# Introduction to Logic

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

April 12, 2023

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Contents

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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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

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

**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。

# Unfaithful Husband Puzzle

## 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?"

# 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Logic vs other Disciplines

- 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 / strength & limitation / paradoxes ...
  Peirce, Frege, Russell, Wittgenstein, Ramsey, Carnap, Quine, Putnam, Kripke, Chomsky, Gödel, Tarski, Turing ...
- Logic vs Mathematics.
  Logicism / Formalism / Intuitionism / Constructivism / Finitism / Structuralism / Homotopy Type Theory
- Logic vs Computer Science.

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

# Logic vs CS

- Computer Architecture.
  Logic gates and digital circuit design $\approx$ Propositional Logic
- Programming Languages.
  LISP $\approx$ $\lambda$-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

# Logic vs other Disciplines

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

《论语》

“工欲善其事，必先利其器。”

《荀子·劝学》

“君子生非异也，善假于物也。”

# Branches of Logic

**Mathematical Logic**

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

**Computational Logic**

- Automata Theory
- Computational Complexity
- Finite Model Theory
- Model Checking
- Lambda Calculus
- Categorical Logic
- Theorem Proving
- Description Logic
- Fixpoint Logic
- Dynamic 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
- Justification Logic
- Hybrid Logic
- Substructural Logic
- Free Logic
- Counterfactual Logic
- Relevance Logic
- Linear Logic
- Quantum Logic
- Paraconsistent Logic
- Intensional Logic
- Partial Logic
- Diagrammatic Logic
- Deontic Logic

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

- 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

Keep yourself open

and don't neglect your larger self

# 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
- Philosophy is to science as pornography is to sex: it is cheaper, easier, and some people seem, bafflingly, to prefer it. — *John Jones*
- 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

# Practice is the sole criterion for testing truth?

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”?

# Why is the Night Sky Dark?

A static, infinitely old universe with an infinite number of stars uniformly distributed in an infinitely large space would be bright rather than dark.

# Newton's Apple

If an apple falls, does the moon also fall?

## 力是改变物体运动状态的原因

- 什么是“运动”?
- 什么是“状态”?
- 什么是“运动状态”?
- 什么是“改变”?
- 什么是“物体”?
- 什么是“原因”?
- 什么是“力”?
- 什么是“定义”?

# Einstein's Train Thought Experiment

- 什么是“同时”?
- 怎么辨别“同时”?
- 怎么测量“时间”?

# Einstein's Elevator Thought Experiment

- The gravitational "force" as experienced locally while standing on a massive body is actually the same as the pseudo-force experienced by an observer in a non-inertial (accelerated) frame of reference.
- Spacetime tells matter how to move; matter tells spacetime how to curve.
- Gravity is not a force that applies via Newton's $2^{\text{nd}}$ Law, but a consequence of the curvature of spacetime caused by the uneven distribution of mass/energy that acts via the geodesic principle, which is the relativistic equivalent of Newton's $1^{\text{st}}$ Law.
- Causality (causally connected regions, finite speed of light)

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

**How to *express* your thoughts <u>precisely</u> and <u>succinctly</u>?**

- Logic builds our minds.
- Not only is the universe stranger than we imagine, it is stranger than we can imagine.
- Logic enlarges our abstract imagination.
- A few lines of reasoning can change the way we see the world.
- Logic is the immune system of the mind!
- Logic can be fun — The music of reason.

What are the extent and limits of reason?

*That which does not kill me, makes me stronger.*

*— F. Nietzsche*

# Contents

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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Readings

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. Y. Manin: A Course in Mathematical Logic for Mathematicians. — $M$

# Readings, Movies and More

参考书 **陈波：《逻辑学导论》**

参考书 郝兆宽、杨睿之、杨跃：《数理逻辑：证明及其限度》

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。

- D. Hofstadter: Gödel, Escher, Bach
- Dangerous Knowledge
- The Imitation Game

- libgen
- sci-hub

# 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

# Aim

- Formalization of an argument ♡
- Demonstration of the validity of an argument ♡
- Object & Meta-language / Syntax & Semantics / Finite & Infinite / Countable & Uncountable / Induction & Recursion / Truth & Proof / Axiomatization / Theory / Soundness / Completeness / Compactness / Elementary Equivalent & Isomorphism / Representability / Definability / Categoricity / Decidability / Complexity / Expressiveness / Succinctness / Interpretability . . . 🚲
- Formal Philosophy 🚲
- Understanding of the nature of mathematics 🚲
- Applications in Math / CS / AI / Linguistics / Cognition / Physics / Information Theory / Game Theory / Social Science . . . 🚲
- Mathematical Logic 🚲

# Outline and Credits

- Critical Thinking ✓
- Propositional Logic ✓
- Term Logic
- Predicate Logic ✓
- Modal Logic
- Set Theory

- Question
- Discussion
- Homework
- Exercises ✓
- Examination ✓
- Presentation
- Paper
- Techniques e.g. LaTeX / Coq ...
- ...

# 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

- Leibniz, Cantor, Frege, Russell, Hilbert, Gödel, Tarski, Turing.
- syntax, semantics, formal system, entailment, proof, consistency, satisfiability, validity, soundness, completeness, compactness, decidability, complexity, expressive power, succinctness. . .
- Philosophy of Logic, Philosophical Logic
- Logicism, Formalism, Intuitionism
- Hilbert's program
- Church-Turing thesis

# 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：小白、大黄和老王。至少有一人是凶手，但不可能三人同时犯罪。如果小白是凶手，那么老王是同犯。如果大黄不是凶手，那么老王也不是。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)

人人都怕德古拉，德古拉只怕我，因此，我就是德古拉。

五、论述题。(10 分)

谈谈你对“希尔伯特规划”的理解。

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Contents

Introduction

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

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Aristotle(384-322 BC)

- 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

# 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 等人认为是 数理逻辑[1]、拓扑学、博弈论、控制论、分形几何、非标准分析、算法信息论、计算主义哲学的先驱。

[1] Wolfgang Lenzen: Leibniz's Logic.

# 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 / Probability

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:**

$$\frac{\begin{array}{c} MAP \\ SIM \end{array}}{SIP} \qquad \frac{\begin{array}{c} \exists X(M = PX) \\ \forall Y(S \neq \overline{M}Y) \end{array}}{\forall Z(S \neq \overline{P}Z)}$$

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.

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

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

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

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

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

1. $\neg\neg A \leftrightarrow A$
2. $\neg(A \leftrightarrow \neg A)$
3. $(A \to B) \leftrightarrow (\neg B \to \neg A)$
4. $\Diamond A \to (A \to B) \to \neg(A \to \neg B)$
5. $A \land \neg A \to B$

1. $(A \to B) \to \Diamond A \to \Diamond B$
2. $(A \to B) \leftrightarrow \neg\Diamond(A \land \neg B)$
3. $\neg\Diamond(A \land \neg A)$

$$\Diamond A \coloneqq \neg(A \to B \land \neg B)$$

# 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 |

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

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

- 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

# George Boole 1815-1864

Figure: self-taught mathematician, philosopher, logician, and poet

- *The Laws of Thought*
- Propositional Logic
  Boolean Algebra
- Logic as Algebra

$$\begin{tikzcd}
\text{proposition} \arrow[r, "\text{Logic}", "deduce"'] \arrow[d, "\text{translate}"'] & \text{proposition} \\
\text{equation} \arrow[r, "calculate", "\text{Algebra}"'] & \text{equation} \arrow[u, "\text{translate}"']
\end{tikzcd}$$

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

# Logic → Truth

“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

# Contents

Introduction

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

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 命题逻辑

- 语言
  原子命题、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 (合式公式 Well-Formed Formula $\mathrm{Wff}$ — Top Down)

1. 原子命题 $p \in \mathrm{Var}$ 是合式公式.
2. 如果 $A$ 是合式公式，则 $(\neg A)$ 也是合式公式.
3. 如果 $A, B$ 是合式公式，则 $(A \wedge B), (A \vee B), (A \rightarrow B), (A \leftrightarrow B)$ 也是合式公式.
4. 除此之外，别无其他合式公式.

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

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

# Example

“大部分人一听到数学的名字就害怕，以至于常常过分夸大自己在数学上的愚蠢。”

— 哈代

1. Lily is beautiful.
2. Lily is not beautiful.
3. Lily is beautiful and Lucy is smart.
4. Lily is beautiful or Lucy is smart.
5. If wishes were wings, then pigs would fly.
6. Lily is beautiful if and only if Lucy is smart.
7. 侠之大者，为国为民。
8. 只要功夫深，铁杵磨成针。/ 砍头不要紧，只要主义真。
9. 只有你请她才来。/ 除非你请，否则她不会来。
10. 没有共产党，就没有新中国。
11. 不是你死，就是我亡。
12. 鱼与熊掌不可兼得。
13. 欲寄君衣君不还，不寄君衣君又寒。寄与不寄间，妾身千万难。

数学无处不在！

# Well-Formed Formula

## Definition (Formula-Building Operators)

$$
\begin{aligned}
f_{\neg}(A) &:= (\neg A) \\
f_{\wedge}(A, B) &:= (A \wedge B) \\
f_{\vee}(A, B) &:= (A \vee B) \\
f_{\rightarrow}(A, B) &:= (A \rightarrow B) \\
f_{\leftrightarrow}(A, B) &:= (A \leftrightarrow B)
\end{aligned}
$$

$$
\boxed{
\begin{aligned}
f_{\neg}(A) &:= \neg A \\
f_{\wedge}(A, B) &:= \wedge AB \\
f_{\vee}(A, B) &:= \vee AB \\
f_{\rightarrow}(A, B) &:= \rightarrow AB \\
f_{\leftrightarrow}(A, B) &:= \leftrightarrow AB
\end{aligned}
}
$$

# 省略括号的约定 ♡

1. 公式最外层的括号可以省略

$$A \to B \equiv (A \to B)$$

2. 连接词的连接强度：$\neg$ 强于 $\wedge, \vee$ 强于 $\to, \leftrightarrow$

$$\neg A \vee B \to C \wedge D \equiv ((\neg A) \vee B) \to (C \wedge D)$$

$$1 + 2 * 3$$

3. 同一连接词反复出现时，右向结合

$$A \to B \to C \to D \equiv A \to (B \to (C \to D))$$

# $\mathrm{Wff}$ — 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 $\mathrm{Wff}$ — Bottom Up)

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

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

# 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 $B \subset U$?

## Bottom Up

$$C_0 := B$$

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

$$C_* := \bigcup_{n \in \mathbb{N}} C_n$$

## Top Down

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

# Bottom Up vs Top Down

## $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。

# Bottom Up vs Top Down

## Example

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

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

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

## Theorem (Bottom Up vs Top Down)

$$C_* = C^*$$

## Proof.

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

# Induction Principle for Formulas

**Theorem (公式上的归纳法)**

令 $P$ 是一个关于合适公式的性质。假设

1. 所有原子命题都有性质 $P$
2. 对任意合式公式 $A, B$，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，所有合式公式都有性质 $P$.

**Proof.**

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

$$P(0) \wedge \forall n(P(n) \to P(n+1)) \to \forall n \in \mathbb{N} P(n)$$

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

# Induction vs Recursion

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

1. If ever you leave milk one day, be sure and leave it the next day as well.
2. Leave milk today.

$\boxed{\text{Leave milk today and read this note again tomorrow.}}$

# 子公式

## Definition (Subformula — Inductive Definition)

The set $\mathrm{Sub}(A)$ of subformulas of a wff $A$ is the smallest set $\Gamma$ 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 (Subformula — Recursive Definition)

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

# 唯一可读性 vs 唯一语法树

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

- 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 five formula-building operators, when restricted to the set of wffs,*

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

**Proof.**

$$
\begin{aligned}
(A \star B) &= (C * D) \\
&\Downarrow \\
A \star B) &= C * D) \\
&\Downarrow \\
A &= C && \text{(Lemma)} \\
&\Downarrow \\
\star &= * \\
&\Downarrow \\
B &= D && \text{(Lemma)}
\end{aligned}
$$

# Contents

Introduction

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

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 真值赋值 ♡

- A truth assignment is a function

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

- Such a truth assignment can be uniquely extended to $\overline{\nu} : \mathrm{Wff} \to \{0, 1\}$ satisfying the following condition[2]:

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 |

[2] Some authors use $[\![\,]\!]_\nu$ for $\overline{\nu}$.

# 真值表 ♡

| $p$ | $\neg p$ |
|---|---|
| 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 |

充分 → 必要

## Example

- If $0 = 1$, then Russell is God.
- Snow is white iff $1 + 1 = 2$.

# Example

## Example

Assume $\nu(p) = 1, \nu(q) = 0, \nu(r) = 1$. Is $\overline{\nu}(p \vee r \to \neg(p \to q)) = 1$?

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

**Remark:** Whenever there is no confusion we omit explicit mention of $\nu$.

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

## Definition (Freely Generated 🚲)

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

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

## Theorem (Recursion Theorem)

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

$$\begin{tikzcd} B \arrow[r] \arrow[dr, "h"'] & C \arrow[d, dashed, "\overline{h}"] \\ & V \end{tikzcd}$$

1. $\overline{h}\restriction_B = h$
2. *for all $f \in \mathcal{F}$ and all $x_1, \ldots, x_n \in C$:*
$$\overline{h}(f(x_1, \ldots, x_n)) = F_f\big(\overline{h}(x_1), \ldots, \overline{h}(x_n)\big)$$

**Remark:**

- $h$ tells you how to color the initial elements in $B$;
- $F_f$ tells you how to convert the color of $\mathbf{x}$ into the color of $f(\mathbf{x})$.

Danger! $F_f$ is saying "green" but $F_g$ is saying "red" for the same point!

# 实质蕴含 — 反直觉?

| $p$ | $q$ | $p \to q$ | If $x$ is rational, then $x^2$ is rational. |
|---|---|---|---|
| 0 | 0 | 0 | $x = \pi$ |
| 0 | 1 | 0 | $x = \sqrt{2}$ |
| 1 | 0 | 0 | |
| 1 | 1 | 1 | $x = \frac{1}{2}$ |

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!

# Find the Error ☺

How do you solve this equation $y + 2 = y$?

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

# Entailment & Validity ♡

If lily is beautiful, then the fact that snow is white implies lily is beautiful.

| $p$ | $q$ | $q \to p$ | $p \to q \to p$ |
|---|---|---|---|
| 0 | 0 | 1 | 1 |
| 0 | 1 | 0 | 1 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 1 | 1 |

There are $2^n$ truth assignments for a set of $n$ atomic propositions.

- Satisfiability: $A$ is *satisfiable* iff there is some truth assignment $\nu$ s.t. $\overline{\nu}(A) = 1$. (We also write $\nu \vDash A$ if $\overline{\nu}(A) = 1$.)
- Entailment: $\Gamma \vDash A$ iff for any truth assignment $\nu$ s.t. $(\text{for all } B \in \Gamma : \nu \vDash B) \implies \nu \vDash A$.
- Validity / Tautology: $A$ is valid $\vDash A$ iff $\emptyset \vDash A$, in other words, for any truth assignment $\nu$, $\nu \vDash A$.
- Logical Equivalence: $A$ and $B$ are logically equivalent $A \equiv B$ iff for any truth assignment $\nu$, $\nu \vDash A \iff \nu \vDash B$.

| 语义蕴含 | 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 | $A \equiv B$ |

| always true<br>tautology | sometimes true/sometimes false<br>contingency(synthetic) | always false<br>contradiction |
|---|---|---|
| valid | invalid | |
| satisfiable | | unsatisfiable |

## Example

Show that $p \wedge (\neg q \vee \neg r) \rightarrow (p \rightarrow \neg q)$ is neither a tautology nor a contradiction.

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

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

# Truth Assignment — set-based version 🚲

A truth assignment for $\mathscr{L}^0$ is a function

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

- $[\![p]\!] := \{v : v(p) = 1\}$
- $[\![\top]\!] := \{0, 1\}^{\mathrm{Var}}$
- $[\![\bot]\!] := \emptyset$
- $[\![\neg A]\!] := \{0, 1\}^{\mathrm{Var}} \setminus [\![A]\!]$
- $[\![A \wedge B]\!] := [\![A]\!] \cap [\![B]\!]$
- $[\![A \vee B]\!] := [\![A]\!] \cup [\![B]\!]$
- $[\![A \to B]\!] := \left(\{0, 1\}^{\mathrm{Var}} \setminus [\![A]\!]\right) \cup [\![B]\!]$

$$A \vDash B \iff [\![A]\!] \subset [\![B]\!]$$

# 什么是“有效论证”? ♡

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

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

是有效的当且仅当 $A_1, \ldots, A_n \vDash B$。

- 如果前提 $A_1, \ldots, A_n$ 为真，那么结论 $B$ 也为真。
- 如果结论 $B$ 为假，那么至少有一个前提 $A_i$ 为假。

**Example**

$$\frac{\begin{array}{c}\text{An integer } x \text{ is even or odd.}\\ \text{If } x \text{ is even, then } x+x \text{ is even.}\\ \text{If } x \text{ is odd, then } x+x \text{ is even.}\end{array}}{\text{Therefore, } x+x \text{ is even.}} \qquad \frac{\begin{array}{c}p \lor q\\ p \to r\\ q \to r\end{array}}{r}$$

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

兴，百姓苦；亡，百姓苦。

# Valid Argument — Example ☹

克莱因是教授，因此，克莱因或者是教授或者是连环杀手。

$$\frac{p}{p \vee q}$$

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。

$$\frac{p \wedge q}{p}$$

# 形式与特例 — 忽略一切无关细节！

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

**Problem (How old is my father?)**

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 (How many apples are there?)**

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。

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

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

# 抽象与形式

## 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。”

- A sculptor subtracts irrelevant wood until a sculpture emerges.
- A mathematician abstracts irrelevant details until a pattern emerges.

# 逻辑形式

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

给定一个逻辑形式，对相同的变元处处以相同的命题替换，得到的命题是这个逻辑形式的特例。

$$p \to p \vee q$$

$$\heartsuit \to \heartsuit \vee$$

$$A \to A \vee B$$

$$A \to A \vee A$$

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

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

# 有效式 / 重言式 — Example

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

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

# 重言式的归谬赋值法

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

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $\underline{\mathbf{0}}$ | | | | | | |
| | 1 | | | | $\mathbf{0}$ | | | | $\underline{\mathbf{0}}$ | | |
| | 1 | | | | $\mathbf{0}$ | | 1 | | $\mathbf{0}$ | | $\underline{\mathbf{0}}$ |
| 1 | 1 | | | 0 | $\mathbf{0}$ | 1 | $\underline{\mathbf{1}}$ | | $\mathbf{0}$ | 1 | $\mathbf{0}$ | 0 |
| 1 | $\underline{\mathbf{1}}$ | 1 | | 0 | $\mathbf{0}$ | 1 | $\mathbf{1}$ | 1 | $\mathbf{0}$ | 1 | $\mathbf{0}$ | 0 |
| 1 | $\mathbf{1}$ | 1 | $\underline{\mathbf{1}}$ | 0 | $\mathbf{0}$ | 1 | $\mathbf{1}$ | 1 | $\mathbf{0}$ | 1 | $\mathbf{0}$ | 0 |
| | | | × | | | | | | | | | |

# 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. $\neg \mathrm{PassLogic} \rightarrow \neg \mathrm{Graduate}$
  2. The store is open unless/except it is sunday. $\mathrm{Open} \leftrightarrow \neg \mathrm{Sunday}$
- **Even If 即使:**[3]
  - I'm going to the party even if it rains. $p \wedge (q \rightarrow p) \equiv p$
  - The Allies would have won even if the U.S. had not entered the war. $p \wedge (\neg q \rightarrow p) \equiv p$
  - The Axis powers would have won if the U.S. had not entered the war. $\neg p \wedge (\neg q \rightarrow p)$

[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.

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

- programmer.

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

# 练习：翻译 — Now it's your turn ✍♡✍

1. I am not good at logic.
2. Either Jones is a fool or he 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.

# Translation — Natural language is not always precise!

1. If I get in the lift then it breaks, and if you get in then the lift breaks. (?)
2. If we both get in the lift, then the lift breaks.(?)[4]

$$\frac{p \vee q \to r}{\overline{(p \to r) \wedge (q \to r)}}$$

$$\frac{p \wedge q \to r}{\overline{(p \to r) \vee (q \to r)}}$$

- I weigh $100$ pounds.
- You weigh $100$ pounds.
- The maximum capacity of the lift is $150$ pounds.

[4]这里不方便细分，$p \to q$. 若细分，或许需要 Linear Logic $p \otimes q \multimap r$.

# Translation

*Philosophy is written in that great book — the universe — which ever is before our eyes, but it can't be understood unless one first learns to comprehend the mathematical language, without which one wanders in vain through a dark labyrinth.*

*— Galileo Galilei*

$$P \wedge E \wedge (\neg C \rightarrow \neg U) \wedge (\neg C \rightarrow W)$$

Figure: Leaning Tower of Pisa

# 练习：有效性判定 — Now it's your turn ✍♡✍

1. $p \lor q \equiv \neg p \to q \equiv (p \to q) \to q$
2. $p \land q \equiv \neg(p \to \neg q)$
3. $p \leftrightarrow q \equiv (p \to q) \land (q \to p)$
4. $p \lor (q \lor r) \equiv (p \lor q) \lor r$
5. $p \land (q \land r) \equiv (p \land q) \land r$
6. $p \lor q \equiv q \lor p$
7. $p \land q \equiv q \land p$
8. $p \land (q \lor r) \equiv (p \land q) \lor (p \land r)$
9. $p \lor (q \land r) \equiv (p \lor q) \land (p \lor r)$
10. $\neg(p \lor q) \equiv \neg p \land \neg q$
11. $\neg(p \land q) \equiv \neg p \lor \neg q$
12. $p \lor (p \land q) \equiv p$
13. $p \land (p \lor q) \equiv p$

1. $\neg\neg p \to p$
2. $p \to \neg\neg p$
3. $p \lor \neg p$
4. $\neg(p \land \neg p)$
5. $p \land \neg p \to q$
6. $(p \to q) \land (\neg p \to q) \to q$
7. $(p \to q) \land (p \to \neg q) \to \neg p$
8. $(\neg p \to q) \land (\neg p \to \neg q) \to p$
9. $(\neg p \to \bot) \to p$
10. $(\neg p \to p) \to p$
11. $((p \to q) \to p) \to p$

$$\frac{\Gamma, A \vDash B}{\Gamma \vDash A \to B} \qquad \frac{A \equiv B}{\vDash A \leftrightarrow B} \qquad \frac{A \lor B \quad \neg B \lor C}{A \lor C}$$

# 无效论证

$$\frac{\begin{array}{c} p \to q \\ \neg p \end{array}}{\neg q} \qquad \frac{\begin{array}{c} p \to q \\ q \end{array}}{p} \qquad \frac{\begin{array}{c} p \vee q \\ p \end{array}}{\neg q}$$

| | | |
|---|---|---|
| 我思故我在<br>我不思 | 我思故我在<br>我在 | 此论证或是有效的或是无效的<br>此论证是有效的 |
| 故我不在 | 故我思 | 此论证不是无效的 |

- 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。

*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*

- Mother: "If you eat your broccoli you can have ice cream."
- Child: "How about if I eat some fish instead?"
- Mother: "You can eat ice cream only if you eat broccoli."

# 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》

$$\frac{\text{主子豢养} \rightarrow \text{是走狗}}{\text{无主子豢养} \rightarrow \text{不是走狗}} \times$$

# 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？”

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

# 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。**"

好货不贱，贱货不好。

## 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。

# Example

- 如果你工作，就能挣钱；如果你赋闲在家，就能悠然自在。你要么工作要么赋闲，总之，你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。你要么工作要么赋闲，总之，你不能悠然自在或者不能挣钱。

$$\frac{\begin{array}{c} p \to r \\ q \to s \end{array}}{p \lor q \to r \lor s} \qquad \frac{\begin{array}{c} p \to \neg s \\ q \to \neg r \end{array}}{p \lor q \to \neg s \lor \neg r}$$

- 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?

# Example

## 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。

$$w \to p, \neg w \to p, w \vee \neg w \vDash p$$

$$w \wedge j \to p, \neg w \wedge c \to p, w \vee \neg w \overset{?}{\vDash} p$$

$$\neg w \wedge j \to \neg p, w \wedge c \to \neg p, w \vee \neg w \overset{?}{\vDash} \neg p$$

$$w \wedge j \to p, \neg w \wedge c \to p, (w \wedge j) \vee (\neg w \wedge c) \vDash p$$

# 鳄鱼悖论

## The Crocodile Dilemma

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

$$x = ? \implies \vDash (x \leftrightarrow r) \to r$$

| $r$ | $(\neg r \leftrightarrow r) \to r$ |
|---|---|
| 0 | 1 |
| 1 | 1 |

$$((r \vee \neg r) \leftrightarrow r) \to r$$

# 怎么得大奖?

**Problem (怎么得大奖? )**

- 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。

$$x = ? \implies \vDash (x \leftrightarrow b \vee s) \rightarrow b$$

| $b$ | $s$ | $(\neg b \wedge \neg s \leftrightarrow b \vee s) \rightarrow b$ | $(\neg s \leftrightarrow b \vee s) \rightarrow b$ | $((s \rightarrow b) \leftrightarrow b \vee s) \rightarrow b$ |
|---|---|---|---|---|
| 0 | 0 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 |

# Gateway to Heaven

**Problem (天堂之路)**

- 你面前有左右两人守卫左右两门。
- 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。
- 怎么问出去天堂的路?

# Gateway to Heaven

**Problem (天堂之路)**

- 你面前有左右两人守卫左右两门。
- 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。
- 怎么问出去天堂的路?

- p: 你说真话
- q: 左门通天堂

$$x = ? \implies \vDash (p \to (x \leftrightarrow q)) \land (\neg p \to (x \leftrightarrow \neg q))$$

| $p$ | $q$ | $p \leftrightarrow q$ | report | $(p \to ((p \leftrightarrow q) \leftrightarrow q)) \land (\neg p \to ((p \leftrightarrow q) \leftrightarrow \neg q))$ |
|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 1 |
| 0 | 1 | 0 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 |
| 1 | 1 | 1 | 1 | 1 |

# Logical Equivalence

- 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.

$$\frac{A \equiv A'}{\neg A \equiv \neg A'}$$

$$\frac{\begin{array}{c} A \equiv A' \\ B \equiv B' \end{array}}{A \star B \equiv A' \star B'}$$

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)**

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

**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 wff $A$ and a sequence $C_1, \ldots, C_n$ of wffs.*

1. *Let $\nu$ be a truth assignment for the set of all atomic propositions. Define $\mu$ 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. $\vDash A \implies \vDash A[C_1/p_1, \ldots, C_n/p_n]$

$$\frac{A}{A[C_1/p_1, \ldots, C_n/p_n]}$$

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

# 等价替换

**Theorem (Equivalent Replacement)**

*Suppose $B \in \mathrm{Sub}(A)$, and $A[C /\!\!/ B]$ arises from the wff $A$ by replacing one or more occurrences of $B$ in $A$ by $C$. Then*

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

**Proof.**

Prove by induction.

# 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.

$$p?$$

- logician.

$$\frac{q \leftrightarrow \neg p}{p \vee q \leftrightarrow p \vee \neg p}\ (p \vee q)[\neg p /\!\!/ q] \qquad \checkmark$$

# Duality

## Theorem

*Let $A$ be a wff 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

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

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Formal Systems

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

# 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.

# Hilbert Formal System = Axiom + Inference Rule

## Axiom Schema

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

## Inference Rule

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

## Lewis Carroll's Paradox: What the Tortoise Said to Achilles ☹

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。

# 什么是“证明”?

$\boxed{\text{This sentence can never be proved.}}$

## Definition (Proof)

A proof from $\Gamma$ is a deduction sequence of wff $(C_1, \ldots, C_n)$ s.t. for $k \leq n$, either

1. $C_k \in \Gamma$, or
2. $C_k$ is an axiom, or
3. for some $i < k$ and $j < k$, $C_i = C_j \to C_k$.

- $\Gamma \vdash A$ iff $A$ is the last member of some deduction sequence from $\Gamma$.
- $A$ is a theorem $\vdash A$ iff $\emptyset \vdash A$.

## A Joke ☺

A mathematician's house is on fire. His wife puts it out with a bucket of water. Then there is a gas leak. The mathematician lights it on fire.

# Aristotle's Four Causes

1. Material Cause: axioms
2. Efficient Cause: inference rules
3. Final Cause: true propositions
4. Formal Cause: algorithm(final cause) = proof

$\boxed{\text{形式因}}$ 动力因 动力因 $\cdots$ 动力因

质料因 $\longrightarrow$ 目的/质料因 $\longrightarrow$ $\cdots$ $\longrightarrow$ 目的/质料因 $\longrightarrow$ 目的因

# Proof vs Understanding / Belief / Knowledge

# Example

## Theorem

$$\vdash p \to p$$

## Proof.

1. $p \to (p \to p) \to p$ A1
2. $(p \to (p \to p) \to p) \to (p \to p \to p) \to p \to p$ A2
3. $(p \to p \to p) \to p \to p$ 1,2 MP
4. $p \to p \to p$ A1
5. $p \to p$ 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**

$$\vdash (\neg p \to p) \to p$$

**Proof.**

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

# Example

## Theorem

$$p \to q, q \to r \vdash p \to r$$

## Proof.

| | | |
|---|---|---|
| 1. | $(q \to r) \to (p \to q \to r)$ | A1 |
| 2. | $q \to r$ | Premise |
| 3. | $p \to q \to r$ | 1,2 MP |
| 4. | $(p \to q \to r) \to (p \to q) \to p \to r$ | A2 |
| 5. | $(p \to q) \to p \to r$ | 4,3 MP |
| 6. | $p \to q$ | Premise |
| 7. | $p \to r$ | 5,6 MP |

# Example — Curry's Paradox ☹

If this sentence is true, then God exists.

$$p \leftrightarrow (p \to q) \vdash q$$

Proof.

1. $p \leftrightarrow (p \to q)$
2. $p \to p \to q$
3. $(p \to p) \to p \to q$
4. $p \to q$
5. $p$
6. $q$

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。

This sentence is false, and God does not exist.

# Curry's Paradox — How to Flirt with a Beauty ☺

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!

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

## 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?

# 布尔代数

| $\bot$ | $\top$ | $\vee$ | $\wedge$ | $\neg$ |
|---|---|---|---|---|
| $0$ | $1$ | $+$ | $\cdot$ | $\overline{\phantom{x}}$ |

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

**Theorem (布尔方程的通解)**

*If* $f(0) \cdot f(1) = 0$, *then*

$$f(x) = 0 \implies x = f(0) + \theta \cdot \overline{f(1)}$$

*where* $\theta \in \{0, 1\}$.

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

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

$$\Downarrow$$

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

$$x = \begin{cases} \neg a \wedge \neg k & \text{if } \theta = 0 \\ a \to k & \text{if } \theta = 1 \end{cases}$$

# 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. $C_i = C_j \rightarrow B$

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

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

$\Gamma \vdash A \rightarrow B$

# Tree Method for Propositional Logic ♡

$$\begin{array}{c} \neg\neg A \\ | \\ A \end{array} \qquad \begin{array}{c} A \to B \\ \diagup \quad \diagdown \\ \neg A \qquad B \end{array} \qquad \begin{array}{c} \neg(A \to B) \\ | \\ A \\ \neg B \end{array}$$

---

$$\begin{array}{c} A \wedge B \\ | \\ A \\ B \end{array} \qquad \begin{array}{c} \neg(A \wedge B) \\ \diagup \quad \diagdown \\ \neg A \qquad \neg B \end{array} \qquad \begin{array}{c} A \vee B \\ \diagup \quad \diagdown \\ A \qquad B \end{array} \qquad \begin{array}{c} \neg(A \vee B) \\ | \\ \neg A \\ \neg B \end{array}$$

---

$$\begin{array}{c} A \leftrightarrow B \\ \diagup \quad \diagdown \\ \begin{array}{c} A \\ B \end{array} \qquad \begin{array}{c} \neg A \\ \neg B \end{array} \end{array} \qquad \begin{array}{c} \neg(A \leftrightarrow B) \\ \diagup \quad \diagdown \\ \begin{array}{c} A \\ \neg B \end{array} \qquad \begin{array}{c} \neg A \\ B \end{array} \end{array}$$

✓

# Examples — 证明 = 闭树

$$\boxed{\dfrac{p \to q \quad \neg q \vee r}{\neg p \vee r}}$$

$p \to q$ ✓

$\neg q \vee r$ ✓

$\neg(\neg p \vee r)$ ✓

$\neg\neg p$

$\neg r$

- $\neg p$ ×
- $q$
  - $\neg q$ ×
  - $r$ ×

# 树形方法指南 ♡

文字 一个原子命题或原子命题的否定。

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.

# 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.

## 证明技巧

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 (Soundness & Completeness Theorem)

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

**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生成使得论证无效的反模型。

# 若无效，“开枝”给出“反模型”的真值赋值

$$\frac{p \lor q \quad r \lor s \quad s \to t}{r} \; ?$$

$$p = 1, \quad t = 1, \quad s = 1 \quad r = 0 \quad q = 1 \text{ (or } 0)$$

$$q = 1, \quad t = 1, \quad s = 1 \quad r = 0 \quad p = 1 \text{ (or } 0)$$

$$p \lor q = 1, \quad r \lor s = 1, \quad s \to t = 1, \quad r = 0$$

**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 ✍♡✍

$$\frac{}{(p \to q) \lor (q \to r)} \qquad \frac{p \to r \quad q \to s}{p \land q \to r \land s} \qquad \frac{p \to r \quad q \to s}{\neg r \lor \neg s \to \neg p \lor \neg q}$$

$$\frac{p \to q \quad \neg p \to q}{q} \qquad \frac{p \to r \quad q \to r}{p \lor q \to r} \qquad \frac{p \to r \quad q \to s}{p \lor q \to r \lor s}$$

$$\frac{(p \to q) \to r}{q \to r} \qquad \frac{(p \to q) \to p}{p} \qquad \frac{p \leftrightarrow (p \to q)}{q}$$

$$\frac{p \to \neg q \to q}{p \to q} \qquad \frac{\neg(p \leftrightarrow q)}{\neg p \leftrightarrow q} \qquad \frac{p \leftrightarrow q}{p \lor q \to p \land q}$$

$$\frac{p \to q}{\neg q \to \neg p} \qquad \frac{p \to q \to r}{p \land q \to r} \qquad \frac{p \land q}{\neg(\neg p \lor \neg q)}$$

# 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。

$$\frac{(p \lor q) \land r}{p \lor (q \land r)}\ ? \qquad \frac{p \lor (q \land r)}{(p \lor q) \land r}\ ? \qquad \frac{p \leftrightarrow (q \to r)}{(p \leftrightarrow q) \to r}\ ? \qquad \frac{(p \leftrightarrow q) \to r}{p \leftrightarrow (q \to r)}\ ?$$

$$\frac{\neg(p \to q \land r) \quad r \to p \land q}{\neg r}\ ? \qquad \frac{p \to q \land r \quad \neg(p \lor q \to r)}{p}\ ?$$

$$\frac{p \to r \quad q \to s \quad p \lor q \quad \neg(r \land s)}{(r \to p) \land (s \to q)}\ ? \qquad \frac{\neg p \to (\neg q \to r)}{\neg(\neg p \to q) \to r}\ ?$$

1. If God does not exist, then it's not the case that *if I pray, my prayers will be answered*; and I don't pray; so God exists.
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.

# 练习：树形证明 — Now it's your turn ✍♡✍

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。

$$\frac{\begin{array}{c} p \to (q \to r) \land (\neg q \to s) \\ \neg r \land \neg s \end{array}}{\neg p}$$

# Natural Deduction

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

---

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

---

$$\frac{\begin{array}{c}[A]^n\\ \vdots \\ B\end{array}}{A \to B}\ [\to^+]^n \qquad \frac{A \to B \quad A}{B}\ [\to^-]$$

---

$$\frac{\begin{array}{c}[A]^n\\ \vdots \\ \bot\end{array}}{\neg A}\ [\neg^+]^n \qquad \frac{\neg\neg A}{A}\ [\neg^-] \qquad \frac{\neg A \quad A}{\bot}\ [\bot^+] \qquad \frac{\bot}{A}\ [\bot^-]$$

# Examples

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

Proof.

$$\dfrac{A \vee B \qquad \begin{array}{c}[A]^1\\ \vdots \\ A\end{array} \qquad \dfrac{\dfrac{[B]^1 \qquad \neg B}{\bot}\,[\bot^+]}{A}\,[\bot^-]}{A}\,[\vee^-]^1$$

$$\boxed{\dfrac{A \to B \qquad \neg B}{\neg A}}$$

Proof.

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

# Examples

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

Proof.

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

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

Proof.

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

# Examples

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

Proof.

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

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

Proof.

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

# Examples

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

Proof.

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

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

Proof.

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

# Valid Arguments & Proof Methods

1. **Direct Proof:**

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

   Example: If $n$ is an odd integer, then $n^2$ is odd.

2. **Backward Reasoning:** 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:**

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

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

# Valid Arguments & Proof Methods

4. **Proof by Cases:**

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

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

5. **Proof by Elimination:**

$$\frac{A \rightarrow B \vee C}{\overline{A \wedge \neg B \rightarrow C}}$$

Example: If $a+bi$ and $c+di$ are complex numbers for which $(a+bi)(c+di) = 1$, then $a \neq 0$ or $b \neq 0$.

6. **Proof by Contradiction:**

$$\frac{\begin{array}{c}\neg A \to B \\ \neg A \to \neg B\end{array}}{A}$$

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

7. **Reductio Ad Absurdum (Refutation by Contradiction):**

$$\frac{\begin{array}{c}A \to B \\ A \to \neg B\end{array}}{\neg A}$$

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.

# Contents

Introduction

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

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Independence

## Definition (Independence)

An axiom $A$ in $\Gamma$ 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

Axiom$1$ and Axiom$2$ 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 Axiom$1$ and Axiom$2$.*

| $\to$ | $0$ | $u$ | $1$ |
|---|---|---|---|
| $0$ | $1$ | $1$ | $1$ |
| $u$ | $0$ | $1$ | $1$ |
| $1$ | $0$ | $u$ | $1$ |

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”?

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

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

## What is “theory”?

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

# Consistency & Satisfiability

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

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

- $\Gamma$ is maximal iff for every wff $A$, either $A \in \Gamma$ or $\neg A \in \Gamma$.
- $\Gamma$ is maximal consistent iff it is both consistent and maximal.
- $\Gamma$ is satisfiable iff $\mathrm{Mod}(\Gamma) \neq \emptyset$.
- $\Gamma$ is finitely satisfiable iff every finite subset of $\Gamma$ is satisfiable.

- If $\Gamma$ is consistent and $\Gamma \vdash A$, then $\Gamma \cup \{A\}$ is consistent.
- $\Gamma \cup \{\neg A\}$ is inconsistent iff $\Gamma \vdash A$.
- If $\Gamma$ is maximal consistent, then $A \notin \Gamma \implies \Gamma \cup \{A\}$ is inconsistent.

# 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 wffs is consistent.*

# Completeness Theorem

## Theorem (Completeness Theorem — Post1921)

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

## Corollary

*Any consistent set of wffs is satisfiable.*

$$\begin{array}{ccc} \Gamma \vDash A & \Longleftrightarrow & \Gamma \vdash A \\ \Updownarrow & & \Updownarrow \\ \Gamma \cup \{\neg A\} & & \Gamma \cup \{\neg A\} \\ \text{unsatisfiable} & \Longleftrightarrow & \text{inconsistent} \end{array}$$

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

## Corollary (Compactness Theorem)

*A set of wffs is satisfiable iff it is finitely satisfiable.*

# Proof of Completeness Theorem

## Proof.

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

$$\Delta_0 := \Gamma$$

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

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

step2. Define a truth assignment that satisfies $\Gamma$.

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

# Compactness Theorem

## Theorem (Compactness Theorem)

*A set of wffs is satisfiable iff it is finitely satisfiable.*

## Corollary

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

## Proof.

$$\begin{aligned}
\Gamma_0 \nvDash A \text{ for any } \Gamma_0 \subset \Gamma &\implies \Gamma_0 \cup \{\neg A\} \text{ is satisfiable for any } \Gamma_0 \subset \Gamma \\
&\implies \Gamma \cup \{\neg A\} \text{ is satisfiable} \\
&\implies \Gamma \nvDash A
\end{aligned}$$

**Remark:** $\Gamma$ is consistent iff every finite subset of $\Gamma$ is consistent.
**Remark:** Compactness does not hold in a language with infinite disjunctions.

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

# 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 $\Gamma$ to a maximal finitely satisfiable set $\Delta$.

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

$$\Delta_0 := \Gamma$$

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

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

part2. Define a truth assignment that satisfies $\Gamma$.

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

# Weak Completeness Theorem

## Lemma

*Let $A$ be a wff 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 $\quad \vDash A \implies \vdash A$

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

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

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

# 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 wff.*

## Theorem (Decidability — Post1921)

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

## Theorem

*If $\Gamma$ is a decidable set of wffs, then the set of logical consequences of $\Gamma$ is recursively enumerable.*

## Model Checking & Satisfiability Checking & Validity Checking[5]

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

Figure: *Hamiltonian Circle(NP)*

Figure: *Eulerian Circle(P)*

Figure: *Graph Coloring(NPC)*

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

# Contents

Introduction

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

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Example ♡

## 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. $\neg(x \wedge \neg y)$

## 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。

# 自指

Alice: Bob is a liar.

Bob: Alice tells the truth.

$$A \leftrightarrow \neg B, B \leftrightarrow A \vdash \bot$$

Problem (Which answer is the correct answer to this question?)

1. *All of the below.*
2. *None of the below.*
3. *All of the above.*
4. *One of the above.*
5. *None of the above.*
6. *None of the above.*

1. $p_1 \leftrightarrow p_2 \land p_3 \land p_4 \land p_5 \land p_6$
2. $p_2 \leftrightarrow \neg p_3 \land \neg p_4 \land \neg p_5 \land \neg p_6$
3. $p_3 \leftrightarrow p_1 \land p_2$
4. $p_4 \leftrightarrow (p_1 \land \neg p_2 \land \neg p_3) \lor (\neg p_1 \land p_2 \land \neg p_3) \lor (\neg p_1 \land \neg p_2 \land p_3)$
5. $p_5 \leftrightarrow \neg p_1 \land \neg p_2 \land \neg p_3 \land \neg p_4$
6. $p_6 \leftrightarrow \neg p_1 \land \neg p_2 \land \neg p_3 \land \neg p_4 \land \neg p_5$

$$1 \land 2 \land 3 \land 4 \land 5 \land 6 \vdash \neg p_1 \land \neg p_2 \land \neg p_3 \land \neg p_4 \land p_5 \land \neg p_6$$

# 数独

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |
| | 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 $3 \times 3$ 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 $1 \leq i, j, n, n' \leq 9$ and $n \neq n'$:

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

# Example: The Wumpus World

- 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.[6]

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

# Shannon — Digital Circuit Design

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

# Shannon — Digital Circuit Design

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

# Circuit-Satisfiability Problem

A one-output boolean combinational circuit is satisfiable if it has a truth assignment that causes the output of the circuit to be $1$.

# McCulloch-Pitts Artificial Neural Network

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

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

# Linearly Separable

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

| $p$ | $q$ | $p \oplus q$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

$$w_1 \cdot 0 + w_2 \cdot 0 + b < 0 \qquad b < 0$$

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

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

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

A simple single-layer perception can't solve nonlinearly separable problems.

# 《三体》— 人列计算机

- 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。

用连续信号模拟离散信号会怎样?

# Neural Network

# Finite State Automaton

| $y_1$ | $y_2$ | $x$ | $y_1^+$ | $y_2^+$ |
|---|---|---|---|---|
| 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 |

$$y_1^+ = \bar{y}_1\bar{y}_2 + \bar{x}\bar{y}_2$$
$$y_2^+ = y_1\bar{y}_2 + xy_1$$

# Reversible Computing — Fredkin Gate: CSWAP

| $c$ | $p$ | $q$ | $x$ | $y$ | $z$ |
|---|---|---|---|---|---|
| 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 | 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.

$$f : (c, p, q) \mapsto (c, c \cdot q + \overline{c} \cdot p, c \cdot p + \overline{c} \cdot q)$$

$\neg \; p = 0 \;\&\; q = 1 \implies z = \overline{c}$

$\wedge \; q = 0 \implies z = c \cdot p$

**Remark:** no entropy changes occur and no heat dissipation is involved.

# Exercise ♡

1. Behind one of the door is a path to freedom, behind the other two doors is an evil dragon.
2. At least one of the three statements is true.
3. At least one of the three statements is false.

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× |

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

# Exercise

## 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。
   $(t_1 \wedge \neg t_2 \wedge \neg t_3) \vee (\neg t_1 \wedge t_2 \wedge \neg t_3) \vee (\neg t_1 \wedge \neg t_2 \wedge t_3)$

# Exercise

## 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. $\neg L$

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

# 练习 — Now it's your turn ✍♡✍

## 谁肯定是凶手?

一起凶杀案有三个嫌疑人：小白、大黄和老王。

1. 至少有一人是凶手，但不可能三人同时犯罪。
2. 如果小白是凶手，那么老王是同犯。
3. 如果大黄不是凶手，那么老王也不是。

## What am I doing?

1. If I'm not playing tennis, I'm watching tennis.
2. If I'm not watching tennis, I'm reading about tennis.
3. I cannot do more than one of these activities at a time.

# 练习 — Now it's your turn ✍♡✍

## 哪个部落的?

一个岛上有 T、F 两个部落，T 部落的居民只说真话，F 部落的居民只说谎。你在岛上遇到了小白、大黄、老王三个土著。

1. 小白："如果老王说谎，我或大黄说的就是真话"。
2. 大黄："如果小白或老王说真话，那么，我们三人中有且只有一人说真话是不可能的"。
3. 老王："小白或大黄说谎当且仅当小白或我说真话"。

# 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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。

$$\frac{\text{人都会死} \quad \text{苏格拉底是人}}{\text{苏格拉底会死}}$$

**A**: All $S$ are $P$.

**E**: No $S$ are $P$.

**I**: Some $S$ are $P$.

**O**: Some $S$ are not $P$.

# 三段论

$$\frac{\begin{array}{c} M - P \\ S - M \end{array}}{S - P} \qquad \frac{\begin{array}{c} P - M \\ S - M \end{array}}{S - P} \qquad \frac{\begin{array}{c} M - P \\ M - S \end{array}}{S - P} \qquad \frac{\begin{array}{c} P - M \\ M - S \end{array}}{S - P}$$

- 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

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。

$$\frac{\begin{array}{c}\text{No } P \text{ are } M \\ \text{Some } S \text{ are not } M\end{array}}{\text{Some } S \text{ are } P}$$

# 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事物，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

$P$ 不周延

| | |
|---|---|
| **A**: <u>All</u> $S$ are $P$. | **E**: <u>No</u> $S$ are $P$. |
| **I**: Some $S$ are $P$. | **O**: Some $S$ are <u>not</u> $P$. |

$P$ 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. 否定前提的数量必须与否定结论的数量相等。
4. 特称前提的数量必须与特称结论的数量相等。(存在谬误)

- 亚里士多德要求满足 $1-3$
- 布尔要求满足 $1-4$

# Deduction/Induction/Abduction/Examplification

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

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

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

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

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

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

# 演绎/归纳/溯因

$$\frac{\text{这个袋子里的豆子都是白的} \quad \text{这些豆子是这个袋子里的}}{\text{这些豆子是白的}} \qquad \frac{M \to P \quad S \to M}{S \to P}$$

$$\frac{\text{这些豆子是白的} \quad \text{这些豆子是这个袋子里的}}{\text{这个袋子里的豆子都是白的}} \qquad \frac{M \to P \quad M \to S}{S \to P}$$

$$\frac{\text{这个袋子里的豆子都是白的} \quad \text{这些豆子是白的}}{\text{这些豆子是这个袋子里的}} \qquad \frac{P \to M \quad S \to M}{S \to P}$$

# 溯因

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。

# Example and Criticism

All men are intelligent
Women are not men
---
Women are not intelligent

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

$MEP$
$SAM$
---
$SEP$

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Contents

Introduction

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Gottlob Frege 1848-1925 (+Charles Peirce 1839-1914)

- *Begriffsschrift, a formal language of pure thought modelled upon that of arithmetic.*
- Predicate Logic. (Relations & Quantifiers)
  (Every boy loves some girl.)

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

- Philosophy of Language.
  The evening star is the morning star. (venus)

$\boxed{\text{Logicism}}$ Mathematics $\rightsquigarrow$ Logic.[7]

[7] Frege: The Foundations of Arithmetic.

# Why Study Predicate Logic?

- Propositional logic assumes the world contains facts.
- Predicate logic assumes the world contains
  - Objects: people, houses, numbers, colors, baseball games, wars, ...
  - Relations: red, round, prime, brother of, bigger than, part of, between, fall in love with, ...
  - Functions: father of, best friend, one more than, plus, ...
- Expressive power.

| **Language** | **Ontological Commitment** | **Epistemological Commit** |
|---|---|---|
| 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: an egg or eternal bliss?

An egg! Because nothing is better than eternal bliss, and an egg is better than nothing.

$$e > 0 > b \implies e > b$$
$$\neg\exists x(x > b) \implies 0 \not> b$$

*"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!"

Modeling
Statistics
Operations Research
Differential Equations
Probability
Cryptology
Number Theory
Knot Theory
Computational Math
Game Theory
Calculus
Numerical Analysis
Graph Theory
Topology
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
Add commutativity
ABELIAN GROUPS
FIELDS
Define =, + & · for equi-valence classes of pairs
Add metric
INTEGERS
Add commutativity
Add multiplicative inverse
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.
MODELS
FORMAL SYSTEMS

?
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

# Emergentism ≠ Reductionism

# Contents

Introduction

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 什么是“量词”?

- All students work hard.
- Some students are asleep.
- At least six students are awake.
- Eight out of ten students are good at logic.
- Nobody knows logic better than Donald Trump.
- Most students love logic.
- There are infinitely many prime numbers.
- There are more PCs than there are Macs.

# 语法 ♡

## Language

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

where

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

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

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

$f^n$ is an $n$-place function symbol.
$P^n$ is an $n$-place predicate symbol.
A $0$-place function symbol $f^0$ is called constant.
A $0$-place predicate symbol $P^0$ is called atomic proposition.

# 项 & 公式 ♡

## Definition (Term)

$$t ::= x \mid c \mid f(t, \ldots, t)$$

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

- $\mathrm{Term}$ is freely generated from $\mathrm{Var}$ by $\mathrm{Fun}$.

## Definition (Well-Formed Formula Wff)

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

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

- $\mathrm{Wff}$ is freely generated from atomic formulas by connective and quantifier operators.

# Syntax

- $A \wedge B := \neg(A \to \neg B)$
- $A \vee B := \neg A \to B$
- $A \leftrightarrow B := (A \to B) \wedge (B \to A)$
- $\exists x A := \neg \forall x \neg A$
- $\bot := A \wedge \neg A$
- $\top := \neg \bot$

- Bottom up and Top down definitions of terms, subterms, wffs and subformulas.
- Induction Principle for terms and wffs.
- Unique readability theorem for terms and wffs.
- Omitting Parenthesis.
  1). outermost parentheses. 2). $\neg, \forall, \exists, \wedge, \vee, \to, \leftrightarrow$ 3). group to the right.

# 自由变元 & 约束变元

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

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

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

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

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

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

# 自由变元 & 约束变元 ♡

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

- In the wff $\forall x A$, we say that the scope of this quantifier $\forall$ is $A$.
- An occurrence of a variable $x$ is bound if it occurs within the scope of $\forall x$, or it is the $x$ in $\forall x$.
- An occurrence of a variable is free if it is not bound.
- The term $t$ is a ground term iff $\mathrm{Var}(t) = \emptyset$.
- $A$ is a sentence iff $\mathrm{Fv}(A) = \emptyset$.
- $A$ is an open formula iff $\mathrm{Bv}(A) = \emptyset$.

Example: $c = d$ is clopen.

# Variable of a Term, Free / Bound Variable of a wff

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

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

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

# 翻译 — 学说“逻辑语”

| | | |
|---|---|---|
| **A** | All $S$ are $P$. | $\forall x(Sx \to Px)$ |
| **E** | No $S$ are $P$. | $\forall x(Sx \to \neg Px)$ |
| **I** | Some $S$ are $P$. | $\exists x(Sx \wedge Px)$ |
| **O** | Some $S$ are not $P$. | $\exists x(Sx \wedge \neg Px)$ |

1. The stupid are cocksure while the intelligent are full of doubt. –Russell

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

2. Every boy loves some girl.

$$\forall x\big(Bx \to \exists y(Gy \wedge Lxy)\big)$$

3. Whoever has a father has a mother.

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

4. Grandfather is father's father.

$$\forall x \forall y\big(Gxy \leftrightarrow \exists z(Fxz \wedge Fzy)\big)$$

5. No man loves children unless he has his own.

$$\forall x\Big(Mx \wedge \exists y \exists z(Cyz \wedge Lxy) \to \exists y Cyx\Big)$$

# 翻译 — 学说“逻辑语”

1. 敌人的敌人是朋友。

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

2. 朋友之间要么都吸烟要么都不吸烟。

$$\forall x \forall y \big(Fxy \to (Sx \leftrightarrow Sy)\big)$$

3. 既没有朋友又没有敌人是寂寞的。

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

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. 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。

$$\forall x \forall y (Fx \wedge Fy \wedge Bxy \to Sxy) \to$$
$$\exists x (Fx \wedge \forall y (Fy \to Bxy)) \to \exists x (Fx \wedge \forall y (Fy \to Sxy))$$

6. 名，可名，非常名。

$$\forall x \Big(\exists y \, \mathrm{Name}(x, y) \wedge \exists z \, \mathrm{Name}(z, x) \to \neg \, \mathrm{Unchanging}(x)\Big)?$$

# 翻译 — 学说“逻辑语”

1. If bad things happen to good people, then God is either not omnipotent or not benevolent.

   $$\exists x \exists y (\mathrm{Bad}(x) \land \mathrm{Good}(y) \land \mathrm{Happen}(x, y)) \to \neg\, \mathrm{Omnip}(g) \lor \neg\, \mathrm{Benev}(g)$$

2. Every book that Alice lends to Bob she steals from Chris.

   $$\forall x \big(Bx \land Laxb \to Saxc\big)$$

3. For every professor, there is a student who likes every book the professor recommends to the student.

   $$\forall x \Big(Px \to \exists y \big(Sy \land \forall z (Bz \land Rxzy \to Lyz)\big)\Big)$$

4. Anyone who befriends any enemy of yours is an enemy of mine.

   $$\forall x \forall y \big(Eyu \land Bxy \to Exm\big)$$

5. When a mathematical or philosophical author writes with a misty profundity, he is talking nonsense. — *Alfred Whitehead*

   $$\forall x \Big(\big(\mathrm{Math}(x) \lor \mathrm{Pilo}(x)\big) \land \mathrm{Write}(x) \to \mathrm{TalkNonsense}(x)\Big)$$

# 翻译 — 学说“逻辑语”

1. If all dancers have knee injuries, then *they* will be frustrated.

$$\forall x\Big(\mathrm{Dancer}(x) \to \mathrm{Knee}(x)\Big) \to \forall x\Big(\mathrm{Dancer}(x) \to \mathrm{Frustrated}(x)\Big)$$

2. If all dancers have knee injuries, then *some of them* will be frustrated.

$$\forall x\Big(\mathrm{Dancer}(x) \to \mathrm{Knee}(x)\Big) \to \exists x\Big(\mathrm{Dancer}(x) \land \mathrm{Frustrated}(x)\Big)$$

3. No dolphin sings unless it jumps.

$$\forall x\Big(\mathrm{Dolphin}(x) \to \neg\,\mathrm{Jumps}(x) \to \neg\,\mathrm{Sings}(x)\Big)$$

4. Alice is the first to have completed the test.

$$Cat \land \forall x(Cxt \land x \neq a \to Baxt)$$

Alice is the oldest daughter of the King.

5. Alice is the second to have completed the test.

$$Cat \land \exists x(Cxt \land x \neq a \land Bxat \land \forall y(Cyt \land y \neq a \land y \neq x \to Bayt))$$

# Example — 究竟有几个神?

有神论 there is at least one god

$$\exists x Gx$$

无神论 there are no gods

$$\neg\exists x Gx$$

泛神论 everything is god

$$\forall x Gx$$

不可知论 there is at most one god

$$\forall x(Gx \to \forall y(Gy \to y = x))$$

一神论 there is exactly one god

$$\exists x(Gx \wedge \forall y(Gy \to y = x))$$

二元神论 there are exactly two gods

$$\exists x(Gx \wedge \exists y(Gy \wedge x \neq y) \wedge \forall z(Gz \to z = x \vee z = y))$$

1. There are at most $1$ elements.

$$\forall x \forall y (x = y)$$

2. There are at least $2$ elements.

$$\exists x \exists y (x \neq y)$$

3. There are exactly $2$ elements.

$$\exists x \exists y \big(x \neq y \wedge \forall z (z = x \vee z = y)\big)$$

4. There are exactly $n$ elements.

$$\exists x_1 \ldots \exists x_n \left( \bigwedge_{1 \leq i < j \leq n} x_i \neq x_j \wedge \forall y \left( \bigvee_{i=1}^{n} y = x_i \right) \right)$$

5. There exists a unique element such that $P$.

$$\exists x Px \wedge \forall y \forall z (Py \wedge Pz \rightarrow y = z)$$

$$\exists x \big(Px \wedge \forall y (Py \rightarrow y = x)\big)$$

$$\exists x \forall y (Py \leftrightarrow y = x)$$

$$\exists! x Px$$

# 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 —"

$$\exists x \forall y ((\mathrm{Song}(x) \leftrightarrow y = x) \wedge \mathrm{Long}(x) \wedge \mathrm{Beautiful}(x))$$

- "Or else what?" said Alice.
- "Or else it doesn't, you know. The name of the song is called 'Haddocks' Eyes.' "

$$\exists! x (\mathrm{Song}(x) \wedge \forall y (\mathrm{Hear}(y, x) \rightarrow \mathrm{Tear}(x, y) \vee \neg \mathrm{Tear}(x, y)))$$

$$\text{call-name}(\mathrm{name}(\text{the-song})) = \text{Haddocks' Eyes}$$

- "Oh, that's the name of the song, is it?" said Alice.

$$\mathrm{name}(\text{the-song}) = \text{Haddocks' Eyes}$$

- "No, you don't understand," the Knight said. "That's what the name is called. The name really is 'The Aged Aged Man.' "

$$\mathrm{name}(\text{the-song}) = \text{The Aged Aged Man}$$

- "Then I ought to have said 'That's what the song is called'?" Alice corrected herself.

$$\mathrm{call\text{-}name}(\text{the-song}) = \text{The Aged Aged Man}$$

- "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!"

$$\mathrm{call\text{-}name}(\text{the-song}) = \text{Ways And Means}$$

- "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'."

$$\text{the-song} = \text{A-sitting On A Gate}$$

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*

$$\forall x(\mathrm{Learn}(x) \land \neg\,\mathrm{Think}(x) \to \mathrm{Lost}(x)) \land$$
$$\forall x(\mathrm{Think}(x) \land \neg\,\mathrm{Learn}(x) \to \mathrm{InDanger}(x))$$

2. He who can, does. He who cannot, teaches. — *Bernard Shaw*

$$\forall xy(\mathrm{Can}(x,y) \to \mathrm{Does}(x,y)) \land \forall xy(\neg\,\mathrm{Can}(x,y) \to \mathrm{Teaches}(x,y))$$

3. Don't interrupt me, while I'm interrupting. — *Churchill*

$$\exists x\,\mathrm{Interrupt}(i,x) \to \forall y \neg\,\mathrm{Interrupt}(y,i)$$

4. There are no shortcuts to any place worth going. — *Beverly Sills*

$$\forall x(\mathrm{Place}(x) \land \mathrm{WorthGo}(x) \to \neg\exists y\,\mathrm{Shortcut}(y,x))$$

5. The old believe everything; the middle-aged suspect everything; the young know everything. — *Oscar Wilde*

$$\forall x(\mathrm{Old}(x) \to \forall y\,\mathrm{Believe}(x,y)) \land$$
$$\forall x(\mathrm{Middle}(x) \to \forall y\,\mathrm{Suspect}(x,y)) \land \forall x(\mathrm{Young}(x) \to \forall y\,\mathrm{Know}(x,y))$$

6. No married man is ever attractive except to his wife. — *Oscar Wilde*

$$\forall x(\mathrm{Man}(x) \land \mathrm{Married}(x) \to \forall y(\mathrm{Attractive}(x,y) \to y = \mathrm{wife}(x)))$$

# 翻译 — 学说“逻辑语”

7. Always two there are: a Master and an Apprentice. — *Yoda*

$$\exists xy(x \neq y \land \forall z(z = x \lor z = y) \land \mathrm{Master}(x) \land \mathrm{Apprentice}(y))$$

8. 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*

$$\exists xy(x \neq y \land \forall z(\mathrm{WayLive}(z) \leftrightarrow z = x \lor z = y) \land$$
$$(\mathrm{WayLive}(x) \to \neg\exists z\, \mathrm{Miracle}(z)) \land (\mathrm{WayLive}(y) \to \forall z\, \mathrm{Miracle}(z)))$$

9. The weakest link in a chain is the strongest because it can break it. — *Stanislaw J. Lec*

$$\forall xy((\mathrm{Link}(x) \land \mathrm{Chain}(y) \land \mathrm{In}(x, y) \land \mathrm{Weakest}(x) \to \mathrm{Break}(x, y))$$
$$\to \mathrm{Strongest}(x))$$

10. To a man who only has a hammer, everything looks like a nail. — *Maslow*

$$\forall x(\mathrm{Man}(x) \land \exists! y(\mathrm{Hammer}(y) \land \mathrm{Has}(x, y)) \to$$
$$\forall z \exists n(\mathrm{Nail}(n) \land \mathrm{LooksLike}(z, n, x)))$$

11. He who refuses to do arithmetic is doomed to talk nonsense. — *John McCarthy*

$$\forall x(\mathrm{Refuse}(x, \text{arithmetic}) \to \Box\, \mathrm{TalkNonsense}(x))$$

1. $\mathrm{Think}(i) \to \exists x(x = i)$ *Descartes*
2. $\exists x(x = i) \vee \neg\exists x(x = i)$ *Shakespeare*
3. $\forall x\big(\mathrm{Month}(x) \to \mathrm{Crueler}(\mathsf{april}, x)\big)$ *Eliot*
4. $\forall x(\neg\,\mathrm{Weep}(x) \to \neg\,\mathrm{See}(x))$ *Hugo*
5. $\forall x\big(\mathrm{Time}(x) \to \mathrm{Better}(t, x)\big) \wedge \forall x\big(\mathrm{Time}(x) \to \mathrm{Better}(x, t)\big)$ *Dickens*
6. $\exists x\Big(\mathrm{Child}(x) \wedge \neg\,\mathrm{Growup}(x) \wedge \forall y\big(\mathrm{Child}(y) \wedge y \neq x \to \mathrm{Growup}(y)\big)\Big)$ *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. $\forall x\big(\mathrm{Problem}(x) \wedge \mathrm{Philo}(x) \wedge \mathrm{Serious}(x) \leftrightarrow x = \mathsf{suicide}\big)$ *Camus*
10. $\forall x\big(\mathrm{Feather}(x) \wedge \mathrm{Perch}(x, \mathsf{soul}) \leftrightarrow x = \mathsf{hope}\big)$ *Dickinson*
11. $\forall x\Big(\mathrm{Enter}(x) \to \forall y\big(\mathrm{Hope}(y) \to \mathrm{Abandon}(x, y)\big)\Big)$ *Dante*
12. $\exists x\forall y\big(\mathrm{For}(y, x) \wedge \mathrm{For}(x, y)\big)$? $\forall x\forall y\big(\mathrm{For}(y, x) \leftrightarrow \mathrm{For}(x, y)\big)$? *Dumas*
13. $\exists x\big(\mathrm{Fear}(\mathsf{we}, x) \leftrightarrow x = \mathsf{fear}\big)$? *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*

# 翻译 — “逻辑语言”到“自然语言”

1. $\forall x\big(\mathrm{Buy}(\mathsf{susan}, x) \to \mathrm{Buy}(\mathsf{alice}, x)\big)$
2. $\forall x\big(\exists y\, \mathrm{Buy}(y, x) \lor \exists z\, \mathrm{Sell}(z, x) \to \neg\, \mathrm{Human}(x)\big)$
3. $\forall x\Big(\big(\mathrm{Petted}(x) \to \mathrm{Jump}(x)\big) \to \mathrm{Dog}(x) \lor \mathrm{Dolphin}(x)\Big)$
4. $\forall x\Big(\mathrm{Girl}(x) \land \mathrm{Love}(\mathsf{bob}, x) \to \forall y\big(\mathrm{PhilosophyBook}(y) \to \mathrm{Read}(x, y)\big)\Big)$

- $\forall x\Big(\mathrm{Minute}(x) \to \exists y\big(\mathrm{Driver}(y) \land \mathrm{Die}(y, x)\big)\Big)$
- $\exists x\Big(\mathrm{Driver}(x) \land \forall y\big(\mathrm{Minite}(y) \to \mathrm{Die}(x, y)\big)\Big)$
- Every minute one driver dies in car accident.
- One driver dies in car accident every minute.

# 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. $\forall x \forall y\Big(\big(Gx \wedge \forall y(By \to \neg Lxy)\big) \wedge \big(Gy \wedge \exists x(Bx \wedge Lyx)\big) \to \neg Lxy\Big)$

# 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. $\forall x \forall y\Big(\big(Gx \wedge \forall y(By \to \neg Lxy)\big) \wedge \big(Gy \wedge \exists x(Bx \wedge Lyx)\big) \to \neg Lxy\Big)$

# 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:

$$\forall \varepsilon > 0 \exists \delta > 0 \big(0 < |x - c| < \delta \to |f(x) - L| < \varepsilon\big)$$

- Divergence

$$\lim_{x\to c} f(x) \uparrow \iff \forall L \in \mathbb{R} \exists \varepsilon > 0 \forall \delta > 0 \big(0 < |x - c| < \delta \wedge |f(x) - L| \geq \varepsilon\big)$$

# 翻译 — “自然语言”到“逻辑语言”

- Continuity

$$\forall x \in I \forall \varepsilon > 0 \exists \delta > 0 \forall y \in I \big( |x - y| < \delta \to |f(x) - f(y)| < \varepsilon \big)$$

$$\forall (x_n)_{n \in \mathbb{N}} \subset I : \lim_{n \to \infty} f(x_n) = f(\lim_{n \to \infty} x_n)$$

- Uniform Continuity

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x y \in I \big( |x - y| < \delta \to |f(x) - f(y)| < \varepsilon \big)$$

- Any two distinct points determine a unique line.

$$\forall x \forall y \Big( \mathrm{Point}(x) \wedge \mathrm{Point}(y) \wedge x \neq y \to \exists! z \, \mathrm{Line}(z) \wedge \mathrm{On}(x, z) \wedge \mathrm{On}(y, z) \Big)$$

# 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. *None but* ripe bananas are edible.
5. *Only* Socrates and Plato are human.
6. *All but* Socrates and Plato are human.
7. Every boy loves *at least* two girls.
8. Adams can't do *every* job right.
9. Adams can't do *any* job right.
10. *Not all* that glitters are gold.
11. Everyone's zipcode within a city has the same first digit.
12. Every farmer who owns a donkey is happy.
13. Every farmer who owns a donkey beats it.
14. All even numbers are divisible by $2$, but *only some* are divisible by $4$.

# 练习：翻译 — Now it's your turn ✍♡✍

15. Behind every great fortune there is a crime. — Balzac
16. If a clown enters the room, then it will be displeased if no person is surprised.
17. Everyone alive $2000BC$ is either an ancestor of nobody alive today or of everyone alive today.
18. John hates all people who do not hate themselves.
19. No barber shaves exactly those who do not shave themselves.
20. Andy and Bob have the same maternal grandmother. $\mathrm{mother}(x, y)$
21. Anyone who loves *two* different girls is Tony.
22. Socrates' wife *has* a face that *only* her mother could love.
23. If dogs are animals, every head of a dog is the head of an animal.
24. *The* detective is *never* happy.
25. Someone *other than the girl* who loves Bob is stupid.
26. Morris only loves *the girl* who loves him.
27. *The one* who loves Alice is *the one* she loves.
28. *The shortest* English speaker loves *the tallest* English speaker.

# 代入

## 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)

$$A[t/x] := \begin{cases} t_1[t/x] = t_2[t/x] & \text{if } A = t_1 = t_2 \\ P(t_1[t/x], \ldots, t_n[t/x]) & \text{if } A = P(t_1, \ldots, t_n) \\ \neg B[t/x] & \text{if } A = \neg B \\ B[t/x] \to C[t/x] & \text{if } A = B \to C \\ \begin{cases} \forall y B[t/x] & \text{if } y \neq x \\ \forall y B & \text{if } y = x \end{cases} & \text{if } A = \forall y B \end{cases}$$

# 可代入[8]

## Definition (Substitutable)

A term $t$ *is substitutable for* $x$ *in* $A$ iff none of the free occurrences of $x$ in $A$ occur in the scope of a quantifier that binds a variable in $t$.

不可代入：$t$ 中的变元代入 $A$ 后变得被约束！

**Counter-example:** $A = \exists y(x \neq y) \quad t = y \quad A[t/x]?$

[8] 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

Introduction

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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* $\forall x\big(T(x) \leftrightarrow A(x)\big)$
2. *materially adequate* $A(s) \leftrightarrow p$
   where ‘$s$’ is the name of a sentence of $\mathscr{L}$, and ‘$p$’ is the translation of this sentence in $\mathscr{L}'$.

# Alfred Tarski 1901-1983

"snow is white" is true iff snow is white.

"I am false."[9]

Model Theory

## 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

[9] Tarski: On the Concept of Truth in Formalized Languages.
Tarski: The Semantic Conception of Truth and the Foundations of Semantics.

# 语法 vs 语义

$in(kim,r123).$

$part_of(r123,cs_building).$

$in(X,Y) \leftarrow$

$\quad part_of(Z,Y) \wedge$

$\quad in(X,Z).$

$in(kim,cs_building)$

无名，天地之始；有名，万物之母。

# 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 | sense of the predicate | thought |
| ↓ | ↓ | ↓ |
| reference/object | concept | truth-value |
| | ↓ | |
| | objects that fall under the concept | |

Table: Frege: 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

A structure over the signature is a pair $\mathcal{M} := (M, [\![\,]\!])$, where $M$ is a non-empty set, and $[\![\,]\!]$ is a mapping which

- assigns to each constant symbol $c$ an element $[\![c]\!] \in M$,
- assigns to each function symbol $f$ an $n$-ary function $[\![f]\!] : M^n \to M$,
- assigns to each predicate symbol $P$ an $n$-ary relation $[\![P]\!] \subset M^n$.

We write $\mathcal{M} = (M, [\![c]\!], [\![f]\!], [\![P]\!])$ or $\left(M, c^{\mathcal{M}}, f^{\mathcal{M}}, P^{\mathcal{M}}\right)$ for convenience.

# Example

# 项的解释

## Definition (Interpretation of Terms)

An interpretation $(\mathcal{M}, \nu)$ is a structure $\mathcal{M}$ with a variable assignment $\nu : \mathrm{Var} \to M$.

We extend $\nu$ to $\overline{\nu} : \mathrm{Term} \to M$ by recursion as follows:

- $\overline{\nu}(x) := \nu(x)$
- $\overline{\nu}(c) := [\![c]\!]$
- $\overline{\nu}(f(t_1, \ldots, t_n)) := [\![f]\!](\overline{\nu}(t_1), \ldots, \overline{\nu}(t_n))$

$$\begin{tikzcd}
\mathrm{Term} \arrow[r, "\overline{\nu}"] \arrow[d, "f"'] & M \arrow[d, "{[\![f]\!]}"] \\
\mathrm{Term} \arrow[r, "\overline{\nu}"'] & M
\end{tikzcd}$$

# 公式的解释

Definition (Interpretation of Formulas $\mathcal{M}, \nu \vDash A$)

- $\mathcal{M}, \nu \vDash t_1 = t_2$ iff $\overline{\nu}(t_1) = \overline{\nu}(t_2)$
- $\mathcal{M}, \nu \vDash P(t_1, \ldots, t_n)$ iff $(\overline{\nu}(t_1), \ldots, \overline{\nu}(t_n)) \in [\![P]\!]$
- $\mathcal{M}, \nu \vDash \neg A$ iff $\mathcal{M}, \nu \nvDash A$
- $\mathcal{M}, \nu \vDash A \to B$ iff $\mathcal{M}, \nu \nvDash A$ or $\mathcal{M}, \nu \vDash B$
- $\mathcal{M}, \nu \vDash \forall x A$ iff for every $a \in M : \mathcal{M}, \nu(a/x) \vDash A$ where

$$\nu(a/x)(y) := \begin{cases} \nu(y) & \text{if } y \neq x \\ a & \text{otherwise} \end{cases}$$

or, $\mathcal{M}, \nu \vDash \forall x A$ iff for all $\nu' \sim_x \nu : \mathcal{M}, \nu' \vDash A$.
where $\nu' \sim_x \nu$ iff for all $y \neq x : \nu'(y) = \nu(y)$.

# 塔斯基的“真”之定义

- $\mathcal{M}, v \vDash \Gamma$ iff for all $A \in \Gamma : \mathcal{M}, v \vDash A$.
- $\Gamma$ is satisfiable iff there exists $\mathcal{M}, v$ s.t. $\mathcal{M}, v \vDash \Gamma$. (Satisfiable)
- $\mathcal{M} \vDash A$ iff for every variable assignment $v : \mathcal{M}, v \vDash A$. (True)
- $\mathcal{M} \vDash \Gamma$ iff for all $A \in \Gamma : \mathcal{M} \vDash A$.
- A structure $\mathcal{M}$ is called a model of $\Gamma$ iff $\mathcal{M} \vDash \Gamma$.
- $\Gamma \vDash A$ iff for all $\mathcal{M}, v : \mathcal{M}, v \vDash \Gamma \implies \mathcal{M}, v \vDash A$. (Entailment)[10]
- $A$ is valid $\vDash A$ iff $\emptyset \vDash A$. (Valid)

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

[10] If we define $\Gamma \vDash^* A$ iff for all $\mathcal{M} : \mathcal{M} \vDash \Gamma \implies \mathcal{M} \vDash A$, then we have $Px \vDash^* \forall x Px$.

# Tarski's Definition of Truth — another version

An interpretation $(\mathcal{M}, \nu)$ is a structure $\mathcal{M}$ with $\nu : \mathrm{Var} \to M$.

We extend $\nu$ to $[\![\,]\!]_\nu$ by recursion as follows:

## Interpretation of Terms

- $[\![x]\!]_\nu := \nu(x)$
- $[\![c]\!]_\nu := [\![c]\!]$
- $[\![f(t_1, \ldots, t_n)]\!]_\nu := [\![f]\!]\big([\![t_1]\!]_\nu, \ldots, [\![t_n]\!]_\nu\big)$

**Remark:** Whenever there is no confusion we omit explicit mention of $\nu$.

## Interpretation of Formulas

- $[\![t_1 = t_2]\!] := \begin{cases} 1 & \text{if } [\![t_1]\!] = [\![t_2]\!] \\ 0 & \text{otherwise} \end{cases}$
- $[\![P(t_1, \ldots, t_n)]\!] := \begin{cases} 1 & \text{if } ([\![t_1]\!], \ldots, [\![t_n]\!]) \in [\![P]\!] \\ 0 & \text{otherwise} \end{cases}$
- $[\![\neg A]\!] := 1 - [\![A]\!]$
- $[\![A \wedge B]\!] := \min\{[\![A]\!], [\![B]\!]\}$
- $[\![A \vee B]\!] := \max\{[\![A]\!], [\![B]\!]\}$
- $[\![A \rightarrow B]\!] := \max\{1 - [\![A]\!], [\![B]\!]\}$
- $[\![A \leftrightarrow B]\!] := 1 - \left|[\![A]\!] - [\![B]\!]\right|$
- $[\![\forall x A]\!]_\nu := \min_{a \in M} \{[\![A]\!]_{\nu(a/x)}\}$
- $[\![\exists x A]\!]_\nu := \max_{a \in M} \{[\![A]\!]_{\nu(a/x)}\}$

**Remark:** Whenever there is no confusion we omit explicit mention of $\nu$.

$$\mathcal{M}, \nu \vDash A \iff [\![A]\!]_\nu = 1$$

# Tarski's Definition of Truth — set-based version

Let $[\![\,]\!]$ map atomic formulas to variable assignments $\mathrm{P}(M^{\mathrm{Var}})$.

- $[\![t_1 = t_2]\!] \coloneqq \{\nu : [\![t_1]\!]_\nu = [\![t_2]\!]_\nu\}$
- $[\![P(t_1, \ldots, t_k)]\!] \coloneqq \{\nu : ([\![t_1]\!]_\nu, \ldots, [\![t_n]\!]_\nu) \in [\![P]\!]\}$

We extend $[\![\,]\!]$ to all wffs by recursion as follows:

1. $[\![\neg A]\!] \coloneqq M^{\mathrm{Var}} \setminus [\![A]\!]$
2. $[\![A \wedge B]\!] \coloneqq [\![A]\!] \cap [\![B]\!]$
3. $[\![A \vee B]\!] \coloneqq [\![A]\!] \cup [\![B]\!]$
4. $[\![A \to B]\!] \coloneqq (M^{\mathrm{Var}} \setminus [\![A]\!]) \cup [\![B]\!]$
5. $[\![\forall x A]\!] \coloneqq \bigcap_{a \in M} \{\nu : \nu(a/x) \in [\![A]\!]\}$
6. $[\![\exists x A]\!] \coloneqq \bigcup_{a \in M} \{\nu : \nu(a/x) \in [\![A]\!]\}$

$$\mathcal{M}, \nu \vDash A \iff \nu \in [\![A]\!]$$

Natural Language

$\Downarrow$ represents

$\boxed{\text{Formal Language (Syntax)}}$

$\Downarrow$ expresses

$\boxed{\text{Theory (calculus } \vdash)}$

interprets $\Downarrow$ $\Uparrow$ characterizes

$\boxed{\text{Models (semantics } \vDash)}$

$\Uparrow$ represents

. . . . . semantic gap

Real World

# Example — 模型

1. $\text{Brother}(\text{Richard}, \text{John})$
2. $\neg\,\text{King}(\text{Richard}) \rightarrow \text{King}(\text{John})$
3. $\text{King}(\text{John}) \wedge \exists x\big(\text{Crown}(x) \wedge \text{OnHead}(x, \text{John})\big)$
4. $\neg\,\text{Brother}\big(\text{leftLeg}(\text{Richard}), \text{John}\big)$
5. $\forall x\big(\text{King}(x) \rightarrow \text{Person}(x)\big)$
6. $\text{leftLeg}(\text{Richard}) = \text{leftLeg}(\text{John})$

# Example — 模型

- $M = \{a_1, a_2, a_3\}$
- $[\![c]\!] = a_3$
- $[\![P]\!] = \{a_1, a_2\}$
- $[\![R]\!] = \{(a_1, a_2), (a_1, a_3), (a_3, a_2), (a_3, a_3)\}$

- $[\![c]\!]$: green point
- $[\![P]\!]$: red circles
- $[\![R]\!]$: arrows

1. $\mathcal{M} \overset{?}{\models} Pc$
2. $\mathcal{M} \overset{?}{\models} Pc \vee Rcc$
3. $\mathcal{M} \overset{?}{\models} \forall x(Px \vee Rxx)$
4. $\mathcal{M} \overset{?}{\models} \exists x \forall y(y = x \vee Rxy)$
5. $\mathcal{M}, v \overset{?}{\models} Rxy \to Rcy$
   where $v(x) = a_1, v(y) = a_3$.

