REVIEWS OF MODERN PHYSICS VOLUME 34, NUMBER 1 JANUARY, 1962

The Perceptron: A Model for Brain Functioning. I*

H. D. BLOCK

Cornell University, Ithaca, New York

THE Perceptron is a self-organizing or adaptive system proposed by Rosenblatt.[1] Its primary purpose is to shed some light on the problem of explaining brain function in terms of brain structure. It also has technological applications as a pattern-recognizing device, but here our emphasis is on the brain function-structure problem. The technological aspects are not completely irrelevant however, since a model, no matter how appealing it may appear from the point of view of structural similarity, must also be judged on the basis of its performance.

In brief a perceptron consists of a *retina* of *sensory units* (for example photocells); these are connected (for example by wires) to *associator units*. The connections are many to many and random. The associator units may be connected to each other or to *response units*. When a *stimulus* is presented to the retina (for example as a pattern of illumination) impulses are conducted from the activated sensory units to the associator units. If the total signal arriving at an associator unit exceeds a certain *threshold* then the associator becomes *active* and sends an impulse to units to which it is connected. The magnitude of the impulse carried by certain connections depends on the past activity of their termini according to certain preassigned *reinforcement rules*. Thus the device changes its internal functional properties. The resultant behavior exhibits, as is shown in the text, interesting aspects of learning, discrimination, generalization, and memory.

We present here a survey of the work to date. We also give in detail the proofs of certain theorems which illustrate some of the methods of analysis and which, in view of their central position, illuminate a wide area. Further elaboration and details are given in the references cited, particularly in Rosenblatt's summary report[1] which presents a detailed and comprehensive exposition of the entire subject.

BACKGROUND AND MOTIVATION

1. Structure and Function

For most of the organs of the body (e.g., the heart, lungs, kidneys, stomach, intestines, liver, spleen, blood stream, bones, skin, peripheral nervous system, etc.) we have some idea of the functions each performs and some explanation of how the structure operates to achieve the function.

Conversely, for most of the functions necessary to sustain life (e.g., locomotion, sensitivity, respiration, reproduction, digestion, nutrition, excretion, etc.) we have some idea of which structures are involved and the manner in which they implement the function.

It is now generally believed that the brain is the principal organ involved in thought, but there is no reasonably precise explanation of how the action of the brain structure produces the "higher functions."

Admittedly there are still open problems for the other organs. We do not know enough to build a real lung. But we do know in a general way that the function of respiration is to replace carbon dioxide in the blood by oxygen. The lung offers a large thin surface area where air can get on one side and blood on the other, and the exchange can take place. Admittedly, for a specialist in this field the really interesting problems start at this point. Nevertheless we take it as evident that at the present time the degree of our ignorance on the brain "function-structure problem" is of a higher order than for the corresponding "respiration-lung problem." To pose the problem precisely, we should now define brain function and brain structure.

2. Brain Function

Since many psychological phenomena have not yet been investigated, any description of "brain function" is necessarily incomplete. Moreover, even those phenomena which have been intensively studied often admit a multiplicity of interpretations.[2,3] It therefore seems pointless to attempt to make a precise definition of "brain function." Clearly there is such a phenomenon. Clearly it is related to perception, memory, discrimination, recognition, association, comparison, learning, communication, reasoning, and attention. An operational definition of these terms might be given in terms of the relation between (1) a sequence of inputs, which might be suitable physical stimuli; e.g., a light pattern on the eye, sound on the ear, pressure or heat on the skin, etc., and (2) a sequence of outputs, which might be the observable response of the subject.

Such a definition might be criticized as being too narrow in that it neglects *thought* which is not triggered by an observable stimulus or displayed in an observable response. This objection might be answered by the argument that, in principle, thought must be accompanied by physical, hence observable, changes some-

---

* Research sponsored by the Office of Naval Research.

[1] F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms* (Spartan Books, Washington, D. C., 1961).

[2] E. G. Boring, H. S. Langfeld, and H. P. Weld, *Foundations of Psychology* (John Wiley & Sons, Inc., New York, 1948).

[3] E. R. Hilgard, *Introduction to Psychology* (Harcourt Brace & Company, Inc., New York, 1957), Chaps. 10–31.

where in the brain and these might be interpreted as the inputs and the outputs.

We are not concerned here with the difficult problems of determining which behavior patterns are *innate* and which ones are *learned.* While questions of innate behavior, such as the navigation systems of certain birds and the mechanisms of their genetic transmission, are clearly formidable, nevertheless most scientists expect that the explanations, when they come, will be based on conventional physics, chemistry, and mathematical analysis and could, in principle, be duplicated by engineering techniques a few orders of ingenuity beyond our present state. For learned behavior on the other hand, we believe, as will be elaborated subsequently, that the physics and chemistry involved may be straightforward, but the organization of information and the mode of operation are based on radically novel principles requiring entirely new concepts of analysis. Thus from the viewpoint of the present paper we would not be particularly concerned with neuronal circuitry yielding an "unconditioned reflex" behavior pattern, such as the pupillary contraction of the eye under bright light. We would, on the other hand, be very interested in learning what physical changes occur in a child's brain when he learns to recognize the letter "A" and how, in precise terms, these changes account for the learning.

Similarly it is possible to arrange circuits so that a machine will perform tasks which appear to have much in common with thinking. For example patter-recognizers,[4,5] chess players,[6] and other "thinking machines" have been built.[7] Although such special purpose machines clarify the nature of the logical problem to be solved, they contribute little to the understanding of how the brain is organized to perform these functions. Even the "heuristic reasoning machines,"[8] while they reveal a great deal about the nature of human reasoning, are several orders of abstraction beyond the basic mechanisms with which we are concerned here.

Let us leave the definition of "brain function" in this rather vague state. These concepts are very slippery and will probably never be formulated in a way that will satisfy everybody.

3. Gross Brain Structure

The gross anatomy of the brain has been well studied.[9–12] Certain regions appear to have well defined functions; in particular, stimulation of points in the post central gyrus causes the subject to feel sensations, while stimulation of the precentral gyrus causes motor action. The body is precisely mapped (with distortions in scale) on these regions, so that the response is quite specific. These facts do not shed any light on our problem, however, since they serve only to move the input and output terminals from the receptor and motor organs onto the brain surface. The main question remains, "what happens between the input and the output?" While other localized areas appear to be concerned with specific functions,[13] the localization usually implies a predominance of function rather than an absolute localization. Furthermore it is also true that there is a certain equi-potentiality involved in brain functions, in which the functions of extirpated parts can be taken over by other parts and the loss of function varies as the mass of brain removed.[14] Indeed large sections of the brain can be removed with no apparent permanent loss in function. The search for specific structures performing specific functions has been generally without success. It seems clear that memory and the other higher functions are distributed in the fine structure of the brain. It is not known however to what extent different functions have structural units in common.[15]

Elsasser[16] says, "When the histologist looks at the brain he sees something which is very reminiscent of large electronic computers. He sees a small number of basic components repeated over and over again. All the complexity lies in the innumerable interconnections, not in the variety of basic components. So far as we know, the brain consists exclusively of neurons. Again, so far as we know, a neuron does nothing but conduct electrochemical pulses from its head end to its tail end. Some of the neurons leave the brain (efferent nerves), others enter it (afferent nerves), but apart from this the head and tail ends of neurons make synaptic connections with other neurons. Thus if one is to study the physiological background of memory one might start with such a model of interconnected

---

[4] W. K. Taylor, Proc. Inst. Elec. Engrs., (London), **106**, 198 (1959).

[5] R. L. Grimsdale, F. H. Sumner, C. J. Tunis, and T. Kilburn, Proc. Inst. Elec. Engrs., **106**, Part B (1959).

[6] R. M. Friedberg, IBM J. Research Develop. **2**, 2 (1958).

[7] *Symposium on the Design of Machines to Simulate the Behavior of the Human Brain,* I.R.E. Trans. on Electronic Computers, **EC-5** (1956).

[8] A. Newell, J. C. Shaw, and H. A. Simon, Psych. Rev. **65**, 151 (1958).

[9] George W. Gray, Sci. Am. **179**, 4 (1948).

[10] J. F. Fulton, *Physiology of the Nervous System* (Oxford University Press, New York, 1943).

