

Xiang Zheng

| | | |
|----------------------------------|--|---|
| CONTACT INFORMATION | G2322, AC1 City University of Hong Kong Kowloon, Hong Kong | Email: xzheng253-c@my.cityu.edu.hk Webpage: https://x-zheng16.github.io Phone: +852 6337 2959 |
| RESEARCH INTERESTS | Reinforcement Learning, Trustworthy AI, AI Red Teaming | |
| EDUCATION | City University of Hong Kong , Hong Kong, China <i>Ph.D. in Computer Science</i> Supervisor: Prof. Cong Wang Fields: Reinforcement Learning, AI Security GPA: 4.00/4.00 | 2020-present |
| | Tsinghua University , Beijing, China <i>M.S. in Control Science and Engineering</i> Supervisor: Prof. Tao Zhang Fields: Robot Learning GPA: 3.55/4.00 | 2016-2019 |
| | Beihang University , Beijing, China <i>B.S. in Automation, Second B.S. in Mathematics</i> Selected into Honors College (1/200) GPA: 3.74/4.00 (Total 222.5 credits) | 2012-2016 |
| RESEARCH & WORK EXPERIENCE | City University of Hong Kong , Hong Kong, China <i>Graduate Research Assistant</i> Supervisor: Prof. Cong Wang Topic: Efficient exploration strategies for reinforcement learning <ul style="list-style-type: none">• Designed constrained intrinsic motivation for unsupervised reinforcement learning. This work is accepted by IJCAI 2024.• Task-aware intrinsic motivation for sparse-reward reinforcement learning. This work is submitted to NeurIPS 2024.• Developed intrinsically motivated adversarial policy against robotic RL agents. This work is accepted by DSN 2024.• Explored intrinsically motivated red teaming for Large Language Models. This work is submitted to NeurIPS 2024. | 2020-present |
| | Xi'an Jiaotong University , Xi'an, China <i>Visiting Researcher</i> Supervised by Prof. Chao Shen Topic: Efficient and robust reinforcement learning | 2022-2023 |

| | |
|--|-------------|
| Huawei , Xi'an, China <i>Senior Researcher</i> Topic: Cloud computing resource scheduling | 2020 |
| Xi'an Jiaotong University , Xi'an, China <i>Visiting Researcher</i> Supervised by Prof. Chao Shen Topic: Adversarial machine learning | 2019 |
| Rich AI Technology , Beijing, China <i>NLP Intern</i> Topic: Text summarization via Transformers | 2019 |
| National Institute of Informatics , Tokyo, Japan <i>Research Intern</i> Supervised by Prof. Tetsunari Inamura Topic: Indoor service robot navigation and control | 2018 |
| The University of New South Wales , Sydney, Australia <i>Student Intern</i> Supervised by Prof. Elias Aboutanios Topic: Spacecraft orbit control | 2016 |

PUBLICATION

Journal

- J-1 Yuxue Cao, Shengjie Wang, **Xiang Zheng**, Wenke Ma, Xinru Xie, Lei Liu, "Reinforcement learning with prior policy guidance for motion planning of dual-arm free-floating space robot," *Aerospace Science and Technology*, vol. 136, pp. 108098, 2023.
Impact Factor 5.6, JCR Rank Q1 (3/34).
- J-2 Shengjie Wang, Yuxue Cao, **Xiang Zheng**, Tao Zhang, "A learning system for motion planning of free-float dual-arm space manipulator towards non-cooperative object," *Aerospace Science and Technology*, vol. 131, pp. 107980, 2022.
Impact Factor 5.6, JCR Rank Q1 (3/34).
- J-3 Shengjie Wang, Yuxue Cao, **Xiang Zheng**, Tao Zhang, "Collision-free trajectory planning for a 6-DoF free-floating space robot via hierarchical decoupling optimization," *IEEE Robotics and Automation Letters (RA-L)*, vol. 7, no. 2, pp. 4953–4960, 2022.
Impact Factor 5.2, JCR Rank Q2 (10/30).

