PADERBORN UNIVERSITY

Department of Electrical Engineering and
Information Technology
Automatic Control Group
https://en.ei.uni-paderborn.de/de/rat/

MASTER PROJECT ET/CE/ESE

Multi-Agent Reinforcement Learning with unkown Players

Background

Multi-Agent Reinforcement Learning (MARL) is a powerful framework to model complex decision-making problems of interacting systems [2]. Agents have to learn from local observations and rewards (feedback) to optimize their joint actions taken in an environment. The resulting feedback loop is shown in Figure 1.

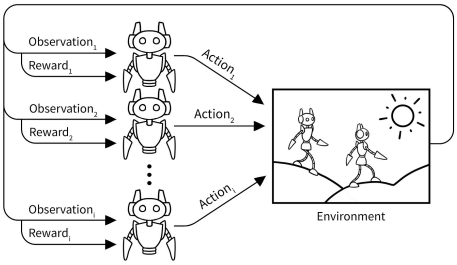

Figure 1: Multi-Agent RL Loop © Justin Terry 2021

A critical assumption for success of MARL in an environment is that all agents are trained together (either centralized or decentralized with some feedback) and then the **same agents** are paired together **during testing and deployment** [3]. However, what happens if an agent that was not part of the training process is paired with the trained agents? This scenario is illustrated Figure 2. A strategy that might have worked well for the trained agents, will likely not work well when paired with an agent that was not part of the training.

Project Goal

The goal of this project is to test Off-Belief Learning [1] for MARL with continuous control in the multi-particle environment from OpenAI [2]. Off-Belief Learning enables agents to learn a conservative strategy that ideally should work well when agents are paired with any possible other unknown players. In this project students should extend [1] to a continuous control scenario and test it with pre-trained agents from known MARL methods.

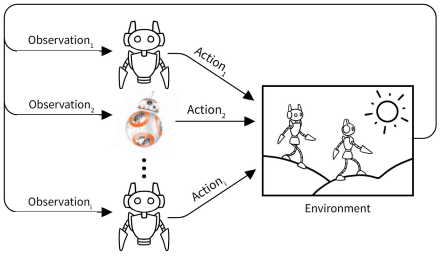

Figure 2: Multi-Agent RL Loop with new Agent

Intermediate Goals

- Literature review of Multi-Agent RL methods and Off-Belief Learning.
- Familiarize with OpenAI multi-particle environment.
- Implement and train different MARL agents.
- Design an Off-Belief Learning Algorithm for continuous control.
- Implement the Off-Belief Learning Algorithm and test it with the previously trained agents (as paired agents).
- Final presentation.

Knowledge

Required:

- Good programming skills (Python).
- Machine learning (gradient based optimization of neural networks), basic system theory.

Not required, but a big +:

- Basic elements of Reinforcement Learning (i.e. explaining an MDP).

References

[1] H. Hu, A. Lerer, B. Cui, L. Pineda, N. Brown, and J. Foerster. Off-belief learning. In *International Conference on Machine Learning*, pages 4369–4379. PMLR, 2021.

[2] R. Lowe, Y. WU, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

[3] A. Redder, A. Ramaswamy, and H. Karl. 3dpg: Distributed deep deterministic policy gradient algorithms for networked multi-agent systems. *https://arxiv.org/abs/2201.00570*, 2023.

Contact: Adrian Redder, (aredder@mail.upb.de), Room: P1.7.09.2