow[1+f]{} & 1 + \overline{N} \end{array}$$

where

$$\tau(x) := \begin{cases} * & \text{if } x = 0 \\ n & \text{if } x = n + 1 \\ \infty & \text{if } x = \infty \end{cases}$$

and f is defined corecursively.

$$f(a) = 0 \text{ if } \alpha(a) = *$$
$$\tau(f(a)) = f(b) \text{ if } \alpha(a) = b$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

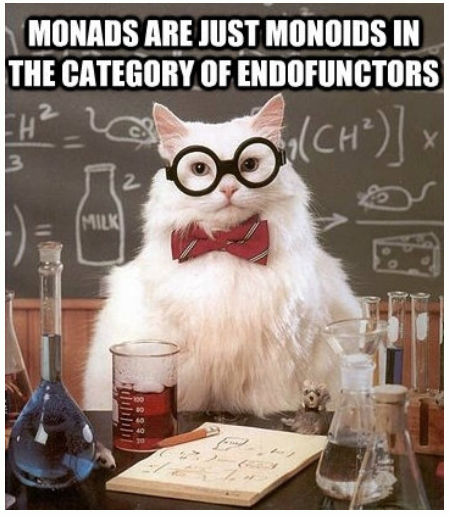
Quantum Computing

Answers to the Exercises

References2039

Monad

Definition (Monad)



A *monad* (T, η, μ) on a category $\mathbf{C}$ consists of

- an endofunctor $T : \mathbf{C} \to \mathbf{C}$,
- a unit natural transformation $\eta : 1_{\mathbf{C}} \to T$,
- a multiplication natural transformation $\mu : T^2 \to T$

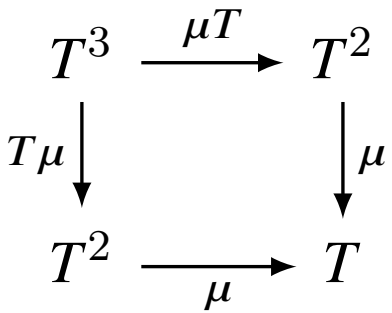
such that:

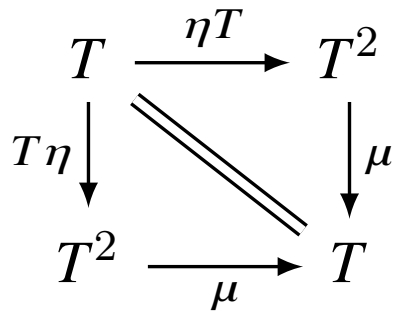
- $\mu \circ T\mu = \mu \circ \mu T$ (as natural transformations $T^3 \to T$);
- $\mu \circ T\eta = \mu \circ \eta T = 1_T$ (as natural transformations $T \to T$).

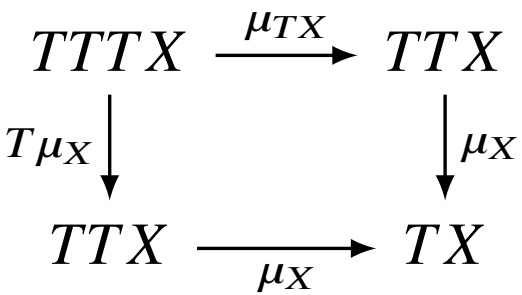
$$\begin{array}{ccc} T^3 & \xrightarrow{\mu T} & T^2 \\ {\scriptstyle T\mu}\downarrow & & \downarrow{\scriptstyle \mu} \\ T^2 & \xrightarrow[\mu]{} & T \end{array} \qquad \begin{array}{ccc} T & \xrightarrow{\eta T} & T^2 \\ {\scriptstyle T\eta}\downarrow & \searrow & \downarrow{\scriptstyle \mu} \\ T^2 & \xrightarrow[\mu]{} & T \end{array}$$

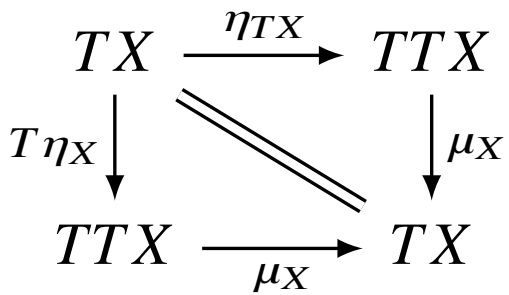
Monad

The first axiom is akin to the associativity in monoids if we think of μ as the monoid's binary operation, and the second axiom is akin to the existence of an identity element. Indeed, a monad on $\mathbf{C}$ can be regarded as a monoid in the category $\mathbf{End}_{\mathbf{C}}$ whose objects are the endofunctors of $\mathbf{C}$ and whose morphisms are the natural transformations between them, with the monoidal structure induced by the composition of endofunctors.









A *comonad* for a category $\mathbf{C}$ is a monad for the opposite category $\mathbf{C}^{\mathrm{op}}$.

Comonad

Definition (Comonad)

A *comonad* (W, ε, δ) on a category $\mathbf{C}$ consists of

- an endofunctor $W : \mathbf{C} \to \mathbf{C}$,
- a counit natural transformation $\varepsilon : W \to 1_{\mathbf{C}}$,
- a multiplication natural transformation $\delta : W \to W^2$

such that:

- $W\delta \circ \delta = \delta W \circ \delta$ (as natural transformations $W \to W^3$);
- $W\varepsilon \circ \delta = \varepsilon W \circ \delta = 1_W$ (as natural transformations $W \to W$).

$$\begin{array}{ccc} W & \xrightarrow{\delta} & W^2 \\ {\scriptstyle\delta}\downarrow & & \downarrow{\scriptstyle\delta W} \\ W^2 & \xrightarrow[W\delta]{} & W^3 \end{array} \qquad \begin{array}{ccc} W & \xrightarrow{\delta} & W^2 \\ {\scriptstyle\delta}\downarrow & \searrow & \downarrow{\scriptstyle\varepsilon W} \\ W^2 & \xrightarrow[W\varepsilon]{} & W \end{array}$$

Remarks

- A monad is a way of extending spaces to include generalized elements and generalized functions of a specific kind.
- A comonad is a way to equip spaces with extra information of a specific kind, and let some morphisms access that information.
- A monad is a consistent choice of formal expressions of a specific kind, together with ways to evaluate them.
- A comonad is a consistent way to construct, from spaces, processes of a specified structure, and give selected strategies or trajectories.

Examples

- Consider $\mathbf{Set}$.
  - $T : A \mapsto \mathrm{P}(A)$, and $T(f) : A \mapsto f(A)$ for object A and morphism f.
  - $\eta_A : A \to \mathrm{P}(A)$ given by $\eta_A(a) \coloneqq \{a\}$.
  - $\mu_A : \mathrm{P}(\mathrm{P}(A)) \to \mathrm{P}(A)$ given by $\mu_A(B) \coloneqq \bigcup B$.
- Consider a monoid $(M, e, \cdot)$, the monoid action $(M \times -, \eta, \mu)$ is a monad on $\mathbf{Set}$
  - $T : A \mapsto M \times A$, and $T(f) : (m, a) \mapsto (m, f(a))$.
  - $\eta_A : A \to M \times A :: a \mapsto (e, a)$.
  - $\mu_A : M \times M \times A \to M \times A :: (m, (n, a)) \mapsto (mn, a)$.
- A preorder $(P, \leq)$ yields a category where endofunctors are monotone functions. Given a monad (T, η, μ), the natural transformations η and μ give that, for $a \in P$, $a \leq Ta$ and $TTa \leq Ta$, since $\eta_a : a \to Ta$, and $\mu_a : TTa \to Ta$. Then $Ta \leq TTa \leq Ta \implies TTa = Ta$. Monads on $(P, \leq)$ are closure operators[34].

[34] A *closure operator* on a poset $(P, \leq)$ is $T : P \to P$ s.t. for $x, y \in P$,

- $x \leq T(x)$
- $x \leq y \to T(x) \leq T(y)$
- $T(T(x)) = T(x)$

Example: List

- $\mathrm{List} : \mathbf{Set} \to \mathbf{Set}$ maps any set X to the set of all finite first-order formal expressions built from elements of X
- For $f : X \to Y$, $\mathrm{List}\, f : \mathrm{List}\, X \to \mathrm{List}\, Y$ acts on expressions like this

$$\mathrm{List}\, f : [x_1] * \cdots * [x_n] \mapsto [f x_1] * \cdots * [f x_n]$$

- $\eta_X : X \to \mathrm{List}(X)$

$$x \mapsto [x]$$

- $\mu_X : \mathrm{List}(\mathrm{List}(X)) \to \mathrm{List}(X)$

$$[[x_1]*\cdots*[x_m]]*\cdots*[[y_1]*\cdots*[y_n]] \mapsto [x_1]*\cdots*[x_m]*\cdots*[y_1]*\cdots*[y_n]$$

- $(\mathrm{List}(X), *, [])$ is a monoid.
- $(\mathrm{List}, \eta, \mu)$ is a monad.

Example: Multiset

- $\mathrm{Multi} : \mathbf{Set} \to \mathbf{Set}$ maps each set X to the set of finite multisets (or bags) from X. For example:

$$\{a : 2, b : 1, c : 1\}$$

- $\eta_X : X \to \mathrm{Multi}(X)$

$$x \mapsto \{x : 1\}$$

- $\mu_X : \mathrm{Multi}(\mathrm{Multi}(X)) \to \mathrm{Multi}(X)$ maps a multiset of multisets to a multiset by taking "unions" which account for multiplicities. For example:

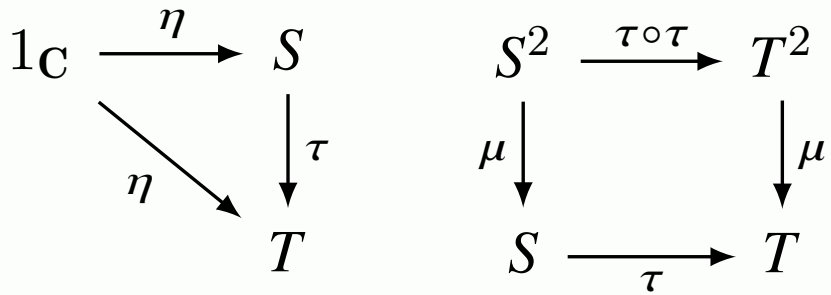
$$\big\{\{a : 1, b : 2\} : 2, \{a : 3, c : 1\} : 1\big\} \mapsto \{a : 5, b : 4, c : 1\}$$

- $(\mathrm{Multi}, \eta, \mu)$ is a monad.

Monad Morphism

Definition (Monad Morphism)

Given a pair of monads over the same base category $S, T : \mathbf{C} \to \mathbf{C}$, a monad morphism is a natural transformation $\tau : S \to T$ such that:



Example (Lists to Multisets)

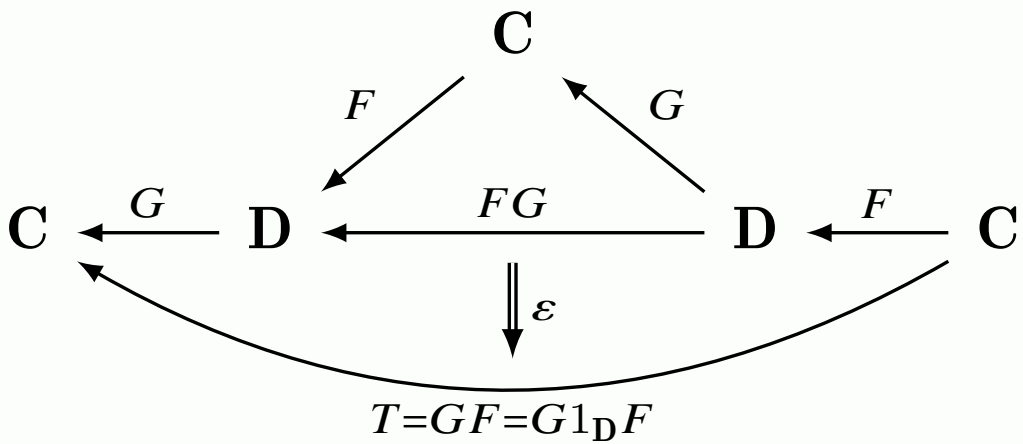
There is a monad morphism $\mathrm{List} \to \mathrm{Multi}$ from the list to the multiset monad, mapping a list to the multiset of elements that appear, with their multiplicities. For example:

$$[a, b, a, c] \mapsto \{a : 2, b : 1, c : 1\}$$

Monads from Adjunctions

Any adjunction $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D}$ gives rise to a monad on $\mathbf{C}$, with

- the endofuntor $T \coloneqq GF$,
- the unit natural transformation $\eta : 1_{\mathbf{C}} \to GF$,
- the multiplication natural transformation $\mu \coloneqq G\varepsilon F : GFGF \to GF$.

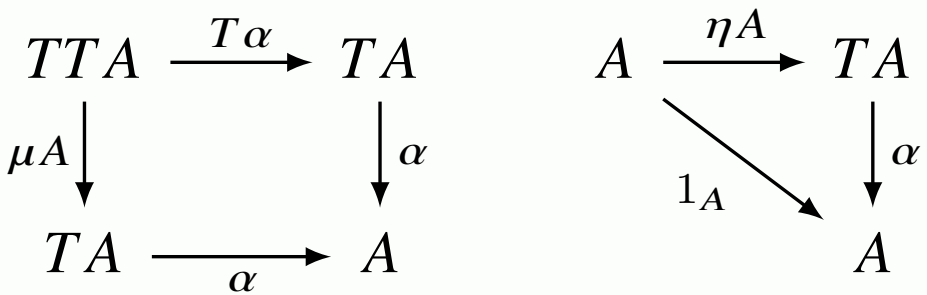


- $(GF, \eta, G\varepsilon F)$ is a monad.
- $(FG, \varepsilon, F\eta G)$ is a comonad.

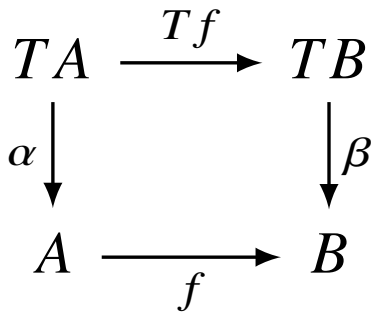
T-Algebra

Definition

Let $T = (T, \eta, \mu)$ be a monad on $\mathbf{C}$. A T-algebra is a pair (A, α) where $A \in \mathbf{C}$ and $\alpha : TA \to A$ s.t.



A homomorphism of T-algebras $f : (A, \alpha) \to (B, \beta)$ is $f : A \to B$ s.t.



The category of T-algebras and T-algebra homomorphisms is denoted $\mathbf{C}^T$. This is called the Eilenberg-Moore category.

Kleisli Category

Definition (Kleisli Category)

The Kleisli Category $\mathbf{C}_T$ of a monad (T, η, μ) on a category $\mathbf{C}$ has

- as objects the object of $\mathbf{C}$
- as morphisms $A \to B$ the morphisms of the form

$$A \to TB$$

- composition of $A \xrightarrow{f} TB$ with $B \xrightarrow{g} TC$ is given by the composition

$$A \xrightarrow{f} TB \xrightarrow{Tg} TTC \xrightarrow{\mu C} TC$$

- the identity morphism on A is

$$A \xrightarrow{\eta_A} TA$$

Proposition

The map

$$A \xrightarrow{f} TB \quad \longmapsto \quad TA \xrightarrow{Tf} TTB \xrightarrow{\mu B} TB$$

constitutes a full and faithful functor $\mathbf{C}_T \to \mathbf{C}^T$.

Theorem

Every monad $T = (T, \eta, \mu)$ arises from an adjunction $\mathbf{C} \underset{U}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{C}^T$.

More precisely, there are natural transformations $(\dot{\eta}, \dot{\varepsilon}) : F \dashv U$ and $T = UF, \eta = \dot{\eta}, \mu = U\dot{\varepsilon}F$.

Proof.

Let the forgetful functor $U(A, \alpha) = A$, and define $FA = (TA, \mu_A)$, and $F(A \xrightarrow{f} B) = Tf$.
It is easy to check that (A, μ_A) is a T-algebra.
Clearly $UF(A) = U(TA, \mu_A) = T(A)$, and we have a natural transformation $\dot{\eta} = \eta : 1_{\mathbf{C}} \to UF$.
Define $\dot{\varepsilon} : FU \to 1_{\mathbf{C}^T}$ by $\dot{\varepsilon}_{(A,\alpha)} = \alpha : TA \to A$.
It is easy to check that $(\dot{\eta}, \dot{\varepsilon}) : F \dashv U$.
For $A \in \mathbf{C}$, $U\dot{\varepsilon}F(A) = U\dot{\varepsilon}_{FA} = U\dot{\varepsilon}_{(TA,\mu_A)} = U\mu_A = \mu_A$. □

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Sheaf

- The topology $\mathcal{O}_X$ of X, i.e. the poset of open subsets of X, ordered by inclusion, can be considered to be a category.
- A presheaf $F : \mathcal{O}_X^{\mathrm{op}} \to \mathbf{Set}$ is a *sheaf* iff it satisfies:
  1. (Locality) For any open cover $\{U_i\}_{i \in I}$ of U, and for $s, t \in F(U)$, and $\forall i : s \restriction_{U_i} = t \restriction_{U_i}$, then $s = t$.
  2. (Gluing) For any open cover $\{U_i\}_{i \in I}$ of U, and for $\{s_i \in F(U_i)\}_{i \in I}$ such that $\forall i, j : s_i \restriction_{U_i \cap U_j} = s_j \restriction_{U_i \cap U_j}$, then there exists some $s \in F(U)$ such that $\forall i : s \restriction_{U_i} = s_i$.
- It "says" roughly that there is a unique way to "glue" together functions that are defined locally. In other words, for sheaves, one can systematically move from the local to the global.
- The category $\mathbf{Sh}(X)$ is the category of sheaves and morphisms are natural transformations between them.

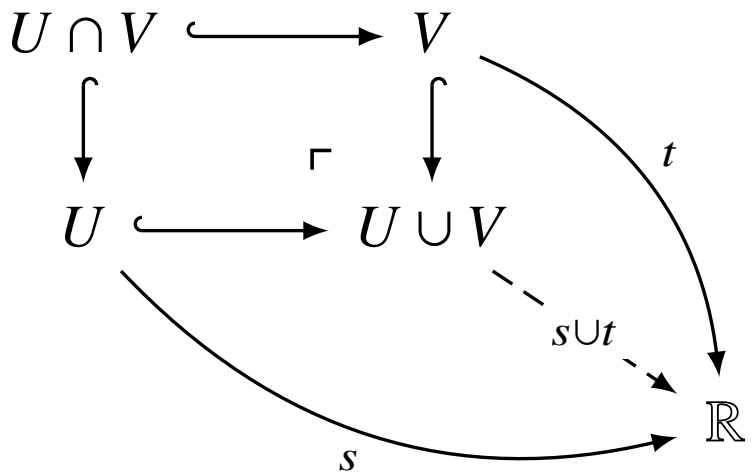
Example

$$C(U, \mathbb{R}) := \{\text{continuous functions } f : U \subset \mathbb{R}^n \to \mathbb{R}\}$$

- There are naturally defined restriction functions: whenever $V \subset U$,

$$\begin{array}{ccc} C(U, \mathbb{R}) & \xrightarrow{\mathrm{res}^U_V} & C(V, \mathbb{R}) \\ f & \longmapsto & f\restriction_V \end{array}$$

  and this assignment is functorial.
- The gluing property: whenever $s \in C(U, \mathbb{R})$ and $t \in C(V, \mathbb{R})$ restrict to the same function $s\restriction_{U\cap V} = t\restriction_{U\cap V} \in C(U \cap V, \mathbb{R})$, there is a unique continuous extension $s \cup t \in C(U \cup V, \mathbb{R})$:



Sheaf — Categorical Reformulations

- A presheaf $F : O_X^{\mathrm{op}} \to \mathbf{Set}$ is a *sheaf* iff for any open cover $\{U_i\}_{i \in I}$ of U, the following diagram is an equalizer:

$$F(U) \xrightarrow{\mathrm{res}^U_{U_i}} \prod_i F(U_i) \underset{\mathrm{res}^{U_j}_{U_i \cap U_j}}{\overset{\mathrm{res}^{U_i}_{U_i \cap U_j}}{\rightrightarrows}} \prod_{i,j} F(U_i \cap U_j)$$

Where

$$\mathrm{res}^U_{U_i} : s \mapsto (s \restriction_{U_i})_{i \in I}$$
$$\mathrm{res}^{U_i}_{U_i \cap U_j} : (s_i)_{i \in I} \mapsto (s_i \restriction_{U_i \cap U_j})_{(i,j) \in I \times I}$$
$$\mathrm{res}^{U_j}_{U_i \cap U_j} : (s_i)_{i \in I} \mapsto (s_j \restriction_{U_i \cap U_j})_{(i,j) \in I \times I}$$

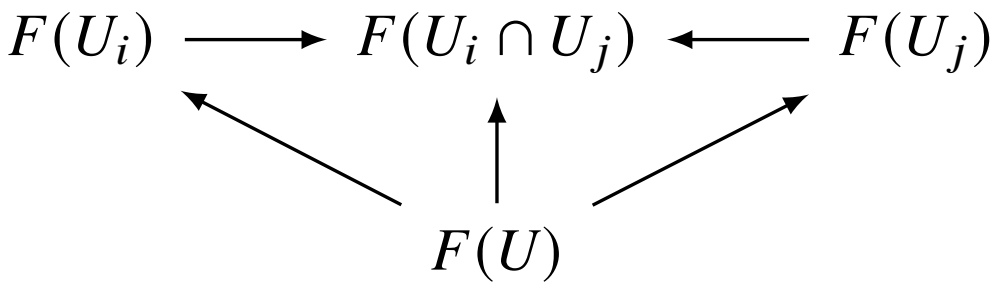
$$\begin{array}{ccccc}
 & & F(U_i) & \xrightarrow{F(U_i \cap U_j \hookrightarrow U_i)} & F(U_i \cap U_j) \\
 & \overset{F(U_i \hookrightarrow U)}{\nearrow} & \uparrow{\scriptstyle \pi_i} & & \uparrow{\scriptstyle \pi_{ij}} \\
F(U) & \xrightarrow{\mathrm{res}^U_{U_i}} & \prod_i F(U_i) & \underset{\mathrm{res}^{U_j}_{U_i \cap U_j}}{\overset{\mathrm{res}^{U_i}_{U_i \cap U_j}}{\rightrightarrows}} & \prod_{i,j} F(U_i \cap U_j) \\
 & \underset{F(U_j \hookrightarrow U)}{\searrow} & \downarrow{\scriptstyle \pi_j} & & \downarrow{\scriptstyle \pi_{ij}} \\
 & & F(U_j) & \xrightarrow[F(U_i \cap U_j \hookrightarrow U_j)]{} & F(U_i \cap U_j)
\end{array}$$

Sheaf — Categorical Reformulations

- Let $\mathbf{I}$ be a category with $\mathrm{ob}(\mathbf{I}) \coloneqq \{U_i : i \in I\} \cup \{U_i \cap U_j : i, j \in I\}$, and the morphisms are the inclusions of $U_i \cap U_j$ in U_i and U_j.

$$U_i \hookleftarrow U_i \cap U_j \hookrightarrow U_j$$

Then U can be taken as a colimit of $\mathbf{I}$.



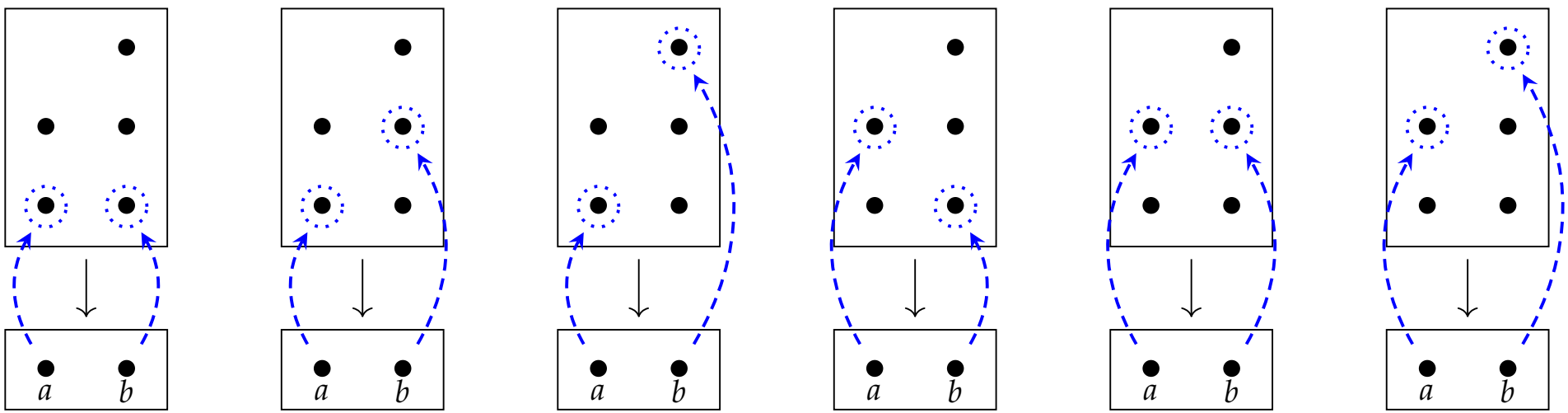
A presheaf $F : \mathcal{O}_X^{\mathrm{op}} \to \mathbf{Set}$ is a *sheaf* iff $\varprojlim_{\mathbf{I}} F \cong F(U)$.

Sheaf — Example

- Let $p\colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ be a continuous function. Define the presheaf $\Gamma_p : \mathcal{O}_X^{\mathrm{op}} \to \mathbf{Set}$ on an arbitrary subset $U \subset X$ by:

$$\Gamma_p(U) \coloneqq \left\{ s : U \to Y : p \circ s(u) = u \text{ for all } u \in U \right\}$$

  Then Γ_p is a sheaf.
- For example, consider the function $p : Y \to X$ that sends each element of Y to the element of X below it, and let $U = \{a, b\} \subset X$. Then $\Gamma_p(U)$ is the set of all ways to choose an element of Y per fiber over U.



