

Perceptron Simulation Experiments*

FRANK ROSENBLATT†

Summary—An experimental simulation program, which has been in progress at the Cornell Aeronautical Laboratory since 1957, is described. This program uses the IBM 704 computer to simulate perceptual learning, recognition, and spontaneous classification of visual stimuli in the perceptron, a theoretical brain model which has been described elsewhere. The paper includes a brief review of the organization of simple perceptrons, and theoretically predicted performance curves are compared with those obtained from the simulation programs, in several types of experiments, designed to study “forced” and “spontaneous” learning of pattern discriminations.

INTRODUCTION

A NUMBER of papers and reports have been published describing the theory of a new brain model called the perceptron. The perceptron is a minimally constrained “nerve net” consisting of logically simplified neural elements, which has been shown to be capable of learning to discriminate and to recognize perceptual patterns [5]–[8]. This paper is concerned with a report of digital simulation experiments which have been carried out on the perceptron, using the IBM 704 computer at the Cornell Aeronautical Laboratory. These experiments are intended to demonstrate the performance of particular systems in typical environmental situations, free from any approximations which have been used in the previously published mathematical analyses. In the simulation programs, the action of every cell and every connection in the network is represented in detail, and visual stimuli are represented by dot patterns corresponding to illuminated points in a retinal mosaic.

Several related experiments have been conducted previously, using a digital computer for the simulation of a nerve net in learning experiments [1], [2], [4]. Rochester and associates, at IBM, have reported on several attempts to simulate the formation of “cell assemblies,” in a model based on the work of Hebb [3]. Hebb proposes that a set of neurons which is repeatedly activated by a particular sensory stimulus becomes organized into a functional unit, which can be triggered as a whole by sensory patterns sufficiently similar to the original one. Hebb’s book, however, does not attempt to specify in a rigorous manner the exact organization or parameters under which the predicted effects would be obtained, so that the IBM group found it necessary to improvise several models and variations of their own, having various degrees of biological plausibility, in an attempt to construct a definite system. The results of these experiments seem ambiguous, not only because

of the uncertain relationship of the final model to the nerve net originally suggested, but also because the phenomenon which was sought after has never been defined in a fashion precise enough so that one might say whether or not it has actually occurred. These experiments illustrate the importance of selecting a suitable measure of performance in work of this type; it is essential that a clearly defined test should be specified for the “learning” which has presumably taken place, or else it is impossible to say either how well a particular system has performed or to compare its performance with any other system, or class of systems, in a systematic fashion.

From this standpoint, the experiments reported by Farley and Clark [1], [2] seem to have been better conceived. In this model, a network of eight randomly connected neurons was simulated. Inputs consisted of stimuli applied to one of two disjunct pairs of “input cells,” and outputs were measured as the activity of two pairs of “output cells.” In later experiments, the size of the network was increased to sixteen cells. It was demonstrated that this system can learn to favor the output from one set of output cells following the presentation of one of the two stimuli, and the alternative output set following presentation of the other stimulus. The problem of generalization was considered only in terms of relatively slight displacements or alterations of the stimulus patterns, and it was suggested that, under these conditions, the response would be most likely to occur which was previously associated to the stimulus having the greatest overlap with the altered stimulus. The problem of generalization to similar but completely disjunct stimuli was not specifically considered. Nonetheless, the process of generalization advocated as a result of these experiments has much in common with our early work on the perceptron. A more thorough consideration of this problem will be published elsewhere [8].

The design of a simulation program for studies of pattern recognition and perceptual generalization in nerve nets should fulfill at least three basic conditions, each of which has been ignored too frequently in previous work along these lines.

- 1) Simulation should not, in general, be attempted without a theoretical analysis of the nerve net in question, sufficient to indicate suitable parameters and rules of organization, and to indicate questions of theoretical interest. The examination of arbitrary networks in the hope that they will yield something interesting, or the simulation of networks which have been specially designed to compute a particular function by a definite algorithmic procedure seem to be about equally lacking in value.

