Chapter 19
Algorithm Table

Zihan Ding

Abstract In this chapter, we summarize the references of some important reinforcement learning algorithms introduced in the book as a table.

Keywords Reinforcement learning · Algorithm · On-policy · Off-policy · Action space

In this chapter, Table 19.1 containing the most popular reinforcement learning algorithms is summarized, especially for those introduced in this book. We hope this will help the readers to refer to the original papers.

Z. Ding (✉)
Imperial College London, London, UK
e-mail: zhding@mail.ustc.edu.cn

H. Dong et al. (eds.), *Deep Reinforcement Learning*,
https://doi.org/10.1007/978-981-15-4095-0_19

Table 19.1 Reinforcement learning algorithms

RL algorithms	Policy	Action space	Year	Paper	Authors
Q-learning	Off-policy	Discrete	1992	Q-learning (Watkins and Dayan 1992)	Cristopher J.C.H Watkins and Peter Dayan
SARSA	On-policy	Discrete	1994	Online Q-learning using connectionist systems (Rummery and Niranjan 1994)	G.A. Rummery and M. Niranjan
DQN	Off-policy	Discrete	2015	Human-level control through deep reinforcement learning (Mnih et al. 2015)	Volodymyr Mnih, et al.
Dueling DQN	Off-policy	Discrete	2015	Dueling network architectures for deep reinforcement learning (Wang et al. 2015)	Ziyu Wang, et al.
Double DQN	Off-policy	Discrete	2016	Deep reinforcement learning with double Q-learning (Van Hasselt et al. 2016)	Hado van Hasselt, et al.
Noisy DQN	Off-policy	Discrete	2017	Noisy networks for exploration (Fortunato et al. 2017)	Meire Fortunato, et al.
Distributed DQN	Off-policy	Discrete	2017	A distributional perspective on reinforcement learning (Bellemare et al. 2017)	Marc G. Bellemare, et al.
Actor-critic (QAC)	On-policy	Discrete or continuous	2000	Actor-critic algorithms (Konda and Tsitsiklis 2000)	Vijay R. Konda and John N. Tsitsiklis
A3C	On-policy	Discrete or continuous	2016	Asynchronous methods for deep reinforcement learning (Mnih et al. 2016)	Volodymyr Mnih, et al.
REINFORCE	On-policy	Discrete or continuous	1988	On the use of backpropagation in associative reinforcement learning (Williams 1988)	Ronald J. Williams

(continued)

Table 19.1 continued

RL algorithms	Policy	Action space	Year	Paper	Authors
DDPG	Off-policy	Continuous	2016	Continuous control with deep reinforcement learning (Lillicrap et al. 2015)	Timothy P. Lillicrap, et al.
TD3	Off-policy	Continuous	2018	Addressing function approximation error in actor-critic methods (Fujimoto et al. 2018)	Scott Fujimoto, et al.
SAC	Off-policy	Discrete or continuous	2018	Soft actor-critic algorithms and applications (Haarnoja et al. 2018)	Tuomas Haarnoja, et al.
TRPO	On-policy	Discrete or continuous	2015	Trust region policy optimization (Schulman et al. 2015)	John Schulman, et al.
PPO	On-policy	Discrete or continuous	2017	Proximal policy optimization algorithms (Schulman et al. 2017)	John Schulman, et al.
DPPO	On-policy	Discrete or continuous	2017	Emergence of locomotion behaviours in rich environments (Heess et al. 2017)	Nicolas Heess, et al.
ACKTR	On-policy	Discrete or continuous	2017	Scalable trust-region method for deep reinforcement learning using Kronecker-factored approximation Wu et al. (2017)	Yuhuai Wu, et al.
CE method	On-policy	Discrete or continuous	2004	The cross-entropy method: A unified approach to Monte Carlo simulation, randomized optimization and machine learning (Rubinstein and Kroese 2004)	R. Rubinstein and D. Kroese

References

Bellemare MG, Dabney W, Munos R (2017) A distributional perspective on reinforcement learning. In: Proceedings of the 34th international conference on machine learning, vol 70, pp 449–458. JMLR.org

Fortunato M, Azar MG, Piot B, Menick J, Osband I, Graves A, Mnih V, Munos R, Hassabis D, Pietquin O, et al. (2017) Noisy networks for exploration. arXiv:170610295

Fujimoto S, van Hoof H, Meger D (2018) Addressing function approximation error in actor-critic methods. arXiv:180209477

Haarnoja T, Zhou A, Hartikainen K, Tucker G, Ha S, Tan J, Kumar V, Zhu H, Gupta A, Abbeel P, et al. (2018) Soft actor-critic algorithms and applications. arXiv:181205905

Heess N, Sriram S, Lemmon J, Merel J, Wayne G, Tassa Y, Erez T, Wang Z, Eslami S, Riedmiller M, et al. (2017) Emergence of locomotion behaviours in rich environments. arXiv:170702286

Konda VR, Tsitsiklis JN (2000) Actor-critic algorithms. In: Advances in neural information processing systems, pp 1008–1014

Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, Silver D, Wierstra D (2015) Continuous control with deep reinforcement learning. arXiv:150902971

Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, Petersen S, Beattie C, Sadik A, Antonoglou I, King H, Kumaran D, Wierstra D, Legg S, Hassabis D (2015) Human-level control through deep reinforcement learning. Nature 518(7540):529–533

Mnih V, Badia AP, Mirza M, Graves A, Lillicrap T, Harley T, Silver D, Kavukcuoglu K (2016) Asynchronous methods for deep reinforcement learning. In: International conference on machine learning (ICML), pp 1928–1937

Rubinstein RY, Kroese DP (2004) The cross-entropy method: a unified approach to monte carlo simulation, randomized optimization and machine learning (Information science and statistics). Springer, New York

Rummery GA, Niranjan M (1994) On-line Q-learning using connectionist systems, vol 37. University of Cambridge, Department of Engineering Cambridge, Cambridge

Schulman J, Levine S, Abbeel P, Jordan M, Moritz P (2015) Trust region policy optimization. In: International conference on machine learning (ICML), pp 1889–1897

Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O (2017) Proximal policy optimization algorithms. arXiv:170706347

Van Hasselt H, Guez A, Silver D (2016) Deep reinforcement learning with double Q-learning. In: Thirtieth AAAI conference on artificial intelligence

Wang Z, Schaul T, Hessel M, Van Hasselt H, Lanctot M, De Freitas N (2015) Dueling network architectures for deep reinforcement learning. arXiv:151106581

Watkins CJ, Dayan P (1992) Q-learning. Mach Learn 8(3–4):279–292

Williams RJ (1988) On the use of backpropagation in associative reinforcement learning. In: Proceedings of the IEEE international conference on neural networks, vol 1, San Diego, pp 263–270

Wu Y, Mansimov E, Grosse RB, Liao S, Ba J (2017) Scalable trust-region method for deep reinforcement learning using Kronecker-factored approximation. In: Advances in neural information processing systems, pp 5279–5288