Sheaf — Examples

1. For any fixed continuous function $p : Y \to X$, there is a sheaf Γ_p defined by

$$U \mapsto \Gamma_p U = \left\{ \begin{array}{ccc} & & Y \\ & \overset{s}{\nearrow} & \downarrow p \\ U & \hookrightarrow & X \end{array} \right\}$$

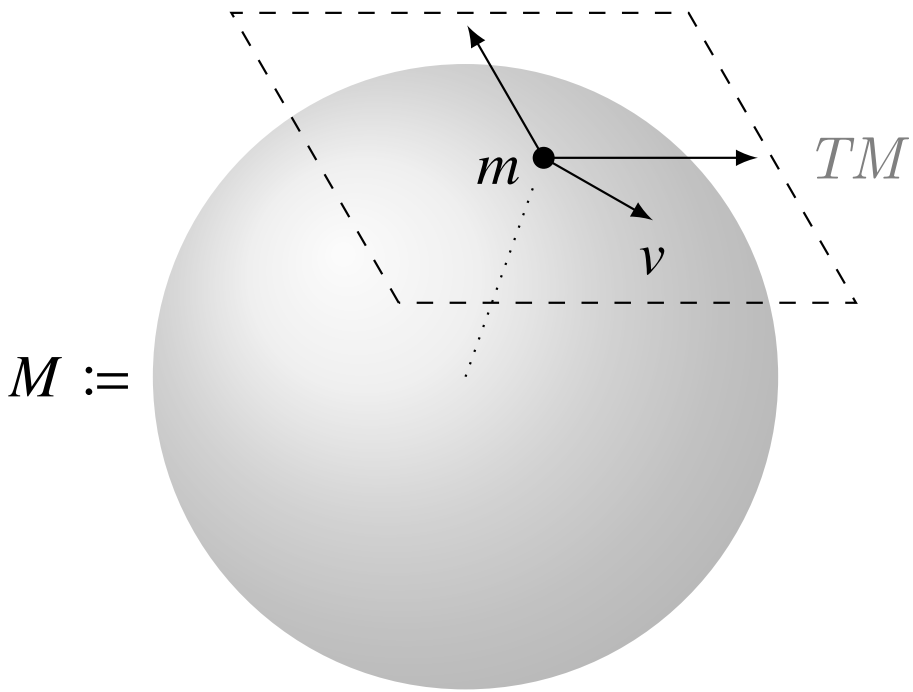
   whose elements are local sections, continuous functions $s : U \to Y$ so that $p \circ s = i : U \hookrightarrow X$.
2. There is a sheaf $U \mapsto \{V \in \mathcal{O} : V \subset U\}$.
3. For any fixed open subset $V \subset X$, there is a sheaf $\mathcal{O}(-, V)$ defined by

$$U \mapsto \mathcal{O}(U, V) = \begin{cases} * & \text{if } U \subset V \\ \emptyset & \text{if } U \not\subset V \end{cases}$$

4. For any fixed space Y, there is a sheaf $U \mapsto C(U, Y)$ of continuous functions $U \to Y$.
5. For $0 \leq k \leq \infty$, there is a sheaf $U \mapsto C^k(U, \mathbb{R})$ of k-times continuously differentiable real-valued functions.

Example — The Sheaf of Vector Fields

A manifold M has a tangent bundle TM, whose points are pairs (m, v), where $m \in M$ and v is a tangent vector emanating from it.



TM comes with a continuous map $p : TM \to M :: (m, v) \mapsto m$.
$\Gamma_p : U \mapsto \{s : U \to TM : s \text{ is continuous and } p \circ s = i : U \hookrightarrow M\}$ is a sheaf. Given an open subset $U \subset M$, an element $v \in \Gamma_p U$ is a vector field which continuously assigns a tangent vector $v(m)$ to each point $m \in U$.

Example: Presheaves that are not sheaves

— Informally, Presheaves are an assignment of some data to the open sets of a space X.

— Sheaves are about creating big(ger) data from small(er) data.

- Let $X = \mathbb{R}$ and consider $F(U) = \{f : U \to \mathbb{R} : \ f \text{ is continuous and bounded}\}$.
- Then if we glue over $\mathbb{R}$ infinitely many functions bounded on subsets of $\mathbb{R}$, we can easily obtain an ubounded function over $\mathbb{R}$.
- Therefore, this is not a sheaf.
- "bounded" is not a local property.

Stalk

- The *stalk* F_x of a presheaf $F : \mathcal{O}_X^{\mathrm{op}} \to \mathbf{Set}$ at $x \in U$ is the set of germs of F at x. Precisely,

$$F_x \coloneqq \left\{ \operatorname{germ}_x s : x \in U, s \in FU \right\}$$

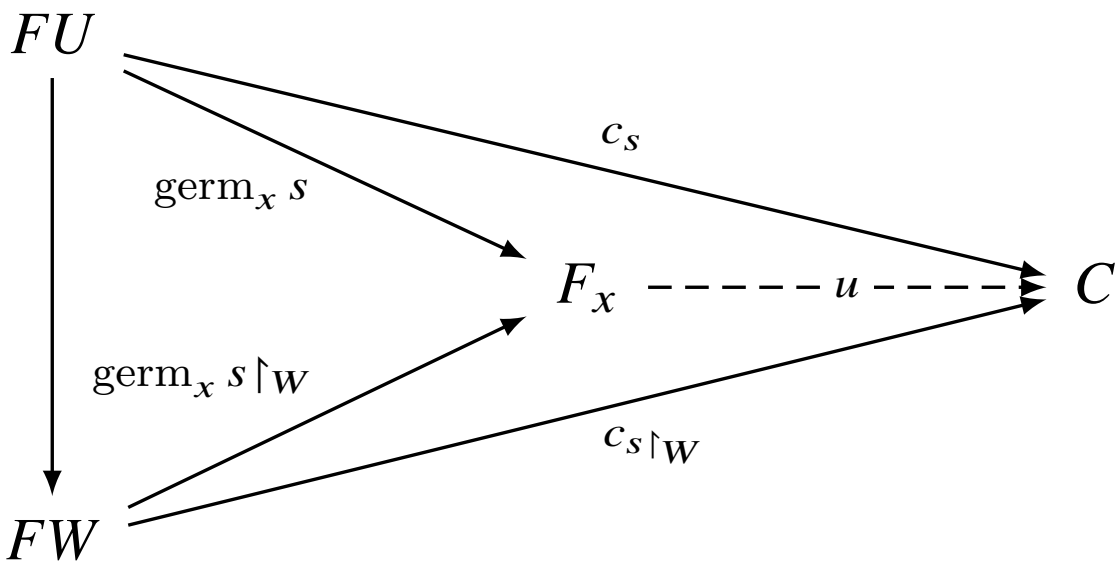
where

$$\operatorname{germ}_x s \coloneqq [s]_x$$

and $s \sim_x t$ iff there is an open set $W \subset U \cap V$ with $x \in W$ and $s \restriction_W = t \restriction_W$ where $s \in FU, t \in FV$.

- Equivalently, $F_x \coloneqq \varinjlim_{U \ni x} FU$.

The function $\operatorname{germ}_x : FU \to F_x$ form a cocone over $F \restriction_{\{U \in \mathcal{O}_X : x \in U\}}$.



Etale Bundle

- Consider the slice category $\mathbf{Top}/X$. An *étale bundle* over X is $(Y, p : Y \to X)$ in $\mathbf{Top}/X$ such that p is a local homeomorphism: that is, for every $y \in Y$, there is an open set $U \ni y$ such that $p(U)$ is open in X and $p\restriction_U : U \to p(U)$ is a homeomorphism.
- The category $\mathbf{Et}(X)$ is the category of étale bundles over X and morphisms are the continous functions between them.
- $\widehat{\mathcal{O}_X} \underset{\Gamma}{\overset{\Lambda}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{Top}/X$

  where $\Gamma : Y \xrightarrow{p} X \mapsto \left\{ s : U \to Y : p \circ s = i : U \hookrightarrow X \right\}$, while $\Lambda : F \mapsto \coprod_{x \in X} F_x = \left\{ \operatorname{germ}_x s : x \in X, s \in FU \right\}$.

  There are natural transformations

  $$\eta_F : F \to \Gamma\Lambda F, \quad \varepsilon_Y : \Lambda\Gamma Y \to Y$$

  for F a presheaf and Y a bundle which are unit and counit.
- If F is a sheaf, η_F is an isomorphism, while if Y is étale, ε_Y is an isomorphism. $\mathbf{Sh}(X) \underset{\longleftarrow}{\overset{\longrightarrow}{\underset{\cong}{}}} \mathbf{Et}(X)$

Proof Sketch.

For $U \in \mathcal{O}_X$,

$$\eta_{FU} : FU \to \Gamma\Lambda FU :: s \mapsto \dot{s}$$

where

$$\dot{s} : U \to \Lambda F :: x \mapsto \operatorname{germ}_x s$$

Given a bundle $Y \to X$, each point of the corresponding étale bundle $\Lambda\Gamma Y$ has the form $\dot{s}x$ for some point $x \in X$ and some actual cross-section $s : U \to Y$ of the given bundle.

$$\varepsilon_Y : \Lambda\Gamma Y \to Y :: \dot{s}x \mapsto sx$$

$$\begin{array}{ccc} \Lambda & \xrightarrow{\Lambda\eta} & \Lambda\Gamma\Lambda \\ & \searrow^{1_\Lambda} & \downarrow{\scriptstyle \varepsilon\Lambda} \\ & & \Lambda \end{array} \qquad \begin{array}{ccc} \Gamma & \xrightarrow{\eta\Gamma} & \Gamma\Lambda\Gamma \\ & \searrow^{1_\Gamma} & \downarrow{\scriptstyle \Gamma\varepsilon} \\ & & \Gamma \end{array}$$

$$\operatorname{germ}_x s \mapsto \operatorname{germ}_x \dot{s} \mapsto \operatorname{germ}_x s \qquad s \mapsto \dot{s} \mapsto s$$

□

Sheafification

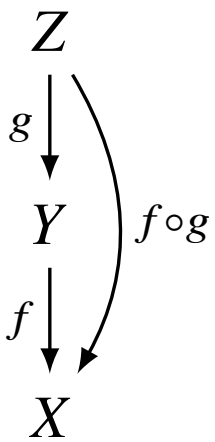
- The inclusion functor
$$i : \mathbf{Sh}(X) \hookrightarrow \widehat{\mathcal{O}_X}$$
has a left adjoint.
- The left adjoint functor
$$\Gamma\Lambda : \widehat{\mathcal{O}_X} \to \mathbf{Sh}(X)$$
is known as the associated sheaf functor, or the sheafification functor. It carries each presheaf F on X to the "best approximation" $\Gamma\Lambda F$ of F by a sheaf.

$$\mathbf{Sh}(X) \underset{i}{\overset{\Gamma\Lambda}{\underset{\hookrightarrow}{\overset{\leftarrow}{\bot}}}} \widehat{\mathcal{O}_X}$$

Sieve

Definition (Sieve)

Given $X \in \mathrm{ob}\,\mathbf{C}$, a *sieve* on X is a collection S of morphisms of $\mathbf{C}$ with codomain X which is downward closed, i.e. if $f \in S$ and $f \circ g$ is defined, then $f \circ g \in S$.



A maximal (or principal) sieve on X, denoted $\uparrow X$, is a set of all morphisms $\mathrm{Hom}(-, X)$

Example: A sieve in $\mathbf{Poset}$ is an upper set. If X is a poset, and $x, y \in X$, then a maximal sieve on x is given by $\uparrow x = \{y \in X : x \leq y\}$

Remark: Given a sieve S on X and a morphism $f : Y \longrightarrow X$, then the pullback of S along f, $f^*S = \{g : \mathrm{cod}\, g = Y,\ f \circ g \in S\}$ is a sieve on Y.

Grothendieck Topos

"It is the topos theme which is this 'bed' or 'deep river' where come to be married geometry and algebra, topology and arithmetic, mathematical logic and category theory, the world of the 'continuous' and that of 'discontinuous' or discrete structures. It is what I have conceived of most broad to perceive with finesse, by the same language rich of geometric resonances, an 'essence' which is common to situations most distant from each other coming from one region or another of the vast universe of mathematical things."

— *Grothendieck*

Grothendieck Topology

Definition (Grothendieck Topology)

A *Grothendieck topology* on a small category $\mathbf{C}$ is a function J which assigns to each object $X \in \mathrm{ob}\,\mathbf{C}$ a collection $J(X)$ of sieves on X such that:

- (maximality) the maximal sieve $\mathrm{Hom}(-, X) \in J(X)$.
- (stability) If $S \in J(X)$ and $f : Y \to X$, then the pullback $f^*S \in J(Y)$.

$$\begin{array}{ccc} f^*S & \longrightarrow & S \\ \downarrow & \lrcorner & \downarrow \\ \mathrm{y}Y & \xrightarrow[\mathrm{y}f]{} & \mathrm{y}X \end{array}$$

- (local character) If $S \in J(X)$ and T is any sieve on X such that, for any $f : Y \to X$ in S, $f^*T \in J(Y)$, then $T \in J(X)$.

Basis for a Grothendieck Topology

Definition (Basis for a Grothendieck Topology)

A basis for a Grothendieck topology on a category $\mathbf{C}$ is a function K assigning to each object $U \in \mathbf{C}$ a collection $K(U)$ of covering families $\{U_i \to U\}_i$ such that,

- $\{1_U : U \to U\} \in K(U)$.
- If $\{f_i : U_i \to U\}_i \in K(U)$, and $g : V \to U$, then there exists a covering family $\{h_j : V_j \to V\}_j \in K(V)$ s.t. each composite $g \circ h_j$ factors through some f_i.

$$\begin{array}{ccc} V_j & \longrightarrow & U_i \\ {\scriptstyle h_j}\downarrow & & \downarrow{\scriptstyle f_i} \\ V & \xrightarrow[g]{} & U \end{array}$$

- If $\{f_i : U_i \to U\}_{i \in I} \in K(U)$, and $\forall i \in I : \{g_{ij} : U_{ij} \to U_i\}_{j \in I_i} \in K(U_i)$, then $\{f_i \circ g_{ij} : U_{ij} \to U\}_{i \in I, j \in I_i} \in K(U)$.

Remark

- (maximality): any set is covered by itself via the identity map.
- (stability): if $\{U_i\}$ covers U, then $\{U_i \cap V\}$ should cover $U \cap V$.
- (local character): if $\{U_i\}$ covers U and $\{V_{ij}\}_{j \in J_i}$ covers U_i for each i, then the collection $\{V_{ij}\}$ for all i and j should cover U.

Every basis K generates a Grothendieck topology J given by:

$$S \in J(U) \iff S \supset B \text{ for some } B \in K(U)$$

Site

Definition (Site)

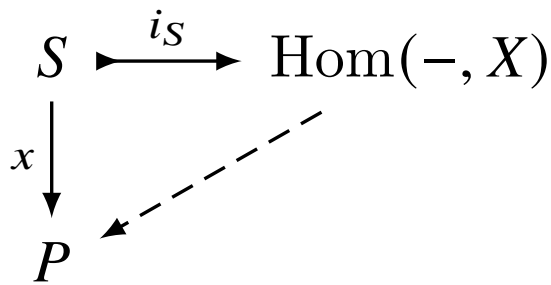
A *site* is a pair $(\mathbf{C}, J)$ where $\mathbf{C}$ is a small category and J is a Grothendieck topology on $\mathbf{C}$.

If $X \in \mathrm{ob}\,\mathbf{C}$ then we call a sieve $S \in J(X)$ a J-covering on X.

Grothendieck Topos

- Given a site $(\mathbf{C}, J)$, a presheaf $P : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$, and a covering sieve S on $X \in \mathrm{ob}\,\mathbf{C}$, a *matching family* for S of elements in P is a function which assigns to each $f : Y \to X$ in S an element $x_f \in P(Y)$ such that $P(g)(x_f) = x_{f \circ g}$ for all $g : Z \to Y$.
- If we view the sieve S as a subobject of the representable $\mathrm{Hom}(-, X)$, then a matching family $(x_f)_{f \in S}$ is precisely a natural transformation $x : S \to P :: f \mapsto x_f$.
- An *amalgamation* of the matching family $(x_f)_{f \in S}$ is an element $x \in P(X)$ such that $P(f)(x) = x_f$ for all $f \in S$.
- Given a site $(\mathbf{C}, J)$, a presheaf P on $\mathbf{C}$ is a *J-sheaf* if every matching family for any J-covering sieve on any object of $\mathbf{C}$ has a unique amalgamation. Equivalently, P is a sheaf iff

$$i_S : \mathrm{Hom}(S, P) \cong \mathrm{Hom}(\mathrm{Hom}(-, X), P)$$



- The category $\mathrm{Sh}(\mathbf{C}, J)$ of sheaves on the site $(\mathbf{C}, J)$ is the full subcategory of $\widehat{\mathbf{C}}$ on the presheaves which are J-sheaves.

Grothendieck Topos

Definition (Grothendieck Topos)

A *Grathendieck topos* is any category equivalent to the category $\mathrm{Sh}(\mathbf{C}, J)$ of sheaves on a site $(\mathbf{C}, J)$.

Theorem

A category $\mathbf{E}$ *is a Grothendieck topos iff there is a small category* $\mathbf{C}$ *and*

$$\mathbf{E} \underset{j_*}{\overset{j^*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \widehat{\mathbf{C}}$$

such that

- $j^* \dashv j_*$
- j_* *is full and faithful*
- j^* *preserves finite limits*

Remark: The category of finite sets, and the category of finite K-sets is an elementary topos but not a Grothendieck topos.

What is a Point of a Topos?

Definition (Point)

A point of a topos $\mathbf{E}$ is a geometric morphism $\mathbf{Set} \to \mathbf{E}$.
A generalized point of a topos $\mathbf{E}$ is a geometric morphism $\mathbf{F} \to \mathbf{E}$.

Example

For any small category $\mathbf{C}$ and any object $A \in \mathbf{C}$, we have a point (f^*, f_*) :

$\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}} \underset{f_*}{\overset{f^*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{Set}$ of the topos $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$, whose inverse image

$f^* = \varepsilon_A : F \mapsto F(A)$ is the evaluation functor at A.

Example

- Given a topological space X, the usual notion of covering gives rise to the Grothendieck topology J_X on the poset category $\mathcal{O}_X$ for a sieve $S = \{U_i \hookrightarrow U : i \in I\}$ on $U \in \mathrm{ob}(\mathcal{O}_X)$,

$$S \in J_X(U) \iff \bigcup_{i \in I} U_i = U$$

- Let $\mathbf{E}_X$ be the topos of sheaves on the site $(\mathcal{O}_X, J_X)$.
- Any element $x \in X$ gives a point

$$\mathbf{E}_X \underset{x_*}{\overset{x^*}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbf{Set}$$

- More generally, any continuous function $Y \xrightarrow{f} X$ gives a point

$$\mathbf{E}_X \underset{f_*}{\overset{f^*}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbf{E}_Y$$

- If $Y = \{\bullet\}$, then a point of X in the usual sense gives a point of $\mathbf{E}_X$.

Example

- Let M be a topological monoid. A continuous action of M on a set X is an action:

$$M \times X \xrightarrow{\alpha} X$$

  which is continuous when X is given the discrete topology.
- The category of continuous actions of M on sets is a Grothendieck topos, $\mathrm{Cont}(M)$.
- The forgetful functor $\mathrm{Cont}(M) \to \mathbf{Set}$ has a right adjoint and preserves finite limits, so determines a point, p.

Subobject Classifier in Topos of Sheaves

$$\Omega(X) = \{S : S \text{ is a sieve on } X\}$$

for $Y \xrightarrow{f} X$, $\Omega_f : \Omega(X) \to \Omega(Y)$ has

$$\Omega_f(S) = \left\{Z \xrightarrow{g} Y : f \circ g \in S\right\}$$

$\top : 1 \to \Omega$ has component $\top_X : \{\bullet\} \to \Omega(X)$ given by

$$\top_X(\bullet) = \mathrm{Hom}(-, X)$$

The subobject classifier Ω in $\mathbf{Sh}(X, \mathcal{O}_X)$ is the sheaf that assigns to $U \in \mathcal{O}_X$ the set of open subsets of U:

$$\Omega : U \mapsto \left\{V \in \mathcal{O}_X : V \subset U\right\}$$

Remark Truth values are open sets. The point is that the truth values in the topos of sheaves $\mathbf{Sh}(X, \mathcal{O}_X)$ on a space $(X, \mathcal{O}_X)$ are the open sets of that space. When someone says "is property P true?", the answer is not yes or no, but "it is true on the open subset U". If this $U = X$, then P is really true; if $U = \emptyset$, then P is really false. But in general, it's just true some places and not others.

Quantifiers $\forall, \exists$ in a Sheaf Topos

$$\begin{array}{ccc} \{x \mid \forall y \varphi\} & \longrightarrow & 1 \\ \downarrow & \lrcorner & \downarrow{\scriptstyle \top^Y} \\ X & \xrightarrow[\hat{\varphi}]{} & \Omega^Y \end{array}$$

where $\hat{\varphi} : X \to \Omega^Y$ is the currying of $\varphi : X \times Y \to \Omega$, and $\top^Y$ is the currying of the composite $1 \times Y \xrightarrow{!} 1 \xrightarrow{\top} \Omega$.

$$\begin{array}{ccc} \{(x, y) \mid \varphi\} & \longrightarrow & X \times Y \\ \downarrow & & \downarrow{\scriptstyle \pi_X} \\ \{x \mid \exists y \varphi\} & \longrightarrow & X \end{array}$$

Morita Equivalence

Definition (Morita Equivalence)

First order theories $\mathbb{T}_1$ and $\mathbb{T}_2$ are called Morita equivalent iff their classifying toposes $\mathbf{E}_{\mathbb{T}_1} \cong \mathbf{E}_{\mathbb{T}_2}$.

Remark

Every first-order theory $\mathbb{T}$ has a unique "classifying topos" $\mathbf{E}_{\mathbb{T}}$, whose "points" can be interpreted as the "models" of the theory $F : \mathbb{T} \to \mathbf{Set}$.

Remark

Every Grothendieck topos can be studied syntactically, geometrically and semantically:

- *syntactically through any "first-order theory" which it classifies,*
- *geometrically through its structure as a category, consisting of objects, arrows and a law of composition for arrows,*
- *semantically through its "points", interpreted as the "models" of any theory which it classifies.*

Point-free Topology — A Bird's Eye View

Point-set Topology

- Point = Element of a set
- Space = A set of points, along with a set of open sets satisfying some specific axioms
- Continuous Maps = A function $f : X \longrightarrow Y$ such that $f^{-1}(U)$ is open for any open set $U \subset Y$

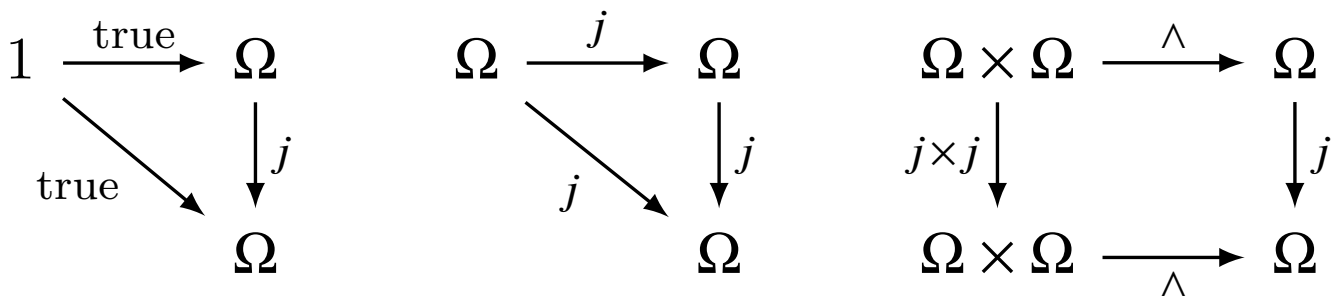
Point-free Topology

- Point = Model of a geometric theory
- Space = The 'World' in which the point lives with other points (i.e. a Grothendieck topos)
- Continuous Maps = A geometric morphism $f : \mathbf{E} \longrightarrow \mathbf{F}$ such that $f^* : \mathbf{F} \longrightarrow \mathbf{E}$ preserves finite limits and small colimits

Definition (Lawvere-Tierney Topology)

A *Lawvere-Tierney topology* on a topos $\mathbf{E}$ is a morphism $j : \Omega \to \Omega$ such that

1. j preserves truth: $j \circ \mathrm{true} = \mathrm{true}$.
2. j is idempotent: $j \circ j = j$.
3. j preserves intersections: $j \circ \wedge = \wedge \circ (j \times j)$.



Remark

1. $A \implies \Box A$
2. $\Box\Box A \implies \Box A$
3. $\Box(A \wedge B) \implies \Box A \wedge \Box B$

Universal Closure Operator

Definition (Universal Closure Operator)

A *universal closure operator* on a topos $\mathbf{E}$ is given by, for each object X, a morphism $\mathrm{cl} : \mathrm{Sub}(X) \to \mathrm{Sub}(X)$ such that

1. $A \leq \mathrm{cl}(A)$
2. $\mathrm{cl}(A) = \mathrm{cl}(\mathrm{cl}(A))$
3. $A \leq B \implies \mathrm{cl}(A) \leq \mathrm{cl}(B)$
4. for $f : Y \to X$ and $A \in \mathrm{Sub}(X)$, $f^*(\mathrm{cl}(A)) = \mathrm{cl}(f^*(A))$

Theorem

The following are equivalent:

1. *A Grothendieck topology on* $\mathbf{C}$.
2. *A Lawvere-Tierney topology on* $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$.
3. *A universal closure operator on* $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$.
4. *A full subcategory* $\mathbf{E}$ *of* $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ *such that the inclusion* $\mathbf{E} \hookrightarrow \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ *has a left adjoint which preserves finite limits.*

$$\mathbf{E} \underset{i}{\overset{F}{\underset{\hookrightarrow}{\overset{\leftarrow}{\bot}}}} \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$$

Definition

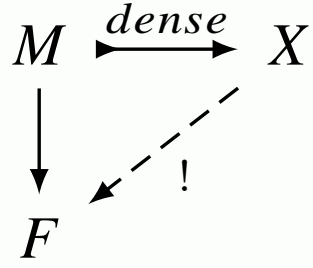
Given a Lawvere-Tierney topology j with associated closure operator on $\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$, a subobject $M \in \mathrm{Sub}(X)$ is called

- *dense* if $\mathrm{cl}(M) = X$.
- *colosed* if $\mathrm{cl}(M) = M$.

A presheaf F is a *j-sheaf* iff for every dense $M \rightarrowtail X$ the induced morphism

$$\mathrm{Hom}(X, F) \to \mathrm{Hom}(M, F)$$

is an isomorphism.



F is called *j-separated* if for every dense $M \rightarrowtail X$ the induced morphism

$$\mathrm{Hom}(X, F) \to \mathrm{Hom}(M, F)$$

is a monomorphism.

