





















- [49] Abolfazl Lavaei, Fabio Somenzi, Sadegh Soudjani, Ashutosh Trivedi, and Majid Zamani. Formal controller synthesis for continuous-space MDPs via model-free reinforcement learning. In *International Conference on Cyber-Physical Systems (ICCPs)*, pages 98–107, 2020.
- [50] Hansol Yoon, Yi Chou, Xin Chen, Eric Frew, and Sriram Sankaranarayanan. Predictive runtime monitoring for linear stochastic systems and applications to geofence enforcement for UAVs. In *International Conference on Runtime Verification*. Springer, 2019.
- [51] Sofie Haesaert and Sadegh Soudjani. Robust dynamic programming for temporal logic control of stochastic systems. *IEEE Transactions on Automatic Control*, 66(6):2496–2511, 2020.
- [52] Xuchu Ding, Stephen L Smith, Calin Belta, and Daniela Rus. Optimal control of Markov decision processes with linear temporal logic constraints. *IEEE Transactions on Automatic Control*, 59(5):1244–1257, 2014.
- [53] Pranav Ashok, Krishnendu Chatterjee, Przemyslaw Daca, Jan Křetínský, and Tobias Meggendorfer. Value iteration for long-run average reward in Markov decision processes. In *International Conference on Computer Aided Verification*, pages 201–221. Springer, 2017.
- [54] George K Atia, Andre Beckus, Ismail Alkhoury, and Alvaro Velasquez. Verifiable planning in expected reward multichain MDPs. *arXiv preprint arXiv:2012.02178*, 2020.
- [55] Javier Garcia and Fernando Fernández. A comprehensive survey on safe reinforcement learning. *Journal of Machine Learning Research*, 16(1):1437–1480, 2015.
- [56] Benjamin Recht. A tour of reinforcement learning: The view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, pages 253–279, 2018.
- [57] Yonathan Efroni, Shie Mannor, and Matteo Pirota. Exploration-exploitation in constrained MDPs. *arXiv preprint arXiv:2003.02189*, 2020.
- [58] Tomáš Brázdil, Krishnendu Chatterjee, Martin Chmelik, Vojtěch Forejt, Jan Křetínský, Marta Kwiatkowska, David Parker, and Mateusz Ujma. Verification of Markov decision processes using learning algorithms. In *Automated Technology for Verification and Analysis (ATVA)*, pages 98–114. Springer, 2014.
- [59] Jie Fu and Ufuk Topcu. Probably approximately correct MDP learning and control with temporal logic constraints. In *Proceedings of Robotics: Science and Systems*, 2014.
- [60] Mohammadhosein Hasanbeig, Yiannis Kantaros, Alessandro Abate, Daniel Kroening, George J Pappas, and Insup Lee. Reinforcement learning for temporal logic control synthesis with probabilistic satisfaction guarantees. In *IEEE Conference on Decision and Control (CDC)*, pages 5338–5343. IEEE, 2019.
- [61] Ryohei Oura, Ami Sakakibara, and Toshimitsu Ushio. Reinforcement learning of control policy for linear temporal logic specifications using limit-deterministic Büchi automata. *IEEE Control Systems Letters*, 4(3):761–766, July 2020.
- [62] Yuqian Jiang, Suda Bharadwaj, Bo Wu, Rishi Shah, Ufuk Topcu, and Peter Stone. Temporal-logic-based reward shaping for continuing reinforcement learning tasks. *Good Systems-Published Research*, 2021.
- [63] Vektor Dewanto, George Dunn, Ali Eshragh, Marcus Gallagher, and Fred Roosta. Average-reward model-free reinforcement learning: a systematic review and literature mapping. *arXiv preprint arXiv:2010.08920*, 2020.
- [64] Yasin Abbasi-Yadkori, Peter Bartlett, Kush Bhatia, Nevena Lazic, Csaba Szepesvari, and Gellert Weisz. POLITEX: Regret bounds for policy iteration using expert prediction. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pages 3692–3702. PMLR, 09–15 Jun 2019.
- [65] Chen-Yu Wei, Mehdi Jafarnia Jahromi, Haipeng Luo, Hiteshi Sharma, and Rahul Jain. Model-free reinforcement learning in infinite-horizon average-reward Markov decision processes. In *International conference on machine learning*, pages 10170–10180. PMLR, 2020.
- [66] Rahul Singh, Abhishek Gupta, and Ness B Shroff. Learning in Markov decision processes under constraints. *arXiv preprint arXiv:2002.12435*, 2020.
- [67] Jan Křetínský, Fabian Michel, Lukas Michel, and Guillermo Perez. Finite-memory near-optimal learning for Markov decision processes with long-run average reward. In *Conference on Uncertainty in Artificial Intelligence*, pages 1149–1158. PMLR, 2020.