Bibliography

1. Sutton, R.S., Barto, A.G.: Introduction to Reinforcement Learning. MIT Press (1998)
2. Bellman, R.: A Markovian decision process. Indiana Univ. Math. J. **6**(4), 679–684 (1957)
3. Markov Decision Process, wikipedia.org. https://en.wikipedia.org/wiki/Markov_decision_process. Accessed Aug 2018
4. Markov Chain, wikipedia.org. https://en.wikipedia.org/wiki/Markov_chain. Accessed Aug 2018
5. Markov Property, wikipedia.org. https://en.wikipedia.org/wiki/Markov_property. Accessed Aug 2018
6. Solving MDPs with dynamic programming, towardsdatascience.com. https://towardsdatascience.com/reinforcement-learning-demystified-solving-mdps-with-dynamic-programming-b52c8093c919. Accessed Aug 2018
7. Dynamic Programming in Python for reinforcement learning, medium.com. https://medium.com/harder-choices/dynamic-programming-in-python-reinforcement-learning-bb288d95288f. Accessed Aug 2018
8. Understanding RL—The Bellman Equation, joshgreaves.com. https://joshgreaves.com/reinforcement-learning/understanding-rl-the-bellman-equations/. Accessed Aug 2018
9. Reinforcement Learning in Finance, Coursera. https://coursera.org/lecture/reinforcement-learning-in-finance. Accessed Aug 2018
10. Deep Reinforcement Learning Demystified, medium.com. https://medium.com/@m.alzantot/deep-reinforcement-learning-demysitifed-episode-2-policy-iteration-value-iteration-and-q-978f9e89ddaa. Accessed Aug 2018
11. Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., Zaremba, W.: OpenAI Gym (2016). https://arxiv.org/abs/1606.01540
12. Docs—Open AI Gym. https://gym.openai.com/docs/. Accessed Aug 2018
13. Keras-RL, GitHub Repository Keras-RL. https://github.com/keras-rl. Accessed Aug 2018
14. Shani, G.: A Survey of Model-Based and Model-Free Methods for Resolving Perceptual Aliasing. Ben-Gurion University (2004)
15. Temporal Difference Learning, Wikipedia. https://en.wikipedia.org/wiki/Temporal_difference_learning. Accessed Aug 2018
16. Sutton, R.: Learning to predict by the methods of temporal differences. Mach. Learn. **3**(1), 9–44 (1988)
17. Eligibility Traces, incompleteideas.net. http://www.incompleteideas.net/book/ebook/node72.html. Accessed Aug 2018
18. SARSA, Wikipedia. https://en.wikipedia.org/wiki/State%E2%80%93action%E2%80%93reward%E2%80%93state%E2%80%93action. Accessed Aug 2018

