



















## REFERENCES

- [1] Eitan Altman. 1999. *Constrained Markov Decision Processes*. Chapman & Hall/CRC.
- [2] Haoyu Bai, David Hsu, Wee Sun Lee, and Vien A. Ngo. 2011. *Monte Carlo value iteration for continuous-state POMDPs*. STAR, Vol. 68. Springer, Heidelberg, 175–191.
- [3] Robert Bell, Yehuda Koren, and Chris Volinsky. 2007. Chasing \$1,000,000: how we won the Netflix progress prize. *Statistical Computer & Graphics* 12, 2 (Dec. 2007), 4–12.
- [4] Richard Bellman. 1957. A Markovian Decision Process. *Journal of Mathematics and Mechanics* 6, 5 (1957), 679–684.
- [5] Craig Boutilier, Thomas Dean, and Steve Hanks. 1999. Decision-theoretic planning: Structural assumptions and computational leverage. *Journal of Artificial Intelligence Research* 11 (1999), 1–94.
- [6] Iadine Chadès, Josie Carwardine, Tara G. Martin, Samuel Nicol, Régis Sabbadin, and Olivier Buffet. 2012. MOMDPs: A solution for modelling adaptive management problems. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*. 267–273.
- [7] Shih-Fen Cheng, Larry Lin, Jiali Du, Hoong Chuin Lau, and Pradeep Varakantham. 2013. An agent-based simulation approach to experience management in theme parks. In *Winter Simulation Conference*. 1527–1538.
- [8] Frits de Nijs, Matthijs T. J. Spaan, and Mathijs M. de Weerdt. 2015. Best-Response Planning of Thermostatically Controlled Loads under Power Constraints. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*. 615–621.
- [9] Frits de Nijs, Matthijs T. J. Spaan, and Mathijs M. de Weerdt. 2018. Preallocation and Planning under Stochastic Resource Constraints. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*.
- [10] Frits de Nijs, Erwin Walraven, Mathijs M. de Weerdt, and Matthijs T. J. Spaan. 2017. Bounding the Probability of Resource Constraint Violations in Multi-Agent MDPs. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*. 3562–3568.
- [11] Richard Dearden, Nir Friedman, and David Andre. 1999. Model Based Bayesian Exploration. In *Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence*. 150–159.
- [12] Michael O’Gordon Duff. 2002. *Optimal learning: Computational procedures for Bayes-adaptive Markov decision processes*. Ph.D. Dissertation. University of Massachusetts Amherst.
- [13] P. C. Gilmore and R. E. Gomory. 1961. A Linear Programming Approach to the Cutting-Stock Problem. *Operations Research* 9, 6 (1961), 849–859.
- [14] Aditya Gopalan and Shie Mannor. 2015. Thompson Sampling for Learning Parameterized Markov Decision Processes. In *Proceedings of The 28th Conference on Learning Theory (Proceedings of Machine Learning Research)*, Peter Grünwald, Elad Hazan, and Satyen Kale (Eds.), Vol. 40. PMLR, Paris, France, 861–898.
- [15] Arthur Guez, David Silver, and Peter Dayan. 2013. Scalable and efficient Bayes-adaptive reinforcement learning based on Monte-Carlo tree search. *Journal of Artificial Intelligence Research* 48 (2013), 841–883.
- [16] Negar Hariri, Bamshad Mobasher, and Robin Burke. 2014. Context adaptation in interactive recommender systems. In *Proceedings of the 8th ACM Conference on Recommender systems*. 41–48.
- [17] Ronald A. Howard. 1966. Information value theory. *IEEE Transactions on Systems Science and Cybernetics* 2, 1 (8 1966), 22–26.
- [18] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. 1998. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* 101, 1-2 (May 1998), 99–134.
- [19] Sham Machandranath Kakade. 2003. *On the Sample Complexity of Reinforcement Learning*. Ph.D. Dissertation. University College London.
- [20] Hanna Kurniawati, David Hsu, and Wee Sun Lee. 2008. SARSOP: Efficient Point-Based POMDP Planning by Approximating Optimally Reachable Belief Spaces. In *Robotics: Science and Systems*. Zurich, Switzerland.
