

REFERENCES

- [1] A. Abate, J.P. Katoen, J. Lygeros, and M. Prandini. 2010. Approximate Model Checking of Stochastic Hybrid Systems. *European Journal of Control* 16, 6 (2010), 624–641.
- [2] Jacob Andreas, Dan Klein, and Sergey Levine. 2017. Modular Multitask Reinforcement Learning with Policy Sketches. In *ICML*, Vol. 70. 166–175.
- [3] Lucian Busoniu, Robert Babuska, Bart De Schutter, and Damien Ernst. 2010. *Reinforcement Learning and Dynamic Programming Using Function Approximators*. Vol. 39. CRC press.
- [4] Kenji Doya. 2000. Reinforcement Learning in Continuous Time and Space. *Neural computation* 12, 1 (2000), 219–245.
- [5] Richard Durrett. 1999. *Essentials of stochastic processes*. Vol. 1. Springer.
- [6] Damien Ernst, Pierre Geurts, and Louis Wehenkel. 2005. Tree-based Batch Mode Reinforcement Learning. *JMLR* 6, Apr (2005), 503–556.
- [7] Jie Fu and Ufuk Topcu. 2014. Probably Approximately Correct MDP Learning and Control With Temporal Logic Constraints. In *Robotics: Science and Systems X*.
- [8] S. Haesaert, S.E.Z. Soudjani, and A. Abate. 2017. Verification of general Markov decision processes by approximate similarity relations and policy refinement. *SIAM Journal on Control and Optimisation* 55, 4 (2017), 2333–2367.
- [9] Mohammadhosein Hasanbeig, Alessandro Abate, and Daniel Kroening. 2018. Logically-Constrained Neural Fitted Q-Iteration. *arXiv preprint arXiv:1809.07823* (2018).
- [10] Mohammadhosein Hasanbeig, Alessandro Abate, and Daniel Kroening. 2018. Logically-Constrained Reinforcement Learning. *arXiv preprint arXiv:1801.08099* (2018).
- [11] Mohammadhosein Hasanbeig, Alessandro Abate, and Daniel Kroening. 2019. Certified Reinforcement Learning with Logic Guidance. *arXiv preprint arXiv:1902.00778* (2019).
- [12] Kurt Hornik. 1991. Approximation Capabilities of Multilayer Feedforward Networks. *Neural networks* 4, 2 (1991), 251–257.
- [13] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. 1989. Multilayer feed-forward networks are universal approximators. *Neural networks* 2, 5 (1989), 359–366.
- [14] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. 2016. Asynchronous Methods for Deep Reinforcement Learning. In *ICML*. 1928–1937.
- [15] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. Human-level Control Through Deep Reinforcement Learning. *Nature* 518, 7540 (2015), 529–533.
- [16] Dirk Ormonieit and Šaunak Sen. 2002. Kernel-based Reinforcement Learning. *Machine learning* 49, 2 (2002), 161–178.
- [17] Martin Riedmiller. 2005. Neural Fitted Q Iteration-First Experiences with a Data Efficient Neural Reinforcement Learning Method. In *ECML*, Vol. 3720. Springer, 317–328.
- [18] Martin Riedmiller and Heinrich Braun. 1993. A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm. In *Neural networks*. IEEE, 586–591.
- [19] Salomon Sickert, Javier Esparza, Stefan Jaax, and Jan Křetínský. 2016. Limit-deterministic Büchi automata for linear temporal logic. In *CAV*. Springer, 312–332.
- [20] Stephen L Smith, Jana Tumulová, Calin Belta, and Daniela Rus. 2011. Optimal path planning for surveillance with temporal-logic constraints. *The International Journal of Robotics Research* 30, 14 (2011), 1695–1708.
- [21] Richard S Sutton. 1996. Generalization in Reinforcement Learning: Successful Examples Using Sparse Coarse Coding. In *NIPS*. 1038–1044.
- [22] Richard S Sutton and Andrew G Barto. 1998. *Reinforcement learning: An introduction*. Vol. 1. MIT press Cambridge.
- [23] Gerald Tesauro. 1995. TD-Gammon: A self-teaching Backgammon Program. In *Applications of Neural Networks*. Springer, 267–285.
- [24] Ilya Tkachev, Alexandru Mereacre, Joost-Pieter Katoen, and Alessandro Abate. 2017. Quantitative model-checking of controlled discrete-time Markov processes. *Information and Computation* 253 (2017), 1–35.
- [25] Hado Van Hasselt and Marco A Wiering. 2007. Reinforcement Learning in Continuous Action Spaces. In *ADPRL*. IEEE, 272–279.