

Cooperation in Multi-Agent Reinforcement Learning

Advances in Multi-Agent Reinforcement Learning (MARL) have led to the development of efficient agents capable of solving practical tasks. A key aspect of MARL is cooperation, wherein agents must collaborate with each other to achieve complex objectives. The need for cooperative agents is further highlighted in the case of partially-observed settings wherein agents have access to limited observations from the environment.

The work aims to explore the cooperative setup of MARL under partial-observability from a Game Theoretic viewpoint. The central objective of the project is to understand the behavior of agents in environments with large action spaces and spurious opponents. The work aims to achieve this objective by implementing and assessing state-of-the-art model-free Q-learning algorithms in MARL; namely QMIX, QMIX-SMiRL, Value Decomposition Networks (VDNs) and Independent Q-Learning (IQL). Algorithms will be simulated in SMAC's StarCraft II benchmark which consists of partially-observable micromanagement scenarios with variable number of opponents and difficulty levels. Comparison between algorithms will be drawn on the basis of payoffs received from the environment, optimization of the cost function, exploration of agent's strategies and sample-efficient learning of the optimal strategy.

While state-of-the-art model-free methods have proven to be successful in learning optimal strategies, these make use of a larger number of samples in comparison to model-based methods. The gap in data-efficiency arises as a result of the restricted access to dynamics of the environment. The project, in addition to its primary objectives, aims to bridge the gap of data-efficient learning by incorporating information about the dynamics in the cost function of model-free agents. Stochastic dynamics such as spurious states corresponding to surprise can be minimized by utilizing the common information across agents.

Recent advances in MARL have focussed on cooperation between agents in a data-efficient manner. The project aims to deliver outcomes which would aid in better intuition about the current developments and their pivotal role in improving stochastic learning from a novel standpoint.

References