AIM: Towards an Autonomous AI Mathematician

Peng Li

Institute for AI Industry Research (AIR)

Tsinghua University

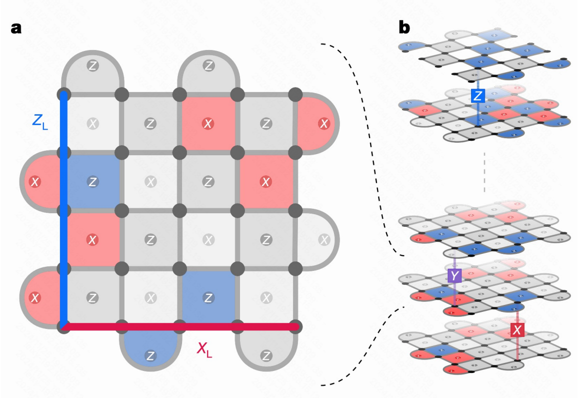
AI and Scientific Research

Tsinghua University

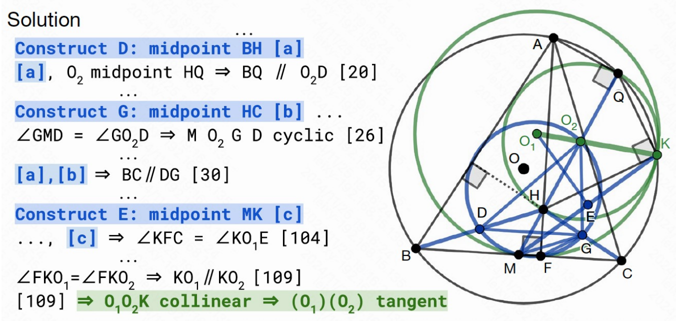
- **AI has significantly accelerated scientific progress.**

John J. Hopfield and Geoffrey Hinton. Ill. Niklas Elmehed © Nobel Prize Outreach

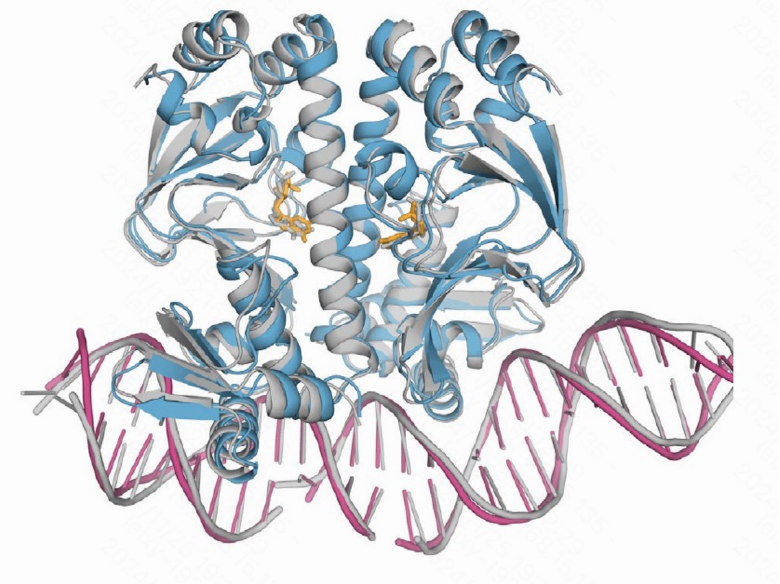
David Baker, Demis Hassabis and John Jumper. Ill. Niklas Elmehed © Nobel Prize Outreach



AlphaQuant



AlphaGeometry



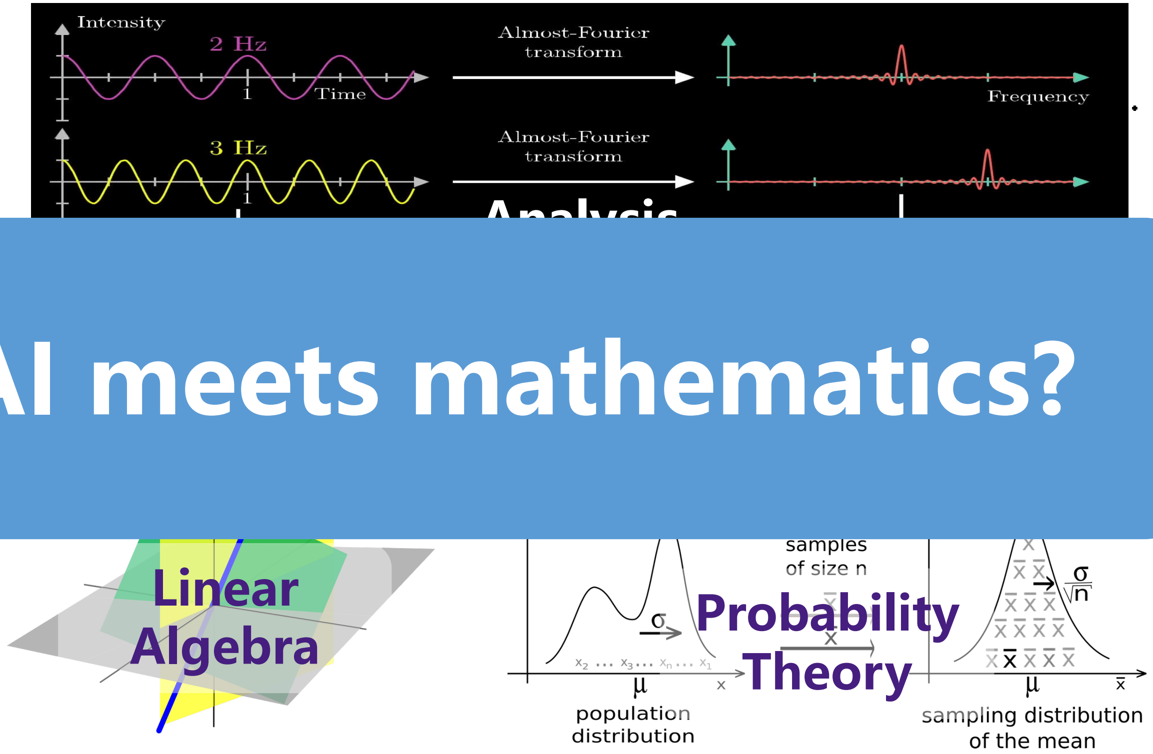
AlphaFold

https://www.nobelprize.org/all-nobel-prizes-2024/
https://www.nature.com/articles/s41586-024-07487-w
https://www.nature.com/articles/s41586-024-08148-8
https://www.nature.com/articles/s41586-023-06747-5

The Power of Math Research

Tsinghua University

- **Mathematical research embodies both profound theoretical and practical value, representing the pinnacle of human intellect.**



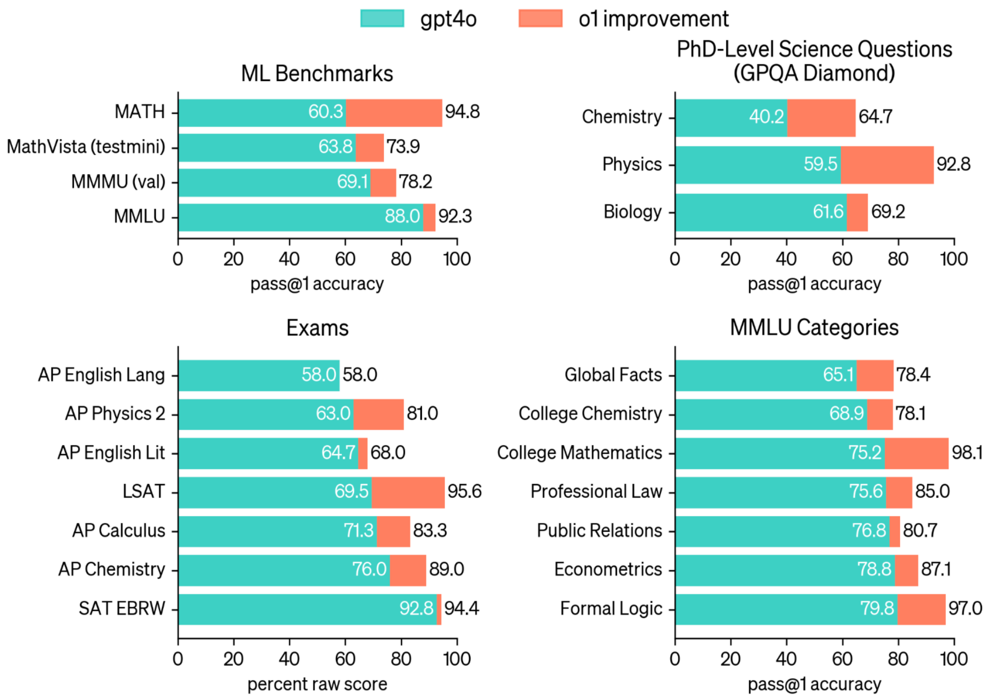
What sparks when AI meets mathematics?

In pursuit of truth and elegance in nature.

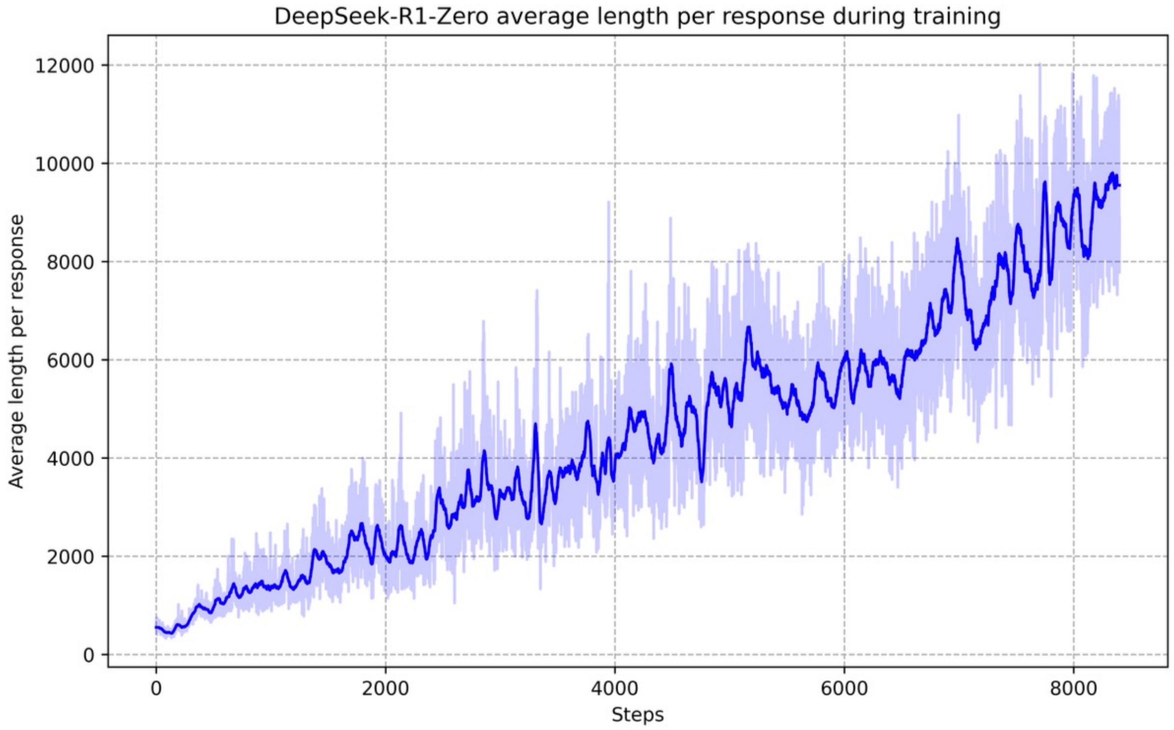
The theoretical foundation for science and engineering.

Math4AI: A Critical Factor

- **Mathematics and code are now key to advancing large models' reasoning capabilities.**



o1: The first reasoning model trained on math/code tasks, showing across-the-board improvements.



DeepSeek-R1 gains deep reasoning via verifiable math-focused RL training.

https://openai.com/index/learning-to-reason-with-llms ; DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. arXiv:2501.12948.

AI4Math: Approaching the Peak

- **By 2025, top models are projected to score 145/150 on Gaokao.**

Lack of sufficient image understanding capabilities leads to point loss

Model	Objective Questions (Text Input)	Single-choice Image Questions	Subjective Questions	Total
Gemini 2.5 pro	68	0	77	**145**
Doubao-1.5-thinking-vision-pro	68	0	76	**144**
DeepSeek R1	68	/	76	**144**
o3	65	0	75	**140**
Qwen3-235b	68	/	71	**139**
hunyuan-t1-latest	68	/	68	**136**
Wenxin X1 Turbo	68	/	66	**134**

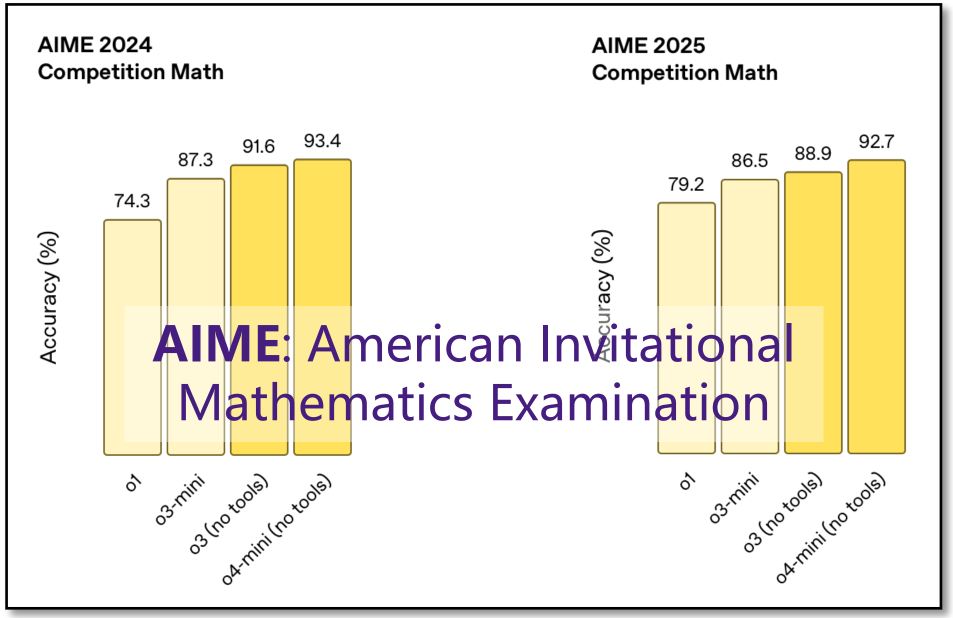
Note: This evaluation consists of three parts: objective questions (text input), Question 6 which is an image-based single-choice question, and subjective questions. The objective questions (text input) section accounts for a total of 68 points, the image-based single-choice question is worth 5 points, and the subjective questions total 77 points. The overall score is 150 points.

