



北京大学数学科学学院

School of Mathematical Sciences  
Peking University

# AI × MATH SYMPOSIUM 2026

人工智能与数学研讨会

MATHEMATICS FOR AI · AI FOR MATHEMATICS

27–28 AUGUST 2026

PEKING UNIVERSITY · BEIJING

主办 中国数学会数学与人工智能专业委员会

承办 北京大学数学科学学院、北京大学国际数学研究中心、国际机器学习研究中心

ABOUT / 关于会议

# AI × Math Symposium 2026

人工智能正以前所未有的速度改变科学研究的工具、方法与组织方式。数学既是人工智能发展的重要理论基础，也正在借助人工智能开拓新的研究范式。

本次论坛围绕“Mathematics for AI”与“AI for Mathematics”两条主线，设置开幕式、主旨报告、专题报告与圆桌对谈，交流数学基础推动人工智能发展，以及人工智能赋能数学发现、证明与研究协作的最新进展。

AT A GLANCE / 会议信息

|                 |                                         |
|-----------------|-----------------------------------------|
| 时间 / DATE       | 2026 年 8 月 27 日（星期四）至 8 月 28 日（星期五）     |
| 地点 / VENUE      | 北京大学智华楼文远堂；北京大学甲乙丙报告厅                   |
| 主线 / THEMES     | Mathematics for AI · AI for Mathematics |
| 主办 / HOST       | 中国数学会数学与人工智能专业委员会                       |
| 承办 / ORGANIZERS | 北京大学数学科学学院、北京大学国际数学研究中心、国际机器学习研究中心      |

CONTENTS / 目录

|    |                      |    |
|----|----------------------|----|
| 01 | 会议概览 / About         | 02 |
| 02 | 会议日程 / Programme     | 03 |
| 03 | 主旨报告 / Keynotes      | 05 |
| 04 | 专题报告 / Invited Talks | 06 |

# DAY 01 · 27 AUGUST

THURSDAY

## Opening & Keynotes / 开幕式与主旨报告 08:30–12:00 · 北京大学智华楼文远堂

| TIME / 时间   | PROGRAMME / 报告                                                                                                                       | SPEAKER / 报告人       |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| 08:30–08:40 | 开幕致辞                                                                                                                                 | 席南华 院士              |
| 08:40–09:00 | 专委会设立情况汇报                                                                                                                            | 刘若川 · 北京大学          |
| 09:00–09:30 | Keynote 1: Math, Coding and Long-horizon Tasks: a Bitter-Lesson Perspective                                                          | 刘征瀛 · Kimi          |
| 09:30–10:00 | Keynote 2: 大模型如何学会数学:从思维链到形式化证明                                                                                                      | 李嘉 · Project numina |
| 10:00–10:20 | 茶歇及合影                                                                                                                                |                     |
| 10:20–10:50 | Keynote 3: From Olympiad Problem Solving to Mathematical Research: Progress, Challenges, and Prospects for Advanced Reasoning in LLM | 涂毅 · 华为小艺           |
| 10:50–11:20 | 人工智能赋能数学研究：范式变革与未来展望                                                                                                                 | 董彬 · 北京大学           |
| 11:20–12:00 | Panel 对谈   AI 与数学的融合：机遇、挑战与协作路径                                                                                                      | 圆桌嘉宾                |

### AFTERNOON · AI FOR MATHEMATICS 14:00–16:00 · 北京大学甲乙丙报告厅

|             |                                                                   |                        |
|-------------|-------------------------------------------------------------------|------------------------|
| 14:00–14:30 | 计算机辅助设计与工程：从一体化到智能化                                               | 陈景润<br>中国科学技术大学        |
| 14:30–15:00 | MechGeo: Autoformalizing and Proving Euclidean Geometry in Lean 4 | 支丽红<br>中国科学院数学与系统科学研究院 |
| 15:00–15:30 | 抽象数学与形式化思维：AI 背景下的数学研究与教育                                         | 陈华一<br>西湖大学            |
| 15:30–16:00 | AI × Combinatorics: A Personal Journey                            | 马杰<br>中国科学技术大学         |
| 16:00–16:20 | 茶歇                                                                |                        |

