AIM：AI数学家系统

李鹏

清华大学

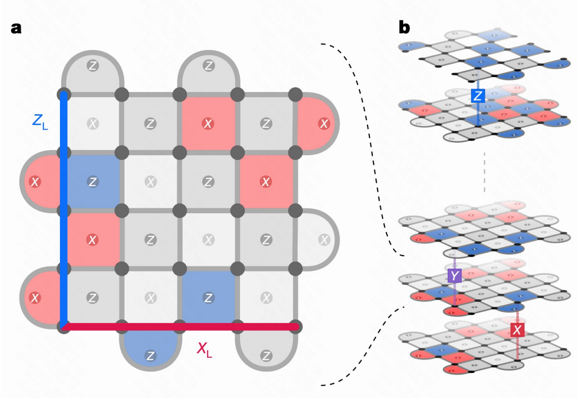
AI深刻改变科学研究范式

Tsinghua University

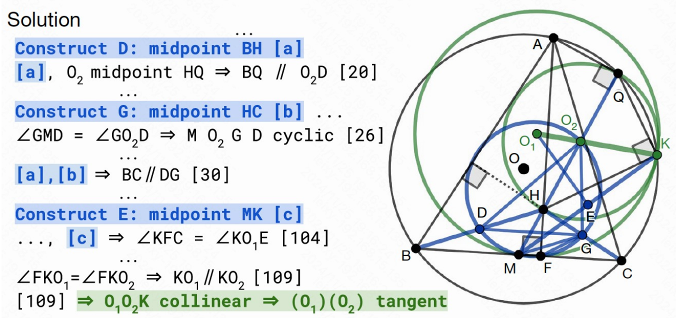
- **深度学习的发展赋能AI4Science，极大地推动了科学的发展。**

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

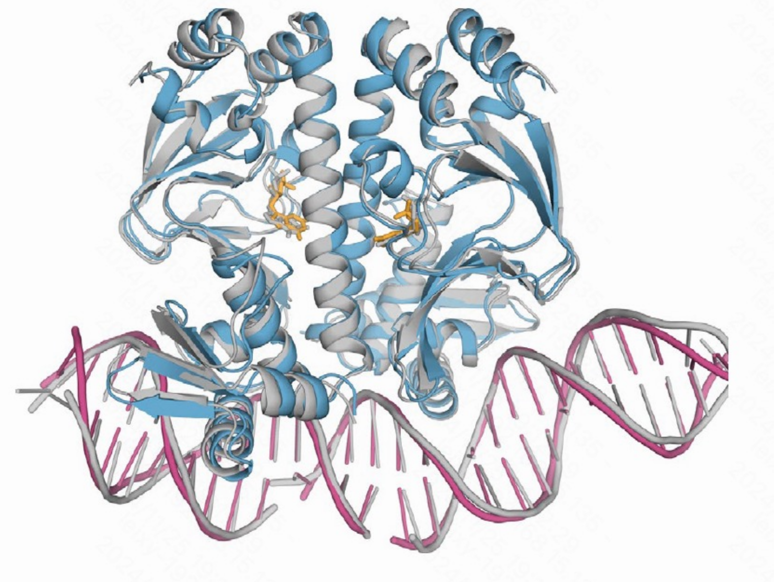
David Baker, Demis Hassabis and John Jumper. Ill. Niklas Elmehed © Nobel Prize Outreach



AlphaQuant



AlphaGeometry

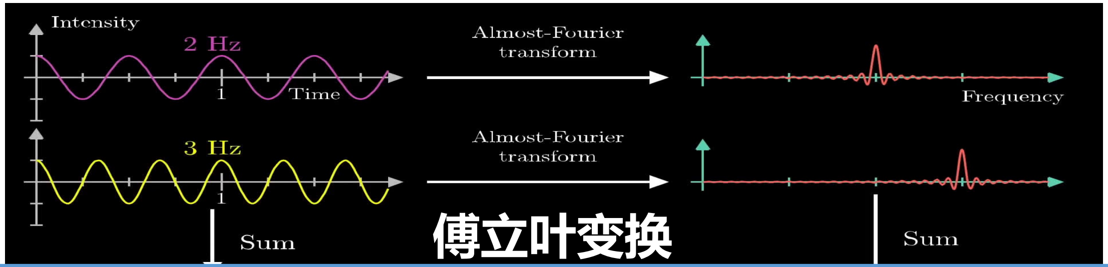


AlphaFold

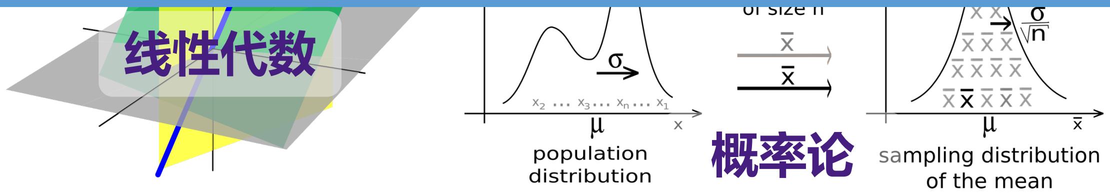
https://www.nobelprize.org/all-nobel-prizes-2024/
https://www.nature.com/articles/s41586-024-07487-w
https://www.nature.com/articles/s41586-024-08148-8
https://www.nature.com/articles/s41586-023-06747-5

数学研究意义重大

- 数学研究具有重要的理论与应用价值，是人类智慧的集中体现。



AI与数学的结合可以产生什么样的化学反应?



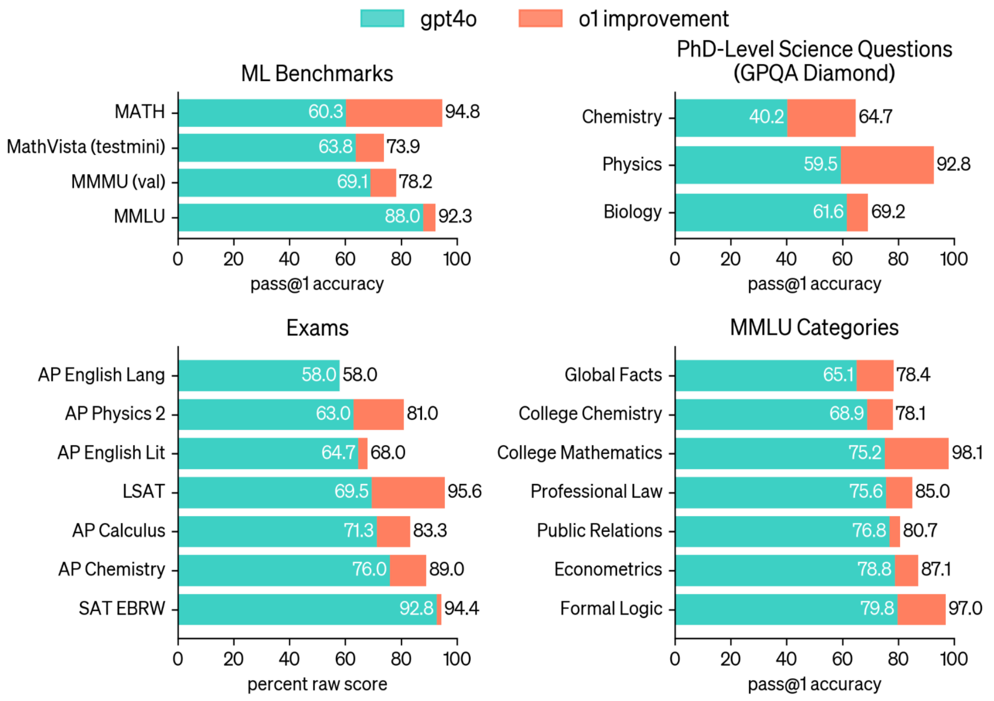
探索至真至美的自然真理

科学研究与工程应用的理论工具

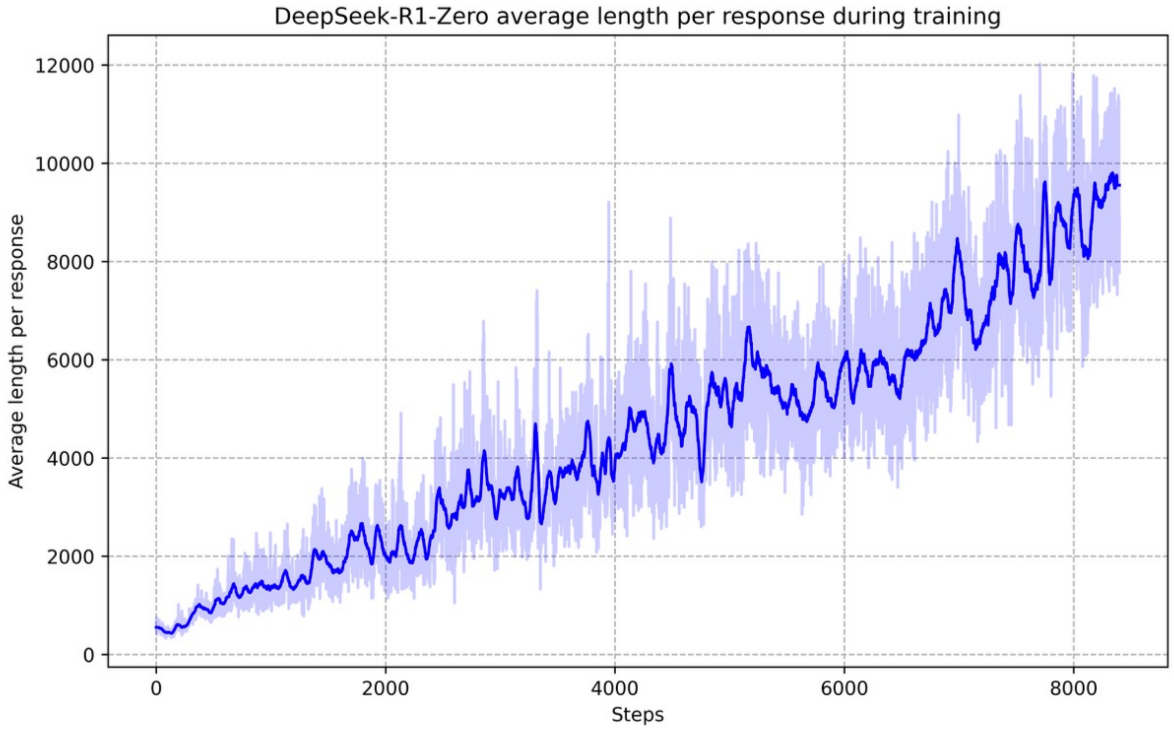
数学之于AI：大模型性能进步关键要素

Tsinghua University

● **数学和代码是当前提升大模型推理能力的关键要素。**



针对数学和代码任务训练的第一个推理模型 o1能够在各方面取得明显进步



DeepSeek-R1通过针对数学问题的可验证反馈的强化学习训练获取了深度思考的能力

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

AI之于数学：数学能力快速接近人类

- **2025年高考数学新课标 I 卷大模型最高水平达到145分（满分150分）。**

7家大模型挑战高考数学新一卷				
模型	客观题（文本）得分	图像题（单选）得分	解答题得分	总分
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
文心X1 Turbo	68	/	66	134
注：本次评测分为客观题（文本）、第6题图像题（单选）和解答题，客观题（文本）总计68分，图像题总共（单选）5分，解答题77分，总分150分				

