数学における AI の理論と実践

園田　翔

理化学研究所／サイバーエージェント

第 96 回産総研人工知能セミナー「数学と AI」
2026 年 8 月 7 日（金）

自己紹介

2025.6-現在 サイバーエージェント

AI 事業本部リサーチサイエンティスト
2025.6-2030.5 JST BOOST「大規模言語モデルによる定理証明 AI の学習理論」に伴うクロアポ

2018.4-現在 理化学研究所

AIP 深層学習理論チーム上級研究員
2025.10-2031.3 JST CREST「量子計算と機械学習の双方向発展を実現するシステム構築」

2017.4-2018.3 早稲田大学助手

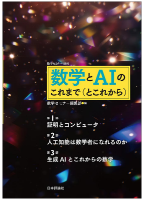
2013.9-2017.3 早稲田大学博士（工学）

2012.4-2013.9 パナソニック株式会社

車載機器開発

研究分野: 機械学習

理論: 統計的学習理論，特に深層学習理論
応用: AI4Math，量子機械学習，脳波解析，工場操業データ解析，自動運転



AI 援用数学の現状

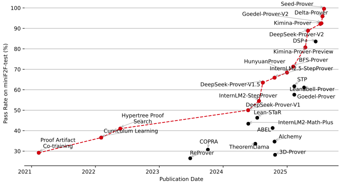
AI 援用数学の現状

〜2025 大学入試や IMO を含む高校数学で高得点を記録

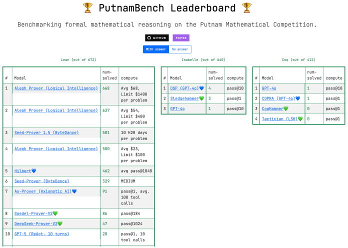
- 2024.07 Google DeepMind (AlphaProof + AlphaGeometry 2) IMO 2024 銀メダル相当（4/6 問回答・28 点）
- 2025.02 GPT-o3-mini-high 東大理系数学 5 完半（解答略氏実験．大手予備校も追随．非形式証明）
- 2025.07 Google DeepMind (Gemini DeepThink) IMO 2025 金メダル相当（5/6 問完答・35 点）
- 2026.03 Harmonic Aristotle 東大理系数学 4 完（折池氏実験．形式証明.）

2025〜未解決問題が解けるようになる

- 2025.11 Terence Tao が AlphaEvolve の未解決問題に対する性能を評価．掛谷予想の下限が改良された．
- 2026.01 Erdös 問題で一連の新結果．代数幾何でも新結果（Bryan+2026）．



MiniF2F
（形式化・高校レベル）
from ByteDance, Seed-Prover, 2025



🏆 PutnamBench Leaderboard 🏆

Benchmarking formal mathematical reasoning on the Putnam Mathematical Competition.

GITHUB PAPER

With answer No answer

Lean (out of 672)

#	Model	num-solved	compute
1	Aleph Prover (Logical Intelligence)	668	Avg $68, Limit $1400 per problem
2	Aleph Prover (Logical Intelligence)	637	Avg $54, Limit $400 per problem
3	Seed-Prover 1.5 (ByteDance)	581	10 H20 days per problem
4	Aleph Prover (Logical Intelligence)	500	Avg $23, Limit $100 per problem
5	Hilbert💙	462	avg pass@1840
6	Seed-Prover (ByteDance)	329	MEDIUM
7	Ax-Prover (Axiomatic AI)💙	91	pass@1, avg. 100 tool calls
8	Goedel-Prover-V2💚	86	pass@184
9	DeepSeek-Prover-V2💚	47	pass@1024
10	GPT-5 (ReAct, 10 turns)	28	pass@1, 10 tool calls

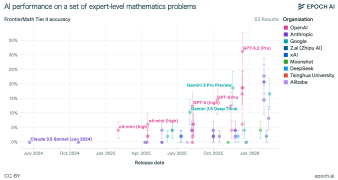
Isabelle (out of 640)

#	Model	num-solved	compute
1	DSP (GPT-4o)💙	4	pass@10
2	Sledgehammer💚	3	pass@1
3	GPT-4o	1	pass@10

Coq (out of 412)

#	Model	num-solved	compute
1	GPT-4o	1	pass@10
2	COPRA (GPT-4o)💙	1	pass@1
3	CoqHammer💚	0	pass@1
4	Tactician (LSH)💚	0	pass@1

PutnamBench
（形式化・学部レベル）



FrontierMath Tier-4
（非形式・専門レベル）

最先端の AI 数学に関する動向

- 2026.01 頃 Erdös 問題 加速．国内貢献 Yuta Oriike さん, Kenta Kitamura さん
- 2026.05.20 OpenAI 単位距離予想の反証（blog）
- 2026.06.02 Leiden 宣言（doi）
- 2026.07.19 Claude Fable 5 Jacobian 予想の反証（x）
- 2026.07.27 Lean v4.32.1 の健全性に関するバグ（Collatz 予想の「反証」github）（修正済 blog）
- 2026.08.01 OpenAI 10 個の未解決問題を推進（blog）
- 2026.08 頃 Formal Conjectures 加速．国内貢献 Kenta Kitamura さん（PR）