- [21] James John Martin. 1967. *Bayesian decision problems and Markov chains*. Wiley, New York.
- [22] Péron Martin, Kai Helge Becker, Peter Bartlett, and Iadine Chadès. 2017. Fast-Tracking Stationary MOMDPs for Adaptive Management Problems. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*. 4531–4537.
- [23] Sylvie C. W. Ong, Shao Wei Png, David Hsu, and Wee Sun Lee. 2010. Planning under Uncertainty for Robotic Tasks with Mixed Observability. *The International Journal of Robotics Research* 29, 8 (2010), 1053–1068.
- [24] Ian Osband and Benjamin Van Roy. 2014. Near-optimal Reinforcement Learning in Factored MDPs. In *Advances in Neural Information Processing Systems* 27, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Eds.), Curran Associates, Inc., 604–612.
- [25] Christos H. Papadimitriou and John N. Tsitsiklis. 1987. The Complexity of Markov Decision Processes. *Mathematics of Operations Research* 12, 3 (1987), 441–450.
- [26] Josep M. Porta, Nikos Vlassis, Matthijs T. J. Spaan, and Pascal Poupart. 2006. Point-based value iteration for continuous POMDPs. *Journal of Machine Learning Research* 7 (11 2006), 2329–2367.
- [27] Pascal Poupart, Aarti Malhotra, Pei Pei, Kee-Eung Kim, Bongseok Goh, and Michael Bowling. 2015. Approximate Linear Programming for Constrained Partially Observable Markov Decision Processes. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*. 3342–3348.
- [28] 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.
- [29] Dana Ron, Yoram Singer, and Naftali Tishby. 1996. The Power of Amnesia: Learning Probabilistic Automata with Variable Memory Length. *Machine Learning* 25 (1996), 117–149.
- [30] Guy Shani, David Heckerman, and Ronen I. Brafman. 2005. An MDP-Based Recommender System. *Journal of Machine Learning Research* 6 (2005), 1265–1295.
- [31] Edward Allan Silver. 1963. *Markovian decision processes with uncertain transition probabilities or rewards*. Ph.D. Dissertation. Massachusetts Institute of Technology.
- [32] Edward J. Sondik. 1971. *The Optimal Control of Partially Observable Markov Processes*. Ph.D. Dissertation. Stanford University.
- [33] Harald Steck. 2013. Evaluation of Recommendations: Rating-prediction and Ranking. In *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys ’13)*. ACM, New York, NY, USA, 213–220.
- [34] Malcolm J. A. Strens. 2000. A Bayesian Framework for Reinforcement Learning. In *Proceedings of the 17th International Conference on Machine Learning*. 943–950.
- [35] Richard S. Sutton and Andrew G. Barto. 2018. *Reinforcement Learning: An introduction* (2 ed.). The MIT Press.
- [36] Liang Tang, Romer Rosales, Ajit Singh, and Deepak Agarwal. 2013. Automatic ad format selection via contextual bandits. In *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management*. 1587–1594.
- [37] Georgios Theodorou, Nikos Vlassis, and Zheng Wen. 2017. An Interactive Points of Interest Guidance System. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion (IUI ’17 Companion)*. ACM, New York, NY, USA, 49–52.
- [38] Bart Thomee, David A. Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. 2016. YFCC100M: the new data in multimedia research. *Commun. ACM* 59, 2 (2016), 64–73.
- [39] William R. Thompson. 1933. On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika* 25, 3/4 (1933), 285–294.
- [40] Erwin Walraven and Matthijs T. J. Spaan. 2018. Column Generation Algorithms for Constrained POMDPs. *Journal of Artificial Intelligence Research* (2018).
- [41] Kirk A. Yost and Alan R. Washburn. 2000. The LP/POMDP Marriage: Optimization with Imperfect Information. *Naval Research Logistics* 47, 8 (2000), 607–619.
- [42] Qi Zhang, Edmund Durfee, and Satinder P. Singh. 2017. Minimizing Maximum Regret in Commitment Constrained Sequential Decision Making. In *Proceedings of the 27th International Conference on Automated Planning and Scheduling*, Laura Barbulescu, Jeremy Frank, Mausam, and Stephen F. Smith (Eds.).