

1. Record Nr.	UNINA9910812040703321
Autore	Sadhu Arup Kumar
Titolo	Multi-agent coordination : a reinforcement learning approach // Arup Kumar Sadhu, Amit Konar
Pubbl/distr/stampa	Hoboken, NJ : , : Wiley : , : IEEE Press, , 2021
ISBN	1-119-69902-9 1-119-69899-5 1-119-69905-3
Descrizione fisica	1 PDF
Disciplina	006.31
Soggetti	Reinforcement learning Multiagent systems
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	PREFACE -- ACKNOWLEDGEMENT -- CHAPTER 1 INTRODUCTION: MULTI-AGENT COORDINATION BY REINFORCEMENT LEARNING AND EVOLUTIONARY ALGORITHMS 1 -- 1.1 INTRODUCTION 2 -- 1.2 SINGLE AGENT PLANNING 3 -- 1.2.1 Terminologies used in single agent planning 4 -- 1.2.2 Single agent search-based planning algorithms 9 -- 1.2.2.1 Dijkstra's algorithm 10 -- 1.2.2.2 A* (A-star) Algorithm 12 -- 1.2.2.3 D* (D-star) Algorithm 14 -- 1.2.2.4 Planning by STRIPS-like language 16 -- 1.2.3 Single agent reinforcement learning 16 -- 1.2.3.1 Multi-Armed Bandit Problem 17 -- 1.2.3.2 Dynamic programming and Bellman equation 19 -- 1.2.3.3 Correlation between reinforcement learning and Dynamic programming 20 -- 1.2.3.4 Single agent Q-learning 20 -- 1.2.3.5 Single agent planning using Q-learning 23 -- 1.3 MULTI-AGENT PLANNING AND COORDINATION 24 -- 1.3.1 Terminologies related to multi-agent coordination 24 -- 1.3.2 Classification of multi-agent system 25 -- 1.3.3 Game theory for multi-agent coordination 27 -- 1.3.3.1 Nash equilibrium (NE) 30 -- 1.3.3.1.1 Pure strategy NE (PSNE) 31 -- 1.3.3.1.2 Mixed strategy NE (MSNE) 33 -- 1.3.3.2 Correlated equilibrium (CE) 36 -- 1.3.3.3 Static game examples 37 -- 1.3.4 Correlation among RL, DP, and GT 39 -- 1.3.5 Classification of MARL 39 -- 1.3.5.1 Cooperative multi-agent reinforcement learning 41 -- 1.3.5.1.1 Static 41 -- Independent

Learner (IL) and Joint Action Learner (JAL)	41
Frequency maximum Q-value (FMQ) heuristic	44
1.3.5.1.2 Dynamic	46
Team-Q	46
Distributed -Q	47
Optimal Adaptive Learning	50
Sparse cooperative Q-learning (SCQL)	52
Sequential Q-learning (SQL)	53
Frequency of the maximum reward Q-learning (FMRQ)	53
1.3.5.2 Competitive multi-agent reinforcement learning	55
1.3.5.2.1 Minimax-Q Learning	55
1.3.5.2.2 Heuristically-accelerated multi-agent reinforcement learning	56
1.3.5.3 Mixed multi-agent reinforcement learning	57
1.3.5.3.1 Static	57
Belief-based Learning rule	57
Fictitious play	57
Meta strategy	58
Adapt When Everybody is Stationary, Otherwise Move to Equilibrium (AWESOME)	60
Hyper-Q	62
Direct policy search based	63
Fixed learning rate	63
Infinitesimal Gradient Ascent (IGA)	63
Generalized Infinitesimal Gradient Ascent (GIGA)	65
Variable learning rate	66
Win or Learn Fast-IGA (WoLF-IGA)	66
GIGA-Win or Learn Fast (GIGA-WoLF)	66
1.3.5.3.2 Dynamic	67
Equilibrium dependent	67
Nash-Q Learning	67
Correlated-Q Learning (CQL)	68
Asymmetric-Q Learning (AQL)	68
Friend-or-Foe Q-learning	70
Negotiation-based Q-learning	71
MAQL with equilibrium transfer	74
Equilibrium independent	76
Variable learning rate	76
Win or Learn Fast Policy hill-climbing (WoLF-PHC)	76
Policy Dynamic based Win or Learn Fast (PD-WoLF)	78
Fixed learning rate	78
