


Autonome Lernende Roboter (ALR)

Prof. Gerhard Neumann

Project Type _____

- Master Thesis
- Bachelor Thesis
- Research Project

Supervisors _____

-  Philipp Becker
-  philipp.becker@kit.edu
-  Fabian Otto
-  Fabian.Otto@de.bosch.com
-  Hanna Ziesche
-  Hanna.Ziesche@de.bosch.com

Difficulty _____

Algorithmic



Math



Application



Improved Trust Regions for Adversarial Imitation Learning

Description

Imitation Learning is an important approach for intuitive teaching and programming of robots and autonomous systems. The benefit over standard reinforcement learning (RL) is, that specifying an explicit reward is not necessary, and a desired behavior can be learned by simply copying the behavior of an expert. A recent family of imitation learning approaches is based on the idea of generative adversarial networks [1] (and various others). Intuitively, a classifier is tasked with distinguishing behavior from the expert and an RL agent and the RL agent is tasked with fooling the discriminator. Yet, due to the adversarial nature those approaches are quite unstable and the schedule of classifier and agent updates needs to be designed with great care. On the agent side, previous works used RL approaches capable of controlling updates using trust regions, such as PPO [3], to improve stability. While the underlying idea is principled, both for standard RL as well as for the adversarial imitation learning setting considered here, these approaches come with various practical downsides - they can be computationally expensive, rely on simplifying and sometimes even unjustified approximations, and are sensitive to seemingly irrelevant implementation details.

We recently introduced a novel approach to such trust regions [2], based on differentiable projections. Compared with PPO, those match or improve performance on standard RL tasks, while being more robust to implementation details and allowing for more fine grained control of updates. The goal of this thesis is to explore the benefit of the trust-region projections for generative adversarial imitation learning. Due to the improved control of the updates and the general stability our projections, this should contribute to the stability of the training, improving convergence, sample efficiency, and performance.

Tasks

The tasks in this project will involve:

- Literature Review: Getting familiar with the basics and reviewing recent adversarial imitation learning literature.
- Implementation: Reimplementing the existing adversarial imitation learning approaches and augmenting the implementation with the trust-region layers.
- Evaluation: Selecting (and potentially implementing) suitable benchmark tasks for the IRL setting. Evaluating performance and stability in contrast to existing methods.

Qualifications

- Background in computer science, mathematics, physics, or similar.
- Experience with basics in machine learning (e. g. deep learning, (Lagrange) optimization, reinforcement learning)
- Experience with programming in Python (PyTorch is a plus).

References

- [1] Jonathan Ho and Stefano Ermon. Generative adversarial imitation learning. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29, pages 4565–4573. Curran Associates, Inc., 2016.
- [2] Fabian Otto, Philipp Becker, Ngo Anh Vien, Hanna Carolin Ziesche, and Gerhard Neumann. Differentiable Trust Region Layers for Deep Reinforcement Learning. In *International Conference on Learning Representations*, 2021.
- [3] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms. In *arXiv preprint*, 2017.