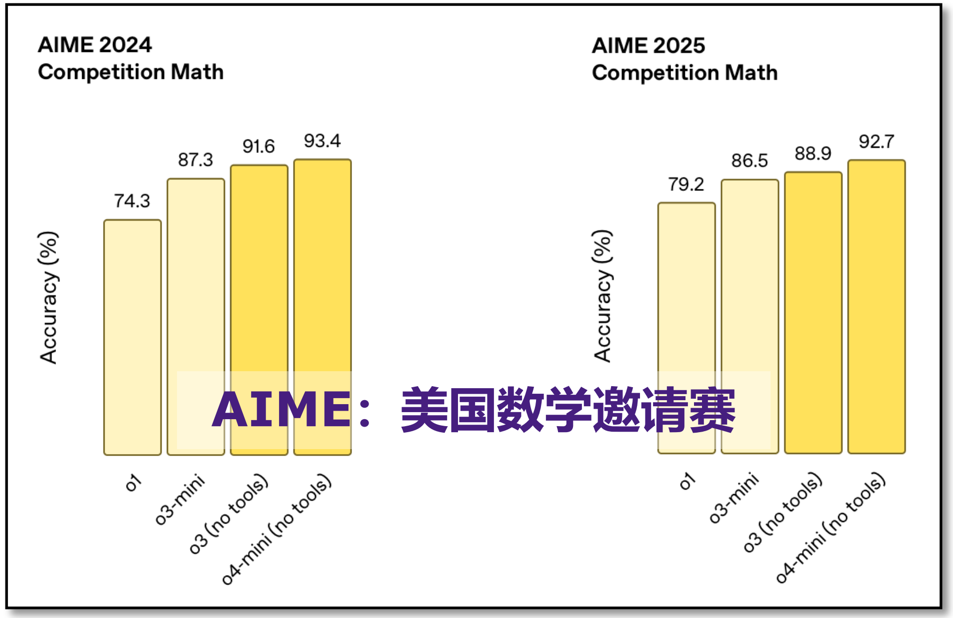
图像理解能力不足导致失分

https://mp.weixin.qq.com/s/E1D_QA5IXwXRc6aSaU_zog?scene=1

AI之于数学：数学能力快速接近人类

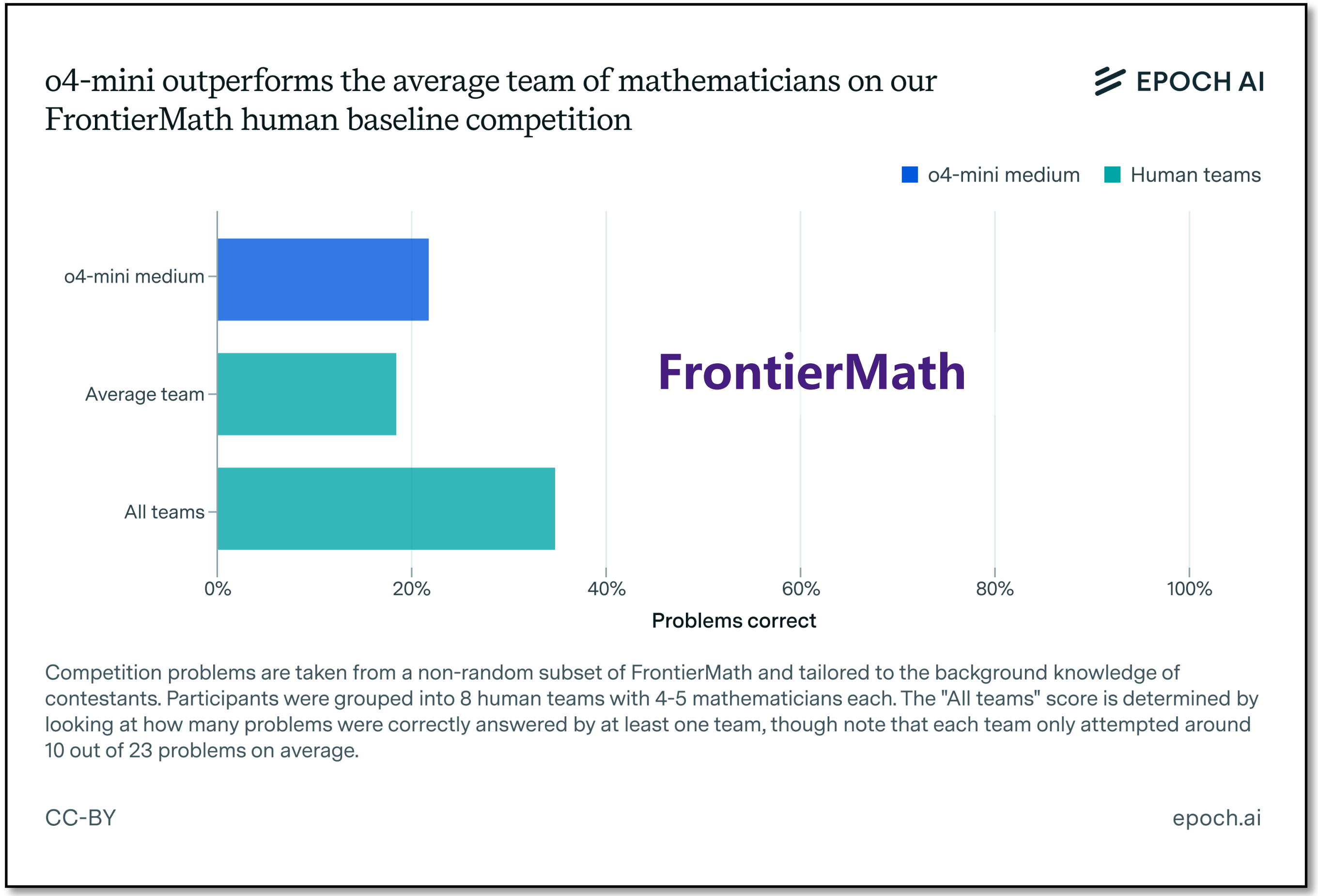
Tsinghua University

- 2025年5月o4-mini在部分前沿数学问题上的表现已经达到专家水平。



"我不想加剧恐慌。但在某些方面，大语言模型的表现已经超越了世界上大多数最优秀的研究生。"

——小野健



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：高中竞赛水平测试集

清華大學 Tsinghua University

- **AIME数据集来自于美国数学邀请赛，赛事面向高中生，是选拔国际数学奥林匹克美国队的赛事之一。**

赛题设置

- 共15道题，每一题的难度不断增大
- 答案均为一个介于0-999之间的整数

局限性

- 答案存在且已知
- 主要考查解题技巧

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：高中竞赛水平测试集

- 当前大模型主要使用可验证的练习题或者数学竞赛题用于训练和测试，与AIME题目类型分布一致性强。

DeepSeek-R1针对AIME2025竞赛的第一题给出的完整正确解答

- **原始题目**：Find the sum of all integer bases b > 9 for which 17_b is a divisor of 97_b
- **题目难度**：结构简单且不涉及高深概念
- **思考长度**：可在简短步骤当中快速完成
- **考查方面**：主要考查解题技解的应用

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}

FrontierMath：挑战性陡升

- **由专业数学家撰写，陶哲轩等均认可FrontierMath具有高挑战性，但仍与研究级数学问题存在显著差距。**

挑战度高

- 题目由专业数学家撰写
- 涵盖现代数学主要分支
- 人类专家一般需要数小时解决一个题目

局限性

- 答案存在且已知
- 难度低于数学研究问题

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.

数学解题 v.s. 数学研究

清華大學 Tsinghua University

差异1：长线思考

数学研究需要比做题显著更多、更加长期且更加深入的有效思考

三维挂谷猜想

- 完整证明总计12个章节、125页正文
- 论文仅呈现了最终证明过程，真实思考和推理量远大于论文呈现结果

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.

数学解题 v.s. 数学研究

清華大學 Tsinghua University

差异2：过程严谨

错误零容忍，正确性需要可验证，而验证代价异常高昂

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

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

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.

三维挂谷猜想证明片段

案例	验证过程	验证耗时
费马大定理	1993年给出初始证明，由一个顶级小组反复审读并指出问题，随后怀尔斯与他的学生泰勒合作修改，1995年才正式发表两篇论文	约2年
庞加莱猜想	2002-2003年分三篇预印本形式在线发布了证明，数学界以非正式审稿方式组织了由多位几何拓扑学家验证、重写和解释其证明的过程	约4-5年

数学解题 v.s. 数学研究

差异3：答案存在性

数学解题答案存在性**已知**且**确定**
数学研究答案存在性**未知**且**不确定**

$P(Q)$ → YES / NO

停机问题

哥德尔

- **第一不完备定理**：任何自洽的形式系统，只要蕴涵皮亚诺算术公理，就可以在其中构造在体系中不能被证明的真命题，因此通过推理演绎不能得到所有真命题（即体系是不完备的）
- **第二不完备定理**：任何逻辑自洽的形式系统，只要蕴涵皮亚诺算术公理，它就不能用于证明其本身的自洽性

