

Artificial Intelligence

A Modern Approach

Third Edition

Stuart J. Russell and Peter Norvig

Contributing writers:

Ernest Davis
Douglas D. Edwards
David Forsyth
Nicholas J. Hay
Jitendra M. Malik
Vibhu Mittal
Mehran Sahami
Sebastian Thrun

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Contents

I Artificial Intelligence

1 Introduction	1
1.1 What Is AI?	1
1.2 The Foundations of Artificial Intelligence	5
1.3 The History of Artificial Intelligence	16
1.4 The State of the Art	28
1.5 Summary, Bibliographical and Historical Notes, Exercises	29
2 Intelligent Agents	34
2.1 Agents and Environments	34
2.2 Good Behavior: The Concept of Rationality	36
2.3 The Nature of Environments	40
2.4 The Structure of Agents	46
2.5 Summary, Bibliographical and Historical Notes, Exercises	59

II Problem-solving

3 Solving Problems by Searching	64
3.1 Problem-Solving Agents	64
3.2 Example Problems	69
3.3 Searching for Solutions	75
3.4 Uninformed Search Strategies	81
3.5 Informed (Heuristic) Search Strategies	92
3.6 Heuristic Functions	102
3.7 Summary, Bibliographical and Historical Notes, Exercises	108
4 Beyond Classical Search	120
4.1 Local Search Algorithms and Optimization Problems	120
4.2 Local Search in Continuous Spaces	129
4.3 Searching with Nondeterministic Actions	133
4.4 Searching with Partial Observations	138
4.5 Online Search Agents and Unknown Environments	147
4.6 Summary, Bibliographical and Historical Notes, Exercises	153
5 Adversarial Search	161
5.1 Games	161
5.2 Optimal Decisions in Games	163
5.3 Alpha-Beta Pruning	167
5.4 Imperfect Real-Time Decisions	171
5.5 Stochastic Games	177

5.6	Partially Observable Games	180
5.7	State-of-the-Art Game Programs	185
5.8	Alternative Approaches	187
5.9	Summary, Bibliographical and Historical Notes, Exercises	189
6	Constraint Satisfaction Problems	202
6.1	Defining Constraint Satisfaction Problems	202
6.2	Constraint Propagation: Inference in CSPs	208
6.3	Backtracking Search for CSPs	214
6.4	Local Search for CSPs	220
6.5	The Structure of Problems	222
6.6	Summary, Bibliographical and Historical Notes, Exercises	227
 III Knowledge, reasoning, and planning		
7	Logical Agents	234
7.1	Knowledge-Based Agents	235
7.2	The Wumpus World	236
7.3	Logic	240
7.4	Propositional Logic: A Very Simple Logic	243
7.5	Propositional Theorem Proving	249
7.6	Effective Propositional Model Checking	259
7.7	Agents Based on Propositional Logic	265
7.8	Summary, Bibliographical and Historical Notes, Exercises	274
8	First-Order Logic	285
8.1	Representation Revisited	285
8.2	Syntax and Semantics of First-Order Logic	290
8.3	Using First-Order Logic	300
8.4	Knowledge Engineering in First-Order Logic	307
8.5	Summary, Bibliographical and Historical Notes, Exercises	313
9	Inference in First-Order Logic	322
9.1	Propositional vs. First-Order Inference	322
9.2	Unification and Lifting	325
9.3	Forward Chaining	330
9.4	Backward Chaining	337
9.5	Resolution	345
9.6	Summary, Bibliographical and Historical Notes, Exercises	357
10	Classical Planning	366
10.1	Definition of Classical Planning	366
10.2	Algorithms for Planning as State-Space Search	373
10.3	Planning Graphs	379

10.4	Other Classical Planning Approaches	387
10.5	Analysis of Planning Approaches	392
10.6	Summary, Bibliographical and Historical Notes, Exercises	393
11	Planning and Acting in the Real World	401
11.1	Time, Schedules, and Resources	401
11.2	Hierarchical Planning	406
11.3	Planning and Acting in Nondeterministic Domains	415
11.4	Multiagent Planning	425
11.5	Summary, Bibliographical and Historical Notes, Exercises	430
12	Knowledge Representation	437
12.1	Ontological Engineering	437
12.2	Categories and Objects	440
12.3	Events	446
12.4	Mental Events and Mental Objects	450
12.5	Reasoning Systems for Categories	453
12.6	Reasoning with Default Information	458
12.7	The Internet Shopping World	462
12.8	Summary, Bibliographical and Historical Notes, Exercises	467
IV	Uncertain knowledge and reasoning	
13	Quantifying Uncertainty	480
13.1	Acting under Uncertainty	480
13.2	Basic Probability Notation	483
13.3	Inference Using Full Joint Distributions	490
13.4	Independence	494
13.5	Bayes' Rule and Its Use	495
13.6	The Wumpus World Revisited	499
13.7	Summary, Bibliographical and Historical Notes, Exercises	503
14	Probabilistic Reasoning	510
14.1	Representing Knowledge in an Uncertain Domain	510
14.2	The Semantics of Bayesian Networks	513
14.3	Efficient Representation of Conditional Distributions	518
14.4	Exact Inference in Bayesian Networks	522
14.5	Approximate Inference in Bayesian Networks	530
14.6	Relational and First-Order Probability Models	539
14.7	Other Approaches to Uncertain Reasoning	546
14.8	Summary, Bibliographical and Historical Notes, Exercises	551
15	Probabilistic Reasoning over Time	566
15.1	Time and Uncertainty	566

