Karlsruhe Institute of Technology

Autonome Lernende Roboter (ALR)
Prof. Gerhard Neumann

Project Type

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

Supervisors

☞ Philipp Becker
@ philipp.becker@kit.edu
☞ Onur Celik
@ celik@kit.edu

Difficulty

Algorithmic

Math

Application

Actor-Critic Methods for Robotics

Description

As trials on real robots and realistic simulators are expensive there is a big need for sample efficient approaches to Reinforcement Learning in robotics. One possible way to achieve this sample efficiency are actor-critic methods. Two recent deep actor-critic methods, based different approaches to the same idea, Maximum A-Posteriori Policy Optimization (MPO) [1] and the Soft Actor-Critic (SAC) [3] have been introduced.
Yet, the authors used different evaluation suites (MPO: DeepMind Control Suite, SAC: OpenAI Gym) and neither compared the approaches to one another. Additionally, both standard test suits are not known for particularly realistic dynamics, thus the evaluation is not very meaningful with regard to the approaches performances in a realistic robotic scenario.
The aim of this thesis is to first re-implement MPO and compare it to SAC on a standard benchmark suite. A very good implementation of SAC is available [a]. For the evaluation on realistic robotic tasks an existing pybullet [2] simulation framework should be extended. Several object manipulation tasks should be realized and both, SAC and MPO, should be evaluated on those.

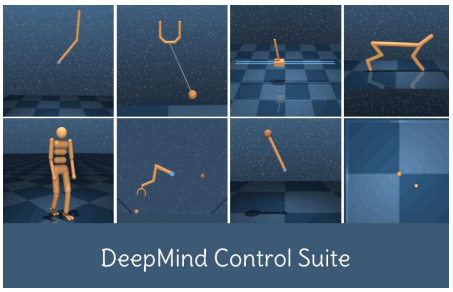

Figure 1: A snapshot from standard benchmark tasks from the deepmind control suite.

[a] https://stable-baselines.readthedocs.io/en/master/modules/sac.html

Tasks

The tasks in this project will involve:

- Implement Maximum A-Posterior Policy Optimization [1]
- Compare to Soft-actor-Critic [3] on standard benchmark suite.
- Evaluate the different components of the algorithms using ablation studies
- Extent to realistic robotic task on existing pybullet [2] simulation framework

References

[1] Abbas Abdolmaleki, Jost Tobias Springenberg, Yuval Tassa, Remi Munos, Nicolas Heess, and Martin Riedmiller. Maximum a posteriori policy optimisation. *arXiv preprint arXiv:1806.06920*, 2018.

[2] Erwin Coumans and Yunfei Bai. Pybullet, a python module for physics simulation for games, robotics and machine learning. `http://pybullet.org`, 2016–2019.

[3] 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.