### AFTERNOON · MATHEMATICS FOR AI 16:20–17:50 · 北京大学甲乙丙报告厅

|             |                                                                                       |             |
|-------------|---------------------------------------------------------------------------------------|-------------|
| 16:20–16:50 | 通过压缩视角理解大语言模型行为：知识获取，扩展定律与上下文学习                                                       | 李建<br>清华大学  |
| 16:50–17:20 | Math4AI: From Learning from Samples to Learning to Sample                             | 焦雨领<br>武汉大学 |
| 17:20–17:50 | Manifold Constraints Unlock Multi-Timescale Momentum for Accelerating LLM Pretraining | 袁坤<br>北京大学  |

## PROGRAMME / 会议日程

## DAY 02 · 28 AUGUST

FRIDAY

Mathematics for AI 09:00–12:50 · 北京大学甲乙丙报告厅

| TIME / 时间   | PROGRAMME / 报告                                                                        | SPEAKER / 报告人      |
|-------------|---------------------------------------------------------------------------------------|--------------------|
| 09:00–09:30 | Deep Generative Learning from Unpaired Data: Mathematical Models and Applications     | 包承龙<br>清华大学        |
| 09:30–10:00 | A Geometric Characterization of the Stationary Point Evolution in Neural Networks     | 孙若愚<br>香港中文大学 (深圳) |
| 10:00–10:30 | 从代数几何视角看大模型表达能力边界                                                                     | 刘勇<br>中国人民大学       |
| 10:30–10:50 | 茶歇                                                                                    |                    |
| 10:50–11:20 | 基于现象的数学理论理解深度学习                                                                       | 许志钦<br>上海交通大学      |
| 11:20–11:50 | 智能的微积分：一个智能体的理论框架<br>The Calculus of Intelligence: A Theoretical Framework for Agents | 袁洋<br>清华大学         |
| 11:50–12:20 | 大模型与智能体如何从经验中学习：从参数泛化到外部记忆优化                                                          | 李健<br>北京师范大学       |

## DINING / 餐饮

8月27日午餐、晚餐、8月28日午餐：中关新园和园。自助餐券在签到时发放。



## Keynotes · Abstracts

### KEYNOTE 01

## Math, Coding and Long-horizon Tasks: a Bitter-Lesson Perspective

刘征瀛 · Kimi

### ABSTRACT / 摘要

In this talk, I will review our explorations in mathematics from Kimina-Prover to Kimi K2 / K2 Thinking, and further to Kimi K3, as well as a bitter lesson we have learned firsthand. After clarifying the direct correspondence between mathematics and coding, and examining existing AI paradigms built upon various harnesses, we will discuss whether current AI paradigms are already sufficient to tackle open mathematical problems—or even extremely challenging ones such as the seven Millennium Prize Problems.

### KEYNOTE 02

## 大模型如何学会数学:从思维链到形式化证明

李嘉 · Project Numina

### ABSTRACT / 摘要

大语言模型最初通过预测下一个词学习语言，却在过去几年中迅速获得了求解数学问题、生成复杂推理乃至构造形式化证明的能力。这一进展并非单纯来自模型规模的增长，而是由高质量数学数据、思维链监督、过程奖励、强化学习、工具调用和测试时计算等方法共同推动。本报告将回顾大模型数学能力的发展历程：从 MATH、Minerva 等早期工作，到代码辅助推理过程奖励模型和大规模合成数据，再到 o1、DeepSeek-R1 所代表的长程推理与强化学习范式。结合 Project Numina 的实践，我们将介绍如何通过 NuminaMath 数据集与开放模型训练数学推理能力，以及如何在 Kimina Prover 中将证明搜索轨迹转化为结合自然语言推理与 Lean 代码的统一推理过程，并利用强化学习提升形式化证明能力。最后，报告将讨论数学 AI 正在经历的关键转变：从解答有标准答案的竞赛题，走向生成可机器验证的证明、形式化已有数学成果，并进一步参与开放问题和前沿数学研究。

