NEURAL NETWORK FUNDAMENTALS WITH GRAPHS, ALGORITHMS, AND APPLICATIONS

N. K. Bose
HRB-Systems Professor of Electrical Engineering
The Pennsylvania State University, University Park

P. Liang
Associate Professor of Electrical Engineering
University of California, Riverside

McGraw-Hill, Inc.
New York St. Louis San Francisco Auckland Bogotá
Caracas Lisbon London Madrid Mexico City Milan Montreal
New Delhi San Juan Singapore Sydney Tokyo Toronto

CONTENTS

List of Figures xiv
List of Tables xix
Preface xxi
List of Acronyms xxix
Glossary of Notations xxxii

I Fundamentals

1 Basics of Neuroscience and Artificial Neuron Models 3
1.1 The Brain as a Neural Network 5
1.2 Basic Properties of Neurons 9
1.2.1 Structure of a Neuron 9
1.2.2 Dendritic Tree 10
1.2.3 Action Potential and Its Propagation 11
1.2.4 Synapses 14
1.2.5 Connection Patterns between Neurons 18
1.2.6 An Example: A Motion Detection Neuron 19
1.3 Neuron Models 21
1.3.1 McCulloch-Pitts Model 21
1.3.2 Neuron Models with Continuous Transfer Characteristics 24
1.3.3 Other Neuron Models 25
1.4 Conclusions and Suggestions 28
Problems 31

2 Graphs 34
2.1 Terminology and Preliminaries 35
2.2 Special Types of Graphs 38
2.3 Directed Graphs 41

2.4 Matrix Representation of Graphs 44
2.4.1 Adjacency Matrix 44
2.4.2 Interconnection Matrix 44
2.5 Topological Invariants 45
2.5.1 Euler and Schlaefli Invariants 46
2.5.2 Genus 47
2.5.3 Thickness 49
2.5.4 Some Other Topological Invariants 51
2.6 Voronoi Diagrams and Delaunay Tessellation 53
2.7 Conclusions and Suggestions 55
Problems 57

3 Algorithms 61
3.1 Computational Complexity: P- and NP-Complete Problems 62
3.2 Shortest-Path and Max-Flow Min-Cut Problems 66
3.2.1 Dijkstra's Shortest-Path Algorithm 66
3.2.2 Max-Flow Min-Cut Algorithm 68
3.3 Interconnection and Routing Algorithms 77
3.3.1 Problem Formulation 77
3.3.2 Minimal Spanning Tree (MST) Algorithms 79
3.3.3 Minimal Fermat Tree (MFT) Problem 82
3.3.4 Traveling Salesperson (TS) Problem 82
3.3.5 Steiner Minimal Tree (SMT) 83
3.4 Placement and Partitioning 92
3.4.1 Placement 92
3.4.2 Partitioning 96
3.5 Parallel Computation 97
3.6 Associative Memory 99
3.6.1 The Linear Associator: Solution by Hebbian Rule 100
3.6.2 The Linear Associator: Solution by Generalized Inverse 101
3.6.3 Implementation of Associative Memory 102
3.7 Conclusions 106
Problems 108

II Feedforward Networks

4 Perceptrons and the LMS Algorithm 119
4.1 Rosenblatt's Perceptron 120
4.1.1 Definitions 122
4.1.2 Linear Separability of Training Patterns 124
4.1.3 Perceptron Learning Algorithms 129
4.1.4 Derivation of the Perceptron Algorithm as Gradient Descent 134
4.1.5 The Perceptron Convergence Theorem 136
4.2 The Widrow-Hoff LMS Algorithm 138
4.3 Order of a Predicate and a Perceptron 142
4.4 Conclusions and Suggestions 147
Problems 148

5 Multilayer Networks 155

5.1 Exact and Approximate Representation Using Feedforward Networks 156
5.1.1 Exact Representation: Kolmogorov's Theorem and Its Consequences 156
5.1.2 Approximate Representations 159
5.2 Fixed Multilayer Feedforward Network Training by Backpropagation 162
5.2.1 Implementation Considerations for Backpropagation 176
5.2.2 Variants of BPA 177
5.2.3 Temporal Signal Recognition and Prediction 179
5.3 Structural Training of Multilayer Feedforward Networks 181
5.3.1 Algorithm for Design Based on VoD 183
5.3.2 Robustness and Size Issues 189
5.4 Unsupervised and Reinforcement Learning 192
5.4.1 Principal Component Analysis Networks 193
5.4.2 Self-Organization in a Perceptual Network 197
5.4.3 Reinforcement Learning 201
5.5 The Probabilistic Neural Network 204
5.6 Conclusions and Suggestions 209
Problems 212

