Adversarial Imitation Learning with Preferences and Movement Primitives

Karlsruhe Institute of Technology

Autonomous Learning Robots(ALR)
Prof. Gerhard Neumann

Project Type

- ☑ Master Thesis
- ☐ Bachelor Thesis
- ☐ Research Project

Supervisors

Aleksandar Taranovic

aleksandar.taranovic@kit.edu

Hongyi Zhou

hongyi.zhou@kit.edu

Difficulty

Algorithmic

Math

Application

Requirements

- ☑ Python
- ☑ Pytorch
- ☑ Knowledge about RL and ML

Description

To facilitate an efficient and enjoyable teaching experience for human teachers, it is important to develop techniques for effectively incorporating and extracting useful information from various forms of feedback. Recently, work has been done in this area [3], resulting in the successful integration of two popular feedback types: demonstrations and preferences.

The most prominent methods in this area focused on learning a step-based reward function based on human demonstration and preference feedback, meanwhile training a policy to maximize the cumulative estimated reward. However, as human preferences usually depend on the entire trajectories, it's more intuitive to formulate a reward function based on episodic performance. Yet, due to the sparsity in episodic reward, most step-based RL algorithms are usually not able to solve tasks with such rewards. Recent episode-based policy searching algorithm with Movement Primitives [2][1] shows promising results in learning tasks with sparse and non-Markovian rewards.

The goal of this thesis is to combine the best of the world of episode-based policy searching and preference-based adversarial imitation learning[3], designing a new algorithm and benchmarking with other state-of-the-art methods in challenging simulation robot control tasks.

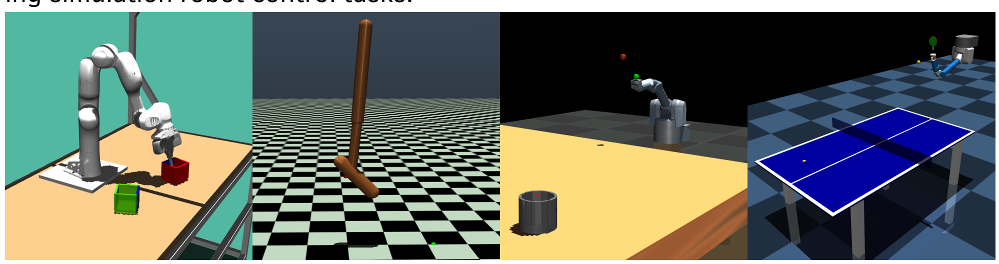

Figure 1: Robot control tasks with sparse and non-Markovian reward [2]

Tasks

The tasks in this project will involve:

- Literature research: The student needs to learn about Preference Learning, (Adversarial) Imitation Learning, and Movement Primitives.
- Implementation and Evaluation: Get familiar with the code base, design and implement a new method that combines existing state of the art approaches.
- Performance Evaluation: The novel method should be analyzed and compared against current state-of-the-art methods in simulation to investigate its performance.

References

[1] Ge Li, Zeqi Jin, Michael Volpp, Fabian Otto, Rudolf Lioutikov, and Gerhard Neumann. Prodmps: A unified perspective on dynamic and probabilistic movement primitives. *IEEE Robotics and Automation Letters*, 2023.

[2] Fabian Otto, Onur Celik, Hongyi Zhou, Hanna Ziesche, Ngo Anh Vien, and Gerhard Neumann. Deep black-box reinforcement learning with movement primitives. In *6th Annual Conference on Robot Learning (CoRL)*, 2022.

[3] Aleksandar Taranovic, Andras Kupcsik, Niklas Freymuth, and Gerhard Neumann. Adversarial imitation learning with preferences. In *International Conference on Learning Representations*, 2023.