# Example — 模型

## Transitive Relation

$$\forall xyz(Rxy \wedge Ryz \rightarrow Rxz)$$

What arrows are missing to make the following a model?

(Add only those arrows that are really needed.)

# Example — 模型

## Euclidean Relation

$$\forall xyz(Rxy \wedge Rxz \rightarrow Ryz)$$

What arrows are missing to make the following a model?

(Add only those arrows that are really needed.)

# 反模型 ♡

若 $A \nvDash B$，则存在反模型$\mathcal{M}$ 使得 $\mathcal{M} \vDash A$ 但 $\mathcal{M} \nvDash B$。

## 构造反模型

$$\frac{\forall x \exists y Rxy}{\exists x \forall y Rxy} \times$$

$$\frac{\exists y \forall x Rxy}{\forall y \exists x Rxy} \times$$

$$\frac{\forall x \exists y Rxy}{\exists y \forall x Rxy} \times$$

# 反模型 ♡

有些动物是猫
有些猫是黄色的

---

有些动物是黄色的

$$\exists x(Ax \wedge Cx)$$
$$\exists x(Cx \wedge Yx)$$

---

$$\exists x(Ax \wedge Yx)$$

$M =$ ,

,

$[\![A]\!] =$

,

$[\![C]\!] =$ ,


$[\![Y]\!] =$

## Problem (Is there a finite counter-model?)

$$\frac{\begin{array}{c}\textit{Everybody loves somebody}\\ \textit{Everybody loves all persons who are loved by his loved ones}\end{array}}{\textit{There is at least a pair of persons who love each other}}$$

## Exercise (构造反模型)

$$\frac{\begin{array}{c}\forall x \exists y Rxy\\ \forall xyz(Rxy \wedge Ryz \to Rxz)\end{array}}{\exists xy(Rxy \wedge Ryx)} \times \qquad \frac{\begin{array}{c}\forall x \exists y Rxy\\ \forall xyz(Rxy \wedge Ryz \to Rxz)\end{array}}{\exists x Rxx} \times$$

$$\cdots \to -2 \to -1 \to 0 \to 1 \to 2 \to 3 \to 4 \to 5 \to \cdots$$

$$(\mathbb{Z}, <)$$

# 爱丽丝梦游仙境 — Lewis Carroll

$$\frac{\forall x(A \to B)}{\forall x(B \to A)} \times$$

- The March Hare: You should say what you mean.
- Alice: I mean what I say — that's the same thing.
- The Mad Hatter: Not the same thing a bit! You might just as well say that "I see what I eat" is the same thing as "I eat what I see"!
- The March Hare: You might just as well say that "I like what I get" is the same thing as "I get what I like"!

# 易犯的“错误”♡

$$\forall x(Bx \to Sx)$$

$$\exists x(Bx \wedge Sx)$$

- $\forall x(Bx \wedge Sx)$
  Everyone is a boy and everyone is smart.
- $\exists x(Bx \to Sx)$
  It is true if there is anyone who is not a boy.

## Is the following sentence true?

All the elephants in this room are purple.

$$\forall x(\mathrm{Elephant}(x) \wedge \mathrm{In}(x, \text{room}) \to \mathrm{Purple}(x))$$

$$\forall x(\mathrm{Elephant}(x) \wedge \mathrm{In}(x, \text{room}) \wedge \mathrm{Purple}(x))$$

# Coincidence Lemma

## Lemma (Coincidence Lemma)

*Assume* $\nu_1, \nu_2 : \mathrm{Var} \to M$*, and for all* $x \in \mathrm{Fv}(A) : \nu_1(x) = \nu_2(x)$*. Then*

$$\mathcal{M}, \nu_1 \vDash A \iff \mathcal{M}, \nu_2 \vDash A$$

- If $A$ is a sentence, then either $\mathcal{M} \vDash A$ or $\mathcal{M} \vDash \neg A$.
- $\mathcal{M} \vDash A \implies \mathcal{M} \vDash \forall x A$
- **Notation**: If $\mathrm{Fv}(A) \subset \{x_1, \ldots, x_n\}$, then we write $\mathcal{M} \vDash A[a_1, \ldots, a_n]$ to mean $\mathcal{M}, \nu \vDash A$ for some (equivalently any) assignment $\nu$ s.t. $\nu(x_i) = a_i$ for $1 \leq i \leq n$.

# 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 wff $A$, then*

$$\mathcal{M}, \nu \vDash A[t/x] \iff \mathcal{M}, \nu(\nu(t)/x) \vDash A$$

$$\begin{tikzcd}
s \arrow[rr, "{[t/x]}"] \arrow[dr, "\overline{\nu(\overline{\nu}(t)/x)}"'] & & s[t/x] \arrow[dl, "\overline{\nu}"] \\
& M &
\end{tikzcd}$$

$$\mathscr{L}_M := \mathscr{L} \cup C_M \text{ where } C_M := \{c_a : a \in M\}$$

$$\mathcal{M}, \nu \vDash A[c_a/x] \iff \mathcal{M}, \nu(a/x) \vDash A$$

We abbreviate $\mathcal{M}, \nu \vDash A[c_a/x]$ by $\mathcal{M}, \nu \vDash A[a]$.

$$\mathcal{M}, \nu \vDash \forall x A \iff \text{for every } a \in M : \mathcal{M}, \nu \vDash A[a]$$

# Equivalent Replacement

## Lemma

*Suppose* $B \in \mathrm{Sub}(A)$, *and* $A[C /\!\!/ B]$ *arises from* $A$ *by replacing one or more occurrences of* $B$ *by* $C$. *Then*

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

# 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* $\mathrm{Var}(t) \cup \mathrm{Var}(s) \subset \{x_1, \ldots, x_n\}$, *and* $A^*$ *arises from the wff* $A$ *by replacing one occurrence of* $t$ *in* $A$ *by* $s$. *Then*

$$\vDash \forall x_1 \ldots x_n (t = s) \to (A \leftrightarrow A^*)$$

$$\mathcal{M} \vDash t = s \implies \mathcal{M} \vDash A \leftrightarrow A^*$$

## Lemma

*Suppose* $\mathrm{Fv}(B) \cup \mathrm{Fv}(C) \subset \{x_1, \ldots, x_n\}$, *and* $A^*$ *arises from the wff* $A$ *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:**

- $\vDash \forall x\big(Px \leftrightarrow Qx\big) \to \big(\forall x Px \leftrightarrow \forall x Qx\big)$
  $\nvDash \big(Px \leftrightarrow Qx\big) \to \big(\forall x Px \leftrightarrow \forall x Qx\big)$
- $\mathcal{M}, v \vDash t = s \not\Longrightarrow \mathcal{M}, v \vDash A \leftrightarrow A^*$
- $\mathcal{M}, v \vDash B \leftrightarrow C \not\Longrightarrow \mathcal{M}, v \vDash A \leftrightarrow A^*$

$$B = Px, \quad C = Py, \quad A = \forall x Px, \quad A^* = \forall x Py$$

# How to Check Validity? — Example

$$\vDash \exists x \forall y Rxy \to \forall y \exists x Rxy$$

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, $\mathcal{M}, \nu \vDash A[t/x]$.

# Contents

Introduction

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Formal Systems

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

# Hilbert System = Axiom + Inference Rule

## Axiom Schema

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.

## Inference Rule

$$\frac{A \quad A \to B}{B}\ [\mathrm{MP}]$$

# 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 $\exists$ |

# Deduction Theorem

## Theorem (Deduction Theorem1)

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

## Inference Rule

$$\frac{A}{\forall x A} \text{ [G]}$$

What if we remove Axiom$9$ and add the rule of generalization [G] 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 $\Gamma \vdash A \to B$.*

# Meta-properties

- Tautology. For $A \in \mathscr{L}^0$, $B_1, \ldots, B_n \in \mathscr{L}^1$:
$$\frac{A}{A[B_1/p_1, \ldots, B_n/p_n]}$$
- reductio ad absurdum / proof by contradiction
$$\frac{\Gamma, A \vdash B \qquad \Gamma, A \vdash \neg B}{\Gamma \vdash \neg A} \qquad \frac{\Gamma, \neg A \vdash B \qquad \Gamma, \neg A \vdash \neg B}{\Gamma \vdash A}$$
- contraposition
$$\frac{\Gamma, A \vdash \neg B}{\Gamma, B \vdash \neg A}$$
- substitution
$$\frac{t = s}{r[t/x] = r[s/x]} \qquad \frac{t = s}{A[t/x] \leftrightarrow A[s/x]}$$
- equivalent replacement
$$\frac{B \leftrightarrow C}{A \leftrightarrow A[C/\!\!/B]}$$
where $A[C/\!\!/B]$ arises from $A$ by replacing one or more occurrences of $B$ in $A$ by $C$.

# Meta-properties

$$\frac{\Gamma, A[t/x] \vdash B}{\Gamma, \forall x A \vdash B}\ \forall L \qquad \frac{\Gamma \vdash A[t/x]}{\Gamma \vdash \exists x A}\ \exists R$$

$$\frac{\Gamma, A \vdash B \qquad x \notin \mathrm{Fv}(\Gamma, B)}{\Gamma, \exists x A \vdash B}\ \exists L \qquad \frac{\Gamma \vdash A \qquad x \notin \mathrm{Fv}(\Gamma)}{\Gamma \vdash \forall x A}\ \forall R$$

$$\frac{\Gamma, A[y/x] \vdash B \quad y \notin \mathrm{Fv}(\Gamma, \exists x A, B)}{\Gamma, \exists x A \vdash B}\ \exists L \qquad \frac{\Gamma \vdash A[y/x] \quad y \notin \mathrm{Fv}(\Gamma, \forall x A)}{\Gamma \vdash \forall x A}\ \forall R$$

$$\frac{\Gamma, A[a/x] \vdash B \quad a \notin \mathrm{Cst}(\Gamma, \exists x A, B)}{\Gamma, \exists x A \vdash B}\ \exists L \qquad \frac{\Gamma \vdash A[a/x] \quad a \notin \mathrm{Cst}(\Gamma, \forall x A)}{\Gamma \vdash \forall x A}\ \forall R$$

$$\frac{\forall x A}{\forall y A[y/x]}\ \text{alphabetic variant}$$

## 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. $(\neg \rightarrow)$ $\quad \Gamma \vdash \neg(A \rightarrow B) \iff \Gamma \vdash A \ \& \ \Gamma \vdash \neg B$
2. $(\neg\neg)$ $\quad \Gamma \vdash \neg\neg A \iff \Gamma \vdash A$
3. $(\neg\forall)$ $\quad \Gamma \vdash \neg\forall x A \iff \Gamma \vdash \neg A[t/x]$
   Unfortunately this is not always possible. Try contraposition, reductio ad absurdum or prove by contradiction...

# 构造性 vs 存在性

$$\frac{\Gamma \vdash A[t/x]}{\Gamma \vdash \exists x A} \qquad \frac{\Gamma, \forall x \neg A \vdash \bot}{\Gamma \vdash \exists x A}$$

Example (There exist two irrational numbers $x, y$ s.t. $x^y$ is rational.)

$$x := \sqrt{2}$$
$$y := \log_2 9$$

Example (The wobbly table problem)

# Intuitionism

- Impredicativism. (*Poincaré*, Russell)
  Vicious circle principle: No entity can be defined only in terms of a totality to which this entity belongs.
- Intuitionism Logic $\rightsquigarrow$ Mathematics $\rightsquigarrow$ Mental construction.
  (Kronecker, *Brouwer*, Heyting, *Kolmogorov*, Weyl)
  - Potential infinity vs actual infinity.
  - To be is to be constructed by intuition.
  - Law of excluded middle.$\times$
  - Non-constructive proof.$\times$

  (There exist two irrational numbers $x$ and $y$ s.t. $x^y$ is rational.)

  $$\sqrt{2} \quad \log_2 9$$

  "God created the integers, all the rest is the work of man."
- Constructive Mathematics. (Bishop, *Martin-Löf*)

# Existence Proofs — Constructive vs Nonconstructive

## The Chomp Game

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。

# Tree Method for Propositional Logic ♡

$$\begin{array}{c} \neg\neg A \\ | \\ A \end{array} \qquad \begin{array}{c} A \to B \\ \diagup \quad \diagdown \\ \neg A \qquad B \end{array} \qquad \begin{array}{c} \neg(A \to B) \\ | \\ A \\ \neg B \end{array}$$

---

$$\begin{array}{c} A \wedge B \\ | \\ A \\ B \end{array} \qquad \begin{array}{c} \neg(A \wedge B) \\ \diagup \quad \diagdown \\ \neg A \qquad \neg B \end{array} \qquad \begin{array}{c} A \vee B \\ \diagup \quad \diagdown \\ A \qquad B \end{array} \qquad \begin{array}{c} \neg(A \vee B) \\ | \\ \neg A \\ \neg B \end{array}$$

---

$$\begin{array}{c} A \leftrightarrow B \\ \diagup \quad \diagdown \\ \begin{array}{c} A \\ B \end{array} \qquad \begin{array}{c} \neg A \\ \neg B \end{array} \end{array} \qquad \begin{array}{c} \neg(A \leftrightarrow B) \\ \diagup \quad \diagdown \\ \begin{array}{c} A \\ \neg B \end{array} \qquad \begin{array}{c} \neg A \\ B \end{array} \end{array}$$

✓

# Tree Method for Predicate Logic — Ground Tree ♡

$$\begin{array}{c} \forall x A \\ | \\ A[t/x] \end{array}$$

where $t$ is a ground term.

$$\begin{array}{c} \exists x A \ \checkmark \\ | \\ A(a) \end{array}$$

where $a$ is a new constant which does not appear on the same branch as $\exists x A$.

---

$$\begin{array}{c} \neg\forall x A \ \checkmark \\ | \\ \exists x \neg A \end{array}$$

$$\begin{array}{c} \neg\exists x A \ \checkmark \\ | \\ \forall x \neg A \end{array}$$

# Tree Method for Predicate Logic ♡

$$\begin{array}{c} A(x) \\ x = y \\ | \\ A[y/\!\!/x] \end{array} \qquad\qquad \begin{array}{c} A(x) \\ y = x \\ | \\ A[y/\!\!/x] \end{array}$$

where $A[y/\!\!/x]$ arises from the wff $A(x)$ by replacing one or more occurrences of $x$ by $y$.

## Contradictions

1. $A, \neg A$
2. $t \neq t$

# 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.

## 证明技巧

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 (Soundness Theorem)**

*If the tree closes, the set is unsatisfiable.*

**Theorem (Completeness Theorem)**

*If a set is unsatisfiable, there exists a closed tree from it.*

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

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

**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

$\boxed{\left\{\forall x \neg P(x), \exists x(P(x) \vee P(f(x)))\right\} \text{ is unsatisfiable.}}$

$$\forall x \neg P(x)$$
$$\exists x(P(x) \vee P(f(x))) \checkmark$$

$$P(a) \vee P(f(a)) \checkmark$$

$$P(a) \qquad\qquad P(f(a))$$

$$\neg P(a) \qquad\qquad \neg P(f(a))$$

$$\times \qquad\qquad \times$$

# Example

有一个人，如果他酗酒，那么所有人都酗酒。 $\boxed{\vdash \exists x\big(A(x) \to \forall x A(x)\big)}$

$$\neg\exists x(A(x) \to \forall x A(x))$$

$$\neg(A(a) \to \forall x A(x)) \checkmark$$

$$A(a)$$

$$\neg\forall x A(x) \checkmark$$

$$\neg A(b)$$

$$\neg(A(b) \to \forall x A(x)) \checkmark$$

$$A(b)$$

$$\neg\forall x A(x)$$

$$\times$$

**Remark:** $\nvdash \exists x Ax \to \forall x Ax$

# 练习：有效性判定 ✍♡✍

$\forall x A \vdash A[t/x]$

$A[t/x] \vdash \exists x A$

$\neg \forall x A \equiv \exists x \neg A$

$\neg \exists x A \equiv \forall x \neg A$

$\forall x (A \wedge B) \equiv \forall x A \wedge \forall x B$

$\forall x A \vee \forall x B \vdash \forall x (A \vee B)$

$\exists x (A \vee B) \equiv \exists x A \vee \exists x B$

$\exists x (A \wedge B) \vdash \exists x A \wedge \exists x B$

$\forall x (A \to B) \vdash \forall x A \to \forall x B$

$\forall x (A \to B) \vdash \exists x A \to \exists x B$

$\forall x (A \leftrightarrow B) \vdash \forall x A \leftrightarrow \forall x B$

$\forall x \forall y A \equiv \forall y \forall x A$

$\exists x \exists y A \equiv \exists y \exists x A$

$\exists x \forall y A \vdash \forall y \exists x A$

$\forall x A \to \exists x B \equiv \exists x (A \to B)$

$\exists x A \to \forall x B \equiv \forall x (A \to B)$

$\vdash \exists x (A \to \forall x A)$

# 练习：有效性判定 ✍♡✍

$x \notin \mathrm{Fv}(A)$ :

| | |
|---|---|
| $A \equiv \forall x A$ | $A \equiv \exists x A$ |
| $\forall x(A \vee B) \equiv A \vee \forall x B$ | $\exists x(A \vee B) \equiv A \vee \exists x B$ |
| $\forall x(A \wedge B) \equiv A \wedge \forall x B$ | $\exists x(A \wedge B) \equiv A \wedge \exists x B$ |
| $\forall x(A \to B) \equiv A \to \forall x B$ | $\exists x(A \to B) \equiv A \to \exists x B$ |
| $\forall x(B \to A) \equiv \exists x B \to A$ | $\exists x(B \to A) \equiv \forall x B \to A$ |

# Remark

$$\forall x B \to A \equiv \exists x (B \to A)$$

$$\operatorname{diam}(X) := \sup \left\{ |x - y| : x, y \in X \right\}$$

$$\big(\forall x \in X \ |x| \leq 1\big) \to \operatorname{diam}(X) \leq 2$$

$$\Updownarrow ?$$

$$\exists x \in X \ \big(|x| \leq 1 \to \operatorname{diam}(X) \leq 2\big) ?$$

# Theorems / Valid Formulas/Arguments ♡

$$\vdash t = t$$

$$\vdash t = s \longrightarrow s = t$$

$$\vdash t = s \longrightarrow s = r \longrightarrow t = r$$

$$t_1 = s_1, \ldots, t_n = s_n \vdash f(t_1, \ldots, t_n) = f(s_1, \ldots, s_n)$$

$$t_1 = s_1, \ldots, t_n = s_n \vdash P(t_1, \ldots, t_n) \leftrightarrow P(s_1, \ldots, s_n)$$

$$t = s \vdash r[t/x] = r[s/x]$$

$$t = s \vdash A[t/x] \leftrightarrow A[s/x]$$

# Theorems / Valid Formulas/Arguments ♡

$x \notin \mathrm{Var}(t)$ :

$$\vdash \exists x(x = t)$$
$$A[t/x] \equiv \exists x(x = t \wedge A)$$
$$A[t/x] \equiv \forall x(x = t \rightarrow A)$$

$y \notin \mathrm{Var}(A)$ :

$$\vdash \forall x(A \leftrightarrow \exists y(y = x \wedge A[y/x]))$$
$$\vdash \forall x(A \leftrightarrow \forall y(y = x \rightarrow A[y/x]))$$

# 练习：树形证明 — Now it's your turn ✍♡✍

$$\frac{\exists x(Ax \to B) \quad \exists x(B \to Ax)}{\exists x(Ax \leftrightarrow B)} \qquad \frac{\exists x \forall y(P(y) \to y = x) \quad \forall x P(f(x))}{\exists x(x = f(x))}$$

$$\frac{\exists x\big(Px \wedge \forall y(Py \to y = x) \wedge Qx\big)}{\exists x \forall y\big((Py \leftrightarrow y = x) \wedge Qx\big)}$$

$$\frac{\exists x\big(Px \wedge \forall y(Py \to y = x)\big) \quad \exists x\big(Qx \wedge \forall y(Qy \to y = x)\big) \quad \neg\exists x(Px \wedge Qx)}{\exists xy\big(x \neq y \wedge (Px \vee Qx) \wedge (Py \vee Qy) \wedge \forall z(Pz \vee Qz \to z = x \vee z = y)\big)}$$

$*\mathbf{54} \cdot \mathbf{43}.\ \vdash: .\alpha, \beta \in 1. \supset: \alpha \cap \beta = \Lambda. \equiv .\alpha \cup \beta \in 2$

*Dem.*

$$\begin{aligned}
&\vdash .*54 \cdot 26. \supset \vdash: .\alpha = \iota' x . \beta = \iota' y. \supset: \alpha \cup \beta \in 2. \equiv .x \neq y. \\
&[*51 \cdot 231] \qquad \equiv .\iota' x \cap \iota' y = \Lambda. \\
&[*13 \cdot 12] \qquad \equiv .\alpha \cap \beta = \Lambda \qquad (1) \\
&\vdash .(1). *11 \cdot 11 \cdot 35. \supset \\
&\qquad \vdash: .(\exists x, y).\alpha = \iota' x . \beta = \iota' y. \supset: \alpha \cup \beta \in 2. \equiv .\alpha \cap \beta = \Lambda \qquad (2) \\
&\vdash .(2). *11 \cdot 54. *52 \cdot 1. \supset \vdash .Prop
\end{aligned}$$

From this proposition it will follow, when arithmetical addition has been defined, that $\mathbf{1} + \mathbf{1} = \mathbf{2}$.

$$\exists_1 x Px \wedge \exists_1 x Qx \wedge \neg\exists x(Px \wedge Qx) \to \exists_2 x(Px \vee Qx)$$

# 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。(非欧几何)

弗雷格 算术是分析的。

$$\vdash 1 + 1 = 2$$

# 罗素[11]

- 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（三段论）打补丁。
- 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。

[11] Bertrand Russell: “Mathematics and the Metaphysicians”. Reprinted in *Mysticism and Logic, and Other Essays*.

# 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。

# Philosophy of Math: Logicism/Intuitionism/Formalism

| **Logicism** | **Intuitionism** | **Formalism** |
|---|---|---|
| $\dfrac{\text{Mathematics}}{\text{Logic}}$ | $\dfrac{\dfrac{\text{Logic}}{\text{Mathematics}}}{\text{Mind}}$ | $\dfrac{\text{Mathematics}}{\text{Game}}$ |
| Realism | Conceptualism | Nominalism |

# David Hilbert 1862-1943

- **Formal** Axiomatization of Geometry.
  The consistency of geometry relative to arithmetic.
  (Klein: Non-Euclidean relative to Euclidean)
  (natural/integer/rational/real/complex)
- Hilbert's $23/24$ problems. $(1^{\text{st}}, 2^{\text{nd}}, 10^{\text{th}}, 24^{\text{th}})$
- Meta-mathematics — Proof Theory.
- Formalism Mathematics $\rightsquigarrow$ Symbolic Game.

- 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”来代替“点、线、面”。
- 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！

# 皮亚诺算术的形式公理化 ☺

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 $X$

$$\forall X \left[ 0 \in X \wedge \forall n (n \in X \to s(n) \in X) \to \forall n \in \mathbb{N} (n \in X) \right]$$

## Hilbert's Program — acquire knowledge about the actual world via a detour through the ideal world

1. Formalization & Axiomatization (Richness): Formalize and axiomatize (recursively) the elementary logic $\mathrm{L}$, the "finitistic" mathematics $\mathrm{F}$ (which is concerned with the 'actual world') and the "infinitistic" mathematics $\mathrm{T}$ (which is concerned with the 'ideal world').
2. Independence: the axioms should be "independent" of one another.
3. Completeness: $(1)$ all valid logical statements can be proved in $\mathrm{L}$; $(2)$ all true mathematical statements can be proved in $\mathrm{T}$.
4. **Consistency**: a finitistic proof that no contradiction can be proved in $\mathrm{T}$.
5. Conservation(Consequence of Consistency $\forall A \in \Pi_1 : \mathrm{T} \vdash A \implies \mathrm{F}, \mathrm{Con}_{\mathrm{T}} \vdash A$): any statement about 'real objects' provable in $\mathrm{T}$ can be proved in $\mathrm{F}$.
6. Decidability (Effectiveness): a mechanical procedure for deciding the validity of any logical statement and the truth of any mathematical statement.
7. Simplicity: a criteria of simplicity, or proof of the greatest simplicity of certain proofs.
8. Categoricity? $\mathrm{T}$ characterizes exactly one model up to isomorphism.

# Hilbert's Program

*"When we consider the axiomatization of logic more closely we soon recognize that the question of the consistency of the integers and of sets is not one that stands alone, but that it belongs to a vast domain of difficult* ***epistemological questions***[12] *which have a specifically mathematical tint: e.g, the problem of the solvability in principle of every mathematical question, the problem of the subsequent checkability of the results of a mathematical investigation, the question of a criterion of simplicity for mathematical proofs, the question of the relationship between content and formalism in mathematics and logic, and finally the problem of the decidability of a mathematical question in a finite number of operations."*

*— Hilbert*

[12] Chaitin: "All of this that's happening now with the computer taking over the world, the digitalization of our society, of information in human society, is the result of a **philosophical question** that was raised by Hilbert at the beginning of the century."

# Leibniz & Hilbert — Dream Shattered...

# 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?[13]

[13] Penelope Maddy: What Do We Want a Foundation to Do?

# 有效论证?

1. Whatever begins to exist, has a cause of its existence. The universe began to exist. Therefore, the universe has a cause of its existence.

$$\frac{\begin{array}{c}\forall x\big(Bx \to \exists y Cyx\big)\\ Bu\end{array}}{\exists y Cyu}$$

2. Alice is looking at Bob, but Bob is looking at Carl. Alice is married, but Carl is not. Is a married person looking at an unmarried person?

$$\frac{\begin{array}{c}Lab\\ Lbc\\ Ma\\ \neg Mc\end{array}}{\exists xy(Mx \land \neg My \land Lxy)}$$

# 有效论证?

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。因此，帕斯卡不会得到救赎。

$$\frac{\begin{array}{c} Eg \to \forall x(\neg Bx \to \neg Sx) \\ \neg Eg \to \neg \exists x Sx \\ \forall x(Wx \to \neg Bx) \end{array}}{Wp \to \neg Sp}$$

$$\frac{\begin{array}{c} \text{Everybody loves my baby} \\ \text{My baby loves nobody but me} \end{array}}{\text{Everybody loves me}}$$

$$\frac{\begin{array}{c} \forall x Lxb \\ \forall x(Lbx \to x = m) \end{array}}{\forall x Lxm}$$

# 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。

$$\frac{\begin{array}{c}\forall x(Px \to Fx \land Sx)\\ \forall x(Wx \to \neg Fx \land \neg Sx)\end{array}}{\forall x(Fx \land \neg Sx \to \neg Px \land \neg Wx)}$$

# 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。"
- For all $y$ there exists $z$ such that $z$ owns $y$.
- For all $x, y, z$ if $z$ owns $y$ and $y$ owns $x$ then $z$ owns $x$.
- 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. Therefore, Colonel West is a criminal.

1. $\forall xyz(\mathrm{American}(x) \land \mathrm{Weapon}(y) \land \mathrm{Hostile}(z) \land \mathrm{Sell}(x, y, z) \to \mathrm{Criminal}(x))$
2. $\exists x(\mathrm{Own}(\mathsf{nono}, x) \land \mathrm{Missile}(x))$
3. $\forall x(\mathrm{Missile}(x) \land \mathrm{Own}(\mathsf{nono}, x) \to \mathrm{Sell}(\mathsf{west}, x, \mathsf{nono}))$
4. $\forall x(\mathrm{Missile}(x) \to \mathrm{Weapon}(x))$
5. $\forall x(\mathrm{Enemy}(x, \mathsf{america}) \to \mathrm{Hostile}(x))$
6. $\mathrm{American}(\mathsf{west})$
7. $\mathrm{Enemy}(\mathsf{nono}, \mathsf{america})$
8. $\mathrm{Criminal}(\mathsf{west})$

1. 所有中国学生和所有日本学生都爱 Gakki。中南大学只有中国学生和日本学生。因此，中南大学的学生都爱 Gakki。

$$\frac{\begin{array}{c}\forall x(Cx \vee Jx \to Lxg)\\ \forall x(Zx \to Cx \vee Jx)\end{array}}{\forall x(Zx \to Lxg)}$$

2. 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。

$$\frac{\begin{array}{c}\forall xCx \to \neg\exists xNx\\ \forall x\neg Cx \to \neg\exists xEx\end{array}}{\exists x(Nx \wedge Ex) \to \exists xCx \wedge \exists x\neg Cx}$$

3. 一个系统是完全的当且仅当所有它能表达的真命题在该系统中都可证。一个系统是一致的当且仅当，存在一个真命题，它在该系统中虽然能表达但不可证。因此，所有不一致的系统都是完全的。

$$\frac{\begin{array}{c}\forall x(Cx \leftrightarrow \forall y(Eyx \wedge Ty \to Pyx))\\ \forall x(Sx \leftrightarrow \exists y(Ty \wedge Eyx \wedge \neg Pyx))\end{array}}{\forall x(\neg Sx \to Cx)}$$

# 有效论证 — “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，那么其包含的事物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. $\forall xy(\mathrm{Man}(x) \land \mathrm{Life}(y,x) \to \exists z(\mathrm{Possible}(z,x,y) \land \neg\,\mathrm{Realize}(z)))$
2. $\forall xy(\mathrm{Man}(x) \land \mathrm{Life}(y,x) \land \neg\exists y'(\mathrm{Life}(y',x) \land y' \neq y) \to$
   $\forall z(\mathrm{Possible}(z,x,y) \land \neg\,\mathrm{Realize}(z) \to \mathrm{UnRealizable}(z)))$
3. $\forall x(\mathrm{UnRealizable}(x) \to \neg\,\mathrm{Meaning}(x))$
4. $\mathrm{Meaning}(u) \to \forall xyz(\mathrm{Contain}(u,x) \land \mathrm{Possible}(z,x,y) \to \mathrm{Meaning}(z))$
5. $\forall x(\mathrm{Learnable}(x) \to \mathrm{Meaning}(x))$
6. $\forall x(\mathrm{Ordered}(x) \to \mathrm{Learnable}(x))$
7. $\mathrm{Ordered}(u)$
8. $\forall x(\mathrm{Man}(x) \to \mathrm{Contain}(u,x))$
9. $\forall xy(\mathrm{Man}(x) \land \mathrm{Life}(y,x) \to \exists y'(\mathrm{Life}(y',x) \land y' \neq y))$

# 练习：树形证明 — Now it's your turn ✍♡✍

1. Some wars are just. No war of aggression is just. Therefore, there are wars that are not wars of aggression.
2. Every non-analytic, meaningful proposition is either verifiable or falsifiable. Philosophical propositions are neither analytic nor verifiable or falsifiable. Therefore, they are meaningless.
3. If dogs are animals, every head of a dog is the head of an animal.
4. *The* present King of France is bald. Bald men are sexy. Hence whoever is a present King of France is sexy.
5. *Only* Russell is a great philosopher. Wittgenstein is a great philosopher who smokes. So Russell smokes.
6. Everyone is afraid of Dracula. Dracula is afraid *only* of me. Therefore, I am Dracula.
7. Everybody loves anybody who loves somebody. Romeo loves Juliet. Therefore, I love you.
8. Everybody loves anybody who loves somebody; hence if somebody loves somebody, everybody loves everybody!
9. Nobody trusts *exactly* those who have no mutual trust with anybody.

10. Whoever stole the goods, knew the safe combination. Someone stole the goods, and *only* Jack knew the safe combination. Hence Jack stole the goods.
11. No girl loves any sexist pig. Caroline is a girl who loves whoever loves her. Henry loves Caroline. Thus Henry isn't a sexist pig.
12. I am a philosopher. A philosopher can *only* be appreciated by philosophers. No philosopher is without some eccentricity. I sing rock. Every eccentric rock singer is appreciated by some girl. Eccentrics are conceited. Therefore, some girl is conceited.
13. Any philosopher admires some logician. Some students admire *only* film stars. No film stars are logicians. Therefore not all students are philosophers.
14. 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.
15. *No one but* Alice and Bette (*who are different people*) admires Carl. All and only those who admire Carl love him. Hence *exactly* two people love Carl.

# Application — Minesweeper

- 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.

$$\forall x(\mathrm{Contain}(x, 1) \rightarrow$$
$$\exists y(\mathrm{Adj}(x, y) \wedge \mathrm{Mine}(y) \wedge \forall z(\mathrm{Adj}(x, z) \wedge \mathrm{Mine}(z) \rightarrow z = y)))$$

- ...

# Application — Example

## Problem

*A man goes to the brook with two measures of* $15$ *pints and* $16$ *pints. How is he to measure* $8$ *pints of water?*

| | |
|---|---|
| States | $(x, y)$: $x$ is the content of the first vessel, $y$ the second. |
| Initial State | $(15, 16)$ |
| Goal State | $\exists x[(8, x) \vee (x, 8)]$ |

There are $8$ possible actions:

| | |
|---|---|
| Fill $1$ | $(x, y) \rightarrow (15, y)$ |
| Empty $1$ | $(x, y) \rightarrow (0, y)$ |
| Fill $2$ | $(x, y) \rightarrow (x, 16)$ |
| Empty $2$ | $(x, y) \rightarrow (x, 0)$ |
| Empty $1$ into $2$ | $(x, y) \wedge x + y \leq 16 \rightarrow (0, y + x)$ |
| Pour $1$ into $2$ | $(x, y) \wedge x + y > 16 \rightarrow (x - (16 - y), 16)$ |
| Empty $2$ into $1$ | $(x, y) \wedge x + y \leq 15 \rightarrow (x + y, 0)$ |
| Pour $2$ into $1$ | $(x, y) \wedge x + y > 15 \rightarrow (15, y - (15 - x))$ |

# Application in Game Theory

## Theorem (Zermelo's Theorem 1913)

*Every finite game of perfect information with no tie is determined.*

## 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.

## 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.

# 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.”

flying horse ⟶

$$B(\iota_x A) := \exists x\big(Ax \wedge \forall y(Ay \to y = x) \wedge Bx\big)$$

- I don't know the author of Sherlock Holmes is Conan Doyle.
  $\neg \mathsf{K}(\iota_x A = c)$ ? $\quad \iota_x A = c \to \mathsf{K}(c = c) \to \mathsf{K}(\iota_x A = c)$ ?
  $$\exists x\big(Ax \wedge \forall y(Ay \to y = x) \wedge \neg \mathsf{K}(x = c)\big)$$
  $$\neg \mathsf{K}\left[\exists x\big(Ax \wedge \forall y(Ay \to y = x) \wedge x = c\big)\right]$$
- 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. $\neg E(\iota_x Fx)$
  $$\exists x\big(Fx \wedge \forall y(Fy \to y = x) \wedge \neg Ex\big) \ ?$$
  $$\neg\exists x\big(Fx \wedge \forall y(Fy \to y = x)\big)$$
  $$Ex \overset{?}{:=} \exists P\big(Px \wedge \exists y \neg Py\big)$$
- Get rid of function symbols.
  $$\mathrm{Bald}(\iota_x\, \mathrm{Father}(x, \mathsf{alice})) \quad vs \quad \mathrm{Bald}(\mathrm{father}(\mathsf{alice}))$$
- Universal Instantiation. $\forall x B \to B(\iota_x A)$ ?
  $$B(\iota_x^y A) := \big(\exists! x A \to \exists x(A \wedge B)\big) \wedge \big(\neg\exists! x A \to B[y/x]\big)$$
  $$\vdash \forall x B \to B(\iota_x^y A)$$

# 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.

$$Z\iota_x\big(Px \wedge \forall y(Py \rightarrow Ryx)\big)\iota_x\big(Cx \wedge \forall y(Py \rightarrow Ryx)\big)$$

# Bertrand Russell 1872-1970

- Russell Paradox.
  ($3^{ed}$ crisis of the Foundations of Mathematics)
- Theory of Descriptions.
  (The present King of France is not bald.)
- Type Theory.
- *Principia Mathematica.*

No barber shaves exactly those who do not shave themselves.[14]

[14] 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。[15]
- 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 $f(x) = x^2$,

$$\frac{df(x)}{dx} = \frac{f(x+dx) - f(x)}{dx} = \frac{(x+dx)^2 - x^2}{dx} = \frac{2xdx + (dx)^2}{dx} = 2x + dx = 2x$$

- 维尔斯特拉斯：$\frac{df(x)}{dx}$ 不是 $df(x)$ 与 $dx$ 的商，$\frac{d}{dx}$ 作为微分运算作用于 $f(x)$,

$$\frac{df(x)}{dx} = \frac{d}{dx}f(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

[15] 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。

— *David Foster Wallace*

## Remark

形式公理化 *vs* 自然齐一性

# 逻辑主义与逻辑实证主义

- 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 $\lambda$-Abstraction

$$v(\iota_x A) = \begin{cases} a & \text{if there is a unique } a \in M : \mathcal{M}, v(a/x) \vDash A \\ \uparrow & \text{otherwise} \end{cases}$$

$$\begin{cases} \mathcal{M}, v \vDash (\lambda x.A)t \iff \mathcal{M}, v \vDash A[t/x] & \text{if } v(t) \downarrow \\ \mathcal{M}, v \nvDash (\lambda x.A)t & \text{if } v(t) \uparrow \end{cases}$$

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.

$$\mathbf{D}(\lambda x(Cx \wedge \neg Lx))$$

# Expressive Limitation of First Order Language

- Most boys are funny.
- Some critics admire only one another.

$$\exists X\Big(\exists x Xx \wedge \forall x(Xx \to Cx) \wedge \forall x \forall y\big(Xx \wedge Axy \to Xy \wedge x \neq y\big)\Big)$$

- 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 $\varepsilon$

$$A(t) \to A(\varepsilon_x A)$$

where $t$ is an arbitrary term.

## $\varepsilon$-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 $\varepsilon$-theorem: Suppose $A$ is a quantifier-free and $\varepsilon$-free wff. Then $\vdash_{\varepsilon} A \implies \vdash_{\mathrm{QF}} A$ in quantifier-free predicate logic.
- Extended first $\varepsilon$-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 $\varepsilon$-theorem: Suppose $A$ is an $\varepsilon$-free wff. Then $\vdash_{\varepsilon} A \implies \vdash A$.

Herbrand's Theorem is a corollary of the extended first $\varepsilon$-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

$$\forall x A \checkmark$$
$$\Big|$$
$$A[x_i/x]$$

where $x_i$ is a new variable.

$$\exists x A \checkmark$$
$$\Big|$$
$$A[f(x_1, \ldots, x_m)/x]$$

where $f$ is a new function and $\{x_1, \ldots, x_m\} = \mathrm{Fv}(\exists x A)$.

---

$$\neg\forall x A \checkmark$$
$$\Big|$$
$$\exists x \neg A$$

$$\neg\exists x A \checkmark$$
$$\Big|$$
$$\forall x \neg A$$

# 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 $\sigma$ is a *unifier* for a set $\Gamma$ of formulas iff for every $A, B \in \Gamma : A\sigma = B\sigma$.
- A unifier $\sigma$ is a *most general unifier* for $\Gamma$ iff for each unifier $\theta$ there exists a substitution $\lambda$ s.t. $\theta = \sigma\lambda$.

$$\begin{tikzcd} \bullet \arrow[d, dashed, "\lambda"'] \arrow[dr, "\theta"] & \\ \bullet \arrow[r, "\sigma"'] & \bullet \end{tikzcd}$$

$$\sigma := \{t_1/x_1, \ldots, t_m/x_m\} \qquad \lambda := \{s_1/y_1, \ldots, s_n/y_n\}$$

$$\sigma\lambda = \left\{t_1\lambda/x_1, \ldots, t_m\lambda/x_m,\ s_1/y_1, \ldots, s_n/y_n\right\} \setminus \left\{s_i/y_i : y_i \in \{x_1, \ldots, x_m\}\right\}$$

- $(A\sigma)\lambda = A(\sigma\lambda)$ and $(t\sigma)\lambda = t(\sigma\lambda)$
- $(\sigma\lambda)\theta = \sigma(\lambda\theta)$

# Example — Tree Method with Unification

$\boxed{\{\forall x P(x), \neg Q(f(a)), \forall x(\neg P(f(x)) \vee Q(x))\} \text{ is unsatisfiable.}}$

$$\forall x P(x) \checkmark$$
$$\neg Q(f(a))$$
$$\forall x(\neg P(f(x)) \vee Q(x)) \checkmark$$
$$P(x_1)$$
$$\neg P(f(x_2)) \vee Q(x_2) \checkmark$$

| $\neg P(f(x_2))$ | $Q(x_2)$ |
|---|---|
| $x_1 = f(x_2)$ | $x_2 = f(a)$ |
| $\times$ | $\times$ |

# Unification — Greedy Unification (incomplete) 🚲

$$\neg P(a) \checkmark$$
$$\neg Q(c) \checkmark$$
$$\neg R(d) \checkmark$$
$$\forall x(((R(d) \vee \neg P(b)) \wedge P(x)) \vee ((R(d) \vee \neg Q(b)) \wedge Q(x))) \checkmark$$
$$((R(d) \vee \neg P(b)) \wedge P(x_1)) \vee ((R(d) \vee \neg Q(b)) \wedge Q(x_1)) \checkmark$$

$$(R(d) \vee \neg P(b)) \wedge P(x_1) \checkmark$$
$$R(d) \vee \neg P(b)$$
$$P(x_1)$$
$$x_1 = a$$
$$\times$$

$$(R(d) \vee \neg Q(b)) \wedge Q(a) \checkmark$$
$$R(d) \vee \neg Q(b) \checkmark$$
$$Q(a)$$

$$R(d)$$
$$\times$$

$$\neg Q(b)$$

Applying unification as soon as a branch can be closed by lead to incompleteness.

# Unification — Final Closure

$\neg P(a)$ ✓

$\neg Q(c)$ ✓

$\neg R(d)$ ✓

$\forall x(((R(d) \vee \neg P(b)) \wedge P(x)) \vee ((R(d) \vee \neg Q(b)) \wedge Q(x)))$ ✓

$((R(d) \vee \neg P(b)) \wedge P(x_1)) \vee ((R(d) \vee \neg Q(b)) \wedge Q(x_1))$ ✓

$(R(d) \vee \neg P(b)) \wedge P(x_1)$ ✓

$R(d) \vee \neg P(b)$ ✓

$P(x_1)$

$R(d)$

×

$\neg P(b)$

$x_1 = b$

×

$(R(d) \vee \neg Q(b)) \wedge Q(x_1)$ ✓

$R(d) \vee \neg Q(b)$ ✓

$Q(x_1)$

$R(d)$

×

$\neg Q(b)$

$x_1 = b$

×

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.

$$\boxed{\vdash \exists x\big(A(x) \to \forall x A(x)\big)}$$

$$\neg\exists x(A(x) \to \forall x A(x)) \checkmark$$

$$\forall x \neg(A(x) \to \forall x A(x)) \checkmark$$

$$\neg(A(x_1) \to \forall x A(x)) \checkmark$$

$$A(x_1)$$

$$\neg\forall x A(x) \checkmark$$

$$\neg A(a)$$

$$x_1 = a$$

$$\times$$

$$\forall x \neg(A(x) \to \forall x A(x))$$

$$\neg(A(a) \to \forall x A(x)) \checkmark$$

$$A(a)$$

$$\neg\forall x A(x) \checkmark$$

$$\neg A(b)$$

$$\neg(A(b) \to \forall x A(x)) \checkmark$$

$$A(b)$$

$$\neg\forall x A(x)$$

$$\times$$

# 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 $\mathrm{PA}$ (proof-theoretical strength of $\mathrm{PA}$)

# Natural Deduction

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

---

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

---

$$\frac{\begin{array}{c}[A]^n\\ \vdots \\ B\end{array}}{A \to B}\ [\to^+]^n \qquad \frac{A \to B \quad A}{B}\ [\to^-]$$

---

$$\frac{\begin{array}{c}[A]^n\\ \vdots \\ \bot\end{array}}{\neg A}\ [\neg^+]^n \qquad \frac{\neg\neg A}{A}\ [\neg^-] \qquad \frac{\neg A \quad A}{\bot}\ [\bot^+] \qquad \frac{\bot}{A}\ [\bot^-]$$

# 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 undischarged 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 undischarged in the derivations ending with $\exists x A, B$ except in $A(a)$.

---

$$\frac{}{t = t}\ [=^+] \qquad\qquad \frac{s = t \quad A(s)}{A[t /\!\!/ s]}\ [=^-]$$

# Remarks

$$\boxed{\dfrac{\forall x \exists y Rxy}{\exists y Ryy}\ [\forall^- ?] \quad \textbf{incorrect} \quad \exists y Rxy[y/x] \times}$$

$$\boxed{\dfrac{\forall x Rxx}{\forall x Rxa}\ ?}$$

$$\dfrac{\dfrac{\forall x Rxx}{Raa}\ [\forall^-]}{\forall x Rxa}\ [\forall^+] \quad \textbf{incorrect}$$

$$\boxed{\dfrac{}{\forall x(Ax \to \forall x Ax)}\ ?}$$

$$\dfrac{\dfrac{\dfrac{[Ax]^1}{\forall x Ax}\ [\forall^+]}{Ax \to \forall x Ax}\ [\to^+]^1}{\forall x(Ax \to \forall x Ax)}\ [\forall^+] \quad \textbf{incorrect}$$

# Remarks

$$\boxed{\frac{\exists x A}{Aa}\ ?}$$

$$\dfrac{\exists x A \qquad \begin{array}{c}[Aa]^1\\ \vdots \\ Aa\end{array}}{Aa}\,[\exists^-]^1$$

**incorrect**

$$\boxed{\frac{\exists x A \qquad \exists x B}{\exists x(A \wedge B)}\ ?}$$

$$\dfrac{\exists x B \qquad \dfrac{\exists x A \qquad \dfrac{\dfrac{[Aa]^1 \qquad [Ba]^2}{Aa \wedge Ba}\,[\wedge^+]}{\exists x(A \wedge B)}\,[\exists^+]}{\exists x(A \wedge B)}\,[\exists^-]^1}{\exists x(A \wedge B)}\,[\exists^-]^2$$

**incorrect**

# Remark: How do we prove universal statements?

- Choose an arbitrary object from a given set and show that this object has the desired property.

$$\frac{A(a)}{\forall x A}\ [\forall^+]$$

  where $a \notin \mathrm{Cst}(\forall x A)$, and $a$ is not in any assumption which is undischarged in the derivation ending with $A(a)$.
- **"random" $\neq$ "arbitrary".**
  If we choose an item randomly, it means that the probability of choosing it is equal to that of choosing any other element. An arbitrary element is a placeholder that acts as a representative of any of the other elements.

# Example

$$\boxed{\dfrac{\forall x(A \to B)}{\exists x A \to \exists x B}}$$

Proof.

$$\dfrac{\dfrac{[\exists x A]^2 \qquad \dfrac{\dfrac{\dfrac{\forall x(A \to B)}{Aa \to Ba} \qquad [Aa]^1}{Ba}}{\exists x B}\,[\exists^-]^1}{\exists x B}}{\exists x A \to \exists x B}\,[\to^+]^2$$

# Examples

$$\boxed{\frac{\forall x(A \to B)}{A \to \forall x B} \quad \text{where } x \notin \mathrm{Fv}(A)}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{\forall x(A \to B)}{A \to Ba}\,[\forall^-] \qquad [A]^1}{Ba}\,[\to^-]}{\forall x B}\,[\forall^+]}{A \to \forall x B}\,[\to^+]^1$$

$$\boxed{\frac{A \to \forall x B}{\forall x(A \to B)} \quad \text{where } x \notin \mathrm{Fv}(A)}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{A \to \forall x B \qquad [A]^1}{\forall x B}\,[\to^-]}{Ba}\,[\forall^-]}{A \to Ba}\,[\to^+]^1}{\forall x(A \to B)}\,[\forall^+]$$

# Examples

$$\boxed{\dfrac{\forall x(A \to B)}{\exists x A \to B} \quad \text{where } x \notin \mathrm{Fv}(B)}$$

Proof.

$$\dfrac{\dfrac{[\exists x A]^1 \qquad \dfrac{[Aa]^2 \qquad \dfrac{\forall x(A \to B)}{Aa \to B}\,[\forall^-]}{B}\,[\to^-]}{B}\,[\exists^-]^2}{\exists x A \to B}\,[\to^+]^3$$

$$\boxed{\dfrac{\exists x A \to B}{\forall x(A \to B)} \quad \text{where } x \notin \mathrm{Fv}(B)}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{[Aa]^1}{\exists x A}\,[\exists^+] \qquad \exists x A \to B}{B}\,[\to^-]}{Aa \to B}\,[\to^+]^1}{\forall x(A \to B)}\,[\forall^+]$$

# Examples

$$\boxed{\dfrac{\exists x \forall y Rxy}{\forall y \exists x Rxy}}$$

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

$$\frac{\text{Everybody loves anybody who loves somebody}}{\text{If somebody loves somebody, everybody loves everybody}}$$

$$\boxed{\frac{\forall xyz(Lyz \to Lxy)}{\exists xyLxy \to \forall xyLxy}}$$

Proof.

$$\dfrac{\dfrac{\dfrac{[\exists xyLxy]^2 \qquad \dfrac{\dfrac{\dfrac{\forall xyz(Lyz \to Lxy)}{Lrj \to Lur} \qquad [Lrj]^1}{Lur} \qquad \dfrac{\forall xyz(Lyz \to Lxy)}{Lur \to Liu}}{Liu}}{Liu}\,[\exists^-]^1}{\forall xyLxy}}{\exists xyLxy \to \forall xyLxy}\,[\to^+]^2$$

# 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

$$\boxed{\begin{array}{c} \forall x \forall y (Sxy \to \exists z Izxy) \\ \forall x \forall y \forall z (Izxy \to Kzx \wedge Kzy) \\ \forall x Sxa \\ \hline \forall x \exists y (Iyxa \wedge Kya) \end{array}}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall xy(Sxy \to \exists z Izxy)}{Sba \to \exists z Izba} \quad \dfrac{\forall x Sxa}{Sba}}{\exists z Izba} \quad \dfrac{\dfrac{\dfrac{[Icba]^1 \quad \dfrac{\forall xyz(Izxy \to Kzx \wedge Izy)}{Icba \to Kcb \wedge Kca}}{Kca}}{Icba \wedge Kca}}{\exists y(Iyba \wedge Kya)}}{\exists y(Iyba \wedge Kya)}\,[\exists^-]^1}{\forall x \exists y (Iyxa \wedge Kya)}$$

# 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.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall x(Dx \to Ax)}{Da \to Aa} \quad [Da]^1}{Aa} \quad \dfrac{}{ha = ha}}{Aa \wedge ha = ha}}{\exists y(Ay \wedge ha = hy)}}{Da \to \exists y(Ay \wedge ha = hy)}\,[\to^+]^1}{\forall x(Dx \to \exists y(Ay \wedge hx = hy))}$$

# Example

$$\frac{\text{如果大鱼比小鱼游得快}}{\text{只要有最大的鱼就有游得最快的鱼}}$$

$$\boxed{\frac{\forall x \forall y (Fx \wedge Fy \wedge Bxy \to Sxy)}{\exists x (Fx \wedge \forall y (Fy \to Bxy)) \to \exists x (Fx \wedge \forall y (Fy \to Sxy))}}$$

Proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[Fa \wedge \forall y (Fy \to Bay)]^1}{Fa \quad \forall y (Fy \to Bay)}}{Fb \to Bab \qquad [Fb]^2}}{Bab}}{Fa \wedge Fb \wedge Bab} \qquad \dfrac{\forall x \forall y (Fx \wedge Fy \wedge Bxy \to Sxy)}{Fa \wedge Fb \wedge Bab \to Sab}}{Sab}}{Fb \to Sab}\,[\to^+]^2}{\forall y (Fy \to Say)}}{Fa \wedge \forall y (Fy \to Say)}}{\exists x (Fx \wedge \forall y (Fy \to Sxy))}\,[\exists^-]^1 \qquad [\exists x (Fx \wedge \forall y (Fy \to Bxy))]^3}{\exists x (Fx \wedge \forall y (Fy \to Sxy))}}{\exists x (Fx \wedge \forall y (Fy \to Bxy)) \to \exists x (Fx \wedge \forall y (Fy \to Sxy))}\,[\to^+]^3$$

# Natural Deduction — another version

$$\frac{A \in \Gamma}{\Gamma \vdash A}\ [\mathrm{I}] \qquad \frac{\Gamma \vdash A \qquad \Gamma \subset \Gamma'}{\Gamma' \vdash A}\ [\mathrm{M}]$$

$$\frac{\Gamma \vdash A \qquad \Gamma \vdash B}{\Gamma \vdash A \wedge B}\ [\wedge^+] \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A}\ [\wedge^-] \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}\ [\wedge^-]$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B}\ [\vee^+] \qquad \frac{\Gamma \vdash A}{\Gamma \vdash B \vee A}\ [\vee^+] \qquad \frac{\Gamma \vdash A \vee B \qquad \Gamma, A \vdash C \qquad \Gamma, B \vdash C}{\Gamma \vdash C}\ [\vee^-]$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B}\ [\to^+] \qquad \frac{\Gamma \vdash A \to B \qquad \Gamma \vdash A}{\Gamma \vdash B}\ [\to^-]$$

$$\frac{\Gamma, A \vdash \bot}{\Gamma \vdash \neg A}\ [\neg^+] \qquad \frac{\Gamma \vdash \neg\neg A}{\Gamma \vdash A}\ [\neg^-] \qquad \frac{\Gamma \vdash \neg A \qquad \Gamma \vdash A}{\Gamma \vdash \bot}\ [\bot^+] \qquad \frac{\Gamma \vdash \bot}{\Gamma \vdash A}\ [\bot^-]$$

# Natural Deduction — another version

$$\frac{\Gamma \vdash A(a) \qquad a \notin \mathrm{Cst}(\Gamma, \forall x A)}{\Gamma \vdash \forall x A}\ [\forall^+] \qquad \frac{\Gamma \vdash \forall x A}{\Gamma \vdash A[t/x]}\ [\forall^-]$$

$$\frac{\Gamma \vdash A[t/x]}{\Gamma \vdash \exists x A}\ [\exists^+] \qquad \frac{\Gamma \vdash \exists x A \qquad \Gamma, A(a) \vdash B \qquad a \notin \mathrm{Cst}(\Gamma, \exists x A, B)}{\Gamma \vdash B}\ [\exists^-]$$

$$\frac{}{\Gamma \vdash t = t}\ [=^+] \qquad \frac{\Gamma \vdash s = t \qquad \Gamma \vdash A(s)}{\Gamma \vdash A[t /\!\!/ s]}\ [=^-]$$

# Examples

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

$$
\dfrac{\dfrac{\dfrac{\dfrac{}{\neg A \to \bot, \neg A \vdash \neg A \to \bot}\,[\mathrm{I}] \qquad \dfrac{}{\neg A \to \bot, \neg A \vdash \neg A}\,[\mathrm{I}]}{\neg A \to \bot, \neg A \vdash \bot}\,[\to^-]}{\neg A \to \bot \vdash \neg\neg A}\,[\neg^+]}{\neg A \to \bot \vdash A}\,[\neg^-]
$$

$\boxed{\text{If } x \notin \mathrm{Fv}(A), \text{ then } \vdash \forall x(A \to B) \to A \to \forall x B}$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\forall x(A \to B) \vdash \forall x(A \to B)}\,[\mathrm{I}]}{\forall x(A \to B) \vdash A \to Ba}\,[\forall^-]}{\forall x(A \to B), A \vdash Ba}\,[\to^-]}{\forall x(A \to B), A \vdash \forall x B}\,[\forall^+]}{\forall x(A \to B) \vdash A \to \forall x B}\,[\to^+]}{\vdash \forall x(A \to B) \to A \to \forall x B}\,[\to^+]
$$

# Sequent Calculus

Axiom

$$\frac{}{A \vdash A}\ [\mathrm{I}]$$

Cut

$$\frac{\Gamma_1 \vdash \Delta_1, A \qquad \Gamma_2, A \vdash \Delta_2}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\ [\mathrm{Cut}]$$

Left structural rules

$$\frac{\Gamma \vdash \Delta}{\Gamma, A \vdash \Delta}\ [\mathrm{WL}]$$

$$\frac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta}\ [\mathrm{CL}]$$

$$\frac{\Gamma_1, A, B, \Gamma_2 \vdash \Delta}{\Gamma_1, B, A, \Gamma_2 \vdash \Delta}\ [\mathrm{PL}]$$

Right structural rules

$$\frac{\Gamma \vdash \Delta}{\Gamma \vdash A, \Delta}\ [\mathrm{WR}]$$

$$\frac{\Gamma \vdash A, A, \Delta}{\Gamma \vdash A, \Delta}\ [\mathrm{CR}]$$

$$\frac{\Gamma \vdash \Delta_1, A, B, \Delta_2}{\Gamma \vdash \Delta_1, B, A, \Delta_2}\ [\mathrm{PR}]$$

# Sequent Calculus

Left logical rules:

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}\ [\wedge L_1]$$

$$\frac{\Gamma, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}\ [\wedge L_2]$$

$$\frac{\Gamma_1, A \vdash \Delta_1 \qquad \Gamma_2, B \vdash \Delta_2}{\Gamma_1, \Gamma_2, A \vee B \vdash \Delta_1, \Delta_2}\ [\vee L]$$

$$\frac{\Gamma_1 \vdash A, \Delta_1 \qquad \Gamma_2, B \vdash \Delta_2}{\Gamma_1, \Gamma_2, A \rightarrow B \vdash \Delta_1, \Delta_2}\ [\rightarrow L]$$

Right logical rules:

$$\frac{\Gamma \vdash A, \Delta}{\Gamma \vdash A \vee B, \Delta}\ [\vee R_1]$$

$$\frac{\Gamma \vdash B, \Delta}{\Gamma \vdash A \vee B, \Delta}\ [\vee R_2]$$

$$\frac{\Gamma_1 \vdash A, \Delta_1 \qquad \Gamma_2 \vdash B, \Delta_2}{\Gamma_1, \Gamma_2 \vdash A \wedge B, \Delta_1, \Delta_2}\ [\wedge R]$$

$$\frac{\Gamma, A \vdash B, \Delta}{\Gamma \vdash A \rightarrow B, \Delta}\ [\rightarrow R]$$

# Sequent Calculus

Left logical rules: | Right logical rules:

$$\frac{\Gamma \vdash A, \Delta}{\Gamma, \neg A \vdash \Delta}\ [\neg L] \qquad \frac{\Gamma, A \vdash \Delta}{\Gamma \vdash \neg A, \Delta}\ [\neg R]$$

$$\frac{\Gamma, A[t/x] \vdash \Delta}{\Gamma, \forall x A \vdash \Delta}\ [\forall L] \qquad \frac{\Gamma \vdash A(a), \Delta \qquad a \notin \mathrm{Cst}(\Gamma, \forall x A, \Delta)}{\Gamma \vdash \forall x A, \Delta}\ [\forall R]$$

$$\frac{\Gamma, A(a) \vdash \Delta \qquad a \notin \mathrm{Cst}(\Gamma, \exists x A, \Delta)}{\Gamma, \exists x A \vdash \Delta}\ [\exists L] \qquad \frac{\Gamma \vdash A[t/x], \Delta}{\Gamma \vdash \exists x A, \Delta}\ [\exists R]$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}\,[\mathrm{I}]}{\vdash \neg A, A}\,[\neg R]}{\vdash \neg A \vee A, A}\,[\vee R_2]}{\vdash A, A \vee \neg A}\,[PR]}{\vdash A \vee \neg A, A \vee \neg A}\,[\vee R_1]}{\vdash A \vee \neg A}\,[CR]$$

$$\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}\,[\mathrm{I}]}{A \wedge B \vdash A}\,[\wedge L_1]}{A \wedge B, \neg A \vdash}\,[\neg L] \qquad \dfrac{\dfrac{\dfrac{}{B \vdash B}\,[\mathrm{I}]}{A \wedge B \vdash B}\,[\wedge L_2]}{A \wedge B, \neg B \vdash}\,[\neg L]}{\dfrac{A \wedge B, \neg A \vee \neg B \vdash}{\neg A \vee \neg B \vdash \neg(A \wedge B)}\,[\neg R]}\,[\vee L]$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}\,[\mathrm{I}]}{A \vdash B, A}\,[WR]}{A, \neg A \vdash B}\,[\neg L] \qquad \dfrac{\dfrac{}{B \vdash B}\,[\mathrm{I}]}{A, B \vdash B}\,[WL]}{A, \neg A \vee B \vdash B}\,[\vee L]}{\neg A \vee B \vdash A \to B}\,[\to R]}{}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{Rab \vdash Rab}\,[\mathrm{I}]}{\forall x Rxb \vdash Rab}\,[\forall L]}{\forall x Rxb \vdash \exists y Ray}\,[\exists R]}{\exists y \forall x Rxy \vdash \exists y Ray}\,[\exists L]}{\exists y \forall x Rxy \vdash \forall x \exists y Rxy}\,[\forall R]}{}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{Aa \vdash Aa}\,[\mathrm{I}]}{\forall x A \vdash Aa}\,[\forall L]}{\forall x A, \neg Aa \vdash}\,[\neg L]}{\neg Aa \vdash \neg \forall x A}\,[\neg R]}{\exists x \neg A \vdash \neg \forall x A}\,[\exists L]$$

# Natural Deduction — constant vs variable

$$\frac{\Gamma \vdash A(a) \qquad a \notin \mathrm{Cst}(\Gamma, \forall x A)}{\Gamma \vdash \forall x A}\ [\forall^+]$$

$$\frac{\Gamma \vdash A[y/x] \qquad y \notin \mathrm{Fv}(\Gamma, \forall x A)}{\Gamma \vdash \forall x A}\ [\forall^+]$$

$$\frac{\Gamma \vdash \exists x A \qquad \Gamma, A(a) \vdash B \qquad a \notin \mathrm{Cst}(\Gamma, \exists x A, B)}{\Gamma \vdash B}\ [\exists^-]$$

$$\frac{\Gamma \vdash \exists x A \qquad \Gamma, A[y/x] \vdash B \qquad y \notin \mathrm{Fv}(\Gamma, \exists x A, B)}{\Gamma \vdash B}\ [\exists^-]$$

# Natural Deduction — another version — constant vs variable

$$\frac{A(a)}{\forall x A}\ [\forall^+] \qquad\qquad \frac{A[y/x]}{\forall x A}\ [\forall^+]$$

where $a \notin \mathrm{Cst}(\forall x A)$, and $a$ is not in any assumption which is undischarged 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 undischarged in the derivation ending with $A[y/x]$.

---

$$\frac{\exists x A \qquad \begin{array}{c}[A(a)]^n\\ \vdots \\ B\end{array}}{B}\ [\exists^-]^n \qquad\qquad \frac{\exists x A \qquad \begin{array}{c}[A[y/x]]^n\\ \vdots \\ B\end{array}}{B}\ [\exists^-]^n$$

where $a \notin \mathrm{Cst}(\exists x A, B)$, and $a$ is not in any assumption which is undischarged 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 undischarged in the derivations ending with $\exists x A, B$ except in $A[y/x]$.

# Sequent Calculus — constant vs variable

$$\frac{\Gamma \vdash A(a), \Delta \qquad a \notin \mathrm{Cst}(\Gamma, \forall x A, \Delta)}{\Gamma \vdash \forall x A, \Delta} \; [\forall R]$$

$$\frac{\Gamma \vdash A[y/x], \Delta \qquad y \notin \mathrm{Fv}(\Gamma, \forall x A, \Delta)}{\Gamma \vdash \forall x A, \Delta} \; [\forall R]$$

$$\frac{\Gamma, A(a) \vdash \Delta \qquad a \notin \mathrm{Cst}(\Gamma, \exists x A, \Delta)}{\Gamma, \exists x A \vdash \Delta} \; [\exists L]$$

$$\frac{\Gamma, A[y/x] \vdash \Delta \qquad y \notin \mathrm{Fv}(\Gamma, \exists x A, \Delta)}{\Gamma, \exists x A \vdash \Delta} \; [\exists L]$$

# Cut-Elimination Theorem

## Theorem (Cut-Elimination Theorem — Gentzen1934)

*If $\Gamma \vdash \Delta$ is provable, then it is provable without use of 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 $\Gamma$ and $\Delta$.*

## 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 $\Theta$ be a set of sentences. If $\Theta$ 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.*

# Contents

Introduction

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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 wff $A$ 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 $\mathrm{T}$ iff there is a wff $A(x_1, \ldots, x_n, y)$ 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 $\mathrm{T}$ iff there is a wff $A$ 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 $\mathrm{Q}$ 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

$$\exists y \big(x = s(0) + s(y)\big) \wedge \forall yz \big(x = y \cdot z \to y = s(0) \vee z = s(0)\big)$$

- 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

$$\exists y_1 y_2 y_3 y_4 \left( x = y_1^2 + y_2^2 + y_3^2 + y_4^2 \right) \quad \text{(Lagrange four-square theorem)}$$

# 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.*

- $\pi$ 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 Structure vs Leibnizian Structure

## Definition (Leibnizian Structure)

A model $\mathcal{M}$ is Leibnizian if any two distinct objects can be distinguished by some property: for $a \neq b$, there is a wff $A$ s.t.,

$$\mathcal{M} \vDash A(a) \wedge \neg A(b)$$

## Definition (Pointwise Definable Structure)

A model $\mathcal{M}$ is pointwise definable if every object of $\mathcal{M}$ is definable in $\mathcal{M}$ without parameters.

$$\mathcal{M} \vDash A(x) \leftrightarrow x = a$$

**Example:**

- $1$ is the unique node that is not pointed at by any node.
- $2$ is the unique node that is pointed at by a node that is not pointed at by any node.
- $\ldots$

# Pointwise Definable Structure vs Leibnizian Structure

**Problem:** Are these notions the same?

- Every pointwise definable structure 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 Structure vs Rigid Structure

## Definition (Rigid Structure)

A structure $\mathcal{M}$ is rigid iff it has no automorphisms other than the trivial automorphism $1_M$.

- Every Leibnizian structure must be rigid, because automorphisms are truth-preserving.
- There are rigid structures 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

$$\mathrm{Re}(a + b\mathrm{i}) = a \qquad \mathrm{Im}(a + b\mathrm{i}) = b$$

  then the expanded structure is Leibnizian.

# Nonrigid Structure as Reduct of Rigid structure

- 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 $v : \mathrm{Var} \to M$.*

1. *For any term $t$, $h\big(\overline{v}(t)\big) = \overline{h \circ v}(t)$*
2. *For any open formula $A$ not containing $=$, $\mathcal{M}, v \vDash A \iff \mathcal{N}, h \circ v \vDash A$*
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".*

## Definition (Elementary Equivalence)

$$\mathcal{M} \equiv \mathcal{N} \text{ iff for any sentence } A : \mathcal{M} \vDash A \iff \mathcal{N} \vDash A$$

$$\mathcal{M} \cong \mathcal{N} \implies \mathcal{M} \equiv \mathcal{N}$$

## 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 $\mathcal{N} \nvDash A_j(b_1, \ldots, b_n)$.

# Substructure

## 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 open formula $A(x_1, \ldots, x_n)$, and any $a_1, \ldots, a_n \in M$,

$$\mathcal{M} \vDash A[a_1, \ldots, a_n] \iff \mathcal{N} \vDash A[a_1, \ldots, a_n]$$

# Example

- $\mathscr{L} = \{0, 1, +, \cdot\}, \mathcal{N} = (\mathbb{N}, 0, 1, +, \cdot), \mathcal{R} = (\mathbb{R}, 0, 1, +, \cdot)$

$$\mathcal{N} \subset \mathcal{R}$$

- $\mathscr{L} = \{<\}, \mathcal{M} = (\mathbb{N}, <), \mathcal{N} = (\{2n : n \in \mathbb{N}\}, <)$

$$h : n \mapsto 2n, \quad h : \mathcal{M} \cong \mathcal{N}, \quad \text{but} \quad \mathcal{M} \not\subset \mathcal{N}$$

- $\mathscr{L} = \{0, +\}$

$$\begin{aligned} \mathcal{M} &= \left(\mathbb{N}, 0^{\mathcal{M}}, +^{\mathcal{M}}\right), && \text{where} \quad 0^{\mathcal{M}} = 0, \quad +^{\mathcal{M}}(a, b) = a + b \\ \mathcal{N} &= \left(\{2^n : n \in \mathbb{N}\}, 0^{\mathcal{N}}, +^{\mathcal{N}}\right), && \text{where} \quad 0^{\mathcal{N}} = 1, \quad +^{\mathcal{N}}(a, b) = a \cdot b \end{aligned}$$

$$h : n \mapsto 2^n, \quad h : \mathcal{M} \cong \mathcal{N} \quad \text{but} \quad \mathcal{M} \not\subset \mathcal{N}.$$

# Example

*quasipolynomial* $2^{O((\log n)^c)}$

# Example

# A Joke

"Let $G_1$ be the group $\ldots$, and $G_2$ be the group $\ldots$ 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 $\ldots$, let $G_2$ be the group $\ldots$. 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

*Let $h$ be an automorphism $h : M \to M$, and $R \subset M^n$ definable in $\mathcal{M}$. Then for any $a_1, \ldots, a_n \in M$,*

$$(a_1, \ldots, a_n) \in R \iff (h(a_1), \ldots, h(a_n)) \in R$$

**Remark:** This corollary is sometimes useful in showing that a given relation 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.

# Example

## Example

The structure $\mathcal{M} := \big(\{a, b, c\}, \{(a, b), (a, c)\}\big)$
where the language is $\mathscr{L} = \{E\}$.

$$\begin{tikzcd} b & a \arrow[l] \arrow[r] & c \end{tikzcd}$$

- $\{b, c\}$ is definable in $\mathcal{M}$: $\exists y E(y, x)$
- $\{b\}$ is not definable in $\mathcal{M}$.

## Example

Consider the vector space $\mathcal{E} := (E, +, f_r)_{r \in \mathbb{R}}$, where $E$ is the universe, $f_r$ is the scalar multiplication by $r$.

- $U := \{\mathbf{x} \in E : |\mathbf{x}| = 1\}$ 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 *Duplicator*, played on two structures $\mathcal{M}$ and $\mathcal{N}$.
Each run of the game has $n$ moves. In each move,

- *Spoiler* picks an element from $\mathcal{M}$ or from $\mathcal{N}$.
- *Duplicator* picks an element from $\mathcal{N}$ or from $\mathcal{M}$.
- *Duplicator* 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 *Duplicator* 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)

$$\mathcal{M} \sim_2 \mathcal{N} \qquad \mathcal{M} \nsim_3 \mathcal{N}$$

$$\mathcal{M} \vDash A \qquad \mathcal{N} \vDash \neg A$$

$$\forall xy \exists z (\neg Exy \to Exz \wedge Eyz)$$

# 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 wff $A(x_1, \ldots, x_n)$ 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

- $(\mathbb{N} \setminus \{0\}, \leq) \subset (\mathbb{N}, \leq) \quad (\mathbb{N} \setminus \{0\}, \leq) \cong (\mathbb{N}, \leq) \quad (\mathbb{N} \setminus \{0\}, \leq) \not\prec (\mathbb{N}, \leq)$
$A(x) := \exists y(y \leq x \wedge \neg y = x) \qquad (\mathbb{N}, \leq) \vDash A[1] \qquad (\mathbb{N} \setminus \{0\}, \leq) \nvDash A[1]$
- $(2\mathbb{Z}, <) \subset (\mathbb{Z}, <) \qquad (2\mathbb{Z}, <) \cong (\mathbb{Z}, <) \qquad (2\mathbb{Z}, <) \not\prec (\mathbb{Z}, <)$
$A(x, y) := \exists z(x < z < y) \qquad (\mathbb{Z}, <) \vDash A[0, 2] \qquad (2\mathbb{Z}, <) \nvDash A[0, 2]$

$$f : \mathcal{M} \prec \mathcal{N} \iff \exists \mathcal{A}\big(f : \mathcal{M} \cong \mathcal{A} \prec \mathcal{N}\big) \iff \exists \mathcal{A}\big(\mathcal{M} \prec \mathcal{A} \cong \mathcal{N}\big)$$

# 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 wff $A(x_1, \ldots, x_n)$, and any $a_1, \ldots, a_n \in M$,

$$f(t^{\mathcal{M}}[a_1, \ldots, a_n]) = t^{\mathcal{N}}[f(a_1), \ldots, f(a_n)]$$

$$\mathcal{M} \vDash A[a_1, \ldots, a_n] \iff \mathcal{N} \vDash A[f(a_1), \ldots, f(a_n)]$$

- $f : \mathcal{M} \prec \mathcal{N}$ iff $f : \mathcal{M} \subset \mathcal{N}$ and for any wff $\exists x A(x_1, \ldots, x_n, x)$ 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 $\mathscr{L}$-structure. $\mathcal{M}_M := (\mathcal{M}, a)_{a \in M}$ is a $\mathscr{L}_M$-structure by interpreting $c_a$ by $a$.
- Let $\mathcal{N}$ be a $\mathscr{L}$-structure, and $X \subset M \ \ \& \ \ f : X \to N$. $(\mathcal{N}, f(a))_{a \in X}$ is a $\mathscr{L}_X$-structure by interpreting $c_a$ by $f(a)$.

$\mathrm{diag}(\mathcal{M}) := \{A : A$ is an atomic or negated atomic sentence of $\mathscr{L}_M$ 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

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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*

# Klein's Erlangen Program vs Logic

## What is Logic?[16]

Logic is the science that investigates the principles of valid reasoning.

$\boxed{\text{what follows from what}}$

*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).

[16] Tarski: What are logical notions?

# Logic as permutation-invariant theory

**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

Introduction

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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 |

$$\frac{\begin{array}{c} p \oplus q \\ p \end{array}}{\neg q} \qquad \frac{\begin{array}{c} p \oplus q \\ \neg p \end{array}}{q} \qquad \frac{p}{p \oplus q}\ ?$$

# Nim Game

## 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 $c_i \coloneqq a_i \oplus b_i$

$$x \oplus (y \oplus z) = (x \oplus y) \oplus z$$

$$x \oplus y = y \oplus x$$

$$x \oplus 0 = x$$

$$x \oplus x = 0$$

$$x \oplus y = x \oplus z \implies y = z$$

# 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

$$v(\#(p,q,r)) = \left\lfloor \frac{v(p)+v(q)+v(r)}{2} \right\rfloor$$

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

| $p$ | $\neg p$ |
|---|---|
| 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

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

| $\nu(p_1),\ldots,\nu(p_n)$ | | $x_1,\ldots,x_n$ | $F_A(x_1,\ldots,x_n)$ | | $\nu(A)$ |
|---|---|---|---|---|---|
| $\nu_1(p_1),\ldots,\nu_1(p_n)$ | $=$ | $0,\ldots,0$ | $F_A(0,\ldots,0)$ | $=$ | $\nu_1(A)$ |
| $\vdots$ | | $\vdots$ | $\vdots$ | | $\vdots$ |
| $\nu_{2^n}(p_1),\ldots,\nu_{2^n}(p_n)$ | $=$ | $1,\ldots,1$ | $F_A(1,\ldots,1)$ | $=$ | $\nu_{2^n}(A)$ |

## Definition

Suppose $A$ is a wff whose atomic propositions are $p_1,\ldots,p_n$.
A truth function $F : \{0,1\}^n \to \{0,1\}$ represented by $A$ is

$$F_A(\nu(p_1),\ldots,\nu(p_n)) = \nu(A)$$

$$A \equiv B \iff F_A = F_B$$

## Theorem (Post1921)

*Every truth function $F : \{0,1\}^n \to \{0,1\}$ can be represented by some wff 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 wff 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 wff 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.

$$\neg A \equiv \bigvee_{i=1}^{m} \bigwedge_{j=1}^{n} \pm p_{ij} \implies A \equiv \neg \left( \bigvee_{i=1}^{m} \bigwedge_{j=1}^{n} \pm p_{ij} \right) \equiv \bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} \mp p_{ij}$$

# CNF Transformation

| subformula | replaced by |
|---|---|
| $A \leftrightarrow B$ | $(\neg A \vee B) \wedge (\neg B \vee A)$ |
| $A \rightarrow B$ | $\neg A \vee B$ |
| $\neg(A \wedge B)$ | $\neg A \vee \neg B$ |
| $\neg(A \vee B)$ | $\neg A \wedge \neg B$ |
| $\neg\neg A$ | $A$ |
| $(A_1 \wedge \cdots \wedge A_n) \vee B$ | $(A_1 \vee B) \wedge \cdots \wedge (A_n \vee B)$ |

# 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 wff containing only connectives from that set.

- $\{\neg, \wedge, \vee\}$
- $\{\neg, \wedge\}; \{\neg, \vee\}; \{\neg, \rightarrow\}; \{\bot, \rightarrow\}$
- $\{\uparrow\}; \{\downarrow\}$
- $\{\wedge, \vee, \rightarrow, \leftrightarrow\}; \{\neg, \leftrightarrow\}$ not adequate.

| $p$ | $\bot$ |
|---|---|
| 0 | 0 |
| 1 | 0 |

$$
\begin{aligned}
\bot &:= p \wedge \neg p \\
p \uparrow q &:= \neg(p \wedge q) \\
p \downarrow q &:= \neg(p \vee q) \\
\neg p &:= p \uparrow p \\
p \wedge q &:= (p \uparrow q) \uparrow (p \uparrow q) \\
p \vee q &:= (p \uparrow p) \uparrow (q \uparrow q)
\end{aligned}
$$

| $p$ | $q$ | $p \uparrow q$ | $p \downarrow q$ |
|---|---|---|---|
| 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 0 |
| 1 | 0 | 1 | 0 |
| 1 | 1 | 0 | 0 |

# 练习

## Exercise

令 $\ell$ 是一个三元连接词：$v(\ell(p,q,r)) = 1 \iff v(p) + v(q) + v(r) = 1$. 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

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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.
- An open formula is in conjuctive normal form (CNF) iff it is a conjuction of clause.
- An open 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

$$\text{skips} \leftarrow \text{long}$$

$$\text{reads} \leftarrow \text{short} \wedge \text{new}$$

$$\text{reads} \leftarrow \text{short} \wedge \text{follow_up} \wedge \text{known}$$

$$\text{skips} \leftarrow \text{short} \wedge \text{follow_up} \wedge \text{unknown}$$

- We want a small and efficient tree
- Ask the question which is most informative

# NNF/CNF/DNF

| subformula | replaced by |
|---|---|
| $A \leftrightarrow B$ | $(\neg A \vee B) \wedge (A \vee \neg B)$ |
| $A \rightarrow B$ | $\neg A \vee B$ |
| $\neg\neg A$ | $A$ |
| $\neg(A \vee B)$ | $\neg A \wedge \neg B$ |
| $\neg(A \wedge B)$ | $\neg A \vee \neg B$ |
| $\neg\forall x A$ | $\exists x \neg A$ |
| $\neg\exists x A$ | $\forall x \neg A$ |

| 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 open formula can be equivalently transformed into CNF/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 $x \notin \mathrm{Fv}(A)$ |
| $A \rightarrow \forall x B$ | $\forall x (A \rightarrow B)$ | where $x \notin \mathrm{Fv}(A)$ |
| $A \rightarrow \exists x B$ | $\exists x (A \rightarrow B)$ | where $x \notin \mathrm{Fv}(A)$ |
| $\forall x A \rightarrow B$ | $\exists x (A \rightarrow B)$ | where $x \notin \mathrm{Fv}(A)$ |
| $\exists x A \rightarrow B$ | $\forall x (A \rightarrow B)$ | 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* $A^{\mathrm{SNF}}$ without existential quantifiers.

