BIBLIOGRAPHY

from the book

Introduction to the Theory of Neural Computation

by

John Hertz, Anders Krogh, and Richard Palmer

(Addison-Wesley, 1991)

The book itself may be ordered directly from Addison-Wesley in the U.S. at 1–800–447–2226 (toll-free) or by mail from *Addison-Wesley, The Advanced Book Program, 350 Bridge Parkway, Suite 209, Redwood City, CA 94065, U.S.A.* The price is $29.95 paperbound (ISBN 0–201–51560–1) or $44.25 hardbound (ISBN 0–201–50395–6). Please quote the ISBN in your order and add local sales tax plus $2.00 (in the U.S.) or $5.00 (overseas airmail) for shipping and handling charges.

Abu-Mostafa, Y.S. (1989). The Vapnik-Chervonenkis Dimension: Information versus Complexity in Learning. *Neural Computation* **1**, 312–317.

Abu-Mostafa, Y.S. and D. Psaltis (1987). Optical Neural Computers. *Scientific American* **256**, March 1987, 88–95.

Ackley, D.H., G.E. Hinton, and T.J. Sejnowski (1985). A Learning Algorithm for Boltzmann Machines. *Cognitive Science* **9**, 147–169. Reprinted in Anderson and Rosenfeld [1988].

Ackley, D.H. and M.S. Littman (1990). Generalization and Scaling in Reinforcement Learning. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 550–557. San Mateo: Morgan Kaufmann.

Ahalt, S.C., A.K. Krishnamurthy, P. Chen, and D.E. Melton (1990). Competitive Learning Algorithms for Vector Quantization. *Neural Networks* **3**, 277–290.

Almeida, L.B. (1987). A Learning Rule for Asynchronous Perceptrons with Feedback in a Combinatorial Environment. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. II, 609–618. New York: IEEE.

Almeida, L.B. (1988). Backpropagation in Perceptrons with Feedback. In *Neural Computers* (Neuss 1987), eds. R. Eckmiller and Ch. von der Malsburg, 199–208. Berlin: Springer-Verlag.

Alspector, J. and R.B. Allen (1987). A Neuromorphic VLSI Learning System. In *Advanced Research in VLSI: Proceedings of the 1987 Stanford Conference*, ed. P. Losleben, 313–349. Cambridge: MIT Press.

Alspector, J., R.B. Allen, V. Hu, and S. Satyanarayana (1988). Stochastic Learning Networks and Their Electronic Implementation. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 9–21. New York: American Institute of Physics.

Amari, S.A. (1977). Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields. *Biological Cybernetics* **27**, 77—87.

Amari, S.A. (1980). Topographic Organization of Nerve Fields. *Bulletin of Mathematical Biology* **42**, 339–364.

Amit, D. (1988). Neural Networks for Counting Chimes. *Proceedings of the National Academy of Sciences, USA* **85**, 2141–2145.

Amit, D. (1989). *Modelling Brain Function.* Cambridge: Cambridge University Press.

Amit, D., H. Gutfreund, and H. Sompolinsky (1985a). Spin-Glass Models of Neural Networks. *Physical Review A* **32**, 1007–1018.

Amit, D., H. Gutfreund, and H. Sompolinsky (1985b). Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks. *Physical Review Letters* **55**, 1530–1533.

Amit, D., H. Gutfreund, and H. Sompolinsky (1987a). Statistical Mechanics of Neural Networks Near Saturation. *Annals of Physics* **173**, 30–67.

Amit, D., H. Gutfreund, and H. Sompolinsky (1987b). Information Storage in Neural Networks with Low Levels of Activity. *Physical Review A* **35**, 2293–2303.

Anderson, C.H. and D.C. Van Essen (1987). Shifter Circuits: A Computational Strategy for Dynamic Aspects of Visual Processing. *Proceedings of the National Academy of Sciences, USA* **84**, 6297–6301.

Anderson, D.Z. (1986). Coherent Optical Eigenstate Memory. *Optics Letters* **11**, 56–58.

Anderson, J.A. (1968). A Memory Model Using Spatial Correlation Functions. *Kybernetik* **5**, 113–119.

Anderson, J.A. (1970). Two Models for Memory Organization. *Mathematical Biosciences* **8**, 137–160.

Anderson, J.A. and M.C. Mozer (1981). Categorization and Selective Neurons. In *Parallel Models of Associative Memory*, eds. G.E. Hinton and J.A. Anderson, 213–236. Hillsdale: Erlbaum.

Anderson, J.A. and E. Rosenfeld, eds. (1988). *Neurocomputing: Foundations of Research*. Cambridge: MIT Press.

Anderson, S., J.W.L. Merrill, and R. Port (1989). Dynamic Speech Categorization with Recurrent Networks. In *Proceedings of the 1988 Connectionist Models Summer School* (Pittsburg 1988), eds. D. Touretzky, G. Hinton, and T. Sejnowski, 398–406. San Mateo: Morgan Kaufmann.

Angéniol, B., G. de La Croix Vaubois, and J.-Y. Le Texier (1988). Self-Organizing Feature Maps and the Travelling Salesman Problem. *Neural Networks* **1**, 289–293.

Arbib, M.A. (1987). *Brains, Machines, and Mathematics*. Berlin: Springer-Verlag.

Bachmann, C.M., L.N. Cooper, A. Dembo, and O. Zeitouni (1987). A Relaxation Model for Memory with High Storage Density. *Proceedings of the National Academy of Sciences, USA* **84**, 7529–7531.

Bachrach, J. (1988). Learning to Represent State. Master's Thesis, University of Massachusetts, Amherst.

Baldi, P. and K. Hornik (1989). Neural Networks and Principal Component Analysis: Learning from Examples Without Local Minima. *Neural Networks* **2**, 53–58.

Barhen, J., S. Gulati, and M. Zak (1989). Neural Learning of Constrained Nonlinear Transformations. *Computer*, June 1989, 67–76.

Barlow, H.B. (1989). Unsupervised Learning. *Neural Computation* **1**, 295–311.

Barto, A.G. and P. Anandan (1985). Pattern Recognizing Stochastic Learning Automata. *IEEE Transactions on Systems, Man, and Cybernetics* **15**, 360–375.

Barto, A.G. and M.I. Jordan (1987). Gradient Following Without Back-Propagation in Layered Networks. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. II, 629–636. New York: IEEE.

Barto, A.G., R.S. Sutton, and C.W. Anderson (1983). Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problems. *IEEE Transactions on Systems, Man, and Cybernetics* **13**, 834–846. Reprinted in Anderson and Rosenfeld [1988].

Barto, A.G., R.S. Sutton, and C.J.C.H. Watkins (1991). Learning and Sequential Decision Making. In *Learning and Computational Neuroscience*, eds. M. Gabriel and J.W. Moore. Cambridge: MIT Press.

Baum, E.B. (1986). Towards Practical "Neural" Computation for Combinatorial Optimization Problems. In *Neural Networks for Computing* (Snowbird 1986), ed. J.S. Denker, 53–58. New York: American Institute of Physics.

Baum, E.B. and D. Haussler (1989). What Size Net Gives Valid Generalization? *Neural Computation* **1**, 151–160.

Baum, E.B. and F. Wilczek (1988). Supervised Learning of Probability Distributions by Neural Networks. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 52–61. New York: American Institute of Physics.

Becker, S. and Y. Le Cun (1989). Improving the Convergence of Back-Propagation Learning with Second Order Methods. In *Proceedings of the 1988 Connectionist Models Summer School* (Pittsburg 1988), eds. D. Touretzky, G. Hinton, and T. Sejnowski, 29–37. San Mateo: Morgan Kaufmann.

Bengio, Y., R. Cardin, and R. De Mori (1990). Speaker Independent Speech Recognition with Neural Networks and Speech Knowledge. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 218–225. San Mateo: Morgan Kaufmann.

Beurle, R.L. (1956). Properties of a Mass of Cells Capable of Regenerating Pulses. *Philosophical Transactions of the Royal Society of London B* **240**, 55–94.

Bienenstock, E.L., L.N. Cooper, and P.W. Munro (1982). Theory for the Development of Neuron Selectivity: Orientation Specificity and Binocular Interaction in Visual Cortex. *Journal of Neuroscience* **2**, 32–48. Reprinted in Anderson and Rosenfeld [1988].

Bilbro, G., R. Mann, T.K. Miller, W.E. Snyder, D.E. Van den Bout, and M. White (1989). Optimization by Mean Field Annealing. In *Advances in Neural Information Processing Systems I* (Denver 1988), ed. D.S. Touretzky, 91–98. San Mateo: Morgan Kaufmann.

Bilbro, G.L. and W. Snyder (1989). Range Image Restoration Using Mean Field Annealing. In *Advances in Neural Information Processing Systems I* (Denver 1988), ed. D.S. Touretzky, 594–601. San Mateo: Morgan Kaufmann.

Binder, K. and D.W. Heerman (1988). *Monte Carlo Simulation in Statistical Mechanics.* Berlin: Springer-Verlag.

Binder, K. and A.P. Young (1986). Spin Glasses: Experimental Facts, Theoretical Concepts, and Open Questions. *Reviews of Modern Physics* **58**, 801–976.

Blakemore, C. and G.F. Cooper (1970). Development of the Brain Depends on the Visual Environment. *Nature* **228**, 477–478.

Block, H.D. (1962). The Perceptron: A Model for Brain Functioning. *Reviews of Modern Physics* **34**, 123–135. Reprinted in Anderson and Rosenfeld [1988].

Blumer, A., A. Ehrenfeucht, D. Haussler, and M. Warmuth (1986). Classifying Learnable Geometric Concepts with the Vapnik-Chervonenkis Dimension. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing* (Berkeley 1986), 273–282. Salem: ACM.