* Original manuscript received by the IRE, July 16, 1959; revised manuscript received, December 14, 1959. This paper was presented at the 1959 IRE National Convention. The work was supported by the Information Systems Branch, Office of Naval Research, under Contract no. Nonr-2831(00), since July, 1957.

† Cornell Aeronautical Lab., Inc., Buffalo, N. Y.

2) Suitable measures of performance must be defined. This means that some task must be set for the system, the outcome of which can be clearly recognized, and, preferably, counted or quantified in some manner. Signal strengths, waiting times for achievement of a criterion, or percentage of correct decisions are examples of suitable measures.

3) Experiments should be designed with suitable controls against trivial or ambiguous results. If we are interested in teaching a device to generalize a response to visual forms, for example, it is essential that a discrimination test should be made involving at least two different responses, to make sure that the system has not simply generalized the desired response universally to all stimuli, regardless of their similarity to one another. Moreover, it is often important to make sure that the cue for the response is the actual *form* of the stimulus, rather than its location on the retina, or some other unintentional source of information. This last condition is often quite tricky to satisfy, and in most of our current work we make use of Born-von Kármán boundary conditions (in which patterns shifted off of one edge of a retinal field re-enter on the opposite side, as in a toroidally connected space) in order to guarantee the logical equivalence of all points in the retinal space. Given such a retinal field, it is sufficient to place each stimulus pattern with equal probability or frequency at all possible locations in the retinal space, in order to guarantee that the illumination of a particular retinal point does not convey any information about which stimulus is present. It should be noted that this condition is not always observed in the experiments reported in this paper, stimuli often being confined to some sub-field of the retina in order to increase the rate of learning. In at least one case (the experiment with the "continuous transducer perceptron" shown in Fig. 8) a discrimination has thus been obtained which would not hold up if the field were uniformly covered with the stimulus patterns.

ORGANIZATION OF A PERCEPTRON

Any perceptron, or nerve net, consists of a network of "cells," or signal generating units, and connections between them. The perceptron is defined by two sets of rules: 1) a set of rules specifying the topological constraints upon the network organization, such as the number of connections to a given unit, or the direction in which connections are made, and 2) a set of rules specifying the dynamic properties of the system, such as thresholds, signal strengths, and memory functions. A "fully random network" would be one in which only the number of cells and the number of connections is specified, each connection being equally likely to originate or terminate on any cell of the system. The topological rules for the organization of a perceptron take the form of constraints applied to such a random network, and it is assumed that all connection properties

other than those specified remain "random," in the sense just indicated.

A simplified version of the known features of a mammalian visual system is shown in Fig. 1, for a comparison with the organization of a perceptron, which will be described presently. At the extreme left we see a mosaic of light-sensitive points, or retina, from which signals are transmitted to the visual projection area, in the cerebral cortex. Several intermediate relay stations exist in a typical biological system, which are not shown here. These connections preserve topological characteristics of the stimulus in a reasonably intact form. Beyond the projection area, however, connections appear to be largely random. Impulses are delivered through a large number of paths to the association areas of the cortex, where local feedback loops are activated, so that activity may persist for some time past the termination of the original visual stimulus. From the association area, signals are transmitted to the motor cortex, which again has a clear topological organization corresponding to the location of muscle groups to be controlled.

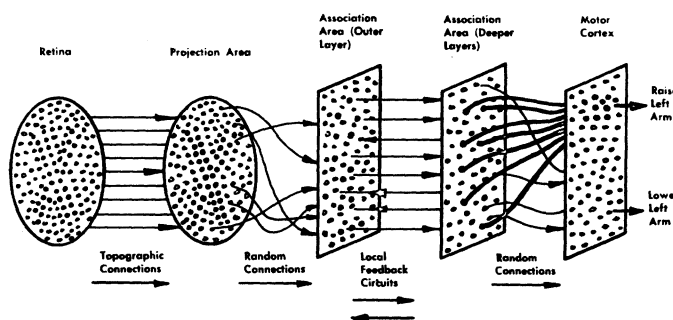


Fig. 1—Organization of a biological brain. (Heavy black areas indicate active cells, responding to the letter X.)

