

Contents

1	Multiagent Systems Under Uncertainty	1
1.1	Motivating Examples	2
1.2	Multiagent Systems	4
1.3	Uncertainty	6
1.4	Applications	7
2	The Decentralized POMDP Framework	11
2.1	Single-Agent Decision Frameworks	11
2.1.1	MDPs	12
2.1.2	POMDPs	13
2.2	Multiagent Decision Making: Decentralized POMDPs	14
2.3	Example Domains	17
2.3.1	Dec-Tiger	17
2.3.2	Multirobot Coordination: Recycling and Box-Pushing	19
2.3.3	Network Protocol Optimization	20
2.3.4	Efficient Sensor Networks	20
2.4	Special Cases, Generalizations and Related Models	21
2.4.1	Observability and Dec-MDPs	21
2.4.2	Factored Models	22
2.4.3	Centralized Models: MMDPs and MPOMDPs	24
2.4.4	Multiagent Decision Problems	25
2.4.5	Partially Observable Stochastic Games	30
2.4.6	Interactive POMDPs	30
3	Finite-Horizon Dec-POMDPs	33
3.1	Optimality Criteria	33
3.2	Policy Representations: Histories and Policies	34
3.2.1	Histories	34
3.2.2	Policies	35
3.3	Multiagent Beliefs	37
3.4	Value Functions for Joint Policies	37

3.5	Complexity	39
4	Exact Finite-Horizon Planning Methods	41
4.1	Backwards Approach: Dynamic Programming	41
4.1.1	Growing Policies from Subtree Policies	41
4.1.2	Dynamic Programming for Dec-POMDPs	44
4.2	Forward Approach: Heuristic Search	46
4.2.1	Temporal Structure in Policies: Decision Rules	46
4.2.2	Multiagent A*	47
4.3	Converting to a Non-observable MDP	48
4.3.1	The Plan-Time MDP and Optimal Value Function	49
4.3.2	Plan-Time Sufficient Statistics	49
4.3.3	An NOMDP Formulation	51
4.4	Other Finite-Horizon Methods	52
4.4.1	Point-Based DP	52
4.4.2	Optimization	52
5	Approximate and Heuristic Finite-Horizon Planning Methods	55
5.1	Approximation Methods	56
5.1.1	Bounded Dynamic Programming	56
5.1.2	Early Stopping of Heuristic Search	57
5.1.3	Application of POMDP Approximation Algorithms	57
5.2	Heuristic Methods	58
5.2.1	Alternating Maximization	58
5.2.2	Memory-Bounded Dynamic Programming	59
5.2.3	Approximate Heuristic-Search Methods	61
5.2.4	Evolutionary Methods and Cross-Entropy Optimization	64
6	Infinite-Horizon Dec-POMDPs	69
6.1	Optimality Criteria	69
6.1.1	Discounted Cumulative Reward	69
6.1.2	Average Reward	70
6.2	Policy Representation	70
6.2.1	Finite-State Controllers: Moore and Mealy	71
6.2.2	An Example Solution for DEC-TIGER	73
6.2.3	Randomization	74
6.2.4	Correlation Devices	74
6.3	Value Functions for Joint Policies	75
6.4	Undecidability, Alternative Goals and Their Complexity	76
7	Infinite-Horizon Planning Methods: Discounted Cumulative Reward	79
7.1	Policy Iteration	79
7.2	Optimizing Fixed-Size Controllers	81
7.2.1	Best-First Search	82
7.2.2	Bounded Policy Iteration	83
7.2.3	Nonlinear Programming	85

7.2.4	Expectation Maximization	85
7.2.5	Reduction to an NOMDP	87
8	Further Topics	91
8.1	Exploiting Structure in Factored Models	91
8.1.1	Exploiting Constraint Optimization Methods	91
8.1.1.1	Coordination (Hyper-)Graphs	91
8.1.1.2	ND-POMDPs	93
8.1.1.3	Factored Dec-POMDPs	95
8.1.2	Exploiting Influence-Based Policy Abstraction	100
8.2	Hierarchical Approaches and Macro-Actions	102
8.3	Communication	104
8.3.1	Implicit Communication and Explicit Communication	105
8.3.1.1	Explicit Communication Frameworks	106
8.3.1.2	Updating of Information States and Semantics	107
8.3.2	Delayed Communication	108
8.3.2.1	One-Step Delayed Communication	108
8.3.2.2	k -Steps Delayed Communication	109
8.3.3	Communication with Costs	111
8.3.4	Local Communication	112
8.4	Reinforcement Learning	113
9	Conclusion	115
	References	117

A Concise Introduction to Decentralized POMDPs

Oliehoek, F.A.; Amato, C.

2016, XX, 134 p. 36 illus., 22 illus. in color., Softcover

ISBN: 978-3-319-28927-4