NEURAL NETWORKS

Igor Vajda and Jiri Grim
Academy of Sciences of the Czech Republic,Czech Republic

Keywords: Neuron, cerebral cortex, artificial intelligence, perceptron, multilayer perceptron, learning, back-propagation algorithm, self-organizing maps, probabilistic approach

Contents

Summary

The method of selective coloring of neurons discovered in 1875 enabled the structure of neural tissue to be shown as a complex network of specific cells—neurons. Neurons occur in a wide variety of shapes and sizes, have a large number of inputs (synapses) at their receptive zones (dendrites), and a single output line (axon) branched into multiple synaptic endings. The microelectrode was the starting point of ideas that led to the binary threshold model of a neuron's electrical activity. In 1949 a physiological rule was proposed for synaptic plasticity as a basic adaptive principle of neural assemblies. However, despite the extensive and detailed knowledge accumulated in the last century, the learning principles of biological neural assemblies still continue to be a strong motivation for the further search for suitable neural network models. About 1960 the perceptron was proposed—the most popular artificial neural network—and the convergence of a heuristic learning rule was proven. Optimism initiated by the perceptron convergence theorem was damped by material published in 1969 demonstrating the fundamental limits of the perceptron networks. A new interest in neural networks came with new approaches and learning algorithms. One of the most influential methods, self-organizing maps, has been proposed to demonstrate the principles of topologically ordered maps in the brain. The widely used back-propagation algorithm was designed to solve the difficult optimization problem in multilayer networks. The present approaches are frequently used as tools to solve practical problems without biological analogies.

1. Introduction. Nervous Systems and Neurons

Advances of the modern neural sciences can be traced back by new research techniques.

The first powerful tool, invented about 1880, was a selective staining of nerve fibers by bichromate silver reaction. This coloring technique was ingeniously applied by researchers, who extensively analyzed the human central nervous system. It was proven that the network of nerve fibers is not a continuous tissue, but actually composed of a vast number of distinct, interconnected cellular units, the neurons.

The detailed investigation of the structure of neural cells was enabled by the invention of the electron microscope around 1940. It was found that all neurons are constructed from the same basic parts. From the bulbous cell body called soma project several rootlike extensions, the dendrites, as well as a single, long, tubular fiber called an axon. At its end, the axon branches into strands and substrands with button-like endings called synapses. The axon of a typical neuron makes a few thousand synapses. The size of the soma of a typical neuron is about 10 μm to 80 μm, while dendrites and axons have a diameter of several μm. The total length of neurons shows great variations, from 0.01 mm to 1 m, depending on their location.

The axon's purpose is the transmission of generated neural activity to other neurons via synapses. The synapses are located either directly at the soma or at the dendrites of the subsequent neurons. It means that the dendrites serve as receptors of incoming signals. At the synapse the two neurons are separated by a tiny gap, only about 0.2 μm wide. In relation to the synapse, the structures are called presynaptic and postsynaptic, such as presynaptic neuron, postsynaptic membrane, and so on.

Nervous signals are of an electrical nature. In the state of inactivity, the interior of the neuron—the protoplasm—is negatively charged against the surrounding neural liquid. This resting potential of about –70 mV is caused by a selective permeability of the cell membrane that is impenetrable for Na^+ ions. The resulting deficiency of positive ions in the protoplasm is responsible for its negative charge.

Signals arriving from synaptic connections may result in a transient weakening, or depolarization, of resting potential. When the polarization is reduced to a critical level of approximately –60 mV, the membrane suddenly loses its impermeability against Na^+ ions, which then enter into the protoplasm and neutralize the negative internal potential of the neuron. The quick discharge, called the spike or action potential, is followed by a gradual recovery of the membrane properties and by the corresponding regeneration of negative resting potential. During this recovery period of several milliseconds, the neuron remains incapable of further excitation.

The discharge, which initially occurs in the cell body, then propagates along the axon to the synapses without decay, because the depolarization of each new segment of the axon is always complete. This all-or-nothing rule, resembling the properties of binary electronic circuits, was a strong motivation for the first simple model of a neuron as a binary threshold unit—proposed in 1943. However, the intensity of a nervous signal can be coded by the output frequency of spikes in the range from 1 Hz to about 100 Hz because the interval between two electrical spikes can take any value longer than the regeneration period.

