

Contents

Preface	xiii
1 Artificial Intelligence	1
1.1 Artificial Intelligence and Intelligence	1
1.1.1 Intelligence	1
1.1.2 Thinking	2
1.1.3 The Turing Test for Thinking	3
1.1.4 The Chinese Room Argument	3
1.1.5 Consciousness and Quantum Mechanics	4
1.1.6 Dualism	4
1.2 Association	5
1.3 Neural Networking	9
1.3.1 Artificial Neural Networks	9
1.3.2 Biological Neural Networks	11
1.4 Symbol Processing	12
1.5 Heuristic Search	14
1.6 The Problems with AI	16
1.7 The New Proposals	16
1.7.1 Real Numbers	17
1.7.2 Picture Processing	18
1.7.3 Memories	18
1.7.4 Quantum Mechanics	19
1.8 The Organization of the Book	20
1.9 Exercises	20
2 Pattern Recognition I	23
2.1 A Simple Pattern Recognition Algorithm	23
2.2 A Short Description of the Neocognitron	26
2.2.1 Detecting Short Lines	27
2.2.2 A Typical Neocognitron	29
2.2.3 Training the Neocognitron	29
2.2.4 Some Results	33
2.3 Recognizing Words	33
2.4 Expanding the Pattern Recognition Hierarchy	43
2.4.1 Hearing	43
2.4.2 Higher Levels	44
2.4.3 The Hierarchy	45

2.4.4	On the Hierarchy	48
2.5	Additional Perspective	48
2.5.1	Other Systems	48
2.5.2	Realism	49
2.5.3	Bigger Problems	49
2.6	Exercises	50
3	Pattern Recognition II	55
3.1	Mathematics, Pattern Recognition, and the Linear Pattern Classifier	55
3.1.1	The Linear Pattern Classifier	55
3.1.2	ADALINEs and MADELINEs	59
3.1.3	Perceptrons	59
3.2	Separating Nonlinearly Separable Classes	59
3.2.1	The Nearest Neighbor Algorithm	59
3.2.2	Learning Vector Quantization Methods	60
3.3	Hopfield Networks	63
3.3.1	The Hopfield Network	63
3.3.2	Storing Patterns	65
3.3.3	The Boltzman Machine	67
3.3.4	Pattern Recognition	72
3.3.5	Harmony	72
3.3.6	Comparison with Human Thinking	72
3.4	Back-Propagation	74
3.4.1	History	74
3.4.2	The Network	74
3.4.3	Computing the Weights	76
3.4.4	Speeding Up Back-Propagation	79
3.4.5	Dealing with Local Minima	80
3.4.6	Using Back-Propagation to Train Hopfield/Boltzman Networks	80
3.5	Pattern Recognition and Curve Fitting	81
3.5.1	Pattern Recognition as Curve Fitting	81
3.5.2	Approximating Real-Valued Functions	84
3.5.3	Overfitting	84
3.6	Associative Memory and Generalization	84
3.6.1	Associative Memory	86
3.6.2	Local and Distributed Representations	90
3.6.3	Reasoning within a Network	91
3.7	Applications of Back-Propagation	92
3.7.1	Interpreting Sonar Returns	93
3.7.2	Reading Text	94
3.7.3	Speech Recognition	95
3.7.4	Detecting Bombs	95
3.7.5	Economic Analysis	96
3.7.6	Learning to Drive	97
3.7.7	DNA Analysis	97
3.8	Additional Perspective	97

3.9	Exercises	98
4	Rule-Based Methods	103
4.1	Introduction	103
4.2	Some Elementary Prolog	103
4.2.1	Stating Facts	104
4.2.2	Syntax	104
4.2.3	Asking Questions	105
4.2.4	Rules	108
4.2.5	Recursion	109
4.2.6	List Processing	110
4.2.7	Other Predicates	114
4.3	Rules and Basic Rule Interpretation Methods	115
4.3.1	A Small Rule-Based System	116
4.3.2	Forward Chaining	118
4.3.3	Backward Chaining	119
4.4	Conflict Resolution	120
4.5	More Sophisticated Rule Interpretation	124
4.5.1	Dealing with Incomplete Data by Asking Questions	124
4.5.2	Other Activation Functions	125
4.5.3	Uncertain Input	126
4.5.4	Extra Facilities for Rule Interpreters	126
4.6	The Famous Expert Systems	127
4.6.1	DENDRAL	127
4.6.2	MYCIN	129
4.6.3	PROSPECTOR	131
4.6.4	ACE	133
4.6.5	XCON	133
4.7	Learning Rules in SOAR	139
4.7.1	A Searching Example	139
4.7.2	The Power Law of Practice	140
4.8	Rules versus Networks	141
4.9	Exercises	143
5	Logic	147
5.1	Standard Form and Clausal Form	147
5.2	Basic Inference Rules	151
5.2.1	Inference Rules	151
5.2.2	Clauses with Variables	152
5.3	Controlling Search	154
5.3.1	The Problem with Blind Searching	154
5.3.2	Proof by Contradiction	155
5.3.3	The Set-of-Support Strategy	156
5.3.4	Weighting	157
5.3.5	Prolog's Strategy	158
5.4	An Example Using Otter	159

