Book Reviews

An Introduction to Neural Networks

James A. Anderson
Cambridge, MA: The MIT Press, 1995
Hardbound, 650 pages, $55.00. ISBN 0-262-01144-1

Reviewed by Joel Davis

The past two decades have seen explosive growth in knowledge about animal nervous systems. At the same time, interest in computers and computing systems has evolved beyond a sequential, von Neuman approach, especially for dealing with ill-posed and "fuzzy" real-world computational problems. Finally, a new generation of cognitive scientists is more amenable to incorporating biological data into their work. Neural networks and neurocomputing represent a confluence of these three disciplines and has enjoyed tremendous recent growth, as readers of this journal are, no doubt, aware.

Anderson's introduction to this area far surpasses the three or four available competitors. Introduction to Neural Networks evolved over several years from course notes, and the integration to topics, perspectives, examples, and applications all suggest a carefully considered approach. Many introductory books in neural nets are designed for engineering students and, therefore, concentrate almost exclusively on a mathematical analysis of available networks and algorithms. Anderson skillfully interweaves a biologically relevant message about real-world processes and how a computational approach can enhance an understanding of the underlying neuroscience. The book concentrates primarily on those approaches that have proven "useful"—primarily for pattern recognition. Anderson limits his mathematical emphasis not so much on the formal analysis of networks as on the use of algorithms. However, a student without some calculus, algebra, familiarity with vectors, and some programming experience may find parts of this text daunting. Some neuroscience and cognitive science experience would also provide a useful background for any prospective student, although the chapters on neuron and synaptic physiology as they pertain to networks are quite clear.

Anderson begins the history of network modeling with Rosenblatt's Perceptron and early gradient descent algorithms such as Widrow and Hoff's ADALINE. Hopfield networks and Boltzmann machines are extremely well covered. Kohonen's adaptive map-forming algorithm is nicely contrasted with a section entitled "Biological Caveat," which contrasts the biological complexity with the "simplicity" suggested by simple adaptive algorithms. Nonlinear Autoassociative neural networks are represented by the brain-state-in-a-box (BSB) model, a system the author is closely associated with. In an approach similar to the other network examples, Anderson describes the theory behind the model and informally explains some of its mathematical properties. He ends with some "real-world" examples.

Extremely well written (humorous at times!) and well organized, anyone willing to read and understand this Introduction will have acquired a strong background in the neural network area that should permit a profitable reading of proliferating journals in this field. Furthermore, students or active researchers who may have heard the commotion and hype generated by the numerous neural network mavens can now judge for themselves whether this approach can provide a "value-added" addition if applied to their own particular area of expertise.

Reprint requests should be sent to Joel Davis, Department of the Navy, Office of Naval Research, Ballston Centre Tower One, 800 North Quincy Street, Arlington, VA 22217-5660.

The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain

Paul M. Churchland
Cambridge, MA: The MIT Press, 1995
Hardbound, 329 pages, $29.95. ISBN 0-262-03224-4

Reviewed by David P. Freidman

By the time you finish *Engine of Reason, Seat of the Soul,* its author, the philosopher of science Paul Churchland, hopes that you will "reconceive at least some of your own mental life in explicitly neurocomputational terms." That's just what I've been doing ever since I finished the book. I assume most of you are like me, and I don't spend that much time reconceiving my life in any terms. That I am doing it now in "explicitly neurocomputational" terms is all the more remarkable because if I have been doing any reconceiving at all, it was in terms of molecular biology. Although I have always preferred the systems level approach in trying to answer the question, "how does the brain work," the pressures to answer that question in molecular terms were beginning to seem all but irresistible. My slide toward regarding gene expression as the holy grail of the nervous system has been halted, at least for now, by Professor Churchland's beautifully written, compellingly argued case that I think instead in terms of neural networks and computational neuroscience. It is in these paradigms, he asserts, that we will find the answers to questions not only about our sensory, motor, and cognitive capacities, but about our moral, social, and creative abilities as well.