When the spike signal arrives at the synapse, special substances called neurotransmitters

(e.g., acetylcholine) are released into the synaptic cleft from tiny vesicles contained in the endplate. The transmitter molecules reach the postsynaptic membrane within about 0.5 ms and modify its conductance for certain ions (Na^+, K^+, Cl^-). The arising flows of ions change the local postsynaptic potential. In case of depolarization, the synapse is termed excitatory; in case of increased polarization, the synapse is called inhibitory, because it counteracts excitation of the neuron. According to Dale's Law, all synaptic endings of an axon are either of an excitatory or an inhibitory nature. There are also significant structural differences between those two types of synapses.

Just as each axon sends synapses to the dendrites and bodies of a number of downstream (efferent) neurons, so is each neuron connected to many upstream (afferent) neurons that transmit their signals to it. The soma of a neuron acts as a kind of summing device that adds the local depolarizing (or polarizing) effects of different synapses. The depolarizing effect of a synapse decays with a characteristic time of 5 ms to 10 ms and, therefore, the excitatory effects of successive spikes may accumulate at the same synapse. A high rate of spike repetition can therefore express high signal intensity. When the total magnitude of depolarization exceeds the threshold of about 10 mV, for example when the potential in the cell body achieves the critical level of –60 mV, the neuron fires, and generates an electrical spike.

In principle, a single synapse can cause a neuron to fire, though the depolarizing contribution of the synapses generally diminishes with the increasing distance of their location from the cell body. On the other hand a synapse placed at a thin dendritic branch is more likely to trigger a spreading depolarization wave with a resulting spike. The influence of a given synapse may depend on many different aspects in a complex way, but the following three factors seem to play a dominant role: the inherent strength of the depolarizing effect of the synapse, its location with respect to the cell body, and the repetition rate of the arriving spikes.

There is a great deal of evidence that the inherent strength of a synapse is not fixed but that it is changing in time dependence on the activity of both the presynaptic and postsynaptic neurons. According to one hypothesis, originally postulated by Donald Hebb, it is assumed that:

When an axon of cell A is near enough to excite cell B, and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A's efficiency as one of the cells firing B is increased (see *Adaptive Systems*). In other words, an active synapse that repeatedly triggers the activation of its postsynaptic neuron will grow in strength while others will gradually weaken. The mechanisms of synaptic plasticity corresponding more or less exactly to the above rule appear to play a dominant role in the complex process of learning.

It should be emphasized that real neurons and neural systems involve many complications that have been ignored in the above selective and purpose-oriented description. To illustrate our simplifications let us remark that

- the structural organization of the human brain is complex, only partly recognized, and still less understood

- at different locations in the brain there are usually different types of neurons with specific properties (e.g., receptor neuron, motor neuron, interneuron)
- there are qualitatively different types of synapses (e.g., axo-somatic, axo-dendritic, axo-axonal, dendro-dendritic) with unknown properties and unclear meanings
- the speed of propagation of the discharge signal along the axon is increased by electrically insulating myelin sheath-segments
- more than 20 chemically different neurotransmitters have been identified
- the excitatory and inhibitory effects of synapses are added to neurons in a complex way, nonlinearly in space and time, many details of which are unclear
- the mechanism of release of the neurotransmitter in the process of synaptical transmission is widely accepted, but discussion about some important details is not complete
- there are many internal structural elements of neurons of which the roles are not fully recognized
- there are mechanisms of a short-term synaptical plasticity (so-called potentiation) with unknown properties and unclear backgrounds and meanings.

In spite of the fact that the neuron is an inaccurate and rather unreliable functional unit, the global performance of the nervous system is surprisingly perfect. The brain is robust and fault tolerant. Many neurons die daily without significant consequence for the brain's performance. It is flexible. It can easily adjust to a new environment by learning. It can deal with information that is fuzzy, probabilistic, noisy, or inconsistent. It is able to perform creatively in a way not previously trained for. It is highly parallel, small, compact, and dissipates little power. The brain outperforms standard computers except in tasks based on simple arithmetic. A good example is the processing of visual information: a one-year-old baby is much better and faster at recognizing objects or faces than even the most advanced AI system running on the fastest supercomputer (see *Biological and Computational Intelligence*).

