

Contents

Foreword	xix
Preface	xx v
Acknowledgments.	xxvii
1 Introduction and Overview	1
1.1 Computational Cognitive Neuroscience	1
1.2 Basic Motivations for Computational Cognitive Neuroscience	2
1.2.1 Physical Reductionism	2
1.2.2 Reconstructionism	3
1.2.3 Levels of Analysis	4
1.2.4 Scaling Issues.	6
1.3 Historical Context	8
1.4 Overview of Our Approach	10
1.5 General Issues in Computational Modeling	11
1.6 Motivating Cognitive Phenomena and Their Biological Bases	14
1.6.1 Parallelism	15
1.6.2 Gradedness	15
1.6.3 Interactivity	17
1.6.4 Competition	17
1.6.5 Learning	18
1.7 Organization of the Book	19
1.8 Further Reading	20
I Basic Neural Computational Mechanisms	21
2 Individual Neurons	23
2.1 Overview	23
2.2 Detectors: How to Think About a Neuron	24
2.2.1 Understanding the Parts of the Neuron Using the Detector Model	26
2.3 The Biology of the Neuron	27
2.3.1 The Axon	29

2.3.2	The Synapse	29
2.3.3	The Dendrite	31
2.4	The Electrophysiology of the Neuron	32
2.4.1	Basic Electricity	32
2.4.2	Diffusion	33
2.4.3	Electric Potential versus Diffusion: The Equilibrium Potential	34
2.4.4	The Neural Environs and Ions	35
2.4.5	Putting It All Together: Integration	37
2.4.6	The Equilibrium Membrane Potential	38
2.4.7	Summary	40
2.5	Computational Implementation of the Neural Activation Function	40
2.5.1	Computing Input Conductances	42
2.5.2	Point Neuron Parameter Values	45
2.5.3	The Discrete Spiking Output Function	45
2.5.4	The Rate Code Output Function	46
2.5.5	Summary	48
2.6	Explorations of the Individual Neuron	48
2.6.1	The Membrane Potential	49
2.6.2	The Activation Output	53
2.6.3	The Neuron as Detector	54
2.7	Hypothesis Testing Analysis of a Neural Detector	58
2.7.1	Objective Probabilities and Example	59
2.7.2	Subjective Probabilities	62
2.7.3	Similarity of V_m and $P(h d)$	65
2.8	The Importance of Keeping It Simple	65
2.9	Self-Regulation: Accommodation and Hysteresis	66
2.9.1	Implementation of Accommodation and Hysteresis	67
2.10	Summary	69
2.11	Further Reading	70
3	Networks of Neurons	71
3.1	Overview	71
3.2	General Structure of Cortical Networks	72
3.3	Unidirectional Excitatory Interactions: Transformations	75
3.3.1	Exploration of Transformations	79
3.3.2	Localist versus Distributed Representations	82
3.3.3	Exploration of Distributed Representations	84
3.4	Bidirectional Excitatory Interactions	85
3.4.1	Bidirectional Transformations	86
3.4.2	Bidirectional Pattern Completion	87
3.4.3	Bidirectional Amplification	89
3.4.4	Attractor Dynamics	92

3.5	Inhibitory Interactions	93
3.5.1	General Functional Benefits of Inhibition	94
3.5.2	Exploration of Feedforward and Feedback Inhibition	95
3.5.3	The k-Winners-Take-All Inhibitory Functions	100
3.5.4	Exploration of kWTA Inhibition	103
3.5.5	Digits Revisited with kWTA Inhibition	104
3.5.6	Other Simple Inhibition Functions	105
3.6	Constraint Satisfaction	106
3.6.1	Attractors Again	108
3.6.2	The Role of Noise	108
3.6.3	The Role of Inhibition	109
3.6.4	Explorations of Constraint Satisfaction: Cats and Dogs	109
3.6.5	Explorations of Constraint Satisfaction: Necker Cube	111
3.7	Summary	112
3.8	Further Reading	114
4	Hebbian Model Learning	115
4.1	Overview	115
4.2	Biological Mechanisms of Learning	116
4.3	Computational Objectives of Learning	118
4.3.1	Simple Exploration of Correlational Model Learning	121
4.4	Principal Components Analysis	122
4.4.1	Simple Hebbian PCA in One Linear Unit	122
4.4.2	Oja's Normalized Hebbian PCA	124
4.5	Conditional Principal Components Analysis	125
4.5.1	The CPCA Learning Rule	127
4.5.2	Derivation of CPCA Learning Rule	128
4.5.3	Biological Implementation of CPCA Hebbian Learning	129
4.6	Exploration of Hebbian Model Learning	130
4.7	Renormalization and Contrast Enhancement	132
4.7.1	Renormalization	133
4.7.2	Contrast Enhancement	134
4.7.3	Exploration of Renormalization and Contrast Enhancement in CPCA	135
4.8	Self-Organizing Model Learning	137
4.8.1	Exploration of Self-Organizing Learning	138
4.8.2	Summary and Discussion	142
4.9	Other Approaches to Model Learning	142
4.9.1	Algorithms That Use CPCA-Style Hebbian Learning	143
4.9.2	Clustering	143
4.9.3	Topography	143
4.9.4	Information Maximization and MDL	144
4.9.5	Learning Based Primarily on Hidden Layer Constraints	144

