Table of Contents

Preface xix

About the Authors xxi

Acknowledgments xxv

1 Introduction 1

1.1 Decision Making 1

1.2 Example Applications 2

1.2.1 Traffic Alert and Collision Avoidance System 2

1.2.2 Unmanned Aircraft Persistent Surveillance 3

1.3 Methods for Designing Decision Agents 4

1.3.1 Explicit Programming 4

1.3.2 Supervised Learning 4

1.3.3 Optimization 5

1.3.4 Planning 5

1.3.5 Reinforcement Learning 5

1.4 Overview 5

1.5 Further Reading 7

References 7

I THEORY 9

2 Probabilistic Models 11

2.1 Representation 11

2.1.1 Degrees of Belief and Probability 12

2.1.2 Probability Distributions 13

2.1.3 Joint Distributions 16

2.1.4 Bayesian Network Representation 17

2.1.5 Conditional Independence 19
2.1.6 Hybrid Bayesian Networks 21
2.1.7 Temporal Models 23
2.2 Inference 25
2.2.1 Inference for Classification 26
2.2.2 Inference in Temporal Models 29
2.2.3 Exact Inference 30
2.2.4 Complexity of Exact Inference 33
2.2.5 Approximate Inference 35
2.3 Parameter Learning 40
2.3.1 Maximum Likelihood Parameter Learning 40
2.3.2 Bayesian Parameter Learning 42
2.3.3 Nonparametric Learning 45
2.4 Structure Learning 46
2.4.1 Bayesian Structure Scoring 46
2.4.2 Directed Graph Search 48
2.4.3 Markov Equivalence Classes 51
2.4.4 Partially Directed Graph Search 52
2.5 Summary 52
2.6 Further Reading 54
References 54

3 Decision Problems 57
3.1 Utility Theory 57
3.1.1 Constraints on Rational Preferences 58
3.1.2 Utility Functions 58

3.1.3 Maximum Expected Utility Principle 59
3.1.4 Utility Elicitation 60
3.1.5 Utility of Money 60
3.1.6 Multiple Variable Utility Functions 61
3.1.7 Irrationality 63
3.2 Decision Networks 64
3.2.1 Evaluating Decision Networks 65
3.2.2 Value of Information 66
3.2.3 Creating Decision Networks 67
3.3 Games 68
3.3.1 Dominant Strategy Equilibrium 69
3.3.2 Nash Equilibrium 70
3.3.3 Behavioral Game Theory 71
3.4 Summary 72
3.5 Further Reading 72
References 74

4 Sequential Problems 77
4.1 Formulation 77
4.1.1 Markov Decision Processes 77
4.1.2 Utility and Reward 78
4.2 Dynamic Programming 79
4.2.1 Policies and Utilities 79
4.2.2 Policy Evaluation 80
4.2.3 Policy Iteration 81
4.2.4 Value Iteration 81

4.2.5 Grid World Example 83
4.2.6 Asynchronous Value Iteration 84
4.2.7 Closed- and Open-Loop Planning 84
4.3 Structured Representations 89
4.3.1 Factored Markov Decision Processes 89
4.3.2 Structured Dynamic Programming 89
4.4 Linear Representations 91
4.5 Approximate Dynamic Programming 93
4.5.1 Local Approximation 93
4.5.2 Global Approximation 96
4.6 Online Methods 99
4.6.1 Forward Search 99
4.6.2 Branch and Bound Search 100
4.6.3 Sparse Sampling 101
4.6.4 Monte Carlo Tree Search 102
4.7 Direct Policy Search 103
4.7.1 Objective Function 104
4.7.2 Local Search Methods 104
4.7.3 Cross Entropy Methods 105
4.7.4 Evolutionary Methods 106
4.8 Summary 108
4.9 Further Reading 108
References 110

5 Model Uncertainty 113
5.1 Exploration and Exploitation 113

5.1.1 Multi-Armed Bandit Problems 113
5.1.2 Bayesian Model Estimation 114
5.1.3 Ad Hoc Exploration Strategies 115
5.1.4 Optimal Exploration Strategies 115
5.2 Maximum Likelihood Model-Based Methods 116
5.2.1 Randomized Updates 117
5.2.2 Prioritized Updates 118
5.3 Bayesian Model-Based Methods 118
5.3.1 Problem Structure 119
5.3.2 Beliefs over Model Parameters 119
5.3.3 Bayes-Adaptive Markov Decision Processes 120
5.3.4 Solution Methods 121
5.4 Model-Free Methods 121
5.4.1 Incremental Estimation 121
5.4.2 Q-Learning 122
5.4.3 Sarsa 123
5.4.4 Eligibility Traces 123
5.5 Generalization 124
5.5.1 Local Approximation 125
5.5.2 Global Approximation 126
5.5.3 Abstraction Methods 128
5.6 Summary 129
5.7 Further Reading 129
References 130