[11] See, M. Singer, in *Histology,* edited by R. O. Greep (Blakiston Company, New York, 1954).

[12] W. G. Walter, *The Living Brain* (W. W. Norton and Company, Inc., New York, 1953).

[13] W. Penfield and T. Rasmussen, *The Cerebral Cortex of Man* (The MacMillan Company, New York, 1957).

[14] K. S. Lashley, *Brain Mechanisms and Intelligence* (University of Chicago Press, Chicago, Illinois, 1929).

[15] K. S. Lashley, Research Publs., of the Assoc. Research Nervous Mental Diseases, **36**, 1 (1958).

[16] W. M. Elsasser, *The Physical Foundation of Biology* (Pergamon Press, New York, 1958), p. 138.

neurons. We do not claim that this model is altogether true, but it is simple, and presents itself on the basis of anatomical data. *There is no anatomical evidence for a storage organ* used to file away the immense amount of information which every person retains in his memory. Also, *brain physiology has not brought to light any evidence for* the existence of the highly complicated *special scanning devices.*"

4. Neurons

The doctrine expressed in this quotation, that the neurons are the functional units of the brain, is largely due to Ramon y Cajal and is now widely held among neurophysiologists.[17] A considerable amount is known about the action of individual neurons.[18–20] A grossly oversimplified description is as follows. When the cell body of a neuron is sufficiently stimulated, an electrochemical pulse travels from the cell body down along the axon, out along the branches to the end feet which impinge (synapse) on the cell bodies of other neurons, thus tending to stimulate or inhibit those neurons. This general description applies to the "internuncial" neurons. The action of a sensory (afferent) neuron differs from this in that its cell body is excited by an external stimulus (for example light impinging on a retinal receptor cell, or pressure on a special-purpose capsule in the skin). The action of a motor (efferent) neuron differs from the general description in that its end feet terminate in a muscle fiber (or a gland); electrochemical impulses arriving at these end feet tend to activate or inhibit the muscle contraction (or the gland output). There are perhaps 10^8 of these input and output neurons, constituting about 1% of the 10^{10} neurons in the brain.

The speed of conduction of the pulses in the neurons varies from about 5 m/sec in the fine neurons up to about 125 m/sec in the large ones. The time for a pulse to be conducted along the length of the neuron is of the order of 3×10^{-4} sec. The time to cross a synapse is of the order of 10^{-3} sec. After a neuron fires there is an absolute refractory period of the order 10^{-2} sec during which the neuron cannot fire again. There is also a relative refractory period, of increased threshold.

We reiterate that the above description is grossly oversimplified. It does, however, furnish a general idea of the manner in which neurons operate.

5. Organization of Neurons in the Brain

To establish detailed anatomical information regarding the connections of the 10^{10} neurons in the brain presents a formidable laboratory task. By ingenious and painstaking techniques, such as microelectrode stimulation, or degeneration and staining, some information has been obtained.[21–23] Some neurons are long, some are short. Some make contacts with nearby neurons, others wander the length of the brain before contacting another neuron. Some neurons connect with only a few others, others contact thousands. A simplified scheme is shown in Fig. 1. In the words of M. Singer, "Almost any type of connection scheme that can be imagined can be found in the brain." It seems impossible to map the entire topology of the neural network. Moreover, even if we accomplished this, we would then face the disheartening task of analyzing the performance of such a network. Now in a digital computer every connection must be exact or the answer can be entirely wrong. If it were also true for the brain that the misplacement or malfunction of a single connection could completely destroy the function, then we could not hope to understand how the brain operates until we have accomplished the impossible tasks of determining the exact wiring diagram of the neural net and analyzing it. However, it is clearly not true that the connections must be exactly right for the brain to function at all. This is proved by the fact that, although neurons do not regenerate, functions which are temporarily lost after extirpation of sections of the brain, are later recovered. Furthermore it seems unlikely that the genes would carry the information to specify every one of 10^{13} connections. It seems more plausible that only certain parameters of growth are specified and the fine connections are grown in a more or less random manner, subject to these constraints. Thus the detailed connection scheme would be unique to each individual. If it is true that individuals, with connection schemes specified only by certain parameters of growth, function

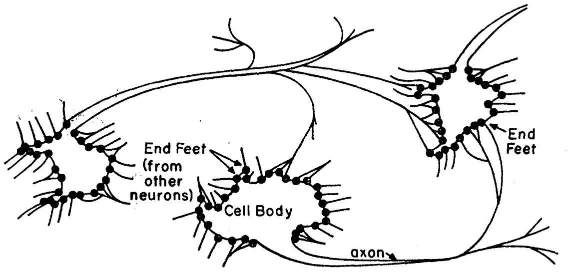


FIG. 1. Neurons (schematic).

---

[17] T. H. Bullock, Science **129**, 997 (1959).

[18] See, Frank Brink, in *Handbook of Experimental Psychology* edited by S. S. Stevens (John Wiley & Sons, Inc., New York, 1951).

[19] J. C. Eccles, *The Physiology of Nerve Cells* (Johns Hopkins Press, Baltimore, Maryland, 1957).

[20] Revs. Modern Phys. (*Biophysical Science*) **31**, 1–598 (1959). See also *Biophysical Science*, edited by J. L. Oncley (John Wiley & Sons, Inc., New York, 1959).

[21] D. A. Sholl, *Organization of the Cerebral Cortex* (Methuen and Company, Ltd., London, 1956).

[22] A. D. Adrian, *The Physical Background of Perception* (Oxford University Press, New York, 1947).

[23] J. C. Eccles, *The Neurophysiological Basis of Mind* (Oxford University Press, New York, 1953).

in similar ways, then there is hope that the performance of such a system might be analyzed in terms of such parameters. This also implies that the operation of the brain is radically different in principle from the logical circuitry of digital computers.[24] The discovery of these principles poses some challenging mathematical questions.

6. The Prospects for a Model

We have seen that our description of brain function is vague and our knowledge of brain structure is very sketchy indeed. It might well be argued that we have not formulated precisely a brain "function-structure" problem at all. Precisely formulated or not, it is clear that the problem exists. Again it is entirely possible that the areas of our ignorance of brain structure and brain function cover items which are essential for an explanation, and, until these are revealed, no understanding is possible. On the other hand, based on the above description of the brain structure, we can, as suggested in the quotation from Elsasser,[16] consider a model of interconnected neurons and study its behavior. Such networks can be arranged to perform any logical function,[25–28] but these arrangements are contrived and do not appear to resemble the biological organization at all. Some early experiments[29–31] were performed on straightforward neural networks, but these systems were very small and simple and the results were difficult to interpret. More plausible, but descriptive models have also been proposed in recent years.[32,33] While the verbal description of the conjectured functioning of such systems is quite attractive, the vagueness in the specifications of these models precluded the possibility of reasonably rigorous analysis or verification. With the Perceptron,[34,35] Rosenblatt offered for the first time a model which was: (a) specified in terms precise enough to permit testing of asserted performance, (b) sufficiently complex to offer the hope that its behavior would be interesting, (c) sufficiently simple to suggest that its performance might be analyzed and predicted, and (d) consistent with the known biological facts. Admittedly the model represents an enormous simplification of even the known brain structure; but if it does not violate the biological constraints (such as the number of units, the organization of connections, the reliability of components, the mechanism of signal transmission, the speed of response, the stability of the performance with respect to component malfunction or extirpation, the capacity for information storage, etc.) and if it exhibits even rudimentary brain functions, then, even if it does not in fact operate in the same manner as the brain does, it still provides at least a possible explanation of how the brain structure, as we know it at this time, *might* be organized to perform these functions.

PERCEPTRONS

7. General Description

The term *Perceptron* refers to a class of theoretical brain models, such as illustrated in Fig. 2.

A stimulus S (for example a pattern of light) is presented to the sensory retina. The illuminated sensory elements send pulses with varying time delays to the associators. Some of the pulses are positive (excitatory) and some are negative (inhibitory). If the algebraic sum of the pulses arriving at an associator in a suitable time interval exceeds a certain threshold (which need not be the same for all associators), that associator sends out pulses as indicated by the arrows to other associators and/or to the response units. Each unit may have its own refractory period. Each connection may have its own transmission time, pulse magnitude and sign, or frequency and phase. The response units also have an activation threshold and may have excitatory and inhibitory connections with some associators and/or each other.