## Theorem (Skolem Normal Form)

*Any formula $A$ can be equisatisfiably transformed into a SNF $A^{\mathrm{SNF}}$, and we have $A^{\mathrm{SNF}} \vDash A$.*

**Remark:** $A^{\mathrm{SNF}} \vDash A$ 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

$$\boxed{A \text{ is satisfiable iff } A^{\mathrm{SNF}} \text{ is satisfiable.}}$$

$$A^{\mathrm{HNF}} := \left(\neg(\neg A)^{\mathrm{SNF}}\right)^{\mathrm{NNF}}$$

$$\boxed{\vDash A \iff \vDash A^{\mathrm{HNF}}}$$

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 ground terms of $\mathscr{L}$. If no constant in $\mathscr{L}$, then add a new constant to $\mathscr{L}$.

# Herbrand Structure

## Definition (Herbrand Structure)

A Herbrand structure for $\mathscr{L}$ is $(H, I)$ s.t.

- $H$ is the Herbrand universe of $\mathscr{L}$.
- for every ground term $t$, $I(t) = t$.

## Theorem

*A formula $A$ is satisfiable iff there is a Herbrand structure satisfying it.*

## Proof.

Assume $A$ is in Skolem normal form, and it is satisfied by some structure $(M, I)$. Then Herbrand structure $(H_A, J) \vDash A$, where for each ground 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 wff $A(x_1, \ldots, x_n)$, the Herbrand expansion over a set $D$ of ground 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 $A^{\mathrm{HNF}}$ 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 wff $A$, $\mathcal{R}(A)$ 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. $\forall x(\forall y(\mathrm{Animal}(y) \to \mathrm{Loves}(x, y)) \to \exists z\, \mathrm{Loves}(z, x))$
2. $\forall x((\exists z\, \mathrm{Animal}(z) \land \mathrm{Kills}(x, z)) \to \forall y \neg\, \mathrm{Loves}(y, x))$
3. $\forall x(\mathrm{Animal}(x) \to \mathrm{Loves}(\mathsf{Jack}, x))$
4. $\mathrm{Kills}(\mathsf{Jack}, \mathsf{Tuna}) \lor \mathrm{Kills}(\mathsf{Curiosity}, \mathsf{Tuna})$
5. $\mathrm{Cat}(\mathsf{Tuna})$
6. $\forall x(\mathrm{Cat}(x) \to \mathrm{Animal}(x))$
7. $\neg\, \mathrm{Kills}(\mathsf{Curiosity}, \mathsf{Tuna})$

# Resolution — Example

1. $\Big(\text{Animal}(F(x)) \vee \text{Loves}(G(x), x)\Big) \wedge \Big(\neg\,\text{Loves}(x, F(x)) \vee \text{Loves}(G(x), x)\Big)$
2. $\neg\,\text{Loves}(y, x) \vee \neg\,\text{Animal}(z) \vee \neg\,\text{Kills}(x, z)$
3. $\neg\,\text{Animal}(x) \vee \text{Loves}(\text{Jack}, x)$
4. $\text{Kills}(\text{Jack}, \text{Tuna}) \vee \text{Kills}(\text{Curiosity}, \text{Tuna})$
5. $\text{Cat}(\text{Tuna})$
6. $\neg\,\text{Cat}(x) \vee \text{Animal}(x)$
7. $\neg\,\text{Kills}(\text{Curiosity}, \text{Tuna})$

# 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

$$\frac{\forall x A}{A[t/x]} \text{ where } t \text{ is a ground term}$$

  - 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 $\infty$ 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. $\quad$ A $\leftrightarrow \neg B$

B. The next statement is false. $\quad$ B $\leftrightarrow \neg C$

C. The first statement is false. $\quad$ C $\leftrightarrow \neg A$

- The above odd cycle is a generalization of the liar paradox $\bullet \circlearrowleft$
- 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 $\Sigma$, 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 $\Delta$ 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 $\nu$ 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 $\Delta$ 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

Introduction

Propositional Logic

Term Logic

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

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Model & Theory & Logical Consequence

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

- $\Gamma \subset \Gamma' \implies \mathrm{Mod}(\Gamma') \subset \mathrm{Mod}(\Gamma)$
- $\mathcal{K} \subset \mathcal{K}' \implies \mathrm{Th}(\mathcal{K}') \subset \mathrm{Th}(\mathcal{K})$
- $\Gamma \subset \mathrm{Th}(\mathrm{Mod}(\Gamma))$
- $\mathcal{K} \subset \mathrm{Mod}(\mathrm{Th}(\mathcal{K}))$
- $\mathrm{Mod}(\Gamma) = \mathrm{Mod}(\mathrm{Th}(\mathrm{Mod}(\Gamma)))$
- $\mathrm{Th}(\mathcal{K}) = \mathrm{Th}(\mathrm{Mod}(\mathrm{Th}(\mathcal{K})))$
- $\mathrm{Cn}(\Gamma) = \mathrm{Th}(\mathrm{Mod}(\Gamma))$
- $\Gamma \subset \Gamma' \implies \mathrm{Cn}(\Gamma) \subset \mathrm{Cn}(\Gamma')$
- $\mathrm{Cn}(\mathrm{Cn}(\Gamma)) = \mathrm{Cn}(\Gamma)$

# Theory & Axiomatization

- A set $\Gamma$ of sentences is a theory iff $\Gamma = \mathrm{Cn}(\Gamma)$.
- A theory $\Gamma$ is complete iff for every sentence $A$, either $A \in \Gamma$ or $\neg A \in \Gamma$.
- A theory $\Gamma$ is finitely axiomatizable iff $\Gamma = \mathrm{Cn}(\Sigma)$ for some finite set $\Sigma$ of sentences.
- A theory $\Gamma$ is axiomatizable iff there is a decidable set $\Sigma$ of sentences s.t. $\Gamma = \mathrm{Cn}(\Sigma)$.
- A class $\mathcal{K}$ of structures is an elementary class $(EC)$ iff $\mathcal{K} = \mathrm{Mod}(A)$ for some sentence $A$.
- A class $\mathcal{K}$ of structures is an elementary class in wider sense $(EC_\Delta)$ iff $\mathcal{K} = \mathrm{Mod}(\Sigma)$ for some set $\Sigma$ of sentences.

# Definitional Extension

## Definition (Definitional Extension)

$\mathscr{L}$ is a first order language, and $\mathrm{T}$ 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 $\mathrm{T}$ to $\mathscr{L}^+$ is a $\mathrm{T}^+$-theory

$$\mathrm{T}^+ = \mathrm{T} \cup \{\delta_s : s \in \mathscr{L}^+ \setminus \mathscr{L}\}$$

where the sentence $\delta_s$ is an explicit definition of $s$ in $\mathscr{L}$.

# Definitionally Equivalent

Let $\mathrm{T}^+$ be a definitional extension of $\mathrm{T}$. 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 $\delta_s$ is the explicit definition of $s$. For $s \in \mathscr{L}$, $Rs = s$.

## Theorem

*If $\mathrm{T}^+$ is a definitional extension of $\mathrm{T}$, then*

- $\mathrm{T} \vdash A \leftrightarrow RIA$ *for any $\mathscr{L}$-formula $A$.*
- $\mathrm{T}^+ \vdash A \leftrightarrow IRA$ *for any $\mathscr{L}^+$-formula $A$.*
- $\mathrm{T}^+$ *is a conservative extension of* $\mathrm{T}$.

## Definition (Definitionally Equivalent)

Let $\mathrm{T}_1$ be a $\mathscr{L}_1$-theory and $\mathrm{T}_2$ be a $\mathscr{L}_2$-theory. Then $\mathrm{T}_1$ and $\mathrm{T}_2$ are said to be *definitionally equivalent* if there is a definitional extension $\mathrm{T}_1^+$ of $\mathrm{T}_1$ to $\mathscr{L}_1 \cup \mathscr{L}_2$ and a definitional extension $\mathrm{T}_2^+$ of $\mathrm{T}_2$ to $\mathscr{L}_1 \cup \mathscr{L}_2$ such that $\mathrm{Cn}(\mathrm{T}_1^+) = \mathrm{Cn}(\mathrm{T}_2^+)$.

# Consistency & Satisfiability

- $\Gamma$ is consistent iff $\Gamma \nvdash \bot$.
- $\Gamma$ is maximal iff for every wff $A$, either $A \in \Gamma$ or $\neg A \in \Gamma$.
- $\Gamma$ is maximal consistent iff it is both consistent and maximal.
- $\Gamma$ is satisfiable iff $\mathrm{Mod}(\Gamma) \neq \emptyset$.
- $\Gamma$ is finitely satisfiable iff every finite subset of $\Gamma$ is satisfiable.

# Soundness Theorem

**Theorem (Soundness Theorem)**

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

**Proof.**

by induction on derivation lengths.

*Truth in a model is preserved under making deductions.*

# Kurt Gödel 1906-1978

"I am unprovable."[17]

- Completeness.
  I think (consistently), therefore I am.
  (Consistency implies existence.)
- Incompleteness.
  1. provable < true
  2. un-self-aware
- Consistency of AC and CH.

[17] 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 wffs is satisfiable.*

$$\begin{array}{ccc}
\Gamma \vDash A & \iff & \Gamma \vdash A \\
\Updownarrow & & \Updownarrow \\
\Gamma \cup \{\neg A\} & & \Gamma \cup \{\neg A\} \\
\text{unsatisfiable} & \iff & \text{inconsistent}
\end{array}$$

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

# Proof of Completeness Theorem — step1

## Lemma (Lindenbaum Lemma)

*Any consistent set $\Theta$ of sentences can be extended to a maximal consistent set $\Delta$ 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 $\Delta$ is Henkin iff $\Delta$ 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 $\Gamma$ is consistent, then there is a Henkin set $\Delta \supset \Gamma$.*

## Proof.

Let $\mathcal{C}$ be a set of new constants. $\mathscr{L}^+ := \mathscr{L} \cup \mathcal{C}$, $\mathscr{L} \cap \mathcal{C} = \emptyset$, $|\mathcal{C}| = |\mathscr{L}|$. Assume $|\mathcal{C}| = \kappa$, and $\mathcal{C} = \{c_\xi : \xi < \kappa\}$. Arrange all formulas of $\mathscr{L}^+$ with at most one free variable in a sequence $\langle A_\xi : \xi < \kappa \rangle$. Let

$$\begin{aligned} \Gamma_0 &:= \Gamma \\ \Gamma_{\xi+1} &:= \Gamma_\xi \cup \{\exists x A_\xi(x) \to A_\xi[c_\beta/x]\} \\ \Gamma_\xi &:= \bigcup_{\alpha<\xi} \Gamma_\alpha \quad \text{if } \xi \text{ is a limit ordinal} \end{aligned}$$

where $c_\beta$ is the first new constant not occurring in $\Gamma_\xi \cup \{A_\xi\}$.

Then $\Theta := \bigcup_{\xi<\kappa} \Gamma_\xi$ is consistent, and we can extend $\Theta$ to a maximal consistent set $\Delta \supset \Theta$ by Lindenbaum lemma.

# Proof of Completeness Theorem — step3

## Lemma (Term Models Lemma)

*If $\Delta$ 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 $\mathrm{Term}$ 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

$$
\begin{tikzcd}
\mathrm{Str} \arrow[r, leftrightarrow, squiggly, "\models"] \arrow[d, "\mathcal{M}\mapsto\mathrm{Th}(\mathcal{M})"'] & \mathrm{Sen} \arrow[d, "A\mapsto[A]"] \\
\mathrm{S}_0 \arrow[r, leftrightarrow, squiggly] \arrow[d, "\mathrm{T}\mapsto\mathrm{T}/\sim"'] \arrow[ur, leftrightarrow, squiggly, "\ni"'] & \mathrm{Lin}_0 \arrow[ur, leftrightarrow, squiggly, "\ni"'] \\
\mathrm{Sto}(\mathrm{Lin}_0) &
\end{tikzcd}
$$

- $\mathrm{Str}$: the class of structures.
- $\mathrm{Sen}$: the set of sentences.
- $\mathrm{S}_0$: the set of complete theories.
- $\mathrm{Lin}_0 = \mathrm{Sen}/\sim$
- $\mathrm{Sto}(M)$: the set of ultrafilters of $M$.

1. The closed subsets of $\mathrm{Sen}$ are precisely the theories.
2. The closed subsets of $\mathrm{S}_0$ 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, $\mathrm{S}_0$ is a Kolmogorov quotient $\mathrm{Str}/\equiv$.
4. For every theory $\mathrm{T}$, $\mathrm{T}/\sim$ is a filter of $\mathrm{Lin}_0$.
5. For every complete theory $\mathrm{T}$, $\mathrm{T}/\sim$ is an ultrafilter of $\mathrm{Lin}_0$.
6. $\mathrm{T} \mapsto \mathrm{T}/\sim : \mathrm{S}_0 \rightarrowtail \mathrm{Sto}(\mathrm{Lin}_0)$ is a homomorphism.
7. The image is dense in $\mathrm{Sto}(\mathrm{Lin}_0)$.
8. The image is a closed subspace of $\mathrm{Sto}(\mathrm{Lin}_0)$.
9. $\mathrm{T} \mapsto \mathrm{T}/\sim : \mathrm{S}_0 \twoheadrightarrow \mathrm{Sto}(\mathrm{Lin}_0)$ iff the topology on $\mathrm{S}_0$ is compact.

# Compactness Theorem

## Theorem (Compactness Theorem)

*A set of wffs 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 structure* $\mathcal{M} \equiv \mathcal{N}$ *but* $\mathcal{M} \not\cong \mathcal{N}$*, where* $\mathcal{N} = (\mathbb{N}, 0, S, +, \cdot)$ *is the standard model of arithmetic.*

## Proof.

By compactness theorem, $\mathrm{Th}(\mathcal{N}) \cup \{c \neq 0, c \neq S0, c \neq SS0, \ldots\}$ has a model $\mathcal{M}$.

# Ultraproduct & Łoś Theorem

## Definition (Ultraproduct)

Suppose $\{\mathcal{M}_i : i \in I\}$ is a set of structures, 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)

$$\prod_{i \in I} \mathcal{M}_i / U \vDash A\left([f_1], \ldots, [f_n]\right) \iff \left\{ i \in I : \mathcal{M}_i \vDash A\big(f_1(i), \ldots, f_n(i)\big) \right\} \in U$$

# 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 $\Gamma$ of wffs 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 $A \in \Gamma$,

$$\begin{aligned} \hat{A} \in U \;\; \& \;\; \hat{A} \subset \{\Delta \in I : \mathcal{M}_\Delta \vDash A\} &\implies \{\Delta \in I : \mathcal{M}_\Delta \vDash A\} \in U \\ &\implies \prod_{\Delta \in I} \mathcal{M}_\Delta / U \;\vDash A \end{aligned}$$

# 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

$$\mathbb{R} \implies \mathbb{R} \cup \{-\infty, +\infty\}$$

$$x \mapsto \left( \frac{x}{1+x^2}, \frac{x^2}{1+x^2} \right)$$

# Nonstandard Analysis

**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

# Nonstandard Analysis

Let $U$ be a nonprinciple ultrafilter on $\mathbb{N}$.

$$\prod_{i \in \mathbb{N}} \mathcal{R}/U \vDash \mathrm{Th}(\mathcal{R})$$

Let $\varepsilon := \left[(1, \frac{1}{2}, \frac{1}{3}, \dots)\right] \in \mathbb{R}$.
For any $n \in \mathbb{N}$,

$$\prod_{i \in \mathbb{N}} \mathcal{R}/U \vDash \varepsilon < \frac{1}{n}$$

**Theorem**

- *$\mathcal{K}$ is $EC_\Delta$ 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

## Theorem

*If a set $\Gamma$ of sentences has arbitrarily large finite models, then it has an infinite model.*

## Proof.

Consider the set $\Gamma \cup \left\{\varphi^{\geq n} : n \in \mathbb{N}\right\}$, where $\varphi^{\geq n} = \exists x_1 \ldots x_n \bigwedge_{1 \leq i < j \leq n} x_i \neq x_j$.

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 structures is not $EC_\Delta$. (Model finiteness is undefinable even by a set of formulas.)
- The class of all infinite structures is $EC_\Delta$ 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 $EC_\Delta$ 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 $EC_\Delta$.

# 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.*

$$\text{Infinite Ramsey Theorem} \xrightarrow{\text{Compactness Theorem}} \text{Finite Ramsey Theorem}$$

## 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*

$$\int_A \tau_A(x)\, d\mu(x) = 1$$

*Equivalently, the mean recurrence time $\langle \tau_A \rangle := \frac{1}{\mu(A)} \int_A \tau_A(x)\, d\mu(x) = \frac{1}{\mu(A)}$.*

# 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 $x$.

# Löwenheim-Skolem Theorem

## Theorem (Downward Löwenheim-Skolem Theorem)

*A consistent set of sentences in a language of cardinality $\lambda$ has a model of cardinality $\leq \lambda$.*

## Theorem (Upward Löwenheim-Skolem Theorem)

*If a set of sentences in a language of cardinality $\lambda$ has an infinite model, then it has models of every cardinality $\geq \lambda$.*

## Proof.

Add $\{c_\xi\}_{\xi<\lambda}$ to $\mathscr{L}$. Let $T' := T \cup \{c_\xi \neq c_\eta : \xi < \eta < \lambda\}$.
Every finite subset $\Sigma \subset T'$ will involve at most a finite number of $c_\xi$.
Hence any infinite model of $T$ can be expanded to a model of $\Sigma$.
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 $c_\xi$ in $\mathcal{M}$ must give different elements. So $\lambda \leq |\mathcal{M}| \leq \lambda$.

# Skolem Paradox? — Models and Reality

- Cantor: $\mathrm{P}(\mathbb{N})$ is uncountable.
- There is a countable model $\mathcal{M} \vDash \mathrm{ZF} \vdash$"$\mathrm{P}(\mathbb{N})$ is uncountable".
- The statement "$\mathrm{P}(\mathbb{N})$ is uncountable" is interpreted in $\mathcal{M}$ as — within $\mathcal{M}$, there is a set $M_1$ that looks like $\mathrm{P}(\mathbb{N})$ 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 $\mathrm{P}(\mathbb{N})$ 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 $\mathrm{ZF}$.

# 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 wff $A(x_1, \ldots, x_n)$ 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 $c_1, \ldots, c_n$ 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 $\mathrm{T}_i$ be a theory in $\mathscr{L}_i$ for $i = 1, 2$. If $\mathrm{T}_1$ and $\mathrm{T}_2$ 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 $\tau$ a set $\mathcal{S}(\tau)$ of sentences, and $\vDash_{\mathcal{L}}$ is a relation between $\tau$-structures 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 $\tau'$-structure, then $\mathcal{M} \vDash_{\mathcal{L}} A \iff \mathcal{M}\restriction_{\tau} \vDash_{\mathcal{L}} A$

$$\mathrm{Mod}^{\tau}_{\mathcal{L}}(A) := \left\{\mathcal{M} \in \tau\text{-structures} : \mathcal{M} \vDash_{\mathcal{L}} A\right\}$$

# 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. $\mathcal{M} \vDash_{\mathcal{L}} B \iff \mathcal{M} \nvDash_{\mathcal{L}} A$; and $\forall A, B \in \mathcal{S}(\tau)\ \exists C \in \mathcal{S}(\tau) : \mathcal{M} \vDash_{\mathcal{L}} C \iff \mathcal{M} \vDash_{\mathcal{L}} A \ \&\ \mathcal{M} \vDash_{\mathcal{L}} B$.
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 $\tau$ assigning $a$ to $c$.
3. (*Renaming*) Let $\pi : \tau \to \tau'$ be a bijection which respects arity, and we extend $\pi$ in a canonical way to $\hat{\pi} : \tau$-structures $\to \tau'$-structures. Then

$$\forall A \in \mathcal{S}(\tau)\ \exists A' \in \mathcal{S}(\tau) : \mathcal{M} \vDash_{\mathcal{L}} A \iff \hat{\pi}(\mathcal{M}) \vDash_{\mathcal{L}} A'$$

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 $\mathcal{M} : (\mathcal{M}, X, b_1, \ldots, b_n) \vDash_{\mathcal{L}} B \iff \mathcal{N} \vDash_{\mathcal{L}} A$ 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 $\tau$ and $A \in \mathcal{S}_1(\tau)$ there is $B \in \mathcal{S}_2(\tau)$ s.t.

$$\mathrm{Mod}^{\tau}_{\mathcal{L}_1}(A) = \mathrm{Mod}^{\tau}_{\mathcal{L}_2}(B)$$

$$\mathcal{L}_1 \sim \mathcal{L}_2 \coloneqq \mathcal{L}_1 \leq \mathcal{L}_2 \ \& \ \mathcal{L}_2 \leq \mathcal{L}_1$$

## 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 $\top, \bot, \neg, \wedge, \vee, \rightarrow, \exists, \forall$.
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 $\top, \bot, \neg, \wedge, \vee, \rightarrow, \exists, \forall$ and infinitary conjunction and infinitary disjunction.

- $\mathrm{T}$ 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.
- $\mathrm{T}$ is a Horn (resp. regular, coherent, geometric) theory iff all the sequents in $\mathrm{T}$ are Horn (resp. regular, coherent, geometric).
- $\mathrm{T}$ 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 $\bot$.
- $\mathrm{T}$ 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 $A \vdash_{\mathbf{x}} \top \qquad A \wedge B \vdash_{\mathbf{x}} A \qquad A \wedge B \vdash_{\mathbf{x}} B \qquad \dfrac{C \vdash_{\mathbf{x}} A \quad C \vdash_{\mathbf{x}} B}{C \vdash_{\mathbf{x}} A \wedge B}$

Disjunction $\bot \vdash_{\mathbf{x}} A \qquad A \vdash_{\mathbf{x}} A \vee B \qquad B \vdash_{\mathbf{x}} A \vee B \qquad \dfrac{A \vdash_{\mathbf{x}} C \quad B \vdash_{\mathbf{x}} C}{A \vee B \vdash_{\mathbf{x}} C}$

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 $\neg a \coloneqq a \rightarrow \bot$.

$$\frac{\text{Intuitionistic}}{\text{Classical}} \approx \frac{\text{Logic of Knowledge}}{\text{Logic of Being}}$$

*Intuitionistic truth* of $A$: knowing how to **prove/solve** $A$

- There is no proof/solution of $\perp$.
- 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

$$(X_1, Y_1) \wedge (X_2, Y_2) \coloneqq (X_1 \times X_2, Y_1 \times Y_2)$$

$$(X_1, Y_1) \vee (X_2, Y_2) \coloneqq \left(X_1 \coprod X_2, Y_1 \coprod Y_2\right)$$

$$(X_1, Y_1) \to (X_2, Y_2) \coloneqq \left(X_2^{X_1}, \left\{f \in X_2^{X_1} : f(Y_1) \subset Y_2\right\}\right)$$

- An interpretation is a map $j$ : atomic propositions $\longrightarrow$ problems. e.g. $j(p) = (X, Y)$, and we denote $j_1(p) = X, j_2(p) = Y$. Specifically, $j(\bot) \coloneqq (\{\emptyset\}, \emptyset)$.
- Two interpretations $j, j'$ are equivalent $j \sim j'$ iff $j_1(p) = j_1'(p)$ for any atomic $p$.
- A wff $A$ 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 wff $A$ 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 $n \in \mathbb{N}$ |
| $\langle m, n \rangle \Vdash x = y$ | if $[\![x]\!] = [\![y]\!]$ and $m \Vdash x$ and $n \Vdash y$ |
| $\langle m, n \rangle \Vdash A \wedge B$ | 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 a valuation 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 $\varepsilon$-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 $f : X \to Y$ between topological spaces is *continuous* iff the inverse image $f^{-1}(V)$ of any open subset $V \subset Y$ is an open subset of $X$.
- 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 $f(U) \subset V$.

$$\begin{tikzcd} U \arrow[r] \arrow[d, "f"'] & f^{-1}(V) \\ f(U) \arrow[r] & V \arrow[u, "f^{-1}"'] \end{tikzcd}$$

# 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 $f^{-1}(U)$ is open in the topology on $X$ for every open subset $U$ 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
$$A \to B := \big((X \setminus A) \cup B\big)^\circ = \bigcup \{U \in \mathcal{O}_X : U \cap A \subset B\}$$
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
$$\neg A = (X \setminus A)^\circ \quad \text{and} \quad A \vee \neg A = A \cup (X \setminus A)^\circ = X \setminus \partial A$$
- $A := (0,1) \cup (1,\infty), \neg A = (-\infty, 0), \neg\neg A = (0,\infty), A \cup \neg A \neq \mathbb{R}, A \subsetneq \neg\neg A$.
- A topological model of intuitionistic propositional logic is $(X, \mathcal{O}_X, [\![\,]\!])$ where $[\![\,]\!] : \mathrm{Var} \to \mathcal{O}_X$.

# Topological Semantics of Intuitionistic Logic

$$
\begin{aligned}
[\![A \wedge B]\!] &:= [\![A]\!] \cap [\![B]\!] \\
[\![A \vee B]\!] &:= [\![A]\!] \cup [\![B]\!] \\
[\![A \to B]\!] &:= \Big( (X \setminus [\![A]\!]) \cup [\![B]\!] \Big)^\circ \\
[\![\bot]\!] &:= \emptyset \\
[\![\neg A]\!] &:= (X \setminus [\![A]\!])^\circ \\
[\![\exists x A]\!] &:= \bigcup_{a \in M} [\![A[a]]\!] \\
[\![\forall x A]\!] &:= \Big( \bigcap_{a \in M} [\![A[a]]\!] \Big)^\circ \\
\mathcal{O}_X \Vdash A &:= [\![A]\!] = X \\
\Gamma \Vdash A &:= \Big( \bigcap_{B \in \Gamma} [\![B]\!] \Big)^\circ \subset [\![A]\!]
\end{aligned}
$$

**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".

## Theorem

*The axiom of choice implies excluded middle.*

## Proof.

Consider an arbitrary proposition $P$. To decide $P$, define

$$A := \{x \in \{0,1\} : P \vee x = 0\}$$
$$B := \{x \in \{0,1\} : P \vee x = 1\}$$

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 structure. The categorical characterizations use second-order logic.
- Categoricity is central to the philosophy of structuralism.

# $\kappa$-Categoricity

**Definition (categoricity / $\kappa$-categoricity)**

- A theory is *categorical* iff it has exactly one model.
- A theory is *$\kappa$-categorical* iff it has exactly one model of cardinality $\kappa$.

**Theorem**

*A theory $\mathrm{T}$ is complete iff for all $\mathcal{M} \vDash \mathrm{T}$ and $\mathcal{N} \vDash \mathrm{T}$: $\mathcal{M} \equiv \mathcal{N}$.*

**Theorem**

*If a theory is categorical, then it is complete.*

**Theorem**

*For a theory with finite models, it is complete iff it is categorical.*

# Dense Linear Ordering without Endpoints

1. $x \not< x$
2. $x < y \rightarrow y < z \rightarrow x < z$
3. $x < y \vee x = y \vee y < x$
4. $x < y \rightarrow \exists z(x < z < y)$
5. $\exists yz(y < x < z)$

## Definition ($\kappa$-categoricity)

A theory is $\kappa$-categorical iff it has a unique model of cardinality $\kappa$.

## Theorem (Cantor)

- *The theory of dense linear orderings without endpoints is $\aleph_0$-categorical.*
- $(\mathbb{R}, <)$ *is the unique complete linear ordering that has a countable dense subset isomorphic to* $(\mathbb{Q}, <)$.

### Theorem (Łoś-Vaught Test)

*If a theory with no finite model is $\kappa$-categorical, then it is complete.*

### Theorem

*The theory $\mathrm{ACF}_p$ of algebraically closed fields of characteristic $p$ (for $p$ prime or $0$) is $\kappa$-categorical for all uncountable cardinals $\kappa$.*

### Corollary

*For $p \in \mathbb{P}$ or $p = 0$, $\mathrm{ACF}_p$ is complete and decidable.*

### Theorem (Morley's Categoricity Theorem)

*If a theory is $\kappa$-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 $\kappa$-categorical for any uncountable $\kappa$.
- The field of complex numbers is $\kappa$-categorical for any uncountable $\kappa$.
- 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 $\mathrm{ACF}_0$.
$(2 \to 3)$ 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 $\mathrm{ACF}_p$, we only need to show $A$ is true in *some* model of $\mathrm{ACF}_p$.
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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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!

# Why Study Modal Logic?

- Modal languages are simple yet expressive languages for talking about relational structures.
- Modal languages provide an internal, local perspective on relational structures.
- Modal languages are not isolated formal systems.
  - Modal vs classical (FOL,SOL), internal vs external perspective.
    In FOL, structures are described from the top point of view. Each object and relation can be named. In modal logic, relational structures are described from an internal perspective, there is no way to mention objects and relations.
  - Relational structures vs Boolean algebra with operators.
    (Jónsson and Tarski's representation theorem.)
- Decidability.
  (seeking a balance between expressiveness and efficiency/complexity)

expressiveness
complexity

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Logics about Modalities

Alice _____ married.

- is possibly (basic modal logic)
- will be (temporal logic)
- is permitted to be (deontic logic)
- is known/believed (to A) 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 $\mathrm{Wff}$)

$$A ::= p \mid \neg A \mid A \wedge A \mid A \vee A \mid A \rightarrow A \mid A \leftrightarrow A \mid \Box A \mid \Diamond A$$

- 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 $\mathrm{T}$. $\Box_{\mathrm{T}} A$
- After the execution of the program $\alpha$, $A$ holds. $[\alpha]A$

# Examples

1. "Ought" implies "can", but it does not imply "is".
$$(\Box p \to \Diamond p) \land \neg(\Box p \to p)$$
2. What must be is, and what is, is possible.
$$(\Box p \to p) \land (p \to \Diamond p)$$
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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Kripke Possible World Semantics[18]

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 $V : \mathrm{Var} \to \mathrm{P}(W)$.

- $\mathcal{M}, w \Vdash p$ 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 $\exists v \in W \Big( Rwv \ \ \& \ \ \mathcal{M}, v \Vdash A \Big)$

[18] 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:

- $[\![p]\!] := V(p)$
- $[\![\neg A]\!] := W \setminus [\![A]\!]$
- $[\![A \wedge B]\!] := [\![A]\!] \cap [\![B]\!]$
- $[\![A \vee B]\!] := [\![A]\!] \cup [\![B]\!]$
- $[\![A \to B]\!] := (W \setminus [\![A]\!]) \cup [\![B]\!]$
- $[\![\Box A]\!] := \Box_R [\![A]\!]$ where $\Box_R X := \{w \in W : R(w) \subset X\}$
- $[\![\Diamond A]\!] := \Diamond_R [\![A]\!]$ where $\Diamond_R X := \{w \in W : R(w) \cap X \neq \emptyset\}$

$A$ $\qquad$ $\Box A$ $\qquad$ $\Diamond A$

$$\mathcal{M}, w \Vdash A \iff w \in [\![A]\!]$$

# 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 $\mathcal{F}$.
- $\Vdash A$ iff $\mathcal{F} \Vdash A$ for every $\mathcal{F}$.

*Truth is in the eye of the beholder.*

## Example

$$\Vdash \Box(A \to B) \to \Box A \to \Box B$$

# Entailment

- local entailment

$$\Gamma \Vdash_{\mathcal{C}} A \;:=\; \forall \mathcal{M} \in \mathcal{C} \forall w \in W \Big( \mathcal{M}, w \Vdash \Gamma \implies \mathcal{M}, w \Vdash A \Big)$$

- global entailment

$$\Gamma \Vdash^{g}_{\mathcal{C}} A \;:=\; \forall \mathcal{M} \in \mathcal{C} \Big( \mathcal{M} \Vdash \Gamma \implies \mathcal{M} \Vdash A \Big)$$

## Example

- $p \nVdash_{\mathcal{C}} \Box p$
- $p \Vdash^{g}_{\mathcal{C}} \Box p$

# Material Implication vs C. I. Lewis' Strict Implication

$$p \mathrel{-\!\!\!\supset} q \coloneqq \Box(p \to q)$$

- $\nVdash A \mathrel{-\!\!\!\supset} B \mathrel{-\!\!\!\supset} A$ ?
- $\nVdash (A \mathrel{-\!\!\!\supset} B) \vee (B \mathrel{-\!\!\!\supset} C)$ ?
- $\nVdash \neg(A \mathrel{-\!\!\!\supset} B) \to (A \wedge \neg B)$ ?
- $\Vdash (A \wedge \neg A) \mathrel{-\!\!\!\supset} B$
- $\Vdash A \mathrel{-\!\!\!\supset} (B \vee \neg B)$
- $\neg\Diamond A \Vdash A \mathrel{-\!\!\!\supset} B$
- $\Box A \Vdash B \mathrel{-\!\!\!\supset} A$

# 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 $W, R \nVdash 5$.
Then $w \Vdash \Diamond p$ but $w \nVdash \Box\Diamond p$.
Now $w \Vdash \Diamond p \implies \exists v : Rwv \;\&\; v \Vdash p$,
and $w \nVdash \Box\Diamond p \implies w \Vdash \Diamond\neg\Diamond p \implies \exists u :$
$Rwu \;\&\; u \Vdash \neg\Diamond p$.
However, $Ruv \implies u \Vdash \Diamond p$.
So we have $Rwv \wedge Rwu$ but $\neg Ruv$.

# Proof ( $\Longrightarrow$ ): Counter-model for $D, T, B, 4, 5$

# Standard Translation

## Definition (Standard Translation)

$$T_x(p) = P(x) \qquad T_y(p) = P(y)$$

$$T_x(\neg A) = \neg T_x(A) \qquad T_y(\neg A) = \neg T_y(A)$$

$$T_x(A \wedge B) = T_x(A) \wedge T_x(B) \qquad T_y(A \wedge B) = T_y(A) \wedge T_y(B)$$

$$T_x(\Box A) = \forall y(Rxy \to T_y(A)) \qquad T_y(\Box A) = \forall x(Ryx \to T_x(A))$$

## 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Tree Method for Modal Logic

$$\begin{array}{c} w \Vdash \Box A \\ | \\ v \Vdash A \end{array}$$

if $Rwv$ is already in the branch.

$$\begin{array}{c} w \nVdash \Box A \\ | \\ Rwv \\ v \nVdash A \end{array}$$

where $v$ is new in the branch.

$$\begin{array}{c} w \Vdash \Diamond A \\ | \\ Rwv \\ v \Vdash A \end{array}$$

where $v$ is new in the branch.

$$\begin{array}{c} w \nVdash \Diamond A \\ | \\ v \nVdash A \end{array}$$

if $Rwv$ is already in the branch.

# Example — Tree Method for Modal Logic

$$\boxed{\Vdash \Box p \wedge \Diamond q \rightarrow \Diamond(p \wedge q)}$$

$$w \Vdash \Box p \wedge \Diamond q$$
$$w \Vdash \neg\Diamond(p \wedge q)$$

$$w \Vdash \Box p$$
$$w \Vdash \Diamond q$$

$$Rwv$$
$$v \Vdash q$$

$$v \Vdash p$$

$$v \Vdash \neg(p \wedge q)$$

$$v \Vdash \neg p \qquad\qquad v \Vdash \neg q$$
$$\times \qquad\qquad\qquad \times$$

# 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

$$\frac{A \quad A \to B}{B}\ [\text{MP}] \qquad\qquad \frac{A}{\Box A}\ [\text{N}]$$

# Intuitionistic Logic vs Modal Logic

$$\mathrm{S4} := K + T + 4$$

$$\mathrm{Grz} := \mathrm{S4} + \Box(\Box(A \to \Box A) \to A) \to A$$

$$\begin{aligned} p^* &:= \Box p \\ (\neg A)^* &:= \Box \neg A^* \\ (A \wedge B)^* &:= A^* \wedge B^* \\ (A \vee B)^* &:= A^* \vee B^* \\ (A \to B)^* &:= \Box(A^* \to B^*) \end{aligned}$$

$$\vdash_{\mathrm{I}} A \iff \vdash_{\mathrm{S4}} A^* \iff \vdash_{\mathrm{Grz}} A^*$$

$$\mathrm{GL} := K + L$$

$$\begin{aligned} p' &:= p \\ (\neg A)' &:= \neg A' \\ (A \wedge B)' &:= A' \wedge B' \\ (A \vee B)' &:= A' \vee B' \\ (A \to B)' &:= A' \to B' \\ (\Box A)' &:= A' \wedge \Box A' \end{aligned}$$

$$\vdash_{\mathrm{Grz}} A \iff \vdash_{\mathrm{GL}} A'$$

$$\vdash_{\mathrm{I}} A \iff \vdash_{\mathrm{GL}} (A^*)'$$

$$\boxed{\Box(\Box A \to A) \to \Box A \vdash_{\mathrm{GL}} \Box A \to \Box\Box A}$$

- $A \wedge \Box A \wedge \Box\Box A \to A \wedge \Box A$
- $A \to \Box(A \wedge \Box A) \to A \wedge \Box A$ $\qquad \Box(p \wedge q) \leftrightarrow \Box p \wedge \Box q$
- $\Box(\Box(A \wedge \Box A) \to A \wedge \Box A) \to \Box(A \wedge \Box A)$ $\qquad$ 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.

# Provability Logic

## 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*

$$\mathrm{GL} \vdash \boxdot\Big(\big(p \leftrightarrow A(p)\big) \wedge \big(q \leftrightarrow A(q)\big)\Big) \to (p \leftrightarrow q)$$

*where $\boxdot A := A \wedge \Box A$.*

## Corollary

*In $\mathrm{GL}$, if $p$ occurs only boxed in $A(p)$, then*

$$\frac{\begin{array}{c} B \leftrightarrow A(B) \\ C \leftrightarrow A(C) \end{array}}{B \leftrightarrow C}$$

## 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.*

$$\mathrm{GL} \vdash B \leftrightarrow A(B)$$

$$\boxed{\text{Uniqueness of Fixpoint + Beth Definability} \implies \text{Existence of Fixpoint}}$$

$$\mathrm{GL} \vdash \neg\Box\bot \leftrightarrow \neg\Box(\neg\Box\bot) \qquad \mathrm{GL} \vdash \top \leftrightarrow \Box\top$$

# 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 $X_i$.*

$$\Gamma \vdash_S A \iff \Gamma \Vdash_C A$$

# 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

$$\Box_a(\Diamond_b \top \wedge \Diamond_c \top)$$

# 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 $\mathcal{F}, V, w \nvDash A$，这与 $\mathcal{F} \vDash A$ 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

$$
\begin{aligned}
[\![\neg A]\!] &\coloneqq W \setminus [\![A]\!] \\
[\![A \wedge B]\!] &\coloneqq [\![A]\!] \cap [\![B]\!] \\
[\![A \vee B]\!] &\coloneqq [\![A]\!] \cup [\![B]\!] \\
[\![A \to B]\!] &\coloneqq \big(W \setminus [\![A]\!]\big) \cup [\![B]\!] \\
[\![\bot]\!] &\coloneqq \emptyset \\
[\![\top]\!] &\coloneqq W \\
[\![\Box A]\!] &\coloneqq [\![A]\!]^\circ \\
[\![\Diamond A]\!] &\coloneqq \overline{[\![A]\!]}
\end{aligned}
$$

where $^\circ$ is the interior operator $A^\circ \coloneqq \bigcup\{U \in \mathcal{O}_W : U \subset A\}$, and $\overline{\phantom{A}}$ is the closure operator $\overline{A} \coloneqq W \setminus (W \setminus A)^\circ$.

$$\mathcal{O}_W \Vdash A \coloneqq [\![A]\!] = W$$

## 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 $W \neq \emptyset$, $N : W \to \mathrm{P}(\mathrm{P}(W))$ is a neighborhood function, and $V : \mathrm{Var} \to \mathrm{P}(W)$ is a assignment function.

## Definition (Truth)

$$\begin{aligned}
\mathcal{M}, w \Vdash p &\iff w \in V(p) \\
\mathcal{M}, w \Vdash \neg A &\iff \mathcal{M}, w \nVdash A \\
\mathcal{M}, w \Vdash A \wedge B &\iff \mathcal{M}, w \Vdash A \ \ \& \ \ \mathcal{M}, w \Vdash B \\
\mathcal{M}, w \Vdash \Box A &\iff [\![A]\!] \in N(w) \\
\mathcal{M}, w \Vdash \Diamond A &\iff W \setminus [\![A]\!] \notin N(w)
\end{aligned}$$

where $[\![A]\!] := \{w : \mathcal{M}, w \Vdash A\}$.

**Remark:**

$$[\![\Box A]\!] = \Box_N [\![A]\!] \quad \text{where} \quad \Box_N X := \left\{w \in W : X \in N(w)\right\}$$
$$[\![\Diamond A]\!] = \Diamond_N [\![A]\!] \quad \text{where} \quad \Diamond_N X := \left\{w \in W : W \setminus X \notin N(w)\right\}$$

# 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 definition196.

$$[\![A]\!] \in N(w) \iff \exists U \in \mathcal{O}_W : w \in U \subset [\![A]\!] \iff w \in [\![A]\!]^\circ$$

# 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 |

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic
- Syntax
- Semantics
- Formal System
- Logic of Knowledge and Action

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Jaakko Hintikka

Figure: Jaakko Hintikka: 1929–2015. Epistemic Logic / Game Semantics / Independence-Friendly Logic

# Information Update

# Logic of Knowledge

- 什么是密码？你知我知。
- 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? ✓✗

# Reasoning about Knowledge

- 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 greatest enemy of knowledge is not ignorance, it is the illusion of knowledge."*

— *Stephen Hawking*

# Reasoning about Knowledge

*"The only good is knowledge and the only evil is ignorance."*

— *Socrates*

$\boxed{\text{know the unknown from the known}}$

- 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."*

— *Confucius*

- 知其然，知其所以然

$$\mathsf{K}\exists x Ax \to \exists x \mathsf{K}Ax$$

# Epistemic Logic

$$\mathcal{M}, w \Vdash \mathsf{K}_i A \quad \text{iff} \quad \forall v \in W\Big(R_i wv \implies \mathcal{M}, v \Vdash A\Big)$$

$w \Vdash p$

$w \Vdash \neg\mathsf{K}_1 p$

$v \Vdash \neg\mathsf{K}_1 p$

$w \Vdash \mathsf{K}_2 p$

$v \Vdash \mathsf{K}_2 \neg p$

$w \Vdash \mathsf{K}_1(\mathsf{K}_2 p \vee \mathsf{K}_2 \neg p)$

$w \Vdash \mathsf{K}_2 \neg\mathsf{K}_1 p$

$v \Vdash \mathsf{K}_2 \neg\mathsf{K}_1 p$

$w \Vdash \mathsf{K}_1\mathsf{K}_2 \neg\mathsf{K}_1 p$

$$w \Vdash p \wedge \neg\mathsf{K}_1 p \wedge \mathsf{K}_2 p \wedge \mathsf{K}_1(\mathsf{K}_2 p \vee \mathsf{K}_2 \neg p) \wedge \mathsf{K}_2 \neg\mathsf{K}_1 p \wedge \mathsf{K}_1\mathsf{K}_2 \neg\mathsf{K}_1 p$$

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 information.

- Common Knowledge:

  everybody in $G$ knows $p$,
  everybody knows that everybody knows,
  and so on.

# Mutual Knowledge

Suppose a group $G \subset \{1 \ldots n\}$ of agents, everyone in $G$ knows $A$:

$$\mathsf{E}_G A \;:=\; \bigwedge_{i \in G} \mathsf{K}_i A$$

$$R_E \;:=\; \bigcup_{i \in G} R_i$$

$$\mathcal{M}, w \Vdash \mathsf{E}_G A \quad \text{iff} \;\; \forall v \in W \Big( R_E w v \implies \mathcal{M}, v \Vdash A \Big)$$

# Distributed Knowledge

**Intuition:** what we know if we put all of our knowledge together.

$$R_D \;:=\; \bigcap_{i \in G} R_i$$

$$\mathcal{M}, w \Vdash \mathsf{D}_G A \quad \text{iff} \;\; \forall v \in W \Big( R_D wv \implies \mathcal{M}, v \Vdash A \Big)$$

$$w_2 \vDash \mathsf{K}_1 p \wedge \neg \mathsf{K}_1 q \wedge \mathsf{K}_2 q \wedge \neg \mathsf{K}_2 p \wedge \mathsf{D}_{\{1,2\}}(p \wedge q)$$

# Common Knowledge

$$\mathsf{E}_G^1 A := \mathsf{E}_G A$$
$$\mathsf{E}_G^{k+1} A := \mathsf{E}_G \mathsf{E}_G^k A$$
$$\mathsf{C}_G A := \bigwedge_{k=1}^{\infty} \mathsf{E}_G^k A$$

$$R^1 := R$$
$$R^{k+1} := R \circ R^k$$
$$R \circ S := \{(x, y) : \exists z (Rxz \wedge Szy)\}$$
$$R^* := \bigcup_{k=1}^{\infty} R^k$$

$$R_C := \left( \bigcup_{i \in G} R_i \right)^*$$

$$\mathcal{M}, w \Vdash \mathsf{C}_G A \text{ iff } \forall v \in W \Big( R_C wv \implies \mathcal{M}, v \Vdash A \Big)$$

## A Hierarchy of States of Knowledge

$$\mathsf{C}_G A \implies \cdots \mathsf{E}_G^k A \implies \cdots \mathsf{E}_G A \implies \bigvee_{i \in G} \mathsf{K}_i A \implies \mathsf{D}_G A \implies A$$

# 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 $i$ 和 $j$ 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 $i$ 和 $j$ 的公共知识么？

$$(0,1) \xleftrightarrow{j} (2,1) \xleftrightarrow{i} \underline{(2,3)} \xleftrightarrow{j} (4,3) \xleftrightarrow{i} (4,5) \xleftrightarrow{j} (6,5) \cdots$$

$$(2,3) \Vdash \neg \mathsf{K}_j \mathsf{K}_i \mathsf{K}_j (x + y \leq 10)$$

$i$ and $j$ commonly know that $j$'s number is odd.

# Coordinated Attack

# Coordinated Attack

# 《三体》— 黑暗森林

## 黑暗森林-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……

$p$: "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

$w \Vdash p$

$w \Vdash \mathsf{K}_2 p$

$w \Vdash \neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1\neg p$

$w \Vdash \mathsf{K}_1(\mathsf{K}_2 p \vee \mathsf{K}_2\neg p)$

$w \Vdash \mathsf{K}_2(\mathsf{K}_2 p \vee \mathsf{K}_2\neg p)$

$w \Vdash \mathsf{E}_{\{1,2\}}(\mathsf{K}_2 p \vee \mathsf{K}_2\neg p)$

$w \Vdash \mathsf{K}_1(\neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1\neg p)$

$w \Vdash \mathsf{K}_2(\neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1\neg p)$

$w \Vdash \mathsf{E}_{\{1,2\}}(\neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1\neg p)$

$w \Vdash \mathsf{C}_{\{1,2\}}(\mathsf{K}_2 p \vee \mathsf{K}_2\neg p)$

$w \Vdash \mathsf{C}_{\{1,2\}}(\neg\mathsf{K}_1 p \wedge \neg\mathsf{K}_1\neg p)$

**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 people are genuine Bayesian rationalists with common priors, and their posteriors are common knowledge, then these posteriors are equal.*

如果两个人有相同的先验知识，则他们不可能对有分歧的后验知识（经过各自的实验获取私人信息）形成公共知识。不管怎么根据进一步的私人证据进行充分的更新和交流，大家都不可能最后 agree to disagree!

# 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 $w$.
- $\mathsf{K}_i : \mathrm{P}(W) \to \mathrm{P}(W)$ 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. $W \setminus \mathsf{K}(A) \subset \mathsf{K}(W \setminus \mathsf{K}(A))$

# 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

$$\mathrm{S5} \ \text{ for } \ \mathsf{C}_G$$

$$+$$

$$\mathsf{C}_G A \leftrightarrow A \wedge \mathsf{E}_G \mathsf{C}_G A$$

$$+$$

$$A \wedge \mathsf{C}_G(A \to \mathsf{E}_G A) \to \mathsf{C}_G A$$

- Distributed Knowledge

$$\mathrm{S5} \ \text{ for } \ \mathsf{D}_G$$

$$+$$

$$\mathsf{K}_i A \to \mathsf{D}_G A \ \text{ when } \ i \in G$$

$$+$$

$$\mathsf{D}_G A \to \mathsf{D}_{G'} A \ \text{ when } \ G \subset G'$$

# Knowledge vs Belief

1. $\mathsf{K}(p \to q) \to \mathsf{K}p \to \mathsf{K}q$
2. $\mathsf{K}p \to p$
3. $\mathsf{K}p \to \mathsf{K}\mathsf{K}p$
4. $\neg\mathsf{K}p \to \mathsf{K}\neg\mathsf{K}p$

1. $\mathsf{K}p \to \mathsf{B}p$
2. $\mathsf{B}p \to \mathsf{K}\mathsf{B}p$
3. $\neg\mathsf{B}p \to \mathsf{K}\neg\mathsf{B}p$
4. $\mathsf{B}p \to \mathsf{B}\mathsf{K}p$ (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. $\mathsf{B}p \to \mathsf{B}\mathsf{B}p$
4. $\neg\mathsf{B}p \to \mathsf{B}\neg\mathsf{B}p$

$$\boxed{\mathsf{B}p \overset{?}{\leftrightarrow} \neg\mathsf{K}\neg\mathsf{K}p}$$

# Moore's Paradox

Are all truths believable?

$$\boxed{\text{It's raining but I don't believe it's raining.}}$$

$$\vdash \neg \mathsf{B}(p \wedge \neg \mathsf{B}p)$$

| | | |
|---|---|---|
| 1. | $\mathsf{B}(p \wedge \neg \mathsf{B}p)$ | Assumption |
| 2. | $\mathsf{B}p \wedge \mathsf{B}\neg \mathsf{B}p$ | $\mathsf{B}(p \wedge q) \to \mathsf{B}p \wedge \mathsf{B}q$ |
| 3. | $\mathsf{B}p$ | |
| 4. | $\mathsf{B}\mathsf{B}p$ | $\mathsf{B}p \to \mathsf{B}\mathsf{B}p$ |
| 5. | $\mathsf{B}\neg \mathsf{B}p$ | |
| 6. | $\mathsf{B}\mathsf{B}p \wedge \mathsf{B}\neg \mathsf{B}p$ | |
| 7. | $\neg(\mathsf{B}\mathsf{B}p \wedge \mathsf{B}\neg \mathsf{B}p)$ | $(\mathsf{B}p \to \neg \mathsf{B}\neg p) \leftrightarrow \neg(\mathsf{B}p \wedge \mathsf{B}\neg p)$ |
| 8. | $\neg \mathsf{B}(p \wedge \neg \mathsf{B}p)$ | |

# Moore's Paradox

Are all truths knowable?

$\boxed{\text{It's rainning but I don't know it's raining.}}$

$$\vdash \neg \mathsf{K}(p \wedge \neg \mathsf{K} p)$$

| | | |
|---|---|---|
| 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. | $\neg \mathsf{K}(p \wedge \neg \mathsf{K} p)$ | |

# Fitch's Paradox of Knowability

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. $\neg\mathsf{K}(p \wedge \neg\mathsf{K}p)$
2. $\Box\neg\mathsf{K}(p \wedge \neg\mathsf{K}p)$ — $\dfrac{A}{\Box A}$
3. $\neg\Diamond\mathsf{K}(p \wedge \neg\mathsf{K}p)$
4. $p \wedge \neg\mathsf{K}p$ — Assumption
5. $\Diamond\mathsf{K}(p \wedge \neg\mathsf{K}p)$ — $p \to \Diamond\mathsf{K}p$
6. $\neg(p \wedge \neg\mathsf{K}p)$
7. $p \to \mathsf{K}p$

# 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. | $\mathsf{B}\bot$ | |

$$\neg \mathsf{K}p \rightarrow \mathsf{K}\neg \mathsf{K}p \ ?$$

# 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. | $\mathsf{K}\neg p_{n+1}$ | ? |

$$\mathsf{K}p \to \mathsf{K}\mathsf{K}p \ ?$$

# Logical Omniscience vs Skepticism

$$\frac{A}{\mathsf{K}A} \qquad \frac{A \to B}{\mathsf{K}A \to \mathsf{K}B} \qquad \frac{A \leftrightarrow B}{\mathsf{K}A \leftrightarrow \mathsf{K}B} \qquad \frac{\mathsf{K}(A \to B) \quad \mathsf{K}A}{\mathsf{K}B}$$

$S_1$ 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: $\mathsf{K}p \coloneqq p \wedge \mathsf{B}p \wedge (\neg p \mathrel{\Box\!\!\to} \neg\mathsf{B}p) \wedge (p \mathrel{\Box\!\!\to} \mathsf{B}p)$
  - Sosa: $\mathsf{K}p \coloneqq p \wedge \mathsf{B}p \wedge (\mathsf{B}p \mathrel{\Box\!\!\to} p) \wedge (p \mathrel{\Box\!\!\to} \mathsf{B}p)$

# Information Update — Muddy Children Problem

## Problem (Muddy Children Problem)

*Consider $k$ of $n$ children get mud on their heads. Each child can see the mud on others but can't see his or her own head. Their father says "at least one is muddy." He then asks the following question repeatedly: "does anyone know whether you have mud on your own head?" Assuming that the children are intelligent, honest, and answer simultaneously, what will happen?*

1. "At least one is muddy. Does anyone. . . ?" $\neg \mathsf{K}_1 m_1 \wedge \neg \mathsf{K}_2 m_2 \wedge \neg \mathsf{K}_3 m_3$
2. "Does anyone. . . ?" $\mathsf{K}_1 m_1 \wedge \mathsf{K}_2 m_2 \wedge \neg \mathsf{K}_3 m_3$
3. $\mathsf{K}_3 \neg m_3$

# Public Announcement Logic

$$A ::= p \mid \neg A \mid A \wedge A \mid \mathsf{K}_i A \mid [A]A$$

$$\mathcal{M}, w \Vdash [B]A \;\text{ iff }\; \mathcal{M}, w \Vdash B \implies \mathcal{M}{\restriction}_B, w \Vdash A$$
$$\mathcal{M}, w \Vdash \langle B \rangle A \;\text{ iff }\; \mathcal{M}, w \Vdash B \;\&\; \mathcal{M}{\restriction}_B, w \Vdash A$$

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!

# Muddy Children Problem

$\mathcal{M}, mmc \Vdash m_1 \wedge m_2 \wedge \neg m_3$

$\mathcal{M}, mmc \Vdash \mathsf{E}_{\{1,2,3\}} P$

$\mathcal{M}, mmc \Vdash \neg \mathsf{C}_{\{1,2,3\}} P$

$\mathcal{M}, mmc \Vdash \neg \mathsf{K}_1 m_1 \wedge \mathsf{K}_1 m_2$

$\mathcal{M}, mmc \Vdash \mathsf{K}_1 \mathsf{K}_3 m_2 \wedge \mathsf{K}_1 \neg \mathsf{K}_2 m_2$

$\mathcal{M}\restriction_P, mcc \Vdash \mathsf{K}_1 m_1$

$\mathcal{M}\restriction_P, mmc \Vdash \langle \neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3 \rangle Q_1 \vee Q_2 \vee Q_3$

$\mathcal{M}\restriction_P, mmm \Vdash \langle \neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3 \rangle \neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3$

$\mathcal{M}\restriction_P \restriction_{\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3}, mmm \Vdash \langle \neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3 \rangle Q_1 \vee Q_2 \vee Q_3$

1. "At least one is muddy." $P := m_1 \vee m_2 \vee m_3$
2. "Does anyone. . . ?" $\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3$ where $Q_i := \mathsf{K}_i m_i \vee \mathsf{K}_i \neg m_i$
3. "Does anyone. . . ?" $Q_1 \wedge Q_2 \wedge \neg Q_3$

$\mathcal{M}, mmc \Vdash [P][\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3][Q_1 \wedge Q_2 \wedge \neg Q_3](\mathsf{K}_1 m_1 \wedge \mathsf{K}_2 m_2 \wedge \mathsf{K}_3 \neg m_3)$

$\mathcal{M}, mmm \Vdash [P][\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3][\neg Q_1 \wedge \neg Q_2 \wedge \neg Q_3](\mathsf{K}_1 m_1 \wedge \mathsf{K}_2 m_2 \wedge \mathsf{K}_3 m_3)$

# Three Cards Puzzle

- Three cards ‘red’, ‘yellow’, ‘green’ are given to three children: 1, 2, 3.
- The children can only see their own cards.
- 2 asks 1: “Do you have the green card?”
- 1 answers: “No”.

# Ages of Three Children

- A census-taker approaches a woman and asks about her children.
- She says "I have three children and the product of their ages is $36$. The sum of their ages is today's date."
- The census-taker complains "I still can't tell."
- The woman replies "Oh, I fogot to mention that my eldest child likes strawberries."

$$\begin{array}{ccc|c} 1 & 1 & 36 & \cancel{38} \\ 1 & 2 & 18 & 21 \\ 1 & 4 & 9 & 14 \\ 1 & 6 & 6 & 13 \\ 2 & 2 & 9 & 13 \\ 2 & 3 & 6 & 11 \\ 3 & 3 & 4 & 10 \end{array} \implies \begin{array}{ccc|c} 1 & 6 & 6 & 13 \\ 2 & 2 & 9 & 13 \end{array} \implies \begin{array}{ccc|c} 2 & 2 & 9 & 13 \end{array}$$

# Birthday Puzzle

**A** and **B** want to know when **C**'s birthday is.
**C** provides a list of $10$ possible dates:

| | | |
|---|---|---|
| $5.15$ | $5.16$ | $5.19$ |
| $6.17$ | $6.18$ | |
| $7.14$ | $7.16$ | |
| $8.14$ | $8.15$ | $8.17$ |

**C** then tells **A** and **B** separately the month and the day of her birthday.

- **A:** I don't know when **C**'s birthday is, but I know that **B** also does not know.
- **B:** At first I didn't know, but now I know.
- **A:** Then I also know it.

# Birthday Puzzle

**A** and **B** want to know when **C**'s birthday is.
**C** provides a list of $10$ possible dates:

$$\begin{array}{lll} 5.15 & 5.16 & 5.19 \\ 6.17 & 6.18 & \\ 7.14 & 7.16 & \\ 8.14 & 8.15 & 8.17 \end{array}$$

**C** then tells **A** and **B** separately the month and the day of her birthday.

- **A:** I don't know when **C**'s birthday is, but I know that **B** also does not know.
- **B:** At first I didn't know, but now I know.
- **A:** Then I also know it.

$$\begin{array}{lll} 5.15 & 5.16 & 5.19 \\ 6.17 & 6.18 & \\ 7.14 & 7.16 & \\ 8.14 & 8.15 & 8.17 \end{array} \Longrightarrow \begin{array}{lll} 7.14 & 7.16 & \\ 8.14 & 8.15 & 8.17 \end{array} \Longrightarrow \begin{array}{ll} 7.16 & \\ 8.15 & 8.17 \end{array} \Longrightarrow 7.16$$

# Russian Cards

## Russian Cards

- From a pack of seven known cards "0123456" $A$ and $B$ each draw three cards and $C$ gets the remaining card.
- How can $A$ and $B$ openly inform each other about their cards, without $C$ learning of any of their cards who holds it?

- Assume $A$' hand is $ijk$ and the remaining cards is $lmno$. Choose one from $ijk$, say $i$, and choose two from $lmno$, say $lm$. Three of the hands are $ijk, ilm, ino$. From $lm$ choose one, say $l$, and from $no$ choose one, say $n$. Two hands are $jln, kmo$. $A$ announces these five hands.
- $B$ announces $C$'s card.

# 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

$$\mathcal{M}, w \Vdash (p \wedge \neg\mathsf{K}_2 p) \wedge \mathsf{K}_1(p \wedge \neg\mathsf{K}_2 p) \wedge [\mathsf{K}_1(p \wedge \neg\mathsf{K}_2 p)]\,\mathsf{K}_1\mathsf{K}_2 p$$

- 哪些“公开宣告”可以使“真”的变成“假”的?
- 哪些“公开宣告”可以使“假”的变成“真”的?
- 哪些“公开宣告”可以保“真”?
- 哪些“公开宣告”可以被“知道”?
- 哪些“公开宣告”可以变成“公共知识”?

- $\langle B\rangle A \leftrightarrow B \wedge [B]A$
- $[B](A \rightarrow C) \leftrightarrow ([B]A \rightarrow [B]C)$
- $[B]p \leftrightarrow (B \rightarrow p)$
- $[B]\neg A \leftrightarrow (B \rightarrow \neg A)$ ?
- $[B]\neg A \leftrightarrow \neg[B]A$ ?
- $[B]\neg A \leftrightarrow (B \rightarrow \neg[B]A)$
- $[B]\mathsf{K}_i A \leftrightarrow (B \rightarrow \mathsf{K}_i(B \rightarrow [B]A))$
- $[B]\mathsf{K}_i A \leftrightarrow (B \rightarrow \mathsf{K}_i[B]A)$
- $[B][C]A \leftrightarrow [B \wedge C]A$ ?
- $[B][C]A \leftrightarrow [B \wedge [B]C]A$
- $\dfrac{A}{[B]A}$ $\quad \dfrac{A(p)}{A(B)}$ ? $\quad \dfrac{A \leftrightarrow B}{[A]C \leftrightarrow [B]C}$ $\quad \dfrac{A \leftrightarrow B}{[C]A \leftrightarrow [C]B}$

# 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

$$\frac{A \quad A \to B}{B}\ [\mathrm{MP}] \qquad\qquad \frac{A}{\mathsf{K}_i A}\ [\mathrm{N}]$$

# 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**

$\mathrm{PAL}$ *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**

$\mathrm{PAL}$ *is exponentially more succinct than modal logic on arbitrary models.*

$$A_0 \coloneqq \top$$

$$A_{n+1} \coloneqq \langle\langle A_n \rangle \Diamond_1 \top \rangle \Diamond_2 \top$$

where $\Diamond_i A \coloneqq \neg \mathsf{K}_i \neg A$ and $\langle B \rangle A \coloneqq \neg [B] \neg A$.

# Announcement and 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

$$A ::= \top \mid p \mid \neg A \mid A \wedge A \mid [\alpha]A$$
$$\alpha ::= a \mid A? \mid \alpha;\alpha \mid \alpha \cup \alpha \mid \alpha^*$$

$$\mathcal{M}, w \Vdash [\alpha]A \;\text{ iff }\; \forall v \in W\Big(R_\alpha wv \implies \mathcal{M}, v \Vdash A\Big)$$
$$\mathcal{M}, w \Vdash \langle\alpha\rangle A \;\text{ iff }\; \exists v \in W\Big(R_\alpha wv \;\&\; \mathcal{M}, v \Vdash A\Big)$$

where

$$\begin{aligned}
R_{A?} &:= \{(w,w) : \mathcal{M}, w \Vdash A\} \\
R_{\alpha;\beta} &:= \big\{(w,v) : \exists u\big(R_\alpha wu \wedge R_\beta uv\big)\big\} \\
R_{\alpha\cup\beta} &:= R_\alpha \cup R_\beta \\
R_{\alpha^*} &:= \bigcup_{n=0}^{\infty} R_{\alpha^n}
\end{aligned}$$

