Kevin Gurney

AN INTRODUCTION TO NEURAL NETWORKS

An introduction to neural networks

An introduction to neural networks

Kevin Gurney
University of Sheffield

London and New York

First published in 1997 by UCL Press

Reprinted 1999

UCL Press Limited
11 New Fetter Lane
London EC4P 4EE

UCL Press Limited is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2005.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

ISBN 0-203-45151-1 Master e-book ISBN

ISBN 0-203-45874-5 (Adobe eReader Format)
ISBNs: 1-85728-673-1 (Print Edition) HB
1-85728-503-4 (Print Edition) PB

Contents

Preface

This book grew out of a set of course notes for a neural networks module given as part of a Masters degree in "Intelligent Systems". The people on this course came from a wide variety of intellectual backgrounds (from philosophy, through psychology to computer science and engineering) and I knew that I could not count on their being able to come to grips with the largely technical and mathematical approach which is often used (and in some ways easier to do). As a result I was forced to look carefully at the basic conceptual principles at work in the subject and try to recast these using ordinary language, drawing on the use of physical metaphors or analogies, and pictorial or graphical representations. I was pleasantly surprised to find that, as a result of this process, my own understanding was considerably deepened; I had now to unravel, as it were, condensed formal descriptions and say exactly *how* these were related to the "physical" world of artificial neurons, signals, computational processes, etc. However, I was acutely aware that, while a litany of equations does not constitute a full description of fundamental principles, without some mathematics, a purely descriptive account runs the risk of dealing only with approximations and cannot be sharpened up to give any formulaic prescriptions. Therefore, I introduced what I believed was just sufficient mathematics to bring the basic ideas into sharp focus.

To allay any residual fears that the reader might have about this, it is useful to distinguish two contexts in which the word "maths" might be used. The first refers to the use of symbols to stand for quantities and is, in this sense, merely a shorthand. For example, suppose we were to calculate the difference between a target neural output and its actual output and then multiply this difference by a constant learning rate (it is not important that the reader knows what these terms mean just now). If t stands for the target, y the actual output, and the learning rate is denoted by a (Greek "alpha") then the output-difference is just $(t-y)$ and the verbose description of the calculation may be reduced to $\alpha(t-y)$. In this example the symbols refer to numbers but it is quite possible they may refer to other mathematical quantities or objects. The two instances of this used here are *vectors* and *function gradients*. However, both these ideas are described at some length in the main body of the text and assume no prior knowledge in this respect. In each case, only enough is given for the purpose in hand; other related, technical material may have been useful but is not considered essential and it is not one of the aims of this book to double as a mathematics primer.

The other way in which we commonly understand the word "maths" goes one step further and deals with the rules by which the symbols are manipulated. The only rules used in this book are those of simple arithmetic (in the above example we have a subtraction and a multiplication). Further, any manipulations (and there aren't many of them) will be performed step by step. Much of the traditional "fear of maths" stems, I believe, from the apparent difficulty in inventing the right manipulations to go from one stage to another; the reader will not, in this book, be called on to do this for him- or herself.

One of the spin-offs from having become familiar with a certain amount of mathematical formalism is that it enables contact to be made with the rest of the neural network literature. Thus, in the above example, the use of the Greek letter α may seem gratuitous (why not use a, the reader asks) but it turns out that learning rates are often denoted by lower case Greek letters and a is not an uncommon choice. To help in this respect, Greek symbols will always be accompanied by their name on first use.

In deciding how to present the material I have started from the bottom up by describing the properties of artificial neurons (Ch. 2) which are motivated by looking at the nature of their real counterparts. This emphasis on the biology is intrinsically useful from a computational neuroscience perspective and helps people from all disciplines appreciate exactly how "neural" (or not) are the networks they intend to use. Chapter 3 moves to networks and introduces the geometric perspective on network function offered by the notion of linear separability in pattern space. There are other viewpoints that might have been deemed primary (function approximation is a favourite contender) but linear separability relates directly to the function of single threshold logic units (TLUs) and enables a discussion of one of the simplest learning rules (the perceptron rule) in Chapter 4. The geometric approach also provides a natural vehicle for the introduction of vectors. The inadequacies of the perceptron rule lead to a discussion of gradient descent and the delta rule (Ch. 5) culminating in a description of backpropagation (Ch. 6). This introduces multilayer nets in full and is the natural point at which to discuss networks as function approximators, feature detection and generalization.

This completes a large section on feedforward nets. Chapter 7 looks at Hopfield nets and introduces the idea of state-space attractors for associative memory and its accompanying energy metaphor. Chapter 8 is the first of two on self-organization and deals with simple competitive nets, Kohonen self-organizing feature maps, linear vector quantization and principal component analysis. Chapter 9 continues the theme of self-organization with a discussion of adaptive resonance theory (ART). This is a somewhat neglected topic (especially in more introductory texts) because it is often thought to contain rather difficult material. However, a novel perspective on ART which makes use of a hierarchy of analysis is aimed at helping the reader in understanding this worthwhile area. Chapter 10 comes full circle and looks again at alternatives to the artificial neurons introduced in Chapter 2. It also briefly reviews some other feedforward network types and training algorithms so that the reader does not come away with the impression that backpropagation has a monopoly here. The final chapter tries to make sense of the seemingly disparate collection of objects that populate the neural network universe by introducing a series of taxonomies for network architectures, neuron types and algorithms. It also places the study of nets in the general context of that of artificial intelligence and closes with a brief history of its research.

The usual provisos about the range of material covered and introductory texts apply; it is neither possible nor desirable to be exhaustive in a work of this nature. However, most of the major network types have been dealt with and, while there are a plethora of training algorithms that might have been included (but weren't) I believe that an understanding of those presented here should give the reader a firm foundation for understanding others they may encounter elsewhere.

