

REFERENCES

- [1] Karl Cobbe, Christopher Hesse, Jacob Hilton, and John Schulman. 2020. Leveraging Procedural Generation to Benchmark Reinforcement Learning. arXiv:1912.01588 [cs.LG]
- [2] Karl Cobbe, Oleg Klimov, Chris Hesse, Taehoon Kim, and John Schulman. 2018. Quantifying Generalization in Reinforcement Learning. arXiv:1812.02341 [cs.LG]
- [3] Will Dabney, André Barreto, Mark Rowland, Robert Dadashi, John Quan, Marc G. Bellemare, and David Silver. 2020. The Value-Improvement Path: Towards Better Representations for Reinforcement Learning. arXiv:2006.02243 [cs.LG]
- [4] Matthew Hausknecht and Peter Stone. 2015. The Impact of Determinism on Learning Atari 2600 Games. In *AAAI Workshop on Learning for General Competency in Video Games*. Austin, Texas, USA. <http://www.cs.utexas.edu/users/ai-lab?hausknecht:aaai15>
- [5] Matteo Hessel, Joseph Modayil, Hado van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Dan Horgan, Bilal Piot, Mohammad Azar, and David Silver. 2017. Rainbow: Combining Improvements in Deep Reinforcement Learning. arXiv:1710.02298 [cs.AI]
- [6] Michael Laskin, Kimin Lee, Adam Stooke, Lerrel Pinto, Pieter Abbeel, and Aravind Srinivas. 2020. Reinforcement Learning with Augmented Data. arXiv:2004.14990 [cs.LG]
- [7] Kimin Lee, Kibok Lee, Jinwoo Shin, and Honglak Lee. 2019. Network Randomization: A Simple Technique for Generalization in Deep Reinforcement Learning. arXiv:1910.05396 [cs.LG]
- [8] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dhharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (February 2015), 529–533. <https://doi.org/10.1038/nature14236>
- [9] Roberta Raileanu, Max Goldstein, Denis Yarats, Ilya Kostrikov, and Rob Fergus. 2020. Automatic Data Augmentation for Generalization in Deep Reinforcement Learning. *arXiv preprint arXiv:2006.12862* (2020).
- [10] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. 2015. Prioritized Experience Replay. arXiv:1511.05952 [cs.LG]
- [11] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal Policy Optimization Algorithms. arXiv:1707.06347 [cs.LG]
- [12] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis. 2017. Mastering the game of Go without human knowledge. *Nature* 550, 7676 (October 2017), 354–359. <https://doi.org/10.1038/nature24270>
- [13] Kingyou Song, Yiding Jiang, Stephen Tu, Yilun Du, and Behnam Neyshabur. 2019. Observational Overfitting in Reinforcement Learning. arXiv:1912.02975 [cs.LG]
- [14] Yuval Tassa, Yotam Doron, Alistair Muldal, Tom Erez, Yazhe Li, Diego de Las Casas, David Budden, Abbas Abdolmaleki, Josh Merel, Andrew Lefrancq, Timothy Lillicrap, and Martin Riedmiller. 2018. DeepMind Control Suite. arXiv:1801.00690 [cs.AI]
- [15] Chiyuan Zhang, Oriol Vinyals, Remi Munos, and Samy Bengio. 2018. A Study on Overfitting in Deep Reinforcement Learning. arXiv:1804.06893 [cs.LG]