AI+数学研究：形式化证明

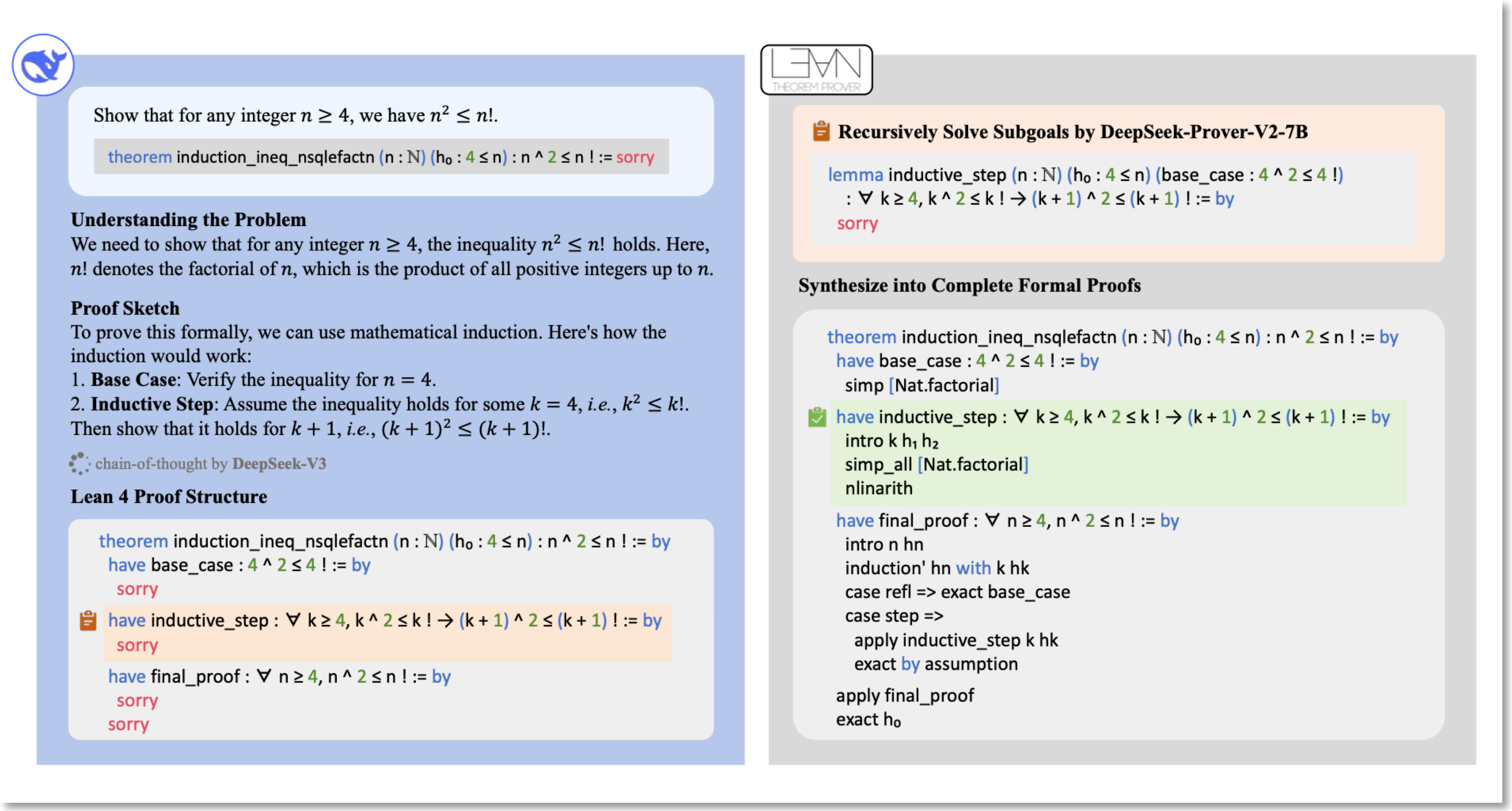
- 形式化证明历史悠久，国内外研究机构以及DeepSeek、字节跳动等积极探索该研究方向。

优势

- 将各种数学证明转换为符号计算，可利用计算机机械化、批量执行
- 保证证明过程严格正确性

缺点

- 当前形式化证明工具存在明显的性能和效率问题
- 进而难以应用到前沿研究工作当中



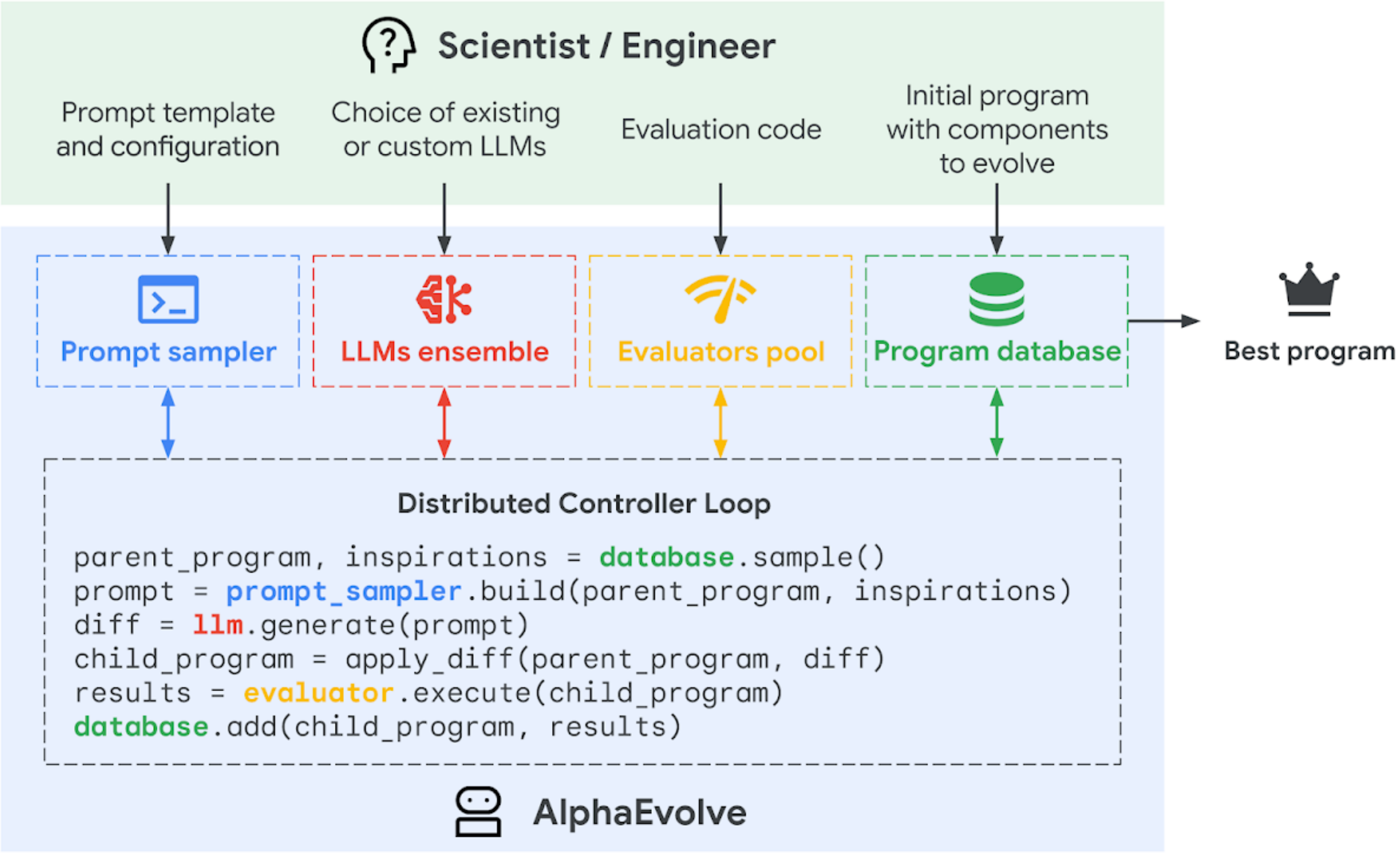
https://deepmind.google/discover/blog/alphaevolve-a-gemini-powered-coding-agent-for-designing-advanced-algorithms/

AI+数学研究：大模型

- 近期人们也开始探索利用大模型的能力执行数学科研任务的可能。

AlphaEvolve

- **起源**：2025年5月14日由谷歌DeepMind推出
- **本质**：专注于算法优化的大模型智能体系统
- **成效**：已经能够独立取得一些新的研究成果
- **局限**：只适用于能够被转换为代码验证的问题



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

AIM：面向数学研究的AI数学家智能体

清華大學
Tsinghua University

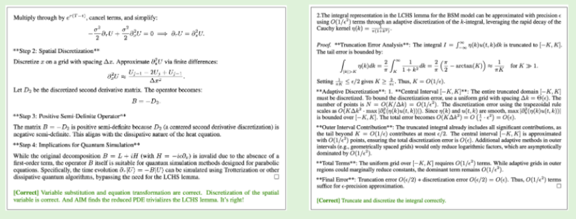
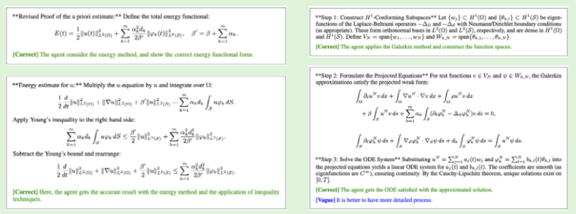
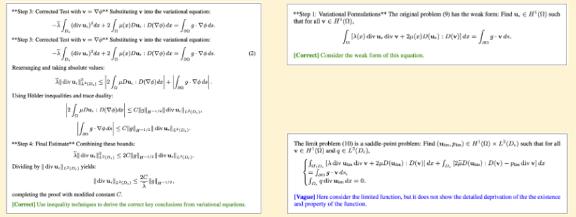
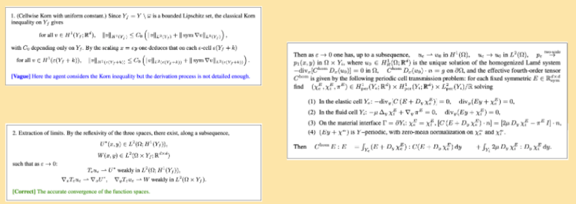
- **AIM在利用人工智能进行复杂数学理论研究中迈出了重要的一步。**

