ANDREW ZHAO

Curriculum Vitae

Department of Automation, Tsinghua University
Citizenship: **Canada**
(+86) 18513245436
✉ andrewzhao112@gmail.com
⌂ *Personal Page*
🎓 *Google Scholar*   *Github*   in *Linkedin*

Research Statement

My research focuses on *reinforcement learning for LLMs*, *reasoning*, *LLM-based agents*, and the *safety* challenges associated with generative AI, with an emphasis on robustness, generalization, diversity, and self-evolving systems requiring minimal human intervention. Recently, I have become interested in the reasoning capabilities of large language models. I aim to advance the development of innovative, reliable, and autonomous AI systems capable of adapting to diverse and dynamic environments.

Education

2021–present	**PhD, Pattern Recognition and Machine Learning**, *Tsinghua University*, Beijing, China. Reinforcement Learning, Agents, Natural Language Processing, Deep Learning
2019–2020	**Master of Science**, *University of Southern California*, Los Angeles.
2012–2017	**Bachelor of Applied Science**, *University of British Columbia*, Vancouver.

Selected Publications

2025 Andrew Zhao, Erle Zhu, Rui Lu, Matthieu Lin, Yong-Jin Liu, and Gao Huang. Self-referencing agents for unsupervised reinforcement learning. *Neural Networks*, page 107448, 2025.

2025 Andrew Zhao, Quentin Xu, Matthieu Lin, Shenzhi Wang, Yong jin Liu, Zilong Zheng, and Gao Huang. Diver-ct: Diversity-enhanced red teaming large language model assistants with relaxing constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence (ORAL)*, 2025.

2025 Andrew Zhao, Yiran Wu, Yang Yue, Tong Wu, Quentin Xu, Yang Yue, Matthieu Lin, Shenzhi Wang, Qingyun Wu, Zilong Zheng, and Gao Huang. Absolute zero: Reinforced self-play reasoning with zero data, 2025.

2024 Andrew Zhao, Daniel Huang, Quentin Xu, Matthieu Lin, Y. Liu, and Gao Huang. Expel: Llm agents are experiential learners. In *Proceedings of the AAAI Conference on Artificial Intelligence (ORAL)*, 2024.

2022 Andrew Zhao, Matthieu Lin, Yangguang Li, Yong-Jin Liu, and Gao Huang. A mixture of surprises for unsupervised reinforcement learning. *Advances in Neural Information Processing Systems*, 2022.

Research Experience

Microsoft Research, Redmond

June,2024 – August,2024	**Research Intern @ Augmented Learning and Reasoning Group**. Worked on LLM-based agent applications for cybersecurity investigations.
Advisor :	**Dr. Jack Stokes**, *Principal RSDE @ MSR Redmond*

Beijing Institute for General Artificial Intelligence (BIGAI)

Jan,2024 – May,2024	**Research Intern @ Natural Language and Conversational AI Lab**. Worked on red teaming LLMs to increase sample diversity using constrained on-policy reinforcement learning and reward shaping.
Advisor :	**Dr. Zilong Zheng**, *Research Scientist @ BIGAI* (*Personal Web-page*)

Peiking University, Beijing

January,2021 **Research Intern**.
– June,2021 Contributed to research projects in embodied AI and robotic research.

Advisor : **Professor Hao Dong**, *Assistant Professor, Department of Electronics Engineering and Computer Science*, Beijing (*Personal Web-page*)

Academic Talks

May/June,2025 ***Hugging Face X AI Insight/Amazon AGI Labs/ByteDance/BAAI/Gaoling School of AI/ERIC Lab/The Network School/Ploutos*** 8 Invited Talks for Absolute Zero Reasoner

August,2024 ***IJCAI 2024*** *AI4Research Workshop* Invited Paper Talk for – ExpeL: LLM Agents are Experiential Learners.

February,2024 ***AAAI 2024*** Oral Presentation of – ExpeL: LLM Agents are Experiential Learners.

January,2024 ***AI TIME Youth PhD Talk*** ExpeL: LLM Agents are Experiential Learners.

February,2023 ***AI TIME Youth PhD Talk*** A Mixture of Surprises for Unsupervised Reinforcement Learning.

Academic Services

2023–present **Academic Reviewer** in NeurIPS 2023/2024, ICLR 2024/2025, AAAI 2025, ICML 2024, AISTATS 2025

All Publications

2025 Andrew Zhao, Erle Zhu, Rui Lu, Matthieu Lin, Yong-Jin Liu, and Gao Huang. Self-referencing agents for unsupervised reinforcement learning. *Neural Networks*, page 107448, 2025.

2025 Andrew Zhao, Quentin Xu, Matthieu Lin, Shenzhi Wang, Yong jin Liu, Zilong Zheng, and Gao Huang. Diver-ct: Diversity-enhanced red teaming large language model assistants with relaxing constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence (ORAL)*, 2025.

2025 Andrew Zhao, Yiran Wu, Yang Yue, Tong Wu, Quentin Xu, Yang Yue, Matthieu Lin, Shenzhi Wang, Qingyun Wu, Zilong Zheng, and Gao Huang. Absolute zero: Reinforced self-play reasoning with zero data, 2025.

2025 Yang Yue, Zhiqi Chen, Rui Lu, Andrew Zhao, Zhaokai Wang, Yang Yue, Shiji Song, and Gao Huang. Does reinforcement learning really incentivize reasoning capacity in llms beyond the base model?, 2025.

2024 Andrew Zhao, Daniel Huang, Quentin Xu, Matthieu Lin, Y. Liu, and Gao Huang. Expel: Llm agents are experiential learners. In *Proceedings of the AAAI Conference on Artificial Intelligence (ORAL)*, 2024.

2024 Shenzhi Wang, Chang Liu, Zilong Zheng, Siyuan Qi, Shuo Chen, Qisen Yang, Andrew Zhao, Chaofei Wang, Shiji Song, and Gao Huang. Avalon's game of thoughts: Battle against deception through recursive contemplation. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2024*. Association for Computational Linguistics, 2024.

2024 Jenny Sheng, Matthieu Lin, Andrew Zhao, Kevin Pruvost, Yu-Hui Wen, Yangguang Li, Gao Huang, and Yong-Jin Liu. Exploring text-to-motion generation with human preference. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1888–1899, June 2024.

2024 Matthieu Lin, Yubin Hu Jenny Sheng, Yangguang Li, Andrew Zhao Lu Qi, Gao Huang, and Yong-Jin Liu. Exploring temporal feature correlation for efficient and stable video semantic segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024.

2022 Andrew Zhao, Matthieu Lin, Yangguang Li, Yong-Jin Liu, and Gao Huang. A mixture of surprises for unsupervised reinforcement learning. *Advances in Neural Information Processing Systems*, 2022.

2022   Rui Lu, Andrew Zhao, Simon S Du, and Gao Huang. Provable general function class representation learning in multitask bandits and mdp. *Advances in Neural Information Processing Systems (SPOTLIGHT)*, 2022.

2021   Haoqi Yuan, Ruihai Wu, Andrew Zhao, Hanwang Zhang, Zihan Ding, and Hao Dong. Dmotion: Robotic visuomotor control with unsupervised forward model learned from videos. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2021.