So far we have made no provision for change (learning or memory) in the system. On this the anatomical, histological, or physiological findings offer no clue. The general belief is that pathways through the network are somehow, as a consequence of being used, facilitated for future conduction. Thus, Hebb[32] says, "When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both of the cells such that A's efficiency, as one of the cells firing B, is increased." This might be brought about in the biological system by the growth of additional end feet, or by chemical changes in the neurons such as the production of enzymes in the cell body which alter the threshold in a small region of the cell body, or by several other plausible means.[36] "Long term memory" in humans, which can survive for a century in spite of severe shocks, must be stored in some fairly permanent

---

[24] J. von Neumann, *The Computer and the Brain* (Yale University Press, New Haven, Connecticut, 1958).

[25] J. T. Culbertson, *Consciousness and Behavior* (William C. Brown Company, Dubuque, Iowa, 1950).

[26] W. Pitts and W. S. McCulloch, Bull. Math. Biophys. **9**, 127 (1947).

[27] W. S. McCulloch and W. Pitts, Bull. Math. Biophys. **5**, 115 (1943).

[28] D. A. Sholl and A. M. Uttley, Nature **171**, 387 (1953).

[29] W. A. Clark and B. G. Farley, Proceedings of the Western Joint Computer Conference, p. 86 (1955).

[30] B. G. Farley and W. A. Clark, I.R.E. Trans. Professional Group on Inform. Theory 4, 76 (1954).

[31] N. Rochester, J. H. Holland, L. H. Haibt, and W. L. Duda, I.R.E. Trans. on Inform. Theory, **IT-2**, 80–93 (1956).

[32] D. O. Hebb, *The Organization of Behavior* (John Wiley & Sons, Inc., New York, 1949).

[33] R. L. Beurle, Trans. Roy. Soc. (London), **B240**, 55 (1956).

[34] F. Rosenblatt, Cornell Aeronautical Laboratory Report No. VG-1196-G-1 (January, 1958).

[35] F. Rosenblatt, Psych. Rev. **65**, 386 (1958).

[36] F. Rosenblatt, Cornell Aeronautical Laboratory, Project PARA Technical Memorandum No. 10 (December, 1959).

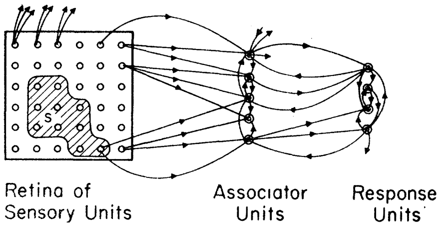


FIG. 2. Organization of Perceptron.

structure; "short-term memory" might be stored by means of a transient state of activity. In any case many of the conjectures are functionally equivalent to the rule that when the two ends of a connection are sequentially active the connection is strengthened, i.e., the pulse it carries is increased. This description of the *reinforcement rule* is intentionally vague; it can be realized in various ways, some of which are given in precise terms below.

Parameters which must be specified to define the perceptron of Fig. 2 are: The number of sensory elements, the number (or probability distribution) of excitatory and inhibitory connections at each level and the geometrical constraints on them, the number of associators and the number of responses; the thresholds, refractory periods, summation intervals, and transmission times. For studying the behavior of such a perceptron we would also have to specify the set of stimulus patterns, the order and times of their presentation, and the observations to be made on the responses. The reinforcement rule must, of course, also be defined.

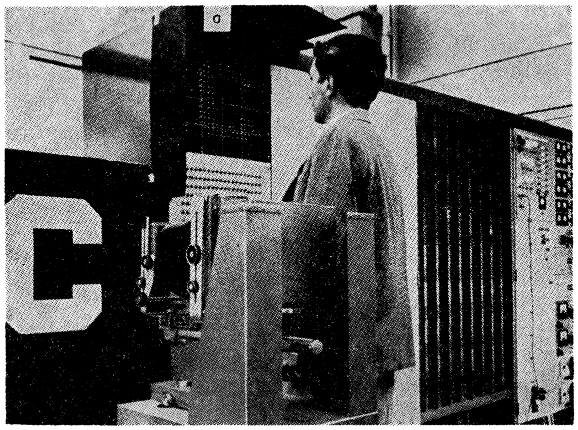
We shall not pursue further here the arguments showing that the above model is consistent with the biological constraints.[1,34]

8. Techniques of Investigation

For studying the behavior of perceptrons, three general techniques are available.

(*a*) *Mathematical analysis.* When it is successful, this approach offers many advantages, such as the predictability of the performance of classes of perceptrons, the effects of variations in the parameters, and so forth. For a model of the complexity of the general perceptron of Fig. 2 the analysis is quite complicated (see Sec. 6 of the paper which follows[37]). For certain simplified cases as in the simple perceptron of Fig. 4 which is discussed later, the analysis is fairly complete. In Sec. 9 we prove some theorems and illustrate the analytical techniques for such systems. In the paper which follows, a more complicated system is analyzed.

(*b*) *Simulation on a digital computer.* The principal advantage of this method is that it can always be done, subject, of course, to time, storage, and cost limitations. A considerable amount of data has been obtained in this way.[1,38] Some of these will be described in Sec. 9 below.



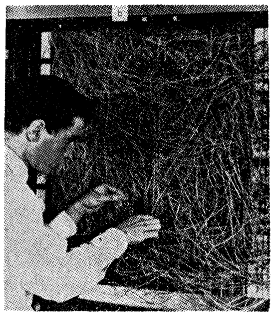


FIG. 3. Mark I Perceptron at Cornell Aeronautical laboratory. (a) Overall view with sensory input at left, association units in center, and control panel and response units at far right. The sensory to associator plugboard, shown in (b) is located behind the closed panel to the right of the operator. The image of the letter "C" on the front panel is a repeater display, for monitoring sensory inputs.

(*c*) *Construction of an actual machine.* This has an enormous advantage in speed over the digital computer, since essentially all the action goes on in parallel simultaneously and the response appears almost immediately, while in the digital simulation all computations are done in sequence. While an actual machine enjoys certain types of flexibility, such as the ease with which the experimenter can vary the stimulus patterns, it is a serious task to change the wiring diagram (in the digital computer this can be generated quickly by a suitable program) and it is impossible to alter certain basic features of the network. There is also the complicating factor of the inexact performance of hardware. A machine of the complexity of Fig. 2 has not yet been built, but one having the organization of Fig. 4 (but with eight binary-response units) has been built, and is known as the Mark I, (Fig. 3).[39–41] The retina is a 20×20 grid of photocells mounted in the picture plane of a camera to which the stimulus pictures are shown. There are 512 associator units and eight binary-response units. Each sensory unit can have up to forty connections to the associator units.

---

[37] H. D. Block, B. W. Knight, Jr., and F. Rosenblatt. Revs. Modern Phys. **34**, 135 (1962).

[38] F. Rosenblatt, Proc. I.R.E. **48**, 301 (1960).

[39] J. C. Hay, F. C. Martin, and C. W. Wightman, Record of I.R.E. 1960 National Convention, Part 2, New York, (1960).

[40] C. W. Wightman, Cornell Aeronautical Laboratory, Project PARA Technical Memorandum No. 4 (February, 1959).

[41] J. C. Hay and A. E. Murray, Cornell Aeronautical Laboratory Report VG-1196-G-5 (February, 1960).

This wiring is normally made according to a table of random numbers. The associator to response connections are varied by motor-driven potentiometers. Many interesting experimental results have been obtained.[39] Some of these are mentioned in Sec. 9(f) below.

For purposes of building a machine having more units and organized more in the direction of Fig. 2, the principal requirement is for inexpensive, compact, low power associator units and connections having the desired variability of weights. Neither precision nor reliability of the components is important. [This is contrasted with the usual engineering situation in which as the number of components increases the reliability of each component must also increase, for the failure of a single element results in the failure of the entire system; the probability of system failure thus increases rapidly as the number of components increases. For the perceptron it is the other way around; cf. the extirpation experiments described in 9(f) below.] Another need is for an inexpensive method of making connections. Again these need not be precise, but there must be a great many of them. Recent developments[42–46] are encouraging and lead to reasonable hope for success.

9. Analysis

a. A Simple Perceptron.

Consider the simple[47] perceptron shown in Fig. 4.