A fundamental question arises about the basic functional principles of the brain, which are robust and reliable enough to guarantee its outstanding properties (see *Adaptive Systems, Biological and Computational Intelligence, Cybernetics,* and *Artificial Intelligence*). Since the 1950s it became increasingly obvious that the functioning of the brain could not be explained without understanding the functional principles of the interconnected neural assemblies—the neural networks. Despite great effort and the many interesting results achieved in the last decades, this problem remains a strong motivation for the further research of different neural network models and learning algorithms.

-
-
-

Bibliography

Haykin S. (1993). *Neural Networks. Comprehensive Foundation*. San Mateo, CA: Morgan Kaufman. [One of the best monographs on neural networks; exhaustive and sufficiently detailed description of the present state of the art in neural networks with extensive bibliography.]

Hebb D.O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley. [Classical work known especially for the first formulation of the fundamental principles of synaptical learning and neural assemblies.]

Kohonen T. (1997). *The Self-Organizing Maps*. New York: Springer Verlag. [Monographs on the well-known and influential topological principle of learning in neural networks proposed by Kohonen.]

Müller B., Reinhardt J. and Strichland M.T. (1995). *Neural Networks. An Introduction*, Second Edition. New York: Springer. [An excellent monograph on neural networks and especially on multilayer perceptrons and related aspects.]

Rosenblatt F. (1962). *Principles of Neurodynamics*. Washington, DC: Spartan Books. [A famous book by the well-known inventor of the perceptron and author of the perceptron convergence theorem.]

Biographical Sketches

Igor Vajda was born in Martin, Czechoslovakia, in 1942. He graduated in mathematics from the Czech Technical University, Prague, in 1965. He received his Ph.D. degree in probability and statistics from Charles University, Prague, in 1968, and the D.Sci. degree in mathematical informatics from the same university in 1990. From 1965 to 1968 he served as a research assistant; from 1968 to 1990 as Director of Research at Institute of Information Theory and Automation, Czech Academy of Sciences, Prague. During 1966–2000 he was, for short periods, Visiting Scientist or Visiting Professor at the Institute of Information Transmission Theory in Moscow, at the University of Rostock, Complutense University of Madrid, University of Maryland in Baltimore, Catholic University of Leuven, Technical University of Budapest, University of Montpellier II, and M. Hernandez University of Alicante. His research interests include information measures and statistical distances, statistical inference about processes, and statistical decisions. Dr. Vajda is a Member of the editorial boards of Statistics and Decisions, Kybernetika, and Applications of Mathematics. In 1988 he was awarded the Prize of the Czechoslovak Academy of Sciences for his part in research on speech coding. The American Journal of Mathematical and Management Sciences awarded him the 1977 Jacob Wolfowitz Prize for a joint paper with L. Gyorfi and E.C. van der Meulen.

Jiri Grim was born in 1944 in Rychnov n.Kn., Czech Republic. He graduated in physical electronics (M.S.-level) from the Czech Technical University, Prague, in 1968. For the year 1966–1967 he obtained a fellowship at the Technical University of Aachen, Germany, from the German Academic Exchange Service (DAAD). He received his Ph.D. degree in computer science from the Czech Academy of Sciences, Prague, in 1981. His current research interests involve statistical decisionmaking, application of finite distribution mixtures, and the probabilistic approach to neural networks. He was awarded the F. de P. Hanika Memorial Award for the best paper in the Socio-Economic Section of the Eleventh European Meeting on Cybernetics and Systems Research, April 21–24, 1992, in Vienna, and the Springer Best Presentation Award at the International ICSC Symposium on Neural Computation, NC'2000, May 23–26, 2000, in Berlin. He is with the Institute of Information Theory and Automation of the Czech Academy of Sciences, Prague. He is a member of the Scientific Council of the Institute since 1996.