





## REFERENCES

- [1] N. Brown and T. Sandholm, "Superhuman ai for heads-up no-limit poker: Libratus beats top professionals," *Science*, vol. 359, no. 6374, pp. 418–424, 2018.
- [2] N. Brown, A. Lerer, S. Gross, and T. Sandholm, "Deep counterfactual regret minimization," in *Proceedings of the International Conference on Machine Learning (ICML)*, 2019, pp. 793–802.
- [3] T. Börgers and R. Sarin, "Learning through reinforcement and replicator dynamics," *Journal of Economic Theory*, vol. 77, no. 1, pp. 1–14, 1997.
- [4] J. Heinrich and D. Silver, "Deep reinforcement learning from self-play in imperfect-information games," *CoRR*, vol. abs/1603.01121, 2016. [Online]. Available: <http://arxiv.org/abs/1603.01121>
- [5] D. Hennes, D. Morrill, S. Omidshafiei, R. Munos, J. Perolat, M. Lanctot, A. Gruslys, J.-B. Lespiau, P. Parmas, E. Duenez-Guzman, and K. Tuyls, "Neural replicator dynamics," 2020.
- [6] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.
- [7] M. Kaisers and K. Tuyls, "Frequency adjusted multi-agent q-learning," in *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2010, p. 309–316.
- [8] —, "Faq-learning in matrix games: Demonstrating convergence near nash equilibria, and bifurcation of attractors in the battle of sexes," in *Proceedings of the Workshop on Interactive Decision Theory and Game Theory*, 2011, p. 36–42.
- [9] S. M. Kakade, "A natural policy gradient," in *Proceedings of the Neural Information Processing Systems conference (NeurIPS)*, vol. 14, 2001, pp. 1–8.
- [10] T. Klos, G. J. van Ahee, and K. Tuyls, "Evolutionary dynamics of regret minimization," in *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, 2010, pp. 82–96.
- [11] J. Maynard Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15–18, 1973.
- [12] L. Panait and K. Tuyls, "Theoretical advantages of lenient q-learners: An evolutionary game theoretic perspective," in *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2007.
- [13] J. Peters and S. Schaal, "Policy gradient methods for robotics," in *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2006, pp. 2219–2225.
- [14] K. Ritzberger and J. W. Weibull, "Evolutionary selection in normal-form games," *Econometrica: Journal of the Econometric Society*, pp. 1371–1399, 1995.
- [15] W. H. Sandholm, *Evolutionary Game Theory*. Springer New York, 2009, pp. 3176–3205.
- [16] —, *Evolutionary Game Theory*. Springer Berlin Heidelberg, 2017, pp. 1–38.
- [17] R. S. Sutton, D. A. McAllester, S. P. Singh, Y. Mansour *et al.*, "Policy gradient methods for reinforcement learning with function approximation," in *Proceedings of the conference on Neural Information Processing Systems (NeurIPS)*, vol. 99, 1999, pp. 1057–1063.
- [18] K. Tuyls, K. Verbeeck, and T. Lenaerts, "A selection-mutation model for q-learning in multi-agent systems," in *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2003, p. 693–700.
- [19] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev *et al.*, "Grandmaster level in starcraft ii using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, 2019.
- [20] J. Weibull, *Evolutionary game theory*. MIT Press, 1995.