Let there be N_s sensory units, N_a associator units, and n stimulus patterns (each stimulus pattern is a specified set of activated retinal points). We denote typical sensory units by s_σ, typical associators by a_μ, and typical stimuli by S_i. Let us represent the connection between s_σ and a_μ by the real number $C_{\sigma\mu}$; in particular the $C_{\sigma\mu}$ might be random numbers having the possible values $+1$, -1, 0. When the stimulus S_i is applied to the retina, the signal

$$\alpha_\mu{}^i = \sum_{\substack{\sigma \\ s_\sigma \in S_i}} C_{\sigma\mu}$$

is transmitted instantly to the associator a_μ. If $\alpha_\mu{}^i \geqq \theta$, where θ is an arbitrary, but fixed real number, the associator a_μ is said to be *active* and instantly transmits a signal v_μ to the response unit. An inactive associator transmits no signal. The total signal arriving at the response unit is $u = \sum'_\mu v_\mu$, where $\sum'$ is taken over the active associator units. If $u > \Theta$, where Θ is an arbitrary but fixed non-negative number, the response output is $+1$. If $u < -\Theta$ the response is -1. If $|u| \leqq \Theta$ the response is 0.

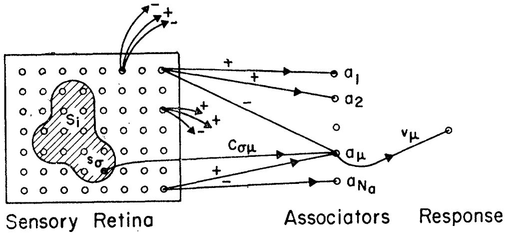


FIG. 4. Simple perceptron.

In this model the connections $C_{\sigma\mu}$ do not change. Therefore $A(S_i)$, the set of associators activated by stimulus S_i, does not change. Thus once the numbers $C_{\sigma\mu}$ have been determined (more will be said about this later) we may disregard the sensory retina altogether and start with the Venn diagram of Fig. 5. Let

$$e_{\mu i} = \begin{cases} 1 & \text{if} \quad a_\mu \in A(S_i) \\ 0 & \text{if} \quad a_\mu \notin A(S_i). \end{cases}$$

The input to the response unit, when stimulus S_i is presented to the retina is then

$$u_i = \sum_\mu v_\mu e_{\mu i}. \tag{1}$$

This model does not use the known biological facts of delay, refractory period, and variability of neurons. We shall see later that these features can be very helpful indeed, but we now show that even without them we get very interesting performance.

b. Discrimination: Learning by Error Correction

Let us assign each stimulus to one of two classes, which we denote by $+1$ and -1. Say stimulus S_i is assigned to class ρ_i, where ρ_i is $+1$ or -1. This dichotomization is then represented by $\rho = (\rho_1, \rho_2, \cdots, \rho_i, \cdots, \rho_n)$. We would like the perceptron, in its terminal state, to give the correct response to each stimulus. From Eq. (1) we see that the response to S_i is correct if and only if

$$\rho_i u_i = \sum_\mu v_\mu e_{\mu i} \rho_i > \Theta. \tag{2}$$

Let B denote the matrix with elements $b_{\mu i} = e_{\mu i} \rho_i$. It may happen that no choice of numbers for y_μ

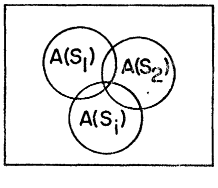


FIG. 5. Associator units.

---

[42] K. R. Shoulders, *Simulation of Neural Networks by Optical-Photographic Methods* (Stanford Research Institute, Menlo Park, California, December, 1959).

[43] K. R. Shoulders, *Research in Microelectronics Using Electron-beam-activated Machining Techniques* (Stanford Research Institute, Menlo Park, California, September, 1960).

[44] B. Widrow, Stanford Electronics Laboratory Technical Report 1553-2, Stanford, California, (1960).

[45] J. K. Hawkins and C. J. Munsey, *A Magnetic Integrator for the Perceptron Program* (Aeroneutronics, Newport Beach, California, 1960).

[46] A. E. Brain, *The Simulation of Neural Elements by Electrical Networks Based on Multi-Aperture Magnetic Cores* (Stanford Research Institute, Menlo Park, California, 1960).

[47] We use the term "simple" here in the colloquial sense. The term is also used with a technical meaning, a precise definition of which is given in Rosenblatt[1]. The Perceptron of Fig. 4 is also "simple" in the technical sense.

($\mu=1,2,\cdots,N_a$) will yield the inequalities

$$\textstyle\sum_\mu y_\mu b_{\mu i} > \Theta, \quad i=1,2,\cdots,n. \tag{3}$$

If this is the case, then the discrimination problem (ρ) cannot possibly be solved (perfectly) by this perceptron, no matter what reinforcing arrangement is used. However, as will be discussed later, for reasonable dichotomies in most systems of interest there will be numbers y_μ satisfying (3) and hence positive numbers ζ_i such that

$$\textstyle\sum_\mu y_\mu b_{\mu i} \equiv \zeta_i > \Theta, \quad i=1,2,\cdots,n. \tag{4}$$

In particular if the Venn diagram (Fig. 5) has, in each set $A(S_i)$, an associator $a_{\mu(i)}$ which is in no other $A(S_j)$, then we can simply set $y_{\mu(i)}=(1+\theta)\rho_i$ for each $i=1,2,\cdots,n$ and the other y_μ's$=0$ and satisfy (4). More generally if the matrix B is of full rank and $N_a \geqq n$ then, by the standard theorem in algebra there will be a solution $y=(y_1,\cdots,y_\mu,\cdots y_{N_a})$ to (4). For the remainder of this section we assume that there is a solution y to (4).

The *error correction procedure* is as follows. A stimulus S_i is shown and the perceptron gives a response. If this response is correct then no reinforcement is made. If the response is incorrect then the v_μ for active associators a_μ is incremented[48,49] by $\eta\rho_i$.[50] The inactive associators are left alone. The initial values are arbitrary, say $(v_1^0,\cdots,v_\mu^0,\cdots,v_{N_a}^0)$.

Suppose the stimuli are shown in an arbitrary sequence, such that each stimulus recurs infinitely often. We shall show that after a certain finite number of steps the machine will thereafter give the correct response to all the stimuli, so that no further changes take place. The proof given here is a distillation of a succession of proofs by Rosenblatt, Joseph, Kesten, and the author.[49,51,52]

(1) Let $\xi=\begin{bmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{bmatrix}$, where the ξ_i are real numbers;

$$F(\xi)=(B\xi,B\xi)\equiv\textstyle\sum_\mu\left(\sum_i b_{\mu i}\xi_i\right)^2$$

and, for $\xi\neq0$, $f(\xi)=F(\xi)/\sum_i\xi_i^2$. Note that $f(\xi)$ is constant along each ray, i.e., $f(\lambda\xi)=f(\xi)$ for any $\lambda\neq0$. Thus all the values that $f(\xi)$ takes on are assumed on the unit sphere $\sum_i\xi_i^2=1$. Since this is a compact set, we have for all $\xi\neq0$, $0\leqq f(\xi)\leqq M$.

Now choose any ξ^* such that $\sum_i\xi_i^{*2}=1$ and each component $\xi_i^*\geqq0$. If $f(\xi^*)=0$ then $B\xi^*=0$ and

$$0=(B\xi^*,y)=(\xi^*,By)=\textstyle\sum_i\xi_i^*\sum_\mu b_{\mu i}y_\mu=\sum_i\xi_i^*\zeta_i>0,$$

by (4), which is a contradiction. Therefore $f(\xi)$ does not vanish on that portion of the unit sphere which lies in the closed first orthant. Hence $f(\xi)$ assumes a positive minimum value on that set, say $m>0$. Since $f(\xi)$ is constant along each ray, it follows that, for any vector $\xi\neq0$ having each component $\xi_i\geqq0$, $f(\xi)\geqq m$.

