Fire Together, Wire Together: From Rosenblatt's Perceptron to the AI Winter and Parallel Distributed Processing

Brenna Ren

ATCS: Neural Networks (Period 4)

Dr. Nelson

September 19, 2025

1 Introduction

Since the dawn of humanity, the optimization of daily functions through tools and machines has driven key innovations from ancient civilizations to modern-day technology. More recently, scientists have been captivated by the dream of equipping machines with human-like learning and perception. However, the journey to widespread artificial intelligence (AI) that we know today has been far from linear. The rise and fall of Frank Rosenblatt's Perceptron and the subsequent "AI winter" epitomize these periods of optimism and skepticism about AI development. Through critiques of individuals like Marvin Minsky and the continued dedication of researchers, the field eventually evolved into the modern era of deep learning and Parallel Distributed Processing (PDP). This paper will explore the historical development of neural networks, tracing their origin from early biological inspirations and Rosenblatt's Perceptron to Marvin Minsky's impact in the context of the AI winter, and finally the development of Multi-layer Perceptrons (MLP) and PDP.

2 Biological Foundations

Long before computational models, researchers were studying how the brain's structure gave rise to thought and learning. A foundational moment came in 1943 when Warren McCulloch, a neurophysiologist at the University of Chicago, and Walter Pitts, a self-taught logician, proposed the binary artificial neuron as a logical model for biological neural networks [McCulloch and Pitts, 1943]. Their work suggested that neurons could be simplified into on-off units, processing information in a logical computational manner.

Building on this, Donald Hebb, a prominent psychologist, significantly influenced the early thinking about how learning might occur in the brain. His 1949 work *The Organization of Behavior* claimed that the persistent or repeated firing of one neuron to another led to metabolic changes that increased the efficiency of the connection [HEBB, 1949]. This concept became the cornerstone for "connectionist" theories, with Rosenblatt explicitly stating that Hebb's work was a source of inspiration for his own theories [Rosenblatt, 1958].

3 Rosenblatt's Perceptron

In July 1958, Frank Rosenblatt (1928-1971), a research psychologist and project engineer at the Cornell Aeronautical Laboratory, revealed his newest invention: the Perceptron [Lefkowitz, 2019]. Rosenblatt later became an associate professor of neurobiology and behavior at Cornell, and his primary motivation was to unravel the mysteries of the brain, particularly to understand the minimum physical requirement of a brain to perform its functions. He envisioned a machine capable of perceiving, recognizing, and identifying its surroundings without human training.

The perceptron was developed as a hypothetical nervous system, a "connectionist" model where information retention took the form of new or modified connections (pathways) between active centers in the system, rather than fixed, coded representations [Rosenblatt, 1958]. Rosenblatt's approach contrasted with "coded memory" theorists, who believed that recognition involved the systematic matching of incoming sensory patterns with stored images. His theory was based on several assumptions, including that the physical connections of the nervous system involved in learning and recognition were largely random at birth and possessed plasticity, or the ability to change and adapt in response to experience and learning. Furthermore, "similar" stimuli tend to form pathways to the same responding cells, and "dissimilar" ones tend to develop connections with different sets of responding cells. Most importantly, positive or negative reinforcement could facilitate or hinder the formation of these connections.

Rosenblatt's Mark I perceptron, demonstrated with an IBM 704 computer, could teach itself to distinguish between punch cards marked on the left from ones marked on the right [Lefkowitz, 2019]. This perceptron was a simple visual pattern binary classifier that operated as a single-layer neural network at its core [Rosenblatt, 1958]. Its organization can be explained as the combination of three parts:

- Sensory units (S-points): These are the input layer, recieving stimuli (*e.g.*, optical patterns for a "photo-percetron"). They respond in an "all-or-nothing" fashion.

- Association units (A-units): These units receive impulses from S-points. They act as a processing layer. Connections between S-points and A-points could be excitatory (increasing activity) or inhibitory (decreasing activity). If the combined effect of the excitatory and inhibitory impulses exceeded a certain threshold, the A-unit would "fire" on an "all-or-nothing" basis.

- Response units (R-units): These are the output layer, receiving impulses from A-units.

All connections within the perceptron were feed-forward, meaning impulses travel in one direction: from the sensory inputs to association units and then to the response units, without looping back. Learning in this system occured by adjusting the activation weights, which Rosenblatt referred to as "values." A value characterized the magnitude of the impulses delivered by an A-unit, and a higher value would indicate a more effective and potent impulse. Periods of activity tended to increase a cell's value, while inactivity may lead to decay.

Rosenblatt quantitatively investigated three types of "value dynamics" for learning:

- Alpha system: Active cells simply gained value with each impulse

- Beta system: Each set of A-units connected to a response (source-set) had a constant rate of gain, distributed based on activity.

- Gamma system: Active cells gained value at the expense of inactive cells within their source-set, maintaining a constant total value for the source-set.

Learning was evaluated by measuring the probability of the correct response for previously seen stimuli (P_r) and the probability of correct generalization for new stimuli of a given class (P_g). This learning was mathematically modeled using complex equations that predicted performance based on physical parameters like the number of excitatory/inhibitory connections, the threshold of A-units, and the number of A-units and R-units in the system.

4 Limitations of the Perceptron & Marvin Minsky's Impact

Despite Rosenblatt's initial optimism and the groundbreaking nature of his work, the Perceptron suffered from significant limitations. Rosenblatt himself acknowledged that "statistical separability alone does not provide a sufficient basis for higher order abstraction" and that "some system, more advanced in principle than the perceptron, seems to be required at this point" [Rosenblatt, 1958]. Tasks requiring the recognition of relationships between stimuli, such as "Name the object left of the square," proved "excessively difficult" for the perceptron.