15.2	Inference in Temporal Models	570
15.3	Hidden Markov Models	578
15.4	Kalman Filters	584
15.5	Dynamic Bayesian Networks	590
15.6	Keeping Track of Many Objects	599
15.7	Summary, Bibliographical and Historical Notes, Exercises	603
16	Making Simple Decisions	610
16.1	Combining Beliefs and Desires under Uncertainty	610
16.2	The Basis of Utility Theory	611
16.3	Utility Functions	615
16.4	Multiatribute Utility Functions	622
16.5	Decision Networks	626
16.6	The Value of Information	628
16.7	Decision-Theoretic Expert Systems	633
16.8	Summary, Bibliographical and Historical Notes, Exercises	636
17	Making Complex Decisions	645
17.1	Sequential Decision Problems	645
17.2	Value Iteration	652
17.3	Policy Iteration	656
17.4	Partially Observable MDPs	658
17.5	Decisions with Multiple Agents: Game Theory	666
17.6	Mechanism Design	679
17.7	Summary, Bibliographical and Historical Notes, Exercises	684
V	Learning	
18	Learning from Examples	693
18.1	Forms of Learning	693
18.2	Supervised Learning	695
18.3	Learning Decision Trees	697
18.4	Evaluating and Choosing the Best Hypothesis	708
18.5	The Theory of Learning	713
18.6	Regression and Classification with Linear Models	717
18.7	Artificial Neural Networks	727
18.8	Nonparametric Models	737
18.9	Support Vector Machines	744
18.10	Ensemble Learning	748
18.11	Practical Machine Learning	753
18.12	Summary, Bibliographical and Historical Notes, Exercises	757
19	Knowledge in Learning	768
19.1	A Logical Formulation of Learning	768

19.2	Knowledge in Learning	777
19.3	Explanation-Based Learning	780
19.4	Learning Using Relevance Information	784
19.5	Inductive Logic Programming	788
19.6	Summary, Bibliographical and Historical Notes, Exercises	797
20	Learning Probabilistic Models	802
20.1	Statistical Learning	802
20.2	Learning with Complete Data	806
20.3	Learning with Hidden Variables: The EM Algorithm	816
20.4	Summary, Bibliographical and Historical Notes, Exercises	825
21	Reinforcement Learning	830
21.1	Introduction	830
21.2	Passive Reinforcement Learning	832
21.3	Active Reinforcement Learning	839
21.4	Generalization in Reinforcement Learning	845
21.5	Policy Search	848
21.6	Applications of Reinforcement Learning	850
21.7	Summary, Bibliographical and Historical Notes, Exercises	853
VI	Communicating, perceiving, and acting	
22	Natural Language Processing	860
22.1	Language Models	860
22.2	Text Classification	865
22.3	Information Retrieval	867
22.4	Information Extraction	873
22.5	Summary, Bibliographical and Historical Notes, Exercises	882
23	Natural Language for Communication	888
23.1	Phrase Structure Grammars	888
23.2	Syntactic Analysis (Parsing)	892
23.3	Augmented Grammars and Semantic Interpretation	897
23.4	Machine Translation	907
23.5	Speech Recognition	912
23.6	Summary, Bibliographical and Historical Notes, Exercises	918
24	Perception	928
24.1	Image Formation	929
24.2	Early Image-Processing Operations	935
24.3	Object Recognition by Appearance	942
24.4	Reconstructing the 3D World	947
24.5	Object Recognition from Structural Information	957

24.6	Using Vision	961
24.7	Summary, Bibliographical and Historical Notes, Exercises	965
25	Robotics	971
25.1	Introduction	971
25.2	Robot Hardware	973
25.3	Robotic Perception	978
25.4	Planning to Move	986
25.5	Planning Uncertain Movements	993
25.6	Moving	997
25.7	Robotic Software Architectures	1003
25.8	Application Domains	1006
25.9	Summary, Bibliographical and Historical Notes, Exercises	1010
VII	Conclusions	
26	Philosophical Foundations	1020
26.1	Weak AI: Can Machines Act Intelligently?	1020
26.2	Strong AI: Can Machines Really Think?	1026
26.3	The Ethics and Risks of Developing Artificial Intelligence	1034
26.4	Summary, Bibliographical and Historical Notes, Exercises	1040
27	AI: The Present and Future	1044
27.1	Agent Components	1044
27.2	Agent Architectures	1047
27.3	Are We Going in the Right Direction?	1049
27.4	What If AI Does Succeed?	1051
A	Mathematical background	1053
A.1	Complexity Analysis and $O()$ Notation	1053
A.2	Vectors, Matrices, and Linear Algebra	1055
A.3	Probability Distributions	1057
B	Notes on Languages and Algorithms	1060
B.1	Defining Languages with Backus-Naur Form (BNF)	1060
B.2	Describing Algorithms with Pseudocode	1061
B.3	Online Help	1062
	Bibliography	1063
	Index	1095