Abelian Category

Definition (Abelian Category)

A category is abelian if

- it has a zero object 0,
- it has all binary products and binary coproducts,
- it has all kernels and cokernels, and
- all monomorphisms and epimorphisms arise as kernels or cokernels, respectively.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

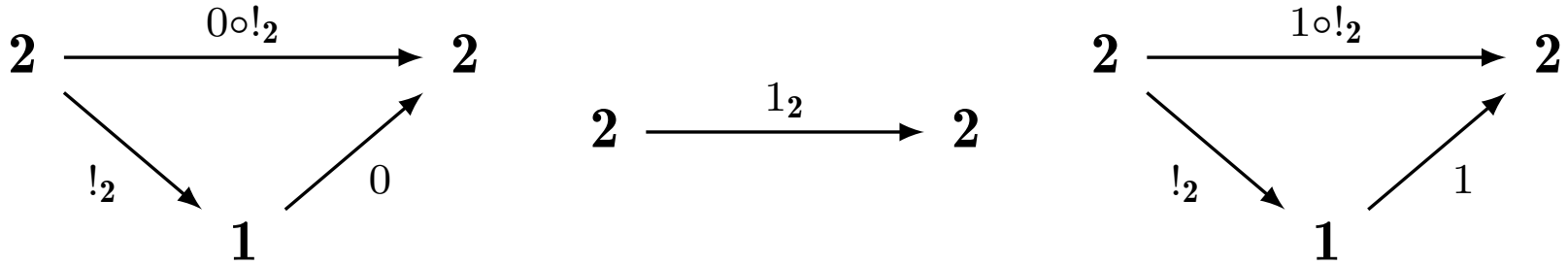
Category of Categories as Foundations CCAF

Lawvere's aim is to provide an axiomatization of category of categories as a foundation for mathematics in first-order language such that:

1. A model of the axioms should be a category.
2. The objects of the model should themselves be categories.
3. The basic properties of category theory can be proved, e.g. functor categories and adjoint functors should be definable, Yoneda's lemma and the adjoint functor theorem should be provable, etc.
4. Sets (all usual mathematical objects) should be definable, and all their usual properties provable.
5. It should be possible to make a distinction between small and large categories, and Grothendieck universes should be models of the theory.

CCAF

- There is a terminal category $\mathbf{1}$.
- There is a category $\mathbf{2}$ that has exactly two functors $0 \neq 1 : \mathbf{1} \to \mathbf{2}$ and three functors $\mathbf{2} \to \mathbf{2}$.



- $\mathbf{2}$ is a universal generator: $\forall F \neq G : \mathbf{A} \to \mathbf{B} \exists f : \mathbf{2} \to \mathbf{A}[Ff \neq Gf]$

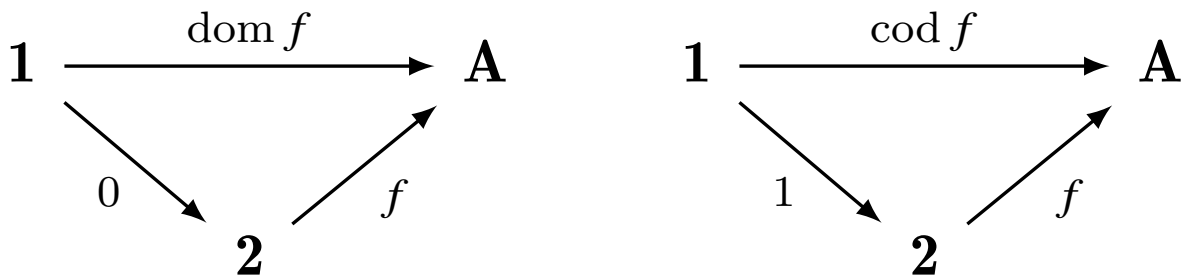
$$\mathbf{2} \xrightarrow{f} \mathbf{A} \underset{G}{\overset{F}{\rightrightarrows}} \mathbf{B}$$

if $\mathbf{C}$ has the same property, then $\exists g : \mathbf{2} \to \mathbf{C} \exists h : \mathbf{C} \to \mathbf{2} : hg = 1_{\mathbf{2}}$.

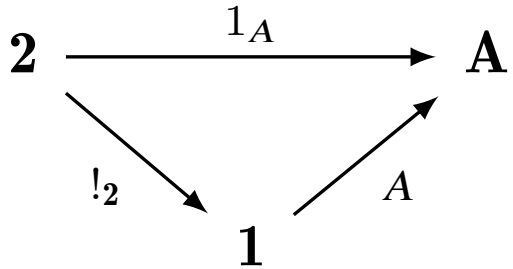
$$\mathbf{2} \underset{h}{\overset{g}{\rightleftarrows}} \mathbf{C}$$

Definition:

- an object A of $\mathbf{A}$ means $A : \mathbf{1} \to \mathbf{A}$.
- a morphism f of $\mathbf{A}$ means $f : \mathbf{2} \to \mathbf{A}$.
- The domain $\mathrm{dom}\, f$ and codomain $\mathrm{cod}\, f$ of f in $\mathbf{A}$ are defined to be the composites of f with $0 : \mathbf{1} \to \mathbf{2}$ and $1 : \mathbf{1} \to \mathbf{2}$.

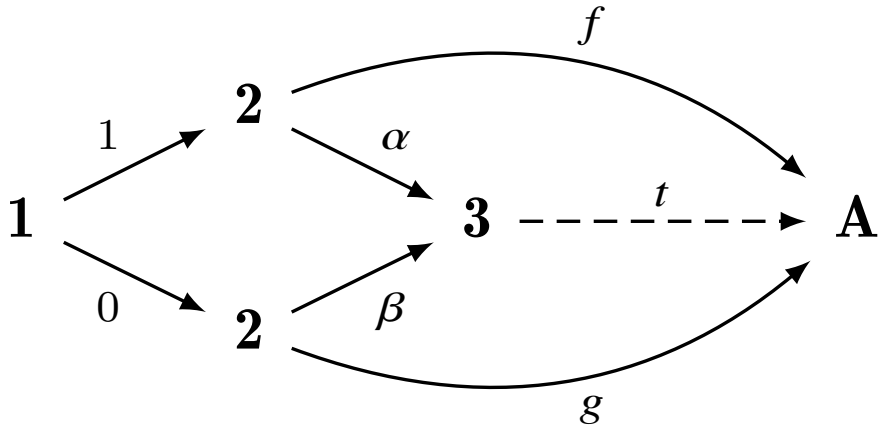


- The identity morphism 1_A on A is defined to be the composite $A \circ !_{\mathbf{2}}$



- A discrete category is defined as a category $\mathbf{C}$ such that each morphism $\mathbf{2} \to \mathbf{C}$ is the identity morphism 1_A for some object $A : \mathbf{1} \to \mathbf{C}$.

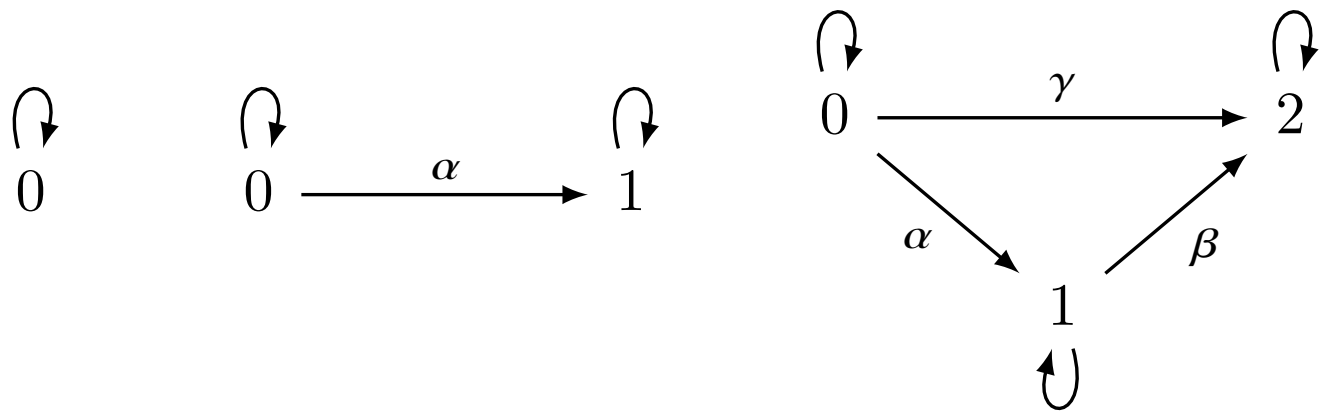
- There is a category $\mathbf{3}$ that is a pushout, and there is a functor $\gamma : \mathbf{2} \to \mathbf{3}$ with $\operatorname{dom}\gamma = \operatorname{dom}\alpha$ and $\operatorname{cod}\gamma = \operatorname{cod}\beta$.



Definition:

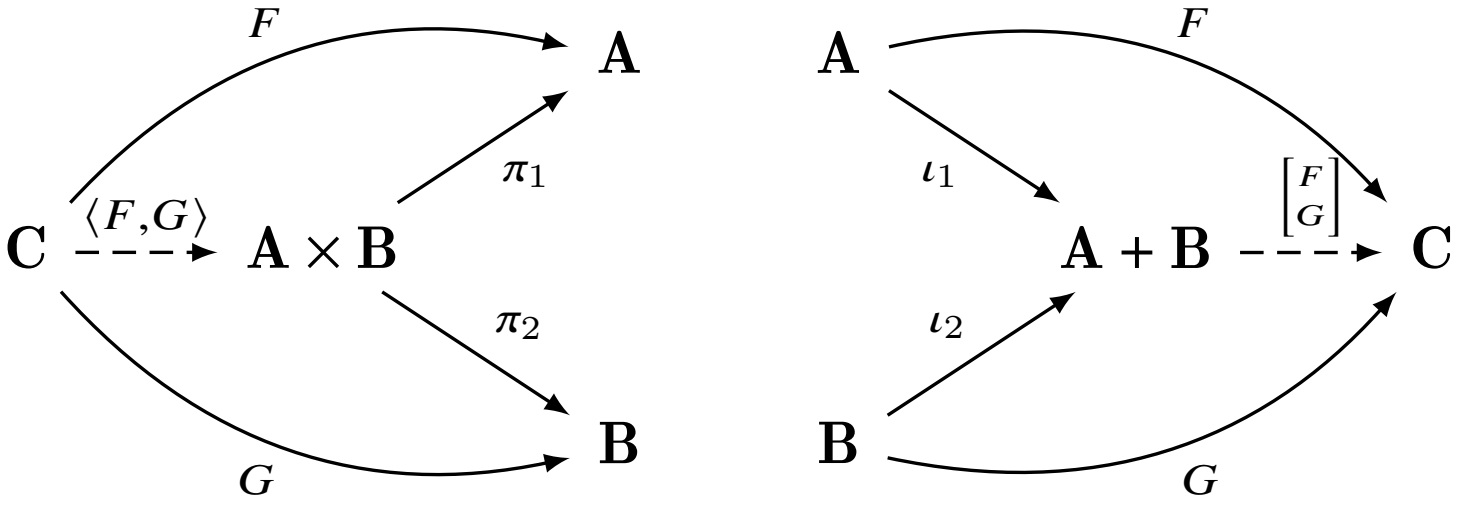
- given any $f : \mathbf{2} \to \mathbf{A}$ and $g : \mathbf{2} \to \mathbf{A}$ with $\operatorname{cod} f = \operatorname{dom} g$, the composition gf in $\mathbf{A}$ is $t\gamma$.

Remark: $\mathbf{1}, \mathbf{2}, \mathbf{3}$ naively look like this:

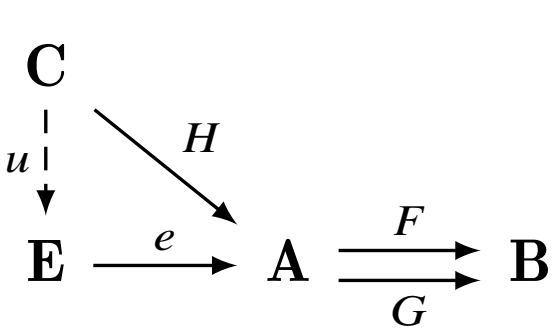


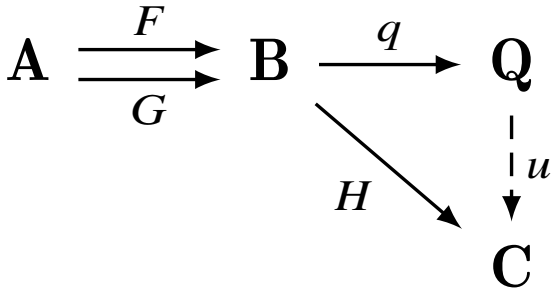
CCAF

- There is an initial category $\mathbf{0}$.
- Every pair of categories $\mathbf{A}$ and $\mathbf{B}$ has a product $\mathbf{A} \times \mathbf{B}$ and a coprodcut $\mathbf{A} + \mathbf{B}$.



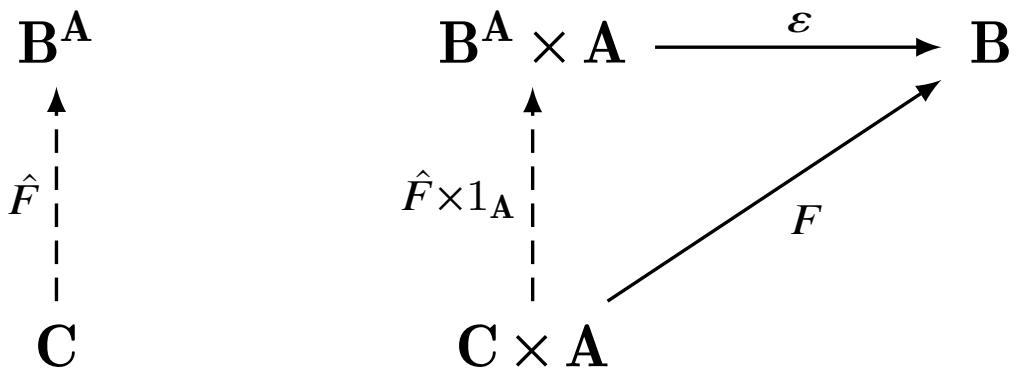
- Every parallel pair of functors $F, G : \mathbf{A} \to \mathbf{B}$ has an equilizer and a coequalizer.





- There is a functor category $\mathbf{B}^{\mathbf{A}}$ from any category $\mathbf{A}$ to any category $\mathbf{B}$.

$$\forall \mathbf{C} \forall F : \mathbf{C} \times \mathbf{A} \to \mathbf{B} \; \exists! \; \hat{F} : \mathbf{C} \to \mathbf{B}^{\mathbf{A}} \left[\varepsilon \circ (\hat{F} \times 1_{\mathbf{A}}) = F \right]$$



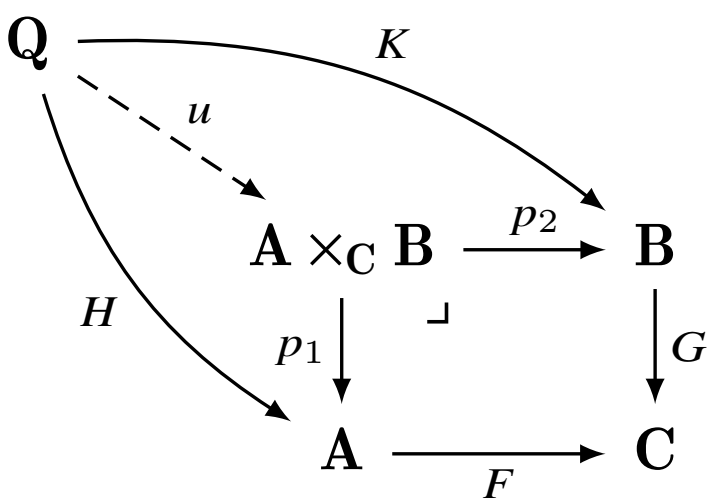
CCAF

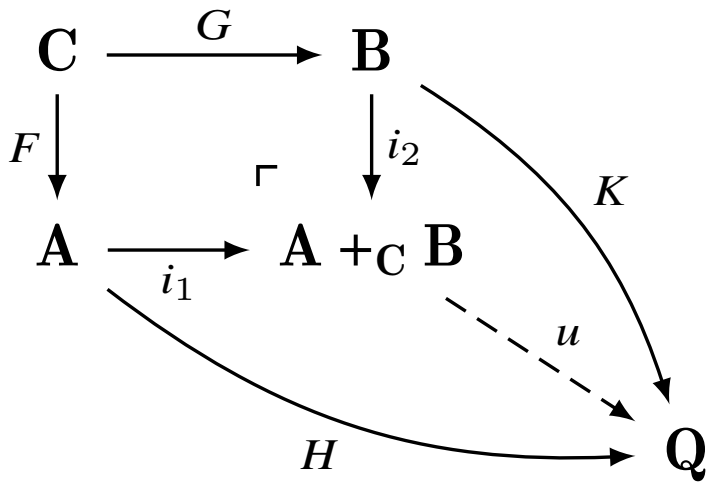
- Every two functors $F : \mathbf{A} \to \mathbf{C}, G : \mathbf{B} \to \mathbf{C}$ have a pullback.

$$\forall H : \mathbf{Q} \to \mathbf{A}, K : \mathbf{Q} \to \mathbf{B}[FH = GK \to \exists! u(p_1 u = H \;\&\; p_2 u = K)]$$

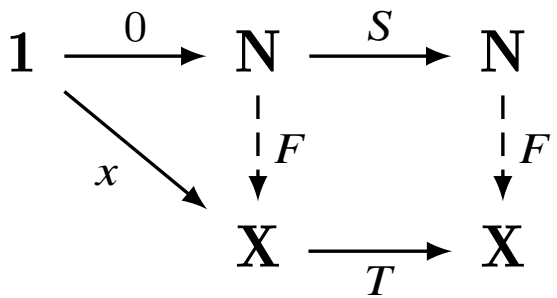
- Every two functors $F : \mathbf{C} \to \mathbf{A}, G : \mathbf{C} \to \mathbf{B}$ have a pushout.

$$\forall H : \mathbf{A} \to \mathbf{Q}, K : \mathbf{B} \to \mathbf{Q}[HF = KG \to \exists! u(ui_1 = H \;\&\; ui_2 = K)]$$





- There is a natural numbers object NNO category.



- Choice. For $F : \mathbf{A} \to \mathbf{B}$ such that $\mathbf{A} \ncong \mathbf{0}$ and $\mathbf{B}$ is discrete, there exists $G : \mathbf{B} \to \mathbf{A}$ such that $FG = 1_{\mathbf{B}}$.

CCAF axioms on $\mathbf{Set}$

- There is a category $\mathbf{Set}$ whose objects and morphisms sastify the ETCS axioms.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

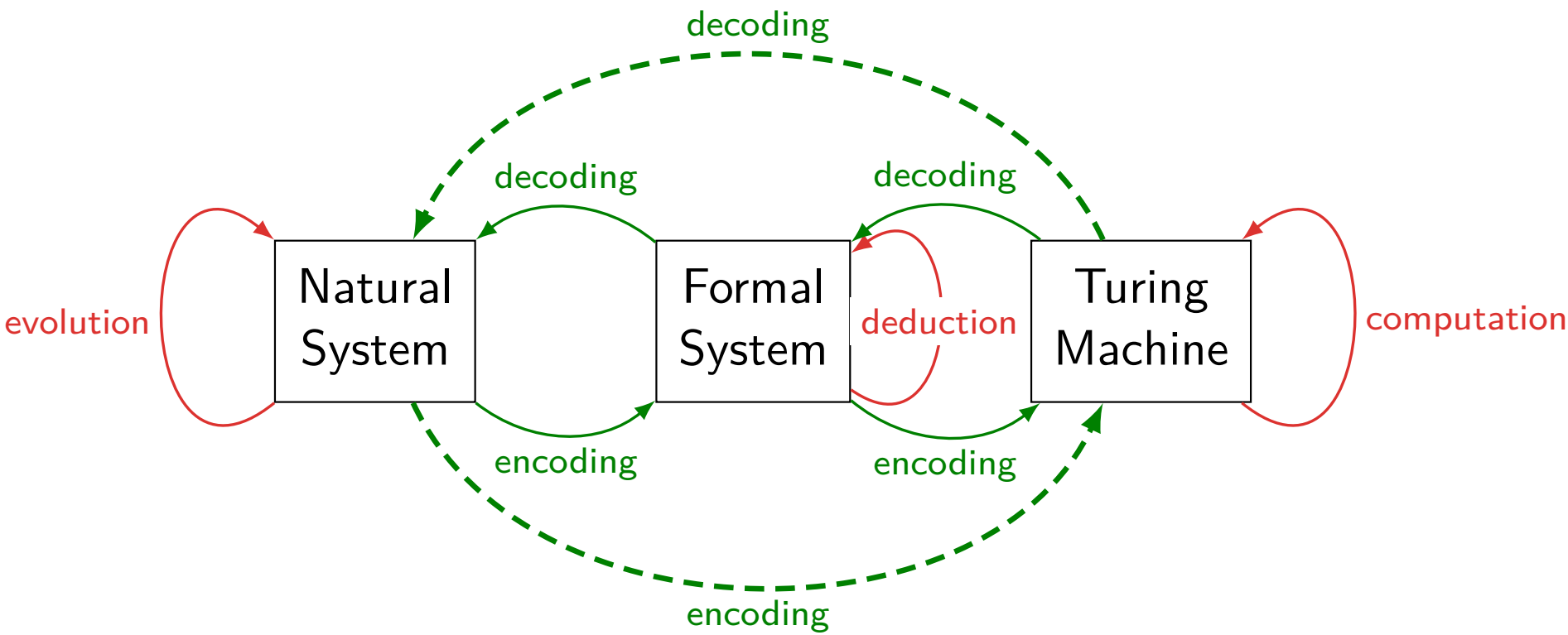
Quantum Computing

Answers to the Exercises

References2039

Rosen's Modeling Relation & Church-Turing Thesis

Is every natural law simulable?

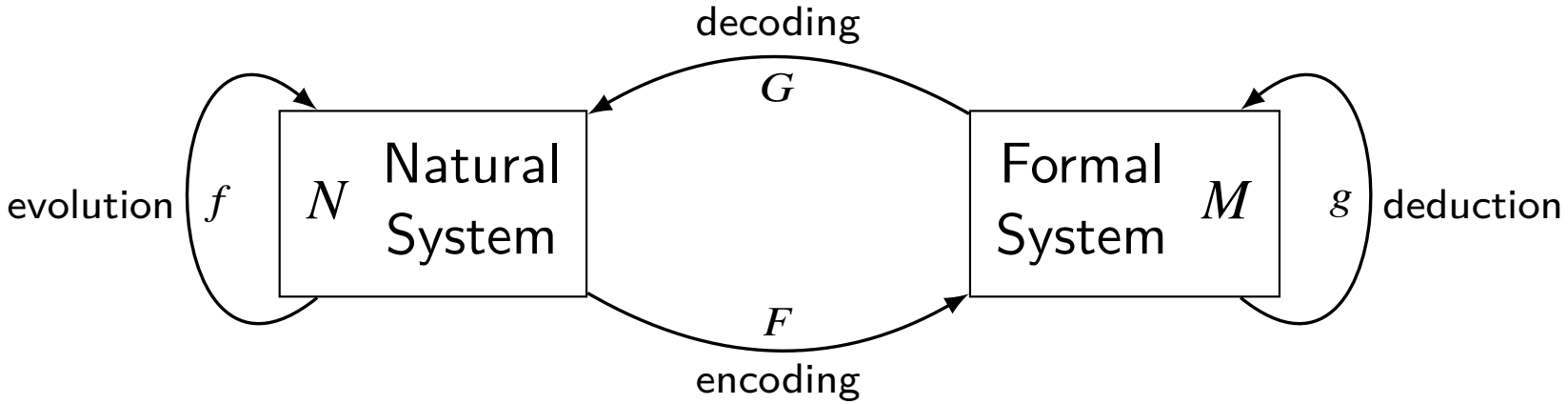


Simulation vs Model

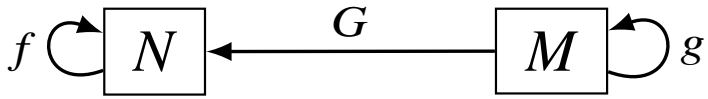
- Simulation describes the trajectories (e.g., curve-fitting)
- Model explains the principle of the dynamics (e.g., Newton)

Rosen's "Simulation" / "Metaphor" / "Model"[Lou09]

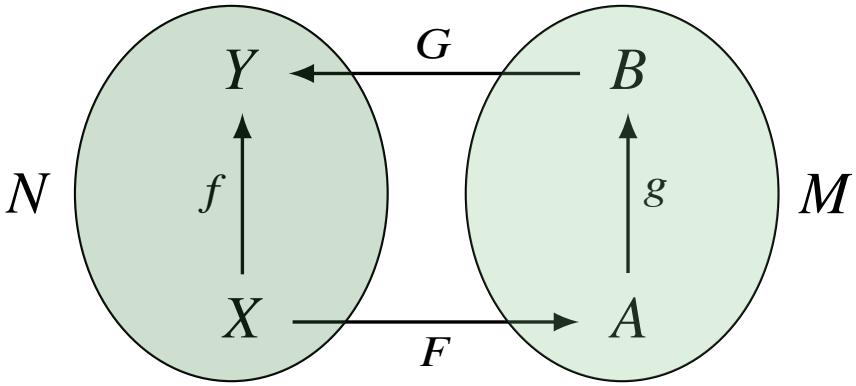
- **Simulation:** M is a simulation of N iff $f = G \circ g \circ F$

