Copyright 1997, All rights reserved.

A Revisionist History of Connectionism

Istvan S. N. Berkeley
Bibliography

According to the standard (recent) history of connectionism (see for example the accounts offered by Hecht-Nielsen (1990: pp. 14-19) and Dreyfus and Dreyfus (1988), or Papert's (1988: pp. 3-4) somewhat whimsical description), in the early days of Classical Computational Theory of Mind (CCTM) based AI research, there was also another allegedly distinct approach, one based upon network models. The work on network models seems to fall broadly within the scope of the term 'connectionist' (see Aizawa 1992), although the term had yet to be coined at the time. These two approaches were "two daughter sciences" according to Papert (1988: p. 3). The fundamental difference between these two 'daughters', lay (according to Dreyfus and Dreyfus (1988: p. 16)) in what they took to be the paradigm of intelligence. Whereas the early connectionists took learning to be fundamental, the traditional school concentrated upon problem solving.

Although research on network models initially flourished along side research inspired by the CCTM, network research fell into a rapid decline in the late 1960's. Minsky (aided and abetted by Papert) is often credited with having personally precipitated the demise of research in network models, which marked the end of the first phase of connectionist research. Hecht-Nielson (1990: pp. 16-17) describes the situation (as it is presented in standard versions of the early history of connectionism) thus,

- The final episode of this era was a campaign led by Marvin Minsky and Seymour Papert to discredit neural network research and divert neural network research funding to the field of "artificial intelligence"....The campaign was waged by means of personal persuasion by Minsky and Papert and their allies, as well as by limited circulation of an unpublished technical manuscript (which was later de-venomized and, after further refinement and expansion, published in 1969 by Minsky and Papert as the book Perceptrons).1

In Perceptrons, Minsky and Papert (1969) argued that there were a number of fundamental problems with the network research program. For example they argued that there were certain tasks, such as the calculation of topological function of connectedness and the calculation of parity, which Rosenblatt's perceptrons2 could not solve. The inability to calculate parity proved to be particularly significant, as this showed that a perceptron could not learn to evaluate the logical function of exclusive-or (XOR). The results of Minsky and Papert's (1969: p. 231-232) analysis lead them to the conclusion that, despite the fact that perceptrons were "interesting" to study, ultimately perceptrons and their possible extensions were a "sterile" direction of research.

The publication of Perceptrons was not the only factor in the decline of network research in the late Sixties and early Seventies, though. A number of apparently significant research successes from the non-network approach, also proved to be influential. Systems such as Bobrow's (1969) STUDENT, Evan's (1969) Analogy program and Quillian's (1969) semantic memory program called the Teachable Language Comprehender, were demonstrated. These systems, which had properties like those associated with the CCTM, did not appear to suffer from the limitations that afflicted network models.3 Indeed, these systems seemed to show considerable promise with respect to emulating aspects of human cognition. Bobrow`s STUDENT program, for example, was designed to solve algebra word problems. In doing this, the program would accept input in (a restricted sub-set of) English. This property of the system lead Minsky (1966: p. 257) to claim that "STUDENT...understands English". Although this is now seen to be highly misleading (see, for

example Dreyfus' 1993: pp. 130-145 critiques of all the systems mentioned above), at the time it was a fairly impressive claim which did broadly seem to be supported by Bobrow's program. Network research, by comparison, had nothing as impressive to offer. Given Minsky and Papert's unfavorable conclusions and the apparent fruitfulness of non-network based approaches, it is not surprising that research into network systems went into decline.

During the 1970s, there was very little work done on connectionist style systems. Almost all the research done in AI concentrated upon the other approach. This is not to say that there was no network research done during this period. A few individuals, most notably Anderson (1972), Kohonen (1972) and Grossberg (1976), did continue to investigate connectionist systems, however network researchers were very much the exception rather than the rule. After a ten year hiatus though, connectionism reappeared on the scene as a significant force. One reason for this resurrection was that a number of technical developments were made which seemed to indicate that Minsky and Papert had been premature to write off such systems .

Minsky and Papert only considered Rosenblatt's perceptrons in their book of the same name. One of the significant limitations to the network technology of the time was that learning rules had only been developed for networks which consisted of two layers of processing units (i.e. input and output layers), with one set of connections between the two layers. However, Minsky and Papert (1969: p. 232) had conjectured (based on what they termed an "intuitive" judgment) that extensions of the perceptron architecture, for example based upon additional layers of units and connections, would be subject to limitations similar to those suffered by one-layer perceptrons. By the early 1980s more powerful learning rules had been developed which enabled multiple-layered networks to be trained. The results that such multiple-layered networks yielded indicated that Minsky and Papert's 'intuitive judgment' was too hasty (see Rumelhart and McClelland 1987: pp. 110-113).4

Another important factor in the renaissance of network models, according to the standard view, was a growing dissatisfaction with the traditional approach. Arguably the most important event in this renaissance was the publication of the two volume work Parallel Distributed Processing by Rumelhart, McClelland et al. (1987).5 Dreyfus and Dreyfus (1988: pp. 34-35) describe the situation thus,

- Frustrated AI researchers, tired of clinging to a research program that Jerry Lettvin characterized in the early 1980s as "the only straw afloat," flocked to the new paradigm [*sic*]. Rumelhart and McClelland's book...sold six thousand copies the day it went onto the market, and thirty thousand are now in print.