Chapter One
Neural networks—an overview

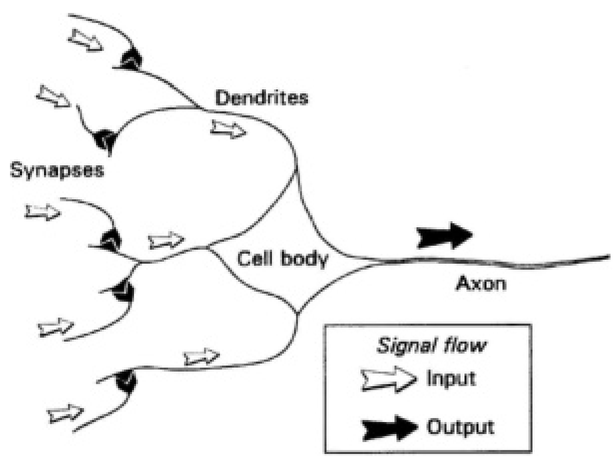

Figure 1.1 Essential components of a neuron shown in stylized form.

The term "Neural networks" is a very evocative one. It suggests machines that are something like brains and is potentially laden with the science fiction connotations of the Frankenstein mythos. One of the main tasks of this book is to demystify neural networks and show how, while they indeed have something to do with brains, their study also makes contact with other branches of science, engineering and mathematics. The aim is to do this in as non-technical a way as possible, although some mathematical notation is essential for specifying certain rules, procedures and structures quantitatively. Nevertheless, all symbols and expressions will be explained as they arise so that, hopefully, these should not get in the way of the essentials: that is, concepts and ideas that may be described in words.

This chapter is intended for orientation. We attempt to give simple descriptions of what networks are and why we might study them. In this way, we have something in mind right from the start, although the whole of this book is, of course, devoted to answering these questions in full.

1.1
What are neural networks?

Let us commence with a provisional definition of what is meant by a "neural network" and follow with simple, working explanations of some of the key terms in the definition.

> A neural network is an interconnected assembly of simple processing elements, *units* or *nodes,* whose functionality is loosely based on the animal neuron. The processing ability of the network is stored in the interunit connection strengths, or *weights,* obtained by a process of adaptation to, or *learning* from, a set of training patterns.

To flesh this out a little we first take a quick look at some basic neurobiology. The human brain consists of an estimated 10^{11} (100 billion) nerve cells or *neurons,* a highly stylized example of which is shown in Figure 1.1. Neurons communicate via electrical signals that are short-lived impulses or "spikes" in the voltage of the cell wall or *membrane.* The interneuron connections are mediated by electrochemical junctions called *synapses,* which are located on branches of the cell referred to as *dendrites*. Each neuron typically receives many thousands of connections from other neurons and is therefore constantly receiving a multitude of incoming signals, which eventually reach the cell body. Here, they are integrated or summed together

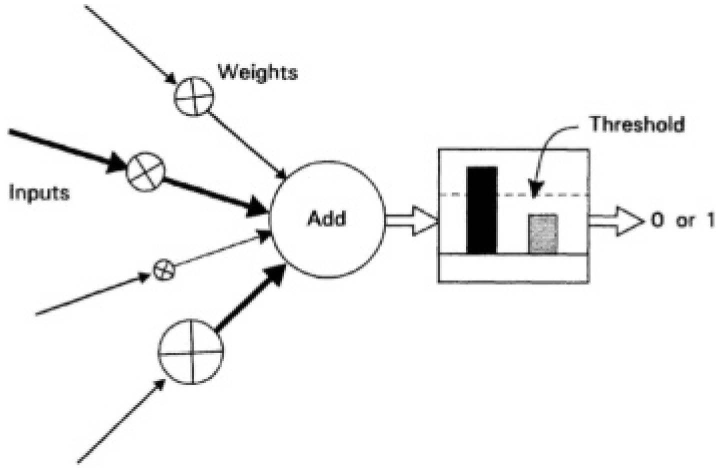

Figure 1.2 Simple artificial neuron.

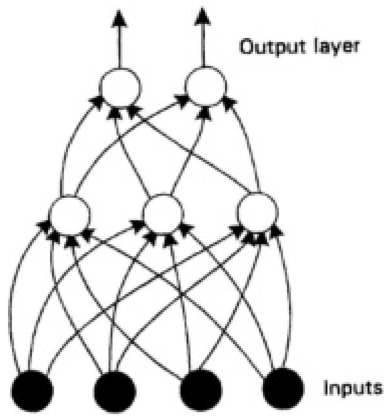

Figure 1.3 Simple example of neural network.

in some way and, roughly speaking, if the resulting signal exceeds some threshold then the neuron will "fire" or generate a voltage impulse in response. This is then transmitted to other neurons via a branching fibre known as the *axon*.

In determining whether an impulse should be produced or not, some incoming signals produce an inhibitory effect and tend to prevent firing, while others are excitatory and promote impulse generation. The distinctive processing ability of each neuron is then supposed to reside in the type—excitatory or inhibitory— and strength of its synaptic connections with other neurons.

It is this architecture and style of processing that we hope to incorporate in neural networks and, because of the emphasis on the importance of the interneuron connections, this type of system is sometimes referred to as being *connectionist* and the study of this general approach as *connectionism*. This terminology is often the one encountered for neural networks in the context of psychologically inspired models of human cognitive function. However, we will use it quite generally to refer to neural networks without reference to any particular field of application.

The artificial equivalents of biological neurons are the nodes or units in our preliminary definition and a prototypical example is shown in Figure 1.2. Synapses are modelled by a single number or *weight* so that each input is multiplied by a weight before being sent to the equivalent of the cell body. Here, the weighted signals are summed together by simple arithmetic addition to supply a node *activation*. In the type of node shown in Figure 1.2—the so-called *threshold logic unit* (TLU)—the activation is then compared with a threshold; if the activation exceeds the threshold, the unit produces a high-valued output (conventionally "1"), otherwise it outputs zero. In the figure, the size of signals is represented by the width of their corresponding arrows, weights are shown by multiplication symbols in circles, and their values are supposed to be proportional to the symbol's size; only positive weights have been used. The TLU is the simplest (and historically the earliest (McCulloch & Pitts 1943)) model of an artificial neuron.