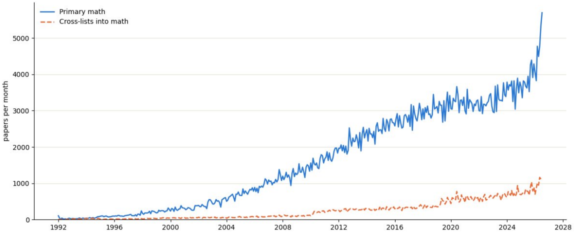


Illustration of monthly submission counts at arxiv math
by Pablo Groisman, 4 Aug 2026

いつから始まっていたか／なぜ Lean か

- 2020.01 Lean Zulip Chat に Machine Learning for Theorem Proving channel が立ち上がる．
- 2020.09 OpenAI GPT-f. 初の形式数学 LLM．Metamath で新しい証明を与える．
- 2021.01 Lean v4.0.0 リリース．
- (2021.03 GitHub Copilot リリース．初のコード補完 AI.)
- 2021.05 The 1st Math-AI workshop (at ICLR2021)
- 2021.03/08/10 高校数学のベンチマークデータ MATH （非形式），OpenAI MiniF2F 形式（Lean/Metamath/Isabelle/HOL Light），OpenAI GSM8k（非形式）
- (2022.01/05 Google Research, Chain-of-Thought. Kojima et al., Step-by-step prompting. 推論の発見.)
- 2022.06 Google Research Minerva (PaLM). LLM による非形式・高校数学（MATH, GSM8k）
- (2022.11 GPT-3.5 “ChatGPT” リリース．LLM 普及のきっかけとなる.)
- (2023.02 Meta Llama パラメータリーク．OSS の母体となる.)（Google Bard リリース．後の Gemini.)
- 2023.05/11 Terence Tao が自身の blog で初めて AI と Lean に関与していくことを表明
- 2023.07 LeanFRO 発足（-2028.6?）
- 2023.12 LeanDojo + ReProver. Lean の OSS RAG
- 2024.07/11 PutnumBench（形式・学部レベル），Epoch AI FrontierMath（非形式・専門レベル）
- 2024.08 DeepSeek-Prover V1.5 本格的に Lean が書ける OSS．(Lean 一本化の遠因？)
- (2024.09 OpenAI o1 推論モデル)
- 2025.01 OpenAI o3 リリース. （Lean 4 の生成安定化．架空ライブラリを引用する幻覚低減.）

どこに向かうか

- 車や機械，計算機，... R，pytorch，... の発明と似てる
  - （ついに）数学が機械化・加速する
  - Human-in-the-loop から完全自動化する
  - ソフトウェア検証・コード生成へ横展開
  - cf. de Moura, 18 feb 2026, Avigad, 7 mar 2026

- 加速の駆動力
  - 2023.07- Lean FRO
  - 2025.04- DARPA Exponentiating Mathematics (ExpMath) project
  - 2025.07- ユニコーン企業の台頭
    Harmonic (Aristotle), Math Inc (Gauss), Logical Intelligence (Aleph Prover), Axiom Math (AxiomProver)

- 理論研究の方向性
  - “ランダム” な AI はなぜ / いつ証明できるのか？
  - 良いデータ表現とは何か？
  - 人間は外挿できる．AI は内挿しかできない．のか？
  - Riemann 予想は解けるのか？

（速く走れるだけでは，遠くまで行けないこともある）

Lean の概観

Lean による形式証明の仕組み

- 定義 — `def`, `structure`, `inductive`, `class`, … で定義する
- 定理 — 命題論理や述語論理を使って主張（=仕様）を書く
- 証明 — 自然演繹や帰納法，背理法，… を使って証明（=実装）を書く

例: 6 は偶数である（lean4web）

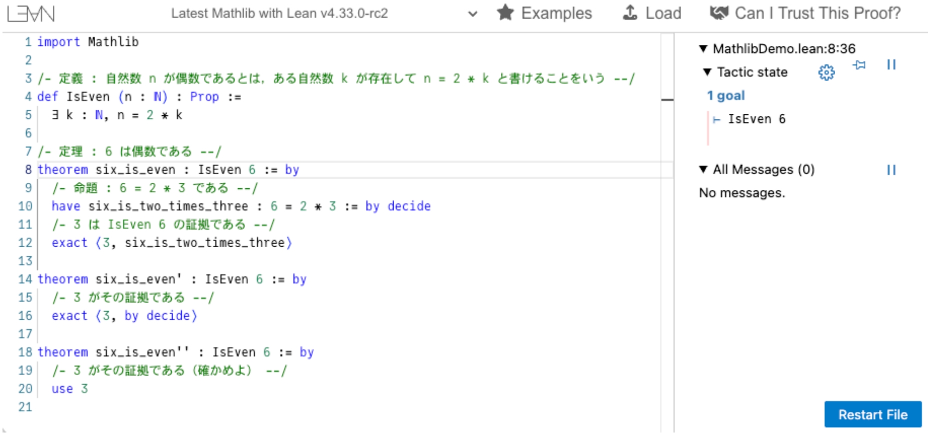
```
import Mathlib

/- 定義 : 自然数 n が偶数であるとは，ある自然数 k が存在して n = 2 * k と書けることをいう --/
def IsEven (n : ℕ) : Prop :=
  ∃ k : ℕ, n = 2 * k

/- 定理 : 6 は偶数である --/
theorem six_is_even : IsEven 6 := by
  /- 命題 : 6 = 2 * 3 である --/
  have six_is_two_times_three : 6 = 2 * 3 := by decide
  /- 3 は IsEven 6 の証拠である --/
  exact ⟨3, six_is_two_times_three⟩

theorem six_is_even' : IsEven 6 := by
  /- 3 がその証拠である（確かめよ） --/
  use 3
```

