



















## REFERENCES

- [1] Yasin Abbasi-Yadkori, Peter Bartlett, Kush Bhatia, Nevena Lazic, Csaba Szepesvari, and Gellert Weisz. 2019. PoliteX: Regret bounds for policy iteration using expert prediction. In *International Conference on Machine Learning*. 3692–3702.
- [2] Marc Bellemare, Joel Veness, and Michael Bowling. 2012. Sketch-based linear value function approximation. In *Advances in Neural Information Processing Systems*. 2213–2221.
- [3] A. Blum and Y. Mansour. 2007. Learning, Regret Minimization, and Equilibria. In *Algorithmic Game Theory*. Cambridge University Press, Chapter 4.
- [4] Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin. 2015. Heads-up limit hold'em poker is solved. *Science* 347, 6218 (2015), 145–149.
- [5] Noam Brown, Adam Lerer, Sam Gross, and Tuomas Sandholm. 2019. Deep Counterfactual Regret Minimization. In *Proceedings of the 36th International Conference on Machine Learning (ICML-19)*. 793–802.
- [6] Noam Brown and Tuomas Sandholm. 2018. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science* 359, 6374 (2018), 418–424.
- [7] Noam Brown and Tuomas Sandholm. 2019. Superhuman AI for multiplayer poker. *Science* 365, 6456 (2019), 885–890.
- [8] Nicolo Cesa-Bianchi and Gabor Lugosi. 2006. *Prediction, learning, and games*. Cambridge university press.
- [9] Gabriele Farina, Christian Kroer, and Tuomas Sandholm. 2019. Regret Circuits: Composability of Regret Minimizers. In *International Conference on Machine Learning*. 1863–1872.
- [10] Yoav Freund and Robert E Schapire. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences* 55, 1 (1997), 119–139.
- [11] Sam Ganzfried and Tuomas Sandholm. 2013. Action translation in extensive-form games with large action spaces: Axioms, paradoxes, and the pseudo-harmonic mapping. In *Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence*.
- [12] Amy Greenwald, Zheng Li, and Casey Marks. 2006. Bounds for Regret-Matching Algorithms. In *ISAIM*.
- [13] Amy Greenwald, Zheng Li, and Casey Marks. 2006. *Bounds for Regret-Matching Algorithms*. Technical Report CS-06-10. Brown University, Department of Computer Science.
- [14] S. Hart and A. Mas-Colell. 2000. A Simple Adaptive Procedure Leading to Correlated Equilibrium. *Econometrica* 68, 5 (2000), 1127–1150.
- [15] Johannes Heinrich and David Silver. 2016. Deep reinforcement learning from self-play in imperfect-information games. *arXiv preprint arXiv:1603.01121* (2016).
- [16] Michael Johanson, Neil Burch, Richard Valenzano, and Michael Bowling. 2013. Evaluating state-space abstractions in extensive-form games. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 271–278.
- [17] Marc Lanctot. 2013. *Monte Carlo Sampling and Regret Minimization for Equilibrium Computation and Decision-Making in Large Extensive Form Games*. Ph.D. Dissertation. Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada.
- [18] Marc Lanctot, Edward Lockhart, Jean-Baptiste Lespiau, Vinicius Zambaldi, Satyaki Upadhyay, Julien Pérolat, Sriram Srinivasan, Finbarr Timbers, Karl Tuyls, Shayegan Omidshafiei, Daniel Hennes, Dustin Morrill, Paul Muller, Timo Ewalds, Ryan Faulkner, János Kramár, Bart De Vylder, Brennan Saeta, James Bradbury, David Ding, Sebastian Borgeaud, Matthew Lai, Julian Schrittwieser, Thomas Anthony, Edward Hughes, Ivo Danihelka, and Jonah Ryan-Davis. 2019. OpenSpiel: A Framework for Reinforcement Learning in Games. *CoRR* abs/1908.09453 (2019). arXiv:cs.LG/1908.09453 <http://arxiv.org/abs/1908.09453>
- [19] Hui Li, Kailiang Hu, Zhibang Ge, Tao Jiang, Yuan Qi, and Le Song. 2018. Double neural counterfactual regret minimization. *arXiv preprint arXiv:1812.10607* (2018).
- [20] Edward Lockhart, Marc Lanctot, Julien Pérolat, Jean-Baptiste Lespiau, Dustin Morrill, Finbarr Timbers, and Karl Tuyls. 2019. Computing Approximate Equilibria in Sequential Adversarial Games by Exploitability Descent. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, 464–470. <https://doi.org/10.24963/ijcai.2019/66>
- [21] Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. 2017. DeepStack: Expert-level artificial intelligence in heads-up no-limit poker. *Science* 356, 6337 (2017), 508–513.
- [22] Dustin Morrill. 2016. *Using Regret Estimation to Solve Games Compactly*. Master's thesis. University of Alberta.
- [23] Shayegan Omidshafiei, Daniel Hennes, Dustin Morrill, Remi Munos, Julien Perolat, Marc Lanctot, Audrunas Gruslys, Jean-Baptiste Lespiau, and Karl Tuyls. 2019. Neural Replicator Dynamics. *arXiv preprint arXiv:1906.00190* (2019).
- [24] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Vedavyas Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy P. Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 7587 (2016), 484–489.
- [25] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharmashan Kumaran, Thore Graepel, et al. 2018. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362, 6419 (2018), 1140–1144.
- [26] Finnegan Southey, Michael Bowling, Bryce Larson, Carmelo Piccione, Neil Burch, Darce Billings, and Chris Rayner. 2005. Bayes's bluff: Opponent modelling in poker. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*. 550–558.
- [27] Sriram Srinivasan, Marc Lanctot, Vinicius Zambaldi, Julien Perolat, Karl Tuyls, Remi Munos, and Michael Bowling. 2018. Actor-Critic Policy Optimization in Partially Observable Multiagent Environments. In *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.). Curran Associates, Inc., 3422–3435.
- [28] Eric Steinberger. 2019. Single Deep Counterfactual Regret Minimization. *arXiv preprint arXiv:1901.07621* (2019).
- [29] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, Laurent Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander S. Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver. 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* (2019). <https://doi.org/10.1038/s41586-019-1724-z>
- [30] Bernhard Von Stengel and Françoise Forges. 2008. Extensive-form correlated equilibrium: Definition and computational complexity. *Mathematics of Operations Research* 33, 4 (2008), 1002–1022.
- [31] Kevin Waugh, Dustin Morrill, James Andrew Bagnell, and Michael Bowling. 2015. Solving games with functional regret estimation. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.
- [32] Kevin Waugh, David Schnizlein, Michael Bowling, and Duane Szafron. 2009. Abstraction pathologies in extensive games. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, 781–788.
- [33] Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. 2008. Regret minimization in games with incomplete information. In *Advances in neural information processing systems*. 1729–1736.