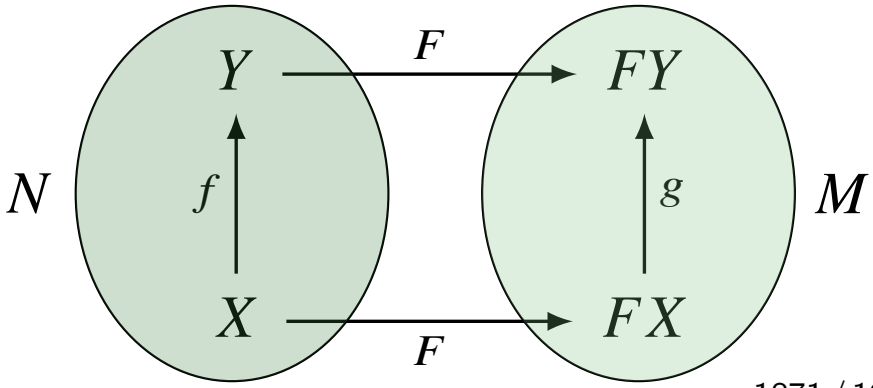


- **Metaphor:** M is a metaphor of N iff there is no encoding arrow

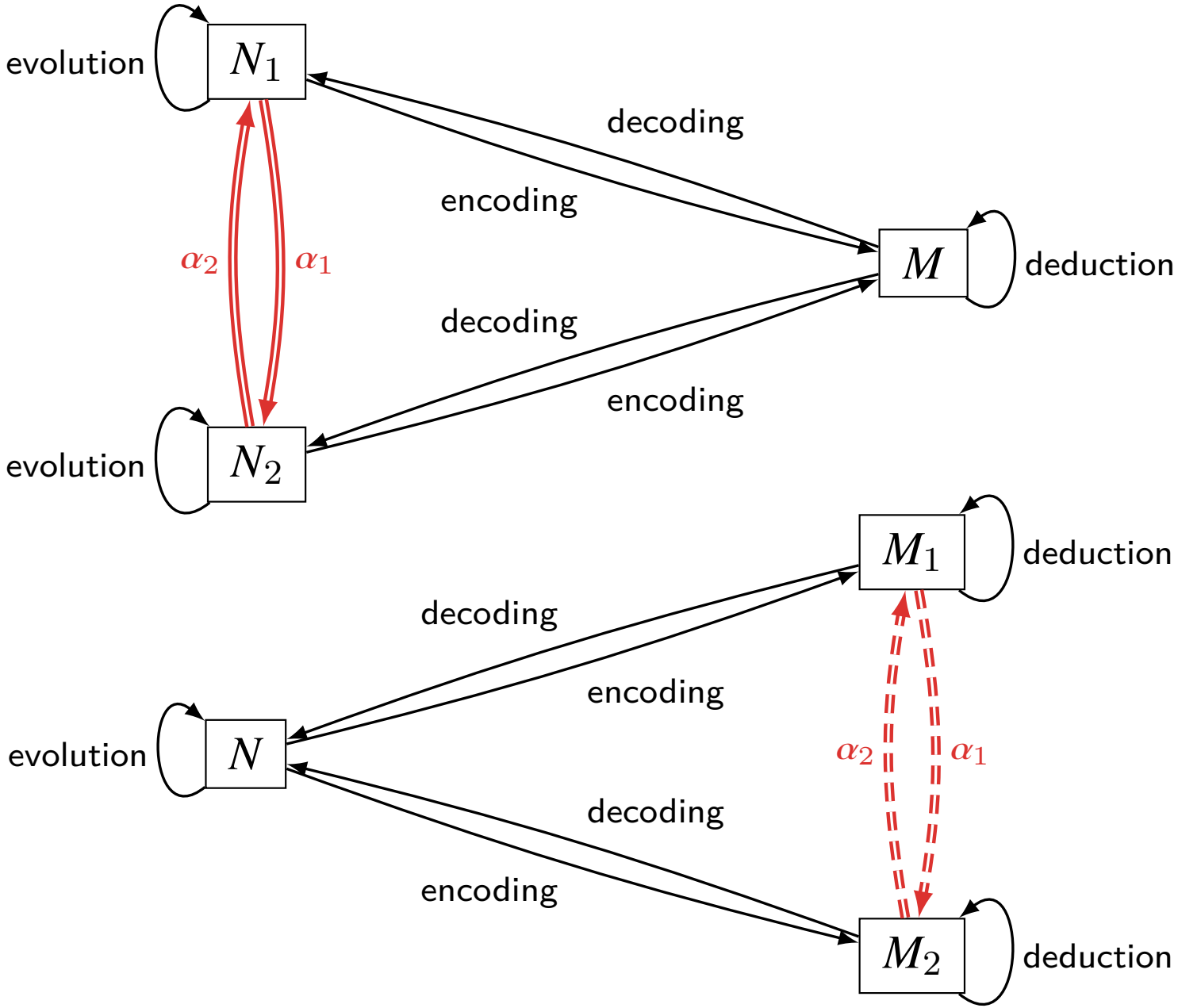


- **Model:** M is a model of N iff M is a simulation of N and $g = Ff$

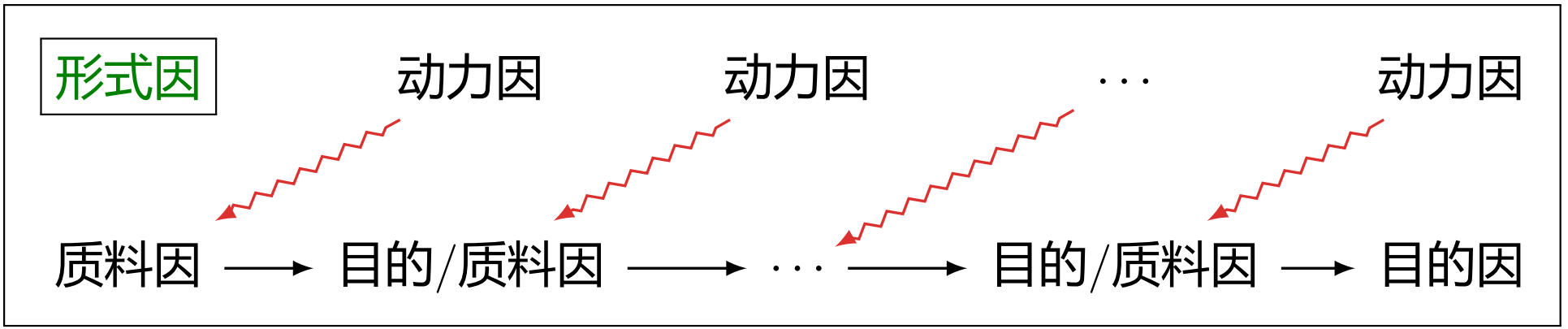




Rosen's "Analogy" as Natural Transformation

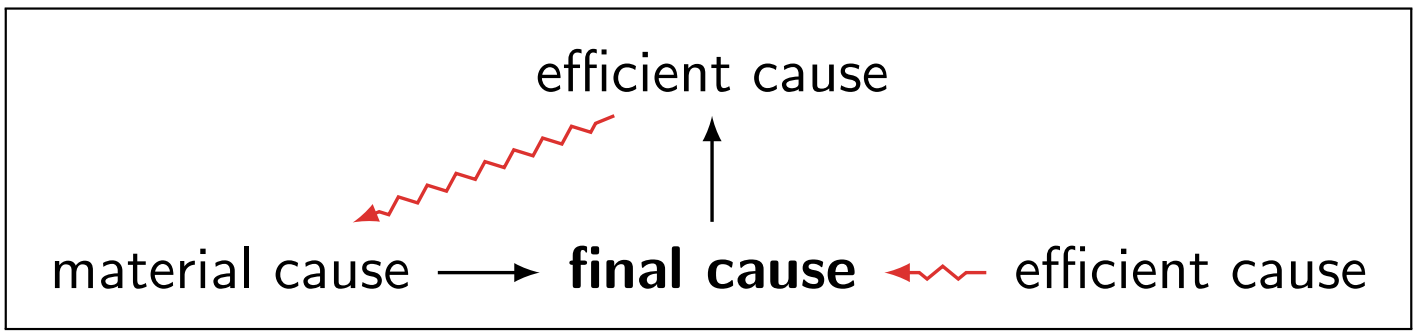


From Aristotle's Four Causes to Rosen's "Life"



*"A living system is a **system closed to efficient causation**, i.e., its every efficient cause is entailed within the system."*

— Robert Rosen





Mechanism or Organism?



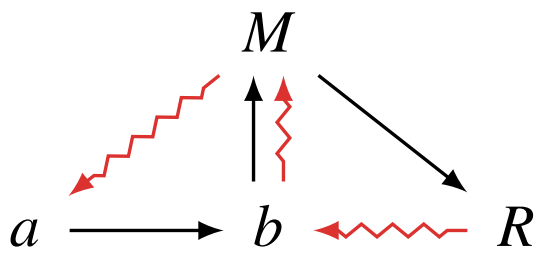
infinite regress?

closure to efficient causation

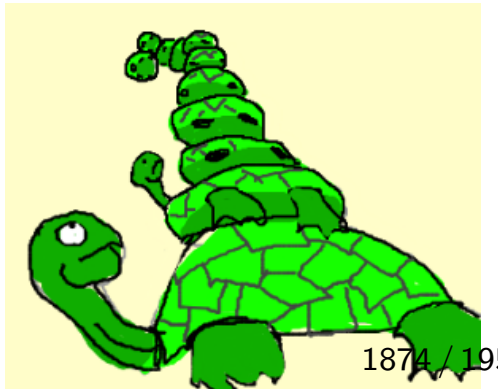
Mr. Why?

1. Mr. Why: "why b?"
2. Rosen: $b = M(a)$
   - 2.1 "because a", this is the "material cause"
   - 2.2 "because M", this is the "efficient cause"
3. Mr. Why: "why M?" — within physics there is not really any answer, other than that this just is a natural law.
4. Rosen: "because R": $R(b) = M$
5. Mr. Why: "why R?"
6. Rosen: "because β": $\beta(M) = R$
7. Mr. Why: "why β?"
8. Rosen: "because M": $\beta \cong b$ and $M(a) = b$

$$A \xrightarrow{M} B \xrightarrow{R} \mathrm{Hom}(A,B) \xrightarrow{\beta} \mathrm{Hom}(B,\mathrm{Hom}(A,B))$$



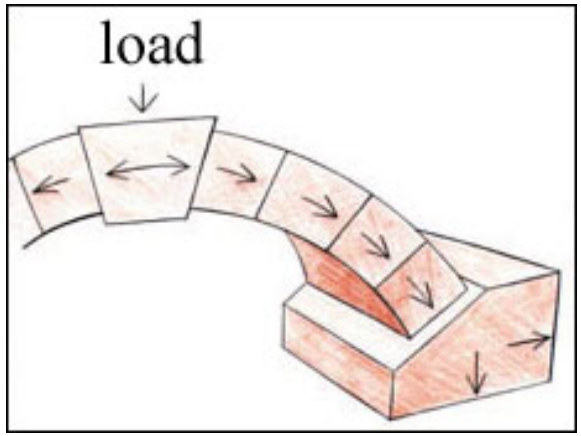
Remark: b is the material cause entailing its own efficient cause M which entails b as its final cause.



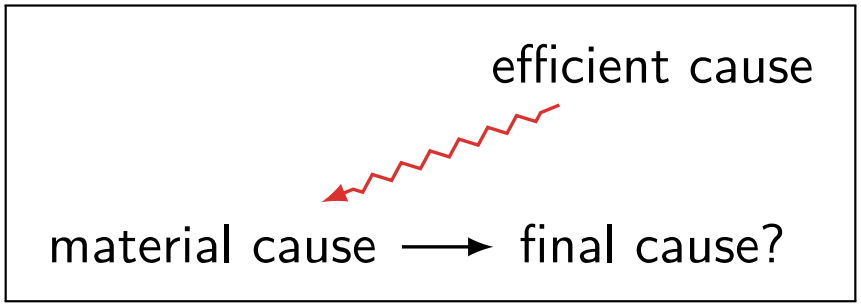
- Marco Polo describes a bridge, stone by stone.
- Kublai Khan: ‘But which is the stone that supports the bridge?’
- Marco Polo: ‘The bridge is not supported by one stone or another, but by the line of the arch that they form.’
- Kublai Khan: ‘Why do you speak to me of the stones? It is only the arch that matters to me.’
- Marco Polo: ‘Without stones there is no arch.’

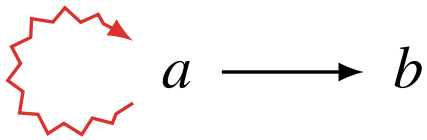
— *Italo Calvino: Invisible Cities*





Example

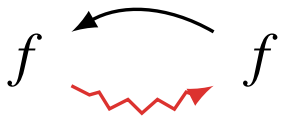




Consider $\chi_{\{a\}} : A \to B$, where $B = \{0, 1\}$, and

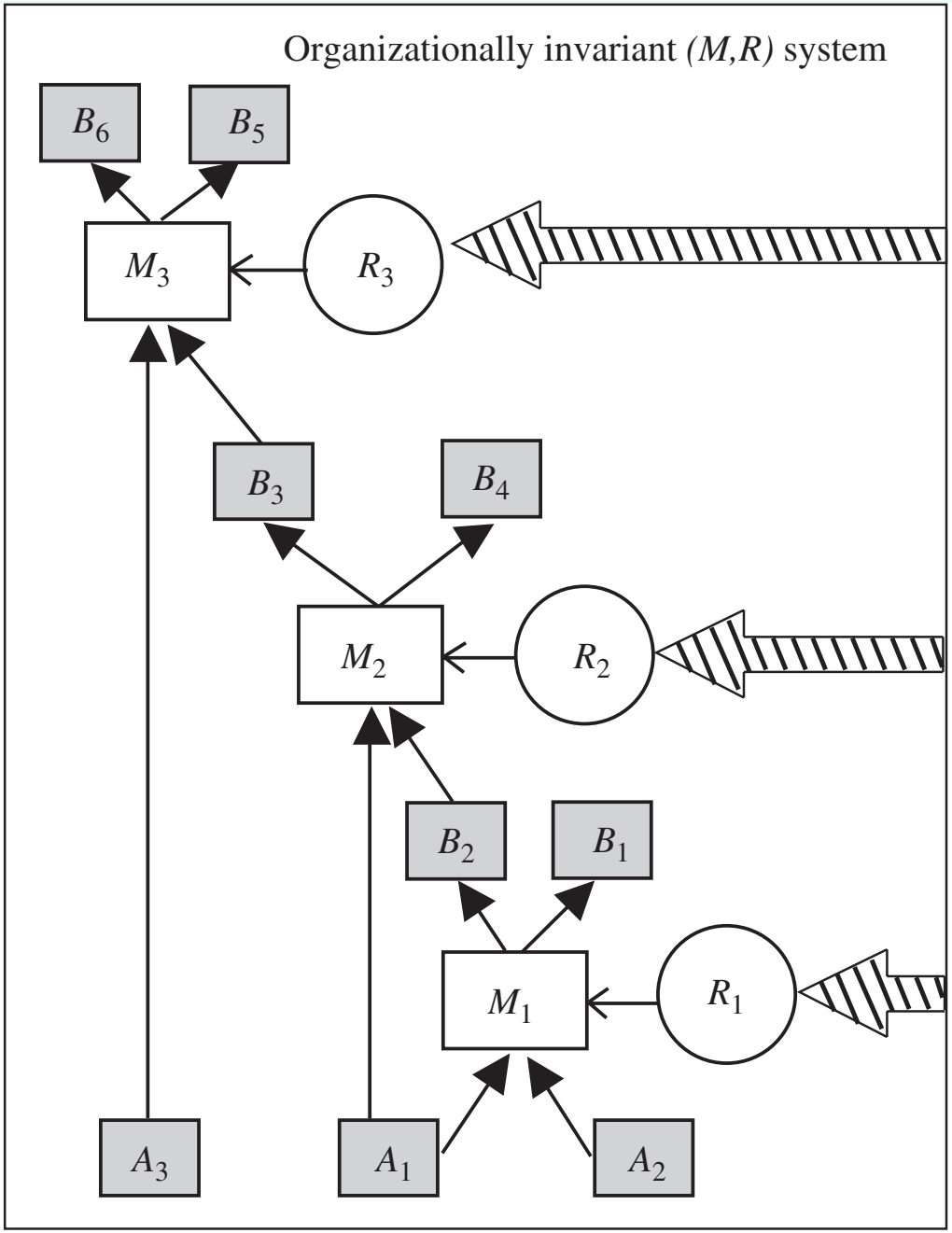
$$\chi_{\{a\}}(x) := \begin{cases} 1 & x = a \\ 0 & \text{otherwise} \end{cases}$$

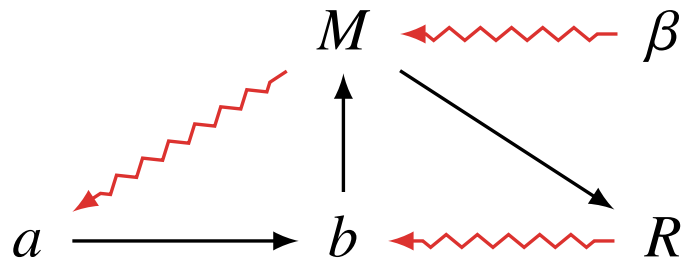
Then $a \cong \chi_{\{a\}}$.

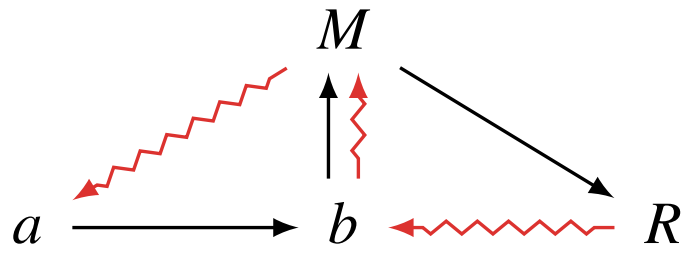


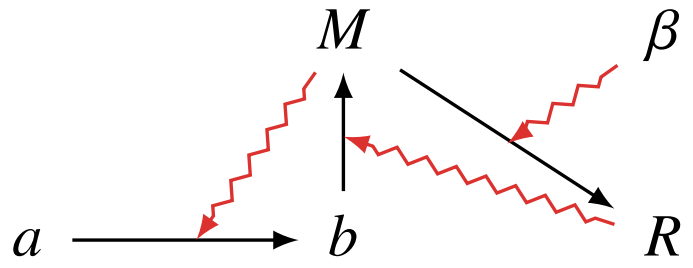
$f(f) = f$ is an impossibility in **Set**.

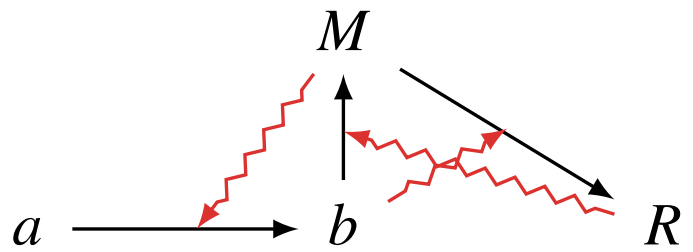
Rosen's Metabolism-Repair (M, R)-System[Ros72]





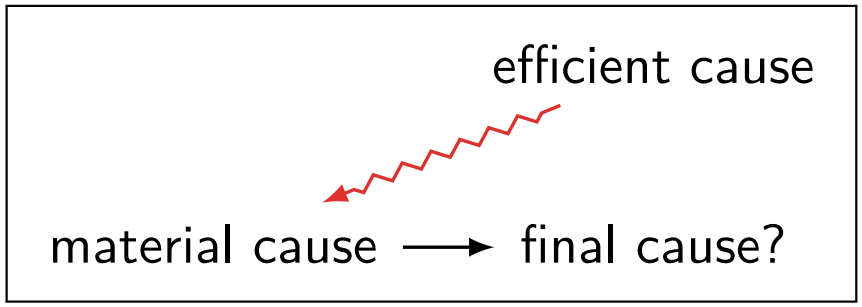




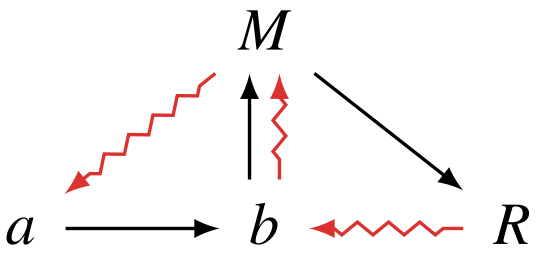


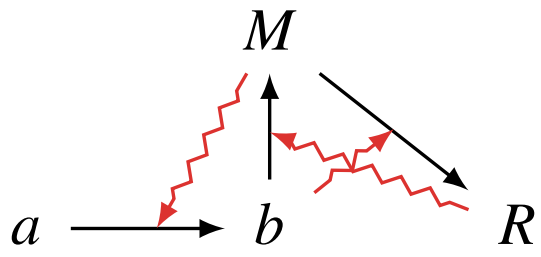
Rosen: "What is Life?"[Luz+09][35]

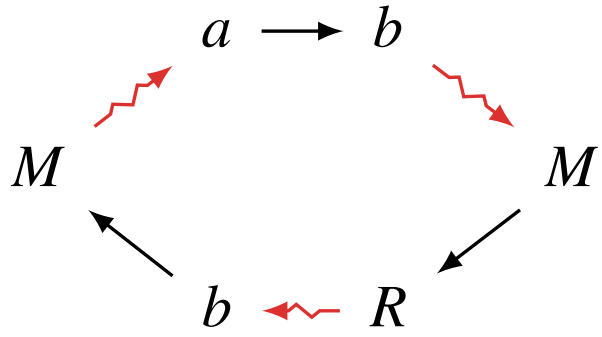
- M: metabolism $M(a) = b$
- R: repair $R(b) = M$
- β: replication $\beta(M) = R$



$$A \xrightarrow{M} B \xrightarrow{R} \mathrm{Hom}(A,B) \xrightarrow{\beta} \mathrm{Hom}(B,\mathrm{Hom}(A,B))$$







Assumption: The evaluation map
$\varepsilon_b : \mathrm{Hom}(B,\mathrm{Hom}(A,B)) \to \mathrm{Hom}(A,B) :: \varepsilon_b(R) = R(b)$ is invertible.

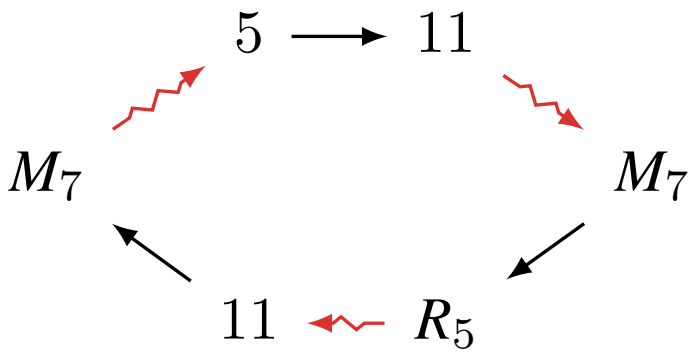
Then $\varepsilon_b^{-1}(M) = R$. Thus, we can set $\beta = \varepsilon_b^{-1}$, i.e., β is determined by b.

[35] Luz Cárdenas *et al*: Closure to efficient causation, computability and artificial life. 2009.

An Arithmetical Example

$$A \xrightarrow{M} B \xrightarrow{R} \mathrm{Hom}(A,B) \xrightarrow{\beta} \mathrm{Hom}(B,\mathrm{Hom}(A,B))$$

- Let $A = B = \mathbb{Z}_{12}$
- $\mathrm{Hom}(A,B) = \{M_n : n \in A\}$ where $M_n(a) \coloneqq na \mod 12$
- $\mathrm{Hom}(B,\mathrm{Hom}(A,B)) = \{R_k : k \in A\}$ where $R_k(b) \coloneqq M_{bk}$
- $\beta(M_n) \coloneqq R_{11n}$
- Let us choose $a = 5, M = M_7$



Remark

Since $\forall R \exists M : \beta(M) = R$, then

$$\hat{\beta}(-, M) = R(-)$$

Given M, there exists R' and b such that $R'(b) = M$.
Take $R \coloneqq G \circ \hat{\beta} \circ \langle 1_B, R' \rangle$.

$$\begin{array}{ccc} B \times \mathrm{Hom}(A,B) & \xrightarrow{\hat{\beta}} & \mathrm{Hom}(A,B) \\ {\scriptstyle \langle 1_B, R' \rangle}\uparrow & & \downarrow{\scriptstyle G} \\ B & \xrightarrow[R]{} & \mathrm{Hom}(A,B) \end{array}$$

$$G(\hat{\beta}(b, M)) = G \circ \hat{\beta} \circ \langle 1_B, R' \rangle(b) = R(b) = \hat{\beta}(b, M)$$

Therefore, $\hat{\beta}(b, M)$ is a fixpoint of G.

