

Yiliu Yang

(765) 421-1120 | yyiliu@cmu.edu | yyiliu.com | linkedin.com/in/yyiliu/

EDUCATION

Carnegie Mellon University

Master of Science in Intelligent Information Systems, School of Computer Science

Pittsburgh, PA

Dec 2027

The Chinese University of Hong Kong

Bachelor of Engineering in Artificial Intelligence

Hong Kong

Jul 2024

GPA: 3.82/4.00, Major GPA: 3.92/4.00, Rank: 1st of 39

SKILLS

Programming: Python, C, C++, SQL

LLM/ML: LLM Agents, PyTorch, RAG, vLLM, Transformers

Engineering: Linux, Git, Docker, Redis, FastAPI

Languages: English (Fluent), Mandarin (Native)

PROFESSIONAL EXPERIENCE

RedotPay

Hong Kong

AI Engineer

Mar 2026 – Aug 2026

- Co-developed and productionized an AI customer-service agent serving approximately 3,500 daily conversations with knowledge-grounded responses, MCP tools, SOP workflows, conversation analytics, and human handoff
- Increased CSAT from 71.6% to 84.0%, and deflection rate from 70.6% to 81.9% versus the legacy Dify bot, saving an estimated 50 support hours per day
- Designed and built a 909-case automated regression framework for tool-using workflows, validating required and prohibited MCP calls with mocked backend states across KYC and SOP scenarios

RESEARCH EXPERIENCE

Multimedia Laboratory, The Chinese University of Hong Kong

Hong Kong

Research Assistant

Mar 2025 – Feb 2026

- Developed *QuadSentinel*, a multi-agent guard system translating natural language policies into machine-checkable rules for robust runtime safety enforcement
- Designed stateful monitoring using top-k semantic predicate retrieval and logical policy verification, reducing runtime overhead by 47% versus the most efficient evaluated method
- Achieved 91.5% accuracy on the AgentHarm safety benchmark across prompt injection, jailbreak, and tool misuse attacks, outperforming prior state-of-the-art methods

ARISE Laboratory, The Chinese University of Hong Kong

Hong Kong

Undergraduate Research Assistant

Feb 2023 – Jun 2024

- Built LogicAsker, a Python framework for rule-based test generation and evaluation, benchmarking 6 LLMs on 5,200 reasoning cases and identifying weaknesses using logical rules
- Implemented in-context learning and fine-tuning experiments using synthetic reasoning data, increasing GPT-4o accuracy on LogicAsker from 91.9% to 97.2% through targeted in-context demonstrations

PUBLICATIONS

Yiliu Yang*, Yilei Jiang*, Qunzhong Wang, et al., *QuadSentinel: Sequent Safety for Machine-Checkable Control in Multi-agent Systems*. arXiv preprint arXiv:2512.16279, 2025

Yuxuan Wan*, Wenxuan Wang*, Yiliu Yang, et al., *LogicAsker: Evaluating and Improving the Logical Reasoning Ability of Large Language Models*. EMNLP 2024 (Main Conference), pp. 2124–2155

LEADERSHIP

International Collegiate Programming Contest (ICPC)

Oct 2020 – Aug 2023

- Led a three-member competitive programming team at CUHK to four regional bronze medals