

Contents

1	Introduction	5
1.1	Evolution of Reinforcement Learning	5
1.1.1	Basic Reinforcement Learning	7
1.1.2	Inverse Reinforcement Learning	26
1.1.3	Meta Reinforcement Learning	29
1.1.4	Hierarchical Reinforcement Learning	35
1.1.5	Multi-Task Reinforcement Learning	40
1.1.6	Multi-agent Reinforcement Learning	45
1.2	Deep Reinforcement Learning	48
1.3	Applications of Reinforcement Learning	50
1.4	Summary	53
2	Reinforcement Learning Foundations	55
2.1	Basic Concepts and Algorithms	55
2.1.1	Basic Concepts	56
2.1.2	Model-based vs Model-free Methods	60
2.1.3	Tabular vs Approximate Methods	62
2.1.4	Value-based vs Policy-based Methods	65
2.2	Basic Reinforcement Learning Methods	66
2.2.1	Dynamic Programming	67
2.2.2	Monte-Carlo Method	70
2.2.3	Policy Gradient Method	76
2.2.4	Deterministic Policy Gradient Method	80
2.2.5	Monte Carlo Policy Gradient Method - REINFORCE	82
2.2.6	REINFORCE with Baseline	84
2.2.7	Actor-Critic Method	87
2.3	Advanced Reinforcement Learning Branches	88
2.3.1	Inverse Reinforcement Learning	89
2.3.2	Meta Reinforcement Learning	98
2.3.3	Hierarchical Reinforcement Learning	101
2.3.4	Multi-Task Reinforcement Learning	105
2.3.5	Multi-agent Reinforcement Learning	106

2.4	Deep Learning Foundations	109
2.4.1	Basic Concepts	110
2.4.2	Network Structures	113
2.5	Advanced Neural Network Components	116
2.5.1	Deep Embeddings	117
2.5.2	Attention	118
2.5.3	Transformers	118
2.5.4	Trust Region Policy Optimization	119
2.6	Deep Reinforcement Learning Foundations	120
2.6.1	Value Networks	120
2.6.2	Policy Networks	121
	References	122
3	Deep Reinforcement Learning Models	127
3.1	Value Function Approximation	127
3.1.1	Deep Q-learning	128
3.1.2	Double DQN	133
3.1.3	More Advanced DQN Methods	138
3.2	Policy Approximation	140
3.2.1	Deep Policy Gradient Method	141
3.2.2	Deep Actor-Critic Method	141
3.2.3	Asynchronous Advantage Actor-Critic	142
3.3	Deep Inverse Reinforcement Learning	146
3.3.1	Maximum Entropy Deep IRL	147
3.3.2	Adversarial Deep IRL	156
3.4	Deep Meta Reinforcement Learning	161
3.4.1	Task Sampling	162
3.4.2	RNN-based Deep MRL	163
3.4.3	Guided Meta-policy Search	167
3.5	Deep Hierarchical Reinforcement Learning	174
3.5.1	Temporal Abstraction and Intrinsic Motivation	175
3.5.2	FeUdal Networks	183
3.6	Deep Multi-task Reinforcement Learning	192
3.6.1	DISTRAL	193
3.6.2	IMPALA	203
3.6.3	Multi-task Deep Reinforcement Learning With Popart	221
3.7	Deep Multi-agent Reinforcement Learning	226
3.7.1	Automatic Learning of Communication Protocols	226
3.7.2	Monotonic Value Function Factorization	239
3.8	Adversarial Reinforcement Learning	257
3.8.1	Robust Adversarial Reinforcement Learning	257
3.8.2	Attacking Deep Reinforcement Learning	270
	References	281

4	Empirical Reinforcement Learning Systems	285
4.1	RL In Robotics	285
4.2	RL In Commerce	288
4.3	RL in Medications	290
4.4	RL In Education	293
4.5	RL In Civilization	295
4.6	RL In Simulations	298
4.7	Deep Reinforcement Learning Systems	300
	References	303
5	Performance Evaluations	307
5.1	Evaluation Metrics	307
5.2	Offline Performance Evaluation	311
5.2.1	On-Policy Evaluation	314
5.2.2	Off-Policy Evaluation	336
5.2.3	Simulation-Based Evaluation	361
5.2.4	Benchmarking	364
5.3	Online Performance Evaluation	365
5.3.1	Real-time Monitoring	366
5.3.2	A/B Testing	368
5.3.3	Testing Interleaving	370
5.4	Performance Emulator	372
5.5	Challenges and Best Practices	375
5.6	System-wise Performance Evaluation	376
6	Advanced Topics	383
6.1	Key Performance Challenges	383
6.1.1	Credit Assignment	384
6.1.2	Memorization	389
6.1.3	Reward Shaping	393
6.1.4	Policy Drift and the Solutions	397
6.1.5	Exploitation and Exploration Tradeoff	421
6.1.6	Interleaving	424
6.1.7	Dynamic Reinforcement Learning	429
6.1.8	Generalization	433
6.1.9	Advanced Optimization Methods	440
6.1.10	Learning Speedup	443
6.2	Responsive Reinforcement Learning	445
6.2.1	Fairness and Personalization	446
6.2.2	Sociality and Trust	451
6.2.3	Privacy Preservation	457
6.2.4	Safety and Security	460

7	Empiriccal Researches Ongoing	465
7.1	Adoptations of Neuroscience	465
7.2	Future of Deep Reinforcement Learning	465
7.3	Hardware Optimization	465
A	Appendix	467
A.1	Math Appendix	467