Universität Paderborn — Fakultät EIM
Fachgebiet Rechnernetze
http://www.upb.de/cs/cn/

Master Thesis

Simulative analysis of multi-agent learning algorithms under lossy & delaying communication

Background

Decentralised decision making with centralised training is a common paradigm for multi-agent learning problems [1]. Usually it is assumed that a simulator of the environment is used, such that local informations (e.g. local states or local policies) are globally available. This assumption is unrealistic for online learning scenarios with physically distributed agents such as those that arise in robotics. In general, exchanging local policies that are potentially parameterised by millions of real values will result in a communication delay even if high data rate communication links are available.

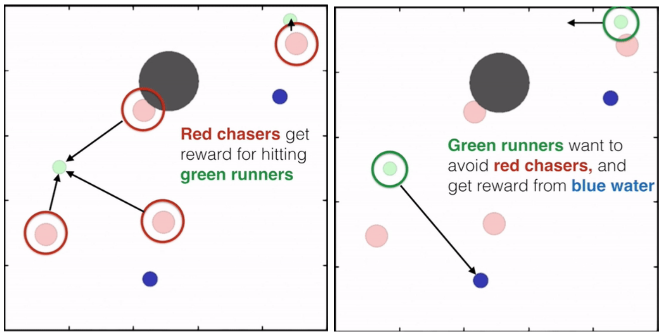

Figure 1: A competitive particle environment used in [1].

Intermediate Goals

- Literature review of multi-agent deep reinforcement learning algorithms.
- Familiarise with openAI gym particle environment and combine it with illustrative communication environments.
- Analyse the effect of the communication environments on the convergence behaviour of [1] and [2].

Thesis Goals

In [2] we proposed an algorithm for continuous decision spaces as an alternative algorithm the one given in [1]. Theoretically, we have shown that both behave asymptotically identical. However their behaviour in finite time is significantly different. Moreover, we extended both algorithms to fully decentralised training to address the problem of only locally available information as discussed above. Specifically, we formulated the algorithm in terms of potentially delayed data received via a communication network. Lossy communication induces that agents have less and "more outdated" data for training in their local experience replays. Moreover, agents can only use the outdated strategies (in form of policies) of other agents. Both of these Age of Information (AoI) effects will affect the convergence behaviour of the multi-agent learning algorithms.

The goal of this thesis is to compare the influence of the above AoI effects on the two multi-agent learning algorithms. We are interested in the rate of convergence of both algorithms as well as how they will behave differently over a finite time horizon. The objective is to perform a simulative analysis of these properties.

Knowledge

Required:

- Good programming skills (Python)
- Machine Learning I

Not required, but a big +:

- Machine Learning II and some basic knowledge in Reinforcement Learning

References

[1] R. Lowe, Y. WU, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[2] A. Redder, A. Ramaswamy, and H. Karl. A theoretical perspective of multi-agent actor-critic algorithms. 2021 (Under review).

Contact: Adrian Redder, aredder@mail.upb.de, Room: O3.170