6 State Uncertainty 133

6.1 Formulation 133
6.1.1 Example Problem 133
6.1.2 Partially Observable Markov Decision Processes 134
6.1.3 Policy Execution 134
6.1.4 Belief-State Markov Decision Processes 134
6.2 Belief Updating 136
6.2.1 Discrete State Filter 136
6.2.2 Linear-Gaussian Filter 138
6.2.3 Particle Filter 138
6.3 Exact Solution Methods 140
6.3.1 Alpha Vectors 140
6.3.2 Conditional Plans 141
6.3.3 Value Iteration 143
6.4 Offline Methods 144
6.4.1 Fully Observable Value Approximation 144
6.4.2 Fast Informed Bound 144
6.4.3 Point-Based Value Iteration 145
6.4.4 Randomized Point-Based Value Iteration 146
6.4.5 Point Selection 147
6.4.6 Linear Policies 149
6.5 Online Methods 149
6.5.1 Lookahead with Approximate Value Function 149
6.5.2 Forward Search 150
6.5.3 Branch and Bound 151
6.5.4 Monte Carlo Tree Search 152

6.6 Summary 155
6.7 Further Reading 155
References 156

7 Cooperative Decision Making 159
7.1 Formulation 159
7.1.1 Decentralized POMDPs 159
7.1.2 Example Problem 161
7.1.3 Solution Representations 162
7.2 Properties 164
7.2.1 Differences with POMDPs 164
7.2.2 Dec-POMDP Complexity 165
7.2.3 Generalized Belief States 165
7.3 Notable Subclasses 166
7.3.1 Dec-MDPs 166
7.3.2 ND-POMDPs 168
7.3.3 MMDPs 169
7.4 Exact Solution Methods 170
7.4.1 Dynamic Programming 170
7.4.2 Heuristic Search 172
7.4.3 Policy Iteration 175
7.5 Approximate Solution Methods 177
7.5.1 Memory-Bounded Dynamic Programming 177
7.5.2 Joint Equilibrium Search 178
7.6 Communication 178
7.7 Summary 180

7.8 Further Reading 180
References 182

II APPLICATION 189

8 Probabilistic Surveillance Video Search 191
8.1 Attribute-Based Person Search 191
8.1.1 Applications 192
8.1.2 Person Detection 193
8.1.3 Retrieval and Scoring 194
8.2 Probabilistic Appearance Model 195
8.2.1 Observed States 195
8.2.2 Basic Model Structure 197
8.2.3 Model Extensions 202
8.3 Learning and Inference Techniques 206
8.3.1 Parameter Learning 207
8.3.2 Hidden State Inference 211
8.3.3 Scoring Algorithm 214
8.4 Performance 217
8.4.1 Search Accuracy 217
8.4.2 Search Timing 220
8.5 Interactive Search Tool 223
8.6 Summary 225
References 227

9 Dynamic Models for Speech Applications 229
9.1 Modeling Speech Signals 229

9.1.1 Feature Extraction 230
9.1.2 Hidden Markov Models 230
9.1.3 Gaussian Mixture Models 231
9.1.4 Expectation-Maximization Algorithm 232
9.2 Speech Recognition 232
9.3 Topic Identification 235
9.4 Language Recognition 236
9.5 Speaker Identification 238
9.5.1 Forensic Speaker Recognition 240
9.6 Machine Translation 242
9.7 Summary 243
References 243

10 Optimized Airborne Collision Avoidance 249
10.1 Airborne Collision Avoidance Systems 249
10.1.1 Traffic Alert and Collision Avoidance System 250
10.1.2 Limitations of Existing System 251
10.1.3 Unmanned Aircraft Sense and Avoid 252
10.1.4 Airborne Collision Avoidance System X 253
10.2 Collision Avoidance Problem Formulation 253
10.2.1 Resolution Advisories 253
10.2.2 Dynamic Model 255
10.2.3 Reward Function 256
10.2.4 Dynamic Programming 258
10.3 State Estimation 259
10.3.1 Sensor Error 259

10.3.2 Pilot Response 260
10.3.3 Time to Potential Collision 260
10.4 Real-Time Execution 261
10.4.1 Online Costs 261
10.4.2 Multiple Threats 262
10.4.3 Traffic Alerts 263
10.5 Evaluation 265
10.5.1 Safety Analysis 265
10.5.2 Operational Suitability and Acceptability 267
10.5.3 Parameter Tuning 271
10.5.4 Flight Test 272
10.6 Summary 273
References 274

11 Multiagent Planning for Persistent Surveillance 277
11.1 Mission Description 277
11.2 Centralized Problem Formulation 278
11.2.1 State Space 278
11.2.2 Action Space 279
11.2.3 State Transition Model 279
11.2.4 Reward Function 280
11.3 Decentralized Approximate Formulations 280
11.3.1 Factored Decomposition 280
11.3.2 Group Aggregate Decomposition 281
11.3.3 Planning 281
11.4 Model Learning 282

11.5 Flight Test 285
11.6 Summary 286
References 289

12 Integrating Automation with Humans 291
12.1 Human Capabilities and Coping 291
12.1.1 Perceptual and Cognitive Capabilities 291
12.1.2 Naturalistic Decision Making 294
12.2 Considering the Human in Design 296
12.2.1 Trust and Value of Decision Logic Transparency 296
12.2.2 Designing for Different Levels of Certainty 300
12.2.3 Supporting Decisions over Long Timescales 305
12.3 A Systems View of Implementation 308
12.3.1 Interface, Training, and Procedures 308
12.3.2 Measuring Decision Support Effectiveness 311
12.3.3 Organization Influences on System Effectiveness 313
12.4 Summary 313
References 314

Index 317