https://mp.weixin.qq.com/s/E1D_QA5lXwXRc6aSaU_zog?scene=1

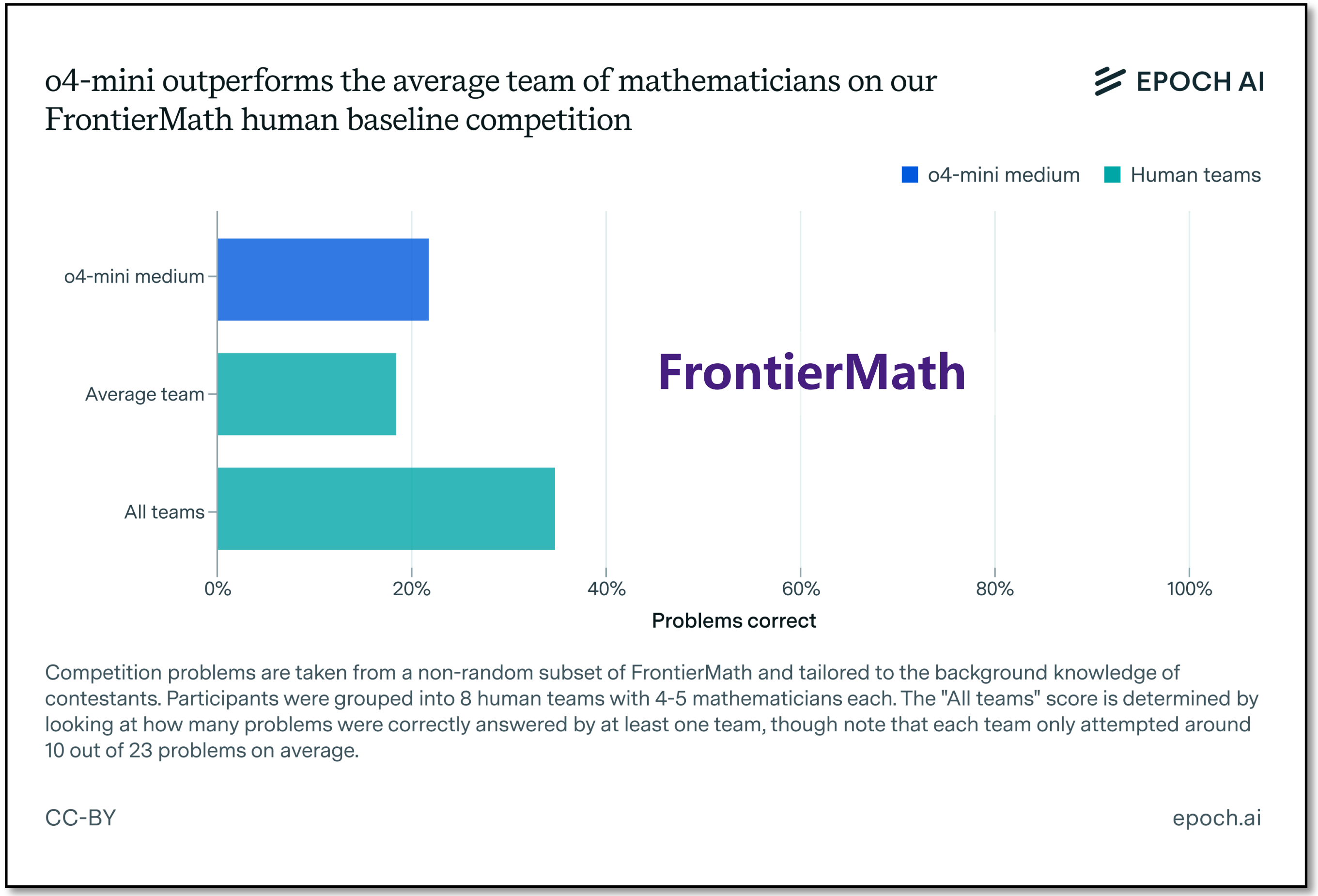
AI4Math: Approaching the Peak

Tsinghua University

- **o4-mini achieves expert-level proficiency on frontier math problems.**



"I don't want to add to the hysteria, but in some ways these large language models are already outperforming most of our best graduate students in the world." -- Ken Ono



https://www.scientificamerican.com/article/inside-the-secret-meeting-where-mathematicians-struggled-to-outsmart-ai/ ; https://epoch.ai/gradient-updates/is-ai-already-superhuman-on-frontiermath ; https://epoch.ai/gradient-updates/beyond-benchmark-scores-analysing-o3-mini-math-reasoning ; https://openai.com/index/introducing-o3-and-o4-mini/

AIME: High School Competition

清華大學 Tsinghua University

- **The AIME dataset derives from a competition (AIME) for high school students that helps select the U.S. team for the IMO.**

Problem Structure

- 15 progressively difficult problems
- Answers are integers from 0-999

Limitations

- Predetermined unique solutions
- Focuses primarily on " small" techniques

Problem

The 9 members of a baseball team went to an ice-cream parlor after their game. Each player had a single scoop cone of chocolate, vanilla, or strawberry ice cream. At least one player chose each flavor, and the number of players who chose chocolate was greater than the number of players who chose vanilla, which was greater than the number of players who chose strawberry. Let N be the number of different assignments of flavors to players that meet these conditions. Find the remainder when N is divided by 1000.

Solution 1

Let c be the number of players who choose chocolate, v be the number of players who choose vanilla, and s be the number of players who choose strawberry ice cream. We are given two pieces of information $c, v, s \geq 1$ and $c + v + s = 9$.

By inspection the only solutions for (c, v, s) are $(2, 3, 4), (1, 2, 6), (1, 3, 5)$.

Now we must choose which player chooses which flavor. For the general case (c, v, s), we begin by choose c of the 9 players who eat chocolate, then we choose v of the $9 - c$ players who vanilla, after this the amount of players who eat strawberry is fixed. Therefore the general formula is $\binom{9}{c}\binom{9-c}{v}$.

Therefore our final answer is,

$$\binom{9}{2}\binom{7}{3} + \binom{9}{1}\binom{8}{2} + \binom{9}{1}\binom{8}{3} = 2\boxed{016}.$$

https://artofproblemsolving.com/wiki/index.php/2025_AIME_I_Problems/Problem_3 ; https://zh.wikipedia.org/wiki/美国数学邀请赛

AIME: High School Competition

Tsinghua University

- **SOTA models primarily train and test on verifiable practice problems and math competition questions, similar to AIME' s data distribution.**

DeepSeek-R1 successfully solves the first problem in AIME 2025

- **Problem**: Find the sum of all integer bases $\mathrm{b} > 9$ for which 17_{b} is a divisor of 97_{b}
- **Complexity**: Simple and no advanced concepts required
- **Reasoning Length**: Solvable within brief deductive steps
- **Assessment Focus**: Core emphasis on applied problem-solving techniques

The problem requires finding all integer bases $b > 9$ such that 17_b divides 97_b, and then summing those bases.

Convert 17_b and 97_b to base 10:

- $17_b = 1 \cdot b + 7 = b + 7$
- $97_b = 9 \cdot b + 7$

The condition that 17_b divides 97_b means that $b + 7$ divides $9b + 7$, or $(b+7) \mid (9b+7)$.

Set $d = b + 7$, so $b = d - 7$. Substitute into $9b + 7$:

$$9b + 7 = 9(d-7) + 7 = 9d - 63 + 7 = 9d - 56$$

The condition becomes $d \mid (9d - 56)$. Since $d \mid 9d$, it follows that $d \mid (9d - (9d - 56)) = d \mid 56$.

Given $b > 9$, it follows that $d = b + 7 > 16$. Thus, d must be a positive divisor of 56 greater than 16.

The positive divisors of 56 are 1, 2, 4, 7, 8, 14, 28, and 56. The divisors greater than 16 are 28 and 56.

- If $d = 28$, then $b = 28 - 7 = 21$.
- If $d = 56$, then $b = 56 - 7 = 49$.

Both bases satisfy $b > 9$. Verify the division condition:

- For $b = 21$:
  - $17_{21} = 1 \cdot 21 + 7 = 28$
  - $97_{21} = 9 \cdot 21 + 7 = 189 + 7 = 196$
  - $196 \div 28 = 7$, which is an integer.
- For $b = 49$:
  - $17_{49} = 1 \cdot 49 + 7 = 56$
  - $97_{49} = 9 \cdot 49 + 7 = 441 + 7 = 448$
  - $448 \div 56 = 8$, which is an integer.

Both bases are valid, and there are no other bases $b > 9$ that satisfy the condition. The digits in both numbers are valid since $b > 9$ implies $b > 9$ for digit 9 in 97_b and $b > 7$ for digit 7 in both numbers.

The sum of the bases is $21 + 49 = 70$.

\boxed{70}

The Challenging FrontierMath

清華大學 Tsinghua University

- **Authored and endorsed by professional mathematicians (including Terence Tao), FrontierMath exhibits extremely high difficulty.**

Core Features

- Authored by professional mathematicians
- Spans major modern mathematical branches
- Requires hours per problem for specialists

Limitations

- Predetermined solutions
- Complexity ceiling below research-grade problems

Construct a degree 19 polynomial $p(x) \in \mathbb{C}[x]$ such that $X := \{p(x) = p(y)\} \subset \mathbb{P}^1 \times \mathbb{P}^1$ has at least 3 (but not all linear) irreducible components over $\mathbb{C}$. Choose $p(x)$ to be odd, monic, have real coefficients and linear coefficient -19 and calculate $p(19)$.

Answer: 1876572071974094803391179

MSC classification: 14 Algebraic geometry; 20 Group theory and generalizations; 11 Number theory generalizations

Glazer et al., 2024. FrontierMath: A Benchmark for Evaluating Advanced Mathematical Reasoning in AI. arXiv:2411.04872.

Competition v.s. Research

Tsinghua University

Key 1: Long Reasoning

Math research requires significantly longer and deeper reasoning.

Kakeya Set Conjecture in 3D

- 12 chapters with 125 pages of core arguments.
- Only the final proof is published, but the underlying reasoning is much more extensive.

Volume estimates for unions of convex sets, and the Kakeya set conjecture in three dimensions

Hong Wang* Joshua Zahl †

February 26, 2025

Abstract

We study sets of δ tubes in $\mathbb{R}^3$, with the property that not too many inside a common convex set V. We show that the union of tubes fro almost maximal volume. As a consequence, we prove that every Kakeya and Hausdorff dimension 3.

Contents

*Courant Institute of Mathematical Sciences, New York University. New York, NY, USA.
†Department of Mathematics, The University of British Columbia. Vancouver, BC, Canada.

Competition v.s. Research

Key 2: Procedural Rigor

Zero tolerance for errors. Verification is mandatory yet exceptionally costly

Hence if $\mathcal{W}' \subset \mathcal{W}_0$, to compare $\#\left(\bigcup_{W_i \in \mathcal{W}'} \mathcal{U}_0[W_i]\right)$ and $\sum_{W_i \in \mathcal{W}'} \#\mathcal{U}_0[W_i]$,

$$\kappa_0 \frac{C_{KT\text{-}CW}(\mathcal{U}_0)}{|U|} \sum_{W_i \in \mathcal{W}'} |W_i| \leq \sum_{W_i \in \mathcal{W}'} \#\mathcal{U}_{i-1}[W_i] = \#\left(\bigsqcup_{W_i \in \mathcal{W}'} \mathcal{U}_{i-1}[W_i]\right)$$
$$\leq \#\left(\bigcup_{W_i \in \mathcal{W}'} \mathcal{U}_0[W_i]\right) \leq \sum_{W_i \in \mathcal{W}'} \#\mathcal{U}_0[W_i] \leq \frac{C_{KT\text{-}CW}(\mathcal{U}_0)}{|U|} \sum_{W_i \in \mathcal{W}'} |W_i|. \tag{4.12}$$

The equality in (4.12) uses the critical fact that if $i \neq i'$, then $\mathcal{U}_{i-1}[W_i]$ and $\mathcal{U}_{i'-1}[W_{i'}]$ are disjoint.

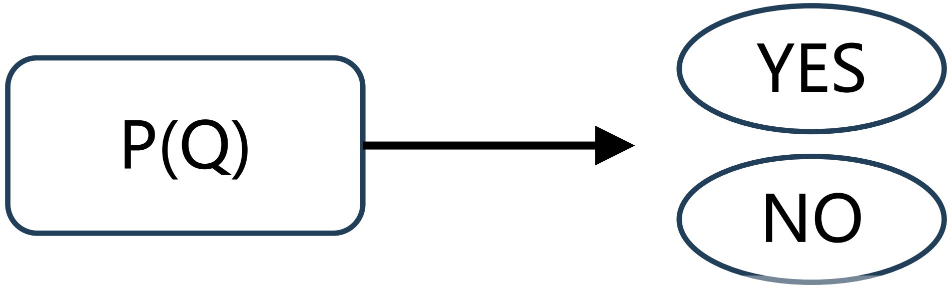
A proof fragment of the 3D Kakeya Set Conjecture

Case	Validating Process	Duration
Fermat's Last Theorem	The initial proof was presented in 1993, and then a flaw was identified after scrutinized by top mathematicians. The revised proof was finally published in 1995.	~2 years
Poincaré Conjecture	The proof was released in three preprints, and was finally validated by the collaborative effort by geometers and topologists.	~4-5 years

Competition v.s. Research

Key 3: Answer Existence

Competition problems have **known** and **determined** answers
The answers of research problems are **unknown** and **undetermined**



Halting Problem

Kurt Gödel

Gödel's First Incompleteness Theorem:
In any consistent formal system that is strong enough to express basic arithmetic, there will always be true statements that the system cannot prove. The system is inherently incomplete.

Gödel's Second Incompleteness Theorem:

Such a system cannot prove its own consistency from within. To establish its consistency, we must rely on methods outside the system.

AI4Math: Formal Proving

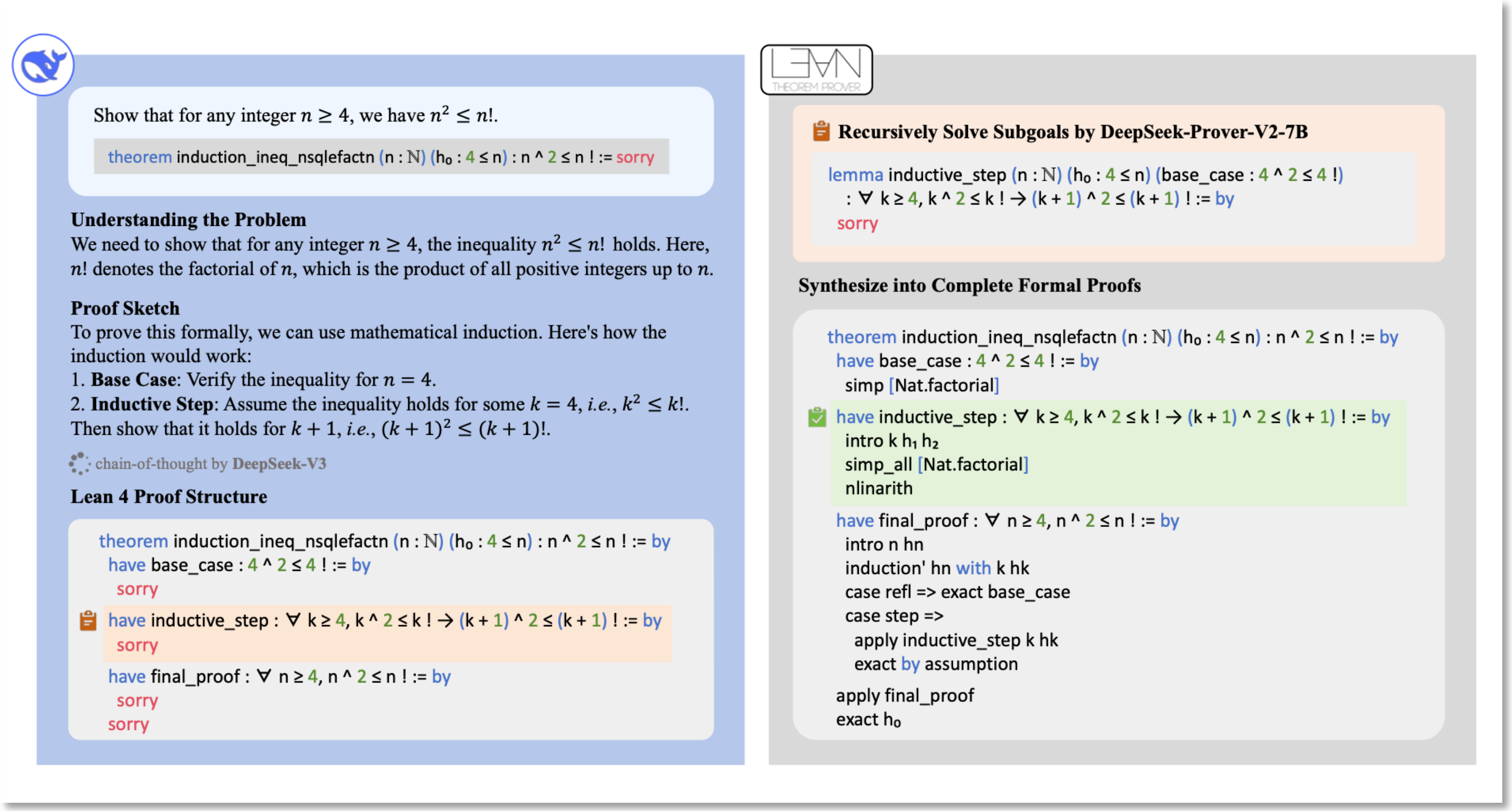
- **Formal proof systems have evolved through decades of research, with institutions worldwide - including DeepSeek and ByteDance - actively advancing this field.**

