REFERENCES

[1] 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 (ICML'17)*. JMLR.org, Sydney, NSW, Australia, 449–458.

[2] Dean Eckles and Maurits Kaptein. 2014. Thompson sampling with the online bootstrap. *CoRR* abs/1410.4009 (2014). arXiv:1410.4009 http://arxiv.org/abs/1410.4009

[3] Dean Eckles and Maurits Kaptein. 2019. Bootstrap Thompson Sampling and Sequential Decision Problems in the Behavioral Sciences. *SAGE Open* 9, 2 (2019), 2158244019851675. https://doi.org/10.1177/2158244019851675 arXiv:https://doi.org/10.1177/2158244019851675

[4] Bradley Efron. 2012. Bayesian inference and the parametric bootstrap. *Ann. Appl. Stat.* 6, 4 (12 2012), 1971–1997. https://doi.org/10.1214/12-AOAS571

[5] Conor F. Hayes, Enda Howley, and Patrick Mannion. 2020. Dynamic Thresholded Lexicograpic Ordering. *Adaptive and Learning Agents Workshop (AAMAS 2020)*.

[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. *arXiv preprint arXiv:2102.00966* (2021). https://arxiv.org/abs/2102.00966

[7] Levente Kocsis and Csaba Szepesvári. 2006. Bandit Based Monte-Carlo Planning. In *Machine Learning: ECML 2006*, Johannes Fürnkranz, Tobias Scheffer, and Myra Spiliopoulou (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 282–293.

[8] Patrick Mannion, Sam Devlin, Jim Duggan, and Enda Howley. 2018. Reward shaping for knowledge-based multi-objective multi-agent reinforcement learning. *The Knowledge Engineering Review* 33 (2018), e23. https://doi.org/10.1017/S0269888918000292

[9] Michael Newton and Adrian Raftery. 1994. Approximate Bayesian Inference by the Weighted Likelihood Bootstrap. *Journal of the Royal Statistical Society Series B-Methodological* 56 (01 1994), 3 – 48.

[10] Roxana Rădulescu, Patrick Mannion, Diederik M. Roijers, and Ann Nowé. 2020. Multi-objective multi-agent decision making: a utility-based analysis and survey. *Autonomous Agents and Multi-Agent Systems* 34, 10 (2020).

[11] 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*.

[12] 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.

[13] Yun Shen, Michael J. Tobia, Tobias Sommer, and Klaus Obermayer. 2014. Risk-Sensitive Reinforcement Learning. *Neural Computation* 26, 7 (2014), 1298–1328.

[14] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature* (2016).

[15] Kristof Van Moffaert, Madalina M Drugan, and Ann Nowé. 2013. Scalarized multi-objective reinforcement learning: Novel design techniques. In *2013 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (AD-PRL)*. IEEE, 191–199.

[16] Joel Veness, Kee Siong Ng, Marcus Hutter, William Uther, and David Silver. 2011. A Monte-Carlo AIXI Approximation. *J. Artif. Int. Res.* 40, 1 (Jan. 2011), 95–142.

[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] Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning* 8, 3-4 (1992), 229–256.