8
Introduction

(1958)
F. Rosenblatt

The perceptron: a probabilistic model for information storage and organization in the brain

Psychological Review 65:386–408

The perceptron created a sensation when it was first described. It was the first precisely specified, computationally oriented neural network, and it made a major impact on a number of areas simultaneously. Rosenblatt was originally a psychologist, and the things that the perceptron was computing were things that a psychologist would think were important. The fact that it was a learning machine that was potentially capable of complex adaptive behavior made it irresistibly attractive to engineers. There were immense practical benefits to be gained from a real learning machine. The device was mathematically complex as well. It and variants of it were challenging to analyze, and it led to some very important general insights into the powers and limitations of learning machines. Much of the later work on perceptrons and successors was done by engineers and physicists, a situation still true today in research on neural networks.

It must be acknowledged that this important paper is hard to read. Perceptrons are described in a number of variations, each labeled with its own arbitrary name. The analysis does not proceed easily. There are so many options, variables, and learning rules introduced that it becomes confusing. The figures are often difficult to follow.

However, if we compare the perceptron with its successors, we see that the machinery Rosenblatt proposed is still used today, in one form or another. Rosenblatt used as his basic architecture one layer of model neurons projecting to another layer of model neurons by way of parallel bundles of connections. He started with a sensory surface—a "retina"—projecting to higher areas. The retina, the first layer, connected to a second layer, an "association area" with random, but localized, connectivity in the simplest perceptron. That is, a number of cells in a region of the retina projected to a single A-unit (Association Unit) in the higher layer. The association unit layer was reciprocally connected to a third layer of R-units (Response Units), that is, the A-unit connected to the R-unit and vice versa. The activation of the appropriate R-unit for a given input pattern or class of input patterns was the goal of the operation of the perceptron. It was deemed undesirable that more than one R-unit be on at a time. To prevent this, a set of reciprocal inhibitory connections was used, so that an R-unit inhibited all the association units that did *not* connect to it. Therefore, when an R-unit was activated, it indirectly suppressed competitors. A system with an extreme form of this behavior is today sometimes called a "Winner-Take-All" system and appears in many current network models. (See Feldman and Ballard, paper 29). The simple perceptron, as described here, was a three-layer device (see figure 2A), but Rosenblatt also considered systems with two layers of association units, a four-layer system (see figure 1).

Probably even more important than the anatomy was what Rosenblatt chose to compute. It was clear from the introductory section that he felt that what earlier theorists of the nervous system had done was fundamentally flawed. He was particularly harsh on theorists who made the brain compute logic functions, which he said amounted to "logical contrivances" (p. 387). He reserved praise for those who were aware that the noise and randomness present in the nervous system were not merely inconveniences because of poor design and construction, but were essential to the kinds of computations brains performed. The language of "symbolic logic and Boolean analysis" was not suitable for the brain.

Rosenblatt felt the appropriate way to view brain operation was as a learning associator. The goal was to couple classification responses to stimuli. Since the environment was noisy and variable, and internal connections in the nervous system were not completely prespecified, it was necessary for the system to learn to make the associations by rewiring itself. The serious technical problem then became not logical realizability, but separability; i.e., whether or not it is possible to separate the input stimuli that need to be discriminated and to make the same response to stimuli that need to be classified together. This meant that the structure of the stimuli had to be investigated. If there are clusters of events in the world that belong together, and if the stimuli coding these events are similar in the well defined sense that they tend to activate the same units and not be "too far" apart, then it is possible to classify them together. The system is very dependent for operation on the internal representation of the world. Perceptrons were found to be rather poor at discriminating arbitrary random patterns, but demonstrably quite good at separating items in a "differentiated environment" where "each response is associated to a distinct class of mutually correlated ... stimuli" (p. 405).

Learning in Rosenblatt's first paper is not analyzed with the depth that was found in later work (see Block, paper 11). The learning rules were largely of a simple reinforcement kind. After a brief preliminary period one of the R-units, randomly determined, would become active. The active R-unit would supress A-units not connected to it. If an A-unit was active when the stimulus was presented, its activity was increased. Then, after learning, if the stimulus appeared again, the A-units it activated would show stronger activity, would drive the R-units harder, and, consequently, the appropriate R-unit would be more likely to respond correctly.

This simplest learning rule was "self-organizing," because a stimulus and a random R-unit became coupled. Rosenblatt also mentioned systems where the responses were "forced." This meant activating the appropriate R-unit and then modifying the activities of the A-units. Rosenblatt showed that the perceptron was capable of learning random associations, but that there was a limit on the capacity, in that if too many random stimuli were associated with a response, accuracy decreased. If, on the other hand, the system was learning structured sets of stimuli, so that members of a class activated similar sets of A-units, then the capacity of the system could increase, and accuracy could potentially become asymptotically high.

It was in this kind of "concept forming" mode that perceptrons displayed their greatest theoretical weaknesses. As much subsequent work showed, it was *very* hard to arrange connections so that stimuli that belong together according to some rule would activate the same units (see Minsky and Papert, paper 13). Spatial arrangements of cells on a retina would not work for many interesting equivalence classes. And simple overlap clustering, as was soon discovered, was not adequate either. But even with severe limitations, the perceptron displayed the invaluable ability to generalize; that is, it could respond appropriately to patterns it had never seen because they were similar to patterns it had seen.

Rosenblatt spent some time discussing the ability of the perceptron to work in the presence of noise and with damaged or missing connections or units. He points out in this context that the memory is distributed, that is, resident in many connection strengths, and therefore resistant to damage.

By the standards of papers only a few years later, the analysis in this first paper is sketchy. There was no proof, or even awareness, of the famous Perceptron Convergence Theorem, which proved that perceptrons could learn many classifications. There were only a few statistical calculations indicating plausibility of learning, but overlooking the detailed limitations on learnability that proved so important later.

Some of the excitement of the discovery of the perceptron still comes through this early work. Here was a brain model that could *do* something. The potential for progress seemed very great. As Rosenblatt put it, "The question may well be raised at this point of where the perceptron's capabilities actually stop.... the system described is sufficient for pattern recognition, associative learning, and such cognitive sets as are necessary for selective attention and selective recall. The system appears to be potentially capable of temporal pattern recognition.... with proper reinforcement it will be capable of trial and error learning, and can learn to emit ordered sequences of responses ..." (p. 404).

Although these claims were ambitious, they were correct in that variants of the basic perceptron architecture were capable of performing them, though perhaps not as easily as it first appeared. It is important to realize that Rosenblatt was quite aware of some more serious computational limitations on the perceptron that he felt might prove very difficult to solve. Specifically, he mentioned "relative judgments" and "symbolic behavior." He mentioned that the perceptron acted in many respects like a brain damaged patient, being literal, inflexible, and unable to handle abstractions.

In fact, these are exactly areas where current neural networks still are inadequate. At present, our networks, even after thirty years of progress, still act "brain damaged." It is an open question as to whether severe modifications of network theory will have to be made to handle these highest cognitive functions.

In spite of the preliminary level of analysis in this early paper, it was clear that Rosenblatt had a good idea of both the strengths and the weaknesses of his approach. If the model had been allowed a less controversial infancy, we might have made more progress than we have done up to now.

As one final point, Rosenblatt uses the word "connectionist" as a descriptor of the perceptron, just as Hebb did for cell assemblies.