# 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).

# First Order Dynamic Logic

## Axiom Schema

1. FOL
2. PDL
3. $\langle x := t \rangle A \leftrightarrow A[t/x]$

## Inference Rule

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

$$\frac{A \to [\alpha^n]B, \quad n \in \omega}{A \to [\alpha^*]B} \text{ [IC]}$$

$$\frac{A}{[\alpha]A} \text{ [N]}$$

$$\frac{A}{\forall x A} \text{ [G]}$$

# 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 1** GCD

```
while x ≠ y do
    if x > y then
        x ← x − y
    else
        y ← 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

$$\frac{}{\{P\}\ \textbf{skip}\ \{P\}}$$

$$\frac{}{\{P[t/x]\}\ x := t\ \{P\}}$$

$$\frac{\{P\}\ \alpha\ \{Q\} \quad \{Q\}\ \beta\ \{R\}}{\{P\}\ \alpha; \beta\ \{R\}}$$

$$\frac{\{B \wedge P\}\ \alpha\ \{Q\} \quad \{\neg B \wedge P\}\ \beta\ \{Q\}}{\{P\}\ \textbf{if}\ B\ \textbf{then}\ \alpha\ \textbf{else}\ \beta\ \{Q\}}$$

$$\frac{P_1 \rightarrow P_2 \quad \{P_2\}\ \alpha\ \{Q_2\} \quad Q_2 \rightarrow Q_1}{\{P_1\}\ \alpha\ \{Q_1\}}$$

$$\frac{\{P \wedge B\}\ \alpha\ \{P\}}{\{P\}\ \textbf{while}\ B\ \textbf{do}\ \alpha\ \{\neg B \wedge P\}}$$

$$\{x = 4 \wedge y = 3\}\ \textbf{if}\ x < y\ \textbf{then}\ z := x; y := y + 1\ \textbf{else}\ z := y; z := z + 1\ \{x = 4 \wedge y = 3 \wedge z = 4\}$$

# 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.

$\vdots$

$$\nVdash A \leftrightarrow \mathsf{G}\neg A$$

## Other Versions of Yablo Paradox

| | | |
|---|---|---|
| always | $A_n \leftrightarrow \forall m > n : \neg A_m$ | $\Vdash \neg\mathsf{G}(A \leftrightarrow \mathsf{G}\neg A)$ |
| sometimes | $A_n \leftrightarrow \exists m > n : \neg A_m$ | $\Vdash \neg\mathsf{G}(A \leftrightarrow \mathsf{F}\neg A)$ |
| almost always | $A_n \leftrightarrow \exists k > n \forall m > k : \neg A_m$ | $\Vdash \neg\mathsf{G}(A \leftrightarrow \mathsf{FG}\neg A)$ |
| infinitely often | $A_n \leftrightarrow \forall k > n \exists m > k : \neg A_m$ | $\Vdash \neg\mathsf{G}(A \leftrightarrow \mathsf{GF}\neg A)$ |

# 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

$$\frac{A \quad A \to B}{B} \qquad \frac{A}{\mathsf{H}A} \qquad \frac{A}{\mathsf{G}A}$$

# Epistemic Temporal Logic

$$w \Vdash \neg \mathsf{K} p \wedge \mathsf{FK} p$$

**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。

$$\frac{\begin{array}{c} p \to \Box p \\ \neg p \to \Box \neg p \end{array}}{\Box p \vee \Box \neg p}$$

$$\frac{\begin{array}{c} \Box(p \to p) \\ \Box(\neg p \to \neg p) \end{array}}{\Box p \vee \Box \neg p} \times$$

| | |
|---|---|
| knowing-whether | $\mathsf{K}A \vee \mathsf{K}\neg A$ |
| knowing-what | $\exists x \mathsf{K}(A \to x = c)$ |
| knowing-how | $\exists \alpha \mathsf{K}[\alpha]A$ |
| knowing-why | $\exists t \mathsf{K}(t : A)$ |

Table: Yanjing Wang: Beyond Knowing That

## 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?

$$\forall xy(\mathsf{K}_x Hy \vee \mathsf{K}_x \neg Hy \to Fy \to Fx)$$

$$\forall x(\mathsf{K}_x Hf \vee \mathsf{K}_x \neg Hf \to x = f)$$

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。

# Proof of God's Existence?

1. God is a being which has every perfection.
2. Existence is a perfection.
3. Hence God exists.

$$E(\iota_x(Px \wedge Ex))$$

# Ontological Argument

- $R$: 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. $\Box \exists x G(x)$

# Pride and Prejudice

*"It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife."*

$$\mathsf{C}_{\mathrm{Human}}\Big(\forall x\Big(\mathrm{Man}(x) \wedge \mathrm{Single}(x) \wedge \mathrm{Fortune}(x) \rightarrow \mathrm{Want}_x\big(\exists y\big(\mathrm{Woman}(y) \wedge \mathrm{Wife}(y,x)\big)\big)\Big)\Big)$$

— *Jane Austen*

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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

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

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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。

# Method of Agreement

$$\frac{\begin{array}{c} ABC \rightarrow xyz \\ ADE \rightarrow xuv \end{array}}{A \rightarrow x}$$

| | 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 → 甲状腺肿大

| | | | |
|---|---|---|---|
| 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 → 失眠

# Method of Difference

$$\frac{\begin{array}{c} ABC \to xyz \\ \overline{A}BC \to \overline{x}yz \end{array}}{A \to x}$$

$$\frac{\begin{array}{ccccc} \text{冰激凌} & \text{沙拉} & \text{面包} & \text{牛奶} & \text{肚子疼} \\ — & \text{沙拉} & \text{面包} & \text{牛奶} & \text{肚子不疼} \end{array}}{\text{冰激凌} \to \text{肚子疼}}$$

$$\frac{\begin{array}{cc} \text{上课} & \text{头疼} \\ \text{不上课} & \text{头不疼} \end{array}}{\text{上课} \to \text{头疼}}$$

$$\begin{tikzcd} \text{上课} \arrow[r] \arrow[rr, bend left, "\times"] & \text{戴眼镜} \arrow[r] & \text{头疼} \end{tikzcd}$$

$$\frac{\begin{array}{ccc} \text{取健康的蜘蛛，} & \text{大吼一声，} & \text{蜘蛛被吓跑了} \\ \text{砍掉蜘蛛的腿，} & \text{大吼一声，} & \text{蜘蛛纹丝不动} \end{array}}{\text{蜘蛛的听觉器官长在腿上。}\ ☹}$$

# Joint Method of Agreement and Difference

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

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

达尔文：环境影响形态

| | 咖啡 | 嗦粉 | 熬夜 | 看书 | 肚子疼 |
|---|---|---|---|---|---|
| 张三 | ● | ● | ● | ● | ● |
| 李四 | ● | | ● | | ● |
| 王五 | | ● | | ● | |
| 赵六 | | | $\boxed{●}$ | ● | |

# Method of Residues

$$\begin{array}{c} ABC \to xyz \\ B \to y \\ C \to z \\ \hline A \to x \end{array} \qquad \begin{array}{c} ABC \to x \\ B \not\to x \\ C \not\to x \\ \hline A \to x \end{array}$$

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。

# Method of Concomitant Variation

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

热胀冷缩、体温表

禁烟人士 吸烟越多越容易患肺癌，所以吸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吸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.

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

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

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

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

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Example

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

## 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，鞠之服罪。

# Example

## Refutation by Analogy

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事物的分界线确定下来，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?

# Example

## Refutation by Analogy

- 人工智能不可能超越人脑，因为，人工智能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 — Does God Exist?

## 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?

# Picking the Right Level of Analogy

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。

作为类比桥梁的 $X$ 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！

# Metaphor: “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”……

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

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

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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

# 非形式谬误

- 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。

三秀才赶考，途遇算命先生，问几人中举？先生竖起一指。

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?[19]

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.

[19] 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 circularity, ambiguity, 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

- 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.

# 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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

References1655

# 一切皆集合

$X$

$X \cap Y$

$X \cup Y$

$\overline{X}$

$X \Delta Y$

$X \setminus Y$

# 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

- $X := \{x : x \in x \to A\}$
- $X \in X \iff X \in X \to A$
- $A$

# 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 $Y = \mathrm{P}(X) := \{u : u \subset X\}$.

$$\forall X \exists Y \forall u (u \in Y \leftrightarrow \forall z (z \in u \to z \in X))$$

- **Union**. For any $X$ there exists a set $Y = \bigcup X$.

$$\forall X \exists Y \forall u (u \in Y \leftrightarrow \exists z (z \in X \wedge u \in z))$$

$$\bigcup_{i \in I} X_i := \bigcup \{X_i : i \in I\}$$

$$\bigcap X := \{u : \forall z (z \in X \to u \in z)\}$$

# 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

$$X \times Y := \{(x, y) : x \in X \wedge y \in Y\}$$

# 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: $[x] := \{y \in X : x \sim y\}$
- quotient set: $X/\!\sim\; := \big\{[x] : x \in X\big\}$

Figure: torus $\mathbb{R}^2/\!\sim$

$$(x, y) \sim (x', y') := (x - x', y - y') \in \mathbb{Z}^2$$

- 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 $R(x, y) \iff \exists A \in \mathcal{E} : x, y \in A$.

# Equivalence Relation — Exercise

## Definition (Ideal)

A subset $I \subset \mathrm{P}(X)$ is an ideal, iff

1. $I \neq \emptyset$
2. $A \subset B \in I \to A \in I$
3. $A, B \in I \to A \cup B \in I$

## 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.

$$f\restriction_A \coloneqq \{(x, y) \in f : x \in A\} \quad (f \circ g)(x) \coloneqq f(g(x))$$

$$f(A) \coloneqq \{f(x) : x \in A\} \quad f^{-1}(A) \coloneqq \{x : f(x) \in A\}$$

# Relation Composition

$$S \circ R := \{(a, c) : \exists b(Rab \wedge Sbc)\}$$

# Exercises

In **Set**,

$$
\begin{aligned}
f : X \rightarrowtail Y \quad & \text{iff} \quad \exists g : Y \to X : gf = 1_X \\
& \text{iff} \quad \forall Z \forall g_1 g_2 : Z \to X : fg_1 = fg_2 \implies g_1 = g_2 \\
f : X \twoheadrightarrow Y \quad & \text{iff} \quad \exists g : Y \to X : fg = 1_Y \\
& \text{iff} \quad \forall Z \forall g_1 g_2 : Y \to Z : g_1 f = g_2 f \implies g_1 = g_2
\end{aligned}
$$

$$
\begin{tikzcd}
X \arrow[r, "f"] \arrow[dr, "1_X"'] & Y \arrow[d, "g"] \\
& X
\end{tikzcd}
\qquad
\begin{tikzcd}
Y \arrow[r, "g"] \arrow[dr, "1_Y"'] & X \arrow[d, "f"] \\
& Y
\end{tikzcd}
$$

$$
\begin{tikzcd}
Z \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] & X \arrow[r, "f"] & Y
\end{tikzcd}
\qquad
\begin{tikzcd}
X \arrow[r, "f"] & Y \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] & Z
\end{tikzcd}
$$

$f : X \rightarrowtail\!\!\!\!\twoheadrightarrow Y$ iff the diagram commutes:

$$
\begin{tikzcd}
X \arrow[r, "f"] \arrow[dr, "1_X"'] & Y \arrow[d, "g"] \arrow[dr, "1_Y"] & \\
& X \arrow[r, "f"'] & Y
\end{tikzcd}
$$

# 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 $g \circ f$,

$$\text{first do } g^{-1}\text{, then do } f^{-1}$$

iii. In symbols, $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

- 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 $\pi$ is the canonical map $x \mapsto [x]$.*

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "\pi"'] & Y \\
X/{\sim} \arrow[ur, dashed, "\bar{f}"'] &
\end{tikzcd}$$

# Order

- pre-order: $x \leq x, x \leq y \wedge y \leq z \rightarrow x \leq z$.
- partial order: pre-order 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 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 $X$.

# ZFC — Axioms

- **Replacement Schema**.
  If a class $F$ is a function, then for every set $X$, $F(X)$ is a set.

$$\forall xyz(A(x,y) \wedge A(x,z) \to y = z) \to \forall X \exists Y \forall y(y \in Y \leftrightarrow \exists x \in X A(x,y))$$

- **Axiom of Regularity**. Every non-empty set has an $\in$-minimal element.

$$\forall X(X \neq \emptyset \to \exists x(x \in X \wedge X \cap x = \emptyset))$$

- **Axiom of Infinity**.

$$\exists X(\emptyset \in X \wedge \forall x(x \in X \to x \cup \{x\} \in X))$$

- **Axiom of Choice** (AC). For any set $X$ of non-empty sets, there exists a choice function $f$ defined on $X$.

$$\forall X \left[\emptyset \notin X \to \exists f : X \to \bigcup X \;\; \forall A \in X \, (f(A) \in A)\right]$$

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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

References1655

# 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") $\mathrm{Card}(x) \;:=\; \mathrm{Ord}(x) \wedge \forall y \in x(|y| \neq |x|)$

$$|M| = |N| \;:=\; \exists f : M \rightarrowtail\!\!\!\!\twoheadrightarrow N \qquad |M| \leq |N| \;:=\; \exists f : M \rightarrowtail N$$

$$|M| := \min\{\alpha \in \mathrm{Ord} : |\alpha| = |M|\}$$

The infinite ordinal numbers that are cardinals are called alephs.

# Ordinal

$$\alpha < \beta := \alpha \in \beta$$

- $\emptyset$ is an ordinal.
- If $\alpha$ is an ordinal and $\beta \in \alpha$, then $\beta$ is an ordinal.
- If $\alpha \neq \beta$ are ordinals and $\alpha \subset \beta$, then $\alpha \in \beta$.
- If $\alpha, \beta$ are ordinals, then either $\alpha \subset \beta$ or $\beta \subset \alpha$.
- $<$ is a linear ordering of the class *Ord*.
- For each $\alpha$, $\alpha = \{\beta : \beta < \alpha\}$.
- 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 $\alpha$, $\alpha \cup \{\alpha\}$ 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.

$$\mathbb{N} := \big\{n : \forall A\big(\emptyset \in A \wedge \forall x \in A(x + 1 \in A) \rightarrow n \in A\big)\big\}$$

- A set $A$ is finite iff $\exists n \in \mathbb{N} : |A| = n$.
- A set $A$ is countable iff $|A| \leq |\mathbb{N}|$.

# Integer $\mathbb{Z}$

$$(m,n) \sim (a,b) := m +_{\mathbb{N}} b = a +_{\mathbb{N}} n$$

$$\mathbb{Z} := \mathbb{N} \times \mathbb{N}/\sim$$

$$\begin{aligned}
0_{\mathbb{Z}} &:= [(0,0)] \\
[(m,n)] \leq_{\mathbb{Z}} [(a,b)] &:= m +_{\mathbb{N}} b \leq_{\mathbb{N}} a +_{\mathbb{N}} n \\
[(m,n)] +_{\mathbb{Z}} [(a,b)] &:= [(m +_{\mathbb{N}} a, n +_{\mathbb{N}} b)] \\
[(m,n)] \cdot_{\mathbb{Z}} [(a,b)] &:= [(m \cdot_{\mathbb{N}} a + n \cdot_{\mathbb{N}} b, m \cdot_{\mathbb{N}} b + n \cdot_{\mathbb{N}} a)] \\
- [(m,n)] &:= [(n,m)]
\end{aligned}$$

$$\mathbb{Z}^+ := \{x \in \mathbb{Z} : x >_{\mathbb{Z}} 0_{\mathbb{Z}}\}$$

$$\exists f : \mathbb{N} \to \mathbb{Z} \quad n \mapsto [(n,0)]$$

# Rational Number $\mathbb{Q}$

$$(m, n) \sim (a, b) := m \cdot_{\mathbb{Z}} b = a \cdot_{\mathbb{Z}} n$$

$$\mathbb{Q} := \mathbb{Z} \times \mathbb{Z}^{+} / \sim$$

$$\begin{aligned}
0_{\mathbb{Q}} &:= [(0_{\mathbb{Z}}, x)] \\
1_{\mathbb{Q}} &:= [(x, x)] \\
[(m, n)] \leq_{\mathbb{Q}} [(a, b)] &:= m \cdot_{\mathbb{Z}} b \leq_{\mathbb{Z}} a \cdot_{\mathbb{Z}} n \\
[(m, n)] +_{\mathbb{Q}} [(a, b)] &:= [(m \cdot_{\mathbb{Z}} b +_{\mathbb{Z}} a \cdot_{\mathbb{Z}} n, n \cdot_{\mathbb{Z}} b)] \\
[(m, n)] \cdot_{\mathbb{Q}} [(a, b)] &:= [(m \cdot_{\mathbb{Z}} a, n \cdot_{\mathbb{Z}} b)] \\
- [(m, n)] &:= [(-m, n)]
\end{aligned}$$

$$\exists f : \mathbb{Z} \to \mathbb{Q} \quad x \mapsto [(x, 1)]$$

# Dedekind's Construction of Real Numbers $\mathbb{R}$

## Definition (Dedekind Cut)

$\mathbb{R}$ is the set of all $x \in \mathrm{P}(\mathbb{Q})$ 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.*

$$\exists f : \mathbb{Q} \rightarrow \mathbb{R} \quad a \mapsto \{b \in \mathbb{Q} : b < a\}$$

$$x \leq_{\mathbb{R}} y \coloneqq x \subset y$$

$$x +_{\mathbb{R}} y \coloneqq \{a +_{\mathbb{Q}} b : a \in x \ \&\ b \in y\}$$

$$-x \coloneqq \{a - b : a < 0 \ \&\ b \in \mathbb{Q} \setminus x\}$$

$$|x| \coloneqq x \cup -x$$

$$x \cdot_{\mathbb{R}} y \coloneqq \begin{cases} \{c : c \leq a \cdot_{\mathbb{Q}} b \ \&\ a \in x \ \&\ b \in y\} & \text{if } x > 0, y > 0 \\ 0 & \text{if } x = 0 \text{ or } y = 0 \\ |x| \cdot_{\mathbb{R}} |y| & \text{if } x < 0, y < 0 \\ -(|x| \cdot_{\mathbb{R}} |y|) & \text{if } x < 0, y > 0 \text{ or } x > 0, y < 0 \end{cases}$$

# 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.*

# 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.*

# 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

$$F_n^2 + F_{n+1}^2 = F_{2n+2} \text{ and } F_{n+1}^2 + 2F_nF_{n+1} = F_{2n+3}$$

**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.

## 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's Paradox

# 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.

$$P(F_1) = P\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} P(E_n) = \prod_{i=1}^{\infty} \frac{9n}{9n+1} = 0$$

$$\forall i : P(F_i) = 0 \implies P\left(\bigcup_{i=1}^{\infty} F_i\right) \leq \sum_{i=1}^{\infty} P(F_i) = 0$$

# 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 (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 $\alpha$.*

# 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

$$\lim_{\xi \to \alpha} \gamma_\xi := \sup \left\{\gamma_\xi : \xi < \alpha\right\}$$

## 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 $\gamma_\alpha$:

$$\gamma_\alpha = \lim_{\xi \to \alpha} \gamma_\xi$$

## Ordinal Addition

Let $\alpha$ and $\beta$ 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 $\alpha$ and $\beta$, and $x \leq y$ in $S$ iff $x \leq y$ in $\alpha$, or $x \leq y$ in $\beta$, or $x \in \alpha$ and $y \in \beta$.

**Remark:** The intuition of $\alpha + \beta$: $\alpha$ followed by $\beta$.

## Ordinal Multiplication

Let $\alpha$ and $\beta$ 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 $\alpha$, or $x \in \alpha_i$ and $y \in \alpha_j$ with $i < j$.

**Remark:** The intuition of $\alpha \cdot \beta$: $\beta$ copies of $\alpha$.

$$(A \otimes B, \leq)$$

$$(a_0, b_0) \leq (a_1, b_1) \iff b_0 < b_1 \text{ or, } b_0 = b_1 \ \wedge \ a_0 < a_1$$

# 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 $k_1, \ldots, k_n \in \mathbb{N}^+$.*

# 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 to 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

$$(7,6) \quad (\omega \cdot 2 + 1, \omega \cdot 2) \quad (\tfrac{19}{8}, \tfrac{39}{16})$$

# 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 of $x$ contains a point of $A$ other than $x$ itself.
- $x$ is a *adherent point* of $A$ iff every neighbourhood of $x$ contains a point of $A$.
- $x$ is an *interior point* of $A$ iff there is an open set $U$ s.t. $x \in U \subset A$.
- The *derived set* $A' := \{x : x \text{ is a limit point of } A\}$
- The *closure* of $A$ is $\overline{A} := \{x : x \text{ is a adherent point of } A\}$
- $\overline{A} = A \cup A'$
- The *interior* of $A$ is $A^\circ := \{x : x \text{ is an interior point of } A\}$
- The *boundary* of $A$ is $\partial A := \overline{A} \setminus A^\circ$
- $A$ is *closed* iff $A = \overline{A}$, or equivalently $A' \subset A$.
- $A$ is *perfect* iff $A = A'$.

$$\boxed{\text{Every perfect set has cardinality } 2^{\aleph_0}.}$$

# 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 $P := A_{\mathrm{rank}(A)}$. 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 $\alpha$ s.t. $a$ is an isolated point of $A_\alpha$.
Let $f(a) := \mu k[A_\alpha \cap J_k = \{a\}]$. 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

$$a_n := \frac{1}{\pi} \int_0^{2\pi} f(t) \cos nt \, dt$$
$$b_n := \frac{1}{\pi} \int_0^{2\pi} f(t) \sin nt \, dt$$

# 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 $A_{\mathrm{rank}(A)} = \emptyset$, then*

$$\forall x \in \mathbb{R} \backslash A \left( \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( a_n \cos nx + b_n \sin nx \right) = 0 \right) \implies \forall n \in \mathbb{N} \left( a_n = b_n = 0 \right)$$

If $f$ is continuous at all but countable points, then it admits a unique trigonometric expansion.

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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

References1655

# 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 $N := \{1, 2, 3, 4\}$.

$$\begin{array}{lcl} \text{apple} & \longleftrightarrow & 1 \\ \text{orange} & \longleftrightarrow & 2 \\ \text{banana} & \longleftrightarrow & 3 \\ \text{grape} & \longleftrightarrow & 4 \end{array}$$

$$|M| = |N| \;:=\; \exists f : M \rightarrowtail\!\!\!\!\twoheadrightarrow N$$

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?

$$\begin{array}{ccccc} \text{apple} & \longleftrightarrow & 1 & \longleftrightarrow & \text{John} \\ \text{orange} & \longleftrightarrow & 2 & \longleftrightarrow & \text{Peter} \\ \text{banana} & \longleftrightarrow & 3 & \longleftrightarrow & \text{Bell} \\ \text{grape} & \longleftrightarrow & 4 & \longleftrightarrow & \text{Emma} \\ & & 5 & \longleftrightarrow & \text{Sam} \end{array}$$

$$\begin{array}{ccc} \text{apple} & \longrightarrow & \text{John} \\ \text{orange} & \longrightarrow & \text{Peter} \\ \text{banana} & \longrightarrow & \text{Bell} \\ \text{grape} & \longrightarrow & \text{Emma} \\ & & \text{Sam} \end{array}$$

$$|M| \leq |N| \;:=\; \exists f : M \rightarrowtail N$$

$$|M| \leq |N| \;:=\; \exists f : N \twoheadrightarrow M$$

$$\begin{array}{ccc} \text{apple} & \longleftarrow & \text{John} \\ \text{orange} & \longleftarrow & \text{Peter} \\ \text{banana} & \longleftarrow & \text{Bell} \\ \text{grape} & \longleftarrow & \text{Emma} \\ & \nwarrow & \text{Sam} \end{array}$$

# The way of comparing the size of finite sets generalizes to infinite sets!

$$|\mathbb{N}| = |\mathbb{Z}|$$

$$\begin{array}{ccr} 0 & \longleftrightarrow & 0 \\ 1 & \longleftrightarrow & 1 \\ 2 & \longleftrightarrow & -1 \\ 3 & \longleftrightarrow & 2 \\ 4 & \longleftrightarrow & -2 \\ 5 & \longleftrightarrow & 3 \\ 6 & \longleftrightarrow & -3 \\ 7 & \longleftrightarrow & 4 \\ 8 & \longleftrightarrow & -4 \\ \vdots & & \vdots \end{array}$$

## 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 $|A| = |\mathbb{N}|$.

$$|\mathbb{N}| = |2^{<\omega}|$$

$$\begin{array}{ccr}
0 & \longleftrightarrow & \epsilon \\
1 & \longleftrightarrow & 0 \\
2 & \longleftrightarrow & 1 \\
3 & \longleftrightarrow & 00 \\
4 & \longleftrightarrow & 01 \\
5 & \longleftrightarrow & 10 \\
6 & \longleftrightarrow & 11 \\
7 & \longleftrightarrow & 000 \\
8 & \longleftrightarrow & 001 \\
\vdots & & \vdots
\end{array}$$

# Countable?

$$|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$$

# Countable?

$$|\mathbb{N}| = |\mathbb{Z} \times \mathbb{N}|$$

# Countable?

$$|\mathbb{N}| = |\mathbb{Z} \times \mathbb{Z}|$$

# $|\mathbb{N}| = |\mathbb{Q}|$

$$
\begin{tikzcd}
-\frac{2}{1} \arrow[d] & -\frac{1}{1} & \frac{0}{1} \arrow[r] & \frac{1}{1} \arrow[d] & \frac{2}{1} \arrow[r] & \frac{3}{1} \arrow[d] \\
-\cancel{\frac{2}{2}} \arrow[d] & -\frac{1}{2} \arrow[u] & \cancel{\frac{0}{2}} \arrow[l] & \frac{1}{2} \arrow[l] & \cancel{\frac{2}{2}} \arrow[u] & \frac{3}{2} \arrow[d] \\
-\frac{2}{3} \arrow[r] & -\frac{1}{3} \arrow[r] & \cancel{\frac{0}{3}} \arrow[r] & \frac{1}{3} \arrow[r] & \frac{2}{3} \arrow[u] & \cancel{\frac{3}{3}} \arrow[d] \\
-\cancel{\frac{2}{4}} & -\frac{1}{4} \arrow[l] & \cancel{\frac{0}{4}} \arrow[l] & \frac{1}{4} \arrow[l] & \cancel{\frac{2}{4}} \arrow[l] & \frac{3}{4} \arrow[l]
\end{tikzcd}
$$

$$|\mathbb{N}| = |\mathbb{Z}| = |2^{<\omega}| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{Z} \times \mathbb{N}| = |\mathbb{Z} \times \mathbb{Z}| = |\mathbb{Q}|$$

# $|\mathbb{N}| = |\mathbb{Q}^+|$

$$
\begin{array}{ccccc}
\frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots \\
\frac{2}{1} & \frac{2}{2} & \frac{2}{3} & \frac{2}{4} & \cdots \\
\frac{3}{1} & \frac{3}{2} & \frac{3}{3} & \frac{3}{4} & \cdots \\
\frac{4}{1} & \frac{4}{2} & \frac{4}{3} & \frac{4}{4} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

$$x \mapsto \frac{1}{\lfloor x \rfloor + 1 - \{x\}}$$

# 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.

$$\begin{array}{cccccc}
0\,. & r_{11} & r_{12} & r_{13} & r_{14} & \dots \\
0\,. & r_{21} & r_{22} & r_{23} & r_{24} & \dots \\
0\,. & r_{31} & r_{32} & r_{33} & r_{34} & \dots \\
0\,. & r_{41} & r_{42} & r_{43} & r_{44} & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

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

$$f : (0,1) \rightarrowtail\!\!\!\!\twoheadrightarrow (0,\infty)$$

$$f : (0,1] \rightarrowtail\!\!\!\!\twoheadrightarrow (0,1)$$

$$f(x) := \begin{cases} \frac{3}{2} - x & \text{for } \frac{1}{2} < x \leq 1 \\ \frac{3}{4} - x & \text{for } \frac{1}{4} < x \leq \frac{1}{2} \\ \frac{3}{8} - x & \text{for } \frac{1}{8} < x \leq \frac{1}{4} \\ \vdots & \end{cases}$$

$$f(x) := \begin{cases} \frac{x}{x+1} & \text{if } \exists n \in \mathbb{N} : x = \frac{1}{n} \\ x & \text{otherwise} \end{cases}$$

# Continuum

# Continuum

**Theorem**

$$|\mathbb{R}| = |\mathbb{R} \times \mathbb{R}|$$

**Proof.**

$$\begin{array}{rlllll} x = & 0.3 & 01 & 2 & 007 & 08\ldots \\ y = & 0.009 & 2 & 05 & 1 & 0003\ldots \end{array}$$

$$z = 0.3\ 009\ 01\ 2\ 2\ 05\ 007\ 1\ 08\ 0003\ \ldots$$

# 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)

$$\left.\begin{array}{l} |M| \leq |N| \\ |N| \leq |M| \end{array}\right\} \Longrightarrow |M| = |N|$$

1. Finite cycles on $2k+2$ distinct elements $(k \geq 0)$
   $m_0 \longrightarrow n_0 \longrightarrow m_1 \longrightarrow \cdots \longrightarrow m_k \longrightarrow n_k$
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 $n_0 \in N \setminus f(M)$
   $n_0 \longrightarrow m_0 \longrightarrow n_1 \longrightarrow m_1 \longrightarrow n_2 \longrightarrow \cdots$

$m_i \mapsto n_i$

# 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 $f(u) = u$.

# 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.

$$C_0 := M \setminus g(N)$$
$$D_0 := f(C_0)$$
$$C_{n+1} := g(D_n)$$
$$D_{n+1} := f(C_n)$$
$$C := \bigcup_{n\in\mathbb{N}} C_n$$
$$f\restriction_C \cup g^{-1}\restriction_{g(N)\setminus C}$$

$$h : S \mapsto (M \setminus g(N)) \cup g(f(S))$$
$$C = \bigcup_{n\to\infty} h^n(\emptyset)$$
$$= \bigcap \left\{ S : (M \setminus g(N)) \cup g(f(S)) \subset S \right\}$$

# 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.

$$M_0 := M, \quad M_1 := g(N), \quad M_{k+2} := g \circ f(M_k)$$

$$M_0 \supset M_1 \supset M_2 \supset \cdots \supset M_k \supset M_{k+1} \supset \cdots$$

$$A := \bigcup_{k=0}^{\infty} (M_{2k+1} \setminus M_{2k+2}) \quad B := \bigcup_{k=0}^{\infty} (M_{2k} \setminus M_{2k+1}) \quad C := \bigcup_{k=1}^{\infty} (M_{2k} \setminus M_{2k+1})$$

$$D := \bigcap_{k=0}^{\infty} M_k$$

$$M = A \cup B \cup D$$

$$M_1 = A \cup C \cup D$$

$$|M_{2k} \setminus M_{2k+1}| = |g \circ f(M_{2k}) \setminus g \circ f(M_{2k+1})| = |M_{2k+2} - M_{2k+3}| \implies |B| = |C|$$

$$|M| = |M_1| = |N|$$

# 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.*
- $(\kappa \cdot \lambda)^{\mu} = \kappa^{\mu} \cdot \lambda^{\mu}$
- $\kappa^{\lambda+\mu} = \kappa^{\lambda} \cdot \kappa^{\mu}$
- $(\kappa^{\lambda})^{\mu} = \kappa^{\lambda \cdot \mu}$
- $\kappa \leq \lambda \implies \kappa^{\mu} \leq \lambda^{\mu}$
- $0 < \lambda \leq \mu \implies \kappa^{\lambda} \leq \kappa^{\mu}$
- $\kappa^{0} = 1$; $1^{\kappa} = 1$; $0^{\kappa} = 0$ *if* $\kappa > 0$.

## 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 $\alpha \times \alpha = \big\{(\xi, \eta) : (\xi, \eta) < (0, \alpha)\big\}$.

Let $\Gamma(\alpha, \beta) := \mathrm{otp}\big\{(\xi, \eta) : (\xi, \eta) < (\alpha, \beta)\big\}$. Then $(\alpha, \beta) < (\gamma, \delta) \iff \Gamma(\alpha, \beta) < \Gamma(\gamma, \delta)$, and $\Gamma(\omega, \omega) = \omega$, $\Gamma(\alpha, \alpha) \geq \alpha$. Assume $\alpha$ 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 $N_\varepsilon(A) := \left\{x \in M : d(x, A) < \varepsilon\right\}$ where $d(x, A) := \inf\limits_{y \in A} d(x, y)$. Let

$$d_H(A, B) := \inf\left\{\varepsilon : A \subset N_\varepsilon(B) \ \& \ B \subset N_\varepsilon(A)\right\}$$

  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 $S(X) = \bigcup\limits_{i=1}^{n} T_i(X)$. Then $S$ is a contraction.
- According to Banach's fixpoint theorem, $\exists X \in \mathcal{M}' : S(X) = X$.
  Furthermore, $\forall Y \in \mathcal{M}' : S(Y) \subset Y \implies X = \bigcap\limits_{k=0}^{\infty} S^k(Y)$.
- 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

- $|\mathcal{C}| = 2^{\aleph_0}$
- $\mathcal{C}$ is perfect.
- $\mathcal{C}$ is *nowhere dense* in $[0, 1]$.

$$\left(\overline{\mathcal{C}}\right)^{\circ} = \emptyset$$

- 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

$$\boxed{\text{A set } A \text{ is a } \textit{fractal} \text{ iff } \dim_H(A) > \dim_T(A).}$$

$$H^d_\varepsilon(A) := \inf\left\{\sum_{k=1}^{\infty} \operatorname{diam}(B_k)^d : A \subset \bigcup_{k=1}^{\infty} B_k \ \ \& \ \ \operatorname{diam}(B_k) \le \varepsilon\right\}$$

$$\dim_H(A) := \inf\left\{d \ge 0 : \lim_{\varepsilon \to 0} H^d_\varepsilon(A) = 0\right\}$$

## 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 $\mathcal{V}$.

# Hilbert's Space-filling Curve

# 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 $\varepsilon$. 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: $2$

# Cardinality of the Permutations of $\mathbb{N}$ — $|\operatorname{Aut}(\mathbb{N})|$

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 $n^{\text{th}}$ 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. $|\operatorname{Aut}(\mathbb{N})| > \aleph_0$.

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 $n$. $|\operatorname{Aut}(\mathbb{N})| \geq 2^{\aleph_0}$.

Proof3. Riemann rearrangement $\left(e.g. \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\right) \implies \mathbb{R} \rightarrowtail \operatorname{Aut}(\mathbb{N})$.

## Theorem (Riemann Rearrangement Theorem)

*Any conditionally convergent series can be rearranged in a permutation to*

1. *converge to any real number;*
2. *diverge to $\infty$ or $-\infty$;*
3. *oscillate finitely or infinitely.*

# 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 $\beta$ s.t. there is an increasing $\beta$-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 $\mathrm{cf}(\alpha) = 1$.
  - $\mathrm{cf}(\omega) = \mathrm{cf}(\omega + \omega) = \mathrm{cf}(\aleph_\omega) = \mathrm{cf}(\aleph_{\aleph_\omega}) = \omega$
  - $\mathrm{cf}(\aleph_1) = \aleph_1$
- An infinite cardinal $\kappa$ is regular iff $\mathrm{cf}(\kappa) = \kappa$.
- It is singular iff $\mathrm{cf}(\kappa) < \kappa$.
- A cardinal $\kappa$ is a strong limit cardinal iff $\forall \lambda < \kappa \big(2^\lambda < \kappa\big)$.
- A cardinal $\kappa$ is inaccessible iff it is a regular strong limit uncountable cardinal.

# 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 $\alpha$ 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

$$V = L$$

$$L \vDash \mathrm{ZF}$$

$$L \vDash V = L$$

$$\mathrm{ZF} + \mathrm{V} = \mathrm{L} \vdash \mathrm{AC} + \mathrm{GCH}$$

An inner model is a transitive class model of $\mathrm{ZF}$ that contains all ordinals.
$L$ is the smallest inner model of $\mathrm{ZF}$.

# 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. $x \in U$.

$$\boxed{U \text{ is a Grothendieck universe iff } U = V_\kappa \text{ for some inaccessible } \kappa.}$$

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 (Axiom I3)

*For some $\lambda$, $V_\lambda$ 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 $\infty$), 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 $j$.

$$\mathcal{E}_\lambda \coloneqq \{j : V_\lambda \prec V_\lambda \ \ \& \ \ j \text{ is non-surjective}\}$$

$$i[j] \coloneqq \bigcup_{\alpha<\lambda} i(j \cap V_\alpha^2)$$

$(\mathcal{E}_\lambda, [])$ is a left-shelf: $i[j[k]] = i[j][i[k]]$

$$\mathrm{crit}(j \circ j) = \mathrm{crit}(j) \quad \text{but} \quad \mathrm{crit}(j[j]) = j(\mathrm{crit}(j)) > \mathrm{crit}(j)$$

$$j_{[n]} := \underbrace{j[j][j]\cdots[j]}_{n \text{ times}}$$

$$\mathrm{Iter}(j) := \left\{ j_{[n]} : n \in \mathbb{N}^+ \right\}$$

$$\boxed{\big(\mathrm{Iter}(j), []\big) \text{ is a left-shelf.}}$$

For $k, k' \in \mathrm{Iter}(j)$, declare $k \equiv_n k' := \forall x \in V_\gamma \big(k(x) \cap V_\gamma = k'(x) \cap V_\gamma\big)$ with $\gamma := \mathrm{crit}(j_{[2^n]})$. Then

$$\boxed{\mathrm{Iter}(j)/{\equiv_n} \text{ is (isomorphic to) the Laver table } A_n.}$$

# Ordinal

$0, 1, 2, 3, \ldots$

$\omega, \omega + 1, \omega + 2, \ldots$

$\omega \cdot 2, (\omega \cdot 2) + 1, (\omega \cdot 2) + 2, \ldots$

$\omega^2, \omega^2 + 1, \omega^2 + 2, \ldots$

$\omega^\omega, \omega^\omega + 1, \omega^\omega + 2, \ldots$

$\omega^{\omega^\omega}, \ldots$

$\varepsilon_0 = \omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}} = \sup\{\omega, \omega^\omega, \omega^{\omega^\omega}, \omega^{\omega^{\omega^\omega}}, \ldots\}$

$\varepsilon_1 = \sup\{\varepsilon_0 + 1, \omega^{\varepsilon_0+1}, \omega^{\omega^{\varepsilon_0+1}}, \omega^{\omega^{\omega^{\varepsilon_0+1}}}, \ldots\} = \sup\{0, 1, \varepsilon_0, {\varepsilon_0}^{\varepsilon_0}, {\varepsilon_0}^{{\varepsilon_0}^{\varepsilon_0}}, \ldots\}$

$\varepsilon_{\alpha+1} = \sup\{\varepsilon_\alpha + 1, \omega^{\varepsilon_\alpha+1}, \omega^{\omega^{\varepsilon_\alpha+1}}, \ldots\} = \sup\{0, 1, \varepsilon_\alpha, \varepsilon_\alpha^{\varepsilon_\alpha}, \varepsilon_\alpha^{\varepsilon_\alpha^{\varepsilon_\alpha}}, \ldots\}$

$\varepsilon_\alpha = \sup\{\varepsilon_\beta : \beta < \alpha\}$ if $\alpha$ 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

$$\varphi_0(\alpha) \coloneqq \omega^\alpha$$

$$\varphi_{\gamma+1}(\alpha) \coloneqq \alpha^{\text{th}} \text{ ordinal s.t. } \varphi_\gamma(\beta) = \beta$$

$$\varphi_\delta(\alpha) \coloneqq \alpha^{\text{th}} \text{ common fixpoint of } \varphi_\gamma \text{ for all } \gamma < \delta$$

$$\Gamma_\alpha \coloneqq \alpha^{\text{th}} \text{ ordinal s.t. } \varphi_\alpha(0) = \alpha$$

$$\boxed{\varepsilon_\alpha = \varphi_1(\alpha)}$$

"inaccessible" cardinals
($ZF$ axioms cannot be used to prove the existence of these sets)
$\kappa = \aleph_\kappa$
First cardinal that can only be named by the ordinal that is as big as itself
$\aleph_\omega$
First cardinal to be preceded by infinitely many cardinals
All sets here and above are ***uncountable***
$\aleph_1$
Second infinite cardinal (first uncountable ordinal)
$\varepsilon_0$
$\omega = \aleph_0$
First infinite cardinal/ordinal ($\mathbb{N}$)
All sets here are ***countable*** (same size as $\mathbb{N}$)
6
All sets here are finite
5
65,536 sets below Stage 5!
The number of sets here has 20,000 digits!
empty set $\varnothing$

$0 = 1$

I0-I3

$n$-huge

superhuge

huge

almost huge

Vopěnka's Principle

extendible

supercompact

strongly compact

superstrong

Woodin

strong

$0^\dagger$ exists

measurable

Ramsey

Rowbottom

Jónsson

$\kappa \longrightarrow (\omega_1)_2^{<\omega}$

$\forall a \in {}^\omega\omega (a^\# \text{ exists})$

$0^\#$ exists

$\kappa \longrightarrow (\omega)_2^{<\omega}$

indescribable

weakly compact

$\alpha$-Mahlo

Mahlo

$\alpha$-inaccessible

inaccessible

weakly inaccessible

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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

References1655

# 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*

## 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

- Well-ordering theorem: Every set can be well-ordered.
- Trichotomy: For any two cardinals $\kappa$ and $\lambda : \kappa < \lambda \vee \kappa = \lambda \vee \kappa > \lambda$.
- For any infinite cardinal $\kappa : \kappa^2 = \kappa$.
- The Cartesian product of any family of non-empty sets is non-empty.
- Every surjective function has a right inverse.
- Hausdorff's Maximal Chain Condition: Each partially ordered set contains a maximal chain.
- Zorn's lemma: If in a partially ordered set $X$ each chain has an upper bound, then $X$ has a maximal element.
- Every vector space has a basis.
- The closed unit ball of the dual of a normed vector space over the reals has an extreme point.
- Tychonoff's theorem: The product of compact topological spaces is compact.
- If a set $\Gamma$ of formulas with $|\mathscr{L}| = \kappa$ is finitely satisfiable, then it has a model with cardinality $\leq \kappa + \aleph_0$.

# AC

## Theorem (Well-Ordering Theorem)

*Every set can be well-ordered.*

## Proof.

Assume $f$ is a choice function for $\mathrm{P}(A) \setminus \{\emptyset\}$. 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.*

# Consistence & Independence

## Theorem (Gödel1938)

$$\mathrm{Con}(\mathrm{ZF}) \to \mathrm{Con}(\mathrm{ZFC} + \mathrm{GCH})$$

## Theorem (Cohen1963)

- $\mathrm{Con}(\mathrm{ZF}) \to \mathrm{Con}(\mathrm{ZF} + \neg\mathrm{AC})$
- $\mathrm{Con}(\mathrm{ZF}) \to \mathrm{Con}(\mathrm{ZFC} + \neg\mathrm{GCH})$

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

$$2^{\aleph_\alpha} = \aleph_{\alpha+1} \qquad \text{(GCH)}$$

$$\Downarrow$$

$$2^\kappa < 2^{\kappa^+} \qquad \text{(WGCH)}$$

$$\Updownarrow$$

$$\kappa < \lambda \implies 2^\kappa < 2^\lambda$$

$$\boxed{\text{“}|X| < |Y| \implies |\mathrm{P}(X)| < |\mathrm{P}(Y)|\text{” is independent of ZFC.}}$$

# Freiling's Axiom of Symmetry

## Freiling's Axiom of Symmetry (AX)

$$\forall f : \mathbb{R} \to \mathbb{R}_{\aleph_0} \exists xy \big[x \notin f(y) \wedge y \notin f(x)\big]$$

## 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
References1655

# Alan Turing 1912-1954

- Universal Turing Machine
- Church-Turing Thesis
- Halting Problem
- Undecidability
- Oracle Machine
- Computable Absolutely Normal Number
- Turing Test
- Morphogenesis
- Good-Turing Smoothing
- Enigma

What is "effective procedure"? — Recursion Theory

# Turing[20]

- Computing is normally done by writing certain symbols on paper.
- We may suppose this paper is divided into squares like a child's arithmetic book.
- The number of symbols which may be printed is finite.
- The behaviour of the computer at any moment is determined by the symbols which he is observing, and his "state of mind" at that moment.
- We will also suppose that the number of states of mind which need to be taken into account is finite.

[20] Turing: On computable numbers, with an application to the Entscheidungsproblem. 1936.

# (Deterministic) Turing Machine

## Definition ((Deterministic) Turing Machine)

A deterministic Turing machine is a triplet $(\Sigma, Q, \delta)$, where $\Sigma$ is a finite alphabet with an identified blank symbol, $Q$ is a finite set of states with identified initial state $q_0$ and finial state $q_f \neq q_0$, and $\delta$, a deterministic transition function

$$\delta : Q \times \Sigma \to \Sigma \times Q \times \{L, R\}$$

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.

$$\delta(q_{23}, 0) = (1, q_{359}, R)$$

# Turing Machine — Example

This machine writes $1$, then moves right forever. It will never halt since the function $\delta$ never maps to the state $q_f$.

$$\Sigma = \{0, 1\}$$
$$Q = (q_0, q_f)$$
$$\delta(q_0, 0) = (1, q_0, R)$$
$$\delta(q_0, 1) = (1, q_0, R)$$

# Turing Machine — Example

This machine will move left then right and so on...

$$\Sigma = \{0, 1\}$$

$$Q = (q_0, q_1, q_f)$$

$$\delta(q_0, 0) = (1, q_1, R)$$

$$\delta(q_1, 0) = (1, q_0, L)$$

$$\delta(q_0, 1) = (1, q_1, R)$$

$$\delta(q_1, 1) = (1, q_0, L)$$

# Turing Machine — Example

The ADD Turing machine takes an input of two unary numbers separated by a $0$, $(1^l 0 1^m)$, then returns the sum of those two numbers, $(1^{l+m})$.

$$\Sigma = \{0, 1\}$$

$$Q = (q_0, q_1, q_2, q_f)$$

$$\delta(q_0, 1) = (1, q_0, R)$$

$$\delta(q_0, 0) = (1, q_1, R)$$

$$\delta(q_1, 1) = (1, q_1, R)$$

$$\delta(q_1, 0) = (0, q_2, L)$$

$$\delta(q_2, 1) = (0, q_f, R)$$

# (Deterministic / Probabilistic) Turing Machine

# (Quantum) Turing Machine

## Thesis (Church-Turing Thesis)

*effective calculable* = *recursive* = *Turing Computable*

||

*representable in* $\mathrm{Q}$ = $\lambda$*-definable*

||

*finite definable* = *Herbrand-Gödel computable*

||

*flowchart (or 'while') computable*

||

*neural network with unbounded tape* = *Conway's 'game of life'*

||

*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?

# 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.*

# Computational Irreducibility vs Free Will

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).

Libet: We are conscious of our free choices only after $300$ms our brain has made them.

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
References1655

# Presburger/Robinson/Peano Arithmetic

- $x + 0 = x$
- $x + y = y + x$
- $(x + y) + z = x + (y + z)$
- $x + z = y + z \rightarrow x = y$
- $x \cdot 0 = 0$
- $x \cdot 1 = x$
- $x \cdot y = y \cdot x$
- $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
- $x \cdot (y + z) = x \cdot y + x \cdot z$
- $P(0) \wedge \forall x(P(x) \rightarrow P(x + 1)) \rightarrow \forall x P(x)$

$$\left.\begin{array}{ll} 1. & s(x) \neq 0 \\ 2. & s(x) = s(y) \rightarrow x = y \\ 3. & y = 0 \vee \exists x(s(x) = y) \\ 4. & x + 0 = x \\ 5. & x + s(y) = s(x + y) \\ 6. & x \cdot 0 = 0 \\ 7. & x \cdot s(y) = (x \cdot y) + x \end{array}\right\} Q$$

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 $\mathrm{PA}^2$

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 $\mathrm{Q}$

$$\pi(x,y) := (x+y)^2 + x + 1$$
$$\pi_1(z) := \mu x[\exists y \le z : \pi(x,y) = z]$$
$$\pi_2(z) := \mu y[\exists x \le z : \pi(x,y) = z]$$
$$\beta(s,i) := \mu x < s\big[\pi_1(s) \equiv x \pmod{1 + (i+1)\cdot\pi_2(s)}\big]$$
$$\beta^*(s,i) := \mu x < s\,[\exists y < s\,\exists z < s : s = \pi(y,z) \ \wedge\ (1 + (\pi(x,i)+1)\cdot z) \mid y]$$

## Lemma

*For every sequence $a_0, \ldots, a_n \in \mathbb{N}$, $\exists s \in \mathbb{N}.\forall i \le n.\beta(s,i) = a_i$.*

## Proof.

$$b := \max\{n, a_0, \ldots, a_n\} \quad d := b! \quad m_i := 1 + (i+1)\cdot d$$

$m_0, \ldots, m_n$ are pairwise coprime.

$c := \mu x\,[\forall i \le n : x \equiv a_i \pmod{m_i}]$ by Chinese Remainder Theorem.

$$s := \pi(c,d)$$
$$x^y := \beta\Big(\mu s\big[\beta(s,0) = 1 \ \wedge\ \forall i < y : \beta(s,i+1) = \beta(s,i)\cdot x\big], y\Big)$$

# 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.

$$M := \prod_{i=1}^{n} m_i \qquad M_i := \frac{M}{m_i}$$

$$n_i := \mu x \left[ x \cdot M_i \equiv 1 \pmod{m_i} \right]$$

$$x \equiv \sum_{i=1}^{n} n_i \cdot M_i \cdot a_i \pmod{m}$$

# 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)

| $\zeta$ | ( | ) | , | $\neg$ | $\rightarrow$ | $\forall$ | $x_k$ | $a_k$ | $f_k^n$ | $P_k^n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $\#\zeta$ | 3 | 5 | 7 | 9 | 11 | 13 | $7+8k$ | $9+8k$ | $11+8(2^n3^k)$ | $13+8(2^n3^k)$ |

$$\langle a_1, \ldots, a_n \rangle \coloneqq \mu x \big[ \beta(x,0) = n \wedge \beta(x,1) = a_1 \wedge \cdots \wedge \beta(x,n) = a_n \big]$$

$$or \quad \langle a_1, \ldots, a_n \rangle \coloneqq \prod_{i=1}^{n} p_i^{a_i+1}$$

# 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 $\mathrm{T}$ 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 wff $A$ s.t. $\mathrm{T} \vdash A \leftrightarrow \neg \Box_{\mathrm{T}} A$?

$$\begin{tikzcd}
 & \text{Gödel Number} \arrow[dr, "\underline{\phantom{x}}"] & \\
\text{Expression} \arrow[ur, "\#"] \arrow[rr, "\ulcorner\urcorner"] & & \text{Gödel Numeral}
\end{tikzcd}$$

meta:

object:

$$\begin{tikzcd}
 & \#\zeta \arrow[dr, "\underline{\phantom{x}}"] & \\
\zeta \arrow[ur, "\#"] \arrow[rr, "\ulcorner\urcorner"] & & \ulcorner \zeta \urcorner
\end{tikzcd}$$

- $\mathrm{Con}_{\mathrm{T}} := \neg\Box_{\mathrm{T}}\bot$
- $\omega$-consistent: $\forall x \Box_{\mathrm{T}} \neg A(\dot{x}) \to \neg\Box_{\mathrm{T}} \exists x A(x)$ for any formula $A$
- 1-consistent: $\forall x \Box_{\mathrm{T}} \neg A(\dot{x}) \to \neg\Box_{\mathrm{T}} \exists x A(x)$ for $A \in \Delta_0$
- $\mathrm{Rfn}_{\Gamma}(\mathrm{T}) : \Box_{\mathrm{T}} A \to A$ for any sentence $A \in \Gamma$
- $\mathrm{RFN}_{\Gamma}(\mathrm{T}) : \forall x \big(\Box_{\mathrm{T}} A(\dot{x}) \to A(x)\big)$ for any wff $A \in \Gamma$
- 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 $g, h$,

$$f(\mathbf{x}, 0) = g(\mathbf{x})$$
$$f(\mathbf{x}, n+1) = h\left(\mathbf{x}, n, f(\mathbf{x}, n)\right)$$

- regular $\mu$-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

- $\mu$-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 $\mu$-operation:

$$\mu x < n[A(x)] \coloneqq \mu x[A(x) \vee x = n]$$

## 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 $\mu$-operation, $\mu$-operation).

## Theorem

*The primitive recursive functions are closed under bounded $\mu$-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

$\boxed{\text{A function/relation is representable in Robinson } \mathrm{Q} \text{ iff it is computable.}}$

(proof sketch.) We have to show that all initial functions are representable, and the representable functions are closed under composition, regular $\mu$-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.

$$f(\mathbf{x}, 0) = g(\mathbf{x})$$
$$f(\mathbf{x}, n+1) = h(\mathbf{x}, n, f(\mathbf{x}, n))$$

we can code the sequence of values of $f$ from $0$ to $y$ by Gödel's $\beta$ function:

$$F(\mathbf{x}, y) = \mu z\big[\beta(z, 0) = g(\mathbf{x}) \ \wedge \ \forall i < y : \beta(z, i+1) = h\big(\mathbf{x}, i, \beta(z, i)\big)\big]$$

$$f(\mathbf{x}, y) = \beta\big(F(\mathbf{x}, y), y\big)$$

# 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

$\varphi_e$ is the $e^{\text{th}}$ partial recursive function[a]:

$$\varphi_e(\mathbf{x}) := U(\mu y T(e, \mathbf{x}, y))$$

[a] Sometimes we write $[\![e]\!]$ or $\{e\}$ for $\varphi_e$.

# 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 $\omega$-consistent Gödelian $\mathrm{T}$ is incomplete.*

## Proof.

Suppose $T$ is represented in $\mathrm{T}$ by $\gamma$.

$$\bar{\varphi}_e(n) \coloneqq \begin{cases} U(\mu y T(e,n,y)) & \text{if } \exists y T(e,n,y) \\ 0 & \text{if } \mathrm{T} \vdash \forall y \neg\gamma(e,n,y) \end{cases}$$

# Enumeration Theorem & s-m-n Theorem

## Theorem (Enumeration Theorem)

*The sequence* $\{\varphi_e^n\}_{e\in\omega}$ *is a partial recursive enumeration of the* $n$*-ary partial recursive functions, in the sense that:*

- *for each* $e$, $\varphi_e^n$ *is a partial recursive function of* $n$ *variables.*
- *if* $\psi$ *is a partial recursive function of* $n$ *variables, then there is* $e$ *s.t.* $\psi = \varphi_e^n$.
- *there is a partial recursive function* $\varphi$ *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 $\psi$ 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)

*$\psi$ is an acceptable numbering iff there is a recursive permutation $h$ s.t.*

$$\psi_e = \varphi_{h(e)}$$

## Theorem (Blum)

*If $\psi$ 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 $\mathrm{Prog}$ 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.

$$\text{target} := [\![\text{spec}]\!](\text{int}, \text{source})$$

$$\begin{aligned}
\text{out} &= [\![\text{source}]\!](\text{in}) \\
&= [\![\text{int}]\!](\text{source}, \text{in}) && \text{(Definition of an interpreter)} \\
&= [\![[\![\text{spec}]\!](\text{int}, \text{source})]\!](\text{in}) && \text{(Definition of a specializer)} \\
&= [\![\text{target}]\!](\text{in})
\end{aligned}$$

- Compiler generation by the second Futamura projection.

$$\text{compiler} := [\![\text{spec}]\!](\text{spec}, \text{int})$$

$$\begin{aligned}
\text{target} &= [\![\text{spec}]\!](\text{int}, \text{source}) \\
&= [\![[\![\text{spec}]\!](\text{spec}, \text{int})]\!](\text{source}) \\
&= [\![\text{compiler}]\!](\text{source})
\end{aligned}$$

- Compiler generator generation by the third Futamura projection.

$$\text{cogen} := [\![\text{spec}]\!](\text{spec}, \text{spec})$$

$$\begin{aligned}
\text{compiler} &= [\![\text{spec}]\!](\text{spec}, \text{int}) \\
&= [\![[\![\text{spec}]\!](\text{spec}, \text{spec})]\!](\text{int}) \\
&= [\![\text{cogen}]\!](\text{int})
\end{aligned}$$

# Futamura Projections

$$
\begin{array}{lclcl}
\text{out} & = & [\![\text{int}]\!](\text{source}, \text{in}) & = & [\![\text{target}]\!](\text{in}) \\
\text{target} & = & [\![\text{spec}]\!](\text{int}, \text{source}) & = & [\![\text{compiler}]\!](\text{source}) \\
\text{compiler} & = & [\![\text{spec}]\!](\text{spec}, \text{int}) & = & [\![\text{cogen}]\!](\text{int}) \\
\text{cogen} & = & [\![\text{spec}]\!](\text{spec}, \text{spec}) & = & [\![\text{cogen}]\!](\text{spec})
\end{array}
$$

- A program $\text{int}$ is an interpreter iff for program $p$ and data $d$:
$$[\![\text{int}]\!](p, d) = [\![p]\!](d)$$
- A program $\text{compiler}$ is a compiler iff for program $p$ and data $d$:
$$[\![[\![\text{compiler}]\!](p)]\!](d) = [\![p]\!](d)$$
- A program $\text{spec}$ 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 $e(y) := s(v, y)$.

$$\varphi^A_{e(y)} = \varphi^A_{s(v,y)} = \varphi^A_{\varphi^A_v(v,y)} = \varphi^A_{f(s(v,y),y)} = \varphi^A_{f(e(y),y)}$$

# 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 $A = \omega$, 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.

# Recursion Theorem

## Theorem (Second Recursion Theorem)

*If $f(x, y)$ is a partial recursive function, there is an index $e$ s.t.*

$$\varphi_e(y) = f(e, y)$$

$$\boxed{\text{Kleene's Fixpoint Theorem} \iff \text{Second Recursion Theorem}}$$

## Theorem (First Recursion Theorem)

*Every partial recursive functional $F(\alpha, x)$ admits a least fixpoint. In other words, there is a partial recursive function $\alpha$ 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* $\mathrm{T}$ *and a number* $n$ *such that* $A$ *admits a proof of length* $\leq n$ *in* $\mathrm{T}'$*, but no proof of length* $\leq h(n)$ *in* $\mathrm{T}$*.*

## Proposition

*If* $\mathrm{T}$ *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 $\mathrm{T}$ radically shortens some proof of some theorem of $\mathrm{T}$.

# Blum's Speed-up Theorem

A Blum complexity measure is a pair $(\varphi, \Phi)$ with $\varphi$ a Gödel numbering of the partial recursive functions $\mathbf{P}^{(1)}$ and a recursive function $\Phi : \mathbb{N} \to \mathbf{P}^{(1)}$ such that

- the domains of $\varphi_i$ and $\Phi_i$ 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 $(\varphi, \Phi)$ 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.

# $m$-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* $\#\operatorname{Th}\mathcal{N} := \{\#A : \mathcal{N} \vDash A\}$ *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.*

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
References1655

# 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) $60\%$ (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.

# 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

# 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?

$X$
$Y$
$X \in Y$
$X \in X$
$X \notin X$

$X = \{Y\}$
$Y = \{X\}$
$X$
$Y$
$X = \{\}$
$Y = \{X, X\}$
$X$
$Y$
$\Longrightarrow$
$X = \{\}$
$Y = \{\}$

$X = \{X, X, X\}$
$\Longrightarrow$
$X = \{\}$

# 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. 31131122211311123113 32...

C. 13211321322113311213212 3...

$$\begin{tikzcd} & A \arrow[dr] & \\ B \arrow[ur] & & C \arrow[ll] \end{tikzcd}$$

# Diagonalization[21]

## 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 $y : 1 \to Y$.*

$$\begin{tikzcd}
X \times Y^X \arrow[r, "\varepsilon"] & Y \arrow[dd, "\alpha"] \\
X \times X \arrow[u, "1_X \times f"] \arrow[ur, "\hat{f}"'] & \\
X \arrow[u, "\Delta"] \arrow[r, "g"'] & Y
\end{tikzcd}$$

[21] 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 $\alpha$ has no fixpoint, then $g$ is not representable by $f$.*
2. *If $g$ is representable by $f$, then $\alpha$ has a fixpoint.*

$$\begin{tikzcd}
X \times X \arrow[r, "f"] & Y \arrow[d, "\alpha"] \\
X \arrow[u, "\Delta"] \arrow[r, "g"'] & Y
\end{tikzcd}$$

$$\alpha\big(f\,(\ulcorner g \urcorner, \ulcorner g \urcorner)\big) = g\,(\ulcorner g \urcorner) = f\,(\ulcorner g \urcorner, \ulcorner g \urcorner)$$

- $\Delta : x \mapsto \langle x, x \rangle$ diagonal
- $f$ evaluation
- $\alpha$ "negation"
- $g$ ($\ulcorner g \urcorner$) fixpoint-(free) transcendence
- $f\,(\ulcorner g \urcorner, \ulcorner g \urcorner)$ self-reference "I have property $\alpha$."

# Lawvere's Fixpoint Theorem

| $f$ | $0$ | $1$ | $2$ | $3$ | $\cdots$ | $t$ | $\cdots$ |
|---|---|---|---|---|---|---|---|
| $0$ | $\alpha f(0,0)$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $f(0,t)$ | $\cdots$ |
| $1$ | $\cdots$ | $\alpha f(1,1)$ | $\cdots$ | $\cdots$ | $\cdots$ | $f(1,t)$ | $\cdots$ |
| $2$ | $\cdots$ | $\cdots$ | $\alpha f(2,2)$ | $\cdots$ | $\cdots$ | $f(2,t)$ | $\cdots$ |
| $3$ | $\cdots$ | $\cdots$ | $\cdots$ | $\alpha f(3,3)$ | $\cdots$ | $f(3,t)$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $t$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\boxed{\begin{array}{c} f(t,t) \\ \parallel \\ \alpha f(t,t) \end{array}}$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ |

# 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 $\alpha$ has no fixpoint, then $g$ is not representable by $f$.*
2. *If $g$ is representable by $f$, then $\alpha$ has a fixpoint.*

$$\begin{tikzcd}
X \times Y \arrow[r, "f"] & Z \arrow[d, "\alpha"] \\
X \arrow[u, "\langle 1_X, \beta \rangle"] \arrow[r, "g"'] & Z
\end{tikzcd}$$

# Lawvere's Fixpoint Theorem — Multi-Valued Version

- A *multi-valued function* $f : X \rightrightarrows Y$ is a function $f : X \to \mathrm{P}(Y)$ 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.

$$\begin{tikzcd}
X \times Y \arrow[r, shift left, "f"] \arrow[r, shift right] & Z \arrow[d, shift right] \arrow[d, shift left, "\alpha"] \\
X \arrow[u, "\langle 1_X, \beta \rangle"] \arrow[r] & Z
\end{tikzcd}$$

$$\beta x = y \implies \alpha(f(x, y)) \cap f(x, y) \neq \emptyset$$

# 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* $\varepsilon$*-representable by* $f$*:* $\exists y \in Y \forall x \in X : d(f(x, y), g(x)) < \varepsilon$, *then* $\alpha$ *has* $\varepsilon$*-fixpoint:* $\exists z \in Z : d(z, \alpha(z)) < \varepsilon$.

$$\begin{tikzcd}
X \times Y \arrow[r, "f"] & Z \arrow[d, "\alpha"] \\
X \arrow[u, "\langle 1_X, \beta \rangle"] \arrow[r, "g"'] & Z
\end{tikzcd}$$

# Example — Grelling/Liar/Quine. . . Paradox

$$\begin{tikzcd} X \times X \arrow[r, "f"] & \mathbf{2} \arrow[d, "\alpha"] \\ X \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{2} \end{tikzcd}$$

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

$$\begin{tikzcd} X \times X \arrow[r, "f"] & \mathbf{2} \arrow[d, "\alpha"] \\ X \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{2} \end{tikzcd}$$

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 | $\cdots$ | $\boxed{?}$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ |

# 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.

$$\begin{tikzcd}
X \times Y \arrow[r, "f"] & Z \arrow[d, "\alpha"] \\
X \arrow[u, "\langle 1_X, \beta \rangle"] \arrow[r, "g"'] & Z
\end{tikzcd}$$

# Example — Yablo Paradox in Linear Temporal Logic(LTL)

$$
\begin{aligned}
n \Vdash A \wedge B \quad &\Longleftrightarrow \quad n \Vdash A \;\&\; n \Vdash B \\
n \Vdash \neg A \quad &\Longleftrightarrow \quad n \nVdash A \\
n \Vdash \bigcirc A \quad &\Longleftrightarrow \quad n+1 \Vdash A \\
n \Vdash \Box A \quad &\Longleftrightarrow \quad \forall m \geq n \implies m \Vdash A
\end{aligned}
$$

$$
\begin{tikzcd}
\mathrm{Sen} \times \mathrm{Sen} \arrow[r, "f"] & \mathbf{2} \arrow[d, "\alpha"] \\
\mathrm{Sen} \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{2}
\end{tikzcd}
$$

$f : (X, Y) \mapsto [\![ X \nleftrightarrow \bigcirc\Box\neg Y ]\!]$ 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.*

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "f"] & \mathbf{2} \arrow[d, "\alpha"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{2}
\end{tikzcd}$$

where

$$f(m,n) = \begin{cases} 1 & \forall p \in \mathbb{P} : p \mid (m! + 1) \to p \leq n \\ 0 & \text{otherwise} \end{cases}$$

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 $\alpha$. 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.*

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "f"] & \mathbf{10} \arrow[d, "\alpha"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{10}
\end{tikzcd}$$

where $f : (m, n) \mapsto r_{mn} :=$ "the $n^{\text{th}}$ digit of the $m^{\text{th}}$ 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 $n^{\text{th}}$ digit of the real named by the $m^{\text{th}}$ 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 | $\cdots$ | $\boxed{?}$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ |

# Example — Cantor's Theorem

**Theorem (Cantor's Theorem)**

$$|X| < |\mathrm{P}(X)|$$

$$\begin{tikzcd}
X \times \mathrm{P}(X) \arrow[r, "f"] & \mathbf{2} \arrow[d, "\alpha"] \\
X \arrow[u, "\langle 1_X, \beta \rangle"] \arrow[r, "g"'] & \mathbf{2}
\end{tikzcd}$$

where $f : (x, y) \mapsto [\![ x \in y ]\!]$ and $\alpha : x \mapsto 1 - x$.

$\beta$ is not surjective.

another proof: assume $h : \mathrm{P}(X) \rightarrowtail X$.

$$\begin{tikzcd}
\mathrm{P}(X) \times \mathrm{P}(X) \arrow[r, "f"] & \mathbf{2} \arrow[d, "\alpha"] \\
\mathrm{P}(X) \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{2}
\end{tikzcd}$$

where $f : (x, y) \mapsto [\![ h(x) \in y ]\!]$, and $\alpha : x \mapsto 1 - x$.

$g$ is representable by $y \coloneqq \{ h(x) : x \subset X \ \& \ h(x) \notin x \}$.

# Example — Cantor's Theorem — another proof

If $|X| \geq |\mathrm{P}(X)|$, then there exists some enumeration $\{S_i\}_{i \in X}$ of $\mathrm{P}(X)$.

$$\begin{tikzcd}
X \times \{S_i\}_{i \in X} \arrow[r, "f"] & \mathbf{2} \arrow[d, "\alpha"] \\
X \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{2}
\end{tikzcd}$$

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 $R := \{x : x \notin S_x\}$ 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 | $\cdots$ | $\boxed{?}$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ |

# Example — Cantor's Theorem

**Theorem (Cantor's Theorem)**

*For* $|Y| \geq 2$,

$$|X| < |Y^X|$$

$$\begin{tikzcd}
X \times X \arrow[r, "f"] & Y \arrow[d, "\alpha"] \\
X \arrow[u, "\Delta"] \arrow[r, "g"'] & Y
\end{tikzcd}$$

where $\alpha$ is the cyclic permutation.

Every $g : X \to Y$ is representable by some $f : X \times X \to Y$ iff $\exists f : X \twoheadrightarrow Y^X$.

$\boxed{\textit{If there exists } f : X \twoheadrightarrow Y^X\textit{, then every } \alpha : Y \to Y \textit{ has a fixpoint.}}$

# 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 $\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)|$.

$$\begin{tikzcd}
\mathbb{R} \times C(\mathbb{R},\mathbb{R}) \arrow[r, "\mathcal{F}"] & \mathbb{R} \arrow[d, "\alpha"] \\
\mathbb{R} \arrow[u, "\langle 1_{\mathbb{R}},\beta\rangle"] \arrow[r, "g"'] & \mathbb{R}
\end{tikzcd}$$

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

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "f"] & \mathbb{N} \arrow[d, "\alpha"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbb{N}
\end{tikzcd}$$

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

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "f_\varphi"] & \{\varphi_n(m)\}_{(m,n)\in\mathbb{N}^2} \arrow[d, "\alpha"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \{\varphi_n(m)\}_{(m,n)\in\mathbb{N}^2}
\end{tikzcd}$$

where $f_\varphi : (m,n) \mapsto \varphi_n(m)$, and $\alpha : \varphi_n(m) \mapsto \min\big(\mathbb{N} \setminus \{\varphi_k(m) : k \leq n\}\big)$

$$g(m) = \min\big(\mathbb{N} \setminus \{\varphi_k(m) : k \leq m\}\big) = \mu n\,[K(n \mid m) > m]$$

$g$ unrepresentable $\implies$ $g$ uncomputable $\implies$ $K$ uncomputable

$$\Sigma(m) := \max\{\varphi_k(0) : k \leq m\} = \max\{n : K(n) \leq m\}$$

## Example — Not every partial recursive function is potentially recursive

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "f_\varphi"] & \{\varphi_n(m)\}_{(m,n)\in\mathbb{N}^2} \arrow[d, "\alpha"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \{\varphi_n(m)\}_{(m,n)\in\mathbb{N}^2}
\end{tikzcd}$$

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.*

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "H"] & \mathbf{2} \arrow[d, "\alpha"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{2}
\end{tikzcd}$$

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 | $\cdots$ | $\boxed{?}$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ |

**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.*

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "H"] & \mathbf{2} \arrow[d, "\alpha"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{2}
\end{tikzcd}$$

where $H : (m, n) \mapsto [\![\varphi_n(m) \downarrow]\!]$, and $\alpha(x) = \begin{cases} 1 & \text{if } x = 0 \\ \uparrow & \text{otherwise} \end{cases}$.

$$H(\ulcorner g \urcorner, \ulcorner g \urcorner) \uparrow$$

$\boxed{\text{There is no perfect anti-virus software.}}$

# 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

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "f_\varphi"] & \{\varphi_n(m)\}_{(m,n)\in\mathbb{N}^2} \arrow[d, "\alpha"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \{\varphi_n(m)\}_{(m,n)\in\mathbb{N}^2}
\end{tikzcd}$$

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. $\times$

# Example — Kleene's Fixpoint Theorem

## Theorem (Kleene's Fixpoint Theorem)

*Given a recursive function $h$, there is an index $e$ s.t.*

$$\varphi_e = \varphi_{h(e)}$$

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "f"] & \{\varphi_n\}_{n \in \mathbb{N}} \arrow[d, "\alpha_h"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \{\varphi_n\}_{n \in \mathbb{N}}
\end{tikzcd}$$

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

| $f$ | $0$ | $1$ | $2$ | $3$ | $\cdots$ | $t$ |
|---|---|---|---|---|---|---|
| $0$ | $\varphi_{h(\varphi_0(0))}$ | $\varphi_{\varphi_0(1)}$ | $\varphi_{\varphi_0(2)}$ | $\varphi_{\varphi_0(3)}$ | $\cdots$ | $\varphi_{\varphi_0(t)}$ |
| $1$ | $\varphi_{\varphi_1(0)}$ | $\varphi_{h(\varphi_1(1))}$ | $\varphi_{\varphi_1(2)}$ | $\varphi_{\varphi_1(3)}$ | $\cdots$ | $\varphi_{\varphi_1(t)}$ |
| $2$ | $\varphi_{\varphi_2(0)}$ | $\varphi_{\varphi_2(1)}$ | $\varphi_{h(\varphi_2(2))}$ | $\varphi_{\varphi_2(3)}$ | $\cdots$ | $\varphi_{\varphi_2(t)}$ |
| $3$ | $\varphi_{\varphi_3(0)}$ | $\varphi_{\varphi_3(1)}$ | $\varphi_{\varphi_3(2)}$ | $\varphi_{h(\varphi_3(3))}$ | $\cdots$ | $\varphi_{\varphi_3(t)}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $t$ | $\varphi_{\varphi_0(0)}$ <br> $\varphi_{h(\varphi_0(0))}$ | $\varphi_{\varphi_1(1)}$ <br> $\varphi_{h(\varphi_1(1))}$ | $\varphi_{\varphi_2(2)}$ <br> $\varphi_{h(\varphi_2(2))}$ | $\varphi_{\varphi_3(3)}$ <br> $\varphi_{h(\varphi_3(3))}$ | $\cdots$ | $\boxed{\varphi_{\varphi_t(t)} \quad \varphi_{h(\varphi_t(t))}}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |

# Example — Kleene's Fixpoint Theorem

## Theorem (Kleene's Fixpoint Theorem)

*Given a recursive function $h$, there is an index $e$ s.t.*

$$\varphi_e = \varphi_{h(e)}$$

## Proof.

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "f"] & \{\varphi_n\}_{n \in \mathbb{N}} \arrow[d, shift right] \arrow[d, shift left, "\alpha_h"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \{\varphi_n\}_{n \in \mathbb{N}}
\end{tikzcd}$$

where $f : (m, n) \mapsto \varphi_{\varphi_n(m)}$, and

$$\alpha_h : \varphi_n \mapsto \left\{\varphi_i : \varphi_i = \varphi_{h(n)}\right\}$$

# Example — $\mathbf{Y}$ Combinator

$$\begin{tikzcd}
\Lambda \times \Lambda \arrow[r, "f"] & \Lambda \arrow[d, "\alpha_y"] \\
\Lambda \arrow[u, "\Delta"] \arrow[r, "g"'] & \Lambda
\end{tikzcd}$$

where $f : (x, y) \mapsto yx$, and $\alpha_y : x \mapsto yx$.

$$g = \lambda x.y(xx)$$

$$gg = \alpha_y(gg)$$

$$\mathbf{Y} := \lambda y.gg = \lambda y.(\lambda x.y(xx))(\lambda x.y(xx))$$

$$\mathbf{Y}h = h(\mathbf{Y}h) = h(h(\mathbf{Y}h)) = \cdots$$

# Example — $\mathbf{Z}$ Combinator

$$\begin{tikzcd}
\Lambda \times \Lambda \arrow[r, "f"] & \Lambda \arrow[d, "\alpha_y"] \\
\Lambda \arrow[u, "\Delta"] \arrow[r, "g"'] & \Lambda
\end{tikzcd}$$

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 — $\Theta$ Combinator

$$\begin{tikzcd}
\Lambda \times \Lambda \arrow[r, "f"] & \Lambda \arrow[d, "\alpha"] \\
\Lambda \arrow[u, "\Delta"] \arrow[r, "g"'] & \Lambda
\end{tikzcd}$$

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 — $\Theta_v$ Combinator

$$\begin{tikzcd}
\Lambda \times \Lambda \arrow[r, "f"] & \Lambda \arrow[d, "\alpha"] \\
\Lambda \arrow[u, "\Delta"] \arrow[r, "g"'] & \Lambda
\end{tikzcd}$$

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 $\lambda$-term $F$ there is a $\lambda$-term $E$ s.t.*

$$F\ulcorner E\urcorner = E$$

$$\begin{tikzcd}
\underline{\Lambda} \times \underline{\Lambda} \arrow[r, "A"] & \Lambda \arrow[d, "\alpha_F"] \\
\underline{\Lambda} \arrow[u, "\Delta"] \arrow[r, "G"'] & \Lambda
\end{tikzcd}$$

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 wff $F(x)$ with one free variable $x$, there exists a sentence $E$ s.t.*

$$\mathrm{Q} \vdash E \leftrightarrow F(\ulcorner E \urcorner)$$

$$\begin{tikzcd}
\mathrm{Lin}_1 \times \mathrm{Lin}_1 \arrow[r, "f"] & \mathrm{Lin}_0 \arrow[d, "\alpha_F"] \\
\mathrm{Lin}_1 \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathrm{Lin}_0
\end{tikzcd}$$

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

$$G(x) := F(d(x))$$

where $d(n) := \begin{cases} \#M(\ulcorner M(x) \urcorner) & \text{if } n = \#M(x) \text{ for } M(x) \in \mathrm{Lin}_1 \\ 0 & \text{otherwise} \end{cases}$

is primitive recursive and is thus represented by some function symbol $d$.

$$E := G(\ulcorner G(x) \urcorner)$$

# 简化一下

$$
\begin{array}{cccccc}
\mathbb{N} & \ulcorner M(x) \urcorner & \ulcorner N(x) \urcorner & \ulcorner F(x) \urcorner & \ulcorner G(x) \urcorner & \cdots \\
\uparrow\downarrow & & & & & \\
\mathrm{Lin}_1 & M(x) & N(x) & F(x) & G(x) & \cdots
\end{array}
$$

- $d(n)$: **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

| $f$ | $\ulcorner M(x) \urcorner$ | $\ulcorner N(x) \urcorner$ | $\cdots$ | $\ulcorner G(x) \urcorner$ |
|---|---|---|---|---|
| $M(x)$ | $F(\ulcorner M(\ulcorner M(x) \urcorner) \urcorner)$ | $M(\ulcorner N(x) \urcorner)$ | $\cdots$ | $M(\ulcorner G(x) \urcorner)$ |
| $N(x)$ | $N(\ulcorner M(x) \urcorner)$ | $F(\ulcorner N(\ulcorner N(x) \urcorner) \urcorner)$ | $\cdots$ | $N(\ulcorner G(x) \urcorner)$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $G(x)$ | $G(\ulcorner M(x) \urcorner)$<br>$F(\ulcorner M(\ulcorner M(x) \urcorner) \urcorner)$ | $G(\ulcorner N(x) \urcorner)$<br>$F(\ulcorner N(\ulcorner N(x) \urcorner) \urcorner)$ | $\cdots$ | $\boxed{G(\ulcorner G(x) \urcorner)}$<br>$\boxed{F(\ulcorner G(\ulcorner G(x) \urcorner) \urcorner)}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |

# Strong Fixpoint Lemma

## Theorem (Strong Fixpoint Lemma — Jeroslow1973)

*For any wff $F(x)$ with one free variable $x$, there is a closed-term $t$ s.t.*

$$\mathrm{T} \vdash t = \ulcorner F(t) \urcorner$$

## Proof.

$$\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "d"] & \mathbb{N} \arrow[d, "\alpha_F"] \\
\mathbb{N} \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbb{N}
\end{tikzcd}$$

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: $E := F(t)$

| $d$ | $0$ | $1$ | $2$ | $\cdots$ | $k$ | $\cdots$ |
|---|---|---|---|---|---|---|
| $0$ | $0$ | $0$ | $0$ | $\cdots$ | $0$ | $\cdots$ |
| $1$ | $\llcorner 1 \lrcorner(0)$ | $\llcorner 1 \lrcorner(1)$ | $\llcorner 1 \lrcorner(2)$ | $\cdots$ | $\llcorner 1 \lrcorner(k)$ | $\cdots$ |
| $2$ | $\llcorner 2 \lrcorner(0)$ | $\llcorner 2 \lrcorner(1)$ | $\llcorner 2 \lrcorner(2)$ | $\cdots$ | $\llcorner 2 \lrcorner(k)$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $k$ | $\llcorner k \lrcorner(0)$ | $\llcorner k \lrcorner(1)$ | $\llcorner k \lrcorner(2)$ | $\cdots$ | $\llcorner k \lrcorner(k)$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ |

| $d$ | $0$ | $1$ | $2$ | $\cdots$ | $k$ | $\cdots$ |
|---|---|---|---|---|---|---|
| $0$ | $\ulcorner F(0) \urcorner$ | $0$ | $0$ | $\cdots$ | $0$ | $\cdots$ |
| $1$ | $\llcorner 1 \lrcorner(0)$ | $\ulcorner F(\llcorner 1 \lrcorner(1)) \urcorner$ | $\llcorner 1 \lrcorner(2)$ | $\cdots$ | $\llcorner 1 \lrcorner(k)$ | $\cdots$ |
| $2$ | $\llcorner 2 \lrcorner(0)$ | $\llcorner 2 \lrcorner(1)$ | $\ulcorner F(\llcorner 2 \lrcorner(2)) \urcorner$ | $\cdots$ | $\llcorner 2 \lrcorner(k)$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $k$ | $\llcorner k \lrcorner(0)$<br>$\ulcorner F(0) \urcorner$ | $\llcorner k \lrcorner(1)$<br>$\ulcorner F(\llcorner 1 \lrcorner(1)) \urcorner$ | $\llcorner k \lrcorner(2)$<br>$\ulcorner F(\llcorner 2 \lrcorner(2)) \urcorner$ | $\cdots$ | $\boxed{\llcorner k \lrcorner(k)}$<br>$\boxed{\ulcorner F(\llcorner k \lrcorner(k)) \urcorner}$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ |

# Fixpoint vs Diagonalization

$$\begin{tikzcd}
X \times X \arrow[r, "f"] & Y \arrow[d, "\alpha"] \\
X \arrow[u, "\Delta"] \arrow[r, "g"'] & Y
\end{tikzcd}$$

| Curry $\mathbf{Y}$ | $\hat{=}$ | $\lambda$-fixpoint | $\hat{=}$ | Gödel | $\hat{=}$ | Kleene | $\hat{=}$ | Russell |
|---|---|---|---|---|---|---|---|---|
| $yx$ | $\hat{=}$ | $N\,(\ulcorner M \urcorner)$ | $\hat{=}$ | $N(\ulcorner M(x) \urcorner)$ | $\hat{=}$ | $\varphi_n(m)$ | $\hat{=}$ | $x \in y$ |
| $xx$ | $\hat{=}$ | $M\,(\ulcorner M \urcorner)$ | $\hat{=}$ | $M(\ulcorner M(x) \urcorner)$ | $\hat{=}$ | $\varphi_n(n)$ | $\hat{=}$ | $x \in x$ |
| $y(xx)$ | $\hat{=}$ | $F\ulcorner M \ulcorner M \urcorner\urcorner$ | $\hat{=}$ | $F(\ulcorner M(\ulcorner M(x) \urcorner) \urcorner)$ | $\hat{=}$ | $h(\varphi_n(n))$ | $\hat{=}$ | $x \notin x$ |
| $\lambda x.y(xx)$ | $\hat{=}$ | $G$ | $\hat{=}$ | $G(x)$ | $\hat{=}$ | $\varphi_t(n)$ | $\hat{=}$ | $x \notin R$ |
| $(\lambda x.y(xx))(\lambda x.y(xx))$ | $\hat{=}$ | $G(\ulcorner G \urcorner)$ | $\hat{=}$ | $G(\ulcorner G(x) \urcorner)$ | $\hat{=}$ | $\varphi_t(t)$ | $\hat{=}$ | $R \notin R$ |

$$\boxed{\text{self-reference} \overset{?}{\Longrightarrow} \text{self-improvement}}$$

# Cantor's Diagonal Argument vs Three-Valued Logic

Let $Z = \{0, u, 1\}$.

$$\begin{tikzcd}
X \times Y \arrow[r, "f"] & Z \arrow[d, "\neg"] \\
X \arrow[u, "\Delta"] \arrow[r, "g"'] & Z
\end{tikzcd}$$

| $z$ | $\neg z$ |
|---|---|
| $0$ | $1$ |
| $u$ | $u$ |
| $1$ | $0$ |

$$\neg f\left(\ulcorner g \urcorner, \ulcorner g \urcorner\right) = g\left(\ulcorner g \urcorner\right) = f\left(\ulcorner g \urcorner, \ulcorner g \urcorner\right) \implies f\left(\ulcorner g \urcorner, \ulcorner g \urcorner\right) = u$$

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]

*The information available to the observer regarding his own state could have absolute limitations, by the laws of nature.*

*— John von Neumann*

$$\begin{tikzcd}
\mathbf{M} \times \mathbf{M} \arrow[r, "f"] & \mathbf{O} \arrow[d, "\alpha"] \\
\mathbf{M} \arrow[u, "\Delta"] \arrow[r, "g"'] & \mathbf{O}
\end{tikzcd}$$

- $\mathbf{M}$: quantum measurements.
- $\mathbf{O}$: possible outcomes of quantum measurements.

If we assume that it is not possible to measure properties without changing them (observer effect: $\alpha$ is fixpoint-free), then there is a limit to self-inspection.

# 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 $\Rightarrow$ 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 $R'_1(g, f, x, y) \coloneqq R'(gx, f(x, y))$.

$$\begin{tikzcd} X \times Y \arrow[r, "f"] & Z \arrow[d, "\alpha"] \\ X \arrow[u, "\langle 1_X, \beta \rangle"] \arrow[r] & Z \end{tikzcd}$$

$$\begin{aligned} R'(\alpha x, x) &\implies R'(\alpha \circ \pi_1 \circ \langle 1_Z, \beta \rangle x, x) \implies R'_1(d\pi_1 x, \pi_1(x, y)) \\ &\implies R'_1(d\pi_1, \pi_1, x, y) \\ R'_1(g, f, x, \beta x) &\implies R'(gx, f(x, \beta x)) \implies R(gx, \alpha(f(x, \beta x))) \\ &\implies R(gx, \alpha \circ f \circ \langle 1_Z, \beta \rangle x) \implies R(gx, dfx) \end{aligned}$$

# General Fixpoint Theorem $\Rightarrow$ 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 $\alpha$ has a fixpoint.*

$$\begin{tikzcd} X \times Y \arrow[r, "f"] & Z \arrow[d, "\alpha"] \\ X \arrow[u, "\langle 1_X, \beta \rangle"] \arrow[r] & Z \end{tikzcd}$$

## Proof.

$$d(f) := \alpha \circ f \circ \langle 1_X, \beta \rangle$$
$$R'(g, f, x, y) := gx = f(x, y)$$
$$R(x, y) := \alpha x = y$$

Since $df$ is representable by $f(-, y)$, we have $R'(df, f, x, y)$, and

$$R'(g, f, x, \beta x) \implies gx = f(x, \beta x) \implies \alpha(gx) = (\alpha \circ f \circ \langle 1_X, \beta \rangle)x$$
$$\implies R(gx, dfx)$$

# General Fixpoint Theorem $\Rightarrow$ 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 $\alpha$ has a fixpoint, i.e. $\exists z \in \alpha(z)$.*

$$\begin{tikzcd}
X \times Y \arrow[r, shift left, "f"] \arrow[r, shift right] & Z \arrow[d, shift right] \arrow[d, shift left, "\alpha"] \\
X \arrow[u, "\langle 1_X, \beta \rangle"] \arrow[r] & Z
\end{tikzcd}$$

## 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.

$$R(X,Y) := f(X) = Y$$

$$R'(X,Y) := \left(\forall x \in Y . x \subset f(x)\right) \wedge \left(\bigcup Y \subset X\right) \wedge \left(\bigcup Y \in Y\right) \wedge \left(f(X) \subset \bigcup Y\right)$$

$$\alpha(X) := \bigcup X$$

Consider $S := \left\{X \in \mathrm{P}(V) : X \subset f(X)\right\}$. 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.

## 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.

$$\begin{aligned} R(x, y) &\coloneqq f(F(x)) = F(y) \\ R'(x, y) &\coloneqq f(F(x)) = F(\langle f(y_n) \rangle_{n \in \omega}) \\ \alpha(x) &\coloneqq \langle f(x_n) \rangle_{n \in \omega} \end{aligned}$$

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. $f(\alpha) = \sup\left\{f(\gamma) : \gamma < \alpha\right\}$ if $\alpha$ is a limit ordinal.

## Theorem (Fixpoint Theorem for Normal Functions)

*Every normal function $f : \mathrm{Ord} \to \mathrm{Ord}$ has a fixpoint.*

## Proof.

Let $\mathcal{A} := \left\{x \in \mathrm{Ord}^\omega : x \text{ is increasing}\right\}$. 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 (Kleene's Fixpoint Theorem)

*Given a recursive function $h$, there is an index $e$ s.t.*

$$\varphi_e = \varphi_{h(e)}$$

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。

You can systematically change an infinite number of programs $\varphi_n \mapsto \varphi_{h(n)}$ 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 $K(x) := \mu e[\varphi_e(0) = x]$

## Theorem (Chaitin's Incompleteness Theorem)

*For any arithmetically sound Gödelian theory* $\mathrm{T}$, $\exists c \forall x : \mathrm{T} \nvdash K(x) > c$.

## Proof.

For any $m$, we can construct:

$$M_n := \text{“find } \mu y\big[\,\mathrm{prf}_{\mathrm{T}}\big(y, K(x) > m\big)\big], \text{output } x\text{”}$$

So there exists a recursive function $f : m \mapsto n$.
By Kleene's fixpoint theorem, there exists $e$ such that

$$M_e = M_{f(e)} = \text{“find } \mu y\big[\,\mathrm{prf}_{\mathrm{T}}\big(y, K(x) > e\big)\big], \text{output } x\text{”}$$

Take $c := e$.

**Remark:** For almost all random strings their randomness cannot be proved.

# Self-Reproducing Program/Quine

There is a program that outputs its own length.

- A Quine is a program which takes no input and outputs its own source code.
- Quines are algorithmic random.

There is a program that outputs its own source code.

## Corollary (Self-Reproducing Program)

*There is a recursive function $\varphi_e$ s.t. $\forall x : \varphi_e(x) = e$.*

## Quine in Python

```
exec(s:='print("exec(s:=%r)"%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 ($\psi$-introspective)

Given a total recursive function $\psi$,

- the $\psi$-*analysis* of $\varphi(x)$ is the code of the computation of $\varphi(x)$ to $\psi(x)$ steps.
- $\varphi$ is $\psi$-*introspective* at $x$ iff $\varphi(x)\downarrow$ and outputs its own $\psi$-analysis.
- $\varphi$ is *totally* $\psi$-*introspective* iff it is $\psi$-introspective at all $x$.

## Corollary

*There is a program that is totally $\psi$-introspective.*

## Proof.

Let $f(n,x) :=$ "the $\psi$-analysis of $\varphi_n(x)$".

# Introspective Program

There is a program that is totally introspective.

$$\boxed{\varphi_{\mathbf{e}} = \varphi_{\mathbf{h}(\mathbf{e})}}$$

| **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.

# Hofstadter — I am a Stange Loop

- Animate entities are those that, at some level of description, manifest a certain type of loopy pattern, which inevitably starts to take form if a system with the inherent capacity of perceptually filtering the world into discrete categories vigorously expands its repertoire of categories ever more towards the abstract.
- This pattern reaches full bloom when there comes to be a deeply entrenched self-representation — a story told by the entity to itself — in which the entity's "I" plays the starring role, as a unitary causal agent driven by a set of desires.

| | |
|---|---|
| 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 第一不完全性定理 | 我不可证 |
| Gödel-Rosser 不完全性定理 | 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** ($\text{AIXI}^S/\text{AIXI}^{t\ell}$) 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

$$u\big[s(t) \oplus \big(\mathrm{switchbit}(t) = 1\big), e(t)\big] > u\big[s(t) \oplus \big(\mathrm{switchbit}(t) = 0\big), e(t)\big]$$

- 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 |

$$\boxed{\varphi_{\mathbf{e}} = \varphi_{\mathbf{h(e)}}}$$

$$\boxed{\text{self-reference} \overset{?}{\Longrightarrow} \text{self-improvement}}$$

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: $\boxed{\text{Learning is to forget!}}$
5. Physical constraint: If we assume that it is not possible to measure properties without changing them (observer effect: $\alpha$ 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.

$$\Omega = \lim_{t\to\infty} \sum_{\ell(p)\le t \ \&\ U(p)\downarrow \text{ within time } t} 2^{-\ell(p)}$$

- 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.

# Jiang ZHANG: Causal Emergence

$$V_t^\pi(æ_{<k}) = Q_t(æ_{<k}\pi(æ_{<k}))$$

$$Q_t(æ_{<k}a_k) = \sum_{e_k \in \mathcal{E}} \mu(e_k \mid \check{æ}_{<k}\check{a}_k)\left[u_t(\check{æ}_{1:k}) + \gamma V_t^{\pi_{t+1}}(æ_{1:k})\right]$$

$$\pi_1^* := \operatorname*{argmax}_\pi V_1^\pi(\epsilon)$$

# The Universal program (warm-up): the petulant child

## The Petulant Child

Consider program $e$:

It searches for a proof in $\mathrm{PA}$ 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 $\mathrm{PA}$ 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. $\mathrm{PA}$ *proves that the set accepted by $e$ is finite.*
2. *For any finite $A \subset \mathbb{N}$, there is a model $\mathcal{M} \vDash \mathrm{PA}$ in which the program $e$ accepts exactly $A$.*
3. *Indeed, for any $A \subset \mathbb{N}$, there is a model $\mathcal{M} \vDash \mathrm{PA}$ 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. $\mathrm{PA}$ *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 $n^{\text{th}}$ number enumerated by $e$ is $f(n)$" are finitely consistent with $\mathrm{PA}$. 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. $\mathrm{PA}$ *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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
References1655

# 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 $2^{\#\mathbb{P}}$ different square-free parts. Furthermore, $s \leq \sqrt{N}$. Hence $N \leq \#\left\{(r, s) : rs^2 \leq N \text{ and } r \text{ is squre-free}\right\} \leq 2^{\#\mathbb{P}}\sqrt{N}$.