The term "network" will be used to refer to any system of artificial neurons. This may range from something as simple as a single node to a large collection of nodes in which each one is connected to every other node in the net. One type of network is shown in Figure 1.3. Each node is now shown by only a circle but weights are implicit on all connections. The nodes are arranged in a layered structure in which each signal emanates from an input and passes via two nodes before reaching an output beyond which it is no longer transformed. This *feedforward* structure is only one of several available and is typically used to place an input pattern into one of several classes according to the resulting pattern of outputs. For example, if the input consists of an encoding of the patterns of light and dark in an image of handwritten letters, the output layer (topmost in the figure) may contain 26 nodes—one for each letter of the alphabet—to flag which letter class the input character is from. This would be done by allocating one output node per class and requiring that only one such node fires whenever a pattern of the corresponding class is supplied at the input.

So much for the basic structural elements and their operation. Returning to our working definition, notice the emphasis on learning from experience. In real neurons the synaptic strengths may, under certain circumstances, be modified so that the behaviour of each neuron can change or adapt to its particular stimulus input. In artificial neurons the equivalent of this is the modification of the weight values. In terms of processing information, there are no computer programs here —the "knowledge" the network has is supposed to be stored in its weights, which evolve by a process of adaptation to stimulus from a set of pattern examples. In one training paradigm called *supervised learning,* used in conjunction with nets of the type shown in Figure 1.3, an input pattern is presented to the net and its response then compared with a target output. In terms of our previous letter recognition example, an "A", say, may be input and the network output compared with the classification code for A. The difference between the two patterns of output then determines how the weights are altered. Each particular recipe for change constitutes a *learning rule,* details of which form a substantial part of subsequent chapters. When the required weight updates have been made another pattern is presented, the output compared with the target, and new changes made. This sequence of events is repeated iteratively many times until (hopefully) the network's behaviour converges so that its response to each pattern is close to the corresponding target. The process as a whole, including any ordering of pattern presentation, criteria for terminating the process, etc., constitutes the *training algorithm*.

What happens if, after training, we present the network with a pattern it hasn't seen before? If the net has learned the underlying structure of the problem domain then it should classify the unseen pattern correctly and the net is said to *generalize* well. If the net does not have this property it is little more than a classification lookup table for the training set and is of little practical use. Good generalization is therefore one of the key properties of neural networks.

1.2
Why study neural networks?

This question is pertinent here because, depending on one's motive, the study of connectionism can take place from differing perspectives. It also helps to know what questions we are trying to answer in order to avoid the kind of religious wars that sometimes break out when the words "connectionism" or "neural network" are mentioned.

Neural networks are often used for statistical analysis and data modelling, in which their role is perceived as an alternative to standard nonlinear regression or cluster analysis techniques (Cheng & Titterington 1994). Thus, they are typically used in problems that may be couched in terms of classification, or forecasting. Some examples include image and speech recognition, textual character recognition, and domains of human expertise such as medical diagnosis, geological survey for oil, and financial market indicator prediction. This type of problem also falls within the domain of classical artificial intelligence (AI) so that engineers and computer scientists see neural nets as offering a style of *parallel distributed computing,* thereby providing an alternative to the conventional algorithmic techniques that have dominated in machine intelligence. This is a theme pursued further in the final chapter but, by way of a brief explanation of this term now, the parallelism refers to the fact that each node is conceived of as operating independently and concurrently (in parallel with) the others, and the "knowledge" in the network is distributed over the entire set of weights, rather than focused in a few memory locations as in a conventional computer. The practitioners in this area do not concern themselves with biological realism and are often motivated by the ease of implementing solutions in digital hardware or the efficiency and accuracy of particular techniques. Haykin (1994) gives a comprehensive survey of many neural network techniques from an engineering perspective.

Neuroscientists and psychologists are interested in nets as computational models of the animal brain developed by abstracting what are believed to be those properties of real nervous tissue that are essential for information processing. The artificial neurons that connectionist models use are often extremely simplified versions of their biological counterparts and many neuroscientists are sceptical about the ultimate power of these impoverished models, insisting that more detail is necessary to explain the brain's function. Only time will tell but, by drawing on knowledge about how real neurons are interconnected as local "circuits", substantial inroads have been made in modelling brain functionality. A good introduction to this programme of *computational neuroscience* is given by Churchland & Sejnowski (1992).

Finally, physicists and mathematicians are drawn to the study of networks from an interest in nonlinear dynamical systems, statistical mechanics and automata theory.[1] It is the job of applied mathematicians to discover and formalize the properties of new systems using tools previously employed in other areas of science. For example, there are strong links between a certain type of net (the Hopfield net—see Ch. 7) and magnetic systems known as spin glasses. The full mathematical apparatus for exploring these links is developed (alongside a series of concise summaries) by Amit (1989).

All these groups are asking different questions: neuroscientists want to know how animal brains work, engineers and computer scientists want to build intelligent machines and mathematicians want to understand the fundamental properties of networks as complex systems. Another (perhaps the largest) group of people are to be found in a variety of industrial and commercial areas and use neural networks to model and analyze large, poorly understood datasets that arise naturally in their workplace. It is therefore important to understand an author's perspective when reading the literature. Their common focal point is, however, neural networks and is potentially the basis for close collaboration. For example, biologists can usefully learn from computer scientists what computations are necessary to enable animals to solve particular problems, while engineers can make use of the solutions nature has devised so that they may be applied in an act of "reverse engineering".

In the next chapter we look more closely at real neurons and how they may be modelled by their artificial counterparts. This approach allows subsequent development to be viewed from both the biological and engineering-oriented viewpoints.

