Chenyang Cao

Phone: 18817831156 | Email: chenyangcao5@gmail.com
Address: Tsinghua SIGS, University Town, 518055, Shenzhen, Guangdong, China

EDUCATION

Tsinghua University   Aug 2022 - Jun 2025

Master of Electronic and Information Engineering   Shenzhen

Research Area: Reinforcement Learning and Robot Learning

Related Courses: Artificial Neural Networks (A), Space Robot Technology (A), Model Predictive Control (A -).

Fudan University   Sep 2018 - Jun 2022

Bachelor of Science   Shanghai

Major: Mathematics and Applied Mathematics

Related courses: Probability Theory (A), Complex Variable Functions (A -), Mathematical Models (A -), Optimization Methods (A -).

HONORS & AWARDS

First Class Scholarship of Tsinghua University   2024

Second Class Scholarship of Tsinghua University   2023

National Inspirational Scholarship   2022

Meritorious Winner of Mathematical Contest in Modeling   2022

PUBLICATIONS

Conference Papers

- **Chenyang Cao**, Zichen Yan, Renhao Lu, Junbo Tan, Xueqian Wang. Offline Goal-Conditioned Reinforcement Learning for Safety-Critical Tasks with Recovery Policy[C]//2024 IEEE International Conference on Robotics and Automation (ICRA).

- Silang Wu, Huayue Liang, **Chenyang Cao**, Houde Liu, Xueqian Wang. A Wristband Device for remote robot arm teleoperation. (Accepted by ROBIO2025)

- **Chenyang Cao**, Yucheng Xin, Silang Wu, Longxiang He, Zichen Yan, Junbo Tan, Xueqian Wang. FOSP: Fine-tuning Safe Policy through World Models. (Accepted by ICLR2025 poster, rating: 8, 8, 6, 6)

Preprint and In-submission

- **Chenyang Cao**, Miguel Rogel-García, Mohamed Nabail, Xueqian Wang, Jonathan Kelly, Nicholas Rhinehart. Residual Reward Model for Preference-based Reinforcement Learning. (submitted to RLC2025)

RESEARCH EXPERIENCE

Residual reward model with human feedback   Sep 2024 - Present

Research Intern, Supervised by Professor Rhinehart   University of Toronto, Toronto

- We explore the effectiveness of using different priors and residual reward models based on human preferences.

- Evaluate through simulation environment and real-world deployment that residual models can improve insufficient priors and **preference-based reinforcement learning** algorithms through human feedback.

Fine-tuning offline safe model-based reinforcement learning in real world   Mar 2024 - Sep 2024

Leader   Tsinghua SIGS, Shenzhen

- A reinforcement learning method is proposed to solve offline-to-online safe problems. Fine-tuning models to generalize to new safe tasks without constraint violations.

- We introduce i) Additionally introduce value function with the **in-sample training** approach for offline RL; ii) Model-based **policy expansion** during offline-to-online fine-tuning; iii) **Reachability guidance** for safety consideration.

- Our method perform well in offline and online stage. We also deploy it in real robot by offline training without sim-to-real. An academic paper is submitted to ICLR2025.

- The videos are shown at this [URL](URL).

Safety obstacle avoidance problem of robotic arms based on offline reinforcement learning   Mar 2023 - Sep 2023

Leader   Tsinghua SIGS, Shenzhen

- An algorithm based on recovery reinforcement learning was designed to extend offline goal-conditioned reinforcement learning to **constrained safety problems**, ensuring the safety of the agent's motion trajectory in offline scenarios.

- A simulation environment with obstacles was constructed, and the performance and safety of the algorithm were verified in the simulation.

- The algorithm was **deployed on a real robot** (Franka robotic arm) to verify its practicality in the real world, and the paper "Offline Goal-Conditioned Reinforcement Learning for Safety-Critical Tasks with Recovery Policy" was accepted by ICRA2024.
- The presentation video is available at Infovaya • Presentation

INTERNSHIP EXPERIENCE

Microsoft Asia   Nov 2023 - Mar 2024

Research Intern DKI   Beijing

- We studied the relationship between **multi-agent and large language models**, and made detailed evaluations of GPT4 on a multi-agent task.
- Discussed the performance capabilities of LLM-based agents in multi-agent cooperative tasks, and designed a **communication framework** to further improve their performance.
- Improve Bing performance for the application.

PROJECT EXPERIENCE

Unrestricted Force Feedback System   Mar 2023 - Jun 2023

Member

- Implementing a vehicle arm platform that can track human hands by gesture recognition and robot navigation.
- Mainly responsible for the construction of the Vrap simulation platform and the derivation of the kinematics model.

Intelligent Quadruped Robot for Tracking and Taking Photos   Oct 2022 - Jan 2023

Member

- Visual recognition for tracking control and planning. The quadruped robot recognized the subject of the object and realized real-time locking of the target, continuous tracking.
- Mainly responsible for visual recognition, using yolov5 network for recognition.

SKILLS, CERTIFICATIONS & OTHERS

- **Skills:** MATLAB, Python, R, LaTex, Office software; Pytorch; Franka's Python interface
- **Certifications:** English (CET-6)，TOFEL：94 (R: 26; L: 23; S: 20; W: 25)

SUMMARY

- I am very interested in **robot learning** and have researched the application of **large language models** in multi-agent cooperation tasks, the combination of **quadruped robots and learning**, and the intelligent decision-making of **reinforcement learning** in various tasks of robotic arms. I also have experience **deploying real robots**.
- I am familiar with classic reinforcement learning algorithms, with in-depth research on **offline goal-conditioned reinforcement learning**. I am able to quickly learn about the innovation and application of reinforcement learning algorithms.
- I graduated from the Department of Mathematics with a bachelor's degree, possessing **good logical reasoning skills and mathematical proof-writing skills**.

REFERENCES

Prof. Xueqian Wang (master superviser)

Tsinghua SIGS, Tsinghua University

(+86) 0755-26031120

E-mail: wang.xq@sz.tsinghua.edu.cn

Google-scholar: scholar.google.com