Therefore we have proved, on the basis of the assumption that there exists a solution to Eq. (4), that there is a constant $m>0$ with the following property. For any nonzero vector ξ having all its components non-negative:

$$0<m\leqq f(\xi)\leqq M.$$

(2) At any time t, let $x_i(t)(i=1,2,\cdots,n)$ be the number of times the machine has incorrectly identified stimulus S_i (and hence has been reinforced, so far, x_i times by amount $\rho_i\eta$). Then $v_\mu(t)=v_\mu^0+\sum_i e_{\mu i}\rho_i\eta x_i(t)$. If the stimulus S_j is now shown to the machine the input to the response unit is

$$u_j(t)=\textstyle\sum_\mu v_\mu^0e_{\mu j}+\sum_\mu\left[\sum_i e_{\mu i}\rho_i\eta x_i(t)e_{\mu j}\right]. \tag{5}$$

Hence

$$\textstyle\sum_\mu\left[\rho_je_{\mu j}\sum_i e_{\mu i}\rho_ix_i(t)\right]\eta=\rho_ju_j(t)-\rho_ju_j^0, \tag{6}$$

where $u_j^0=\sum_\mu v_\mu^0e_{\mu j}$.

Reinforcement occurs at this stage if and only if the response determined by (5) is incorrect, i.e., if $\rho_ju_j(t)\leqq\Theta$. Thus if reinforcement occurs we have, from (6),

$$\textstyle\sum_\mu\left[b_{\mu j}\sum_i b_{\mu i}x_i(t)\right]\leqq(\Theta-\rho_ju_j^0)/\eta\equiv D_j. \tag{7}$$

Suppose that reinforcement takes place. Let us consider the change in $F(\xi)$ as ξ goes from

$$P_1=(x_1,x_2,\cdots,x_j,\cdots,x_n)$$

to

$$P_2=(x_1,x_2,\cdots,x_j+h,\cdots,x_n),$$

for $0\leqq h\leqq1$. We have

$$\left.\frac{\partial F}{\partial\xi_j}\right]_{P_1}=2\textstyle\sum_\mu\left(b_{\mu j}\sum_i b_{\mu i}x_i\right),$$

$$\begin{aligned}\left.\frac{\partial F}{\partial\xi_j}\right]_{P_2}&=2\textstyle\sum_\mu b_{\mu j}\sum_i b_{\mu i}(x_i+h\delta_{ji})\\&=2\textstyle\sum_\mu b_{\mu j}\sum_i b_{\mu i}x_i+2h\sum_\mu b_{\mu j}^2,\end{aligned}$$

---

[48] R. D. Joseph, Cornell Aeronautical Laboratory Report No. VG-1196-G-7. See also Ph.D. Thesis, Cornell University, Ithaca, New York, 1961.

[49] F. Rosenblatt, Cornell Aeronautical Laboratory Report No. VG-1196-G-4 (February, 1960).

[50] Other schemes for the amount of reinforcement have been investigated.[1,48] In particular, the system in which the inactive connections suffer a decrement in such a way that the sum of the input connections at each response unit remains constant is called a "γ system" in contrast to the "α system" described in the text. The "Γ system" on the other hand, conserves the sum of the output connections at each associator. Another system which has been studied is the "λ system" in which the v_μ cannot exceed a certain bound. Another modification, in which each connection v_μ suffers a decay proportional to v_μ will be used in the paper which follows. In the interest of simplicity of presentation we do not go into any of these here.

[51] R. D. Joseph, Cornell Aeronautical Laboratory, Project PARA Technical Memorandum No. 12 (May, 1960).

[52] R. D. Joseph, Cornell Aeronautical Laboratory, Project PARA Technical Memorandum No. 13 (July, 1960).

and in the interval from P_1 to P_2 we have, using (7)

$$\frac{\partial F}{\partial \xi_j} \leqq 2 \sum_\mu b_{\mu j} \sum_i b_{\mu i} x_i + 2hN_j \leqq 2D_j + 2hN_j,$$

where N_j is the number of associators activated by S_j. Therefore the change in $F(\xi)$ as ξ_j is varied from x_j to x_j+1 is

$$\Delta F = \int_{x_j}^{x_j+1} \frac{\partial F}{\partial \xi_j} \partial \xi_j \leqq 2 \int_0^1 (D_j + N_j h) dh$$
$$\leqq 2D_j + N_j \leqq \max_j (2D_j + N_j) \equiv D. \quad (8)$$

The total change in F from the beginning of the training, $x^0 = (0,0,\cdots,0)$ to the state where

$$x = (x_1, x_2, \cdots, x_i, \cdots, x_n)$$

is

$$F(x) - F(0) = F(x) = \sum \Delta F \leqq D \sum_i x_i.$$

Hence

$$0 < m \leqq f(x) = \frac{F(x)}{\sum_i x_i^2} \leqq \frac{D \sum_i x_i}{\sum_i x_i^2} \leqq \frac{Dn}{\sum_i x_i},$$

where the last inequality follows from Schwartz's inequality:

$$n \sum_i x_i^2 = \sum_i 1^2 \sum_i x_i^2 \geqq (\sum_i x_i)^2.$$

Therefore

$$\sum_i x_i \leqq (n/m\eta) \max_i [2(\Theta - \rho_i u_i^0) + \eta N_i]. \quad (9)$$

After at most this many corrections there will be no more; i.e., the machine will thereafter give only correct responses. It has learned the dichotomy.

The above can be generalized so that instead of the corrections all having the same magnitude η each time, they have magnitudes $h_1, h_2, \cdots, h_\nu, \cdots$, where the h_ν are bounded and

$$\sum_{\nu=1}^{\infty} h_\nu$$

diverges. The analysis is analogous, with $\eta x_i(t)$ in Eq. (5) replaced by $X_i(t)$, the absolute magnitude of reinforcement applied up to time t as a result of incorrect responses to stimulus S_i. Analogous to (8) we get

$$\Delta F \leqq 2(\Theta - \rho_j u_j^0) h_\nu + h_\nu^2 N_j \leqq (C + Nh) h_\nu,$$

where $C = \max_j 2(\Theta - \rho_j u_j^0)$, $N = \max_j N_j$, $h =$ the maximum h_ν used to date. Then, summing as before, we get after K corrections

$$m \leqq \frac{F(X)}{\sum_i X_i^2} \leqq \frac{(C+Nh) \sum h_\nu}{\sum_i X_i^2} \leqq \frac{(C+Nh) n \sum h_\nu}{(\sum_i X_i)^2} = \frac{(C+Nh) n}{\sum_{\nu=1}^{K} h_\nu}.$$

Hence

$$\sum_{\nu=1}^{K} h_\nu \leqq \frac{(C+Nh) n}{m}. \quad (9')$$

Thus the process terminates if

$$\sum_{\nu=1}^{\infty} h_\nu$$

diverges and the h_ν are bounded. This last condition can clearly be weakened to the condition that for arbitrarily large values of K:

$$\frac{\underset{(\nu=1,\cdots,K)}{\max(h_\nu)}}{\sum_{\nu=1}^{K} h_\nu} \leq \frac{rm}{Nn} \quad \text{where } r<1;$$

for then we get from (9′)

$$\sum_{\nu=1}^{K} h_\nu \leqq \frac{C}{m(1-r)}.$$

Various modifications and generalizations of the fundamental theorem expressed in Eq. (9) have been obtained.[49,51,52] We confine ourselves here to the following remark.

The condition for the existence of a solution is, from (3):

$$\operatorname{sgn} \sum_\mu e_{\mu i} y_\mu = \rho_i.$$

Hence, with a fixed Venn diagram, the number of dichotomies which the machine will be able to learn is equal to the number of orthants (in n space) which can be entered by linear combinations of the N_a row vectors of the matrix $e_{\mu i}$. An upper bound for this number[53] shows that it may be considerably less than the 2^n possible dichotomizations. However most of these dichotomizations are "unreasonable" and could not be learned by humans either. As an example of the power of the machine, the Mark I was shown the twenty horizontal bars (4×20 retinal units) and the twenty vertical bars (20×4 retinal units) with the dichotomy being that alternate bars were in opposite classes. This should be a difficult dichotomy, since bars with the greatest overlap are in opposite classes (compare the discussion below). The Mark I learned this dichotomy perfectly after seeing 214 stimuli, requiring 600 sec of reinforcement in all. (In this experiment the reinforcement rule was to hold on the reinforcement until the sign of the response changed).[49] It would be interesting to compare this performance with that of a human subject on the same problem.

---

[53] R. D. Joseph and L. Hay, Cornell Aeronautical Laboratory, Project PARA Technical Memorandum No. 8 (1960).

c. Forced Learning.