1.3 Summary

Artificial neural networks may be thought of as simplified models of the networks of neurons that occur naturally in the animal brain. From the biological viewpoint the essential requirement for a neural network is that it should attempt to capture what we believe are the essential information processing features of the corresponding "real" network. For an engineer, this correspondence is not so important and the network offers an alternative form of parallel computing that might be more appropriate for solving the task in hand.

The simplest artificial neuron is the threshold logic unit or TLU. Its basic operation is to perform a weighted sum of its inputs and then output a "1" if this sum exceeds a threshold, and a "0" otherwise. The TLU is supposed to model the basic "integrate-and-fire" mechanism of real neurons.

1.4 Notes

1. It is not important that the reader be familiar with these areas. It suffices to understand that neural networks can be placed in relation to other areas studied by workers in these fields.

References

Aleksander, I. 1965. Fused logic element which learns by example. *Electronics Letters* **1,** 73–4.

Aleksander, I. 1980. Whatever happened to cybernetics? Technical Report N/S/103, Department of Electrical Engineering, Brunel University.

Aleksander, I. & R.C.Albrow 1968. Microcircuit learning nets: some tests with handwritten numerals. *Electronics Letters* **4,** 406–7.

Aleksander, I. & H.Mamdani 1968. Microcircuit learning nets: improved recognition by means of pattern feedback. *Electronics Letters* **4,** 425–6.

Aleksander, I. & T.J.Stonham 1979. Guide to pattern recognition using random-access memories. *Computers and Digital Techniques* **2,** 29–40.

Aleksander, I., W.V.Thomas, P.A.Bowden 1984. WISARD: a radical step forward in image recognition. *Sensor Review* **4,** 120–24.

Amit, D.J. 1989. *Modelling brain function: the world of attractor neural networks.* Cambridge: Cambridge University Press.

Amit, D.J. & H.Gutfreund 1985. Spin-glass models of neural networks. *Physical Review A* **32,** 1007–18.

Anderson, A. & E.Rosenfeld (eds) 1988. *Neurocomputing: foundations of research.* Cambridge, MA: MIT Press.

Anderson, J.A. 1972. A simple neural network generating an interactive memory. *Mathematical Biosciences* **14,** 197–220.

Austin, J. 1987a. ADAM: A distributed associative memory for scene analysis. In *1st IEEE International Conference on Neural Networks,* vol. IV, 285–92, San Diego.

Austin, J. 1987b. *The designs and application of associative memories for scene analysis.* PhD thesis, Department of Electrical Engineering, Brunel University.

Baba, N. 1989. A new approach for finding the global minimum of error function of neural networks. *Neural Networks* **2,** 367–73.

Banquet, J.P. & S.Grossberg 1987. Probing cognitive processes through the structure of event related potentials during learning: an experimental and theoretical analysis. *Applied Optics* **26,** 4931–46.

Barto, A.G. 1985. Learning by statistical cooperation of self-interested neuron-like computing elements. *Human Neurobiology* **4,** 229–56.

Barto, A.G. 1992. Reinforcement learning and adaptive. In *Handbook of intelligent control* D.A.White & D.A.Sofge (eds), 469–91. New York: Van Nostrand Reinhold.

Barto, A.G. & P.Anandan 1985. Pattern-recognizing stochastic learning automata. *IEEE Transactions on Systems, Man and Cybernetics* **SMC-15,** 360–75.

Barto, A.G. & M.I.Jordan 1987. Gradient following without backpropagation in layered networks. In *1st IEEE International Conference on Neural Networks,* vol. II, San Diego.

Barto, A.G., R.S.Sutton, C.Anderson 1983. Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man and Cybernetics* **SMC-13,** 834–6.

Battiti, R. 1992. First- and second-order methods for learning: between steepest descent and Newton's method. *Neural Computation* **4,** 141–66.

Baum, E. & D.Haussler 1989. What size net gives valid generalisation? *Neural Computation* **1,** 151–60.

Bezdek, J.C. & N.R.Pal 1995. A note on self-organizing semantic maps. *IEEE Transactions on Neural Networks* **6,** 1029–36.

Bledsoe, W.W. & C.L.Bisson 1962. Improved memory matrices for the n-tuple pattern recognition method. *IRE Transactions on Electronic Computers* **EC-11,** 414–15.

Bledsoe, W.W. & I.Browning 1959. Pattern recognition and reading by machines. In *Proceedings of the Eastern Joint Computer Conference,* 225–32.

Bolouri, H., P.Morgan, K.Gurney 1994. Design, manufacture and evaluation of a scalable high-performance neural system. *Electronics Letters* **30,** 426.

Bonhoeffer, T. & A.Grinvald 1991. Iso-orientation domains in cat visual cortex are arranged in pin wheel-like patterns. *Nature* **353,** 429–31.

Broomhead, D.S. & D.Lowe 1988. Multivariable functional interpolation and adaptive networks. *Complex Systems* **2,** 321–55.

Bruce, V. & P.Green 1990. *Visual perception: physiology, psychology and ecology,* 2nd edn. Hove: Erlbaum.

Bullock, T.H., R.Orkand, A.Grinnell 1977. *Introduction to nervous systems.* San Francisco: Freeman. (Neural coding—Ch. 6, Sect. B.)

Burke, L.I. 1991. Clustering characterisation of adaptive resonance. *Neural Networks* **4,** 485–91.

Carpenter, G.A. & S.Grossberg 1987a. ART-2: self-organization of stable category recognition codes for analog input patterns. *Applied Optics* **26,** 4919–30.

Carpenter, G.A. & S.Grossberg 1987b. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing* **37,** 54–115.

Carpenter, G.A. & S.Grossberg 1988. The ART of adaptive pattern recognition by a self-organizing neural network. *Computer* **21,** 77–90.

Carpenter, G.A. & S.Grossberg 1990. ART 3: hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks* **3,** 129–52.

