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 [1]. A key aspect of MARL is cooperation [2], 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 [3, 4]. 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 [5] in MARL; namely QMIX [6], QMIX-SMiRL [7], Value Decomposition Networks (VDNs) [8] and Independent Q-Learning (IQL) [9]. Algorithms will be simulated in SMAC's StarCraft II [10] 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

[1] Lucian Buşoniu, Robert Babuška, and Bart De Schutter. Multi-agent reinforcement learning: An overview. In *Innovations in multi-agent systems and applications-1*, pages 183–221. Springer, 2010.

[2] Liviu Panait and Sean Luke. Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-Agent Systems*, 11(3):387–434, November 2005.

[3] Yoav Shoham and Kevin Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations.* Cambridge University Press, 2008.

[4] Ann Nowé, Peter Vrancx, and Yann-Michaël De Hauwere. Game theory and multi-agent reinforcement learning. In *Reinforcement Learning*, pages 441–470. Springer, 2012.

[5] Jakob N Foerster. *Deep multi-agent reinforcement learning*. PhD thesis, University of Oxford, 2018.

[6] Tabish Rashid, Mikayel Samvelyan, Christian Schroeder De Witt, Gregory Farquhar, Jakob Foerster, and Shimon Whiteson. Qmix: Monotonic value function factorisation for deep multi-agent reinforcement learning. *arXiv preprint arXiv:1803.11485*, 2018.

[7] Glen Berseth, Daniel Geng, Coline Devin, Dinesh Jayaraman, Chelsea Finn, and Sergey Levine. Smirl: Surprise minimizing rl in entropic environments. 2019.

[8] Peter Sunehag, Guy Lever, Audrunas Gruslys, Wojciech Marian Czarnecki, Vinicius Zambaldi, Max Jaderberg, Marc Lanctot, Nicolas Sonnerat, Joel Z Leibo, Karl Tuyls, et al. Value-decomposition networks for cooperative multi-agent learning. *arXiv preprint arXiv:1706.05296*, 2017.

[9] Ming Tan. Multi-agent reinforcement learning: Independent vs. cooperative agents. In *Proceedings of the tenth international conference on machine learning*, pages 330–337, 1993.

[10] Mikayel Samvelyan, Tabish Rashid, Christian Schroeder de Witt, Gregory Farquhar, Nantas Nardelli, Tim GJ Rudner, Chia-Man Hung, Philip HS Torr, Jakob Foerster, and Shimon Whiteson. The starcraft multi-agent challenge. *arXiv preprint arXiv:1902.04043*, 2019.