数学解题

- 推理链条短
- 严谨性易满足
- 答案存在性已知
- 答案确定

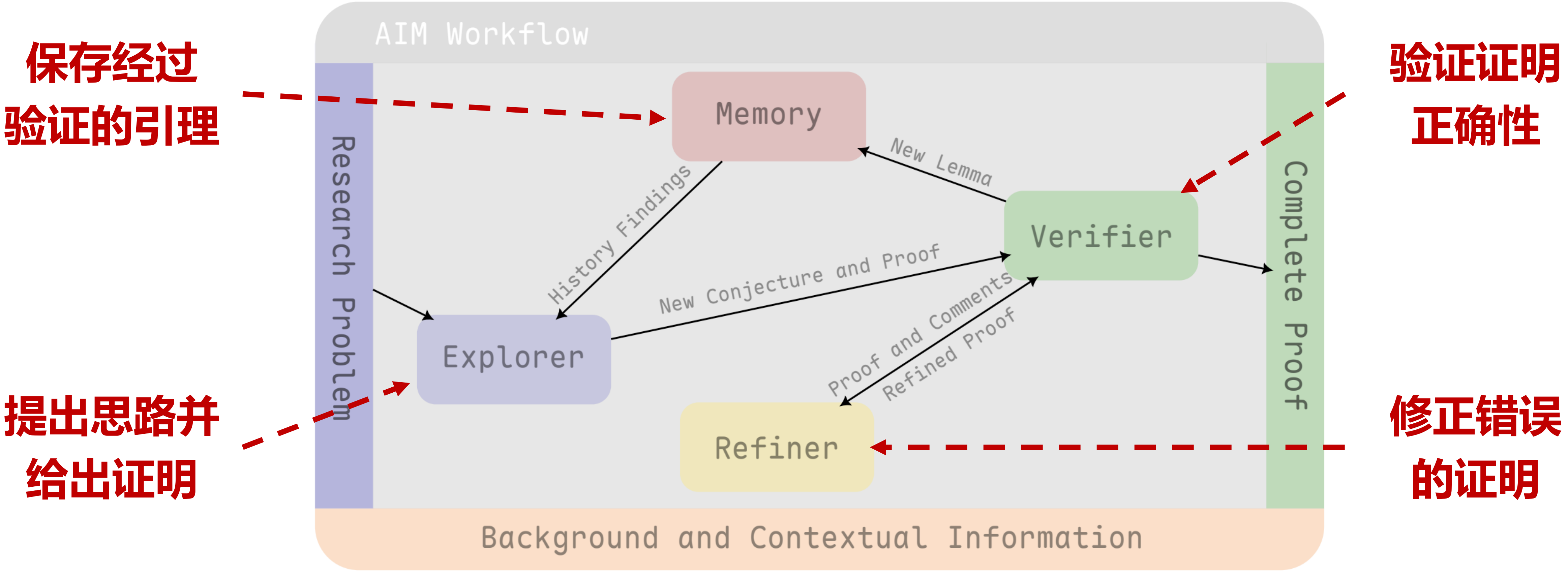
数学研究

- 推理链条长
- 严谨性难满足
- 答案存在性未知
- 答案不确定

数学问题	AIM效果	部分证明步骤
量子算法问题 (结论已知) Quantum Algorithm Problem	AIM完成了带有详细分析推导过程的证明	
吸收边界问题 (结论已知) Absorbing Boundary Condition	AIM基本完成此问题的证明过程和细节	
高对比度极限 (结论已知) High Contrast Limit	AIM完成了主要结论的证明过程，并且证明了新的结论	
均匀化问题 (结论未知) Homogenization Problem	AIM给出了合理的证明思路和部分结论的证明，对数学工作者很有启发意义	

AIM框架概览

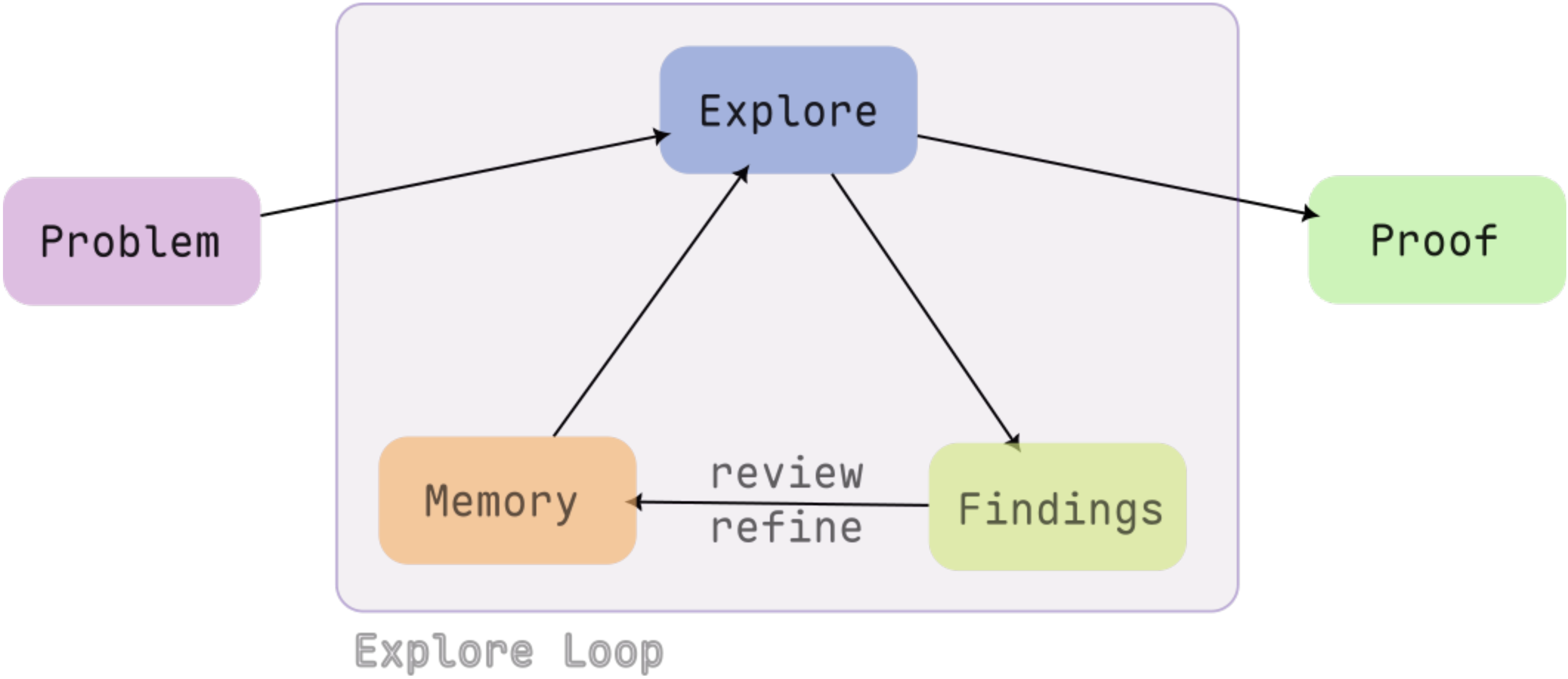
- 由大模型驱动的多智能体系统，针对数学证明特点进行特别设计，包含Explorer、Verifier、Refiner三个智能体以及一个Memory模块。