If we take $\beta = b$, then

$$Ma = b$$
$$Rb = M$$
$$bM = R$$

By substitution, we get

$$((Ma)M)(Ma) = M$$

Let

$$G := \lambda x.((xa)x)(xa)$$

then $GM = M$.
We have $\mathbf{Y}G = G(\mathbf{Y}G)$ with the $\mathbf{Y}$ combinator.
Then, M, b, R are fully determined by a.

$$M = \mathbf{Y}G$$
$$b = \mathbf{Y}Ga$$
$$R = (\mathbf{Y}Ga)(\mathbf{Y}G)$$

Ouroboros Equation $f(f) = f$ [Let+06]

$$A \xrightarrow{M} B \xrightarrow{R} \mathrm{Hom}(A,B) \xrightarrow{\beta} \mathrm{Hom}(B,\mathrm{Hom}(A,B))$$

$M(a) = b$ with $M \in \mathrm{Hom}(A,B)$

$R(b) = M$ with $R \in \mathrm{Hom}(B,\mathrm{Hom}(A,B))$

$\beta(M) = R$ with $\beta \in \mathrm{Hom}(\mathrm{Hom}(A,B),\mathrm{Hom}(B,\mathrm{Hom}(A,B)))$

$C_0 := A$

$C_1 := B$

$C_2 := \mathrm{Hom}(C_0, C_1)$

$C_n := \mathrm{Hom}(C_{n-2}, C_{n-1})$

$f_0 := M$

$f_1 := R$

$f_2 := \beta$

$c_0 := a$

$c_1 := b = f_0(c_0)$

$c_2 := M = f_1(c_1) = f_0$

$c_3 := R = f_2(c_2) = f_1$

$c_{n+1} := f_n(c_n) = f_{n-1}$

$$C_0 \xrightarrow{f_0} C_1 \xrightarrow{f_1} C_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} C_n \xrightarrow{f_n} C_{n+1} = \mathrm{Hom}(C_{n-1}, C_n)$$

$$c_0 \longmapsto c_1 \longmapsto c_2 \longmapsto \cdots \longmapsto c_n \longmapsto c_{n+1} = f_{n-1} = f_n(f_{n-2})$$

Ouroboros Equation $f(f) = f$

$$C_0 \xrightarrow{f_0} C_1 \xrightarrow{f_1} C_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} C_n \xrightarrow{f_n} C_{n+1} = \mathrm{Hom}(C_{n-1}, C_n)$$

$$c_0 \longmapsto c_1 \longmapsto c_2 \longmapsto \cdots \longmapsto c_n \longmapsto c_{n+1} = f_{n-1} = f_n(f_{n-2})$$

$$C_\infty \coloneqq \lim_{n\to\infty} C_n = \mathrm{Hom}(C_\infty, C_\infty)$$

$$f_\infty \coloneqq \lim_{n\to\infty} f_n \implies f_\infty(f_\infty) = f_\infty$$

$$f_\infty \in \mathrm{Hom}(C_\infty, C_\infty) = C_\infty$$

$$f_n(c_n) = f_{n-1} \implies \varepsilon_{c_n}(f_n) = f_{n-1}$$

$$p_n \coloneqq \varepsilon_{c_{n-1}} : C_n \leftarrow C_{n+1} :: p_{n+1}(f_n) = f_{n-1}$$

$$C_1 \xleftarrow{p_1} C_2 \xleftarrow{p_2} \cdots \xleftarrow{p_{n-1}} C_n \xleftarrow{p_n} C_{n+1} \xleftarrow{p_{n+1}} C_{n+2} \xleftarrow{p_{n+2}} \cdots$$

$$c_1 \xmapsto{p_1} f_0 \xmapsto{p_2} \cdots \xmapsto{p_{n-1}} f_{n-2} \xmapsto{p_n} f_{n-1} \xmapsto{p_{n+1}} f_n \xmapsto{p_{n+2}} \cdots$$

$$C^\infty \coloneqq \varprojlim(C_n, p_n) = \left\{(c_1, c_2, \dots) : c_n \in C_n \text{ and } p_n(c_{n+1}) = c_n \text{ for all } n\right\}$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory
- Categories
- Cartesian Closed Categories
- Functors
- Natural Transformations
- Equivalence of Categories
- Product and Functor Category
- Limits and Colimits
- Construct New from Old
- Topos
- ETCS
- Yoneda Lemma
- Adjunctions
- Categorical Logic
- Chu Space
- Kan Extension
- Monoidal Categories
- Enriched Categories
- Internal Category
- Profunctor
- Algebra & Coalgebra
- Monad
- Sheaves
- CCAF
- Rosen's (M, R)-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References2039

Definition (Para)

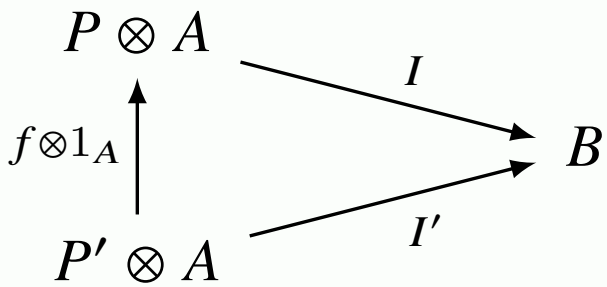
Given a symmetric monoidal category $\mathbf{C}$, let $\mathbf{Para}(\mathbf{C})$ be the bicategory whose:

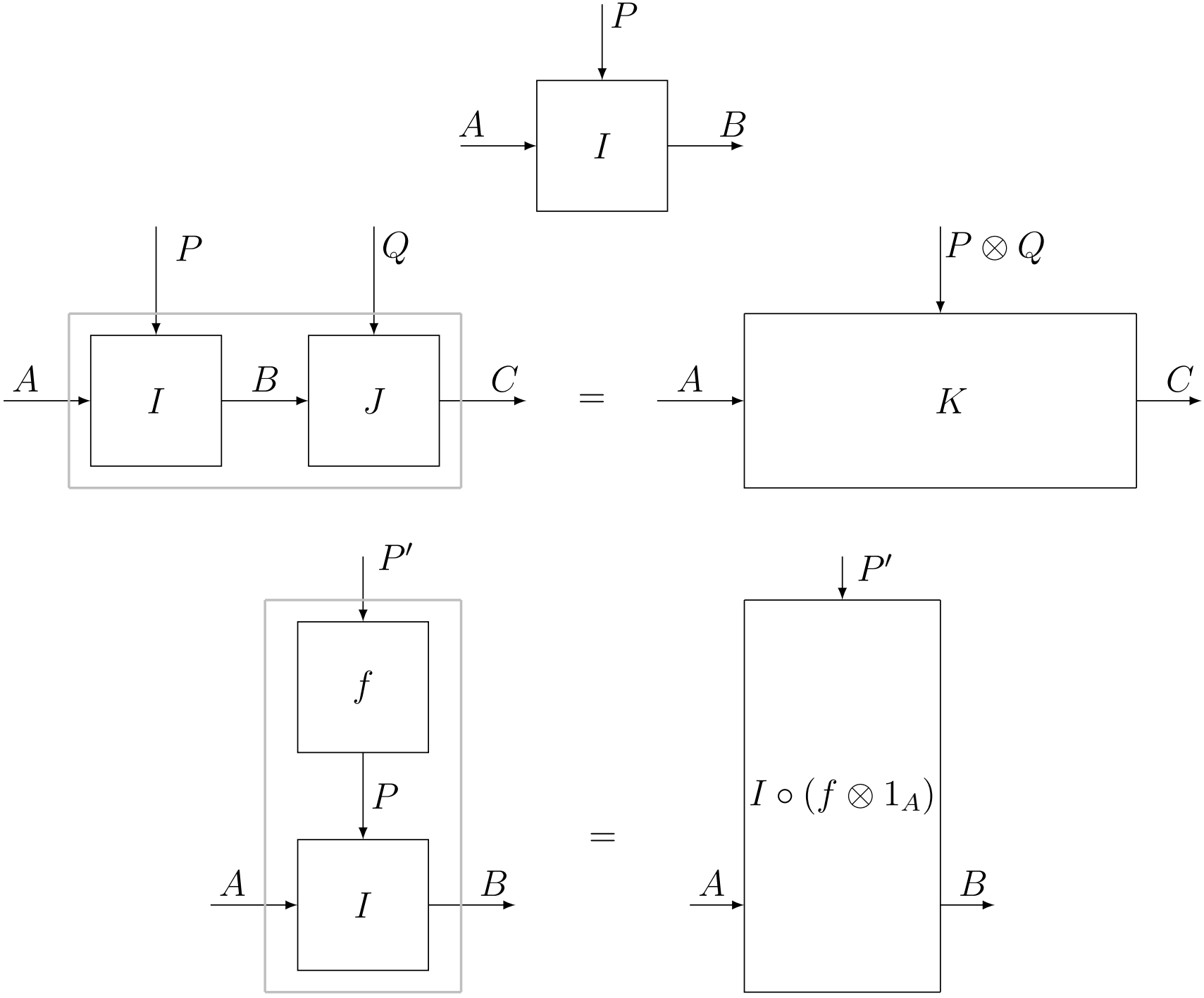
- Objects are the objects of $\mathbf{C}$;
- Morphisms $A \to B$ are pairs (P, I), where $I : P \otimes A \to B$ is a morphism of $\mathbf{C}$,

  and the composition $A \xrightarrow{(P,I)} B \xrightarrow{(Q,J)} C$ is $(Q \otimes P, K)$ where K is the composite

$$(Q \otimes P) \otimes A \xrightarrow{\alpha_{Q,P,A}} Q \otimes (P \otimes A) \xrightarrow{1_Q \otimes I} Q \otimes B \xrightarrow{J} C$$

- Identity on object A is $(I_{\mathbf{C}}, 1_A)$;
- 2-cell $(P, I) \to (P', I') : A \to B$ are morphisms $f : P' \to P$ s.t.




P
A
I
B
P
Q
A
I
B
J
C
=
P ⊗ Q
A
K
C
P′
f
P
A
I
B
=
P′
I ∘ (f ⊗ 1_A)
A
B

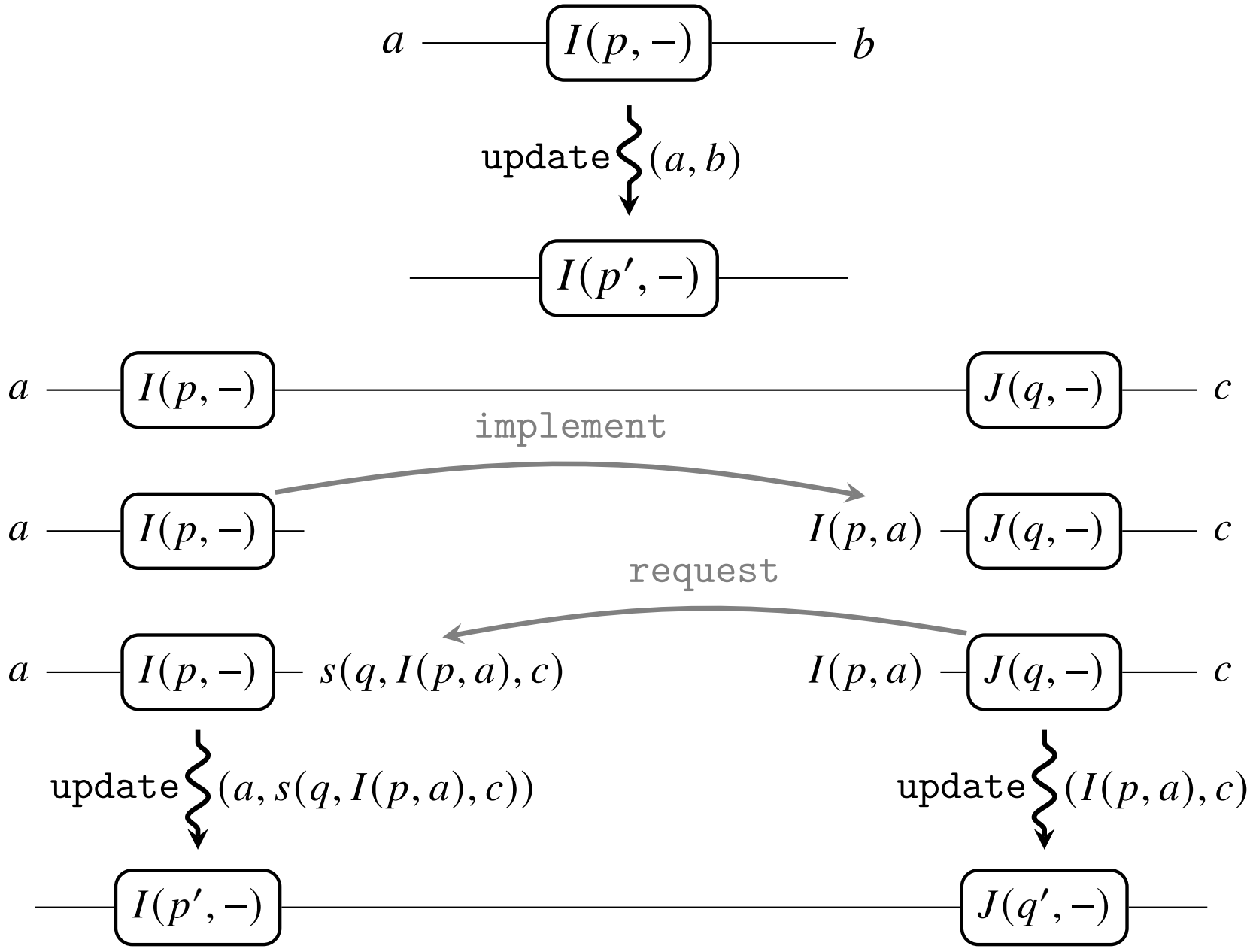
Learner

Definition (Learner)

A learner $A \to B$ between two sets A and B is a tuple (P, I, U, r):

1. P is a parameter space of some functions from A to B.
2. I is the implement map $I : P \times A \to B$ describing the functions in P.
3. U is the update map $U : P \times A \times B \to P$, where $U(p, a, b)$ sends a closer to b than the map p did.
4. r is the request map $r : P \times A \times B \to A$. The new element $r(p, a, b) = a'$ in A is such that $I(p, a')$ will be closer to b than $I(p, a)$ was.

$$I\colon P \times A \to B \qquad \text{implement}$$
$$U\colon P \times A \times B \to P \qquad \text{update}$$
$$r\colon P \times A \times B \to A \qquad \text{request}$$



A request function allows an update function to be defined for the composite $J(q, I(p, -))$.

The Category of Learners $\mathbf{Learn}(\mathbf{A}, \mathbf{B})$ between A and B

Given two learners $A \underset{(P',I',U',r')}{\overset{(P,I,U,r)}{\rightrightarrows}} B$, a map between (P, I, U, r) and (P', I', U', r') is a function $f : P \to P'$ s.t.

1. f is surjective.
2. f preserves implementations: $I'(f(p), a) = I(p, a)$
3. f preserves updates: $U'(f(p), a, b) = f\big(U(p, a, b)\big)$
4. f preserves requests: $r'(f(p), a, b) = r(p, a, b)$

$$\begin{array}{ccc} P \times A & \xrightarrow{I} & B \\ {\scriptstyle f\times 1_A}\downarrow & \nearrow_{I'} & \\ P' \times A & & \end{array} \qquad \begin{array}{ccc} P \times A \times B & \xrightarrow{(U,r)} & P \times A \\ {\scriptstyle f\times 1_A\times 1_B}\downarrow & & \downarrow{\scriptstyle f\times 1_A} \\ P' \times A \times B & \xrightarrow[(U',r')]{} & P' \times A \end{array}$$

- $\mathbf{Learn}(\mathbf{A}, \mathbf{B})$ is the category with objects the learners $A \to B$, and with morphisms the maps between learners.
- **Theorem:** $\mathbf{Learn}(\mathbf{A}, \mathbf{B})$ is a topos.
- The map between learners generates an equivalence relation on learners.

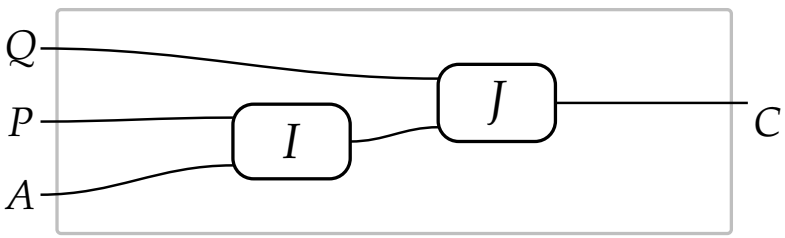
The Category of Learners **Learn**

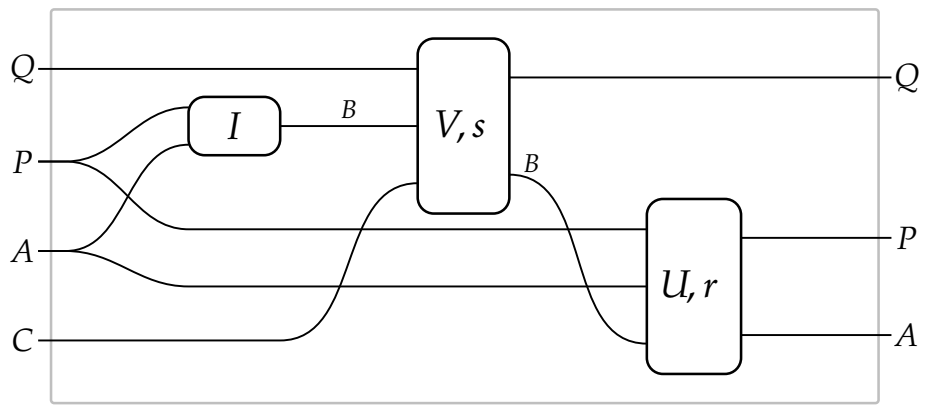
Learn is the symmetric monoidal category with objects the sets and morphisms the equivalence classes of learners. The composite learner $A \xrightarrow{(P,I,U,r)} B \xrightarrow{(Q,J,V,s)} C$ is $(P \times Q, I * J, U * V, r * s)$ where

$$(I * J)(p, q, a) := J(q, I(p, a))$$

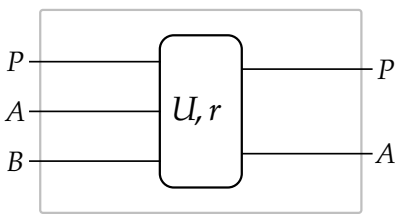
$$(U * V)(p, q, a, c) := \Big(U(p, a, s(q, I(p, a), c), V(q, I(p, a), c))\Big)$$

$$(r * s)(p, q, a, c) := r(p, a, s(q, I(p, a), c))$$

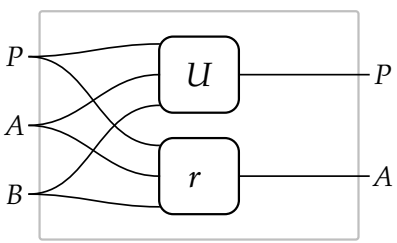




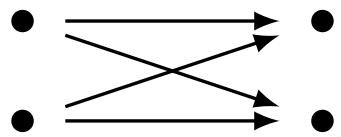
where



$:=$



Deep Neural Network



- There is a symmetric monoidal functor
$$L : \mathbf{Para} \to \mathbf{Learn}$$
which is the identity on objects and sends a parameterized function $I : P \times \mathbb{R}^m \to \mathbb{R}^n$ to the learner $\mathbb{R}^m \xrightarrow{(P,I,U,r)} \mathbb{R}^n$.
- For example, the implement map for $p = (w, b) \in P$ is
$$I(p, x) \coloneqq \sigma\left(w^{\mathsf{T}} x + b\right)$$
$$U(p, x, y) \coloneqq p - \eta \nabla_p \frac{1}{2}\left\|I(p, x) - y\right\|^2$$
$$r(p, x, y) \coloneqq x - \nabla_x \frac{1}{2}\left\|I(p, x) - y\right\|^2$$
- A deep neural network with k layers is the composition of k learners.
$$\mathbb{R}^{a_0} \longrightarrow \mathbb{R}^{a_1} \longrightarrow \mathbb{R}^{a_2} \longrightarrow \cdots \longrightarrow \mathbb{R}^{a_k}$$

Let $\mathbf{NNet}$ be a category where

- the objects are natural numbers and,
- a morphism $m \longrightarrow n$ is a list of natural numbers $(a_0, a_1, \ldots, a_k)$ with $a_0 = m$ and $a_k = n$.

composition is given by concatenation and this category is symmetric monoidal using the + operation in $\mathbb{N}$.

Remark: The idea is that a morphism $(a_0, a_1, \ldots, a_k)$ represents a neural network with $k - 1$ hidden layers and number of neurons given by the a_i.

Theorem

Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a differentiable function. Then there is a symmetric monoidal functor

$$F : \mathbf{NNet} \to \mathbf{Para}$$

which

- *sends a natural number n to vector space $\mathbb{R}^n$ and,*
- *sends a list $(a_0, a_1, \ldots, a_k)$ to the parameterized map*

$$\mathbb{R}^m \times \mathrm{GL}(a_1, a_2) \times \cdots \times \mathrm{GL}(a_{k-1}, a_{k-1}) \to \mathbb{R}^n$$

which is the alternating composite of the linear maps given by the parameters and the extension of σ to the corresponding vector spaces.

Remark: Now we have a machine for building neural networks compositionally.

$$\mathbf{NNet} \xrightarrow{F} \mathbf{Para} \xrightarrow{L} \mathbf{Learn}$$

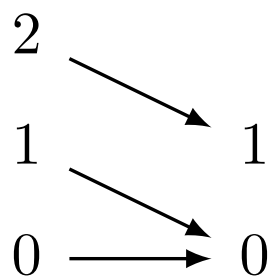
Simplex Category

The simplex category Δ consists of

- objects: linearly ordered sets of the form $[n]$ for $n \geq 0$, where $[n] := \{0 < 1 < 2 < \cdots < n-1 < n\}$.
- morphisms: $f : [m] \to [n]$ which is nondecreasing $i \leq j \implies f(i) \leq f(j)$.

All morphisms can be built out of primitive injections/surjections

- $\delta_i : [n] \to [n+1]$ injection skipping i
- $\sigma_i : [n] \to [n-1]$ surjection repeating i



What is Compositionality?

- Category Theory is the study of compositionality.
- Compositionality is the ability to compose objects, and the ability to work with an object *after intentionally forgetting how it was built.*

Compositionality

The meaning of a whole should only depend on the meanings of its parts and how they are fitted together.

1. Ability to build systems by composing them out of smaller subsystems
2. Ability to reason about the resulting system in terms of its components

Remark: Our models of systems that have 1 but not 2 are not compositional.

- Neural networks as differentiable functions are compositional
- Neural networks as generative/discriminative models are not compositional

Remark

- Functorial semantics and the principle of compositionality often go hand-in-hand.
- The former prompts us to model behavior using a functor between syntax and semantics categories.
- The latter encourages us to take things one at a time: *To model a huge system,* compositionality tells us, *it's enough to model smaller pieces of it and then stick those pieces together.*

The process of doing category theory

1. We see an interesting structure somewhere.
2. We express it categorically, that is, using only objects and morphisms in some relevant category.
3. We look for that structure in other categories.
4. We use it to prove things that will immediately be true in many different categories.
5. We investigate how it transports between categories via functors.
6. We explore how to generalize it, perhaps by relaxing some conditions.
7. We investigate what universal properties it has.
8. We ask whether it has generalizations to higher dimensions via categorification.

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Topological Space & Metric Space

Definition (Topological Space)

A *topological space* $(X, \mathcal{O}_X)$ is a set X with a family $\mathcal{O}_X \subset \mathrm{P}(X)$ of subsets of X which contains $\emptyset$ and X, and is closed under finite intersections and arbitrary unions.

Definition (Metric Space)

A *metric space* (X, d) is a set X with a metric $d : X \times X \to \mathbb{R}^{\geq 0}$ s.t.

1. $d(x, y) = 0 \leftrightarrow x = y$
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leq d(x, y) + d(y, z)$

Remark: 拓扑空间用包含关系表示远近, 度量空间用距离表示远近.

Manifold

Definition (Manifold)

An *m-manifold* is a Hausdorff space with a countable basis such that every point has an open neighborhood homeomorphic to an open neighborhood in Euclidean space $\mathbb{R}^m$.

Example

- A point is a 0-dimensional manifold.
- The empty set is an n-manifold for any n.
- There are essentially two types of 1-dimensional manifolds: an open line and something closed like a circle.
- It is important that the ends of the curved line be "open". $(0, 1)$ is a 1-manifold, while $[0, 1]$ is not.

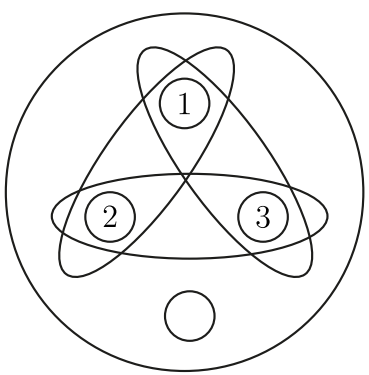
Some Topological Property

- A is dense iff every nonempty open set $U \in \mathcal{O}_X$ intersects A. Equivalently: $\mathrm{cl}(A) = X$.
- A is nowhere dense iff $(\mathrm{cl}(A))^\circ = \emptyset$. Equivalently: $\left(\overline{A}\right)^\circ$ is dense.
- A topological space $(X, \mathcal{O}_X)$ is connected iff the only clopen sets are $\emptyset$ and X. Equivalently:

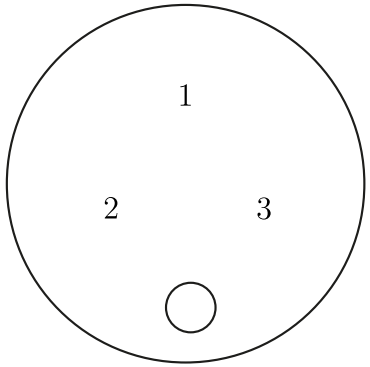
$$\forall A, B \subset X : \left(X = A \uplus B \implies A \cap \mathrm{cl}(B) \neq \emptyset \text{ or } B \cap \mathrm{cl}(A) \neq \emptyset \right)$$

- A topological space $(X, \mathcal{O}_X)$ is compact iff every open cover of X has a finite subcover.
- A topological space $(X, \mathcal{O}_X)$ is an Alexandroff space iff $\mathcal{O}_X$ is closed under arbitrary intersections, i.e., $\bigcap \mathcal{S} \in \mathcal{O}_X$ for any $\mathcal{S} \subset \mathcal{O}_X$.

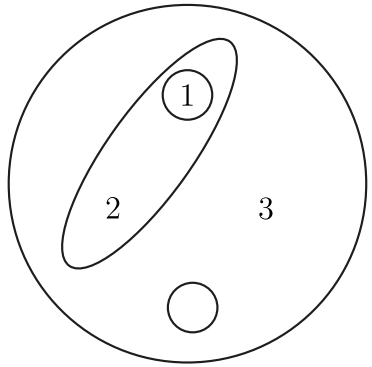
Some (non)examples of topological spaces



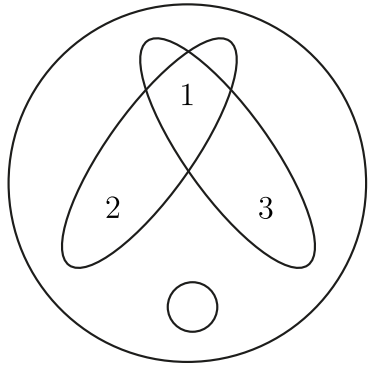
Discrete Topology



Trivial Topology



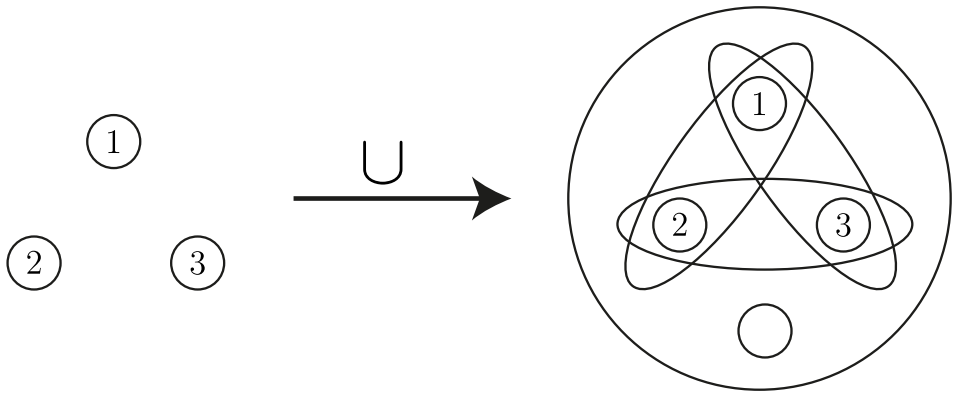
$\{\{1\}, \{1,2\}, \{1,2,3\}, \emptyset\}$



Not a topology

Figure: Some topologies over the finite set $\{1, 2, 3\}$.

- A collection $\mathcal{B}$ of open subsets of X is a basis for the topology of X iff every open subset is the union of some collection of elements of $\mathcal{B}$.



- A topological space is second countable if it admits a countable basis.