### KEYNOTE 03

## From Olympiad Problem Solving to Mathematical Research: Progress, Challenges, and Prospects for Advanced Reasoning in LLM

涂毅 · 华为小艺

### ABSTRACT / 摘要

近年来，大模型的数学能力正从标准化评测与竞赛解题，迈向开放、前沿且需要专家审查的研究型数学。本报告结合大模型高阶推理的研发实践，梳理与讨论这一能力跃迁背后的关键技术（数据与后训练、长程搜索与自我纠错、工具使用与形式化验证，以及支撑规模化推理的高效 AI 系统）。

相较于边界清晰、答案确定的竞赛问题，数学研究要求模型面对开放问题和未知路径，完成文献理解、猜想生成、反例搜索与严格证明。近期，以 GPT、Claude 为代表的前沿模型已在部分开放问题上提出新思路并取得可验证进展，AI 正从“问题求解者”向“研究伙伴”演进。本报告将进一步探讨数学研究者与 AI 研究者如何共建基础设施与协作范式，推动 AI for Mathematics 与 Mathematics for AI 的双向融合。

## Invited Talk Abstracts

### INVITED TALK 08

## Deep Generative Learning from Unpaired Data: Mathematical Models and Applications

包承龙 · 清华大学

### ABSTRACT / 摘要

Many deep learning methods rely on paired training data. However, in many scientific applications, only unpaired observations from different domains are available, making it challenging to learn their underlying relationships. In this talk, I will present our recent work on developing mathematical models for deep generative learning from unpaired data. We formulate unpaired learning as a probabilistic inference problem and develop generative frameworks to recover joint distributions from marginal observations. I will demonstrate these ideas through applications in scientific imaging, particularly cryo-electron microscopy, and conclude with future directions for generative AI under limited supervision.

### INVITED TALK 09

## 抽象数学与形式化思维：AI 背景下的数学研究与教育

陈华一 · 西湖大学

### ABSTRACT / 摘要

人工智能正在改变数学知识的获取、计算与证明方式，也促使我们重新思考数学研究与教育的目标。本报告将结合西湖大学大一数学教学的实践，讨论抽象概念、形式语言与严格证明在数学训练中的作用，并探讨 AI4Math 背景下人机协作可能为数学研究与教育带来的机遇与挑战。

### INVITED TALK 10

## 计算机辅助设计与工程：从一体化到智能化

陈景润 · 中国科学技术大学

### ABSTRACT / 摘要

全流程工业设计过程需要迭代式的几何设计与工程仿真。对工程师来说，实际的业务痛点有：（1）设计与仿真流程切换以及人工干预；（2）操作门槛高且耗时，需要大量的行业知识。从应用与计算数学角度看，这对计算方法提出了新的要求：

（1）如何将物理模型的适定性等知识嵌入到工业设计仿真流程中去；（2）如何提高计算方法的鲁棒性以实现工业设计流程的顺畅流转。本次报告首先介绍随机特征方法，旨在消除仿真中的网格划分步骤，更好地实现设计与仿真的衔接；

再介绍工业软件智能体，内嵌行业知识，实现软件操作的自动化、智能化。

### INVITED TALK 01

## Math4AI: From Learning from Samples to Learning to Sample

焦雨领 · 武汉大学

### ABSTRACT / 摘要

Artificial intelligence has undergone a fundamental paradigm shift over the past decade, evolving from learning from samples in statistical learning to learning to sample in modern generative AI. This transition has not only transformed the design of learning algorithms, but has also raised profound mathematical questions concerning function/data representation, statistical learning, optimization, and stochastic generation. This talk presents the mathematical foundations of modern artificial intelligence through this evolution.

