

- [31] Ranjit Nair, Milind Tambe, Makoto Yokoo, David Pynadath, and Stacy Marsella. 2003. Taming decentralized POMDPs: Towards efficient policy computation for multiagent settings. In *Proceedings of the 18th international joint conference on Artificial intelligence*. Morgan Kaufmann Publishers Inc., Acapulco, Mexico, 705–711.
- [32] Frans Adriaan Oliehoek. 2013. Sufficient Plan-Time Statistics for Decentralized POMDPs. In *Proceedings of the 23rd international joint conference on Artificial Intelligence*. AAAI Press, Beijing, China, 302–308.
- [33] Frans A Oliehoek and Christopher Amato. 2016. *A concise introduction to decentralized POMDPs* (1st ed.). Springer.
- [34] Frans A Oliehoek, Matthijs TJ Spaan, Bas Terwijn, Philipp Robbel, and João V Messias. 2017. The MADP toolbox: an open source library for planning and learning in (multi-) agent systems. *The Journal of Machine Learning Research* 18, 1 (2017), 3112–3116.
- [35] Frans A Oliehoek, Matthijs TJ Spaan, and Nikos Vlassis. 2008. Optimal and approximate Q-value functions for decentralized POMDPs. *Journal of Artificial Intelligence Research* 32 (2008), 289–353.
- [36] S. Omidshafiei, D.-K. Kim, M. Liu, G. Tesaro, M. Riemer, C. Amato, M. Campbell, and J. P. How. 2018. Learning to Teach in Cooperative Multiagent Reinforcement Learning. *ArXiv e-prints* (May 2018). arXiv:cs.MA/1805.07830
- [37] Shayegan Omidshafiei, Jason Pazis, Christopher Amato, Jonathan P How, and John Vian. 2017. Deep Decentralized Multi-task Multi-Agent Reinforcement Learning under Partial Observability. In *Proceedings of the 34th International Conference on Machine Learning*, Vol. 70. PMLR, Sydney, Australia, 2681–2690.
- [38] Aravind Rajeswaran, Vikash Kumar, Abhishek Gupta, Giulia Vezzani, John Schulman, Emanuel Todorov, and Sergey Levine. 2018. Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations. In *Proceedings of Robotics: Science and Systems*. Pittsburgh, Pennsylvania, USA. <https://doi.org/10.15607/RSS.2018.XIV.049>
- [39] Sarvapali D Ramchurn, Feng Wu, Wenchao Jiang, Joel E Fischer, Steve Reece, Stephen Roberts, Tom Rodden, Chris Greenhalgh, and Nicholas R Jennings. 2016. Human-agent collaboration for disaster response. *Autonomous Agents and Multi-Agent Systems* 30, 1 (2016), 82–111.
- [40] Stéphane Ross and Drew Bagnell. 2010. Efficient reductions for imitation learning. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. 661–668.
- [41] Stéphane Ross, Geoffrey Gordon, and Drew Bagnell. 2011. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. 627–635.
- [42] John Schulman, Philipp Moritz, Sergey Levine, Michael Jordan, and Pieter Abbeel. 2016. High-Dimensional Continuous Control Using Generalized Advantage Estimation. In *International Conference on Learning Representations*. San Juan, Puerto Rico.
- [43] Sven Seuken and Shlomo Zilberstein. 2007. Improved memory-bounded dynamic programming for decentralized POMDPs. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*. AUAI Press, Vancouver, BC, Canada, 344–351.
- [44] Kaushik Subramanian, Charles L Isbell Jr, and Andrea L Thomaz. 2016. Exploration from demonstration for interactive reinforcement learning. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 447–456.
- [45] Wen Sun, Arun Venkatraman, Geoffrey J Gordon, Byron Boots, and J Andrew Bagnell. 2017. Deeply aggravated: Differentiable imitation learning for sequential prediction. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 3309–3318.
- [46] Daniel Szer, François Charpillet, and Shlomo Zilberstein. 2005. MAA*: a heuristic search algorithm for solving decentralized POMDPs. In *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*. AUAI Press, Edinburgh, Scotland, 576–583.
- [47] Ardi Tampuu, Tabet Matiisen, Dorian Kodelja, Ilya Kuzovkin, Kristjan Korjus, Juhan Aru, Jaan Aru, and Raul Vicente. 2017. Multiagent cooperation and competition with deep reinforcement learning. *PLoS one* 12, 4 (2017), e0172395.
- [48] Hado Van Hasselt, Arthur Guez, and David Silver. 2016. Deep Reinforcement Learning with Double Q-Learning. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. AAAI Press, Phoenix, Arizona, USA, 2094–2100.
- [49] Ruben Van Parys, Maarten Verbandt, Marcus Kotzé, Jan Swevers, Herman Bruyninckx, Johan Philips, and Goele Pipeleers. 2018. Flexible Multi-Agent System for Distributed Coordination, Transportation & Localisation. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, Stockholm, Sweden, 1832–1834.
- [50] Shiyong Wang, Jiafu Wan, Daqiang Zhang, Di Li, and Chunhua Zhang. 2016. Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. *Computer Networks* 101 (2016), 158–168.
- [51] Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning* 8, 3-4 (1992), 229–256.
- [52] E. Zhan, S. Zheng, Y. Yue, L. Sha, and P. Lucey. 2018. Generative Multi-Agent Behavioral Cloning. *ArXiv e-prints* (March 2018). arXiv:1803.07612
- [53] X. Zhang and H. Ma. 2018. Pretraining Deep Actor-Critic Reinforcement Learning Algorithms With Expert Demonstrations. *ArXiv e-prints* (Jan. 2018). arXiv:cs.AI/1801.10459