Some Topological Property

- A function $f : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ is Continuous iff $f^{-1}(U) \in \mathcal{O}_X$ for every $U \in \mathcal{O}_Y$.
  - Every function $f : (X, \mathrm{P}(X)) \to (Y, \mathcal{O}_Y)$ from the discrete topoloty to any topology is continuous.
  - Every function $f : (X, \mathcal{O}_X) \to (Y, \{\emptyset, Y\})$ from any topology to the indiscrete topoloty is continuous.
- Arbitrary intersections of topologies is a topology.
- The topologies on X form a complete lattice.
- Every family $\mathcal{T} \subset \mathrm{P}(X)$ has a unique smallest topology containing it.

$$\mathcal{O}_{\mathcal{T}} := \bigcap \{\mathcal{O} \supset \mathcal{T} : \mathcal{O} \text{ is a topology}\}$$

- The forgetful functor $U : \mathbf{Top} \to \mathbf{Set}$ and the Discrete and Indiscrete functors $\mathbf{Set} \to \mathbf{Top}$ form an adjunction.

$$\mathrm{Disc} \dashv U \dashv \mathrm{Indisc}$$

  It means U has to respect all limits and colimits.
- We can compute (co)limits in $\mathbf{Top}$ by first computing the (co)limit of underlying sets, then choosing the smallest/largest topologies making the limiting arrows continuous.

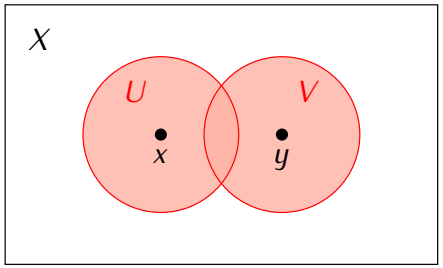
Measure Algebra

- A Measure Algebra $(X, \mathcal{A}_X)$ is a set X with a family $\mathcal{A}_X \subset \mathrm{P}(X)$ of subsets of X which contains $\emptyset$ and X, and is closed under finite intersections, countable unions, and complementation.
- A function $f : (X, \mathcal{A}_X) \to (Y, \mathcal{A}_Y)$ is measurable iff $f^{-1}(U) \in \mathcal{A}_X$ for every $U \in \mathcal{A}_Y$.
  - Every function $f : (X, \mathrm{P}(X)) \to (Y, \mathcal{O}_Y)$ from the discrete algebra to any measure algebra is measurable.
  - Every function $f : (X, \mathcal{O}_X) \to (Y, \{\emptyset, Y\})$ from any measure algebra to the indiscrete algebra is measurable.
- The measure algebras on X is a complete lattice.
- We also have the adjuntion:

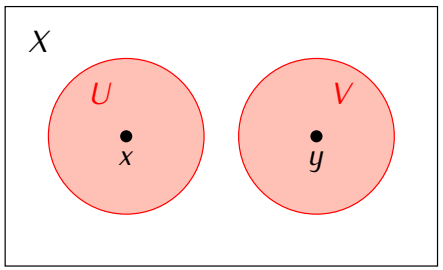
$$\mathrm{Disc} \dashv U \dashv \mathrm{Indisc}$$

Separation Axioms

- X is T_0, or Kolmogorov, iff any two distinct points in X are topologically distinguishable.
  — two points of X are topologically indistinguishable iff they have exactly the same neighborhoods.
  — In an indiscrete space, any two points are topologically indistinguishable.
- X is T_1, or accessible or Fréchet, iff for any two distinct points in X, each of them has a neighbourhood that is not a neighbourhood of the other.

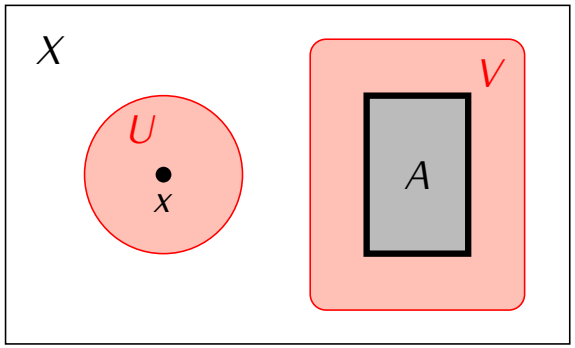


- X is T_2, or Hausdorff, iff any two distinct points in X are separated by neighbourhoods.

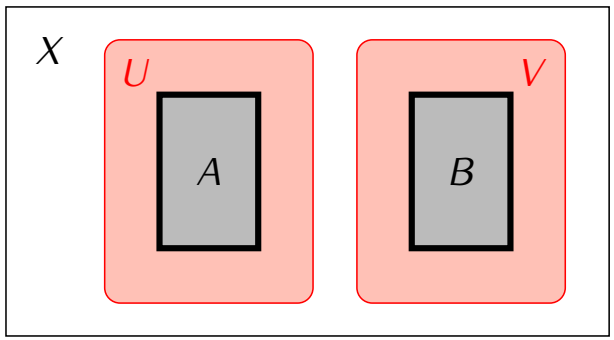


Separation Axioms

- X is regular, iff, given any point x and closed set A such that $x \notin A$, they are separated by neighbourhoods.
- X is T_3, or regular Hausdorff, iff it is both T_0 and regular.



- X is normal, iff, any two disjoint closed subsets of X are separated by neighbourhoods.
- X is T_4, or normal Hausdorff, iff it is both T_1 and normal.



Vector Space

Definition (Vector Space)

A vector space over a field $\mathbb{K}$ is a set V with an element $0 \in V$, addition $+ : V \times V \to V$, and scalar multiplication $\cdot : \mathbb{K} \times V \to V$ s.t. for all $a, b \in \mathbb{K}$ and $u, v, w \in V$:

1. $(u + v) + w = u + (v + w)$
2. $u + v = v + u$
3. $v + 0 = v$
4. there exists a $-v \in V$ s.t $v + (-v) = 0$
5. $(ab)v = a(bv)$
6. $1v = v$
7. $a(u + v) = au + av$
8. $(a + b)v = av + bv$

Normed Vector Space

Definition (Normed Vector Space)

A *normed vector space* $(V, \|\cdot\|)$ is a vector space over a field $\mathbb{K}$ with a *norm* $\|\cdot\| : V \to [0, \infty)$ s.t. for all $a \in \mathbb{K}$, and $u, v \in V$:

1. $\|av\| = |a|\|v\|$
2. $\|u + v\| \leq \|u\| + \|v\|$
3. $\|v\| = 0 \to v = 0$

Example: the ℓ^p norm of vector $\mathbf{x} = (x_1, \ldots, x_n)$:

$$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

- If $\|\cdot\|$ is a norm on V, then $d(u, v) = \|u - v\|$ defines a metric on V.
- A metric space is *complete* iff for every Cauchy sequence $\{x_n\}$, $\lim_{n\to\infty} \|x - x_n\| = 0$.
- A *Banach space* is a complete normed vector space.

Inner Product Space

Definition (Inner Product Space)

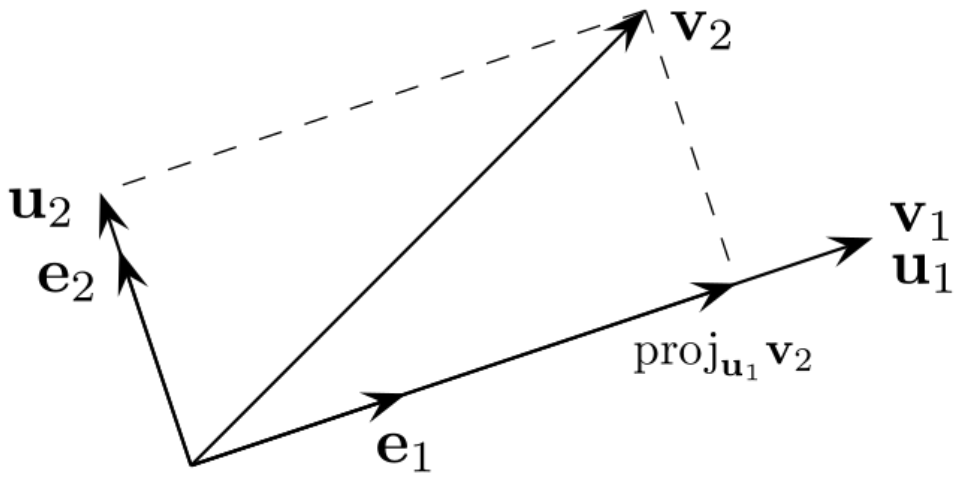
An *inner product space* $(V, \langle\cdot|\cdot\rangle)$ is a vector space over a field $\mathbb{K}$ with an *inner product* $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{K}$ s.t. for all $a, b \in \mathbb{K}$ and $u, v, w \in V$:

1. $\langle u|av + bw\rangle = a\langle u|v\rangle + b\langle u|w\rangle$
2. $\langle u|v\rangle = \overline{\langle v|u\rangle}$
3. $\langle v|v\rangle \geq 0$
4. $\langle v|v\rangle = 0 \to v = 0$

- The vectors u, v are called *orthogonal* iff $\langle u|v\rangle = 0$.
- A basis $\{e_1, \ldots, e_n\}$ is called *orthogonal* iff $\langle e_i|e_j\rangle = 0$ for all $i \neq j$. It is called *orthonormal* iff in addition $\langle e_i|e_i\rangle = 1$ for all i.
- A *Hilbert space* is a real or complex inner product space that is complete in the norm $\|v\| = \sqrt{\langle v|v\rangle}$.
- The trace of an operator A acting on a Hilbert space is

$$\mathrm{Tr}(A) = \sum_{e_i} \langle e_i|A|e_i\rangle \qquad \text{where } \{|e_i\rangle\} \text{ is any orthonormal basis.}$$

Gram-Schmidt Orthogonalization



$$u_1 = v_1 \qquad e_1 = \frac{u_1}{\|u_1\|}$$

$$u_2 = v_2 - \text{proj}_{u_1}(v_2) \qquad e_2 = \frac{u_2}{\|u_2\|}$$

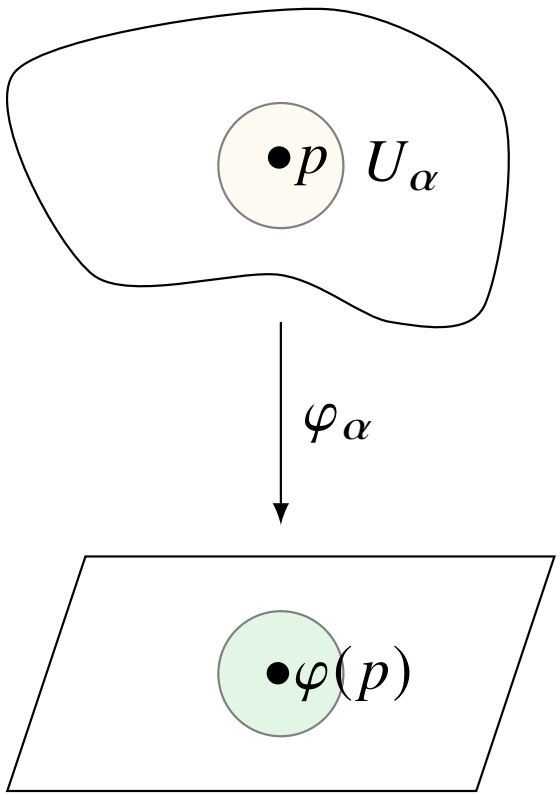
$$u_k = v_K - \sum_{i=1}^{k-1} \text{proj}_{u_i}(v_k) \qquad e_k = \frac{u_k}{\|u_k\|}$$

where $\text{proj}_u(v) = \frac{\langle v|u\rangle}{\langle u|u\rangle}u$.

拓扑流形

Definition (拓扑流形)

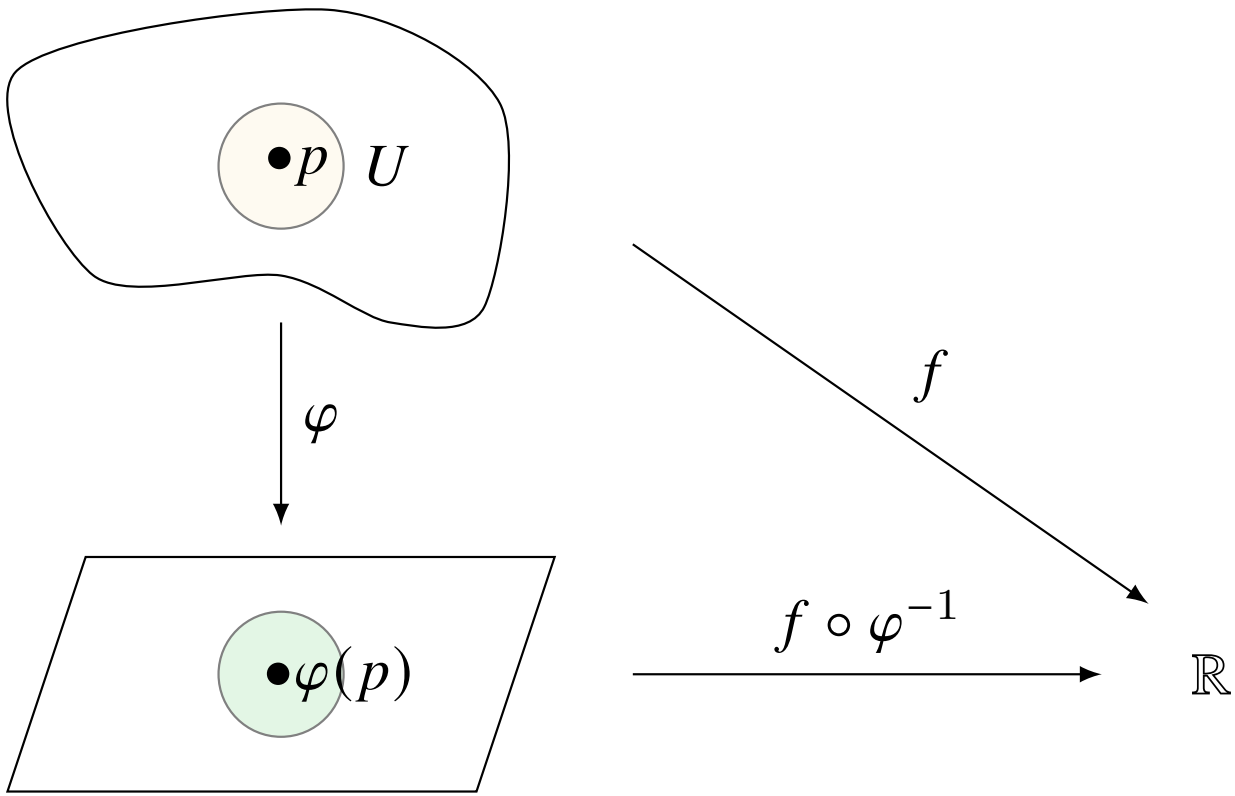
若 Hausdorff 空间 M 有可数开覆盖 $M = \bigcup_{\alpha} U_\alpha$ (U_α 为开集), 且每一个 U_α 都同胚于 $\mathbb{R}^n$ 的一个开子集, $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha) \subset \mathbb{R}^n$, 则称 M 是一个 n 维的拓扑流形, 称 $(U_\alpha, \varphi_\alpha)$ 是一个坐标卡 (chart).



光滑函数

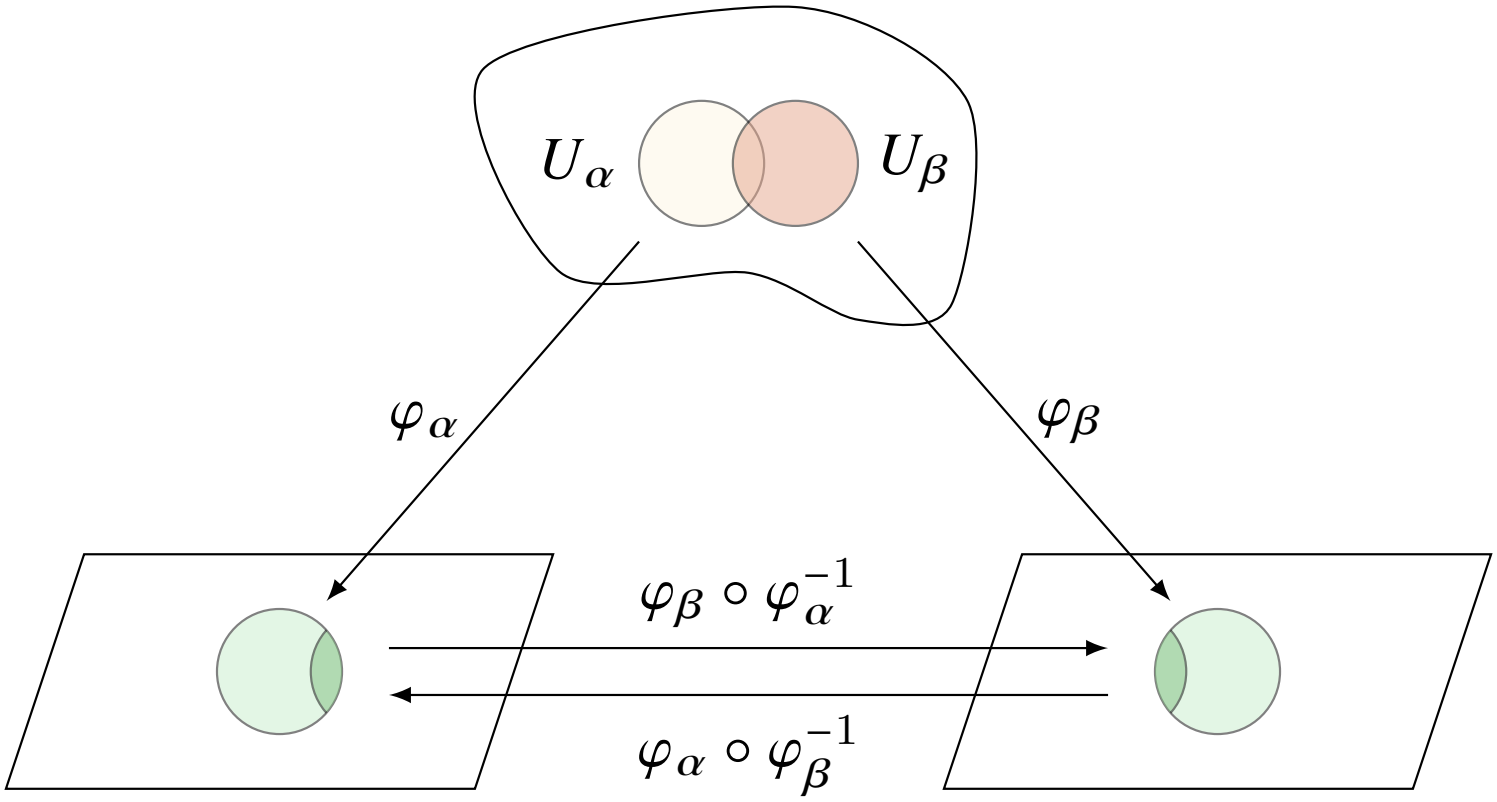
我们称函数 $f : M \to \mathbb{R}$ 在点 $p \in U$ 是光滑的 C^∞, 如果 $f \circ \varphi^{-1} : \varphi(U) \to \mathbb{R}$ 在 $\varphi(p)$ 是光滑的.

$$\mathbb{R}^n \supset \varphi(U) \xrightarrow{\varphi^{-1}} U \xrightarrow{f} \mathbb{R}$$



微分流形

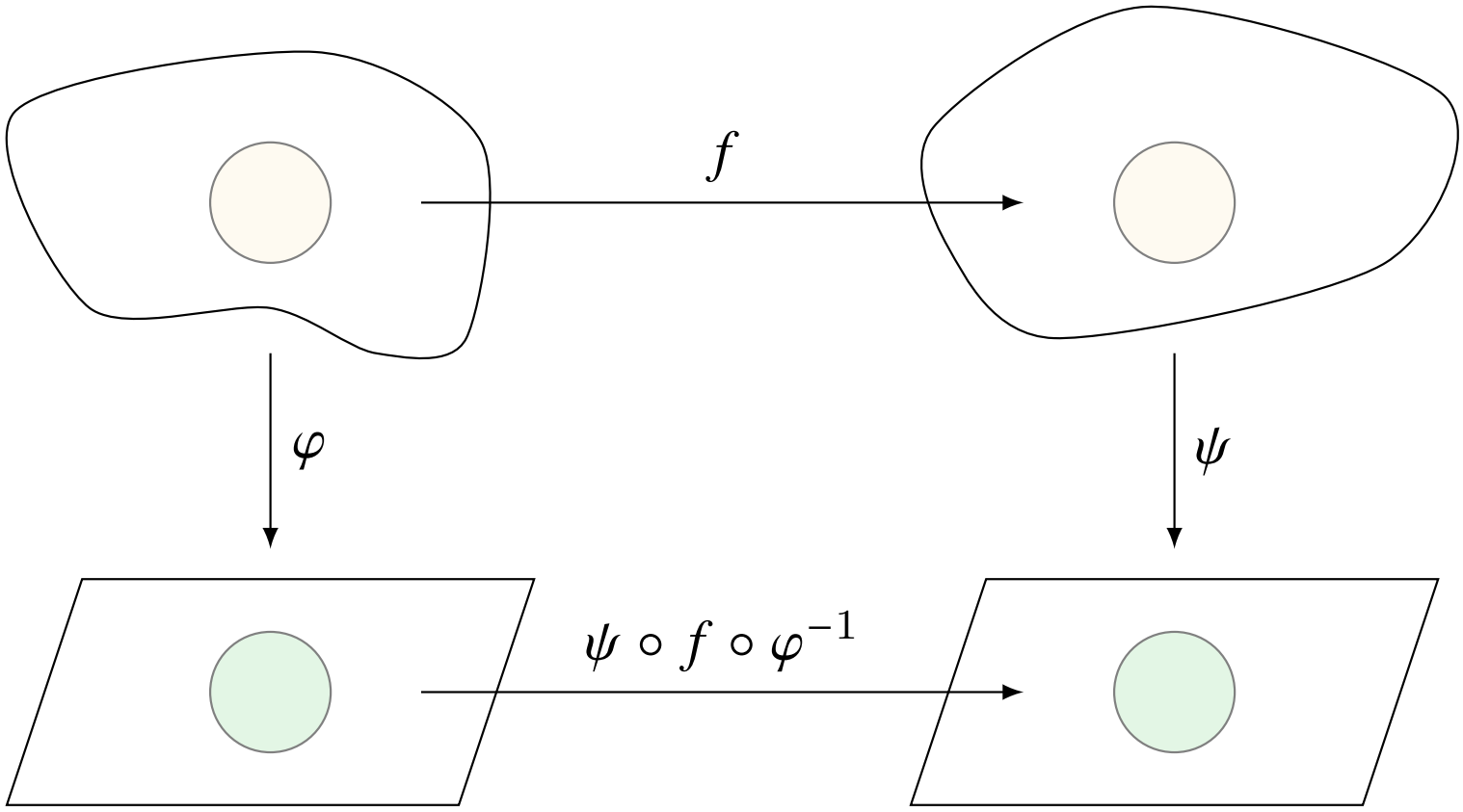
设 $(U_\alpha, \varphi_\alpha)$, (U_β, φ_β) 是 M 的两个坐标卡, 若 $U_\alpha \cap U_\beta \neq \emptyset$, 则 $U_\alpha \cap U_\beta$ 中的点有两套坐标, 同胚映射 $\varphi_\alpha \circ \varphi_\beta^{-1}$ 建立了两套坐标之间的变换.



此时, 称覆盖了 M 的坐标卡集 $(U_\alpha, \varphi_\alpha)_\alpha$ 为图册 (atlas).
若所有 $\varphi_\alpha \circ \varphi_\beta^{-1}$ 都是光滑函数, 则称 M 为微分流形.

微分同胚

若在 M 的每一点 $p \in M$ 都有 p 的坐标卡 (U, φ) 和 $f(p)$ 的坐标卡 (V, ψ) 使得 $\psi \circ f \circ \varphi^{-1} : \varphi(U) \to \psi(V)$ 光滑, 则称 $f : M \to N$ 光滑.
若 $f : M \to N$ 是光滑的双射, 则称 f 是 M 到 N 的微分同胚.



