Discrete Mathematics: Lecture 15

Part III. Mathematical Logic

proposition, truth value, propositional constant/variable, negation, conjunction, disjunction, implication, bi-implication, formula, precedence, translation

Xuming He
Associate Professor

School of Information Science and Technology
ShanghaiTech University

Spring Semester, 2025

Notes by Prof. Liangfeng Zhang and Xuming He

Mathematical Logic

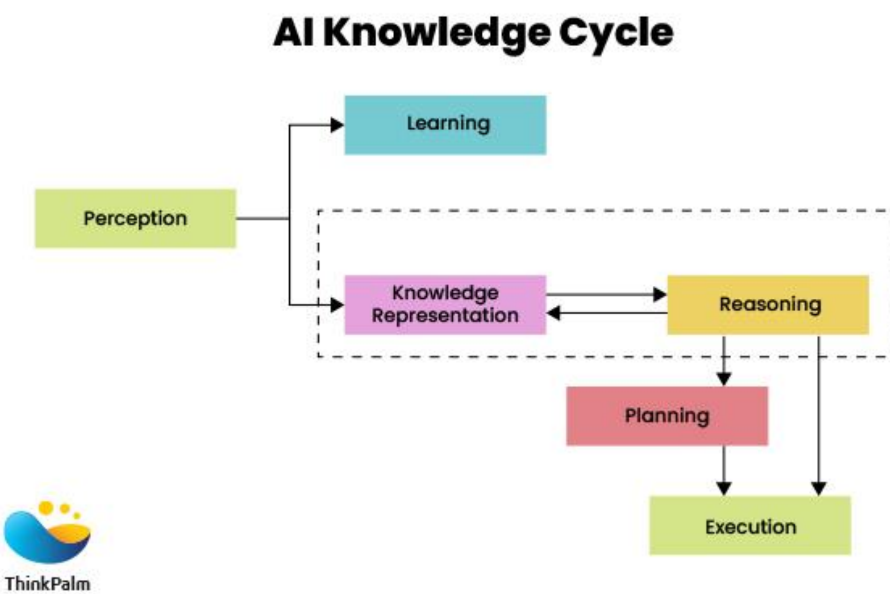
Logic: the study of reasoning, the basis of all mathematical reasoning.

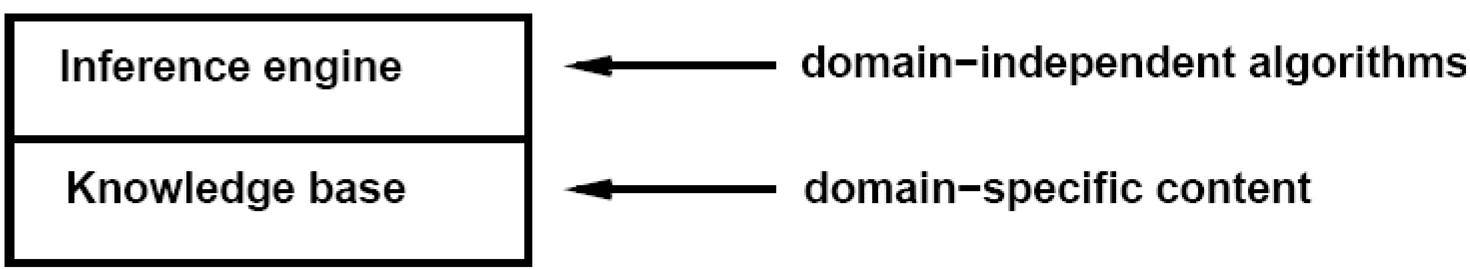
Mathematical logic: the mathematical study of reasoning and the study of mathematical reasoning //foundation of mathematics

- Leibniz: introduced the idea of mathematical logic in "Dissertation on the Art of Combinations" in 1666
  - universal system of reasoning: reasoning based on **symbols+calculations**
- **Contributors**: Boole, De Morgan, Frege, Peano, Russell, Hilbert, Gödel,…
- **Areas**: (1) set theory, (2) proof theory, (3) recursion theory, (4) model theory, and their foundation (5) propositional logic and predicate logic
- **Our focus**: propositional logic and predicate logic

