Changzhi Yan

Homepage Github cz.yan20@gmail.com

Education

Hubei University of Technology Sep 2012 – Jun 2016
B.E. in Automation GPA: 3.73 / 4.00

- Ranking: 1 / 133

Tsinghua University Sep 2017 – Jun 2020
M.E. in Control Engineering GPA: 3.43 / 4.00

- Supervisor: Xueqian Wang

Experience

Full-time Assistant Researcher Jul 2020 – Jul 2021
Tsinghua Shenzhen International Graduate School

- Lectured on *Reinforcement Learning Introduction* in course *Flight Control System* (Course No. 80250833)
- Researched on reinforcement learning in *AI and Robot Center*
- Led the reinforcement learning team

Projects

Optimality Guarantee for AWR Oct 2024 – Dec 2024

- Motivated the advantage–weighted regression (AWR, an iterative regression-based RL algorithm) as an approximate optimization of a constrained policy search problem and derived its policy update rule
- Developed the optimality guarantee for the AWR algorithm in the tabular setting: under some conditions, AWR guarantees to output near-optimal policies
- Derived an upper bound on the sub-optimality of the policy output by AWR upon termination, which has no dependence on the size of the state and action space
- This is a pure theory work conducted exclusively on my own in my gap years

Goal-specific Transferable Skills in Latent Space Jul 2020 – Oct 2020

- Proposed a hierarchical RL method that enabled sample-efficient training for agents to perform goal-based robotic tasks
- Proposed an algorithm to train latent variable conditioned policies to be diverse and distinguishable in the goal-based sparse reward environment (latent embedding, low-level hierarchy)
- Proposed a method to train a meta-controller that modulates and interpolates low-level base policies to solve downstream tasks with high sample efficiency (transferring skills, high-level hierarchy)
- Tools Used: Python, Tensorflow, MuJoCo

Goal-based Robotic Tasks Oct 2019 – Jan 2020

- Integrated the Soft Actor-Critic (SAC) algorithm with the hindsight experience replay (HER) buffer to train the robotic manipulator to perform a collection of goal-based robotics tasks in gym benchmarks
- Implemented an algorithm called SAC-HER and demonstrated the necessity of incorporating hindsight experience replay buffer to train goal-conditioned policies that perform goal-based tasks
- Built a wrapper that converts gym goal-based robotic environments to rllab environments
- Tools Used: Python, Tensorflow, MuJoCo

Control of Free-Floating Robots to Capture Targets May 2018 – Jul 2018

- Trained a robotic arm to capture targets in the microgravity environment using the Soft Q-learning (SQL) algorithm with heuristic reward shaping
- Built the robotic arm model in the V-REP simulator and its interaction protocol with Python
- Tools Used: Python, Tensorflow, V-REP

Preprints and Publications

- **Changzhi Yan.** (2024). On Global Optimality Guarantee for Advantage-Weighted Regression. (Preprint)
- **C. Yan**, Q. Zhang, Z. Liu, X. Wang and B. Liang, "Control of Free-Floating Space Robots to Capture Targets Using Soft Q-Learning," *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Kuala Lumpur, Malaysia, 2018, pp. 654-660, doi: 10.1109/ROBIO.2018.8665049.

Awards

Outstanding Undergraduate Thesis Award in Hubei Province	2016
The First Prize Scholarship at Hubei University of Technology	2013 – 2015
National Scholarship	2015
7th National College Students Mathematical Competition, Second Prize	2015
2014 Contemporary Undergraduate Mathematical Contest in Modeling, Third Prize	2014
National Encouragement Scholarship	2013