Carpenter, G.A. & S.Grossberg 1992. A self-organizing neural network for supervised learning, recognition and prediction. *IEEE Communications Magazine,* 38–49.

Carpenter, G.A., S.Grossberg, C.Mehanian 1989. Invariant recognition of cluttered scenes by a self-organizing ART architecture: CORT-X boundary segmentation. *Neural Networks* **2,** 169–81.

Carpenter, G.A., S.Grossberg, J.H.Reynolds 1991a. ARTMAP: supervised real-time learning and classification of nonstationary data by a self-organising neural network. *Neural Networks* **4,** 565–88.

Carpenter, G.A., S.Grossberg, D.B.Rosen 1991b. ART2 A: an adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks* **4,** 493– 504.

Carpenter, G.A., S.Grossberg, D.B.Rosen 1991c. Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks* **4,** 759–71.

Carpenter, G.A., S.Grossberg, D.B.Rosen 1991d. A neural network realization of fuzzy ART. Technical Report CAS/CNS-91-021, Boston University.

Caudell, T.P., S.D.G.Smith, R.Escobedo, M.Anderson 1994. NIRS: large scale ART-1 architectures for engineering design retrieval. *Neural Networks* **7,** 1339– 50.

Caudill, M. 1991. Expert networks. *Byte,* 108–16.

Cheng, B. & D.M.Titterington 1994. Neural networks: a review from a statistical perspective. *Statistical Science* **9,** 2–54.

Churchland, P.S. & T.J.Sejnowski 1992. *The computational brain.* Cambridge, MA: MIT Press (Bradford Books).

Clark, A. 1990. *Microcognition: philosophy, cognitive science and parallel distributed processing.* Cambridge, MA: MIT Press (Bradford Books).

Conners, B.W. & M.J.Gutnick 1990. Intrinsic firing patterns of diverse neocortical neurons. *Trends in Neuroscience* **13,** 99–104.

Connolly, M. & D.Van Essen 1984. The representation of the visual field in parvocellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. *Journal of Comparative Neurology* **226,** 544–64.

Cover, T.M. 1965. Geometrical and statistical properties of systems of linear inequalities with applications on pattern recognition. *IEEE Transactions on Electronic Computers* **EC-14,** 326–34.

Cvitanović, P. 1984. *Universality in chaos.* Bristol: Adam Hilger.

Davis, G.E., W.E.Lowell, G.L.Davis 1993. A neural network that predicts psychiatric length of stay. *MD Computing* **10,** 87–92.

Desimone, R. 1992. Neural circuits for visual attention in the primate brain. In *Neural networks for vision and image processing,* G.A.Carpenter & S.Grossberg (eds), 343–64. Cambridge, MA: MIT Press.

Dreyfus, H.L. 1979. *What computers can't do—the limits of artificial intelligence.* New York: Harper and Row.

Durbin, R. & G.Mitchison 1990. A dimension reduction framework for understanding cortical maps. *Nature* **343,** 644–7.

Fahlman, S.E. & C.Lebiere 1990. The cascade-correlation learning architecture. In *Advances in neural information processing systems,* D.S.Touretzky (ed.), vol. 2, 875–82. San Mateo, CA: Morgan Kaufmann.

Farmer, D., T.Toffoli, S.Wolfram 1983. Cellular automata—proceedings of an interdisciplinary workshop, Los Alamos. In *Physica,* vol. 10D (special volume). Amsterdam: North-Holland.

Fletcher, R. & C.M.Reeves 1964. Function minimisation by conjugate gradients. *Computer Journal* **7,** 149–54.

Fukushima, K. 1975. Cognitron: a self-organizing multilayered neural network. *Biological Cybernetics* **20,** 121–36.

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. 1988. A neural network for visual pattern recognition. *Computer* **21,** 65–75.

Fukushima, K. 1989. Analysis of the process of visual pattern recognition by the neocognitron. *Neural Networks* **2,** 413–20.

Funahashi, K. 1989. On the approximate realization of continuous mappings by neural networks. *Neural Networks* **2,** 183–92.

Gallant, S.I. 1988. Connectionist expert systems. *Communications of the ACM* **31,** 152–69.

Georgiopoulous, M., G.L.Heileman, J.Huang 1991. Properties of learning related pattern diversity in ART1. *Neural Networks* **4,** 751–7.

Goodhill, G.J. 1993. Topography and ocular dominance—a model exploring positive correlations. *Biological Cybernetics* **69,** 109–18.

Gorse, D. & J.G.Taylor 1989. An analysis of noisy RAM and neural nets. *Physica D* **34,** 90–114.

Graf, H.P., W.Hubbard, L.D.Jackel, P.G.N.deVegvar 1987. A CMOS associative memory chip. In *1st IEEE International Conference on Neural Networks,* vol. III, 469–77, San Diego.

Gray, R.M. 1984. Vector quantization. *IEEE ASSP Magazine* **1,** 4–29.

Grossberg, S. 1973. Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics* **52,** 217–57.

Grossberg, S. 1976a. Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. *Biological Cybernetics* **23,** 121–34.

Grossberg, S. 1976b. Adaptive pattern classification and universal recoding: II. feedback, expectation, olfaction, illusions. *Biological Cybernetics* **23,** 187–202.

Grossberg, S. 1980. How does a brain build a cognitive code? *Psychological Review* **87,** 1–51.

Grossberg, S. 1987. Competitive learning: from interactive activation to adaptive resonance. *Cognitive Science* **11,** 23–63.

Grossberg, S. 1988. Nonlinear neural networks: principles, mechanisms, and architectures. *Neural Networks* **1,** 17–61.

Gullapalli, V. 1990. A stochastic reinforcement learning algorithm for learning real-valued functions. *Neural Networks* **3,** 671–92.

Gurney, K.N. 1989. *Learning in networks of structured hypercubes.* PhD thesis, Department of Electrical Engineering, Brunel University. Available as Technical Memorandum CN/R/144.