Application: Logic-based Symbolic AI

- Logic (Knowledge-Based) AI
  - Knowledge base
    - set of sentences in a formal language to represent knowledge about the “world”
  - Inference engine
    - answers any answerable question following the knowledge base

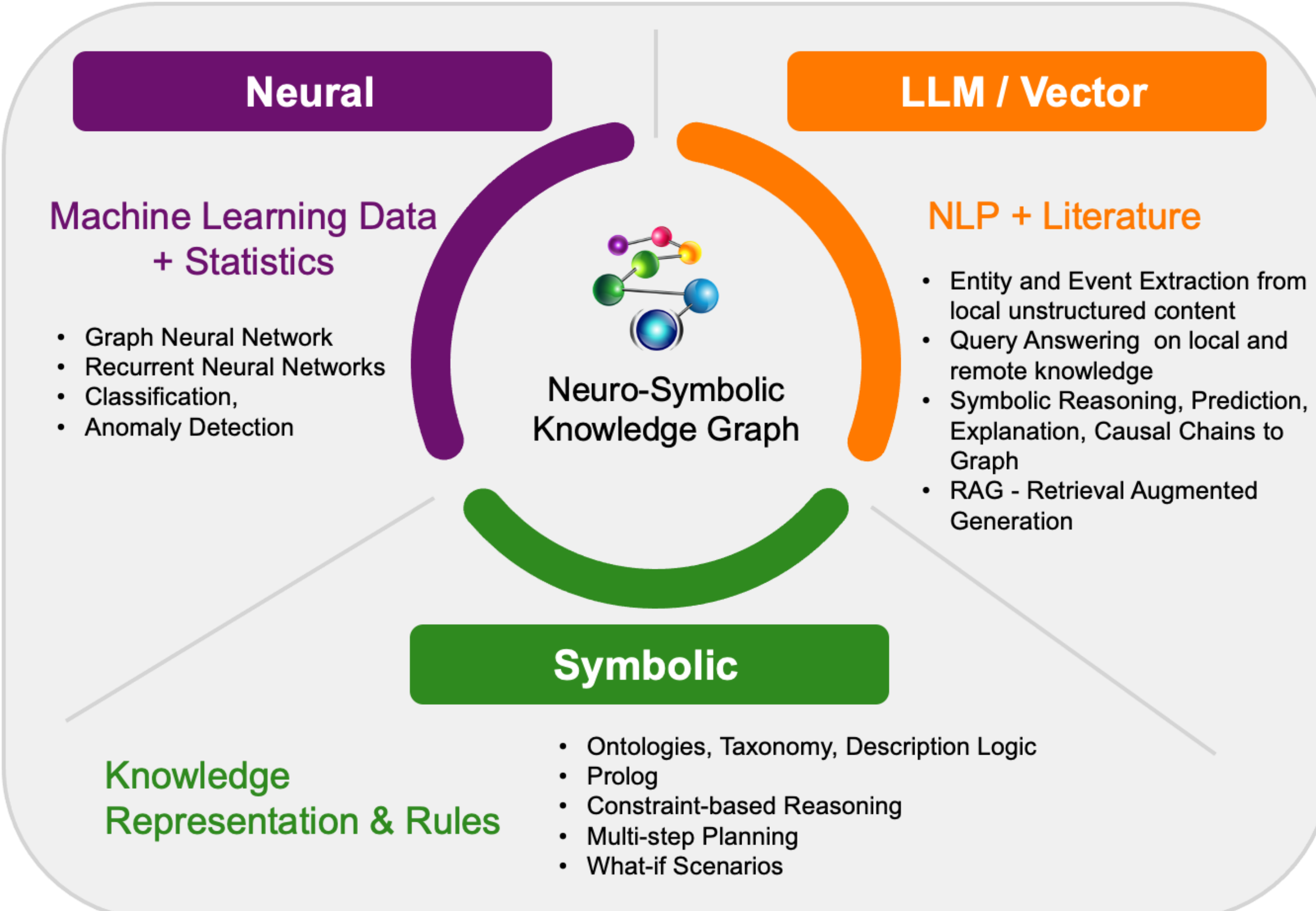




Application: Relational Databases

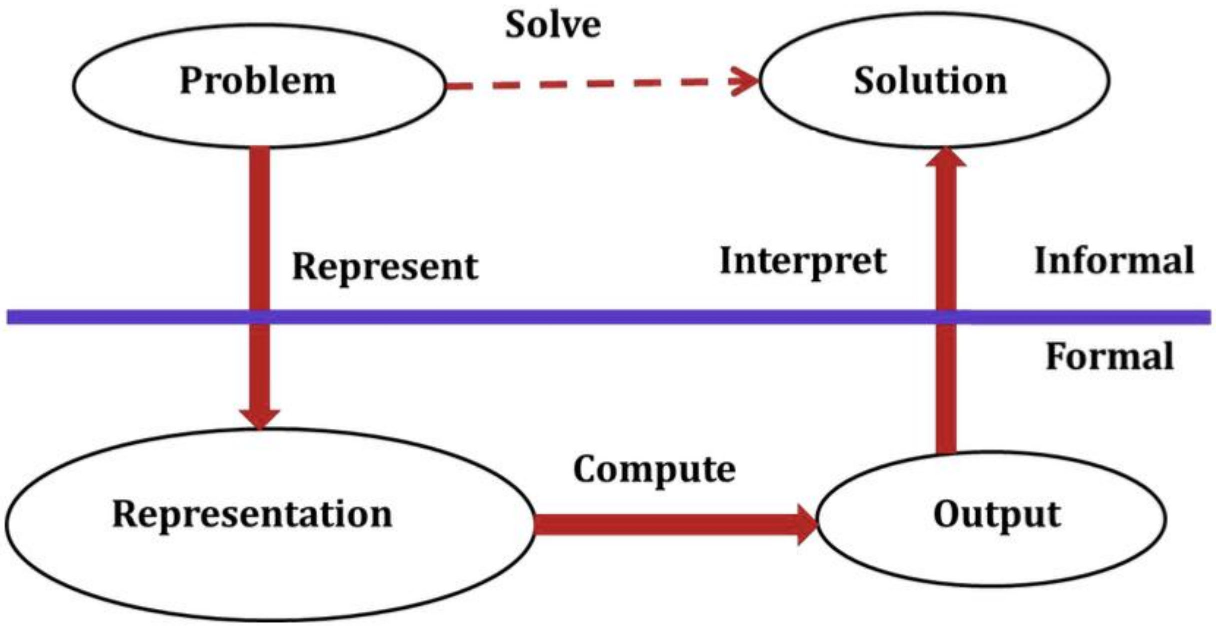
- Syntax: ground relational sentences, e.g., *Sibling*(*Ali*,*Bo*)
- Semantics: sentences in the DB are true, everything else is false
  - Query language (SQL etc.) typically some variant of first-order logic
  - Often augmented by first-order rule languages, e.g., Datalog
- Knowledge graphs (roughly: relational DB + ontology of types and relations)
  - Google Knowledge Graph: 5 billion entities, 500 billion facts, >30% of queries
  - Facebook network: 2.8 billion people, trillions of posts, maybe quadrillions of facts

Application: Reasoning in LLM



https://allegrograph.com/what-is-neuro-symbolic-ai/

Overview of Modeling Pipeline



Proposition

DEFINITION: A **proposition** (命题) is a declarative sentence that is either true or false.

- Lower-case letters represent propositions: $p, q, r, \ldots$
- **truth value** (真值)**:** The truth value of p is true (**T**) if p is a true proposition. The truth value of p is false (**F**) if p is a false proposition.

EXAMPLE:

- $\sqrt{2}$ is irrational. (**T**)
- $1 + 1 = 2.$ (**T**)
- $1 + 1 = 0.$ (**F**)
- $\left(x^2\right)' = 2x.$ (**T**)

Proposition

Simple Proposition(简单命题): cannot be broken into 2 or more propositions

- $\sqrt{2}$ is irrational.

Compound Proposition(复合命题): not simple

- 2 is rational and $\sqrt{2}$ is irrational.

Propositional Constant(命题常项): a concrete proposition

- Every even integer $n > 2$ is the sum of two primes.

