

- Dazeley, Fredrik Heintz, et al. 2021. A practical guide to multi-objective reinforcement learning and planning. *arXiv preprint arXiv:2103.09568* (2021).
- [9] Kazuyuki Hiraoka, Manabu Yoshida, and Taketoshi Mishima. 2009. Parallel reinforcement learning for weighted multi-criteria model with adaptive margin. *Cognitive neurodynamics* 3, 1 (2009), 17–24.
- [10] Ammar Jalalimanesh, Hamidreza Shahabi Haghighi, Abbas Ahmadi, Hossein Hejazian, and Madjid Soltani. 2017. Multi-objective optimization of radiotherapy: distributed Q-learning and agent-based simulation. *Journal of Experimental & theoretical artificial intelligence* 29, 5 (2017), 1071–1086.
- [11] Aviral Kumar, Xue Bin Peng, and Sergey Levine. 2019. Reward-conditioned policies. *arXiv preprint arXiv:1912.13465* (2019).
- [12] Pablo Alvarez Lopez, Michael Behrisch, Laura Bieker-Walz, Jakob Erdmann, Yun-Pang Flötteröd, Robert Hilbrich, Leonhard Lücken, Johannes Rummel, Peter Wagner, and Evamarie Wießner. 2018. Microscopic Traffic Simulation using SUMO, In The 21st IEEE International Conference on Intelligent Transportation Systems. *IEEE Intelligent Transportation Systems Conference (ITSC)*. <https://elib.dlr.de/124092/>
- [13] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (2015), 529.
- [14] Hossam Mossalam, Yannis M. Assael, Diederik M. Roijers, and Shimon Whiteson. 2016. Multi-Objective Deep Reinforcement Learning. *CoRR* abs/1610.02707 (2016). [arXiv:1610.02707](https://arxiv.org/abs/1610.02707) <http://arxiv.org/abs/1610.02707>
- [15] Takayuki Osa, Joni Pajarinen, Gerhard Neumann, J Andrew Bagnell, Pieter Abbeel, and Jan Peters. 2018. An algorithmic perspective on imitation learning. *arXiv preprint arXiv:1811.06711* (2018).
- [16] Simone Parisi, Matteo Pirota, and Jan Peters. 2017. Manifold-based multi-objective policy search with sample reuse. *Neurocomputing* 263 (2017), 3–14.
- [17] Simone Parisi, Matteo Pirota, Nicola Smacchia, Luca Bascetta, and Marcello Restelli. 2014. Policy gradient approaches for multi-objective sequential decision making. In *2014 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2323–2330.
- [18] Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. 2018. Film: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.
- [19] Diederik M Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. 2013. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research* 48 (2013), 67–113.
- [20] Diederik Marijn Roijers, Shimon Whiteson, and Frans A Oliehoek. 2015. Computing convex coverage sets for faster multi-objective coordination. *Journal of Artificial Intelligence Research* 52 (2015), 399–443.
- [21] Juergen Schmidhuber. 2019. Reinforcement Learning Upside Down: Don't Predict Rewards—Just Map Them to Actions. *arXiv preprint arXiv:1912.02875* (2019).
- [22] Wen Sun, J Andrew Bagnell, and Byron Boots. 2018. Truncated horizon policy search: Combining reinforcement learning & imitation learning. *arXiv preprint arXiv:1805.11240* (2018).
- [23] Peter Vamplew, Richard Dazeley, Adam Berry, Rustam Issabekov, and Evan Dekker. 2011. Empirical evaluation methods for multiobjective reinforcement learning algorithms. *Machine learning* 84, 1-2 (2011), 51–80.
- [24] Kristof Van Moffaert and Ann Nowé. 2014. Multi-objective reinforcement learning using sets of pareto dominating policies. *The Journal of Machine Learning Research* 15, 1 (2014), 3483–3512.
- [25] Runzhe Yang, Xingyuan Sun, and Karthik Narasimhan. 2019. A Generalized Algorithm for Multi-Objective Reinforcement Learning and Policy Adaptation. In *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 14610–14621.
- [26] Luisa M Zintgraf, Timon V Kanthers, Diederik M Roijers, Frans Oliehoek, and Philipp Beu. 2015. Quality assessment of MORL algorithms: A utility-based approach. In *Benelearn 2015: proceedings of the 24th annual machine learning conference of Belgium and the Netherlands*.
- [27] Eckart Zitzler and Lothar Thiele. 1999. Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE transactions on Evolutionary Computation* 3, 4 (1999), 257–271.
- [28] Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M Fonseca, and Viviane Grunert Da Fonseca. 2003. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation* 7, 2 (2003), 117–132.