Yunpeng Qing

- Email: qingyunpeng@zju.edu.cn • Github: github/Plankson • Homepage: plankson.github.io
- Google Scholar: https://scholar.google.com/citations?user=-RvDI44AAAAJ

Research Interest: Reinforcement Learning, Offline Reinforcement Learning

EDUCATION

R.A., Visual Intelligence and Pattern Analysis Lab, Zhejiang University, Hangzhou, China Sept. 2021 - Present

B.S., Computer Science and Technology, Zhejiang University, Hangzhou, China Sept. 2019 - Jun. 2023

- *GPA*: 3.94/4.0; *Comprehensive score ranking*: 19/159
- *Main Course*: Calculus A I (96), Calculus A II (93), Fundamentals of Data Structures (100), Numerical Analysis(96), Operating System (97), Computing Vision (92), Artificial Intelligence (96), Artificial Intelligence Security (96).
- *Award*: Third-class Scholarship (2 consecutive years), Excellent Graduation Thesis.

PUBLICATIONS

(* indicates Equal Contribution)

A Survey on Explainable Reinforcement Learning: Concepts, Algorithms, and Challenges [Paper] [GitHub]

Yunpeng Qing, Shunyu Liu, Jie Song, Huiqiong Wang, Mingli Song.

Major Revision in *ACM Computing Survey*. (IF/JCR: 16.6/Q1)

- We propose a new RL-based taxonomy for current Explainable Reinforcement Learning (XRL) works to make up for the shortcomings of lacking RL-based architecture in the XRL community. The taxonomy is based on the explainability of different target parts of the reinforcement learning framework: model, reward, state, and task.

Advantage-Aware Policy Optimization for Offline Reinforcement Learning [Paper]

Yunpeng Qing, Shunyu Liu, Jingyuan Cong, Kaixun Chen, Yihe Zhou, Mingli Song.

Under review in *International Conference On Machine Learning 2024*.

- We propose an Advantage-Aware Policy Optimization (A2PO) framework for offline reinforcement learning to solve the constraint conflict issue when offline datasets are collected from multiple behavior policies. Experiments conducted on the various datasets of the D4RL benchmark demonstrate that A2PO yields results superior to state-of-the-art counterparts.

Curricular Subgoals for Inverse Reinforcement Learning [Paper] [Code]

Shunyu Liu*, **Yunpeng Qing***, Shuqi Xu, Hongyan Wu, Jiangtao Zhang, Jingyuan Cong, Tianhao Chen, Yunfu Liu, Mingli Song.

Under review in *IEEE Transaction on Intelligent Transportation Systems*. (IF/JCR: 8.5/Q1)

- We propose a Curricular Subgoal-based Inverse Reinforcement Learning (CSIRL) framework explicitly disentangling a task with several subgoals to guide imitation in solving the error propagation problem. Experiments on D4RL and self-driving benchmarks demonstrate that D4RL yields results superior to the state-of-the-art counterparts, as well as better explainability.

Centralized Advising with Decentralized Pruning Framework for Multi-Agent Reinforcement Learning [Paper] [Code]

Yihe Zhou, Shunyu Liu, **Yunpeng Qing**, Kaixuan Chen, Tongya Zheng, Yanhao Huang Jie Song, Mingli Song.

Under review in *International Conference On Machine Learning 2024*.

- We introduce a novel Centralized Advising and Decentralized Pruning (CADP) framework for multi-agent reinforcement learning, realizing efficacious message exchange for training to fully utilize global information. Experiments on StarCraft II and Google Research Football benchmarks show the superiority to the state-of-the-art counterparts.

RESEACH EXPERIENCE

Participation in School-Enterprise Cooperation Project. China May 2022 - May 2023

- *Project name*: Decision-Making for Autonomous Vehicles in Urban Road Scenarios based on Inverse Reinforcement Learning.
- I have designed a novel inverse reinforcement learning algorithm that automatically generates subgoals based on the agent uncertainty to solve complex navigation tasks. Additionally, I have been responsible for generating expert data, implementing the entire framework in Python, and conducting experimental comparisons with other inverse reinforcement learning baselines.

Participation in National Key Research and Development Project May 2023 - present

- *Project name*: Power Grid Control for Human-in-the-loop Hybrid Reinforcement Learning
- I have designed a power grid-based environment and utilized reinforcement learning baselines on it to train expert policy. The environment is about adjusting the output power of the generator to achieve the target power of the specified section.