## 大模型与智能体如何从经验中学习：从参数泛化到外部记忆优化

李健 · 北京师范大学

### ABSTRACT / 摘要

大模型与智能体的学习正在由参数空间逐步扩展到上下文与外部经验空间。预训练和微调将数据中的知识压缩至模型参数，上下文学习与检索增强使模型能够在推理阶段利用示例和外部知识，智能体进一步将交互轨迹转化为可持续更新的记忆、技能与 workflow。这一演进提出了一个共同的基础问题：如何从有限且含噪的经验中提取可泛化信息，并避免参数空间中的噪声拟合以及外部经验空间中的冗余积累和局部过拟合？本报告将沿上述演进脉络，梳理从预训练与微调、上下文学习与检索增强，到记忆、技能和工作流持续优化的代表性方法，并结合参数学习的优化动力学与泛化理论、上下文学习的内在机制，以及外部经验的表示、压缩与离散优化，分析不同学习方式的能力边界与泛化问题。在此基础上，报告将讨论参数学习与外部经验学习的统一视角，并提出当学习对象由模型权重扩展为可检索、可编辑、可组合的经验结构时，学习理论需要回答的新问题。

## 通过压缩视角理解大语言模型行为：知识获取，扩展定律与上下文学习

李建 · 清华大学

### ABSTRACT / 摘要

大语言模型 (LLMs) 在众多任务中展现出卓越能力，但其内在机制及扩展定律、幻觉现象等相关行为仍缺乏系统性理论解释。本研究算法信息论重新审视压缩与预测的经典关系，从而深入理解大语言模型的行为机制。通过运用 Kolmogorov 结构函数并将 LLM 压缩理解为双部分编码过程，我们详细揭示了模型如何随参数数量和数据规模增长获取并存储信息——从普遍存在的句法模式到逐渐稀疏的知识要素。受此理论视角及 Heap 定律与 Zipf 定律启发的自然假设驱动，我们提出一个简化但具有代表性的分层数据生成框架——句法-知识模型。在贝叶斯框架下，研究表明该模型中的预测与压缩过程可自然推导出大语言模型的多样化学习行为与扩展规律。特别地，我们的理论分析为数据与模型扩展定律、训练与微调中的知识获取动态、大模型的上下文学习能力，以及 LLM 事实性知识幻觉现象提供了直观的理论的解释。我们在大语言模型上的实验结果验证了理论预测的合理性。

## 从代数几何视角看大模型表达能力边界

刘勇 · 中国人民大学

### ABSTRACT / 摘要

当前学界对大模型表达能力的分析多从函数构建的角度展开，本报告引入代数几何视角，重点依托热带几何理论对大模型基础架构开展机制性分析。报告针对混合专家 (MoE) 模型与 Transformer 架构的表达能力进行了系统性剖析，为大模型底层机制的研究提供了全新的理论视角与分析路径。

## AI × Combinatorics: A Personal Journey

马杰 · 中国科学技术大学

### ABSTRACT / 摘要

In this talk, I will explore the evolving interaction between AI and combinatorics from a personal perspective. I will first review some recent breakthroughs in combinatorics driven by AI. I will then share my own journey as a combinatorialist through the rapid development of AI. Finally, I will discuss my perspectives on the role of AI in combinatorics (and beyond).

## INVITED TALK 03

## A Geometric Characterization of the Stationary Point Evolution in Neural Networks

孙若愚 · 香港中文大学 (深圳)

## ABSTRACT / 摘要

We investigate the geometric structure of stationary points that arise in the loss landscape of two-layer neural networks. We focus on the phenomenon of “neuron splitting” where duplicating a hidden neuron yields an affine set of stationary points in a wider network. We provide a comprehensive classification of all stationary points due to neuron splitting, determining under what conditions they constitute local minima or saddle points. Our characterization hinges on a per-neuron curvature object we term the “inner Hessian” matrix, which is the difference of the Hessian and Gauss-Newton matrix. Our analysis reveals that the definiteness of the inner Hessian and the choice of splitting coefficients jointly dictate the local geometry of the plateau. We show that “splitting” a local minimum can yield either a mixture of local minima and saddles or an all-saddle plateau, with a concrete sure-saddle region identified under mild assumptions. In contrast, splitting a saddle point always produces a plateau of saddle points. Our results elucidate when and how model expansion preserves or alters the nature of stationary points. These findings offer new geometric insights into the fine-grained geometrical structure of neural network landscape.