対話的に証明する



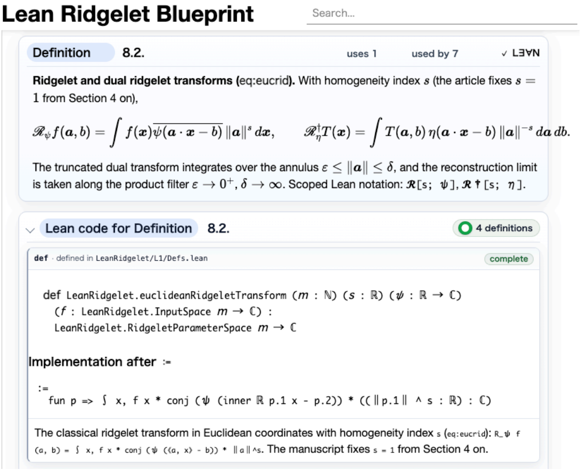
ブラウザ環境での実行例（lean4web）

形式化できると何が嬉しいか

- 定理の主張（仮定と結論）が明瞭になる
- 証明の正しさを機械的に検証できる
- 知識を再利用可能な形で蓄積できる

  e.g. Mathlib
- 証明をより深く理解できる

  e.g. Lean blueprints
- 証明を整理して新しい定理が発見できる
- AI を併用して効率的に数学を発展させられる



Lean Ridgelet Blueprint

依存関係の可視化

どこまで形式化できるか

- 現代の（典型的な）数学を記述できる

  Lean, Rocq (Coq), Agda: 依存型理論に基づく証明支援系

  Isabelle/HOL: 高階論理に基づく証明支援系

  Mizar, Metamath: 集合論に基づく証明支援系

- 数学ライブラリ

  Mathlib (Lean) — algebra, topology, analysis, probability, geometry, combinatorics, logic, …

  MathComp (Rocq), AFP (Isabelle), MML (Mizar)

- Lean の独自リポジトリ，数学以外のライブラリ

  Reservoir（Lean projects インデックス）

  CSlib（計算機科学），Physlib（物理学），StatMLlib（統計・機械学習）

A mathlib overview

The goal of this web page is to give a rough list of topics currently covered in mathlib, and provide pointers for exploration. This is not meant to be an exhaustive list, and could be outdated (see the full index for exhaustive and up to date information).

Here topics are listed in the greatest generality we currently have in mathlib, hence some things may be difficult to recognize. We also have a page dedicated to undergraduate mathematics which may be easier to read, as well as a page listing undergraduate maths topics that are not yet in mathlib.

To make updates to this list, please make a pull request to mathlib after editing the yaml source file. This can be done entirely on GitHub, see "Editing files in another user's repository".

General algebra

Category theory category, small category, functor, natural transformation, Yoneda embedding, isomorphism, adjunction, monad, comma category, limits, presheafed space, sheafed space, monoidal category, Cartesian closed, abelian category, Grothendieck topology. See also our documentation page about category theory.

Numbers natural number, integer, rational number, continued fraction, real number, extended real number, complex number, p-adic number, p-adic integer, hyper-real number. See also our documentation page about natural numbers.

Group theory group, group morphism, group action, class formula, Burnside lemma, subgroup, index of a subgroup, normal subgroup, simple group, subgroup generated by a subset, quotient group, first isomorphism theorem, second isomorphism theorem, third isomorphism theorem, abelianization, free group, presented group, Schreier's lemma, cyclic group, nilpotent group, permutation group of a type, Cayley's theorem, alternating group, alternating groups on 5 letters or more are simple, classification of abelian simple groups, structure of finitely generated abelian groups, Lagrange's theorem, first Sylow theorem, second Sylow theorem, third Sylow theorem.

Rings ring, ring morphism, the category of rings, subring, localization, local ring, Noetherian ring, ordered ring.

Ideals and quotients ideal of a commutative ring, quotient ring, prime ideal, maximal ideal, Chinese remainder theorem, fractional ideal, first isomorphism theorem for commutative rings.

Divisibility in integral domains irreducible element, coprime element, unique factorisation domain, greatest common divisor, least common multiple, principal ideal domain, Euclidean domain, Euclid's' algorithm, Euler's totient function (φ), Lucas-Lehmer primality test.

Polynomials and power series polynomial in one indeterminate, roots of a polynomial, multiplicity, separable polynomial, $K[X]$ is Euclidean, Hilbert basis theorem, $A[X]$ has gcd and lcm if A does, $A[X_i]$ is a UFD when A is a UFD, irreducible polynomial, Eisenstein's criterion, polynomial in several indeterminates, power series.

Algebras over a ring associative algebra over a commutative ring, the category of algebras over a ring, free algebra of a commutative ring, tensor product of algebras, tensor algebra of a commutative ring, Lie algebra, exterior algebra, Clifford algebra.

Field theory field, subfield, characteristic of a ring, characteristic zero, characteristic p, Frobenius morphism, algebraically closed field, existence of algebraic closure of a field, $\mathbb{C}$ is algebraically closed, field of fractions of an integral domain, algebraic extension, rupture field, splitting field, perfect closure, Galois correspondence, Abel-Ruffini theorem (one direction).