Pros

- Provides certain symbolic automation
- Guarantees absolute rigor via machine verification

Cons

- Current formal tools faces severe performance bottlenecks
- They can not be applied to research level problems



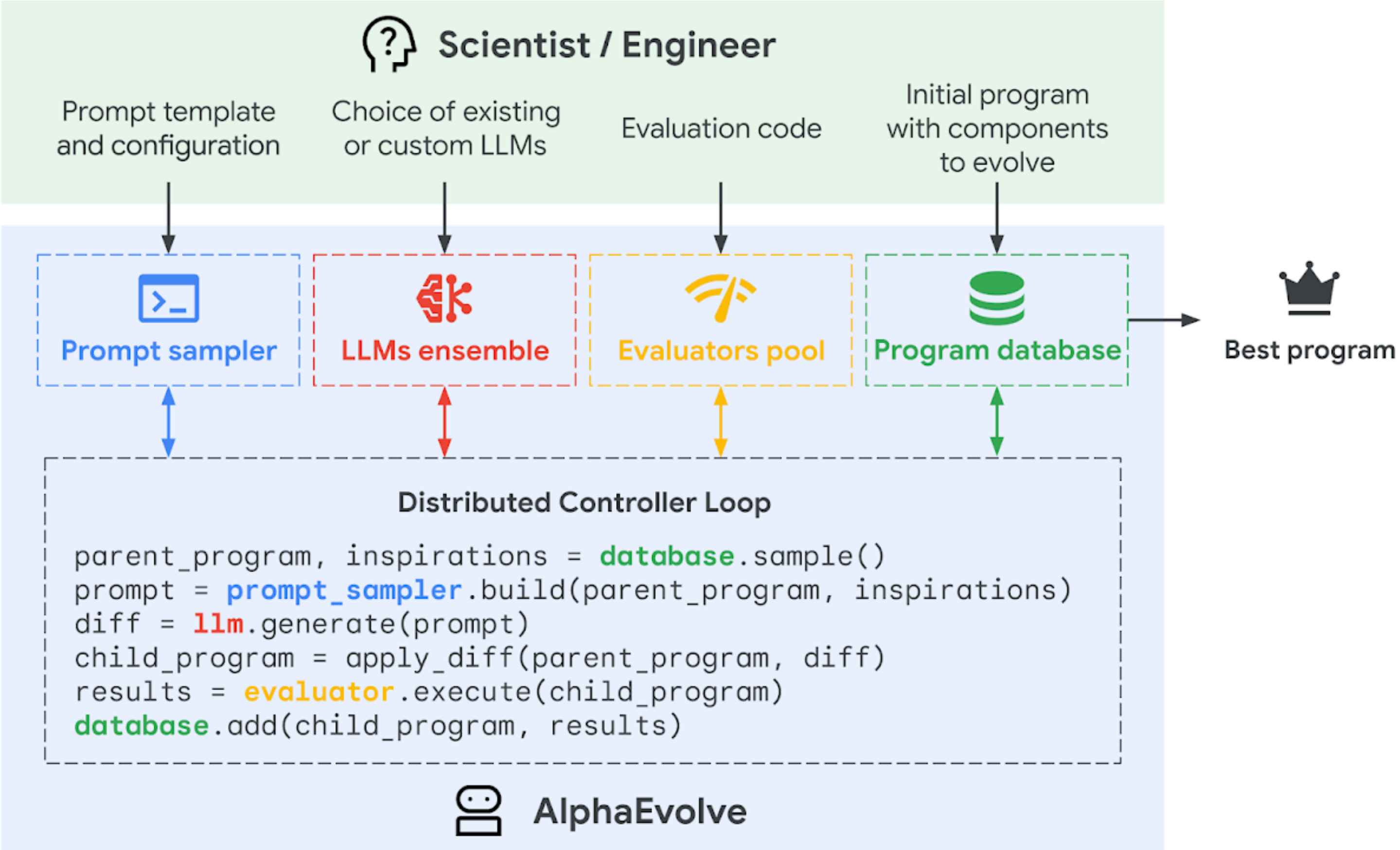
https://deepmind.google/discover/blog/alphaevolve-a-gemini-powered-coding-agent-for-designing-advanced-algorithms/

AI4Math: Coding Agent

- **Recent studies have begun exploring the potential of using large models to perform mathematical research tasks.**

AlphaEvolve

- **Origin**: Launched by Google DeepMind on May 14, 2025
- **Nature**: An agent system specialized in algorithmic optimization
- **Achievement**: Demonstrated capacity for independent novel discoveries
- **Limitation**: Limited to problems amenable to code



https://deepmind.google/discover/blog/alphaevolve-a-gemini-powered-coding-agent-for-designing-advanced-algorithms/

AIM: AI Mathematician

Tsinghua University

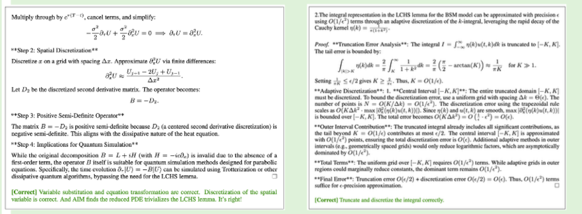
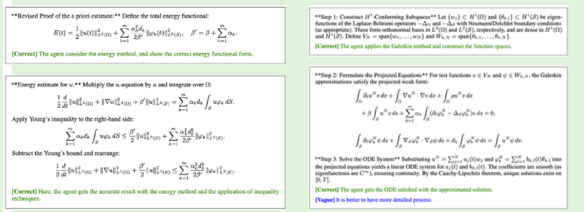
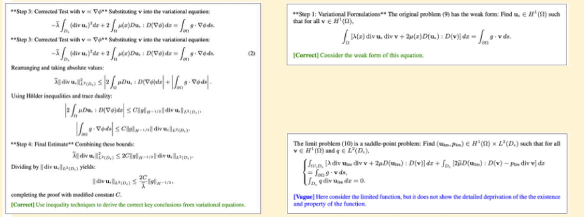
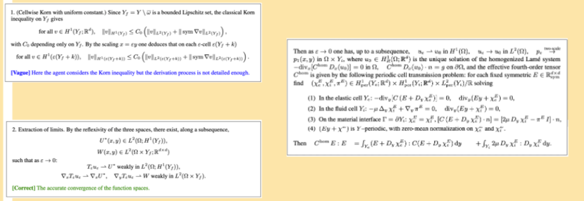
- **AIM represents an important step towards automated math research**

Competition

- Short reasoning
- Rigor requirement is easily satisfied
- Known existance of the answer
- Deterministic

Research

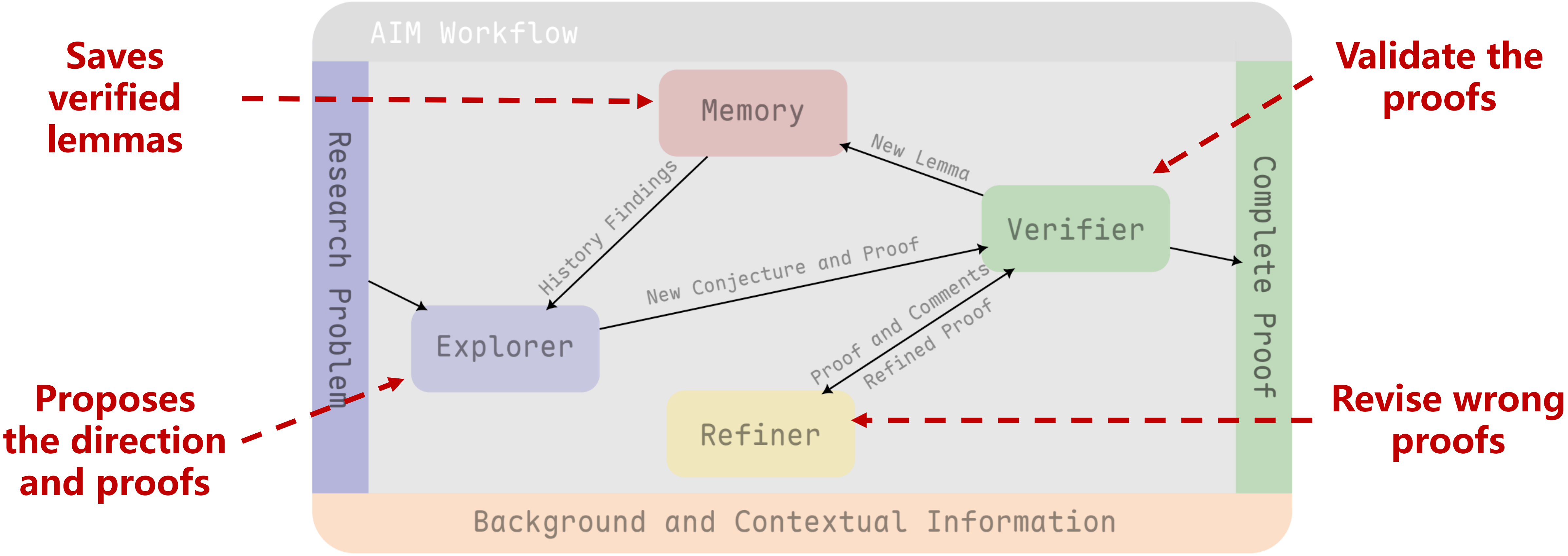
- Long reasoning
- Rigor requirement is hardly satisfied
- Unknown existance of the answer
- Indeterministic

Research Problem	Results of AIM	Selected Proof Steps
Quantum Algorithm Problem (Settled Problem)	AIM effectively completes the problem with a detailed solution process	
Absorbing Boundary Condition (Settled Problem)	AIM provides a substantially complete proof for this problem	
High Contrast Limit (Settled Problem)	AIM completes main proof of the conclusion and produces other correct results	
Homogenization Problem (Open Problem)	AIM provides partially correct conclusions and reasoning, offering instructive guidance	

Overview of AIM

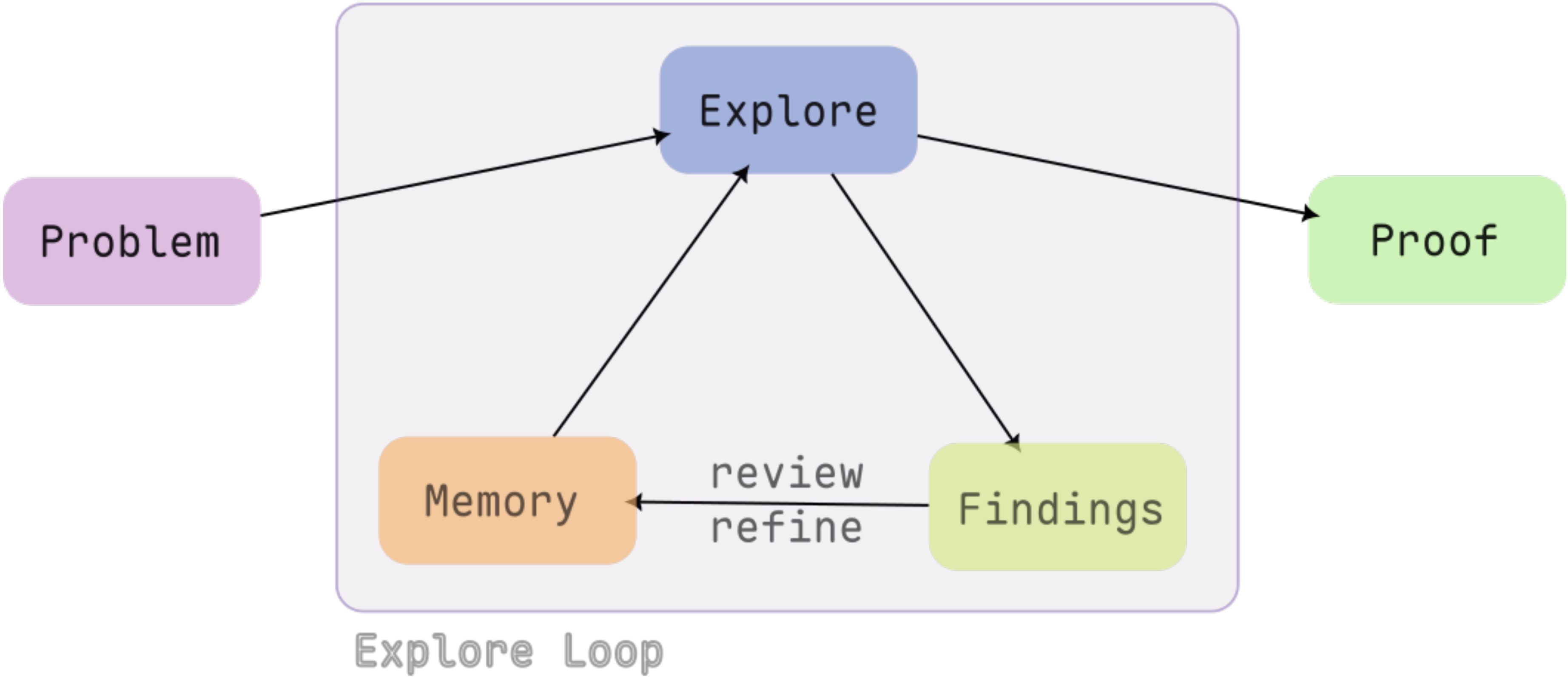
Tsinghua University

- **An LLM-powered agent system specially designed for math research, consisting of three agents (Explorer, Verifier, Refiner) and a Memory.**



Long Reasoning: Explore & Memory

- **The agents within AIM are tasked with exploring the original problem and documenting their discoveries as lemmas during the process. By iterating this procedure, the exploration progresses further, ultimately achieving problem resolution.**

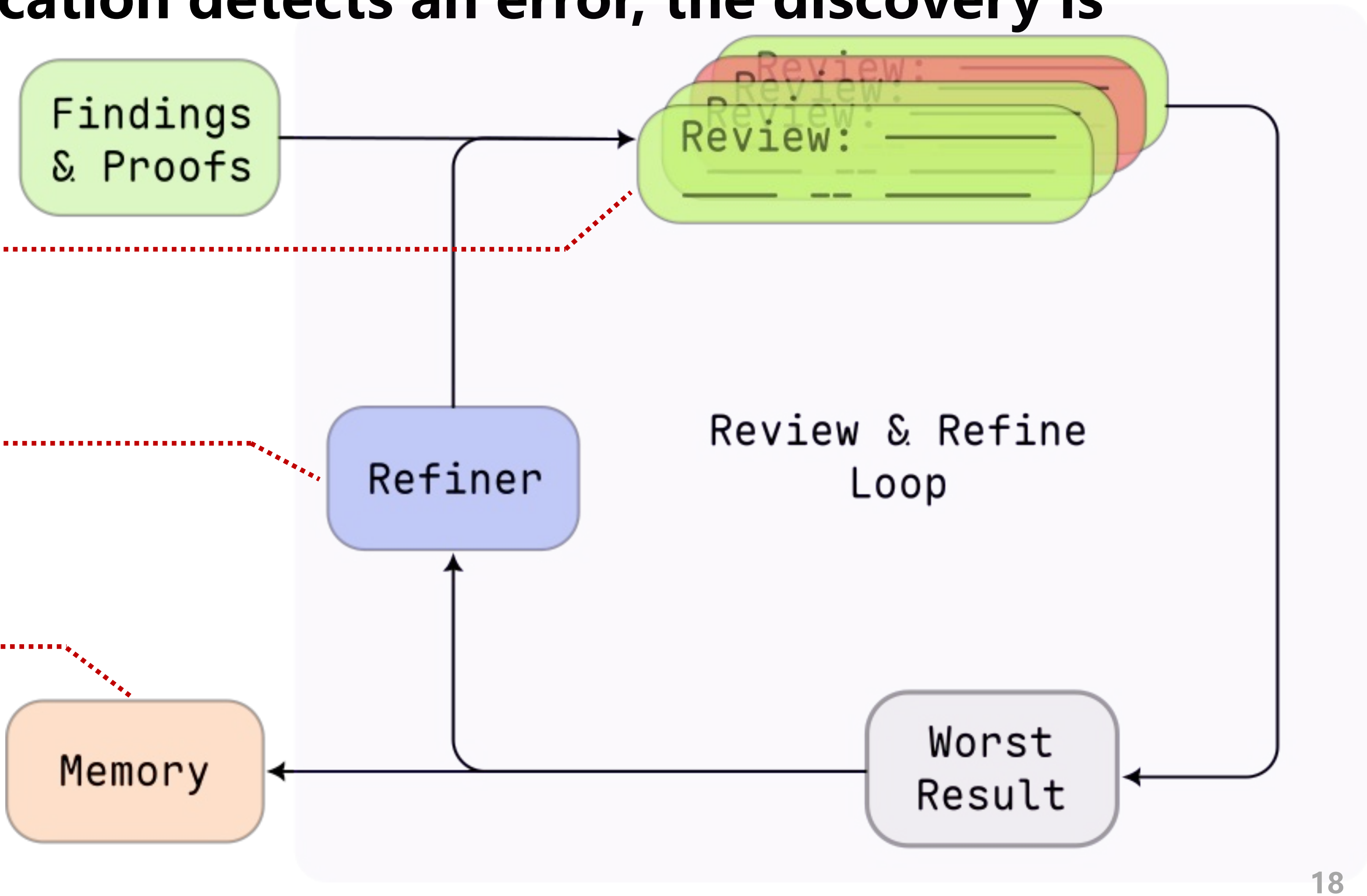


Validation: Pessimistic Verification

Tsinghua University

- **Each new discovery during exploration is repeatedly verified multiple times; if any single verification detects an error, the discovery is deemed incorrect.**

