

Xi Chen

1578 Neil Ave ◊ Columbus, Ohio 43210

chen.10183@buckeyemail.osu.edu

cx441000319.github.io

EDUCATION

The Ohio State University

Columbus, OH, USA

Ph.D. Candidate in Computer Science and Engineering

2020 – Present

• Advisors: Prof. Andrew Perrault and Prof. Zhihui Zhu

• RESEARCH INTERESTS: **Reinforcement Learning; Large Language Model; AI Agent**

Xidian University

Xi'an, Shaanxi, China

Bachelor in Software Engineering

2016 – 2020

• GPA: 3.9/4.0; Rank: 15/445 (Top 3%)

• Class of Software Engineering for Talented (40 out of 445 students)

CURRENT RESEARCH PROJECTS

Motivation. We hypothesize that the current RLVR setting is constrained by limited exploration.

Ongoing. We are developing a new fine-tuning algorithm that augments RLVR with an SFT phase to bring in more exploration.

PUBLICATIONS

- Understanding Learned Representations and Action Collapse in Visual Reinforcement Learning

Xi Chen, Zhihui Zhu, and Andrew Perrault. *RLC 2025*.

- To understand how an RL agent with image observations represents the environment, we apply linear probing on hidden-layer activations to recover vectorized states. That not only reveals where and how the vectorized states are learned, but also which part of the model causes an action-repetition failure mode.

- To better understand the reason for the failure mode, we compute relevant model metrics and analyze them in the context of the algorithm's design, revealing that the exploration schedule is the cause. Based on this finding, we propose a simple rule that enables the RL agent to escape from that failure mode.

- **Key Takeaway:** Learn to diagnose failures via internal metrics and trace them to algorithm design.

- The Distributional Reward Critic Framework for Reinforcement Learning Under Perturbed Rewards

Xi Chen, Zhihui Zhu, and Andrew Perrault. *AAAI 2025*.

- To handle perturbations that alter the optimal policy, we reformulate the reward regression as a classification task, outperforming baselines under certain assumptions about the perturbation structure.

- To remove these strict assumptions, we leverage training metrics to infer the perturbation structure, yielding a more general method that matches the performance of the original approach.

- **Key Takeaway:** Learn to use optimization metrics as proxies for critical algorithm parameters.

- Using Reinforcement Learning for Multi-Objective Cluster-Level Optimization of Non-Pharmaceutical Interventions for Infectious Disease

Xueqiao Peng, Jiaqi Xu, **Xi Chen**, Dinh Song An Nguyen, Andrew Perrault. *ML4H 2023*.

AWARDS

Eleanor Quinlan Graduate Teaching Award — CSE Department, The Ohio State University

2025

TECHNICAL SKILLS

ML Framework: PyTorch

RL & LLM: PPO; RLHF; DPO; RLVR