M. Sewak, *Deep Reinforcement Learning*,
https://doi.org/10.1007/978-981-13-8285-7

19. Diving deep into reinforcement learning, freecodecamp.org. https://medium.freecodecamp.org/diving-deeper-into-reinforcement-learning-with-q-learning-c18d0db58efe. Accessed Aug 2018
20. Tokic, M.: Adaptive epsilon-greedy exploration in reinforcement learning based on value differences. In: KI 2010: Advances in Artificial Intelligence, pp. 203–210. Springer, Berlin (2010)
21. Imaddabbura, Bandit Algorithms, github.io. https://imaddabbura.github.io/blog/data%20science/2018/03/31/epsilon-Greedy-Algorithm.html. Accessed Aug 2018
22. Sewak, M., Rezaul Karim, Md., Pujari, P.: Practical Convolutional Neural Networks: Implement Advanced Deep Learning Models Using Python. Packt Publishing (2018). ISBN: 1788392302, 9781788392303.
23. Sewak, M., Sahay, S.K., Rathore, H.: An overview of deep learning architecture of deep neural networks and autoencoders. In: International Conference on Intelligent Computing, 25–27 Oct 2018. Amrita University, Proceedings in Journal of Computational and Theoretical Nanoscience (2018)
24. Sewak, M., Sahay, S.K., Rathore, H.: Comparison of deep learning and the classical machine learning algorithm for the malware detection. In: 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, SNPD (2018)
25. Sewak, M., Sahay, S.K., Rathore, H.: An investigation of a deep learning-based malware detection system. In: Proceedings of the 13th International Conference on Availability, Reliability and Security, ARES 2018, pp. 26:1–26:5 (2018)
26. "OpenAI Universe", GitHub Repository. https://github.com/openai/universe
27. "OpenAI Retro", GitHub Repository. https://github.com/openai/retro
28. Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., Zaremba, W.: Openai gym (2016)
29. "DeepMind Lab". https://deepmind.com/blog/open-sourcing-deepmind-lab/
30. "DeepMind Control Suite", GitHub Repository. https://github.com/deepmind/dm_control
31. Johnson, M., Hofmann, K., Hutton, T., Bignell, D.: The Malmo platform for artificial intelligence experimentation. In: Kambhampati, S. (ed.) Proceedings 25th International Joint Conference on Artificial Intelligence, pp. 42–46. AAAI Press, Palo Alto, California USA (2016). https://github.com/Microsoft/malmo
32. Duan, Y., Chen, X., Houthooft, R., Schulman, J., Abbeel, P.: Benchmarking deep reinforcement learning for continuous control. In: Proceedings of the 33rd International Conference on Machine Learning (ICML) (2016)
33. "DeepMind's TRFL", GitHub Repository. https://github.com/deepmind/trfl
34. "OpenAI Baselines", GitHub Repository. https://github.com/openai/baselines
35. Plappert, M.: "Keras-RL", GitHub Repository (2016). https://keras-rl.https://github.com/keras-rl/keras-rl
36. Caspi, I., Leibovich, G., Novik, G., Endrawis, S.: Reinforcement learning coach (2017). https://doi.org/10.5281/zenodo.1134899
37. Liang, E., Liaw, R., Nishihara, R., Moritz, P., Fox, R., Goldberg, K., Gonzalez, J.E., Jordan, M.I., Stoica, I.: RLlib: abstractions for distributed reinforcement learning. In: International Conference on Machine Learning (ICML) (2018)
38. DQN, deepmind.com. https://deepmind.com/research/dqn/. Accessed Aug 2018
39. AlphaGo, deepmind.com. https://deepmind.com/research/alphago/. Accessed Aug 2018
40. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M.A.: Playing Atari with Deep Reinforcement Learning (2013). https://arxiv.org/abs/1312.5602
41. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D.: Human-level control through deep reinforcement learning. Nature **518** (2015). 10.1038/nature14236

