

Fundamentals of Computational Neuroscience

**Philip S. Ulinski
The University of Chicago**

Contents

Part I -- Neural Computation

Chapter 1 -- Neurons as Computational Devices

Part II -- Single Neuron Computation

Chapter 2 -- Neurons as Cables

Chapter 3 -- Compartmental Models

Chapter 4 -- Electrotonic Structure

Chapter 5 -- Electrotonic Filtering

Part I -- Neural Computation

Chapter 1

Neurons as Computational Devices

INTRODUCTION

Marine invertebrates can orient to a visual stimulus that appears suddenly in their environment. Male flies use visual cues to accurately follow female flies as they make sharp, aerial turns. Newts and salmon use olfactory cues to return to their home territories. Canaries and finches can identify one member of their species from another using the auditory cues in their songs. Baby gibbons learn to jump from one branch to another for distances equal to many times their body. Humans can learn to use their sense of touch to discriminate the letters of the alphabet embossed in Braille, to do calculus and speak and understand many languages.

The fundamental premise of contemporary neuroscience is that cells in the nervous systems of multicellular organisms interact to produce this range of behaviors, as well as many others. Various metaphors have been introduced over the centuries to conceptualize how the nervous system produces this amazing array of behaviors. Early speculations considered, for example, that the elements of the nervous system were thin pipes that conducted fluids. However, a theme that developed during the 20th century was that the cells in the nervous system function in analogy to human-made computing devices. In this metaphor, the nervous system is a computational device and nerve cells, or neurons, are involved in sequences of complex computations. Neurons interact with each other through synapses, which can be either excitatory or inhibitory.

Individual neurons receive inputs from many, neighboring neurons. The inputs are integrated and result in electrical signals that are sent, in turn, to other neurons. Neurons are, in many instances, grouped into functional units, or circuits that are, then, the fundamental computational units of the nervous system. Circuits, in turn, can be combined to form large, interacting networks of neurons. An overall goal in neuroscience is to understand how the properties of individual neurons and neuronal circuits determine organismal behavior. This approach to thinking about the function of the nervous system has come to be called *computational neurobiology* or *computational neuroscience* and is the subject of this book. This chapter reviews the history of the concept that neurons function as computational devices, setting the stage for detailed discussions in the following chapters.

THE NEURON DOCTRINE

The idea that the nervous system is composed of cells similar to those present in other body tissues was controversial until the 20th century. Shepherd (1991) has produced a readable history of the dispute. Robert Hooke in 1665 and Anton von Leuwenhoek in 1674 used the earliest light microscopes to study the structure of plant and animal tissues. Over a century later, Jan Purkinje (1837) in Czechoslovakia was the first to describe cell-like elements in the nervous system. Figure 1-1 is a drawing of the cerebellum from his paper showing the small granule cells and the large "corpuscles" that now bear Purkinje's name. At about the same time, Matthias Schleiden (1838) conducted extensive microscopic examinations of

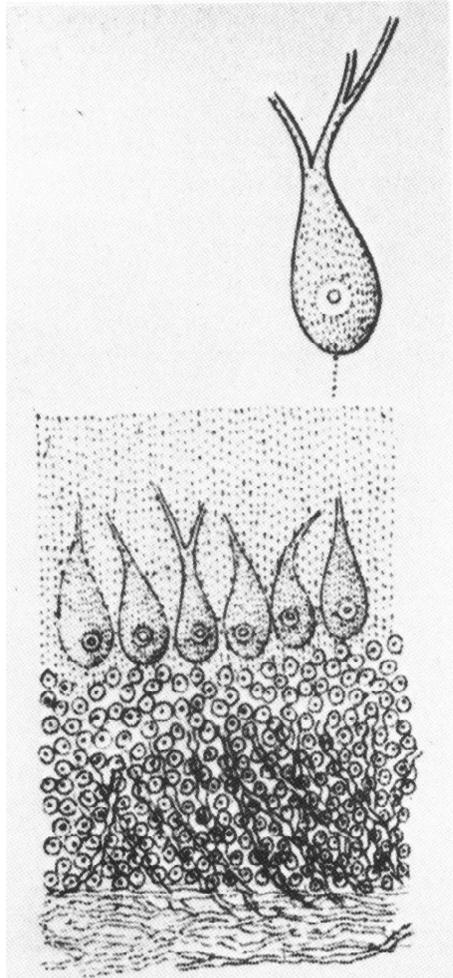


Figure 1-1. Purkinje cells. This is a drawing of cells in the cerebellum done by Purkinje. The top figure is of a single, isolated Purkinje cell showing the cell body and parts of the cell's dendrites. The bottom drawing shows parts of several Purkinje cells situated among many, small granule cells. From Purkinje (1837).

Plants and suggested that plant tissues are composed of fundamental units, or cells. He extended this idea to include animal tissues in collaboration with Theodeus Schwann (1839), who was studying spinal cord tissue. Subsequent workers described cells in various regions of the nervous system. However, the methods available for staining nervous tissue at that time were not able to clearly delineate the structural elements of nervous tissue. It quickly



Figure 1-2. Network of neurons. This drawing suggests that neurons in the spinal cord are interconnected with one another. A number of motoneurons are interconnected through their dendrites. Arrows show the direction of nerve impulses emanating from the neural net. From Kölliker (1867).

became clear that nerve cells have long, thin processes which came to be called dendrites and axons, but it was difficult to understand the origin and termination of these processes. Workers such as Albrecht von Kölliker (1849) and Joseph von Gerlach (1872) argued that the nervous system is composed of an interconnecting network of processes. They presented drawings of nerve cells in which the dendrites of one cell anastomose with the dendrites of other cells (Fig. 1-2). It appeared that nervous tissue might

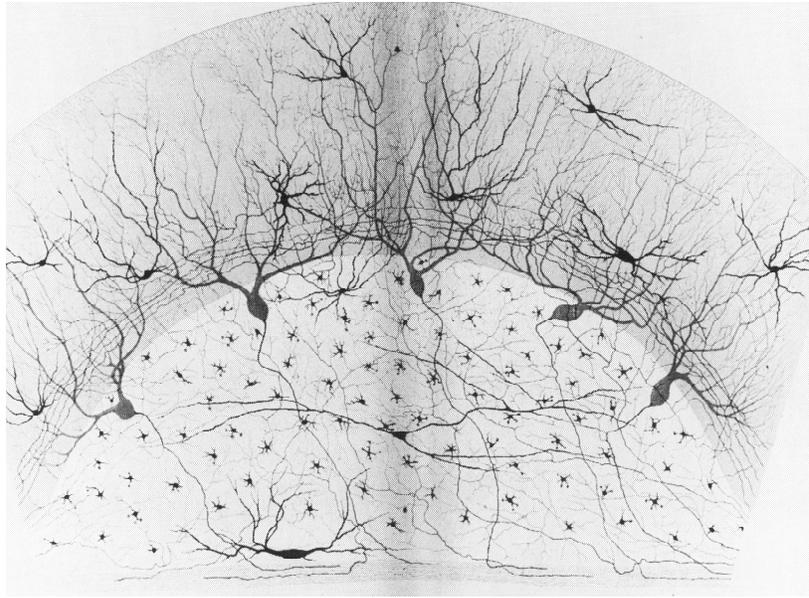


Figure 1-3. Nerve cells in the cerebellum. This drawing by Camillio Golgi shows several neurons in the cerebellum. The large neurons are Purkinje cells. Notice that, as compared to the drawing in Figure 1-1, this drawing provides more complete views of Purkinje cells. From Golgi (1886).

be an exception to the cell doctrine in that it was an interconnected reticulum instead of an aggregation of discrete cellular units.