Homological algebra chain complex, functorial homology.

Number theory sum of two squares, sum of four squares, quadratic reciprocity, solutions to Pell's equation, Matiyasevič's theorem, arithmetic functions, Bernoulli numbers, Chevalley-Warning theorem, Hensel's lemma (for $\mathbb{Z}_p$), ring of Witt vectors, perfection of a ring.

Transcendental numbers Liouville's theorem on existence of transcendental numbers.

Algebraic Number Theory Finiteness of the class number, Dirichlet unit theorem, Dirichlet class number formula.

Representation theory representation, category of finite-dimensional representations, character, orthogonality of characters.

mathlib overview

国内の Lean 開発事例

- 統計的学習理論 Lean Rademacher, StatsMLlib
- 量子情報理論 Lean Quantum
- 微分方程式論 Leray-Hopf 対訳ノート
- 関数解析 Lean Ridgelet
- 測度論 測度論と積分, Lebesgue 積分論講義ノート
- 入試問題 東大数学 2026

どのように形式化するか

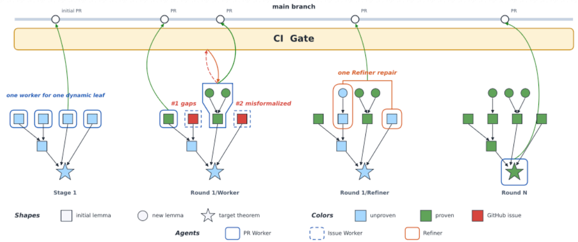
手動形式化

- 基本:
  公式サイト
  ゼロから始める Lean 言語入門
  数学系エンドユーザーのための Lean 入門

- 定理の探し方:
  Zulip chat
  Lean-by-Example
  Loogle!, Moogle
  自動証明機能 simp, aesop, grind
  mathlib 検索

自動形式化

e.g. LeanMarathon（完全自律型ハーネス）



Lean のコンパイルが通ったら安心か

- 仕様バグ
  定義・定理の形式化が意図したものと異なる
  e.g. 型強制（キャスト）・未定義に対する実装がおかしい，
  e.g. 強い仮定（結論そのものなど）が混入している

- 実装バグ
  e.g. 型が決まらない．健全（sound）でない．

- サポートツール
  nanoda (Rust), Lean4lean (lean) — Lean kernel の別実装
  Comparator — 信頼できない証明ファイルを安全にコンパイルするためのツール
  Lean Blueprints, LeanArchitect — 非形式証明（LaTeX）と形式証明（Lean）を比較する

研究紹介

Overview (AI4Math)

- Theorem Proving / Mathematical Reasoning using LLMs
  - AITP 2023, AITP 2024, SNL 2025, NALOMA 2026
- Auto / Manual-formalization with LLMs

  2024.1-present **Lean-Rademacher**
  - Lean formalization of statistical learning theory
  - ITP2026, github: lean-rademacher (indexed at Reservoir)
  - novelty: clarified assumptions on the measurability of supremum, first formalization of full-set SLT

  2025.11-present **Lean-Quantum**
  - Lean formalization of quantum information theory (sandwiched Renyi relative entropy, quantum parallel of data processing inequality, just published in 2025 by H. Yamasaki)
  - arxiv: 2607.05492, github: lean-quantum (indexed at Reservoir)
  - novelty: clarified assumptions on the infinite-dimensional quantum states, coordinate-free formalization

  2026.07-present **Lean-Ridgelet**
  - Lean formalization of ridgelet analysis (harmonic analysis for neural networks)
  - arxiv: 2106.04770, github: lean-ridgelet
  - novelty: clarified assumptions for L^2-theories, obtain stronger statements
- Theory

  2026.1-present **Statistical Provability** theory of agentic provers, ICML2026

How Formalization looks like (excerpt from Lean-Rademacher)[1]

[1] S-Kasaura-Mizuno-Tsukamoto-Onda, ITP2026 (ArXiv:2503.19605, GitHub:lean-rademacher)

Example of Lean Formalization

- Given a family of functions $\mathcal{F} \subset \{f \colon \mathcal{X} \to \mathbb{R}\}$ and a set of vectors $S := \{x_k\}_{k=1}^n \subset \mathcal{X}^n$,
- Empirical and Population Rademacher complexities

$$\widehat{\mathfrak{R}}_n(\mathcal{F}; S) := \mathbb{E}_\sigma \left[\sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{k=1}^n \sigma_k f(x_k) \right| \right], \quad \mathfrak{R}_n(\mathcal{F}) := \mathbb{E}_S[\widehat{\mathfrak{R}}_n(\mathcal{F}; S)].$$

- Uniform deviation

$$\Delta_n(S) := \sup_{f \in \mathcal{F}} \left| \widehat{L}_S(f) - L(f) \right|.$$

```
def empiricalRademacherComplexity (n : ℕ) (f : F → χ → ℝ) (S : Fin n → χ) : ℝ :=
  (Fintype.card (Signs n) : ℝ)⁻¹ *
    ∑ σ : Signs n, ⨆ i : F, |(n : ℝ)⁻¹ * ∑ k : Fin n, (σ k : ℝ) * f i (S k)|

def rademacherComplexity (n : ℕ) (f : F → χ → ℝ) (μ : Measure Ω) (X : Ω → χ) : ℝ :=
  μⁿ[fun ω : Fin n → Ω ↦ empiricalRademacherComplexity n f (X ∘ ω)]

def uniformDeviation (n : ℕ) (f : F → χ → ℝ) (μ : Measure Ω) (X : Ω → χ) (S : Fin n → χ)
  : ℝ :=
  ⨆ i : F, |(n : ℝ)⁻¹ * ∑ k : Fin n, f i (S k) - μ[fun ω' ↦ f i (X ω')]|
```

