2
Introduction

(1943)
Warren S. McCulloch and Walter Pitts

A logical calculus of the ideas immanent in nervous activity

Bulletin of Mathematical Biophysics 5:115–133

It would be possible to write a whole book about the important aspects of this paper and the influence it has had. It is an attempt to understand what the nervous system might actually be *doing*, given primitive computing elements that are abstractions of the properties of neurons and their connections as they were known in 1943.

McCulloch and Pitts only had three references in the paper, all on mathematical logic: Carnap, *The Logical Syntax of Language*; Hilbert and Ackermann, *Foundations of Theoretical Logic*; and Whitehead and Russell, *Principia Mathematica.* In case it was thought that McCulloch did not know his physiology, a 1947 paper, "Modes of functional organization of the cerebral cortex," is about 3 pages long and has 101 references, all physiological. In 1947, McCulloch was talking about physiology; in the 1943 paper, McCulloch and Pitts are discussing what the physiology is computing—hence logic.

McCulloch and Pitts list five assumptions governing the operation of neurons. The assumptions describe what has become known as the 'McCulloch-Pitts' neuron, which is familiar to all computer scientists, often by that name. The McCulloch-Pitts neuron is a binary device; that is, it can be in only one of two possible states. Each neuron has a fixed threshold. The neuron can receive inputs from excitatory synapses, all of which have identical weights. It also can receive inputs from inhibitory synapses, whose action is absolute; that is, if the inhibitory synapse is active, the neuron cannot turn on. There is a time quantum for integration of synaptic inputs, based loosely on the physiologically observed synaptic delay.

The mode of operation of the McCulloch-Pitts neuron is simple. During the time quantum, the neuron responds to the activity of its synapses, which reflect the state of the presynaptic cells. If no inhibitory synapses are active, the neuron adds its synaptic inputs and checks to see if the sum meets or exceeds its threshold. If it does, the neuron then becomes active. If it does not, the neuron is inactive.

As an example, suppose we have a simple unit with two excitatory inputs, a and b, and with a threshold of one. Synapses connected to active cells contribute one unit. Suppose at the start of the ith time quantum, both a and b are inactive. Then, at the start of the $(i + 1)$th time quantum, the unit is inactive since the sum of two inactive synapses is zero. If a is active but b is inactive, the state of the unit at the start of the $(i + 1)$th time quantum is active since the inactive synapse plus the active synapse is one, which equals the threshold. Similarly, if b is active and a is inactive, or if both are active, the cell becomes active. This threshold-one unit with two excitatory inputs is performing the logical operation INCLUSIVE OR on its inputs, and becomes active only if a OR b OR BOTH a AND b is active.

If the unit was given a threshold of two, it would compute the logic function AND, since it would only become active if *a* AND *b* are active.

The McCulloch-Pitts neuron is therefore performing simple threshold logic. As McCulloch and Pitts put it in their Introduction, "The 'all-or-none' law of nervous activity is sufficient to insure that the activity of any neuron may be represented as a proposition. Physiological relations existing among nervous activities correspond, of course, to relations among the propositions; and the utility of the representation depends upon the identity of these relations with those of the logic of propositions" (p. 117).

As was obvious to McCulloch and Pitts, the network of connections between the simple propositions can give rise to very complex propositions. The central result of the paper is that any finite logical expression can be realized by McCulloch-Pitts neurons. This was an exciting result, since it showed that simple elements connected in a network could have immense computational power. Since the elements were based on neurophysiology, it suggested that the brain was potentially a powerful logic and computational device. The formal results about networks of logical elements also influenced others when they were thinking about the potential of digital computers, for example, John von Neumann (see paper 7).

Unfortunately, the formal proofs and the notation in this early paper are exceptionally difficult. A much better introduction to McCulloch-Pitts neurons, which also extends their results and places them in the perspective of later work in automata theory and theory of computation, is *Computation*, by Marvin Minsky (1967). This clearly written book develops a theory of computation directly from McCulloch-Pitts neurons.

One point that was clear to McCulloch and Pitts—so obvious that it was not specifically mentioned—was that a single neuron was simple, and that computational power came because simple neurons were embedded in an interacting nervous *system*. If we do not count Williams James, this paper describes perhaps the first true connectionist model, since it has simple computing elements, arranged partly in parallel, doing powerful computation with appropriately constructed connection strengths.

The question to ask next is whether McCulloch-Pitts neurons are correct approximations to real neurons, that is, whether they are a good brain model. Given the state of neurophysiology in 1943, when the ionic and electrical basis of neural activity was unclear, the approximations McCulloch and Pitts made were much more supportable than they are now. The dominant feature of observed cell activity was the 'all-or-none' action potential. It was not possible to make intracellular recordings, so it was hard to see that the postsynaptic potentials, due to presynaptic activation actually extended over a good many milliseconds, were graded, and that neurons were acting a lot more like voltage to frequency converters than like simple logic elements. (See the paper on *Limulus*, paper 23, for a more realistic model of a neural system.)

McCulloch and Pitts themselves were aware of the many continuous phenomena occurring in the nervous system, and even in the Introduction to the 1943 paper they comment on the potential importance of continuous changes in threshold brought about by adaptation and learning. It is worth mentioning this only because even now we find extended discussions of McCulloch-Pitts neurons in the scientific literature suggesting that they are adequate brain models and useful approximations to neurophysiology. This is not correct. Neurons, except in some special cases, are not simple computing devices realizing the propositions of formal logic. Binary neurons can be, however, useful approximations of underlying continuous processes in some special cases.

The immense theoretical influence of this paper was not among neuroscientists but among computer scientists. The history of this work is encouraging for theoreticians. It is not necessary to be correct in detail, or even in the original domain of application, to create an enduring work of great importance. It is possible to buy McCulloch-Pitts neurons at your local Radio Shack store, in the form of logic circuits.

A bibliographic note: McCulloch's papers are collected in a highly recommended book, *Embodiments of Mind* (1965).

References

W. McCulloch (1965), *Embodiments of Mind*, Cambridge, MA: MIT Press.

M. Minsky (1967), *Computation: Finite and Infinite Machines*, New York: Prentice-Hall.