

- [29] OSBAND I, VAN ROY B, WEN Z. Generalization and exploration via randomized value functions[J]. Proceedings of the 33rd International Conference on International Conference on Machine Learning, 2014, 48(1): 2377-2386.
- [30] MUNOS R, STEPLETON T, HARUTYUNYAN A, et al. Safe and efficient off-policy reinforcement learning[C]// Advances in Neural Information Processing Systems. Spain: NIPS Press, 2016: 1054-1062.
- [31] FRANÇOIS-LAVET V, FONTENEAU R, ERNST D. How to discount deep reinforcement learning: Towards new dynamic strategies[C]// Proceedings of the Workshops at the Advances in Neural Information Processing Systems. Montreal, Canada: NIPS Press, 2015: 1107-1160.
- [32] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning, U. S. Patent Application 15/217,758[P]. 2017-1-26.
- [33] MNIH V, BADIA A P, MIRZA M, et al. Asynchronous methods for deep reinforcement learning[C]// International Conference on Machine Learning. New York City: ICML press, 2016: 1928-1937.
- [34] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[J/OL]. <https://arxiv.org/abs/1707.06347>.
- [35] HEES N, SRIRAM S, LEMMON J, et al. Emergence of locomotion behaviors in rich environments[J/OL]. <https://arxiv.org/abs/1707.02286>.
- [36] FOERSTER J, NARDELLI N, FARQUHAR G, et al. Stabilizing experience replay for deep multi-agent reinforcement learning[J]. International Conference on Machine Learning, 2017, 70(3): 1146-1155.
- [37] CIOSEK K, WHITESON S. Offer: Off environment reinforcement learning[J]. AAAI Conference on Artificial Intelligence, 2017.
- [38] TESAURO G. Extending q-learning to general adaptive multi-agent systems[J]. Advances in Neural Information Processing Systems, 2004, 16(4): 871-878.
- [39] TAN M. Multi-Agent Reinforcement Learning, Independent vs. Cooperative Agents[C]// Proceedings of the Tenth International Conference on Machine Learning. MA, USA: ICML Press, 1993: 330-337.
- [40] SHOHAM Y, LEYTON K. Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations[M]. New York: Cambridge University Press, 2009.
- [41] ZAWADZKI E, LIPSON A, LEYTON K. Empirically evaluating multiagent learning algorithms[J/OL]. <https://arxiv.org/abs/1401.8074>.
- [42] YANG Y, LUO R, LI M, et al. Mean Field Multi-Agent Reinforcement Learning [J/OL]. <https://arxiv.org/abs/1802.05438>.
- [43] PALMER G, TUYLS K, BLOEMBERGEN D, et al. Lenient multi-agent deep reinforcement learning[C]// Proceedings of the 17th International Conference on Autonomous Agents and Multi-agent Systems. Swede: AAMAS press, 2018: 443-451.
- [44] ZHENG Y, MENG Z, HAO J, et al. Weighted Double Deep Multiagent Reinforcement Learning in Stochastic Cooperative Environments[C]// Pacific Rim International Conference on Artificial Intelligence. Springer, Cham: PRICAI press, 2018: 421-429.
- [45] TAMPUU A, MATIISEN T, KODELJA D, et al. Multiagent cooperation and competition with deep reinforcement learning [J]. Plus One, 2017, 12(4): e0172395.
- [46] SONG J, REN H, SADIGH D, et al. Multi-agent generative adversarial imitation learning [J/OL]. <https://arxiv.org/abs/1807.09936>.
- [47] WAI H T, YANG Z, WANG Z, et al. Multi-Agent Reinforcement Learning via Double Averaging Primal-Dual Optimization [J/OL]. <https://arxiv.org/abs/1806.00877>.
- [48] ABOUEHAF M, GUEAIEB W. Multi-agent reinforcement learning approach based on reduced value function approximations[C]// 2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS). Canada: IEEE Press, 2017: 111-116.
- [49] QI S, ZHU S C. Intent-aware Multi-agent Reinforcement Learning[J/OL]. <https://arxiv.org/abs/1803.02018>.
- [50] RAILEANU R, DENTON E, SZLAM A, et al. Modeling Others using Oneself in Multi-Agent Reinforcement Learning[J/OL]. <https://arxiv.org/abs/1802.09640>.
- [51] RABINOWITZ N, PERBET F, SONG H, et al. Machine Theory of Mind[J/OL]. <https://arxiv.org/abs/1802.07740>.
- [52] OMIDSHAFIEI S, KIM D, LIU M, et al. Learning to Teach in Cooperative Multiagent Reinforcement Learning [J/OL]. <https://arxiv.org/abs/1805.07830>.
- [53] GU S, LILLICRAP T, SUTSKEVER I, et al. Continuous deep q-learning with model-based acceleration [C]// International Conference on Machine Learning. New York City: ICML Press, 2016: 2829-2838.
- [54] DUAN Y, CHEN X, HOUTHOOFT R, et al. Benchmarking deep reinforcement learning for continuous control[C]// International Conference on Machine Learning. New York City: ICML Press, 2016: 1329-1338.
- [55] KOFINAS P, DOUNIS A I, VOUIROS G A. Fuzzy Q-Learning for multi-agent decentralized energy management in microgrids [J]. Applied Energy, 2018, 219(3): 53-67.
- [56] CHEN W, ZHOU K, CHEN C. Real-time bus holding control on a transit corridor based on multi-agent reinforcement learning [C]// 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). Brazil: IEEE Press, 2016: 100-106.
- [57] VIDHATE D A, KULKARNI P. Cooperative multi-agent reinforcement learning models (CMRLM) for intelligent traffic control[C]// 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM). India: IEEE Press, 2017: 325-331.