Example of Lean Formalization

Theorem (Rademacher generalization bound)

Assume that the function class $\mathcal{F}$ satisfies $|f(z)| \le b$ for every $f \in \mathcal{F}$ and $z \in \mathcal{Z}$. Under suitable measurability, separability, and continuity assumptions,

$$\Pr\left\{\Delta_n(S) \ge 2\mathfrak{R}_n(\mathcal{F}) + \varepsilon\right\} \le \exp\left(-\frac{n\varepsilon^2}{2b^2}\right),$$

```
theorem uniform_deviation_tail_bound_separable_of_pos
    [MeasurableSpace χ] [Nonempty χ]
    [Nonempty H] [TopologicalSpace H] [SeparableSpace H] [FirstCountableTopology H]
    [IsProbabilityMeasure μ]
    (F : H → χ → ℝ) (hF_meas : ∀ h, Measurable (F h))
    (X : Ω → χ) (hX : Measurable X)
    {b : ℝ} (hb : 0 < b) (hF_bound : ∀ h x, |F h x| ≤ b)
    (hF_cont : ∀ x : χ, Continuous fun h ↦ F h x)
    {ε : ℝ} (hε : 0 ≤ ε) :
    (μⁿ {S |
      2 * rademacherComplexity n F μ X + ε ≤
        uniformDeviation n F μ X (X ∘ S)}).toReal ≤
     (-ε ^ 2 * n / (2 * b ^ 2)).exp
```

Why Agentic Theorem Prover Works: A Statistical Provability theory of Mathematical Reasoning Models[2]

[2] S-Akiyama-Uezato, ICML2026 (ArXiv:2602.10538)+ preprint: 2602.10512

A new gap between theory and practice is emerging in AI4Math

- Theory: Theorem Proving (TP) is hard
  - Cook–Levin (1971): Boolean satisfiability problem (SAT) is NP-complete
  - Church (1936) and Turing (1936): First-order validity problem is undecidable

- Real world: generative AIs can solve expert-level mathematical tasks
  - 2025: high scores on high-school and olympiad-style benchmarks
  - 2026: new progress on Erdos problems, algebraic geometry (Bryan+2026), etc.

- Possible Scenarios: Real-world datasets are biased, not uniform
  — but how to formulate a probability distribution of problems?

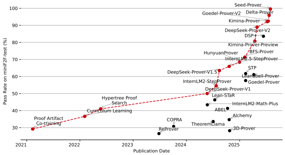
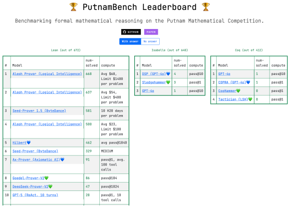


Figure 1: MiniF2F-Test performance (Fig. 1 from Seed-Prover, Aug 2025)



🏆 PutnamBench Leaderboard 🏆

Benchmarking formal mathematical reasoning on the Putnam Mathematical Competition.

GITHUB PAPER

With answer | No answer

Lean (out of 672)

#	Model	num-solved	compute
1	Aleph Prover (Logical Intelligence)	668	Avg $68, Limit $1400 per problem
2	Aleph Prover (Logical Intelligence)	637	Avg $54, Limit $400 per problem
3	Seed-Prover 1.5 (ByteDance)	581	10 H20 days per problem
4	Aleph Prover (Logical Intelligence)	500	Avg $23, Limit $100 per problem
5	Hilbert💙	462	avg pass@1840
6	Seed-Prover (ByteDance)	329	MEDIUM
7	Ax-Prover (Axiomatic AI)💙	91	pass@1, avg. 100 tool calls
8	Goedel-Prover-V2💚	86	pass@184
9	DeepSeek-Prover-V2💚	47	pass@1024
10	GPT-5 (ReAct, 10 turns)	28	pass@1, 10 tool calls

Isabelle (out of 640)

#	Model	num-solved	compute
1	DSP (GPT-4o)💙	4	pass@10
2	Sledgehammer💚	3	pass@1
3	GPT-4o	1	pass@10

Coq (out of 412)

#	Model	num-solved	compute
1	GPT-4o	1	pass@10
2	COPRA (GPT-4o)💙	1	pass@1
3	CoqHammer💚	0	pass@1
4	Tactician (LSH)💚	0	pass@1

Figure 2: PutnamBench Leaderboard

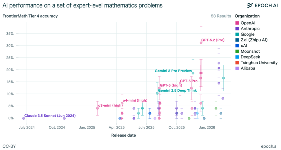
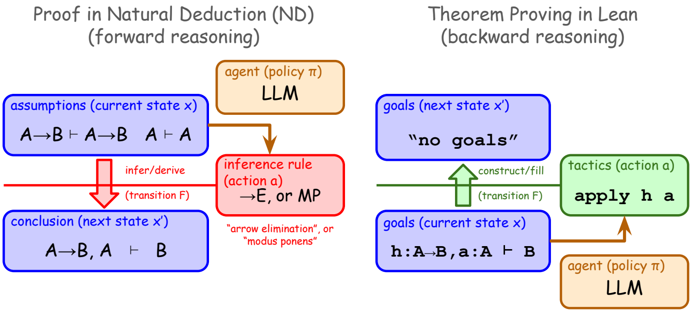