## 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 $\prod_{k=0}^{n-1} F_k = F_n - 2$, 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

$$N_{a,b} := \{an + b : n \in \mathbb{Z}\}$$

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 $f(n) := \#\{k \in \mathbb{P} : p \mid n\}$, and $P := \prod_{p \in \mathbb{P}} p$. 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 $P := \prod_{i=2}^{n} p_i$. 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* $\mathrm{T}$, $\exists c \forall x : \mathrm{T} \nvdash K(x) > c$.

"given $n$, find $\mu y\big[\,\mathrm{prf}_{\mathrm{T}}\big(y, K(x) > n\big)\big]$, 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 $\mu y\big[\,\mathrm{prf}_{\mathrm{T}}\big(y, K(x) > e\big)\big]$, ouput $x$" (Berry Paradox)

## Theorem (Chaitin)

*For any arithmetically sound Gödelian* $\mathrm{T}$, $\#\big\{x : \mathrm{T} \vdash K(x) > \ell(x)\big\} < \infty$.

**Remark:** For almost all random strings their randomness cannot be proved.

# Incompressibility vs Incompleteness vs Intelligence

- $P(x) := \{p \in \mathcal{X}^* : \exists t \forall k \geq t\, (p(x_{1:k}) = x_{k+1})\}$
- $P(A) := \bigcap_{x \in A} P(x)$
- $P_n := P(\{x : K(x) \leq n\})$

- $\forall n \exists p \in P_n : K(p) \overset{+}{\leq} n + O(\log n)$
- $\forall n : p \in P_n \implies K(p) \overset{+}{\geq} n$

## Theorem (Legg)

*For any arithmetically sound Gödelian* $\mathrm{T}$, $\exists c \forall n \geq c \forall p : \mathrm{T} \nvdash p \in P_n$.

"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)

$$
\begin{aligned}
E_t &:= \{p : U(p) \downarrow \text{ in at most } t \text{ steps}\} \\
E &:= E_\infty \\
\Omega^t &:= \sum_{p \in E_t} 2^{-\ell(p)} \\
\Omega &:= \Omega^\infty \\
t(n) &:= \mu t[\Omega^t \geq \Omega_{1:n}]
\end{aligned}
$$

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 $\Omega$.

# 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:** $\Omega$ is a "philosopher's stone". Knowing $\Omega$ 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 $\Omega$

## Theorem (Randomness of $\Omega$)

$$\exists c \forall n : K(\Omega_{1:n}) \geq n - c$$

## 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*

$$L(n, x_0, x_1, \ldots, x_m) = R(n, x_0, x_1, \ldots, x_m)$$

*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,
hence there exists $L(y, x_0, x_1, \ldots, x_m) = R(y, x_0, x_1, \ldots, x_m)$ s.t.

$$\langle n, k\rangle \in A \iff \exists!\langle x_1, \ldots, x_m\rangle\big(L(n, k, x_1, \ldots, x_m) = R(n, k, x_1, \ldots, x_m)\big)$$

Thereby, $L(n, k, x_1, \ldots, x_m) = R(n, k, x_1, \ldots, x_m)$ has exactly one solution $x_1, \ldots, x_m$ if $\Omega_n^k = 1$, and it has no solution if $\Omega_n^k = 0$.

## Theorem (Chaitin $\Omega$ Incompleteness)

*For any arithmetically sound Gödelian* $\mathrm{T}$, $\mathrm{T}$ *can determine at most finitely many (scattered) bits of* $\Omega$.

## Proof.

Assume $\mathrm{T}$ can provide infinitely many bit of $\Omega$.
Any $k$ different bits $i_1, i_2, \ldots, i_k$ of $\Omega$ give us a covering $A_k$ of measure $2^{-k}$ which includes $\Omega$.

$$A_k \coloneqq \left\{ s_1 \Omega_{i_1} \cdots s_k \Omega_{i_k} 2^\omega : \forall 1 \leq j \leq k \left( s_j \in 2^{<\omega} \ \& \ \ell(s_j) = i_j - i_{j-1} - 1 \right) \right\}$$

$$\mu(A_k) = \frac{2^{i_k - k}}{2^{i_k}} = 2^{-k}$$

Thereby,

$$\forall k \left[ \mu(A_k) \leq 2^{-k} \ \ \& \ \ \Omega \in A_k \right]$$

which contradicts the Martin-Löf randomness of $\Omega$.

## Definition (Busy Beaver)

$$\Sigma(n) := \max\{x : K(x) \leq n\}$$
$$\sigma(n) := \mu t \left[\Omega^t_{1:n} = \Omega_{1:n}\right]$$

## 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 $\sigma(n) := \mu t \left[\Omega^t_{1:n} = \Omega_{1:n}\right]$ 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* $\mathrm{T}$, $\exists n \forall x : \mathrm{T} \nvdash \sigma(n) \leq x$.

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
References1655

- 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 wff $F(x)$ 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 $\mathrm{T}$ that is a primitively recursive axiomatised extension of $\mathrm{PA}$ 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。
- 无名，天地之始；有名，万物之母。
- 故常无，欲以观其妙；常有，欲以观其徼。
- 此两者，同出而异名，同谓之玄。
- 玄之又玄，众妙之门。

The theory that can be formulated can't be the ultimate theory. The formulated theory of categories evolves, and its projection on reality changes. The unformulatable ultimate theory is the truth of universe. The formulated theory is the basis to describe all the matter. In search of the unformulatable ultimate theory, we give meaning to life. Within the formulated theory, we study its limits. The gap between the formulatable and the unformulatable is a mystery. 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* $\mathrm{T}$ *is consistent,* $\mathrm{T} \nvdash G$
2. *if* $\mathrm{T}$ *is* $\omega$*-consistent,* $\mathrm{T} \nvdash \neg G$

## Gödel's First Incompleteness Theorem

$$F(x) := \neg\Box x$$

$$G = \text{“I am not provable.”}$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \neg\Box G$$

## Proof.

$G \leftrightarrow \neg\Box G \implies \Box G \to \neg G \implies \Box(\Box G \to \neg G) \implies \Box G \to \Box\neg G \implies$
$\Box G \to \Box(G \wedge \neg G) \implies \neg\Box\bot \to \neg\Box G$

# 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**

$$\Box^R x := \exists y \big[ \mathrm{prf}_{\mathrm{T}}(y, x) \wedge \neg \exists z < y \ \mathrm{prf}_{\mathrm{T}}\big(z, N(x)\big)\big]$$

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 $\mathrm{truth}$ Theorem

**Theorem (Tarski's Undefinability of $\mathrm{truth}$ 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 wff $F(x, y)$, there is a wff $G(x)$ 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 wff $F(x, y, z)$, there is a wff $G(x)$ 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

## Yablo Paradox

$$F(x, y, z) \coloneqq y > x \to \neg\Box z$$

$$\mathrm{T} \vdash \forall x\big(Y(x) \leftrightarrow \forall y > x \neg\Box Y(y)\big)$$

- $\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \forall x \neg\Box Y(x)$
- $\mathrm{T} \vdash \forall x\big(Y(x) \leftrightarrow \mathrm{Con}_{\mathrm{T}}\big)$
- $\mathrm{T} \vdash \forall x \forall y\big(Y(x) \leftrightarrow Y(y)\big)$

# Provability Conditions

## 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

$$F(x) := \Box x \to A$$

$$G = \text{“If I am provable, then } A\text{.”}$$

## 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.”

# Gödel's Second Incompleteness Theorem

## Theorem (Gödel's Second Incompleteness Theorem)

*For any Gödelian* $\mathrm{T} \supset \mathrm{PA}$, $\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \neg\Box\,\mathrm{Con}_{\mathrm{T}}$.

## Proof.

$$\mathrm{T} \vdash \Box(\Box\bot \to \bot) \to \Box\bot \implies \mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \neg\Box\,\mathrm{Con}_{\mathrm{T}}$$

$$\mathrm{PA} \nvdash \neg\Box_{\mathrm{PA}}\,\mathrm{Con}(\mathrm{PA})$$

$$\mathrm{PA} \nvdash \mathrm{Con}(\mathrm{PA})$$

$$\mathrm{PA}^* \vdash \neg\,\mathrm{Con}(\mathrm{PA}^*) \text{ where } \mathrm{PA}^* = \mathrm{PA} + \neg\,\mathrm{Con}(\mathrm{PA})$$

**Remark:** The second incompleteness theorem does not imply that the consistency of a system $\mathrm{T}$ can only be proved in a stronger system.

# Gödel's Second Incompleteness Theorem

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \mathrm{Con}(\mathrm{T} + \neg\,\mathrm{Con}_{\mathrm{T}})$$

Proof.

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \neg\Box\,\mathrm{Con}_{\mathrm{T}}$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \neg\Box(\neg\,\mathrm{Con}_{\mathrm{T}} \to \bot)$$
$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \mathrm{Con}(\mathrm{T} + \neg\,\mathrm{Con}_{\mathrm{T}})$$

*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.

$$\mathrm{T} \vdash G \leftrightarrow \neg\Box G$$

$$\mathrm{T} \vdash G \to (\Box G \to \bot)$$

$$\mathrm{T} \vdash \Box G \to \Box(\Box G \to \bot)$$

$$\mathrm{T} \vdash \Box G \to \Box\bot$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \neg\Box G$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \implies \mathrm{T} \vdash \neg\Box G$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \implies \mathrm{T} \vdash G$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \implies \mathrm{T} \vdash \Box G$$

$$\mathrm{T} \nvdash \mathrm{Con}_{\mathrm{T}}$$

$$\boxed{\text{second incompleteness} \implies \text{Löb}}$$

## Löb's Theorem

$$\mathrm{T} \vdash \Box A \to A \iff \mathrm{T} \vdash A$$

## Proof.

Assume $\mathrm{T} \nvdash A$.

Then $\mathrm{T} + \neg A$ is consistent.

$$\mathrm{T} + \neg A \nvdash \mathrm{Con}(\mathrm{T} + \neg A)$$

$$\mathrm{T} + \neg A \nvdash \neg\Box(\neg A \to \bot)$$

$$\mathrm{T} + \neg A \nvdash \neg\Box A$$

$$\mathrm{T} \nvdash \Box A \to A$$

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 $\mathrm{T}$.
2. For any Gödelian $\mathrm{T} \supset \mathrm{PA}$, $\top$ is the only fixpoint of $\Box x$ up to the logical equivalence in $\mathrm{T}$.

# Surprise Exam Paradox vs Second Incompleteness Theorem

$$\boxed{\mathrm{T} \nvdash \mathrm{Con}_{\mathrm{T}}}$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \to \forall x \neg \Box\big(K(x) > c\big) \qquad \text{(Chaitin)}$$

$$\mathrm{T} \vdash \mathrm{Con}_{\mathrm{T}} \implies \mathrm{T} \vdash \forall x \in \mathcal{X}^{c+1} \neg \Box\big(K(x) > c\big)$$

$$\mathrm{T} \vdash \forall x \in \mathcal{X}^{c+1} \big[K(x) \leq c \to \Box\big(K(x) \leq c\big)\big] \qquad (\Sigma_1\text{-complete})$$

$$m \coloneqq \left|\left\{x \in \mathcal{X}^{c+1} : K(x) > c\right\}\right|$$

$$\mathrm{T} \vdash 1 \leq m \leq 2^{c+1}$$

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: $A \to \Box A$
- $\omega$-completeness: $\forall x \Box A(\dot{x}) \to \Box \forall x A(x)$

$$\omega\text{-complete} \implies \omega\text{-consistent} \implies 1\text{-consistent} \implies \text{consistent}$$

## Theorem

*For any Gödelian* $\mathrm{T} \supset \mathrm{PA}$*, the following are equivalent in* $\mathrm{T}$*.*

1. $\neg\,\mathrm{Con}_{\mathrm{T}}$
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 $\mathrm{Q}$ / $\mathrm{PA}$ / $\mathrm{ZFC}$

# Parikh Sentences

## 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.

$$\mathrm{prflen}_{\mathrm{T}}(m) := \text{“the length of the proof encoded by } m\text{”}$$

$$\Box^n x := \exists m\big(\,\mathrm{prf}_{\mathrm{T}}(m,x) \wedge \mathrm{prflen}_{\mathrm{T}}(m) < n\big)$$

$$F(x) := \neg\Box^n x$$

$$G = \text{“I have no proof of myself shorter than } n\text{.”}$$

$$\neg\Box G \implies G \implies \Box G$$

# 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

$\boxed{\mathrm{Q} \text{ is incomplete.}}$

**Theorem ($\Sigma_1$-completeness of Robinson Arithmetic $\mathrm{Q}$)**

*For any Gödelian $\mathrm{T} \supset \mathrm{Q}$, and any sentence $A \in \Sigma_1$, $\mathrm{Q} \vdash A \to \Box A$.*

**Theorem (Strong Undecidability of $\mathrm{Q}$)**

*If $\mathrm{T} \cup \mathrm{Q}$ is consistent, then $\mathrm{T}$ 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.

$\boxed{\text{First order logic } \mathcal{L}_{\omega\omega} \text{ is undecidable.}}$

**Theorem (Church1936, Turing1936)**

*The set of valid sentences is recursively enumerable but undecidable.*

# Undecidability

## 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 structures 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 wff $A$ 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.*

# Undecidability

## 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 $10^{\text{th}}$ 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

$$\underbrace{\Gamma_r \cup \Gamma_m \cup \Gamma_b \vdash A \qquad\qquad \mathcal{M}_r \vDash \Gamma_r}_{\Big\Downarrow ?}$$

$$\mathcal{M}_r \vDash A$$

$$\frac{\Gamma_r \cup \Gamma_r' \vdash A \quad \mathcal{M}_r \vDash \Gamma_r \cup \Gamma_r'}{\mathcal{M}_r \vDash A}$$

where $\Gamma_r' \subset \Gamma_m \cup \Gamma_b$

# When you come across the Hydra — “natural” vs “ad-hoc”

## 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* $\mathrm{PA}$ *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

$$\begin{aligned} m_0 &\coloneqq m \\ m_k &\coloneqq G_{k+1}(m_{k-1}) \\ f_G(m) &\coloneqq \mu k[m_k = 0] \end{aligned}$$

$$\begin{aligned} m_0 &= 266 = 2^{2^{2+1}} + 2^{2+1} + 2^1 \\ m_1 &= G_2(m_0) = 3^{3^{3+1}} + 3^{3+1} + 2 \approx 10^{38} \\ m_2 &= G_3(m_1) = 4^{4^{4+1}} + 4^{4+1} + 1 \approx 10^{616} \\ m_3 &= G_4(m_2) = 5^{5^{5+1}} + 5^{5+1} \approx 10^{10000} \end{aligned}$$

# 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 $\alpha$, 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 $f_\alpha$ with $\alpha < \varepsilon_0$ is computable, and provably total in $\mathrm{PA}$.
- If $f$ is computable and provably total in $\mathrm{PA}$, then $\exists \alpha < \varepsilon_0 : f < f_\alpha$. Hence $f_{\varepsilon_0}$ is not provably total in $\mathrm{PA}$.

# 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 $\mathrm{T} : \exists n \forall m : \mathrm{T} \nvdash \Sigma(n) \leq m$

# Paris-Harrington Theorem vs Ramsey Theorem

$$[A]^n := \{X \subset A : |X| = n\}$$

$$\kappa \to (\lambda)^n_m := \forall F : [\kappa]^n \to m\Big(\exists H \subset \kappa\Big(|H| = \lambda \ \wedge \ \exists i \in m\Big([H]^n \subset F^{-1}(i)\Big)\Big)\Big)$$

Ramsey theorem: $\forall mn \in \omega : \aleph_0 \to (\aleph_0)^n_m$

$$s \to (k_0, \ldots, k_{m-1})^n_m := \forall F : [s]^n \to m\left(\bigvee_{i=0}^{m-1} \exists H \subset s\Big(|H| = k_i \ \wedge \ [H]^n \subset F^{-1}(i)\Big)\right)$$

$$s \underset{*}{\to} (k)^n_m := \forall F : [s]^n \to m\Big(\exists H \subset s\Big(|H| \geq \min(H) \ \wedge \ |H| \geq k \ \wedge \ \exists i \in m\Big([H]^n \subset F^{-1}(i)\Big)\Big)\Big)$$

$$\mathrm{ZFC} \vdash \forall mnk \exists s \left(s \underset{*}{\to} (k)^n_m\right)$$

$$\mathrm{PA} \nvdash \forall mnk \exists s \left(s \underset{*}{\to} (k)^n_m\right)$$

$$\forall \alpha < \varepsilon_0 \, (f_\alpha < f_R) \quad \text{where} \quad f_R(m, n, k) := \mu s \left[s \underset{*}{\to} (k)^n_m\right]$$

# Leibniz — Hilbert — Gödel — Turing $\cdots$

- $G_{\mathrm{T}}$ := "This sentence cannot be proved in the formal axiomatic system $\mathrm{T}$"
- We humans can easily see that $G_{\mathrm{T}}$ must be true.
- Since any AI is a FAS $\mathrm{T}$, no AI can prove $G_{\mathrm{T}}$. — *Penrose*
- Therefore there are things humans, but no AI system can do.
- $P$ := "Penrose can't consistently assert that this sentence is true"
- Penrose cannot assert $P$, 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*

02/1530

# Math-Matter-Mind (Penrose)

$$\begin{tikzcd}
\text{Mind} \arrow[rr, Rightarrow] & & \text{Math} \arrow[dl, Rightarrow] \\
& \text{Matter} \arrow[ul, Rightarrow] &
\end{tikzcd}$$

$$\begin{tikzcd}
\text{Mind} & & \text{Math} \\
& ? \arrow[ul, Rightarrow] \arrow[ur, Rightarrow] \arrow[d, Rightarrow] & \\
& \text{Matter} &
\end{tikzcd}$$

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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

References1655

# 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.

# Syntax

## Language

$$\mathscr{L}^{=} := \{=, (, )\} \cup \mathrm{Var} \cup \mathrm{Fun}$$

where

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

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

$f^n$ is an $n$-place function symbol.
A $0$-place function symbol $f^0$ is called constant.

# Term & Formula

## Term

$$t ::= x \mid c \mid f(t, \ldots, t)$$

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

## Well-Formed Formula Wff

$$A ::= s = t$$

where $s, t \in \mathrm{Term}$.

# Semantics

A structure 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 $[\![f]\!] : M^n \to M$.
We write $\mathcal{M} = (M, [\![c]\!], [\![f]\!])$ or $\left(M, c^{\mathcal{M}}, f^{\mathcal{M}}\right)$ for convenience.
An interpretation $(\mathcal{M}, \nu)$ is a structure $\mathcal{M}$ with a variable assignment $\nu : \mathrm{Var} \to M$.
We extend $\nu$ to $\overline{\nu} : \mathrm{Term} \to M$ by recursion as follows:

## Assignment over Terms

- $\overline{\nu}(x) = \nu(x)$
- $\overline{\nu}(c) = [\![c]\!]$
- $\overline{\nu}(f(t_1, \ldots, t_n)) = [\![f]\!](\overline{\nu}(t_1), \ldots, \overline{\nu}(t_n))$

$$\begin{tikzcd}
\mathrm{Term} \arrow[r, "\overline{\nu}"] \arrow[d, "f"'] & M \arrow[d, "{[\![f]\!]}"] \\
\mathrm{Term} \arrow[r, "\overline{\nu}"'] & M
\end{tikzcd}$$

# Semantics

- $\mathcal{M}, v \vDash s = t$ iff $\overline{v}(s) = \overline{v}(t)$. (Satisfaction)
- $\mathcal{M} \vDash s = t$ iff for all $v : \mathcal{M}, v \vDash s = t$. (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

$$\frac{}{t = t}\ [\text{refl}] \qquad \frac{s = t}{t = s}\ [\text{symm}]$$

$$\frac{r = s \qquad s = t}{r = t}\ [\text{trans}] \qquad \frac{s = t}{r(\ldots s \ldots) = r(\ldots t \ldots)}\ [\text{rep}]$$

$$\frac{r(x_1, \ldots, x_n) = s(x_1, \ldots, x_n)}{r[t_1/x_1, \ldots, t_n/x_n] = s[t_1/x_1, \ldots, t_n/x_n]}\ [\text{subst}]$$

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 $\mathrm{T}$ iff $s = t$ is the last member of some deduction sequence from $\mathrm{T}$.

# Meta-Theorems

**Theorem (Soundness & Completeness)**

$$\mathrm{T} \vdash s = t \iff \mathrm{T} \vDash s = t$$

- Determining Validity? Undecidable!

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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

References1655

# 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. $\forall xy \exists z : xz = y$
3. $\forall xy \exists z : zx = y$

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. $\forall x \exists y : x \cdot y = y \cdot x = e$

Structure

- $(\mathbb{Z}, 0, +)$
- $(\mathbb{Q} \setminus \{0\}, 1, \times)$
- 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

$$\begin{array}{ccccccc} \dagger & \clubsuit & \divideontimes & \oint & \blacklozenge & \nabla & \$ \\ \hline \divideontimes & \blacklozenge & \nabla & \clubsuit & \oint & \$ & \dagger \end{array} \iff \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 3 & 5 & 6 & 2 & 4 & 7 & 1 \end{array} = (1,3,6,7)(2,5,4)$$

$$\{e, r, r^2, m, mr, mr^2\}$$

*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 $\mathscr{L} = \{0, 1, \oplus, -, \odot, ^{-1}\}$

- $\forall xy : x \odot y = y \odot x$
- $\forall x \neq 0 : x \odot x^{-1} = x^{-1} \odot x = 1$

# Boolean Algebra

Boolean Algebra $\mathscr{L} = \{0, 1, +, \cdot, ^{-}\}$

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

$$\begin{tikzcd}
\text{proposition} \arrow[r, "\text{Logic}", "deduce"'] \arrow[d, "\text{translate}"'] & \text{proposition} \\
\text{equation} \arrow[r, "calculate", "\text{Algebra}"'] & \text{equation} \arrow[u, "\text{translate}"']
\end{tikzcd}$$

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

$$(\mathrm{P}(A), \emptyset, A, \cup, \cap, \overline{\phantom{x}})$$

$$x \oplus y \coloneqq (x \cdot \overline{y}) + (\overline{x} \cdot y)$$

$$x = y \iff x \oplus y = 0$$

# 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

$$f(\mathbf{z}) := \left(\sum_{i=1}^{k} f_i(1, \ldots, 1, \mathbf{z})\right) \cdot \ldots \cdot \left(\sum_{i=1}^{k} f_i(0, \ldots, 0, \mathbf{z})\right)$$

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

$\mathbf{A}: x \cdot \overline{y} = 0;$ $\mathbf{E}: x \cdot y = 0;$ $\mathbf{I}: v = v \cdot x \cdot y;$ $\mathbf{O}: v = v \cdot x \cdot \overline{y}$

$$\frac{MAP \quad SAM}{SAP}$$

$$\underbrace{m \cdot \overline{p} = 0 \qquad s \cdot \overline{m} = 0}$$

$$\Downarrow \textit{Reduction}$$

$$m \cdot \overline{p} + s \cdot \overline{m} = 0$$

$$\Downarrow \textit{Elimination}$$

$$(1 \cdot \overline{p} + s \cdot 0) \cdot (0 \cdot \overline{p} + s \cdot 1) = 0$$

$$\Downarrow \textit{Expansion}$$

$$(1 \cdot 0 + 1 \cdot 0) \cdot (0 \cdot 1 + 1 \cdot 1) \cdot s \cdot p + \ldots + (1 \cdot 1 + 0 \cdot 0) \cdot (0 \cdot 1 + 0 \cdot 1) \cdot \overline{s} \cdot \overline{p} = 0$$

$$\Downarrow$$

$$s \cdot \overline{p} = 0$$

$$s \cdot \overline{p} = s \cdot 1 \cdot \overline{p} = s \cdot (m + \overline{m}) \cdot \overline{p} = s \cdot m \cdot \overline{p} + s \cdot \overline{m} \cdot \overline{p} = 0 + 0 = 0$$

# 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** : $x \cdot \overline{y} \neq 0$.

$$\frac{\begin{array}{c} PAM \\ SOM \end{array}}{SOP}$$

$$p \cdot \overline{m} = s \cdot \overline{m} \cdot p + \overline{s} \cdot \overline{m} \cdot p = 0$$

$$s \cdot \overline{m} = s \cdot \overline{m} \cdot p + s \cdot \overline{m} \cdot \overline{p} \neq 0$$

$$\Downarrow$$

$$p \cdot \overline{m} + s \cdot \overline{m} = s \cdot \overline{m} \cdot \overline{p} \neq 0$$

$$\Downarrow$$

$$s \cdot \overline{m} \cdot \overline{p} + s \cdot m \cdot \overline{p} = s \cdot \overline{p} \neq 0$$

$$\frac{\begin{array}{c} MEP \\ MIS \end{array}}{SOP}$$

$$m \cdot p = s \cdot m \cdot p + \overline{s} \cdot m \cdot p = 0$$

$$m \cdot s = s \cdot m \cdot p + s \cdot m \cdot \overline{p} \neq 0$$

$$\Downarrow$$

$$m \cdot p + m \cdot s = s \cdot m \cdot \overline{p} \neq 0$$

$$\Downarrow$$

$$s \cdot m \cdot \overline{p} + s \cdot \overline{m} \cdot \overline{p} = s \cdot \overline{p} \neq 0$$

# Propositional Logic vs Boolean Algebra

$$
\begin{aligned}
(\bot)^* &:= 0 \\
(\top)^* &:= 1 \\
(p)^* &:= p \\
(\neg A)^* &:= \overline{A^*} \\
(A \vee B)^* &:= A^* + B^* \\
(A \wedge B)^* &:= A^* \cdot B^*
\end{aligned}
$$

$$
\begin{aligned}
(0)^\star &:= \bot \\
(1)^\star &:= \top \\
(p)^\star &:= p \\
\left(\overline{A}\right)^\star &:= \neg A^\star \\
(A + B)^\star &:= A^\star \vee B^\star \\
(A \cdot B)^\star &:= A^\star \wedge B^\star
\end{aligned}
$$

$$\frac{A \vdash B}{\overline{\overline{A^* \leq B^*}}} \qquad \frac{A \leq B}{\overline{\overline{A^\star \vdash B^\star}}}$$

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

# Boolean Algebra vs Propositional Logic

## Exercise

Alice, Ben, Charlie, and Diane are considering going to a Halloween party.

1. If Alice goes then Ben won't go and Charlie will.
2. If Ben and Diane go, then either Alice or Charlie (but not both) will go.
3. If Charlie goes and Ben does not, then Diane will go but Alice will not.

$$A \rightarrow \neg B \wedge C \qquad A \cdot (B + \overline{C}) = 0$$

$$B \wedge D \rightarrow (A \wedge \neg C) \vee (\neg A \wedge C) \qquad B \cdot D \cdot (\overline{A} \cdot \overline{C} + A \cdot C) = 0$$

$$\neg B \wedge C \rightarrow \neg A \wedge D \qquad \overline{B} \cdot C \cdot (A + \overline{D}) = 0$$

# General Solution?[22]

## Exercise — Save Yourself

You can say one sentence. If you lie I will hang you. If you tell the truth I will shoot you.

$$x = ? \implies \vDash (\neg x \to h) \land (x \to s) \land (s \leftrightarrow \neg h) \to \neg h \land \neg s$$

$$\vDash (\neg h \to h) \land (h \to s) \land (s \leftrightarrow \neg h) \to \neg h \land \neg s$$

## Problem (General Solution?)

$$x = ? \implies \vDash A(x)$$

[22] Frank Markham Brown: Boolean Reasoning — The Logic of Boolean Equations.

# General Solution?

$$x^5 - x - 1 = 0 \implies x = ?$$

There is no solution in radicals to general polynomial equations of degree five or higher with arbitrary coefficients.

$$ax^2 + bx + c = 0 \implies x = ?$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

# General Solution of Boolean Equation

## Theorem (General Solution of Boolean Equation)

*Assume* $f(x) = 0$ *is consistent (it has at least one solution, i.e.* $f(0) \cdot f(1) = 0$*), then*

$$f(x) = 0$$

$$\Updownarrow$$

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

$$\Updownarrow$$

$$x = f(0) + \theta \cdot \overline{f(1)}$$

*where* $\theta \in \{0, 1\}$.

# 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.

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

# 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.

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

## Solution

$$(a \cdot x + \overline{a} \cdot \overline{x}) \cdot \overline{k} = 0 \quad \implies \quad x = \overline{a} \cdot \overline{k} + \theta \cdot (\overline{a} + k)$$

$$\vDash (a \leftrightarrow \neg a \wedge \neg k) \to k$$

$$\vDash (a \leftrightarrow (a \to k)) \to k$$

# General Solution of Boolean Equation

## Theorem (General Solution of Boolean Equation)

*Given the Boolean function* $f : \{0,1\}^n \to \{0,1\}$, *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*

$$\begin{aligned} f_n &\coloneqq f \\ f_{i-1}(x_1, \ldots, x_{i-1}) &\coloneqq f_i(x_1, \ldots, x_{i-1}, 0) \cdot f_i(x_1, \ldots, x_{i-1}, 1) \end{aligned}$$

*then*

$$\begin{aligned} f_0 &= 0 \\ f_i(x_1, \ldots, x_{i-1}, 0) \le x_i &\le \overline{f_i(x_1, \ldots, x_{i-1}, 1)} \qquad (i = 1, \ldots, n) \end{aligned}$$

*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 $h : M \rightarrowtail\!\!\!\!\rightarrow N$, then $\mathcal{M}$ and $\mathcal{N}$ are isomorphic ($\mathcal{M} \cong \mathcal{N}$).

# Example — Lindenbaum Algebra of Propositional Logic

## Lindenbaum Algebra of Propositional Logic

$$\mathrm{Lin} := \left(\mathrm{Wff}/{\sim}, 0, 1, +, \cdot, {}^{-}\right) \qquad \mathbf{2} := \Big(\{0,1\}, 0, 1, \max, \min, 1 - \Big)$$

where

$$\begin{aligned}
A \sim B &:= \vdash A \leftrightarrow B \\
[A] &:= \big\{B \in \mathrm{Wff} : A \sim B\big\} \\
\mathrm{Wff}/{\sim} &:= \big\{[A] : A \in \mathrm{Wff}\big\} \\
0 &:= [\bot] \\
1 &:= [\top] \\
\overline{[A]} &:= [\neg A] \\
[A] + [B] &:= [A \vee B] \\
[A] \cdot [B] &:= [A \wedge B]
\end{aligned}$$

$$\mathrm{Lin} \overset{?}{\cong} \mathbf{2}$$

If $\nu$ 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.

$$\begin{tikzcd}
\mathrm{Wff} \arrow[d, "A \mapsto [A]"'] \arrow[dr, "\nu"] & \\
\mathrm{Lin} \arrow[r, "h"'] & \mathbf{2}
\end{tikzcd}$$

# 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 $\mathcal{B} = (B, 0, 1, +, \cdot, \overline{\phantom{x}})$ 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 $\iff$
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 $F_\omega$ of the set of natural numbers.
- Since $F_\omega$ 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 $\mathrm{Sto}(B) := \{w : w \text{ is an ultrafilter on } B\}$. Define a map $h : B \to \mathrm{P}(\mathrm{Sto}(B))$ by

$$x \mapsto \{w \in \mathrm{Sto}(B) : x \in w\}$$

Then

$$\mathcal{B} \cong (h(B), \emptyset, \mathrm{Sto}(B), \cup, \cap, \setminus)$$

$$\boxed{\vDash A \implies \vdash A}$$

$$\begin{gathered}
\nvdash A \\
\Downarrow \\
[A] \neq [\top] \\
\Downarrow \\
[\neg A] \neq [\bot] \\
\Downarrow \\
h([\neg A]) \neq \emptyset \\
\Downarrow \\
\exists w \in \mathrm{Sto}(\mathrm{Lin})\big([\neg A] \in w\big) \\
\Downarrow \\
\chi_w([\neg A]) = 1
\end{gathered}$$

# Compactness in Topology $\iff$ Compactness in Logic

## Compactness in Topology $\implies$ Compactness in Logic

Let $\mathbf{2} := \{0, 1\}$ be the discrete topology.
For any wff $A$, let $\mathrm{Mod}(A) := \left\{\nu \in \mathbf{2}^{\mathrm{Var}} : \nu \vDash A\right\}$.
Then $\{\mathrm{Mod}(A) : A \in \mathrm{Wff}\}$ is a basis for a topology on $\mathbf{2}^{\mathrm{Var}}$.
Let $\tau$ be the topology on $\mathbf{2}^{\mathrm{Var}}$ generated by $\mathrm{Mod}(A)$. Then $(\mathbf{2}^{\mathrm{Var}}, \tau)$ is the product topology.
By Tychonoff Theorem, $(\mathbf{2}^{\mathrm{Var}}, \tau)$ is a compact, Hausdorff space.
It can be shown that $\mathrm{Mod}(A)$ is clopen in $\mathbf{2}^{\mathrm{Var}}$.
By hypothesis, for each finite $\Gamma_0 \subset \Gamma$, $\mathrm{Mod}(\Gamma_0) \neq \emptyset$. That is to say, $\{\mathrm{Mod}(A) : A \in \Gamma\}$ has the Finite Intersection Property. By the compactness of $(\mathbf{2}^{\mathrm{Var}}, \tau)$, $\mathrm{Mod}(\Gamma) \neq \emptyset$.

## Compactness in Logic $\implies$ Compactness in Topology

$\Gamma$ is unsatisfiable iff $\{\mathrm{Mod}(\neg A) : A \in \Gamma\}$ covers $\mathbf{2}^{\mathrm{Var}}$.
Given any cover $\{\mathrm{Mod}(A)\}_{A \in \Gamma}$ of $\mathbf{2}^{\mathrm{Var}}$. Then $\Gamma^* := \{\neg A : A \in \Gamma\}$ is unsatisfiable. By compactness theorem, some finite subset $\Delta \subset \Gamma^*$ is unsatisfiable. The set $\{\mathrm{Mod}(A)\}_{\neg A \in \Delta} \subset \{\mathrm{Mod}(A)\}_{A \in \Gamma}$ is a finite subcover for $\mathbf{2}^{\mathrm{Var}}$. So $(\mathbf{2}^{\mathrm{Var}}, \tau)$ is compact.

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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

References1655

# Church 1903-1995

- Lambda Calculus
- Church-Turing Thesis
- Undecidability
- Church-Rosser Theorem
- Frege-Church Ontology

# Lambda Calculus

$\mathscr{L} = \{\lambda, .\}$

## Definition ($\lambda$-Terms)

$$\Lambda ::= x \mid \Lambda\Lambda \mid \lambda x.\Lambda$$

**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)

$$\mathrm{Fv}(\Lambda) := \begin{cases} \{x\} & \text{if } \Lambda = x \\ \mathrm{Fv}(M) \cup \mathrm{Fv}(N) & \text{if } \Lambda = MN \\ \mathrm{Fv}(M) \setminus \{x\} & \text{if } \Lambda = \lambda x.M \end{cases}$$

# 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

$$\lambda x.M \stackrel{\alpha}{=} \lambda y.M[y/x] \quad \text{if } y \text{ does not occur in } M.$$

$$(\lambda x.M)N \stackrel{\beta}{=} M[N/x]$$

$$\lambda x.Mx \stackrel{\eta}{=} M \quad \text{if } x \notin \mathrm{Fv}(M)$$

# $\lambda$-definability

## Definition (Church Numeral)

$$\begin{aligned} \underline{n} &:= \lambda f x.f^n x \\ f^0 x &:= x \\ f^{n+1} x &:= f(f^n x) \end{aligned}$$

## Definition ($\lambda$-definability)

An $n$-ary function $f(x_1, \ldots, x_n)$ is $\lambda$-definable iff there is a $\lambda$-term $F$ s.t. for all $a_1, \ldots, a_n$,

$$F\underline{a_1} \ldots \underline{a_n} \overset{\beta}{=} \underline{f(a_1, \ldots, a_n)}$$

$\boxed{\text{A function } f \text{ is computable iff it is } \lambda\text{-definable.}}$

$$\begin{aligned} \mathbf{succ} &:= \lambda n f x.f(nfx) \\ \mathbf{add} &:= \lambda m n f x.mf(nfx) \\ \mathbf{mult} &:= \lambda m n f.m(nf) \\ \mathbf{exp} &:= \lambda m n.nm \end{aligned}$$

$$\begin{tikzcd} \mathbb{N}^* \arrow[r, "f"] \arrow[d, "-"'] & \mathbb{N} \arrow[d, "-"] \\ \Lambda^* \arrow[r, "F"'] & \Lambda \end{tikzcd}$$

# Combinator

## Definition (Combinator)

A $\lambda$-term $M$ is called a combinator iff $\mathrm{Fv}(M) = \emptyset$.

$$\mathbf{K} = \lambda xy.x$$
$$\mathbf{S} = \lambda xyz.xz(yz)$$
$$\mathbf{I} = \lambda x.x$$
$$\omega = \lambda x.xx$$
$$\Omega = \omega\omega$$
$$\mathbf{Y} = \lambda y.(\omega(\lambda x.y(xx)))$$

$$\mathbf{F} = \lambda xy.y$$
$$\mathbf{T} = \mathbf{K}$$
$$\mathbf{B} = \mathbf{S}(\mathbf{KS})\mathbf{K}$$
$$\mathbf{C} = \mathbf{S}(\mathbf{BBS})(\mathbf{KK})$$
$$\mathbf{W} = \mathbf{SS}(\mathbf{SK})$$
$$\mathbf{D} = \mathbf{SII}$$
$$\mathbf{L} = \mathbf{D}(\mathbf{BDD})$$
$$\mathbf{neg} = \lambda x.x\mathbf{FT}$$
$$\mathbf{and} = \lambda xy.xy\mathbf{F}$$
$$\mathbf{or} = \lambda xy.x\mathbf{T}y$$
$$\mathbf{iszero} = \lambda x.x(\lambda y.\mathbf{F})\mathbf{T}$$

$$\iota = \lambda x.x\mathbf{SK}$$

$$\mathbf{K} = \iota(\iota(\iota\iota))$$
$$\mathbf{S} = \iota(\iota(\iota(\iota\iota)))$$
$$\mathbf{I} = \iota\iota$$

# 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 $\Theta$ 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: $\Theta O$, $O\Theta$, $UU$, $SLL$.
6. Mocking birds: $OI$, $LI$, $SII$.
7. $\Theta M$ is egocentric. $\Theta M = M(\Theta M) = \Theta M(\Theta M)$.
8. $\Theta K$ 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 $\lambda$-term $F$ there is a $\lambda$-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 $\lambda$-term $C(f, \vec{x})$, there exists a $\lambda$-term $M$ s.t. for all $\lambda$-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 $M$ $\beta$-reduces to $N$ in zero or more steps.

## Definition ($\beta$-nf)

A term is in $\beta$ normal form iff it can't be $\beta$-reduced.
A term $M$ has a $\beta$ normal form iff it $\beta$ reduces to some $N$ that is in $\beta$-nf.

## Theorem (Church-Rosser Theorem)

*Let $\twoheadrightarrow$ denote either $\twoheadrightarrow_\beta$ or $\twoheadrightarrow_{\beta\eta}$. Suppose $M, P, Q$ are $\lambda$-terms s.t. $M \twoheadrightarrow P$ and $M \twoheadrightarrow Q$. Then there exists a $\lambda$-term $N$ s.t. $P \twoheadrightarrow N$ and $Q \twoheadrightarrow N$.*

$$\begin{tikzcd}
 & M \arrow[dl, two heads] \arrow[dr, two heads] & \\
P \arrow[dr, dashed, two heads] & & Q \arrow[dl, dashed, two heads] \\
 & N &
\end{tikzcd}$$

# Second Fixpoint Theorem in Lambda Calculus

We write $\ulcorner M \urcorner := \underline{\#M}$ to denote the Church numeral representing the Gödel number of $M$.

## Theorem (Second Fixpoint Theorem in Lambda Calculus)

*For every $\lambda$-term $F$ there is a $\lambda$-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 $\lambda$-term $D$ s.t. for all $\underline{n}$,*

$$D\underline{n} = \begin{cases} \underline{0} & \textit{if term with Gödel number } n \textit{ has a } \beta\textit{-nf} \\ \underline{1} & \textit{otherwise} \end{cases}$$

## 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 $\beta$-nf $\implies D\ulcorner X \urcorner = \underline{0} \implies G\ulcorner X \urcorner = \mathbf{\Omega} \implies X$ has no $\beta$-nf
$X$ has no $\beta$-nf $\implies D\ulcorner X \urcorner = \underline{1} \implies G\ulcorner X \urcorner = \mathbf{I} \implies X$ has a $\beta$-nf

# Undecidability

## Theorem (Church1936)

*There is no $D$ s.t. for all $M$,*

$$DM = \begin{cases} \mathbf{T} & \textit{if } M \textit{ has a normal form} \\ \mathbf{F} & \textit{otherwise} \end{cases}$$

## 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 $\Omega$ 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 $B := \{M : M\ulcorner M \urcorner \in A\}$.
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 $Q \in \Lambda \setminus A$.

$$G := \lambda n.\mathbf{iszero}(Dn)QP$$

$$G \in B \iff D\ulcorner G \urcorner = \underline{0} \implies G\ulcorner G \urcorner = Q \implies G\ulcorner G \urcorner \notin A \implies G \notin B$$
$$G \notin B \iff D\ulcorner G \urcorner = \underline{1} \implies G\ulcorner G \urcorner = P \implies G\ulcorner G \urcorner \in A \implies G \in B$$

## 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)

$$C ::= x \mid \mathbf{K} \mid \mathbf{S} \mid (CC)$$

## Reduction

$$\mathbf{K}MN = M$$
$$\mathbf{S}MNL = ML(NL)$$

- $ex \simeq \varphi_e(x)$
- $\varphi_k(x, y) \simeq x$
- $\varphi_s(x, y, z) \simeq \varphi_{\varphi_x(z)}(\varphi_y(z))$

# Combinatory Completeness

## Proposition (Combinatory Completeness)

*For every $\lambda$-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* $\mathrm{CL}$ *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

$$\frac{x : A \in \Gamma}{\Gamma \vdash x : A}\ \text{var} \qquad \frac{}{\Gamma \vdash * : 1}\ \text{unit}$$

$$\frac{\Gamma \vdash t : A \qquad \Gamma \vdash u : B}{\Gamma \vdash \langle t, u \rangle : A \times B}\ \times^{+} \qquad \frac{\Gamma \vdash t : A \times B}{\Gamma \vdash \pi_1 t : A}\ \times^{-} \qquad \frac{\Gamma \vdash t : A \times B}{\Gamma \vdash \pi_2 t : B}\ \times^{-}$$

$$\frac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \to B}\ \text{abs} \qquad \frac{\Gamma \vdash t : A \to B \qquad \Gamma \vdash u : A}{\Gamma \vdash tu : B}\ \text{app}$$

## 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 $\beta/\eta$-reduction correspond to?

$$\dfrac{\dfrac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \to B} \qquad \Gamma \vdash u : A}{\Gamma \vdash (\lambda x.t)u : B} \quad \overset{\text{cut}}{\Longrightarrow} \quad \Gamma \vdash t[u/x] : B$$

$$\Gamma \vdash t : A \to B \quad \overset{\text{expansion}}{\Longrightarrow} \quad \dfrac{\dfrac{\Gamma, x : A \vdash t : A \to B \qquad \Gamma, x : A \vdash x : A}{\Gamma, x : A \vdash tx : B}}{\Gamma \vdash \lambda x.tx : A \to B}$$

# Example

$$
\dfrac{\dfrac{\dfrac{\dfrac{[x : A \to B \to C]^3 \qquad [z : A]^1}{xz : B \to C} \qquad \dfrac{[y : A \to B]^2 \qquad [z : A]^1}{yz : B}}{xz(yz) : C}}{\lambda z.xz(yz) : A \to C}\,[\to^+]^1}{\lambda y.\lambda z.xz(yz) : (A \to B) \to A \to C}\,[\to^+]^2}{\lambda x.\lambda y.\lambda z.xz(yz) : (A \to B \to C) \to (A \to B) \to A \to C}\,[\to^+]^3
$$

- $\mathbf{I} := \lambda x.x : A \to A$
- $\mathbf{K} := \lambda x.\lambda y.x : A \to B \to A$
- $\mathbf{S} := \lambda x.\lambda y.\lambda z.xz(yz) : (A \to B \to C) \to (A \to B) \to A \to C$
- $\mathbf{B} : (B \to C) \to (A \to B) \to A \to C$
- $\mathbf{C} : (A \to B \to C) \to B \to A \to C$
- $\mathbf{W} : (A \to A \to B) \to A \to B$

## 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 $\beta$-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 $t \to t'$, then the derivation of $\Gamma \vdash t : A$ can only mention subformulas of $A$ and subformulas of assumptions in $\Gamma$.*

**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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# What is a Formalization of Math?

$$\begin{tikzcd}
\boxed{\begin{array}{c}\text{A formal}\\ \text{deduction}\\ \text{system}\end{array}} \arrow[r, "\textit{Semantics}"] & \boxed{\begin{array}{c}\text{Class of objects}\\ \text{directly described}\\ \text{by the system}\end{array}} & \boxed{\text{Math}} \arrow[l, "\textit{Encoding}"'] \\
\textit{Syntax} \arrow[u] & &
\end{tikzcd}$$

$$\text{First Order Logic} \xrightarrow{\textit{express}} ZFC \xrightarrow{\textit{describe}} V$$

$$\text{Intensional } MLTT \xrightarrow{\textit{express}} HoTT \xrightarrow{\textit{describe}} \infty\text{–}\mathbf{Grpd}$$

- set theory
- category theory
- homotopy type theory

# Type Theory vs Set Theory

Set Theory

Logic

$$\wedge, \vee, \rightarrow, \neg, \forall, \exists$$

Set

$$\times, +, \rightarrow, \prod, \sum$$

$x \in A$ is a proposition

Type Theory

Types

$$\times, +, \rightarrow, \prod, \sum$$

Logic

$$\wedge, \vee, \rightarrow, \neg, \forall, \exists$$

$x : A$ is a typing judgment

# 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 $\Gamma$ |
| $\Gamma \vdash a : A$ | $a$ 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 $\pi \circ s = 1_A$.

# 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

$$\frac{}{\Gamma \vdash 1 : \mathsf{type}}\ \text{1-F}$$

$$\frac{}{\Gamma \vdash * : 1}\ \text{1-I}$$

$$\frac{\Gamma \vdash x : 1}{\Gamma \vdash x = * : 1}\ \text{1-C}$$

# Rules for dependent product type

$$\frac{A : \mathsf{type} \quad x : A \vdash B(x) : \mathsf{type}}{\prod_{x:A} B(x) : \mathsf{type}} \prod\text{-F}$$

$$\frac{x : A \vdash fx : B(x)}{\lambda x.fx : \prod_{x:A} B(x)} \prod\text{-I}$$

$$\frac{a : A \quad f : \prod_{x:A} B(x)}{fa : B(a)} \prod\text{-E}$$

$$\frac{a : A \quad x : A \vdash fx : B(x)}{(\lambda x.fx)\, a = fa : B(a)} \prod\text{-C}$$

**Remark:** We write $A \to B$ instead of $\prod_{x:A} B$ if $x$ is not free in $B$.

# Rules for dependent sum type

$$\frac{A : \text{type} \quad x : A \vdash B(x) : \text{type}}{\sum_{x:A} B(x) : \text{type}} \sum\text{-F}$$

$$\frac{a : A \quad b : B(a)}{(a, b) : \sum_{x:A} B(x)} \sum\text{-I}$$

$$\frac{p : \sum_{x:A} B(x) \quad x : A, y : B(x) \vdash c(x, y) : C((x, y))}{E(c, p) : C(p)} \sum\text{-E}$$

**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 $C(p)$.

$$\frac{a : A \quad b : B(a) \quad x : A, y : B(x) \vdash c(x, y) : C((x, y))}{E(c, (a, b)) = c(a, b) : C((a, b))} \sum\text{-C}$$

# Derived Rules

$$\pi_1(p) \coloneqq E(\lambda xy.x, p)$$
$$\pi_2(p) \coloneqq E(\lambda xy.y, p)$$

$$\frac{A : \mathsf{type} \quad x : A \vdash B(x) : \mathsf{type}}{\sum_{x:A} B(x) : \mathsf{type}} \ \sum\text{-F}$$

$$\frac{a : A \quad b : B(a)}{(a, b) : \sum_{x:A} B(x)} \ \sum\text{-I}$$

$$\frac{p : \sum_{x:A} B(x)}{\pi_1(p) : A} \ \sum\text{-E}_1 \qquad \frac{p : \sum_{x:A} B(x)}{\pi_2(p) : B(\pi_1(p))} \ \sum\text{-E}_2$$

$$\frac{a : A \quad b : B(a)}{\pi_1(a, b) = a : A} \ \sum\text{-C}_1 \qquad \frac{a : A \quad b : B(a)}{\pi_2(a, b) = b : B(a)} \ \sum\text{-C}_2$$

**Remark:** We write $A \times B$ instead of $\sum_{x:A} B$ if $x$ is not free in $B$.

# Rules for coproduct type

$$\frac{A : \mathsf{type} \quad B : \mathsf{type}}{A + B : \mathsf{type}} \; +\text{-F}$$

$$\frac{a : A}{\iota_1(a) : A + B} \; +\text{-I}_1 \qquad \frac{b : B}{\iota_2(b) : A + B} \; +\text{-I}_2$$

$$\frac{c : A + B \quad x : A \vdash f(x) : C(\iota_1(x)) \quad y : B \vdash g(y) : C(\iota_2(y))}{D(f, g, c) : C(c)} \; +\text{-E}$$

$$\frac{a : A \quad x : A \vdash f(x) : C(\iota_1(x)) \quad y : B \vdash g(y) : C(\iota_2(y))}{D(f, g, \iota_1(a)) = f(a) : C(\iota_1(a))} \; +\text{-C}_1$$

$$\frac{b : B \quad x : A \vdash f(x) : C(\iota_1(x)) \quad y : B \vdash g(y) : C(\iota_2(y))}{D(f, g, \iota_2(b)) = g(b) : C(\iota_2(b))} \; +\text{-C}_2$$

# Rules for identity type

**Notation:** Sometimes we write $x =_A y$ for $\mathsf{Id}_A(x, y)$.

$$\frac{A : \mathsf{type} \quad a, b : A}{a =_A b : \mathsf{type}} \ \mathsf{Id\text{-}F}$$

$$\frac{a : A}{\mathsf{refl}_a : a =_A a} \ \mathsf{Id\text{-}I}$$

$$\frac{p : a =_A b \qquad \begin{array}{c} x, y : A, z : x =_A y \vdash B(x, y, z) : \mathsf{type} \\ x : A \vdash d(x) : B(x, x, \mathsf{refl}_x) \end{array}}{J_d(a, b, p) : B(a, b, p)} \ \mathsf{Id\text{-}E}$$

$$\frac{a : A \qquad \begin{array}{c} x, y : A, z : x =_A y \vdash B(x, y, z) : \mathsf{type} \\ x : A \vdash d(x) : B(x, x, \mathsf{refl}_x) \end{array}}{J_d(a, a, \mathsf{refl}_a) = d(a) : B(a, a, \mathsf{refl}_a)} \ \mathsf{Id\text{-}C}$$

# Remark: Id-E 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 $\mathsf{refl}_x$. 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 $P(x, y, p) \coloneqq y =_A x$ and $Q(x, y, p) \coloneqq y =_A z \to x =_A z$. By path induction, we may reduce to the cases $P(x, x, \mathsf{refl}_x) \coloneqq x =_A x$ and $Q(x, x, \mathsf{refl}_x) \coloneqq x =_A z \to x =_A z$.

# Logic in HoTT

| Logical Connectives | Interpretation in HoTT |
|---|---|
| $\bot$ | $0$ |
| $\top$ | $1$ |
| $A \wedge B$ | $A \times B$ |
| $A \vee B$ | $\lVert A + B \rVert$ |
| $A \to B$ | $A \to B$ |
| $A \leftrightarrow B$ | $A \simeq B$ |
| $\neg A$ | $A \to 0$ |
| $\forall_{x:A} B(x)$ | $\prod_{x:A} B(x)$ |
| $\exists_{x:A} B(x)$ | $\lVert \sum_{x:A} B(x) \rVert$ |
| $\exists!_{x:A} B(x)$ | $\mathrm{contr}\left(\sum_{x:A} B(x)\right)$ |

# 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 $a : A$.

$$\begin{tikzcd} A \arrow[r] \arrow[dr, "f"'] & \|A\| \arrow[d, dashed, "\overline{f}"] \\ & B \end{tikzcd}$$

**Remarks:** The first rule means that if $A$ is inhabited, so is $\|A\|$.
The second ensures that $\|A\|$ is a mere proposition.
**Remark:** $\|A\| := \prod_{X:\mathrm{Prop}} (A \to X) \to X$ satisfies the rules.

# Translation

Every farmer who owns a donkey beats it.

$$\forall x \forall y (Fx \wedge Dy \wedge Oxy \to Bxy)$$

$$b : \prod_{z : \sum_{x:F} \sum_{y:D} Oxy} B(\pi_1(z), \pi_1(\pi_2(z)))$$

- 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

$$\mathrm{contr}(A) \coloneqq \sum_{x:A} \prod_{y:A} x =_A y$$

$$\mathrm{prop}(A) \coloneqq \prod_{x,y:A} \mathrm{contr}\left(x =_A y\right) \qquad \left(\text{equivalently, } \prod_{x,y:A} x =_A y\right)$$

$$\mathrm{set}(A) \coloneqq \prod_{x,y:A} \mathrm{prop}\left(x =_A y\right)$$

$$\mathrm{groupoid}(A) \coloneqq \prod_{x,y:A} \mathrm{set}\left(x =_A y\right)$$

$$n+1\text{-}\mathrm{groupoid}(A) \coloneqq \prod_{x,y:A} n\text{-}\mathrm{groupoid}\left(x =_A y\right)$$

$$\mathrm{Prop} \coloneqq \sum_{A:U} \mathrm{prop}(A) \qquad \mathrm{Set} \coloneqq \sum_{A:U} \mathrm{set}(A)$$

# 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 $\overset{?}{=}$ 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 $\alpha$ from $p$ to $q$ |
| $\theta : \alpha =_{p=_{a=_A b} q} \beta$ | ... |

- 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 | $\cdots$ |
|---|---|---|---|---|
| $\bullet$ | $\bullet \longrightarrow \bullet$ | $\bullet \underset{\Downarrow}{\overset{\longrightarrow}{\underset{\longrightarrow}{\longrightarrow}}} \bullet$ | $\bullet \underset{\Rightarrow}{\overset{\longrightarrow}{\underset{\longrightarrow}{\longrightarrow}}} \bullet$ | Globes |
| $\bullet$ | $\bullet \longrightarrow \bullet$ | $\begin{tikzcd} & \bullet \arrow[dr] & \\ \bullet \arrow[ur] \arrow[rr] & & \bullet \end{tikzcd}$ | $\begin{tikzcd} & \bullet \arrow[dr] \arrow[ddr] & \\ & & \bullet \\ \bullet \arrow[uur] \arrow[urr, dashed] \arrow[rr] & & \bullet \arrow[u] \end{tikzcd}$ | Simplices |

# Equivalence

$$\begin{aligned}
\mathrm{contr}(A) &:\equiv \sum_{x:A} \prod_{y:A} x =_A y \\
\mathrm{fib}_f(y) &:\equiv \sum_{x:A} f(x) =_A y \\
\mathrm{equiv}(f) &:\equiv \prod_{y:B} \mathrm{contr}\left(\mathrm{fib}_f(y)\right) \\
A \simeq B &:\equiv \sum_{f:A\to B} \mathrm{equiv}(f)
\end{aligned}$$

- 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
  $$\mathrm{biinv}(f) \coloneqq \left(\sum_{g:B\to A} g \circ f \sim \mathrm{id}_A\right) \times \left(\sum_{h:B\to A} f \circ h \sim \mathrm{id}_B\right)$$
- isomorphism
  $$\mathrm{iso}(f) \coloneqq \sum_{g:B\to A} \left[\left(\prod_{x:A} gfx =_A x\right) \times \left(\prod_{y:B} fgy =_B y\right)\right]$$
  $$A \cong B \coloneqq \sum_{f:A\to B} \mathrm{iso}(f)$$
- $\mathrm{equiv}(f) \simeq \mathrm{biinv}(f) \simeq \mathrm{iso}(f)$

# 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 $\implies$ 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.

# 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

$$\prod_{A,B:\mathrm{Prop}} A = B \simeq (A \leftrightarrow B)$$

- Paths are isomorphisms for sets

$$\prod_{A,B:\mathrm{Set}} A = B \simeq (A \cong B)$$

# 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 $x : A$.

$$fx : \sum_{b:B(x)} C(x,b) \qquad (\prod\text{-E})$$

$$\pi_1(fx) : B(x) \qquad (\sum\text{-}E_1)$$

$$\pi_2(fx) : C(x, \pi_1(fx)) \qquad (\sum\text{-}E_2)$$

$$\lambda x.\pi_1(fx) : \prod_{x:A} B(x) \qquad (\prod\text{-I})$$

$$(\lambda x.\pi_1(fx))x = \pi_1(fx) : B(x) \qquad (\prod\text{-C})$$

$$\pi_2(fx) : C(x, (\lambda x.\pi_1(fx))x) \qquad (\text{substitution})$$

$$\lambda x.\pi_2(fx) : \prod_{x:A} C(x, (\lambda x.\pi_1(fx))x) \qquad (\prod\text{-I})$$

$$\big(\lambda x.\pi_1(fx), \lambda x.\pi_2(fx)\big) : \sum_{g:\prod_{x:A} B(x)} \prod_{x:A} C(x, g(x)) \qquad (\sum\text{-I})$$

where $g \coloneqq \lambda x.\pi_1(fx)$. By $\prod$-I, 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 $B \to A$ with fibers $B(x)$ |
| $x : A \vdash b(x) : B(x)$<br>$0, 1$ | conditional proof<br>$\bot, \top$ | family of elements<br>$\emptyset, \{\emptyset\}$ | section<br>$\emptyset, \{\bullet\}$ |
| $A + B$ | $A \vee B$ | 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 $B \to A$ |
| $\prod_{x:A} B(x)$ | $\forall_{x:A} B(x)$ | product | space of sections of fibration $B \to A$ |
| $p : x =_A y$ | proof of equality | $x = y$ | path from $x$ to $y$ |
| $\sum_{x,y:A} x =_A y$ | equality relation | $\{(x, x) : x \in A\}$ | path space $A^I$ |

| 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 $\Omega$ | subobject classifier (elementary topos) |
| universe type $U$ | object classifier ($\infty$-topos) |
| natural numbers $\mathbb{N}$ | natural numbers object |
| coequalizer type $\mathrm{coeq}(f, g)$ | coequalizer |

# Contents

Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# 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 $\Delta$ | 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 — A Gentle Introduction.
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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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:
  - $\forall AB \in \mathrm{ob}(\mathbf{C}) \forall f : A \to B \left[ f \circ 1_A = f = 1_B \circ f \right]$
  - $\forall ABCD \in \mathrm{ob}(\mathbf{C}) \forall fgh : A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D \left[ h \circ (g \circ f) = (h \circ g) \circ f \right]$

$$\begin{tikzcd}
A \arrow[r, "f"] & B \\
A \arrow[u, "1_A"] \arrow[ur, "f"] \arrow[r, "f"'] & B \arrow[u, "1_B"']
\end{tikzcd}$$

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[rr, bend left, "g\circ f"] \arrow[rrr, bend left=50, "h\circ(g\circ f)"] \arrow[rrr, bend right=50, "(h\circ g)\circ f"'] & B \arrow[r, "g"] \arrow[rr, bend right, "h\circ g"'] & C \arrow[r, "h"] & D
\end{tikzcd}$$

# Associativity

$$\begin{tikzcd}
A \arrow[r, "\text{father of}"] \arrow[rr, bend right, "\text{grandfather of}"'] \arrow[rrr, bend right=40, "\text{great-grandfather of}"'] & B \arrow[r, "\text{father of}"] & C \arrow[r, "\text{father of}"] & D
\end{tikzcd}$$

$$\begin{tikzcd}
A \arrow[r, "\text{father of}"] \arrow[rrr, bend right=40, "\text{great-grandfather of}"'] & B \arrow[r, "\text{father of}"] \arrow[rr, bend right, "\text{grandfather of}"'] & C \arrow[r, "\text{father of}"] & D
\end{tikzcd}$$

# Commutative?

$$\begin{tikzcd}
\overset{A}{\boxed{\text{a child}}} \arrow[r, "\text{has as parents}"] \arrow[dr, "\text{has as mother}"'] & \overset{B}{\boxed{\begin{array}{l}\text{a pair } (w,m) \\ \text{where } w \text{ is a} \\ \text{woman and } m \\ \text{is a man}\end{array}}} \arrow[d, "w"] \arrow[dl, phantom, "\circlearrowright"] \\
& \overset{C}{\boxed{\text{a woman}}}
\end{tikzcd}$$

$$\begin{tikzcd}
A \arrow[r, "\text{sister of}"] \arrow[d, "\text{husband of}"'] \arrow[dr, phantom, "\overset{?}{\circlearrowright}"] & C \arrow[d, "\text{husband of}"] \\
B \arrow[r, "\text{brother of}"'] & D
\end{tikzcd}$$

# 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 category* is a category having at most one morphism from one object to another.
- A *poset category* is a preorder where if there's a morphism $f : x \to y$ and $g : y \to x$, then $x = y$.
- A *monoid category* $(M, \cdot, e)$ 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 category* is a category that has only one object and in which every morphism is an isomorphism.

# Groups as Categories

- A *group category* 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 |
|---|---|---|
| **Set**/**FinSet** | sets / finite sets | total functions |
| **Par** | sets | partial functions |
| **Rel** | sets | relations |
| **Preord** | preorders | monotone functions |
| **Poset** | partial order sets | monotone functions |
| **Graph** | directed graphs | graph homomorphisms |
| **Type** | types | recursive functions |
| **Mon** | monoids | homomorphisms |
| **Grp**/**AbGrp** | groups / abelian groups | homomorphisms |
| **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 |
| **Top** | topological spaces | continuous functions |
| **Diff** | smooth manifolds | smooth maps |

# 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?

$$\begin{tikzcd}
 & \bullet & \\
\bullet \arrow[ur, "d"] & \bullet \arrow[u, "e"] & \bullet \arrow[ul, "f"'] \\
 & \bullet \arrow[ul, "a"] \arrow[u, "b"] \arrow[ur, "c"'] &
\end{tikzcd}$$

no equation

$$\begin{tikzcd}
 & \bullet & \\
\bullet \arrow[ur, "d"] \arrow[rr, phantom, "\circlearrowleft"] & \bullet \arrow[u, "e"] & \bullet \arrow[ul, "f"'] \arrow[ll, phantom, "\circlearrowleft"] \\
 & \bullet \arrow[ul, "a"] \arrow[u, "b"] \arrow[ur, "c"'] &
\end{tikzcd}$$

$da = eb = fc$

What equations do you need to make the following graphs into preorders?

$$\begin{tikzcd}
\bullet \arrow[r, "f"] \arrow[d, "g"'] \arrow[dr, phantom, "\circlearrowleft"] & \bullet \arrow[d, "h"] \\
\bullet \arrow[r, "i"'] & \bullet
\end{tikzcd}$$

$hf = ig$

$$\begin{tikzcd}
\bullet \arrow[r, "f"] \arrow[d, "g"'] & \bullet \arrow[d, "h"] \\
\bullet & \bullet
\end{tikzcd}$$

no equation

$$\begin{tikzcd}
\bullet \arrow[r, "h"] \arrow[d, "f"'] \arrow[dr, phantom, "\circlearrowleft"] & \bullet \\
\bullet \arrow[ur, "g"'] &
\end{tikzcd}$$

$h = gf$

$$\begin{tikzcd}
 \\
a \arrow[loop above, "f"]
\end{tikzcd}$$

$f = 1_a$

# Category vs Graph

- From a graph $G$ we can get the "free category on $G$" $\mathrm{Free}(G)$, 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.

# Example — Drawings

There exists a category $\mathbf{Draw}$ in which:

- An object is a black-and-white drawing, that is a function $\alpha : \mathbb{R}^2 \to \mathbf{Bool}$.
- 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 $\alpha$ is the identity map on $\mathbb{R}^2$.
- Composition is given by function composition.

# Some Morphisms in $\mathbf{Draw}$

- 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.
$$\mathrm{rot}_{\theta} : (x, y) \mapsto (x \cos\theta + y \sin\theta, y \cos\theta - x \sin\theta)$$
- Roto-translations.
$$\mathrm{rotra}_{\theta,a,b} : (x, y) \mapsto (x \cos\theta + y \sin\theta + a, y \cos\theta - x \sin\theta + b)$$
- 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 $\mathrm{rot}_{\theta}$.

# The Category of Typed Programming Language

- A category of typed programming language is a category whose objects are datatypes, 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 $A \to B$ represents programs. Each term of type $A \to B$ is a program that transforms its argument of type $A$ into an output of type $B$.

$$\begin{tikzcd}
\mathrm{String} \arrow[loop above] \arrow[rr, "g \circ f"] \arrow[dr, "f"'] & & \mathrm{Boolean} \arrow[loop above] \\
& \mathrm{Integer} \arrow[ur, "g"'] \arrow[loop below] &
\end{tikzcd}$$

# 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 $f : A \to B$ of $\mathbf{C}$ a morphism $f^{\mathrm{op}} : B \to A$ in $\mathbf{C}^{\mathrm{op}}$, with the same identities and a composition defined (when possible) by $f^{\mathrm{op}} \circ g^{\mathrm{op}} := (g \circ f)^{\mathrm{op}}$.

$$\begin{tikzcd} A \arrow[d, "f"] \\ B \arrow[d, "g"] \\ C \end{tikzcd} \quad \leftrightsquigarrow \quad \begin{tikzcd} A \\ B \arrow[u, "f^{\mathrm{op}}"'] \\ C \arrow[u, "g^{\mathrm{op}}"'] \end{tikzcd}$$

# 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 $\varphi$ is true in a category $\mathbf{C}$ iff the dual statement $\varphi^{\mathrm{op}}$ 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 $f = g \circ \hat{f}$.

$$\begin{tikzcd} A \arrow[r, hook, "i"] \arrow[dr, "f"'] & B \arrow[d, "\overline{f}"] \\ & C \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "\hat{f}"] \arrow[dr, "f"'] & B \arrow[d, "g"] \\ & C \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "\text{lifting}"] \arrow[dr, "\text{composition}"'] & B \arrow[d, "\text{extension}"] \\ & C \end{tikzcd}$$

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 $h = g \circ f$?

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[d, "g"] \\ & C \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

# 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.

$$A = I + \frac{B}{2} - 1$$

$$A = 7 + \frac{8}{2} - 1 = 10$$

# 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. $\exists g : B \to A : gf = 1_A$.

$$\begin{tikzcd} B \arrow[r, "g"] \arrow[dr, "1_B"'] & A \arrow[d, "f"] \\ & B \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "1_A"'] & B \arrow[d, "g"] \\ & A \end{tikzcd} \qquad \begin{tikzcd} A \arrow[loop left, "g \circ f"] \arrow[loop right, "1_A"] \arrow[r, bend left, "f"] & B \arrow[l, bend left, "g"] \arrow[loop right, "1_B"] \arrow[loop left, "f \circ g"] \end{tikzcd}$$

# 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 |

$$\circlearrowright \quad \begin{array}{ccc} 0 & \longleftrightarrow & 0 \\ 90 & \longleftrightarrow & 1 \\ 180 & \longleftrightarrow & 2 \\ 270 & \longleftrightarrow & 3 \end{array} \qquad \begin{array}{ccc} 0 & \longleftrightarrow & 0 \\ 90 & \longleftrightarrow & 3 \\ 180 & \longleftrightarrow & 2 \\ 270 & \longleftrightarrow & 1 \end{array} \quad \circlearrowleft$$

# 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.

$$\begin{tikzcd}
\text{monic} & \text{bimorphism} \arrow[l] \arrow[r] & \text{epic} \\
\text{injection} \arrow[u] & \text{bijection} \arrow[l] \arrow[r] \arrow[u] & \text{surjection} \arrow[u] \\
\text{section} \arrow[u] & \text{isomorphism} \arrow[l] \arrow[r] \arrow[u] & \text{retraction} \arrow[u]
\end{tikzcd}$$

# 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 $\mathbf{Top}$.

$$f : [0, 1) \to \left\{(x, y) : x^2 + y^2 = 1\right\}$$

- 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

$$
\begin{tikzcd}
\boxed{\text{Items consumed at the table}} \arrow[r, "\text{dining}", "\cong"'] \arrow[d, "\text{kind}"'] \arrow[ddr, bend right, "f"'] & \boxed{\text{Empty plates left on the table}} \arrow[d, "\overline{\text{kind}}"] \arrow[ddl, bend left, "\overline{f}"] \\
\boxed{\text{Kind of iterms}} \arrow[r, "\text{shape}"] \arrow[dr, "\text{price}"'] & \boxed{\text{Shapes of plates}} \arrow[dl, "\overline{\text{price}}"] \\
& \boxed{\text{Amount of money}}
\end{tikzcd}
$$

$$\sum_k \text{price}(k) \cdot (\text{size of the stack of } \textit{kind} \text{ over } k) = \sum_x x \cdot (\text{size of the stack of } f \text{ over } x)$$

$$\sum_k \overline{\text{price}}(s) \cdot (\text{size of the stack of } \overline{\text{kind}} \text{ over } s) = \sum_x x \cdot (\text{size of the stack of } \overline{f} \text{ over } x)$$

- 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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

$$\begin{tikzcd} A \arrow[loop left, "1_A"] \arrow[r, bend left, "f"] & B \arrow[l, bend left, "g"] \arrow[loop right, "1_B"] \end{tikzcd}$$

- 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 $\begin{tikzcd} A \arrow[loop above, "1_A"] \end{tikzcd}$ and $A \xrightarrow{f} B$ .
  - $A$ is initial: there are unique arrows $\begin{tikzcd} B \arrow[loop above, "1_B"] \end{tikzcd}$ and $B \xrightarrow{g} A$ .

— And there is a unique isomorphism between them.

# Examples

- In $\mathbf{Set}$, any one-element set is terminal. The empty set 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 $\mathbf{Preord}$, the greatest element $\top$ is terminal and the least element $\perp$ is initial.
- In $\mathbf{Top}$, the one-element topological space is terminal and the empty topological space 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 is terminal.
- In $\mathbf{FdVect}$, The zero vector space $\{\mathbf{O}\}$ 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.

# Product

$$\begin{tikzcd} & A \\ C \arrow[ur, bend left, "f"] \arrow[r, dashed, "u"] \arrow[dr, bend right, "g"'] & A \times B \arrow[u, "\pi_1"'] \arrow[d, "\pi_2"] \\ & B \end{tikzcd} \qquad \begin{tikzcd} & A \\ C \arrow[ur, bend left, "f"] \arrow[r, dashed, "\langle f,g \rangle"] \arrow[dr, bend right, "g"'] & A \times B \arrow[u, "\pi_1"'] \arrow[d, "\pi_2"] \\ & B \end{tikzcd}$$

## 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

$$\forall C \forall f g : A \xleftarrow{f} C \xrightarrow{g} B \; \exists! u : C \to A \times B \big[ f = \pi_1 \circ u \;\; \& \;\; g = \pi_2 \circ u \big]$$

Example: $\min\{x, y\}$ in $(\mathbb{R}, \leq)$. $x \cap y$ in $(\mathrm{P}(A), \subset)$. $\gcd(x, y)$ in $(\mathbb{N}, |)$.

$A \wedge B \vdash A \qquad A \wedge B \vdash B \qquad \dfrac{C \vdash A \quad C \vdash B}{C \vdash A \wedge B}$

# Example — Product

$$
\begin{tikzcd}
\overset{C_1}{\boxed{\text{a piece of furniture}}} \arrow[r, "\text{has}"] & \overset{D_1}{\boxed{\text{a width}}} \arrow[r, "\text{is, in inches}"] & \overset{A}{\boxed{\text{a number}}} \\
\overset{C=C_1\times C_2}{\boxed{\begin{array}{l}\text{a pair } (f,s) \text{ where} \\ f \text{ is a piece of fur-} \\ \text{niture and } s \text{ is a} \\ \text{space in the house}\end{array}}} \arrow[u, "f"] \arrow[d, "s"'] \arrow[rr, dashed, "u"] \arrow[dr, phantom, "\circlearrowleft"] & & \overset{A\times B}{\boxed{\begin{array}{l}\text{a pair of numbers} \\ (w_1, w_2)\end{array}}} \arrow[u, "w_1"'] \arrow[d, "w_2"] \\
\overset{C_2}{\boxed{\text{a space in the house}}} \arrow[r, "\text{has}"] & \overset{D_2}{\boxed{\text{a width}}} \arrow[r, "\text{is, in inches}"] & \overset{B}{\boxed{\text{a number}}}
\end{tikzcd}
$$

- 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 $(A, \leq) \times (B, \leq)$?

$$
\begin{tikzcd}
(A,\leq),(B,\leq) \arrow[r, "\text{underlying set}"] \arrow[d, dashed, "\text{product}"'] & A, B \arrow[d, "\text{product}"] \\
\boxed{\begin{array}{c}(A,\leq)\times(B,\leq)\\ \cong ?\\ (A\times B,\leq)\end{array}} \arrow[r, dashed, shift left, "\text{underlying set}"] & A\times B \arrow[l, dashed, shift left, "\text{construct } \leq"]
\end{tikzcd}
\qquad
\begin{tikzcd}
 & (A,\leq) \\
(A\times B,\leq) \arrow[ur, "\pi_1"] \arrow[dr, "\pi_2"'] & \\
 & (B,\leq)
\end{tikzcd}
$$

since $\pi_1, \pi_2$ 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

$$(a_1, b_1) \leq (a_2, b_2) \coloneqq a_1 \leq a_2 \wedge b_1 \leq b_2$$

and this makes $(A \times B, \leq)$ the product.

$$(A, \leq) \times (B, \leq) \cong (A \times B, \leq)$$

$$(a_1, b_1) \leq (a_2, b_2) \coloneqq a_1 \leq a_2 \wedge b_1 \leq b_2$$

# 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 $A, B$.

$$\begin{tikzcd} & A \\ X \arrow[ur, "f"] \arrow[dr, "g"'] & \\ & B \end{tikzcd}$$

- a morphism from $\begin{tikzcd} & A \\ X \arrow[ur, "f_1"] \arrow[dr, "g_1"'] & \\ & B \end{tikzcd}$ to $\begin{tikzcd} & A \\ Y \arrow[ur, "f_2"] \arrow[dr, "g_2"'] & \\ & B \end{tikzcd}$ is a map $X \xrightarrow{h} Y$ in $\mathbf{C}$

  which 'preserves the structure'. $\begin{tikzcd} & & A \\ X \arrow[r, "h"] \arrow[urr, bend left, "f_1"] \arrow[drr, bend right, "g_1"'] & Y \arrow[ur, "f_2"] \arrow[dr, "g_2"'] & \\ & & B \end{tikzcd}$

- then the terminal object of $\mathbf{C}_{AB}$ is the product $A \times B$.

# Coproduct

$$\begin{tikzcd}
A \arrow[dr, "\iota_1"'] \arrow[drr, bend left, "f"] & & \\
& A + B \arrow[r, dashed, "u"] & C \\
B \arrow[ur, "\iota_2"] \arrow[urr, bend right, "g"'] & &
\end{tikzcd}
\qquad
\begin{tikzcd}
A \arrow[dr, "\iota_1"'] \arrow[drr, bend left, "f"] & & \\
& A + B \arrow[r, dashed, "\begin{bmatrix} f \\ g \end{bmatrix}"] & C \\
B \arrow[ur, "\iota_2"] \arrow[urr, bend right, "g"'] & &
\end{tikzcd}$$

## 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

$$\forall C \forall f g : A \xrightarrow{f} C \xleftarrow{g} B \; \exists! u : A + B \to C \big[ f = u \circ \iota_1 \;\; \& \;\; g = u \circ \iota_2 \big]$$

Example: $\max\{x, y\}$ in $(\mathbb{R}, \leq)$. $x \cup y$ in $(\mathrm{P}(A), \subset)$. $\mathrm{lcm}(x, y)$ in $(\mathbb{N}, |)$.

$A \vdash A \vee B \qquad B \vdash A \vee B \qquad \dfrac{A \vdash C \quad B \vdash C}{A \vee B \vdash C}$

# 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 $\leq : (A + B) \times (A + B) \to \mathbf{Bool}$ is given by

$$
\begin{aligned}
((0, a), (0, b)) &\mapsto a \leq_A b \\
((1, a), (1, b)) &\mapsto a \leq_B b \\
((0, -), (1, -)) &\mapsto \top \\
((1, -), (0, -)) &\mapsto \bot
\end{aligned}
$$

# 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 $A, B$.

$$\begin{tikzcd} A \arrow[dr, "f"] & \\ & X \\ B \arrow[ur, "g"'] & \end{tikzcd}$$

- a morphism from $\begin{tikzcd} A \arrow[dr, "f_1"] & \\ & X \\ B \arrow[ur, "g_1"'] & \end{tikzcd}$ to $\begin{tikzcd} A \arrow[dr, "f_2"] & \\ & Y \\ B \arrow[ur, "g_2"'] & \end{tikzcd}$ is a map $X \xrightarrow{h} Y$ in $\mathbf{C}$

  which 'preserves the structure'. $\begin{tikzcd} A \arrow[dr, "f_1"'] \arrow[drr, bend left, "f_2"] & & \\ & X \arrow[r, "h"] & Y \\ B \arrow[ur, "g_1"] \arrow[urr, bend right, "g_2"'] & & \end{tikzcd}$

- then the initial object of $\mathbf{C}_{AB}$ is the coproduct $A + B$.