Conference

- C-1 **Xiang Zheng**, Xingjun Ma, Chao Shen, and Cong Wang, "Constrained

- intrinsic motivation for reinforcement learning,” in *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI)*, 2024.
- C-2 **Xiang Zheng**, Xingjun Ma, Shengjie Wang, Xinyu Wang, Chao Shen, and Cong Wang, “Toward evaluating robustness of reinforcement learning with adversarial policy,” in *Proceedings of the 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, 2024.
Acceptance Rate 20.7%
- C-3 Shengjie Wang, Fengbo Lan, **Xiang Zheng**, Yuxue Cao, Oluwatosin Oseni, Haotian Xu, Tao Zhang, Yang Gao, “A policy optimization method towards optimal-time stability,” in *Proceedings of the 7th Conference on Robot Learning (CoRL)*, 2023.
- C-4 Shengjie Wang, **Xiang Zheng**, Yuxue Cao, Tao Zhang, “A Multi-Target Trajectory planning of a 6-DoF free-floating space robot via reinforcement learning,” in *Proc. of the 34th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021.
- C-5 Shengjie Wang, Yuxue Cao, **Xiang Zheng**, Tao Zhang, “An end-to-end trajectory planning strategy for free-floating space robots,” in *Proc. of the 40th Chinese Control Conference (CCC)*, 2021.
- C-6 Shihao Zhao, Xingjun Ma, Xiang Zheng, James Bailey, Jingjing Chen, Yu-Gang Jiang, “Clean-label backdoor attacks on video recognition models,” in *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
Acceptance Rate 22.1%
- C-7 Xiang Zheng, Ziwei Wang, Tao Zhang, “Robust finite-time attitude tracking control for nonlinear quadrotor with uncertainties and delays,” in *Proc. of IEEE International Conference on Robotics and Biomimetics (RO-BIO)*, 2017.

| | | |
|--------------------|-------------------------------------|------------------|
| AWARDS & HONORS | IJCAI Travel Grant | 2024 |
| | IJCAI Organization and AIJ Division | |
| | Amount: USD 500 | |
| | Conference Grant | 2024 |
| | City University of Hong Kong | |
| | Amount: HKD 10,000 | |
| | DSN Student Travel Grant | 2024 |
| | DSN Student Travel Awards Committee | |
| | Amount: USD 1,500 | |
| | Research Activities Fund | 2022-2023 |
| | City University of Hong Kong | |

| | | |
|--------------------------|---|-------------------------------|
| | Amount: HKD 96,000 | |
| | Institutional Research Tuition Grant City University of Hong Kong Amount: HKD 168,384 | 2020-2024 |
| | CityU Presidential PhD Scholarship City University of Hong Kong Amount: HKD 1,353,120 | 2020-2024 |
| TEACHING & EXPERIENCE | Teaching Assistant, City University of Hong Kong CS5293, Topics on Information Security | Semester B 2023/24 |
| | Teaching Assistant, City University of Hong Kong CS4394, Information Security and Management | Semester A 2023/24 |
| | Teaching Assistant, City University of Hong Kong CS4293, Topics in Cybersecurity CS5293, Topics on Information Security | Semester B 2021/22 |
| | Teaching Assistant, City University of Hong Kong CS4394, Information Security and Management CS5294, Information Security Technology Management | Semester A 2021/22 |
| | Teaching Assistant, City University of Hong Kong CS4293, Topics in Cybersecurity CS6290, Privacy-enhancing Technologies | Semester B 2020/21 |
| | Teaching Assistant, City University of Hong Kong CS2310 / CS2311, Computer Programming | Semester A 2020/21 |
| MISC. | Language: Python, C/C++ Framework: PyTorch Simulator: MuJoCo Platform: Tianshou, Stable Baselines, CleanRL, TRL Environment: Linux/Unix, macOS and Windows Fluent in English and Chinese | |