方向导数

- 对 $U \subset \mathbb{R}^n$, 光滑函数 $f : U \to \mathbb{R}$, 点 $p \in U$, 和向量 $\mathbf{v} \in \mathbb{R}^n$, 方向导数定义为

$$D_{\mathbf{v}} f|_p := \lim_{t \to 0} \frac{f(p + t\mathbf{v}) - f(p)}{t}$$

若 $\mathbf{v} = \mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, 则

$$D_{\mathbf{e}_i} f|p = \left.\frac{\partial f}{\partial x_i}\right|_p$$

当 $\mathbf{v} = (a_1, \ldots, a_n)$ 时,

$$D_{\mathbf{v}} f|_p = \sum_{k=1}^{n} a_k \left.\frac{\partial f}{\partial x_k}\right|_p$$

- Curve: 一条曲线是一个光滑映射 $\gamma : I \to M$.
- 若 $\gamma(0) = p \in U \subset M$, 光滑函数 $f : U \to \mathbb{R}$, 则 f 沿着 γ 的导数为

$$X(f) := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(\gamma(t))$$

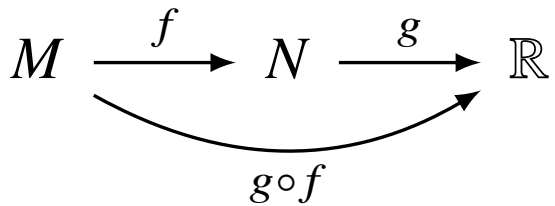
切空间

- Derivation: 点 $p \in U \subset M$ 处的导子 (Derivation) 是满足 Leibniz 律的线性映射 $X : C^\infty(U) \to \mathbb{R}$:

$$X(fg) = fX(g) + X(f)g$$

- 切空间 $T_pM :=$ {点 $p \in U \subset M$ 处的导子}
- $\left\{ \left.\frac{\partial}{\partial x_1}\right|_p, \ldots, \left.\frac{\partial}{\partial x_n}\right|_p \right\}$ 是 $T_p\mathbb{R}^n$ 的一组基.
- Derivative: 设 $f \in C^\infty(M, N)$, 且 $f(p) = q$. 则 f 在 p 点的导数 $Df|_p : T_pM \to T_qN$ 为

$$Df|_p(X)(g) := X(g \circ f)$$



- 链式法则 (Chain Rule): 对流形 M, N, P, 和 $f \in C^\infty(M, N)$, $g \in C^\infty(N, P)$, 点 $p \in M$, $q = f(p) \in N$, 易证,

$$D(g \circ f)|_p = Dg|_q \circ Df|_p$$

坐标系变换

- 若 f 是微分同胚, 则 $Df|_p$ 是线性同构, 且 $(Df|_p)^{-1} = D(f^{-1})|_{f(p)}$.
- 给定坐标卡 (U, φ), 其中 $\varphi = (x_1, \ldots, x_n)$, 定义

$$\left.\frac{\partial}{\partial x_i}\right|_p := (D\varphi|_p)^{-1} \left(\left.\frac{\partial}{\partial x_i}\right|_{\varphi(p)} \right) \in T_pM$$

- $\left\{ \left.\frac{\partial}{\partial x_1}\right|_p, \ldots, \left.\frac{\partial}{\partial x_n}\right|_p \right\}$ 是 T_pM 的一组基.
- 假设 p 点附近还有其它坐标卡的坐标 $(y_1, \ldots, y_n)$. 那么有 a_k 使得

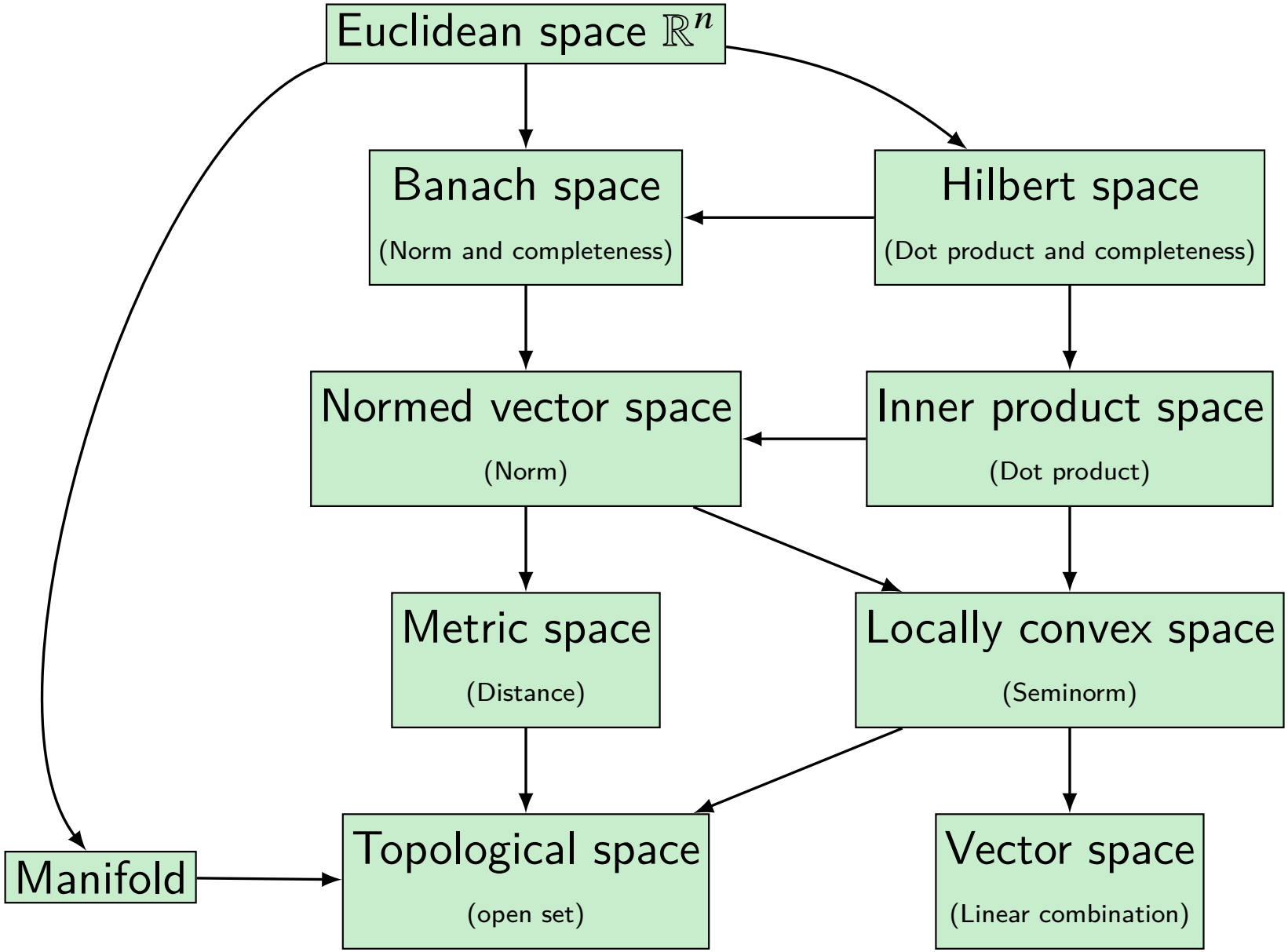
$$\left.\frac{\partial}{\partial y_i}\right|_p = \sum_{k=1}^{n} a_k \left.\frac{\partial}{\partial x_k}\right|_p$$

把 x_k 看作 $(y_1, \ldots, y_n)$ 的函数, 则

$$\left.\frac{\partial}{\partial y_i}\right|_p (x_k) = \frac{\partial x_k}{\partial y_i}(p) = a_k$$

$$\left.\frac{\partial}{\partial y_i}\right|_p = \sum_{k=1}^{n} \frac{\partial x_k}{\partial y_i}(p) \left.\frac{\partial}{\partial x_k}\right|_p$$

Spaces


Euclidean space $\mathbb{R}^n$
Banach space
(Norm and completeness)
Hilbert space
(Dot product and completeness)
Normed vector space
(Norm)
Inner product space
(Dot product)
Metric space
(Distance)
Locally convex space
(Seminorm)
Manifold
Topological space
(open set)
Vector space
(Linear combination)

Linear Operator

Definition

A linear operator A is

- *normal* iff $AA^\dagger = A^\dagger A$.
- *unitary* iff $AA^\dagger = A^\dagger A = I$.
- *self-adjoint* iff $A = A^\dagger$.
- a *projection* iff $A = A^\dagger$ and $AA = A$.
- *bounded* iff $\exists a \geq 0 \forall v : \|Av\| \leq a\|v\|$.

The adjoint of a linear operator A is the function $A^\dagger$ s.t. for every u, v:

$$\langle Au|v\rangle = \langle u|A^\dagger v\rangle$$

In terms of matrices, $A^\dagger = \overline{A^\mathsf{T}}$.

- General linear group $\mathrm{GL}(n, \mathbb{R})$ of order n is the group of $n \times n$ invertible matrices with entries in $\mathbb{R}$.

$$\mathrm{GL}(n, \mathbb{R}) \coloneqq \left\{ M \in \mathbf{Mat}_{n \times n}(\mathbb{R}) : \det(M) \neq 0 \right\}$$

- Orthogonal group

$$\mathrm{O}(n, \mathbb{R}) \coloneqq \left\{ M \in \mathbf{Mat}_{n \times n}(\mathbb{R}) : MM^{\mathsf{T}} = M^{\mathsf{T}}M = I \right\}$$

- Special linear group

$$\mathrm{SL}(n, \mathbb{R}) \coloneqq \left\{ M \in \mathbf{Mat}_{n \times n}(\mathbb{R}) : \det(M) = 1 \right\}$$

- Special orthogonal group

$$\mathrm{SO}(n, \mathbb{R}) \coloneqq \left\{ M \in \mathbf{Mat}_{n \times n}(\mathbb{R}) : MM^{\mathsf{T}} = M^{\mathsf{T}}M = I \text{ and } \det(M) = 1 \right\}$$

- Euclidean group

$$\mathrm{E}(n, \mathbb{R}) \coloneqq \left\{ \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \in \mathbf{Mat}_{(n+1) \times (n+1)}(\mathbb{R}) : R \in \mathrm{O}(n, \mathbb{R}) \text{ and } t \in \mathbb{R}^n \right\}$$

- Special Euclidean group $\mathrm{SE}(n, \mathbb{R})$: replace $R \in \mathrm{O}(n, \mathbb{R})$ with $R \in \mathrm{SO}(n, \mathbb{R})$ in the above definition.

$\mathrm{GL}(n, \mathbb{R})$	General linear group	invertible linear transformations
$\mathrm{SL}(n, \mathbb{R})$	Special linear group	preserve volume form
$\mathrm{O}(n, \mathbb{R})$	Orthogonal group	preserve length of vectors
$\mathrm{SO}(n, \mathbb{R})$	Special orthogonal group	rotations
$\mathrm{E}(n, \mathbb{R})$	Euclidean group	preserve distances and angles
$\mathrm{SE}(n, \mathbb{R})$	Special Euclidean group	rigid motions

Remark: The groups $\mathrm{SE}(2, \mathbb{R})$ and $\mathrm{SE}(3, \mathbb{R})$ are particular important in robotics because they represent the roto-translations of the plane and 3D space, respectively.

Example: The group $\mathrm{SE}(n, \mathbb{R})$ induces a transformation on the points of $\mathbb{R}^n$. We call this an action.

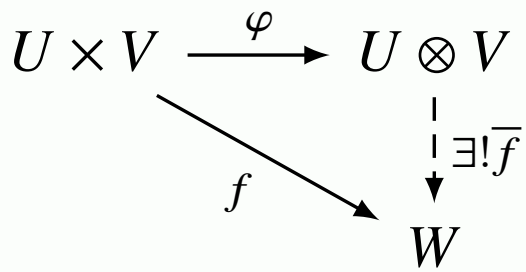
$$\mathrm{apply} : \mathrm{SE}(n, \mathbb{R}) \times \mathbb{R}^n \to \mathbb{R}^n :: ((R, t), p) \mapsto Rp + t$$

Given a roto-translation and a point, it returns the roto-translated point.

$$\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p \\ 1 \end{bmatrix} = \begin{bmatrix} Rp + t \\ 1 \end{bmatrix}$$

Definition (Tensor Product)

Suppose U and V are vector spaces over a field $\mathbb{K}$. Then a tensor product of U and V is a vector space $U \otimes V$ over $\mathbb{K}$ with a bilinear map $\varphi : U \times V \to U \otimes V :: (u, v) \mapsto u \otimes v$ having the "universal property":



Given two linear maps $A : U \to X$ and $B : V \to Y$ between vector spaces, the tensor product of the two linear maps A and B is a linear map $A \otimes B : U \otimes V \to X \otimes Y$ defined by

$$(A \otimes B)(u \otimes v) = Au \otimes Bv$$

$$W = X \otimes Y \qquad f : (u, v) \mapsto Au \otimes Bv \qquad \overline{f} = A \otimes B$$

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{m1}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

Tensor = Multilinear Map

$$U \otimes V \cong U^{**} \otimes V^{**} \cong (U^* \otimes V^*)^* \cong \mathrm{Hom}^2(U^* \times V^*, \mathbb{F})$$

where this last space is the space of bilinear maps $U^* \times V^* \longrightarrow \mathbb{F}$.
The isomorphism maps an element $u \otimes v$ of the LHS to the bilinear map

$$(\alpha, \beta) \mapsto \alpha(u) \cdot \beta(v)$$

If we fix V, the (k, l)-tensor space is $T \coloneqq V^{\otimes k} \otimes (V^*)^{\otimes l}$.

$$V^{\otimes k} \otimes (V^*)^{\otimes l} \cong \mathrm{Hom}^{k+l}\left((V^*)^k \times V^l\,, \mathbb{F}\right)$$

Then $v_1 \otimes \cdots \otimes v_k \otimes \alpha^1 \otimes \cdots \otimes \alpha^l$ is identified with the multilinear map

$$(\beta^1, \ldots, \beta^k, w_1, \ldots, w_l) \mapsto \beta^1(v_1) \cdots \beta^k(v_k) \cdot \alpha^1(w_1) \cdots \alpha^l(w_l)$$

Take $\{\partial_\mu\}$ as the basis on V, and $\{\mathrm{d}x^\mu\}$ as the basis on V^*, we expand T as

$$T = T^{\mu_1 \ldots \mu_k}_{\nu_1 \ldots \nu_l} \partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_k} \otimes \mathrm{d}x^{\nu_1} \otimes \cdots \otimes \mathrm{d}x^{\nu_l}$$

The coordinate transformation rule is

$$\bar{T}^{\mu'_1 \ldots \mu'_k}_{\nu'_1 \ldots \nu'_l} = \frac{\partial \bar{x}^{\mu'_1}}{\partial x^{\mu_1}} \cdots \frac{\partial \bar{x}^{\mu'_k}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial \bar{x}^{\nu'_1}} \cdots \frac{\partial x^{\nu_l}}{\partial \bar{x}^{\nu'_l}} T^{\mu_1 \ldots \mu_k}_{\nu_1 \ldots \nu_l}$$

The Postulates of Quantum Mechanics[NC10]

I. A pure state of a system in quantum mechanics is represented in terms of a normalized vector $|\psi\rangle$ in a separable complex Hilbert space.

II. The time evolution of the state of a closed quantum system from t_0 to t_1 is described by a unitary transformation: $|\psi_{t_1}\rangle = U|\psi_{t_0}\rangle$.

III. A quantum measurement is described by an observable, A, a self-adjoint linear operator acting on the Hilbert space. The possible outcomes of the measurement correspond to the eigenvalues a of the observable. The observable has a spectral decomposition $A = \sum_a aP_a$, where $P_a = \sum_i |e_i\rangle\langle e_i| \, \delta_{a_i a}$ is the projector onto the subspace spanned by all the eigenvectors that produce the same eigenvalue a. If the system is in a pure state $|\psi\rangle$ immediately before the measurement then the probability of obtaining an eigenvalue a of an observable A is $p(a) = \langle\psi|P_a|\psi\rangle$, and the state of the system after the measurement is $\frac{P_a|\psi\rangle}{\sqrt{\langle\psi|P_a|\psi\rangle}}$.

IV. The Hilbert space of a composite system is the tensor product of the state spaces of the component systems.

量子不可克隆定理 The No-Cloning Theorem

Theorem (The No-Cloning Theorem)

If there is a unitary operator U and two quantum states $|\phi\rangle$ and $|\psi\rangle$, and U takes $\phi\rangle \otimes |0\rangle$ to $\phi\rangle \otimes |\phi\rangle$ and $\psi\rangle \otimes |0\rangle$ to $\psi\rangle \otimes |\psi\rangle$, then either $\phi = \psi$ or $\phi \perp \psi$.

Proof.

$$\begin{aligned}
\langle\phi|\psi\rangle &= \langle\phi|\psi\rangle\langle 0|0\rangle \\
&= \big(\langle\phi|\langle 0|\big)\big(|\psi\rangle|0\rangle\big) \\
&= \big(\langle\phi|\langle 0|\big)U^\dagger U\big(|\psi\rangle|0\rangle\big) \\
&= \big(\langle\phi|\langle\phi|\big)\big(|\psi\rangle\psi\rangle\big) \\
&= \langle\phi|\psi\rangle\langle\phi|\psi\rangle \\
&= \langle\phi|\psi\rangle^2
\end{aligned}$$

□

Remark: 非正交的量子态不能被复制. 经典信息可以被复制, 因为, 经典信息的不同状态可以被认为是正交的量子态.

量子不可删除定理 The No-Deleting Theorem

It is impossible to delete one of two copies of two different and nonorthogonal quantum states by the same unitary operation.

Theorem (The No-Deleting Theorem)

If there is a unitary operator U and two quantum states $|\phi\rangle$ and $|\psi\rangle$, and U takes $\phi\rangle \otimes |\phi\rangle$ to $\phi\rangle \otimes |0\rangle$ and $\psi\rangle \otimes |\psi\rangle$ to $\psi\rangle \otimes |0\rangle$, then either $\phi = \psi$ or $\phi \perp \psi$.

Remark: The no-cloning and the no-deleting theorems point to the conservation of quantum information.

Physical Concept	Mathematical Representation	
	Classical mechanics	Quantum mechanics
state	point	vector
state space	set of points (phase space)	Hilbert space
property	function on points	operator on vectors

Tensor Product / Inner Product / Outer Product

- the standard basis states: $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
- the vector representation of a 1-qubit: $|\psi\rangle = a|0\rangle + b|1\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$

  where $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$.
- n-qubit: $|\psi\rangle = \sum_{x \in \{0,1\}^n} a_x |x\rangle$ with $\sum_{x \in \{0,1\}^n} |a_x|^2 = 1$.
- for $|\phi\rangle = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}, |\psi\rangle = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$, the tensor product of two qubits is

$$|\phi\rangle \otimes \psi\rangle = |\phi\rangle|\psi\rangle = |\phi\psi\rangle = \begin{bmatrix} a_0 b_0 \\ a_0 b_1 \\ a_1 b_0 \\ a_1 b_1 \end{bmatrix}$$

- the conjugate transpose of $|\phi\rangle$ is $\langle\phi| = |\phi\rangle^\dagger = \begin{bmatrix} \overline{a_0} & \overline{a_1} \end{bmatrix}$
- the inner product: $\langle\phi||\psi\rangle = \langle\phi|\psi\rangle = \sum_i \overline{a_i} b_i = \overline{a_0} b_0 + \overline{a_1} b_1$
- the outer product: $|\phi\rangle\langle\psi| = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \begin{bmatrix} \overline{b_0} & \overline{b_1} \end{bmatrix} = \begin{bmatrix} a_0\overline{b_0} & a_0\overline{b_1} \\ a_1\overline{b_0} & a_1\overline{b_1} \end{bmatrix}$

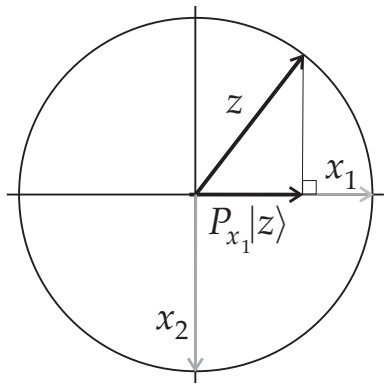
Projector & Probability Amplitude & Measurement

The projector $P_{x_i} : H \rightarrow L_{x_i}$ projects any point $|z\rangle$ in Hilbert space H into the subspace L_{x_i}. It is constructed from the outer product $|x_i\rangle\langle x_i|$, i.e.,

$$P_{x_i}|z\rangle = (|x_i\rangle\langle x_i|)|z\rangle = |x_i\rangle\langle x_i|z\rangle = \langle x_i|z\rangle|x_i\rangle$$

The inner product $\langle x_i|z\rangle$ can be interpreted as the probability amplitude of transiting to state $|x_i\rangle$ from state $|z\rangle$. The probability of transiting to state $|x_i\rangle$ from state $|z\rangle$ is

$$p(x_i) = \|P_{x_i}|z\rangle\|^2 = \langle z|P_{x_i}|z\rangle = |\langle x_i|z\rangle|^2$$



The state vector $|z\rangle$ can be expressed in terms of the basis states as $|z\rangle = \sum_i \langle x_i|z\rangle|x_i\rangle$.

The measurement of an n-qubit quantum state:

$|\psi\rangle = \sum_{x\in\{0,1\}^n} a_x|x\rangle \xrightarrow{\text{measurement}} |x\rangle$ with probability $|a_x|^2 = \overline{a_x}a_x$.

Classic Probability vs Quantum Probability

Classic	**Quantum**
Each unique outcome is a member of a set of points called the Sample space	Each unique outcome is an orthonormal vector from a set that spans a Vector space
Each event is a subset of the sample space	Each event is a subspace of the vector space
State is a probability function P defined on subsets of the sample space	State is a unit length vector $\|z\rangle$: $p(A) = \|p_A\|z\rangle\|^2$
$P(A \wedge B) = P(B \wedge A)$	$\|P_B P_A\|z\rangle\|^2 \neq \|P_A P_B\|z\rangle\|^2$
$P(B \mid A) = \frac{P(A \wedge B)}{P(A)}$	$P(B \mid A) = \frac{\|P_B P_A\|z\rangle\|^2}{\|P_A\|z\rangle\|^2}$
$P(A) = \sum_i P(A \mid B_i)P(B_i)$	Law of total probability violated

Quantum Gates & Universality

- What is a quantum gate? A unitary operator on a number of qubits.
  - 经典世界只有一个非平凡的单比特门 — 非门.
  - 任何酉算子都定义一个量子门. 所以有很多非平凡的单量子比特门.
- A set of gates is universal iff for any unitary matrix U and any $\varepsilon > 0$, there is some circuit $\widetilde{U}$ built out of the set of gates such that

$$\left\| U - \widetilde{U} \right\| < \varepsilon$$

In other words,

$$\sup_{|\psi\rangle} \left\| \left(U - \widetilde{U} \right) |\psi\rangle \right\| < \varepsilon$$

NOT Gate & $\sqrt{\text{NOT}}$ Gate

- NOT Gate

$$\text{NOT} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\text{NOT}\,|0\rangle = |1\rangle$$
$$\text{NOT}\,|1\rangle = |0\rangle$$
$$\text{NOT}\,\frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

- $\sqrt{\text{NOT}}$ Gate

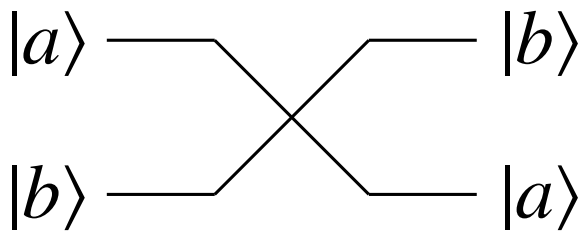
$$\sqrt{\text{NOT}} = \frac{1}{2}\begin{bmatrix} 1+\mathrm{i} & 1-\mathrm{i} \\ 1-\mathrm{i} & 1+\mathrm{i} \end{bmatrix}$$

$$\sqrt{\text{NOT}} \cdot \sqrt{\text{NOT}} = \text{NOT}$$
$$\sqrt{\text{NOT}} \cdot \sqrt{\text{NOT}}^{\dagger} = I$$

SWAP Gate & Hadamard Gate

- SWAP Gate

$$\text{SWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



- Hadamard Gate

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \qquad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$H = |+\rangle\langle 0| + |-\rangle\langle 1| = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$|\psi\rangle = a|0\rangle + b|1\rangle = \frac{a+b}{\sqrt{2}}|+\rangle + \frac{a-b}{\sqrt{2}}|-\rangle$$

$$H(|\psi\rangle) = a|+\rangle + b|-\rangle = \frac{a+b}{\sqrt{2}}|0\rangle + \frac{a-b}{\sqrt{2}}|1\rangle$$

Pauli Gates

Pauli Gates

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Phase Shift Gate

$$R_\phi = \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\phi} \end{bmatrix}$$

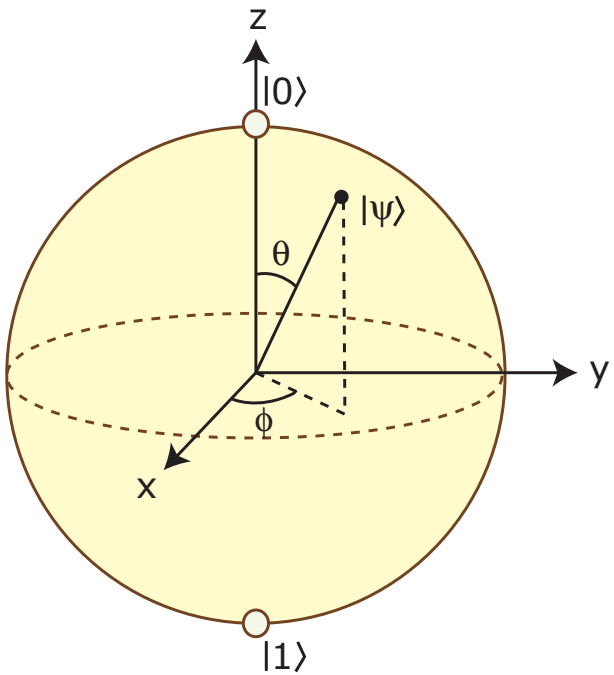
- $H = \frac{X+Z}{\sqrt{2}}$
- $H^2 = HH^\dagger = X^2 = Y^2 = Z^2 = -\mathrm{i}XYZ = I$
- For $x \in \{0,1\}^n$,

$$H^{\otimes n}|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y \in \{0,1\}^n} (-1)^{x \cdot y} |y\rangle$$

where $x \cdot y := \sum_{i=1}^{n} x_i y_i \bmod 2$.

1-Qubit Quantum State & Bloch Sphere

$$|\psi\rangle = a|0\rangle + b|1\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle = \cos\frac{\theta}{2}|0\rangle + (\cos\phi + i\sin\phi)\sin\frac{\theta}{2}|1\rangle$$



Points on the surface can also be expressed in Cartesian coordinates as

$$(x, y, z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

Starting from $|0\rangle$, any state can be reached by first rotating about y by angle θ and then about z by angle ϕ.

2×2 unitary matrix = rotation on Bloch sphere

The density matrix of $|\psi\rangle$ is

$$\rho = |\psi\rangle\langle\psi| = \begin{bmatrix} \cos^2\frac{\theta}{2} & e^{-i\phi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} & \sin^2\frac{\theta}{2} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1+\cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & 1-\cos\theta \end{bmatrix}$$

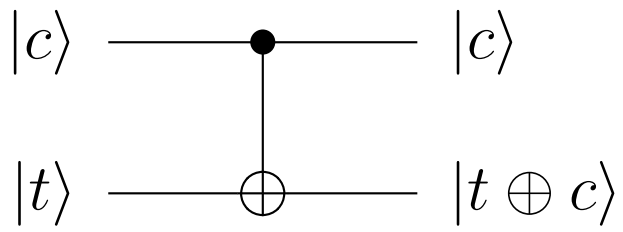
$$= \frac{1}{2}(I + xX + yY + zZ)$$

CNOT Gate

Definition

A quantum state $\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$ is a product state iff it can be expressed as a tensor product $\psi_1\rangle \otimes \cdots \otimes \psi_n\rangle$ of n 1-qubit states. Otherwise, it is entangled.

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$



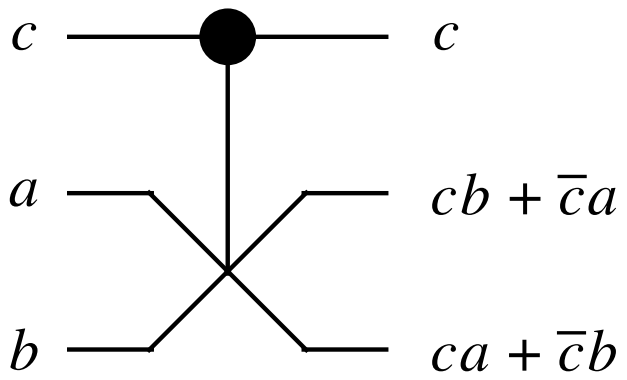
In general, one can define controlled versions of any unitary gate U as

$$CU = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$$

$\text{CNOT}(|+\rangle \otimes |0\rangle) = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}^{\mathsf{T}} = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ is an entangled state which can't be separated as tensor product.