6 Complexity of Learning Using Feedforward Networks 219

6.1 Learnability in ANN 219
6.1.1 The Problem of Loading 221
6.1.2 Using an Appropriate Network to Get Around Intractability 228
6.2 Generalizability of Learning 231
6.2.1 VC Dimension and Generalization 232
6.2.2 Sufficient Conditions for Valid Generalization in Feedforward Networks 237
6.2.3 Necessary Conditions for Valid Generalization in Feedforward Networks 238
6.2.4 Discussions and Ways to Improve Generalization 240
6.3 Space Complexity of Feedforward Networks 245
6.3.1 Order of a Function and the Complexity of a Network 247
6.3.2 High Connectivity in Analog Neural Computations 248
6.4 Summary and Discussion 250
Problems 252

7 Adaptive-Structure Networks 254

7.1 Growth Algorithms 255
7.1.1 The Upstart Algorithm 257
7.1.2 Learning by Divide and Conquer 259
7.1.3 Other Growth Algorithms 265

7.2 Networks with Nonlinear Synapses and Nonlinear Synaptic Contacts 270
7.2.1 Quasi-Polynomial Synapses and Product Synaptic Contacts 273
7.2.2 Generalization of Learning and Hardware Considerations 276
7.3 Conclusions and Suggestions 278
Problems 281

III Recurrent Networks

8 Symmetric and Asymmetric Recurrent Networks 287
8.1 Symmetric Hopfield Networks and Associative Memory 289
8.1.1 Convergence Proofs 292
8.1.2 Computation in a Network and Minimum Cuts in a Graph 294
8.1.3 Capacity and Spurious Memory 300
8.1.4 Correlated Patterns 304
8.1.5 Hopfield Networks with Variations in the Connection Weights 306
8.1.6 Bidirectional Associative Memory 307
8.2 Symmetric Networks with Analog Units 310
8.2.1 Analog Hopfield Networks 310
8.2.2 Convergence Proof 314
8.2.3 Relation between Stable States of Discrete and Analog Hopfield Networks 315
8.2.4 Cellular Neural Networks 316
8.3 Seeking the Global Minimum: Simulated Annealing 318
8.3.1 Simulated Annealing in Optimization 319
8.3.2 Stochastic Networks: Applying Simulated Annealing to Hopfield Networks 323
8.4 A Learning Algorithm for the Boltzmann Machine 324
8.4.1 Learning the Underlying Structure of an Environment 324
8.4.2 The Learning Procedure 328
8.4.3 Mean Field Theory and the Deterministic Boltzmann Machine 330
8.5 Asymmetric Recurrent Networks 331
8.5.1 Phase Transition from Stationary to Chaotic 331
8.5.2 Spatial and Temporal Patterns 332
8.5.3 Learning in Asymmetric Networks: Recurrent Backpropagation 335
8.6 Summary and Discussion 340
Problems 341

9 Competitive Learning and Self-Organizing Networks 343
9.1 Unsupervised Competitive Learning 344
9.1.1 Two Phases of Competitive Learning 346

9.1.2 Using a Competitive Learning Network for Associative Memory 349
9.2 Adaptive Resonant Networks 352
9.2.1 The ART1 Clustering Algorithm 352
9.2.2 The ART1 Network 355
9.3 Self-Organizing Feature Maps 360
9.3.1 The Kohonen Map 361
9.3.2 Analysis of Kohonen Maps 366
9.3.3 Adaptive and Learning Vector Quantization 375
9.3.4 Two-Dimensional Topographic Maps 381
9.3.5 A Multilayer Self-Organizing Feature Map 385
9.4 Hybrid Learning 391
9.4.1 Counterpropagation Network 391
9.4.2 Regularizing Networks and Radial Basis Functions 395
9.5 Summary and Discussion 398
Problems 401

IV Applications of Neural Networks

10 Neural Network Approaches to Solving Hard Problems 407
10.1 The Traveling Salesperson Problem 409
10.2 Multitarget Tracking 412
10.3 Time Series Prediction 415
10.4 Talking Network and Phonetic Typewriter 417
10.4.1 Speech Generation 417
10.4.2 Speech Recognition 418
10.5 Autonomous Vehicle Navigation 422
10.6 Handwritten Digit Recognition 425
10.7 Image Compression by a Multilayer Feedforward Structure Trained through Backpropagation 430
10.8 Character Retrieval Using the Discrete Hopfield Network 435
10.9 Visual Processing Networks 437
10.10 Conclusion and Discussion 443

References 447

Appendix A Basis of Gradient-Based Optimization Methods 463
A.1 The Gradient Descent Method 464
A.2 Newton's Method 467
A.3 The Conjugate Gradient Method 468
A.4 Constrained Optimization 469
Bibliography 470

Index 471