In the *forced-learning* reinforcement rule the v_μ of each active associator is incremented by $\eta\rho_i$ each time S_i is shown, regardless of the machine's response. The input to the response unit when stimulus S_j is presented to the retina is again given by

$$u_j(t)=\mu_j{}^0+\eta\sum_\mu e_{\mu j}\sum_i e_{\mu i}\rho_i x_i(t), \tag{10}$$

where $x_i(t)$ now denotes the number of times the stimulus S_i has been shown, up to time t. Suppose that at some time T, $x_i(T)=p_iT$, for $i=1, 2, \cdots, n$. The vector

$$p=\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_i \\ \vdots \\ p_n \end{bmatrix}, \quad p_i\geqq 0, \quad \sum_i p_i=1,$$

corresponds to the relative frequencies of occurrence of the various stimuli

$$\begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_i \\ \vdots \\ S_n \end{bmatrix}.$$

The condition for the response to S_j to be correct is $\rho_j u_j>\Theta$, i.e.,

$$\rho_j u_j{}^0+\eta T\sum_\mu \rho_j e_{\mu j}\sum_i e_{\mu i}\rho_i p_i>\Theta. \tag{11}$$

Inequality (11) will hold in general for large T, if and only if

$$\sum_\mu \rho_j e_{\mu j}\sum_i e_{\mu i}\rho_i p_i>0. \tag{12}$$

Thus, if the perceptron can learn this dichotomy under forced learning, for some choice of p, we can take $y_\mu=\sum_i e_{\mu i}\rho_i p_i$ to satisfy (3). Therefore it will learn under error correction. Conversely if the machine can learn the dichotomy under error correction then from (5) starting with zero initial values, the x_i obtained by the error correction procedure will yield a frequency vector

$$p_i=x_i/\sum_j x_j$$

satisfying (12). To summarize: If the perceptron can learn a dichotomy under forced learning, it will learn it under error correction. If it can learn it under error correction it can learn it under forced learning for some choices of the frequency vector, but not in general (see below) for others. Thus the error correction method is a more generally effective method than forced learning.

Let $\mathfrak{N}$ denote the matrix whose elements are $n_{ij}=n_{ji}=\sum_\mu e_{\mu j}e_{\mu i}$. Note that n_{ij} is the number of associators activated by both S_i and S_j; i.e., the number of elements in $A(S_i)\cap A(S_j)$. The matrix B^TB has for its i–jth element $\rho_i n_{ij}\rho_j$. The desired dichotomy having been specified, we can relabel the stimuli without loss of generality, so that all those in the $+1$ category come before all those in the -1 category. Then the matrix B^TB has the appearance

$$U=\left[\begin{array}{ccc|cc} n_{11} & n_{12} & n_{13} & -n_{14} & -n_{15} \\ n_{21} & n_{22} & n_{23} & -n_{24} & -n_{25} \\ n_{31} & n_{32} & n_{33} & -n_{34} & -n_{35} \\ \hline -n_{41} & -n_{42} & -n_{43} & n_{44} & n_{45} \\ -n_{51} & -n_{52} & -n_{53} & n_{54} & n_{55} \end{array}\right], \qquad \text{where} \quad U=B^TB,$$

and where we have assumed, for purposes of illustration that there are three stimuli in the first class and two in the second. Clearly if there is any negative entry in U then there is some frequency vector p for which (12), which now reads

$$\sum_i U_{ji}p_i>0, \tag{13}$$

will fail to hold for some j. If the stimuli are "equally likely" to occur, ($p_i=1/n$) then (13) is the requirement that each row sum of U is positive; or roughly, that each stimulus has a greater intersection (in the associator set) with its own class than with the opposite class.

If the stimuli are presented in random order, with p_i the probability of occurrence of S_i, then, in (10), $x_i(t)$ and $u_i(t)$ are random variables, with the expected value of $u_j(T)$ again given by the left side of (11). Using Tchebycheff's inequality, rigorous bounds on the probability of error as a function of time can be obtained.[48] By using a normal approximation, an estimate of the learning curves (Probability of correct response vs t) have been found.[48] The success of the perceptron at this type of learning has led to its being applied as an engineering pattern-recognizing device.

The entries of the $\mathfrak{N}$ matrix, for a given set of stimuli, stem from the connections $C_{\sigma\mu}$ and the threshold θ. By fixing these suitably we could "gimmick" the $\mathfrak{N}$ matrix to fit our needs. [For example by taking $C_{\sigma1}=1$ for those sensory units s_σ which are activated by S_1; $C_{\sigma1}=-1$ if s_σ is not activated by S_1; and setting $\theta_1=\frac{1}{2}\sum_\sigma(C_{\sigma1}+|C_{\sigma1}|)$ we can be sure that associator a_1 responds to stimulus S_1 and only to S_1. If we now want an associator a_μ to respond to S_1 and to S_2 and to no other stimulus we could use a preliminary layer of associators and the connection scheme (Fig. 6), where a_2 responds only to S_2 and $\theta_\mu=1$. Similarly we can deal with the other logical functions. Thus in fact, using the preliminary layer of n associators, we can fix the number of elements in each subset of the Venn

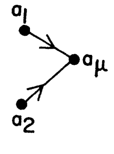


FIG. 6. Associators and connection scheme.

diagram (Fig. 5), and thus control the $\mathfrak{N}$ matrix.] However we take the position that we want the perceptron to react equally well when the set of stimuli $S_1, \cdots, S_n$ are not specified in advance. Thus an $\mathfrak{N}$ matrix that leads to interesting results for a preassigned set of stimuli may be poor for a different set. For this reason the $C_{\sigma\mu}$ are taken at random. The resulting entries in the $\mathfrak{N}$ matrix become random variables which have been studied[1] and tabulated in some detail.[54] From these data it is possible to select design parameters and to predict performance of perceptrons constructed in this way.[55] These results have been verified and elaborated by simulation testing of Perceptrons[1,56] as well as by experiments on Mark I.[39] (Actually, by the use of a reinforcement scheme applied to the $C_{\sigma\mu}$ analogous to that used on the v_μ, the perceptron can improve its sensory connection scheme for a given problem.[1] We shall not go into the analysis of such systems here).

d. Generalization.

The simple perceptron we are here analyzing exhibits a good deal of generalization, i.e., correct response to a stimulus it has not seen before. Examination of inequality (13) reveals the reason for this. If the stimulus has considerable overlap with some of the members of the same class and very little overlap with members of the opposite class it will give the correct response on the basis of having seen the like stimuli. To illustrate with an extreme example suppose the $\mathfrak{N}$ matrix has the following appearance

$$\left[\begin{array}{ccccc|ccccc}
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1
\end{array}\right] \tag{14}$$

In the matrix each stimulus overlaps (in the associator set) with its two nearest neighbors but with no others. (The retina in this case is conceived as being cylindrical; patterns which go off the right edge reappear at the corresponding point at the left edge. In this way each stimulus has exactly two nearest neighbors.) Then, starting with zero initial values, once the machine sees one stimulus and is reinforced it will give the correct response to its two neighboring stimuli even if it has not been shown these previously. While this type of $\mathfrak{N}$ matrix is a gross oversimplification of those obtained in practice, it illustrates a basic mechanism which is operating here as we show next.

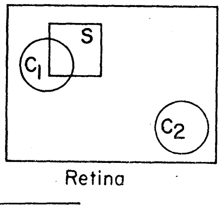


FIG. 7. Stimulus pattern of circles and squares.

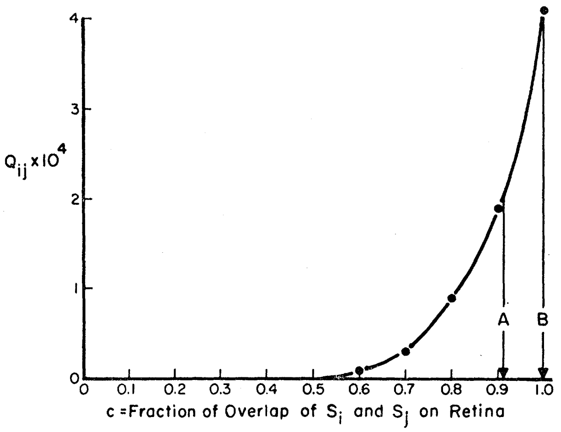


FIG. 8. Q_{ij} as a function of c.