This general plan of organization has been considerably simplified in the perceptron. First of all, we will eliminate the projection area, and assume that the retinal points are directly coupled to association cells, or "A units." The number of input connections to each A unit is specified, but the locations of the origin points for the connections are selected at random from the set of sensory points. Each A unit receives some number, x , of excitatory connections, and some number, y , of inhibitory connections. The connection system from the sensory to association system is a many-to-many system. An excitatory connection from an illuminated sensory point is assumed to transmit a unit positive signal, while an inhibitory connection carries a unit negative signal. Each A unit has a fixed threshold, θ , and is triggered to deliver an output pulse if the algebraic sum (α) of the signals received from the $x+y$ input connections is equal to or greater than θ . A further simplification is introduced at the output side of the association system. Instead of delivering its output signals at

random to a large number of "motor area" cells, the cells of the association system are connected to one or more binary response units, which are turned to their "1" state if they receive a positive signal from the association system, or to their "0" state if they receive a negative signal. The magnitude of the output signal generated by an active A unit is called the "value" of that unit, and is represented by the symbol v . The values of the units are stochastic variables, which change as a function of the history of the system. The organization of a simple perceptron with a single binary response is shown in Fig. 2. The total signal delivered by the set of A units is equal to $\sum \alpha_i v_i$ where α_i is equal to 1 if unit a_i is active, and 0 if a_i is inactive, and v_i is the current value of unit a_i . Note that there are two feedback lines from the response unit (or R unit) to the set of A units. These feedbacks control the "reinforcement," or changes in value, of the A units. In general, if the response $R=1$ occurs, active A units will gain in value, while if the response $R=0$ occurs, active units will lose in value. The value of the A unit thus acts as the memory variable for the system. It has been shown to be desirable to further modify the values of the A units by the rule that if some subset of units gains or loses in value, then the remainder of the units must change in the opposite direction just sufficiently to balance out the net change to zero. Thus, one unit can only gain parasitically, at the expense of the other units, and the total value of all of the A units is kept equal to zero at all times. A perceptron with this property is called a "gamma system." The theory of such systems has been considered in detail elsewhere [5], [6].

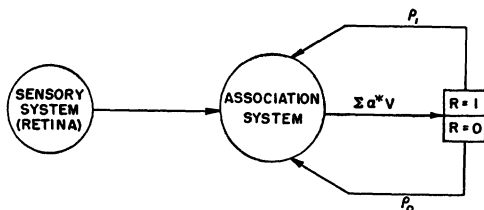


Fig. 2—Organization of a simple perceptron.

DESCRIPTION OF SIMULATION PROGRAMS

Fig. 3 shows the organization of a typical simulation program, for the study of perceptron performance in an environment of visual forms. Actually, four basically different programs have so far been written with a number of variations of each, but the two programs which were used for most of the experiments reported here are both organized in the manner illustrated. The third program involves more direct methods of computation rather than true simulation, while the fourth program (designed to study "cross-coupled systems," in which A units may be connected to one another as well as to S points and R units) has proven too slow to be used successfully.¹

The simulation programs have four main tasks, each of which is actually performed by a separate, self-sufficient program, which is stored on tape, and called into the computer by a supervisory routine. The supervisory routine reads instruction cards provided by the experimenter, which provide information on parameters, and control the sequence of subprograms performed in the course of the experiment. When each subprogram has been completed, control is passed back to the supervisory routine, which reads the next card for further instructions. In a typical experiment, the sequence is as follows:

