Zhaoyi Zhou

(412) 759-4397 • zhaoyiz@andrew.cmu.edu • https://zhaoyizhou1123.github.io/

EDUCATION

Carnegie Mellon University August 2024 – May 2029 (Expected)
Ph.D. in Electrical and Computer Engineering Pittsburgh, PA, USA
Advisor: Andrea Zanette

- Main research areas**:** Foundation models, reinforcement learning

Tsinghua University September 2020 - June 2024
B.E. in Computer Science and Technology (Yao Class) Beijing, China

- Selected coursework**:** Machine Learning, Theory of Computation, Game Theory, Algorithm Design
- GPA**:** 3.98/4.00

PUBLICATIONS AND PREPRINTS

- Accelerating Unbiased LLM Evaluation via Synthetic Feedback.
  Zhaoyi Zhou, Yuda Song, and Andrea Zanette.
  International Conference on Machine Learning (ICML) 2025. (Under review)
- Free from Bellman Completeness: Trajectory Stitching via Model-based Return-conditioned Supervised Learning.
  Zhaoyi Zhou, Chuning Zhu, Runlong Zhou, Qiwen Cui, Abhishek Gupta, and Simon S. Du.
  International Conference on Learning Representations (ICLR) 2024. (31% acceptance rate)
  (ORAL presentation at Foundation Models for Decision Making (FMDM) workshop of NeurIPS 2023.)
- Convergence Rates for Localized Actor-Critic in Networked Markov Potential Games.
  Zhaoyi Zhou, Zaiwei Chen, Yiheng Lin, and Adam Wierman.
  Conference on Uncertainty in Artificial Intelligence (UAI) 2023. (31% acceptance rate)

SELECTED HONORS

- Yao Award, Recognition Prize (presented by Dean Andrew C. Yao of IIIS), Tsinghua University, 2023;
- Scholarship of Academic Excellence, Tsinghua University, 2023 & 2022;
- Scholarship of Scientific Innovation, Tsinghua University, 2023 & 2022;
- Scholarship of Comprehensive Excellence, Tsinghua University, 2021.

RESEARCH EXPERIENCE

Unbiased Acceleration of LLM Evaluation via Synthetic Feedback Aug 2024 – Feb 2025
Carnegie Mellon University Pittsburgh, USA

- Design Control Variates Evaluation, an unbiased LLM evaluation method with reduced human annotation cost by incorporating synthetic preferences.
- Further improve the human annotation saving effect by finetuning synthetic evaluators.
- Propose the human annotation saving ratio as an easy-to-measure metrics to assess and predict the performance of Control Variates Evaluation.

University of Washington February 2023 - August 2023
Research Intern. Supervisors: Simon Shaolei Du & Abhishek Gupta Seattle, USA
Led research in offline reinforcement learning (RL).

- Discovered freedom from Bellman completeness requirement as an advantage of return-conditioned supervised learning (RCSL) over dynamic-programming (DP) based offline RL methods.
- Conducted both theoretical and empirical analysis on explicit examples to demonstrate strength of RCSL over DP-based methods in near-deterministic environments given expert dataset.
- Theoretically proved that RCSL cannot do trajectory stitching, a limitation of RCSL observed in previous literature.
- Developed model-based return-conditioned supervised learning (MBRCSL) framework, which enables trajectory stitching of RCSL while avoiding Bellman completeness requirements.
- Tested MBRCSL on offline RL benchmarks, in which MBRCSL outperforms state-of-the-art model-based and model-free offline RL algorithms.

California Institute of Technology February 2022 - February 2023
Research Intern. Supervisor: Adam Wierman Remote
Led research in networked multi-agent RL.

- Introduced the class of networked Markov potential games as relaxation of Markov potential games (MPG), encompassing practical examples such as Markov congestion game.
- Designed localized actor-critic algorithm with linear function approximation, overcoming the curse of dimensionality.
- Derived the first finite-sample bound for multi-agent competitive games that is independent of the number of agents.

PROFESSIONAL ACTIVITIES

Paper Reviewer: ICLR 2025, ICML 2025.

OTHERS

Software: Python, C/C++, Go, Verilog, MATLAB
Languages: English (Fluent), German (Fluent), Chinese (Native)