Bohr, H., J. Bohr, S. Brunak, R.M.J. Cotterill, B. Lautrup, L. Nørskov, O.H. Olsen, and S.B. Petersen (1988). Protein Secondary Structure and Homology by Neural Networks: The α-Helices in Rhodopsin. *FEBS Letters* **241**, 223–228.

Bourland, H. and Y. Kamp (1988). Auto-Association by Multilayer Perceptrons and Singular Value Decomposition. *Biological Cybernetics* **59**, 291–294.

Brandt, R.D., Y. Wang, A.J. Laub, and S.K. Mitra (1988). Alternative Networks for Solving the Travelling Salesman Problem and the List-Matching Problem. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. II, 333–340. New York: IEEE.

Brunak, S. and B. Lautrup (1989). Liniedeling med et Neuralt Netværk. *Skrifter for Anvendt Matematik og Lingvistik* **14**, 55–74.

Brunak, S. and B. Lautrup (1990). *Neural Networks: Computers with Intuition.* Singapore: World Scientific.

Bryson, A.E. and Y.-C. Ho (1969). *Applied Optimal Control.* New York: Blaisdell.

Buhmann, J. and K. Schulten (1987). Noise-Driven Temporal Association in Neural Networks. *Europhysics Letters* **4**, 1205–1209.

Buhmann, J. and K. Schulten (1988). Storing Sequences of Biased Patterns in Neural Networks with Stochastic Dynamics. In *Neural Computers* (Neuss 1987), eds. R. Eckmiller and Ch. von der Malsburg, 231–242. Berlin: Springer-Verlag.

Burr, D.J. (1988). An Improved Elastic Net Method for the Travelling Salesman Problem. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 69–76. New York: IEEE.

Caianiello, E.R. (1961). Outline of a Theory of Thought and Thinking Machines. *Journal of Theoretical Biology* **1**, 204–235.

Canning, A. and E. Gardner (1988). Partially Connected Models of Neural Networks. *Journal of Physics A* **21**, 3275–3284.

Carpenter, G.A. and S. Grossberg (1987a). A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine. *Computer Vision, Graphics, and Image Processing* **37**, 54–115.

Carpenter, G.A. and S. Grossberg (1987b). ART2: Self-Organization of Stable Category Recognition Codes for Analog Input Patterns. *Applied Optics* **26**, 4919–4930.

Carpenter, G.A. and S. Grossberg (1988). The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network. *Computer*, March 1988, 77–88.

Casdagli, M. (1989). Nonlinear Prediction of Chaotic Time Series. *Physica* **35D**, 335–356.

Cater, J.P. (1987). Successfully Using Peak Learning Rates of 10 (and Greater) in Back-Propagation Networks with the Heuristic Learning Algorithm. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. II, 645–651. New York: IEEE.

Caudill, M. (1989). *Neural Networks Primer.* San Francisco: Miller Freeman.

Chauvin, Y. (1989). A Back-Propagation Algorithm with Optimal Use of Hidden Units. In *Advances in Neural Information Processing Systems I* (Denver 1988), ed. D.S. Touretzky, 519–526. San Mateo: Morgan Kaufmann.

Cleeremans, A., D. Servan-Schreiber, and J.L. McClelland (1989). Finite State Automata and Simple Recurrent Networks. *Neural Computation* **1**, 372–381.

Cohen, M.A. and S. Grossberg (1983). Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks. *IEEE Transactions on Systems, Man, and Cybernetics* **13**, 815–826.

Cohen, M.S. (1986). Design of a New Medium for Volume Holographic Information Processing. *Applied Optics* **25**, 2228–2294.

Coolen, A.C.C. and C.C.A.M. Gielen (1988). Delays in Neural Networks. *Europhysics Letters* **7**, 281–285.

Cooper, L.N. (1973). A Possible Organization of Animal Memory and Learning. In *Collective Properties of Physical Systems* (24th Nobel Symposium), eds. B. Lundqvist and S. Lundqvist, 252–264. New York: Academic Press. Reprinted in Anderson and Rosenfeld [1988].

Cortes, C. and J.A. Hertz (1989). A Network System for Image Segmentation. In *International Joint Conference on Neural Networks* (Washington 1989), vol. I, 121–127. New York: IEEE.

Cortes, C., A. Krogh, and J.A. Hertz (1987). Hierarchical Associative Networks. *Journal of Physics A* **20**, 4449–4455.

Cottrell, G.W., P. Munro, and D. Zipser (1987). Learning Internal Representations from Gray-Scale Images: An Example of Extensional Programming. In *Ninth Annual Conference of the Cognitive Science Society* (Seattle 1987), 462–473. Hillsdale: Erlbaum.

Cottrell, M. and J.C. Fort (1986). A Stochastic Model of Retinotopy: A Self Organizing Process. *Biological Cybernetics* **53**, 405–411.

Cover, T.M. (1965). Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. *IEEE Transactions on Electronic Computers* **14**, 326–334.

Cowan, J.D. and D.H. Sharp (1988a). Neural Nets and Artificial Intelligence. *Dædalus, Proceedings of the American Academy of Arts and Sciences* **117**, 85–121.

Cowan, J.D. and D.H. Sharp (1988b). Neural Nets. *Quarterly Reviews of Biophysics* **21**, 365–427.

Cragg, B.G. and H.N.V. Temperley (1954). The Organization of Neurones: A Cooperative Analogy. *EEG and Clinical Neurophysiology* **6**, 85–92.

Cragg, B.G. and H.N.V. Temperley (1955). Memory: The Analogy with Ferromagnetic Hysteresis. *Brain* **78**, II,304–316.

Crick, F. (1989). The Recent Excitement About Neural Networks. *Nature* **337**, 129–132.

Crisanti, A. and H. Sompolinsky (1987). Dynamics of Spin Systems with Randomly Asymmetric Bonds: Langevin Dynamics and a Spherical Model. *Physical Review A* **36**, 4922–4939.

Crisanti, A., D.J. Amit, and H. Gutfreund (1986). Saturation Level of the Hopfield Model for Neural Network. *Europhysics Letters* **2**, 337–341.

Cybenko, G. (1988). Continuous Valued Neural Networks with Two Hidden Layers Are Sufficient. Technical Report, Department of Computer Science, Tufts University, Medford, MA.

Cybenko, G. (1989). Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals, and Systems* **2**, 303–314.

Dehaene, S., J.-P. Changeux, and J.-P. Nadal (1987). Neural Networks That Learn Temporal Sequences by Selection. *Proceedings of the National Academy of Sciences, USA* **84**, 2727–2731.

Denker, J. (1986). Neural Network Refinements and Extensions. In *Neural Networks for Computing* (Snowbird 1986), ed. J.S. Denker, 121–128. New York: American Institute of Physics.

Denker, J., D. Schwartz, B. Wittner, S. Solla, R. Howard, L. Jackel, and J. Hopfield (1987). Large Automatic Learning, Rule Extraction, and Generalization. *Complex Systems* **1**, 877–922.

Derrida, B., E. Gardner, and A. Zippelius (1987). An Exactly Soluble Asymmetric Neural Network Model. *Europhysics Letters* **4**, 167–173.

Derthick, M. (1984). Variations on the Boltzmann Machine. Technical Report CMU–CS–84–120, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

DeSieno, D. (1988). Adding a Conscience to Competitive Learning. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 117–124. New York: IEEE.

Devijver, P.A. and J. Kittler (1982). *Pattern Recognition: A Statistical Approach*. London: Prentice-Hall.

Diederich, S. and M. Opper (1987). Learning of Correlated Patterns in Spin-Glass Networks by Local Learning Rules. *Physical Review Letters* **58**, 949–952.

Duda, R.O. and P.E. Hart (1973). *Pattern Classification and Scene Analysis*. New York: Wiley.

Durbin, R. and D. Willshaw (1987). An Analogue Approach to the Travelling Salesman Problem Using an Elastic Net Method. *Nature* **326**, 689–691.

Elman, J.L. (1990). Finding Structure in Time. *Cognitive Science* **14**, 179–211.

Elman, J.L. and D. Zipser (1988). Learning the Hidden Structure of Speech. *Journal of the Acoustical Society of America* **83**, 1615–1626.

Fahlman, S.E. (1989). Fast-Learning Variations on Back-Propagation: An Empirical Study. In *Proceedings of the 1988 Connectionist Models Summer School* (Pittsburg 1988), eds. D. Touretzky, G. Hinton, and T. Sejnowski, 38–51. San Mateo: Morgan Kaufmann.

Fahlman, S.E. and C. Lebiere (1990). The Cascade-Correlation Learning Architecture. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 524–532. San Mateo: Morgan Kaufmann.

Farhat, N.H. (1987). Optoelectronic Analogs of Self-Programming Neural Nets: Architectures and Methods for Implementing Fast Stochastic Learning by Simulated Annealing. *Applied Optics* **26**, 5093–5103.

Farhat, N.H., D. Psaltis, A. Prata, and E. Paek (1985). Optical Implementation of the Hopfield Model. *Applied Optics* **24**, 1469–1475. Reprinted in Anderson and Rosenfeld [1988].

Farmer, D. and J. Sidorowich (1987). Predicting Chaotic Time Series. *Physical Review Letters* **59**, 845–848.

Farmer, D. and J. Sidorowich (1988). Exploiting Chaos to Predict the Future and Reduce Noise. In *Evolution, Learning, and Cognition*, ed. W.C. Lee, 277–330. Singapore: World Scientific.