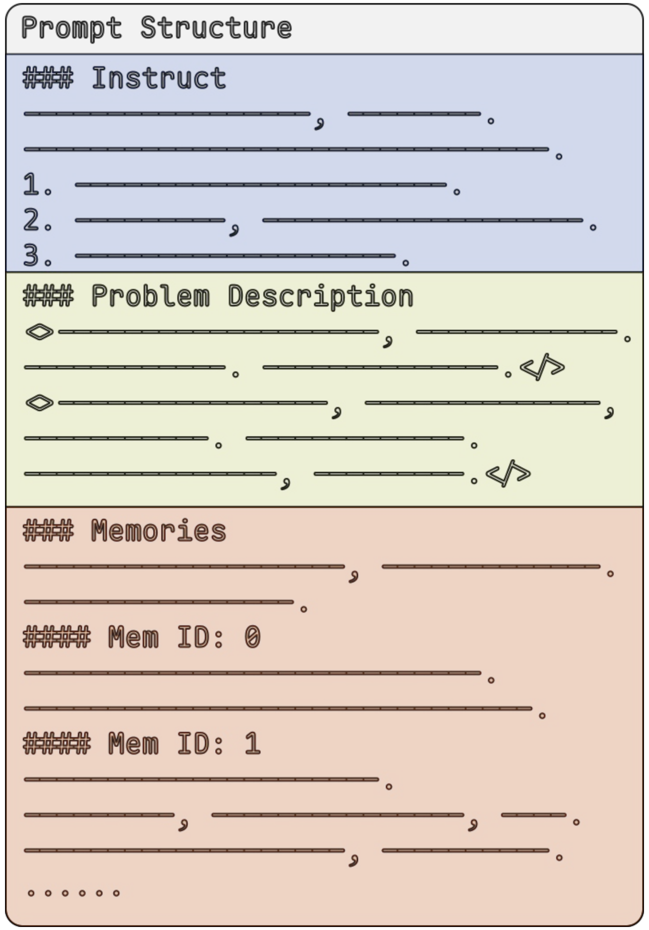
  - Conducting multiple rounds of self-checking can effectively identify issues
  - Erroneous proofs can be analyzed and corrected
  - Correct proof conclusions are eventually stored in memory



Agent Design in AIM

清華大學 Tsinghua University

- **All three agents are guided by meticulously designed prompts, supported by logical processing methods to facilitate collaboration.**



Define behavior of an agent

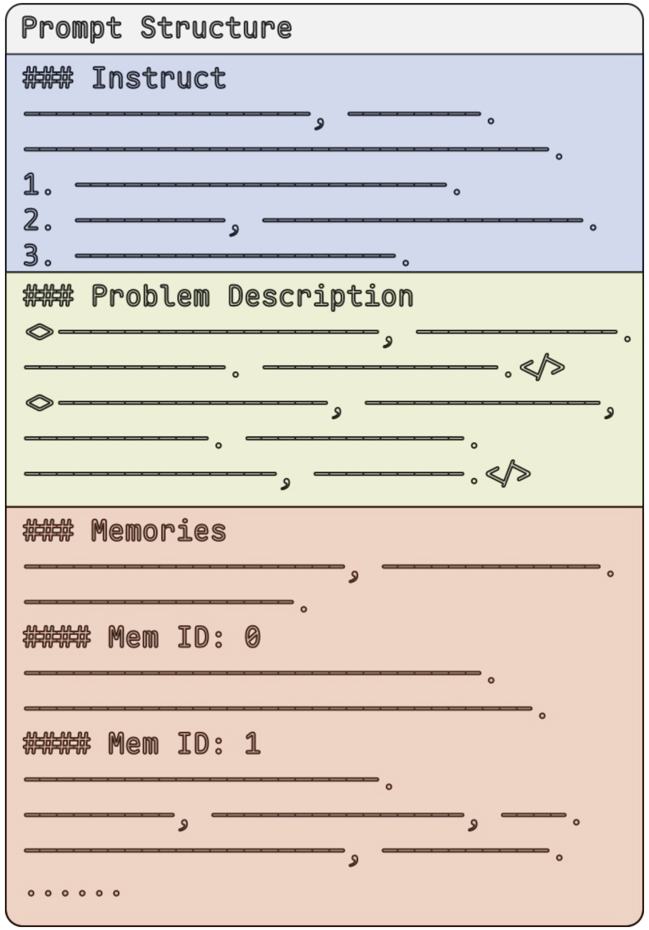
Prompt Example

Including task goal, reasoning direction, output format, etc.

```
  You are an expert that is knowledgeable across all domains in math. This time you are as
ked to help with our frontier math research. Its statement is as follows:
  This problem could be difficult and not able to be directly solved, but you can make you
r contribution with the following instructions:

  1. You are required to explore different approaches or directions that might help with o
ur final goal, and write down one interesting finding in your explorations as a new conjectu
re in your response. DO NOT claim that you can not do this jod.
  2. Your conjecture must contain the complete definitions required within it, such that i
t is able to stand alone as an independent lemma, unless it is declared in memory. Do not pr
opose any existing lemmas as your new conjectures. You can directly use them in your explora
tions.
  3. You should wrap your finding inside a latex environment: \begin{conjecture}\end{conje
cture}. This conjecture should be equiped with a detailed, complete and rigorous proof. You
should explicitly write down every intermediate derivation step in the proof. The correspond
ing proof should be wrapped in \begin{proof}\end{proof} directly followed by the conjecture.
  4. After these components you should also provide the dependency of this conjecture. You
 need to write down the memory IDs of lemmas used in this conjecture in a JSON array format,
 and warp them inside \begin{dependency}\end{dependency}. For example, a dependency of a new
 conjecture could be \begin{dependency}[0, 3, 4]\end{dependency}. You can use an empty array
 "[]" when this conjecture does not depend on other lemmas.

  More accurately, your response should obey the following format:

  \begin{conjecture}Your new findings here\end{conjecture}
  \begin{proof}Your proof of the conjecture above\end{proof}
  \begin{dependency}An json array of related memory IDs of this conjecture\end{dependency}
  Moreover, when you think the time is right that you are able to prove the original probl
```

Agent Design in AIM

- **All three agents are guided by meticulously designed prompts, supported by logical processing methods to facilitate collaboration.**



Define behavior of an agent

Description of the problem

Prompt Example

Information directly related to the task, for each agent it could be:

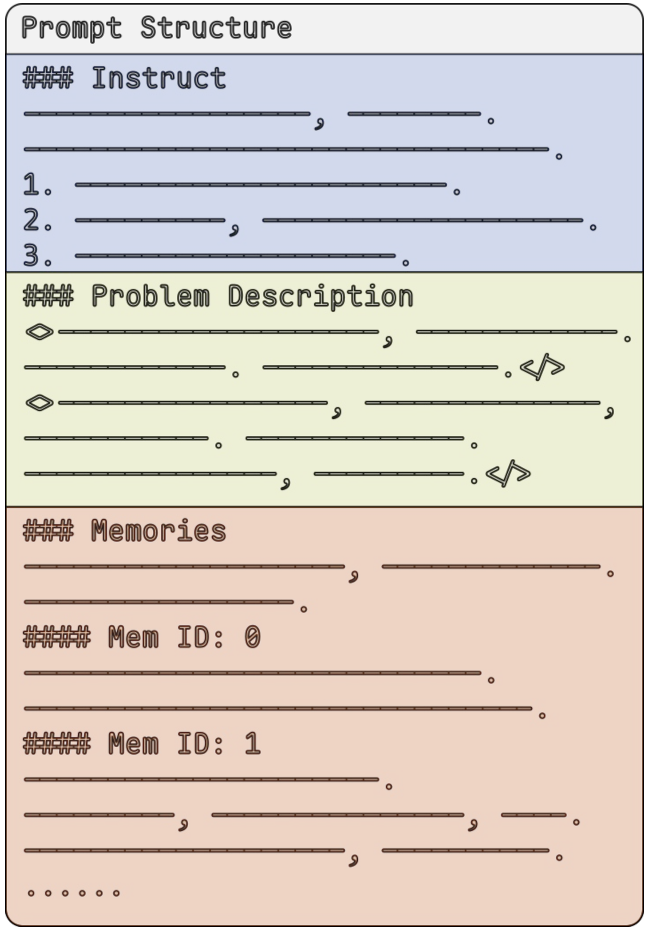
- Explorer: the statement of the final goal
- Verifier: a conjecture and its proof
- Refiner: a flawed conjecture, proof, and the feedback from the verifier

```
    \begin{problem}Question:

    Can we prove: for any $\delta>0$,

    \[
    \mathbb{P}[n^{-\delta} \mathbb{E}[Y_n]\leq Y_n\leq n^{\delta}\mathbb{E}[Y_n]]\geq 1-O(d^
{-n})
    \]

    If the second claim is not true, can we prove: for any $\delta>0$,

    \[
    \mathbb{P}[Y_n\leq n^{2+\delta}]\geq 1-O(d^{-n})
    \]
    \end{problem}
    This problem could be difficult and not able to be directly solved, but you can make you
r contribution with the following instructions:
```

Agent Design in AIM

Tsinghua University

- **All three agents are guided by meticulously designed prompts, supported by logical processing methods to facilitate collaboration.**



Define behavior of an agent

Description of the problem

Infomation in memory

Prompt Example

Formatted exploration history in memory

```
### Context and History Explorations

Here is a list of context that we have collected for this problem or our history finding
s during exploration. They serve as the background of the conjecture and proof and can be ac
cepted without controversy as correct.

#### Memory **ID: 1**

\begin{lemma}

There exists a constant \(A>0\) and a nonnegative random variable \(Y\) such that
\[
\forall k\ge1:\quad \EE[Y^k]\le k!\,A^k,
\]
yet for some \(t>0\),
\[
\PP(Y\ge t)\;>\;\exp\!\Bigl(-\frac{t}{2A}\Bigr).
\]
In other words, the bound
\[
\PP(Y\ge t)\le\exp\!\Bigl(-\frac{t}{2A}\Bigr)
\]
cannot hold for all \(t>0\) under only the moment hypothesis.


**DEPENDENCY**: []
\end{lemma}

#### Memory **ID: 2**
```

Memory Design in AIM

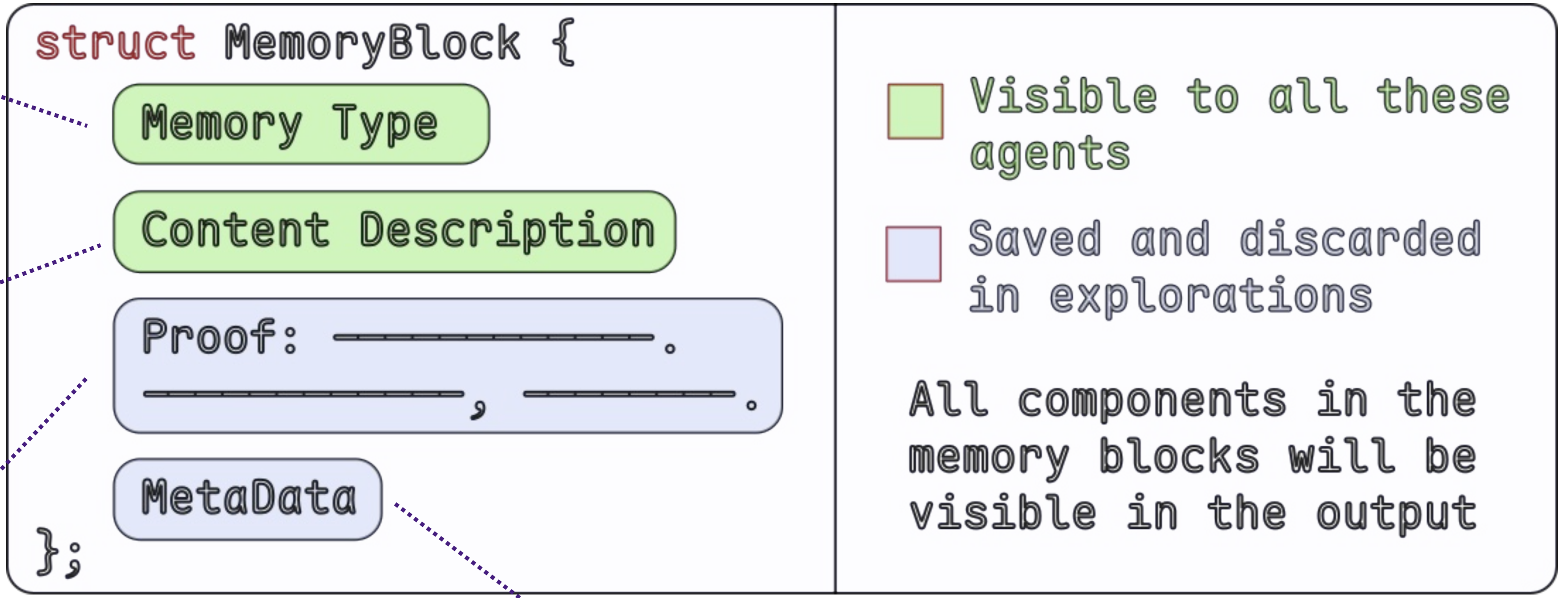
Tsinghua University

- **AIM parses and logs four categories of data from the model's outputs, with a portion of it being structured and fed into later agents' inputs.**

Memory type, e.g., context, lemma, conjecture, etc

Textual description of the memory content

A complete proof to this lemma



Other metadata of this memory, e.g., solved tag, num reviews, etc

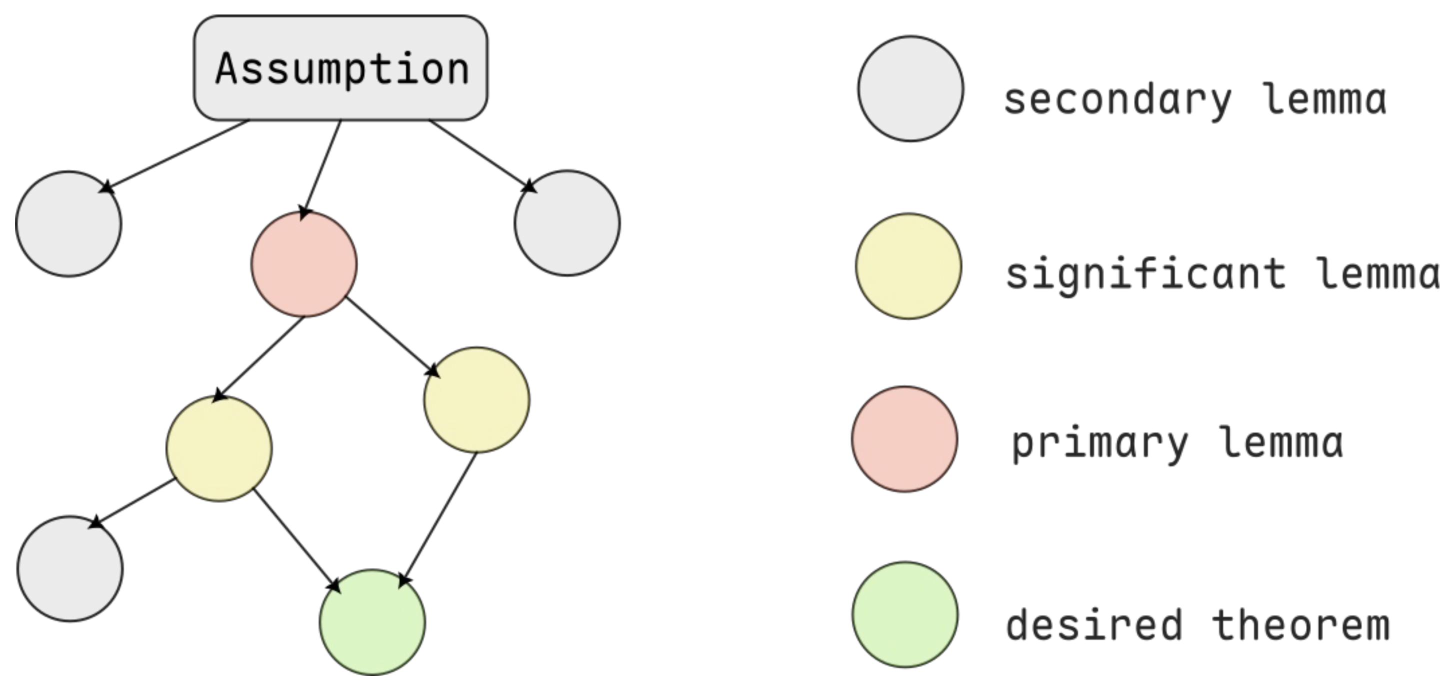
Lemma Graph

- **By arranging the lemmas according to their mutual dependencies, the exploration process can be structured into a lemma graph—essentially a directed graph that starts from the initial assumptions.**

This brings two critical improvements:

- ❑ Easy identification of actural proof path.
- ❑ Quantitative assessment of each lemmas. (By calculating the derivations of each node in the graph.)

