Email : ycsun1531@gmail.com
Yanchao Sun
Website : https://ycsun2017.github.io

Research Interests

- *Foundation Models:* building and utilizing data-centric foundation models in various domains
- *Reinforcement Learning:* sample efficiency, robustness, and adaptability in sequential decision making
- *Trustworthy Machine Learning:* trustworthiness and adversarial robustness of deep neural networks

Work Experience

- **AI Research Scientist** — JPMorgan Chase & Co., New York
  Supervisor: Dr. Sumitra Ganesh — *Jun 2023 – Present*
  - **Applied Research on Large Language Models and Foundation Models.**
    improving LLM-powered autonomous agents with multi-task learning, and building foundation models with financial dadta.

- **Ph.D. Student, Graduate Assistant** — University of Maryland, College Park
  Advisor: Dr. Furong Huang — *Sep 2018 – May 2023*
  - **Thesis: Towards Robust and Adaptable Real-World Reinforcement Learning**
    established systematic theory and methods for the robustness of RL agents against adversarial attacks, achieving state-of-the-art robustness in multiple benchmarks;
    proposed a series of multi-task learning and transfer learning algorithms to improve the adaptability of reinforcement learning methods.

- **Research Intern** — Microsoft Research, Redmond
  Supervisor: Dr. Shuang Ma — *Jun 2022 – Aug 2022*
  - **Pretraining Representation for Reinforcement Learning Tasks.**
    proposed a self-supervised pretraining framework that works for various downstream control tasks, based on a transformer backbone.

- **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.

- **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.

Education

- **University of Maryland, College Park** — Maryland, U.S.A.
  Ph.D. in Computer Science. — *Sep 2018 – May 2023*

- **Sichuan University** — Chengdu, China
  B.S. in Computer Science and Technology. — *Sep 2014 – Jun 2018*

Publications

14. Yao Wei, **Yanchao Sun**, Ruijie Zheng, Sai Vemprala, Rogerio Bonatti, Shuhang Chen, Ratnesh Madaan, Zhongjie Ba, Ashish Kapoor and Shuang Ma. "Is Imitation All You Need? Generalized Decision-Making with Dual-Phase Training". ICCV 2023.

13. **Yanchao Sun**, Shuang Ma, Ratnesh Madaan, Rogerio Bonatti, Furong Huang, and Ashish Kapoor. "SMART: Self-supervised Multi-task pretrAining with contRol Transformers". ICLR 2023 (***Spotlight***).

12. **Yanchao Sun**, Ruijie Zheng, Parisa Hassanzadeh, Yongyuan Liang, Soheil Feizi, Sumitra Ganesh and Furong Huang. "Certifiably Robust Multi-Agent Reinforcement Learning against Adversarial Communication". ICLR 2023.

11. Yuancheng Xu, **Yanchao Sun**, and Furong Huang. "Exploring and Exploiting Decision Boundary Dynamics for Adversarial Robustness". ICLR 2023.

10. Yongyuan Liang*, **Yanchao Sun***, Ruijie Zheng, and Furong Huang. "Efficient Adversarial Training without Attacking: Worst-Case-Aware Robust Reinforcement Learning". (*Equal Contribution.) NeurIPS 2022.

9. Jifeng Hu, **Yanchao Sun**, Hechang Chen, Sili Huang, Haiyin Piao, Yi Chang, and Lichao Sun. "Distributional Reward Estimation for Effective Multi-agent Deep Reinforcement Learning". NeurIPS 2022.

8. Kaiwen Yang, **Yanchao Sun**, Jiahao Su, Fengxiang He, Xinmei Tian, Furong Huang, Tianyi Zhou, and Dacheng Tao. "Adversarial Auto-Augment with Label Preservation: A Representation Learning Principle Guided Approach". NeurIPS 2022 (***Spotlight***).

7. **Yanchao Sun**, Ruijie Zheng, Yongyuan Liang, and Furong Huang. "Who Is the Strongest Enemy? Towards Optimal and Efficient Evasion Attacks in Deep RL". ICLR 2022. (***Best Paper Award*** at the NeurIPS 2021 SafeRL Workshop.)

6. **Yanchao Sun**, Ruijie Zheng, Xiyao Wang, Andrew Cohen, and Furong Huang. "Transfer RL across Observation Feature Spaces via Model-Based Regularization" ICLR 2022.

5. **Yanchao Sun**, Da Huo, and Furong Huang. "Vulnerability-Aware Poisoning Mechanism for Online RL with Unknown Dynamics". ICLR 2021.