Feldman, J.A. and D.H. Ballard (1982). Connectionist Models and Their Properties. *Cognitive Science* **6**, 205–254. Reprinted in Anderson and Rosenfeld [1988].

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, vol. I. New York: Wiley.

Fischer, K.H. and J.A. Hertz (1990). *Spin Glasses*. Cambridge: Cambridge University Press.

Földiák, P. (1989). Adaptive Network for Optimal Linear Feature Extraction. In *International Joint Conference on Neural Networks* (Washington 1989), vol. I, 401–405. New York: IEEE.

Franzini, M.A. (1987). Speech Recognition with Back Propagation. In *Proceedings of the Ninth Annual Conference of the IEEE Engineering in Medicine and Biology Society* (Boston 1987), 1702–1703. New York: IEEE.

Frean, M (1990). The Upstart Algorithm: A Method for Constructing and Training Feedforward Neural Networks. *Neural Computation* **2**, 198–209.

Fu, Y. and P.W. Anderson (1986). Application of Statistical Mechanics to NP-Complete Problems in Combinatorial Optimization. *Journal of Physics A* **19**, 1605–1620.

Fukushima, K. (1975). Cognitron: A Self-Organizing Multilayered Neural Network. *Biological Cybernetics* **20**, 121–136.

Fukushima, K. (1980). Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. *Biological Cybernetics* **36**, 193–202.

Fukushima, K., S. Miyake, and T. Ito (1983). Neocognitron: A Neural Network Model for a Mechanism of Visual Pattern Recognition. *IEEE Transactions on Systems, Man, and Cybernetics* **13**, 826–834. Reprinted in Anderson and Rosenfeld [1988].

Gallant, S.I. (1986). Optimal Linear Discriminants. In *Eighth International Conference on Pattern Recognition* (Paris 1986), 849–852. New York: IEEE.

Gardner, E. (1987). Maximum Storage Capacity in Neural Networks. *Europhysics Letters* **4**, 481–485.

Gardner, E. (1988). The Space of Interactions in Neural Network Models. *Journal of Physics A* **21**, 257–270.

Gardner, E. and B. Derrida (1988). Optimal Storage Properties of Neural Network Models. *Journal of Physics A* **21**, 271–284.

Gardner, E. and B. Derrida (1989). Three Unfinished Works on the Optimal Storage Capacity of Networks. *Journal of Physics A* **22**, 1983–1994.

Gardner, E., H. Gutfreund, and I. Yekutieli (1989). The Phase Space of Interactions in Neural Networks with Definite Symmetry. *Journal of Physics A* **22**, 1995–2008.

Garey, M.R. and D.S. Johnson (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York: Freeman.

Geman, S and D. Geman (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741. Reprinted in Anderson and Rosenfeld [1988].

Geszti, T. (1990). *Physical Models of Neural Networks*. Singapore: World Scientific.

Geszti, T. and F. Pázmándi (1987). Learning Within Bounds and Dream Sleep. *Journal of Physics A* **20**, L1299–L1303.

Glauber, R.J. (1963). Time-Dependent Statistics of the Ising Model. *Journal of Mathematical Physics* **4**, 294–307.

Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading: Addison-Wesley.

Golomb, D., N. Rubin, and H. Sompolinsky (1990). Willshaw Model: Associative Memory with Sparse Coding and Low Firing Rates. *Physical Review A* **41**, 1843–1854.

Gorman, R.P. and T.J. Sejnowski (1988a). Analysis of Hidden Units in a Layered Network Trained to Classify Sonar Targets. *Neural Networks* **1**, 75–89.

Gorman, R.P. and T.J. Sejnowski (1988b). Learned Classification of Sonar Targets Using a Massively-Parallel Network. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **36**, 1135–1140.

Graf, H.P., L.D. Jackel, R.E. Howard, B. Straughn, J.S. Denker, W. Hubbard, D.M. Tennant, and D. Schwartz (1986). VLSI Implementation of a Neural Network Memory with Several Hundreds of Neurons. In *Neural Networks for Computing* (Snowbird 1986), ed. J.S. Denker, 182–187. New York: American Institute of Physics.

Graf, D.H. and W.R. LaLonde (1988). A Neural Controller for Collision-Free Movement of General Robot Manipulators. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 77–84. New York: IEEE.

Gray, R.M. (1984). Vector Quantization. *IEEE ASSP Magazine*, April 1984, 4–29.

Grossberg, S. (1967). Nonlinear Difference-Differential Equations in Prediction and Learning Theory. *Proceedings of the National Academy of Sciences, USA* **58**, 1329–1334.

Grossberg, S. (1968a). Some Nonlinear Networks Capable of Learning a Spatial Pattern of Arbitrary Complexity. *Proceedings of the National Academy of Sciences, USA* **59**, 368–372.

Grossberg, S. (1968b). Some Physiological and Biochemical Consequences of Psychological Postulates. *Proceedings of the National Academy of Sciences, USA* **60**, 758–765.

Grossberg, S. (1969). Embedding Fields: A Theory of Learning with Physiological Implications. *Journal of Mathematical Psychology* **6**, 209–239.

Grossberg, S. (1972). Neural Expectation: Cerebellar and Retinal Analogs of Cells Fired by Learnable or Unlearned Pattern Classes. *Kybernetik* **10**, 49–57.

Grossberg, S. (1976a). Adaptive Pattern Classification and Universal Recoding: I. Parallel Development and Coding of Neural Feature Detectors. *Biological Cybernetics* **23**, 121–134. Reprinted in Anderson and Rosenfeld [1988].

Grossberg, S. (1976b). Adaptive Pattern Classification and Universal Recoding: II. Feedback, Expectation, Olfaction, Illusions. *Biological Cybernetics* **23**, 187–202.

Grossberg, S. (1980). How Does the Brain Build a Cognitive Code? *Psychological Review* **87**, 1–51. Reprinted in Anderson and Rosenfeld [1988].

Grossberg, S. (1987a). *The Adaptive Brain*, 2 vols. Amsterdam: Elsevier.

Grossberg, S. (1987b). Competitive Learning: From Interactive Activation to Adaptive Resonance. *Cognitive Science* **11**, 23–63.

Grossman, T. (1990). The CHIR Algorithm for Feed Forward Networks with Binary Weights. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 516–523. San Mateo: Morgan Kaufmann.

Grossman, T., R. Meir, and E. Domany (1989). Learning by Choice of Internal Representations. *Complex Systems* **2**, 555–575.

Gutfreund, H. (1988). Neural Networks with Hierarchically Correlated Patterns. *Physical Review A* **37**, 570–577.

Gutfreund, H. and M. Mézard (1988). Processing of Temporal Sequences in Neural Networks. *Physical Review Letters* **61**, 235–238.

Gutzmann, K. (1987). Combinatorial Optimization Using a Continuous State Boltzmann Machine. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. III, 721–734. New York: IEEE.

Györgyi, G. (1990). Inference of a Rule by a Neural Network with Thermal Noise. *Physical Review Letters* **64**, 2957–2960.

Györgyi, G. and N. Tishby (1990). Statistical Theory of Learning a Rule. In *Neural Networks and Spin Glasses*, eds. W.K. Theumann and R. Koeberle. Singapore: World Scientific.

Hanson, S.J. and L. Pratt (1989). A Comparison of Different Biases for Minimal Network Construction with Back-Propagation. In *Advances in Neural Information Processing Systems I* (Denver 1988), ed. D.S. Touretzky, 177–185. San Mateo: Morgan Kaufmann.

Harp, S.A., T. Samad, and A. Guha (1990). Designing Application-Specific Neural Networks Using the Genetic Algorithm. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 447–454. San Mateo: Morgan Kaufmann.

Hartman, E.J., J.D. Keeler, and J.M. Kowalski (1990). Layered Neural Networks with Gaussian Hidden Units As Universal Approximations. *Neural Computation* **2**, 210–215.

Hebb, D.O. (1949). *The Organization of Behavior.* New York: Wiley. Partially reprinted in Anderson and Rosenfeld [1988].

Hecht-Nielsen, R. (1987a). Combinatorial Hypercompression. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. II, 455–461. New York: IEEE.

Hecht-Nielsen, R. (1987b). Counterpropagation Networks. *Applied Optics* **26**, 4979–4984.

Hecht-Nielsen, R. (1988). Applications of Counterpropagation Networks. *Neural Networks* **1**, 131–139.

Hecht-Nielsen, R. (1989). Theory of the Backpropagation Neural Network. In *International Joint Conference on Neural Networks* (Washington 1989), vol. I, 593–605. New York: IEEE.

Hegde, S.U., J.L. Sweet, and W.B. Levy (1988). Determination of Parameters in a Hopfield/Tank Computational Network. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. II, 291–298. New York: IEEE.

van Hemmen, J.L., L.B. Ioffe, R. Kühn, and M. Vaas (1990). Increasing the Efficiency of a Neural Network through Unlearning. *Physica* **163A**, 386–392.

van Hemmen, J.L. and R. Kühn (1986). Nonlinear Neural Networks. *Physical Review Letters* **57**, 913–916.

van Hemmen, J.L. and R.G. Palmer (1979). The Replica Method and a Solvable Spin Glass Model. *Journal of Physics A* **12**, 563–580.

Hertz, J.A. (1989). A Gauge Theory in Computational Vision: A Model for Outline Extraction. *Physica Scripta* **39**, 161–167.

Hertz, J.A. (1990). Statistical Dynamics of Learning. Preprint 90/34 S, Nordita, Copenhagen, Denmark.

Hertz, J.A., G. Grinstein, and S. Solla (1986). Memory Networks with Asymmetric Bonds. In *Neural Networks for Computing* (Snowbird 1986), ed. J.S. Denker, 212–218. New York: American Institute of Physics.

