

# SHENAO ZHANG

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

## EDUCATION

### Northwestern University

Ph.D. student in IEMS (Industrial Engineering & Management Sciences)  
Advisor: Prof. Zhaoran Wang

Sep. 2023 - Present

Evanston, IL

### Georgia Institute of Technology

M.S. in ECE (Electrical and Computer Engineering), GPA: 3.81/4.00  
Advisor: Prof. Tuo Zhao and Prof. Bo Dai

May 2020 - May. 2022

Atlanta, GA

### South China University of Technology

B.Eng. in EE (Electronic and Information Engineering, Innovation Class)

Aug. 2016 - May 2020

Guangzhou, China

### University of California, Berkeley

Visiting student at the Department of EECS, GPA: 3.90/4.00

Jan. 2019 - May 2019

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] Huaijie Wang\*, Shibo Hao\*, Hanze Dong, **Shenao Zhang**, Yilin Bao, Ziran Yang, Yi Wu, "Offline Reinforcement Learning for LLM Multi-Step Reasoning", *Preprint*. [\[PDF\]](#)
- [14] 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*. [\[PDF\]](#)
- [13] **Shenao Zhang**, Zhihan Liu, Boyi Liu, Yufeng Zhang, Yingxiang Yang, Yongfei Liu, Liyu Chen, Tao Sun, Zhaoran Wang, "Reward-Augmented Data Enhances Direct Preference Alignment of LLMs", *Preprint*. [\[PDF\]](#)
- [12] **Shenao Zhang**, Donghan Yu, Hiteshi Sharma, Ziyi Yang, Shuohang Wang, Hany Hassan, Zhaoran Wang, "Self-Exploring Language Models: Active Preference Elicitation for Online Alignment", *ICML 2024 AutoRL Workshop* (**Best Paper Award**). [\[PDF\]](#)
- [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*. [\[PDF\]](#)
- [10] Xiaoyu Chen, **Shenao Zhang**, Pushi Zhang, Li Zhao, Jianyu Chen, "Asking Before Action: Gather Information in Embodied Decision Making with Language Models", *Preprint*. [\[PDF\]](#)

## PROCEEDINGS

- [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<sup>†</sup>, Tuo Zhao<sup>†</sup>, “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

---

- |                                                                                                                                    |                                                                               |
|------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| <b>Google</b><br><i>Student Researcher</i>                                                                                         | <i>Dec. 2024 - Now</i>                                                        |
| • Worked on LLMs that reason by adapting with reward model signals to reflect and backtrack.                                       |                                                                               |
| <b>Microsoft GenAI</b><br><i>Student Researcher</i>                                                                                | <i>Jan. 2024 - June 2024</i><br><i>Advisor: Donghan Yu</i>                    |
| • Worked on active preference elicitation for online alignment <a href="#">[12]</a> .                                              |                                                                               |
| <b>ByteDance Seed</b><br><i>Research Intern</i>                                                                                    | <i>June 2024 - Sep. 2024</i><br><i>June 2023 - Aug. 2023</i>                  |
| • Worked on RL with LLM policy prior <a href="#">[11]</a> and reward-conditioned augmentation for alignment <a href="#">[13]</a> . |                                                                               |
| <b>Microsoft Research (MSR), Asia</b><br><i>Research Intern</i>                                                                    | <i>Feb. 2023 - May 2023</i><br><i>Advisor: Li Zhao</i>                        |
| • Worked on autonomous LLM agents that actively gather information <a href="#">[10]</a> .                                          |                                                                               |
| <b>Tencent AI Lab</b><br><i>Research Intern</i>                                                                                    | <i>Aug. 2019 - Sep. 2020</i><br><i>Advisors: Li Shen, Lei Han and Li Shen</i> |
| • Worked on visual attention representation <a href="#">[1]</a> and multi-agent RL <a href="#">[2]</a> .                           |                                                                               |

## TEACHING EXPERIENCE

---

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

## PROFESSIONAL SERVICE

---

**Conference Review:** NeurIPS 2020-24, ICLR 2022-24, AISTATS 2022-24, ICML 2022-24, COLM 2024.  
**Journal Review:** Neurocomputing, Transactions on Pattern Analysis and Machine Intelligence (TPAMI).

## HONORS AND AWARDS

---

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