Propositional Variable(命题变项): a variable that represents any proposition

- Lowercase letters denote proposition variables: $p, q, r, s, \ldots$
  - Truth value is not determined until it is assigned a concrete proposition

Propositional Logic(命题逻辑): the area of logic that deals with propositions

Proposition

- Every even integer $n > 2$ is the sum of two primes.
  - Goldbach’s conjecture
  - A proposition whose truth value is not known now
- 5 is rational and $\sqrt{5}$ is irrational.
  - A proposition that is not atomic
- What time is it?
  - Not a proposition. It’s not declarative.
- Do not smoke!
  - Not a proposition. It’s not declarative.
- $x + 1 = 2$.
  - Not a proposition. It’s neither true nor false

Negation ($\neg$)

DEFINITION: Let p be any proposition.

- The **negation** (否定) of p is the statement "It is not the case that p"
- notation: $\neg p$; read as "not p"
- **truth table** (真值表):

p	$\neg p$
T	**F**
F	**T**

EXAMPLE:

- p:" Snow is black."
  - $\neg p$: "It is not the case that snow is black."
  - $\neg p$: "Snow is not black."
  - $\neg p$ is not equal to "Snow is white."

Negation ($\neg$)

DEFINITION: Let p be any proposition.

- The **negation** (否定) of p is the statement "It is not the case that p"
- notation: $\neg p$; read as "not p"
- **truth table** (真值表):

p	$\neg p$
T	**F**
F	**T**

EXAMPLE:

- p: "Amy's smartphone has at least 32 GB of memory. "
  - $\neg p$: "It is not the case that Amy's smartphone has at least 32 GB of memory. "
  - $\neg p$: "Amy's smartphone does not have at least 32 GB."
  - $\neg p$: "Amy's smartphone has less than 32 GB."

Conjunction ($\wedge$)

DEFINITION: Let p, q be any propositions.

- The **conjunction** (合取) of p and q is the statement "p and q"
- Notation: $p \wedge q$; read as "p and q"
- Truth table:

p	q	$p \wedge q$
T	**T**	**T**
T	**F**	**F**
F	**T**	**F**
F	**F**	**F**

EXAMPLE:

- p: "$3 < 5$."; q: "$3^5 < 5^3$."
  - $p \wedge q$: "$3 < 5$ and $3^5 < 5^3$." (**F**)

Disjunction (∨)

DEFINITION: Let p, q be any propositions.

- The **disjunction** (析取) of p and q is the statement "p or q"
- Notation: $p \vee q$; read as "p or q"
- Truth table:

p	q	$p \vee q$
T	**T**	**T**
T	**F**	**T**
F	**T**	**T**
F	**F**	**F**

EXAMPLE:

- p: "$e^{i\pi} = -1$."; q: "Dog can fly."
  - $p \vee q$: "$e^{i\pi} = -1$ or dog can fly." **(T)**

Implication (→)

DEFINITION: Let p, q be any propositions.

- The conditional statement $p \rightarrow q$ (蕴涵) is the proposition "if p, then q."
  - p: hypothesis*;* q*:* conclusion*;* read as "p implies q", or "if p, then q"
- Truth table:

p	q	$p \rightarrow q$
T	**T**	**T**
T	**F**	**F**
F	**T**	**T**
F	**F**	**T**

EXAMPLE*:*

- p:"you get 100 on the final"; q:"you will receive A+"
  - $p \rightarrow q$:"If you get 100 on the final, then you will receive A+."
  - This is false when you get 100 on the final but don't receive A+.
    - i.e., when p is true but q is false.

Bi-Implication ($\leftrightarrow$)

DEFINITION: Let p, q be any propositions.