Hertz, J.A., G. Grinstein, and S. Solla (1987). Irreversible Spin Glasses and Neural Networks. In *Heidelberg Colloquium on Glassy Dynamics* (Heidelberg 1986), eds. J.L. van Hemmen and I. Morgenstern, 538–546. Berlin: Springer-Verlag.

Hertz, J.A., A. Krogh, and G.I. Thorbergsson (1989). Phase Transitions in Simple Learning. *Journal of Physics A* **22**, 2133–2150.

Herz, A., B. Sulzer, R. Kühn, and J.L. van Hemmen (1989). Hebbian Learning Reconsidered: Representation of Static and Dynamic Objects in Associative Neural Nets. *Biological Cybernetics* **60**, 457–467.

Hinton, G.E. (1986). Learning Distributed Representations of Concepts. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (Amherst 1986), 1–12. Hillsdale: Erlbaum.

Hinton, G.E. (1989). Deterministic Boltzmann Learning Performs Steepest Descent in Weight Space. *Neural Computation* **1**, 143–150.

Hinton, G.E. and T.J. Sejnowski (1983). Optimal Perceptual Inference. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Washington 1983), 448–453. New York: IEEE.

Hinton, G.E. and T.J. Sejnowski (1986). Learning and Relearning in Boltzmann Machines. In *Parallel Distributed Processing*, vol. 1, chap. 7. *See* Rumelhart, McClelland, et al. [1986].

Hopfield, J.J. (1982). Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences, USA* **79**, 2554–2558. Reprinted in Anderson and Rosenfeld [1988].

Hopfield, J.J. (1984). Neurons with Graded Responses Have Collective Computational Properties Like Those of Two-State Neurons. *Proceedings of the National Academy of Sciences, USA* **81**, 3088–3092. Reprinted in Anderson and Rosenfeld [1988].

Hopfield, J.J. (1987). Learning Algorithms and Probability Distributions in Feed-Forward and Feed-Back Networks. *Proceedings of the National Academy of Sciences, USA* **84**, 8429–8433.

Hopfield, J.J., D.I. Feinstein, and R.G. Palmer (1983). "Unlearning" Has a Stabilizing Effect in Collective Memories. *Nature* **304**, 158–159.

Hopfield, J.J. and D.W. Tank (1985). "Neural" Computation of Decisions in Optimization Problems. *Biological Cybernetics* **52**, 141–152.

Hopfield, J.J. and D.W. Tank (1986). Computing with Neural Circuits: A Model. *Science* **233**, 625–633.

Hopfield, J.J. and D.W. Tank (1989). Neural Architecture and Biophysics for Sequence Recognition. In *Neural Models of Plasticity*, eds. J.H. Byrne and W.O. Berry, 363–377. San Diego: Academic Press.

Hornik, K., M. Stinchcombe, and H. White (1989). Multilayer Feedforward Networks Are Universal Approximators. *Neural Networks* **2**, 359–366.

Hsu, K, D. Brady, and D. Psaltis (1988). Experimental Demonstration of Optical Neural Computers. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 377–386. New York: American Institute of Physics.

Huang, K. (1987). *Statistical Mechanics*. New York: Wiley.

Huang, W.Y. and R.P. Lippmann (1988). Neural Net and Traditional Classifiers. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 387–396. New York: American Institute of Physics.

Hubel, D.H. and T.N. Wiesel (1962). Receptive Fields, Binocular Interaction, and Functional Architecture in the Cat's Visual Cortex. *Journal of Physiology (London)* **160**, 106–154.

Hueter, G.J. (1988). Solution of the Travelling Salesman Problem with an Adaptive Ring. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 85–92. New York: IEEE.

Hush, D.R. and J.M. Salas (1988). Improving the Learning Rate of Back-Propagation with the Gradient Reuse Algorithm. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 441–447. New York: IEEE.

Jacobs, R.A. (1988). Increased Rates of Convergence Through Learning Rate Adaptation. *Neural Networks* **1**, 295–307.

Johnson, D.S., C.R. Aragon, L.A. McGeoch, and C. Schevon (1989). Optimization by Simulated Annealing: An Experimental Evaluation; Part I, Graph Partitioning. *Operations Research* **37**, 865–891. Parts II and III are expected to appear in 1990.

Jolliffe, I.T. (1986). *Principal Component Analysis*. New York: Springer-Verlag.

Jordan, M.I. (1986). Attractor Dynamics and Parallelism in a Connectionist Sequential Machine. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (Amherst 1986), 531–546. Hillsdale: Erlbaum.

Jordan, M.I. (1989). Serial Order: A Parallel, Distributed Processing Approach. In *Advances in Connectionist Theory: Speech*, eds. J.L. Elman and D.E. Rumelhart. Hillsdale: Erlbaum.

Kahng, A.B. (1989). Travelling Salesman Heuristics and Embedding Dimension in the Hopfield Model. In *International Joint Conference on Neural Networks* (Washington 1989), vol. I, 513–520. New York: IEEE.

Kammen, D.M. and A.L. Yuille (1988). Spontaneous Symmetry-Breaking Energy Functions and the Emergence of Orientation Selective Cortical Cells. *Biological Cybernetics* **59**, 23–31.

Kandel, E.R. and J.H. Schwartz (1985). *Principles of Neural Science* (2nd ed.). New York: Elsevier.

Kanter, I. and H. Sompolinsky (1987). Associative Recall of Memory Without Errors. *Physical Review A* **35**, 380–392.

Kerszberg, M. and A. Zippelius (1990). Synchronization in Neural Assemblies. *Physica Scripta* **T33**, 54–64.

Kinzel, W. and M. Opper (1990). Dynamics of Learning. In *Physics of Neural Networks*, eds. E. Domany, J.L. van Hemmen, and K. Schulten, vol. I. Berlin: Springer-Verlag.

Kirkpatrick, S., C.D. Gelatt Jr., and M.P. Vecchi (1983). Optimization by Simulated Annealing. *Science* **220**, 671–680. Reprinted in Anderson and Rosenfeld [1988].

Kirkpatrick, S. and G. Toulouse (1985). Configuration Space Analysis of Travelling Salesman Problems. *Journal de Physique (Paris)* **46**, 1277–1292.

Kleinfeld, D. (1986). Sequential State Generation by Model Neural Networks. *Proceedings of the National Academy of Sciences, USA* **83**, 9469–9473.

Kleinfeld, D. and H. Sompolinsky (1989). Associative Network Models for Central Pattern Generators. In *Methods in Neuronal Modeling: From Synapses to Networks*, eds. C. Koch and I. Segev, 195–246. Cambridge: MIT Press.

Klopf, A.H. (1982). *The Hedonistic Neuron: A Theory of Memory, Learning, and Intelligence*. Washington: Hemisphere.

Koch, C., J. Luo, C. Mead, and J. Hutchinson (1988). Computing Motion Using Resistive Networks. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 422–431. New York: American Institute of Physics.

Koch, C., J. Marroquin, and A. Yuille (1986). Analog "Neuronal" Networks in Early Vision. *Proceedings of the National Academy of Sciences, USA* **83**, 4263–4267.

Kohonen, T. (1974). An Adaptive Associative Memory Principle. *IEEE Transactions on Computers* **C-23**, 444–445.

Kohonen, T. (1982). Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics* **43**, 59–69. Reprinted in Anderson and Rosenfeld [1988].

Kohonen, T. (1989). *Self-Organization and Associative Memory* (3rd ed.). Berlin: Springer-Verlag.

Kohonen, T., G. Barna, and R. Chrisley (1988). Statistical Pattern Recognition with Neural Networks: Benchmarking Studies. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 61–68. New York: IEEE.

Kohonen, T., K. Mäkisara, and T. Saramäki (1984). Phonotopic Maps—Insightful Representation of Phonological Features for Speech Recognition. In *Proceedings of the Seventh International Conference on Pattern Recognition* (Montreal 1984), 182–185. New York: IEEE.

Kramer, A.H. and A. Sangiovanni-Vincentelli (1989). Efficient Parallel Learning Algorithms for Neural Networks. In *Advances in Neural Information Processing Systems I* (Denver 1988), ed. D.S. Touretzky, 40–48. San Mateo: Morgan Kaufmann.

Krauth, W. and M. Mézard (1989). The Cavity Method and the Travelling-Salesman Problem. *Europhysics Letters* **8**, 213–218.

Krishnaiah, P.R. and L.N. Kanal, eds. (1982). *Classification, Pattern Recognition, and Reduction of Dimensionality*. Handbook of Statistics, vol. 2. Amsterdam: North Holland.

Krogh, A. and J.A. Hertz (1988). Mean Field Analysis of Hierarchical Associative Networks with Magnetization. *Journal of Physics A* **21**, 2211–2224.

Krogh, A. and J.A. Hertz (1990). Hebbian Learning of Principal Components. In *Parallel Processing in Neural Systems and Computers* (Düsseldorf 1990), eds. R. Eckmiller, G. Hartmann, and G. Hauske, 183–186. Amsterdam: Elsevier.

Krogh, A., G.I. Thorbergsson, and J.A. Hertz (1990). A Cost Function for Internal Representations. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 733–740. San Mateo: Morgan Kaufmann.

Kullback, S. (1959). *Information Theory and Statistics*. New York: Wiley.

Kushner, H.J. and D.S. Clark (1978). *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. New York: Springer-Verlag.

Lapedes, A. and R. Farber (1986a). A Self-Optimizing, Nonsymmetrical Neural Net for Content Addressable Memory and Pattern Recognition. *Physica* **22D**, 247–259.

