XUHAN HUANG

xuhanhuang@link.cuhk.edu.cn | https://xuhan-huang.github.io/

EDUCATION

The Chinese University of Hong Kong, Shenzhen *Sep 2022 - present*
Applied Mathematics *(Expected 2025)*
CGPA: 3.88/4.00, rank 5/219 in School of Science and Engineering, rank 4/77 in Math

The Chinese University of Hong Kong *Jan 2025 - May 2025*
Visiting Student

PUBLICATIONS

*Equal contribution

[1] **LLMs for Mathematical Modeling: Towards Bridging the Gap between Natural and Mathematical Languages**
Xuhan Huang, Qingning Shen, Yan Hu, Anningzhe Gao, Benyou Wang
Findings of the Association for Computational Linguistics (NAACL), 2025. [paper] [code]

[2] **Federated Linear Dueling Bandits**
Xuhan Huang, Yan Hu, Zhiyan Li, Zhiyong Wang, Benyou Wang, Zhongxiang Dai
AAAI Conference on Artificial Intelligence (AAAI), 2026. [paper]

[3] **Re:Form—Reducing Human Priors in Scalable Formal Software Verification with RL in LLMs**
Chuanhao Yan*, Fengdi Che*, **Xuhan Huang***, Xu Xu*, Xin Li*, Yizhi Li*, Xingwei Qu*, Jingzhe Shi, Chenghua Lin, Yaodong Yang, Binhang Yuan, Hang Zhao, Yu Qiao, Bowen Zhou, Jie Fu.
Preprint, 2025. [paper] [code]

[4] **Differentiable Evolutionary Reinforcement Learning**
Sitao Cheng*, Tianle Li*, **Xuhan Huang***, Xunjian Yin, Difan Zou.
Preprint, 2025. [paper] [code]

[5] **VeriEquivBench: An Equivalence Score for Ground-Truth-Free Evaluation of Formally Verifiable Code**
Lingfei Zeng*, Fengdi Che*, **Xuhan Huang**, Fei Ye, Xu Xu, Binhang Yuan, Jie Fu
Preprint, 2025. [paper] [code]

[6] **CALM Before the STORM: Unlocking Native Reasoning for Optimization Modeling**
Zhengyang Tang*, Zihan Ye*, Chenyu Huang*, **Xuhan Huang**, Chengpeng Li, Sihang Li, Guanhua Chen, Ming Yan, Zizhuo Wang, Hongyuan Zha, Dayiheng Liu, Benyou Wang.
Preprint, 2025. [paper]

RESEARCH EXPERIENCE

Shanghai AI Laboratory *Shanghai, China*
*Research Intern. Advised by **Prof. Jie Fu*** *May 2025 – Present*

Topic: Large Language Models for Formal Reasoning (Training & Evaluation)

- **ReForm – RL for Formal Verifiable Code Synthesis**
  - **Method:** Designed a scalable data curation pipeline and a Reinforcement Learning (RL) framework integrating feedback from the Dafny verifier. Proposed the *Spec Superiority Rate* to mitigate reward hacking and a *Diversity Score* to enhance specification diversity during post-training.
  - **Outcome:** Developed Re:Form, enabling a 0.5B model to surpass proprietary baselines. The RL-trained model achieved SOTA on *DafnyComp* and exhibited strong generalization on out-of-domain tasks.
- **VeriEquivBench – Ground-Truth-Free Evaluation**
  - **Method:** Identified reliability bottlenecks in existing ground-truth matching benchmarks; constructed VeriEquivBench of 2,389 complex algorithmic problems and proposed the *Equivalence Score*, a formally grounded metric ensuring zero false positives in evaluation.
  - **Outcome:** Demonstrated that SOTA LLMs struggle with rigorous formal verification, establishing a new standard for evaluating reliable coding agents.
- **Self-Play for Data-Efficient Reasoning (Ongoing)**
  - **Method: Leading the development** of a self-evolution framework to address data scarcity in formal reasoning. Designed a dual-model "Conjecturer-Solver" system that autonomously generates a progressive training curriculum for the *code-to-specification* task.