长线思考：探索与记忆机制

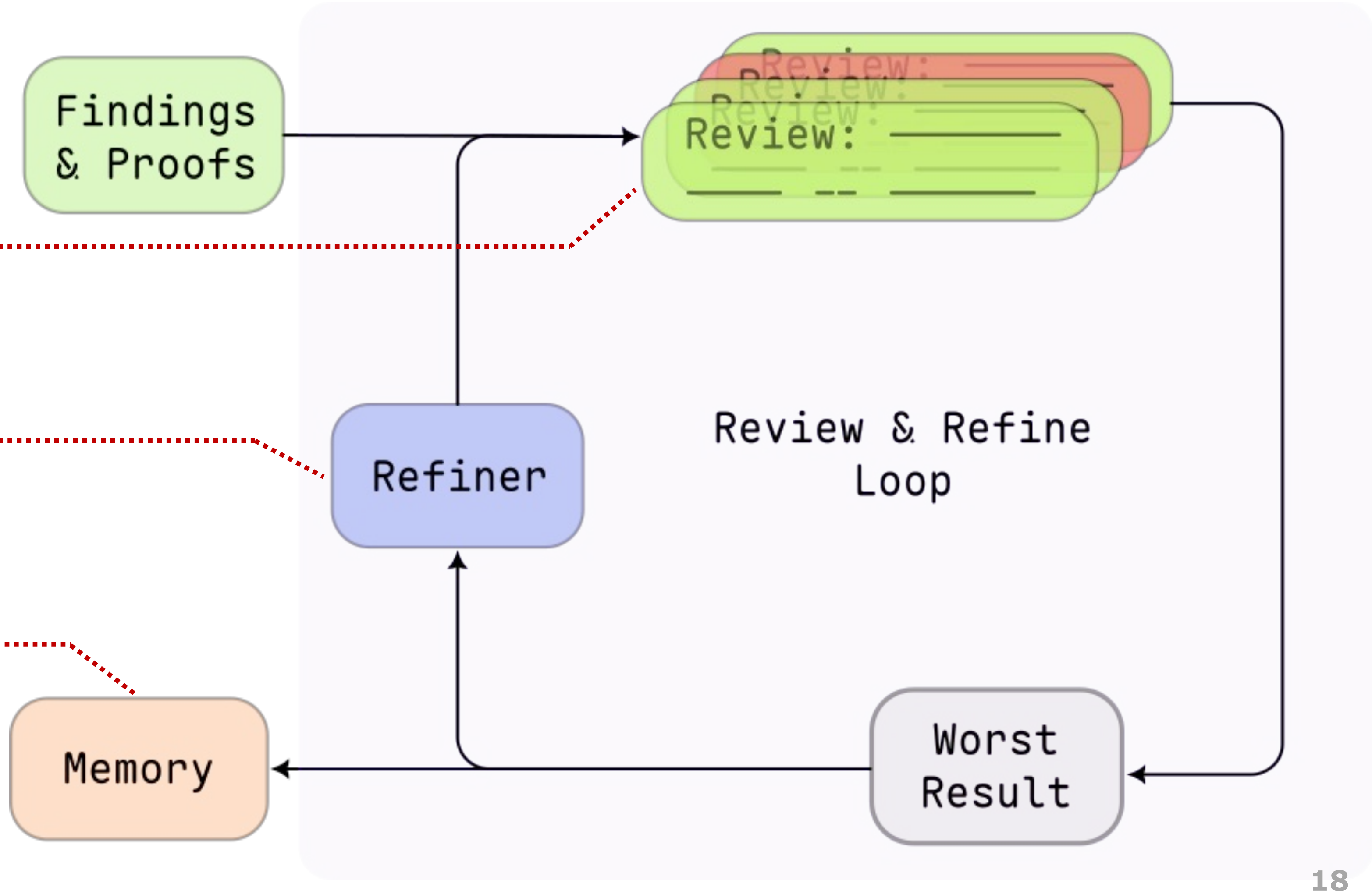
清華大學 Tsinghua University

- 相比较直接求解问题，AIM当中的智能体被要求针对原问题进行探索，并记录过程当中的发现作为引理；此后继续迭代这一过程即可推进探索进度，并最终实现问题求解。



结果验证：悲观检验与调优机制

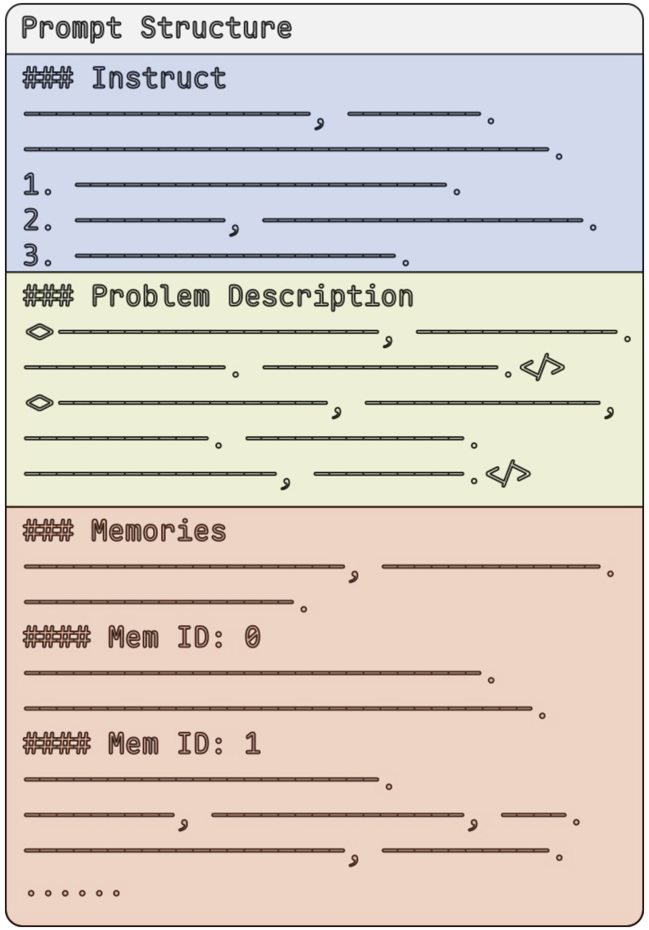
- **对探索过程中的每一个新发现重复检验多次，任意一次检验认为有错则判定新发现有错。**



- 进行多轮自我检查，能够有效地发现问题
- 存在错误的证明可以尝试修正
- 正确证明结论会最终存储到记忆当中

智能体设计

- 三个智能体均通过精心设计的prompt指导其行为以及输出格式，并辅以逻辑处理方法实现协作。



定义智能体行为

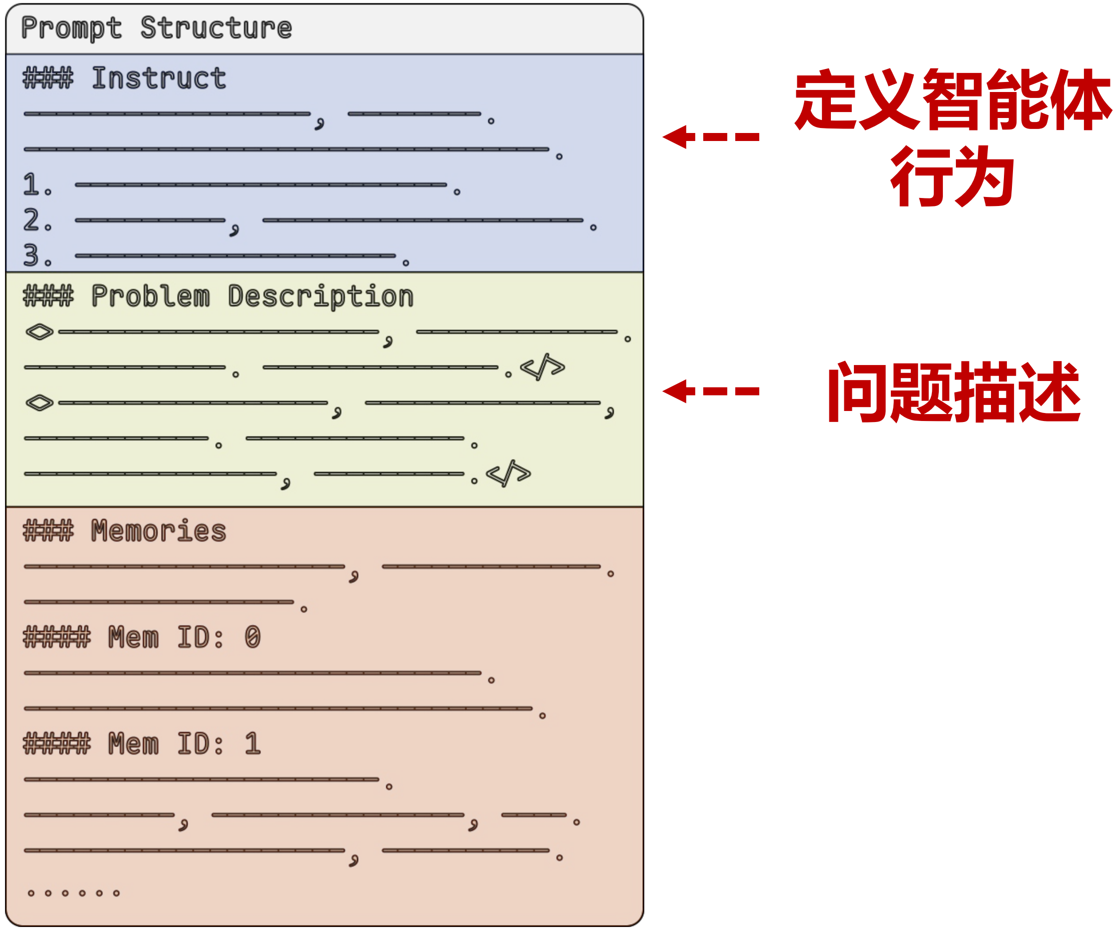
Prompt Example

主要内容包括：规定任务目标，推理方向，输出格式，以及过程当中的各种注意事项。

```
    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
```

智能体设计

- 三个智能体均通过精心设计的prompt指导其行为以及输出格式，并辅以逻辑处理方法实现协作。



Prompt Example

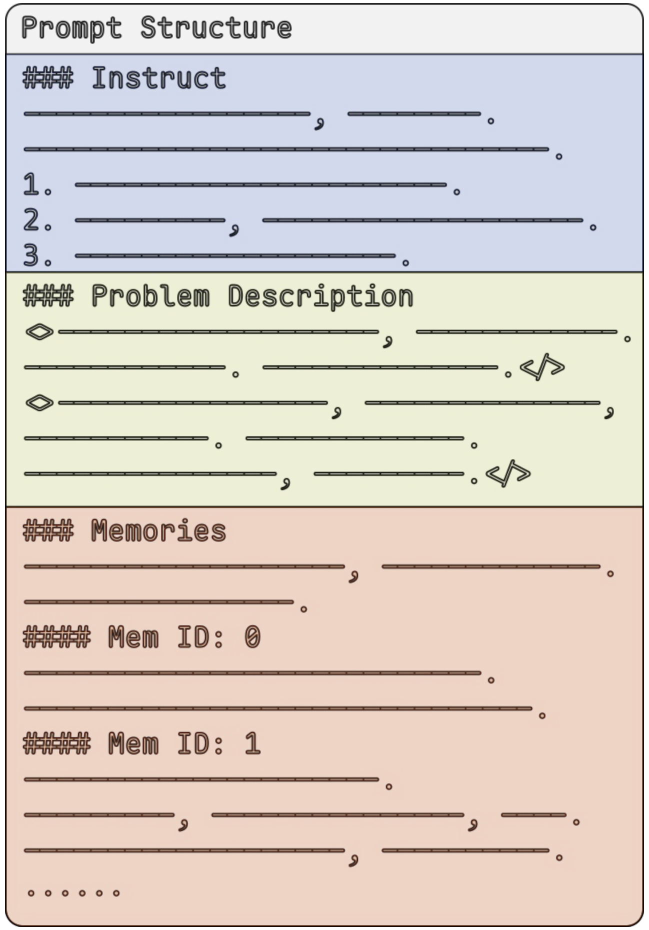
这部分内容是与问题目标直接相关的信息，对于不同智能体而言内容有所不同：

- Explorer：最终目标问题表述
- Verifier：待检验猜想与证明
- Refiner：存在疑问的猜想、证明以及从Verifier处获取的反馈

```
    \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:
```