Lapedes, A. and R. Farber (1986b). Programming a Massively Parallel, Computation Universal System: Static Behavior. In *Neural Networks for Computing* (Snowbird 1986), ed. J.S. Denker, 283–298. New York: American Institute of Physics.

Lapedes, A. and R. Farber (1987). Nonlinear Signal Processing Using Neural Networks: Prediction and System Modelling. Technical Report LA–UR–87–2662, Los Alamos National Laboratory, Los Alamos, NM.

Lapedes, A. and R. Farber (1988). How Neural Nets Work. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 442–456. New York: American Institute of Physics.

Lawler, E.L. (1976). *Combinatorial Optimization: Networks and Matroids.* New York: Holt-Rinehart-Winston.

Lawler, E.L., J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, eds. (1985). *The Travelling Salesman Problem.* Chichester: Wiley.

Le Cun, Y. (1985). Une Procédure d'Apprentissage pour Réseau à Seuil Assymétrique. In *Cognitiva 85: A la Frontière de l'Intelligence Artificielle des Sciences de la Connaissance des Neurosciences* (Paris 1985), 599–604. Paris: CESTA.

Le Cun, Y., B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation* **1**, 541–551.

Le Cun, Y., B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel (1990). Handwritten Digit Recognition with a Back-Propagation Network. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 396–404. San Mateo: Morgan Kaufmann.

Le Cun, Y., J.S. Denker, and S.A. Solla (1990). Optimal Brain Damage. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 598–605. San Mateo: Morgan Kaufmann.

von Lehman, A., E.G. Paek, P.F. Liao, A. Marrakchi, and J.S. Patel (1988). Factors Influencing Learning by Back-Propagation. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 335–341. New York: IEEE.

Liang, F.M. (1983). Ph.D. Thesis. Stanford University. See Appendix H of *The TEXBook*, D.E. Knuth (Reading: Addison Wesley 1986).

Lin, S. and B.W. Kernighan (1973). An Effective Heuristic Algorithm for the Travelling Salesman Problem. *Operations Research* **21**, 498–516.

Linsker, R. (1986). From Basic Network Principles to Neural Architecture. *Proceedings of the National Academy of Sciences, USA* **83**, 7508–7512, 8390–8394, 8779–8783.

Linsker, R. (1988). Self-Organization in a Perceptual Network. *Computer*, March 1988, 105–117.

Lippmann, R.P. (1987). An Introduction to Computing with Neural Nets. *IEEE ASSP Magazine*, April 1987, 4–22.

Lippmann, R.P. (1989). Review of Neural Networks for Speech Recognition. *Neural Computation* **1**, 1–38.

Lister, R. (1990). Segment Reversal and the TSP. In *International Joint Conference on Neural Networks* (Washington 1990), vol. 1, 424–427. Hillsdale: Erlbaum.

Little, W.A. (1974). The Existence of Persistent States in the Brain. *Mathematical Biosciences* **19**, 101–120.

Little, W.A. and G.L. Shaw (1975). A Statistical Theory of Short and Long Term Memory. *Behavioral Biology* **14**, 115–133. Reprinted in Anderson and Rosenfeld [1988].

Little, W.A. and G.L. Shaw (1978). Analytic Study of the Memory Storage Capacity of a Neural Network. *Mathematical Biosciences* **39**, 281–290.

Luenberger, D.G. (1986). *Linear and Nonlinear Programming.* Reading: Addison-Wesley.

Ma, S.-K. (1985). *Statistical Mechanics.* Philadelphia: World Scientific.

MacKay, D.J.C. and K.D. Miller (1990). Analysis of Linsker's Simulation of Hebbian Rules. *Neural Computation* **2**, 173–187.

Mackey, M.C. and L. Glass (1977). Oscillation and Chaos in Physiological Control Systems. *Science* **197**, 287.

Makram-Ebeid, S., J.-A. Sirat, and J.-R. Viala (1989). A Rationalized Back-Propagation Learning Algorithm. In *International Joint Conference on Neural Networks* (Washington 1989), vol. II, 373–380. New York: IEEE.

von der Malsburg, Ch. (1973). Self-Organization of Orientation Sensitive Cells in the Striate Cortex. *Kybernetik* **14**, 85–100. Reprinted in Anderson and Rosenfeld [1988].

von der Malsburg, Ch. and J.D. Cowan (1982). Outline of a Theory for the Ontogenesis of Iso-Orientation Domains in Visual Cortex. *Biological Cybernetics* **45**, 49–56.

Mandelbrot, B.B. (1982). *The Fractal Geometry of Nature*. San Francisco: Freeman.

Marchand, M., M. Golea, and P. Ruján (1990). A Convergence Theorem for Sequential Learning in Two-Layer Perceptrons. *Europhysics Letters* **11**, 487–492.

Marr, D. (1969). A Theory of Cerebellar Cortex. *Journal of Physiology (London)* **202**, 437–470.

Marr, D. (1970). A Theory for Cerebral Neocortex. *Proceedings of the Royal Society of London B* **176**, 161–234.

Marr, D. (1971). Simple Memory: A Theory for Archicortex. *Philosophical Transactions of the Royal Society of London B* **262**, 23–81.

Marr, D. (1982). *Vision*. San Francisco: Freeman. Partially reprinted in Anderson and Rosenfeld [1988].

Marr, D. and T. Poggio (1976). Cooperative Computation of Stereo Disparity. *Science* **194**, 283–287. Reprinted in Anderson and Rosenfeld [1988].

Mattis, D. (1976). Solvable Spin Systems with Random Interactions. *Physics Letters* **56A**, 421–422.

Mazaika, P.K. (1987). A Mathematical Model of the Boltzmann Machine. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. III, 157–163. New York: IEEE.

McClelland, J.L. and J.L. Elman (1986). Interactive Processes in Speech Perception: The TRACE Model. In *Parallel Distributed Processing*, vol. 2, chap. 15. *See* McClelland, Rumelhart, et al. [1986].

McClelland, J.L. and D.E. Rumelhart (1988). *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises*. Cambridge: MIT Press.

McClelland, J.L., D.E. Rumelhart, and the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 2: Psychological and Biological Models*. Cambridge: MIT Press.

McCulloch, W.S. and W. Pitts (1943). A Logical Calculus of Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics* **5**, 115–133. Reprinted in Anderson and Rosenfeld [1988].

McEliece, R.J., E.C. Posner, E.R. Rodemich, and S.S. Venkatesh (1987). The Capacity of the Hopfield Associative Memory. *IEEE Transactions on Information Theory* **33**, 461–482.

McInerny, J.M., K.G. Haines, S. Biafore, and R. Hecht-Nielsen (1989). Back Propagation Error Surfaces Can Have Local Minima. In *International Joint Conference on Neural Networks* (Washington 1989), vol. II, 627. New York: IEEE.

Mead, C. (1989). *Analog VLSI and Neural Systems*. Reading: Addison Wesley.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller (1953). Equation of State Calculations for Fast Computing Machines. *Journal of Chemical Physics* **21**, 1087–1092.

Mézard, M. and J.-P. Nadal (1989). Learning in Feedforward Layered Networks: The Tiling Algorithm. *Journal of Physics A* **22**, 2191–2204.

Mézard, M. and G. Parisi (1985). Replicas and Optimization. *Journal de Physique Lettres (Paris)* **46**, 771–778.

Mézard, M. and G. Parisi (1986). A Replica Analysis of the Travelling Salesman Problem. *Journal de Physique (Paris)* **47**, 1285–1296.

Mézard, M. and G. Parisi (1988). The Euclidean Matching Problem. *Journal de Physique (Paris)* **49**, 2019–2025.

Mézard, M., G. Parisi, and M.A. Virasoro (1987). *Spin Glass Theory and Beyond*. Singapore: World Scientific.

Miller, G.F., P.M. Todd, and S.U. Hegde (1989). Designing Neural Networks Using Genetic Algorithms. In *Proceedings of the Third International Conference on Genetic Algorithms* (Arlington 1989), ed. J.D. Schaffer, 379–384. San Mateo: Morgan Kaufmann.

Minsky, M.L. (1967). *Computation: Finite and Infinite Machines*. Englewood Cliffs: Prentice-Hall.

Minsky, M.L. and S.A. Papert (1969). *Perceptrons*. Cambridge: MIT Press. Partially reprinted in Anderson and Rosenfeld [1988].

Mitchison, G.J. and R.M. Durbin (1989). Bounds on the Learning Capacity of Some Multi-Layer Networks. *Biological Cybernetics* **60**, 345–356.

Montana, D.J. and L. Davis (1989). Training Feedforward Networks Using Genetic Algorithms. In *Eleventh International Joint Conference on Artificial Intelligence* (Detroit 1989), ed. N.S. Sridharan, 762–767. San Mateo: Morgan Kaufmann.

Moody, J. and C. Darken (1988). Learning with Localized Receptive Fields. In *Proceedings of the 1988 Connectionist Models Summer School* (Pittsburg 1988), eds. D. Touretzky, G. Hinton, and T. Sejnowski, 133–143. San Mateo: Morgan Kaufmann.

Moody, J. and C. Darken (1989). Fast Learning in Networks of Locally-Tuned Processing Units. *Neural Computation* **1**, 281–294.

Moore, B. (1988). ART1 and Pattern Clustering. In *Proceedings of the 1988 Connectionist Models Summer School* (Pittsburg 1988), eds. D. Touretzky, G. Hinton, and T. Sejnowski, 174–185. San Mateo: Morgan Kaufmann.

Mozer, M.C. (1989). A Focused Back-Propagation Algorithm for Temporal Pattern Recognition. *Complex Systems* **3**, 349–381.