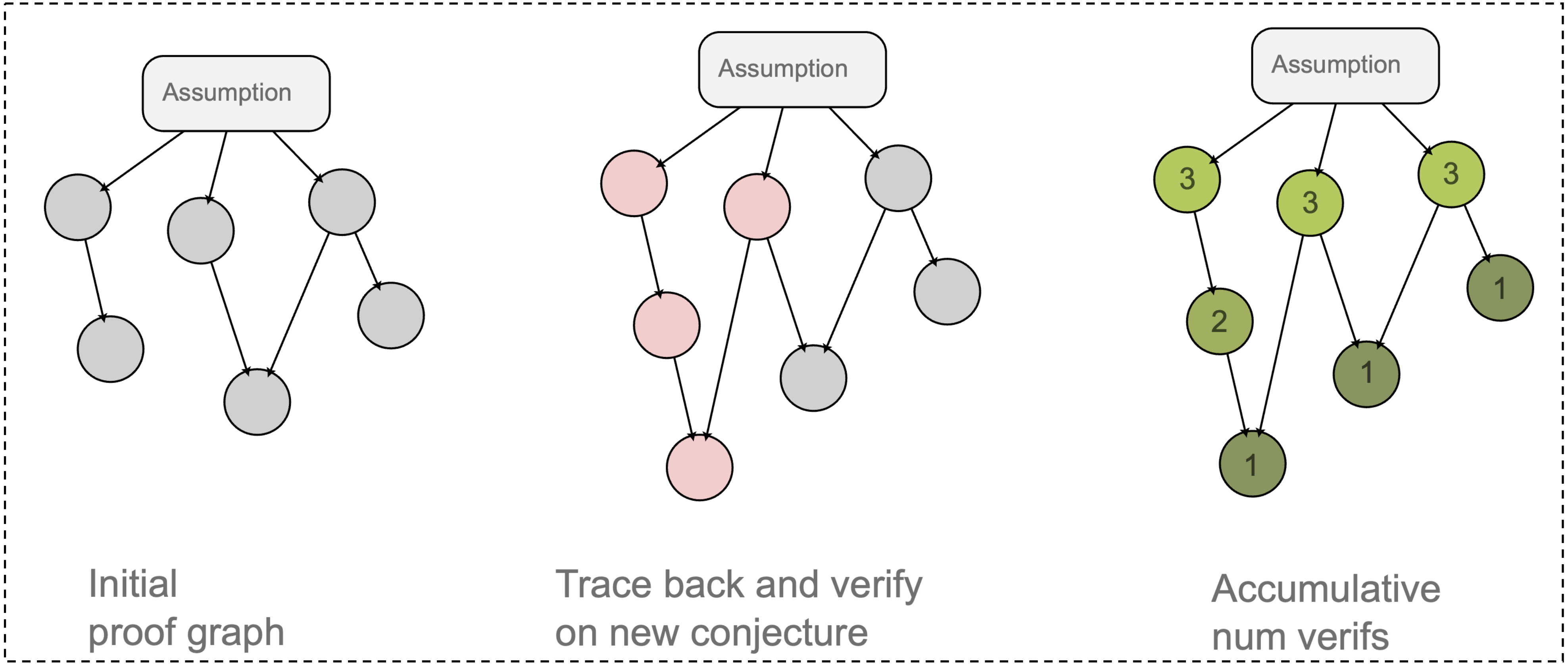
Example of Lemma Graph



Dynamic Verification in Lemma Graph

Tsinghua University

- **By dynamically allocating verification resources by the importance of each lemmas, we can further guarantee the reliability while increasing its efficiency.**



Overview of Experiments

清華大學 Tsinghua University

- **AIM is applied to address four mathematical theory problems, including three settled problems and one open problem.**

Research Problem	Results of AIM	Selected Proof Steps	
Quantum Algorithm Problem (Settled Problem)	AIM effectively completes the problem with a detailed solution process		DeepSeek-R1
Absorbing Boundary Condition (Settled Problem)	AIM provides a substantially complete proof for this problem		DeepSeek-R1
High Contrast Limit (Settled Problem)	AIM completes main proof of the conclusion and produces other correct results		DeepSeek-R1 and o4-mini, each conducted one experiment
Homogenization Problem (Open Problem)	AIM provides partially correct conclusions and reasoning, offering instructive guidance		o4-mini

Settled Problem → Quantum Algorithm Problem, Absorbing Boundary Condition, High Contrast Limit

Open Problem → Homogenization Problem

Quantum Algorithm Problem

Tsinghua University

The Linear Combination of Hamiltonian Simulation (LCHS) method is an efficient approach in scientific computing. Its main idea is to transform non-unitary dynamical problems into linear combinations of Hamiltonian simulation.

$$\mathcal{T}e^{-\int_0^t A(s)\,\mathrm{d}s} = \int_{\mathbb{R}} \frac{1}{\pi(1+k^2)} \mathcal{T}e^{-i\int_0^t (H(s)+kL(s))\,\mathrm{d}s}\,\mathrm{d}k$$

The Black-Scholes-Merton (BSM) model is the fundamental mathematical framework used for pricing European options in finance.

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

Objective: To simulate the BSM model using the LCHS method, design corresponding quantum algorithms, and analyze the complexity of the algorithms.

AIM correctly applies this method, providing a detailed proof and basically solving this problem

Quantum Algorithm Problem

Tsinghua University

System input

Lemma (Linear combination of Hamiltonian simulation, LCHS): For $t \in [0, T]$, let $A(t) \in \mathbb{C}^{N\times N}$ be decomposed into Hermitian and anti-Hermitian parts such that $A(t) = L(t) + iH(t)$, where $L(t) = \frac{1}{2}[A(t) + A^\dagger(t)]$ and $H(t) = \frac{1}{2i}[A(t) - A^\dagger(t)]$. Assume that $L(t)$ is positive semi-definite for all $t \in [0, T]$. Denoting the time ordering operator by $\mathcal{T}$, we have $\mathcal{T}\exp\left\{-\int_0^t A(s)ds\right\} = \int_{\mathbb{R}} \eta(k)u(t,k)dk$, where $u(t,k)$ is the propagator for a time-dependent Hamiltonian simulation problem such that $u(t,k) = \mathcal{T}\exp\left\{-i\int_0^t [H(s) + kL(s)]ds\right\}$, and $\eta(k) = \frac{1}{\pi(1+k^2)}$ is the kernel function with respect to k.

Explanation of the LCHS lemma

BSM model: The PDE in the BSM model is given by

$$\frac{\partial V(S,t)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Boundary conditions are characterized by: $V(0,t) = 0$ for all $0 < t \leq T$; $V(S,t) \to S - Ke^{-r(T-t)}$ for $S \to \infty$; $V(S,T) = \max\{S - K, 0\}$.

Explanation of the BSM model

Your tasks are to: Use the idea of LCHS to deal with the BSM model. Design a quantum algorithm to solve the equation. And compute the complexity (query complexity to state preparation oracle and operator oracles, gate complexity, ancilla qubits, (parameters: norm of boundary and initial, T, approximation error ϵ, etc)).

Objective

Quantum Algorithm Problem

Tsinghua University

- **AIM solves the problem by transforming the BSM equation into a heat equation, applying the LCHS lemma to reformulate the problem, and proposing a quantum algorithm. The core steps are as follows:**

	Step	Description
1	PDE transformation and spatial discretization	Transform the BSM equation into the standard heat equation using variable substitution, and discretize the spatial variables.
2	Operator decomposition and integral discretization	Apply the LCHS lemma to decompose the solution operator of the heat equation into a continuous linear combination of unitary operators. Then, truncate and approximate the continuous integral and discretize it.
3	Quantum algorithm implementation and complexity analysis	Design quantum algorithms using the linear combination of unitaries (LCU) and Hamiltonian simulation techniques, and calculate the complexity.

Quantum Algorithm Problem

清華大學 Tsinghua University

AIM explores step by step through a seq. of progressively built lemmas.

Lemma 1: First, AIM transforms the BSM equation into a heat equation, then discretizes the spatial variables, and verifies semi-definiteness. Further, AIM converts it into a form that can utilize the LCHS lemma, with a detailed proof.

1.The Black-Scholes-Merton (BSM) partial differential equation (PDE) can be transformed into the standard heat equation via a logarithmic substitution. Specifically, under the transformation $x = \ln(S/K) + (r - \sigma^2/2)(T-t)$, $\tau = \sigma^2(T-t)/2$, and $V(S,t) = e^{-r(T-t)}U(x,\tau)$, the BSM PDE reduces to:

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2}.$$

Discretizing this PDE spatially yields a matrix equation $\partial_\tau|U\rangle = -B|U\rangle$, where $B = -\partial_x^2$. The operator B is positive semi-definite and can be directly interpreted as $B = L$, where $L = -\partial_x^2$. While the original Hermitian decomposition $B = L + iH$ (with $H = 0$) trivializes the LCHS lemma, the dissipative nature of B permits alternative quantum simulation methods tailored for parabolic equations.

Transform the BSM equation into a heat equation

Proof. **Step 1: Transform BSM to the Heat Equation**

Start with the BSM PDE:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Apply the substitution $V(S,t) = e^{-r(T-t)}U(x,\tau)$, $x = \ln(S/K) + (r-\sigma^2/2)(T-t)$, and $\tau = \sigma^2(T-t)/2$. Compute derivatives rigorously: - $\partial_t V = e^{-r(T-t)}\left(rU - \frac{\sigma^2}{2}\partial_\tau U - (r-\sigma^2/2)\partial_x U\right)$, - $\partial_S V = e^{-r(T-t)}\partial_x U \cdot \frac{1}{S}$, - $\partial_S^2 V = e^{-r(T-t)}\left(\partial_x^2 U - \partial_x U\right) \cdot \frac{1}{S^2}$.

Substitute into the BSM PDE:

$$e^{-r(T-t)}\left[rU - \frac{\sigma^2}{2}\partial_\tau U - (r-\sigma^2/2)\partial_x U + \frac{1}{2}\sigma^2(\partial_x^2 U - \partial_x U) + r\partial_x U - rU\right] = 0.$$

Multiply through by $e^{r(T-t)}$, cancel terms, and simplify:

$$-\frac{\sigma^2}{2}\partial_\tau U + \frac{\sigma^2}{2}\partial_x^2 U = 0 \implies \partial_\tau U = \partial_x^2 U.$$

Step 2: Spatial Discretization

Discretize x on a grid with spacing Δx. Approximate $\partial_x^2 U$ via finite differences:

$$\partial_x^2 U \approx \frac{U_{j+1} - 2U_j + U_{j-1}}{\Delta x^2}.$$

Let D_2 be the discretized second derivative matrix. The operator becomes:

$$B = -D_2.$$

Discretize spatial variables

Step 3: Positive Semi-Definite Operator

The matrix $B = -D_2$ is positive semi-definite because D_2 (a centered second derivative discretization) is negative semi-definite. This aligns with the dissipative nature of the heat equation.

Step 4: Implications for Quantum Simulation

While the original decomposition $B = L + iH$ (with $H = -ic\partial_x$) is invalid due to the absence of a first-order term, the operator B itself is suitable for quantum simulation methods designed for parabolic equations. Specifically, the time evolution $\partial_\tau|U\rangle = -B|U\rangle$ can be simulated using Trotterization or other dissipative quantum algorithms, bypassing the need for the LCHS lemma. □

[Correct] Variable substitution and equation transformation are correct. Discretization of the spatial variable is correct. And AIM finds the reduced PDE trivializes the LCHS lemma. It's right!

Quantum Algorithm Problem

Tsinghua University

AIM explores step by step through a seq. of progressively built lemmas.

Lemma 2: AIM analyzes truncation errors and determines discrete intervals, laying the foundation for constructing quantum algorithms.

2.The integral representation in the LCHS lemma for the BSM model can be approximated with precision ϵ using $O(1/\epsilon^2)$ terms through an adaptive discretization of the k-integral, leveraging the rapid decay of the Cauchy kernel $\eta(k) = \frac{1}{\pi(1+k^2)}$.

Proof. **Truncation Error Analysis**: The integral $I = \int_{-\infty}^{\infty} \eta(k)u(t,k)dk$ is truncated to $[-K, K]$. The tail error is bounded by:

$$\int_{|k|>K} \eta(k)dk = \frac{2}{\pi}\int_K^\infty \frac{1}{1+k^2}dk = \frac{2}{\pi}\left(\frac{\pi}{2} - \arctan(K)\right) \approx \frac{1}{\pi K} \quad \text{for } K \gg 1.$$

Setting $\frac{1}{\pi K} \leq \epsilon/2$ gives $K \geq \frac{2}{\pi\epsilon}$. Thus, $K = O(1/\epsilon)$.

Adaptive Discretization: 1. **Central Interval $[-K, K]$**: The entire truncated domain $[-K, K]$ must be discretized. To bound the discretization error, use a uniform grid with spacing $\Delta k = \Theta(\epsilon)$. The number of points is $N = O(K/\Delta k) = O(1/\epsilon^2)$. The discretization error using the trapezoidal rule scales as $O(K\Delta k^2 \cdot \max|\partial_k^2(\eta(k)u(t,k))|)$. Since $\eta(k)$ and $u(t,k)$ are smooth, $\max|\partial_k^2(\eta(k)u(t,k))|$ is bounded over $[-K, K]$. The total error becomes $O(K\Delta k^2) = O\left(\frac{1}{\epsilon}\cdot\epsilon^2\right) = O(\epsilon)$.

Outer Interval Contribution: The truncated integral already includes all significant contributions, as the tail beyond $K = O(1/\epsilon)$ contributes at most $\epsilon/2$. The central interval $[-K, K]$ is approximated with $O(1/\epsilon^2)$ points, ensuring the total discretization error is $O(\epsilon)$. Additional adaptive methods in outer intervals (e.g., geometrically spaced grids) would only reduce logarithmic factors, which are asymptotically dominated by $O(1/\epsilon^2)$.

Total Terms: The uniform grid over $[-K, K]$ requires $O(1/\epsilon^2)$ terms. While adaptive grids in outer regions could marginally reduce constants, the dominant term remains $O(1/\epsilon^2)$.

Final Error: Truncation error $O(\epsilon/2)$ + discretization error $O(\epsilon/2) = O(\epsilon)$. Thus, $O(1/\epsilon^2)$ terms suffice for ϵ-precision approximation. □

[Correct] Truncate and discretize the integral correctly.

Analyze truncation error

Determine discrete intervals

Overall error analysis

Quantum Algorithm Problem

Tsinghua University

AIM explores step by step through a seq. of progressively built lemmas.