The limitations of Rosenblatt's perceptron were brought into focus by Marvin Minsky (1927-2016), a pivotal figure in AI development. Minsky, who earned his Ph.D. in mathematics from Princeton University in 1954, was a co-founder of the MIT Artificial Intelligence Laboratory in the early 1960s, alongside John McCarthy [Winston, 2019]. Minsky's contributions spanned cognitive psychology, neural networks, automata theory, and symbolic AI. He was awarded the Turing Award in 1969 for his central role in creating, shaping, promoting, and advancing the field of AI.

In the 1960s, Minsky, working alongside Seymour Papert, published their book, *Perceptrons: An Introduction to Computational Geometry.* This book delivered a strong critique of Rosenblatt's work, detailing what perceptrons could, and more importantly, could not do. Their work highlighted the fundamental issue behind perceptrons with a single hidden layer: they could only classify linearly separable data. This means that they could only solve problems where a single straight line (or plane in higher dimensions) could separate the different categories of input. The Perceptron struggled with more complex, non-linear problems, such as distinguishing an XOR logical function, where the data points cannot be separated by a single straight line. Minsky and Papert demonstrated that Rosenblatt's perceptron was inherently limited in its computational capabilities, unable to solve more complex problems (like the XOR problem) deemed crucial for true intelligence. At the time, research funding and public expectations for AI were high, so their critique greatly discouraged research on neural networks in the 1970s, ushering in an era known as the "AI Winter" [Toosi et al., 2021]. Federal funding for AI declined, and many researchers shifted their focus away from neural networks toward symbolic AI approaches.

5 Early Development of MLPs and PDP

Despite Minsky's critique, the core idea of neurally-inspired computing remained alive as researchers explored networks with multiple layers, leading to MLPs. Unlike its predecessors, MLPs can distinguish data that is not linearly separable. They consist of an input layer that receives the initial data, one or more hidden layers that extract features from the data, and an output layer that produces the final result or prediction. In addition, neurons are fully connected, meaning each neuron in one layer connects with a specific activation weight to another neuron in the subsequent layer.

A crucial development for MLPs was the adoption of nonlinear activation functions. Since a linear activation function would reduce any number of layers to a simple two-layer input-output model (because a linear function of a linear function is still a linear function), MLPs require nonlinear activation functions to enable effective learning across multiple layers. Common nonlinear activation functions include sigmoid functions, such as the hyperbolic tangent and logistic function.

However, the breakthrough that truly enabled the effective training of MLPs was the backpropagation algorithm, which allowed a multi-layered network to learn by efficiently adjusting the activation weights based on the error in its output. Backpropagation was independently developed multiple times. The earliest published instance was by Finnish mathematician Seppo Linnainmaa in his 1970 master's thesis (published 1976), where he introduced the reverse mode of automatic differentiation in a purely math-based context [Linnainmaa, 1976]. Paul Werbos, an American, later developed backpropagation for his Harvard University PhD work independently in 1971, though his work was not widely published until 1982. However, it was ultimately David E. Rumelhart (along with Geoffrey E. Hinton and Ronald J. Williams) who popularized backpropagation in 1986 through their work *Learning representations by back-propagating errors*, which laid out more foundational work for the connectionist model and applied the back propagation algorithm to multi-layer networks [Rumelhart et al., 1986]. That same year, Rumelhart, alongside James McClelland, published *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* [Rumelhart and McClelland, 1986]. This work was central to the broader movement of Parallel Distributed Processing (PDP), which presented the idea that the mind is composed of many interconnected, simple processing units that simultaneously activate (rather than in a single, sequential line) to process information. Re-emerging interest in neurally-inspired mechanisms led to the reprinting of Minsky and Papert's *Perceptrons* with a new chapter addressing contemporary developments, signifying a shift in perspective where the original critique became a catalyst for deeper understanding.

6 Conclusion

The journey of neural networks, from Frank Rosenblatt's Perceptron to the more sophisticated Multi-layer Perceptron, demonstrates persistence toward scientific inquiry. Inspired by the fundamental understanding of how the human brain organizes and stores information, Rosenblatt laid the groundwork connectionist models with his Mark I perceptron. Inherent limitations of the single-layer perceptron from its inability to solve non-linearly separable problems was revealed by Marvin Minsky and Seymour Papert in their book *Perceptrons*. This work contributed to the "AI Winter" that saw a significant decline in funding and enthusiasm for neural network research. Ultimately, this period of skepticism spurred the development of more advanced architectures, such as the Multi-layer Perceptron and backpropagation algorithm popularized by researchers like Rumelhart. These advancements, coupled with Parallel Distributed Processing, allowed neural networks to overcome their early constraints, leading to the powerful deep learning systems that transform society today.

References

[HEBB, 1949] HEBB, D. O. (1949). *The organization of behavior: A neuropsychological theory.* John Wiley & Sons.

[Lefkowitz, 2019] Lefkowitz, M. (2019). Professor's perceptron paved the way for ai – 60 years too soon.

[Linnainmaa, 1976] Linnainmaa, S. (1976). Taylor expansion of the accumulated rounding error. *BIT*, 16(2):146–160.

[McCulloch and Pitts, 1943] McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133.

[Rosenblatt, 1958] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408.

[Rumelhart et al., 1986] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536.

[Rumelhart and McClelland, 1986] Rumelhart, D. E. and McClelland, J. L. (1986). Parallel distributed processing, volume 1: Explorations in the microstructure of cognition: Foundations. *Parallel distributed processing*.

[Toosi et al., 2021] Toosi, A., Bottino, A., Saboury, B., Siegel, E., and Rahmim, A. (2021). A brief history of ai: How to prevent another winter (a critical review). *PET Clinics*, 16.

[Winston, 2019] Winston, P. H. (2019). Marvin minsky.