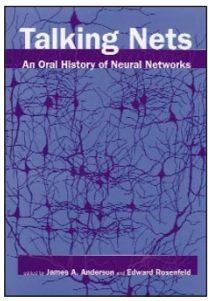
Hawking nets



Talking Nets: An Oral History of Neural Networks

Edited by James A. Anderson and Edward Rosenfeld
MIT Press, 1998. $39.95 softcover, pp 500
ISBN 0-262-01167-0

Reviewed by Peter Dayan

There is a familiar (though distinctly incomplete) view of the history of neural nets that goes like this: it was born in the MIT-influenced era of cybernetics and systems theory; it was mercilessly strangled by Minsky and Papert in the late 60's, but kept alive by the plucky few in cosmopolitan places such as Europe and Boston University; it was resurrected by the independent discovery of a piece of standard mathematics called the chain rule by a group of computational psychologists at the University of California, San Diego (UCSD); and, finally, it blossomed unstoppably as the second-best way to solve almost any signal processing or statistics problem, at least in an era of free computing power. This book consists of a set of somewhat Rogerian interviews conducted half after the facts with many of the victors under this interpretation of the field.

For those who have come across the protagonists, there is an undeniable, though voyeuristic, fascination in these brief and gossipy memoirs. The interviews were mostly very deftly conducted and edited, allowing people to reveal their characters with only a minimum of questions. We get to read a host of interesting stories about the scientific histories of some of the key players in the field and the excitement of the early discoveries (and the difficulties, such as trying to choose between lead pencils on the basis of their electrical resistance). And indeed, who could dislike a field one of whose activists admits in public that the insight behind one of the most cited results is "so trivial, it's embarrassing," or another that his (rejected) tenure case was supported by letters from fifty distinguished scientists from a variety of disciplines, a number of which said that he "deserved a Nobel prize and ... [was] a genius," or, from one who presumably lapped at the government funding trough, that "I could stand up and present things that were well known in the field and would be considered old stuff; yet to those at ARPA [a US governmental funding body], it was all new," or, from another, that he would like to disinter Helmholtz to share his enthusiasm about his eponymous machine.

Unfortunately, though, the book suffers from two key problems. First is the coverage—many of importance were simply omitted. Of course, there is limited space (although using some of it up with those who have only a fuzzy connection to the field shows questionable taste). However, the book would have been greatly enriched by the opinions of the angels and ghosts who stalk the book, such as Amari and Minsky, and also of some of the younger generation, such as the former students from the UCSD days. The second problem comes from the light hand of the interviewing and the lack of an editorial stance. Interviewees make conflicting claims as to the intellectual history and priority of the ideas. These are all printed without any comment, which is unhelpful to readers who are not up on the minutiae of credit assignment. Equally, the impressions of the researchers about such critical things as the availability of funding could usefully have been supplemented with the facts.

What certainly comes out clearly in the book is some basic tensions, on at least two levels. One comes from the wide range of parent disciplines—to what extent are neural nets about modeling neural or psychological results, and to what extent are they really engineering tools? The book is rather weighted towards the latter—indeed, some fields, such as biophysics, that have the closest relationships to the biological side of neural nets are barely mentioned. Of course, there is no need for these views to be in opposition, but the role that neural nets should play as a theoretical wing of neuroscience has tended to be compromised by the perception that the field is dominated by engineering abstraction and a reluctance to engage the scientific data. This book, particularly with its strong emphasis on commercialization in the field, will not alter such a perception.

A second level of tension highlights the youth of the field. Although there are some standard techniques, groups are still only at the early stages of exploring a very wide range of options, particularly for modeling the brain. There is not an accepted armamentarium of well understood methods, as could be found in a more mature discipline. For instance, apart from an early burst of enthusiasm, the field has been quite slow to adopt recurrently connected networks with rich dynamical behavior. This is mostly because rigorous and convenient methods of controlling and adapting such networks are not readily available. However, this leaves unanswered a very basic question about the sorts of models on which to focus. Something similar is true about how to capture the apparent stochasticity of neural activity. Of course some things have been decided—and indeed, Lettvin, Cowan and Arbib all discuss one of the critical shifts in the field away from the two-state (or possibly three-state) logical Turing machine model that seized McCulloch and Pitts, although it was ultimately rejected by the latter.

There is surprisingly little prognostication in the interviews for the future. Perhaps the closest they come is that students are repeatedly exhorted to study the more mathematical, statistical and theoretical aspects of the field. Bernie Widrow, who was one of the early pioneers of learning and adaptation in engineering, tells a great story from when he attended the first of a series of major neural network conferences in the mid-eighties. During the question period after a talk, people started talking about him as if he were an historical figure, "You know what I was? I was like a dead man. I was a man who'd died, who was sitting up on a cloud somewhere, looking down on Earth, watching what happened after he died." After introducing himself to the assemblage, people crowded round "want[ing] to touch me to see that I was alive." Widrow goes on to describe his more recent work on getting a simulated truck to learn how to back up into a parking space. His story is a metaphor for the version of the field captured by Anderson and Rosenfeld—creation in a romantic era of big problems and unknown boundaries, death, and then rebirth as literally prosaic science, no longer the stuff of oral legend.

Peter Dayan is at the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 45 Carleton Steet, Cambridge, Massachusetts 02142, USA
email: dayan@psyche.mit.edu