A breakthrough came when the Italian anatomist, Camillio Golgi (1873), introduced a new method of staining tissue. His method depended upon fixing the tissue with agents such as osmium tetroxide and then soaking the tissue in silver salts. It turns out that the "Golgi method" impregnates or fills membrane-bound spaces with silver precipitates. Thus, Golgi was able to identify the Golgi apparatus in several different types of cells. Figure 1-3, for example, shows a drawing of cells in the cerebellum. Comparison with Purkinje's drawing in Figure 1-1 emphasizes the significance of Golgi's contribution. When applied to nervous tissue, Golgi methods typically

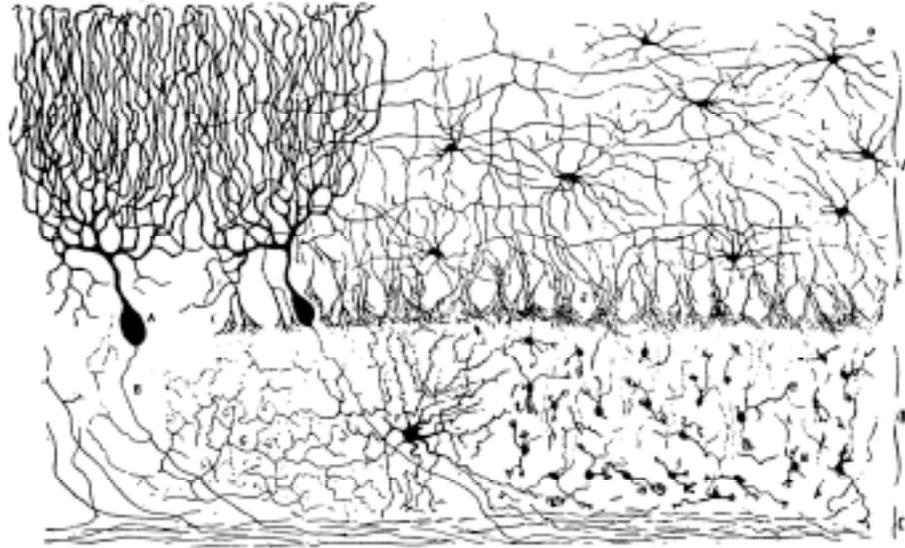


Figure 1-4. Illustration of cells in the cerebellum by Santiago Ramón y Cajal. This drawing by Cajal shows several type of cells in the cerebellum. From Cajal (1888).

impregnate only a fraction of the nerve cells. Individual cells can be impregnated in their entirety and the processes of nerve cells could be clearly traced in Golgi preparations. The Spanish microscopist, Santiago Ramón y Cajal, applied Golgi's methods to a wide range of nervous tissues from both invertebrates and vertebrates. His first paper, published in 1889, dealt with the cerebellum of birds and contains a drawing (Fig. 1-4) of the cerebellum that can be compared with Golgi's drawing in Figure 1-3. Cajal's work is summarized in a two volume work that was originally published in Spanish, but has recently been translated into English (Swanson and Swanson, 1995). The scientific community was impressed with the variety of shapes and forms of nerve cells visible in Cajal's drawings. This variety was not consistent with an undifferentiated reticulum of neural processes, and cells appeared to be discrete units in Golgi preparations. Wilhelm von Waldeyer (1891) summarized

the new knowledge about nervous system structure in a series of six review articles. He coined the term "neuron" for nerve cells in this review and founded the neuron doctrine, which held that nervous tissue was no exception to the cell doctrine. Golgi was eclipsed by Cajal's work and grew embittered. Ironically, the two men shared the Nobel Prize for Physiology in 1906.

While microscopists were debating the structure of nervous tissue, physiologists were learning that nervous tissues shows electrical activity. Luigi Galvani (1791) showed that stimulation of the brain of a frog could elicit leg movements, which he attributed to "animal electricity". The Italian, Nobili, invented a "galvanometer" that permitted accurate measurements of electrical currents in 1825. By 1840, Carlo Matteuci had used the galvanometer to measure a steady current from the cut end of a living muscle fiber. Slightly later, Emil Du Bois Raymond (1843) showed this current results from an electrical event, now recognized as an action potential, and the German physician/physicist, Hermann von Helmholtz (1850) showed that these events propagate with a finite velocity along nerve fibers. It was natural to combine the finding of electrical activity in the nervous system with the reticularist position which depicted the nervous system as an interconnected network of elements that could serve for electrical communication within the brain.

The discovery of electrical activity in the brain had considerable philosophical import because it supported the idea that living tissues could be explained in terms of physiochemical phenomena. Physicists and mathematicians were developing the basis of the theory of electricity and magnetism during this period of time, so it was possible to suggest that the behavior of nervous tissue could be described mathematically. The theoretical

basis underlying battery potentials that was developed by Nernst (1888) could, for example, be applied to understanding the resting membrane potential that is characteristic of living cells. Bernstein (1912) suggested that biological membranes are selectively permeable to ions and used the equation developed by Nernst to predict the potential difference across the membrane due to differences in the concentration of potassium ions inside and outside of cells. More detailed measurements of action potentials were made by Erlanger and Gasser (1937) using a variety of nerve fibers. These studies involved using voltmeters to measure the voltage changes induced in muscle or nerve fibers by electrical stimulation of tissue. They were adequate to characterize the basic electrical properties of nervous tissue, but were not helpful in understanding their mechanistic basis. One limitation was that electrical potentials were made from outside of the cell as an extracellular recording instead of being made across the membrane of the cell as an intracellular recording. This limitation was removed by the development of microelectrodes that permitted at least large nerve cells to be impaled with a thin recording device that measured the potential difference between the inside and outside of the nerve cell. A second limitation was that the electrical currents that produced voltage changes could not be measured. This limitation was removed when K. S. Cole (1949) developed an electrical circuit, called the voltage clamp, that permits the control of cellular membrane potentials and permitted the measurement of currents in nerve cells.