Suppose that we are interested in discriminating squares from circles, all of unit area. In Fig. 7 it is clear that the circle C_1 in the upper left of the retina has many more sensory points in common with the square S than it has with the other circle C_2. Why should the machine tend to classify the first circle C_1 with the second circle C_2 rather than with the square S, with which it has the greater retinal overlap?

To answer this, let us use some typical data. Suppose the connections $C_{\sigma\mu}$ are made by taking at each associator 5 inhibitory and 5 excitatory inputs connected randomly to the retina. Further suppose that $\theta = 5$ and the total retina is of area 10/3. Then the probability Q_{ij} of the associator being activated by both of two stimuli S_i and S_j is a function of the retinal overlap c of the stimuli, S_i and S_j as shown in Fig. 8 (the data for Fig. 8 are taken from the tables[54] mentioned earlier).

With an extremely fine lattice spacing on the retina the overlap between any one of the circles and any one of the squares is less than 91%, as can be verified by elementary geometry. Thus for a circle and a square we are always to the left of point A in Fig. 8, and the probability of an associator being activated by both is less than 2×10^{-4}. If the circles can be displaced by very small amounts we will have close to 100% retinal overlap between a circle and a nearby circle; consequently we are near point B in Fig. 8 with the probability of the associator being activated by both circles greater than 4×10^{-4}. If there are, say, 3000

---

[54] F. Rosenblatt, Cornell Aeronautical Laboratory Report No. VG-1196-G-6 (May, 1960).

[55] R. D. Joseph, I.R.E. 1960 Convention Record, 2, New York, (1960).

[56] F. Rosenblatt, Proc. I.R.E. 48, 301 (1960).

associators, then the expected number of associators activated by any specified circle and square is less than 0.6. On the other hand, each circle has neighbors for which the expected number of associators activated by it and a specified neighbor is greater than 1.2. Therefore it is plausible that the matrix $\mathfrak{N}$ should have an appearance not unlike that indicated in Eq. (14) and the possibility for generalization is present. We reiterate that this is a gross oversimplification of what happens in the actual examples that have been dealt with in practice, but it illustrates a mechanism which makes some dichotomies "natural" and others "unnatural."

The effect operating here depends on the maximal overlap between stimuli of opposite classes being less than the maximal overlap between members of the same class. This effect can be sharpened, for example, by replacing the solid stimulus figures by their boundaries (with some small width). For example the ratio of the amount of overlap between any circular ring and a square ring to the overlap between neighboring circular rings clearly tends to zero with the ring thickness, instead of the 91% of the solid figures in the above example. The process of "contour extraction" can be realized by simple neuronal circuitry. For example if the output of each sensory unit is nullified by the stimulation of its four nearest neighbors then the figure will be replaced by its boundary. Another arrangement is to take account of a relatively long absolute refractory period and small rapid random motions of the retina, such as the human eye makes. After a short instant, the figure will again be replaced by its contour. Other similar pattern property filters[25,57] can be used to organize information at the sensory level, resulting in enhancing the effect under discussion.

The performance of the perceptron can also be improved by modifying the two-valued nature of the output of the associators by taking into account the magnitude of the input in excess of the threshold.[48,58] We shall not go into this here.

The above discussion is concerned with "perfect" or 100% performance. The perceptron will emit a response in any case and, under much more general circumstances, the performance will be "better than chance." By putting several "better than chance" machines in parallel and using, say, a "majority decision" rule for the final response, considerable improvements in performance are obtained. Often, however, better performance results from combining all the associators of the parallel systems into a larger single set of associators. "Learning curves," giving the probability of a correct response as a function of training time, have been obtained by analysis as indicated above and also by simulation and by Mark I experiments.[1,48] The results justified the approximations of the analysis as well as indicating the ability of the machine.

The simple perceptron discussed here generalizes on the basis of retinal overlap. The four layer system which will be described and analyzed in the paper which follows[37] generalizes also on the basis of temporal contiguity.

e. *Spontaneous Organization.*

The *spontaneous-organization* program consists of showing the machine stimuli, letting it compute its own response, and reinforce in accordance with that response. The only contact between the experimenter and the machine is the presentation of the stimuli. Although it is true that for certain special cases and with certain modifications[38] the simple perceptrons described here do make interesting dichotomies, they do not do so in general.[56,59,60] The four-layer systems described in the next paper (and more generally the cross coupled systems[1]) do make interesting spontaneous classifications, as will be shown there. Since it is the spontaneous classification that corresponds to the machine having an "original concept," we see that we shall have to look to the paper which follows for this.

f. *Psychological Testing.*

While it is possible on the basis of the analysis indicated to estimate the performance of simple perceptrons, experimentation with Mark I has suggested various "psychological" experiments on that machine. We cite some of these results.[39] Figure 9 compares the "forced learning" with the "error-correction" procedure. The same perceptron is "trained" by each method and "tested" at various times. During the testing, of course, no reinforcement is applied.

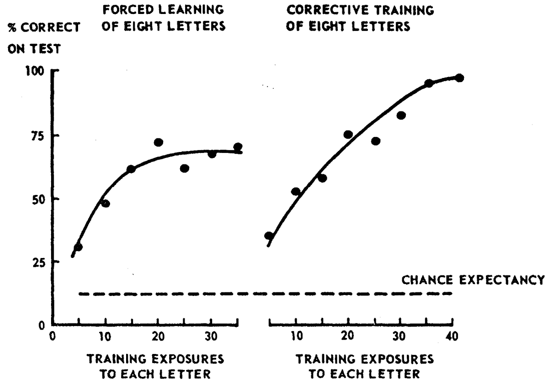


FIG. 9. Learning curves for eight letter identification task (each letter upright, but in a variety of locations).

---

[57] M. Babcock, A. Inselberg, L. Löfgren, H. von Foerster, P. Weston, and G. Zopf, Tech. Rept. No. 2, University of Illinois, 1960.

[58] R. D. Joseph, Cornell Aeronautical Laboratory, Project PARA Technical Memorandum No. 11 (March, 1960).

[59] See, F. Rosenblatt, in *The Mechanization of Thought Processes* (Her Majesty's Stationery Office, London, England, 1959).

[60] F. Rosenblatt, Cornell Aeronautical Laboratory, Project PARA Technical Memorandum No. 2 (October, 1958).

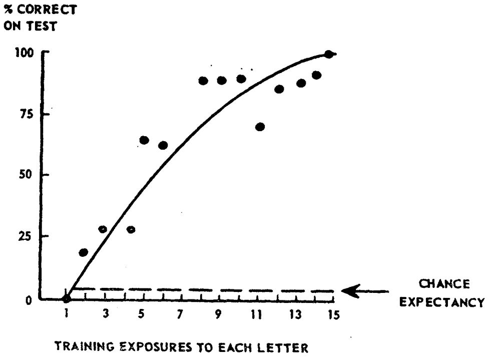


FIG. 10. Learning curve for 26 letters, each in standard position; corrective training.

Figure 10 shows a learning curve for recognition of all 26 letters of the alphabet in one standard position. For this experiment the coding of the outputs of five binary-response units was selected in a quasi-optimal manner.

To study the effects of random noise in the stimulus pattern, the target stimuli were taken as the letters E and X, with a small amount of retinal shift allowed. Noise in the target display is illustrated in Fig. 11. The effect of the noise on performance is illustrated in Figs. 12(a) and 12(b). An additional disturbance, in which the trainer makes errors (when deciding what reinforcement to apply, he misidentifies the stimuli) at random 30% of the time, is introduced in Figs. 12(c) and 12(d). In Fig. 12(c) the perceptron does better than the trainer and in fact rises to the same level of performance as in Fig. 12(a).

In Fig. 13 damage to the machine is simulated by removing association units at random from a perceptron already trained on E–X discrimination. The decline in performance is gradual rather than sudden and varies with the amount removed. Since the memory does not operate by comparison with a stored file of patterns, but rather is distributed throughout the structure we get behavior here analogous to Lashleys' law of "Equipotentiality and Mass Action."

