SHENAO ZHANG

shenao@u.northwestern.edu | shenao-zhang.github.io

EDUCATION

Northwestern University *Sep. 2023 - Present*
Ph.D. student in IEMS (Industrial Engineering & Management Sciences) *Evanston, IL*
Advisor: Prof. Zhaoran Wang

Georgia Institute of Technology *May 2020 - May. 2022*
M.S. in ECE (Electrical and Computer Engineering), GPA: 3.81/4.00 *Atlanta, GA*
Advisor: Prof. Tuo Zhao and Prof. Bo Dai

South China University of Technology *Aug. 2016 - May 2020*
B.Eng. in EE (Electronic and Information Engineering, Innovation Class) *Guangzhou, China*

University of California, Berkeley *Jan. 2019 - May 2019*
Visiting student at the Department of EECS, GPA: 3.90/4.00 *Berkeley, CA*

RESEARCH INTERESTS

My research centers around Large Language Models (LLMs) and Reinforcement Learning (RL). I'm currently interested in two key areas: aligning LLMs efficiently by discovering valuable data with minimal cost, and enhancing LLMs and language agents with advanced reasoning capabilities. The ultimate goal of my research is to build systems that self-improve by actively synthesizing data and learning to reason to achieve super-human intelligence. Previously, I developed data-efficient decision-making algorithms with applications to robotic and multi-agent systems.

PREPRINTS

[15] Han Zhong*, Yutong Yin*, **Shenao Zhang**, Xiaojun Xu, Yuanxin Liu, Yifei Zuo, Zhihan Liu, Boyi Liu, Sirui Zheng, Hongyi Guo, Liwei Wang, Mingyi Hong, Zhaoran Wang, "BRiTE: Bootstrapping Reinforced Thinking Process to Enhance Language Model Reasoning", *Preprint*. [PDF]

[14] Huaijie Wang*, Shibo Hao*, Hanze Dong, **Shenao Zhang**, Yilin Bao, Ziran Yang, Yi Wu, "Offline Reinforcement Learning for LLM Multi-Step Reasoning", *ICLR Workshop on Reasoning and Planning for LLMs*, 2025. [PDF]

[13] Zhihan Liu, **Shenao Zhang**, Yongfei Liu, Boyi Liu, Yingxiang Yang, Zhaoran Wang, "DSTC: Direct Preference Learning with Only Self-Generated Tests and Code to Improve Code LMs", *Preprint*. [PDF]

[12] **Shenao Zhang***, Zhihan Liu*, Boyi Liu, Yufeng Zhang, Yingxiang Yang, Liyu Chen, Tao Sun, Zhaoran Wang, "Reward-Augmented Data Enhances Direct Preference Alignment of LLMs", *ICLR DATA-FM Workshop*, 2025. [PDF]

[11] **Shenao Zhang***, Sirui Zheng*, Shuqi Ke, Zhihan Liu, Wanxin Jin, Jianbo Yuan, Yingxiang Yang, Hongxia Yang, Zhaoran Wang, "How Can LLM Guide RL? A Value-Based Approach", *Preprint*. [PDF]

PROCEEDINGS

[10] **Shenao Zhang**, Donghan Yu, Hiteshi Sharma, Ziyi Yang, Shuohang Wang, Hany Hassan, Zhaoran Wang, "Self-Exploring Language Models: Active Preference Elicitation for Online Alignment", *Transactions on Machine Learning Research (TMLR)*, *ICML AutoRL Workshop* (Best Paper Award), 2024. [PDF]

[9] Zhihan Liu*, Miao Lu*, **Shenao Zhang**, Boyi Liu, Hongyi Guo, Yingxiang Yang, Jose Blanchet, Zhaoran Wang, "Provably Mitigating Overoptimization in RLHF: Your SFT Loss is Implicitly an Adversarial Regularizer", *Neural Information Processing Systems (NeurIPS)*, 2024. [PDF]