Two Englishmen, Hodgkin and Huxley (1952), then used the voltage clamp to analyze the mechanisms underlying the generation and propagation of action potentials in the giant axon of squid. The biochemical nature of cellular membranes was not fully understood at this time, but Hodgkin and Huxley

postulated that the membranes of electrically active cells contained elements that they called "particles". They hypothesized that membrane particles were altered when an action potential is produced. They used the voltage clamp to characterize the currents that flow through the axon's membrane during an action potential. Experimental methods showed these currents involved sodium and potassium ions flowing through the membrane. Hodgkin and Huxley then applied concepts from chemical kinetics to explain these currents in terms of the behavior of the postulated membrane particles.

Hodgkin and Huxley's findings provided a compelling explanation of how action potentials can propagate along the elongate cellular processes that the early microscopists had identified in neural tissue. However, it was not clear what happens at the ends of the processes. (See Cowan and Kandel (2001) for a history of this topic.) The reticularists argued that neurons are physically connected via their processes so that electrical current could simply flow from one neuron to another through an anastomotic network. The neuron doctrine, by contrast, held that individual neurons are separate structures, and postulated that nerve cells form contacts of some kind between themselves (and between nerve cells and other targets such as muscle fibers or gland cells). If the neuron doctrine is true, how does information get from one neuron to another across the gaps that separate neighboring nerve cells ? The distinguished British physiologist, Charles Stead Sherrington, named these contact points "synapses" from the Greek word for "contacts" (1897). There was increasing evidence that synaptic transmission between neurons involves a chemically mediated state. Otto Loewi (1921), for example, studied the control of cardiac activity in frogs by the vagus nerve. He showed that electrical stimulation of the vagus nerve caused a release of an unknown chemical that

could modulate heart rate. It was subsequently demonstrated that this chemical is the small molecule, acetylcholine.

The existence of a chemical step in the transmission of information across synapses was convincingly demonstrated by Bernard Katz and his colleagues (del Castillo and Katz, 1954) using the neuromuscular junction, the synapse between the axons of motoneurons and muscle fibers. An elegant analysis provided strong evidence that packets, or "quanta", of acetylcholine were released by motoneuron axons and evoked electrical potentials in muscle fibers. They performed a statistical analysis of the amplitudes of these potentials and showed there is a random, or stochastic, component to synaptic transmission. The nature of quantal transmission was clarified when electron microscopy was applied to the nervous system by de Robertis and Bennett (1955) and Palay and Palade (1955). It was immediately obvious that axon terminals contain membrane bound vesicles that serve as the anatomical substrate for physiological quanta of neurotransmitters, like acetylcholine, that mediate the chemical step in synaptic transmission. It turns out, however, that both chemical and electrical synapses exist in the nervous system. Most synapses in multicellular animals are chemical synapses, but a significant minority of synapses in multicellular animals involve a direct flow of electrical current between nerve cells via intercellular connections called gap junctions that provide a continuity between two nerve cells. The nature of both chemical and electrical synapses was summarized by John Eccles (1964) who was a leading figure in the development of concepts of how neurons interact with each other. The electronics needed to record electrical potentials across the membranes of neurons in the central nervous had been developed by the 1950's. Eccles and his many collaborators applied this technology to a variety

of neurons. This body of work clarified how chemical synapses mediate both excitatory and inhibitory potentials in neurons.

Wilfrid Rall, a student of Eccles, developed a quantitative framework for understanding the interactions that occur when a given neuron is subjected to the activity of many individual synapses. (See Rall (2001) for a history.) Rall applied the mathematical theory of the flow of current in electrical cables to neuronal processes. This was a key step in establishing the metaphor that the nervous system is essentially a computational device and will be the starting point of our detailed discussion in Chapter 2. William Thomson (who was later designated Lord Kelvin) derived an equation, the cable equation, that provides a quantitative description of current flow in electrical cables. Rall used cable theory to explain how the voltage changes that result from the activity of chemical synapses propagate along dendrites. He then used an early digital computer to model interactions between synapses in motoneurons and account for experimental data being collected by Eccles and others.

The neuron doctrine had, thus, been firmly established by the 1970's and a coherent concept of the basic processes involved in the function of individual neurons was rapidly being established. At the same time, the idea that neurons interact with each other to form functional circuits was being developed.

NEURAL CIRCUITS

The earliest drawings of neural tissue resemble drawings of other tissues of the body in that they provide accurate representations of what is visible

when the tissue is viewed through the microscope. However, Ramón y Cajal and his student, Raphael Lorente de Nó, took the analysis a step further and developed the concept of a neural circuit. Many of Cajal's diagrams show groups of neurons and suggest functional interactions between them. The early physiologists understood that electrical impulses can actually propagate in both directions along a neuronal process such as an axon. The idea that impulses proceed in one direction from neuron to neuron requires some mechanism of imposing a unidirectionality on neural activation. Cajal argued for such a mechanism as the "law of dynamic polarization" and proposed how information flows in neural circuits by including small arrows in some of his diagrams (Fig. 1-5). Cajal also developed the idea that a specific neural structure contains a fixed repertoire of neuronal types that interact with each other in a stereotyped manner. He surveyed the structure of the retina in a wide range of vertebrates and argued that the retinas of all of the vertebrates contain the same populations of neurons that interact in a circuit with an organization that is relatively constant across species.

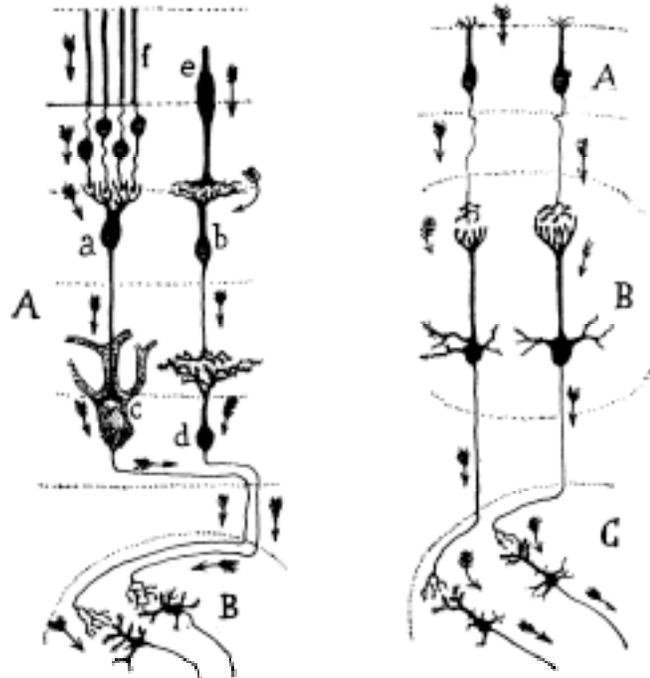


Figure 1-5. Examples of information flow in neural circuits. These drawings by Cajal show the hypothesized flow of information in two neural circuits. A. Cells in the retina showing information flowing from photoreceptors (e,f), to bipolar cells (a,b) to ganglion cells (c,d) through the optic nerve and to neurons in the brain. B. Cells in the olfactory system showing the flow of information from olfactory receptors (A,B), to mitral cells in the olfactory bulb (B) to cells in the olfactory cortex (C). Arrows in both diagrams show the flow of information. From Cajal (1989).