5.4.1	The Problem	159
5.5	The Usefulness of Predicate Calculus	163
5.6	Other Reasoning Methods	163
5.7	Exercises	164
6	Complex Architectures	167
6.1	The Basic Human Architecture	167
6.2	Flow of Control	169
6.3	The Virtual Symbol Processing Machine Proposal	172
6.4	Mental Representation and Computer Representation	173
6.4.1	A Problem with Symbolic Representation	173
6.4.2	Symbol Grounding as a Solution	175
6.4.3	Structure and Operations on Structures	176
6.5	Storing Sequential Events	177
6.5.1	The Symbolic Solution	177
6.5.2	Neural Solutions	178
6.6	Structuring Individual Thoughts	181
6.6.1	The Symbolic Methods	181
6.6.2	Neural Methods	183
6.7	Frames and Scripts	187
6.7.1	Schemas and Frames	187
6.7.2	Scripts	189
6.8	Exercises	190
7	Case-Based and Memory-Based Reasoning	193
7.1	Condensed versus Uncondensed Knowledge	193
7.1.1	Arguments For Condensed Knowledge	195
7.1.2	Arguments Against Condensed Knowledge	196
7.1.3	Problems with Condensed Representations	197
7.2	Memory-Based Reasoning	199
7.2.1	A Simple Example	199
7.2.2	MBRtalk	200
7.2.3	A HERBIE Solution to Reading	201
7.2.4	JOHNNY	202
7.2.5	PACE	204
7.3	Case-Based Reasoning	204
7.3.1	Case-Based Reasoning in People	204
7.3.2	CHEF	207
7.4	Other Case-Based Programs	211
7.5	Exercises	211
8	Problem Solving and Heuristic Search	213
8.1	The 8-Puzzle	213
8.1.1	The Blind Search Methods	214
8.1.2	Heuristic Searches	216
8.1.3	Other Methods	220

8.2	A Geometry Theorem Prover	220
8.3	Symbolic Integration and Heuristic Search	228
8.3.1	SAINT	228
8.3.2	A Symbolic Program to Learn Integration	231
8.3.3	A Partial Back-Propagation Solution	231
8.4	Other Heuristic Programs	234
8.5	Exercises	235
9	Game Playing	237
9.1	General Game Playing Techniques	237
9.1.1	Minimax	237
9.1.2	More Sophisticated Searching Methods	240
9.1.3	Using Experience	241
9.2	Checkers	244
9.2.1	Rote Learning	245
9.2.2	Generalization Learning	247
9.2.3	Samuel's Later Work	250
9.2.4	Chinook	251
9.3	Backgammon	252
9.3.1	Berliner's BKG Program	252
9.3.2	Backgammon using Back-Propagation	253
9.3.3	A Second Back-Propagation Approach	255
9.3.4	Temporal Difference Learning	257
9.4	Exercises	258
10	Natural Language Processing	261
10.1	Formal Languages	262
10.2	The Transition Network Grammar	267
10.2.1	A Simple Transition Network	267
10.2.2	A Prolog Implementation	269
10.2.3	A Neural Analog	271
10.2.4	Syntax is not Enough	275
10.3	Semantics-Based Methods	276
10.3.1	Semantic Grammar	276
10.3.2	Conceptual Dependency Notation	277
10.4	Scripts and Short Stories	287
10.5	A Neural-Network-Based Approach	292
10.6	Defining Words by the Way they are Used	298
10.7	A Recurrent Network for Sentences	302
10.8	Neural-Based Scripts	305
10.9	Learning the Past Tense of Verbs	310
10.9.1	Over-Regularization	312
10.9.2	The Rumelhart and McClelland Network	312
10.9.3	The Classical Rule-Based Model	315
10.9.4	A Hybrid Model	317
10.10	Other Positions on Language	317

10.11 Exercises	319
Afterword	325
A Appendix A	327
A.1 A Derivation of Back-Propagation	327
A.1.1 The Delta Rule	327
A.1.2 The Generalized Delta Rule, or Back-Propagation	329
Glossary	331
Bibliography	337
Index	357