Yanchao Sun

Email : ycs@umd.edu
Phone : +1 (240) 413-8873
Website : https://ycsun2017.github.io

Education

- **University of Maryland, College Park** — Maryland, U.S.A.
  Ph.D. in Computer Science; Advisor: Dr. Furong Huang — *Sep 2018 – Present*
- **Sichuan University** — Chengdu, China
  B.S. in Computer Science and Technology — *Sep. 2014 – Jun 2018*

Research Interests

- *Robustness of reinforcement learning agents* against adversarial attacks (training-time and test-time robustness)
- *Knowledge transfer in reinforcement learning* (in-domain and cross-domain transfer learning)
- *Sample efficiency and computational efficiency* of reinforcement learning algorithms
- *Representation learning for control problems* driven by theoretical understandings

Publications

1. **Yanchao Sun**, Ruijie Zheng, Yongyuan Liang, and Furong Huang. "Who Is the Strongest Enemy? Towards Optimal and Efficient Evasion Attacks in Deep RL". Proceedings of the 10th International Conference on Learning Representations (**ICLR** 2022). (**Best Paper Award** at the NeurIPS 2021 SafeRL Workshop.)

2. **Yanchao Sun**, Ruijie Zheng, Xiyao Wang, Andrew Cohen, and Furong Huang. "Transfer RL across Observation Feature Spaces via Model-Based Regularization". Proceedings of the 2th International Conference on Learning Representations (**ICLR** 2022).

3. **Yanchao Sun**, Da Huo, and Furong Huang. "Vulnerability-Aware Poisoning Mechanism for Online RL with Unknown Dynamics". Proceedings of the 9th International Conference on Learning Representations (**ICLR** 2021).

4. **Yanchao Sun**, Xiangyu Yin, and Furong Huang. "TempLe: Learning Template of Transitions for Sample Efficient Multi-task RL". Proceedings of the 35th AAAI Conference on Artificial Intelligence (**AAAI** 2021).

5. **Yanchao Sun** and Furong Huang. "Can Agents Learn by Analogy? An Inferable Model for PAC Reinforcement Learning". Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (**AAMAS** 2020).

6. Jingling Li, **Yanchao Sun**, Jiahao Su, Taiji Suzuki and Furong Huang. "Understanding Generalization in Deep Learning via Tensor Methods". Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS** 2020).

7. **Yanchao Sun**, Cong Qian, Ning Yang and Philip S. Yu. "Collaborative Inference of Coexisting Information Diffusions". Proceedings of the IEEE 17th International Conference on Data Mining (**ICDM** 2017).

Preprints & Workshops

1. **Yanchao Sun**, Ruijie Zheng, Parisa Hassanzadeh, Yongyuan Liang, Soheil Feizi, Sumitra Ganesh and Furong Huang. "Provably Robust Multi-agent Reinforcement Learning against Adversarial Communication". Under review.

2. Yongyuan Liang*, **Yanchao Sun***, Ruijie Zheng, and Furong Huang. "Efficiently Improving the Robustness of RL Agents against Strongest Adversaries". (*Equal contribution.) Oral presentation at the NeurIPS 2021 SafeRL Workshop.

Research Experience

- **Research Assistant** — University of Maryland, College Park
  Advisor: Dr. Furong Huang — *Jun 2019 – Present*
  - **Adversarial Robustness of Deep Reinforcement Learning**
    established a systematical understanding of the robustness of RL agents against adversarial attacks, including both training-time attacks and test-time attacks;
    proposed effective and efficient algorithms for evaluating and improving the robustness of any RL agents.
  - **Sample efficient Multi-task Reinforcement Learning**
    proposed the first PAC-MDP method for multi-task reinforcement learning that could be applied to tasks with varying state/action space.
  - **Understanding Dynamics of Adversarial Training**
    related robustness of a deep network with the movement of decision boundary, and proposed a re-weighting algorithm to improve adversarial training.
- **Research Intern** — Microsoft Research, Redmond
  Supervisor: Dr. Shuang Ma — *(Upcoming) Jun 2022 – Aug 2022*
  - **Pretraining Representation for Reinforcement Learning Tasks.** (Ongoing project.)
- **AI Research Summer Associate** — JPMorgan Chase & Co., New York
  Supervisor: Dr. Sumitra Ganesh — *Jun 2021 – Aug 2021*
  - **Robustifying Agents in a Communicative Multi-agent System.**
    studied the emergence of adversarial communication in a multi-agent system and how to make agents robust against adversarial communication by defensive information sharing and selective information usage.
- **Machine Learning Research Intern** — Unity Technologies, San Francisco
  Supervisor: Dr. Andrew Cohen — *May 2020 – Aug 2020*
  - **Cross-domain Transfer RL with Model Regularizers.**
    designed an algorithm that utilizes model-based regularizers to transfer a learned policy to a new task with different observation space, contributed to the ML-Agents toolkit.

Honors and Awards

- WiML Travel Funding by NSF Grant, 2019
- Dean's Fellowship, University of Maryland, College Park, 2018
- Outstanding Graduates of Sichuan University, 2017
- Special Award of Wang Wen Guo Scholarship, Wuyuzhang Honors College, 2016
- Excellent Student Cadre of Sichuan University, 2016
- National Endeavor Scholarship, China, 2016
- The **1st Prize** of Blue Bridge Cup National C/C++ Programming Contest, Sichuan Province, 2016
- National Scholarship, China, 2015
- Excellent Student of Sichuan University, 2015
- The **1st Prize** of The Seventh Chinese Mathematics Competitions, Sichuan Province, 2015

Professional Services

- Reviewer of International Conference on Machine Learning (ICML), 2020, 2021, 2022
- Reviewer of International Conference on Learning Representations (ICLR), 2021, 2022
- Reviewer of Advances in Neural Information Processing Systems (NeurIPS), 2021