– **Progress:** Architected and implemented the complete iterative training loop within the `veRL` framework; currently optimizing the curriculum generation strategy.

The Chinese University of Hong Kong, Shenzhen *Shenzhen, China*
*Research Assistant. Advised by **Prof. Benyou Wang** & **Prof. Zhongxiang Dai*** *Dec 2023 – Present*

Topic 1: Theoretical Machine Learning (Bandit Algorithms)

- **Federated Linear Dueling Bandits**
  - **Method:** Proposed the `FLDB-OGD` algorithm for the federated linear dueling bandit setting, utilizing Online Gradient Descent to reduce communication rounds while maintaining sub-linear regret. Led the complete theoretical analysis and **solely** implemented the simulation environment.
  - **Outcome:** Accepted to **AAAI 2026**.

Topic 2: Large Language Models for Mathematical & Optimization Modeling

- **MAMO – Benchmark for Mathematical Modeling**
  - **Method: Independently** developed MAMO, a comprehensive benchmark evaluating the translation of natural language into formal mathematical models. Designed an automatic evaluation pipeline that leverages external solvers to execute the generated formal models for rigorous correctness verification.
  - **Outcome:** Published in **NAACL 2025 (Findings)**; work was featured by *MIT Technology Review China*.
- **DERL – Differentiable Evolutionary Reinforcement Learning**
  - **Method:** Proposed DERL, a bi-level framework for automated reward discovery that optimizes a Meta-Optimizer via reinforcement learning to generate structured reward functions, effectively approximating the meta-gradient of task success.
  - **Outcome:** Achieved SOTA performance on ALFWorld and ScienceWorld benchmarks, significantly outperforming heuristic baselines in out-of-distribution scenarios and enabling autonomous agent self-improvement.
- **CALM – Optimization Modeling & Reasoning**
  - **Method:** Leveraged expert-guided intervention to identify reasoning flaws and curated corrected trajectories for Supervised Fine-Tuning; further optimized the model using RL to enhance model formulation.
  - **Outcome:** Developed STORM, a 4B-parameter model achieving 68.9% accuracy across five optimization benchmarks, matching the performance of the 671B DeepSeek-R1 model.

TEACHING EXPERIENCE

Teaching Assistant for Honours Ordinary Differential Equations
School of Science and Engineering, CUHK(SZ) *Sep 2024 – Dec 2024*
Work: Delivered weekly lectures on ODE exercises (1 hour/week) and held office hours (1 hour/week).

Teaching Assistant for Honours Calculus
School of Science and Engineering, CUHK(SZ) *Jan 2024 – May 2024*
Work: Delivered lectures on Calculus and Analysis (2 hours/week) and held office hours (1 hour/week).

Student Seminar Speaker *Summer 2023, Summer 2025*
Work: Delivered technical lectures for undergraduate students on advanced topics:

- **Reinforcement Learning (2025):** Covered PPO derivation and the `veRL` training framework.
- **Real Analysis (2023):** Covered continuity, limits, and differentiation in Euclidean space ($\mathbb{R}^N$).

AWARDS AND SCHOLARSHIPS

Undergraduate Research Award *2024, 2025*

Bowen Scholarship and Half Tuition Scholarship *2022 - 2025*

The Academic Performance Scholarship (Class C, 4% in school) for two academic years *2023 - 2025*

The Academic Performance Scholarship (Class B, 2% in school) *2022 - 2023*

Achieved first prize in the first round of the *Mathematical Olympiad Competition* for two academic years *2020, 2021*

Achievements in *National Olympiad in Informatics*
Won first prize in the Second Round of the China Standard Software (CSP) Competition 2020, Senior Group *2020*
Won second prize in the National Olympiad in Informatics in Provinces (NOIP) 2020 *2020*

SKILLS

- **Programming & Tools:** Python (PyTorch, NumPy, Pandas, `veRL`), C++ (Algorithms), LaTeX, Git.
- **Mathematical Expertise:** Real Analysis, Optimization, Advanced Linear Algebra, ODE/PDE.