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

Difficulty

Algorithmic

Math

Application

Reinforcement Learning with Human Preferences

Description

In order to enable humans to teach robots new skills in a natural and comfortable manner, robots need to be capable of learning from diverse types of feedback. Preferences are a common type of feedback, in which human teachers are presented with two options and asked to select the better one. In recent years, preference learning has demonstrated that it is possible to successfully teach robots new skills using this method. However, most of this work has been evaluated in simulation with simulated feedback. In order to generate this feedback, we often resort to using proxy reward functions, which have been shown to not be the best model of human behavior. Recent studies have shown that different factors need to be considered in order to better model human behavior [1].
Collecting real human feedback would enable better analysis of existing preference learning methods and provide valuable insights to adapt them so that they perform better in more realistic situations, even beyond the lab setup.

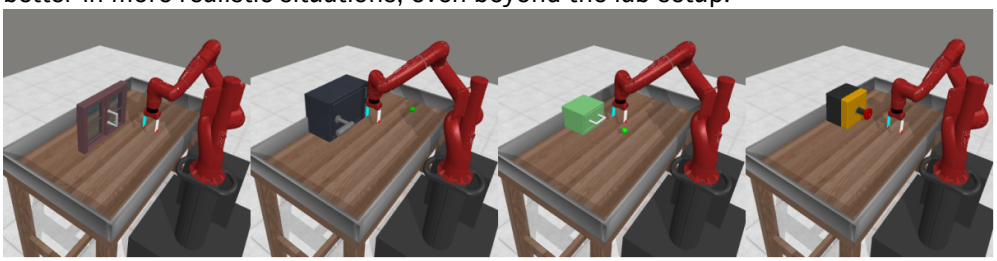

Figure 1: Different tasks from metaworld benchmark [3]

In this project, we want to investigate how humans interact with simulated and real robots in manipulations tasks. The goal is to extend an existing method for learning from preferences [2] that will efficiently collect human feedback in simulated and physical robot manipulation tasks. Collected data will be analyzed and compared to different simulated human profiles.

Tasks

The tasks in this project will involve:

- Literature Research: Getting familiar with preference learning.
- Implementation and data collection in simulation: Design and implement an extention of [2] for collecting data on simulated robot manipulation tasks. Collect and analyze collected human feedback data.
- Implementation and data collection on a real robot: Transfer and adapt the extension from the previous task to a real robot. Collect and analyze collected data.

References

[1] Kimin Lee, Laura Smith, Anca Dragan, and Pieter Abbeel. B-pref: Benchmarking preference-based reinforcement learning. *NeurIPS Datasets and Benchmarks Track*, 2021.

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

[3] Tianhe Yu, Deirdre Quillen, Zhanpeng He, Ryan Julian, Avnish Narayan, Hayden Shively, Adithya Bellathur, Karol Hausman, Chelsea Finn, and Sergey Levine. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. *Conference on Robot Learning (CoRL)*, 2019.