

REFERENCES

- [1] Leon Barrett and Srin Narayanan. 2008. Learning all optimal policies with multiple criteria. In *Proceedings of the 25th international conference on Machine learning*. 41–47.
- [2] Marc G Bellemare, Will Dabney, and Rémi Munos. 2017. A distributional perspective on reinforcement learning. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 449–458.
- [3] Richard Bellman. 1957. *Dynamic programming*. Courier Corporation.
- [4] Daniel Bryce, William Cushing, and Subbarao Kambhampati. 2007. Probabilistic planning is multi-objective. *Arizona State University, Tech. Rep. ASU-CSE-07-006* (2007).
- [5] Peichen Gong. 1992. Multiobjective dynamic programming for forest resource management. *Forest Ecology and Management* 48, 1 (1992), 43–54. [https://doi.org/10.1016/0378-1127\(92\)90120-X](https://doi.org/10.1016/0378-1127(92)90120-X)
- [6] Conor F. Hayes, Mathieu Reymond, Diederik M. Roijers, Enda Howley, and Patrick Mannion. 2021. Risk-Aware and Multi-Objective Decision Making with Distributional Monte Carlo Tree Search. In: *Proceedings of the Adaptive and Learning Agents workshop at AAMAS 2021* (2021).
- [7] Conor F. Hayes, Mathieu Reymond, Diederik M. Roijers, Enda Howley, and Patrick Mannion. 2021 In Press. Distributional Monte Carlo Tree Search for Risk-Aware and Multi-Objective Reinforcement Learning. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, Vol. 2021. IFAAMAS.
- [8] Conor F. Hayes, Roxana Rădulescu, Eugenio Bargiacchi, Johan Källström, Matthew Macfarlane, Mathieu Reymond, Timothy Verstraeten, Luisa M. Zintgraf, Richard Dazeley, Fredrik Heintz, Enda Howley, Athirai A. Irissappane, Patrick Mannion, Ann Nowé, Gabriel Ramos, Marcello Restelli, Peter Vamplew, and Diederik M. Roijers. 2021. A Practical Guide to Multi-Objective Reinforcement Learning and Planning. *arXiv preprint arXiv:2103.09568* (2021).
- [9] Conor F. Hayes, Timothy Verstraeten, Diederik M. Roijers, Enda Howley, and Patrick Mannion. 2021. Dominance Criteria and Solution Sets for the Expected Scalarised Returns. In *Proceedings of the Adaptive and Learning Agents workshop at AAMAS 2021*.
- [10] Conor F. Hayes, Timothy Verstraeten, Diederik M. Roijers, Enda Howley, and Patrick Mannion. 2021. Expected Scalarised Returns Dominance: A New Solution Concept for Multi-Objective Decision Making. *arXiv preprint arXiv:2106.01048* (2021).
- [11] Michael Painter, Bruno Lacerda, and Nick Hawes. 2020. Convex Hull Monte-Carlo Tree-Search. In *Proceedings of the Thirtieth International Conference on Automated Planning and Scheduling, Nancy, France, October 26–30, 2020*. AAAI Press, 217–225.
- [12] Mathieu Reymond, Conor F. Hayes, Diederik M. Roijers, Denis Steckelmacher, and Ann Nowé. 2021. Actor-Critic Multi-Objective Reinforcement Learning for Non-Linear Utility Functions. *Multi-Objective Decision Making Workshop (MODEM 2021)* (2021).
- [13] Diederik M. Roijers, Denis Steckelmacher, and Ann Nowé. 2018. Multi-objective Reinforcement Learning for the Expected Utility of the Return. In *Proceedings of the Adaptive and Learning Agents workshop at FAIM 2018*.
- [14] 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.
- [15] Peter Vamplew, Cameron Foale, and Richard Dazeley. 2021. The impact of environmental stochasticity on value-based multiobjective reinforcement learning. In *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-021-05859-1>
- [16] Peter Vamplew, Benjamin J Smith, Johan Kallstrom, Gabriel Ramos, Roxana Radulescu, Diederik M Roijers, Conor F Hayes, Fredrik Heintz, Patrick Mannion, Pieter JK Libin, et al. 2021. Scalar reward is not enough: A response to Silver, Singh, Precup and Sutton (2021). *arXiv preprint arXiv:2112.15422* (2021).
- [17] Weijia Wang and Michèle Sebag. 2012. Multi-objective Monte-Carlo Tree Search (*Proceedings of Machine Learning Research, Vol. 25*), Steven C. H. Hoi and Wray Buntine (Eds.). PMLR, Singapore Management University, Singapore, 507–522.
- [18] DJ White. 1982. Multi-objective infinite-horizon discounted Markov decision processes. *Journal of mathematical analysis and applications* 89, 2 (1982), 639–647.
- [19] Elmar Wolfstetter. 1999. *Topics in Microeconomics: Industrial Organization, Auctions, and Incentives*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511625787>