Smolensky (1988) describes how "...recent meetings [i.e. those circa 1988] of the Cognitive Science Society have begun to look like connectionist pep rallies.". Hecht-Nielsen explicitly (1990: p. 19) describes those who came 'flocking' to the new connectionism as 'converts'. The religious analogy is not insignificant here. Just as it is often the case that religious converts seek to vilify other belief systems, so the converts to connectionism often attempted to emphasize what they believed to be the fundamental differences between the connectionist and the CCTM based approach. Of course, such an environment is highly conducive to the development of myths.

So, the history of connectionism as commonly characterized, is a history which, apart from the early years, has been marked by a struggle with the approach which had roots in the assumptions underlying the CCTM. Many recent descriptions of the relationship between the approaches dwell almost exclusively upon the putative differences between them. For example, Schneider (1987), Churchland (1989), Smolensky (1991), Sterelny (1990), Cummins (1991), Tienson (1991), Bechtel and Abrahamsen (1991), Fodor and Pylyshyn (1988) and Hecht-Nielsen (1991) all portray the two approaches as being in direct competition with one another. Given the standardly told story of the history of connectionism, such an antagonistic relationship between the two approaches is far from

surprising. The standard version of this history also suggests that certain episodes (such as the publication and circulation of Perceptrons) were marked by a certain guile and personal crusading on the part of the anti-connectionist camp. Connectionism is usually portrayed as a field of research which was unfairly retarded early on, but which, due to the publication of The PDP Volumes and the empirical inadequacies of the alternative, has only comparatively recently begun to bloom. This kind of perspective fits well with the view that connectionism provides the basis of some kind of substantial alternative to the assumptions underlying the CCTM. Unfortunately, this version of history is highly selective, partial and in certain respects, down right misleading.

As a matter of historical fact, in the early days of AI research, a number of high profile researchers in the field worked with *both* approaches. Even Papert (1988: p. 10) for example, did work on network models. Another example is von Neumann, who worked with McCulloch-Pitts nets and showed that such nets could be made reliable and (moderately) resistant to damage by introducing redundancy (i.e. having several units do the job of one). In fact, von Neumann published quite extensively on the topic of networks (see von Neumann 1951, 1956 and 1966), although his name is most often associated with classical systems.

There were a number of significant results which came to light in the 1940's and 1950's, with respect to network models. Arguably the most important of these was McCulloch and Pitt's (1943) demonstration that networks of simple interconnected binary units (which they called 'formal neurons'), when supplemented by indefinitely large memory stores, were computationally equivalent to a Universal Turing Machine. Later, Rosenblatt (1958) developed an improved version of the units employed by McCulloch and Pitts. Both McCulloch and Pitt's formal neurons and Rosenblatt's units had threshold activation functions (by contrast, most modern connectionist units have continuous activations).6 The innovation which Rosenblatt made was to develop modifiable continuously valued connections (i.e. weights) between the units. This enabled networks of these units to be effectively trained. In particular, Rosenblatt's training procedure was supervised and such that the system learned only when it made a 'mistake' with respect to the desired output for a particular input pattern. Rosenblatt called networks of his units 'Perceptrons'.

The significance of Rosenblatt's innovation became clear when he (1962) demonstrated the Perceptron Convergence Theorem. This theorem holds that if there is a set of weighted connections of a perceptron, such that the perceptron gives the desired responses for a set of stimulus patterns, then after a finite number of presentations of the stimulus-response pairs and applications of the training procedure, the perceptron will converge upon that set of weights which would enable it to respond correctly to each stimulus in the set.7

Marvin Minsky, so often portrayed as a villain in the standard version of the history of connectionism, has also made significant contributions to network research. In 1951 Minsky, in conjunction with Dean Edmonds, constructed a machine known as the SNARC (Rumelhart and Zipster (1987: pp. 152-154)). The SNARC was the first 'learning' machine and was constructed along what would now be thought of as connectionist principles, according to Hecht-Nielson (1990: p. 15). Indeed, his work with the SNARC formed the basis of Minsky's Ph.D. dissertation. Minsky (1954) even included the phrase 'neural nets' in the title of his dissertation. According to Minsky (personal communication, 1994) it wasn't until "...around 1955, largely at the suggestion of my friend Ray Solomonoff....[that] I moved toward the direction of heuristic serial problem solving.". That is to say, Minsky's interest in network based system in fact predates his interest in CCTM based systems.