42. Schaul, T., Quan, J., Antonoglou, I., Silver, D.: Prioritized Experience Replay (2015). https://arxiv.org/abs/1511.05952
43. Lin, L.-J.: Reinforcement learning for robots using neural networks. Carnegie Mellon University, Ph.D. Thesis Lin:1992: RLR:168871 (1992)
44. Seno, T.: Welcome to deep reinforcement learning, towardsdatascience.com. https://towardsdatascience.com/welcome-to-deep-reinforcement-learning-part-1-dqn-c3cab4d41b6b . Accessed Aug 2018
45. Juliani, A.: Simple reinforcement learning with tensorflow, medium.com. https://medium.-com/@awjuliani/simple-reinforcement-learning-with-tensorflow-part-4-deep-q-networks-and-beyond-8438a3e2b8df. Accessed Aug 2018
46. van Hasselt, H., Guez, A., Silver, D.: Deep Reinforcement Learning with Double Q-Learning (2015). https://arxiv.org/abs/1509.06461
47. Wang, Z., de Freitas, N., Lanctot, M.: Dueling Network Architectures for Deep Reinforcement Learning (2015). https://arxiv.org/abs/1511.06581
48. Sutton, R.S., McAllester, D., Singh, S., Mansour, Y.: Policy gradient methods for reinforcement learning with function approximation. In: Proceedings of the 12th International Conference on Neural Information Processing Systems, NIPS'99, pp. 1057–1063 (1999)
49. Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., Riedmiller, M.: Deterministic policy gradient algorithms. In: Proceedings of the 31st International Conference on Machine Learning, pp. 387–395 (2014)
50. Silver, D.: Lecture 7—Policy Gradient, University College London. http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching_files/pg.pdf. Accessed Aug 2018
51. Li, F.-F., Johnson, J., Yeung, S.: Lecture 14, CS231—Stanford. http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture14.pdf. Accessed Aug 2018
52. Williams, R.J.: A class of gradient-estimating algorithms for reinforcement learning in neural networks. In: Proceedings of the IEEE First International Conference on Neural Networks, San Diego, CA (1987)
53. Williams, R.J.: Simple statistical gradient-following algorithms for connectionist reinforcement learning. Mach. Learn. **8**(3), 229–256 (1992)
54. Mnih, V., Badia, A.P., Mirza, M., Graves, A., Lillicrap, T.P., Harley, T., Silver, D., Kavukcuoglu, K.: Asynchronous methods for deep reinforcement learning (2016). https://arxiv.org/abs/1602.01783
55. Degris, T., Pilarski, P.M., Sutton, R.S.: Model-free reinforcement learning with continuous action in practice. In: 2012 American Control Conference (ACC), pp. 2177–2182 (2012)
56. Bhatnagar, S., Sutton, R.S., Ghavamzadeh, M., Lee, M.: Natural actor-critic algorithms. Automatica **45**(11), 2471–2482 (2009)
57. Sutton, R.S., Barto, A.G.: Reinforcement learning—an introduction. In: Adaptive Computation and Machine Learning. MIT Press (1998)
58. Wu, Y., Mansimov, E., Liao, S., Radford,.A., Schulman, J.: Openai baselines: Acktr a2c (2017). https://blog.openai.com/baselines-acktr-a2c
59. Konda, V.R., Tsitsiklis, J.N.: On actor-critic algorithms. SIAMJ. Control Optim. **42**(4), 1143–1166 (2003)
60. Abadi, M., et al.: Model Sub-classing, TensorFlow Guide: High Level API—Keras (2019). https://www.tensorflow.org/guide/keras#model_subclassing
61. Abadi, M., et al.: Functional API, TensorFlow Guide: High Level API—Keras (2019). https://www.tensorflow.org/guide/keras#model_subclassing
62. Abadi, M., et al.: Gradient Tapes, TensorFlow Tutorial: Automatic Differentiation and Gradient Tapes (2019). https://www.tensorflow.org/tutorials/eager/automatic_differentiation
63. Abadi, M., et al.: apply_gradient(), TensorFlow API Docs: tf.train.Optimizer Class (2019). https://www.tensorflow.org/api_docs/python/tf/train/Optimizer

64. Yuan, R.: Deep reinforcement learning: playing CartPole through asynchronous advantage actor critic (A3C) with tf.keras and eager execution, Medium.com (2018). https://medium.com/tensorflow/deep-reinforcement-learning-playing-cartpole-through-asynchronous-advantage-actor-critic-a3c-7eab2eea5296
65. Daoust, M.: A3C Blog Post, GitHub repository: TensorFlow/Models/Research (2018). https://github.com/tensorflow/models/tree/master/research/a3c_blogpost
66. Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., Riedmiller, M.: Deterministic policy gradient algorithms. In: Proceedings of the 31st International Conference on Machine Learning, Proceedings of Machine Learning Research, vol. 32, pp. 387–395, Beijing, China, 22–24 Jun 2014
67. Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D.: Continuous control with deep reinforcement learning (2015). https://arxiv.org/abs/1509.02971
68. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift (2015). https://arxiv.org/abs/1502.03167
69. Weng, L.: Policy Gradient Algorithms (2018). https://lilianweng.github.io/lil-log/2018/04/08/policy-gradient-algorithms.html#off-policy-policy-gradient. Accessed Jan 2019

Index

M. Sewak, *Deep Reinforcement Learning*,
https://doi.org/10.1007/978-981-13-8285-7