# Coproduct & Product

$$\begin{tikzcd} X_i \arrow[r, "\iota_i"] \arrow[dr, bend right, "f_i"'] & \coprod_{i\in I} X_i \arrow[d, dashed, "u"] \\ & Y \end{tikzcd} \qquad \begin{tikzcd} X \arrow[d, dashed, "u"'] \arrow[ddr, bend left, "f_i"] & \\ \prod_{i\in I} Y_i \arrow[dr, "\pi_i"'] & \\ & Y_i \end{tikzcd}$$

$$\mathrm{Hom}\left(\coprod_{i\in I} X_i, Y\right) \cong \prod_{i\in I} \mathrm{Hom}(X_i, Y) \qquad u \mapsto (u\iota_i)_{i\in I}$$

$$\mathrm{Hom}\left(X, \prod_{i\in I} Y_i\right) \cong \prod_{i\in I} \mathrm{Hom}(X, Y_i) \qquad u \mapsto (\pi_i u)_{i\in I}$$

**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 $\coprod_{i=1}^{m} X_i \to \prod_{j=1}^{n} Y_j$?

$$\begin{tikzcd}
X_1 \arrow[dr, "\iota_1"'] \arrow[drr, bend left, "f_1"] & & \\
& X_1 + X_2 \arrow[r, dashed, "f"] & Y_1 \times Y_2 \\
X_2 \arrow[ur, "\iota_2"] \arrow[urr, bend right, "f_2"'] & &
\end{tikzcd}$$

$$\begin{tikzcd}
& & Y_1 \\
X_1 \arrow[r, dashed, "f_1"] \arrow[urr, bend left, "f_{11}"] \arrow[drr, bend right, "f_{12}"'] & Y_1 \times Y_2 \arrow[ur, "\pi_1"] \arrow[dr, "\pi_2"'] & \\
& & Y_2
\end{tikzcd}
\qquad
\begin{tikzcd}
& & Y_1 \\
X_2 \arrow[r, dashed, "f_2"] \arrow[urr, bend left, "f_{21}"] \arrow[drr, bend right, "f_{22}"'] & Y_1 \times Y_2 \arrow[ur, "\pi_1"] \arrow[dr, "\pi_2"'] & \\
& & Y_2
\end{tikzcd}$$

$$f_1 = \langle f_{11}, f_{12} \rangle \quad f_2 = \langle f_{21}, f_{22} \rangle \quad f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$$

# What are the morphisms $\coprod_{i=1}^{m} X_i \to \prod_{j=1}^{n} Y_j$?

$$\mathrm{Hom}\left(\coprod_{i=1}^{m} X_i, \prod_{j=1}^{n} Y_j\right) \cong \prod_{i=1}^{m} \mathrm{Hom}\left(X_i, \prod_{j=1}^{n} Y_j\right) \cong \prod_{i=1}^{m} \prod_{j=1}^{n} \mathrm{Hom}(X_i, Y_j)$$

## 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 & Cartesian Closed Category ($\mathbf{CCC}$)

- An *exponential* of objects $A, B$ is an object $B^A$ with a morphism $\varepsilon : B^A \times A \to B$ s.t.

$$\forall C \forall f : C \times A \to B \; \exists! \; \hat{f} : C \to B^A \left[ \varepsilon \circ (\hat{f} \times 1_A) = f \right]$$

$$\begin{tikzcd} B^A & B^A \times A \arrow[r, "\varepsilon"] & B \\ C \arrow[u, dashed, "\hat{f}"] & C \times A \arrow[u, dashed, "\hat{f} \times 1_A"] \arrow[ur, "f"'] & \end{tikzcd}$$

- **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.

# Curry-Howard-Lambek Isomorphism

- **Objects** types/formulas $\top \mid A \times B \mid A \to B$
- **Morphisms** terms/proofs $1_A \mid ! \mid g \circ f \mid \langle f, g \rangle \mid \pi_1 \mid \pi_2 \mid \hat{f} \mid \varepsilon$
  $f : A \to B \iff A \vdash B$

$$\frac{}{1_A : A \vdash A} \qquad \frac{}{! : A \vdash \top} \qquad \frac{f : A \vdash B \qquad g : B \vdash C}{g \circ f : A \vdash C}$$

$$\frac{f : C \vdash A \qquad g : C \vdash B}{\langle f, g \rangle : C \vdash A \times B} \qquad \frac{}{\pi_1 : A \times B \vdash A} \qquad \frac{}{\pi_2 : A \times B \vdash B}$$

$$\frac{f : A \times B \vdash C}{\hat{f} : A \vdash B \to C} \qquad \frac{}{\varepsilon : (A \to B) \times A \vdash B}$$

# Curry-Howard-Lambek Isomorphism

| **Logic** | **Type Theory** | **Category Theory** |
|---|---|---|
| Formula | Type | Object |
| Proof | Term/Program | Morphism |
| false $\perp$ | 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 $\rightarrow$ | function type $\rightarrow$ | exponential $B^A$ |
| cut-elimination | $\beta$-reduction | composition $\circ$ |
| modus ponens | application $\mathrm{app}$ | evaluation $\varepsilon$ |

$$\begin{tikzcd}
\text{Logic} \arrow[rr, leftrightarrow] \arrow[dr, leftrightarrow] & & \text{Computation} \arrow[dl, leftrightarrow] \\
& \text{Categories} &
\end{tikzcd}$$

*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 $\mathrm{Hom}$ | 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# Functor

## Definition (Functor)

A functor $F : \mathbf{C} \to \mathbf{D}$ between categories $\mathbf{C}$ and $\mathbf{D}$ is a mapping of objects to objects and morphisms to morphisms, in such a way that:

- $F(f : A \to B) = Ff : FA \to FB$
- $F(1_A) = 1_{FA}$
- $F(g \circ f) = Fg \circ Ff$

identity functor $1_{\mathbf{C}}$ and composition of functors

$$\begin{tikzcd} A \arrow[d, "f", "1_{\mathbf{C}}"'] \\ B \end{tikzcd} = \begin{tikzcd} A \arrow[d, "f"] \\ B \end{tikzcd} \qquad \begin{tikzcd} A \arrow[d, "f", "G \circ F"'] \\ B \end{tikzcd} = \begin{tikzcd} G(FA) \arrow[d, "G(Ff)"] \\ G(FB) \end{tikzcd}$$

# 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 $\mathbf{C}$.

$$\begin{tikzcd} A \arrow[d, "f"'] \arrow[r, phantom, "\overset{\Delta_X}{\Longrightarrow}"] & X \arrow[d, "1_X"] \\ B & X \end{tikzcd}$$

## 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 $\Delta_X$ of shape $\mathbf{I}$ in $\mathbf{C}$ where all objects are copies of $X$ and all morphisms are copies of $1_X : X \to X$.

$$\begin{aligned}(\Delta_X)i &:= X \text{ for } i \in \mathbf{I} \\ (\Delta_X)f &:= 1_X \text{ for } f \in \mathbf{I} \\ (\Delta f)i &:= f \text{ for } f \in \mathrm{Hom}_{\mathbf{C}}(X, Y) \text{ and } i \in \mathbf{I}\end{aligned}$$

**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

$$\begin{tikzcd} & X \arrow[dl, "1_X"'] \arrow[d, "\Delta"] \arrow[dr, "1_X"] & \\ X & X \times X \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & X \end{tikzcd}.$$

**Remark:** The diagonal morphisms in $\mathbf{Cat}$ are diagonal functors.

# 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 $\gamma$ 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" $f^{-1}(t) := f(1-t)$.
- 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 $S^1 = \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\right\}$.

## 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 $\pi_1$ 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

$$\pi_1(r) \cdot \pi_1(i) = \pi_1(ri) = \pi_1(1_{S^1}) = 1_{\pi_1(S^1)}$$

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: $\forall T \in \mathbf{C} \forall a : T \to A \forall s : T \to S[h \circ a = i \circ s \to p \circ a = s]$

$$\begin{tikzcd} T \arrow[r, "a"] \arrow[d, "s"'] & A \arrow[d, "h"] \arrow[dl, "p"'] \\ S \arrow[r, "i"'] & B \end{tikzcd}$$

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 $\neq$ 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 $S \xrightarrow{i} B$.

# 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* $d(pq) = d(p)q + pd(q)$

# **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.

$$q_y = \sum_{x \in f^{-1}(y)} p_x$$

$$\begin{tikzcd} & 1 \arrow[dl, squiggly, "p"'] \arrow[dr, squiggly, "q"] & \\ X \arrow[rr, "f"'] & & Y \end{tikzcd}$$

- 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 $p$.

$$D_{\mathsf{KL}}(p\|q) := \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}$$

$$\begin{tikzcd}
 & 1 \arrow[dl, squiggly, "p"'] \arrow[dr, squiggly, "q"] & \\
X \arrow[rr, bend right, "f"'] & & Y \arrow[ll, squiggly, "s"']
\end{tikzcd}$$

$$f \circ p = q$$
$$f \circ s = 1_Y$$

- Think of $f$ as a 'measurement process' and $s$ as a 'hypothesis' about the state in $X$ given the measurement in $Y$.
- $D_{\mathsf{KL}}(p\|s \circ q)$: 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 $s$.

# 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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:

$$\begin{tikzcd}
FA \arrow[r, "Ff"] \arrow[d, "\alpha_A"'] \arrow[dr, phantom, "\circlearrowleft"] & FB \arrow[d, "\alpha_B"] \\
GA \arrow[r, "Gf"'] & GB
\end{tikzcd}$$

$$\begin{tikzcd}
FA \arrow[rr, "Ff"] \arrow[dd, "\alpha_A"'] & & FB \arrow[dd, "\alpha_B"] \\
& A \arrow[ul, "F"'] \arrow[dl, "G"] \arrow[r, "f"] & B \arrow[ur, "F"] \arrow[dr, "G"'] \\
GA \arrow[rr, "Gf"'] & & GB
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbf{C} \arrow[r, bend left, "F"] \arrow[r, bend right, "G"'] \arrow[r, phantom, "\Downarrow\alpha"] & \mathbf{D}
\end{tikzcd}$$

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.

$$\begin{tikzcd}
FA \arrow[r, "Ff"] \arrow[d, shift right, "\alpha_A"'] & FB \arrow[d, shift right, "\alpha_B"] \\
GA \arrow[r, "Gf"'] \arrow[u, shift right, "\alpha_A^{-1}"'] & GB \arrow[u, shift right, "\alpha_B^{-1}"']
\end{tikzcd}$$

$$Ff = \alpha_B^{-1} \circ Gf \circ \alpha_A$$
$$Gf = \alpha_B \circ Ff \circ \alpha_A^{-1}$$

# Remark

$$\begin{tikzcd}
\mathbf{C} \arrow[r, bend left=60, "F"] \arrow[r, bend right=60, "G"'] \arrow[r, phantom, "\Downarrow\alpha \quad \Uparrow\alpha^{-1}"] & \mathbf{D} \arrow[loop above, "1_F"] \arrow[loop below, "1_G"]
\end{tikzcd}$$

- 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.

# Contents


Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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).

$$\begin{tikzcd}
A \arrow[r, shift left, "f_1"] \arrow[r, shift right, "f_2"'] \arrow[d, dash, "\cong"'] & B \arrow[d, dash, "\cong"] \\
A' \arrow[r, shift left, "f_3"] \arrow[r, shift right, "f_4"'] \arrow[ur, crossing over] & B'
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{C} \arrow[r, shift left, "F"] & \mathbf{D} \arrow[l, shift left, "G"]
\end{tikzcd}
\qquad
\begin{tikzcd}
X \arrow[r, "g"] & Y
\end{tikzcd}$$

$$F := \begin{cases} A & \mapsto X \\ A' & \mapsto X \\ B & \mapsto Y \\ B' & \mapsto Y \\ f_i & \mapsto g \end{cases} \qquad G := \begin{cases} X & \mapsto A \\ Y & \mapsto B \\ g & \mapsto f_1 \end{cases} \qquad \begin{aligned} FG &\cong 1_{\mathbf{D}} \\ GF &\cong 1_{\mathbf{C}} \end{aligned}$$

# Examples

- The following categories are equivalent, but not isomorphic.

- $\mathrm{CPL}$ and $\mathrm{IPL}$ are not equivalent.

$$\begin{tikzcd} \mathrm{CPL} \arrow[r, bend left, "\neg\neg"] & \mathrm{IPL} \arrow[l, bend left, "i"] \end{tikzcd}$$

$$A \equiv_{\mathrm{CPL}} i(\neg\neg(A)) \quad \text{but} \quad A \not\equiv_{\mathrm{IPL}} \neg\neg(i(A))$$

# 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, $\alpha$ 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, $\beta_A$ is an isomorphism for every $A$. Naturality of $\beta_A$ follows from functoriality of $F$ and naturality of $\alpha$.

# Contents


Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# Whiskering

$$\begin{tikzcd} \mathbf{C} \arrow[r, "F"] & \mathbf{D} \arrow[r, bend left, "G_1"] \arrow[r, bend right, "G_2"'] \arrow[r, phantom, "\Downarrow\beta"] & \mathbf{E} \end{tikzcd} \qquad \begin{tikzcd} \mathbf{C} \arrow[r, bend left, "G_1"] \arrow[r, bend right, "G_2"'] \arrow[r, phantom, "\Downarrow\beta"] & \mathbf{D} \arrow[r, "F"] & \mathbf{E} \end{tikzcd}$$

The prewhiskering of $\beta$ by $F$, denoted $\beta F \coloneqq \beta * 1_F : G_1 F \to G_2 F$ (resp. the postwhiskering of $\beta$ by $F$, denoted $F\beta \coloneqq 1_F * \beta : F G_1 \to F G_2$) is defined as follows.

$$\begin{tikzcd} \mathbf{C} \arrow[r, bend left, "G_1 F"] \arrow[r, bend right, "G_2 F"'] \arrow[r, phantom, "\Downarrow\beta F"] & \mathbf{E} \end{tikzcd} \qquad \begin{tikzcd} \mathbf{C} \arrow[r, bend left, "F G_1"] \arrow[r, bend right, "F G_2"'] \arrow[r, phantom, "\Downarrow F\beta"] & \mathbf{E} \end{tikzcd}$$

$$(\beta F)_A \coloneqq \beta_{FA}$$
$$(F\beta)_A \coloneqq F\beta_A$$

# Vertical and Horizontal Composition

$$\begin{tikzcd}
\mathbf{C} \arrow[r, bend left=60, "F"] \arrow[r, bend right=60, "H"'] \arrow[r, "G"] \arrow[r, phantom, shift left=3, "\Downarrow\alpha"] \arrow[r, phantom, shift right=3, "\Downarrow\beta"] & \mathbf{D}
\end{tikzcd}
\implies
\begin{tikzcd}
\mathbf{C} \arrow[r, bend left=60, "F"] \arrow[r, bend right=60, "H"'] \arrow[r, phantom, "\Downarrow\beta\circ\alpha"] & \mathbf{D}
\end{tikzcd}$$

$$(\beta \circ \alpha)_A \coloneqq \beta_A \circ \alpha_A$$

$$\begin{tikzcd}
\mathbf{C} \arrow[r, bend left=60, "F_1"] \arrow[r, bend right=60, "F_2"'] \arrow[r, phantom, "\Downarrow\alpha"] & \mathbf{D} \arrow[r, bend left=60, "G_1"] \arrow[r, bend right=60, "G_2"'] \arrow[r, phantom, "\Downarrow\beta"] & \mathbf{E}
\end{tikzcd}
\implies
\begin{tikzcd}
\mathbf{C} \arrow[r, bend left=60, "G_1F_1"] \arrow[r, bend right=60, "G_2F_2"'] \arrow[r, phantom, "\Downarrow\beta*\alpha"] & \mathbf{E}
\end{tikzcd}$$

$$\begin{tikzcd}
G_1F_1A \arrow[r, "G_1\alpha_A"] \arrow[d, "\beta_{F_1A}"'] \arrow[dr, "(\beta*\alpha)_A"] & G_1F_2A \arrow[d, "\beta_{F_2A}"] \\
G_2F_1A \arrow[r, "G_2\alpha_A"'] & G_2F_2A
\end{tikzcd}$$

# Middle Four Interchange

## Theorem (Middle Four Interchange)

$$\begin{tikzcd}
\mathbf{C} \arrow[r, bend left=60, "F_1"] \arrow[r, bend right=60, "H_1"'] \arrow[r, "G_1" description] \arrow[r, phantom, shift left=3, "\Downarrow\alpha_1"] \arrow[r, phantom, shift right=3, "\Downarrow\beta_1"] & \mathbf{D} \arrow[r, bend left=60, "F_2"] \arrow[r, bend right=60, "H_2"'] \arrow[r, "G_2" description] \arrow[r, phantom, shift left=3, "\Downarrow\alpha_2"] \arrow[r, phantom, shift right=3, "\Downarrow\beta_2"] & \mathbf{E}
\end{tikzcd}$$

$$(\beta_2 * \beta_1) \circ (\alpha_2 * \alpha_1) = (\beta_2 \circ \alpha_2) * (\beta_1 \circ \alpha_1)$$

# 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 $1_{(A,B)} := (1_A, 1_B)$.
- composition $(f', g') \circ (f, g) := (f' \circ f, g' \circ g)$.

$$
\begin{tikzcd}
\mathbf{C} \\
A \arrow[d, "f"] \\
A' \arrow[d, "f'"] \\
A''
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{D} \\
B \arrow[d, "g"] \\
B' \arrow[d, "g'"] \\
B''
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{C} \times \mathbf{D} \\
(A, B) \arrow[d, "(f,g)"] \\
(A', B') \arrow[d, "(f',g')"] \\
(A'', B'')
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{C} \times \mathbf{D} \\
(A, B) \arrow[dd, "(f' \circ f, g' \circ g)"] \\
\\
(A'', B'')
\end{tikzcd}
$$

# 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

$$\begin{tikzcd}
\mathbf{C} \arrow[r, bend left=60, "F"] \arrow[r, bend right=60, "H"'] \arrow[r, "G"] \arrow[r, phantom, shift left=3, "\Downarrow\alpha"] \arrow[r, phantom, shift right=3, "\Downarrow\beta"] & \mathbf{D}
\end{tikzcd}
\implies
\begin{tikzcd}
\mathbf{C} \arrow[r, bend left=60, "F"] \arrow[r, bend right=60, "H"'] \arrow[r, phantom, "\Downarrow\beta\circ\alpha"] & \mathbf{D}
\end{tikzcd}
\qquad
\begin{tikzcd}
FA \arrow[r, "(\beta\circ\alpha)_A"] \arrow[d, "Ff"'] & HA \arrow[d, "Hf"] \\
FB \arrow[r, "(\beta\circ\alpha)_B"'] & HB
\end{tikzcd}$$

$$(\beta \circ \alpha)_A \coloneqq \beta_A \circ \alpha_A$$

- 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 $\mathbf{1} := \overset{1}{\bullet}\circlearrowleft$ .
- $\mathbf{Cat}$ 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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.

$$\begin{tikzcd}
C \arrow[d, dashed, "u"'] \arrow[dr, "h"] & & \\
E \arrow[r, "e"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B
\end{tikzcd}$$

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 $\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \arrow[r, "h"] & C \end{tikzcd}$ s.t. $hf = hg$.
- Let $\mathbf{C}$ be a category and take $\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \end{tikzcd}$. An *coequalizer* of $f$ and $g$ is an object $Q$ with a morphism $B \xrightarrow{q} Q$ s.t. $\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \arrow[r, "q"] & Q \end{tikzcd}$ is a cofork, and for any cofork $\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \arrow[r, "h"] & C \end{tikzcd}$, there exists a unique morphism $Q \xrightarrow{u} C$ s.t.

$$\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \arrow[r, "q"] \arrow[dr, "h"'] & Q \arrow[d, dashed, "u"] \\ & & C \end{tikzcd}$$

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

$$\{\bullet\} \underset{\bullet \mapsto 2\pi}{\overset{\bullet \mapsto 0}{\rightrightarrows}} [0, 2\pi] \xrightarrow{\theta \mapsto (\cos\theta, \sin\theta)} \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\right\}$$

## 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
$$\begin{tikzcd} & B \arrow[d, "g"] \\ A \arrow[r, "f"] & C \end{tikzcd}$$
is an object $P$ with $P \xrightarrow{p_1} A$ and $P \xrightarrow{p_2} B$ s.t.
$$\begin{tikzcd} P \arrow[r, "p_2"] \arrow[d, "p_1"'] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$
, and for any
$$\begin{tikzcd} Q \arrow[r, "q_2"] \arrow[d, "q_1"'] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$
there is a unique morphism $Q \xrightarrow{u} P$ s.t.
$$\begin{tikzcd} Q \arrow[dr, dashed, "u"] \arrow[drr, bend left, "q_2"] \arrow[ddr, bend right, "q_1"'] & & \\ & P \arrow[r, "p_2"] \arrow[d, "p_1"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, "g"] \\ & A \arrow[r, "f"'] & C \end{tikzcd}$$
.

# Pullback

$$\begin{tikzcd} P \arrow[r, "p_2"] \arrow[d, "p_1"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$

In $\mathbf{Set}$, $P := A \times_C B := \{(a,b) \in A \times B : fa = gb\} = \coprod_{c \in C} f^{-1}(c) \times g^{-1}(c)$.

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 $i : B \hookrightarrow C$.

$$\begin{tikzcd} f^*B \arrow[r] \arrow[d, "f^*g"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd} \qquad \begin{tikzcd} f^*B \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, hook, "i"] \\ A \arrow[r, "f"'] & C \end{tikzcd} \qquad \begin{tikzcd} f^{-1}B \arrow[r, "f\restriction_{f^{-1}B}"] \arrow[d, hook, "i"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, hook, "i"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$

In this case $\{(a,b) \in A \times B : fa = b\} \cong \{a \in A : fa \in B\} = f^*B$.

# Example — Pullback

$$\begin{tikzcd}
\overset{A \times_C B}{\boxed{\text{a customer that is wealthy and loyal}}} \arrow[r, "\text{is}"] \arrow[d, "\text{is}"'] & \overset{B}{\boxed{\text{a loyal customer}}} \arrow[d, "\text{is}"] \\
\overset{A}{\boxed{\text{a wealthy customer}}} \arrow[r, "\text{is}"'] & \overset{C}{\boxed{\text{a customer}}}
\end{tikzcd}$$

$$\begin{tikzcd}
\overset{A \times_C B}{\boxed{\text{a dog whose owner is a woman}}} \arrow[r, "\text{has as owner}"] \arrow[d, "\text{is}"'] & \overset{B}{\boxed{\text{a woman}}} \arrow[d, "\text{is}"] \\
\overset{A}{\boxed{\text{a dog}}} \arrow[r, "\text{has as owner}"'] & \overset{C}{\boxed{\text{a person}}}
\end{tikzcd}$$

# products & equalizers $\implies$ pullbacks

## Theorem

*In a category with products and equalizers, given* $A \xrightarrow{f} C \xleftarrow{g} B$ *, consider the diagram:*

$$\begin{tikzcd}
E \arrow[dr, "e"] & & \\
& A \times B \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & B \arrow[d, "g"] \\
& A \arrow[r, "f"'] & C
\end{tikzcd}$$

*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*

$$\begin{tikzcd}
E \arrow[r, "\pi_2 \circ e"] \arrow[d, "\pi_1 \circ e"'] \arrow[dr, phantom, "\lrcorner"] & B \arrow[d, "g"] \\
A \arrow[r, "f"'] & C
\end{tikzcd}$$

*is a pullback.*

- pullbacks & terminal objects $\implies$ products

  $A \times B \cong A \times_1 B$

$$\begin{tikzcd} A \times B \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, "!_B"] \\ A \arrow[r, "!_A"'] & 1 \end{tikzcd}$$

- pullbacks & products $\implies$ equalizers

  The equalizer $E \overset{e}{\rightarrowtail} A \underset{g}{\overset{f}{\rightrightarrows}} B$ is constructed as the following pullback,

$$\begin{tikzcd} E \arrow[r, "h"] \arrow[d, "e"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, "\Delta"] \\ A \arrow[r, "\langle f,g \rangle"'] & B \times B \end{tikzcd}$$

## Theorem

$$\begin{tikzcd} A \times_C B \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$

*If $g$ is monic, so is $p$.*

## Theorem

$$\begin{tikzcd} F \arrow[r] \arrow[d] & E \arrow[r] \arrow[d] & D \arrow[d] \\ A \arrow[r] & B \arrow[r] & C \end{tikzcd}$$

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 $\begin{tikzcd} C \arrow[r, "g"] \arrow[d, "f"'] & B \\ A & \end{tikzcd}$ is an object $P$ with

$A \xrightarrow{i_1} P$ and $B \xrightarrow{i_2} P$ s.t. $\begin{tikzcd} C \arrow[r, "g"] \arrow[d, "f"'] & B \arrow[d, "i_2"] \\ A \arrow[r, "i_1"'] & P \end{tikzcd}$, and for any $\begin{tikzcd} C \arrow[r, "g"] \arrow[d, "f"'] & B \arrow[d, "j_2"] \\ A \arrow[r, "j_1"'] & Q \end{tikzcd}$ there is

a unique morphism $P \xrightarrow{u} Q$ s.t. $\begin{tikzcd} C \arrow[r, "g"] \arrow[d, "f"'] & B \arrow[d, "i_2"] \arrow[ddr, bend left, "j_2"] & \\ A \arrow[r, "i_1"'] \arrow[drr, bend right, "j_1"'] & P \arrow[dr, dashed, "u"] \arrow[ul, phantom, "\ulcorner", very near start] & \\ & & Q \end{tikzcd}$.

# Pushout

$$\begin{tikzcd} C \arrow[r, "g"] \arrow[d, "f"'] & B \arrow[d, "i_2"] \\ A \arrow[r, "i_1"'] \arrow[ur, phantom, "\ulcorner"] & A +_C B \end{tikzcd}$$

In $\mathbf{Set}$, $A +_C B := (A + B)/{\sim}$, where $\sim$ is the smallest equivalence relation on $A + B$ such that for all $x \in C : fx = gx$.

$$\begin{tikzcd} A \cap B \arrow[r, hook] \arrow[d, hook] & B \arrow[d, hook] \\ A \arrow[r, hook] \arrow[ur, phantom, "\ulcorner"] & A \cup B \end{tikzcd}$$

$$A +_{A \cap B} B = A \cup B$$

# Example — Pushout

$$\begin{tikzcd}
\left\{(x,y,0) \in \mathbb{R}^3 : x^2+y^2=1\right\} \arrow[r] \arrow[d] & \left\{(x,y,z) \in \mathbb{R}^3 : \begin{array}{l} x^2+y^2+z^2=1 \\ z \geq 0 \end{array}\right\} \arrow[d, hook] \\
\left\{(x,y,z) \in \mathbb{R}^3 : \begin{array}{l} x^2+y^2+z^2=1 \\ z \leq 0 \end{array}\right\} \arrow[r, hook] \arrow[ur, phantom, "\ulcorner"] & \left\{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2=1\right\}
\end{tikzcd}$$

## Theorem

$$\begin{tikzcd} C \arrow[r, "g"] \arrow[d, "f"'] \arrow[dr, phantom, "\ulcorner"] & B \arrow[d, "j"] \\ A \arrow[r, "i"'] & A +_C B \end{tikzcd}$$

*If $f$ is epic, so is $j$.*

## Theorem

$$\begin{tikzcd} C \arrow[r] \arrow[d] & B \arrow[r] \arrow[d] & A \arrow[d] \\ D \arrow[r] & E \arrow[r] & F \end{tikzcd}$$

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.*

# 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 a morphism $f : A \to B$ is the equalizer of $f$ and the zero morphism $0_{AB}$.

$$\ker(f) = \mathrm{eq}(f, 0_{AB})$$

- In a category with zero morphism, the *cokernel* $\mathrm{coker}(f)$ of a morphism $f : A \to B$ is the coequalizer of $f$ and the zero morphism $0_{AB}$.

$$\mathrm{coker}(f) = \mathrm{coeq}(f, 0_{AB})$$

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:

$$\begin{tikzcd} \ker(f) \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & 0 \arrow[d] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

- In a category with a terminal object $1$, the *cokernel* $\operatorname{coker}(f)$ of a morphism $f : A \to B$ is the pushout:

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d] & B \arrow[d] \\ 1 \arrow[r] \arrow[ur, phantom, "\ulcorner", very near start] & \operatorname{coker}(f) \end{tikzcd}$$

# 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.

$$\begin{tikzcd}
 & D_i \arrow[d, "Df"] \\
C \arrow[ur, "c_i"] \arrow[r, "c_j"'] & D_j
\end{tikzcd}$$

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 $\theta$ s.t.

$$\begin{tikzcd}
C \arrow[d, "\theta"'] \arrow[dr, "c_i"] & \\
C' \arrow[r, "c'_i"'] & D_i
\end{tikzcd}$$

Then we have a category $\mathbf{Cone}(D)$ of 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.

$$\begin{tikzcd}
D_i \arrow[dr, "c_i"] \arrow[dd, "Df"'] & \\
& C \\
D_j \arrow[ur, "c_j"'] &
\end{tikzcd}$$

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 $\theta$ s.t.

$$\begin{tikzcd}
& C \arrow[d, "\theta"] \\
D_i \arrow[ur, "c_i"] \arrow[dr, "c'_i"'] & \\
& C'
\end{tikzcd}$$

Then we have a category $\mathbf{Cocone}(D)$ of cones over $D$.

# Limit

- A *limit* $\left(\varprojlim D, p\right)$ for a diagram $D : \mathbf{I} \to \mathbf{C}$ is a terminal object in $\mathbf{Cone}(D)$. 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 $i \in \mathbf{I}$.

$$\begin{tikzcd}
 & & D_i \\
C \arrow[urr, "c_i"] \arrow[r, dashed, "u"] \arrow[drr, "c_j"'] & \varprojlim D \arrow[ur, "p_i"] \arrow[dr, "p_j"'] & \\
 & & D_j
\end{tikzcd}$$

**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 $\mathbf{Cocone}(D)$. 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 $i \in \mathbf{I}$.

$$\begin{tikzcd}
D_i \arrow[dr, "p_i"'] \arrow[drr, "c_i"] & & \\
& \varinjlim D \arrow[r, dashed, "u"] & C \\
D_j \arrow[ur, "p_j"] \arrow[urr, "c_j"'] & &
\end{tikzcd}$$

**Remark:** There is an isomorphism natural in $C$.

$$\mathbf{C}^{\mathbf{I}}(D, \Delta_C) \cong \mathbf{C}(\varinjlim D, C)$$

# Inverse/Direct Limit

- Inverse Limit

$$\begin{tikzcd}
\varprojlim D \arrow[d] \arrow[dr] \arrow[drr] \arrow[drrr] & & & \\
D_0 & D_1 \arrow[l, "f_0"'] & D_2 \arrow[l, "f_1"'] & \cdots \arrow[l, "f_2"']
\end{tikzcd}$$

- Direct Limit

$$\begin{tikzcd}
D_0 \arrow[r, "f_0"] \arrow[d] & D_1 \arrow[r, "f_1"] \arrow[dl] & D_2 \arrow[r, "f_2"] \arrow[dll] & \cdots \arrow[dlll] \\
\varinjlim D & & &
\end{tikzcd}$$

# Inverse/Direct Limit — Example

- Inverse Limit

$$\begin{tikzcd}
0 & X \arrow[l, dashed] \arrow[dl] \arrow[dl] \arrow[d] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \\
1 & 0.1 \arrow[l] & 0.01 \arrow[l] & \cdots \arrow[l]
\end{tikzcd}$$

- Direct Limit

$$\begin{tikzcd}
0.9 \arrow[r] \arrow[d] \arrow[dr] & 0.99 \arrow[r] \arrow[d] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] & 0.999 \arrow[r] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] & \cdots \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \arrow[dl] \\
1 \arrow[r, dashed] & X
\end{tikzcd}$$

# Inverse/Direct Limit — Example

$$\begin{tikzcd}
\varprojlim x \arrow[d] \arrow[dr] \arrow[drr] \arrow[drrr] & & & \\
x_0 & x_1 \arrow[l] & x_2 \arrow[l] & \cdots \arrow[l]
\end{tikzcd}$$

Let $\mathbb{R}$ be the category whose objects are real numbers and which has a unique morphism $x \to y$ iff $x \leq y$. A functor $x : \mathbb{Z}_{\leq 0} \to \mathbb{R}$ 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 "$\varepsilon$-$\delta$" 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** $\varprojlim D$ | $\mathbf{I}$ | **Colimit** $\varinjlim D$ |
|---|---|---|
| terminal object | $\emptyset$ | initial object |
| binary product | $\bullet \quad \bullet$ | binary coproduct |
| equalizor | $\bullet \rightrightarrows \bullet$ | coequalizor |
| inverse limit | $\bullet \leftarrow \bullet \leftarrow \bullet \leftarrow \cdots$ | |
| | $\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \cdots$ | direct limit |
| pullback | $\begin{tikzcd} \bullet \arrow[dr] & \\ & \bullet \\ \bullet \arrow[ur] & \end{tikzcd}$ | |
| | $\begin{tikzcd} & \bullet \\ \bullet \arrow[ur] \arrow[dr] & \\ & \bullet \end{tikzcd}$ | pushout |

# Exercise

Suppose we want to take the colimit of the diagram.

$$\begin{tikzcd}
 & B \arrow[r] \arrow[d] & Z \\
A \arrow[r] \arrow[d] & Y & \\
X & &
\end{tikzcd}
\qquad
\begin{tikzcd}
 & B \arrow[r] \arrow[d] & Z \arrow[d] \\
A \arrow[r] \arrow[d] & Y \arrow[r] \arrow[d] \arrow[ur, phantom, "\ulcorner"] & D \\
X \arrow[r] & C \arrow[ur, phantom, "\ulcorner"] &
\end{tikzcd}
\qquad
\begin{tikzcd}
 & B \arrow[r] \arrow[d] & Z \arrow[d] \\
A \arrow[r] \arrow[d] & Y \arrow[r] \arrow[d] \arrow[ur, phantom, "\ulcorner"] & D \arrow[d] \\
X \arrow[r] & C \arrow[r] \arrow[ur, phantom, "\ulcorner"] & P \arrow[ur, phantom, "\ulcorner"]
\end{tikzcd}$$

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 $\implies$ 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 $\alpha : i \to j \in \operatorname{mor} \mathbf{I}$.

$$\begin{tikzcd}
C \arrow[d, dashed, "u"'] \arrow[dr, "c"] & & & \\
E \arrow[r, tail, "e"] & \prod\limits_{i \in \operatorname{ob}\mathbf{I}} D_i \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] \arrow[d, "\pi_i"'] \arrow[dr, "\pi_j"] & \prod\limits_{\alpha : i \to j \in \operatorname{mor}\mathbf{I}} D_j \arrow[d, "\pi_\alpha"] & \\
& D_i \arrow[r, "D\alpha"'] \arrow[ur, phantom, "?"] & D_j &
\end{tikzcd}$$

but generally $D\alpha \circ \pi_i \neq \pi_j$

Take any $c : C \to \prod\limits_{i \in \operatorname{ob}\mathbf{I}} D_i$. 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 $c : C \to \prod\limits_{i \in \operatorname{ob}\mathbf{I}} D_i$, 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 $E = \varprojlim D$.

# Coproducts & Coequalizers $\implies$ 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 $g_\alpha := g \circ \iota_\alpha = \iota_{\operatorname{dom}(\alpha)}$.

$$\begin{tikzcd}
\coprod_{\alpha \in \operatorname{mor} \mathbf{I}} D_{\operatorname{dom}(\alpha)} \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & \coprod_{i \in \operatorname{ob} \mathbf{I}} D_i \arrow[r, two heads, "q"] \arrow[dr, "h"] & \varinjlim D \arrow[d, dashed, "u"] \\
D_{\operatorname{dom}(\alpha)} \arrow[u, "\iota_\alpha"] \arrow[ur, "\iota_{\operatorname{dom}(\alpha)}"] \arrow[r, "D_\alpha"'] \arrow[ur, phantom, shift right=4ex, "?"] & D_{\operatorname{cod}(\alpha)} \arrow[u, "\iota_{\operatorname{cod}(\alpha)}"'] & C
\end{tikzcd}$$

### 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 $D$,

$$(C, c) \text{ is a limit over } D \Longrightarrow (FC, Fc) \text{ is a limit over } FD : \mathbf{I} \to \mathbf{D}$$

- 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 $D$,

$$(C, c) \text{ is a limit over } D \Longleftarrow (FC, Fc) \text{ is a limit over } FD : \mathbf{I} \to \mathbf{D}$$

- 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 $F\left(\varprojlim D\right) \cong \varprojlim(FD)$.

# Examples

$$\begin{tikzcd} P \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d] \\ A \arrow[r] & C \end{tikzcd}$$

preserves $\Downarrow$

$$\begin{tikzcd} FP \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & FB \arrow[d] \\ FA \arrow[r] & FC \end{tikzcd}$$

$$\begin{tikzcd} P \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d] \\ A \arrow[r] & C \end{tikzcd}$$

reflects $\Uparrow$

$$\begin{tikzcd} FP \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & FB \arrow[d] \\ FA \arrow[r] & FC \end{tikzcd}$$

$$\begin{tikzcd} P \arrow[r, dashed] \arrow[d, dashed] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d] \\ A \arrow[r] & C \end{tikzcd}$$

creates $\Uparrow$

$$\begin{tikzcd} Q \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & FB \arrow[d] \\ FA \arrow[r] & FC \end{tikzcd}$$

## 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 $\kappa$ copies of $B$, where $\kappa$ is the cardinality of the set of arrows in $\mathbf{C}$.
Thus, we can construct $2^\kappa$ distinct cones $(A, \{A \to B\}_\kappa)$ by combining $f$ and $g$ in every possible way.
Thus, we have $2^\kappa$ 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

$$F_0 := 1$$
$$F_1 := 1$$
$$F_n := F_{n-1} + F_{n-2}$$

$$\gcd(F_m, F_n) = F_{\gcd(m,n)}$$

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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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.

# 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 $f, g : A \to B$ and $h : B \to C$, if $f \sim_{AB} g$ then $hf \sim_{AC} hg$.
  - for $f : A \to B$ and $g, h : B \to C$, if $g \sim_{BC} h$ then $gf \sim_{AC} hf$.
- 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,

$$\mathrm{Hom}_{\mathbf{C}/\sim}(A, B) = \mathrm{Hom}_{\mathbf{C}}(A, B)/\sim_{AB}$$

# Slice/Coslice Category

## Definition (Slice/Coslice 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.

$$\begin{tikzcd} & B \arrow[dl, "f"'] \arrow[d, "g"] \\ A & B' \arrow[l, "f'"] \end{tikzcd}$$

- 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.

$$\begin{tikzcd} & B \arrow[d, "g"] \\ A \arrow[ur, "f"] \arrow[r, "f'"'] & B' \end{tikzcd}$$

# 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": $f/\mathbf{C} := - \circ f : B/\mathbf{C} \to A/\mathbf{C}$.
- A new metaphor functor $F : B/\mathbf{C} \to A/\mathbf{C}$ is created by a natural transformation $\alpha$ from the base-of-metaphor functor $f/\mathbf{C}$.

$$\begin{tikzcd}
A/\mathbf{C} \arrow[r, bend right, "-\circ f"'] \arrow[r, bend left, "F"] \arrow[r, phantom, "\Uparrow\alpha"] & B/\mathbf{C}
\end{tikzcd}$$

$$\begin{tikzcd}
F(b) \arrow[d] \arrow[dr, phantom, "\circlearrowleft"] & b \circ f \arrow[l, "\alpha_b"'] \arrow[d] \\
F(b') & b' \circ f \arrow[l, "\alpha_{b'}"]
\end{tikzcd}$$

# Metaphor Example — "Wing is like Sail"

| The Ground — A Japanese Poem |
| --- |
| An ant<br>pull a wing of a butterfly.<br>O<br>It is like yachting. |

$$\begin{tikzcd}
F(\text{yacht}) = \text{pull} \arrow[d] \arrow[dr, phantom, "\circlearrowleft"] & \text{yacht} \circ f \arrow[l, "\alpha_{\text{yacht}}"'] \arrow[d] \\
F(\text{over}) = \text{on} & \text{over} \circ f \arrow[l, "\alpha_{\text{over}}"]
\end{tikzcd}$$

# 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

$$\begin{tikzcd} & B \arrow[dl, "f"'] \arrow[d, "g"] \\ A & B' \arrow[l, "f'"] \end{tikzcd} \quad \text{to} \quad \begin{tikzcd} & B \arrow[dl, "hf"'] \arrow[d, "g"] \\ A' & B' \arrow[l, "hf'"] \end{tikzcd}$$

- 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 $\overline{\mathbf{C}} : \mathbf{C} \to \mathbf{Cat}/\mathbf{C}$.

$$\begin{tikzcd} \mathbf{C} \arrow[r, "\overline{\mathbf{C}}"] \arrow[dr, "\mathbf{C}/-"'] & \mathbf{Cat}/\mathbf{C} \arrow[d, "U_{\mathbf{C}}"] \\ & \mathbf{Cat} \end{tikzcd}$$

where $\overline{\mathbf{C}} : A \mapsto (\mathbf{C}/A, U_A)$ and $\overline{\mathbf{C}} : h \mapsto$

$$\begin{tikzcd} \mathbf{C}/A \arrow[rr, "\mathbf{C}/h"] \arrow[dr, "U_A"'] & & \mathbf{C}/A' \arrow[dl, "U_{A'}"] \\ & \mathbf{C} & \end{tikzcd}$$

# 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.

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "a"'] & B \arrow[d, "b"] \\ A' \arrow[r, "f'"'] & B' \end{tikzcd}$$

The composition is given by

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "a"'] & B \arrow[d, "b"] \\ A' \arrow[r, "f'"] \arrow[d, "a'"'] & B' \arrow[d, "b'"] \\ A'' \arrow[r, "f''"] & B'' \end{tikzcd}$$

# 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.

$$\begin{tikzcd} A \arrow[r, "f"] & B \arrow[d, "b"] \\ A' \arrow[u, "a"] \arrow[r, "f'"'] & B' \end{tikzcd}$$

The composition is given by

$$\begin{tikzcd} A \arrow[r, "f"] & B \arrow[d, "b"] \\ A' \arrow[u, "a"] \arrow[r, "f'"] & B' \arrow[d, "b'"] \\ A'' \arrow[u, "a'"] \arrow[r, "f''"] & B'' \end{tikzcd}$$

**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.

$$\begin{tikzcd} FA \arrow[r, "f"] \arrow[d, "Fa"'] & GB \arrow[d, "Gb"] \\ FA' \arrow[r, "f'"'] & GB' \end{tikzcd}$$

- 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.

$$\begin{tikzcd} A \arrow[r, "u"] \arrow[dr, "f"'] & GB \arrow[d, dashed, "Gg"] \\ & GB' \end{tikzcd}$$

- 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.

$$\begin{tikzcd} A & GB \arrow[l, "u"'] \\ & GB' \arrow[u, dashed, "Gg"'] \arrow[ul, "f"] \end{tikzcd}$$

# 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.

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "f'"'] & GB \arrow[d, dashed, "Gg"] \\ & GB' \end{tikzcd}$$

- 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.

$$\begin{tikzcd} A & GB \arrow[l, "f"'] \\ & GB' \arrow[u, dashed, "Gg"'] \arrow[ul, "f'"] \end{tikzcd}$$

- 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 $G \downarrow A$.

# 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.

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "1_A"'] & GB \arrow[d, "Gg"] \\
A \arrow[r, "f'"'] & GB'
\end{tikzcd}$$

- 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 $\mathbf{D}$.

$$\begin{tikzcd} X & X \times Y \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] & Y \\ & Z \arrow[ul, "f"] \arrow[u, dashed, "h"] \arrow[ur, "g"'] & \end{tikzcd} \qquad \begin{tikzcd} (X,Y) & \Delta(X \times Y) \arrow[l, "(\pi_1,\pi_2)"'] \\ & \Delta(Z) \arrow[ul, "(f,g)"] \arrow[u, dashed, "\Delta(h)"'] \end{tikzcd}$$

Take $\mathbf{C}$ to be the product category $\mathbf{D} \times \mathbf{D}$, and let $\Delta$ 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 $\Delta$ 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 $\Delta$ to $D$.
- Dually, the colimit $\varinjlim D$ is a universal morphism from $D$ to $\Delta$.

# Contents


Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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:

$$\begin{tikzcd} M \arrow[r, "!"] \arrow[d, tail, "m"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, tail, "\top"] \\ X \arrow[r, dashed, "\varphi"'] & \Omega \end{tikzcd}$$

**Remark:** The arrow $\top \circ !$: $M \xrightarrow{!} 1 \overset{\top}{\rightarrowtail} \Omega$ is often denoted as $\top_M : M \to \Omega$.

# 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 $\mathbf{Set}^{\rightarrow}$ 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.

# 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.

$$\begin{aligned} \top(A) &= \{\bullet\} \\ (F \times G)A &= FA \times GA \\ G^F A &= \mathrm{Nat}(\mathbf{C}(-, A) \times F, G) \qquad \text{(Yoneda)} \end{aligned}$$

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

$$\begin{tikzcd}
R \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \in_A \arrow[d, tail] \\
A \times B \arrow[r, dashed, "1_A \times \chi_R"'] & A \times \mathrm{P}(A)
\end{tikzcd}$$

## 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 $\varepsilon : A \times \Omega^A \to \Omega$.

$$\begin{tikzcd}
\in_A \arrow[r, "!"] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\top"] \\
A \times \Omega^A \arrow[r, "\varepsilon"'] & \Omega
\end{tikzcd}$$

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. $\varepsilon \circ (1_A \times \hat{f}) = f$.

$$\begin{tikzcd}
A \times B \arrow[r, "1_A \times \hat{f}"] \arrow[dr, "f"'] & A \times \Omega^A \arrow[d, "\varepsilon"] \\
& \Omega
\end{tikzcd}
\qquad
\begin{tikzcd}
R \arrow[r, dashed] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] \arrow[rr, bend left, "!"] & \in_A \arrow[r, "!"] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\top"] \\
A \times B \arrow[r, "1_A \times \hat{f}"'] \arrow[rr, bend right, "f"'] & A \times \Omega^A \arrow[r, "\varepsilon"'] & \Omega
\end{tikzcd}$$

# 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 $(f^*, f_*)$: $\mathbf{F} \underset{f_*}{\overset{f^*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{E}$ , 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.*

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, two heads, "e"'] & & Y \\ & A \arrow[ur, tail, "m"'] & \end{tikzcd}$$

## 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 $f$.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "f"'] & Y \arrow[d, "q"] \\ Y \arrow[r, "p"'] \arrow[ur, phantom, "\ulcorner"] & P \end{tikzcd} \qquad \begin{tikzcd} X \arrow[d, dashed, "e"'] \arrow[dr, "f"] & & \\ A \arrow[r, tail, "m"] & Y \arrow[r, shift left, "p"] \arrow[r, shift right, "q"'] & P \end{tikzcd}$$

# Epi-Mono Factorization

The epi-mono factorization is unique up to a unique isomorphism.

$$\begin{tikzcd}
X \arrow[rr, "f"] \arrow[dr, two heads, "e"] \arrow[ddr, bend right, two heads, "e'"'] & & Y \\
& A \arrow[ur, tail, "m"'] \arrow[d, dashed, "u"] & \\
& A' \arrow[uur, bend right, tail, "m'"'] &
\end{tikzcd}$$

Given an epi-mono factorization, we call $A$ the image of $f$, and denote it as $\operatorname{Im} f$.

# 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 $y$.

$$\begin{tikzcd}
 & & \overset{B}{\boxed{\begin{array}{l}\text{a pair } (w,m) \\ \text{where } w \text{ is a} \\ \text{woman and } m \\ \text{is a man}\end{array}}} \arrow[d, "w"] \\
\overset{A}{\boxed{\text{a child}}} \arrow[r, "\text{has as parents}"] \arrow[d, "e"'] \arrow[dr, "\text{has as mother } f"] \arrow[dr, phantom, shift right=3ex, "\circlearrowleft"] \arrow[dr, phantom, shift left=3ex, "\circlearrowleft"] & & \\
\underset{\operatorname{Im} f}{\boxed{\text{a mother}}} \arrow[r, hook, "\text{is}"] & & \overset{C}{\boxed{\text{a woman}}}
\end{tikzcd}$$

# $\subset$ and $\in$

- Given subobjects $m$ and $m'$, define $m \subset_X m'$ (or $M \subset_X M'$) as follows

$$m \subset_X m' := \exists f : m \to m' \in \mathbf{C}/X$$

$$\begin{tikzcd} M \arrow[rr, "f"] \arrow[dr, tail, "m"'] & & M' \arrow[dl, tail, "m'"] \\ & X & \end{tikzcd}$$

- $m \sim m' := m \subset_X m' \;\&\; m' \subset_X m$
- $\mathrm{Sub}(X) := \{[m] : m \text{ is monic with } \mathrm{cod}(m) = X\}$
- $[m] \subset_X [m'] := m \subset_X m'$
- In terms of 'generalized elements'[23] of an object $X$, $z : Z \to X$, one can define a *local membership* relation,

$$z \in_X M := \exists f : Z \to M[z = mf]$$

$$\begin{tikzcd} Z \arrow[r, "f"] \arrow[dr, "z"'] & M \arrow[d, "m"] \\ & X \end{tikzcd}$$

[23] 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:

$$\begin{tikzcd}
f^*M \arrow[r] \arrow[d, tail, "f^*m"'] \arrow[dr, phantom, "\lrcorner", very near start] & M \arrow[d, tail, "m"] \\
X \arrow[r, "f"'] & Y
\end{tikzcd}$$

# 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 $\mathbf{C}^{\mathrm{op}}$,

$$\begin{tikzcd} FA \arrow[r, "Ff"] \arrow[d, hook] \arrow[dr, phantom, "\circlearrowleft"] & FB \arrow[d, hook] \\ GA \arrow[r, "Gf"'] & GB \end{tikzcd}$$

# Predicate

$$\begin{tikzcd}
Z \arrow[dr, dashed] \arrow[drr, bend left] \arrow[ddr, bend right, "x"'] & & \\
& M \arrow[r] \arrow[d, mapsto] & 1 \arrow[d, mapsto, "\top"] \\
& X \arrow[r, dashed, "\varphi"'] & \Omega
\end{tikzcd}$$

$$\varphi x = \top \iff x \in_X M$$

**Remark:** A predicate $\varphi$ is the character of subobject $M$ of universe $X$.

# $\bot$ and $\neg$

$$\begin{tikzcd}
Z \arrow[dr, dashed] \arrow[drr, bend left] \arrow[ddr, bend right, "x"'] & & \\
& 1 \arrow[r] \arrow[d, mapsto, "\bot"] & 1 \arrow[d, mapsto, "\top"] \\
& \Omega \arrow[r, dashed, "\neg"'] & \Omega
\end{tikzcd}$$

where $\bot$ is defined by the pullback,

$$\begin{tikzcd}
0 \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & 1 \arrow[d, mapsto, "\top"] \\
1 \arrow[r, dashed, "\bot"'] & \Omega
\end{tikzcd}$$

$$\neg x = \top \iff x = \bot$$

## Remark

- $\bot : 1 \to \Omega$ *is the character of the unique arrow* $0 \to 1$.
- $\neg : \Omega \to \Omega$ *is the character of* $\bot : 1 \to \Omega$.

# $\wedge$

$$\begin{tikzcd}
Z \arrow[dr, dashed] \arrow[drr, bend left] \arrow[ddr, bend right, "\langle x,y\rangle"'] & & \\
& 1 \arrow[r] \arrow[d, mapsto, "\langle\top,\top\rangle"] & 1 \arrow[d, mapsto, "\top"] \\
& \Omega\times\Omega \arrow[r, dashed, "\wedge"'] & \Omega
\end{tikzcd}$$

$$x \wedge y = \top \iff \langle x, y\rangle = \langle \top, \top\rangle$$

## Remark

- $\wedge : \Omega \times \Omega \to \Omega$ *is the character of* $\langle \top, \top\rangle : 1 \to \Omega \times \Omega$.

# $\vee$

$$\begin{tikzcd}
Z \arrow[dr, dashed] \arrow[drr, bend left] \arrow[ddr, bend right, "\langle x,y\rangle"'] & & \\
& U \arrow[r] \arrow[d, tail] & 1 \arrow[d, tail, "\top"] \\
& \Omega \times \Omega \arrow[r, dashed, "\vee"'] & \Omega
\end{tikzcd}$$

where

$$\begin{tikzcd}
\Omega + \Omega \arrow[rr, "\begin{bmatrix}\langle \top, 1_\Omega\rangle \\ \langle 1_\Omega, \top\rangle\end{bmatrix}"] \arrow[dr, two heads] & & \Omega \times \Omega \\
& U \arrow[ur, tail] &
\end{tikzcd}$$

$$x \vee y = \top \iff \langle x, y\rangle \in_{\Omega\times\Omega} U$$

## Remark

- $\vee : \Omega \times \Omega \to \Omega$ *is the character of the image (the mono part of the epi-mono factorization) of the arrow* $\begin{bmatrix}\langle \top, 1_\Omega\rangle \\ \langle 1_\Omega, \top\rangle\end{bmatrix} : \Omega + \Omega \to \Omega \times \Omega$.

# $\rightarrow$

$$\begin{tikzcd}
Z \arrow[dr, dashed] \arrow[drr, bend left] \arrow[ddr, bend right, "\langle x,y \rangle"'] & & \\
& \leq_\Omega \arrow[r] \arrow[d, "e"] & 1 \arrow[d, "\top"] \\
& \Omega \times \Omega \arrow[r, dashed, "\rightarrow"'] & \Omega
\end{tikzcd}$$

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* $\wedge, \pi_1 : \Omega \times \Omega \rightarrow \Omega$.

# $\forall$

$$\begin{tikzcd}
Y \arrow[dr, dashed] \arrow[drr, bend left] \arrow[ddr, bend right, "\hat{f}"'] & & \\
& 1 \arrow[r] \arrow[d, mapsto, "\widehat{\top_X}"] & 1 \arrow[d, mapsto, "\top"] \\
& \Omega^X \arrow[r, dashed, "\forall_X"'] & \Omega
\end{tikzcd}$$

where $\hat{f} : Y \to \Omega^X$ is the currying of $Y \times X \xrightarrow{f} \Omega$, and $\widehat{\top_X}$ is the currying of the composite $1 \times X \xrightarrow{!} 1 \xrightarrow{\top} \Omega$.

$$\text{Obviously,} \quad \forall_X \circ \hat{f} = \top_Y \iff \varepsilon \circ (\hat{f} \times 1_X) = \top_{Y \times X}$$

## Theorem

*Let $f : Y \times X \to \Omega$ and $g : Y \to \Omega$ be morphisms in $\mathbf{E}$. Then*

$$\frac{g \circ \pi_1 \subset_{Y \times X} f}{g \subset_Y \forall_X \circ \hat{f}}$$

# Kripke-Joyal Semantics

$$\begin{tikzcd}
U \arrow[dr, dashed, "u"] \arrow[drr, bend left, "!"] \arrow[ddr, bend right, "a"'] & & \\
& \{x : \varphi(x)\} \arrow[r] \arrow[d, mapsto, "m"] & 1 \arrow[d, mapsto, "\top"] \\
& A \arrow[r, dashed, "\varphi"'] & \Omega
\end{tikzcd}$$

## 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* $a : U \to A$ *a generalized element of* $A \in \mathbf{E}$*, and* $\varphi(x), \psi(x)$ *wff with a free variable* $x$ *of sort* $A$*. 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* $B$*, then*

- $U \Vdash \exists y \varphi(a, y)$ *iff there exists an epic* $f : V \twoheadrightarrow U$ *and a generalized element* $b : V \to B$ *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* $b : V \to B$*,* $V \Vdash \varphi(a \circ f, b)$*.*

## 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 $\Omega$.*

$$\mathrm{Sub}_{\mathbf{C}}(X) \cong \mathrm{Hom}_{\mathbf{C}}(X, \Omega)$$

$$\mathrm{Sub}_{\mathbf{Set}}(X) \cong \mathrm{P}(X)$$

## 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.

# $\mathrm{Sub}(X)$

## 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 $A, B \in \mathbf{C}$,

$$
\begin{aligned}
(F \wedge G)(A) &\coloneqq FA \cap GA \\
(F \vee G)(A) &\coloneqq FA \cup GA \\
1 &\coloneqq X \\
0 &\coloneqq \Delta_{\emptyset} \\
(\neg F)(A) &\coloneqq \Big\{ a \in X(A) : \forall f : B \to A \Big( Ffa \notin FB \Big) \Big\} \\
(F \to G)(A) &\coloneqq \Big\{ a \in X(A) : \forall f : B \to A \Big( Ffa \in FB \implies Ffa \in GB \Big) \Big\}
\end{aligned}
$$

**Remark:** $\neg F \vee F$ may not be the same as $X$.

# 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 $\mathrm{Hom}$ and $\mathrm{Sub}$ 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 $\Omega^X$ 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 |
|---|---|
| | $A \vee \neg A = \top$ |
| $A \to \neg\neg A$ | $\neg\neg A \to A$ |
| $\neg A \wedge \neg B \leftrightarrow \neg(A \vee B)$ | |
| $\neg A \vee \neg B \to \neg(A \wedge B)$ | $\neg(A \vee B) \to \neg A \wedge \neg B$ |
| $\forall x \neg A \leftrightarrow \neg\exists x A$ | |
| $\exists x A \to \neg\forall x \neg A$ | $\neg\forall x \neg A \to \exists x A$ |
| $\forall x A \to \neg\exists x \neg A$ | $\neg\exists x \neg A \to \forall x A$ |

$$
\begin{tikzcd}
 & \Omega \arrow[loop above, "\neg"] & \\
1 \arrow[r, shift left, "\top"] \arrow[r, shift right, "\bot"'] & \Omega & \Omega \times \Omega \arrow[l, shift right, "\wedge"'] \arrow[l, "\vee"] \arrow[l, shift left, "\leftarrow"]
\end{tikzcd}
$$

| 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

# 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.

$$\begin{tikzcd} & \text{Theory} \arrow[dl] \arrow[d] \arrow[dr] & \\ \text{Model1} & \text{Model2} & \text{Model3} \end{tikzcd} \qquad \begin{tikzcd} \text{Logic} \arrow[r, bend left] & \text{Geometry} \arrow[l, bend left] \end{tikzcd}$$

- 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# Elementary Theory of the Category of Sets $\mathrm{ETCS}$

## Definition (Well-pointed Topos)

A topos is called *well-pointed* iff

1. extensionality: for $A \underset{g}{\overset{f}{\rightrightarrows}} B$ , if $\forall x : 1 \to A[fx = gx]$ then $f = g$.
2. non-triviality: $0 \not\cong 1$.

**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 ($\mathrm{ETCS}$ — Lawvere)

$\mathrm{ETCS}$ 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 $\mathrm{ETCS}$: 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

$$\begin{tikzcd} S_x \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "x"] \\ S \arrow[r, "f"'] & X \end{tikzcd}$$

## Theorem

$\mathrm{ETCS} + \mathrm{Replacement}$ *is bi-interpretable with* $\mathrm{ZFC}$.

# 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

*"Set theory should not be based on membership, as in* $\mathrm{ZFC}$*, 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 $\mathrm{ZFC}$: $2 \in 3$, $5 \subset 12$, $1 = \mathrm{P}(0)$
- Junk Theorems in $\mathrm{ETCS}$: $\mathrm{dom}(5) = \mathrm{dom}(7)$

# The Category $\mathbf{Set}^{\circlearrowleft}$ of Endomaps of Sets

- An object of $\mathbf{Set}^{\circlearrowleft}$ is any set $A \in \mathbf{Set}$ equipped with an endomap.
- A morphism $f : A^{\circlearrowleft s} \to B^{\circlearrowleft t}$ is a map satisfying $f \circ s = t \circ f$.

$$\begin{tikzcd} A \arrow[r, "s"] \arrow[d, "f"'] & A \arrow[d, "f"] \\ B \arrow[r, "t"'] & B \end{tikzcd}$$

- Some questions can be asked about a particular object $A^{\circlearrowleft s}$ of $\mathbf{Set}^{\circlearrowleft}$.
  - accessibility: Given $a \in A^{\circlearrowleft s}$, does $\exists b : s(b) = a$?
  - convergence to equilibrium: Given $a \in A^{\circlearrowleft s}$, does $\exists n : s^{n+1}(a) = s^n(a)$?
- Example: let $C_n$ be

$$\begin{tikzcd} 0 \arrow[r] & 1 \arrow[r] & 2 \arrow[r] & \cdots \arrow[r] & n-1 \arrow[lllll, bend right] \end{tikzcd}$$

then $C_n \to B^{\circlearrowleft t}$ corresponds to elements in $B^{\circlearrowleft t}$ having period $n$.

# Dynamical Systems

- A discrete dynamical system is a set $X$ with an 'evolution function' $X \circlearrowleft f$, 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 $\mathbf{Set} \circlearrowleft$ 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 $f : \mathbb{R}^{\geq 0} \to \mathbf{Top}$.

# 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:

$$\begin{tikzcd}
1 \arrow[r, "0"] \arrow[dr, "x"'] & N \arrow[r, "s"] \arrow[d, dashed, "f"] & N \arrow[d, dashed, "f"] \\
& X \arrow[r, "t"'] & X
\end{tikzcd}$$

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?

# Recursion Lemma

$$\begin{tikzcd} 1 \arrow[r, "0"] \arrow[dr, "x"'] & N \arrow[r, "s"] \arrow[d, dashed, "f"] & N \arrow[d, dashed, "f"] \\ & X \arrow[r, "t"'] & X \end{tikzcd} \qquad \frac{1 \xrightarrow{x} X \text{ in } \mathbf{Set}}{N \circlearrowleft_s \xrightarrow{f} X \circlearrowleft_t \text{ in } \mathbf{Set}^{\circlearrowleft}} \qquad \begin{tikzcd} & & N \\ N \arrow[urr, bend left, "k"] \arrow[r, dashed, "f"] \arrow[drr, bend right, "g"'] & N \times A \arrow[ur, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ & & A \end{tikzcd}$$

## Theorem (Recursion Lemma)

*Given $h : N \times A \to A$ and $1 \xrightarrow{a} A$, there exists a unique $N \xrightarrow{g} A$ s.t.*

$$g(0) = a$$

$$g(n+1) = h(n, g(n))$$

## Proof.

Let $X := N \times A$, $t(n, y) := \langle n+1, h(n, y) \rangle$, $x := \langle 0, a \rangle$. Then $\exists! f : N \to X$:

$$f(0) = x = \langle 0, a \rangle$$

$$f(n+1) = t \circ f(n)$$

Let $k := \pi_1 \circ f$ and $g := \pi_2 \circ f$. Then $f(0) = \langle k(0), g(0) \rangle = \langle 0, a \rangle$ and $f(n+1) = \langle k(n+1), g(n+1) \rangle = t \circ \langle k(n), g(n) \rangle = \langle k(n)+1, h(k(n), g(n)) \rangle$. 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. $f \circ s = t \circ f$

$$\begin{tikzcd}
1 \arrow[r, "a"] \arrow[dr, "b"'] & A \arrow[r, "s"] \arrow[d, dashed, "f"] & A \arrow[d, dashed, "f"] \\
& B \arrow[r, "t"'] & B
\end{tikzcd}$$

The *Natural Numbers Object* NNO is an initial object of the Peano category.

# NNO — Example

- Addition

$$\begin{tikzcd} 1 \arrow[r, "0"] \arrow[dr, "m"'] & N \arrow[r, "s"] \arrow[d, dashed, "f_m"] & N \arrow[d, dashed, "f_m"] \\ & N \arrow[r, "s"'] & N \end{tikzcd}$$

$$\begin{tikzcd} 1 \arrow[r, "0"] \arrow[dr, "\widehat{1_N}"'] & N \arrow[r, "s"] \arrow[d, dashed, "\hat{+}"] & N \arrow[d, dashed, "\hat{+}"] \\ & N^N \arrow[r, "s^N"'] & N^N \end{tikzcd}$$

$$f_m(0) = m$$

$$f_m(s(n)) = s(f_m(n))$$

$$m + 0 = m$$

$$m + s(n) = s(m + n)$$

- Multiplication

$$\begin{tikzcd} 1 \arrow[r, "0"] \arrow[dr, "0"'] & N \arrow[r, "s"] \arrow[d, dashed, "f_m"] & N \arrow[d, dashed, "f_m"] \\ & N \arrow[r, "m+(-)"'] & N \end{tikzcd}$$

$$\begin{tikzcd} 1 \arrow[r, "0"] \arrow[dr, "\widehat{0}"'] & N \arrow[r, "s"] \arrow[d, dashed, "\hat{\times}"] & N \arrow[d, dashed, "\hat{\times}"] \\ & N^N \arrow[r, "(m+(-))^N"'] & N^N \end{tikzcd}$$

$$f_m(0) = 0$$

$$f_m(s(n)) = m + f_m(n)$$

$$m \times 0 = 0$$

$$m \times s(n) = m + m \times n$$

# NNO — Example

- Iteration

$$
\begin{tikzcd}
1 \arrow[r, "0"] \arrow[dr, "\widehat{1_X}"'] & N \arrow[r, "s"] \arrow[d, dashed, "\hat{f}"] & N \arrow[d, dashed, "\hat{f}"] \\
& X^X \arrow[r, "t^X"'] & X^X
\end{tikzcd}
\quad \Longrightarrow \quad
\begin{tikzcd}
1 \times X \arrow[r, "0\times 1_X"] \arrow[d, "\cong"'] & N \times X \arrow[r, "s\times 1_X"] \arrow[d, dashed, "f"] & N \times X \arrow[d, dashed, "f"] \\
X \arrow[r, "1_X"'] & X \arrow[r, "t"'] & X
\end{tikzcd}
$$

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

$$\begin{tikzcd}
X \arrow[r, "\langle 0!, 1_X \rangle"] \arrow[dr, "g"'] & N \times X \arrow[r, "s \times 1_X"] \arrow[d, dashed, "f"] & N \times X \arrow[d, dashed, "f"] \\
& Y \arrow[r, "h"'] & Y
\end{tikzcd}$$

$$f(0, x) = g(x)$$
$$f(sn, x) = h(f(n, x))$$

# 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 $M \overset{m}{\rightarrowtail} N$ 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 $M \overset{m}{\cong} N$.

$$\begin{tikzcd}
1 \arrow[r, "z"] \arrow[dr, "0"'] & M \arrow[r, "r"] \arrow[d, tail, "m"] & M \arrow[d, tail, "m"] \\
& N \arrow[r, "s"'] & N
\end{tikzcd}$$

# 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.

$$\begin{tikzcd}
1 \arrow[dr, "\iota_1"'] \arrow[drr, bend left, "0"] & & \\
& 1 + N \arrow[r, dashed, "\begin{bmatrix} 0 \\ s \end{bmatrix}"] & N \\
N \arrow[ur, "\iota_2"] \arrow[urr, bend right, "s"'] & &
\end{tikzcd}$$

2. $N \underset{1_N}{\overset{s}{\rightrightarrows}} N \xrightarrow{!_N} 1$ is a coequalizer.

# Skeleton

## Definition (Skeleton)

Given a category $\mathbf{C}$, a *skeleton* of $\mathbf{C}$ is a full subcategory containing exactly one objects from each isomorphism class of objects of $\mathbf{C}$.

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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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.

| | |
|---|---|
| $\boxed{\text{arithmetic in } \mathbb{Z}}$ = truth | arithmetic = $\boxed{\text{true in } \mathbb{Z}}$ |
| truth is absolute | truth is contextual |

**Remark:** Math isn't about absolute truth; it's about different contexts in which different things can be true.

# Context: What is a Circle?

- What is a circle? — all the points the same distance from a center?
- What is a circle of radius $4$? $L^2$ norm? $L^1$ norm?

*"We are often interested not just in whether or not something is true, but in where it is true."*

*— John Baez*

# 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 $f_* := \mathrm{Hom}(A, f) : \mathrm{Hom}(A, X) \to \mathrm{Hom}(A, Y) :: g \mapsto fg$.

$$\begin{tikzcd}
X \arrow[r, "f"] & Y \\
\mathrm{Hom}(A, X) \arrow[r, "\mathrm{Hom}(A,f)"] & \mathrm{Hom}(A, Y) \\
g \arrow[r, mapsto, "f \circ -"] & fg
\end{tikzcd}
\qquad
\begin{tikzcd}
 & A \arrow[dl, "g"'] \arrow[dr, "f \circ g"] & \\
X \arrow[rr, "f"'] & & Y
\end{tikzcd}$$

- 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 $f^* := \mathrm{Hom}(f, A) : \mathrm{Hom}(X, A) \to \mathrm{Hom}(Y, A) :: g \mapsto gf$.

$$\begin{tikzcd}
Y \arrow[r, "f"] & X \\
\mathrm{Hom}(X, A) \arrow[r, "\mathrm{Hom}(f,A)"] & \mathrm{Hom}(Y, A) \\
g \arrow[r, mapsto, "- \circ f"] & gf
\end{tikzcd}
\qquad
\begin{tikzcd}
 & A & \\
X \arrow[ur, "g"] & & Y \arrow[ul, "g \circ f"'] \arrow[ll, "f"]
\end{tikzcd}$$

# Hom Functor

$$\begin{tikzcd}
A \arrow[r, "g \in h^A X"] \arrow[d, "h^A f"] \arrow[dr, "(h^A f)g = fg \in h^A Y"'] & X \arrow[d, "f \in \mathrm{Hom}(X,Y)"] \\
 & Y
\end{tikzcd}$$

$$\begin{tikzcd}
A & X \arrow[l, "g \in h_A X"'] \\
 & Y \arrow[u, "f \in \mathrm{Hom}(X,Y)"'] \arrow[ul, "(h_A f)g = gf \in h_A Y"] \arrow[u, "h_A f"]
\end{tikzcd}$$

# Natural Transformation between $\mathrm{Hom}$ Functors

$$\alpha \in \mathrm{Nat}\left(h^X, h^Y\right) \text{ for } h^X, h^Y \in \mathbf{Set}^{[\mathbf{C},\mathbf{Set}]}$$

$$\begin{tikzcd}
h^X A \arrow[rr, "h^X f = f \circ"] \arrow[dd, "\alpha_A"'] & & h^X B \arrow[dd, "\alpha_B"] \\
& A \arrow[ul, "h^X"'] \arrow[dl, "h^Y"] \arrow[r, "f"] & B \arrow[ur, "h^X"] \arrow[dr, "h^Y"'] & \\
h^Y A \arrow[rr, "h^Y f = f \circ"'] & & h^Y B
\end{tikzcd}$$

$$f \circ \alpha_A = \alpha_B \circ f$$

$\mathrm{Hom}(-,-) : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Set}$ maps $A, B \in \mathbf{C}$ to $\mathrm{Hom}(A,B)$, and maps $f : A' \to A, g : B \to B'$ to
$\mathrm{Hom}(f,g) : \mathrm{Hom}(A,B) \to \mathrm{Hom}(A',B') :: h \mapsto ghf$.

$$
\begin{tikzcd}
 & h \arrow[rr, mapsto, "-\circ f"] \arrow[dd, mapsto, "g\circ -"'] & & hf \arrow[dd, mapsto, "g\circ -"] \\
B \arrow[d, "g"'] & & & \\
B' & & & \\
 & gh \arrow[rr, mapsto, "-\circ f"'] & & ghf
\end{tikzcd}
$$

$$
\begin{tikzcd}
\mathrm{Hom}(A,B) \arrow[r, "\mathrm{Hom}(f,B)"] \arrow[d, "\mathrm{Hom}(A,g)"'] & \mathrm{Hom}(A',B) \arrow[d, "\mathrm{Hom}(A',g)"] \\
\mathrm{Hom}(A,B') \arrow[r, "\mathrm{Hom}(f,B')"'] & \mathrm{Hom}(A',B')
\end{tikzcd}
$$

$$
\begin{tikzcd}
A' \arrow[r, "f"] & A
\end{tikzcd}
$$

# 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 $\psi(a)_B(f) \coloneqq Ff(a)$.
By construction, $\varphi(\psi(a)) = a$. It remains to verify that $\psi(\varphi(\alpha)) = \alpha$.
$\psi(\varphi(\alpha))_B(f) = \psi(\alpha_A(1_A))_B(f) = Ff(\alpha_A(1_A)) = \alpha_B(f)$

$$\begin{tikzcd}
\mathbf{C}(A,A) \arrow[r, "\mathbf{C}(A,f)"] \arrow[d, "\psi(a)_A"'] & \mathbf{C}(A,B) \arrow[d, "\psi(a)_B"] \\
FA \arrow[r, "Ff"'] & FB
\end{tikzcd}$$

$$\begin{tikzcd}
1_A \arrow[r, mapsto] \arrow[d, mapsto] & f \arrow[d, mapsto] \\
a \arrow[r, mapsto] & \psi(a)_B(f) = Ff(a)
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbf{C}(A,A) \arrow[r, "\mathbf{C}(A,f)"] \arrow[d, "\alpha_A"'] & \mathbf{C}(A,B) \arrow[d, "\alpha_B"] \\
FA \arrow[r, "Ff"'] & FB
\end{tikzcd}$$

# Proof Sketch of Yoneda Lemma

## Proof continued.

To show that $\varphi$ is natural in $F$, suppose given a natural transformation $\beta : F \to G$.

$$\begin{tikzcd}
\mathbf{Set}^{\mathbf{C}}(\mathbf{C}(A,-),F) \arrow[r, "\beta\circ -"] \arrow[d, "\varphi_F"'] & \mathbf{Set}^{\mathbf{C}}(\mathbf{C}(A,-),G) \arrow[d, "\varphi_G"] \\
FA \arrow[r, "\beta_A"'] & GA
\end{tikzcd}$$

$$\beta_A(\varphi_F(\alpha)) = \beta_A(\alpha_A(1_A)) = (\beta \circ \alpha)_A(1_A) = \varphi_G(\beta \circ \alpha)$$

To show that $\varphi$ is natural in $A$, suppose given $f : A \to B$.

$$\begin{tikzcd}
\mathbf{Set}^{\mathbf{C}}(\mathbf{C}(A,-),F) \arrow[r, "-\circ\mathbf{C}(f,-)"] \arrow[d, "\varphi_A"'] & \mathbf{Set}^{\mathbf{C}}(\mathbf{C}(B,-),F) \arrow[d, "\varphi_B"] \\
FA \arrow[r, "Ff"'] & FB
\end{tikzcd}$$

$Ff(\varphi_A(\alpha)) = Ff(\alpha_A(1_A)) = \alpha_B \circ \mathbf{C}(A,f)(1_A) = \alpha_B(f) =$
$\alpha_B \circ \mathbf{C}(f,-)_B(1_B) = (\alpha \circ \mathbf{C}(f,-))_B(1_B) = \varphi_B(\alpha \circ \mathbf{C}(f,-))$

$$
\begin{tikzcd}
h^X A \arrow[rr, "h^X f = f\circ"] \arrow[dd, "\alpha_A"'] & & h^X B \arrow[dd, "\alpha_B"] \\
& A \arrow[ul, "h^X"'] \arrow[r, "f"] \arrow[dl, "F"] & B \arrow[ur, "h^X"] \arrow[dr, "F"'] & \\
FA \arrow[rr, "Ff = f\circ"'] & & FB
\end{tikzcd}
$$

$$
\begin{tikzcd}
\mathrm{Hom}(A,A) \arrow[rr, "\mathrm{Hom}(A,f)"] \arrow[dd, "\psi(a)_A"'] \arrow[dr, phantom, "\ni"] & & \mathrm{Hom}(A,B) \arrow[dd, "\psi(a)_B"] \arrow[dl, phantom, "\in"] \\
& 1_A \arrow[r, mapsto] \arrow[d, mapsto] & f \arrow[d, mapsto] & \\
& a \arrow[r, mapsto] & \psi(a)_B(f) = Ff(a) & \\
FA \arrow[rr, "Ff"'] \arrow[ur, phantom, "\ni"] & & FB \arrow[ul, phantom, "\in"]
\end{tikzcd}
$$

$$\begin{tikzcd}
\mathbf{C}^{\mathrm{op}} \times \mathbf{Set}^{\mathbf{C}} \arrow[r, bend left, "\mathrm{Hom}(\mathsf{y}(-),\bullet)"] \arrow[r, bend right, "\bullet(-)"'] \arrow[r, phantom, "\Downarrow\varphi"] & \mathbf{Set}
\end{tikzcd}$$

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

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".

$$\text{for example} \quad \begin{bmatrix} a & b & c \\ a' & b' & c' \end{bmatrix} \mapsto \begin{bmatrix} a + 2a' & b + 2b' & c + 2c' \\ a' & b' & c' \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 + 2 \cdot 0 & 0 + 2 \cdot 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$\text{we have} \quad \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ a' & b' & c' \end{bmatrix} \mapsto \begin{bmatrix} a + 2a' & b + 2b' & c + 2c' \\ a' & b' & c' \end{bmatrix}$$

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 $\alpha_m$ 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 $\lambda$ 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,* $\mathsf{y}$ *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 $f \coloneqq \alpha_A(1_A)$.

$$\begin{tikzcd}
 & 1_A \arrow[rr, mapsto] \arrow[dd, mapsto] & & \alpha_A(1_A) \arrow[dd, mapsto] \\
X \arrow[d, "g"'] & & \mathbf{C}(A, A) \arrow[r, "\alpha_A"] \arrow[d, "g^*"'] & \mathbf{C}(A, B) \arrow[d, "g^*"] \\
A & & \mathbf{C}(X, A) \arrow[r, "\alpha_X"'] & \mathbf{C}(X, B) \\
 & 1_A \circ g \arrow[rr, mapsto] & & \alpha_X(1_A \circ g) = \alpha_A(1_A) \circ g \\
 & & & \alpha_X(g) = f \circ g
\end{tikzcd}$$

# 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 $A \cong B$.

# 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:

$$\begin{tikzcd}
1_A \arrow[rrr, mapsto] \arrow[ddd, mapsto] & & & g \arrow[ddd, mapsto] \\
& \mathrm{Hom}(A, A) \arrow[r, "\cong"] \arrow[d, "f\circ -"'] & \mathrm{Hom}(B, A) \arrow[d, "f\circ -"] & \\
& \mathrm{Hom}(A, B) \arrow[r, "\cong"] & \mathrm{Hom}(B, B) & \\
f \arrow[rrr, mapsto] & & & 1_B = f \circ g
\end{tikzcd}$$

similarly,
$g \circ f = 1_A$

# 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, $\mathbf{C} = \mathbf{FVect}$ 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* $\mathsf{y}$ *preserves all limits that exist in* $\mathbf{C}$*.*

## Proof.

Suppose $(L, \lambda)$ is a limit of $D : \mathbf{I} \to \mathbf{C}$. The Yoneda embedding maps $D$ to the diagram $\mathsf{y}D : \mathbf{I} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$ defined by $(\mathsf{y}D)_i = \mathsf{y}D_i = \mathrm{Hom}(-, D_i)$, and it maps $(L, \lambda)$ to $(\mathsf{y}L, \mathsf{y}\lambda)$ on $\mathsf{y}D$ defined by $(\mathsf{y}\lambda)_i = \mathsf{y}\lambda_i = \mathrm{Hom}(-, \lambda_i)$.
To see that $(\mathsf{y}L, \mathsf{y}\lambda)$ is a limit cone on $\mathsf{y}D$, consider a cone $(M, \mu)$ on $\mathsf{y}D$. 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 $\varphi_A : MA \to \mathrm{Hom}(A, L) = (\mathsf{y}L)A$ forms a unique factorization $\varphi : M \to \mathsf{y}L$.