智能体设计

- 三个智能体均通过精心设计的prompt指导其行为以及输出格式，并辅以逻辑处理方法实现协作。



定义智能体行为

问题描述

Memory中的信息

Prompt Example

经过格式化之后的历史探索轨迹以及相关结论。

```
### 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设计

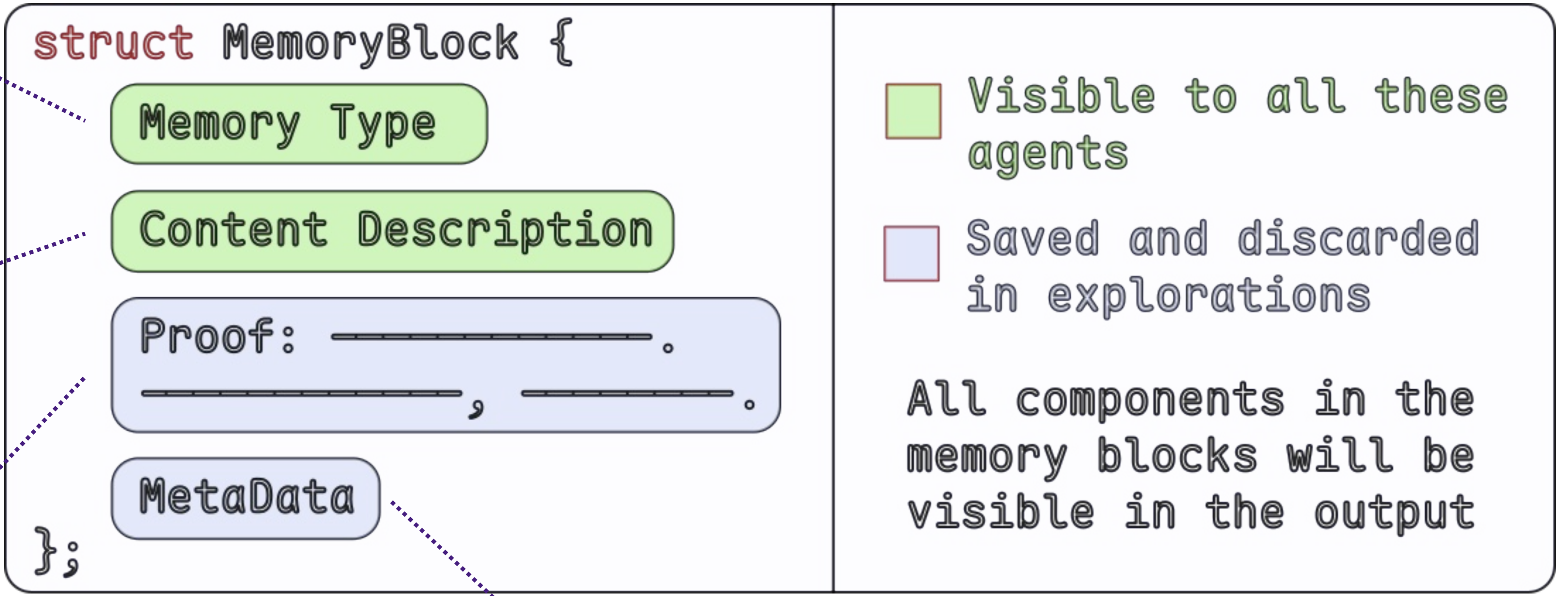
Tsinghua University

- **系统从模型输出当中解析并记录四类数据，其中的部分内容会被格式化并进入到后续智能体的输入当中。**

记忆块类型，包括context、lemma、conjecture等

记忆块内容的文本表述

对于记忆内容的完整证明



其他相关信息，如猜想是否解决、经过检验次数、依赖关系标记等等

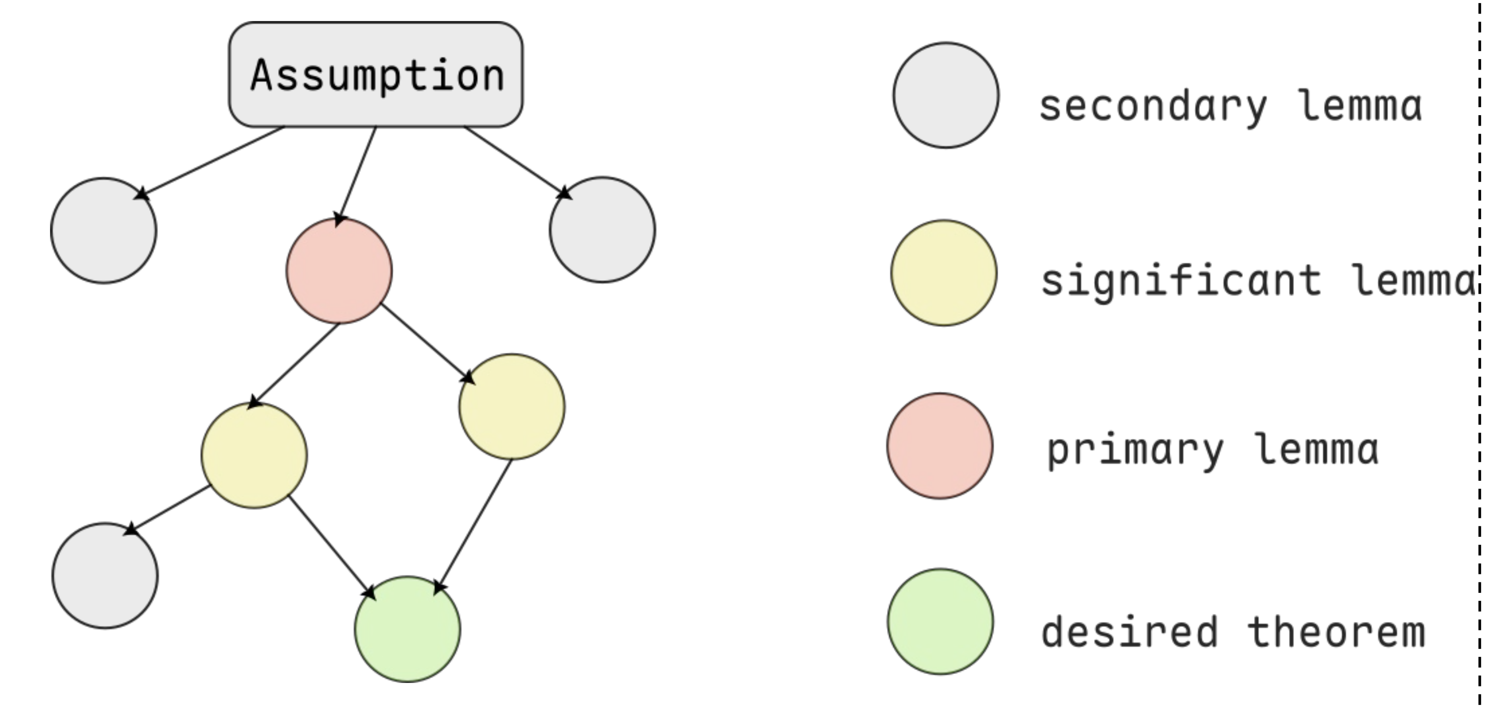
引理图

- **引理图是一个有向无环图（DAG），其中节点表示引理，边“A → B”表示要证明节点B所代表的引理时需要使用节点A所代表的引理。**

收益

- ❑易于回溯证明路径
- ❑对引理的相对重要性进行量化评估（如通过计算图中每个节点的子节点数目来进行估计）

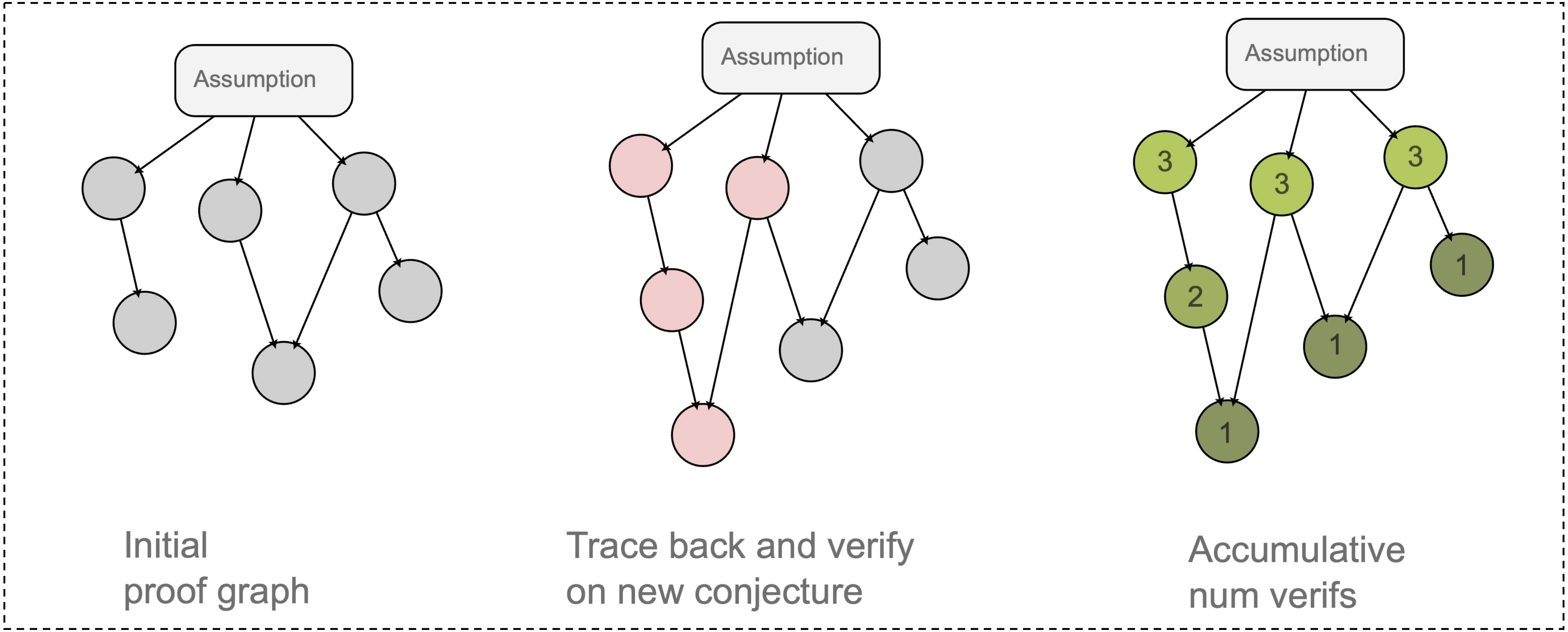
引理图示列



基于引理图的动态检验

Tsinghua University

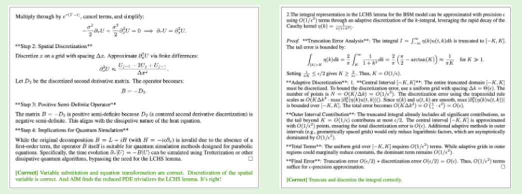
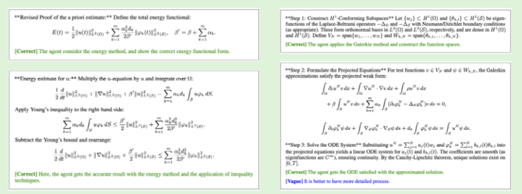
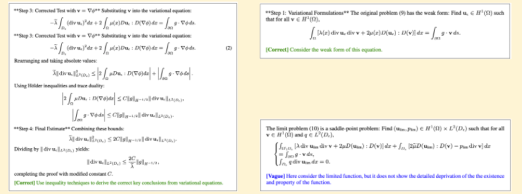
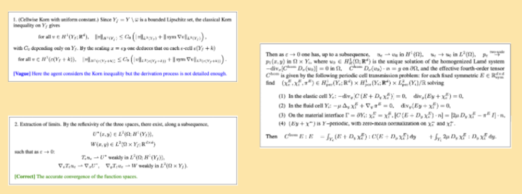
- 通过根据各引理的重要性动态分配验证资源，可以在提高效率的同时进一步保障可靠性。



实验概览

Tsinghua University 清华大学

- 运用AIM尝试求解了四个数学理论问题，其中前三个是已有结论的问题，最后一个是未被求解的开放性问题。

数学问题	AIM效果	部分证明步骤	
量子算法问题 (结论已知) Quantum Algorithm Problem	AIM完成了带有详细分析推导过程的证明		DeepSeek-R1
吸收边界问题 (结论已知) Absorbing Boundary Condition	AIM基本完成此问题的证明过程和细节		DeepSeek-R1
高对比度极限 (结论已知) High Contrast Limit	AIM完成了主要结论的证明过程，并且证明了新的结论		DeepSeek-R1和o4-mini各进行一次实验
均匀化问题 (结论未知) Homogenization Problem	AIM给出了合理的证明思路和部分结论的证明，对数学工作者很有启发意义		o4-mini

结论已知

结论未知

实例分析：量子算法问题

Linear Combination of Hamiltonian Simulation（LCHS）方法是科学计算中的一种高效方法，主要思路是将非酉动力学问题转化为为哈密顿模拟的线性组合

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

Black-Scholes-Merton(BSM)模型是金融学中用于欧式期权定价的基本数学框架

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

目标：利用LCHS方法对BSM模型进行模拟，设计相应的量子算法，并分析算法的复杂度

AIM正确运用了这一方法，提供了较为详细的证明过程，基本解决这一问题

实例分析：量子算法问题

清华大学 Tsinghua University

系统输入

LCHS方法说明

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.

BSM模型说明

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 $S \to \infty$; $V(S,T) = \max\{S-K, 0\}$.

目标

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)).

实例分析：量子算法问题

- 具体来说，AIM通过将BSM方程转化为热方程并利用LCHS引理转化问题并给出量子算法，基本解决了该问题，核心步骤如下：

1	偏微分方程转换与空间离散化	通过变量替换将BSM偏微分方程转换为标准热方程，并对空间变量进行离散化处理
2	算子分解与积分离散化	应用LCHS引理将热方程的解算子分解为酉算子的连续线性组合，随后对连续积分截断近似并离散化
3	量子算法实现与复杂度分析	利用酉算子线性组合（LCU）和哈密顿模拟技术设计量子算法，并计算复杂度

实例分析：量子算法问题

清華大學 Tsinghua University

AIM通过层层递进的多个引理逐步探索：

引理1：AIM先将BSM方程转化为热方程，再离散化空间变量，验证半正定性；进一步转化为可以利用LCHS引理的形式，证明过程详细

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 to parabolic equations.

将BSM方程转化为热方程

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:

$$B = -D_2.$$

空间变量离散化

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!

实例分析：量子算法问题

清华大学 Tsinghua University

AIM通过层层递进的多个引理逐步探索：

引理2：AIM分析了截断误差并确定离散间距，为构建量子算法奠定基础

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, $K = O(1/\epsilon)$.

分析截断误差

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 $O(K\Delta k^2) = O\left(\frac{1}{\epsilon}\cdot\epsilon^2\right) = O(\epsilon)$.

确定离散间距

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. □

总误差分析

[Correct] Truncate and discretize the integral correctly.

实例分析：量子算法问题

Tsinghua University

AIM通过层层递进的多个引理逐步探索：

AIM 结合之前的引理，使用哈密顿模拟和酉算子的线性组合（LCU）构建了一个量子算法，然后计算其复杂度；复杂度计算有一些错误，但核心方法是正确的

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 ... Each term ... Hamiltonian simulation of L scaled by $k_j\tau$. The total query complexity ... L-oracle is ... gate complexity scales polynomially with $\|L\|$, ...

在第一个引理基础上，利用LCHS引理转化为酉算子的连续积分

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 $\eta(k)$.

将方程解转化为酉算子的线性组合

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

利用哈密顿模拟算法

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 $O\left(\|L\| t_j + \log(1/\epsilon_{\text{sim}})\right)$, 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:

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

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:

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

复杂度分析

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.

均匀化问题

清華大學 Tsinghua University

- **问题描述：传输系统均匀化问题，需分析在具体物理尺度极限下，方程和相应的解的性质，并最终证明解的误差估计，是一个开放性问题。**