Continuing in the same spirit, Lorente de Nó completed a seminal analysis of the structure the hippocampus (Lorente de Nó, 1934) in mammals. We will examine Lorente de Nó's work in some details because it provides the basis for a great deal of important work in neuroscience. The hippocampus is situated on the medial surface of the cerebral hemispheres of mammals. It is important in complex behaviors such as memory and the formation of a cognitive map (a general understanding on the part of the animal of its position in its environment). A striking feature of the hippocampus, like other regions of the cerebral cortex, is that it contains alternating layers of densely packed cell bodies, dendrites and bundles or fascicles of axons. Cells within a given layer

have characteristic morphological features. The most obvious cells in the hippocampus are pyramidal cells, whose somata often have a pyramidal configuration, prompting the general name. Most pyramidal cells have two dendritic arbors. A basal dendritic tree extends towards the ventricular surface and an apical tree extends towards the pial or external surface of the brain. Dendrites in both the basal and apical trees are covered by small protuberances or *dendritic spines*. Axons of pyramidal cells in one cortical area may have collaterals in the immediate vicinity of the cell, but also travel to other regions of the brain. They are *projection neurons* as opposed to purely *local circuit neurons*. The axons of pyramidal cells can terminate in other regions of the cortex on the same side of the brain (*association projections*), in cortical regions on the opposite side of the brain (*commissural projections*) or leave the cortex entirely.

Although all pyramidal cells have important features in common, they also show significant anatomical variations. The hippocampal formation can be divided into several major areas, each of which contains pyramidal cells with a characteristic anatomy. These differences were first detailed by Ramón y Cajal (1895) and then by Lorente de Nó (1934) using the Golgi method to impregnate pieces of mouse and rat hippocampus. Figure 1-6 is a drawing from Cajal's work. The neurons to the right are pyramidal cells from the hippocampal area known as CA3. The hippocampus derives its name from its vague resemblance to a cornucopia or Ammon's horn - *cornu ammonis*. CA3 is one of four fields (CA1, CA2, CA3 and CA4). Another region, the *dentate gyrus*, is named because of a fanciful resemblance to teeth. The CA fields and the dentate gyrus all contain three distinct layers. Pyramidal cells in the CA fields have somata in the second or pyramidal cell layer of these areas -- the *stratum*

pyramidale. They have relatively large somata, a single apical dendrite and several basal dendrites. The apical dendrite is noteworthy in that it bears several large, thorny excrescences close to its soma. They are from the axons of granule cells -- a modified form of pyramidal cell located in the dentate gyrus (Fig. 1-7). Granule cell axons bear large swellings, or varicosities, located in tandem on their parent axons. The varicosities give the axons a mossy appearance and the axons are, thus, known as *mossy fibers*. The axons belong to the *granule cells* -- a modified form of pyramidal cell located in the dentate gyrus (Fig. 1-7). The varicosities on the granule cell axons contact the excrescences on the dendrites of the CA3 pyramidal cells.

Axons of the CA3 pyramidal cells give rise to several collaterals, some of which (known as the *Schaffer collaterals* after the German anatomist who first described them) course to the left in the drawing and intersect the apical dendrite of a second variety of hippocampal pyramidal cell located in hippocampal field CA1. This pyramidal cell is more slender than the CA3 pyramidal cell (Fig. 1-6). The CA1 pyramidal cells receive synapses from the Schaffer collaterals some distance from their somata at a point where the apical dendrites begin to branch into a tuft of daughter branches that reach the surface of the hippocampus.

The somata of both sets of pyramidal cells are packed together in a tight layer within the hippocampus and their dendrites are in precise alignment. Axon systems -- such as the mossy fibers or the Schaffer collaterals -- running through the hippocampus consequently intersect

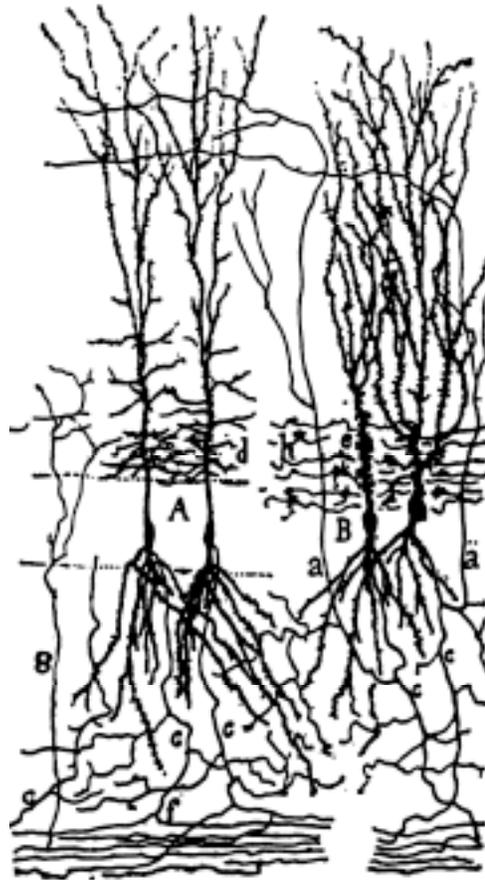


Figure 1-6. Hippocampal pyramidal cells. This drawing by Ramón y Cajal shows pyramidal cells from the CA3 and CA1 regions of the hippocampus. A pair of CA3 pyramidal cells (with somata labeled B) are shown on the right. Mossy fibers (e) approach the CA3 pyramidal cells from the right and form large presynaptic elements that contact the proximal shafts of the apical dendrites of the pyramidal cells. The axons of the CA3 pyramids (a) ascend in the cortex and course to the left, where they enter the CA1 region and intersect the distal apical dendrites of a pair of CA1 pyramidal cells. These axons are the Schaffer collaterals. From S. Ramón y Cajal (1995).

dendrites of hippocampal pyramidal cells in specific regions, establishing a very ordered set of interconnections. They are so tightly packed that they form discernible layers in the CA fields. The mossy fibers form the *stratum lucidum* of field CA3 and the Schaffer collaterals form the *stratum lacunosum* in field CA1. Synapses effected by both the mossy fibers and the Schaffer collaterals are excitatory and cause a flow of ions