## INVITED TALK 04

## 基于现象的数学理论理解深度学习

许志钦 · 上海交通大学

## ABSTRACT / 摘要

深度学习的基础研究是重要的研究方向，其中一类重要的研究方法是现象驱动的数学理论研究。通过对重要现象的发现和挖掘，并发展相关的数学理论，可以使基础研究更好的服务于神经网络的理解和使用。本报告将介绍一些重要的现象和相关理论，包括频率原则、凝聚现象等，理解深度学习的原理，从最简单的感知机的学习到大模型的记忆和推理。参考《深度学习现象导论》。

## INVITED TALK 06

## Manifold Constraints Unlock Multi-Timescale Momentum for Accelerating LLM Pretraining

袁坤 · 北京大学

## ABSTRACT / 摘要

Pretraining accounts for a large fraction of the total computational cost in LLM training. However, noise-dominant gradients and the highly ill-conditioned loss landscape bring severe challenges. Although modern adaptive optimizers such as AdamW and Muon have achieved great success in large-scale pretraining, their nearly isotropic updates along different Hessian eigen-directions are suboptimal: progress along flat directions (eigen-directions of small eigenvalues), which dominate the final loss reduction, remains relatively slow. To address this limitation, we propose a multi-timescale momentum method with manifold constraints ( $M^3$ ) that enhances training dynamics along flat directions, yielding steady acceleration in pretraining. Our flat-directional multi-timescale momentum pairs a slow-decay component, which enhances the signal-to-noise ratio (SNR), with a fast-decay component that adapts rapidly to local curvature, harnessing their complementary strengths. Crucially, we employ a manifold constraint technique to avert parameter inflation and excessively rapid effective learning rate decay that would otherwise arise from a naive combination. Extensive experiments show that  $M^3$  significantly accelerates both Muon and AdamW across diverse architectures (dense, MoE), model sizes (120M--2.3B parameters), and learning-rate schedules (cosine, wsd). Theoretically, we verify the acceleration effect and provide insight into the design principles underlying the flat-direction multi-scale momentum.

## 智能的微积分：一个智能体的理论框架

The Calculus of Intelligence: A Theoretical Framework for Agents

袁洋 · 清华大学

### ABSTRACT / 摘要

智能体正在成为写代码、写作与规划等很多领域的基础工具。它们能极大放大个人能力，但要构建百倍、千倍规模的系统，仅靠经验式智能体框架仍然不够。本报告提出“智能的微积分”（COIN）：用格罗滕迪克拓扑斯（Grothendieck topos）描述类型化任务拆解，用自由单子构造可检查的行动蓝图，并通过层论与单子组合完成结果重组。这个框架意味着，通过正确地组织智能体，我们可以构建超大规模系统。

Agents are becoming foundational tools across many domains, including coding, writing, and planning. They can dramatically amplify individual capabilities, but heuristic agent frameworks alone are insufficient for building systems at hundredfold or thousandfold scale. This report introduces the Calculus of Intelligence (COIN): a framework that uses Grothendieck topos to model typed task decomposition, free monads to construct verifiable action blueprints, and sheaf theory together with monadic composition to integrate results. The framework suggests that, by organizing agents correctly, we can build systems of unprecedented scale.

## MechGeo: Autoformalizing and Proving Euclidean Geometry in Lean 4

支丽红 · 中国科学院数学与系统科学研究院

### ABSTRACT / 摘要

We present MechGeo, a Mathlib-native agentic framework for faithful autoformalization and certified proof construction in Euclidean geometry. It combines GeoFormalizer, which translates informal problems into Lean 4 and repairs formal statements, with GeoProver, which integrates geometric reasoning, selective algebraization, and symbolic computation. All proofs and counterexamples are checked by the Lean kernel.

On 43 historical IMO geometry problems, MechGeo proves 29 statements generated by GeoFormalizer, refutes the remaining 14 with Lean-verified counterexamples, and proves all 14 after expert repair. Together with IMO 2026 Problem 2, this yields faithful formalizations and kernel-checked proofs for 44 IMO geometry problems. On the 14 geometry statements in LEAP's Lean-IMO-Bench, MechGeo proves 12, refutes the remaining two, and proves both repaired statements.

Joint work with Hao Shen, Junyu Guo, Tian Cui, and Yuxuan Xiao.