Non-Stationary Converging Policies (NSCP)	78
Extended Optimal Response Learning (EXORL)	79
1.3.6 Coordination and planning by MAQL	80
1.3.7 Performance analysis of MAQL and MAQL-based coordination	81
1.4 COORDINATION BY OPTIMIZATION ALGORITHM	83
1.4.1 Particle Swarm Optimization (PSO) Algorithm	84
1.4.2 Firefly Algorithm (FA)	87
1.4.2.1 Initialization	87
1.4.2.2 Attraction to Brighter Fireflies	87
1.4.2.3 Movement of Fireflies	88
1.4.3 Imperialist Competitive Algorithm (ICA)	89
1.4.3.1 Initialization	89
1.4.3.2 Selection of Imperialists and Colonies	89
1.4.3.3 Formation of Empires	89
1.4.3.4 Assimilation of Colonies	90
1.4.3.5 Revolution	91
1.4.3.6 Imperialistic Competition	91
1.4.3.6.1 Total Empire Power Evaluation	91
1.4.3.6.2 Reassignment of Colonies and Removal of Empire	92
1.4.3.6.3 Union of Empires	92
1.4.4 Differential evolutionary (DE) algorithm	93
1.4.4.1 Initialization	93
1.4.4.2 Mutation	93
1.4.4.3 Recombination	93
1.4.4.4 Selection	93
1.4.5 Offline optimization	94
1.4.6 Performance analysis of optimization algorithms	94
1.4.6.1 Friedman test	94
1.4.6.2 Iman-Davenport test	95
1.5 SCOPE OF THE Book	95
1.6 SUMMARY	98
References	98
CHAPTER 2 IMPROVE CONVERGENCE SPEED OF MULTI-AGENT Q-LEARNING FOR COOPERATIVE TASK-PLANNING	107
2.1 INTRODUCTION	108
2.2 LITERATURE REVIEW	112
2.3 PRELIMINARIES	114
2.3.1 Single agent Q-learning	114
2.3.2 Multi-agent Q-learning	115
2.4 PROPOSED MULTI-AGENT Q-LEARNING	118
2.4.1 Two useful properties	119
2.5 PROPOSED FCMQL ALGORITHMS AND THEIR CONVERGENCE ANALYSIS	120
2.5.1 Proposed FCMQL algorithms	120
2.5.2 Convergence analysis of the proposed FCMQL algorithms	121
2.6 FCMQL-BASED COOPERATIVE MULTI-AGENT PLANNING	122
2.7 EXPERIMENTS AND RESULTS	123
2.8 CONCLUSIONS	130
2.9 SUMMARY	131
2.10 APPENDIX 2.1	131
2.11 APPENDIX 2.2	135
References	152
CHAPTER 3 CONSENSUS Q-LEARNING FOR MULTI-AGENT COOPERATIVE PLANNING	157
3.1 INTRODUCTION	158
3.2 PRELIMINARIES	159
3.2.1 Single agent Q-learning	159
3.2.2 Equilibrium-based multi-agent Q-learning	160
3.3 CONSENSUS	161
3.4 PROPOSED CONSENSUS Q-LEARNING AND PLANNING	162
3.4.1 Consensus Q-learning	162

-- 3.4.2 Consensus-based multi-robot planning	164 -- 3.5
EXPERIMENTS AND RESULTS	165 -- 3.5.1 Experimental setup 165 --
3.5.2 Experiments for CoQL	165 -- 3.5.3 Experiments for consensus-
based planning	166 -- 3.6 CONCLUSIONS 168 -- 3.7 SUMMARY 168 --
References	168 -- CHAPTER 4 AN EFFICIENT COMPUTING OF
CORRELATED EQUILIBRIUM FOR COOPERATIVE Q-LEARNING BASED	MULTI-AGENT PLANNING 171 -- 4.1 INTRODUCTION 172 -- 4.2
SINGLE-AGENT Q-LEARNING AND EQUILIBRIUM BASED MAQL	175 --
4.2.1 Single Agent Q learning	175 -- 4.2.2 Equilibrium based MAQL
175 -- 4.3 PROPOSED COOPERATIVE MULTI-AGENT Q-LEARNING AND	PLANNING 176 -- 4.3.1 Proposed schemes with their applicability 176
-- 4.3.2 Immediate rewards in Scheme-I and -II	177 -- 4.3.3 Scheme-I
induced MAQL	178 -- 4.3.4 Scheme-II induced MAQL 180 -- 4.3.5
Algorithms for scheme-I and II	182 -- 4.3.6 Constraint QL-I/ QL-II(C
.....	183 -- 4.3.7 Convergence 183 --
Multi-agent planning	185 -- 4.4 COMPLEXITY ANALYSIS 186 -- 4.4.1
Complexity of Correlated Q-Learning	187 -- 4.4.1.1 Space Complexity
187.	