- $\mathrm{Hom}(A, \varprojlim D_i) \cong \varprojlim \mathrm{Hom}(A, D_i)$
- $\mathrm{Hom}(\varinjlim D_i, A) \cong \varprojlim \mathrm{Hom}(D_i, A)$

# 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 $\mathbf{G}(-,g) : x \mapsto gx$.
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[24] 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 $\mathrm{Aut}(X)$.

$$\mathrm{Aut}(X) \to \mathrm{End}(X) \to \mathrm{Hom}(X, Y)$$

$$\frac{\text{Space}}{\text{Transformation group}} \sim \frac{\text{Category}}{\text{Algebra of mappings}}$$

[24] 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 $(A, \alpha)$ is a *representation* of $F$.

$$\begin{tikzcd}
h_X A \arrow[rr, "h_X f"] \arrow[dd, "\alpha_A"'] & & h_X B \arrow[dd, "\alpha_B"] \\
& A \arrow[ul, "h_X"'] \arrow[dl, "F"] \arrow[r, "f"] & B \arrow[ur, "h_X"] \arrow[dr, "F"'] & \\
F A \arrow[rr, "F f"'] & & F B
\end{tikzcd}$$

# 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 $Ff(a) = b$.

$$\begin{tikzcd}
\mathrm{End}(A) \arrow[rr, "h_A f"] \arrow[dd, "\alpha_A"'] & & h_A B \arrow[dd, "\alpha_B"] \\
& A \arrow[ul, "h_A"'] \arrow[dl, "F"] \arrow[r, "f"] & B \arrow[ur, "h_A"] \arrow[dr, "F"] & \\
FA \arrow[rr, "Ff"'] & & FB
\end{tikzcd}$$

- If $(A, \alpha)$ 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

$\boxed{\text{Find } x \text{ s.t. } x^2 + 1 = 0.}$

## 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*

# Category of Elements

## Definition

- 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 $F$.

## Theorem

*A covariant set-valued functor is representable iff its category of elements has an initial object. Dually, 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 $A$.

$$\begin{tikzcd}
A & PA \arrow[r, "\alpha_A"] \arrow[d, "Pf"'] & \mathbf{D}(FA, B) \arrow[d, "\mathbf{D}(Ff, B)"] \\
A' \arrow[u, "f"'] & PA' \arrow[r, "\alpha_{A'}"'] & \mathbf{D}(FA', B)
\end{tikzcd}$$

Such an $\alpha$ can also be considered as $\{\alpha_A(a) : FA \to B\}_{(A,a) \in \int_{\mathbf{C}} P}$. Then

$$\begin{tikzcd}
A & FA = F\pi_P(A, a) \arrow[dr, "\alpha_A(a)"] & \\
 & & B \\
A' \arrow[uu, "f"] & FA' = F\pi_P(A', a') \arrow[uu, "Ff"] \arrow[ur, "\alpha_{A'}(a')"'] &
\end{tikzcd}$$

This means that $(B, \alpha_A)$ 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,

$$R_{\mathsf{y}}(B)(A) = \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}(\mathsf{y}A, B) \cong B(A)$$

this means that $R_{\mathsf{y}}$ 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.

$$P \cong \varinjlim \left( \int_{\mathbf{C}} P \xrightarrow{\pi_P} \mathbf{C} \xrightarrow{\mathsf{y}} \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}} \right)$$

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 $h : A \to B$ it suffices to give one $\theta : \mathrm{y}A \to \mathrm{y}B$ 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 $\mathrm{y}A \to \mathrm{y}B$ than $A \to B$?
- 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 $\mathsf{y}X = \mathrm{Hom}_{\mathbf{C}}(-, X)$
5. Opening the walls and displaying its new perspectives = finding a subcategory $\mathbf{A}$ of $\mathbf{C}$ such that: $\mathsf{y}_{|\mathbf{A}} : \mathbf{C} \to \mathbf{Set}^{\mathbf{A}^{\mathrm{op}}} :: X \mapsto \mathrm{Hom}_{\mathbf{A}}(-, X)$ 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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

$$\theta_{A,B} : \mathbf{D}(FA, B) \xrightarrow{\cong} \mathbf{C}(A, GB)$$

that is natural in both $A$ and $B$.

Naturality of $\theta$ means that for $f : A \to A'$ and $g : B \to B'$,

$$\begin{tikzcd}
\mathbf{D}(FA', B) \arrow[r, "\theta_{A',B}"] \arrow[d, "\mathbf{D}(Ff,B)"'] & \mathbf{C}(A', GB) \arrow[d, "\mathbf{C}(f,GB)"] \\
\mathbf{D}(FA, B) \arrow[r, "\theta_{A,B}"'] & \mathbf{C}(A, GB)
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{D}(FA, B) \arrow[r, "\theta_{A,B}"] \arrow[d, "\mathbf{D}(FA,g)"'] & \mathbf{C}(A, GB) \arrow[d, "\mathbf{C}(A,Gg)"] \\
\mathbf{D}(FA, B') \arrow[r, "\theta_{A,B'}"'] & \mathbf{C}(A, GB')
\end{tikzcd}$$

$$
\begin{tikzcd}
\mathbf{C}^{\mathrm{op}} \times \mathbf{D} \arrow[r, bend left, "\mathrm{Hom}_{\mathbf{D}}(F(-),\bullet)"] \arrow[r, bend right, "\mathrm{Hom}_{\mathbf{C}}(-,G(\bullet))"'] \arrow[r, phantom, "\Downarrow\theta"] & \mathbf{Set}
\end{tikzcd}
$$

For any $A, B \in \mathbf{C}$, and $f : A \to A'$, $g : B \to B'$, $h : FA \to B$,

$$
\begin{tikzcd}
\mathrm{Hom}_{\mathbf{D}}(FA, B) \arrow[rr, "\mathrm{Hom}_{\mathbf{D}}(Ff,g)"] \arrow[dd, "\theta_{AB}"'] \arrow[dr, phantom, "\ni"] & & \mathrm{Hom}_{\mathbf{D}}(FA', B') \arrow[dd, "\theta_{A'B'}"] \arrow[dl, phantom, "\in"] \\
& h \arrow[r, mapsto] \arrow[d, mapsto] & g \circ h \circ Ff \arrow[d, mapsto] \\
& \theta_{AB} h \arrow[r, mapsto] & \begin{array}{c} \theta_{A'B'}(g \circ h \circ Ff) \\ = \\ Gg \circ \theta_{AB} h \circ f \end{array} \\
\mathrm{Hom}_{\mathbf{C}}(A, GB) \arrow[rr, "\mathrm{Hom}_{\mathbf{C}}(f,Gg)"'] \arrow[ur, phantom, "\ni"] & & \mathrm{Hom}_{\mathbf{C}}(A', GB') \arrow[ul, phantom, "\in"]
\end{tikzcd}
$$

# 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* $f : A \to A'$, $g : B \to B'$, $h^\sharp : FA \to B$, $k^\sharp : FA' \to B'$,

$$\begin{tikzcd} A \arrow[r, "h^\flat"] \arrow[d, "f"'] & GB \arrow[d, "Gg"] \\ A' \arrow[r, "k^\flat"'] & GB' \end{tikzcd} \iff \begin{tikzcd} FA \arrow[r, "h^\sharp"] \arrow[d, "Ff"'] & B \arrow[d, "g"] \\ FA' \arrow[r, "k^\sharp"'] & B' \end{tikzcd}$$

## Proof.

$$\begin{tikzcd} \mathbf{D}(FA', B') \arrow[r, "\flat"] \arrow[d, "-\circ Ff"'] & \mathbf{C}(A', GB') \arrow[d, "-\circ f"] \\ \mathbf{D}(FA, B') \arrow[r, "\flat"'] & \mathbf{C}(A, GB') \end{tikzcd} \qquad \begin{tikzcd} \mathbf{D}(FA, B) \arrow[r, "\flat"] \arrow[d, "g\circ -"'] & \mathbf{C}(A, GB) \arrow[d, "Gg\circ -"] \\ \mathbf{D}(FA, B') \arrow[r, "\flat"'] & \mathbf{C}(A, GB') \end{tikzcd}$$

$$k^\flat \circ f = (k^\sharp \circ Ff)^\flat \qquad Gg \circ h^\flat = (g \circ h^\sharp)^\flat$$

# Left/Right Adjoint

## Example

$$\frac{A \wedge B \vdash C}{\overline{\overline{A \vdash B \to C}}}$$

$$F_B : X \mapsto X \wedge B \qquad G_B : X \mapsto B \to X \qquad F_B \dashv G_B$$

## Example

$$\frac{A \cap B \subset C}{\overline{\overline{A \subset B \to C}}} \qquad \text{where } B \to C := \overline{B} \cup C$$

$$F_B : X \mapsto X \cap B \qquad G_B : X \mapsto B \to X \qquad F_B \dashv G_B$$

## Example

$$\frac{A \setminus B \subset C}{\overline{\overline{A \subset B \cup C}}}$$

$$F_B : X \mapsto X \setminus B \qquad G_B : X \mapsto B \cup X \qquad F_B \dashv G_B$$

# Left/Right Adjoint

## Example $\mathbf{Cost} := ([0, \infty], \geq, +, 0)$

$$\frac{A + B \geq C}{\overline{A \geq B \to C}}$$

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 $z \in \mathbb{Z}, r \in \mathbb{R}$:

$$\frac{r \leq i(z)}{\overline{\lceil r \rceil \leq z}} \qquad \frac{i(z) \leq r}{\overline{z \leq \lfloor r \rfloor}} \qquad \lceil\rceil \dashv i \dashv \lfloor\rfloor$$

# 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 $B \to C$.

# Left/Right Adjoint

## 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 $\mathrm{Disc}$ 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 $\mathrm{Indisc}$ 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 $\mathrm{Disc}$ also has a left adjoint $\pi_0$ 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

$$\mathbf{1} := \begin{tikzcd} \overset{1}{\bullet} \arrow[loop above] \end{tikzcd}$$

## Example

# 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.

$$\frac{c(A) \subset B}{\overline{\overline{A \subset i(B)}}}$$

# 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 $\mathrm{P}(X)$.

$$\frac{i\,U \subset S}{\overline{\overline{U \subset S^\circ}}}$$

$$\begin{tikzcd} \mathcal{O}_X \arrow[r, shift left, "i"] \arrow[r, phantom, "\perp"] & \mathrm{P}(X) \arrow[l, shift left, "(\,)^\circ"] \end{tikzcd}$$

# $0_{\mathbf{C}} \dashv !_{\mathbf{C}} \dashv 1_{\mathbf{C}}$

- The terminal object of $\mathbf{CAT}$ is the category $\mathbf{1}$ containing just one object and one morphism $\begin{tikzcd} 1_\bullet \arrow[loop below] \\ \bullet \end{tikzcd}$ .
- 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:

$$\frac{1_\bullet : !_{\mathbf{C}} A \to \bullet}{!_A : A \to 1_{\mathbf{C}}\bullet} \qquad \text{similarly,} \qquad \frac{1_\bullet : \bullet \to !_{\mathbf{C}} A}{!_A : 0_{\mathbf{C}}\bullet \to A}$$

$$\mathbf{1}(!_{\mathbf{C}} A, \bullet) \cong \mathbf{C}(A, 1_{\mathbf{C}}\bullet) \qquad\qquad \mathbf{C}(0_{\mathbf{C}}\bullet, A) \cong \mathbf{1}(\bullet, !_{\mathbf{C}} A)$$

$$0_{\mathbf{C}} \dashv !_{\mathbf{C}} \dashv 1_{\mathbf{C}}$$

# $+ \dashv \Delta \dashv \times$

## Theorem

1. *$\Delta$ 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. *$\Delta$ has a left adjoint iff $\mathbf{C}$ has binary coproducts, and the left adjoint is $+ : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$.*

$$\mathbf{C} \times \mathbf{C}(\Delta X, (A, B)) \cong \mathbf{C}(X, A) \times \mathbf{C}(X, B) \cong \mathbf{C}(X, A \times B)$$

$$\mathbf{C}(A + B, X) \cong \mathbf{C}(A, X) \times \mathbf{C}(B, X) \cong \mathbf{C} \times \mathbf{C}((A, B), \Delta X)$$

$$\begin{tikzcd}
X \arrow[d, "f"'] & X \times Y \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] \arrow[d, "f \times g"] & Y \arrow[d, "g"] \\
A & A \times B \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & B
\end{tikzcd}
\quad
\begin{tikzcd}
X \arrow[d, "f"'] & X \times X \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] \arrow[d, "f \times g"] & X \arrow[d, "g"] \\
A & A \times B \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & B
\end{tikzcd}
\quad
\frac{(f, g) : \Delta X \to (A, B)}{\overline{f \times g : X \to A \times B}}$$

$$\begin{tikzcd}
X & X + Y \arrow[l, "\iota_1"'] \arrow[r, "\iota_2"] & Y \\
A \arrow[u, "f"] & A + B \arrow[l, "\iota_1"] \arrow[r, "\iota_2"'] \arrow[u, "f+g"'] & B \arrow[u, "g"']
\end{tikzcd}
\quad
\begin{tikzcd}
X & X + X \arrow[l, "\iota_1"'] \arrow[r, "\iota_2"] & X \\
A \arrow[u, "f"] & A + B \arrow[l, "\iota_1"] \arrow[r, "\iota_2"'] \arrow[u, "f+g"'] & B \arrow[u, "g"']
\end{tikzcd}
\quad
\frac{f + g : A + B \to X}{\overline{(f, g) : (A, B) \to \Delta X}}$$

# Cartesian Closed Category

## Theorem (Cartesian Closed Category)

*A category $\mathbf{C}$ is cartesian closed iff, the following functors have right adjoints:*

$$
\begin{aligned}
!_{\mathbf{C}} &: \mathbf{C} \to \mathbf{1} \\
\Delta &: \mathbf{C} \to \mathbf{C} \times \mathbf{C} \\
(- \times A) &: \mathbf{C} \to \mathbf{C}
\end{aligned}
$$

*Here $!_{\mathbf{C}}$ is the unique functor from $\mathbf{C}$ to the terminal category $\mathbf{1}$ and $\Delta$ is the diagonal functor $\Delta A = \langle A, A \rangle$, and the right adjoint of $- \times A$ is exponentiation by $A$.*

# $\vee \dashv \Delta \dashv \wedge$

$\vee \dashv \Delta$

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

$$\mathbf{Prop}(A \vee B, C) \cong \mathbf{Prop} \times \mathbf{Prop}((A, B), \Delta C)$$

$\Delta \dashv \wedge$

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

$$\mathbf{Prop} \times \mathbf{Prop}(\Delta C, (A, B)) \cong \mathbf{Prop}(C, A \wedge B)$$

# $\varinjlim \dashv \Delta \dashv \varprojlim$

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

$$\begin{tikzcd} \mathbf{C} \arrow[r, shift left, "F"] \arrow[r, phantom, "\perp"] & \mathbf{D} \arrow[l, shift left, "G"] \arrow[r, shift left, "F'"] \arrow[r, phantom, "\perp"] & \mathbf{E} \arrow[l, shift left, "G'"] \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} \mathbf{C} \arrow[r, shift left, "F'F"] \arrow[r, phantom, "\perp"] & \mathbf{E} \arrow[l, shift left, "G'G"] \end{tikzcd}$$

**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* $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}$ *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.*

$$\begin{tikzcd} F \arrow[r, "F\eta"] \arrow[dr, "1_F"'] & FGF \arrow[d, "\varepsilon F"] \\ & F \end{tikzcd} \qquad \begin{tikzcd} G \arrow[r, "\eta G"] \arrow[dr, "1_G"'] & GFG \arrow[d, "G\varepsilon"] \\ & G \end{tikzcd}$$

**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, $\mathrm{Hom}$-sets, infinite conditions, etc. Thus, anything defined by adjunctions such as free groups, product spaces, quantifiers, ... can be defined equationally.

| $X \underset{g}{\overset{f}{\rightleftarrows}} Y$ | **Topology** | **Category** | $\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D}$ |
|---|---|---|---|
| $fg = 1_Y$<br>$gf = 1_X$<br>equality | homeomorphism between spaces | isomorphism<br>$\mathbf{C} \cong \mathbf{D}$ | $FG = 1_{\mathbf{D}}$<br>$GF = 1_{\mathbf{C}}$<br>equality |
| $fg \cong 1_Y$<br>$gf \cong 1_X$<br>homotopy | homotopy equivalence between spaces | equivalence<br>$\mathbf{C} \simeq \mathbf{D}$ | $FG \cong 1_{\mathbf{D}}$<br>$GF \cong 1_{\mathbf{C}}$<br>natural isomorphism |
| | | adjunction<br>$F \dashv G$ | $\eta : 1_{\mathbf{C}} \to GF$<br>$\varepsilon : FG \to 1_{\mathbf{D}}$<br>natural transformation |

*Every sufficiently good analogy is yearning to become a functor.*

*— John Baez*

functor < adjunction < equivalence < isomorphism < equality

# 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.

$$\begin{tikzcd} F \arrow[r, "F\eta"] \arrow[dr, "1_F"'] & FGF \arrow[d, "\varepsilon F"] \\ & F \end{tikzcd} \qquad \begin{tikzcd} G \arrow[r, "\eta G"] \arrow[dr, "1_G"'] & GFG \arrow[d, "G\varepsilon"] \\ & G \end{tikzcd}$$

## 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* $\eta$ *and counit* $\varepsilon : FG \xrightarrow{FG\gamma^{-1}} FGFG \xrightarrow{F\eta^{-1}G} FG \xrightarrow{\gamma} 1_{\mathbf{D}}$.

Let $\theta : \mathbf{D}(F-,-) \to \mathbf{C}(-,G-)$ 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 $\theta$ from $\eta$ as follows, for $g : FA \to B$,

$$\theta_{A,B}g = \theta_{A,B}(g \circ 1_{FA}) = Gg \circ \theta_{A,FA}(1_{FA}) = Gg \circ \eta_A$$

$$\begin{tikzcd}
\mathbf{D}(FA,FA) \arrow[r, "\theta_{A,FA}"] \arrow[d, "\mathbf{D}(FA,g)"'] & \mathbf{C}(A,GFA) \arrow[d, "\mathbf{C}(A,Gg)"] \\
\mathbf{D}(FA,B) \arrow[r, "\theta_{A,B}"'] & \mathbf{C}(A,GB)
\end{tikzcd}$$

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 $\theta^{-1}$ from $\varepsilon$ as follows, for $f : A \to GB$,

$$\theta^{-1}_{A,B} f = \theta^{-1}_{A,B}(1_{GB} \circ f) = \theta^{-1}_{GB,B} 1_{GB} \circ Ff = \varepsilon_B \circ Ff$$

$$\begin{tikzcd}
\mathbf{D}(FGB, B) \arrow[d, "\mathbf{D}(Ff,B)"'] & \mathbf{C}(GB, GB) \arrow[l, "\theta^{-1}_{GB,B}"'] \arrow[d, "\mathbf{C}(f,GB)"] \\
\mathbf{D}(FA, B) & \mathbf{C}(A, GB) \arrow[l, "\theta^{-1}_{A,B}"]
\end{tikzcd}$$

## Theorem

- *Functors* $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}$ *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.* $f = Gg \circ \eta_A$.

$$\begin{tikzcd} A \arrow[r, "\eta_A"] \arrow[dr, "f"'] & G(FA) \arrow[d, dashed, "Gg"] \\ & GB \end{tikzcd}$$

- *Functors* $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}$ *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.* $g = \varepsilon_B \circ Ff$.

$$\begin{tikzcd} B & F(GB) \arrow[l, "\varepsilon_B"'] \\ & FA \arrow[ul, "g"] \arrow[u, dashed, "Ff"'] \end{tikzcd}$$

## Lemma

*Given* $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}$ *, with counit* $\varepsilon : FG \to 1_{\mathbf{D}}$*,*

- $G$ *is faithful iff* $\varepsilon_B$ *is an epimorphism for all* $B$*.*
- $G$ *is full and faithful iff* $\varepsilon_B$ *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 $r \dashv i$:

$$\mathbf{C} \underset{i}{\overset{r}{\underset{\hookrightarrow}{\overset{\longleftarrow}{\perp}}}} \mathbf{D}$$

# 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.

$\eta$ and $\varepsilon$ may not satisfy the triangle identity. Define $\varepsilon_0$ to be

$$\begin{tikzcd}
FGFG \arrow[r, "F\eta^{-1}G"] & FG \arrow[d, "\varepsilon"] \\
FG \arrow[u, "FG\varepsilon^{-1}"] \arrow[r, "\varepsilon_0"'] & 1_{\mathbf{D}}
\end{tikzcd}$$

the following diagram commutes,

$$\begin{tikzcd}
F \arrow[rr, "\varepsilon^{-1}F"] \arrow[dr, "F\eta"] \arrow[ddr, bend right, "1_F"'] & & FGF \arrow[dr, "F\eta GF"] & \\
& FGF \arrow[rr, "FG\varepsilon^{-1}F"] \arrow[dl, "\varepsilon_0 F"] & & FGFGF \arrow[dl, "F\eta^{-1}GF"] \\
F & & FGF \arrow[ll, "\varepsilon F"] &
\end{tikzcd}$$

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 $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}$ 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 fixed points:
$$\mathbf{C}_{\text{fix}} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\simeq_{\perp}}}}} \mathbf{D}_{\text{fix}}$$

# Fixpoint Equivalence of an Adjunction

$$\begin{tikzcd}
\mathbf{C} \arrow[r, shift left, "F"] \arrow[r, phantom, "\perp"] & \mathbf{D} \arrow[l, shift left, "G"] \\
\mathbf{C}_{\mathrm{fix}} \arrow[u, hook] \arrow[r, shift left, "F"] \arrow[r, phantom, "\simeq\perp"] & \mathbf{D}_{\mathrm{fix}} \arrow[u, hook] \arrow[l, shift left, "G"]
\end{tikzcd}$$

Proof.

- for $A \in \mathbf{C}_{\mathrm{fix}} \hookrightarrow \mathbf{C}$, since $\eta_A$ is an isomorphism in $\mathbf{C}$, $\varepsilon_{FA}$ is an isomorphism in $\mathbf{D}$.
- for $B \in \mathbf{D}_{\mathrm{fix}} \hookrightarrow \mathbf{D}$, since $\varepsilon_B$ is an isomorphism in $\mathbf{D}$, $\eta_{GB}$ is an isomorphism in $\mathbf{C}$.

$$\begin{tikzcd}
FA \arrow[r, "F\eta_A"] \arrow[dr, "1_{FA}"'] & FGFA \arrow[d, "\varepsilon_{FA}"] \\
& FA
\end{tikzcd}
\qquad
\begin{tikzcd}
GB \arrow[r, "\eta_{GB}"] \arrow[dr, "1_{GB}"'] & GFGB \arrow[d, "G\varepsilon_B"] \\
& GB
\end{tikzcd}$$

# 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*

$$\begin{tikzcd}
\mathbf{C} \arrow[r, shift left, "F"] \arrow[r, phantom, "\perp"] & \mathbf{D} \arrow[l, shift left, "G"] \\
\mathbf{C}_{\mathrm{fix}} \arrow[u, hook] \arrow[r, shift left, "F"] \arrow[r, phantom, "\simeq\perp"] & \mathbf{D}_{\mathrm{fix}} \arrow[u, hook] \arrow[l, shift left, "G"]
\end{tikzcd}$$

*"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 $\mathbf{C}_{\mathrm{fix}} \simeq \mathbf{D}_{\mathrm{fix}}$ 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 $n \in \mathbb{P}$.

$$\begin{tikzcd}
\mathbb{N} \arrow[r, shift left, "F"] \arrow[r, phantom, "\perp"] & \mathbb{N} \arrow[l, shift left, "G"] \\
\mathbb{N} \arrow[u, hook] \arrow[r, shift left, "F"] \arrow[r, phantom, "\simeq\perp"] & \mathbb{P} \arrow[u, hook] \arrow[l, shift left, "G"]
\end{tikzcd}$$

# 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 $\mathrm{cosk} = R \circ T$:

$$\mathrm{sk}(n) = \begin{cases} n & n \in N_{\text{even}} \\ n - 1 & n \in N_{\text{odd}} \end{cases} \quad \text{and} \quad \mathrm{cosk}(n) = \begin{cases} n + 1 & n \in N_{\text{even}} \\ n & n \in N_{\text{odd}} \end{cases}$$

# 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 $+ \dashv \Delta \dashv \times$

$$\mathbf{C} \times \mathbf{C}(\Delta X, (A, B)) \cong \mathbf{C}(X, A \times B)$$

$$\eta : 1_{\mathbf{C}} \to \times \circ \Delta$$

$$\varepsilon : \Delta \circ \times \to 1_{\mathbf{C} \times \mathbf{C}}$$

$$\begin{tikzcd}
 & A \arrow[dl, "\Delta"'] \arrow[dd, "\eta"] \\
(A, A) \arrow[dr, "\times"'] & \\
 & A \times A
\end{tikzcd}
\qquad
\begin{tikzcd}
(A \times B, A \times B) \arrow[dd, "\varepsilon"', "(\pi_1, \pi_2)"] & \\
 & A \times B \arrow[ul, "\Delta"'] \\
(A, B) \arrow[ur, "\times"'] &
\end{tikzcd}$$

$$\mathbf{C}(A + B, X) \cong \mathbf{C} \times \mathbf{C}((A, B), \Delta X)$$

$$\eta : 1_{\mathbf{C} \times \mathbf{C}} \to \Delta \circ +$$

$$\varepsilon : + \circ \Delta \to 1_{\mathbf{C}}$$

$$\begin{tikzcd}
 & (A, B) \arrow[dl, "+"'] \arrow[dd, "(\iota_1, \iota_2)"', "\eta"] \\
A + B \arrow[dr, "\Delta"'] & \\
 & (A + B, A + B)
\end{tikzcd}
\qquad
\begin{tikzcd}
A + A \arrow[dd, "\varepsilon"'] & \\
 & (A, A) \arrow[ul, "+"'] \\
A \arrow[ur, "\Delta"'] &
\end{tikzcd}$$

# Example — Left/Right Adjoint & Unit/Counit
# $- \times B \dashv (-)^B$

$$\mathrm{Hom}(A \times B, C) \cong \mathrm{Hom}(A, C^B)$$

$$\eta_A : A \to (A \times B)^B$$

$$\varepsilon_A : A^B \times B \to A$$

$$\begin{tikzcd}
 & A \arrow[dl, "-\times B"'] \arrow[dd, "\eta"] \\
A \times B \arrow[dr, "(-)^B"'] & \\
 & (A \times B)^B
\end{tikzcd}
\qquad
\begin{tikzcd}
A^B \times B \arrow[dd, "\varepsilon"'] & \\
 & A^B \arrow[ul, "-\times B"'] \\
A \arrow[ur, "(-)^B"'] &
\end{tikzcd}$$

$$\frac{A \wedge B \vdash C}{\overline{A \vdash B \to C}}$$

$$\eta_A : A \to B \to A \wedge B$$

$$\varepsilon_A : (B \to A) \wedge B \to A$$

# 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 $(GB, \varepsilon_B)$ is a universal morphism from $F$ to $B$.

$$\begin{tikzcd} B & F(GB) \arrow[l, "\varepsilon_B"'] \\ & FA \arrow[u, dashed, "Ff"'] \arrow[ul, "g"] \end{tikzcd}$$

- 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 $(FA, \eta_A)$ is a universal morphism from $A$ to $G$.

$$\begin{tikzcd} A \arrow[r, "\eta_A"] \arrow[dr, "f"'] & G(FA) \arrow[d, dashed, "Gg"] \\ & GB \end{tikzcd}$$

$$\boxed{\text{Universality} \equiv \text{Adjunctions}}$$

# Characterizations of Adjunction $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}$

- 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

$$\eta : 1_{\mathbf{C}} \to GF \qquad \varepsilon : FG \to 1_{\mathbf{D}}$$

$$\begin{tikzcd} F \arrow[r, "F\eta"] \arrow[dr, "1_F"'] & FGF \arrow[d, "\varepsilon F"] \\ & F \end{tikzcd} \qquad \begin{tikzcd} G \arrow[r, "\eta G"] \arrow[dr, "1_G"'] & GFG \arrow[d, "G\varepsilon"] \\ & G \end{tikzcd}$$

that are natural transformations.

- Universal Property
$(FA, \eta_A)$ is an initial object of the comma category $A \downarrow G$.

$$\begin{tikzcd} A \arrow[r, "\eta_A"] \arrow[dr, "f"'] & G(FA) \arrow[d, dashed, "Gg"] \\ & GB \end{tikzcd}$$

# Example — Extension Problem

$$\begin{tikzcd} A \arrow[r, "\eta_A"] \arrow[dr, "f"'] & G(FA) \arrow[d, dashed, "G\overline{f}"] \\ & GB \end{tikzcd}$$

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 (Continuity)

- *Right adjoints preserve limits.*
- *Left adjoints preserve colimits.*

## Proof.

Given $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}$ , assume limits of shape $\mathbf{I}$ exist in both $\mathbf{C}$ and $\mathbf{D}$.

Define $F_*(D) := FD$ and $G_*(D) := GD$. Then it's easy to see that

$$\mathbf{C}^{\mathbf{I}} \underset{G_*}{\overset{F_*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}^{\mathbf{I}}$$

Since the diagram of right adjoints of the functors in a commutative square commutes up to natural isomorphism, then

$$\begin{tikzcd} \mathbf{C} \arrow[r, "F"] \arrow[d, "\Delta"'] & \mathbf{D} \arrow[d, "\Delta"] \\ \mathbf{C}^{\mathbf{I}} \arrow[r, "F_*"'] & \mathbf{D}^{\mathbf{I}} \end{tikzcd} \implies \begin{tikzcd} \mathbf{C} & \mathbf{D} \arrow[l, "G"'] \\ \mathbf{C}^{\mathbf{I}} \arrow[u, "\varprojlim"] & \mathbf{D}^{\mathbf{I}} \arrow[l, "G_*"] \arrow[u, "\varprojlim"'] \end{tikzcd}$$

$$G \varprojlim D \cong \varprojlim GD$$

# Example

- Right adjoints preserve limits.

$$\frac{C \to A \wedge B}{\overline{(C \to A) \wedge (C \to B)}} \; - \wedge C \dashv C \to -$$

$$\frac{\forall x (Ax \wedge Bx)}{\overline{\forall x Ax \wedge \forall x Bx}} \; \exists \dashv \pi^* \dashv \forall$$

- Left adjoints preserve colimits.

$$\frac{(A \vee B) \wedge C}{\overline{(A \wedge C) \vee (B \wedge C)}} \; - \wedge C \dashv C \to -$$

$$\frac{\exists x (Ax \vee Bx)}{\overline{\exists x Ax \vee \exists x Bx}} \; \exists \dashv \pi^* \dashv \forall$$

## 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# Left/Right Adjoint

$A$ $\qquad$ $\Box A$ $\qquad$ $\Diamond A$

$\Box A$ 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 $[R] : \mathrm{P}(Y) \to \mathrm{P}(X)$:

$$\frac{F_R(A) \subset B}{\overline{\overline{A \subset [R]B}}}$$

The definition of $[R]$ which satisfies this condition is:

$$[R]B := \big\{x : \forall y \big(Rxy \to y \in B\big)\big\}$$

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 $f : X \to Y$, consider
$f^* : \mathrm{P}(Y) \to \mathrm{P}(X) :: B \mapsto \{x \in X : fx \in B\}$.
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, $f^*B(x) = B(fx) =: (Bf)(x)$.

$$\begin{tikzcd} f^*B \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, hook, "i"] \\ X \arrow[r, "f"'] & Y \end{tikzcd}$$

Then $f^*$ has both a left and a right adjoint $\exists_f, \forall_f : \mathrm{P}(X) \to \mathrm{P}(Y)$.

$$\frac{A \subset f^*B}{\exists_f A \subset B} \qquad \frac{f^*B \subset A}{B \subset \forall_f A} \qquad \text{i.e.} \qquad \frac{A \vdash_X Bf}{\exists_f A \vdash_Y B} \qquad \frac{Bf \vdash_X A}{B \vdash_Y \forall_f A}$$

$$\boxed{\exists_f \dashv f^* \dashv \forall_f}$$

The unique functions satisfying these conditions can be defined as

$$\exists_f A := \bigcap \{B : A \subset f^*B\} = fA = \big\{y \in Y : \exists x\big(fx = y \wedge x \in A\big)\big\}$$

$$\forall_f A := \bigcup \{B : f^*B \subset A\} = \{y \in Y : f^*y \subset A\} = \big\{y \in Y : \forall x\big(fx = y \to x \in A\big)\big\}$$

# Quantifiers are Adjoints

$$\frac{A \subset f^{-1}B}{\overline{\overline{fA \subset B}}} \qquad \frac{A \vdash_X Bf}{\overline{\overline{\exists_f A \vdash_Y B}}}$$

$$\begin{tikzcd}
\mathrm{P}(X) \arrow[r, shift left=3, "\exists_f"] \arrow[r, shift right=3, "\forall_f"'] & \mathrm{P}(Y) \arrow[l, "f^*"'] \arrow[l, phantom, shift left=3, "\perp"] \arrow[l, phantom, shift right=3, "\perp"]
\end{tikzcd}$$

# Remark

$$\begin{tikzcd}
\{Ax\}_{x\in X} \cong \mathrm{P}(X) \arrow[r, shift left=3, "\exists_f"] \arrow[r, shift right=3, "\forall_f"'] & f^* \arrow[l, shift right=3, "\perp"'] \arrow[l, shift left=3, "\perp"] & \mathrm{P}(\bullet) = \{\emptyset, \bullet\} \cong \{\perp, \top\} \arrow[l, equal]
\end{tikzcd}$$

The set $\mathrm{P}(X) = \{A : A \subset X\}$ can be identified with the set of predicates $Ax$. The corresponding subset is $\{x \in X : Ax = \top\} = f^{-1}(\top)$.

$$\exists_f : A \mapsto \exists x A$$
$$\forall_f : A \mapsto \forall x A$$

## Quantifiers are Adjoints $\exists_\pi \dashv \pi^* \dashv \forall_\pi$

Let $f$ be the projection $\pi : X \times Y \to X$. Hence $\pi^* : \mathrm{P}(X) \to \mathrm{P}(X \times Y)$.

$$\begin{tikzcd} \pi^* B \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, hook, "i"] \\ X \times Y \arrow[r, "\pi"'] & X \end{tikzcd}$$

By the pullback, $\pi^* B(x, y) = B\pi(x, y)$.
Then $\pi^*$ 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 wff $A$. Let $[\![A]\!] := \{(\mathbf{b}, a) : \mathcal{M} \vDash A[\mathbf{b}, a]\}$. 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: $\longrightarrow$ | Heyting implication |
| existential quantification: $\exists$ | left adjoint to pullback |
| universal quantification: $\forall$ | right adjoint to pullback |

# Diagonalization[Law69][25]

## 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. $y = f \circ x$.

$$\begin{tikzcd} X \arrow[r, "f"] & Y \\ 1 \arrow[u, dashed, "x"] \arrow[ur, "y"'] & \end{tikzcd}$$

## 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 $x : 1 \to X$:

$$g \circ x = f \circ \langle x, y \rangle$$

$$\begin{tikzcd} X \times Y \arrow[r, "f"] & Z \\ 1 \arrow[u, "\langle x,y \rangle"] \arrow[r, "x"'] & X \arrow[u, "g"'] \end{tikzcd}$$

## 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* $y : 1 \to Y$.

$$\begin{tikzcd} X \times X \arrow[r, "f"] & Y \arrow[d, "\alpha"] \\ X \arrow[u, "\Delta"] \arrow[r, "g"'] & Y \end{tikzcd}$$

[25] 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.

$$g = f \circ (1_X \times y) \circ i$$

$$\begin{tikzcd} X \arrow[r, "i", "\cong"'] \arrow[rrr, bend right, "g"'] & X \times 1 \arrow[r, "1_X \times y"] & X \times Y \arrow[r, "f"] & Z \end{tikzcd}$$

## 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.*

$$\begin{tikzcd} X \times X \arrow[rrr, "f"] & & & Y \arrow[d, "\alpha"] \\ X \arrow[u, "\Delta"] \arrow[r, "i", "\cong"'] & X \times 1 \arrow[r, "1_X \times y"] & X \times X \arrow[r, "f"] & Y \\ 1 \arrow[u, "y"] \arrow[r, "y"'] & X \arrow[ur, "\Delta"'] & & \end{tikzcd}$$

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* is a category $\mathbf{C}$ with a functor $\otimes : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$.
- A *pointed magmoidal category* 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 $\Delta_{\mathbf{C}}$.

$$\begin{tikzcd}
X \arrow[r, "\Delta_X"] \arrow[d, "f"'] & X \otimes X \arrow[d, "f \otimes f"] \\
Y \arrow[r, "\Delta_Y"'] & Y \otimes Y
\end{tikzcd}$$

## 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* $z : T \to Z$.

$$
\begin{tikzcd}
X \arrow[r, "\Delta_X"] & X \otimes X \arrow[r, "1_X \otimes \beta"] & X \otimes Y \arrow[r, "f"] & Z \arrow[d, "\alpha"] \\
T \arrow[u, "\dot{x}"] \arrow[r, "\Delta_T"] \arrow[d, "\dot{x}"'] & T \otimes T \arrow[r, "\dot{x} \otimes y"] \arrow[d, "\dot{x} \otimes \dot{x}"'] & X \otimes Y \arrow[r, "f"] & Z \\
X \arrow[r, "\Delta_X"'] & X \otimes X \arrow[ur, "1_X \otimes \beta"'] & &
\end{tikzcd}
$$

$$
\begin{aligned}
\alpha \circ f \circ (1_X \otimes \beta) \circ \Delta_X \circ \dot{x} &= f \circ (\dot{x} \otimes y) \circ \Delta_T = f \circ (1_X \otimes \beta) \circ (\dot{x} \otimes \dot{x}) \circ \Delta_T \\
&= f \circ (1_X \otimes \beta) \circ \Delta_X \circ \dot{x}
\end{aligned}
$$

# 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 $\mathrm{Mod}(\mathbb{T}, \mathbf{C})$.

# Functorial Semantics

**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 |

## Definition (Groups in a Category)

Let $\mathbf{C}$ be a category with finite products. A group in $\mathbf{C}$ consists of an object $G$ and morphisms

$$\begin{tikzcd} G \times G \arrow[r, "m"] & G & G \arrow[l, "i"'] \\ & 1 \arrow[u, "e"'] & \end{tikzcd}$$

such that,

1. $m$ is associative,

$$\begin{tikzcd} (G \times G) \times G \arrow[r, "\cong"] \arrow[d, "m \times 1_G"'] & G \times (G \times G) \arrow[d, "1_G \times m"] \\ G \times G \arrow[r, "m"'] & G & G \times G \arrow[l, "m"] \end{tikzcd}$$

2. $e$ is a unit,

$$\begin{tikzcd} 1 \times G \arrow[d, "e \times 1_G"'] & G \arrow[l, "\cong"'] \arrow[r, "\cong"] \arrow[d, "1_G"] & G \times 1 \arrow[d, "1_G \times e"] \\ G \times G \arrow[r, "m"'] & G & G \times G \arrow[l, "m"] \end{tikzcd}$$

3. $i$ is an inverse,

$$\begin{tikzcd} G \times G \arrow[d, "1_G \times i"'] & G \arrow[l, "\Delta"'] \arrow[r, "\Delta"] \arrow[d, "!_G"] & G \times G \arrow[d, "i \times 1_G"] \\ G \times G \arrow[r, "m"'] & 1 \arrow[d, "e"] & G \times G \arrow[l, "m"] \\ & G & \end{tikzcd}$$

# Groups in a Category

## Definition

A *group homomorphism* from $G$ to $H$ is a morphism $f : G \to H$ in $\mathbf{C}$ s.t.

$$\begin{tikzcd} G \times G \arrow[r, "f \times f"] \arrow[d, "m_1"'] & H \times H \arrow[d, "m_2"] \\ G \arrow[r, "f"'] & H \end{tikzcd} \qquad \begin{tikzcd} G \arrow[rr, "f"] & & H \\ & 1 \arrow[ul, "e_1"] \arrow[ur, "e_2"'] & \end{tikzcd} \qquad \begin{tikzcd} G \arrow[r, "f"] \arrow[d, "i_1"'] & H \arrow[d, "i_2"] \\ G \arrow[r, "f"'] & H \end{tikzcd}$$

# Groups in a Category

- A group in the usual sense is a model in $\mathbf{Set}$.

$$\mathbb{T} \to \mathbf{Set}$$

- If $\mathbf{C}$ is the category of algebraic varieties, a group in $\mathbf{C}$ is an algebraic group.
- If $\mathbf{C} = \mathbf{Top}$, a group in $\mathbf{C}$ is a topological group.
- If $\mathbf{C} = \mathbf{Diff}$, a group in $\mathbf{C}$ is a Lie group.
- If $\mathbf{C} = \mathbf{Grp}$, a group in $\mathbf{C}$ is an abelian group.

Suppose $\circ$ is a group homomorphism $(G, \bullet) \times (G, \bullet) \to (G, \bullet)$. Then

$$(a \bullet b) \circ (c \bullet d) = (a \circ c) \bullet (b \circ d)$$

$$a \bullet b = (a \circ 1) \bullet (1 \circ b) = (a \bullet 1) \circ (1 \bullet b) = a \circ b = (1 \bullet a) \circ (b \bullet 1) = (1 \circ b) \bullet (a \circ 1) = b \bullet a$$

# 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 ... |
|---|---|
| $A^n \to 2$ | provably equivalent wffs of $n$ variables |
| $A^n \to 2 \times 2$ | provably equivalent tuples of wffs of $n$ variables |
| $A^n \to A$ | provably equivalent terms of $n$ variables |
| $1 \to 2$ | sentences |
| $1 \to A$ | constant terms |
| $\top : 1 \to 2$ | sentences provable |
| $\bot : 1 \to 2$ | sentences refutable |
| $2^n \to 2$ | propositional operations e.g., $\neg : 2 \to 2$ |

- 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 $\neg : 2 \to 2 : \neg\varphi \neq \varphi$ for all $\varphi : 1 \to 2$.
- A theory is *complete* iff $\mathrm{Hom}(1,2) = \{\top, \bot\}$.

# Gödel Encoding

$$\begin{tikzcd} & \text{Gödel Number} \arrow[dr, "-"] & \\ \text{Expression} \arrow[ur, "\#"] \arrow[rr, "\ulcorner\urcorner"] & & \text{Gödel Numeral} \end{tikzcd}$$

meta:

object:

$$\begin{tikzcd} & \#\zeta \arrow[dr, "-"] & \\ \zeta \arrow[ur, "\#"] \arrow[rr, "\ulcorner\urcorner"] & & \ulcorner\zeta\urcorner \end{tikzcd}$$

$$\ulcorner\zeta\urcorner = \underline{\#\zeta} = s^{\#\zeta}0$$

meta-language

$\Uparrow ? \Downarrow$

object-language

$$\begin{tikzcd} \text{Formalization} \arrow[rr, bend left, equal] & & \text{Arithmetization} \arrow[ll, bend left, equal] \end{tikzcd}$$

## Definition (Gödel Encoding)

Gödel encoding is an injective map $\ulcorner\urcorner : \mathrm{Hom}(A^n, 2) \to \mathrm{Hom}(1, A)$.

# Undefinability of $\mathrm{sat}$

## 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 $a : 1 \to A$: $\mathbb{T} \vdash \mathrm{sat}\langle a, \ulcorner \varphi \urcorner \rangle \leftrightarrow \varphi a$.

$$\begin{tikzcd}
A \times A \arrow[r, "\mathrm{sat}"] & 2 \\
1 \arrow[u, "\langle a, \ulcorner \varphi \urcorner \rangle"] \arrow[r, "a"'] & A \arrow[u, "\varphi"']
\end{tikzcd}$$

**Remark:** This is exactly the condition for weak point-surjectivity!

## Theorem (Undefinability of $\mathrm{sat}$)

*If $\mathbb{T}$ is consistent, then $\mathrm{sat}$ is not definable in $\mathbb{T}$.*

$$\begin{tikzcd}
A \times A \arrow[r, "\mathrm{sat}"] & \mathbf{2} \arrow[d, "\neg"] \\
A \arrow[u, "\Delta"] \arrow[r] & \mathbf{2}
\end{tikzcd}$$

## Proof.

If $\mathrm{sat}$ is definable in $\mathbb{T}$, then $\neg$ has a fixpoint.

# Truth Predicate

"snow is white" 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., $\mathrm{truth} \circ \ulcorner\varphi\urcorner = \varphi$.

$$\begin{tikzcd}
A \arrow[rr, "\mathrm{truth}"] & & 2 \\
1 \arrow[u, "\ulcorner\varphi\urcorner"] \arrow[urr, "\mathrm{truth}_*\ulcorner\varphi\urcorner = \mathrm{truth} \circ \ulcorner\varphi\urcorner = \varphi"'] & & \\
& & \arrow[ll, shift left, "\ulcorner\urcorner"] \arrow[ull, shift left, "\mathrm{truth}_*"]
\end{tikzcd}$$

# Tarski's Undefinability of $\mathrm{truth}$ Theorem

## Definition (Substitution)

"*substitution*" is definable in $\mathbb{T}$ iff there is $\mathrm{subst} : A \times A \to A$ such that $\mathbb{T} \vdash \mathrm{subst}\langle a, \ulcorner\varphi\urcorner\rangle = \ulcorner\varphi a\urcorner$.

$$\begin{tikzcd}
A \times A \arrow[r, "\mathrm{subst}"] & A \\
1 \arrow[u, "\langle a, \ulcorner\varphi\urcorner\rangle"] \arrow[ur, "\ulcorner\varphi a\urcorner"'] &
\end{tikzcd}$$

## Theorem (Tarski's Undefinability of $\mathrm{truth}$ Theorem)

*If $\mathbb{T}$ is consistent and "substitution" is definable in $\mathbb{T}$, then $\mathrm{truth}$ is not definable in $\mathbb{T}$.*

$$\begin{tikzcd}
A \times A \arrow[r, "\mathrm{subst}"] & A \arrow[r, "\mathrm{truth}"] & 2 \\
1 \arrow[u, "\langle a, \ulcorner\varphi\urcorner\rangle"] \arrow[ur, bend left, "\ulcorner\varphi a\urcorner"] \arrow[urr, bend left, "\varphi a"] \arrow[rr, "a"'] & & A \arrow[u, "\varphi"']
\end{tikzcd}$$

$$\mathrm{sat} = \mathrm{truth} \circ \mathrm{subst}$$

# 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 $\varphi : 1 \to 2$,

$$\mathbb{T} \vdash \varphi \iff \mathbb{T} \vdash \mathrm{prov}(\ulcorner \varphi \urcorner)$$

## 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, $\mathrm{truth} = \mathrm{prov}$.

$$\begin{tikzcd} A \arrow[r, "\mathrm{prov}"] & 2 \\ 1 \arrow[u, "\ulcorner \varphi \urcorner"] \arrow[ur, "\varphi"'] & \end{tikzcd}$$

# 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 | $\xrightarrow{\text{logic interpretation}}$ | Model |
| HOL | $\xrightarrow{\text{internal language}}$ | Topos |
| IHOL | $\xrightarrow{\text{Kleene realizability}}$ | 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 |
| $f : \mathbb{N} \to \mathbb{N}$ | 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 $p$ |

# 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 $n \Vdash_{\Sigma} \bot \iff \varphi_n(n) \uparrow$.
- 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 $\Sigma \to \Sigma$ has a fixpoint.
- It immediately follows that negation is not definable on $\Sigma$ 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 $\varphi_e$ $\Vdash_\Sigma$-realizes $h$.

# Scott's Reflexive Domain[SV84][26]

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 $\perp$, 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 fixed point, namely the supremum of the monotone sequence $\coprod_n f^n(\perp)$.
- **CCPo**: the subcategory of **CPo** obtained by restricting the morphisms to be continuous.

[26] J. Soto-Andrade & F. J. Varela: Self-Reference and Fixed Points: A Discussion and an Extension of Lawvere's Theorem.

# Examples of $\mathbf{CPo}$

- 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 $\Sigma$, 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 fixed points
- Continuity serves as an '*intrinsic* approximation' to computability.

# Reflexive Domain

For $D \in \mathrm{ob}(\mathbf{CPo})$, define

$$D_1 \coloneqq D$$

$$D_{n+1} \coloneqq [D_n, D_n]$$

$$i_n : D_n \to D_{n+1}$$

$$j_n : D_{n+1} \to D_n$$

$$i_1(x) \coloneqq \mathrm{con}(x) : D \to D, \text{ where } \mathrm{con}(x)(y) \coloneqq x$$

$$j_1(f) \coloneqq f(\bot)$$

$$i_{n+1}(x_{n+1}) \coloneqq i_n \circ x_{n+1} \circ j_n$$

$$j_{n+1}(x_{n+2}) \coloneqq j_n \circ x_{n+2} \circ i_n$$

$$\begin{tikzcd} D_{n+1} \arrow[r, "i_{n+1}(x_{n+1})"] \arrow[d, "j_n"'] & D_{n+1} \\ D_n \arrow[r, "x_{n+1}"'] & D_n \arrow[u, "i_n"'] \end{tikzcd} \qquad \begin{tikzcd} D_{n+1} \arrow[r, "x_{n+2}"] & D_{n+1} \arrow[d, "j_n"] \\ D_n \arrow[u, "i_n"] \arrow[r, "j_{n+1}(x_{n+2})"'] & D_n \end{tikzcd}$$

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 $D^\infty$ is reflexive.

$$F \circ G = 1_{[D^\infty \to D^\infty]}$$
$$G \circ F = 1_{D^\infty}$$

## Theorem

- *In the category* $\mathbf{CPo}$, *we have* $D^\infty \cong [D_\infty, D^\infty]$.
- *In the category* $\mathbf{CCPo}$, *we have* $D^\infty \cong [D^\infty, D^\infty]$.

# Models of $\lambda$-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:
$$\begin{aligned} [\![x]\!]_\rho &\coloneqq \rho(x) \\ [\![MN]\!]_\rho &\coloneqq [\![M]\!]_\rho \cdot [\![N]\!]_\rho \\ [\![\lambda x.M]\!]_\rho &\coloneqq \mathrm{lam}(f) \text{ where } f : d \mapsto [\![M]\!]_{\rho(d/x)} \end{aligned}$$

**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 $\lambda$-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.

$$\begin{tikzcd}
Y \arrow[r, "j"] & X \arrow[r, "f"] & X \arrow[r, "i"] & Y \\
1 \arrow[u, "y"] \arrow[ur] \arrow[urr] \arrow[urrr, "y"'] & & &
\end{tikzcd}$$

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 $\mathrm{dom}\, f = \mathrm{dom}\, g$
3. $\overline{g \circ \overline{f}} = \overline{g} \circ \overline{f}$ whenever $\mathrm{dom}\, f = \mathrm{dom}\, g$
4. $\overline{g} \circ f = f \circ \overline{g \circ f}$ whenever $\mathrm{dom}\, g = \mathrm{cod}\, f$

## Example

- Every category admits the trivial restriction operator $\overline{f} = 1_{\mathrm{dom}\, f}$.
- 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:

$$\begin{tikzcd}
X \times U \arrow[r, "\tau"] & Y \\
X \times T \arrow[u, dashed, "1_X \times h"] \arrow[ur, "f"'] &
\end{tikzcd}$$

**Remark:** In the special case when $h$ is uniquely determined we say that $\tau$ 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 $U^U \underset{\mathrm{app}}{\overset{\mathrm{lam}}{\rightleftarrows}} U$ s.t. $\mathrm{app} \circ \mathrm{lam} = 1_{U^U}$.
- A reflexive object is said to be extensional when also $\mathrm{lam} \circ \mathrm{app} = 1_U$.
- In a Turing category with Turing object $U$, the Turing morphism is defined by taking an embedding-retraction pair $(\mathrm{lam}, \mathrm{app}) : U^U \triangleleft U$,

$$\tau : U \times U \xrightarrow{1_U \times \mathrm{app}} U \times U^U \xrightarrow{\varepsilon} U$$

**Remark:** Viewed as a model of lambda calculus, the equation $\mathrm{app} \circ \mathrm{lam} = 1_{U^U}$ of a reflexive object represents $\beta$-reduction $(\lambda x.M)N = M[N/x]$, while the equation $\mathrm{lam} \circ \mathrm{app} = 1_U$ of an extensional reflexive object represents $\eta$-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.

$$\begin{tikzcd}
1 \times U \arrow[r, "\tau"] & X \\
1 \times X \arrow[u, dashed, "1\times\mathrm{lam}"] \arrow[ur, "\pi_2"'] &
\end{tikzcd}$$

Let $\mathrm{app} := \langle !, 1_X \rangle \circ \tau$. Then

$$\mathrm{app} \circ \mathrm{lam} = 1_X$$

## Theorem (Recognition Theorem)

*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.

"$2 \Rightarrow 1$": For arbitrary objects $X, Y \in \mathbf{C}$, by assumption we have $(\mathrm{lam}_X, \mathrm{app}_X) : X \triangleleft U$ and $(\mathrm{lam}_Y, \mathrm{app}_Y) : Y \triangleleft U$.

$$X \times U \xrightarrow{\tau} Y \quad = \quad X \times U \xrightarrow{\mathrm{lam}_X \times 1_U} U \times U \xrightarrow{\tau} U \xrightarrow{\mathrm{app}_Y} Y$$

# 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.

$$
\begin{tikzcd}
 & U \times U \times U \arrow[r, "1\times\tau"] & U \times U \arrow[d, "f"] \\
U(\times 1) \arrow[ur, bend left, "1\times e"] \arrow[r, "1\times\ulcorner g\urcorner"] \arrow[dr, bend right, "1\times e"'] & U \times U(\times 1) \arrow[u, "1\times\Delta"] \arrow[r, "g"] \arrow[d, dashed, "1\times 1\times\ulcorner g\urcorner"'] & U \\
 & U \times U \times U \arrow[r, "1\times\tau"'] & U \times U \arrow[u, "\tau"']
\end{tikzcd}
$$

$$g := f(1 \times \tau)(1 \times \Delta) \qquad e := \tau(\ulcorner g \urcorner \times \ulcorner g \urcorner)$$

# 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 $\sum_f \dashv f^*$.*

$$\begin{tikzcd} A \times_B Y \arrow[r] \arrow[d, "f^*b"'] \arrow[dr, phantom, "\lrcorner", very near start] & Y \arrow[d, "b"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

## 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 $g \in \mathrm{Hom}_{\mathbf{C}/B}(\sum_f a, b)$.

$$\begin{tikzcd} X \arrow[r, "g"] \arrow[d, "a"'] & Y \arrow[d, "b"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

By definition of $f^*b$, the following diagram is a pullback.

$$\begin{tikzcd} X \arrow[dr, dashed, "u"] \arrow[drr, bend left, "g"] \arrow[ddr, bend right, "a"'] & & \\ & A \times_B Y \arrow[r, "p"] \arrow[d, "f^*b"'] \arrow[dr, phantom, "\lrcorner", very near start] & Y \arrow[d, "b"] \\ & A \arrow[r, "f"'] & B \end{tikzcd}$$

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 "$\Rightarrow$".

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 $\mathbf{C}/B$.

$$\begin{tikzcd}
 & A \times_B Y \arrow[r] \arrow[dl, "g"'] \arrow[d, "f^*b"'] \arrow[dr, "f \times_B b"] & Y \arrow[d, "b"] \\
X \arrow[r, "a"'] & A \arrow[r, "f"'] & B
\end{tikzcd}$$

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

$$\mathrm{Hom}_{\mathbf{C}/B}(f \times_B b, fa) \cong \mathrm{Hom}_{\mathbf{C}/B}(b, (fa)^f)$$

## proof sketch of "$\Rightarrow$" 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

$$\mathrm{Hom}_{\mathbf{C}/B}(f, f) \cong \mathrm{Hom}_{\mathbf{C}/B}(1_B, f^f)$$

$$\begin{tikzcd}
\prod_f X \arrow[r] \arrow[d, "\prod_f a"'] \arrow[dr, phantom, "\lrcorner"] & X^f \arrow[d, "a^f"] \arrow[ddr, bend left, "(fa)^f"] & \\
B \arrow[r, "s"'] \arrow[drr, bend right, "1_B"'] & A^f \arrow[dr, "f^f"] & \\
& & B
\end{tikzcd}$$

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 "$\Rightarrow$" 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 $g \mapsto u$.

$$
\begin{tikzcd}
 & A \times_B Y \arrow[r] \arrow[dl, red, "g"'] \arrow[d, "f^*b"] \arrow[dr, "f \times_B b"] & Y \arrow[dr, dashed, red, "u"] \arrow[drr, red, bend left, "h"] \arrow[ddr, "b"] & & \\
X \arrow[r, "a"'] & A \arrow[rr, "f"'] & & B \arrow[r, "s"'] \arrow[d, "\prod_f a"] & A^f \arrow[d, "a^f"] \arrow[dr, "f^f"] \\
 & & & \prod_f X \arrow[r] & X^f \arrow[rr, bend right, "(fa)^f"'] & & B \\
 & & & & & B \arrow[urr, bend right, "1_B"'] &
\end{tikzcd}
$$

## proof sketch of "$\Leftarrow$".

The terminal object in $\mathbf{C}/B$ is $1_B$.
The pullback of $f$ and $b$ in $\mathbf{C}$ corresponds to a product in $\mathbf{C}/B$.

$$\begin{tikzcd} P \arrow[r] \arrow[d, "f^*b"'] \arrow[dr, "f\times_B b"] & X \arrow[d, "b"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

$$f \times_B b = f \circ f^* b = \sum\nolimits_f f^* b$$

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

$$\begin{tikzcd}[column sep=huge] \{Ax\}_{x\in X} \cong \mathrm{P}(X) \arrow[r, shift left=3, "\exists_f"] \arrow[r, shift right=3, "\forall_f"'] \arrow[r, phantom, shift left=1.5, "\perp"] \arrow[r, phantom, shift right=1.5, "\perp"] & \mathrm{P}(\bullet) = \{\emptyset, \bullet\} \cong \{\bot, \top\} \arrow[l, "f^*"'] \end{tikzcd}$$

The set $\mathrm{P}(X) = \{A : A \subset X\}$ can be identified with the set of predicates $Ax$. The corresponding subset is $\{x \in X : Ax = \top\} = f^{-1}(\top)$.

$$\exists_f : A \mapsto \exists x A$$

$$\forall_f : A \mapsto \forall x A$$

$$\begin{tikzcd}[column sep=huge] \mathbf{Set}/X \arrow[r, shift left=3, "\Sigma_f"] \arrow[r, shift right=3, "\prod_f"'] \arrow[r, phantom, shift left=1.5, "\perp"] \arrow[r, phantom, shift right=1.5, "\perp"] & \mathbf{Set}/Y \arrow[l, "f^*"'] \end{tikzcd}$$

$$f^* : \{By\}_{y\in Y} \mapsto \{B(fx)\}_{x\in X}$$

$$\textstyle\sum_f : \{Ax\}_{x\in X} \mapsto \left\{\sum_{x\in f^{-1}(y)} Ax\right\}_{y\in Y}$$

$$\textstyle\prod_f : \{Ax\}_{x\in X} \mapsto \left\{\prod_{x\in f^{-1}(y)} Ax\right\}_{y\in Y}$$

# The Poset of Partitions

# Subset Logic vs Partition Logic

$$\begin{tikzcd}
 & \{a,b,c\} & \\
\{a,b\} \arrow[ur] & \{a,c\} \arrow[u] & \{b,c\} \arrow[ul] \\
\{a\} \arrow[u] \arrow[ur] & \{b\} \arrow[u] \arrow[ul] \arrow[ur] & \{c\} \arrow[u] \arrow[ul] \\
 & \emptyset \arrow[ul] \arrow[u] \arrow[ur] &
\end{tikzcd}
\qquad
\begin{tikzcd}
 & \{\{a\},\{b\},\{c\}\} & \\
\{\{a,b\},\{c\}\} \arrow[ur] & \{\{a\},\{b,c\}\} \arrow[u] & \{\{b\},\{a,c\}\} \arrow[ul] \\
 & \{\{a,b,c\}\} \arrow[ul] \arrow[u] \arrow[ur] &
\end{tikzcd}$$

- 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 $Q$.

$$\sim_{P \wedge Q} \coloneqq \sim_P \cap \sim_Q$$

- The join $P \vee Q$ is the finest partition that is coarser than $P$ and $Q$.

$$\sim_{P \vee Q} \coloneqq \left(\sim_P \cup \sim_Q\right)^*$$

  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 $X$: $f^*(P) \coloneqq \{f^*(S) : S \in P \;\&\; f^*(S) \neq \emptyset\}$.
- 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:** $f(a \vee b) \ncong f(a) \vee f(b)$ 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 $\mathrm{P}(A)$ to $\mathrm{P}(B)$ and the function $Y \mapsto Y^\dagger$ from $\mathrm{P}(B)$ to $\mathrm{P}(A)$ 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

$$\forall a, b : f(a) = b \iff a = g(b)$$

- 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. $F^* \lhd G \implies \mathrm{Aut}(F/K) \cong G/F^*$.

# Example: Formal Concept Analysis / Chu Space

- A formal context (or *Chu space*) is a triple $(X, Y, \vDash)$, 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 $(A, X, \vDash)$ to $(B, Y, \Vdash)$ is a pair of functions $(f, g)$ with $f : A \to B$ and $g : Y \to X$ s.t.

$$\forall a \in A \forall y \in Y : a \vDash gy \iff fa \Vdash y$$

$$\begin{tikzcd}
a \arrow[r, "f"] \arrow[d, "\vDash"'] \arrow[dr, phantom, "\iff"] & fa \arrow[d, "\Vdash"] \\
gy & y \arrow[l, "g"]
\end{tikzcd}$$

- 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 $B$".

| $\vDash$ | 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 $(X, Y, \vDash)$ 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 $(X, Y, \vDash)$ is

$$\mathbf{B}(X, Y, \vDash) := \left\{(A, B) \in 2^X \times 2^Y : A^{\uparrow} = B \;\;\&\;\; B^{\downarrow} = A\right\}$$

- formal concept = fixpoint of $\uparrow\downarrow$

$$\mathbf{B}(X, Y, \vDash) \cong \mathrm{Fix}(\uparrow\downarrow) \cong \mathrm{Fix}(\downarrow\uparrow)$$

where $\mathrm{Fix}(f) = \{x : fx = x\}$.

## formal concepts = maximal rectangles = complete bipartite subgraph

| $\vDash$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
|---|---|---|---|---|
| $x_1$ | 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 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).
- Subconcept $(A, B) \leq (A', B')$ if $A \subset A'$, or equivalently if $B \supset B'$.
- The set of concepts $\mathbf{B}(X, Y, \vDash)$ 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 $B \to B' := B^{\downarrow} \subset B'^{\downarrow}$ for $B, B' \subset Y$.
- Armstrong rules for attribute implication.

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

# 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 $(X, \mathcal{O}_X, \in)$, 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 $(X, \mathcal{O}_X, \in)$ and $(Y, \mathcal{O}_Y, \in)$.
- 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. $(A, X, \vDash) \xrightarrow{(f,g)} (B, Y, \vDash)$ is a Chu transform.
- A topological space is $T_0$ if $(X, \mathcal{O}_X, \in)$ is separable.

# Chu Spaces — Examples

$$(X, Y, \vDash)$$

| | $X$ | $Y$ |
|---|---|---|
| Formal Concept | objects | attributes |
| Logic | models | formulas |
| Information System | messages | contents |
| Game | strategies | counterstrategies |
| Topology | points | open sets |

# 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 $(A, X, \vDash)$ 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 $(A, X, \vDash) \to (B, Y, \Vdash)$ are pairs of morphisms $f : A \to B$ and $g : Y \to X$ s.t.

$$\begin{tikzcd} A \otimes Y \arrow[r, "1_A \otimes g"] \arrow[d, "f \otimes 1_Y"'] \arrow[dr, phantom, "\circlearrowleft"] & A \otimes X \arrow[d, "\vDash"] \\ B \otimes Y \arrow[r, "\Vdash"'] & K \end{tikzcd}$$

- 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 $(A, X, \vDash\colon A \otimes X \to K)$ to $(X, A, \vDash^\circ\colon X \otimes A \to K)$.
- On morphisms, it takes $(f, g) : (A, X, \vDash) \to (B, Y, \Vdash)$ to $(g, f) : (Y, B, \Vdash^\circ) \to (X, A, \vDash^\circ)$.

# Remarks

$$(A, X, \vDash) \xrightarrow{(f,g)} (B, Y, \Vdash)$$

$$\begin{tikzcd}
a \arrow[r, "f"] \arrow[d, "\vDash"'] \arrow[dr, phantom, "\Longleftrightarrow"] & fa \arrow[d, "\Vdash"] \\
gy & y \arrow[l, "g"]
\end{tikzcd}$$

- 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.

## 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.

$$\mathsf{y} : A \mapsto (I_A, O_A, \vDash)$$

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,

$$\mathsf{y} : A \xrightarrow{h} B \mapsto (I_A, O_A, \vDash) \xrightarrow{(f,g)} (I_B, O_B, \vDash)$$

$$\begin{tikzcd} \bullet \arrow[r, "e_{in}"] & A \arrow[r, "h"] \arrow[rr, bend left] & B \arrow[r, "e_{out}"] & \bullet \end{tikzcd}$$

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 $(A, X, \vDash)$, where $A$ and $X$ are sets and $\vDash\colon A \times X \to 2$ is a relation.
  A morphism from $(A, X, \vDash)$ to $(B, Y, \Vdash)$ is a pair of functions $(f, g)$ with $f : A \to B$ and $g : Y \to X$ s.t.,
  $$\forall a \in A \forall y \in Y : a \vDash gy \implies fa \Vdash y$$
  $$\begin{tikzcd} a \arrow[r, "f"] \arrow[d, "\vDash"'] \arrow[dr, phantom, "\implies"] & fa \arrow[d, "\Vdash"] \\ gy & y \arrow[l, "g"] \end{tikzcd}$$
- 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 $(A, X, \vDash) \to (B, Y, \Vdash)$ is a reduction of $(B, Y, \Vdash)$ to $(A, X, \vDash)$.
  $$\forall a \in A \forall y \in Y : a \vDash gy \implies fa \Vdash y$$
  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 $\pi$.
- $E$ 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 $\pi$.
- $g : L \to E$ 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 $(\mathbb{R}^2, E, \vDash)$.
- Problem $B$: To decide whether a given point lies on a given line $l$ is the problem $(\pi, L, \Vdash)$.
- 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 $(\mathbb{R}^2, [E]^2, \vDash) \to (\pi, [L]^2, \Vdash)$.

# 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.

$$\begin{aligned}
A_D &:= A \text{ for atomic } A \\
(A \wedge B)_D &:= \exists xu \forall yv [A_D(x, y) \wedge B_D(u, v)] \\
(A \vee B)_D &:= \exists zxu \forall yv [(z = 0 \rightarrow A_D(x, y)) \wedge (z \neq 0 \rightarrow B_D(u, v))] \\
(A \rightarrow B)_D &:= \exists fg \forall xy [A_D(x, gxy) \rightarrow B_D(fx, y)] \\
(\exists z A(z))_D &:= \exists zx \forall y A_D(x, y, z) \\
(\forall z A(z))_D &:= \exists f \forall zy A_D(fz, y, z)
\end{aligned}$$

- 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 $A = (A, X, R)$ and $B = (B, Y, S)$:

$$A \otimes B \coloneqq \left(A \times B, Y^A \times X^B, \otimes\right)$$

where

$$(a, b) \otimes (f, g) \coloneqq R(a, gb) \wedge S(b, fa)$$

- The internal-hom of $A = (A, X, R)$ and $B = (B, Y, S)$:

$$A \multimap B \coloneqq \left(B^A \times X^Y, A \times Y, \multimap\right)$$

where

$$(f, g) \multimap (a, y) \coloneqq R(a, gy) \to S(fa, y)$$

- The linear negation of $A = (A, X, R)$:

$$A^* \coloneqq (X, A, R^*) \quad \text{where} \quad R^*(x, a) \coloneqq \neg R(a, x)$$

- The unit and the dualizing object:

$$I = (1, 1, =) \qquad \bot = (1, 1, \emptyset)$$

- 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 $A = (A, X, R)$ and $B = (B, Y, S)$ are defined by the pullback $P$ and $Q$:

$$\begin{tikzcd} P \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & Y^A \arrow[d, "(\check{S})^A"] \\ X^B \arrow[r, "(\check{R})^B"'] & K^{A \otimes B} \end{tikzcd} \qquad \begin{tikzcd} Q \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & B^A \arrow[d, "(\hat{S})^A"] \\ X^Y \arrow[r, "(\check{R})^Y"'] & K^{A \otimes Y} \end{tikzcd}$$

$$A \otimes B := (A \otimes B, P, \otimes) \qquad A \multimap B := (Q, A \otimes Y, \multimap)$$

**Example:** In $\mathbf{Chu}_2(\mathbf{Set})$, we have

$$(a \otimes b) \otimes (f, g) := R(a, gb) \quad (\text{also} = S(b, fa))$$
$$(f, g) \multimap (a \otimes y) := R(a, gy) \quad (\text{also} = S(fa, y))$$

- The linear negation of $A = (A, X, R)$:

$$A^* := (X, A, R^\circ) \quad \text{where} \quad R^\circ(x, a) := R(a, x)$$

- The unit and the dualizing object:

$$I = (I, K, \lambda_K : I \otimes K \to K) \qquad \bot = (K, I, \rho_K : K \otimes I \to K)$$

# Adjoint Functor Theorem for Posets

## 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

$$z \geq \bigwedge \{y \in Y : g(z) \leq g(y)\} = f(g(z))$$

# General Adjoint Functor Theorem

## 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 $f = Gg_i \circ f_i$.*

$$\begin{tikzcd} A \arrow[r, "f_i"] \arrow[dr, "f"'] & GB_i \arrow[d, "Gg_i"] \\ & GB \end{tikzcd}$$

# Special Adjoint Functor Theorem

## 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 $A \underset{g}{\overset{f}{\rightrightarrows}} B$ 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# Kan Extension

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & & \mathbf{E} \\
& \mathbf{D} \arrow[ur, dashed, "?"'] &
\end{tikzcd}$$

- 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 $\gamma = \alpha K \circ \eta$.

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & \arrow[d, phantom, "\Downarrow\eta"] & \mathbf{E} \\
& \mathbf{D} \arrow[ur, dashed, "\mathrm{Lan}_K F"'] &
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & \arrow[d, phantom, "\Downarrow\gamma"] & \mathbf{E} \\
& \mathbf{D} \arrow[ur, "G"'] &
\end{tikzcd}
=
\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & \arrow[d, phantom, "\Downarrow\eta"] & \mathbf{E} \\
& \mathbf{D} \arrow[ur, bend left, "\mathrm{Lan}_K F"] \arrow[ur, bend right, "G"'] \arrow[ur, phantom, "\overset{\alpha}{\Searrow}"] &
\end{tikzcd}$$

# 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 $\delta = \varepsilon \circ \beta K$.

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & \arrow[d, phantom, "\Uparrow\varepsilon"] & \mathbf{E} \\
& \mathbf{D} \arrow[ur, dashed, "\mathrm{Ran}_K F"'] &
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & \arrow[d, phantom, "\Uparrow\delta"] & \mathbf{E} \\
& \mathbf{D} \arrow[ur, "G"'] &
\end{tikzcd}
\quad = \quad
\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & \arrow[d, phantom, "\Uparrow\varepsilon"] & \mathbf{E} \\
& \mathbf{D} \arrow[ur, bend left, "\mathrm{Ran}_K F"] \arrow[ur, bend right, "G"'] \arrow[ur, phantom, "\Leftarrow\beta"] &
\end{tikzcd}$$

# 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

$$\begin{tikzcd}
\mathbf{1} \arrow[rr, "*"] \arrow[dr, "A"'] & \arrow[d, phantom, "\Downarrow x"] & \mathbf{Set} \\
& \mathbf{C} \arrow[ur, "F"'] &
\end{tikzcd}$$

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 $A : \mathbf{1} \to \mathbf{C}$.

$$\begin{tikzcd}
\mathbf{1} \arrow[rr, "*"] \arrow[dr, "A"'] & \arrow[d, phantom, "\Downarrow 1_A"] & \mathbf{Set} \\
& \mathbf{C} \arrow[ur, dashed, "\mathrm{Hom}(A,-)"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{1} \arrow[rr, "*"] \arrow[dr, "A"'] & \arrow[d, phantom, "\Downarrow 1_A"] & \mathbf{Set} \\
& \mathbf{C} \arrow[ur, bend left, "\mathrm{Hom}(A,-)"] \arrow[ur, bend right, "F"'] \arrow[ur, phantom, "\Downarrow \psi(x)"] &
\end{tikzcd}$$

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 $K^* := - \circ K$.*

$$\begin{tikzcd}
\mathbf{E}^{\mathbf{C}} \arrow[rr, bend left=60, "\mathrm{Lan}_K"] \arrow[rr, bend right=60, "\mathrm{Ran}_K"'] \arrow[rr, phantom, "\perp"] \arrow[rr, phantom, "\perp"'] & & \mathbf{E}^{\mathbf{D}} \arrow[ll, "-\circ K"']
\end{tikzcd}$$

## Proof.

By the Yoneda Lemma, any pair $(G, \gamma)$, 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 $(\mathrm{Lan}_K, \eta)$ yields a natural isomorphism.

$$\mathbf{E}^{\mathbf{D}}(\mathrm{Lan}_K F, -) \cong \mathbf{E}^{\mathbf{C}}(F, - \circ K)$$

# Computing Left Kan Extension

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & \arrow[d, phantom, "\Downarrow\eta"] & \mathbf{E} \\
& \mathbf{D} \arrow[ur, dashed, "\mathrm{Lan}_K F"'] &
\end{tikzcd}$$

If $\mathbf{E}$ is cocomplete, then the left Kan extension $\mathrm{Lan}_K F$ exists and is:

$$\mathrm{Lan}_K F(d) := \varinjlim \left( K \downarrow d \xrightarrow{\pi_{K\downarrow d}} \mathbf{C} \xrightarrow{F} \mathbf{E} \right) \quad \text{for} \quad d \in \mathbf{D}$$

with the natural transformation $\eta$ extracted from colimiting cocones in $\mathbf{E}$.

$$\begin{tikzcd}
Ka \arrow[rr] \arrow[dr] \arrow[ddr, bend right] & & Kc \arrow[dl] \arrow[ddl, bend left] \\
& Kb \arrow[d] & \\
& d &
\end{tikzcd}
\quad \xmapsto{\pi_{K\downarrow d}} \quad
\begin{tikzcd}
a \arrow[rr] \arrow[dr] & & c \arrow[dl] \\
& b &
\end{tikzcd}$$

# Computing Right Kan Extension

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & & \mathbf{E} \\
& \mathbf{D} \arrow[u, Rightarrow, "\varepsilon"] \arrow[ur, dashed, "\mathrm{Ran}_K F"'] &
\end{tikzcd}$$

If $\mathbf{E}$ is complete, then the right Kan extension $\mathrm{Ran}_K F$ exists and is:

$$\mathrm{Ran}_K F(d) := \varprojlim \left( d \downarrow K \xrightarrow{\pi_{d\downarrow K}} \mathbf{C} \xrightarrow{F} \mathbf{E} \right) \quad \text{for} \quad d \in \mathbf{D}$$

with the natural transformation $\varepsilon$ extracted from limiting cones in $\mathbf{E}$.

$$\begin{tikzcd}
Ka \arrow[rr] \arrow[dr] & & Kc \arrow[dl] \\
& Kb & \\
& d \arrow[u] \arrow[uul, bend left] \arrow[uur, bend right] &
\end{tikzcd}
\quad \xmapsto{\pi_{d\downarrow K}} \quad
\begin{tikzcd}
a \arrow[rr] \arrow[dr] & & c \arrow[dl] \\
& b &
\end{tikzcd}$$

# 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 $\mathrm{Ran}_K : \mathbf{Rel}^{\mathbf{C}} \to \mathbf{Rel}^{\mathbf{D}}$.

$$
\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & & \mathbf{Rel} \\
& \mathbf{D} \arrow[uu, Rightarrow, "\eta"] \arrow[ur, bend right, "G"'] \arrow[ur, bend left, "\mathrm{Lan}_K F"] \arrow[ur, phantom, "\overset{\alpha}{\Searrow}"] &
\end{tikzcd}
\qquad
\frac{\alpha : \mathrm{Lan}_K F \implies G}{\eta : F \implies G \circ K}
$$

$$
\begin{tikzcd}
\mathbf{Rel}^{\mathbf{C}} \arrow[rr, bend left, "\mathrm{Lan}_K"] \arrow[rr, bend right, "\mathrm{Ran}_K"'] \arrow[rr, phantom, "\perp"] \arrow[rr, phantom, "\perp"] & & \mathbf{Rel}^{\mathbf{D}} \arrow[ll, "-\circ K"']
\end{tikzcd}
$$

$$
\begin{tikzcd}
\mathbf{C} \arrow[rr, "F"] \arrow[dr, "K"'] & & \mathbf{Rel} \\
& \mathbf{D} \arrow[uu, Rightarrow, "\varepsilon"] \arrow[ur, bend right, "G"'] \arrow[ur, bend left, "\mathrm{Ran}_K F"] \arrow[ur, phantom, "\overset{\beta}{\Nwarrow}"] &
\end{tikzcd}
\qquad
\frac{\delta : G \circ K \implies F}{\beta : G \implies \mathrm{Ran}_K F}
$$

# 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 |
|---|---|
| Ilsa | Italian1 |
| Klaus | Italian2 |
| Jörg | Italian3 |
| Sabine | Italian4 |

| Italians |
|---|
| Italian1 |
| Italian2 |
| Italian3 |
| 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* $\varinjlim D$.

$$\begin{tikzcd} \mathbf{I} \arrow[rr, "D"] \arrow[dr, "!"'] & \arrow[d, phantom, "\Downarrow\eta"] & \mathbf{C} \\ & \mathbf{1} \arrow[ur, "C"'] & \end{tikzcd} \quad = \quad \begin{tikzcd} \mathbf{I} \arrow[rr, "D"] \arrow[rr, bend right, "\Delta_C"'] & \arrow[d, phantom, "\Downarrow\eta"] & \mathbf{C} \end{tikzcd}$$

2. *Dually, the right Kan extension* $\mathrm{Ran}_! D$ *defines the limit* $\varprojlim D$.

$$\begin{tikzcd} \mathbf{I} \arrow[rr, "D"] \arrow[dr, "!"'] & \arrow[d, phantom, "\Uparrow\varepsilon"] & \mathbf{C} \\ & \mathbf{1} \arrow[ur, "C"'] & \end{tikzcd} \quad = \quad \begin{tikzcd} \mathbf{I} \arrow[rr, "D"] \arrow[rr, bend right, "\Delta_C"'] & \arrow[d, phantom, "\Uparrow\varepsilon"] & \mathbf{C} \end{tikzcd}$$

# 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 $(G, \eta)$ is a left Kan extension of the identity functor $1_{\mathbf{C}}$ along $F$ and $(F, \varepsilon)$ is a right Kan extension of the identity functor $1_{\mathbf{D}}$ along $G$.*

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "1_{\mathbf{C}}"] \arrow[dr, "F"'] & \arrow[d, phantom, "\Downarrow\eta"] & \mathbf{C} \\
& \mathbf{D} \arrow[ur, dashed, "G \cong \mathrm{Lan}_F\, 1_{\mathbf{C}}"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{D} \arrow[rr, "1_{\mathbf{D}}"] \arrow[dr, "G"'] & \arrow[d, phantom, "\Uparrow\varepsilon"] & \mathbf{D} \\
& \mathbf{C} \arrow[ur, dashed, "F \cong \mathrm{Ran}_G\, 1_{\mathbf{D}}"'] &
\end{tikzcd}$$

   *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 $\eta$.*

# Contents


Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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 $\alpha$, 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 $\lambda$ and $\rho$, 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 $\alpha, \lambda$ and $\rho$ are identities.

