Jiayi Yuan

(540) 255-9188 yuancarrieyjy@gmail.com yuanjiayiy.github.io

AI researcher working on applying and improving learning algorithms for cross-domain intelligent systems

Education

Master of Science, University of Washington, Seattle, Washington Sept 2022- Dec 2024 (Expected)

- Major in Computational Linguistics
- GPA 4.0/4.0 | Advisor: Shane Steinert-Threlkeld and Natasha Jaques

Bachelor of Science, Carnegie Mellon University, Pittsburgh, Pennsylvania Aug 2018- May 2021

- Major in Neuroscience and Computer Science
- GPA 3.5/4.0 | University Honor

Research Experience

Paul G. Allen Center for Computer Science, University of Washington, Seattle, WA, June 2023 - Present

Working towards developing and applying multiagent reinforcement learning algorithm in various problem spaces such as robotics and climate change

- **CASHER**: Crowdsourcing and Amortizing Human Effort for Real-to-Sim-to-Real, a pipeline for scaling up data collection and learning generalist policies, by leveraging crowdsource digital twins of real-world scenes using 3D reconstruction techniques and collecting large-scale data in these simulation scenes, rather than in the real-world. Accepted by Data Generation for Robotics Workshop at RSS 2024.
- **InvestESG**: Investing in Environmental, social, and governance (ESG) disclosure, a multiagent reinforcement learning benchmark environment to analyze the impact of ESG disclosure mandate on corporate investment in climate efforts as a sequential social dilemma environment. Accepted by NeurIPS 2024 Workshop on Tackling Climate Change with Machine Learning. Under review for ICLR 2025.
- Engineered an approach of fine-tuning pre-trained foundational models with asynchronous human feedbacks onto the WidowX 250s Robot Arm on short-horizon tasks.

Work Experience

Amazon AGI, Seattle, WA, July 2021 - July 2024

Software Development Engineer II

- Designed, prototyped, and engineered the iDPS CLI, an intuitive command-line interface optimized for initiating and monitoring large-scale dataset processing jobs. This innovation drastically reduced development effort by 80% while enhancing the efficiency of dataset processing.
- Led the data preparation team to develop the first data quality reporting tool for AGI, Data Quality Report Tool, to automate the labor-intensive training data inspection process, reducing development time by 95%.
- Pioneered the development of annotation inconsistency checker tools for Alexa NLU utterance data. Employed advanced SentenceBERT semantic encoding techniques to detect inconsistently labeled data in training and test sets. This tool not only increased model accuracy by 25% but also reduced the time annotators needed to resolve inconsistencies by 50%.
- Demonstrated my work and provided tutorials within the Amazon AGI organization, reaching over 300 engineers and scientists. Received distinguished Org-Level Peer Recognition Awards in both 2022 and 2023, acknowledging exceptional contributions to cutting-edge software engineering initiatives.

Skills

Robots: WidowX250s, Franka Emika

Programming Languages: Python, Java, C/C++, TypeScript, MATLAB, Bash, SQL, R, Prolog, OCaml

Data Science and Machine Learning: PyTorch, JAX, Apache Spark, OpenCV, Tensorflow, Scikit-learn, NumPy

Robotics and Reinforcement Learning: ROS, Gym, MoveIt, Gazebo, IsaacSim, PettingZoo, SB3

Miscellaneous Tools: AWS, CUDA, GCP, Docker, Git, Conda, Slurm.