

REFERENCES

- [1] Elias Bareinboim, Andrew Forney, and Judea Pearl. 2015. Bandits with unobserved confounders: A causal approach. In *Advances in Neural Information Processing Systems*. 1342–1350.
- [2] E. Bareinboim and J. Pearl. 2016. Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences* 113 (2016), 7345–7352. Issue 27.
- [3] Anthony Cassandra, Michael L Littman, and Nevin L Zhang. 1997. Incremental pruning: A simple, fast, exact method for partially observable Markov decision processes. In *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence (UAI)*. Morgan Kaufmann, 54–61.
- [4] Hsien-Te Cheng. 1988. *Algorithms for partially observable Markov decision processes*. Ph.D. Dissertation. University of British Columbia.
- [5] Nicolás Della Penna, Mark D Reid, and David Balduzzi. 2016. Compliance-Aware Bandits. *arXiv preprint arXiv:1602.02852* (2016).
- [6] A. Forney, J. Pearl, and E. Bareinboim. 2017. Counterfactual Data-Fusion for Online Reinforcement Learners. In *Proceedings of the 34th International Conference on Machine Learning (Proceedings of Machine Learning Research)*, Doina Precup and Yee Whye Teh (Eds.), Vol. 70. 1156–1164.
- [7] J.Y. Halpern. 1998. Axiomatizing Causal Reasoning. In *Uncertainty in Artificial Intelligence*, G.F. Cooper and S. Moral (Eds.), Morgan Kaufmann, San Francisco, CA, 202–210. Also, *Journal of Artificial Intelligence Research* 12:3, 17–37, 2000.
- [8] Finnian Lattimore, Tor Lattimore, and Mark D Reid. 2016. Causal Bandits: Learning Good Interventions via Causal Inference. In *Advances in Neural Information Processing Systems*. 1181–1189.
- [9] Michael L Littman. 1994. The witness algorithm: Solving partially observable Markov decision processes. (1994).
- [10] J. Pearl. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York.
- [11] Martin L Puterman. 2014. *Markov decision processes: discrete stochastic dynamic programming*. Wiley & Sons.
- [12] Diederik M Roijers, Peter Vamplew, Shimon Whiteson, Richard Dazeley, et al. 2013. A Survey of Multi-Objective Sequential Decision-Making. *J. Artif. Intell. Res. (JAIR)* 48 (2013), 67–113.
- [13] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 7587 (2016), 484–489.
- [14] Satinder P Singh, Tommi S Jaakkola, and Michael I Jordan. 1994. Learning Without State-Estimation in Partially Observable Markovian Decision Processes. In *ICML*. 284–292.
- [15] Richard D Smallwood and Edward J Sondik. 1973. The optimal control of partially observable Markov processes over a finite horizon. *Operations Research* 21, 5 (1973), 1071–1088.
- [16] Chris Urmson, Joshua Anhalt, Drew Bagnell, Christopher Baker, Robert Bittner, MN Clark, John Dolan, Dave Duggins, Tugrul Galatali, Chris Geyer, et al. 2008. Autonomous driving in urban environments: Boss and the urban challenge. *Journal of Field Robotics* 25, 8 (2008), 425–466.
- [17] Junzhe Zhang and Elias Bareinboim. 2017. Transfer Learning in Multi-Armed Bandits: A Causal Approach. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*.
- [18] Shlomo Zilberstein. 2015. Building Strong Semi-autonomous Systems. In *Proceedings of the Twenty-Ninth AAAI Conference*. AAAI Press, 4088–4092.