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 the course *Flight Control System* (Course No. 80250833).
- Researched reinforcement learning in *Intelligent Robot and Control Center*.
- Led the reinforcement learning team.

Projects

Optimality Guarantee for Advantage-Weighted Regression (AWR) Oct 2024 – Dec 2024

- Formulated AWR as an approximate solution to a constrained policy search problem.
- Derived AWR policy update rule and established its near-optimality guarantees in the tabular setting.
- Proved an upper bound on the suboptimality of AWR's output policy, which is independent of the size of the state and action space.
- Conducted all theoretical analysis independently during my gap years.
- Resulted in a preprint: *On Optimality Guarantee for Advantage-Weighted Regression (2024)*.

Goal-Specific Skill Discovery and Transfer (Hierarchical RL) Jul 2020 – Oct 2020

- Formulated an information-theoretic objective for learning diverse, distinguishable, and goal-specific skills in multi-goal, sparse-reward environments.
- Derived a variational lower bound to enable tractable optimization; developed the Diverse Goal-Specific Skill Learning (DGSL) algorithm to maximize the lower bound for learning admissible skills.
- Proposed an extension to DGSL for hierarchical RL to improve sample efficiency in solving complex downstream tasks.
- Built a meta-controller capable of composing and interpolating discovered skills as the hierarchical policy.
- Resulted in a preprint: *Learning Diverse Goal-Specific Skills via Latent Embedding (2020)*.
- Tools Used: Python, TensorFlow, MuJoCo

Multi-Goal Robotic Manipulation Oct 2019 – Jan 2020

- Integrated Soft Actor-Critic (SAC) with Hindsight Experience Replay (HER) to train a robotic manipulator on multi-goal tasks.
- Implemented SAC-HER and empirically validated the necessity of using hindsight goal-relabeling for learning goal-conditioned policies in sparse-reward environments.
- Built a wrapper to convert the gym multi-goal robotic environments into rllab-compatible interfaces.
- 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 in simulation using Soft Q-Learning with heuristic reward shaping.
- Built the robotic arm model and interaction protocol in V-REP.
- Tools Used: Python, TensorFlow, V-REP

Preprints and Publications

- **Changzhi Yan.** (2024). On Optimality Guarantee for Advantage-Weighted Regression. (Preprint)
- **Changzhi Yan.** (2020). Learning Diverse Goal-Specific Skills via Latent Embedding. (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, China	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