# Monoidal Category

$$
\begin{tikzcd}
A \otimes (B \otimes (C \otimes D)) \arrow[r, "\alpha_{A,B,C\otimes D}"] \arrow[d, "1_A \otimes \alpha_{B,C,D}"'] & (A \otimes B) \otimes (C \otimes D) \arrow[r, "\alpha_{A\otimes B,C,D}"] & ((A \otimes B) \otimes C) \otimes D \\
A \otimes ((B \otimes C) \otimes D) \arrow[rr, "\alpha_{A,B\otimes C,D}"'] & & (A \otimes (B \otimes C)) \otimes D \arrow[u, "\alpha_{A,B,C} \otimes 1_D"']
\end{tikzcd}
$$

$$
\begin{tikzcd}
A \otimes (I \otimes B) \arrow[rr, "\alpha_{A,I,B}"] \arrow[dr, "1_A \otimes \lambda_B"'] & & (A \otimes I) \otimes B \arrow[dl, "\rho_A \otimes 1_B"] \\
& A \otimes B &
\end{tikzcd}
$$

# Example

- $(\mathbf{Set}, \times, \{\bullet\})$
- $(\mathbf{Set}, \coprod, \emptyset)$
- $(\mathbf{Vect}, \otimes, \mathbb{R})$
- $(\mathbf{Vect}, \oplus, \mathbf{1})$
- Any monoid can be thought of as a discrete monoidal category.
  - $(\mathbb{R}, +, 0)$
  - $(\mathbb{R}, \cdot, 1)$
  - $(\mathrm{list}(X), *, [])$
  - $(\mathbf{Bool}, \wedge, \top)$

# State

## Definition (State)

A state of an object $A$ in a monoidal category is a morphism $I \to A$.

## Example

- In $\mathbf{Set}$, points of a set $A$ are morphisms $\{\bullet\} \to A$, which correspond to elements of $A$.
- In $\mathbf{Hilb}$, points of a Hilbert space $H$ are morphisms $\mathbb{C} \to H$, 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": $A \to I$.

- The identity $1_I : I \to I$ 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

$$(f \otimes 1_z) \,;\, (1_x \otimes g)$$

$$(f \,;\, g) \otimes (h \,;\, k) = (f \otimes h) \,;\, (g \otimes k)$$

# 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* | | $[\ ]$ | $0 \to 1$ |
| *copy* | | $\begin{bmatrix} 1 & 1 \end{bmatrix}$ | $1 \to 2$ |
| *discard* | | $[\ ]$ | $1 \to 0$ |
| *swap* | | $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ | $2 \to 2$ |
| *identity* |  | $[1]$ | $1 \to 1$ |
| *scalar* | | $[a]$ | $1 \to 1$ |

# 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 $\sigma$ 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.,

$$\begin{tikzcd}
A \otimes (B \otimes C) \arrow[r, "\alpha_{A,B,C}"] \arrow[d, "\sigma_{A,B\otimes C}"'] & (A \otimes B) \otimes C \arrow[r, "\sigma_{A,B}\otimes 1_C"] & (B \otimes A) \otimes C \arrow[d, "\alpha^{-1}_{B,A,C}"] \\
(B \otimes C) \otimes A & B \otimes (C \otimes A) \arrow[l, "\alpha_{B,C,A}"] & B \otimes (A \otimes C) \arrow[l, "1_B\otimes\sigma_{A,C}"]
\end{tikzcd}$$

$$\begin{tikzcd}
(A \otimes B) \otimes C \arrow[r, "\alpha^{-1}_{A,B,C}"] \arrow[d, "\sigma_{A\otimes B,C}"'] & A \otimes (B \otimes C) \arrow[r, "1_A\otimes\sigma_{B,C}"] & A \otimes (C \otimes B) \arrow[d, "\alpha_{A,C,B}"] \\
C \otimes (A \otimes B) & (C \otimes A) \otimes B \arrow[l, "\alpha^{-1}_{C,A,B}"] & (A \otimes C) \otimes B \arrow[l, "\sigma_{A,C}\otimes 1_B"]
\end{tikzcd}$$

# 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

$(\mathrm{Mat}, \rightarrow, +, 0)$ is a symmetric monoidal preorder, where $\mathrm{Mat}$ is the set of all collections of atoms and molecules, sometimes called *materials*.

For example, given these chemical reactions:

$$\mathrm{C} + \mathrm{O_2} \rightarrow \mathrm{CO_2}$$
$$\mathrm{CO_2} + \mathrm{NaOH} \rightarrow \mathrm{NaHCO_3}$$
$$\mathrm{NaHCO_3} + \mathrm{HCl} \rightarrow \mathrm{H_2O} + \mathrm{NaCl} + \mathrm{CO_2}$$

we have

$$\mathrm{C} + \mathrm{O_2} + \mathrm{NaOH} + \mathrm{HCl} \rightarrow \mathrm{CO_2} + \mathrm{H_2O} + \mathrm{NaCl}$$

# 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

$$\multimap : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{C}$$

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 $A \otimes B \cong B \otimes A$.

# Examples

- 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.
- 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 $(\mathbf{C}, \otimes, I, \multimap, \perp)$ is a symmetric monoidal closed category $\mathbf{C}$ with a *dualizing* object $\perp$, such that if we set $A^* := A \multimap \perp$, the canonical morphism $A \to A^{**}$ is an isomorphism.

**Remark:** The operation $(-)^*$ induces a contravariant dualizing functor

$$* : \mathbf{C} \to \mathbf{C}^{\mathrm{op}}$$

such that

- $\mathrm{Hom}(A, B) \cong \mathrm{Hom}(B^*, A^*)$
- $\mathrm{Hom}(A \otimes B, C) \cong \mathrm{Hom}(A, B \multimap C)$
- $A \multimap B \cong (A \otimes B^*)^*$
- $I \cong \perp^*$

**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[27][Gir11]

$$A ::= p \mid 0 \mid 1 \mid \bot \mid \top \mid A^{\bot} \mid A \oplus A \mid A \,\&\, A \mid A \otimes A \mid A \parr A \mid A \multimap A \mid !A \mid ?A$$

| | | | |
|---|---|---|---|
| $\oplus$ | 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 |

[27] Philip Wadler: A Taste of Linear Logic.

# Negation

| | |
|---|---|
| $(p)^\perp = p^\perp$ | $(p^\perp)^\perp = p$ |
| $(A \otimes B)^\perp = A^\perp \parr B^\perp$ | $(A \parr B)^\perp = A^\perp \otimes B^\perp$ |
| $(A \oplus B)^\perp = A^\perp \,\&\, B^\perp$ | $(A \,\&\, B)^\perp = A^\perp \oplus B^\perp$ |
| $(1)^\perp = \perp$ | $(\perp)^\perp = 1$ |
| $(0)^\perp = \top$ | $(\top)^\perp = 0$ |
| $(!A)^\perp = ?(A^\perp)$ | $(?A)^\perp = !(A^\perp)$ |

# Linear Logic — A Logic of Resources

- $A \multimap B \coloneqq A^{\perp} \parr B$: consume $A$ to produce $B$.
- $A \parr B$: 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 | $A, A \to B \vdash A \wedge B$ | $A, A \multimap B \nvdash A \otimes B$ | $!A \vdash A \otimes !A$ |
| 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 $\qquad \dfrac{}{\Gamma \vdash \Delta, \top}\ \top R$

$$\dfrac{}{\Gamma, 0 \vdash \Delta}\ 0L \qquad \text{no } 0R \text{ rule}$$

$$\dfrac{\Gamma \vdash \Delta}{\Gamma, 1 \vdash \Delta}\ 1L \qquad \dfrac{}{\vdash 1}\ 1R$$

$$\dfrac{}{\bot \vdash}\ \bot L \qquad \dfrac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \bot}\ \bot R$$

$$\dfrac{\Gamma \vdash \Delta, A}{\Gamma, A^{\perp} \vdash \Delta}\ {}^{\perp}L \qquad \dfrac{\Gamma, A \vdash \Delta}{\Gamma \vdash \Delta, A^{\perp}}\ {}^{\perp}R$$

$$\dfrac{\Gamma, !A, !A \vdash \Delta}{\Gamma, !A \vdash \Delta}\ !C \qquad \dfrac{\Gamma \vdash \Delta}{\Gamma, !A \vdash \Delta}\ !W \qquad \dfrac{\Gamma, A \vdash \Delta}{\Gamma, !A \vdash \Delta}\ !D \qquad \dfrac{!\Gamma \vdash ?\Delta, A}{!\Gamma \vdash ?\Delta, !A}\ !P$$

$$\dfrac{\Gamma \vdash \Delta, ?A, ?A}{\Gamma \vdash \Delta, ?A}\ ?C \qquad \dfrac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, ?A}\ ?W \qquad \dfrac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, ?A}\ ?D \qquad \dfrac{!\Gamma, A \vdash ?\Delta}{!\Gamma, ?A \vdash ?\Delta}\ ?P$$

# 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

$$
\begin{aligned}
A \to B &= !A \multimap B \\
A \wedge B &= A \,\&\, B \\
A \vee B &= !A \oplus !B \\
\top &= \top \\
\bot &= 0 \\
\neg A &= !A \multimap 0 \\
\Gamma \vdash A &= !\Gamma \vdash A
\end{aligned}
$$

**Remark:** an alternative embedding:

$$A \wedge B = !A \otimes !B$$

$$\begin{tikzcd}
\boxed{\begin{array}{c}\text{Cartesian closed category}\\ \text{Non-linear functions}\\ !A \multimap B,\ \&\ ,\oplus\end{array}} \arrow[r, bend left, "\mathrm{U}"] & \boxed{\begin{array}{c}\text{Monoidal closed category}\\ \text{Linear functions}\\ A \multimap B, \otimes,\ \parr\end{array}} \arrow[l, bend left, "!"]
\end{tikzcd}$$

$$\begin{aligned}
A \multimap B &\equiv A^{\perp} \parr B \equiv (A \otimes B^{\perp})^{\perp} \\
A \otimes B &\equiv (A \multimap B^{\perp})^{\perp} \\
A \parr B &\equiv (A^{\perp} \otimes B^{\perp})^{\perp} \equiv A^{\perp} \multimap B \\
A \,\&\, B &\equiv (A^{\perp} \oplus B^{\perp})^{\perp} \\
!(A \,\&\, B) &\equiv !A \otimes !B \\
?(A \oplus B) &\equiv ?A \parr ?B
\end{aligned}$$

- $\&$ : product
- $\oplus$: coproduct
- $\otimes$: tensor product

| | $\otimes$ | $\parr$ | $\multimap$ |
|---|---|---|---|
| $\otimes$ | | $A \otimes B \equiv (A^\perp \parr B^\perp)^\perp$ | $A \otimes B \equiv (A \multimap B^\perp)^\perp$ |
| $\parr$ | $A \parr B \equiv (A^\perp \otimes B^\perp)^\perp$ | | $A \parr B \equiv A^\perp \multimap B$ |
| $\multimap$ | $A \multimap B \equiv (A \otimes B^\perp)^\perp$ | $A \multimap B \equiv A^\perp \parr B$ | |

| | |
|---|---|
| $\dfrac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}$ | $\dfrac{f : \Gamma \otimes A \otimes B \otimes \Delta \longrightarrow C}{f \circ (1_\Gamma \otimes \sigma_{B,A} \otimes 1_\Delta) : \Gamma \otimes B \otimes A \otimes \Delta \longrightarrow C}$ |
| $\dfrac{}{A \vdash A}$ | $\dfrac{}{1_A : A \longrightarrow A}$ |
| $\dfrac{\Gamma \vdash A \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}$ | $\dfrac{f : \Gamma \longrightarrow A \quad g : A \otimes \Delta \longrightarrow B}{g \circ (f \otimes 1_\Delta) : \Gamma \otimes \Delta \longrightarrow B}$ |
| $\dfrac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \otimes B}$ <br><br> $\dfrac{\Gamma, A, B \vdash C}{\Gamma, A \otimes B \vdash C}$ | $\dfrac{f : \Gamma \longrightarrow A \quad g : \Delta \longrightarrow B}{f \otimes g : \Gamma \otimes \Delta \longrightarrow A \otimes B}$ <br><br> $\dfrac{f : (\Gamma \otimes A) \otimes B \longrightarrow C}{f \circ \alpha_{\Gamma,A,B} : \Gamma \otimes (A \otimes B) \longrightarrow C}$ |
| $\dfrac{\Gamma, A \vdash B}{\Gamma \vdash A \multimap B}$ <br><br> $\dfrac{\Gamma \vdash A \multimap B \quad \Delta \vdash A}{\Gamma, \Delta \vdash B}$ | $\dfrac{f : \Gamma \otimes A \longrightarrow B}{\hat{f} : \Gamma \longrightarrow (A \multimap B)}$ <br><br> $\dfrac{f : \Gamma \longrightarrow (A \multimap B) \quad g : \Delta \longrightarrow A}{\varepsilon_{A,B} \circ (f \otimes g) : \Gamma \otimes \Delta \longrightarrow B}$ |

# 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 — $\mathbf{Bayes}$

**Definition**: The dagger category $\mathbf{Bayes}$ 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 $(f \otimes g)(xy \mid ab) = f(x \mid a) g(y \mid b)$.

## Definition

A morphism $f : A \to B$ in a dagger category is:

- the *adjoint* of $g : B \to A$ 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 $ff^\dagger = 1_B$;
- 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 $g : A \to B$ (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

$$\begin{tikzcd}
A \arrow[r, "\rho_A^{-1}"] \arrow[d, "1_A"'] & A \otimes I \arrow[r, "1_A \otimes \eta_A"] & A \otimes (A^* \otimes A) \arrow[d, "\alpha_{A,A^*,A}"] \\
A & I \otimes A \arrow[l, "\lambda_A"] & (A \otimes A^*) \otimes A \arrow[l, "\varepsilon_A \otimes 1_A"]
\end{tikzcd}$$

$$\begin{tikzcd}
A^* \arrow[r, "\lambda_{A^*}^{-1}"] \arrow[d, "1_{A^*}"'] & I \otimes A^* \arrow[r, "\eta_A \otimes 1_{A^*}"] & (A^* \otimes A) \otimes A^* \arrow[d, "\alpha^{-1}_{A^*,A,A^*}"] \\
A^* & A^* \otimes I \arrow[l, "\rho_{A^*}"] & A^* \otimes (A \otimes A^*) \arrow[l, "1_{A^*} \otimes \varepsilon_A"]
\end{tikzcd}$$

# 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. $\xrightarrow{A} \;=\; \xleftarrow{A^*}$ .
- 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

$$\eta_V : \mathbb{K} \to V^* \otimes V :: 1 \mapsto \sum_{i=1}^{n} \overline{e}_i \otimes e_i$$

$$\varepsilon_V : V \otimes V^* \to \mathbb{K} :: e_i \otimes \overline{e}_j \mapsto \overline{e}_j(e_i)$$

  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 $\eta_V$ and $\varepsilon_V$ 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 $\eta_V$ 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

$$\begin{tikzcd} A^* \otimes A \arrow[r, "1_{A^*} \otimes f"] & A^* \otimes B \\ I \arrow[u, "\eta_A"] \arrow[ur, "\ulcorner f \urcorner"'] & \end{tikzcd} \qquad \begin{tikzcd} & I \\ A \otimes B^* \arrow[ur, "\llcorner f \lrcorner"] \arrow[r, "f \otimes 1_{B^*}"'] & B \otimes B^* \arrow[u, "\varepsilon_B"'] \end{tikzcd}$$

**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 $(m_{ij})$ the matrix of $f$ in bases $\{e_i^V\}_{i=1}^n$ and $\{e_j^W\}_{j=1}^m$:

$$\ulcorner f \urcorner : \mathbb{K} \to V^* \otimes W :: 1 \mapsto \sum_{ij} m_{ij} \cdot \overline{e}_i^V \otimes e_j^W$$

$$\llcorner f \lrcorner : V \otimes W^* \to \mathbb{K} :: e_i^V \otimes \overline{e}_j^W \mapsto m_{ij}$$

# Name & Coname

$$I \xrightarrow{\ulcorner f \urcorner} A^* \otimes B \qquad\qquad A \otimes B^* \xrightarrow{\llcorner f \lrcorner} I$$

- 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:

$$\begin{tikzcd} I \arrow[r, "\varepsilon_A^\dagger"] \arrow[dr, "\eta_A"'] & A \otimes A^* \arrow[d, "\sigma_{A,A^*}"] \\ & A^* \otimes A \end{tikzcd}$$

## 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 $\phi : I_{\mathbf{D}} \to F I_{\mathbf{C}}$
- a natural transformation $\phi_{A,B} : FA \otimes_{\mathbf{D}} FB \to F(A \otimes_{\mathbf{C}} B)$

satisfying the following conditions:

$$\begin{tikzcd}
(FA \otimes_{\mathbf{D}} FB) \otimes_{\mathbf{D}} FC \arrow[r, "\alpha_{FA,FB,FC}"] \arrow[d, "\phi_{A,B} \otimes 1_{FC}"'] & FA \otimes_{\mathbf{D}} (FB \otimes_{\mathbf{D}} FC) \arrow[d, "1_{FA} \otimes \phi_{B,C}"] \\
F(A \otimes_{\mathbf{C}} B) \otimes_{\mathbf{D}} FC \arrow[d, "\phi_{A \otimes_{\mathbf{C}} B, C}"'] & FA \otimes_{\mathbf{D}} F(B \otimes_{\mathbf{C}} C) \arrow[d, "\phi_{A, B \otimes_{\mathbf{C}} C}"] \\
F((A \otimes_{\mathbf{C}} B) \otimes_{\mathbf{C}} C) \arrow[r, "F\alpha_{A,B,C}"'] & F(A \otimes_{\mathbf{C}} (B \otimes_{\mathbf{C}} C))
\end{tikzcd}$$

$$\begin{tikzcd}
I_{\mathbf{D}} \otimes_{\mathbf{D}} FA \arrow[r, "\phi \otimes 1_{FA}"] \arrow[d, "\lambda_{FA}"'] & FI_{\mathbf{C}} \otimes_{\mathbf{D}} FA \arrow[d, "\phi_{I_{\mathbf{C}}, A}"] \\
FA & F(I_{\mathbf{C}} \otimes_{\mathbf{C}} A) \arrow[l, "F\lambda_A"]
\end{tikzcd}
\qquad
\begin{tikzcd}
FA \otimes_{\mathbf{D}} I_{\mathbf{D}} \arrow[r, "1_{FA} \otimes \phi"] \arrow[d, "\rho_{FA}"'] & FA \otimes_{\mathbf{D}} FI_{\mathbf{C}} \arrow[d, "\phi_{A, I_{\mathbf{C}}}"] \\
FA & F(A \otimes_{\mathbf{C}} I_{\mathbf{C}}) \arrow[l, "F\rho_A"]
\end{tikzcd}$$

# Monoidal Functors

- If $\phi$ 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

$$\begin{tikzcd}
FA \otimes_{\mathbf{D}} FB \arrow[r, "\sigma_{FA,FB}"] \arrow[d, "\phi_{A,B}"'] & FB \otimes_{\mathbf{D}} FA \arrow[d, "\phi_{B,A}"] \\
F(A \otimes_{\mathbf{C}} B) \arrow[r, "F\sigma_{A,B}"'] & F(B \otimes_{\mathbf{C}} A)
\end{tikzcd}$$

- 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 $(F, \phi)$ and $(G, \psi)$ 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

$$\begin{tikzcd}
FA \otimes_{\mathbf{D}} FB \arrow[r, "\eta_A \otimes_{\mathbf{D}} \eta_B"] \arrow[d, "\phi_{A,B}"'] & GA \otimes_{\mathbf{D}} GB \arrow[d, "\psi_{A,B}"] \\
F(A \otimes_{\mathbf{C}} B) \arrow[r, "\eta_{A \otimes_{\mathbf{C}} B}"'] & G(A \otimes_{\mathbf{C}} B)
\end{tikzcd}
\qquad
\begin{tikzcd}
I_{\mathbf{D}} \arrow[d, "\phi"'] \arrow[dr, "\psi"] & \\
FI_{\mathbf{C}} \arrow[r, "\eta_{I_{\mathbf{C}}}"'] & GI_{\mathbf{C}}
\end{tikzcd}$$

## 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 $\mathbf{SymmMonCat}$.

# 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.

# 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: $\mathbf{Mat}$

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

$$A_1 \oplus A_2 := \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

- symmetries are permutation matrices

# Contents


Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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

$$
\begin{tikzcd}
\mathbf{C}(A,B) \otimes (\mathbf{C}(B,C) \otimes \mathbf{C}(C,D)) \arrow[rr, "\alpha"] \arrow[d, "1 \otimes \circ_{BCD}"'] & & (\mathbf{C}(A,B) \otimes \mathbf{C}(B,C)) \otimes \mathbf{C}(C,D) \arrow[d, "\circ_{ABC} \otimes 1"] \\
\mathbf{C}(A,B) \otimes \mathbf{C}(B,D) \arrow[r, "\circ_{ABD}"'] & \mathbf{C}(A,D) & \mathbf{C}(A,C) \otimes \mathbf{C}(C,D) \arrow[l, "\circ_{ACD}"] \\
\mathbf{C}(A,B) \otimes \mathbf{C}(B,B) \arrow[r, "\circ_{ABB}"] & \mathbf{C}(A,B) & \mathbf{C}(A,A) \otimes \mathbf{C}(A,B) \arrow[l, "\circ_{AAB}"'] \\
\mathbf{C}(A,B) \otimes I \arrow[u, "1 \otimes \mathrm{id}_B"] \arrow[ur, "\rho"'] & & I \otimes \mathbf{C}(A,B) \arrow[u, "\mathrm{id}_A \otimes 1"'] \arrow[ul, "\lambda"]
\end{tikzcd}
$$

## 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

$$\begin{tikzcd}
\mathbf{C}(A,B) \otimes \mathbf{C}(B,C) \arrow[r, "\circ_{A,B,C}"] \arrow[d, "F_{A,B} \otimes F_{B,C}"'] & \mathbf{C}(A,C) \arrow[d, "F_{A,C}"] \\
D(FA,FB) \otimes D(FB,FC) \arrow[r, "\circ_{FA,FB,FC}"'] & D(FA,FC)
\end{tikzcd}$$

$$\begin{tikzcd}
 & I \arrow[dl, "\mathrm{id}_A"'] \arrow[dr, "\mathrm{id}_{FA}"] & \\
\mathbf{C}(A,A) \arrow[rr, "F_{A,A}"'] & & \mathbf{D}(FA,FA)
\end{tikzcd}$$

# Examples

- A $\mathbf{Set}$-category is a locally small category.
- A $\mathbf{Vect}$-category is a linear category: hom-sets are vector spaces and multiplication is bilinear.
- A $\mathbf{Bool}$-category is a preorder.

$$x \leq y \iff \mathrm{Hom}(x, y) = \top$$

- Any preorder gives a $\mathbf{Bool}$-category.

# Self-Enrichment

- Any cartesian closed category $\mathbf{V}$ can be viewed as self-enriched. This is because every external $\mathrm{Hom}$-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

$$\mathrm{Hom}(I, [A, A]) \cong \mathrm{Hom}(I \otimes A, A)$$

# $\mathrm{Hom}$ Functors

- $\mathrm{Hom}$ Functor
$$\mathrm{Hom}(-,-) : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Set}$$
- $\mathbf{V}$-enriched $\mathrm{Hom}$ functor
$$\mathrm{Hom}(-,-) : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{V}$$
- internal $\mathrm{Hom}$ 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 $x \leq_X y := d(x, y) = 0$, then $\leq_X$ is a partial order.
- A Lawvere metric space can be viewed as a category enriched in the monoidal poset $\mathbf{Cost} := ([0, \infty], \geq, +, 0)$, where the tensor product is $\otimes = +$, and the identity is $I = 0$.

$$\mathrm{Hom}(x, y) := d(x, y) \in \mathrm{ob}(\mathbf{Cost})$$

## 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?)

$$d'(A, B) := \sup_{a \in A} \inf_{b \in B} d(a, b) \text{ for } A, B \subset X$$

then $(\mathrm{P}(X), d')$ is a Lawvere metric space.

# Enriched Categories

enriched categories

monoidal categories

# 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 $[0, \infty]$ ... |

# Contents


Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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:

$$\begin{tikzcd}
 & x \arrow[dl, "\alpha_A"'] \arrow[dr, "\alpha_B"] & \\
F(A,A) \arrow[dr, "F(1_A,f)"'] & & F(B,B) \arrow[dl, "F(f,1_B)"] \\
 & F(A,B) &
\end{tikzcd}$$

# 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 $A \in \mathbf{C} : \alpha_A = \pi_A \circ u$.

$$\begin{tikzcd} & x \arrow[dl, "\alpha_A"'] \arrow[d, dashed, "u"] \\ F(A,A) & \mathrm{end} \arrow[l, "\pi_A"] \end{tikzcd}$$

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 $F, G : \mathbf{C} \to \mathbf{D}$:

$$\mathrm{Nat}(F,G) = \int_{A \in \mathbf{C}} \mathrm{Hom}_{\mathbf{D}}(FA, GA)$$

## 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:

$$\begin{tikzcd}
 & F(A,B) \arrow[dl, "F(1_A,f)"'] \arrow[dr, "F(f,1_A)"] & \\
F(A,A) \arrow[dr, "\alpha_A"'] & & F(B,B) \arrow[dl, "\alpha_B"] \\
 & x &
\end{tikzcd}$$

## 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 $A \in \mathbf{C} : \alpha_A = u \circ \iota_A$.

$$\begin{tikzcd} F(A,A) \arrow[r, "\iota_A"] \arrow[dr, "\alpha_A"'] & \mathrm{coend} \arrow[d, dashed, "u"] \\ & x \end{tikzcd}$$

We usually write $\mathrm{coend}(F) = \displaystyle\int^{A \in \mathbf{C}} F(A,A)$ or just $\displaystyle\int^{\mathbf{C}} F$

## 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* $\mathrm{Hom}$*-functors* $\mathrm{Hom} : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Set}$.
- *Profunctors are generalised relations.*
  *A relation is a* $\mathbf{V} := \{0 < 1\}$*-enriched profunctor.*
- *Profunctors* $F : \mathbf{C} \nrightarrow \mathbf{D}$ *correspond to functors* $\hat{F} : \mathbf{D} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}$.
- *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* $\overline{F} : \mathbf{C} \nrightarrow \mathbf{D}$.

# Profunctor

- 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

$$F : \mathbf{X}^{\mathrm{op}} \times \mathbf{Y} \to \mathbf{Bool}$$

- Given $x \in X, y \in Y$, if $F(x, y) = \top$ we say $x$ *can be obtained given* $y$.
- The requirement that $F$ is monotone says that

$$x' \leq_X x \ \ \& \ \ y \leq_Y y' \implies F(x, y) \leq_{\mathbf{Bool}} F(x', y')$$

- 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 $y'$.

# 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, $(G \circ F)(a, q) = 24$

# 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

$$\begin{aligned}
& \mathrm{Hom}_{\mathbf{Set}}\left(\int^{X \in \mathbf{C}} \mathrm{Hom}(X, A) \times FX, Y\right) \\
\cong\ & \int_{X \in \mathbf{C}} \mathrm{Hom}_{\mathbf{Set}}(\mathrm{Hom}(X, A) \times FX, Y) \\
\cong\ & \int_{X \in \mathbf{C}} \mathrm{Hom}_{\mathbf{Set}}(\mathrm{Hom}(X, A), \mathrm{Hom}_{\mathbf{Set}}(FX, Y)) \\
\cong\ & \mathrm{Nat}(\mathrm{Hom}(-, A), \mathrm{Hom}_{\mathbf{Set}}(F-, Y)) \\
\cong\ & \mathrm{Hom}_{\mathbf{Set}}(FA, Y) && \text{(Yoneda Lemma)}
\end{aligned}$$

The $\mathrm{Hom}$-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:

$$\int_{x \in \mathbb{R}} f(x)\delta(x-a)\,dx = f(a)$$

$$\int^{X \in \mathbf{C}} FX \times \mathrm{Hom}(X,A) \cong FA$$

## 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# Monad

## Definition (Monad)

A *monad* $(T, \eta, \mu)$ 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 $T \to T$).

$$\begin{tikzcd} T^3 \arrow[r, "\mu T"] \arrow[d, "T\mu"'] & T^2 \arrow[d, "\mu"] \\ T^2 \arrow[r, "\mu"'] & T \end{tikzcd} \qquad \begin{tikzcd} T \arrow[r, "\eta T"] \arrow[d, "T\eta"'] \arrow[dr, equal] & T^2 \arrow[d, "\mu"] \\ T^2 \arrow[r, "\mu"'] & T \end{tikzcd}$$

# Monad

The first axiom is akin to the associativity in monoids if we think of $\mu$ 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.

$$\begin{tikzcd} T^3 \arrow[r, "\mu T"] \arrow[d, "T\mu"'] & T^2 \arrow[d, "\mu"] \\ T^2 \arrow[r, "\mu"'] & T \end{tikzcd} \qquad \begin{tikzcd} T \arrow[r, "\eta T"] \arrow[d, "T\eta"'] \arrow[dr, equal] & T^2 \arrow[d, "\mu"] \\ T^2 \arrow[r, "\mu"'] & T \end{tikzcd}$$

$$\begin{tikzcd} TTTX \arrow[r, "\mu_{TX}"] \arrow[d, "T\mu_X"'] & TTX \arrow[d, "\mu_X"] \\ TTX \arrow[r, "\mu_X"'] & TX \end{tikzcd} \qquad \begin{tikzcd} TX \arrow[r, "\eta_{TX}"] \arrow[d, "T\eta_X"'] \arrow[dr, equal] & TTX \arrow[d, "\mu_X"] \\ TTX \arrow[r, "\mu_X"'] & TX \end{tikzcd}$$

A *comonad* for a category $\mathbf{C}$ is a monad for the opposite category $\mathbf{C}^{\mathrm{op}}$.

# Comonad

## Definition (Comonad)

A *comonad* $(W, \varepsilon, \delta)$ 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 $W \to W$).

$$\begin{tikzcd} W \arrow[r, "\delta"] \arrow[d, "\delta"'] & W^2 \arrow[d, "\delta W"] \\ W^2 \arrow[r, "W\delta"'] & W^3 \end{tikzcd} \qquad \begin{tikzcd} W \arrow[r, "\delta"] \arrow[d, "\delta"'] \arrow[dr, equal] & W^2 \arrow[d, "\varepsilon W"] \\ W^2 \arrow[r, "W\varepsilon"'] & W \end{tikzcd}$$

# 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 $\eta_A(a) := \{a\}$.
  - $\mu_A : \mathrm{P}(\mathrm{P}(A)) \to \mathrm{P}(A)$ given by $\mu_A(B) := \bigcup B$.
- Consider $\mathbf{Set}$ and a monoid $(M, e, \cdot)$, the functor $M \times - : \mathbf{Set} \to \mathbf{Set}$ has a monad structure with
  - $T : A \mapsto M \times A$, 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 $(T, \eta, \mu)$, the natural transformations $\eta$ and $\mu$ 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[28].

[28] 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

- $\mathrm{list} : \mathbf{Set} \to \mathbf{Set}$ which maps any set $A$ to the set of all finite first-order formal expressions built from elements of $A$
- for $f : A \to B$, $\mathrm{list}\, f : \mathrm{list}\, A \to \mathrm{list}\, B$ acts on expressions like so

$$\mathrm{list}\, f : [x_1] * \cdots * [x_n] \mapsto [f x_1] * \cdots * [f x_n]$$

- $\eta_A : A \to \mathrm{list}(A)$

$$x \mapsto [x]$$

- $\mu_A : \mathrm{list}(\mathrm{list}(A)) \to \mathrm{list}(A)$

$$[[x_1]*\cdots*[x_m]]*\cdots*[[y_1]*\cdots*[y_n]] \mapsto [x_1]*\cdots*[x_m]*\cdots*[y_1]*\cdots*[y_n]$$

$(\mathrm{list}, \eta, \mu)$ is a monad.
$(\mathrm{list}(A), *, [])$ is a monoid.

# Monads from Adjunctions

## Monads from Adjunctions

Any adjunction $\mathbf{C} \underset{G}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{D}$ gives rise to a monad on $\mathbf{C}$, with

- the endofuntor $T := GF$,
- the unit natural transformation $\eta : 1_{\mathbf{C}} \to GF$,
- the multiplication natural transformation $\mu := G\varepsilon F : GFGF \to GF$.

$$
\begin{tikzcd}
 & & \mathbf{C} \arrow[dl, "F"'] & & \\
\mathbf{C} & \mathbf{D} \arrow[l, "G"'] & & \mathbf{D} \arrow[ul, "G"'] \arrow[ll, "FG"'] \arrow[d, phantom, "\Downarrow\varepsilon"] & \mathbf{C} \arrow[l, "F"'] \arrow[llll, bend left, "T=GF=G1_{\mathbf{D}}F"]
\end{tikzcd}
$$

- $(GF, \eta, G\varepsilon F)$ 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 $(A, \alpha)$ where $A \in \mathbf{C}$ and $\alpha : TA \to A$ s.t.

$$\begin{tikzcd} TTA \arrow[r, "T\alpha"] \arrow[d, "\mu A"'] & TA \arrow[d, "\alpha"] \\ TA \arrow[r, "\alpha"'] & A \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "\eta A"] \arrow[dr, "1_A"'] & TA \arrow[d, "\alpha"] \\ & A \end{tikzcd}$$

A homomorphism of $T$-algebras $f : (A, \alpha) \to (B, \beta)$ is $f : A \to B$ s.t.

$$\begin{tikzcd} TA \arrow[r, "Tf"] \arrow[d, "\alpha"'] & TB \arrow[d, "\beta"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

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 $(T, \eta, \mu)$ 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* $\mathbf{C}_T \to \mathbf{C}^T$.

# NNO — Initial $T$-Algebra

- An initial $T$-algebra is an $T$-algebra $(I, \iota)$ such that given any other $T$-algebra $(A, \alpha)$ there exists a unique $T$-algebra homomorphism $f : (I, \iota) \to (A, \alpha)$.
- A natural numbers object NNO is an initial $T$-algebra for the endofunctor $TA = 1 + A$.

$$\begin{tikzcd} 1 + N \arrow[r, "Tf"] \arrow[d, "\iota"'] & 1 + A \arrow[d, "\alpha"] \\ N \arrow[r, "f"'] & A \end{tikzcd}$$

$$\iota = \begin{bmatrix} 0 \\ s \end{bmatrix} : 1 + N \to N.$$

# Initial $T$-Algebra and Fixed Points

If we lift the initial algebra $(I, \iota)$, we get a new algebra $(TI, T\iota)$. And there must be a unique morphism $f : (I, \iota) \to (TI, T\iota)$.
The morphism $f$ is the inverse of $\iota$.

$$\begin{tikzcd}
TI \arrow[r, "Tf"] \arrow[d, "\iota"'] & TTI \arrow[r, "T\iota"] \arrow[d, "T\iota"] & TI \arrow[d, "\iota"] \\
I \arrow[r, dashed, "f"'] & TI \arrow[r, "\iota"'] & I
\end{tikzcd}$$

since

$$\iota \circ f = 1_I$$

$$f \circ \iota = T\iota \circ Tf = T(\iota \circ f) = T(1_I) = 1_{TI}$$

which means that $\iota$ is an isomorphism.

$$TI \cong I$$

$I$ is a fixed point of $T$.

# Algebra/Coalgebra for an Endofunctor

- For a category $\mathbf{C}$ and endofunctor $F$, an algebra of $F$ is an object $A$ in $\mathbf{C}$ and a morphism $\alpha : FA \to A$.
- For a category $\mathbf{C}$ and endofunctor $F$, a coalgebra of $F$ is an object $A$ in $\mathbf{C}$ and a morphism $\alpha : A \to FA$.
- An initial algebra for an endofunctor $F$ on a category $\mathbf{C}$ is an initial object in the category of algebras of $F$.
- A terminal coalgebra for an endofunctor $F$ on a category $\mathbf{C}$ is a terminal object in the category of coalgebras of $F$.
- **Theorem:** The initial algebra $(\mu F, \iota)$ is the least fixed point of $F$ in that there is a map from $\mu F$ to any other fixed point.
- **Theorem:** The terminal algebra $(\nu F, \tau)$ is the greatest fixed point of $F$ in that there is a map to $\nu F$ from any other fixed point.

## Theorem

*Every monad $T = (T, \eta, \mu)$ arises from an adjunction* $\mathbf{C} \underset{U}{\overset{F}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{C}^T$ .

*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 $(A, \mu_A)$ 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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 $s \cup t \in C(U \cup V,\mathbb{R})$:

$$\begin{tikzcd} U \cap V \arrow[r, hook] \arrow[d, hook] & V \arrow[d, hook] \arrow[ddr, bend left, "t"] & \\ U \arrow[r, hook] \arrow[drr, bend right, "s"'] \arrow[dr, phantom, "\ulcorner"] & U \cup V \arrow[dr, dashed, "s\cup t"] & \\ & & \mathbb{R} \end{tikzcd}$$

# Sheaf — Categorical Reformulations

- A presheaf $F : \mathcal{O}_X^{\mathrm{op}} \to \mathbf{Set}$ is a *sheaf* iff for any open cover $(U_i)_{i \in I}$ of $U$, the following diagram is an equalizer:

$$\begin{tikzcd} F(U) \arrow[r, tail, "\mathrm{res}^U_{U_i}"] & \prod_i F(U_i) \arrow[r, shift left, "\mathrm{res}^{U_i}_{U_i \cap U_j}"] \arrow[r, shift right, "\mathrm{res}^{U_j}_{U_i \cap U_j}"'] & \prod_{i,j} F(U_i \cap U_j) \end{tikzcd}$$

Where

$$\mathrm{res}^U_{U_i} : s \mapsto (s \restriction_{U_i})_{i \in I}$$

$$\mathrm{res}^{U_i}_{U_i \cap U_j} : (s_i)_{i \in I} \mapsto (s_i \restriction_{U_i \cap U_j})_{(i,j) \in I \times I}$$

$$\mathrm{res}^{U_j}_{U_i \cap U_j} : (s_i)_{i \in I} \mapsto (s_j \restriction_{U_i \cap U_j})_{(i,j) \in I \times I}$$

$$\begin{tikzcd} & F(U_i) \arrow[r, "F(U_i \cap U_j \hookrightarrow U_i)"] & F(U_i \cap U_j) \\ F(U) \arrow[ur, "F(U_i \hookrightarrow U)"] \arrow[r, tail, "\mathrm{res}^U_{U_i}"] \arrow[dr, "F(U_j \hookrightarrow U)"'] & \prod_i F(U_i) \arrow[u, "\pi_i"'] \arrow[d, "\pi_j"] \arrow[r, shift left, "\mathrm{res}^{U_i}_{U_i \cap U_j}"] \arrow[r, shift right, "\mathrm{res}^{U_j}_{U_i \cap U_j}"'] & \prod_{i,j} F(U_i \cap U_j) \arrow[u, "\pi_{ij}"'] \arrow[d, "\pi_{ij}"] \\ & F(U_j) \arrow[r, "F(U_i \cap U_j \hookrightarrow U_j)"'] & F(U_i \cap U_j) \end{tikzcd}$$

# Sheaf — Categorical Reformulations

- Let $\mathbf{I}$ be a category with $\mathrm{ob}(\mathbf{I}) := \{U_i : i \in I\} \cup \{U_i \cap U_j : i, j \in I\}$, 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 $\mathbf{I}$.

$$\begin{tikzcd}
F(U_i) \arrow[r] & F(U_i \cap U_j) & F(U_j) \arrow[l] \\
& F(U) \arrow[ul] \arrow[u] \arrow[ur] &
\end{tikzcd}$$

A presheaf $F : \mathcal{O}_X^{\mathrm{op}} \to \mathbf{Set}$ is a *sheaf* iff $\varprojlim_{\mathbf{I}} F \cong F(U)$.

# Sheaf — Example

- For a fixed space $Y$, define the presheaf $\Gamma_p$ on an arbitrary subset $U \subset X$ by:

$$\Gamma_p(U) = \left\{ s : U \to Y : p \circ s(u) = u \text{ for all } u \in U \right\}$$

- 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 $0 \leq k \leq \infty$, there is a sheaf $U \mapsto C^k(U, \mathbb{R})$ of $k$-times continuously differentiable real-valued functions on $X$.
2. For any fixed space $Y$, there is a sheaf $U \mapsto C(U, Y)$ of continuous functions $U \to Y$.
3. There is a sheaf $U \mapsto \{V \in \mathcal{O} : V \subset U\}$.
4. For any fixed open subset $V \subset X$, there is a sheaf $\mathcal{O}(-, V)$ defined by

$$U \mapsto \mathcal{O}(U, V) = \begin{cases} * & \text{if } U \subset V \\ \emptyset & \text{if } U \not\subset V \end{cases}$$

5. For any fixed continuous function $p : Y \to X$, there is a sheaf $\Gamma_p$ defined by

$$U \mapsto \Gamma_p U = \left\{ \begin{tikzcd} & Y \arrow[d, "p"] \\ U \arrow[ur, dashed, "s"] \arrow[r, hook] & X \end{tikzcd} \right\}$$

whose elements are local sections, continuous functions $s : U \to Y$ so that $p \circ s = i : U \hookrightarrow X$.

# Sheaf — Example

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.

# 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,
$$F_x := \left\{ \operatorname{germ}_x s : x \in U, s \in FU \right\}$$
where
$$\operatorname{germ}_x s := [s]_x$$
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, $F_x := \varinjlim_{U \ni x} FU$.

The function $\operatorname{germ}_x : FU \to F_x$ form a cocone over $F\restriction_{\{U \in \mathcal{O}_X : x \in U\}}$.

$$\begin{tikzcd}
FU \arrow[dd] \arrow[dr, "\operatorname{germ}_x s"] \arrow[drr, "c_s"] & & \\
& F_x \arrow[r, dashed, "u"] & C \\
FW \arrow[ur, "\operatorname{germ}_x s\restriction_W"] \arrow[urr, "c_{s\restriction_W}"'] & &
\end{tikzcd}$$

# Etale Bundle

- Consider the slice category $\mathbf{Top}/X$. An *étale bundle* over $X$ is $(Y, f : Y \to X)$ in $\mathbf{Top}/X$ such that $f$ is a local homeomorphism: that is, for every $y \in Y$, there is an open set $U \ni y$ such that $f(U)$ is open in $X$ and $f\restriction_U : U \to f(U)$ is a homeomorphism.
- The category $\mathbf{Et}(X)$ is the category of étale bundles over $X$ and morphisms are the continous functions between them.
- $\widehat{\mathcal{O}_X} \underset{\Gamma}{\overset{\Lambda}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{Top}/X$

  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, $\eta_F$ is an isomorphism, while if $Y$ is étale, $\varepsilon_Y$ is an isomorphism. $\mathbf{Sh}(X) \underset{\longleftarrow}{\overset{\longrightarrow}{\cong}} \mathbf{Et}(X)$

## Proof Sketch.

For $U \in \mathcal{O}_X$,

$$\eta_{FU} : FU \to \Gamma\Lambda FU :: s \mapsto \dot{s}$$

where

$$\dot{s} : U \to \Lambda F :: x \mapsto \mathrm{germ}_x\, s$$

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.

$$\varepsilon_Y : \Lambda\Gamma Y \to Y :: \dot{s}x \mapsto sx$$

$$\begin{tikzcd} \Lambda \arrow[r, "\Lambda\eta"] \arrow[dr, "1_\Lambda"'] & \Lambda\Gamma\Lambda \arrow[d, "\varepsilon\Lambda"] \\ & \Lambda \end{tikzcd} \qquad \begin{tikzcd} \Gamma \arrow[r, "\eta\Gamma"] \arrow[dr, "1_\Gamma"'] & \Gamma\Lambda\Gamma \arrow[d, "\Gamma\varepsilon"] \\ & \Gamma \end{tikzcd}$$

$$\mathrm{germ}_x\, s \mapsto \mathrm{germ}_x\, \dot{s} \mapsto \mathrm{germ}_x\, s \qquad s \mapsto \dot{s} \mapsto s$$

# 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.

$$\begin{tikzcd} \mathbf{Sh}(X) \arrow[r, hook, shift right, "i"'] \arrow[r, phantom, "\perp"] & \widehat{\mathcal{O}_X} \arrow[l, shift right, "\Gamma\Lambda"'] \end{tikzcd}$$

# 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 closed under pre-composition, i.e. if $f \in S$ and $f \circ g$ is defined, then $f \circ g \in S$.

$$\begin{tikzcd} Z \arrow[r, "g"] \arrow[rr, bend right, "f \circ g"'] & Y \arrow[r, "f"] & X \end{tikzcd}$$

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 $f : Y \to X$, 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*

## Definition (Grothendieck Topology)

A *Grothendieck topology* on a small category $\mathbf{C}$ is a function $J$ which assigns to each object $X \in \operatorname{ob}\mathbf{C}$ 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 $f^*S \in J(Y)$.

$$\begin{tikzcd} f^*S \arrow[r] \arrow[d, mapsto] \arrow[dr, phantom, "\lrcorner", very near start] & S \arrow[d, mapsto] \\ \mathrm{y}Y \arrow[r, "\mathrm{y}f"'] & \mathrm{y}X \end{tikzcd}$$

- (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 $T \in J(X)$.

## 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 $U$.*

## 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 $f_i$.

$$\begin{tikzcd} V_j \arrow[r] \arrow[d, "h_j"'] & U_i \arrow[d, "f_i"] \\ V \arrow[r, "g"'] & U \end{tikzcd}$$

- 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 $\{f_i \circ g_{ij} : U_{ij} \to U\}_{i \in I, j \in I_i} \in K(U)$.

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 $X \in \operatorname{ob} \mathbf{C}$ 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

$$i_S : \mathrm{Hom}(S, P) \cong \mathrm{Hom}(\mathrm{Hom}(-, X), P)$$

$$\begin{tikzcd} S \arrow[r, tail, "i_S"] \arrow[d, "x"'] & \mathrm{Hom}(-, X) \arrow[dl, dashed] \\ P & \end{tikzcd}$$

- 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*

$$\begin{tikzcd} \mathbf{E} \arrow[r, hook, shift right, "j_*"'] \arrow[r, phantom, "\perp"] & \widehat{\mathbf{C}} \arrow[l, shift right, "j^*"'] \end{tikzcd}$$

*such that*

- $j^* \dashv j_*$
- $j_*$ *is full and faithful*
- $j^*$ *preserves finite limits*

# 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 $(f^*, f_*)$ :

$$\mathbf{Set}^{\mathbf{C}^{\mathrm{op}}} \underset{f_*}{\overset{f^*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} \mathbf{Set}$$

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

$$\begin{tikzcd} \mathbf{E}_X \arrow[r, shift right, "x_*"'] \arrow[r, phantom, "\perp"] & \mathbf{Set} \arrow[l, shift right, "x^*"'] \end{tikzcd}$$

- More generally, any continuous function $Y \xrightarrow{f} X$ gives a point

$$\begin{tikzcd} \mathbf{E}_X \arrow[r, shift right, "f_*"'] \arrow[r, phantom, "\perp"] & \mathbf{E}_Y \arrow[l, shift right, "f^*"'] \end{tikzcd}$$

- If $Y = \{\bullet\}$, then a point of $X$ in the usual sense gives a point of $\mathbf{E}_X$.

# 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 $f : X \to Y$ such that $f^{-1}(U)$ is open for any open set $U \subset Y$

## Point-set 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 $f : \mathbf{E} \to \mathbf{F}$ such that $f^* : \mathbf{F} \to \mathbf{E}$ 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: $j \circ \wedge = \wedge \circ (j \times j)$.

$$\begin{tikzcd} 1 \arrow[r, "\mathrm{true}"] \arrow[dr, "\mathrm{true}"'] & \Omega \arrow[d, "j"] \\ & \Omega \end{tikzcd} \qquad \begin{tikzcd} \Omega \arrow[r, "j"] \arrow[dr, "j"'] & \Omega \arrow[d, "j"] \\ & \Omega \end{tikzcd} \qquad \begin{tikzcd} \Omega \times \Omega \arrow[r, "\wedge"] \arrow[d, "j \times j"'] & \Omega \arrow[d, "j"] \\ \Omega \times \Omega \arrow[r, "\wedge"'] & \Omega \end{tikzcd}$$

## 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 $\overline{(\cdot)} : \mathrm{Sub}(X) \to \mathrm{Sub}(X)$ such that

1. $A \leq \overline{A}$
2. $\overline{A} = \overline{\overline{A}}$
3. $A \leq B \implies \overline{A} \leq \overline{B}$
4. for $f : Y \to X$ and $A \in \mathrm{Sub}(X)$, $f^*(\overline{A}) = \overline{f^*(A)}$

## 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.*

$$\begin{tikzcd} \mathbf{E} \arrow[r, hook, shift right, "i"'] \arrow[r, phantom, "\perp"] & \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}} \arrow[l, shift right, "F"'] \end{tikzcd}$$

## 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 $\overline{M} = X$.
- *colosed* if $\overline{M} = M$.

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.

$$\begin{tikzcd} M \arrow[r, tail, "dense"] \arrow[d] & X \arrow[dl, dashed, "!"] \\ F & \end{tikzcd}$$

$F$ 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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 $\mathbf{2} \to \mathbf{2}$.

$$\begin{tikzcd} \mathbf{2} \arrow[rr, "0 \circ !_{\mathbf{2}}"] \arrow[dr, "!_{\mathbf{2}}"'] & & \mathbf{2} \\ & \mathbf{1} \arrow[ur, "0"'] & \end{tikzcd} \qquad \begin{tikzcd} \mathbf{2} \arrow[r, "1_{\mathbf{2}}"] & \mathbf{2} \end{tikzcd} \qquad \begin{tikzcd} \mathbf{2} \arrow[rr, "1 \circ !_{\mathbf{2}}"] \arrow[dr, "!_{\mathbf{2}}"'] & & \mathbf{2} \\ & \mathbf{1} \arrow[ur, "1"'] & \end{tikzcd}$$

- $\mathbf{2}$ is a universal generator: $\forall F \neq G : \mathbf{A} \to \mathbf{B} \exists f : \mathbf{2} \to \mathbf{A}[Ff \neq Gf]$

$$\begin{tikzcd} \mathbf{2} \arrow[r, "f"] & \mathbf{A} \arrow[r, shift left, "F"] \arrow[r, shift right, "G"'] & \mathbf{B} \end{tikzcd}$$

if $\mathbf{C}$ has the same property, then $\exists g : \mathbf{2} \to \mathbf{C} \exists h : \mathbf{C} \to \mathbf{2} : hg = 1_{\mathbf{2}}$.

$$\begin{tikzcd} \mathbf{2} \arrow[r, shift left, "g"] & \mathbf{C} \arrow[l, shift left, "h"] \end{tikzcd}$$

# CCAF

**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 $1 : \mathbf{1} \to \mathbf{2}$.

$$\begin{tikzcd} \mathbf{1} \arrow[rr, "\mathrm{dom}\, f"] \arrow[dr, "0"'] & & \mathbf{A} \\ & \mathbf{2} \arrow[ur, "f"'] & \end{tikzcd} \qquad \begin{tikzcd} \mathbf{1} \arrow[rr, "\mathrm{cod}\, f"] \arrow[dr, "1"'] & & \mathbf{A} \\ & \mathbf{2} \arrow[ur, "f"'] & \end{tikzcd}$$

- The identity morphism $1_A$ on $A$ is defined to be the composite $A \circ !_{\mathbf{2}}$

$$\begin{tikzcd} \mathbf{2} \arrow[rr, "1_A"] \arrow[dr, "!_{\mathbf{2}}"'] & & \mathbf{A} \\ & \mathbf{1} \arrow[ur, "A"'] & \end{tikzcd}$$

- 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 $A : \mathbf{1} \to \mathbf{C}$.

# CCAF

- 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 $\operatorname{cod}\gamma = \operatorname{cod}\beta$.

$$\begin{tikzcd}
 & \mathbf{2} \arrow[dr, "\alpha"] \arrow[drr, bend left, "f"] & & \\
\mathbf{1} \arrow[ur, "1"] \arrow[dr, "0"'] & & \mathbf{3} \arrow[r, dashed, "t"] & \mathbf{A} \\
 & \mathbf{2} \arrow[ur, "\beta"'] \arrow[urr, bend right, "g"'] & &
\end{tikzcd}$$

**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:

$$\begin{tikzcd}
0 \arrow[loop above]
\end{tikzcd}
\qquad
\begin{tikzcd}
0 \arrow[loop above] \arrow[r, "\alpha"] & 1 \arrow[loop above]
\end{tikzcd}
\qquad
\begin{tikzcd}
0 \arrow[loop above] \arrow[rr, "\gamma"] \arrow[dr, "\alpha"'] & & 2 \arrow[loop above] \\
 & 1 \arrow[ur, "\beta"'] \arrow[loop below] &
\end{tikzcd}$$

# 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 $\mathbf{A} + \mathbf{B}$.

$$\begin{tikzcd}
 & \mathbf{A} \\
\mathbf{C} \arrow[ur, bend left, "F"] \arrow[r, dashed, "\langle F,G \rangle"] \arrow[dr, bend right, "G"'] & \mathbf{A} \times \mathbf{B} \arrow[u, "\pi_1"'] \arrow[d, "\pi_2"] \\
 & \mathbf{B}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{A} \arrow[dr, "\iota_1"'] \arrow[drr, bend left, "F"] & & \\
 & \mathbf{A} + \mathbf{B} \arrow[r, dashed, "\begin{bmatrix} F \\ G \end{bmatrix}"] & \mathbf{C} \\
\mathbf{B} \arrow[ur, "\iota_2"] \arrow[urr, bend right, "G"'] & &
\end{tikzcd}$$

- Every parallel pair of functors $F, G : \mathbf{A} \to \mathbf{B}$ has an equilizer and a coequalizer.

$$\begin{tikzcd}
\mathbf{C} \arrow[d, dashed, "u"'] \arrow[dr, "H"] & & \\
\mathbf{E} \arrow[r, "e"] & \mathbf{A} \arrow[r, shift left, "F"] \arrow[r, shift right, "G"'] & \mathbf{B}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{A} \arrow[r, shift left, "F"] \arrow[r, shift right, "G"'] & \mathbf{B} \arrow[r, "q"] \arrow[dr, "H"'] & \mathbf{Q} \arrow[d, dashed, "u"] \\
 & & \mathbf{C}
\end{tikzcd}$$

# CCAF

- There is a functor category $\mathbf{B}^{\mathbf{A}}$ from any category $\mathbf{A}$ to any category $\mathbf{B}$.

$$\forall \mathbf{C} \forall F : \mathbf{C} \times \mathbf{A} \to \mathbf{B}\ \exists!\ \hat{F} : \mathbf{C} \to \mathbf{B}^{\mathbf{A}}\ \left[\varepsilon \circ (\hat{F} \times 1_{\mathbf{A}}) = F\right]$$

$$\begin{tikzcd}
\mathbf{B}^{\mathbf{A}} \\
\mathbf{C} \arrow[u, dashed, "\hat{F}"]
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{B}^{\mathbf{A}} \times \mathbf{A} \arrow[r, "\varepsilon"] & \mathbf{B} \\
\mathbf{C} \times \mathbf{A} \arrow[u, dashed, "\hat{F} \times 1_{\mathbf{A}}"] \arrow[ur, "F"'] &
\end{tikzcd}$$

# 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.

$$\forall H : \mathbf{A} \to \mathbf{Q}, K : \mathbf{B} \to \mathbf{Q}[HF = KG \to \exists! u(ui_1 = H \;\&\; ui_2 = K)]$$

$$\begin{tikzcd}
\mathbf{Q} \arrow[dr, dashed, "u"] \arrow[drr, bend left, "K"] \arrow[ddr, bend right, "H"'] & & \\
& \mathbf{A} \times_{\mathbf{C}} \mathbf{B} \arrow[r, "p_2"] \arrow[d, "p_1"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbf{B} \arrow[d, "G"] \\
& \mathbf{A} \arrow[r, "F"'] & \mathbf{C}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{C} \arrow[r, "G"] \arrow[d, "F"'] & \mathbf{B} \arrow[d, "i_2"] \arrow[ddr, bend left, "K"] & \\
\mathbf{A} \arrow[r, "i_1"'] \arrow[drr, bend right, "H"'] & \mathbf{A} +_{\mathbf{C}} \mathbf{B} \arrow[ul, phantom, "\ulcorner", very near start] \arrow[dr, dashed, "u"] & \\
& & \mathbf{Q}
\end{tikzcd}$$

# CCAF

- There is a natural numbers object NNO category.

$$\begin{tikzcd}
\mathbf{1} \arrow[r, "0"] \arrow[dr, "x"'] & \mathbf{N} \arrow[r, "S"] \arrow[d, dashed, "F"] & \mathbf{N} \arrow[d, dashed, "F"] \\
& \mathbf{X} \arrow[r, "T"'] & \mathbf{X}
\end{tikzcd}$$

- 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 $\mathrm{ETCS}$ axioms.

# Contents


Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# Rosen's Simulation vs 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

# Rosen's Modeling Relation & Church-Turing Thesis

Simulation vs Model

Every natural law is simulable?

# 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 $\beta$": $\beta(M) = R$
7. Mr. Why: "why $\beta$?"
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 $\mathbf{Set}$.

# Rosen's Metabolism-Repair $(M, R)$-System[Ros72]

# Rosen: "What is Life?"[Luz+09][29]

- $M$: metabolism $M(a) = b$
- $R$: repair $R(b) = M$
- $\beta$: 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., $\beta$ is determined by $b$.

[29] 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 $M_n(a) := na \mod 12$
- $\mathrm{Hom}(B,\mathrm{Hom}(A,B)) = \{R_k : k \in A\}$ where $R_k(b) := M_{bk}$
- $\beta(M_n) := R_{11n}$
- Let us choose $a = 5, M = M_7$

$$\begin{tikzcd}
 & 5 \arrow[r] & 11 \arrow[dr, squiggly] & \\
M_7 \arrow[ur, squiggly] & & & M_7 \arrow[dl] \\
 & 11 \arrow[ul] & R_5 \arrow[l, squiggly] &
\end{tikzcd}$$

# 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 $R := G \circ \hat{\beta} \circ \langle 1_B, R' \rangle$.

$$\begin{tikzcd}
B \times \mathrm{Hom}(A, B) \arrow[r, "\hat{\beta}"] & \mathrm{Hom}(A, B) \arrow[d, "G"] \\
B \arrow[u, "\langle 1_B, R' \rangle"] \arrow[r, "R"'] & \mathrm{Hom}(A, B)
\end{tikzcd}$$

$$G(\hat{\beta}(b, M)) = G \circ \hat{\beta} \circ \langle 1_B, R' \rangle(b) = R(b) = \hat{\beta}(b, M)$$

Therefore, $\hat{\beta}(b, M)$ is a fixpoint of $G$.

# Rosen's $(M, R)$-System in terms of $\lambda$-calculus[MLS09]

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 $\beta \in \mathrm{Hom}(\mathrm{Hom}(A,B),\mathrm{Hom}(B,\mathrm{Hom}(A,B)))$

$$C_0 := A \qquad C_1 := B \qquad C_2 := \mathrm{Hom}(C_0, C_1) \qquad C_n := \mathrm{Hom}(C_{n-2}, C_{n-1})$$

$$f_0 := M \qquad f_1 := R \qquad f_2 := \beta$$

$$c_0 := a \qquad c_1 := b = f_0(c_0) \qquad c_2 := M = f_1(c_1) = f_0 \qquad c_3 := R = f_2(c_2) = f_1 \qquad c_{n+1} := f_n(c_n) = f_{n-1}$$

$$C_0 \xrightarrow{f_0} C_1 \xrightarrow{f_1} C_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} C_n \xrightarrow{f_n} C_{n+1} = \mathrm{Hom}(C_{n-1}, C_n)$$

$$c_0 \longmapsto c_1 \longmapsto c_2 \longmapsto \cdots \longmapsto c_n \longmapsto c_{n+1} = f_{n-1} = f_n(f_{n-2})$$

# Ouroboros Equation $f(f) = f$

$$C_0 \xrightarrow{f_0} C_1 \xrightarrow{f_1} C_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} C_n \xrightarrow{f_n} C_{n+1} = \mathrm{Hom}(C_{n-1}, C_n)$$

$$c_0 \longmapsto c_1 \longmapsto c_2 \longmapsto \cdots \longmapsto c_n \longmapsto c_{n+1} = f_{n-1} = f_n(f_{n-2})$$

$$C_\infty := \lim_{n\to\infty} C_n = \mathrm{Hom}(C_\infty, C_\infty)$$

$$f_\infty := \lim_{n\to\infty} f_n \implies f_\infty(f_\infty) = f_\infty$$

$$f_\infty \in \mathrm{Hom}(C_\infty, C_\infty) = C_\infty$$

$$f_n(c_n) = f_{n-1} \implies \varepsilon_{c_n}(f_n) = f_{n-1}$$

$$p_n := \varepsilon_{c_{n-1}} : C_n \leftarrow C_{n+1} :: p_{n+1}(f_n) = f_{n-1}$$

$$C_1 \xleftarrow{p_1} C_2 \xleftarrow{p_2} \cdots \xleftarrow{p_{n-1}} C_n \xleftarrow{p_n} C_{n+1} \xleftarrow{p_{n+1}} C_{n+2} \xleftarrow{p_{n+2}} \cdots$$

$$c_1 \xleftarrow{p_1} f_0 \xleftarrow{p_2} \cdots \xleftarrow{p_{n-1}} f_{n-2} \xleftarrow{p_n} f_{n-1} \xleftarrow{p_{n+1}} f_n \xleftarrow{p_{n+2}} \cdots$$

$$C^\infty := \varprojlim (C_n, p_n) = \left\{ (c_1, c_2, \dots) : c_n \in C_n \text{ and } p_n(c_{n+1}) = c_n \text{ for all } n \right\}$$

# Contents


Introduction

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

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
- Kan Extensions
- Monoidal Categories
- Enriched Categories
- End/Coend
- Monads
- Sheaves
- CCAF
- Rosen's $(M, R)$-System
- Category Theory in Machine Learning

Quantum Computing

Answers to the Exercises

References1655

# 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 $I : P \times A \to P$ 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.

| | |
|---|---|
| $I \colon P \times A \to B$ | implement |
| $U \colon P \times A \times B \to P$ | update |
| $R \colon P \times A \times B \to A$ | request |

A request function allows an update function to be defined for the composite $J(q, I(p,-))$.

# The Category of Learners $\mathbf{Learn(A,B)}$ 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: $R'(f(p), a, b) = R(p, a, b)$

$$\begin{tikzcd} P \times A \arrow[r, "I"] \arrow[d, "f\times 1_A"'] & B \\ P' \times A \arrow[ur, "I'"'] & \end{tikzcd} \qquad \begin{tikzcd} P \times A \times B \arrow[r, "(U,R)"] \arrow[d, "f\times 1_A\times 1_B"'] & P \times A \arrow[d, "f\times 1_A"] \\ P' \times A \times B \arrow[r, "(U',R')"'] & P' \times A \end{tikzcd}$$

- $\mathbf{Learn(A,B)}$ is the category with objects the learners $A \to B$, and with morphisms the maps between learners.
- **Theorem:** $\mathbf{Learn(A,B)}$ is a topos.
- The map between learners generates an equivalence relation on learners.

# The Category of Learners $\mathbf{Learn}$

$\mathbf{Learn}$ 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

- Given a parameter space $P$ and differentiable function $I : P \times \mathbb{R}^m \to \mathbb{R}^n$, we can define a learner $\mathbb{R}^m \xrightarrow{(P,I,U,R)} \mathbb{R}^n$ .
- For example, the implement map for $\theta = (w, b) \in P$ is

$$I(\theta, x) \coloneqq \sigma\left(w^\top x + b\right)$$

$$U(\theta, x, y) \coloneqq \theta - \eta \nabla_\theta \frac{1}{2}\left\|I(\theta, x) - y\right\|^2$$

$$R(\theta, x, y) \coloneqq x - \nabla_x \frac{1}{2}\left\|I(\theta, x) - y\right\|^2$$

- A deep neural network with $k$ layers is the composition of $k$ learners.

$$\mathbb{R}^{N_0} \longrightarrow \mathbb{R}^{N_1} \longrightarrow \mathbb{R}^{N_2} \longrightarrow \cdots \longrightarrow \mathbb{R}^{N_k}$$

# 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

# 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Metric Space & Topological Space

## Definition (Metric Space)

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. $d(x, z) \leq d(x, y) + d(y, z)$

## 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 (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 $\mathbb{R}^m$.

# Some (non)examples of topological spaces

Figure: Some topologies over the finite set $\{1, 2, 3\}$.

- A collection $\mathscr{B}$ 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 $\mathscr{B}$.

- A topological space is second countable if it admits a countable basis.

# 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

$$\mathrm{Tr}(A) = \sum_{e_i} \langle e_i|A|e_i\rangle \qquad \text{where } \{|e_i\rangle\} \text{ is any orthonormal basis.}$$

# 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 $\mathbb{R}$.

$$\mathrm{GL}(n, \mathbb{R}) := \left\{ M \in \mathbf{Mat}_{n \times n}(\mathbb{R}) : \det(M) \neq 0 \right\}$$

- Orthogonal group

$$\mathrm{O}(n, \mathbb{R}) := \left\{ M \in \mathbf{Mat}_{n \times n}(\mathbb{R}) : MM^{\mathsf{T}} = M^{\mathsf{T}}M = I \right\}$$

- Special linear group

$$\mathrm{SL}(n, \mathbb{R}) := \left\{ M \in \mathbf{Mat}_{n \times n}(\mathbb{R}) : \det(M) = 1 \right\}$$

- Special orthogonal group

$$\mathrm{SO}(n, \mathbb{R}) := \left\{ M \in \mathbf{Mat}_{n \times n}(\mathbb{R}) : MM^{\mathsf{T}} = M^{\mathsf{T}}M = I \text{ and } \det(M) = 1 \right\}$$

- Euclidean group

$$\mathrm{E}(n, \mathbb{R}) := \left\{ \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \in \mathbf{Mat}_{(n+1) \times (n+1)}(\mathbb{R}) : R \in \mathrm{O}(n, \mathbb{R}) \text{ and } t \in \mathbb{R}^n \right\}$$

- 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":

$$\begin{tikzcd} U \times V \arrow[r, "\varphi"] \arrow[dr, "f"'] & U \otimes V \arrow[d, dashed, "\exists! \overline{f}"] \\ & W \end{tikzcd}$$

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

$$(A \otimes B)(u \otimes v) = Au \otimes Bv$$

$W = X \otimes Y \qquad f : (u, v) \mapsto Au \otimes Bv \qquad \overline{f} = A \otimes B$

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{m1}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

# 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.

# 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:** Quantum states cannot be cloned. It's impossible to measure a qubit in two different ways (even, indirectly, by using a copy trick).

| 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

$$|\phi\rangle \otimes \psi\rangle = |\phi\rangle|\psi\rangle = |\phi\psi\rangle = \begin{bmatrix}a_0b_0\\a_0b_1\\a_1b_0\\a_1b_1\end{bmatrix}$$

- 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 $P_{x_i} : H \to L_{x_i}$ 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 $\|z\rangle$: $p(A) = \|p_A\|z\rangle\|^2$ |
| $P(A \wedge B) = P(B \wedge A)$ | $\|P_B P_A\|z\rangle\|^2 \neq \|P_A P_B\|z\rangle\|^2$ |
| $P(B \mid A) = \dfrac{P(A \wedge B)}{P(A)}$ | $P(B \mid A) = \dfrac{\|P_B P_A\|z\rangle\|^2}{\|P_A\|z\rangle\|^2}$ |
| $P(A) = \sum_i P(A \mid B_i)P(B_i)$ | Law of total probability violated |

# Quantum Gates & Universality

- A quantum gate: Unitary operation on a number of qubits.
- A set of gates is universal iff for any unitary matrix $U$ and any $\varepsilon > 0$, there is some circuit $\tilde{U}$ built out of the set of gates such that

$$\left\| U - \tilde{U} \right\| < \varepsilon$$

In other words,

$$\sup_{|\psi\rangle} \left\| \left( U - \tilde{U} \right) |\psi\rangle \right\| < \varepsilon$$

# NOT Gate & $\sqrt{\text{NOT}}$ Gate

- NOT Gate

$$\text{NOT} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\text{NOT}\,|0\rangle = |1\rangle$$
$$\text{NOT}\,|1\rangle = |0\rangle$$

- $\sqrt{\text{NOT}}$ Gate

$$\sqrt{\text{NOT}} = \frac{1}{2}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}$$

$$\sqrt{\text{NOT}} \cdot \sqrt{\text{NOT}} = \text{NOT}$$
$$\sqrt{\text{NOT}} \cdot \sqrt{\text{NOT}}^{\dagger} = I$$

# 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

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \qquad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$H = |+\rangle\langle 0| + |-\rangle\langle 1| = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$|\psi\rangle = a|0\rangle + b|1\rangle = \frac{a+b}{\sqrt{2}}|+\rangle + \frac{a-b}{\sqrt{2}}|-\rangle$$

# Pauli Gates

## Pauli Gates

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

## Phase Shift Gate

$$R_\phi = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}$$

- $H = \frac{X+Z}{\sqrt{2}}$
- $H^2 = HH^\dagger = X^2 = Y^2 = Z^2 = -iXYZ = I$
- 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 $\theta$ and then about $z$ by angle $\phi$.

$2\times 2$ 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.

$$\mathrm{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

In general, one can define controlled versions of any unitary gate $U$ as

$$CU = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$$

$\mathrm{CNOT}(|+\rangle \otimes |0\rangle) = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}^{\mathsf{T}} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$ is an entangled state which can't be separated as tensor product.

# Fredkin Gate: CSWAP

| $c$ | $p$ | $q$ | $x$ | $y$ | $z$ |
|---|---|---|---|---|---|
| 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 | 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.

$$f : (c, p, q) \mapsto (c, c \cdot q + \overline{c} \cdot p, c \cdot p + \overline{c} \cdot q)$$

$$\text{CSWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$\neg \; p = 0 \;\&\; q = 1 \implies z = \overline{c}$

$\wedge \; q = 0 \implies z = c \cdot p$

# Toffoli Gate: CCNOT or $D(\frac{\pi}{2})$

| $c_1$ | $c_2$ | $t$ | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|---|---|
| 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.
$f : (c_1, c_2, t) \mapsto \big(c_1, c_2, t \oplus c_1 \cdot c_2\big)$

$$\text{CCNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$\neg \; c_1 = c_2 = 1 \implies x_3 = \bar{t}$

$\wedge \; t = 0 \implies x_3 = c_1 \cdot c_2$

# Universal Quantum Gates

- Fredkin gate $\mathrm{CSWAP}$ is universal for classical computation, but not universal for quantum computation.
- Toffoli gate $\mathrm{CCNOT}$ 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

$$D(\theta) : |a, b, c\rangle \mapsto \begin{cases} i\cos\theta\,|a, b, c\rangle + \sin\theta\,|a, b, 1-c\rangle & \text{for } a = b = 1 \\ |a, b, c\rangle & \text{otherwise} \end{cases}$$

- 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 $|\psi_i\rangle \to |\psi_f\rangle = \sum_{x\in\{0,1\}^n} b_x|x\rangle$ 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 $f : \{0,1\}^n \to \{0,1\}$ 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 $\mathrm{CNOT}$ | After $H$ on $1^{\mathrm{st}}$ 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

Propositional Logic

Term Logic

Predicate Logic

Modal Logic

Induction, Analogy, Fallacy

Set Theory

Recursion Theory

Equational Logic

Homotopy Type Theory

Category Theory

Quantum Computing

Answers to the Exercises

References1655

# Answers to the Exercises — Translation

1. $\forall x \forall y (Hxy \to Uxy)$
2. $\forall x \big(\neg Sx \to \exists y (Eyx \land \neg Sy)\big)$
3. $\forall x (Mx \lor Wx \to Ax)$
4. $\forall x \big(\neg (Rx \land Bx) \to \neg Ex\big)$
5. $Hs \land Hp \land \forall x (Hx \to x = s \lor x = p)$
6. $\forall x (x \neq s \land x \neq p \to Hx)$
7. $\forall x \big(Bx \to \exists y \exists z (Gy \land Gz \land y \neq z \land Lxy \land Lxz)\big)$
8. $\neg \forall x (Jx \to Dx)$
9. $\forall x (Jx \to \neg Dx)$
10. $\neg \forall x \big(\mathrm{Glitter}(x) \to \mathrm{Gold}(x)\big)$
11. $\forall xyz (\mathrm{City}(z) \land \mathrm{In}(x, z) \land \mathrm{In}(y, z) \to \mathrm{first}(\mathrm{zip}(x)) = \mathrm{first}(\mathrm{zip}(y))$
12. $\forall x \forall y (Fx \land Dy \land Oxy \to Hx)$
13. $\forall x \forall y (Fx \land Dy \land Oxy \to Bxy)$
14. $\forall x (Ex \to Dx2) \land \exists x (Ex \land Dx4) \land \exists x (Ex \land \neg Dx4)$

# Answers to the Exercises — Translation

15. $\forall x(Fx \to \exists y(Cy \land Byx))$
16. $\forall x(Cx \land Ex \to \neg\exists y(Py \land Sy) \to Dx)$
17. $\forall x\big(Bx \to \neg\exists y(Ty \land Axy) \lor \forall y(Ty \to Axy)\big)$
18. $\forall x(\neg Hxx \to Hjx)$
19. $\neg\exists x\big(Bx \land \forall y(Sxy \leftrightarrow \neg Syy)\big)$
20. $\exists x\exists y\exists z(Mxa \land Myb \land Mzx \land Mzy) \land \forall u\forall v\forall x\forall y(Mua \land Mvb \land Mxu \land Myv \to x = y)$ $\qquad m(m(a)) = m(m(b))$
21. $\forall x\exists y\exists z(Gy \land Gz \land y \neq z \land Lxy \land Lxz \to x = t)$
22. $\forall x\big(L(x, f(w(s))) \to x = m(w(s))\big)$
23. $\forall x(Dx \to Ax) \to \forall x\big(\exists y(Dy \land Hxy) \to \exists y(Ay \land Hxy)\big)$
    $\forall x(Dx \to Ax) \to \forall x(Dx \to \exists y(Ay \land hx = hy))$
24. $\exists x\big(Dx \land \forall y(Dy \to y = x) \land \neg\exists z(Tz \land Hxz)\big)$
25. $\exists x\big(Gx \land Lxb \land \forall y(Gy \land Lyb \to y = x) \land \exists y(y \neq x \land Sy)\big)$
26. $\exists x(Gx \land Lxm \land \forall y(Gy \land Lym \to y = x) \land Lmx \land \forall y(Lmy \to y = x))$
27. $\exists x(Lxa \land \forall y(Lya \to y = x) \land \exists y(Lay \land \forall z(Laz \to z = y) \land y = x))$
28. $\exists x\Big(Ex \land \forall z(Ez \to Tzx) \land \exists y\big(Ey \land \forall z(Ez \to Tyz) \land Lxy\big)\Big)$

1. $\exists x Jx, \neg\exists x(Ax \land Jx) \vdash \exists x \neg Ax$
2. $\forall x(\neg Ax \land Mx \to Vx \lor Fx), \forall x(Px \to \neg Ax \land \neg Vx \land \neg Fx) \vdash \forall x(Px \to \neg Mx)$
3. $\forall x(Dx \to Ax) \to \forall x\big(\exists y(Dx \land Hxy) \to \exists y(Ay \land Hxy)\big)$
4. $\exists x \forall y\big((Ky \leftrightarrow y = x) \land Bx\big), \forall x(Bx \to Sx) \vdash \forall x(Kx \to Sx)$
5. $\forall x(Px \to x = r), Pw \land Sw \vdash Sr$
6. $\forall x Fxd, \forall x(Fdx \to x = i) \vdash i = d$
7. $\forall x \forall y(\exists z Lyz \to Lxy), Lrj \vdash Liu$
8. $\forall x \forall y(\exists z Lyz \to Lxy) \vdash \exists x \exists y Lxy \to \forall x \forall y Lxy$
9. $\neg\exists x \forall y(Txy \leftrightarrow \neg\exists z(Tyz \land Tzy))$
10. $\forall x(Sx \to Kx), \exists x Sx, \forall x(Kx \to x = j) \vdash Sj$
11. $\neg\exists x \exists y(Gx \land Sy \land Lxy), Gc \land \forall x(Lxc \to Lcx), Lhc \vdash \neg Sh$
12. $Pi, \forall x \forall y(Px \land Axy \to Py), \neg\exists x(Px \land \neg Ex), Ri, \forall x(Ex \land Rx \to \exists y(Gy \land Axy)), \forall x(Ex \to Cx) \vdash \exists x(Gx \land Cx)$
13. $\forall x(Px \to \exists y(Ly \land Axy)), \exists x(Sx \land \forall y(Axy \to Fy)), \neg\exists x(Fx \land Lx) \vdash \neg\forall x(Sx \to Px)$
14. $\forall x \forall y(Sxy \to \exists z Izxy), \forall x \forall y \forall z(Izxy \to Kzx \land Kzy), \forall x Sxa \vdash \forall x \exists y(Iyxa \land Kya)$
15. $Aac \land Abc \land a \neq b \land \forall x(Axc \to x = a \lor x = b), \forall x(Axc \leftrightarrow Lxc) \vdash \exists x \exists y(Lxc \land Lyc \land x \neq y \land \forall z(Lzc \to z = x \lor z = y))$

# 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.

[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.

[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.

[Lou09] A. H. Louie. *A Synthetic Continuation in Relational Biology*. Berlin, Boston: De Gruyter, 2009.

# References II


[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.

[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.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[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.

# References III


[San20] Paulo Guilherme Santos. *Diagonalization in Formal Mathematics*. Springer Spektrum, Wiesbaden, 2020.

[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.

[Svo18] Karl Svozil. *Physical (A)Causality Determinism, Randomness and Uncaused Events*. Fundamental Theories of Physics, 192. Cham: Springer International Publishing, 2018.

[Uni13] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. Institute for Advanced Study: https://homotopytypetheory.org/book, 2013.

# Thank