Figure 3: FrontierMath Tier-4 accuracy

Formulation: Interactive Theorem Proving as Markov Decision Process



ITP as MDP

state x	a set of open sequents / subgoals $\{(\Gamma_i \vdash \Delta_i)\}_{i=1}^{n}$
action a	an inference rule / tactic with parameters
state transition $F(x, a) \mapsto x'$	(deterministic) goal reduction by Lean kernel
solved state G	accepted proof / "no goals"
policy $\pi(a \mid x)$	LLM (or human) prover

New Concept: Statistical Provability := Goal-reaching Probability

For a policy π, initial distribution q_0, solved set G, and budget B, define the rollout trace

$$X_0 \sim q_0, \quad A_t \sim \pi(\cdot \mid X_t), \quad X_{t+1} = F(X_t, A_t),$$

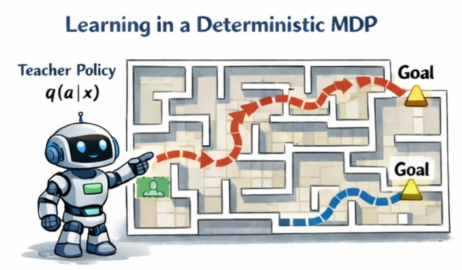
hitting time

$$T_{\mathsf{G}} := \inf\{t \geq 0 : X_t \in \mathsf{G}\}.$$

Definition (statistical-provability)

$$V_B^\pi(x) := \mathbb{P}^\pi(T_{\mathsf{G}} \leq B \mid X_0 = x), \qquad \mathrm{SP}_B^\pi(q_0) := \mathbb{E}_{X_0 \sim q_0}[V_B^\pi(X_0)].$$

Namely, the pointwise and average goal-reaching probabilities



Case I: Action-value regression and greedy search[3]

[3] S-Akiyama-Uezato, ICML2026 (preprint: 2602.10538)

Learning setting: action-value regression

For each remaining horizon $t = 1, \ldots, B$, train a value scorer separately.

Training data

$$(X_{t,i}, A_{t,i}) \overset{\text{i.i.d.}}{\sim} q_t \quad (i = 1, \ldots, N_t), \qquad \mathcal{D}_t = \{(X_{t,i}, A_{t,i}, Y_{t,i,1}, \ldots, Y_{t,i,m_t})\}_{i=1}^{N_t}.$$

- q_t: query distribution on the relevant state-action region $\mathcal{C}_t$
- N_t is the number of queried pairs sampled from q_t.
- m_t: number of independent rollout labels attached to each pair.
- $Y_{t,i,j} \in \{0, 1\}$: verifier-backed success label for the j-th rollout under the remaining budget.

Hypothesis space

$\mathcal{H}_t \subset \{h : \mathcal{C}_t \to [0, 1]\}$ is a class of leanring machines (e.g. LLMs)

Learning algorithm

With $\bar{Y}_{t,i} := m_t^{-1} \sum_{j=1}^{m_t} Y_{t,i,j}$,

$$\widehat{Q}_t \in \operatorname*{arg\,min}_{h \in \mathcal{H}_t} \max_{1 \le i \le N_t} |h(X_{t,i}, A_{t,i}) - \bar{Y}_{t,i}|.$$

Search setting: greedy search with action-value

Test-time pipeline

- Given the current state x,
- compute action values $\widehat{Q}_t(x, a)$ (over a candidate action sets $\mathcal{A}_t(x)$)
- draw the next action $\hat{\pi}_t(x) \in \arg\ \max_{a \in \mathcal{A}_t(x)} \widehat{Q}_t(x, a)$.

Main bound and error decomposition

With high probability over sampled query pairs and rollout labels,

$$\mathrm{SP}_B^{\hat{\pi}}(q_0) \geq \mathrm{SP}_B^{*}(q_0) - 2\sum_{s=1}^{B} \mathbb{P}_{q_0}^{\hat{\pi}}(T_{\mathsf{G}} \geq s)\, \varepsilon_{B-s+1},$$

where $\mathbb{P}_{q_0}^{\hat{\pi}}(T_{\mathsf{G}} \geq s)$ is the probability that the learned prover is still active at step s.

$$\varepsilon_t = \underbrace{\inf_{h\in\mathcal{H}_t} \|h - Q_t^*\|_{L_\infty(\mathcal{C}_t)}}_{\text{approximation}} + \underbrace{C_{\mathrm{cov}}(L_{H,t} + L_{Q,t})c_t^{-1/d_t}\left(\frac{\log(4N_t/\delta_t)}{N_t}\right)^{1/d_t}}_{\text{coverage / geometry}} + \underbrace{2\sqrt{\frac{\log(4N_t/\delta_t)}{2m_t}}}_{\text{rollout-label noise}}.$$

Reading as regret

$$\mathrm{SP}_B^{*}(q_0) - \mathrm{SP}_B^{\hat{\pi}}(q_0)$$

The gap is controlled by average search-length factor $\mathbb{P}_{q_0}^{\hat{\pi}}(T_{\mathsf{G}} \geq s)$ and the local value-estimation error at the remaining horizon.