According to Professor Churchland, the answer to all of our questions about brain function can be found by asking those questions in the context of neural networks, parallel distributed processing, combinatorial processing, vector coding, backpropagation, and the transformations taking place in the "hidden layers" of the processing array. These concepts don't quite match the metaphorical beauty of Sherrington's enchanted loom, but they pack far more power to create the kinds of testable hypotheses that drive research.

To accomplish his goal of transforming our thinking, Professor Churchland uses the first half of this book as a crash course in computational neuroscience. He describes an array of findings that build a progressively stronger case that brain functions, from the simplest sensory discriminations to scientific creativity, moral reasoning, and social insight, can be explained as a continuum of progressively more complex sets of vector codes and neural activation spaces. Just as our remarkably rapid and accurate ability to recognize faces can be understood as sets of vector codes in the neural activation space of temperoparietal cortex, so too can the control of movement be explained as another kind of vector code in the neural activation space of the agranular motor fields of the frontal cortex. More than that, the very process of inductive reasoning can be explained in these same neurocomputational terms. Novel observations are represented by partial or degraded vectors (because we can't actually perceive and measure everything that is going on, typically knowing only the manipulation and the response of some dependent variable). The clever scientist completes these partial vectors by matching them to some prototypical set of vectors representing an already known or imagined phenomenon. Thus, the well established vector codes representing a pattern of waves splashing against a seawall might well have served as the prototype vector that led to the recognition that light too displayed wave-like behavior. No doubt anticipating our skeptical reception of his far-reaching proposal, Professor Churchland supports his assertions with a rich and lucid prose, highlighted by vivid descriptions and clarifying analogies.

Although always logical and mostly fun to confront (I mentally argued with the text throughout most of the second half of the book), the overall theory of consciousness that Churchland tries to build weakens fast after the case for explaining cognitive processes in terms of neural networks is presented. Given the dearth of existing knowledge about phenomenon like consciousness or attention, this has to be expected. But after an authoritative and well-documented beginning, the relative lack of data handicaps the theoretical structure in the succeeding chapters and makes them seem weaker simply by contrast. In the first half of this book, we bridge many gaps in or omissions of knowledge with confidence arising from otherwise strong data and rich explanations. At the outset, for example, I was not particularly concerned that all of cellular neurobiology, including synaptic transmission, second messenger systems, and, yes, even gene expression, was reduced to the single two-word phrase "synaptic weight." As the book progressed, and conceptual terrain got ever more difficult, however, my ability to follow Churchland with such forbearance waned. His description of consciousness relied more and more on sparse data spanning progressively larger gaps in our knowledge. Thus, in the first half of this book, I could keep up with the rapid progress Professor Churchland made as he guided me along the well-constructed path that began our intellectual journey. In the second half, however, my willingness to follow waned, and gave way to a more begrudging pace increasingly slowed because I could no longer hold at bay the doubts raised by the large areas of neurobiology left out of his explanations or because the data he did use were just not sufficient for the task.

But, so what? If my survey of arbitrarily selected neuroscientists is any indication, there are a lot of us out there who don't know much about neural networks. Churchland convinced me that we need to understand them, and this book is a good place to start. Those not versed in the terminology and conceptual milieu of neural networks will find a useful education packed into a small space. People already network-knowledgeable, by contrast, probably won't have much to learn from the first half of the book. The second half, however, should be appreciated by almost everyone interested in consciousness. There is a focused review of some key philosophical analyses of consciousness that will provide unique perspectives and insights for most neuroscientists. In addition, Professor Churchland pushes his paradigm to the edges of its envelope of credibility in exploring a set of human activities he believes will be bettered by the application of neural network architecture to complex computational problems.

But no matter how much you know about neural networks when you start, if you are interested in the fundamental question,"how does the brain work," you will have much to savor in this book. It is well crafted by a philosopher who obviously loves neuroscience. He gives us all a gift by reminding us of the exhilaration inherent in the awesome task of trying to understand how the brain works.

