

REFERENCES

- [1] L. C. Baird III. Residual Algorithms: Reinforcement Learning with Function Approximation. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 30–37, 1995.
- [2] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling. The Arcade Learning Environment: An Evaluation Platform for General Agents. *Journal of Artificial Intelligence Research*, 47:253–279, 06 2013.
- [3] M. G. Bellemare, J. Veness, and M. Bowling. Investigating Contingency Awareness using Atari 2600 Games. In *Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI)*, pages 864–871, 2012.
- [4] M. G. Bellemare, J. Veness, and M. Bowling. Sketch-Based Linear Value Function Approximation. In *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, pages 2222–2230, 2012.
- [5] M. Bowling, N. Burch, M. Johanson, and O. Tammelin. Heads-up Limit Hold'em Poker is Solved. *Science*, 347(6218):145–149, January 2015.
- [6] M. Campbell, A. J. H. Jr., and F.-H. Hsu. Deep Blue. *Artificial Intelligence*, 134(1&A2):57 – 83, 2002.
- [7] A. Defazio and T. Graepel. A Comparison of Learning Algorithms on the Arcade Learning Environment. *CoRR*, abs/1410.8620, 2014.
- [8] D. A. Ferrucci, E. W. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A. Kalyanpur, A. Lally, J. W. Murdock, E. Nyberg, J. M. Prager, N. Schlaefer, and C. A. Welty. Building Watson: An Overview of the DeepQA Project. *AI Magazine*, 31(3):59–79, 2010.
- [9] A. Gionis, P. Indyk, and R. Motwani. Similarity Search in High Dimensions via Hashing. In *Proceedings of International Conference on Very Large Data Bases (VLDB)*, pages 518–529, 1999.
- [10] G. J. Gordon. Reinforcement Learning with Function Approximation Converges to a Region. In *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, pages 1040–1046, 2000.
- [11] X. Guo, S. P. Singh, H. Lee, R. L. Lewis, and X. Wang. Deep Learning for Real-Time Atari Game Play Using Offline Monte-Carlo Tree Search Planning. In *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, pages 3338–3346, 2014.
- [12] M. J. Hausknecht and P. Stone. Deep Recurrent Q-Learning for Partially Observable MDPs. *CoRR*, abs/1507.06527, 2015.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, pages 1106–1114, 2012.
- [14] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [15] Y. Liang, M. C. Machado, E. Talvitie, and M. H. Bowling. State of the Art Control of Atari Games Using Shallow Reinforcement Learning. *CoRR*, abs/1512.01563, 2015.
- [16] N. Lipovetzky, M. Ramirez, and H. Geffner. Classical Planning with Simulators: Results on the Atari Video Games. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.
- [17] M. C. Machado, S. Srinivasan, and M. Bowling. Domain-Independent Optimistic Initialization for Reinforcement Learning. In *AAAI Workshop on Learning for General Competency in Video Games*, 2015.
- [18] F. S. Melo, S. P. Meyn, and M. I. Ribeiro. An Analysis of Reinforcement Learning with Function Approximation. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 664–671, 2008.
- [19] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller. Playing Atari With Deep Reinforcement Learning. In *NIPS Deep Learning Workshop*, 2013.
- [20] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level Control through Deep Reinforcement Learning. *Nature*, 518(7540):529–533, 02 2015.
- [21] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, K. Kavukcuoglu. Asynchronous Methods for Deep Reinforcement Learning. *CoRR*, abs/1602.01783, 2016.
- [22] G. A. Rummery and M. Niranjan. On-line Q-Learning using Connectionist Systems. CUED/F-INFENG/TR 166, Cambridge University Engineering Dept., September 1994.
- [23] J. Schaeffer, N. Burch, Y. Björnsson, A. Kishimoto, M. Müller, R. Lake, P. Lu, and S. Sutphen. Checkers is Solved. *Science*, 317(5844):1518–1522, 2007.
- [24] J. Schulman, S. Levine, P. Abbeel, M. I. Jordan, and P. Moritz. Trust Region Policy Optimization. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 1889–1897, 2015.
- [25] N. Sprague. Parameter Selection for the Deep Q-learning Algorithm ((Extended Abstract)). In *Proceedings of the Multidisciplinary Conference on Reinforcement Learning and Decision Making (RLDM)*, 2015.
- [26] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.
- [27] C. Szepesvári. *Algorithms for Reinforcement Learning*. Synthesis lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool, 2010.
- [28] G. Tesauro. Temporal Difference Learning and TD-Gammon. *Communications of the ACM*, 38(3):58–68, Mar. 1995.
- [29] H. van Hasselt, A. Guez, and D. Silver. Deep Reinforcement Learning with Double Q-learning. *CoRR*, abs/1509.06461, 2015.
- [30] J. Veness, M. G. Bellemare, M. Hutter, A. Chua, and G. Desjardins. Compress and Control. In *Proceedings of the Conference on Artificial Intelligence (AAAI)*, 2015.
- [31] C. J. C. H. Watkins and P. Dayan. Technical Note: Q-Learning. *Machine Learning*, 8(3-4), May 1992.