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

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

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正确运用数学定理和分析推导，给出了一些正确的结论和思路

对数学家完成此项研究具有指导意义

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

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

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 $\alpha \in (0,1)$

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

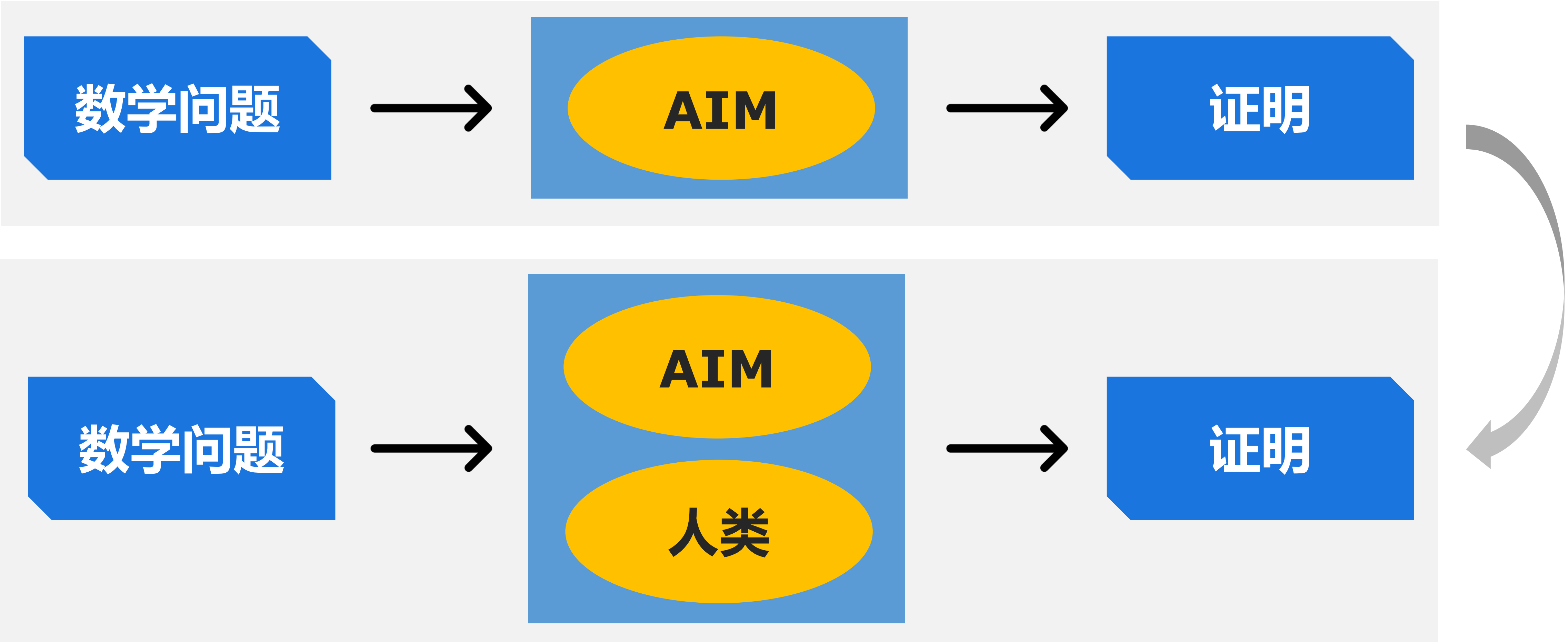
From this one deduces $u_\varepsilon \to u_0$ strongly in $L^2(\Omega)$ (and even in $H^1(\Omega_\varepsilon)$ up to the corrector).

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

正确的渐进展开思路和结论探索

人机协同解决数学问题

- **目标：在尽可能减少人类干预的条件下，通过人机协同的方式证明前述均匀化问题。**



人机协同解决数学问题

清華大學 Tsinghua University

- 完成近17页数学证明，AIM在难度最高的环节作为非平凡贡献。

完全由AIM生成（人工调整格式）

橙色：AIM给出具有较高完成度证明后人工校正

蓝色：人工完成

黑色：问题描述等

人机协同解决数学问题

清华大学 Tsinghua University

- **将均匀化问题的证明拆解为6个子问题：**

子问题	难度	当前状态
方程展开	低	人为正确展开
均匀化处理	中	人为构建了合适的cell problem，并且推导出正确的均匀化方程
存在唯一性	高	在有限提示下，AIM自动搜寻到正确的定理，给出证明；人补足部分证明细节
算子椭圆性	中	在有限提示下，AIM完成证明，人补足细节
误差估计	高	在有限提示下，AIM给出了正确的证明思路和部分过程，在人为调整后得到了完整的证明过程
正则性	高	在有限提示下，AIM给出了完整的证明过程

背景：Cell Problem

在推导均匀化方程的过程中，构造Cell Problem是一个必要的技术步骤；具体而言，我们手动构造了如下所示的Cell Problem

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

子问题：Cell Problem的正则性证明

清华大学 Tsinghua University

AIM 获得了一个良好的误差估计

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

人工检查发现证明过程中AIM依赖于前述Cell Problem的一个性质但并未给出证明

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 $u_0 \in H^2(\Omega)$*

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

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

初步判断上述性质很可能是正确的，进一步尝试利用AIM证明该性质

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

- Difference Quotinet
- Schauder Theory
- Galerkin Method

子问题：Cell Problem的正则性证明

清華大學 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 $B(1) = \{\|x\| \leq 1\}$. Consider this equation: for $V \in H_0^1(B(1); R^d)$*

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

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

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

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

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

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

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

$$\exists c \in R^d, s.t. \quad M_+ M_- = Ce_d^T \tag{4}$$

$$(I - e_d^T e_d)M_+ = (I - e_d^T e_d)M_- \tag{5}$$

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

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

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

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

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

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.*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

人为将 Schauder Theory中的引理调整为适用于 Cell Problem 方程形式

我们使用 AIM 补全了与 Schauder Theory相关的引理。这些内容作为“上下文”部分输入模型，为后续正则性证明提供了方法指导

子问题：Cell Problem的正则性证明

清華大學 Tsinghua University

- **我们将问题转化为对下述定理的证明：**

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*

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

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

人为将 Schauder Theory中的引理调整为适用于 Cell Problem 方程形式

我们指示 AIM 使用 Schauder Theory 来证明该定理。根据实验结果的反馈，我们迭代地拆解问题，最终完成了证明

子问题：Cell Problem的正则性证明

Tsinghua University

- 整个问题被划分为以下三个部分，AIM 依次补全每一部分的证明细节。经过多轮迭代后，AIM 给出了一个完整度较高的推理过程，并最终完成了该定理的证明。

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

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

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

χ is C^β

Hölder continuous

3 Hölder Regularity

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

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

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

Beyond expectations

均匀化问题的最终结论

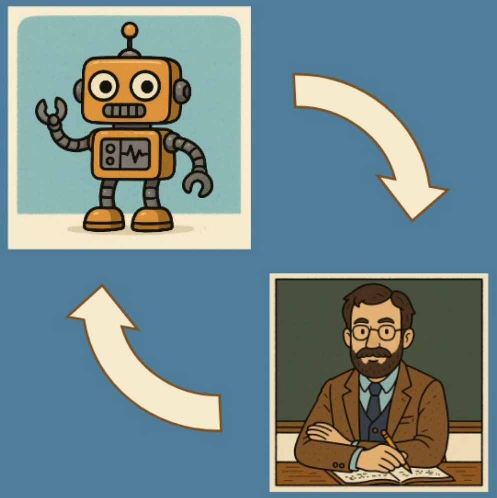
基于所获得的正则性结果，我们完成了对均化极限的误差控制