Gurney, K.N. 1992a. Training nets of hardware realisable sigma-pi units. *Neural Networks* **5,** 289–303.

Gurney, K.N. 1992b. Training nets of stochastic units using system identification. *Neural Networks* **6,** 133–45.
Gurney, K.N. 1992c. Training recurrent nets of hardware realisable sigma-pi units. *International Journal of Neural Systems* **3,** 31–42.
Gurney, K.N. 1992d. Weighted nodes and RAM-nets: a unified approach. *Journal of Intelligent Systems* **2,** 155–86.
Gurney, K.N. 1995. Towards a theory of neural-processing complexity. In *Proceedings of IPCAT'95,* Liverpool, UK.
Gurney, K.N. & M.J.Wright 1992a. Digital nets and intelligent systems. *Journal of Intelligent Systems* **2,** 1–10. (Special issue on advances in digital neural networks.)
Gurney, K.N. & M.J.Wright 1992b. A self-organising neural network model of image velocity encoding. *Biological Cybernetics* **68,** 173–81.
Hart, P.E., R.O.Duda, M.T.Einaudi 1978. A computer-based consultation system for mineral exploration. Technical Report, SRI International.
Hartline, H.K. 1934. *Journal of Cell and Comparative Physiology* **5,** 229.
Hartline, H.K. 1940. The nerve messages in the fibers of the visual pathway. *Journal of the Optical Society of America* **30,** 239–47.
Haykin, S. 1994. *Neural networks: a comprehensive foundation.* New York: Macmillan.
Hebb, D.O. 1949. *The organization of behaviour.* New York: John Wiley.
Heims, S.J. 1982. *John von Neumann and Norbert Weiner—from mathematics to the technologies of life and death.* Cambridge, MA: Academic Press.
Heywood, M. & P.Noakes 1995. A framework for improved training of sigma-pi networks. *IEEE Transactions on Neural Networks* **6,** 893–903.
Hinton, G.E. 1987. Connectionist learning principles. Technical Report CMU-CS-87–115, Carnegie-Mellon University. (Reproduced in *Artificial Intelligence* **40,** 185–234, 1989.)
Hinton, G.E., T.J.Sejnowski, D.Ackley 1984. Boltzmann machines: constraint satisfaction networks that learn. Technical Report CMU-CS-84–119, Carnegie-Mellon University.
Hodges, A. 1985. *Alan Turing—the enigma of intelligence.* London: Counterpoint (Unwin).
Hodgkin, A.L. & A.L.Huxley 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology (London)* **117,** 500–44.
Hopfield, J.J. 1982. Neural networks and physical systems with emergent collective computational properties. *Proceedings of the National Academy of Sciences of the USA* **79,** 2554–88.
Hopfield, J.J. 1984. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences of the USA* **81,** 3088–92.
Hopfield, J.J. & D.W.Tank 1985. Neural computation of decisions in optimization problems. *Biological Cybernetics* **52,** 141–52.
Hornik, K., M.Stinchcombe, H.White 1989. Multilayer feedforward networks are universal approximators. *Neural Networks* **2,** 359–66.
Huang, Z. & A.Kuh 1992. A combined self-organizing feature map and multilayer perceptron for isolated word recognition. *IEEE Transactions on Signal Processing* **40,** 2651–7.
Hubel, D. & T.N.Wiesel 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology* **160,** 106–54.
Hubel, D. & T.N.Wiesel 1974. Sequence regularity and geometry of orientation columns in the monkey striate cortex. *Journal of Comparative Neurology* **158,** 267–94.
Hui, T., P.Morgan, K.Gurney, H.Bolouri 1991. A cascadable 2048-neuron VLSI artificial neural network with on-board learning. In *Artificial neural networks 2,* I.Aleksander & J.Taylor (eds), 647–51. Amsterdam: Elsevier.
Hung, C.A. & S.F.Lin 1995. Adaptive Hamming net: a fast-learning ART1 model without searching. *Neural Networks* **8,** 605–18.
Jacobs, R.A. 1988. Increased rates of convergence through learning rate adaptation. *Neural Networks* **1,** 295–307.
Jagota, A. 1995. Approximating maximum clique with a Hopfield net. *IEEE Transactions on Neural Networks* **6,** 724–35.
Jones, A.J. 1993. Genetic algorithms and their applications to the design of neural networks. *Neural Computing and Applications* **1,** 32–45.
Jones, D.S. 1979. *Elementary information theory.* Oxford: Clarendon Press.
Kan, W.K. & I.Aleksander 1987. A probabilistic logic neuron network for associative learning. In *1st IEEE International Conference on Neural Networks,* vol. II, 541–8, San Diego.
Kandel, E.F., J.H.Schwartz, T.J.Jessell 1991. *Principles of neural science,* 3rd edn. Amsterdam: Elsevier.
Karhunan, J. & J.Joutsensalo 1995. Generalization of principal component analysis, optimization problems and neural networks. *Neural Networks* **8,** 549–62.
Kauffman, S.A. 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology* **22,** 437–67.
Kendall, M. 1975. *Multivariate analysis.* London: Charles Griffin.
Kim, E.J. & Y.Lee 1991. Handwritten Hangul recognition using a modified neocognitron. *Neural Networks* **4,** 743–50.
Kirkpatrick, S., C.D.Gelatt, M.P.Vechi 1983. Optimization by simulated annealing. *Science* **230,** 671–9.
Koch, C. & I.Segev (eds) 1989. *Methods in neuronal modeling.* Cambridge, MA: MIT Press (Bradford Books).
Koch, C., T.Poggio, V.Torre 1982. Retinal ganglion cells: a functional interpretation of dendritic morphology. *Philosophical Transactions of the Royal Society B* **298,** 227–64.
Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics* **43,** 59–69.
Kohonen, T. 1984. *Self-organization and associative memory.* Berlin: Springer-Verlag.
Kohonen, T. 1988a. Learning vector quantization. *Neural Networks* **1,** suppl. 1, 303.
Kohonen, T. 1988b. The 'neural' phonetic typewriter. *Computer* **21,** 11–22.
Kohonen, T. 1990. The self-organizing map. *Proceedings of the IEEE* **78,** 1464–80.

