

Xinye Li

leeasdf0123@gmail.com | asdf0123.github.io

EDUCATION

Harbin Institute of Technology

Aug. 2022 – Jun. 2026

B. Eng in Software Engineering

Weihai, China

- GPA: 3.77/4.0
- **National Scholarship**, 2023 (Top 0.2 % Nationwide) | Taihu Future Science and Technology Scholarship, 2023 (Top 1.6 %) | First Grade Scholarship, 2022 (Top 3 %)
- **“Excellent Student Model”** Honorary Title, 2024 (Top 0.5%)

PUBLICATIONS

Xinye Li, Zunwen Zheng, Qian Zhang, Dekai Zhuang, Jiabao Kang, Liyan Xu, Qingbin Liu, Xi Chen, Zhiying Tu, Dianhui Chu, Dianbo Sui. ScEdit: Script-based Assessment of Knowledge Editing. In *Findings of the Association for Computational Linguistics: ACL 2025*.

Jiabao Kang, **Xinye Li**, Liyan Xu, Qingbin Liu, Xi Chen, Zhiying Tu, Dianhui Chu, Dianbo Sui. Exploring Deductive and Inductive Reasoning Capabilities of Large Language Models in Procedural Planning. Accepted by *EMNLP 2025 Findings*.

Zecheng Wang, **Xinye Li**, Zhanyue Qin, Chunshan Li, Zhiying Tu, Dianhui Chu, Dianbo Sui. Can We Debias Multimodal Large Language Models via Model Editing? In *Proceedings of ACMML 2024*.

Xinye Li, Mingqi Wan, Dianbo Sui. LLMs@XLLM25: An Empirical Study of LLM for Structural Reasoning. In *Proceedings of ACL 2025 Workshop XLLM Shared Task*.

Qian Zhang, **Xinye Li**, Zhanyue Qin, Qingbin Liu, Junxian Cai, Xi Chen, Bolin Zhang, Zhiying Tu, Dianhui Chu, Xiaoyan Yu, Dianbo Sui. Editing the Moving World: Model Editing for Video LLMs. Under Review.

RESEARCH EXPERIENCE

Research Assistant in HKUST(GZ) HPML Lab

Apr. 2025 - Present

GUI Agent, Knowledge

HKUST (GZ), RA

- Conducting research on GUI Agents, focusing on GUI grounding and challenging dataset construction.
- Assisting in research efforts related to Knowledge.

Research on Knowledge Editing

Sep. 2024 - Feb. 2025

Knowledge Editing, Script, NLG

Accepted at ACL 2025, First Author

- Designed a script-based assessment framework to evaluate the performance of Knowledge Editing in complex real-world reasoning and generation tasks.
- Developed a rigorous benchmark accompanied by comprehensive experiments on existing methods, along with in-depth analyses of metric correlations and discussions on future directions for advancement.

Research on Inductive & Deductive Reasoning

Sep. 2024 - Dec. 2024

Inductive & Deductive Reasoning, Procedural Planning

Under Review at EMNLP 2025, Second Author

- Investigated the deductive and inductive capabilities of LLMs in procedural planning.
- Introduced an effective method with multiple sampling to enhance the inductive reasoning capabilities of LLMs in procedural planning.

Debiasing Multimodal LLMs via Model Editing

Jan. 2024 – Jun. 2024

MLLM, LLM Debiasing, Model Editing

Accepted at ACMML 2024, Second Author

- Introduced a novel benchmark for debiasing editing in MLLM, evaluating the reliability, locality, and generality of model editing based debiasing methods across IC and VQA tasks.
- Conducted comprehensive research on the application of model editing methods for debiasing, involving editing 2 modules (LLM and Vision) and 4 types of biases.

ACHIEVEMENTS

CP (Competitive Programming)

Aug. 2022 – Dec. 2024

Algorithm Contest player (OIer/ACMer)

Contestant

- **Bronze Medal**, The 2024 ACM-ICPC Asia East Continent Final Contest (**EC-Final**)
- **Silver Medal**, The 2023 ACM-ICPC Asia Hangzhou Regional Contest
- **Bronze Medal**, The 2023 ACM-ICPC Asia Nanjing Regional Contest
- **Gold Medal**, The 2024 Weihai Collegiate Programming Contest
- **Bronze Medal**, The 2023 CCF Collegiate Computer Systems & Programming Contest

Other Contests

- **Second Prize**, The 19th “Challenge Cup”: 2024 ‘Open Call for Solutions’ Special Program (Team Leader)
- **Second Prize**, Huawei CodeCraft Contest 2024 (**9th** in the preliminary round)
- **Second Prize**, The 6th Global Campus AI Algorithm Elite Competition

PROJECTS

Game Arena for Evaluating Sequential Reasoning of LLMs

Nov. 2023 - Feb. 2024

LLM Evaluation, Exhaustive Algorithm

Student Researcher

- A system based on card game evaluating LLM’s sequential reasoning & decision-making ability; An exhaustive algorithm was simultaneously proposed to record the game data in real time.

Reproduction of R1 for Mathematical Reasoning

Mar. 2025 – Apr. 2025

Mathematical Reasoning, GRPO, vLLM, RL

Student Researcher

- Reproduced the R1 project to enhance mathematical reasoning of LLMs. Utilized an 8x 4090 setup for training with GRPO and vLLM for inference, achieving competitive scores on GSM8K (65.7) and Math500 (48.0).

Activating Puzzle Reasoning in Qwen2.5-VL 7B via SFT

May. 2025 – May. 2025

Multimodal LLMs, SFT, Knowledge Distillation, Visual Reasoning, LLaMA Factory

Competitor

- Activated puzzle reasoning in the Qwen2.5-VL 7B model via a pipeline of data augmentation, knowledge distillation, and supervised fine-tuning (SFT). Significantly improved performance on visual puzzles, boosting in-domain accuracy from 0.21 to 0.7.

SKILLS

Languages: English (CET4: 619, CET6: 577), Chinese, Cantonese (learning)

Programming Languages: C++, C, Python, Java, Tex, SQL

Frameworks: PyTorch, Sklearn, Numpy, Pandas, Matplotlib

Developer Tools: Git, Google Cloud Platform, Huggingface, Github, Jupyter Notebook, Vs Code, Postman

WORKING

HITWH ACM Club

Aug. 2023 - Present

Student Club

Club leader & Teaching Assistant

- Teaching assistant in Prof. Kaikun Dong’s course *Problem-Oriented Advanced Programming*
- Organizer and Problem Contributor of the ACM Freshman Contest