

REFERENCES

- [1] Pieter Abbeel, Adam Coates, and Andrew Y. Ng. 2010. Autonomous Helicopter Aerobatics through Apprenticeship Learning. *Int. J. Rob. Res.* 29, 13 (Nov. 2010), 1608–1639.
- [2] Joshua Achiam, David Held, Aviv Tamar, and Pieter Abbeel. 2017. Constrained Policy Optimization.
- [3] Eitan Altman. 1999. *Constrained Markov Decision Processes*. Taylor & Francis.
- [4] Yinlam Chow, Mohammad Ghavamzadeh, Lucas Janson, and Marco Pavone. 2017. Risk-Constrained Reinforcement Learning with Percentile Risk Criteria. *J. Mach. Learn. Res.* 18, 1 (Jan. 2017), 6070–6120.
- [5] Yan Duan, Xi Chen, Rein Houthooft, John Schulman, and Pieter Abbeel. 2016. Benchmarking Deep Reinforcement Learning for Continuous Control (*Proceedings of Machine Learning Research, Vol. 48*), Maria Florina Balcan and Kilian Q. Weinberger (Eds.). PMLR, New York, New York, USA, 1329–1338.
- [6] William Fedus, Carles Gelada, Yoshua Bengio, Marc G. Bellemare, and Hugo Larochelle. 2019. Hyperbolic Discounting and Learning over Multiple Horizons. arXiv:1902.06865 [stat.ML]
- [7] The garage contributors. 2019. Garage: A toolkit for reproducible reinforcement learning research.
- [8] Javier García and Fernando Fernández. 2012. Safe Exploration of State and Action Spaces in Reinforcement Learning. 45, 1 (Sept. 2012), 515–564.
- [9] Javier García and Fernando Fernández. 2015. A Comprehensive Survey on Safe Reinforcement Learning. *Journal of Machine Learning Research* 16, 42 (2015), 1437–1480.
- [10] Chris Gaskett. 2003. Reinforcement learning under circumstances beyond its control. (2003). <https://researchonline.jcu.edu.au/632/>
- [11] Clement Gehring and Doina Precup. 2013. Smart Exploration in Reinforcement Learning Using Absolute Temporal Difference Errors. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-Agent Systems* (St. Paul, MN, USA) (AAMAS '13). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1037–1044.
- [12] Peter Geibel and Fritz Wysotzki. 2005. Risk-Sensitive Reinforcement Learning Applied to Control under Constraints. *J. Artif. Int. Res.* 24, 1 (July 2005), 81–108.
- [13] Sham Kakade and John Langford. 2002. Approximately Optimal Approximate Reinforcement Learning. In *Proceedings of the Nineteenth International Conference on Machine Learning (ICML '02)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 267–274.
- [14] Jan Leike, Miljan Martic, Victoria Krakovna, Pedro A. Ortega, Tom Everitt, Andrew LeFrancq, Laurent Orseau, and Shane Legg. 2017. AI Safety Gridworlds.
- [15] Horia Mania, Aurelia Guy, and Benjamin Recht. 2018. Simple random search provides a competitive approach to reinforcement learning.
- [16] Alberto Maria Metelli, Matteo Papini, Nico Montali, and Marcello Restelli. 2020. Importance Sampling Techniques for Policy Optimization. *Journal of Machine Learning Research* 21, 141 (2020), 1–75. <http://jmlr.org/papers/v21/20-124.html>
- [17] Santiago Paternain, Miguel Calvo-Fullana, Luiz F. O. Chamon, and Alejandro Ribeiro. 2019. Safe Policies for Reinforcement Learning via Primal-Dual Methods. arXiv:1911.09101 [eess.SY]
- [18] Tom Schaul, Diana Borsa, Joseph Modayil, and Razvan Pascanu. 2019. Ray Interference: a Source of Plateaus in Deep Reinforcement Learning.
- [19] John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. 2015. Trust Region Policy Optimization. In *International conference on machine learning*. 1889–1897.
- [20] John Schulman, Philipp Moritz, Sergey Levine, Michael I. Jordan, and Pieter Abbeel. 2016. High-Dimensional Continuous Control Using Generalized Advantage Estimation. In *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.).
- [21] Richard S. Sutton and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction*. A Bradford Book, Cambridge, MA, USA.
- [22] Chen Tessler, Daniel J. Mankowitz, and Shie Mannor. 2018. Reward Constrained Policy Optimization.
- [23] Eiji Uchibe and Kenji Doya. 2009. Constrained Reinforcement Learning from Intrinsic and Extrinsic Rewards. In *Theory and Novel Applications of Machine Learning*, Meng Joo Er and Yi Zhou (Eds.). IntechOpen, Rijeka, Chapter 11.
- [24] Lu Wen, Jingliang Duan, Shengbo Eben Li, Shaobing Xu, and H. Peng. 2020. Safe Reinforcement Learning for Autonomous Vehicles through Parallel Constrained Policy Optimization. *ArXiv abs/2003.01303* (2020).