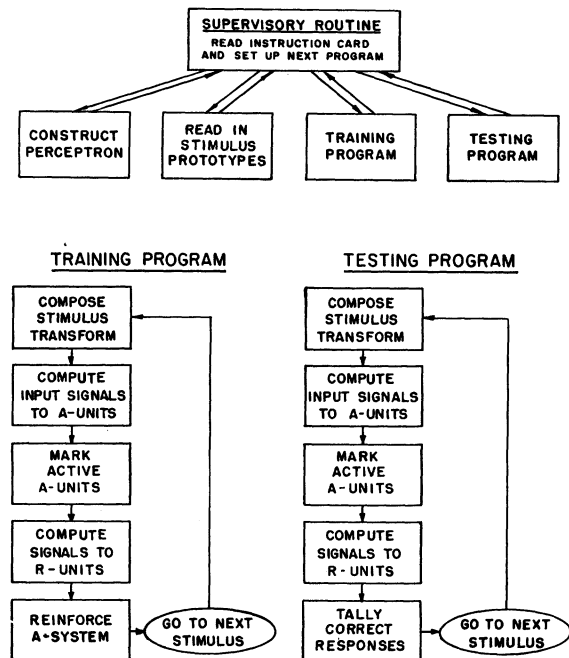


Fig. 3—Flow diagram for simulation program.

1) The perceptron construction routine is called into the core memory, and reads in a set of parameters describing the perceptron to be constructed. These parameters include the number of A units, the number of excitatory and inhibitory connections to each unit, the thresholds of the units, the number of R units, the number of R units connected to each A unit, the decay rate for A unit values (which decay with time in some models) and a random number to be used for priming the pseudo-random-number generator used to control the choice of connections. The program then selects for each A unit a set of $x+y$ sensory points to be assigned as origins for the input connections. This is done by generating a random number modulo N_s (the number of sensory points) for each connection. This number is used to locate one of the N_s storage locations in which

¹ The cross-coupled system was successfully simulated, and predicted effects obtained in December, 1959, using an improved program. Results will be reported in later publications.

the state of each sensory point is indicated when a "visual" pattern is presented.² The perceptron construction routine prepares a table listing all of these connections. In the first simulation program this table was stored on tape; but in the second program, by cutting down the admissible number of A units and connections, it was possible to store the entire table in the core memory, saving a factor of about five in running time of the program. The R units to which each A unit is connected are similarly assigned at random in each of the first two programs, which permit multiple output connections from each A unit. Since, in practice, all experiments have been concerned only with simple binary discrimination problems, more recent programs have been designed with only one R unit, to which all A units are connected. In the second program, it is also possible to assign an initial random distribution of values to the A units, although in most experiments it is assumed that the values start out uniformly from zero.

2) The second stage in the experiment calls for reading a set of "prototype stimulus patterns" into the memory of the computer. These patterns consist of actual dot images of the stimuli to be used, punched as patterns of holes in IBM cards. Thus, if it is planned to teach the perceptron the first four letters of the alphabet, we would read in the images of the letters A, B, C, and D, which are stored for future reference by later routines. These prototypes are never altered, but are used by the stimulus transformation routines which are included in the remaining two programs, to construct variously displaced, rotated, or contracted patterns which are the stimuli actually "shown" to the perceptron.

3) Having constructed the connection tables and read in the prototype stimuli, the computer is ready to begin the actual learning experiment. This consists of an alternation between the two remaining programs, one of which attempts to "teach" the perceptron to recognize the stimulus patterns, while the other evaluates the performance of the perceptron at intervals specified by the control cards. For example, in a typical experiment, the discrimination of the letters "E" and "X," the procedure is as follows. First, a control card calls for the training program to show ten different transformations of the letter "E" (the first stimulus). Each of these is generated by applying a vertical and lateral shift of random magnitudes between zero units of retinal distance and a maximum shift specified by the control card,