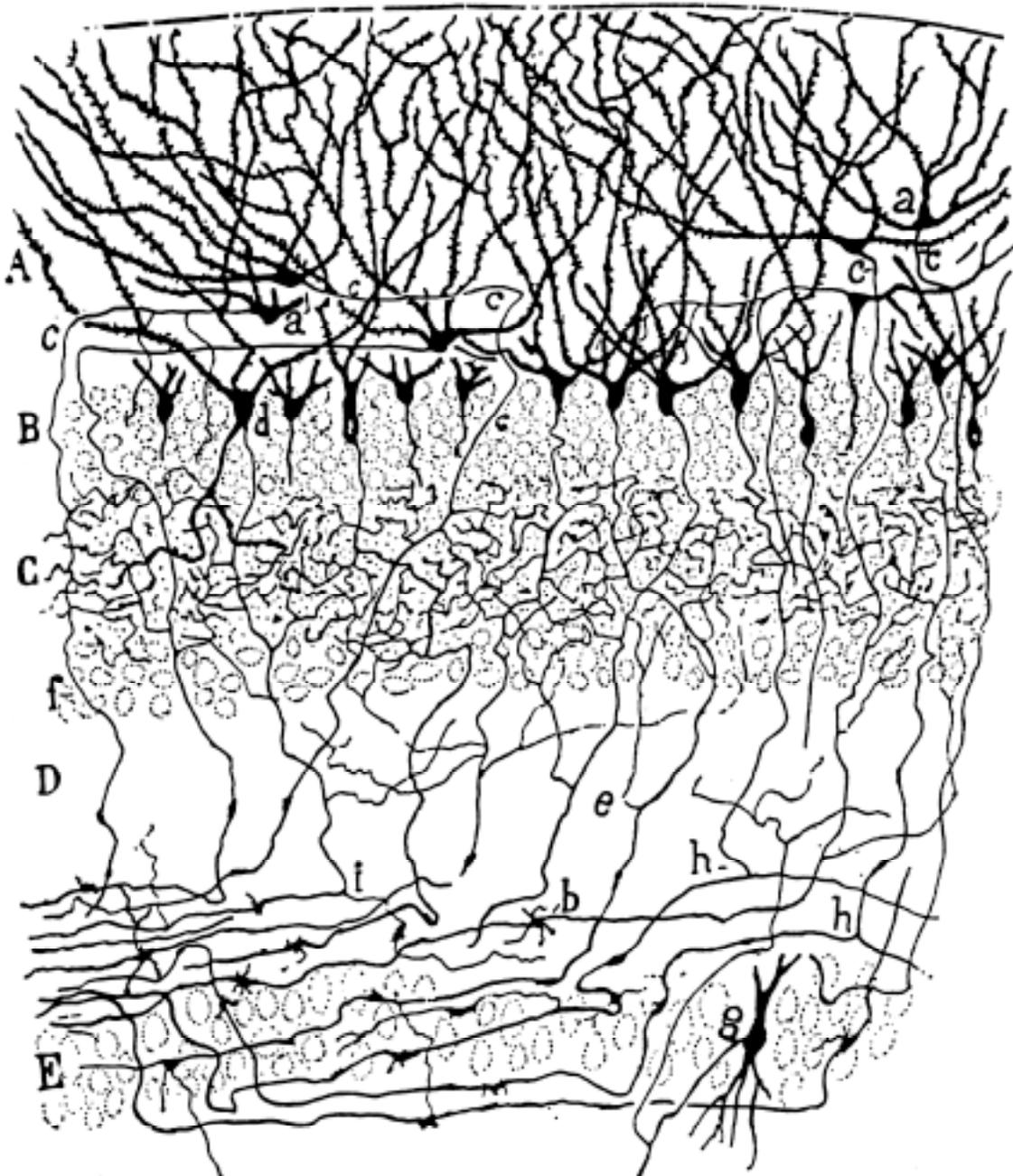


Figure 1-7. Dentate gyrus granule cells. This drawing by Ramón y Cajal shows a population of granule cells from the dentate gyrus region of the hippocampus. The granule cells have somata arranged in a layer of densely packed cells at the top of the figure. Their dendrites extend towards the top of the figure where they abut the surface of the brain that forms the inner surface of the hippocampal fissure. Their axons course towards the bottom of the figure where they turn and form the mossy fibers of the CA3 region. From S. Ramón y Cajal (1995).

into the dendrites of the CA3 or CA1 pyramidal cells, respectively, when activated.

Cajal suggested that the granule cells, CA3 pyramidal cells and CA1 pyramidal cells interact to form a defineable circuit within the hippocampus. Axons that course across the dendritic fields of granule cells originate in a cortical region that is situated adjacent to the hippocampus. This is the *entorhinal cortex*, which is known to receive several different kinds of sensory information. The granule cell axons course into CA3 and intersect the proximal dendrites of the CA3 pyramidal cells. The axons of the CA3 pyramids, in turn, course into CA1 and intersect the dendrites of the CA1 pyramids. Both the CA3 and CA1 pyramidal cells are the sources of projections away from the hippocampus to other regions of the brain such as the septum and hypothalamus. This is an oversimplified picture of neural circuits within the hippocampus because we are ignoring several kinds of inhibitory cells known to be important in the organization of the hippocampus. Nonetheless, even this basic account of neuronal interactions in the hippocampus raise a number of specific questions. Why, for example, do the granule cell axons make synapses upon the proximal dendrites of pyramidal cells while the axons of the CA3 pyramids make synapses upon the distal dendrites of the CA1 pyramidal cells ? Do the differences in the size and shape of the CA3 and CA1 pyramids have functional consequences ? How do all of the small differences in the anatomy of these three kinds of cells determine the overall properties of hippocampal circuits ?

COMPUTATION IN NEURAL CIRCUITS

The hippocampal tri-synaptic circuit is an example of groups of two or more neurons that can be reasonably regarded as neuronal circuits. It demonstrates the general feature that the formation of circuits often depends upon the details of the anatomy and physiology of the neurons. There is, for example, a striking tendency for axons in the hippocampus to run within distinct layers and contact specific domains of their target neurons. We will now consider how such circuits can be regarded as implementing "computations".

McCulloch-Pitts neurons

The idea that neural circuits may be conceived as carrying out computations results from pioneering work done by Warren McCulloch and Walter Pitts and published as a 1943 paper entitled "A logical calculus of the ideas immanent in nervous activity". McCulloch and Pitts (1943) were influenced by the drawings of neuronal circuits that had been prepared by Ramón y Cajal and Lorente de Nó. They also drew from the field of symbolic or mathematical logic that had recently been developed by Alfred Whitehead, Bertram Russel and Rudolph Carnap. This branch of mathematics set out a series of axioms and rules for carrying out logical arguments or syllogisms. That is, assuming that one or more propositions are either true or false, mathematical logic provides rules by which complex arguments depending upon these propositions can be evaluated as either true or false.