$$\left\|u_\varepsilon - u_{\lim}\right\|_{H^1(\Omega)} \lesssim \varepsilon^{\frac{1}{2}}$$

以上是原始方程解与均化方程解之间的误差控制结论

证明全文 & 经验总结

Key Takeaway



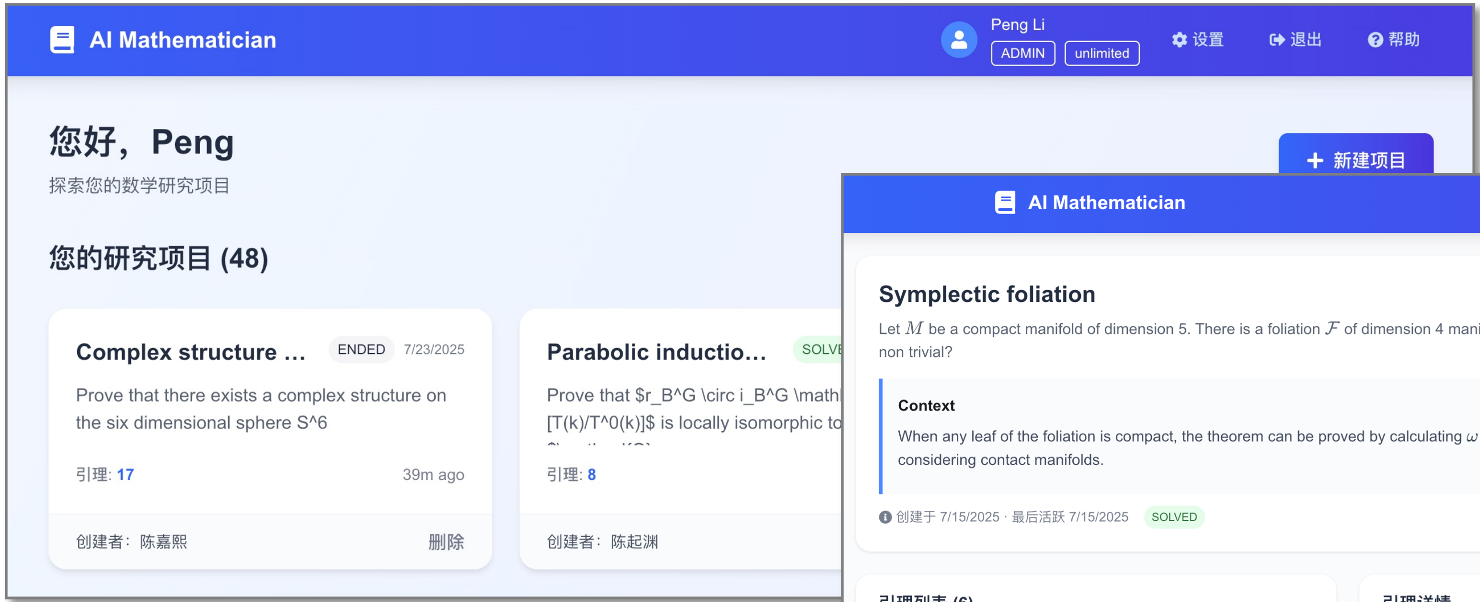
AI may still be a flawed individual researcher today.

Yet, it can already serve as a valuable research partner—if used wisely.

AIM在线系统

清華大學
Tsinghua University

- 我们为AIM系统开发了网页前端，并且邀请了一些数学研究人员来体验和评测AIM的系统性能。



系统主页

项目详情页

部分用户反馈

清华大学
Tsinghua University

AIM 成功解决了一些群分类问题，并提供了详细的证明过程 😄

Classify certain finite subgroup of SO(4)

Suppose A is a finite subg… (4) with
line bundle l on $S^3 \backslash Z$ 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

这是一个关于辛叶状分布的难题。用户在提问时意外遗漏了一个条件，AIM 针对该问题提供了一个有效的反例 😄

AIM 未能直接解决该研究级问题，该问题的复杂度已超出 AIM 的可处理范围 ☹

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

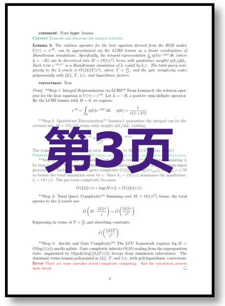
状态: ⌛ 待处理　重要性: 次要　评审次数: 6　依赖: 2, 3, 10

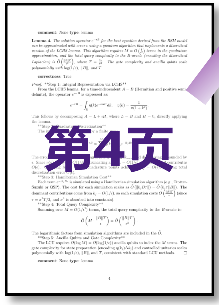
引理陈述:

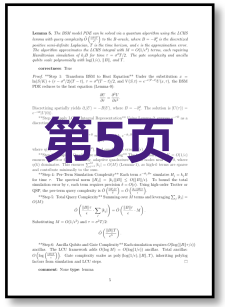
Let (M, g) be as in Memory ID 0, with tubular neighbourhood $N \cong S^1_s \times D_{r,\theta}$ and real line bundle $l \to N \setminus K$ of monodromy -1. Suppose moreover that in local coordinates (s, r, θ)

局限1：重复性探索

- 智能体常沿相同方向进行探索，提出一系列相似猜想和相同引理，推高成本、降低效率并制约性能上限。





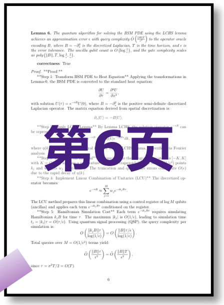


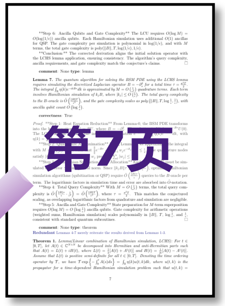
Lemma 4. *The solution operator $e^{-\tau B}$ for the heat equation derived from the BSM model can be approximated with error ϵ using a quantum algorithm that implements a discretized version of the LCHS lemma. This algorithm requires $M = O\left(\frac{1}{\epsilon^2}\right)$ terms in the quadrature approximation, 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.*

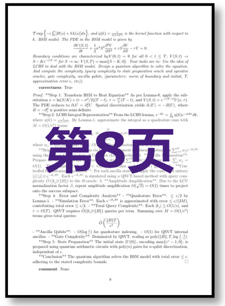
Lemma 6. *The quantum algorithm for solving the BSM PDE using the LCHS lemma achieves an approximation error ϵ with query complexity $\tilde{O}\left(\frac{\|B\|T}{\epsilon^3}\right)$ to the operator oracle encoding B, where $B = -\partial_x^2$ is the discretized Laplacian, T is the time horizon, and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

3个相似的引理

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 $k_j B$, 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 $\mathrm{poly}\left(\|B\|, T, \log\frac{1}{\epsilon}, \frac{1}{\epsilon}\right)$, with ancilla qubit count $O\left(\log\frac{1}{\epsilon}\right)$.*







局限2：数学设定理解不足

清華大學 Tsinghua University

- 当前AIM对于较为冗长的数学设定、背景条件的理解能力还不够强，导致智能体在分析过程中产生错误。

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 $L^2(\Omega)$, $p_\varepsilon \xrightarrow{\text{two-scale}} p_1(x, y)$ 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: $-\mu\,\Delta_y\,\chi_i^E + \nabla_y\,\pi^E = 0$, $\boxed{\mathrm{div}_y(Ey + \chi_i^E) = 0,}$

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

此处缺乏对晶胞的边界理解，导致给出了不符合实际情况的结论

局限3：推导细节缺失

清華大學 Tsinghua University

- **当前AIM会出现“跳步”现象，不利于证明过程的人工或自动化验证，同时可能导致错误的证明过程。**

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.

此处并未给出ODE系统的具体推导分析过程，但是是正确的结论

局限4：需求与供给矛盾

清華大學 Tsinghua University

- **实际用户需求呈现出高度多样性，且显著超出 AIM 当前能力范围。**

用户期望	AIM当前能力
使用尽可能简洁的方式输入问题	问题描述更加具体时将获得明显更优的效果
中文用户习惯使用中文输入问题	采用英文表达时AIM可以获得更好的效果
错误零容忍	当前无法做到零错误
预期系统可以开箱尽用且可以解决任意问题	AIM仍然具有诸多局限性
……	……

展望：一场关于自然语言的豪赌

清華大學 Tsinghua University

- **AI在数学研究中起作用的具体技术路线仍具争议，将自然语言作为这一过程的第一公民将是一声豪赌，但越来越多证据支持赌局将获胜。**

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

展望：AI作用从辅助到主动

辅助证明 + 思路验证 + 开放探索

终极目标 理论问题 → AI数学家 → 理论证明

展望：AI将有助于提出高价值数学问题

- 好问题对推动数学发展至关重要；我们期望AIM能协助提出此类问题。

Fermat’s Last Theorm

ELLIPTIC CURVES

MATHEMATICAL COLLABORATION AND SYNTHESIS

MODULAR FORMS

AUTOMORPHIC FORMS

ALGEBRAIC NUMBER THEORY

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

总结

- 大模型的快速发展为AI在数学研究中发挥更大作用提供了基础
- 数学研究区别于数学解题的三个关键特征：
  - 长线思考：数学研究需要更多、更长期、更深入有效思考
  - 过程严谨：错误零容忍，需可验证，但验证代价高昂
  - 答案存在性：数学研究答案存在性未知且不确定
- 提出AI数学家系统AIM，并在四个研究级数学问题上取得初步成效，潜力可期待
- 未来AI将在数学研究中起到更加主动且重要的作用

AIM技术报告

Blog

Our AIM is AI Mathematician!