[8] Zhihan Liu*, Hao Hu*, **Shenao Zhang***, Hongyi Guo, Shuqi Ke, Boyi Liu, Zhaoran Wang, "Reason for Future, Act for Now: A Principled Framework for Autonomous LLM Agents with Provable Sample

Efficiency", *International Conference on Machine Learning (ICML)*, 2024. [PDF]

[7] Feng Gao*, Liangzhi Shi*, **Shenao Zhang**, Zhaoran Wang, Yi Wu, "Adaptive-Gradient Policy Optimization: Enhancing Policy Learning in Non-Smooth Differentiable Simulations", *International Conference on Machine Learning (ICML)*, 2024. [PDF]

[6] **Shenao Zhang**, Boyi Liu, Zhaoran Wang†, Tuo Zhao†, "Model-Based Reparameterization Policy Gradient: Theory and Practical Algorithms", *Neural Information Processing Systems (NeurIPS)*, 2023. [PDF].

[5] Zhihan Liu*, Miao Lu*, Wei Xiong*, Han Zhong, Hao Hu, **Shenao Zhang**, Sirui Zheng, Zhuoran Yang, Zhaoran Wang, "Maximize to Explore: One Objective Function Fusing Estimation, Planning, and Exploration", *Neural Information Processing Systems (NeurIPS)* (Spotlight), 2023. [PDF].

[4] **Shenao Zhang**, Wanxin Jin, Zhaoran Wang, "Adaptive Barrier Smoothing for First-Order Policy Gradient with Contact Dynamics", *International Conference on Machine Learning (ICML)*, 2023. [PDF]

[3] **Shenao Zhang**, "Conservative Dual Policy Optimization for Efficient Model-Based Reinforcement Learning", *Neural Information Processing Systems (NeurIPS)*, 2022. [PDF].

[2] **Shenao Zhang**, Li Shen, Lei Han, Li Shen, "Learning Meta Representation for Agents in Multi-Agent Reinforcement Learning", *Conference on Lifelong Learning Agents (CoLLAs)* (Oral), 2023. [PDF]

[1] **Shenao Zhang**, Li Shen, Zhifeng Li, Wei Liu, "Structure-Regularized Attention for Deformable Object Representation", *NeurIPS Workshop on Object Representations for Learning and Reasoning*, 2020. [PDF]

INTERNSHIP EXPERIENCE

Google *Dec. 2024 - Now*
Student Researcher
- Worked on why, when, and how LLMs should deliberately reason, e.g., reflect, from an adaptive RL view.

Microsoft GenAI *Jan. 2024 - June 2024*
Student Researcher *Advisor: Donghan Yu*
- Worked on active preference elicitation for online alignment [10].

ByteDance Seed *June 2024 - Sep. 2024*
Research Intern *June 2023 - Aug. 2023*
- Worked on RL with LLM policy prior [11] and reward-conditioned augmentation for alignment [12].

Microsoft Research (MSR), Asia *Feb. 2023 - May 2023*
Research Intern *Advisor: Li Zhao*
- Worked on autonomous LLM agents that actively gather information [15].

Tencent AI Lab *Aug. 2019 - Sep. 2020*
Research Intern *Advisors: Li Shen, Lei Han and Li Shen*
- Worked on visual attention representation [1] and multi-agent RL [2].

TEACHING EXPERIENCE

Head TA of the graduate course CS 7648: Interactive Robot Learning (Fall 2021) at Georgia Tech.

PROFESSIONAL SERVICE

Conference Review: NeurIPS 20-24, ICLR 22-25, AISTATS 22-25, ICML 22-25, COLM 24-25.
Journal Review: Neurocomputing, TPAMI, TMLR.

HONORS AND AWARDS

Meshy Fellowship Finalist	*2025*
NeurIPS Top Reviewer	*2024*
NeurIPS Scholar Award	*2022-2023*
ICML Travel Award	*2023*
Georgia Tech Level A Premier Merit-Based Scholarship	*2020-2021*
Outstanding Freshman Scholarship (Awarded to 30 among 6,500 students)	*2016*