McCulloch and Pitts proposed that neurons function as logical devices that combine inputs. A given neuron receives inputs from two or more neighboring

neurons. Each input is binary because an axon is either active or inactive at any point in time. McCulloch and Pitts assign the input, I , a value of 1 if it is active and a value of 0 if it is inactive. Inputs can be weighted and made more or less effective. The neuron sums up the values of all of the weighted inputs at each point in time. It has a threshold, θ , and generates an action potential if the sum of the weighted inputs exceeds θ . The output, O , of the neuron is also binary, either a 1 or a 0. We can make this more formal for the simple case of a neuron that has two inputs, I_1 and I_2 . Both I_1 and I_2 can be either 1 or 0. The inputs are weighted by numbers, w_1 and w_2 , which can be positive if the input is excitatory and negative if the input is inhibitory. Thus, the output of the neuron is $O = 1$ if $w_1 I_1 + w_2 I_2 \geq \theta$ and $O = 0$ if $w_1 I_1 + w_2 I_2 < \theta$.

To see how a McCulloch-Pitts neuron can compute logical functions, consider the function "and" or "+". Suppose we have two propositions, A and B . Each proposition can be either true or false. Given the truth values of A and B , we want to know if $A + B$ is true. This happens only when both A and B are true. To construct a McCulloch-Pitts neuron that computes "and", we give the neuron two inputs, set $w_1 = w_2 = 1$ and $\theta = 2$. This neuron will generate an output $O = 1$ if, and only if, the two inputs are 1. It is easy to construct a McCulloch-Pitts neuron that computes the logical function "inclusive or" which is true if either one, or both, of the propositions are

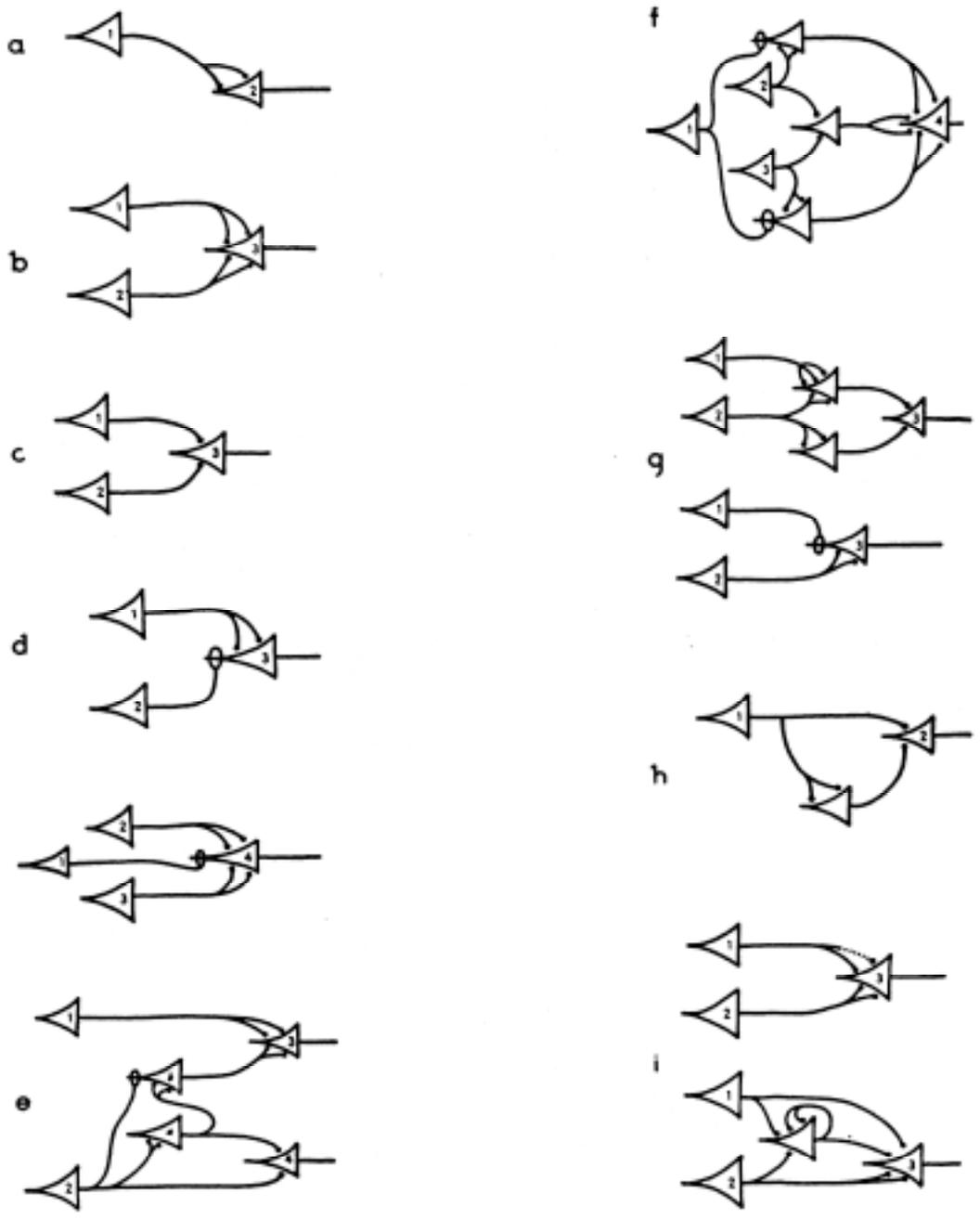


Figure 1-15. Networks of McCulloch-Pitts neurons. These are sketches of networks formed from two or more McCulloch-Pitts neurons.

true. In this case, we can simply reset the threshold at $\theta = 1$. The neuron generates an output of $O = 1$ if either one, or both, of the inputs has a value of 1 because the weighted sum is then above the threshold. These are logical functions that can be computed or evaluated using a single neuron. In their paper, McCulloch and Pitts show that a wide range of logical functions can be computed if networks of two or more neurons are constructed (1-8). They, in fact, provide a mathematical proof that a network of such neurons can evaluate any possible logical function.

The 1943 McCulloch and Pitts paper attracted a great deal of attention at the time and led to two sets of developments. One expanded into the field that was eventually called AI or artificial attentions. The second gave rise to the concept of neurons as computational devices. We will follow the development of McCulloch-Pitts neurons into artificial intelligence and artificial neural networks in this section, and come back to the evolution of McCulloch-Pitts neurons into concepts of biological neural computation in the last section of this chapter.

