

THE LOGIC OF ADAPTIVE BEHAVIOR

Knowledge Representation and Algorithms for
Adaptive Sequential Decision Making under Uncertainty
in First-Order and Relational Domains

Martijn van Otterlo

Department of Computer Science
Katholieke Universiteit Leuven, Belgium

IOS
Press

Amsterdam • Berlin • Oxford • Tokyo • Washington, DC

Contents

Preface	vii
CHAPTER 1 Introduction	1
1.1. Science and Engineering of Adaptive Behavior	3
1.1.1. Artificial Intelligence	3
1.1.2. Constructing Artificial Behavior	4
1.1.3. The Reinforcement Learning Paradigm	7
1.2. You Can Only Learn What You Can Represent	8
1.2.1. Generalization, Abstraction and Representation Formation	10
1.2.2. CANTOR: Representing the World in Snapshots	11
1.2.3. BOOLE: Representing the World in Twenty Questions	12
1.2.4. FREGE: Representing the World in Terms of Objects and Relations	14
1.2.5. The World Might be Larger than We See	16
1.3. About the Contents and Structure of this Book	17
1.3.1. Main Theme of This Book	17
1.3.2. A Road Map	24
1.3.3. Other Main Themes and Contributions	28
I Learning Sequential Decision Making under Uncertainty	29
CHAPTER 2 Markov Decision Processes: Concepts and Algorithms	31
2.1. Learning Sequential Decision Making	33
2.2. A Formal Framework	38
2.2.1. Markov Decision Processes.	38
2.2.2. Policies	40
2.2.3. Optimality Criteria and Discounting	41
2.3. Value Functions and Bellman Equations	42
2.4. Solving Markov decision processes	44
2.5. Dynamic Programming: Model-Based Solution Techniques	46
2.5.1. Fundamental DP Algorithms	46
2.5.2. Efficient DP Algorithms	50
2.6. Reinforcement Learning: Model-Free Solution Techniques	52
2.6.1. Temporal Difference Learning	54

2.6.2. Monte Carlo Methods	57
2.6.3. Efficient Exploration and Value Updating	58
2.7. Beyond the Markov Assumption	62
2.7.1. Partially Observable Markov Decision Processes	64
2.8. Discussion	66
CHAPTER 3 Generalization and Abstraction in Markov Decision Processes	69
3.1. From Algorithmic to Representational	71
3.1.1. Algorithmic Aspects	73
3.1.2. Fundamental Problems of Huge State Spaces	74
3.1.3. Representational Aspects	75
3.2. The Essence of Abstraction	76
3.2.1. Knowledge Representation	77
3.2.2. Definitions and Theories of Abstraction	79
3.2.3. Representation Change	79
3.3. Abstraction in the MDP Setting	81
3.3.1. Dimensions of MDP Abstractions	83
3.3.2. The PIAGET-Principle	87
3.3.3. Representations in MDP Abstractions	90
3.4. ABSTRACTION TYPE I: State Spaces	92
3.4.1. Model-Based State Abstractions	96
3.4.2. Model-Free State Abstractions	99
3.5. ABSTRACTION TYPE II: Factored Markov Decision Processes	99
3.5.1. Structured Representation	100
3.5.2. Structured Algorithms	102
3.6. ABSTRACTION TYPE III: Value Function Approximation	106
3.6.1. Fundamentals of Value Function Approximation	109
3.6.2. Architectures for VFA	113
3.7. ABSTRACTION TYPE IV: Searching in Policy Space	127
3.8. ABSTRACTION TYPE V: Hierarchical and Temporal Abstraction	129
3.8.1. Semi-Markov Decision Processes	131
3.8.2. Fixed Hierarchical Abstractions	132
3.8.3. Model-Minimization for SMDPs	136
3.8.4. Dynamic Hierarchical Abstractions	137
3.9. An Abstraction Case Study: Fingerprint Recognition	140
3.9.1. Reinforcement Learning for Minutiae Detection	141
3.9.2. Experimental Results	142
3.9.3. Benefits of Various Abstractions	143
3.10. Discussion	144
II Sequential Decisions in the First-Order Setting	151
CHAPTER 4 Reasoning, Learning and Acting in Worlds with Objects	153
4.1. The World Consists of Objects	159
4.1.1. Objects are Omnipresent and Indispensable	159
4.1.2. A Relational Domain: BLOCKS WORLD	163