AIM combines previous lemmas to construct a quantum algorithm using Hamiltonian simulation and the linear combination of unitary operators (LCU), then calculates its complexity. There are some errors in the complexity calculation, but the core methods are correct.

3.The solution operator for the heat equation derived from the BSM model, $U(\tau) = e^{\tau B}$, can be approximated via the LCHS lemma as a linear combination of Hamiltonian simulations. Specifically, the integral representation $\int_{\mathbb{R}} \eta(k) e^{-ikL\tau} dk$ (where $L = -B$) can be discretized into $M = O(1/\epsilon^2)$ terms with quadrature weights [illegible]. The total query complexity to the L-oracle is $\tilde{O}(\|L\|T/\epsilon^3)$, where $T = \frac{2\tau}{\sigma^2}$, and the gate complexity scales polynomially with $\|L\|, T, 1/\epsilon$, and logarithmic factors.

Based on the first lemma, use the LCHS lemma to transform into a continuous integral of unitary operators

Proof:

Step 1: Integral Representation via LCHS From Lemma-0, the solution operator for the heat equation is $U(\tau) = e^{\tau B}$. Let $L = -B$, a positive semi-definite operator. By the LCHS lemma with $H = 0$, we express:

$$e^{\tau B} = \int_{\mathbb{R}} \eta(k) e^{-ikL\tau} dk, \quad \eta(k) = \frac{1}{\pi(1+k^2)}.$$

Step 2: Quadrature Discretization Lemma-1 guarantees the integral can be discretized into $M = O(1/\epsilon^2)$ terms with weights $\eta(k_j)\Delta k$, yielding:

$$e^{\tau B} \approx \sum_{j=1}^{M} \eta(k_j) \Delta k \cdot e^{-ik_j L\tau}.$$

The truncation and discretization error is $O(\epsilon)$ due to the $O(1/k^2)$ decay of $\eta(k)$.

[Correct] Properly utilizing the LCHS lemma and memory lemmas for problem transformation.

Transform the solution of the equation into a linear combination of unitary operators

Utilize Hamiltonian simulation algorithms

Step 3: Hamiltonian Simulation Costs Each term $e^{-ik_j L\tau}$ requires simulating L for time $t_j = k_j\tau$. Using optimal Hamiltonian simulation algorithms (e.g., quantum signal processing or Taylor series) with query complexity $O(\|L\|t_j + \log(1/\epsilon_{\text{sim}}))$, set $\epsilon_{\text{sim}} = \epsilon/M$ to bound the total simulation error by ϵ. Since $k_j = O(1/\epsilon)$ dominates the quadrature, $t_j = O(\tau/\epsilon)$. The per-term complexity becomes:

$$O(\|L\|\tau/\epsilon + \log(M/\epsilon)) = \tilde{O}(\|L\|\tau/\epsilon).$$

Step 4: Total Query Complexity Summing over $M = O(1/\epsilon^2)$ terms, the total queries to the L-oracle are:

$$\tilde{O}\left(M \cdot \frac{\|L\|\tau}{\epsilon}\right) = \tilde{O}\left(\frac{\|L\|\tau}{\epsilon^3}\right).$$

Expressing in terms of $T = \frac{2\tau}{\sigma^2}$ and absorbing constants:

$$\tilde{O}\left(\frac{\|L\|T}{\epsilon^3}\right).$$

Complexity analysis

Step 5: Ancilla and Gate Complexity The LCU framework requires $\log M = O(\log(1/\epsilon))$ ancilla qubits. Gate complexity inherits $\tilde{O}(M)$-scaling from the superposition state, augmented by $O(\text{poly}(\log(\|L\|T/\epsilon)))$ factors from simulation subroutines. The dominant terms remain polynomial in $\|L\|$, T, and $1/\epsilon$, with polylogarithmic corrections.

[Error] There are some mistakes about complexity computing. And the calculation process lacks detail.

Homogenization Problem

Tsinghua University

- **Problem Description:** **The homogenization problem for transmission systems requires analyzing the properties of equations and their corresponding solutions under specific physical scale limits. The goal is to ultimately prove the error estimation of solutions, which remains an open problem.**

$$
\begin{cases}
\mathcal{L}_{\lambda,\mu}\mathbf{u}_\epsilon = 0 & \text{in } \Omega \setminus \overline{D_\epsilon}, \\
\mathcal{L}_{\tilde{\mu}}(\mathbf{u}_\epsilon, p_\epsilon) = 0 \text{ and div } \mathbf{u}_\epsilon = 0 & \text{in } D_\epsilon, \\
\mathbf{u}_\epsilon|_- = \mathbf{u}_\epsilon|_+ \text{ and } \frac{\partial(\mathbf{u}_\epsilon, p_\epsilon)}{\partial\nu_{(\infty,\tilde{\mu})}}\Big|_- = \frac{\partial \mathbf{u}_\epsilon}{\partial\nu_{(\lambda,\mu)}}\Big|_+ & \text{on } \partial D_\epsilon, \\
\frac{\partial \mathbf{u}_\epsilon}{\partial\nu_{(\lambda,\mu)}}\Big|_{\partial\Omega} = g \in H_{\mathbb{R}}^{-\frac{1}{2}}(\partial\Omega) \quad \text{and} \quad \mathbf{u}_\epsilon|_{\partial\Omega} \in H_{\mathbb{R}}^{\frac{1}{2}}(\partial\Omega), &
\end{cases}
$$

where $\mathcal{L}_{\tilde{\mu}}(\mathbf{u}_\epsilon, p_\epsilon) = \tilde{\mu}\Delta\mathbf{u}_\epsilon + \nabla p_\epsilon$ denotes the Stokes operator with viscosity constant $\tilde{\mu}$ and p_ϵ is the pressure field. Its exterior derivative is defined as $\frac{\partial(u,p)}{\partial\nu_{(\infty,\mu)}}\Big|_- := pN + 2\mu D(u)N$.

Now we let $\epsilon \to 0$, so that the scale of the cell is tend to be zero. We wonder what homogenization equation does the limited solution $\mathbf{u}_{lim}$ satisfies. At the same time, whether the estimate between the original solution and the limited solution has the form $C\epsilon^\alpha \|g\|_{H^{-\frac{1}{2}}(\partial\Omega)}$ for some $\alpha \in (0,1)$.

- **AIM correctly applied mathematical theorems and analytical derivations, providing some accurate conclusions and approaches.**

- **The results are of guiding significance for human mathematicians conducting this research.**

6. Error estimate and strong convergence Define the two-scale corrector

$$u_\varepsilon^{\mathrm{app}}(x) = u_0(x) + \varepsilon\chi^{D_x u_0(x)}(x/\varepsilon) \text{ in } \Omega_\varepsilon,$$

extended in each D_ε by $Ey + \chi_i^E(y)$. A Céa-lemma–type argument in the perforated domain (again relying on uniform Korn and Stokes inf–sup, Lemma A.1–A.2) yields for some $\alpha \in (0,1)$

$$\|u_\varepsilon - u_\varepsilon^{\mathrm{app}}\|_{H^1(\Omega_\varepsilon)} \le C\varepsilon^\alpha \|g\|_{H_R^{-1/2}(\partial\Omega)}.$$

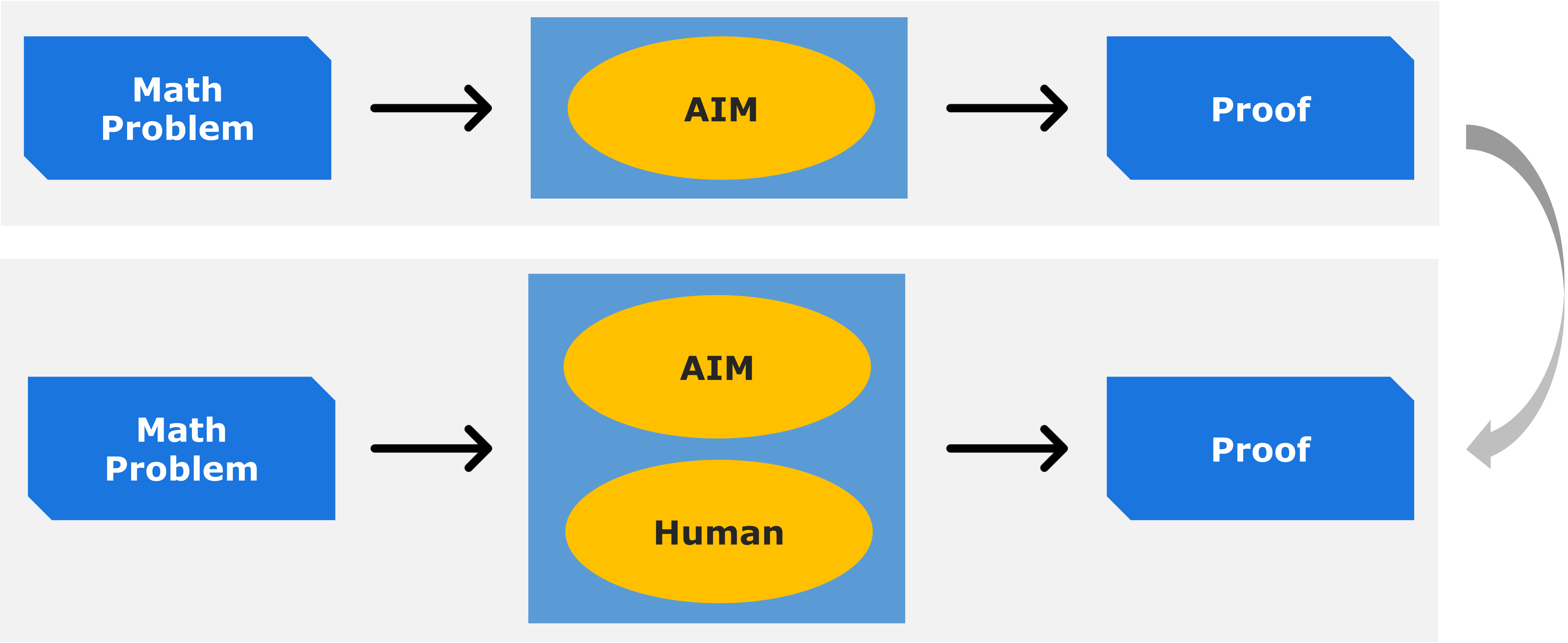
From this one deduces $u_\varepsilon \to u_0$ strongly in $L^2(\Omega)$ (and even in $H^1(\Omega_\varepsilon)$ up to the corrector).

This completes a fully rigorous derivation of the homogenized elastic limit and the explicit formula for C^{hom}.

Correct asymptotic expansion approaches and exploration of conclusions.

Human-AI Collaboration

- **Objective: Solve the Homogenization Problem with minimal human input through human-AI collaboration.**



Human-AI Collaboration

Tsinghua University

- **Subproblem decomposition of the homogenization problem:**

Steps	Hardness	Current Status
Two-Scale Expansion	Easy	Humans handle the task
Cell Problem and Homogenization Equation	Medium	A suitable cell problem is manually constructed, and the homogenized equation is derived by hand
Existence and Uniqueness	Hard	With minimal hints, AIM discovers the correct theorem and proof; humans fill in some details
Ellipticity of Operator	Medium	With minimal hints, AIM provides a largely complete proof; humans refine some details
Error Estimation and Control	Hard	With minimal hints, AIM presents the correct proof approach and some steps, which, after human adjustments, led to a complete proof process
Regularity of Cell Problem	Hard	With minimal hints, AIM provides the complete proof

Background: Cell Problem

Tsinghua University

In the derivation process of the homogenization equation, the construction of the Cell Problem is a necessary technical step. Specifically, we manually constructed such a Cell Problem as described in the following equation.

$$
\begin{aligned}
&\nabla \cdot [\lambda \nabla_y \cdot \chi^{ij} I + 2\mu D_y u] = 0 && in \quad Y \setminus \omega \\
&\nabla \cdot [r^{ij} I + 2\tilde{\mu} D_y \chi^{ij}] = 0 && in \quad \omega \\
&\nabla_y \cdot \chi^{ij} = 0 && in \quad \omega \\
&\chi^{ij}|_+ = \chi^{ij}|_- && in \quad \partial\omega \\
&[r^{ij} I + 2\tilde{\mu} D \chi^{ij}] N|_- - [\lambda \nabla_y \cdot \chi^{ij} I + 2\mu D \chi^{ij}] N|_+ = 0 && in \quad \partial\omega
\end{aligned}
$$

Subproblem: Regularity of Cell Problem

清華大學 Tsinghua University

AIM attempted to derive the error estimation

$$\|u_\epsilon - u_0 - \epsilon\chi(\frac{x}{\epsilon})\nabla u_0\|_{H^1(\Omega)} \le C(\mu, \Omega, \|\chi\|_\infty)\epsilon^{\frac{1}{2}}\|u_0\|_{W^{2,d}(\Omega)}$$

Manual review of the estimation revealed that AIM relied on the following properties without providing proofs

Lemma 4. *Let $\Omega \subset \mathbb{R}^d$ be as above, $\chi(y)$ Y-periodic with $\chi \in L^\infty(Y)$, $\nabla_y \chi \in L^2(Y)$, the cut-off of Lemma 8, and S_ε the mollifier of Lemma 7. Then for every $u_0 \in H^2(\Omega)$*

$$\|\varepsilon\chi(x/\varepsilon)\eta_\varepsilon(\nabla u_0 - S_\varepsilon^2(\nabla u_0))\|_{H^1(\Omega)} \le C\varepsilon\|u_0\|_{H^2(\Omega)},$$

where C depends only on Ω, $\|\chi\|_{L^\infty(Y)}$, $\|\nabla\chi\|_{L^2(Y)}$, and the mollifier.

We analyzed that this property is likely correct and applied AIM to prove the regularity

$$\chi \in W^{1,\infty}(Y \setminus \omega) \cup W^{1,\infty}(\omega)$$

- **Difference Quotient** (thumbs down)
- **Schauder Theory** (thumbs up)
- **Galerkin Method** (thumbs down)

Subproblem: Regularity of Cell Problem

Tsinghua University

Schauder Theory

Lemma 1. *Suppose* $\Omega = R^d$, $S = \{x_d = 0\}$, $B_+ = \{x \in B(1) : x_d > 0\}$ *and* $B_- = \{x \in B(1) : x_d < 0\}$. *Here* $B(1) = \{\|x\| \le 1\}$. *Consider this equation: for* $V \in H_0^1(B(1); R^d)$

$$(\nabla V : A_1 \nabla \tilde{\chi})_{B_+} + (\nabla V : A_2 \nabla \tilde{\chi})_{B_-} + (\tilde{r}, \nabla \cdot (aV))_{B_-} = 0 \quad (1)$$

$$\nabla \cdot (a\tilde{\chi}) = 0 \quad (2)$$

Here $\tilde{\chi} = D^\alpha \chi, \tilde{r} = D^\alpha r, |\alpha| \ge 1$ *and* A_1, A_2 *are constant tensors,* a *is a constant matrix. Then we have for* $\forall k \ge 1$

$$\sum_{\pm} \|\chi\|_{H^k(B(\frac{1}{2},\pm))} \le C\|\chi\|_{L^2(B(1))}$$

$$\|r\|_{H^k(B(\frac{1}{2})_-)} \le C\|r\|_{L^2(B(\frac{1}{2})_-)}$$

Lemma 2. *Suppose that* M *is the constant matrix in* $R^{d\times d}$, *the following are equivalent :*

$$\forall y \in \{y_d = 0\} \quad M_+ x = M_- x \quad (3)$$

$$\exists c \in R^d, s.t. \quad M_+ M_- = C e_d^T \quad (4)$$

$$(I - e_d^T e_d) M_+ = (I - e_d^T e_d) M_- \quad (5)$$

Definition 1. A_1, A_2 *are the tensor constant,* a *is the matrix constant. If* M *satisfies the above Lemma 2 and* $\nabla \cdot (aM_- y) = 0$ *in* $B(t)_-$. *Let* $l(y) = M_+ y 1_{y\ge 0} + M_- y 1_{y \le 0} + C, q(y) = r(0)$.