Networks with modifiable synapses

The networks of neurons that McCulloch and Pitts described were static in the sense that the weights of the synapses (the w_i 's) were fixed. The network was, thus, designed to compute one specific function. However, subsequent work considered networks that could learn to compute specific functions (or be trained to carry out specific tasks). This work was inspired by a set of ideas that were explicated by Donald Hebb at McGill University in Montreal. Hebb's ideas were brought together in 1949 in an important book called *The Organization of Behavior* (Hebb, 1949). Hebb was reacting, in part, to

a group of psychologists, such as B. F. Skinner, who argued that attempts to reduce behavior to physiological processes are inherently misleading. Hebb built explicitly on the work done by Lorente de Nó and McCulloch and Pitts. One of his ideas was that many behaviors involve groups of neurons that he called *cell assemblies*. He hypothesized that any frequently repeated stimulus would lead to the eventual development of a sensory cell assembly or diffuse structure comprising cells throughout the brain. The sensory assembly would be capable of acting briefly as a closed system that delivers facilitation to other assemblies that have a motor function. Hebb gives an example of visual perception based on what was known or believed at the time about the anatomy of the visual cortex. The German anatomist, Brodmann, had divided the cerebral cortex in mammals into a series of distinct areas. Brodmann assigned these areas numbers and his Areas 17, 18, 19 and 20 were thought to be involved in vision. Figure 1 – 9 is from Hebb's book and illustrates his idea of how cell assemblies might be organized. It was well established that Area 17 embodies a topographic map of visual space. Thus, the cells in regions A, B and C of Area 17 would respond preferentially to stimuli at different loci in visual space. Cells in the three regions send axons in different patterns to cells in Area 18. The patterns are such that cells D through H in Area 18 would respond to different combinations of visual stimuli. Hebb suggested that assemblies of cells such as these might be responsible for the perception of stimuli such as lines oriented in particular directions or angles in the visual scene.

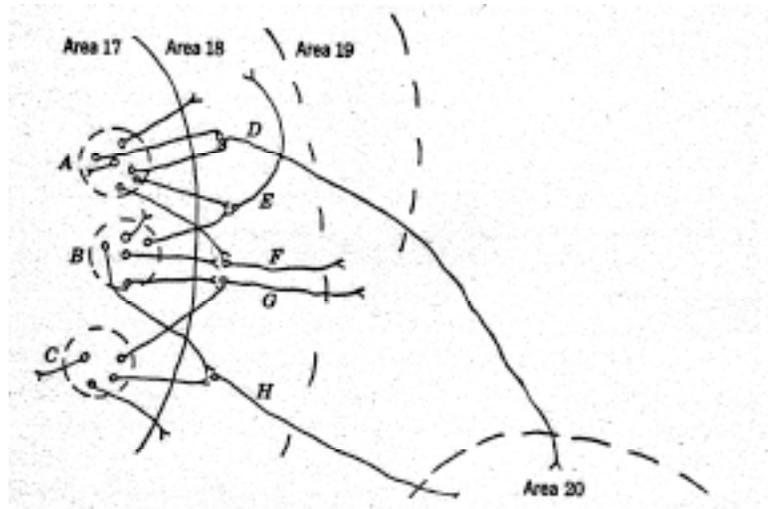


Figure 1 - 9. Cell assemblies in the visual cortex.

Hebb also considered how such cell assemblies could be formed and suggested that they might develop through repeated stimulus patterns. He suggested that any two cells, or systems of cells, that are repeatedly active at the same time would tend to become "associated", so that activity in one cell facilitates activity in the other. Hebb did not know how this process of association might be achieved at the cellular level, but he did suspect that it involved strengthening synapses. The idea has come to be called Hebb's postulate and synapses that are strengthened in this way are "Hebbian" synapses. The mechanisms underlying this process are still not completely known and remain an active area of research. However, the concept of modifiable, or plastic, synapses was key to the development of artificial neural networks.

Perceptrons

The first step towards artificial neural networks was made by F. Rosenblatt who introduced networks that he called perceptrons in a 1958 paper (Rosenblatt, 1958). Figure 1-10 is taken from Rosenblatt's paper and shows the general structure of a perceptron. You should notice that this figure is similar to the figure from Hebb's book that is reproduced in Figure 1-9. The idea of the perceptron is to build a network that can recognize patterns. The perceptron, thus, receives inputs from a retina and produces responses that correspond to individual patterns. The retinal inputs feed into a projection area that corresponds to Area 17 in Hebb's diagram. The projection area feeds into an association area, that corresponds to Area 18 in Hebb's diagram. The projection and association areas each contain a number of units. Each unit effects a linear combination of its inputs, just as is done in McCulloch-Pitts neurons. The problem is that it is not immediately obvious how to set the synaptic weights in the projection and association areas so that the perceptron can discriminate inputs. The difference between McCulloch-Pitts networks and perceptrons is that Rosenblatt took advantage of Hebb's idea of modifiable synapses and suggested that the synaptic weights could be modified in some systematic way so that the network would effectively learn how to discriminate patterns.

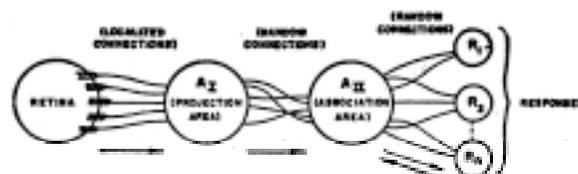


Figure 1 - 10. A perceptron. From Rosenblatt (1958).

Rosenblatt was able to show that perceptrons could, in fact, compute a number of interesting functions. His finding generated considerable enthusiasm and fostered the idea that it could be possible to construct artificial networks that were capable of “intelligent” computations. It, unfortunately, turned out that perceptrons are actually limited in the computations that they can effect. This was demonstrated in an important book, *Perceptrons*, published by Minsky and Papert in 1969. They (Minsky and Papert, 1969) showed that perceptrons are not able to compute exclusive or. That is, they can not evaluate the function either A or B, but not both A and B. The nascent artificial intelligence community subsequently lost confidence in artificial neural networks.

