

REFERENCES

- [1] Christopher Amato and Frans A Oliehoek. 2015. Scalable Planning and Learning for Multiagent POMDPs. In *29th AAAI Conference on Artificial Intelligence*.
- [2] Thomas Anthony, Zheng Tian, and David Barber. 2017. Thinking Fast and Slow with Deep Learning and Tree Search. In *Advances in Neural Information Processing Systems*.
- [3] Craig Boutilier. 1996. Planning, Learning and Coordination in Multiagent Decision Processes. In *Proceedings of the 6th conference on Theoretical aspects of rationality and knowledge*. Morgan Kaufmann Publishers Inc.
- [4] Daniel Claes, Frans Oliehoek, Hendrik Baier, and Karl Tuyls. 2017. Decentralised Online Planning for Multi-Robot Warehouse Commissioning. In *Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS.
- [5] Daniel Claes, Philipp Robbel, Frans A Oliehoek, Karl Tuyls, Daniel Hennes, and Wiebe Van der Hoek. 2015. Effective Approximations for Multi-Robot Coordination in Spatially Distributed Tasks. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS.
- [6] Jakob Foerster, Gregory Farquhar, Triantafyllos Afouras, Nantas Nardelli, and Shimon Whiteson. 2018. Counterfactual Multi-Agent Policy Gradients. *32th AAAI Conference on Artificial Intelligence* (2018).
- [7] Jayesh K Gupta, Maxim Egorov, and Mykel Kochenderfer. 2017. Cooperative Multi-Agent Control using Deep Reinforcement Learning. In *International Conference on Autonomous Agents and Multiagent Systems*. Springer.
- [8] Erwan Lecarpentier, Guillaume Infantes, Charles Lesire, and Emmanuel Rachelson. 2018. Open Loop Execution of Tree-Search Algorithms. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. IJCAI Organization, 2362–2368. <https://doi.org/10.24963/ijcai.2018/327>
- [9] Frans A Oliehoek, Julian FP Kooij, and Nikos Vlassis. 2008. The Cross-Entropy Method for Policy Search in Decentralized POMDPs. *Informatica* 32, 4 (2008), 341–357.
- [10] Diego Perez Liebana, Jens Dieskau, Martin Hunermund, Sanaz Mostaghim, and Simon Lucas. 2015. Open Loop Search for General Video Game Playing. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*. ACM.
- [11] Thomy Phan, Lenz Belzner, Thomas Gabor, and Kyrill Schmid. 2018. Leveraging Statistical Multi-Agent Planning with Emergent Value Function Approximation. In *Proceedings of the 17th Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS.
- [12] Thomy Phan, Lenz Belzner, Marie Kiermeier, Markus Friedrich, Kyrill Schmid, and Claudia Linnhoff-Popien. 2018. Memory Bounded Open-Loop Planning in Large POMDPs Using Thompson Sampling. *33th AAAI Conference on Artificial Intelligence* (2018).
- [13] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. 2017. Mastering the Game of Go without Human Knowledge. *Nature* 550, 7676 (2017).
- [14] Richard S Sutton. 1988. Learning to Predict by the Methods of Temporal Differences. *Machine learning* 3, 1 (1988).
- [15] Richard S Sutton and Andrew G Barto. 1998. *Introduction to Reinforcement Learning*. Vol. 135. MIT Press Cambridge.
- [16] Ming Tan. 1993. Multi-Agent Reinforcement Learning: Independent versus Cooperative Agents. In *Proceedings of the 10th International Conference on International Conference on Machine Learning*. Morgan Kaufmann Publishers Inc.
- [17] Rene Vidal, Omid Shakernia, H Jin Kim, David Hyunchul Shim, and Shankar Sastry. 2002. Probabilistic Pursuit-Evasion Games: Theory, Implementation, and Experimental Evaluation. *IEEE transactions on robotics and automation* 18, 5 (2002).
- [18] Ari Weinstein and Michael L Littman. 2013. Open-Loop Planning in Large-Scale Stochastic Domains. In *27th AAAI Conference on Artificial Intelligence*.