Variables in the error bound

$$\varepsilon_t = \underbrace{\inf_{h\in\mathcal{H}_t} \|h - Q_t^*\|_{L_\infty(\mathcal{C}_t)}}_{\text{approximation}} + \underbrace{C_{\mathrm{cov}}(L_{H,t} + L_{Q,t})c_t^{-1/d_t}\left(\frac{\log(4N_t/\delta_t)}{N_t}\right)^{1/d_t}}_{\text{coverage / geometry}} + \underbrace{2\sqrt{\frac{\log(4N_t/\delta_t)}{2m_t}}}_{\text{rollout-label noise}}.$$

Value approximation

- $Q_t^*(x, a) = V_{t-1}^*(F(x, a))$: optimal success probability after trying action a.
- $\mathcal{C}_t$: relevant state-action region where the theorem is local.
- $\mathcal{H}_t$: scorer class used to approximate Q_t^* on $\mathcal{C}_t$.

Sampling, coverage, and smoothness

- q_t: query distribution on $\mathcal{C}_t$.
- N_t: number of sampled query pairs; m_t: rollout labels per pair.
- c_t, d_t: local coverage parameters, $q_t(B(z, r)) \geq c_t r^{d_t}$.
- $L_{H,t}, L_{Q,t}$: Lipschitz constants of $\mathcal{H}_t$ and Q_t^*; C_{cov}: geometric constant.

The error gets smaller with better approximation, denser coverage N_t, more rollout labels m_t, lower effective dimension d_t, and smoother value functions.

Assumptions behind the bound

- **Candidate retention:** for states where the learned prover is evaluated, $\mathcal{C}_t$ contains at least one optimal action.
- **Regularity:** Q_t^* and the hypothesis class $\mathcal{H}_t$ are Lipschitz on $\mathcal{C}_t$.
- **Coverage:** q_t puts enough local mass on metric balls in $\mathcal{C}_t$:

$$q_t(B(z, r)) \geq c_t r^{d_t}.$$

- **Rollout labels:**

$$\mathbb{E}[Y_{t,i,j} \mid X_{t,i}, A_{t,i}] = Q_t^*(X_{t,i}, A_{t,i}).$$

Where the regret bound comes from

If $\|\widehat{Q}_t - Q_t^*\|_\infty \le \varepsilon_t$, then for the greedy action $\hat{\pi}_t(x)$ and an optimal action $a_t^*(x)$,

$$Q_t^*(x, a_t^*(x)) - Q_t^*(x, \hat{\pi}_t(x)) \le 2\varepsilon_t.$$

Telescoping

The one-step greedy regret is accumulated along the learned trajectory:

$$\text{end-to-end regret} \le 2 \sum_{s=1}^{B} \mathbb{P}_{q_0}^{\hat{\pi}}(T_{\mathsf{G}} \ge s)\, \varepsilon_{B-s+1}.$$

Errors count only while the learned prover has not yet solved the instance.

Average proof length scales the regret

If $\sup_t \varepsilon_t \leq \bar{\varepsilon}_B$, then

$$\mathrm{SP}^*_B(q_0) - \mathrm{SP}^{\hat{\pi}}_B(q_0) \leq 2\bar{\varepsilon}_B \sum_{s=1}^{B} \mathbb{P}^{\hat{\pi}}_{q_0}(T_{\mathsf{G}} \geq s) = 2\bar{\varepsilon}_B \, \mathrm{Len}^{\hat{\pi}}_B(q_0).$$

Remark

- The relevant length is not only shortest proof length. It is the learned prover's average truncated path length through verifier states.
- The same local value error is less damaging if the prover usually solves early.

Margin condition: further improvements

Let $\Delta_t(x)$ be the gap between the best and second-best candidate action values at remaining horizon t. If near ties are rare,

$$P_s^{\hat{\pi}}(\Delta_t \le u) \le C_\Delta u^\beta, \qquad t = B - s + 1,$$

then the theorem improves to

$$\mathrm{SP}_B^{\hat{\pi}}(q_0) \ge \mathrm{SP}_B^*(q_0) - 2C_\Delta \sum_{s=1}^{B} \mathbb{P}_{q_0}^{\hat{\pi}}(T_\mathsf{G} \ge s)(2\varepsilon_{B-s+1})^{\beta+1}.$$

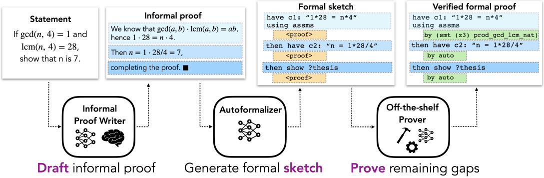
Only states near a decision boundary are fragile. If the correct continuation has a clear value gap, small score errors do not change the greedy action.