More advanced perceptual problems, such as figure ground determination, relations among objects in complex fields ('the square is inside the circle,' 'the tree is behind the dog') and so on, are beyond the capacity of the simple perceptron considered here, but it seems possible that the richer models will be able to perform these functions.[1]

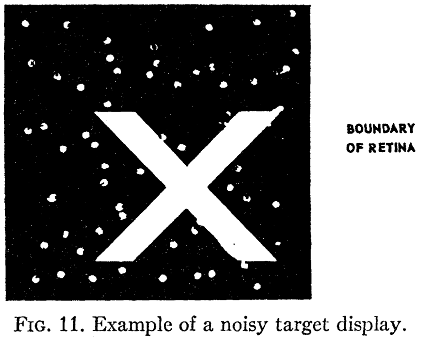


FIG. 11. Example of a noisy target display.

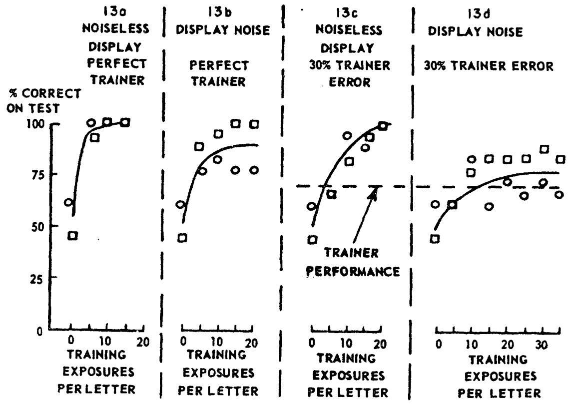


FIG. 12. Effects of noisy display and imperfect trainer on learning of "E"–"X" discrimination.

g. Stimulus Modalities.

We have described the stimuli as visual patterns on a retina. Other interpretations of the input patterns are equally possible.

If, for example, each sensory unit has as its source of activation, the output of some "property filter,"[57] then a "stimulus pattern" on the retina represents a listing of the presence or absence of the various properties. Similarly as far as the logic and functioning of the simple perceptron are concerned the stimulus patterns could represent, for example, the magnitudes of the Fourier components of a sound wave or the combinations of taste sensors activated by particular food preparations. The analysis would be similar to that given above. Indeed a simple perceptron had considerable success in selecting medical diagnoses where the sensory input was the patient's coded clinical signs and symptoms.[61]

By putting several perceptrons in parallel one may, with cross connections, obtain *conditioned reflexes*, *association* between stimulus patterns of different modalities and so on.[1]

The perceptron might also be used as the perceptual input to the first stage of a "heuristic logic" machine.[8]

10. Current Research

In all of the problems discussed so far, we have been concerned with an *instantaneous* pattern. If on the other hand our interest is in the properties of a temporal sequence such as is involved, e.g., in speech recognition or the sequence of nerve impulses being fed back to the brain as a muscle movement is performed, then the

---

[61] A. E. Murray, Cornell Aeronautical Laboratory Report VE-1446-G-1 (November, 1960).

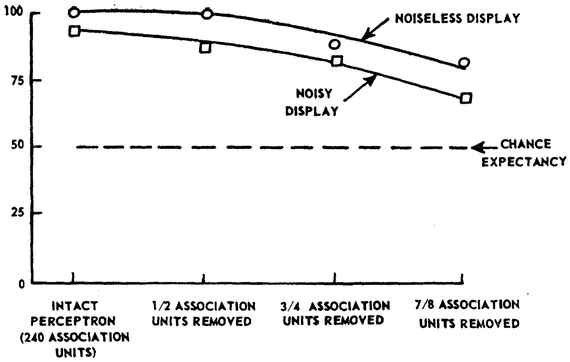


FIG. 13. Effect of association unit removal on trained "E"–"X" discrimination.

simple perceptron of Fig. 4 is no longer adequate. We shall find certain temporal effects in the paper which follows, but for others it is necessary to introduce time delays into the system.[1] A speech recognizing perceptron which utilizes such delays is currently being built at Cornell University.

Other activities now in progress[1] include quantitative studies of cross-coupled and multi-layer systems (by means of analysis and digital simulation), studies of selective attention mechanisms, the effects of geometric constraints on network organization, new types of reinforcement rules, and attempts at relating this research to biological data. Work is also in progress on development of electrolytic and other low-cost integrating devices and additional electronic components necessary for the construction of large-scale physical models.

It is clear that we are still far from the point of understanding how the brain functions. It is equally clear, we believe, that a promising road is open for further investigation.

REVIEWS OF MODERN PHYSICS VOLUME 34, NUMBER 1 JANUARY, 1962

Analysis of a Four-Layer Series-Coupled Perceptron. II*

H. D. BLOCK, B. W. KNIGHT, JR., AND F. ROSENBLATT

Cornell University, Ithaca, New York

1. INTRODUCTION

THE preceding paper[1] presented motivation and background for the general subject of perceptrons and gave some analysis and results for a simple three-layer perceptron. While it has been shown there that it is possible to associate any arbitrary set of responses to an arbitrary set of stimuli in a simple three-layer perceptron, such a perceptron characteristically requires a large representative sample of each kind of pattern (e.g., letters "A" and "B"), covering all parts of the retina, before it will recognize an arbitrarily positioned stimulus which is similar to one which it has seen before. In other words, a three-layer perceptron has no concept of "similarity" based on any criterion other than the intersections of sets of retinal elements. In a previous paper,[2] Rosenblatt has shown that a "cross-coupled perceptron," in which A units are connected to one another by modifiable connections, should tend to develop an improved similarity criterion for generalizing responses from one stimulus to another when exposed to a suitably organized environment. In this paper a simpler network, consisting of four layers of units but without cross coupling, is analyzed in a more rigorous fashion, and is shown to possess the same property.

The perceptron of the present paper is "self-organizing" in the sense that during the training period the experimenter does not tell the machine the category of each stimulus. As the analysis below will show, the only contact between the experimenter and the machine is the presentation of the stimuli.

2. THE MODEL

The model to be analyzed here is a four-layer perceptron of the schematic type S–A^{I}–A^{II}–R, as indicated in Fig. 1.

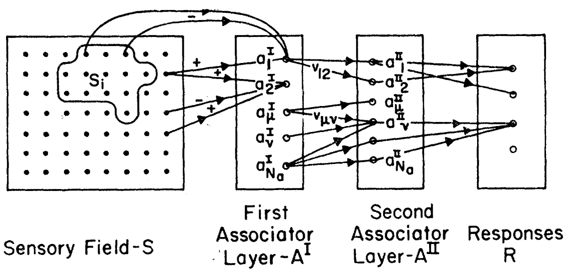


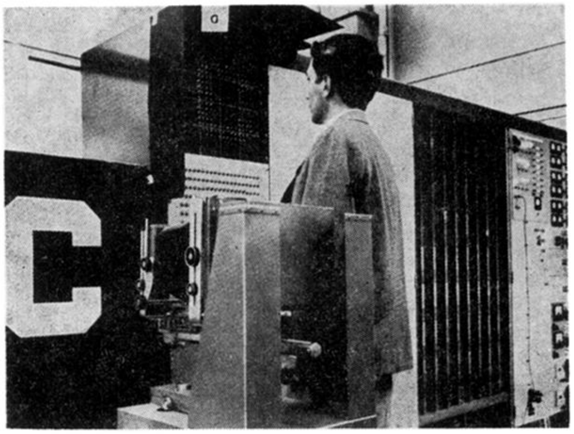
FIG. 1. Organization of four-layer series-coupled perceptron.

---

* Research sponsored by the Office of Naval Research.

[1] H. D. Block, Revs. Modern Phys. 34, 123 (1962).

[2] See, F. Rosenblatt, in *Self-Organizing Systems*, edited by M. Yovits and S. Cameron (Pergamon Press, New York, 1960).



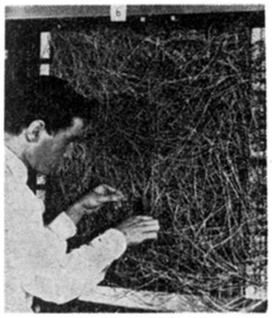


FIG. 3. Mark I Perceptron at Cornell Aeronautical labortory. (a) Overall view with sensory input at left, association units in center, and control panel and response units at far right. The sensory to associator plugboard, shown in (b) is located behind the closed panel to the right of the operator. The image of the letter "C" on the front panel is a repeater display, for monitoring sensory inputs.