4.9.6 Generative Models	145
4.10 Summary	145
4.11 Further Reading	146
5 Error-Driven Task Learning	147
5.1 Overview	147
5.2 Exploration of Hebbian Task Learning	148
5.3 Using Error to Learn: The Delta Rule	150
5.3.1 Deriving the Delta Rule	152
5.3.2 Learning Bias Weights	152
5.4 Error Functions, Weight Bounding, and Activation Phases	154
5.4.1 Cross Entropy Error	154
5.4.2 Soft Weight Bounding	155
5.4.3 Activation Phases in Learning	156
5.5 Exploration of Delta Rule Task Learning	156
5.6 The Generalized Delta Rule: Backpropagation	158
5.6.1 Derivation of Backpropagation	160
5.6.2 Generic Recursive Formulation	161
5.6.3 The Biological Implausibility of Backpropagation	162
5.7 The Generalized Recirculation Algorithm	162
5.7.1 Derivation of GeneRec	163
5.7.2 Symmetry, Midpoint, and CHL	165
5.8 Biological Considerations for GeneRec	166
5.8.1 Weight Symmetry in the Cortex	166
5.8.2 Phase-Based Activations in the Cortex	167
5.8.3 Synaptic Modification Mechanisms	168
5.9 Exploration of GeneRec-Based Task Learning	170
5.10 Summary	171
5.11 Further Reading	172
6 Combined Model and Task Learning, and Other Mechanisms	173
6.1 Overview	173
6.2 Combined Hebbian and Error-driven Learning	173
6.2.1 Pros and Cons of Hebbian and Error-Driven Learning	174
6.2.2 Advantages to Combining Hebbian and Error-Driven Learning	175
6.2.3 Inhibitory Competition as a Model-Learning Constraint	175
6.2.4 Implementation of Combined Model and Task Learning	176
6.2.5 Summary	177
6.3 Generalization in Bidirectional Networks	178
6.3.1 Exploration of Generalization	179
6.4 Learning to Re-represent in Deep Networks	181
6.4.1 Exploration of a Deep Network	183

6.5	Sequence and Temporally Delayed Learning	186
6.6	Context Representations and Sequential Learning	187
6.6.1	Computational Considerations for Context Representations	188
6.6.2	Possible Biological Bases for Context Representations	189
6.6.3	Exploration: Learning the Reber Grammar	189
6.6.4	Summary	193
6.7	Reinforcement Learning for Temporally Delayed Outcomes	193
6.7.1	The Temporal Differences Algorithm	195
6.7.2	Phase-Based Temporal Differences	198
6.7.3	Exploration of TD: Classical Conditioning	199
6.8	Summary	202
6.9	Further Reading	202
II	Large-Scale Brain Area Organization and Cognitive Phenomena	203
7	Large-Scale Brain Area Functional Organization	205
7.1	Overview	205
7.2	General Computational and Functional Principles	206
7.2.1	Structural Principles	206
7.2.2	Dynamic Principles	210
7.3	General Functions of the Cortical Lobes and Subcortical Areas	211
7.3.1	Cortex	211
7.3.2	Limbic System	212
7.3.3	The Thalamus	212
7.3.4	The Basal Ganglia, Cerebellum, and Motor Control	213
7.4	Tripartite Functional Organization	214
7.4.1	Slow Integrative versus Fast Separating Learning	214
7.4.2	Active Memory versus Overlapping Distributed Representations	215
7.5	Toward a Cognitive Architecture of the Brain	216
7.5.1	Controlled versus Automatic Processing	217
7.5.2	Declarative/Procedural and Explicit/Implicit Distinctions	218
7.6	General Problems	219
7.6.1	The Binding Problem for Distributed Representations of Multiple Items	220
7.6.2	Representing Multiple Instances of the Same Thing	222
7.6.3	Comparing Representations	222
7.6.4	Representing Hierarchical Relationships	222
7.6.5	Recursion and Subroutine-like Processing	223
7.6.6	Generalization, Generativity, and Abstraction	224
7.6.7	Summary of General Problems	224
7.7	Summary	225