- The biconditional statement $p \leftrightarrow q$ (等值) is the proposition "p if and only if q."
  - read as "p if and only if q"
- Truth table:

p	q	$p \leftrightarrow q$
T	**T**	**T**
T	**F**	**F**
F	**T**	**F**
F	**F**	**T**

EXAMPLE*:*

- p:"you can take the flight"; q:"you buy a ticket"
  - $p \leftrightarrow q$:"You can take the flight if and only if you buy a ticket."
  - False when $(p, q) = (\mathbf{T}, \mathbf{F})$ or $(\mathbf{F}, \mathbf{T})$

Well-Formed Formulas (合式公式)

DEFINITION: recursive definition of **well-formed formulas** (**WFFs**)

① propositional constants (**T, F**) and propositional variables are WFFs

② If A is a WFF, then $\neg A$ is a WFF

③ If A, B are WFFs, then $(A \land B), (A \lor B), (A \rightarrow B), (A \leftrightarrow B)$ are WFFs

④ WFFs are results of finitely many applications of ①, ② , and ③

EXAMPLE:

- $\Big(\big((p \rightarrow q) \land (q \rightarrow r)\big) \leftrightarrow (p \rightarrow r)\Big)$ WFF
- $\neg p \lor q \lor$ not WFF
- $\neg(p \land q) \rightarrow (r \land s)$ WFF
- $(p \land q)\neg r$ not WFF

Use tree structure to check

REMARK: well-formed formulas=propositional formulas=formulas

Precedence of Logical Operators

Precedence (priority) (优先级)：

① Formulas inside ( ) are computed firstly

② Different connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$ (Decreasing Precedence)

③ Same connectives: from left to the right

EXAMPLE:

- $\neg p \wedge q$:

  $(\neg p) \wedge q$

- $\neg(p \wedge q)$:

  First $p \wedge q$, then $\neg$

- $p \vee q \wedge r$:

  $p \vee (q \wedge r)$

- $\big((p \rightarrow q) \wedge (q \rightarrow r)\big) \leftrightarrow (p \rightarrow r)$:

  $$\underline{\underline{\big(\underline{(p \rightarrow q)} \wedge \underline{(q \rightarrow r)}\big)} \leftrightarrow \underline{(p \rightarrow r)}}$$

From Natural Language to WFFs

The Method of Translation:

- Introduce symbols p, q, r ... to represent simple propositions
- Connect the symbols with logical connectives to obtain WFFs

EXAMPLE:

- "It is not the case that snow is black."
  - p:"Snow is black" Translation: $\neg p$
- "π and e are both irrational."
  - p:"π is irrational."; q:"e is irrational." Translation: $p \wedge q$
- "If π is irrational, then 2π is irrational"
  - p:"π is irrational"; q:"2π is irrational" Translation: $p \to q$
- "$e^\pi > \pi^e$ if and only if $\pi > e \ln \pi$."
  - p:"$e^\pi > \pi^e$"; q:"$\pi > e \ln \pi$" Translation: $p \leftrightarrow q$

Remark: it is better to choose the simple proposition to be affirmative sentence.

Example

- $\left(\sqrt{2}\right)^{\sqrt{2}}$ is rational or irrational. (ambiguity in natural language)
  - p: "$\left(\sqrt{2}\right)^{\sqrt{2}}$ is rational"; q: "$\left(\sqrt{2}\right)^{\sqrt{2}}$ is irrational"
  - **Explanation 1**: $\left(\sqrt{2}\right)^{\sqrt{2}}$ cannot be neither rational nor irrational.
    - Translation 1: $p \vee q$
      - We agree that $p \vee q$ is the correct translation of "$\left(\sqrt{2}\right)^{\sqrt{2}}$ is rational or irrational ."
  - **Explanation 2**: $\left(\sqrt{2}\right)^{\sqrt{2}}$ cannot be both rational and irrational.
    - Translation 2: $(p \wedge \neg q) \vee (q \wedge \neg p)$
      - We agree that $(p \wedge \neg q) \vee (q \wedge \neg p)$ is the translation of "$\left(\sqrt{2}\right)^{\sqrt{2}}$ is rational or irrational, but not both."