Reprint requests should be sent to David P. Freidman, Department of Physiology and Pharmacology, The Bowman Gray School of Medicine, Wake Forest University, Medical Center Boulevard, Winston-Salem, NC 27157-1010.

Dynamic Patterns: The Self-Organization of Brain and Behavior

J. A. Scott Kelso
Cambridge, MA: The MIT Press, 1995
Hardbound, 334 pages, $49.95. ISBN 0-262-11200-0

Reviewed by A. J. van Opstal

Professor Kelso's book offers a thorough view of his scientific ideas and on his experimental work as it has developed over the past two decades. He presents the intriguing hypothesis that the mechanisms underlying mental activities, perception, and sensorimotor control may all be understood on the basis of spontaneous pattern formation in complex systems.

The metaphor of pattern formation is not only applicable to the brain but, as the author already points out in the first chapter, may be used in the inanimate world as well, as soon as we are willing to treat the systems under study as complex nonlinear dynamical systems that are away from equilibrium. I assume that nobody would be uncomfortable with the notion that indeed the brain is a complex system far away from equilibrium. An interesting fact from nonlinear dynamical systems theory is that quite complicated behavior already emerges by rather moderately complex, and, in fact, often quite simple, systems. Famous examples are the quadratic logistic map and the Lorentz system, which are often put forward in the popular literature on chaotic systems. In his book, Kelso offers a theoretical framework for applying such models in order to understand brain functioning.

The author goes to great length in introducing the underlying basic concepts of such systems to the nonspecialist reader in a nontechnical way. Thus, one is made familiar with the notion of the phase plane, different types of attractors (equilibrium states of the system), intermittency (noisy bursting behavior intermixed with rather regular epochs), bifurcations (sudden qualitative changes in the system's behavior), and others in a concise and elegant way.

Throughout, the author tries to link these ideas with experimental data (mostly behavioral data from human subjects) obtained over the years and that, in his view, betray an underlying low-dimensional dynamical system capable of undergoing rapid phase transitions. The prime example, treated extensively, concerns the bistable performance of coordinated oscillating movements of the two index fingers, in which basically two "patterns" emerge: in-phase motion and counter-phase motion. Kelso and his co-workers interpret this behavior as the result of a dynamical system with two attractor states. The bistability emerges as a sudden phase transition, in which the system switches from one pattern to the other. Kelso emphasizes the important observation, predicted by his model, that such a transition is usually preceded by an increase in the fluctations of the movement phase. The simple model also allows for asymmetric stability of the two attractor states by controlling only two free parameters.

The details of the theory that back up this interpretation are nicely exposed and are made very clear, even to a reader with a limited mathematical background. Overall, I think the author has succeeded very well in bringing these difficult concepts to a broad audience, and I can certainly recommend the book for getting acquainted with these novel and interesting ideas. The illustrations are usually to the point and certainly enhance the reader's understanding of the underlying theory. Also, the author's style makes the book a pleasant read.

Despite all these positive remarks, I nevertheless also have some points of critique. In several instances the author seems to get carried away by his apparent success in explaining the experimental data. Often, this attitude even takes the form of a certain arrogance towards other approaches which, in my view, is not (and cannot be) fully justified. For example, Kelso makes it very clear from the start (in Chapter 2) that he thinks cybernetic theories of brain function are no good because ultimately such models will not be able to explain the mechanisms of brain functioning. The main reason is, of course, that these models often have some "control signal" that is used as a reference by the model; for example, in a feedback control loop. If one pursues such models towards the origin of the "control signal," this would lead to an infinite regress. Somewhere along the line, such models would inevitably break down.

This may be so (who denies it, anyway?) but does not necessarily mean, in my opinion, that these models are therefore a useless exercise to begin with. For example, cybernetic models of the oculomotor system, developed most notably by David Robinson and many other researchers in the field, have inspired the oculomotor community for over two decades. I am convinced that this effort has not been in vain, despite Kelso's objections. Many of the concepts stemming from these cybernetic models have actually been tested experimentally and into great detail.

One nice example from the oculomotor field (and