8 Perception and Attention	227
8.1 Overview	227
8.2 Biology of the Visual System	228
8.2.1 The Retina	228
8.2.2 The LGN of the Thalamus	230
8.2.3 Primary Visual Cortex: V1	230
8.2.4 Two Visual Processing Streams	232
8.2.5 The Ventral Visual Form Pathway: V2, V4, and IT	233
8.2.6 The Dorsal Where/Action Pathway	233
8.3 Primary Visual Representations	234
8.3.1 Basic Properties of the Model	235
8.3.2 Exploring the Model	237
8.3.3 Summary and Discussion	240
8.4 Object Recognition and the Visual Form Pathway	241
8.4.1 Basic Properties of the Model	243
8.4.2 Exploring the Model	246
8.4.3 Summary and Discussion	255
8.5 Spatial Attention: A Simple Model	257
8.5.1 Basic Properties of the Model	258
8.5.2 Exploring the Simple Attentional Model	261
8.5.3 Summary and Discussion	268
8.6 Spatial Attention: A More Complex Model	269
8.6.1 Exploring the Complex Attentional Model	269
8.6.2 Summary and Discussion	272
8.7 Summary	272
8.8 Further Reading	273
9 Memory	275
9.1 Overview	275
9.2 Weight-Based Memory in a Generic Model of Cortex	277
9.2.1 Long-Term Priming	278
9.2.2 AB-AC List Learning	282
9.3 The Hippocampal Memory System	287
9.3.1 Anatomy and Physiology of the Hippocampus	287
9.3.2 Basic Properties of the Hippocampal Model	289
9.3.3 Explorations of the Hippocampus	293
9.3.4 Summary and Discussion	296
9.4 Activation-Based Memory in a Generic Model of Cortex	298
9.4.1 Short-Term Priming	298
9.4.2 Active Maintenance	299
9.4.3 Robust yet Rapidly Updatable Active Maintenance	303
9.5 The Prefrontal Cortex Active Memory System	305

9.5.1	Dynamic Regulation of Active Maintenance	306
9.5.2	Details of the Prefrontal Cortex Model	307
9.5.3	Exploring the Model	310
9.5.4	Summary and Discussion.	312
9.6	The Development and Interaction of Memory Systems	314
9.6.1	Basic Properties of the Model	314
9.6.2	Exploring the Model	315
9.6.3	Summary and Discussion.	317
9.7	Memory Phenomena and System Interactions	318
9.7.1	Recognition Memory	318
9.7.2	Cued Recall.	319
9.7.3	Free Recall	319
9.7.4	Item Effects.	320
9.7.5	Working Memory.	320
9.8	Summary	320
9.9	Further Reading	321
10	Language	323
10.1	Overview	323
10.2	The Biology and Basic Representations of Language	325
10.2.1	Biology	325
10.2.2	Phonology	327
10.3	The Distributed Representation of Words and Dyslexia	329
10.3.1	Comparison with Traditional Dual-Route Models	330
10.3.2	The Interactive Model and Division of Labor	331
10.3.3	Dyslexia	331
10.3.4	Basic Properties of the Model	333
10.3.5	Exploring the Model	335
10.3.6	Summary and Discussion	341
10.4	The Orthography to Phonology Mapping	341
10.4.1	Basic Properties of the Model	343
10.4.2	Exploring the Model	344
10.4.3	Summary and Discussion	349
10.5	Overregularization in Past-Tense Inflectional Mappings	350
10.5.1	Basic Properties of the Model	352
10.5.2	Exploring the Model	353
10.5.3	Summary and Discussion	357
10.6	Semantic Representations from Word Co-occurrences and Hebbian Learning	358
10.6.1	Basic Properties of the Model	360
10.6.2	Exploring the Model	361
10.6.3	Summary and Discussion	365
10.7	Sentence-Level Processing	365

