

## EDUCATION

### Tsinghua University

Beijing, China

Dual Major: Computer Science and Technology &amp; Economics and Finance

2022 - 2026

- **GPA:** 3.9/4.0, **Ranked** 1-st (4.0/4.0) for the past year.
- **Selected A Courses:** Data Mining (A+), Ordinary Differential Equation (A+), Game Theory and Mechanism Design, Probability and Statistics, Artificial Intelligence, Advanced Linear Algebra, Discrete Mathematics, Computational Humanities and Social Sciences, Computer Networks, Computer Graphics, Calculus A(2), etc.
- **Academic Interests:** ML, NLP (AI for Math, ML for NLP).

## RESEARCH EXPERIENCES

### Princeton Language and Intelligence (PLI), Princeton | Remote

2024.11 - Present

- Research intern, advised by Prof. Chi Jin. Worked with Dr. Yong Lin.
- Working on automated theorem proving with formal language. Core member of Goedel project.
- I built high-performance training (SFT & RL) and inference infra, and studied algorithms for multi-stage continue-training and continuable RL training.

### Rose Spatiotemporal Lab, UCSD | San Diego, CA, USA

2024.07 - 2024.11

- Research intern, advised by Prof. Rose Yu. Worked with Dr. Yadi Cao.
- Led **adapting while learning** project that proposes a new learning paradigm for efficient scientific agents, allowing the models to self-evolve and act adaptively.

### Tsinghua NLP Lab | Beijing, China

2023.07 - 2024.06

- Research intern, advised by Prof. Zhiyuan Liu. Worked with Prof. Xin Cong, Dr. Yujia Qin.
- Led **OpenAct & OpenAgent** (ACL Main) and **Reasoning Prism** (in submission) projects, member of **XAgent** (GitHub 8.5k stars) team.

## PUBLICATIONS

1. **Bohan Lyu\***, Yadi Cao\*, Duncan Watson-Parris, Leon Bergen, Taylor Berg-Kirkpatrick, Rose Yu. **Adapting While Learning: Grounding LLMs for Scientific Problems with Intelligent Tool Usage Adaptation**. *International Conference on Machine Learning (ICML) 2025, AAAI Fall Symposium Series (Oral)* 2024, featured at *Agentic AI Summit 2025 @ Berkeley RDI*.
2. Yong Lin\*, Shange Tang\*, **Bohan Lyu\***, Ziran Yang\*, Jui-Hui Chung\*, Haoyu Zhao\*, Lai Jiang\*, Yihan Geng\*, Jiawei Ge, Jingruo Sun, Jiayun Wu, Jiri Gesi, Ximing Lu, David Acuna, Kaiyu Yang, Hongzhou Lin, Yejin Choi, Danqi Chen, Sanjeev Arora, Chi Jin. **Goedel-Prover-V2: Scaling Formal Theorem Proving with Scaffolded Data Synthesis and Self-Correction**. *AI4Math @ ICML (Oral)* 2025.
3. **Bohan Lyu\***, Xin Cong\*, Heyang Yu, Pan Yang, Yujia Qin, Yining Ye, Yaxi Lu, Zhong Zhang, Yukun Yan, Yankai Lin, Zhiyuan Liu, Maosong Sun. **Enhancing LLM's Capabilities in Open Domains via Autonomous Tool Integration**. *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL Main)* 2025.
4. Yong Lin\*, Shange Tang\*, **Bohan Lyu\***, Jiayun Wu, Hongzhou Lin, Kaiyu Yang, Jia Li, Mengzhou Xia, Danqi Chen, Sanjeev Arora, Chi Jin. **Goedel-Prover: A Frontier Model for Open-Source Automated Theorem Proving**. *Conference on Language Modeling (COLM)* 2025.
5. **Bohan Lyu\***, Siqiao Huang\*, Zichen Liang\*, Qi-An Sun, Jiaming Zhang. **SURGE: On the Potential of Large Language Models as General-Purpose Surrogate Code Executors**. *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP Main, top 0.3% meta score)* 2025.

## HIGHLIGHTED PROJECTS

### Goedel Prover Series | V1 → COLM 2025, V2 → AI4MATH @ ICML 2025 (Oral)

- **PI:** Chi Jin, Sanjeev Arora, Danqi Chen, Yejin Choi.
- **Background:** LLM-based automated math theorem proving with formal language.
- **Methods:** 1. *Verifier-guided self-correction* where model learns to correct its own answer based on compiler feedback, 2. *Scaffolded Learning* that synthesizes appropriately difficult questions to provide better learning signals, and 3. *Model Averaging* that boosts model's output diversity and enables multi-stage continue-training.
- **Results:** Our 8B model outperforms the 80× bigger DeepSeek-Prover-V2-671B; our 32B model achieved and has since maintained as open-source SOTA, solving 3 IMO/USAMO and 39 Putnam problems that have never been solved with LEAN.

### Adapt wile Learning for Scientific Agents | ICML 2025, AAAI FSS 2024 (Oral)

- **PI:** Rose Yu, Taylor Berg-Kirkpatrick, Leon Bergen, Duncan Watson-Parris.
- **Background:** Current LLM-based Agents can't learn from tools and over-rely on tools.
- **Methods:** 1. *World Knowledge Learning:* LLMs internalize scientific knowledge by learning from solutions generated with the assistance of tools, and 2. *Tool Usage Adaptation:* Train LLMs to maintain direct reasoning for easy problems while resorting to tools adaptively for hard ones.
- **Results:** Our 8B models achieve 29.11% higher answer accuracy and 12.72% better tool usage accuracy in 6 scenarios, surpassing GPT-4o and Claude-3.5-Sonnet.

## AWARDS

- |   |         |
|---|---------|
| • <b>National Scholarship</b> (Top 0.4% nation-wide)                                    | 2025.10 |
| • <b>Comprehensive Excellence Award</b> of Tsinghua University                          | 2025.10 |
| • <b>Spark Scientific and Technological Innovation Fellowship</b> (Top 1% in Tsinghua)  | 2024.05 |
| • <b>Scientific and Technological Innovation Excellence Scholarship</b>                 | 2024.11 |
| • <b>No. 1/636</b> , Baidu Inc. Data Mining Competition (Loan Default Prediction)       | 2024.06 |
| • <b>First Prize</b> , National College Student Mathematical Modeling Contest (Beijing) | 2023.10 |
| • <b>Second Place &amp; Newcomer Prize</b> , Tsinghua University's Challenge Cup        | 2024.04 |
| • <b>Best Paper of Popularity</b> , Contest of Scientific Communication in Tsinghua     | 2024.05 |
| • <b>Academic Advancement Program</b> , Excellent Program                               | 2024.11 |

## ACADEMIC SERVICES

**Vice President** of *the Student Association for Science and Technology*, SEM, Tsinghua University.

**Reviewer:** ICLR 2026, ARR Feb./May/July 2025, ICLR 2025, AI4MATH @ ICML 2024, LLMAgents @ ICLR 2024.

**Volunteer:** ICML 2025, EMNLP 2025, NeurIPS 2025.

## SKILLS

**Languages:** English (TOEFL 107), Chinese.

**Programming:** Python, C/C++, System Verilog, LEAN4.

**Tools:** Ray, DeepSpeed, vllm, PyTorch, Docker, etc.