Munro, P. (1987). A Dual Back-Propagation Scheme for Scalar Reward Learning. In *The Ninth Annual Conference of the Cognitive Science Society* (Seattle 1987), 165–176. Hillsdale: Erlbaum.

Nadal, J.-P., G. Toulouse, J.-P. Changeux, and S. Dehaene (1986). Networks of Formal Neurons and Memory Palimpsests. *Europhysics Letters* **1**, 535–542.

Narendra, K. and M.A.L. Thathachar (1989). *Learning Automata: An Introduction*. Englewood Cliffs: Prentice-Hall.

Nasrabadi, N.M. and Y. Feng (1988). Vector Quantization of Images Based upon the Kohonen Self-Organizing Feature Maps. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 101–108. New York: IEEE.

Nasrabadi, N.M. and R.A. King (1988). Image Coding Using Vector Quantization: A Review. *IEEE Transactions on Communications* **36**, 957–971.

Nass, M.M. and L.N. Cooper (1975). A Theory for the Development of Feature Detecting Cells in Visual Cortex. *Biological Cybernetics* **19**, 1–18.

Naylor, J. and K.P. Li (1988). Analysis of a Neural Network Algorithm for Vector Quantization of Speech Parameters. *Neural Networks Supplement* **1**, 310.

von Neumann, J. (1956). Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components. In *Automata Studies*, eds. C.E. Shannon and J. McCarthy, 43–98. Princeton: Princeton University Press.

Niranjan, M. and F. Fallside (1990). Neural Networks and Radial Basis Functions in Classifying Static Speech Patterns. *Computer Speech and Language* **4**, 275–289.

Nishimori, H., T. Nakamura, and M. Shiino (1990). Retrieval of Spatio-Temporal Sequence in Asynchronous Neural Network. *Physical Review A* **41**, 3346–3354.

Nowlan, S.J. (1988). Gain Variation in Recurrent Error Propagation Networks. *Complex Systems* **2**, 305–320.

Oja, E. (1982). A Simplified Neuron Model As a Principal Component Analyzer. *Journal of Mathematical Biology* **15**, 267–273.

Oja, E. (1989). Neural Networks, Principal Components, and Subspaces. *International Journal of Neural Systems* **1**, 61–68.

Oja, E. and J. Karhunen (1985). On Stochastic Approximation of the Eigenvectors and Eigenvalues of the Expectation of a Random Matrix. *Journal of Mathematical Analysis and Applications* **106**, 69–84.

Opper, M., W. Kinzel, J. Kleinz, and R. Nehl (1990). On the Ability of the Optimal Perceptron to Generalize. *Journal of Physics A* **23**, L581–L586.

Orland, H. (1985). Mean-Field Theory for Optimization Problems. *Journal de Physique Lettres (Paris)* **46**, 763–770.

Owens, A.J. and D.L. Filkin (1989). Efficient Training of the Back Propagation Network by Solving a System of Stiff Ordinary Differential Equations. In *International Joint Conference on Neural Networks* (Washington 1989), vol. II, 381–386. New York: IEEE.

Packard, N.H., J.P Crutchfield, J.D. Farmer, and R.S. Shaw (1980). Geometry from a Time Series. *Physical Review Letters* **45**, 712–716.

Palmer, R.G. (1988). Statistical Mechanics Approaches to Complex Optimization Problems. In *The Economy As an Evolving Complex System*, eds. P.W. Anderson, K.J. Arrow, and D. Pines. SFI Studies in the Sciences of Complexity, proc. vol. 5, 177–193. Redwood City: Addison-Wesley.

Palmer, R.G. (1989). Neural Nets. In *Lectures in the Sciences of Complexity*, ed. D.L. Stein. SFI Studies in the Sciences of Complexity, lect. vol. 1, 439–461. Redwood City: Addison-Wesley.

Papadimitriou, C.H. and K. Steiglitz (1982). *Combinatorial Optimization: Algorithms and Complexity.* Englewood Cliffs: Prentice-Hall.

Parga, N. and M.A. Virasoro (1986). The Ultrametric Organization of Memories in a Neural Network. *Journal de Physique (Paris)* **47**, 1857–1864.

Parisi, G. (1986). Asymmetric Neural Networks and the Process of Learning. *Journal of Physics A* **19**, L675–L680.

Parker, D.B. (1985). Learning Logic. Technical Report TR–47, Center for Computational Research in Economics and Management Science, Massachusetts Institute of Technology, Cambridge, MA.

Parker, D.B. (1987). Optimal Algorithms for Adaptive Networks: Second Order Back Propagation, Second Order Direct Propagation, and Second Order Hebbian Learning. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. II, 593–600. New York: IEEE.

Parks, M. (1987). Characterization of the Boltzmann Machine Learning Rate. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. III, 715–719. New York: IEEE.

Pearlmutter, B.A. (1989a). Learning State Space Trajectories in Recurrent Neural Networks. In *International Joint Conference on Neural Networks* (Washington 1989), vol. II, 365–372. New York: IEEE.

Pearlmutter, B.A. (1989b). Learning State Space Trajectories in Recurrent Neural Networks. *Neural Computation* **1**, 263–269.

Pearlmutter, B.A. and G.E. Hinton (1986). G-Maximization: An Unsupervised Learning Procedure for Discovering Regularities. In *Neural Networks for Computing* (Snowbird 1986), ed. J.S. Denker, 333–338. New York: American Institute of Physics.

Peretto, P. (1984). Collective Properties of Neural Networks: A Statistical Physics Approach. *Biological Cybernetics* **50**, 51–62.

Peretto, P. (1988). On Learning Rules and Memory Storage Abilities of Asymmetrical Neural Networks. *Journal de Physique (Paris)* **49**, 711–726.

Peretto, P. and J.J. Niez (1986). Collective Properties of Neural Networks. In *Disordered Systems and Biological Organization* (Les Houches 1985), eds. E. Bienenstock, F. Fogelman-Soulié, and G. Weisbuch, 171–185. Berlin: Springer-Verlag.

Pérez, R., L. Glass, and R. Shlaer (1975). Development of Specificity in the Cat Visual Cortex. *Journal of Mathematical Biology* **1**, 275–288.

Personnaz, L., I. Guyon, and G. Dreyfus (1985). Information Storage and Retrieval in Spin-Glass-Like Neural Networks. *Journal de Physique Lettres (Paris)* **46**, 359–365.

Personnaz, L., I. Guyon, and G. Dreyfus (1986). Collective Computational Properties of Neural Networks: New Learning Mechanisms. *Physical Review A* **34**, 4217–4228.

Peterson, C. and J.R. Anderson (1987). A Mean Field Theory Learning Algorithm for Neural Networks. *Complex Systems* **1**, 995–1019.

Peterson, C., S. Redfield, J.D. Keeler, and E. Hartman (1990). An Optoelectronic Architecture for Multilayer Learning in a Single Photorefractive Crystal. *Neural Computation* **2**, 25–34.

Peterson, C. and B. Söderberg (1989). A New Method for Mapping Optimization Problems onto Neural Networks. *International Journal of Neural Systems* **1**, 3–22.

Pineda, F.J. (1987). Generalization of Back-Propagation to Recurrent Neural Networks. *Physical Review Letters* **59**, 2229–2232.

Pineda, F.J. (1988). Dynamics and Architecture for Neural Computation. *Journal of Complexity* **4**, 216–245.

Pineda, F.J. (1989). Recurrent Back-Propagation and the Dynamical Approach to Adaptive Neural Computation. *Neural Computation* **1**, 161–172.

Plaut, D., S. Nowlan, and G. Hinton (1986). Experiments on Learning by Back Propagation. Technical Report CMU–CS–86–126, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Poggio, T. and F. Girosi (1990). Regularization Algorithms for Learning That Are Equivalent to Multilayer Networks. *Science* **247**, 978–982.

Poggio, T., V. Torre, and C. Koch (1985). Computational Vision and Regularization Theory. *Nature* **317**, 314–319.

Pomerleau, D.A. (1989). ALVINN: An Autonomous Land Vehicle in a Neural Network. In *Advances in Neural Information Processing Systems I* (Denver 1988), ed. D.S. Touretzky, 305–313. San Mateo: Morgan Kaufmann.

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling (1986). *Numerical Recipes.* Cambridge: Cambridge University Press.

Qian, N. and T.J. Sejnowski (1988a). Predicting the Secondary Structure of Globular Proteins Using Neural Network Models. *Journal of Molecular Biology* **202**, 865–884.

Qian, N. and T.J. Sejnowski (1988b). Learning to Solve Random-Dot Stereograms of Dense Transparent Surfaces with Recurrent Back-Propagation. In *Proceedings of the 1988 Connectionist Models Summer School* (Pittsburg 1988), eds. D. Touretzky, G. Hinton, and T. Sejnowski, 435–443. San Mateo: Morgan Kaufmann.

Rao, C.R. and S.K. Mitra (1971). *Generalized Inverse of Matrices and Its Applications.* New York: Wiley.

Ramanujam, J. and P. Sadayappan (1988). Optimization by Neural Networks. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. II, 325–332. New York: IEEE.

Rashevsky, N. (1938). *Mathematical Biophysics.* Chicago: University of Chicago Press.

Rescorla, R.A. and A.R. Wagner (1972). A Theory of Pavlovian Conditioning: The Effectiveness of Reinforcement and Nonreinforcement. In *Classical Conditioning II: Current Research and Theory*, eds. A.H. Black and W.F. Prokasy, 64–69. New York: Appleton-Century-Crofts.

Ricotti, L.P., S. Ragazzini, and G. Martinelli (1988). Learning of Word Stress in a Sub-Optimal Second Order Back-Propagation Neural Network. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 355–361. New York: IEEE.