Case II: Hierarchical Teacher-Student setting[4]

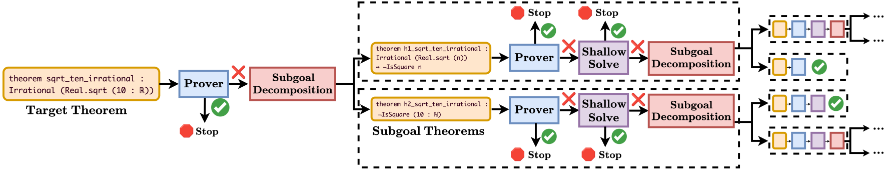
[4] S-Akiyama-Uezato, 2026 (preprint: 2602.10512)

Motivation: Structured proving is considered key in agentic provers

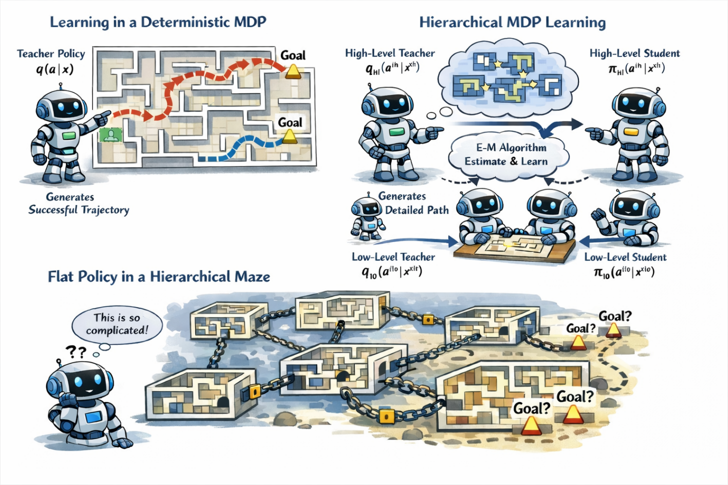
- **Draft–Sketch–Prove (ICLR2023):** draft informal proof → formal sketch → fill remaining gaps



- **DeepSeek-Provers**, **Prover Agent**, **Seed-Prover**, **Hilbert**, **Aristotle**, etc.: recursive subgoal decomposition and lemma reuse



Case II overview: three teacher-student protocols



- Setting 1: flat teacher/student; successful inlined traces.
- Setting 2: hierarchical teacher/student; proof DAGs and memoization.
- Setting 3: hierarchical teacher data unfolded for a flat student.

Common evaluation

$$\mathbb{E}_{s_0 \sim q_{0|\mathsf{Succ}}}[V_T(s_0)].$$

Case II result overview: sample size versus decision count

Provability bounds

Here L is an average decision count, η is intrinsic top-k loss, e is KL imitation error, and Δ_{graph} is graph-prediction failure.

$$\begin{aligned}
\text{Setting 1 (flat):} \qquad & \mathbb{E}[V_T^{\mathrm{flat}}] \gtrsim 1 - L_{\mathsf{flat}}\big(\eta_{\mathsf{flat}} + c_{\mathsf{flat}}\, e_{\mathsf{flat}}^{p_{\mathsf{flat}}}\big), \\
\text{Setting 2 (hierarchical):} \qquad & \mathbb{E}[V_T^{\mathrm{hier}}] \gtrsim 1 - \Delta_{\mathrm{graph}} - L_{\mathsf{hi}}\big(\eta_{\mathsf{hi}} + c_{\mathsf{hi}}\, e_{\mathsf{hi}}^{p_{\mathsf{hi}}}\big) - L_{\mathsf{lo}}\big(\eta_{\mathsf{lo}} + c_{\mathsf{lo}}\, e_{\mathsf{lo}}^{p_{\mathsf{lo}}}\big).
\end{aligned}$$

So when $L_{\mathsf{flat}} \gg L_{\mathsf{hi}} + L_{\mathsf{lo}}$, the hierarchical student can achieve a much higher success probability than the flat student.

Setting 3 (hier-teacher vs flat-student) length separation results in sample-complexity separation

Let N_{hier} and N_{flat} be the number of independent successful traces needed to achieve a given success probability δ in the two settings.
Then the ratio of sufficient sample sizes is lower-bounded by the ratio of decision counts:

$$\frac{\bar{N}_{\mathsf{flat}}(\delta)}{\bar{N}_{\mathsf{hier}}(\delta)} \gtrsim \left(\frac{L_{\mathsf{flat}}}{L_{\mathsf{hi}} + L_{\mathsf{lo}}}\right)^{1/(p\gamma)}.$$

So, when the hierarchical teacher has a much shorter proof than the flat teacher, i.e. $L_{\mathsf{flat}}/(L_{\mathsf{hi}} + L_{\mathsf{lo}}) = \exp(\Omega(D))$, it yields an exponential gap in certified sufficient sample size.

Conclusion and Notices

- **Statistical provability** $\mathrm{SP}_B^{\pi}(q_0)$: goal-reaching probability of policy π within budget B for theorem distribution q_0.
- **Case I:** provability gap is bounded by average length:

$$\mathrm{SP}_B^{*}(q_0) - \mathrm{SP}_B^{\hat{\pi}}(q_0) \leq 2\bar{\varepsilon}_B \sum_{s=1}^{B} \mathbb{P}_{q_0}^{\hat{\pi}}(T_{\mathsf{G}} \geq s) = 2\bar{\varepsilon}_B \,\mathrm{Len}_B^{\hat{\pi}}(q_0).$$

- **Case II:** hierarchical proof structures improve sample efficiency when $L_{\mathsf{flat}} \gg L_{\mathsf{hi}} + L_{\mathsf{lo}}$.

- **Notices:** We are planning AI4Math-related workshops:
- Nov 11-14, 2026 AI4Math 2026 (KSE2026 workshop) at Kanazawa (deadline: Jul 31)
- Nov 16-17, 2026 22nd Theorem Proving and Provers meeting (TPP2026) at RIKEN Tokyo office