Karlsruhe Institute of Technology

Autonome Lernende Roboter (ALR)
Prof. Gerhard Neumann

Project Type

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

Supervisors

☞ Fabian Otto
@ fabian.otto@bosch.com
☞ Onur Celik
@ celik@kit.edu

Difficulty

Algorithmic

Math

Application

All You Need is State-Dependent Deep Reinforcement Learning?

Description

Reinforcement Learning (RL) has seen significant performance improvements in recent years and was able to solve sophisticated learning tasks. Recent algorithmic advances [2] leverage state-dependent exploration, i. e., for Gaussian parametrizations, the covariance of the policy is also state-dependent. However, for standard benchmarking tasks, such as from the Deepmind control suite [3] (Fig. 1) or the OpenAI gym environments [1] (Fig. 2), these advances in exploration did not show many benefits. Hence, this prompts the question if the exploration problem in these environments is easily manageable.
The aim of this thesis is to investigate, when state-dependent exploration matters. To this extend, we will first examine the environments from the Deepmind control suite and OpenAI gym and consequently design robot-related tasks for which the exploration is a major factor to solve the task.
The student will get familiar with and directly apply state-of-the-art methods in deep RL as well as investigates different exploration strategies. Python knowledge is highly recommended.

Figure 1: Deepmind Humanoid task [3].

Figure 2: OpenAI Ant task [1].

Tasks

The tasks in this project will involve:

- Literature review: Get to know the RL algorithms and their exploration strategies.
- Benchmarking Tasks: Get to know the tasks from Deepmind control suite [3] and OpenAI Gym [1].
- Creativity: Design environments for which state-dependent exploration matters.
- Evaluation: Evaluate different RL algorithms with state-dependent and state-independent exploration on those tasks.

References

[1] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym, 2016.

[2] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *arXiv preprint arXiv:1801.01290*, 2018.

[3] Yuval Tassa, Saran Tunyasuvunakool, Alistair Muldal, Yotam Doron, Siqi Liu, Steven Bohez, Josh Merel, Tom Erez, Timothy Lillicrap, and Nicolas Heess. dm_control: Software and tasks for continuous control, 2020.