a rotation between zero degrees and a specified maximum, and a size somewhere between a specified lower and upper bound. Random numbers generated by the routine determine the exact transformation to be applied to each stimulus, and a new image is composed. The control card then specifies that the response "1" is to be reinforced as the appropriate response for the letter "E." The program accordingly calculates the signals received by each A unit from the transformed stimulus, determines which A units are active, and reinforces the units according to the rules for reinforcement of the $R=1$ condition, for the gamma system, *i.e.*, each active A unit gains an increment in value, while the inactive units lose a compensating amount. In the second of the simulation programs, it is also possible for the stimulus to persist for a designated number of cycles, undergoing a random walk during this time, consisting of unit displacements, rotations, or size changes from the position in which it first appeared. This procedure is characteristic of the "forced learning mode" of experiment, which is the only mode possible for the first simulation program. In this mode, the desired response is turned on, or forced, by the training program at the same time that a stimulus is presented. The second program is also designed to permit a "spontaneous learning mode," in which stimuli occur in a random sequence, and the response spontaneously occurring upon presentation of the stimulus is reinforced, regardless of whether or not it is the response ultimately desired. Most of the experiments to be described in this paper were performed in the forced learning mode. After having presented the ten transformations of the letter "E" which were called for, and reinforced the response $R=1$ for each transformation, control is returned to the supervisory routine, which reads the next control card. In this typical experiment, we next call for ten transformations of the letter "X," to be associated to the response $R=0$. This procedure is carried out in the same manner as before.

We now switch to the testing program, which composes a series of stimulus transformations in the same manner as the training program, and goes through an identical set of calculations to determine the active A units in each case. Instead of reinforcing the association units, however, this program merely records the response, and checks it against the desired response for correctness. If the response is correct, it increments a tally of correct responses. Typically, we may look at twenty transformations of the "E" and twenty transformations of the "X," determining in each case the percentage of correct responses ($R=1$ or $R=0$, respectively). During this procedure, a running description of the responses of the system, numbers of active units, and other analytic data, are printed out by the computer. We may now present another ten E's and another ten X's, reinforcing the system as before, then test the performance once more, to find out whether this addi-

² In each of the first two simulation programs, multiple connections from the same A unit to the same S point are prohibited. In the second program, an inverse constraint was originally employed, fixing the number of connections originating from each sensory point, and assigning termini at random in the association system. This was later modified by the addition of a scheme to obtain, as nearly as possible, uniform numbers of inputs to each A unit as well as fixed numbers of outputs from the sensory units. These variations have not seriously affected the performance of the program, but it appears that somewhat better performance is obtained with the numbers of inputs to the A units is kept uniform.

tional training has improved the performance, and thus continue alternating between training and testing programs indefinitely. It is also possible to reverse the assigned responses in the middle of the experiment, thus reversing previous learning. In order to obtain unambiguous comparisons of performance in different parts of the training series, the testing series are generally "primed" with the same random number to guarantee that the same stimulus transformations will be used on each repetition of the program. The training programs, on the other hand, continue to select stimuli at random, independently of what has gone before. A comparison of the organization of the training and testing programs is presented in the flow diagrams in Fig. 3.

The two main simulation programs total about 5000 words each. The first program was designed to handle up to 1000 A units, and a 72 by 72 sensory mosaic. It was found that this large sensory system presented stimuli with a fineness of grain considerably better than the limits of discrimination of a thousand-unit perceptron, and at the same time, required an excessive amount of time for stimulus transformations, since each illuminated point in the stimulus must be transformed individually into its image point. The second program reduced the retina to a 20 by 20 mosaic, and limited the number of A units to 500. For the first system, the computing time averaged about 15 seconds per stimulus cycle, while in the second system the time was cut to about 3 seconds per cycle. Subsequent improvements in programming techniques indicate that it should be possible to reduce the computing time still further—say to about one second per cycle—for perceptrons of the size allowed by the second program. At the same time, however, analytic developments have suggested a way of actually calculating the exact performance of a given perceptron of the type discussed above, provided all possible stimuli are known, and a matrix of g coefficients, describing the interactions of each pair of stimuli, is computed for the particular network in question. This technique is discussed in the appendix to [7], and is the method employed in the third of our simulation programs for the analysis of spontaneous learning in infinite perceptrons. In that program, the response of the system is obtained analytically, rather than simulated, but the sequence of stimuli is governed by a series of random numbers generated by the program. We will consider some of the results of this program later in this paper.