Riedel, U., R. Kühn, and J.L. van Hemmen (1988). Temporal Sequences and Chaos in Neural Nets. *Physical Review A* **38**, 1105–1108.

Ritter, H. and K. Schulten (1986). On the Stationary State of Kohonen's Self-Organizing Sensory Mapping. *Biological Cybernetics* **54**, 99–106.

Ritter, H. and K. Schulten (1988a). Extending Kohonen's Self-Organizing Mapping Algorithm to Learn Ballistic Movements. In *Neural Computers* (Neuss 1987), eds. R. Eckmiller and Ch. von der Malsburg, 393–406. Berlin: Springer-Verlag.

Ritter, H. and K. Schulten (1988b). Convergence Properties of Kohonen's Topology Conserving Maps: Fluctuations, Stability, and Dimension Selection. *Biological Cybernetics* **60**, 59–71.

Ritter, H. and K. Schulten (1988c). Kohonen's Self-Organizing Maps: Exploring Their Computational Capabilities. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 109–116. New York: IEEE.

Robinson, A.J. and F. Fallside (1988). Static and Dynamic Error Propagation Networks with Application to Speech Coding. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 632–641. New York: American Institute of Physics.

Rohwer, R. (1990). The "Moving Targets" Training Algorithm. In *Advances in Neural Information Processing Systems II* (Denver 1989), ed. D.S. Touretzky, 558–565. San Mateo: Morgan Kaufmann.

Rohwer, R. and B. Forrest (1987). Training Time-Dependence in Neural Networks. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. II, 701–708. New York: IEEE.

Romeo, F.I. (1989). Simulated Annealing: Theory and Applications to Layout Problems. Ph.D. Thesis, University of California at Berkeley Memorandum UCB/ERL–M89/29.

Rose, D. and V.G. Dobson, eds. (1985). *Models of the Visual Cortex.* Chichester: Wiley.

Rosenblatt, F. (1962). *Principles of Neurodynamics.* New York: Spartan.

Rubner, J. and K. Schulten (1990). Development of Feature Detectors by Self-Organization. *Biological Cybernetics* **62**, 193–199.

Rubner, J. and P. Tavan (1989). A Self-Organizing Network for Principal-Component Analysis. *Europhysics Letters* **10**, 693–698.

Rumelhart, D.E., G.E. Hinton, and R.J. Williams (1986a). Learning Representations by Back-Propagating Errors. *Nature* **323**, 533–536. Reprinted in Anderson and Rosenfeld [1988].

Rumelhart, D.E., G.E. Hinton, and R.J. Williams (1986b). Learning Internal Representations by Error Propagation. In *Parallel Distributed Processing*, vol. 1, chap. 8. *See* Rumelhart, McClelland, et al. [1986]. Reprinted in Anderson and Rosenfeld [1988].

Rumelhart, D.E., J.L. McClelland, and the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1: Foundations.* Cambridge: MIT Press.

Rumelhart, D.E. and D. Zipser (1985). Feature Discovery by Competitive Learning. *Cognitive Science* **9**, 75–112. Also reprinted in *Parallel Distributed Processing*, vol. 1, chap. 5. *See* Rumelhart, McClelland, et al. [1986].

Saad, D. and E. Marom (1990a). Learning by Choice of Internal Representations—An Energy Minimization Approach. Preprint, Faculty of Engineering, Tel Aviv University, Ramat-Aviv, Israel.

Saad, D. and E. Marom (1990b). Training Feed Forward Nets with Binary Weights via a Modified CHIR Algorithm. Preprint, Faculty of Engineering, Tel Aviv University, Ramat-Aviv, Israel.

Salamon, P., J.D. Nulton, J. Robinson, J. Petersen, G. Ruppeiner, and L. Liao (1988). Simulated Annealing with Constant Thermodynamic Speed. *Computer Physics Communications* **49**, 423–428.

Sanger, T.D. (1989a). Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network. *Neural Networks* **2**, 459–473.

Sanger, T.D. (1989b). An Optimality Principle for Unsupervised Learning. In *Advances in Neural Information Processing Systems I* (Denver 1988), ed. D.S. Touretzky, 11–19. San Mateo: Morgan Kaufmann.

Sato, M. (1990). A Real Time Learning Algorithm for Recurrent Analog Neural Networks. *Biological Cybernetics* **62**, 237–241.

Scalettar, R. and A. Zee (1988). Emergence of Grandmother Memory in Feed Forward Networks: Learning with Noise and Forgetfulness. In *Connectionist Models and Their Implications: Readings from Cognitive Science*, eds. D. Waltz and J.A. Feldman, 309–332. Norwood: Ablex.

Schwartz, D.B., V.K. Samalam, S.A. Solla, and J.S. Denker (1990). Exhaustive Learning. *Neural Computation* **2**, 371–382.

Scofield, C.L. (1988). Learning Internal Representations in the Coulomb Energy Network. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 271–276. New York: IEEE.

Sejnowski, T.J., P.K. Kienker, and G. Hinton (1986). Learning Symmetry Groups with Hidden Units: Beyond the Perceptron. *Physica* **22D**, 260–275.

Sejnowski, T.J. and C.R. Rosenberg (1987). Parallel Networks that Learn to Pronounce English Text. *Complex Systems* **1**, 145–168.

Sherrington, D. and S. Kirkpatrick (1975). Solvable Model of a Spin Glass. *Physical Review Letters* **35**, 1792–1796.

Shimohara, K., T. Uchiyama, and Y. Tokunaga (1988). Back-Propagation Networks for Event-Driven Temporal Sequence Processing. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 665–672. New York: IEEE.

Sietsma, J. and R.J.F. Dow (1988). Neural Net Pruning—Why and How. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 325–333. New York: IEEE.

Silverman, R.H. and A.S. Noetzel (1988). Time-Sequential Self-Organization of Hierarchical Neural Networks. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 709–714. New York: American Institute of Physics.

Simard, P.Y., M.B. Ottaway, and D.H. Ballard (1989). Analysis of Recurrent Backpropagation. In *Proceedings of the 1988 Connectionist Models Summer School* (Pittsburg 1988), eds. D. Touretzky, G. Hinton, and T. Sejnowski, 103–112. San Mateo: Morgan Kaufmann.

Simic, P.D. (1990). Statistical Mechanics As the Underlying Theory of "Elastic" and "Neural" Optimizations. *Network* **1**, 89–103.

Sirat, J.-A. and J.-P. Nadal (1990). Neural Trees: A New Tool for Classification. Preprint, Laboratoires d'Electronique Philips, Limeil-Brevannes, France.

Sivilotti, M.A., M.A. Mahowald, and C.A. Mead (1987). Real-Time Visual Computations Using Analog CMOS Processing Arrays. In *Advanced Research in VLSI: Proceedings of the 1987 Stanford Conference*, ed. P. Losleben, 295–312. Cambridge: MIT Press. Reprinted in Anderson and Rosenfeld [1988].

Smolensky, P. (1986). Information Processing in Dynamical Systems: Foundations of Harmony Theory. In *Parallel Distributed Processing*, vol. 1, chap. 6. *See* Rumelhart, McClelland, et al. [1986].

Soffer, B.H., G.J. Dunning, Y. Owechko, and E. Marom (1986). Associative Holographic Memory with Feedback Using Phase-Conjugate Mirrors. *Optics Letters* **11**, 118–120.

Solla, S.A. (1989). Learning and Generalization in Layered Neural Networks: The Contiguity Problem. In *Neural Networks from Models to Applications* (Paris 1988), eds. L. Personnaz and G. Dreyfus, 168–177. Paris: I.D.S.E.T.

Solla, S.A., E. Levin, and M. Fleisher (1988). Accelerated Learning in Layered Neural Networks. *Complex Systems* **2**, 625–639.

Sompolinsky, H. (1987). The Theory of Neural Networks: The Hebb Rules and Beyond. In *Heidelberg Colloquium on Glassy Dynamics* (Heidelberg 1986), eds. J.L. van Hemmen and I. Morgenstern, 485–527. Berlin: Springer-Verlag.

Sompolinsky, H., A. Crisanti, and H.J. Sommers (1988). Chaos in Random Neural Networks. *Physical Review Letters* **61**, 259–262.

Sompolinsky, H. and I. Kanter (1986). Temporal Association in Asymmetric Neural Networks. *Physical Review Letters* **57**, 2861–2864.

Soukoulis, C.M., K. Levin, and G.S. Grest (1983). Irreversibility and Metastability in Spin-Glasses. I. Ising Model. *Physical Review B* **28**, 1495–1509.

Specht, D.F. (1990). Probabilistic Neural Networks. *Neural Networks* **3**, 109–118.

Steinbuch, K. (1961). Die Lernmatrix. *Kybernetik* **1**, 36–45.

Stokbro, K., D.K. Umberger, and J.A. Hertz (1990). Exploiting Neurons with Localized Receptive Fields to Learn Chaos. Preprint 90/28 S, Nordita, Copenhagen, Denmark.

Stornetta, W.S., T. Hogg, and B.A. Huberman (1988). A Dynamical Approach to Temporal Pattern Processing. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 750–759. New York: American Institute of Physics.

Sutton, R.S. (1984). Temporal Credit Assignment in Reinforcement Learning. Ph. D. Thesis, University of Massachusetts, Amherst.

Sutton, R.S. (1988). Learning to Predict by the Methods of Temporal Differences. *Machine Learning* **3**, 9–44.

Sutton, R.S. and A.G. Barto (1991). Time Derivative Models of Pavlovian Reinforcement. In *Learning and Computational Neuroscience*, eds. M. Gabriel and J.W. Moore. Cambridge: MIT Press.