4. **Yanchao Sun**, Xiangyu Yin, and Furong Huang. "TempLe: Learning Template of Transitions for Sample Efficient Multi-task RL". AAAI 2021.

3. **Yanchao Sun** and Furong Huang. "Can Agents Learn by Analogy? An Inferable Model for PAC Reinforcement Learning". AAMAS 2020.

2. Jingling Li, **Yanchao Sun**, Jiahao Su, Taiji Suzuki and Furong Huang. "Understanding Generalization in Deep Learning via Tensor Methods". AISTATS 2020.

1. **Yanchao Sun**, Cong Qian, Ning Yang and Philip S. Yu. "Collaborative Inference of Coexisting Information Diffusions". ICDM 2017.

Manuscripts Under Submission

10. Yongyuan Liang, **Yanchao Sun**, Ruijie Zheng, Xiangyu Liu, Tuomas Sandholm, Furong Huang and Stephen Marcus McAleer. "Adapting Robust Reinforcement Learning to Handle Temporally-Coupled Perturbations".

9. Sili Huang, **Yanchao Sun**, Jifeng Hu, Siyuan Guo, Bo Yang, Hechang Chen, Yi Chang and Lichao Sun. "Learning Generalizable Agents via Saliency-guided Features Decorrelation".

8. Siyuan Guo, **Yanchao Sun**, Jifeng Hu, Sili Huang, Hechang Chen, haiyin piao, Lichao Sun and Yi Chang. "A Simple Unified Uncertainty-Guided Framework for Offline-to-Online Reinforcement Learning".

7. Jifeng Hu, **Yanchao Sun**, Sili Huang, Siyuan Guo, Hechang Chen, Li Shen, Lichao Sun, Yi Chang and Dacheng Tao. "Instructed Diffuser with Temporal Condition Guidance for Offline Reinforcement Learning."

6. Xiangyu Liu, Souradip Chakraborty, **Yanchao Sun** and Furong Huang. "Rethinking Adversarial Policies: A Generalized Attack Formulation and Provable Defense in Multi-Agent RL".

5. Lilin Zhang, Ning Yang, **Yanchao Sun** and Philip S Yu. "Provable Unrestricted Adversarial Training without Compromise with Generalizability".

4. Ruijie Zheng, Xiyao Wang, **Yanchao Sun**, Shuang Ma, Jieyu Zhao, Huazhe Xu, Hal Daumé III and Furong Huang. "TACO: : Temporal Latent Action-Driven Contrastive Loss for Visual Reinforcement Learning".

3. Xiyao Wang, Ruijie Zheng, **Yanchao Sun**, Ruonan Jia, Wichayaporn Wongkamjan, Huazhe Xu and Furong Huang. "COPlanner: Plan to Roll Out Conservatively but to Explore Optimistically for Model-Based RL".

2. Yuancheng Xu, Chenghao Deng, **Yanchao Sun**, Ruijie Zheng, Xiyao Wang, Jieyu Zhao and Furong Huang. "Equal Long-term Benefit Rate: Adapting Static Fairness Notions to Sequential Decision Making".

1. Zhi Zhang, Haochen Zhang, **Yanchao Sun**, Han Liu, Furong Huang and Oscar Hernan Madrid Padilla. "A Stochastic PAC-Bayes Algorithm For Lifelong Reinforcement Learning".

Honors and Awards

- **Outstanding Research Assistant Award** (top 2%), University of Maryland, College Park, 2022
- **Best Paper Award** at the NeurIPS SafeRL Workshop, 2021
- Dean's Fellowship, University of Maryland, College Park, 2018
- 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

- Co-organizer of the PerDream Workshop at ICCV 2023
- Co-organizer of the 1st Reincarnating RL Workshop at ICLR 2023
- Program Committee of NeurIPS 2022 Deep RL Workshop
- Reviewer of International Conference on Learning Representations (ICLR), 2021, 2022, 2023
- Reviewer of Advances in Neural Information Processing Systems (NeurIPS), 2021, 2022, 2023
- Reviewer of International Conference on Machine Learning (ICML), 2020, 2021, 2022, 2023

Invited Talks

3. Invited talk at the Robotics Session at the 18th Coordinated Science Laboratory Student Conference (CSLSC), University of Illinois at Urbana-Champaign, 2023.

2. Talk at the MSR NYC Seminar, 2023

1. Talk at the Machine Learning Seminar Series, University of Maryland, 2019