4.4.1.2 Time Complexity	187 -- 4.4.2 Complexity of the proposed
algorithms	188 -- 4.4.2.1 Space Complexity 188 -- 4.4.2.2 Time
Complexity	188 -- 4.4.3 Complexity comparison 189 -- 4.4.3.1 Space
complexity	190 -- 4.4.3.2 Time complexity 190 -- 4.5 SIMULATION
AND EXPERIMENTAL RESULTS	191 -- 4.5.1 Experimental platform 191
-- 4.5.1.1 Simulation	191 -- 4.5.1.2 Hardware 192 -- 4.5.2
Experimental approach	192 -- 4.5.2.1 Learning phase 193 -- 4.5.2.2
Planning phase	193 -- 4.5.3 Experimental results 194 -- 4.6
CONCLUSION	201 -- 4.7 SUMMARY 202 -- 4.8 APPENDIX 203 --
References	209 -- CHAPTER 5 A MODIFIED IMPERIALIST COMPETITIVE
ALGORITHM FOR MULTI-AGENT STICK- CARRYING APPLICATION	213
-- 5.1 INTRODUCTION	214 -- 5.2 PROBLEM FORMULATION FOR
MULTI-ROBOT STICK-CARRYING	219 -- 5.3 PROPOSED HYBRID
ALGORITHM	222 -- 5.3.1 An Overview of Imperialist Competitive
Algorithm (ICA)	222 -- 5.3.1.1 Initialization 222 -- 5.3.1.2 Selection of
Imperialists and Colonies	223 -- 5.3.1.3 Formation of Empires 223 --
5.3.1.4 Assimilation of Colonies	223 -- 5.3.1.5 Revolution 224 --
5.3.1.6 Imperialistic Competition	224 -- 5.3.1.6.1 Total Empire Power
Evaluation	225 -- 5.3.1.6.2 Reassignment of Colonies and Removal of
Empire	225 -- 5.3.1.6.3 Union of Empires 226 -- 5.4 AN OVERVIEW OF
FIREFLY ALGORITHM (FA)	226 -- 5.4.1 Initialization 226 -- 5.4.2
Attraction to Brighter Fireflies	226 -- 5.4.3 Movement of Fireflies 227
-- 5.5 PROPOSED IMPERIALIST COMPETITIVE FIREFLY ALGORITHM	227
-- 5.5.1 Assimilation of Colonies	229 -- 5.5.1.1 Attraction to Powerful
Colonies	230 -- 5.5.1.2 Modification of Empire Behavior 230 -- 5.5.1.3
Union of Empires	230 -- 5.6 SIMULATION RESULTS 232 -- 5.6.1
Comparative Framework	232 -- 5.6.2 Parameter Settings 232 -- 5.6.3
Analysis on Explorative Power of ICFA	232 -- 5.6.4 Comparison of
Quality of the Final Solution	233 -- 5.6.5 Performance Analysis 233 --
5.7 COMPUTER SIMULATION AND EXPERIMENT	240 -- 5.7.1 Average
total path deviation (ATPD)	240 -- 5.7.2 Average Uncovered Target
Distance (AUTD)	241.
5.7.3 Experimental Setup in Simulation Environment	241 -- 5.7.4
Experimental Results in Simulation Environment	242 -- 5.7.5
Experimental Setup with Khepera Robots	244 -- 5.7.6 Experimental
Results with Khepera Robots	244 -- 5.8 CONCLUSION 245 -- 5.9
SUMMARY	247 -- 5.10 APPENDIX 5.1 248 -- References 249 --
CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS	255 -- 6.1
CONCLUSIONS	256 -- 6.2 FUTURE DIRECTIONS 257.

Agent Coordination. Chapter 1 introduces fundamentals of the multi-robot coordination. Chapter 2 offers two useful properties, which have been developed to speed-up the convergence of traditional multi-agent Q-learning (MAQL) algorithms in view of the team-goal exploration, where team-goal exploration refers to simultaneous exploration of individual goals. Chapter 3 proposes the novel consensus Q-learning (CoQL), which addresses the equilibrium selection problem. Chapter 4 introduces a new dimension in the literature of the traditional correlated Q-learning (CQL), in which correlated equilibrium (CE) is computed partly in the learning and the rest in the planning phases, thereby requiring CE computation once only. Chapter 5 proposes an alternative solution to the multi-agent planning problem using meta-heuristic optimization algorithms. Chapter 6 provides the concluding remarks based on the principles and experimental results acquired in the previous chapters. Possible future directions of research are also examined briefly at the end of the chapter."--