Szu, H. (1986). Fast Simulated Annealing. In *Neural Networks for Computing* (Snowbird 1986), ed. J.S. Denker, 420–425. New York: American Institute of Physics.

Takens, F. (1981). Detecting Strange Attractors In Turbulence. In *Dynamical Systems and Turbulence*. Lecture Notes in Mathematics, vol. 898 (Warwick 1980), eds. D.A. Rand and L.-S. Young, 366–381. Berlin: Springer-Verlag.

Takeuchi, A. and S. Amari (1979). Formation of Topographic Maps and Columnar Microstructures in Nerve Fields. *Biological Cybernetics* **35**, 63–72.

Tank, D.W. and J.J. Hopfield (1986). Simple "Neural" Optimization Networks: An A/D Converter, Signal Decision Circuit, and a Linear Programming Circuit. *IEEE Transactions on Circuits and Systems* **33**, 533–541.

Tank, D.W. and J.J. Hopfield (1987a). Neural Computation by Concentrating Information in Time. *Proceedings of the National Academy of Sciences, USA* **84**, 1896–1900.

Tank, D.W. and J.J. Hopfield (1987b). Concentrating Information in Time: Analog Neural Networks with Applications to Speech Recognition Problems. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. IV, 455–468. New York: IEEE.

Taylor, W.K. (1956). Electrical Simulation of Some Nervous System Functional Activities. In *Information Theory* (London 1985), ed. C. Cherry, 314–328. London: Butterworths.

Tesauro, G. (1986). Simple Neural Models of Classical Conditioning. *Biological Cybernetics* **55**, 187–200.

Tesauro, G. (1990). Neurogammon Wins Computer Olympiad. *Neural Computation* **1**, 321–323.

Tesauro, G. and B. Janssens (1988). Scaling Relationships in Back-Propagation Learning. *Complex Systems* **2**, 39–44.

Tesauro, G. and T.J. Sejnowski (1988). A "Neural" Network That Learns to Play Backgammon. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 442–456. New York: American Institute of Physics.

Thakoor, A.P., A. Moopenn, J. Lambe, and S.K. Khanna (1987). Electronic Hardware Implementations of Neural Networks. *Applied Optics* **26**, 5085–5092.

Ticknor, A.J. and H. Barrett (1987). Optical Implementations of Boltzmann Machines. *Optical Engineering* **26**, 16–21.

Tishby, N., E. Levin, and S.A. Solla (1989). Consistent Inference of Probabilities in Layered Networks: Predictions and Generalization. In *International Joint Conference on Neural Networks* (Washington 1989), vol. II, 403–410. New York: IEEE.

Toulouse, G., S. Dehaene, and J.-P. Changeux (1986). Spin Glass Model of Learning by Selection. *Proceedings of the National Academy of Sciences, USA* **83**, 1695–1698.

Touretzky, D.S. and D.A. Pomerleau (1989). What's Hidden in the Hidden Layers? *BYTE*, August 1989, 227–233.

Tsodyks, M.V. and M.V. Feigel'man (1988). The Enhanced Storage Capacity in Neural Networks with Low Activity Level. *Europhysics Letters* **6**, 101–105.

Van den Bout, D.E. and T.K. Miller (1988). A Travelling Salesman Objective Function That Works. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. II, 299–303. New York: IEEE.

Van den Bout, D.E. and T.K. Miller (1989). Improving the Performance of the Hopfield-Tank Neural Network Through Normalization and Annealing. *Biological Cybernetics* **62**, 129–139.

Vapnik, V.N. (1982). *Estimation of Dependences Based on Empirical Data.* Berlin: Springer-Verlag.

Vapnik, V.N. and A.Y. Chervonenkis (1971). On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities. *Theory of Probability and Its Applications* **16**, 264–280.

Vogl, T.P., J.K. Mangis, A.K. Rigler, W.T. Zink, and D.L. Alkon (1988). Accelerating the Convergence of the Back-Propagation Method. *Biological Cybernetics* **59**, 257–263.

Wagner, K. and D. Psaltis (1987). Multilayer Optical Learning Networks. *Applied Optics* **26**, 5061–5076.

Waibel, A. (1989). Modular Construction of Time-Delay Neural Networks for Speech Recognition. *Neural Computation* **1**, 39–46.

Waibel, A., T. Hanazawa, G. Hinton, K. Shikano, and K. Lang (1989). Phoneme Recognition Using Time-Delay Neural Networks. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **37**, 328–339.

Watrous, R.L. (1987). Learning Algorithms for Connectionist Networks: Applied Gradient Methods of Nonlinear Optimization. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. II, 619–627. New York: IEEE.

Webster (1988). *Webster's Ninth New Collegiate Dictionary.* Springfield: Merriam-Webster.

Weisbuch, G. and F. Fogelman-Soulié (1985). Scaling Laws for the Attractors of Hopfield Networks. *Journal de Physique Lettres (Paris)* **46**, 623–630.

Werbos, P. (1974). Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. Thesis, Harvard University.

Werbos, P.J. (1987). Building and Understanding Adaptive Systems: A Statistical/Numerical Approach to Factory Automation and Brain Research. *IEEE Transactions on Systems, Man, and Cybernetics* **17**, 7–20.

Werbos, P.J. (1988). Generalization of Backpropagation with Application to a Recurrent Gas Market Model. *Neural Networks* **1**, 339–356.

Whitley, D. and T. Hanson (1989). Optimizing Neural Networks Using Faster, More Accurate Genetic Search. In *Proceedings of the Third International Conference on Genetic Algorithms* (Arlington 1989), ed. J.D. Schaffer, 391–396. San Mateo: Morgan Kaufmann.

Widrow, B. (1962). Generalization and Information Storage in Networks of Adaline "Neurons". In *Self-Organizing Systems 1962* (Chicago 1962), eds. M.C. Yovits, G.T. Jacobi, and G.D. Goldstein, 435–461. Washington: Spartan.

Widrow, B., N.K. Gupta, and S. Maitra (1973). Punish/Reward: Learning with a Critic in Adaptive Threshold Systems. *IEEE Transactions on Systems, Man, and Cybernetics* **3**, 455–465.

Widrow, B. and M.E. Hoff (1960). Adaptive Switching Circuits. In *1960 IRE WESCON Convention Record*, part 4, 96–104. New York: IRE. Reprinted in Anderson and Rosenfeld [1988].

Wiener, N. (1948). *Cybernetics, or Control and Communication in the Animal and the Machine*. New York: Wiley.

Williams, R.J. (1987). A Class of Gradient-Estimating Algorithms for Reinforcement Learning in Neural Networks. In *IEEE First International Conference on Neural Networks* (San Diego 1987), eds. M. Caudill and C. Butler, vol. II, 601–608. New York: IEEE.

Williams, R.J. (1988a). On the Use of Back-Propagation in Associative Reinforcement Learning. In *IEEE International Conference on Neural Networks* (San Diego 1988), vol. I, 263–270. New York: IEEE.

Williams, R.J. (1988b). Towards a Theory of Reinforcement-Learning Connectionist Systems. Technical Report NU–CCS–88–3, College of Computer Science, Northeastern University, Boston, MA.

Williams, R.J. and J. Peng (1989). Reinforcement Learning Algorithms As Function Optimizers. In *International Joint Conference on Neural Networks* (Washington 1989), vol. II, 89–95. New York: IEEE.

Williams, R.J. and D. Zipser (1989a). A Learning Algorithm for Continually Running Fully Recurrent Neural Networks. *Neural Computation* **1**, 270–280.

Williams, R.J. and D. Zipser (1989b). Experimental Analysis of the Real-Time Recurrent Learning Algorithm. *Connection Science* **1**, 87–111.

Willshaw, D.J., O.P. Buneman, and H.C. Longuet-Higgins (1969). Non-Holographic Associative Memory. *Nature* **222**, 960–962. Reprinted in Anderson and Rosenfeld [1988].

Willshaw, D.J. and C. von der Malsburg (1976). How Patterned Neural Connections Can Be Set Up by Self-Organization. *Proceedings of the Royal Society of London B* **194**, 431–445.

Wilson, G.V. and G.S. Pawley (1988). On the Stability of the Travelling Salesman Problem Algorithm of Hopfield and Tank. *Biological Cybernetics* **58**, 63–70.

Wilson, H.R. and J.D. Cowan (1973). A Mathematical Theory of the Functional Dynamics of Cortical and Thalamic Nervous Tissue. *Kybernetik* **13**, 55–80.

Winograd, S. and J.D. Cowan (1963). *Reliable Computation in the Presence of Noise*. Cambridge: MIT Press.

Winters, J.H. and C. Rose (1989). Minimum Distance Automata in Parallel Networks for Optimum Classification. *Neural Networks* **2**, 127–132.

Wittgenstein, L. (1958). *Philosophical Investigations*. 2nd ed., translated by G.E.M. Anscombe. Oxford: Blackwell.

Wittner, B.S. and J.S. Denker (1988). Strategies for Teaching Layered Networks Classification Tasks. In *Neural Information Processing Systems* (Denver 1987), ed. D.Z. Anderson, 850–859. New York: American Institute of Physics.

Yuille, A.L., D.M. Kammen, and D.S. Cohen (1989). Quadrature and the Development of Orientation Selective Cortical Cells by Hebb Rules. *Biological Cybernetics* **61**, 183–194.

Zak, M. (1988). Terminal Attractors for Addressable Memory in Neural Networks. *Physics Letters* **133A**, 18–22.

Zak, M. (1989). Terminal Attractors in Neural Networks. *Neural Networks* **2**, 259–274.