Artificial neural nets

Interest in artificial neural networks was rekindled in by several workers who developed a family of networks with an architecture that was different from that of perceptrons. Like perceptrons, the new generation of artificial neural nets have an input layer (the projection area of the perceptrons) and an output layer. However, they also have an intermediate or hidden layer

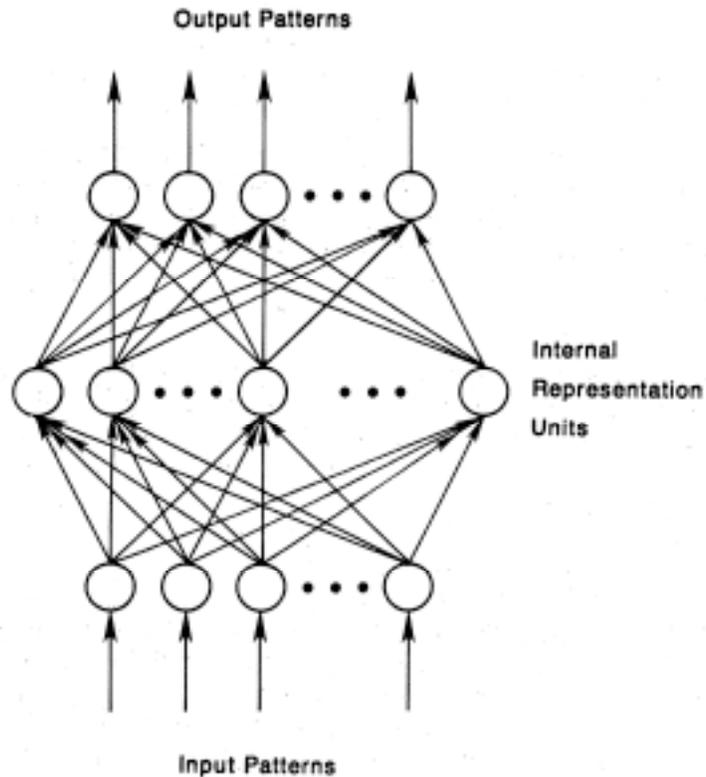


Figure 1 -12. An artificial neural network.

that is positioned between the inputs and outputs. Each hidden unit receives inputs from all of the input units and provides inputs to all of the output units. In contrast to McCulloch-Pitts neurons and perceptron units, the hidden units use a sigmoidal function to combine the inputs from the first layer of the network. If the inputs to a given hidden unit are I_1, I_2, \dots, I_n , then the output, O , from that unit is given by

$$(1) \quad O = \frac{1}{1 + \exp(-\sum_i w_i I_i)}$$

where the w_i 's are synaptic weights. In addition to introducing a new architecture for the networks, workers in the area of artificial neural networks introduced a method (the delta rule) of selecting the synaptic weights. The procedure is to present the network with a specific pattern of inputs. The idea

is that the output should represent the trial input. This will, generally, not happen, so the synaptic weights are adjusted and the stimulus is presented again. This process is repeated again, typically thousands of times, until the output corresponds to the input. The entire process can then be carried out again with a new input. The network can eventually be trained in this fashion to recognize each item in a family of input patterns and, hopefully, also be able to generalize and recognize novel inputs.

Artificial neural networks have proven to be very useful tools in areas of inquiry as disparate as economics and astrophysics. However, any sense of biological reality was lost along the way so that artificial neural networks are mathematical tools. They are useful in modeling some aspects of neural function and behavior, but do not give insights into how real neurons process real inputs. Artificial neural nets will, then, not be discussed further in this book. The interested reader is referred to texts such as Haykin (1999).

COMPUTATION IN BIOLOGICAL CIRCUITS

In this last section of Chapter 1, we return to the McCulloch – Pitts paper and trace the impact of their ideas on concepts of computation in real neurons. It is important to emphasize that McCulloch-Pitts neurons differ in several important ways from real, biological neurons. The first is that they do calculations in discrete time steps. All of the inputs to the neuron are activated at the same time, the weighted sum is calculated and the output is generated in the next time step. Inputs to real neurons are not necessarily synchronized in this way. Second, the outputs of real neurons are not really digital. Axons have the all-or-none property of either generating an action

potential or not generating an action potential, but neurons that exceed their firing thresholds often generate trains of action potentials in fairly complicated patterns.

The Eccles neuron

Despite these differences from real neurons, the concepts incorporated in McCulloch-Pitts neurons have considerable heuristic value and bear some resemblance to the concepts of neuronal function that were being developed in the 1950's when it became possible to obtain intracellular records from neurons with microelectrodes. Many of these ideas were summarized by Sir John Eccles in two books, *The Physiology of Nerve Cells* (1957) and *The Physiology of Synapses* (1964). We will refer to this model of neuronal function as the "Eccles neuron". To take a simple example, consider a neuron that has a soma, two dendrites and an axon. The dendrites receive synapses from two axons (originating from two different, neighboring neurons) on the two dendrites. The synapses become active when action potentials invade the axons. They function by releasing neurotransmitter that binds to receptor proteins on the neuron. The receptors open and generate ionic currents that flow into the dendrites. Currents then flow down the dendrites towards the soma of the neuron and either depolarize or hyperpolarize the soma membrane. Since some of the current leaks out of the membrane as it flows towards the soma, synapses that are distant from the soma will be more effective than those that are near the soma. Inputs are, thus, weighted according to the distance of the synapse from the soma. The soma membrane acts as a linear threshold device - as in McCulloch-Pitts neurons - and generates an action potential if the electrical potential at the soma exceeds some threshold value.

Section II of this book (single neuron computation) will essentially consider the extent to which the concept of individual neurons as linear devices is accurate. It turns out that the Eccles model of neuronal calculation is sometimes accurate. However, neurons typically can carry out more complex, non-linear computations.

What the frog's eye tells the frog's brain

McCulloch and Pitts had a direct influence on biological implementations of their idea through a collaboration with Jerome Lettvin and H. R. Maturana at MIT. This group was interested in understanding the responses of neurons in the visual system to visual stimuli. The basic idea of their work was published in an influential article (Lettvin et al., 1959) entitled "What the frog's eye tells the frog's brain". By 1959 there was already a large literature dealing with recordings from the retina and structures in the visual system to visual stimuli. However, the stimuli that had been used were principally flashes of light. Lettvin and his colleagues argued that the visual system most likely evolved to respond to complex visual stimuli that were important in the natural behavior of the animal instead of simple flashes of light. They, consequently, designed a series of experiments in which they recorded from axons in the optic nerve as it leaves the back of the frog's eye. These are axons of the retinal ganglion cells which send information from the eye to the brain. Recordings were made while the animal was presented with stimuli that were selected to resemble objects that might occur in the animal's natural environment. One stimulus, for example, was a small, round object chosen because it would resemble a fly or other insect that would be a suitable prey for the frog. The result was that

different individual ganglion cell axons responded to different stimuli. This suggested that retinal ganglion cells are able to recognize or compute specific features of visual stimuli that are behaviorally important. Maturana studied the anatomy of frog retinal ganglion cells and discovered that he could establish tentative correlations between the shape and size of ganglion cells and their physiological properties. Circuits of neurons in the retina, thus, could be viewed as carrying out behaviorally significant computations on the patterns of light that fall upon the retinal surface. We will follow the further development of this approach in Part III of this book.