Kohonen, T., K.Mäkisara, T.Saramäki 1984. Phontopic maps—insightful representation of phonological features for speech recognition. In *Proceedings of Seventh International Conference on Pattern Recognition,* 182–5, Montreal, Canada.

Kosko, B. 1992. *Neural networks and fuzzy systems.* Englewood Cliffs, NJ: Prentice Hall.

Kuffler, S.W., J.G.Nicholls, A.R.Martin 1984. *From neuron to brain: a cellular approach to the function of the nervous system,* 2nd edn. Sunderland, MA: Sinauer Associates.

Lee, Y., S.Oh, M.Kim 1991. The effect of initial weights on premature saturation in back-propagation learning. In *International Joint Conference on Neural Nets,* vol. 1, Seattle.

Linsker,R. 1986. From basic network principles to neural architecture. *Proceedings of the National Academy of Sciences of the USA* **83,** 7508–12, 8390–4, 8779–83. (Series of three articles.)

Linsker, R. 1988. Self-organization in a perceptual network. *Computer* **21,** 105–17.

Lippmann, R.P. 1987. An introduction to computing with neural nets. *IEEE ASSP Magazine,* 4–22.

Little, W.A. 1974. The existence of persistent states in the brain. *Mathematical Biosciences* **19,** 101–20.

Makhoul, J., A.El-Jaroudi, R.Schwartz 1989. Formation of disconnected decision regions with a single hidden layer. In *International Joint Conference on Neural Nets,* vol. 1, 455–60, Seattle.

Mandelbrot, B.B. 1977. *The fractal geometry of nature.* New York: Freeman.

Marr, D. 1982. *Vision.* New York: Freeman.

Martland, D. 1987. Auto-associative pattern storage using synchronous Boolean nets. In *1st IEEE International Conference on Neural Networks,* vol. III, San Diego.

Maxwell, T., C.L.Giles, Y.C.Lee 1987. Generalization in neural networks: the contiguity problem. In *1st IEEE International Conference on Neural Networks,* vol. II, 41–5, San Diego.

McCulloch, W. & W.Pitts 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* **7,** 115–33.

McDermott, J. 1982. R1: a rule-based configurer of computer systems. *Artificial Intelligence* **19,** 39–88.

McEliece, R.J., E.C.Posner, E.R.Rodemich, S.S.Venkatesh 1987. The capacity of the Hopfield associative memory. *IEEE Transactions on Information Theory* **IT-33,** 461–82.

Milligan, D.K. 1988. Annealing in RAM-based learning networks. Technical Report CN/R/142, Department of Electrical Engineering, Brunel University.

Minsky, M. & S.Papert 1969. *Perceptrons.* Cambridge, MA: MIT Press.

Murre, J.M.J. 1995. Neurosimulators. In *The handbook of brain theory and neural networks,* M.A.Arbib (ed.), 634–9. Cambridge, MA: MIT Press.

Myers, C.E. & I.Aleksander 1988. Learning algorithms for probabilistic neural nets. In *1st INNS Annual Meeting,* Boston.

Nabhan, T.M. & A.Y.Zomaya 1994. Toward generating neural network structures for function approximation. *Neural Networks* **7,** 89–99.

Narendra, K.S. & M.A.L.Thathacar 1974. Learning automata—a survey. *IEEE Transactions on Systems, Man and Cybernetics* **SMC-4,** 323–34.

Newell, A. & H.A.Simon 1976. Computer science as empirical enquiry: symbols and search. *Communications of the ACM* **19,** 113–26.

Nolfi, S. & D.Parisi 1995. 'Genotypes' for neural networks. In *The handbook of brain theory and neural networks,* M.A.Arbib (ed.), 431–4. Cambridge, MA: MIT Press.

Obermayer, K., H.Ritter, K.Schulten 1990. A principle for the formation of the spatial structure of cortical feature maps. *Proceedings of the National Academy of Sciences of the USA* **87,** 8345–9.

Oja, E. 1982. A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology* **15,** 267–73.

Parker, D.B. 1982. Learning-logic. Technical Report 581–64, Office of Technology Licensing, Stanford University.

Plumbley, M.D. 1993. Efficient information transfer and anti-Hebbian neural networks. *Neural Networks* **6,** 823–33.

Poggio, T. & F.Girosi 1990a. Networks for approximation and learning. *Proceedings of the IEEE* **78,** 1481–97.

Poggio, T. & F.Girosi 1990b. Regularization algorithms for learning that are equivalent to multilayer networks. *Science* **247,** 978–82.

Powell, M.J.D. 1987. Radial basis functions for multivariable interpolation: a review. In *Algorithms for approximation,* J.C.Mason & M.G.Cox (eds). Oxford: Clarendon Press.

Raj, R. 1988. Foundations and grand challenges of artificial intelligence. *AI Magazine* **9,** 9–21.

Rail, W. 1957. Membrane time constant of motoneurons. *Science* **126,** 454.

Rall, W. 1959. Branching dendritic trees and motoneuron membrane resistivity. *Experimental Neurology* **2,** 503–32.

Reed, R. 1993. Pruning algorithms—a survey. *IEEE Transactions on Neural Networks* **4,** 740–7.

Refenes, A.N., A.Zapranis, G.Francis 1994. Stock performance modelling using neural networks: a comparative study with regression models. *Neural Networks* **7,** 375–88.

Rich, E. & K.Knight 1991. *Artificial intelligence.* New York: McGraw-Hill.

Ritter, H. & T.Kohonen 1989. Self-organizing semantic maps. *Biological Cybernetics* **61,** 241–54.