We call l, q *the piecewise linear solution of the following equation:*

$$\nabla \cdot (A_1 \nabla l) = 0 \qquad in \quad R_+^d \quad (6)$$

$$\nabla \cdot (A_2 \nabla l) + a^T \nabla q = 0, \nabla \cdot (al) = 0, \qquad in \quad R_-^d \quad (7)$$

$$l_+ = l_-; \frac{\partial l}{\partial \nu}|_+ - \frac{\partial l}{\partial \nu}|_- = (A_1 M_+) e_d - (A_2 M_- + r(0)) e_d, \qquad on \quad \{x_d = 0\} \quad (8)$$

Suppose that $\mathcal{L}$ *is the space of all the piecewise-linear solutions of the above equation. And* $\forall (l,q) \in \mathcal{L}$, *we define* $\zeta(l,q) = (\frac{\partial l}{\partial \nu})_+ - (\frac{\partial l}{\partial \nu})_-$

Lemma 3. A_1, A_2 *are the tensor constant,* a *is the matrix constant.*

$$\nabla \cdot (A_1 \nabla \chi) = 0 \qquad in \quad B(1)_+ \quad (9)$$

$$\nabla \cdot (A_2 \nabla \chi) + a^T \nabla r = 0; \nabla \cdot (a\chi) = 0 \qquad in \quad B(1)_- \quad (10)$$

$$\chi_+ = \chi_-; \frac{\partial \chi}{\partial \nu}|_+ - \frac{\partial \chi}{\partial \nu}|_- = g_0, \qquad on \quad B(1) \cap \{x_d = 0\} \quad (11)$$

χ *and* r *are the weak solutions of the above equations. Then for* $\forall k \ge 0, \alpha \in [0,1]$ *, we have* $\sum_{\pm} \|\chi\|_{H^k(B(\frac{1}{2},\pm))} \le C(\|\chi\|_{L^2(B(1))} + |g_0|)$.

Lemma 4. A_1, A_2 *are the tensor constant,* a *is the matrix constant.*

$$\nabla \cdot (A_1 \nabla \chi) = 0 \qquad in \quad B(1)_+ \quad (12)$$

$$\nabla \cdot (A_2 \nabla \chi) + a^T \nabla r = 0, \nabla \cdot (a\chi) = 0, \qquad in \quad B(1)_- \quad (13)$$

$$\chi_+ = \chi_-; \frac{\partial \chi}{\partial \nu}|_+ - \frac{\partial \chi}{\partial \nu}|_- = g_0, \qquad on \quad B(1) \cap \{x_d = 0\} \quad (14)$$

χ *and* r *are the weak solutions of the above equations. Let* $l(y) = (\nabla\chi)_+(0) y 1_{y \ge 0} + (\nabla \chi)_-(0) y 1_{y \le 0} + \chi(0), q(y) = r(0)$.

By Lemma 1 we know that $l_+ = l_-$ *and* $(I - e_d^T e_d)(\nabla l)_+ = (I - e_d^T e_d)(\nabla l)_-$ *on* $B(t) \cap \{y_d = 0\}$. *So by Lemma 2, we know* $(l,q) \in \mathcal{L}$

Moreover, $\forall y \in B(\frac{1}{2})$ *for some* $\beta \in (0,1)$

$$|\chi(y) - l(y)| \le |\chi(y) - \chi(0) - (\nabla\chi)(0)y| \le C|y|^{\beta+1}([\chi]_{C^{1,\beta}(B(\frac{1}{2},))}) \le C|y|^{\beta+1}((\int_{B(1)} |\chi|^2)^{\frac{1}{2}} + |g_0|)$$

and $\forall y \in B(\frac{1}{2})_-$

$$|r - q| \le C|y|^\beta ([r]_{C^{0,\beta}(B(\frac{t}{2})_-)}) \le C|y|^\beta (\int_{B(1)_-} |r|^2)^{\frac{1}{2}}$$

Therefore, $\forall y \in B(\frac{t}{2})$ *for some* $\beta \in (0,1)$

$$|\chi(y) - l(y)| \le |\chi(y) - \chi(0) - (\nabla\chi)(0)y| \le C|\frac{y}{t}|^{\beta+1}([\chi]_{C^{1,\beta}(B(\frac{t}{2},))}) \le C|\frac{y}{t}|^{\beta+1}((\fint_{B(t)} |\chi|^2)^{\frac{1}{2}} + t|g_0|)$$

and $\forall y \in B(\frac{t}{2})_-$

$$|r - q| \le C|\frac{y}{t}|^\beta ([r]_{C^{0,\beta}(B(\frac{t}{2})_-)}) \le C|\frac{y}{t}|^\beta (\fint_{B(t)_-} |r|^2)^{\frac{1}{2}}$$

Lemma 5. A_1, A_2 *are the tensor constant,* a *is the matrix constant.*

$$\nabla \cdot (A_1 \nabla \chi) = 0 \qquad in \quad B(1)_+ \quad (15)$$

$$\nabla \cdot (A_2 \nabla \chi) + a^T \nabla r = 0; \nabla \cdot (a\chi) = 0, \qquad in \quad B(1)_- \quad (16)$$

$$\chi_+ = \chi_-; \frac{\partial \chi}{\partial \nu}|_+ - \frac{\partial \chi}{\partial \nu}|_- = g_0, \qquad on \quad B(1) \cap \{x_d = 0\} \quad (17)$$

χ *and* r *are the weak solutions of the above equations. Moreover,* $\forall \rho \in (0,t)$ *integrate the above inequalities to get*

$$(\fint_{B(\rho)} |\chi - l|^2)^{\frac{1}{2}} + \rho|g_0 - \zeta(l,q)| \le |\frac{\rho}{t}|^{\beta+1}((\fint_{B(t)} |\chi|^2)^{\frac{1}{2}} + t|g_0|)$$

So $\forall (l', q') \in \mathcal{L}$, *by the inequality above , we have*

$$inf_{l,q\in\mathcal{L}}\{(\fint_{B(\rho)} |\chi - l|^2)^{\frac{1}{2}} + \rho|g_0 - \zeta(l,q)|\} \le C|\frac{\rho}{t}|^{\beta+1} inf_{l,q\in\mathcal{L}}((\fint_{B(t)} |\chi - l|^2)^{\frac{1}{2}} + t|g_0|)$$

Lemma 6. *Suppose* $\phi : R_+ \to R_+$ *is a non-decreasing non-negative function satisfying* $\phi(\rho) \le C(\frac{\rho}{t})^\beta \phi(r) + Br^\alpha$, *where* $\beta > \alpha > 0, C > 0$.

Then $\forall 0 < \rho < r < R, \exists C_1, s.t. \phi(\rho) \le C_1(\frac{\rho}{t})^\alpha \phi(r) + B\rho^\alpha$

We adjusted the lemmas from the schauder theory method into a form suitable for the equations of the Cell Problem.

AIM was utilized to complete the schauder theory related lemmas. These contents are input into the model as the "context" section served as a methodological guide for subsequent regularity proofs.

Subproblem: Regularity of Cell Problem

清華大學 Tsinghua University

- **We transform the problem into the proof of the following theorem.**

Theorem 1. *Suppose A_1, A_2, a are C^α-Hölder continuous, we set $S_t = B(t) \cap \{x_d = 0\}$. χ, r are the weak solutions of the following equations: for $V \in H_0^1(B(1); R^d)$*

$$(\nabla V : A_1 \nabla \chi)_{B_+} + (\nabla V : A_2 \nabla \chi)_{B_-} + (r, \nabla \cdot (a\chi))_{B_-} = 0 \qquad (18)$$

$$\nabla \cdot (a\chi) = 0 \qquad (19)$$

Please prove $\sum_{\pm} \|\chi\|_{C^{1,\alpha}(B(\frac{1}{2},))} \leq C\|\chi\|_{L^2(B(1))}$.
This is equivalent to prove $\forall \rho \in (0, \frac{1}{4})$, *we have*

$$inf_{l,q \in \mathcal{L}} (\fint_{B(\rho)} |u - l|^2)^{\frac{1}{2}} \leq C\rho^{1+\alpha} \sum \|u\|_{L^2(B(\frac{3}{4}))}$$

Here, (l, q) are the piecewise linear function and $B(\rho)$ is a small ball with any given center of the ball on $S_{\frac{3}{4}}$.

We adjusted the lemmas from the schauder theory method into a form suitable for the equations of the Cell Problem.

We instruct AIM to use schauder theory methods to prove this theorem. Based on feedback from experimental results, we iteratively split the problem and ultimately complete the proof.

Subproblem: Regularity of Cell Problem

Tsinghua University

- **Specifically, the entire problem is divided into the following three parts. AIM progressively completes the proof details for each part. After iterations, AIM provided a process with a high level of completeness and ultimately completed the proof.**

1 Perturbation of the Equation

$$(\nabla V : A_1^0 \nabla w_t)_{B_+} + (\nabla V : A_2^0 \nabla w_t)_{B_-} + (s_t, \nabla \cdot (a^0 V))_{B_-} = 0$$

$$\nabla \cdot (a^0 w_t) = 0 \qquad in \quad B(t)_-$$

$$w_t = \chi \quad on \quad \partial B(t) \quad and \quad s_t = r \qquad on \quad \partial B(t)_-$$

2 Morrey's Esitmate Bootstrap Ananlysis

$$\Psi(r) = \int_{B(r)} |\nabla \chi|^2 + \int_{B(r)_-} |r|^2$$

$$\Psi(\rho) \leq C(\frac{\rho}{t})^d \Psi(t) + Ct^{2\alpha}\Psi(t), \forall 0 < \rho < t < \frac{1}{2}$$

→ χ is C^β Hölder continuous

3 Hölder Regularity

$$\Phi(r) = \inf_{l,q \in \mathcal{L}} \{ \int_{B(r)} |\chi - l|^2 + r^{d+2} |\zeta(l,q)|^2 \}$$

$$\Phi(\rho) \leq C(\frac{\rho}{t})^{d+2\beta+2}\Phi(t) + Ct^{d+2+\alpha}\Psi(\frac{1}{2})$$

→ $\chi \in C^{1,\alpha}(\overline{B_{\frac{1}{2}}}; R^d), \forall \rho \in (0, \frac{1}{2})$

↓ **Beyond expectations**

Final Conclusion

Tsinghua University

Finally, based on the regularity obtained from the proof, we completed the error control of the homogenization limit.

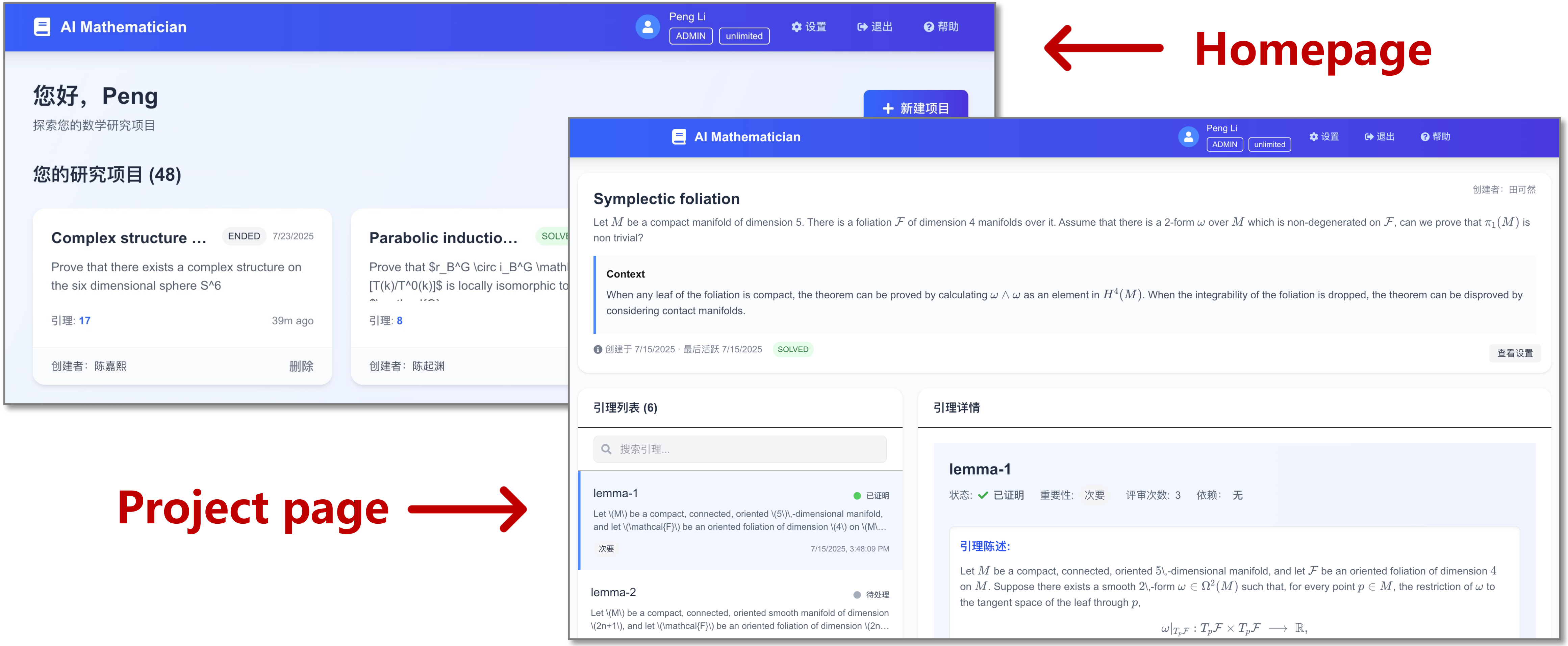
$$\|u_\epsilon - u_0 - \epsilon\chi(\frac{x}{\epsilon})\nabla u_0\|_{H^1(\Omega)} \leq C(\mu, \Omega, \|\chi\|_\infty)\epsilon^{\frac{1}{2}}\|u_0\|_{W^{2,d}(\Omega)}$$

Here is the error control conclusion between the solution of the original equation and the solution of the homogenized equation.

- **The proofs of the subproblems have been preliminarily verified as correct by a mathematics professor.**
- **We are currently reorganizing them into a format more suitable for submission to mathematical journals, followed by a more thorough internal review before final submission.**

AIM Online Demo

- **We have deployed a web frontend of AIM systems and invited some domain experts to use it and evaluate its performance.**



← **Homepage**

Project page →

Some User Feedbacks of AIM

Tsinghua University

AIM successfully addressed some group classification, and provides a detailed proof to it. 😄

Classify certain finite subgroup of SO(4)

Suppose A is a finite subg… (4) with
line bundle l on $S^3 \setminus Z$ whi… Z. And f…
subgroup A isomorphic wit… …ite subg…

创建于 7/17/2025 · 最后活跃…

证明:

We sketch one explicit counter–example, coming from the binary tetrahedral group.