10.7.1 Basic Properties of the Model	367
10.7.2 Exploring the Model	370
10.7.3 Summary and Discussion	375
10.8 Summary	376
10.9 Further Reading	377
11 Higher-Level Cognition	379
11.1 Overview	379
11.2 Biology of the Frontal Cortex	384
11.3 Controlled Processing and the Stroop Task	385
11.3.1 Basic Properties of the Model	387
11.3.2 Exploring the Model	388
11.3.3 Summary and Discussion	391
11.4 Dynamic Categorization/Sorting Tasks	392
11.4.1 Basic Properties of the Model	395
11.4.2 Exploring the Model	397
11.4.3 Summary and Discussion	402
11.5 General Role of Frontal Cortex in Higher-Level Cognition	403
11.5.1 Functions Commonly Attributed to Frontal Cortex	403
11.5.2 Other Models and Theoretical Frameworks	407
11.6 Interacting Specialized Systems and Cognitive Control	408
11.7 Summary	409
11.8 Further Reading	410
12 Conclusions	411
12.1 Overview	411
12.2 Fundamentals	411
12.3 General Challenges for Computational Modeling	413
12.3.1 Models Are Too Simple	414
12.3.2 Models Are Too Complex	417
12.3.3 Models Can Do Anything	418
12.3.4 Models Are Reductionistic	418
12.3.5 Modeling Lacks Cumulative Research	419
12.4 Specific Challenges	419
12.4.1 Analytical Treatments of Learning	419
12.4.2 Error Signals	420
12.4.3 Regularities and Generalization	420
12.4.4 Capturing Higher-Level Cognition	421
12.5 Contributions of Computation to Cognitive Neuroscience	421
12.5.1 Models Help Us to Understand Phenomena	421
12.5.2 Models Deal with Complexity	422
12.5.3 Models Are Explicit	423

12.5.4 Models Allow Control	423
12.5.5 Models Provide a Unified Framework	423
12.6 Exploring on Your Own	424
III Simulator Details	425
A Introduction to the PDP++ Simulation Environment	427
A.1 Overview	427
A.2 Downloading and Installing the Software	427
A.3 Overall Structure of PDP++	428
A.4 Buttons and Menu Commands	429
A.5 Edit Dialogs	430
A.6 Control Panels	430
A.7 Specs	431
A.8 Networks and NetViews	431
A.8.1 NetView	431
A.9 Environments and EnviroViews	431
A.10 Processes	432
A.10.1 Process Control Panels	432
A.10.2 Statistics	432
A.11 Logs	433
A.11.1 TextLog	433
A.11.2 GraphLog	433
A.11.3 GridLog	433
A.12 Scripts	434
B Tutorial for Constructing Simulations in PDP++	435
B.1 Overview	435
B.2 Constructing a Basic Simulation	436
B.2.1 Creating a Project	436
B.2.2 Recording a Script	436
B.2.3 Creating a Network and Layers	437
B.2.4 Creating Projections and Connections	438
B.2.5 Specifying Layer Activity Levels	439
B.2.6 Creating an Environment	440
B.2.7 Creating Training Process	442
B.2.8 Creating a Training Log and Running the Model	443
B.3 Examining the Script	443
B.3.1 Object Paths	444
B.3.2 Object-Oriented Function Calls	444
B.3.3 Assigning Member Values	444

B.3.4 Scoped Variables	445
B.3.5 Running Processes from the Script	445
B.3.6 Saving Files from the Script	446
B.3.7 Compiling and Running the Script	446
B.4 Creating a Testing Process	447
B.4.1 Monitoring Unit Activities	447
B.4.2 Creating a New Statistic for Testing	448
B.4.3 Automatically Testing during Training	449
B.5 Writing Script Code to Create an Environment	450
B.5.1 Setting a Stopping Criterion for Training	451
B.6 Creating an Overall Control Panel	452
B.7 Creating SRN Context Layers	453
C Leabra Implementation Reference	455
C.1 Overview	455
C.2 Pseudocode	456
C.3 Connection-Level Variables	456
C.4 Unit-Level Variables	459
C.5 Layer-Level Variables	463
C.6 Process-Level Variables	465
References	467
Author Index	485
Subject Index	491