M2 internship: Reinforcement learning techniques for Multi-agent path finding with imperfect information

Jilles Dibangoye, Ocan Sankur, François Schwarzentruber

September 26, 2022

In missions such as search and rescue, nuclear plant decommissioning, automated farming, it is convenient to use several agents (robots, UAVs). The core problem is called multi-agent path finding, in which each agent starts in an initial position and has to reach a final target position, without colliding with the other agents. We face the issues of cooperation in a imperfect information setting. The problem has been partially defined in [QSS21], where uncertainty on the map is considered. In this setting, agents start moving in the graph but they are uncertain about the presence of some of the edges. They discover the presence or absence of these edges as they move in the graph, and communicate this information with each other in order to build an optimal execution. Figure 1 shows a graph where nodes are positions; the edges marked with ? represent the initial uncertainty of the agents.

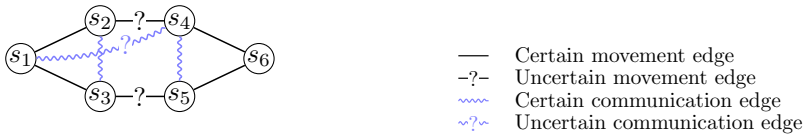

Figure 1: Graph of a given environement.

As shown in [QSS21], the problem is algorithmically challenging. As the agents need to adapt their strategies to the environment, it is natural to consider reinforcement learning (RL) for solving the multi-agent path finding.

In this M2 internship, we will design RL solutions for multi-agent path finding in the imperfect information case. We will gradually augment the level of difficulty.

1. First we will study the simplest setting in which all agents perfectly and instantly communicate. The sole source of uncertainty is the outcome of an action. We will directly use standard techniques from [SB98].
2. Second, we will introduce delay in the communication (see [SOV08] and [OS12]).
3. Finally, we will challenge the uncertainty on the map as proposed in [QSS21]. For that, we will use techniques from [DABC16] and [BDM+18].

The M2 internship will be located in Irisa, Rennes and will be co-supervised by Jilles Dibangoye (INSA Lyon), Ocan Sankur (IRISA), François Schwarzentruber (IRISA).

This internship can lead to a PhD funded by the ANR project epiRL (epistemic reinforcement learning).

References

[BDM+18] Guillaume Bono, Jilles Steeve Dibangoye, Laëtitia Matignon, Florian Pereyron, and Olivier Simonin. Cooperative multi-agent policy gradient. In Michele Berlingerio, Francesco Bonchi, Thomas Gärtner, Neil Hurley, and Georgiana Ifrim, editors, *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2018, Dublin, Ireland, September 10-14, 2018, Proceedings, Part I*, volume 11051 of *Lecture Notes in Computer Science*, pages 459–476. Springer, 2018.

[DABC16] Jilles Steeve Dibangoye, Christopher Amato, Olivier Buffet, and François Charpillet. Optimally solving dec-pomdps as continuous-state mdps. *J. Artif. Intell. Res.*, 55:443–497, 2016.

[OS12] Frans Adriaan Oliehoek and Matthijs T. J. Spaan. Tree-based solution methods for multiagent pomdps with delayed communication. In Jörg Hoffmann and Bart Selman, editors, *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, July 22-26, 2012, Toronto, Ontario, Canada*. AAAI Press, 2012.

[QSS21] Arthur Queffelec, Ocan Sankur, and François Schwarzentruber. Planning for connected agents in a partially known environment. In Luiza Antonie and Pooya Moradian Zadeh, editors, *Proceedings of the 34th Canadian Conference on Artificial Intelligence, Canadian AI 2021, online, May 2021*. Canadian Artificial Intelligence Association, 2021.

[SB98] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning - an introduction*. Adaptive computation and machine learning. MIT Press, 1998.

[SOV08] Matthijs T. J. Spaan, Frans A. Oliehoek, and Nikos Vlassis. Multiagent planning under uncertainty with stochastic communication delays. In Jussi Rintanen, Bernhard Nebel, J. Christopher Beck, and Eric A. Hansen, editors, *Proceedings of the Eighteenth International Conference on Automated Planning and Scheduling, ICAPS 2008, Sydney, Australia, September 14-18, 2008*, pages 338–345. AAAI, 2008.