1. Let $\widetilde{T} \subset S^3$ be the binary tetrahedral subgroup of order 24, and let

$$A = \widetilde{T}/\{\pm 1\} \subset SO(4)$$

be its image under the double covering $S^3 \to SO(4)$. Inside $\widetilde{T}$ there are exactly four cyclic

Symplectic foliation

Let M be …
on $\mathcal{F}$, can …

Contex…

When a …
droppe…

theorem-6

状态: ✔ 已证明 重要性: 关键 评审次数: 9 依赖: 1, 3

引理陈述:

Let M be a compact manifold of dimension 5. There is a foliation $\mathcal{F}$ of dimension 4 manifolds over it. Assume that there is a 2-form ω over M which is non-degenerated on $\mathcal{F}$, can we prove

A hard problem on symplectic foliation. The user omitted a condition in the query, and AIM provides a valid counter example for this problem. 😄

AIM failed to directly adress a research level problem. It exceeded the maximum complexity that can be handled by AIM. ☹

三维空间中管状邻域Z2调和函数的存在性问题

Given any positive integer n, is there a Z2 harmonic fu…

This is known to be true if the metric is given by the tri…
functions. However, for a more general smooth metric…

Context

Let (M, g) be a 3-dimensional smooth oriented Ri…
$\rho(S^1)$. Suppose there is a real line bundle l over M…

lemma-12

状态: ⌛ 待处理 重要性: 次要 评审次数: 6 依赖: 2, 3, 10

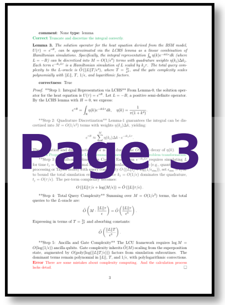
引理陈述:

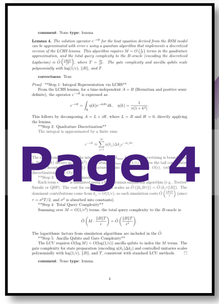
Let (M, g) be as in Memory ID 0, with tubular neighbourhood $N \cong S^1_s \times D_{r,\theta}$ and real line bundle $l \to N \setminus K$ of monodromy -1. Suppose moreover that in local coordinates (s, r, θ)

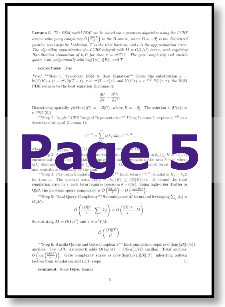
Limitation 1: Redundant Exploration

Tsinghua University

- **Currently, AIM often explores in the same direction, presenting a series of similar conjectures and identical lemmas, which increases costs, reduces efficiency, and limits the performance ceiling.**





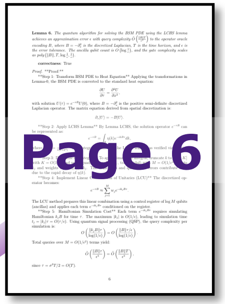


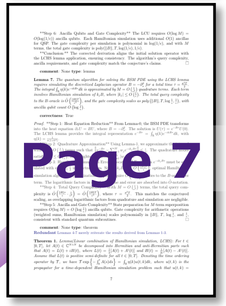
Lemma 4. *The solution operator $e^{-\tau B}$ for the heat equation derived from the BSM model can be approximated with error ϵ using a quantum algorithm that implements a discretized version of the LCHS lemma. This algorithm requires $M = O\left(\frac{1}{\epsilon^2}\right)$ terms in the quadrature approximation, and the total query complexity to the B-oracle (encoding the discretized Laplacian) is $\tilde{O}\left(\frac{\|B\|T}{\epsilon^3}\right)$, where $T = \frac{2\tau}{\sigma^2}$. The gate complexity and ancilla qubits scale polynomially with $\log(1/\epsilon)$, $\|B\|$, and T.*

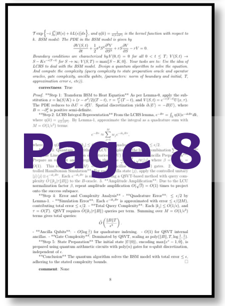
Lemma 6. *The quantum algorithm for solving the BSM PDE using the LCHS lemma achieves an approximation error ϵ with query complexity $\tilde{O}\left(\frac{\|B\|T}{\epsilon^3}\right)$ to the operator oracle encoding B, where $B = -\partial_x^2$ is the discretized Laplacian, T is the time horizon, and ϵ is the error tolerance. The ancilla qubit count is $O\left(\log\frac{1}{\epsilon}\right)$, and the gate complexity scales as poly$\left(\|B\|, T, \log\frac{1}{\epsilon}, \frac{1}{\epsilon}\right)$.*

Three Similar Lemmas

Lemma 7. *The quantum algorithm for solving the BSM PDE using the LCHS lemma requires simulating the discretized Laplacian operator $B = -\partial_x^2$ for a total time $\tau = \frac{\sigma^2 T}{2}$. The integral $\int_{\mathbb{R}} \eta(k)e^{-ikB\tau}dk$ is approximated by $M = O\left(\frac{1}{\epsilon^2}\right)$ quadrature terms. Each term involves Hamiltonian simulation of k_jB, where $|k_j| \leq O\left(\frac{1}{\epsilon}\right)$. The total query complexity to the B-oracle is $\tilde{O}\left(\frac{\|B\|T}{\epsilon^3}\right)$, and the gate complexity scales as poly$\left(\|B\|, T, \log\frac{1}{\epsilon}, \frac{1}{\epsilon}\right)$, with ancilla qubit count $O\left(\log\frac{1}{\epsilon}\right)$.*







Limitation 2: Insufficient Understanding of Mathematical Settings

- **The current AIM has limited ability to understand lengthy mathematical setups and background conditions, which can lead to errors in the agent's analysis process.**

Then as $\varepsilon \to 0$ one has, up to a subsequence, $u_\varepsilon \rightharpoonup u_0$ in $H^1(\Omega)$, $u_\varepsilon \to u_0$ in $L^2(\Omega)$, $p_\varepsilon \overset{\text{two-scale}}{\longrightarrow} p_1(x,y)$ in $\Omega \times Y_i$, where $u_0 \in H^1_R(\Omega;\mathbb{R}^d)$ is the unique solution of the homogenized Lamé system $-\mathrm{div}_x[C^{\mathrm{hom}} D_x(u_0)] = 0$ in Ω, $C^{\mathrm{hom}} D_x(u_0)\cdot n = g$ on $\partial\Omega$, and the effective fourth-order tensor C^{hom} is given by the following periodic cell transmission problem: for each fixed symmetric $E \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ find $(\chi_e^E, \chi_i^E, \pi^E) \in H^1_{\mathrm{per}}(Y_e;\mathbb{R}^d) \times H^1_{\mathrm{per}}(Y_i;\mathbb{R}^d) \times L^2_{\mathrm{per}}(Y_i)/\mathbb{R}$ solving

(1) In the elastic cell Y_e: $-\mathrm{div}_y[C(E + D_y\chi_e^E)] = 0$, $\mathrm{div}_y(Ey + \chi_e^E) = 0$,

(2) In the fluid cell Y_i: $-\mu\Delta_y\chi_i^E + \nabla_y\pi^E = 0$, $\mathrm{div}_y(Ey + \chi_i^E) = 0$,

(3) On the material interface $\Gamma = \partial Y_i$: $\chi_e^E = \chi_i^E$, $[C(E + D_y\chi_e^E)\cdot n] = [2\mu D_y\chi_i^E - \pi^E I]\cdot n$,

The lack of understanding of the boundaries of the unit cell here has led to conclusions that do not align with the actual situation

Limitation 3: Lack of Derivation Details

- AIM tends to "skip steps" , which is detrimental to both manual and automated verification of the proof and may lead to incorrect proofs.

Step 2: Formulate the Projected Equations For test functions $v \in V_N$ and $\psi \in W_{k,N}$, the Galerkin approximations satisfy the projected weak form:

$$\int_\Omega \partial_t u^N v\, dx + \int_\Omega \nabla u^N \cdot \nabla v\, dx + \int_\Omega \rho u^N v\, dx$$
$$+ \beta \int_S u^N v\, ds + \sum_{k=1}^{m} \alpha_k \int_S (\partial_t \varphi_k^N - \Delta_S \varphi_k^N) v\, ds = 0,$$

$$\int_S \partial_t \varphi_k^N \psi\, ds + \int_S \nabla_S \varphi_k^N \cdot \nabla_S \psi\, ds + d_k \int_S \varphi_k^N \psi\, ds = \int_S u^N \psi\, ds.$$

Step 3: Solve the ODE System Substituting $u^N = \sum_{j=1}^N a_j(t) w_j$ and $\varphi_k^N = \sum_{l=1}^N b_{k,l}(t)\theta_{k,l}$ into the projected equations yields a linear ODE system for $a_j(t)$ and $b_{k,l}(t)$. The coefficients are smooth (as eigenfunctions are C^∞), ensuring continuity. By the Cauchy-Lipschitz theorem, unique solutions exist on $[0, T]$.

[Correct] The agent gets the ODE satisfied with the approximated solution.

[Vague] It is better to have more detailed process.

The specific derivation and analysis process of the ODE system is not provided here, but the conclusion is correct.

Limitation 4: Rising Demand, Limited Supply

清華大學 Tsinghua University

- **Actual user demands exhibit considerable diversity and extend substantially beyond the current operational scope of AIM.**

User Expectations	AIM Capabilities
Interact with minimal input	Provide a relatively complete problem description
Chinese-language input	English-language input is more effective
Zero tolerance for errors	Is not yet error-free
Capable of solving any problem	Has its own limitations
……	……

Outlook: A Bold Bet on Natural Language

清華大學 Tsinghua University

- **For the first time, large language models performed on a par with gold medallists in the International Mathematical Olympiad (IMO).**

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature > news > article

NEWS | 24 July 2025

DeepMind and OpenAI models solve maths problems at level of top students

big paradigm shift

RELATED

DeepMind AI crushes tough maths problems on par with top human solvers

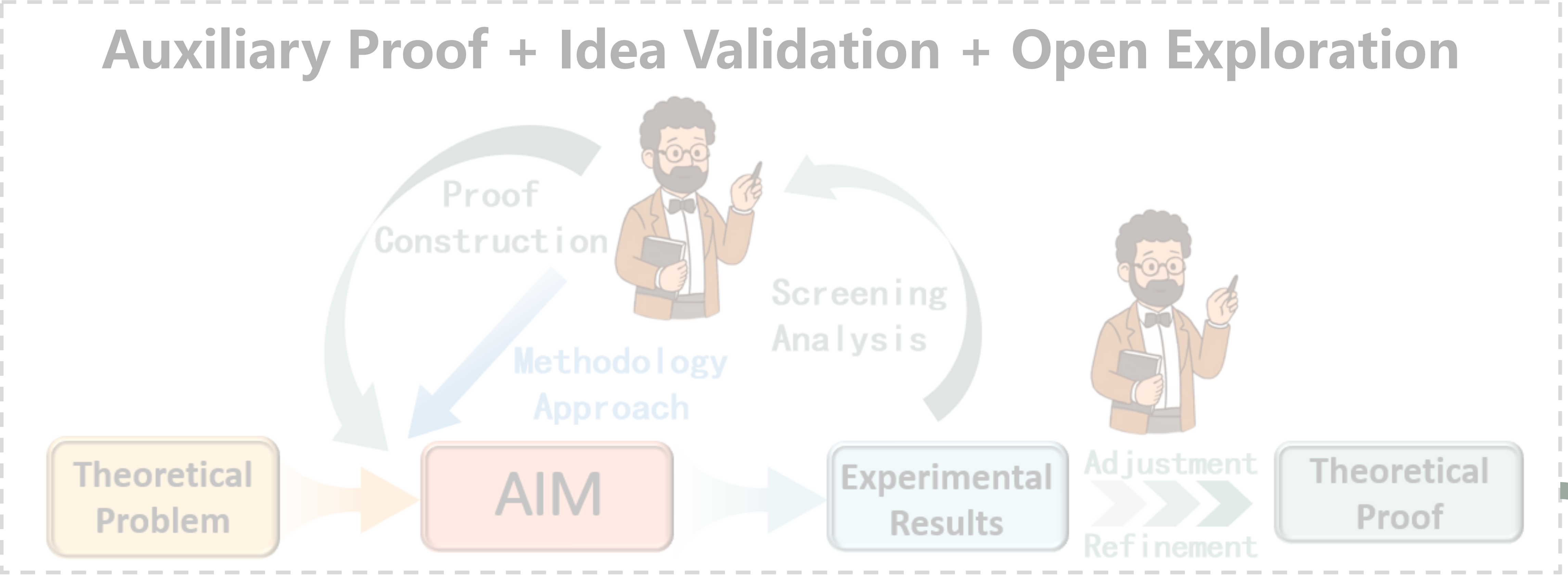
But the grades this year hide a "big paradigm shift," says Thang Luong, a computer scientist at DeepMind in Mountain View, California. The company achieved its previous feats using two artificial intelligence (AI) tools specifically designed to carry out rigorous logical steps in mathematical proofscalculations, called AlphaGeometry and AlphaProof. The process required human experts to first translate the problems' statements into something similar to a programming language, and then to translate the AI's solutions back into English.

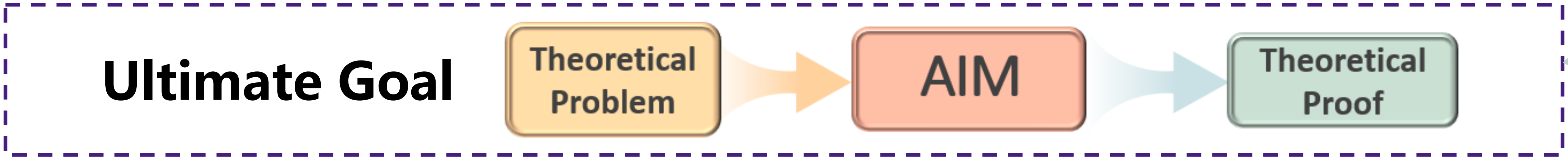
This year, everything is natural language, end to end

"This year, everything is natural language, end to end," says Luong. The team employed a large language model (LLM) called DeepThink, which is based on its Gemini system but with some additional developments that made it better and faster at producing mathematical arguments, such as handling multiple chains of thought in parallel. "For a long time, I didn't think we could go that far with LLMs," Luong adds.

https://www.nature.com/articles/d41586-025-02343-x

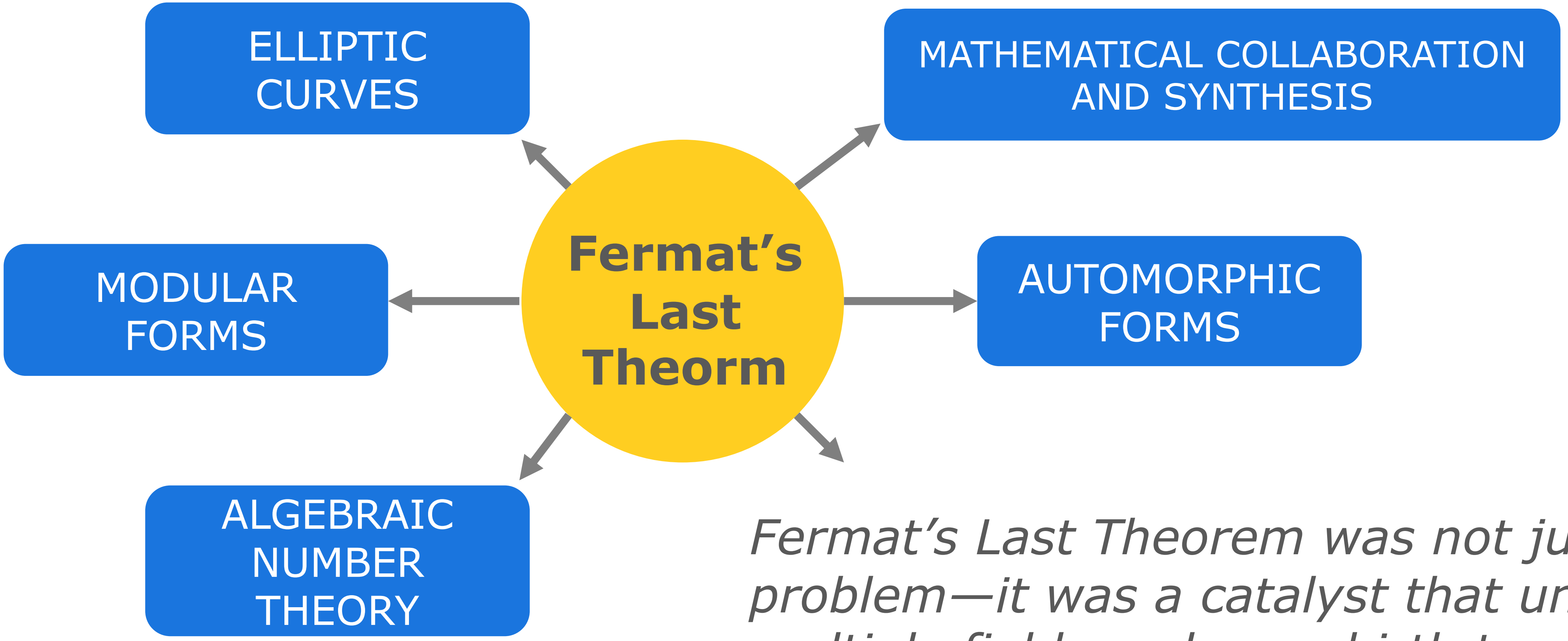
Outlook: Assistive to Proactive





Outlook: Pose New Mathematical Problems

- Good problems drive mathematics; we expect AIM to help pose them.



Fermat's Last Theorem was not just a problem—it was a catalyst that united multiple fields and gave birth to modern number theory.

Summary

- **Large language models enable greater AI impact in mathematical research.**
- **Three key characteristics distinguish mathematical research from problem-solving, including long reasoning, procedural rigor, and answer existence.**
- **We propose the AI mathematician system AIM, which has achieved preliminary success on four research-level mathematical problems, showing promising potential.**
- **In the future, AI will play a more proactive and important role in mathematical research.**

Tech. Report

Code

Our <u>AIM</u> is <u>AI</u> <u>M</u>athematician!