4.1.3. Representing a World of Objects and Relations	165
4.2. Representation and Inference in First-Order Domains	171
4.2.1. First-Order Logic	172
4.2.2. Fragments and Extensions of FOL	177
4.2.3. First-Order Abstraction and Generalization	190
4.3. Learning in First-Order Domains	193
4.3.1. Obtaining Logical Abstractions	193
4.3.2. Inductive Logic Programming	195
4.3.3. Statistical Relational Learning	203
4.4. Acting in First-Order Domains	206
4.4.1. Formalizing and Modeling First-Order Domains	206
4.4.2. Two Characteristic Systems	211
4.4.3. Beyond Basic Action Theories	218
4.5. Learning Sequential Decision Making in Relational Domains	219
4.5.1. Lifting the MDP Framework to First-Order Domains	219
4.5.2. The PIAGET-Principle in First-Order Domains	231
4.5.3. Learning and Representation Tasks in Relational RL	240
4.5.4. What is Relational RL?: Different Viewpoints	242
4.6. Conclusions	243
CHAPTER 5 Model-Free Algorithms for Relational MDPs	247
5.1. Model-Free Relational Reinforcement Learning	248
5.1.1. Sampling and Structural Induction	250
5.1.2. Representations, and Value Functions vs. Policies	251
5.2. CARCASS: A Model-Free, Value-Based Approach	252
5.2.1. Relational Abstractions over RMDPs	252
5.2.2. <i>Q</i> -Learning for CARCASSs	257
5.2.3. Indirect Value Learning for CARCASSs using Approximate Models	259
5.2.4. Analysis and Experiments	261
5.2.5. Discussion	270
5.3. A Survey of Model-Free, Value-Based Approaches	271
5.3.1. Value-Based Learning on Fixed Abstraction Levels	271
5.3.2. Value-Based Learning using Dynamic Generalization	276
5.3.3. Discussion of Model-Free, Value-Based Techniques	283
5.4. GREY: Evolutionary Policy Search in Relational Domains	285
5.4.1. Evolutionary Search and ILP	285
5.4.2. GREY's Anatomy	287
5.4.3. Experimental Evaluation	291
5.5. A Survey of Policy-Based Model-Free Relational RL	297
5.5.1. Evolutionary Policy Search	297
5.5.2. Policy Search as Classification	298
5.5.3. Policy Gradient Approaches	300
5.6. Discussion	301
CHAPTER 6 Model-Based Algorithms for Relational MDPs	305
6.1. Intensional Dynamic Programming in Five Easy Steps	309
6.1.1. STEP I: Classical Dynamic Programming	310

6.1.2. STEP II: Replacing Tables by Sets	313
6.1.3. STEP III: Set-Based Value Functions	317
6.1.4. STEP IV: Set-Based Dynamic Programming	321
6.1.5. STEP V: Intensional Dynamic Programming	324
6.2. A Relational State Description Language	347
6.2.1. Abstract States	347
6.2.2. Abstract Actions	349
6.2.3. Rewards	350
6.2.4. Domain Theory and Constraint Handling	351
6.2.5. Markov Decision Programs, Value Functions and Policies	353
6.3. REBEL: Value Iteration for Markov Decision Programs	355
6.3.1. Overlaps, Regression and Weakest Preconditions	356
6.3.2. Combination and First-Order Decision-Theoretic Regression	361
6.3.3. Maximization: Computing Abstract State Values	364
6.3.4. Relational Bellman Backup Operator	365
6.3.5. Experiments	366
6.4. Logic Programming meets Dynamic Programming	374
6.4.1. Tabling	374
6.4.2. Policy Induction in REBEL	376
6.4.3. Other Extensions and Domain Theories	379
6.5. A Survey of Model-Based Approaches	381
6.5.1. Methods for exact IDP in First-Order Domains	381
6.5.2. Approximate Model-Based Methods for First-Order MDPs	386
6.5.3. Beyond the Markov Assumption	391
6.6. Discussion	392

III Implications, Challenges and Conclusions 395

CHAPTER 7 Sapience, Models and Hierarchy	397
7.1. Scaling Up	399
7.1.1. Extending Mental States	399
7.1.2. Declarative versus Procedural Representations	400
7.1.3. Learning vs. Reasoning	401
7.1.4. Examples of Existing Formalisms	402
7.2. Characterizing Sapient Agents	403
7.2.1. Cognitive Agents	404
7.2.2. Learning in Cognitive Agents	407
7.2.3. The Social Environment	409
7.2.4. Discussion of the Sapient Model of Agents	410
7.3. A Survey of Hierarchies, Models, Guidance and Transfer	412
7.3.1. Learning World Models	412
7.3.2. Bias, Guidance and Heuristics	415
7.3.3. Hierarchies	416
7.3.4. Transfer	417
7.3.5. Multi-Agent Approaches	419

7.4. Discussion	420
CHAPTER 8 Conclusions and Future Directions	423
8.1. Conclusions and Reflections	424
8.1.1. Main Argument	424
8.1.2. Contributions	428
8.1.3. Dimensions of First-Order MDPs and Solution Algorithms	429
8.2. Future Challenges	433
8.2.1. Upgrading the Complete Spectrum of RL Methods	433
8.2.2. Techniques Developed in this Book	433
8.2.3. Representational Aspects	435
8.2.4. Algorithmic Aspects	436
8.2.5. Theory	436
8.2.6. Agents, Cognitive Architectures, Reasoning and Transfer	439
8.2.7. Applications, Actions, Robots and Activities	439
8.2.8. Benchmarks and Toolboxes	441
8.2.9. Beyond the Markov Assumption	441
8.3. Concluding Remarks	442
Bibliography	443
List of Acronyms	477
Author Index	479