It is also the case that in the early phase of connectionist research, there was relatively little antagonism between the two approaches. The difference was rather one of attitude. Minsky (personal communication, 1994) characterizes the situation as follows,

- ...Nilsson [a network researcher from Stanford] was a good mathematician, as were we, so this attitudinal split had no important effect on what both sets of pioneers actually did; both groups

did in fact try to understand why each method worked on some problems but not on others.

These facts are perhaps somewhat surprising, given the malevolent role ascribed to Minsky in the standard histories of connectionism. Perhaps, it might be conjectured, the adversarial relationship between the approaches derives from Minsky and Papert's critique of networks in Perceptrons. If this is the case for some though, this adversarial perspective does not seem to be shared by Minsky himself. Even long after the publication of Perceptrons, Minsky continued to do theoretical work upon network models. In 1972 for example, Minsky (1972: p. 55) published a proof that showed that "Every finite state machine is equivalent to, and can be 'simulated' by, some neural net". Indeed, Minsky does not endorse the adversarial view of the relation between the approaches even today. Consider the following remark by Minsky (1990),

- Why is there so much excitement about Neural Networks today, and how is this related to research on Artificial Intelligence? Much has been said, in the popular press as though these were conflicting activities. This seems exceedingly strange to me, because both are parts of the same enterprise.

These facts serve to show that the supposed distinction between the two approaches, at least in the early days of network research, were not as sharp as some commentators would have us believe (C.f. Dreyfus and Dreyfus (1988)). Furthermore, there seem to be grounds for wondering just who is responsible for the putative conflict between the approaches. Although he is frequently 'demonised' in the connectionist literature, it does not seem to be Minsky!

The responsibility for the antagonistic relation between the approaches, and the consequently partial standard history, does not straightforwardly lie with any one individual or group. It is rather the consequence of a number of factors. It is certainly the case that the authors of the PDP Volumes must take some of the responsibility. For example, McClelland, Rumelhart and Hinton (1987: p. 11) remark that

- PDP models...hold out the hope of offering computationally sufficient and psychologically accurate mechanistic accounts of the phenomena of human cognition which have eluded successful explication in conventional computational formalisms...

Such remarks are fairly clearly antagonistic to advocates of the more traditional approach. There are many other similar examples which can be found in the PDP Volumes.

It is also the case that the authors of the PDP Volumes make a number of claims about the relationship between their systems and the ones discussed by Minsky and Papert in Perceptrons which are not entirely accurate. Examples of misleading claims can be found in Rumelhart, Hinton and McClelland (1986: p. 65), Rumelhart and McClelland (1986: p. 113) and Rumelhart, Hinton and Williams (1986: p. 361), for example. Minsky and Papert's responses to these specific claims are in the epilogue of the third edition of Perceptrons (1988). Of course, the authors of the PDP Volumes were not alone in misunderstanding Minsky and Papert's work. Minsky (personal communication, 1994) describes the situation thus,

- It would seem that Perceptrons has much the same role as The Necronomicon -- that is, often cited but never read.

It is by no means the case though that the responsibility for the adversarial relationship between connectionism and approaches which share assumptions with the CCTM belongs just to the authors

of the PDP Volumes. In fact, Rumelhart (personal communication, 1994) still considers his work as part of the more general enterprise of AI. He also believes that the 'AI is dead' talk which arose just after the publication of the PDP Volumes, was mistaken. Undoubtedly, the emergence of 'new' connectionism was accompanied by a certain amount of jumping on the proverbial connectionist bandwagon. It is almost certainly the case that a number of the new 'converts' to connectionism made claims which were far too strong and thereby engendered the wrath of some of the advocates of the other approach. This too is likely to have encouraged an antagonistic relation between the two approaches. It is also certainly the case that some of the antagonism between the approaches can be traced backed to Fodor and Pylyshyn's (1988) paper.

Although it would be possible to pursue this theme in much greater detail, I hope that the above is sufficient to make it clear that this putative antagonism between CCTM and connectionist approaches to studying the mind is, for the most part, a comparatively recent phenomenon. It is interesting and (I believe) significant to note that some of the major figures in the fields (e.g. Rumelhart and Minsky) do not subscribe to this view of the relationship.

Notes

1) Some of the hostility described in this account is confirmed by Papert (1988: pp.4-5).

2) Perceptron based systems were, arguably, the flag-ship variety of network systems at the time.

3) It is worth noting that all the systems mentioned here were developed by Minsky's own graduate students, according to Dreyfus (1993: p. 149). For a more detailed overview of each of these programs and the way they were evaluated, see Dreyfus (1993: pp. 130-145).

4) For a more detailed account of the work which underwrote the rebirth of connectionism, as well as a more detailed account of network research during the 1970s and early 1980s, see McClelland, Rumelhart and Hinton (1987: pp. 41-44).

5) It is now standard practice to refer to this work by the title The PDP Volumes.

6) See the discussion of activation functions in Dr. Ish's Introduction to Connectionism, for further details.

7) For a more detailed account of the history of network models, see Cowan and Sharp (1988).