Wenjie Qiu

echo-inception@live.com +1-314-203-5508 Piscataway, NJ 08854, United States

floatingsong.com Scholar LinkedIn GitHub

Research Interests

My research interests involve reinforcement learning (**RL**), preference learning (**PL**), and large language models (**LLMs**). I am interested in programmatic RL for building interpretable policies. Additionally, I explore the application of RL algorithms across various complex, real-world domains, including Embodied AI and Agentic problems.

I am interested in how LLMs can be leveraged to enhance RL training (**LLMs for RL**). Inspired by recent advancements in preference-based RL algorithms (e.g., **RLHF, DPO, CPL**) that have demonstrated significant success in LLMs, I am working on constructing contrastive data from unlabeled demonstrations.

Education

- **Rutgers University** *Sep 2020 - Jan 2026*
  Ph.D. in Computer Science New Brunswick, NJ, USA
  - Advisor: *Prof. He Zhu*
  - Dissertation: *From Reasoning to Acting: A Program Synthesis Approach to Reinforcement Learning* [pdf]
  - Research Interests: Reinforcement Learning, Large Language Models, Program Synthesis, and Preference Learning
- **Washington University in St. Louis** *Aug 2018 - May 2020*
  M.S. in Computer Science St. Louis, MO, USA
  - Advisor: *Prof. Ning Zhang*
  - Thesis: *A Performance Analysis of Hardware-assisted Security Technologies* [pdf]
  - Research Interests: System Security, Software Security
- **Central China Normal University** *Sep 2014 - Jun 2018*
  B.E. in Electronic Information Engineering Wuhan, Hubei, China
  - Coursework: Math, Physics, Communication Engineering, Computer Science, and English Literature

Professional Experience

- **Amazon** *Feb 2026 - Now*
  Applied Scientist (Ads) New York, NY, USA
  - Agent Platform and Sequential Modeling
- **Amazon** *Sep 2025 - Dec 2025*
  Applied Scientist Intern (Finance Technology) Bellevue, WA, USA
  - Post-training on Large Language Model with focus on long-context finance tabular question-answering
  - Applied RL algorithms (GSPO, GRPO, DAPO) on Qwen3-4B-Thinking, improved the benchmarking results by 6%
- **LinkedIn** *May 2025 - Aug 2025*
  PhD AI/ML Intern (Core AI) Mountain View, CA, USA
  - Implemented recruiter simulator for LinkedIn Hiring Agent (LiHA), synthesizing high-quality conversational data
  - Proposed evaluation metrics for judging LiHA conversations, identified a design flaw of the Agent system
- **Samsung Research America** *May 2022 - Aug 2022*
  Research Intern (AI Center) Mountain View, CA, USA
  - Reproduced baselines of BEHAVIOR benchmark, a complex set of EAI problems running in ai2thor
  - Implemented programmatically interpretable controllers for robots in iGibson simulation environment
  - Utilized programmatic controllers for complex object detection, robot navigation and manipulation problems
- **Stevens Institute of Technology** *Jun 2019 - Aug 2019*
  Research Intern (S3 Lab) Hoboken, NJ, USA
  - Built a Docker-based microservice backend for Automatic Verification of Temporal Alignment
  - Exploited the disassembling algorithms and workflow of Radare2 – a reverse-engineering framework

- **Rutgers University** *Jan 2021 - May 2025*
  Research Assistant (Rutgers Automated Reasoning Lab) New Brunswick, NJ, USA
  - Research with focus on Reinforcement Learning, Program Synthesis, Neurosymbolic Programming, Embodied AI, Preference Learning, World Models, and Large Language Models
- **Rutgers University** *Sep 2020 - Dec 2020*
  Teaching Assistant (Department of Computer Science) New Brunswick, NJ, USA
  - Computer Architecture (CS 211)
- **Washington University in St. Louis** *Jan 2019 - May 2020*
  Research Assistant (Computer Security & Privacy Laboratory) St. Louis, MO, USA
  - Research on program behaviors in a trusted execution environment
  - Research on emulation of and fuzzing on embedded operating systems

PUBLICATIONS (*equal contribution)

Conference Papers:

[ICLR'26] Preference-based Policy Optimization from Sparse-reward Offline Dataset [pdf]
Wenjie Qiu, Guofeng Cui, Shicheng Liu, Yuanlin Duan, and He Zhu

[NeurIPS'25] Explainable Reinforcement Learning from Human Feedback to Improve Alignment [pdf]
Shicheng Liu, Siyuan Xu, **Wenjie Qiu**, Hangfan Zhang, and Minghui Zhu

[NeurIPS'25] Learning from Demonstrations via Capability-Aware Goal Sampling [pdf] [code]
Yuanlin Duan, Yuning Wang, **Wenjie Qiu**, and He Zhu

[PLDI'24] Program Synthesis of Intelligent Agents from Rewards [pdf] [code]
Guofeng Cui*, Yuning Wang*, **Wenjie Qiu***, and He Zhu

[NeurIPS'23] Instructing Goal-Conditioned Reinforcement Learning Agents with Temporal Logic Objectives [pdf] [code]
Wenjie Qiu*, Wensen Mao*, and He Zhu

[ICLR'22] Programmatic Reinforcement Learning without Oracles **[Spotlight]** [pdf] [code]
Wenjie Qiu and He Zhu

Recent Submissions (Under Review):

- FinLCR: Scaling Financial Long-Context Reasoning via Curriculum Reinforcement Learning (ACL)
- Compositional Policy Optimization with Language Models (AAAI)

HONORS & AWARDS

- **Spotlight Presentation** *2022*
  International Conference on Learning Representations Top 3%
- **Boya Scholarship** *2016*
  Central China Normal University Top 3%

ACADEMIC SERVICE

- **Conference Reviewer:** NeurIPS 2024-2025, ICLR 2025-2026, ICML 2026, AISTATS 2025-2026, NeuS 2025
- **Journal Reviewer:** TIST, TNNLS

TECHNICAL SKILLS

- **ML Frameworks:** PyTorch, NumPy, FAISS, ai2thor
- **RL Frameworks:** OpenAI spinningup, OpenAI gym/gymnasium, stable-baselines3, MuJoCo
- **LLM Frameworks:** Hugging Face Transformers, lighteval, verl, vLLM
- **Tools:** Git, Docker, Kubernetes, Ray, CUDA, Weights & Biases, AWS, LaTeX
- **Programming Languages:** Python, C, C++, MATLAB, Bash, SQL
- **Natural Languages:** English (Fluent), Chinese (Native)