Example

- You are eligible to be President of the U.S.A. only if you are at least 35 years old, were born in the U.S.A, or at the time of your birth both of your parents were citizens, and you have lived at least 14 years in the country.
  - e: “You are eligible to be President of the U.S.A.,”
  - a: “You are at least 35 years old,”
  - b: “You were born in the U.S.A,”
  - p: “At the time of your birth, both of your parents were citizens,”
  - r: “You have lived at least 14 years in the U.S.A.”
    - Translation: $e \rightarrow (a \wedge (b \vee p) \wedge r)$

Example

- A comes to the party if and only if B doesn't come, but, if B comes, then C doesn't come and D comes.
  - a: "A comes to the party."
  - b: "B comes to the party."
  - c: "C comes to the party."
  - d: "D comes to the party."
    - Translation: $(a \leftrightarrow \neg b) \wedge \big(b \rightarrow (\neg c \wedge d)\big)$

- A sufficient condition for A coming to the party is that, if B does not come, then at least one of C and D must come.
  - a, b, c, d defined as above.
    - Translation: $(\neg b \rightarrow (c \vee d)) \rightarrow a$

Example

- **System Specifications**: Determine if there is a system that satisfies all of the following requirements.
  1. The diagnostic message is stored in the buffer or it is retransmitted.
  2. The diagnostic message is not stored in the buffer.
  3. If the diagnostic message is stored in the buffer, then it's retransmitted.
     - s: "The diagnostic message is stored in the buffer"
     - r: "The diagnostic message is retransmitted"
       - $s \vee r; \neg s; s \rightarrow r$
         - There is a system that satisfies 1, 2 and 3. ($s = \mathbf{F}, r = \mathbf{T}$)
  4. Add one more requirement "The diagnostic message is not retransmitted"
     - $s \vee r; \neg s; s \rightarrow r; \neg r$
       - There is no system that satisfies 1, 2, 3 and 4.

Logic Reasoning in LLMs

PROOF OR BLUFF? EVALUATING LLMS ON 2025 USA MATH OLYMPIAD

Ivo Petrov[2], Jasper Dekoninck[1], Lyuben Baltadzhiev[2], Maria Drencheva[2], Kristian Minchev[2]
Mislav Balunović[1,2], Nikola Jovanović[1], Martin Vechev[1,2]
[1]ETH Zurich [2]INSAIT, Sofia University "St. Kliment Ohridski"
https://matharena.ai/
https://github.com/eth-sri/matharena

ABSTRACT

Recent math benchmarks for large language models (LLMs) such as MathArena indicate that state-of-the-art reasoning models achieve impressive performance on mathematical competitions like AIME, with the leading model, GEMINI-2.5-PRO, achieving scores comparable to top human competitors. However, these benchmarks evaluate models solely based on final numerical answers, neglecting rigorous reasoning and proof generation which are essential for real-world mathematical tasks. To address this, we introduce the first comprehensive evaluation of full-solution reasoning for challenging mathematical problems. Using expert human annotators, we evaluated several state-of-the-art reasoning models on the six problems from the 2025 USAMO within hours of their release. Our results reveal that all tested models struggled significantly: only GEMINI-2.5-PRO achieves a non-trivial score of 25%, while all other models achieve less than 5%. Through detailed analysis of reasoning traces, we identify the most common failure modes and find several unwanted artifacts arising from the optimization strategies employed during model training. Overall, our results suggest that current LLMs are inadequate for rigorous mathematical reasoning tasks, highlighting the need for substantial improvements in reasoning and proof generation capabilities.

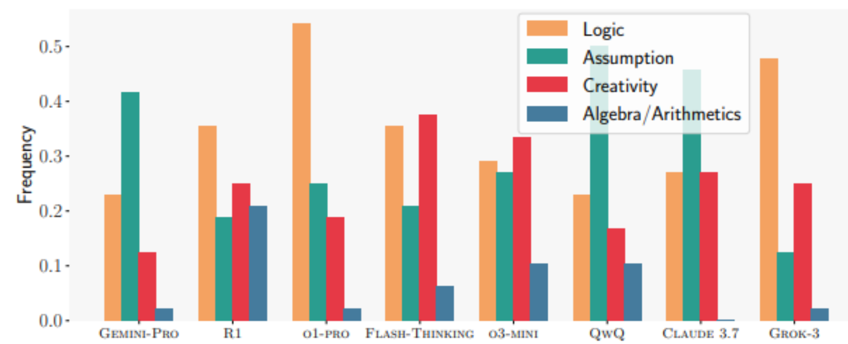


Figure 2: Distribution of first encountered failure mode.

https://arxiv.org/pdf/2503.21934