Rosenblatt, F. 1962. *Principles of neurodynamics.* New York: Spartan Books.

Rosin, P.L. & F.Fierens 1995. Improving neural net generalisation. In *International Geoscience and Remote Sensing Symposium,* Florence.

Rumelhart, D. & D.Zipser 1985. Feature discovery by competitive learning. *Cognitive Science* **9,** 75–112.

Rumelhart, D.E., G.E.Hinton, R.J.Williams 1986a. Learning representations by back-propagating errors. *Nature* **323,** 533–6.

Rumelhart, D.E., J.L.McCllelland, The PDP Research Group 1986b. *Parallel distributed processing,* vol. 1, ch. 9. Cambridge, MA: MIT Press (Bradford Books).

Rumelhart, D.E., J.L.McCllelland, The PDP Research Group 1986c. *Parallel distributed processing,* vol. 1, ch. 5. Cambridge, MA: MIT Press (Bradford Books).

Rumelhart, D.E., J.L.McCllelland, The PDP Research Group 1986d. *Parallel distributed processing,* vol. 1, ch. 2. Cambridge, MA: MIT Press (Bradford Books).

Sánchez-Sinencio, E. & R.W.Newcomb 1992a. Guest editorial for special issue on neural network hardware . *IEEE Transactions on Neural Networks* **3,** 345–518.

Sánchez-Sinencio, E. & R.W.Newcomb 1992b. Guest editorial for special issue on neural network hardware. *IEEE Transactions on Neural Networks* **4,** 385–541.

Sanger, T. 1989. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks* **2,** 459–73.

Shawe-Taylor, J. 1992. Threshold network learning in the presence of equivalence. In *Advances in neural information processing systems,* J.E.Moody, S.J.Hanson, R.P.Lippmann (eds), vol. 4, 879–86. San Mateo, CA: Morgan Kaufmann.

Shepherd, G.M. 1978. Microcircuits in the nervous system. *Scientific American* **238,** 92–103.

Smolensky, P. 1988. On the proper treatment of connectionism. *Behavioural and Brain Sciences* **11,** 1–74.

Steiger, U. 1967. Über den Feinbau des Neuopils im Corpus Pedunculatum der Waldemeise. *Zeitschrift fur Zellfbrschung* **81,** 511–36. (As reproduced in Bullock, T.H. et al. 1977. *Introduction to nervous systems.* San Francisco: Freeman .)

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society* **B36,** 111–33.

Sun, R. & L.Bookman 1993. How do symbols and networks fit together? *AI Magazine,* 20–23. (A report on the workshop on integrating neural and symbolic processes sponsored by the American Association for Artificial Intelligence (AAAI).)

Sun, R. & L.A.Bookman (eds) 1994. *Computational architectures integrating neural and symbolic processes.* The Kluwer International Series in Engineering and Computer Science 292. Norwell, MA: Kluwer Academic.

Tanenbaum, A.S. 1990. *Structured computer organization.* Englewood Cliffs, NJ: Prentice Hall.

Thompson, R.F. 1993. *The brain: a neuroscience primer.* New York: Freeman.

Thorndike, E.L. 1911. *Animal learning.* New York: Macmillan. (For a modern discussion, see Hergenhahn, B.R. 1988. *An introduction to theories of learning.* Englewood Cliffs, NJ: Prentice Hall.)

Thornton, C.J. 1992. *Techniques in computational learning.* London: Chapman and Hall Computing.

Turing, A.M. 1937. On computable numbers with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* **42,** 230–65.

von der Malsburg, C. 1973. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik* **14,** 85–100.

von Neumann, J. 1987. First draft of a report on the EDVAC. In *Papers of John von Neumann on computing and computer theory,* vol. 12 in the Charles Babbage Institute Reprint Series for the History of Computing, W.Aspray & A.Burks (eds). Cambridge, MA: MIT Press.

Walker, C.C. & W.R.Ashby 1966. On temporal characteristics of behaviour in certain complex systems. *Kybernetik* **3,** 100–8.

Werbos, P. 1974. *Beyond regression: new tools for prediction and analysis in the behavioural sciences.* PhD thesis, Harvard University.

Weymare, N. & J.P.Martens 1994. On the initialization and optimization of multilayer perceptrons. *IEEE Transactions on Neural Networks* **5,** 738–51.

Widrow, B. & M.E.Hoff, Jr 1960. Adaptive switching circuits. In *1960 IRE WESCON convention record,* 96–104. New York: IRE. (Reprinted in Anderson, A. & E.Rosenfeld (eds) 1988. *Neurocomputing—foundations of research.* Cambridge, MA: MIT Press .)

Widrow, B. & S.D.Stearns 1985. *Adaptive signal processing.* Englewood Cliffs, NJ: Prentice-Hall.

Widrow, B., R.G.Winter, R.A.Baxter 1987. Learning phenomena in layered neural networks. In *1st IEEE International Conference on Neural Networks,* vol. II, 411–29, San Diego.

Wieland, A. & R.Leighton 1987. Geometric analysis of neural network capabilities. In *1st IEEE International Conference on Neural Networks,* vol. III, San Diego.

Williams, R.J. 1987. Reinforcement-learning connectionist systems. Technical Report NU-CCS-87–3, Northeastern University, Boston.

Willshaw, D.J., O.P.Buneman, H.C.Longuet-Higgins 1969. Non-holographic associative memory. *Nature* **222,** 960–62.

Willshaw, D.J. & C.von der Malsburg 1976. How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society B* **194,** 431–45.

Winston, P.H. 1984. *Artificial intelligence.* Reading, MA: Addison-Wesley.

Yu, X.H., G.A.Chen, S.X.Cheng 1995. Dynamic learning rate optimization of the backpropagation algorithm. *IEEE Transactions on Neural Networks* **6,** 669–77.

Zadeh, L. 1965. Fuzzy sets. *Information and Control* **8,** 338–53.