Fredkin Gate: CSWAP

c	a	b	c	x	y
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	1	0
1	1	0	1	0	1
1	1	1	1	1	1



transmit the first bit unchanged and
swap the last two bits iff the first bit is 1.

$$f : (c, a, b) \mapsto (c, cb + \overline{c}a, ca + \overline{c}b)$$

$$\text{CSWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\neg \; a = 0 \;\; \& \;\; b = 1 \implies y = \overline{c}$$

$$\wedge \; b = 0 \implies y = ca$$

Toffoli Gate: CCNOT or $D(\frac{\pi}{2})$

x	y	t	x	y	z
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	1
1	1	1	1	1	0

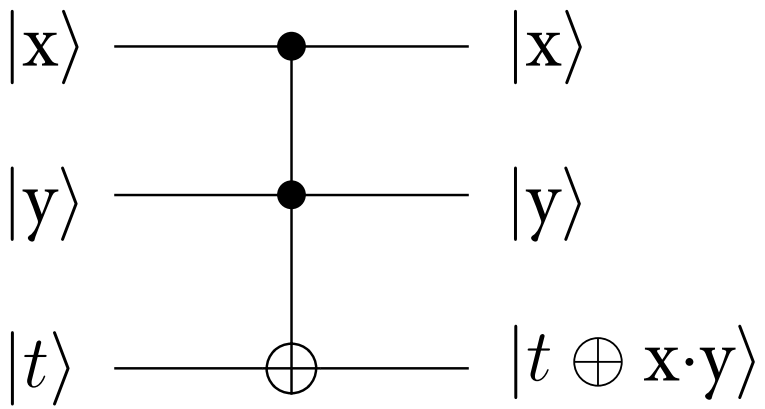


Figure: if the first two bits are 1, it inverts the third bit, otherwise all bits stay the same.

$f : (x, y, t) \mapsto \big(x, y, t \oplus xy\big)$

$$\text{CCNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$\neg \; x = y = 1 \implies z = \bar{t}$

$\wedge \; t = 0 \implies z = xy$

Universal Quantum Gates

- Fredkin gate CSWAP is universal for classical computation, but not universal for quantum computation.
- Toffoli gate CCNOT is universal for classical computation, but not universal for quantum computation.
- $\{\mathrm{CNOT}, H, R_{\frac{\pi}{4}}\}$ is universal for quantum computation.
- Deutsch gate $D(\theta)$ is universal for quantum computation.
- Toffoli gate and Hadamard gate $\{\mathrm{CCNOT}, H\}$ constitute a universal set of quantum gates.

Deutsch Gate

$$D(\theta) : |x, y, z\rangle \mapsto \begin{cases} \mathrm{i} \cos\theta \, |x, y, z\rangle + \sin\theta \, |x, y, 1 - z\rangle & \text{for } x = y = 1 \\ |x, y, z\rangle & \text{otherwise} \end{cases}$$

- Transformations on qubits are reversible.
- Qubit transformations are operators on vector spaces. And an operator defined on an n-dim vector space (e.g. n-qubit space) that acts on n-dim vectors (e.g. n-qubits) can only spit out n-dim vectors.

Quantum Algorithms

I. **Initialization** Build an initial state $|\psi_i\rangle = \sum_{x\in\{0,1\}^n} a_x|x\rangle$.

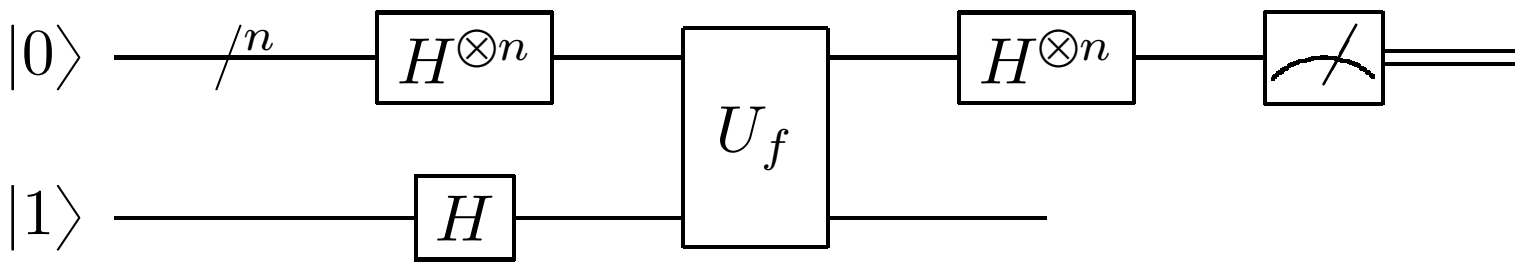
   For example, $\frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle$ (uniform superposition) can be build from $|0\rangle^{\otimes n}$ by application of Hadamard gate $H^{\otimes n}$.

II. **Transformations** Transform $|\psi_i\rangle \longrightarrow |\psi_f\rangle = \sum_{x\in\{0,1\}^n} b_x|x\rangle$ through a sequence of elementary quantum gates.

III. **Measurement** Extract information by quantum measurement of $|\psi_f\rangle$.

The "balanced vs constant" Problem

- given a function $f : \{0,1\}^n \rightarrow \{0,1\}$ that is either constant (same output for all x) or balanced ($f(x)$ is equal to 0 for exactly half of the possible values of x).
- classically, we need to query the function $2^{n-1} + 1$ times to be sure whether the function is constant or balanced.
- but quantumly, the Deutsch-Jozsa Algorithm requires only 1 oracle call.—measuring the first n-qubits allows us to determine with certainty whether the function is constant (measure all zeros) or balanced (measure at least one 1).



Deutsch-Jozsa Algorithm

1. prepare the initial state $|\psi_0\rangle = |0\rangle^{\otimes n}|1\rangle$
2. apply $H^{\otimes n} \otimes H$

$$|\psi_1\rangle = \left(H^{\otimes n} \otimes H\right)|\psi_0\rangle = \left(\frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle\right) \otimes |-\rangle$$

3. apply f as a quantum oracle $U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$

$$|\psi_2\rangle = U_f|\psi_1\rangle = \left(\frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle\right) \otimes |-\rangle$$

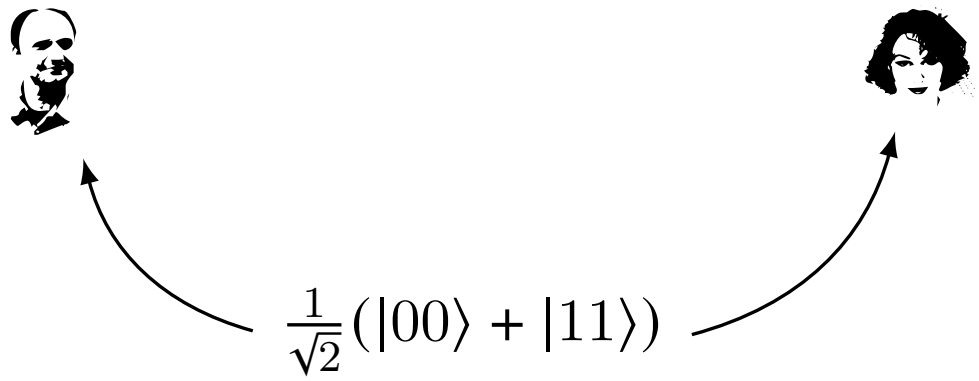
4. apply $H^{\otimes n} \otimes I$

$$|\psi_3\rangle = \left(H^{\otimes n} \otimes I\right)|\psi_2\rangle = \left(\frac{1}{2^n} \sum_{y\in\{0,1\}^n} \sum_{x\in\{0,1\}^n} (-1)^{f(x)+x\cdot y}|y\rangle\right) \otimes |-\rangle$$

5. examine the probability of measuring $|y\rangle = |0\rangle^{\otimes n}$

$$\left|\frac{1}{2^n} \sum_{x\in\{0,1\}^n} (-1)^{f(x)}\right|^2 = \begin{cases} 1 & \text{if } f(x) \text{ is constant} \\ 0 & \text{if } f(x) \text{ is balanced} \end{cases}$$

Teleportation & Superdense Coding



Teleportation Use a shared entanglement and two bits of classical information to transfer one qubit.

Superdense Coding Use a shared entanglement and one qubit of quantum information to transfer two classical bits.

Bell States

$$|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \qquad |\Phi^-\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

$$|\Psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}} \qquad |\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

Quantum Teleportation

Alice and Bob share $|\Phi^+\rangle$. Alice wants to teleport $|\psi\rangle = a|0\rangle + b|1\rangle$ to Bob.

The joint state is

$$|\psi\rangle|\Phi^+\rangle = \frac{|\Phi^+\rangle\otimes(a|0\rangle+b|1\rangle)+|\Phi^-\rangle\otimes(a|0\rangle-b|1\rangle)+|\Psi^+\rangle\otimes(a|1\rangle+b|0\rangle)+|\Psi^-\rangle\otimes(a|1\rangle-b|0\rangle)}{2}.$$

Alice measures her two qubits in the Bell basis $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle$. Then the joint state would collapse to one of the four states with equal probability.

- $|\Phi^+\rangle \otimes (a|0\rangle + b|1\rangle)$
- $|\Phi^-\rangle \otimes (a|0\rangle - b|1\rangle)$
- $|\Psi^+\rangle \otimes (a|1\rangle + b|0\rangle)$
- $|\Psi^-\rangle \otimes (a|1\rangle - b|0\rangle)$

Alice transmits two classical bits to Bob that indicate which of the above four states the system is in.

Bob uses this classical information to apply a correction to his qubit.

$$\begin{array}{ccc} a|0\rangle + b|1\rangle & I & a|0\rangle + b|1\rangle \\ a|0\rangle - b|1\rangle & Z & a|0\rangle + b|1\rangle \\ a|1\rangle + b|0\rangle & X & a|0\rangle + b|1\rangle \\ a|1\rangle - b|0\rangle & ZX & a|0\rangle + b|1\rangle \end{array}$$

Superdense Coding

By applying a quantum gate to $|\Phi^+\rangle$, Alice can transform $|\Phi^+\rangle$ into any of the four Bell states.

Alice's Bits	Initial State	Operation	Final State
00	$\|\Phi^+\rangle$	I	$\|\Phi^+\rangle$
01	$\|\Phi^+\rangle$	X	$\|\Psi^+\rangle$
10	$\|\Phi^+\rangle$	Z	$\|\Phi^-\rangle$
11	$\|\Phi^+\rangle$	ZX	$\|\Psi^-\rangle$

Bob's correction:

Initial State	After CNOT	After H on 1^{st} qubit
$\|\Phi^+\rangle$	$\|+\rangle\|0\rangle$	$\|00\rangle$
$\|\Psi^+\rangle$	$\|+\rangle\|1\rangle$	$\|01\rangle$
$\|\Phi^-\rangle$	$\|-\rangle\|0\rangle$	$\|10\rangle$
$\|\Psi^-\rangle$	$\|-\rangle\|1\rangle$	$\|11\rangle$

Quantum Kolmogorov Complexity

Definition (Quantum Kolmogorov Complexity — Vitányi's Version)

The quantum Kolmogorov complexity of $|x\rangle$ with respect to quantum Turing machine M is

$$K^Q(x) = \min_p \left\{ \ell(p) + \lceil -\log \|\langle z|x\rangle\|^2 \rceil : M(p) = |z\rangle \right\}$$

Definition (Quantum Kolmogorov Complexity — Müller's Version)

Given a QTM M and a finite error $\delta > 0$, the finite-error quantum Kolmogorov complexity of a qubit string $|x\rangle$ is

$$K^Q_\delta(x) = \min_p \left\{ \ell(p) : \|x - M(p)\|_{\mathrm{tr}} < \delta \right\}$$

and the approximate-scheme quantum Kolmogorov complexity of $|x\rangle$ is

$$K^Q(x) = \min_p \left\{ \ell(p) : \forall k \in \mathbb{N} : \|x - M(p,k)\|_{\mathrm{tr}} < \frac{1}{k} \right\}$$

where $\|\cdot\|_{\mathrm{tr}}$ is the trace norm, i.e. $\|\rho - \sigma\|_{\mathrm{tr}} := \frac{1}{2} \operatorname{Tr}\left(\sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)}\right)$.

Quantum Logic

- Any closed linear subspace of — or, equivalently, any projection operator on — a Hilbert space corresponds to a proposition.
- The conjunction $\wedge$ is identified with the intersection of two subspaces. For propositions p, q and their associated closed linear subspaces L_p, L_q: $L_{p \wedge q} = L_p \cap L_q$.
- The disjunction $\vee$ is identified with the closure of the linear span $\oplus$ of the subspaces corresponding to the two propositions. $L_{p \vee q} = L_p \oplus L_q = \{ax + by : a, b \in \mathbb{C}, x \in L_p, y \in L_q\}$.
- The negation $\neg$ is identified with operation of taking the orthogonal subspace $\perp$. $L_{\neg p} = L_p^{\perp} = \{x : \forall y \in L_p : \langle x|y\rangle = 0\}$.
- The implication $\rightarrow$ is identified with the subset relation. $p \rightarrow q \iff L_p \subset L_q$.
- A trivial true statement $\top$ is represented by the entire Hilbert space H. $L_{\top} = H$.
- An absurd statement $\perp$ is represented by the zero vector 0. $L_{\perp} = 0$.

- De Morgan's Law

$$U^\perp \cap V^\perp = (U \oplus V)^\perp$$

$$U^\perp \oplus V^\perp = (U \cap V)^\perp$$

- Law of Double Negation

$$(V^\perp)^\perp = V$$

- Law of Excluded Middle

$$V \oplus V^\perp = H$$

- Law of Non-Contradiction

$$V \cap V^\perp = \{0\}$$

- Law of Contrapositive

$$U \subset V \iff V^\perp \subset U^\perp$$

- In **FdHilb**, $U \subset V \implies V \cap (U \oplus W) = U \oplus (V \cap W)$.
- In **Hilb**, $U \subset V \implies U = V \cap (U \oplus V^\perp)$.

Distributivity Fails

Let A, B, C be three distinct states in $\mathbb{C}^2$, then:

- the meet of any two of them is $\{0\}$;
- the join of any two of them is the whole space $\mathbb{C}^2$.

$$(A \cap B) \oplus C = C \neq \mathbb{C}^2 = (A \oplus C) \cap (B \oplus C)$$

$$(A \oplus B) \cap C = C \neq \{0\} = (A \cap C) \oplus (B \cap C)$$

Contents

Introduction

Induction, Analogy, Fallacy

Term Logic

Propositional Logic

Predicate Logic

Modal Logic

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References2039

Answers to the Exercises — Translation

1. $\neg E$
2. $\neg E \rightarrow \neg U$
3. $U \rightarrow F$
4. $\neg U \wedge \neg M$
5. $D \rightarrow M$
6. $\neg S \rightarrow B$
7. $(W \rightarrow T) \rightarrow \neg S$
8. $\neg L \rightarrow (\neg S \rightarrow P)$
9. $(A \rightarrow B) \rightarrow (W \rightarrow B)$
10. $\neg A \wedge \neg B \rightarrow M \rightarrow N$
11. $\neg E \rightarrow (P \wedge \neg K \rightarrow A)$
12. $(\neg R \rightarrow B) \wedge (R \rightarrow P \vee S)$
13. $(S \wedge \neg C \rightarrow B) \wedge (\neg(S \wedge \neg C) \rightarrow P \vee S)$
14. $(E \rightarrow H) \wedge \neg(H \rightarrow E)$

Answers to the Exercises — Translation

1. $\forall x \forall y (Hxy \rightarrow Uxy)$
2. $\forall x \big(\neg Sx \rightarrow \exists y (Eyx \wedge \neg Sy)\big)$
3. $\forall x (Mx \vee Wx \rightarrow Ax)$
4. $\neg \forall x (Jx \rightarrow Dx)$
5. $\forall x (Jx \rightarrow \neg Dx)$
6. $\forall x \big(\neg (Rx \wedge Bx) \rightarrow \neg Ex\big)$
7. $Hs \wedge Hp \wedge \forall x (Hx \rightarrow x = s \vee x = p)$
8. $\forall x (x \neq s \wedge x \neq p \rightarrow Hx)$
9. $\forall x \big(Bx \rightarrow \exists y \exists z (Gy \wedge Gz \wedge y \neq z \wedge Lxy \wedge Lxz)\big)$
10. $\neg \forall x \big(\mathsf{Glitter}(x) \rightarrow \mathsf{Gold}(x)\big)$
11. $\forall xyz (\mathsf{City}(z) \wedge \mathsf{In}(x,z) \wedge \mathsf{In}(y,z) \rightarrow \mathsf{first}(\mathsf{code}(x)) = \mathsf{first}(\mathsf{code}(y))$
12. $\forall x \forall y (Bx \wedge Gy \wedge Fyx \rightarrow Hx)$
13. $\forall x \forall y (Bx \wedge Gy \wedge Fyx \rightarrow Cxy)$
14. $\forall x (Ex \rightarrow Dx2) \wedge \exists x (Ex \wedge Dx4) \wedge \exists x (Ex \wedge \neg Dx4)$

Answers to the Exercises — Translation

15. $\forall x(Bx \wedge \forall y(Gy \rightarrow Lxy) \rightarrow \neg\exists z(Gz \wedge Lzx))$
16. $\forall x(Cx \rightarrow \exists y(Iyx \wedge \forall z(Izx \rightarrow Lzy)))$
17. $\forall x(Dx \rightarrow Ax) \rightarrow \forall x\big(\exists y(Dy \wedge Hxy) \rightarrow \exists y(Ay \wedge Hxy)\big)$
    $\forall x(Dx \rightarrow Ax) \rightarrow \forall x(Dx \rightarrow \exists y(Ay \wedge hx = hy))$
18. $\forall x\big(\neg D(x, f(w(s))) \rightarrow x = m(w(s))\big)$
19. $\forall x\exists y\exists z(Gy \wedge Gz \wedge y \neq z \wedge Lxy \wedge Lxz \rightarrow x = d)$
20. $\forall x(\neg Lxx \rightarrow Lhx)$
21. $\exists x(Gx \wedge Lxq \wedge \forall y(Gy \wedge Lyq \rightarrow y = x) \wedge Lqx \wedge \forall y(Lqy \rightarrow y = x))$
22. $\exists x(Lxw \wedge \forall y(Lyw \rightarrow y = x) \wedge \exists y(Lwy \wedge \forall z(Lwz \rightarrow z = y) \wedge x \neq y))$
23. $\exists x\Big(Bx \wedge \forall z(Bz \rightarrow Txz) \wedge \exists y\big(Gy \wedge \forall z(Gz \rightarrow Tzy) \wedge Lxy\big)\Big)$
24. $\forall x(Cx \wedge Ex \rightarrow \neg\exists y(Py \wedge Sy) \rightarrow Dx)$
25. $\forall x\big(Bx \rightarrow \neg\exists y(Ty \wedge Axy) \vee \forall y(Ty \rightarrow Axy)\big)$
26. $\forall x\Big(Mx \wedge \exists y\exists z(Cyz \wedge Lxy) \rightarrow \exists yCyx\Big)$
27. $\exists x\exists y\exists z(Mxa \wedge Myb \wedge Mzx \wedge Mzy) \wedge \forall u\forall v\forall x\forall y(Mua \wedge Mvb \wedge Mxu \wedge Myv \rightarrow x = y)$ $\qquad m(m(a)) = m(m(b))$
28. $\exists x\big(Gx \wedge Lxb \wedge \forall y(Gy \wedge Lyb \rightarrow y = x) \wedge \exists y(y \neq x \wedge Sy)\big)$

Answers to the Exercises — Validity

1. $\exists x Jx, \neg\exists x(Ax \wedge Jx) \vdash \exists x \neg Ax$
2. $\forall x(Cx \vee Ax \rightarrow Lxa), \forall x(Zx \rightarrow Cx \vee Ax) \vdash \forall x(Zx \rightarrow Lxa)$
3. $\forall x(\neg Ax \wedge Mx \rightarrow Vx \vee Fx), \forall x(Px \rightarrow \neg Ax \wedge \neg Vx \wedge \neg Fx) \vdash \forall x(Px \rightarrow \neg Mx)$
4. $\forall x(Dx \rightarrow Ax) \rightarrow \forall x\big(\exists y(Dy \wedge Hxy) \rightarrow \exists y(Ay \wedge Hxy)\big)$
5. $\forall x(Px \rightarrow x = a), Pb \wedge Sb \vdash Sa$
6. $Lab, Lbc, Ma, \neg Mc \vdash \exists x \exists y(Mx \wedge \neg My \wedge Lxy)$
7. $\forall x Lxa, \forall x(Lax \rightarrow x = i) \vdash \forall x Lxi$
8. $\forall x \forall y(\exists z Lyz \rightarrow Lxy) \vdash Lrj \rightarrow Lum$
9. $\neg\exists x\big(Bx \wedge \forall y(Sxy \leftrightarrow \neg Syy)\big)$
10. $\forall x \forall y(Fx \wedge Fy \wedge Bxy \rightarrow Sxy) \rightarrow \exists x(Fx \wedge \forall y(Fy \rightarrow Bxy)) \rightarrow \exists x(Fx \wedge \forall y(Fy \rightarrow Sxy))$
11. $\forall x(Sx \rightarrow Kx), \exists x Sx, \forall x(Kx \rightarrow x = a) \vdash Sa$

Answers to the Exercises — Validity

12. $\neg\exists x\exists y(Gx \wedge Sy \wedge Lxy), Ga \wedge \forall x(Lxa \rightarrow Lax), Lba \vdash \neg Sb$
13. $\forall xCx \rightarrow \neg\exists xNx, \forall x\neg Cx \rightarrow \neg\exists xEx \vdash \exists x(Nx \wedge Ex) \rightarrow \exists xCx \wedge \exists x\neg Cx$
14. $\neg\exists x\forall y(Txy \leftrightarrow \neg\exists z(Tyz \wedge Tzy))$
15. $Ri \wedge Pi, \forall x\forall y(Px \wedge Axy \rightarrow Py), \neg\exists x(Px \wedge \neg Ex), \forall x(Ex \wedge Rx \rightarrow \exists y(Gy \wedge Axy)), \forall x(Ex \rightarrow Sx) \vdash \exists x(Gx \wedge Sx)$
16. $\exists x\forall y\big((Ky \leftrightarrow y = x) \wedge Bx\big), \forall x(Bx \rightarrow Sx) \vdash \forall x(Kx \rightarrow Sx)$
17. $Aac \wedge Abc \wedge a \neq b \wedge \forall x(Axc \rightarrow x = a \vee x = b), \forall x(Axc \leftrightarrow Lxc) \vdash \exists x\exists y(Lxc \wedge Lyc \wedge x \neq y \wedge \forall z(Lzc \rightarrow z = x \vee z = y))$
18. $\forall x\forall y(Sxy \rightarrow \exists zIzxy), \forall x\forall y\forall z(Izxy \rightarrow Kzx \wedge Kzy), \forall xSxa \vdash \forall x\exists y(Iyxa \wedge Kya)$

References I

[FST20] Miho Fuyama, Hayato Saigo, and Tatsuji Takahashi. "A category theoretic approach to metaphor comprehension: Theory of indeterminate natural transformation". In: *Biosystems* 197 (2020), p. 104213. ISSN: 0303-2647. DOI: https://doi.org/10.1016/j.biosystems.2020.104213. URL: https://www.sciencedirect.com/science/article/pii/S0303264720301052.

[Gir11] Jean-Yves Girard. "The Blind Spot: Lectures on Logic". In: European Mathematical Society, 2011.

[Law69] F. William Lawvere. "Diagonal arguments and cartesian closed categories". In: *Category theory, homology theory and their applications II*. Springer, 1969, pp. 134–145.

References II

[Let+06] J. C. Letelier et al. "Organizational invariance and metabolic closure: analysis in terms of (M,R) systems". In: *Journal of Theoretical Biology* 238.4 (2006), pp. 949–961. DOI: 10.1016/j.jtbi.2005.07.007.

[Lou09] A. H. Louie. *More Than Life Itself — A Synthetic Continuation in Relational Biology*. Berlin, Boston: De Gruyter, 2009. ISBN: 9783110321944. DOI: 10.1515/9783110321944. URL: https://doi.org/10.1515/9783110321944.

[Luz+09] María Luz Cárdenas et al. "Closure to efficient causation, computability and artificial life". In: *Journal of Theoretical Biology* 263.1 (2009), pp. 79–92. ISSN: 0022-5193. DOI: https://doi.org/10.1016/j.jtbi.2009.11.010. URL: https://www.sciencedirect.com/science/article/pii/S0022519309005360.

References III

[MLS09] Matteo Mossio, Giuseppe Longo, and John Stewart. "A computable expression of closure to efficient causation". In: *Journal of Theoretical Biology* 257.3 (2009), pp. 489–498. ISSN: 0022-5193. DOI: https://doi.org/10.1016/j.jtbi.2008.12.012. URL: https://www.sciencedirect.com/science/article/pii/S0022519308006498.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511976667.

References IV

[Ros72] Robert Rosen. "Some relational cell models: the metabolism-repair system". In: *Foundations of Mathematical Biology*. Ed. by Robert Rosen. Academic Press, 1972, pp. 217–253. ISBN: 978-0-12-597202-4. DOI: https://doi.org/10.1016/B978-0-12-597202-4.50011-6. URL: https://www.sciencedirect.com/science/article/pii/B9780125972024500116.

[San20] Paulo Guilherme Santos. *Diagonalization in Formal Mathematics*. Springer Spektrum, Wiesbaden, 2020. DOI: https://doi.org/10.1007/978-3-658-29111-2.

[SV84] Jorge Soto-Andrade and Francisco J. Varela. "Self-reference and fixed points: A discussion and an extension of Lawvere's Theorem". In: *Acta Applicandae Mathematica* 2 (1984), pp. 1–19.

References V

[Svo18] Karl Svozil. *Physical (A)Causality Determinism, Randomness and Uncaused Events*. Fundamental Theories of Physics, 192. Cham: Springer International Publishing, 2018. ISBN: 3-319-70815-5.

[Sza18] Jochen Szangolies. "Epistemic Horizons and the Foundations of Quantum Mechanics". In: *Foundations of Physics* 48.12 (2018), pp. 1669–1697. DOI: 10.1007/s10701-018-0221-9.

[Uni13] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. Institute for Advanced Study: https://homotopytypetheory.org/book, 2013.

Thank 