THEORETICAL PREDICTIONS AND PROBLEMS

Before considering the results of the simulation experiments, let us review the main predictions coming from the theory of the perceptron (see [5]–[7]). The simulation experiments were designed in part to verify these predictions, and in part to study problems which were suggested by the theoretical investigations.

Fig. 4 shows a set of theoretical performance curves for perceptrons of three different sizes, in the problem of discriminating a square from a circle. The broken curves (for P_r) show the probability of giving the correct response to a stimulus which is identical in position, size, etc., to one which was shown previously, during the training period. The horizontal axis indicates the number of stimuli of each class (squares and circles) which were presented during the training period. The solid curves indicate the probability of correct response to *any* square or circle, regardless of whether it was used as a training stimulus or not. Note that both sets of curves approach the same asymptotes as the number of training stimuli becomes large. The first task of the simulation program was to check the general character of these learning curves for typical stimulus material, such as letters of the alphabet or geometric patterns. In particular, it was essential to determine whether the rates of learning agreed with the predicted rates, at least to a reasonable approximation.

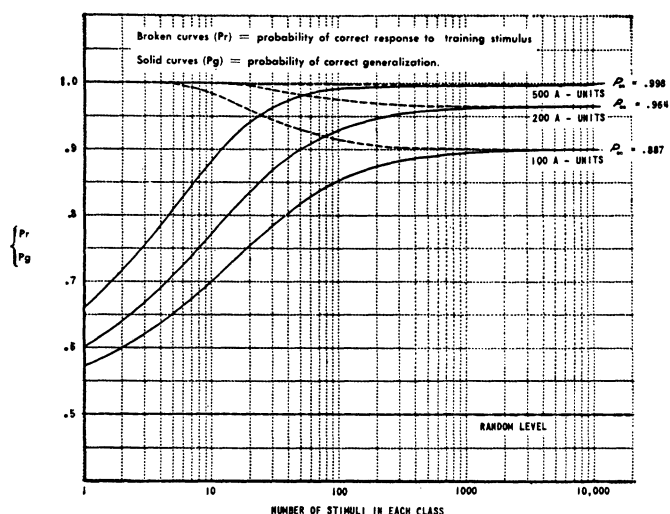


Fig. 4—Learning curves for three typical perceptrons.

A second problem concerned the effect of particular types of transformations, such as shifting of stimuli, rotations, or size changes, upon the learning curves. The original theory did not distinguish among these types of transformations, and it was important to find out whether the system would work equally well for all of them. While sufficient demonstrations have now been made of performance under shifting and rotation conditions, the problem of size changes remains a serious one, with a number of special cases. One such special case involves the assignment of different responses to two stimuli, one of which could be considered a "part" of the other, such as a small circle which could be completely imbedded in a larger one, or the letter "F," which can be considered as an "E" with the lower bar missing. It was predicted that such discriminations would be possible only with a mixture of excitatory and inhibitory in-

put connections, excitatory connections alone being insufficient.³

Related to the problem of size variation in the stimuli is the problem of frequency variation, *i.e.*, some kinds of stimuli being more frequent than others. The response assigned to the more frequent stimulus type will generally tend to dominate the response assigned to the less frequent type, unless the system is designed in such a way as to minimize interaction between different classes of stimuli. The extent of this frequency bias was one of the problems originally set for the simulation programs, but a systematic investigation has not yet been completed.

A different problem area concerns the performance of linear systems. At one stage of the perceptron program, we were particularly interested in systems in which no threshold at all was employed in the A units, the output simply being equal to αv (the algebraic product of the input signal and the stored value) rather than α^*v , as in the model described above. The values were to be augmented by a quantity equal to α if $R=1$, and diminished by α if $R=0$. It can easily be shown that in such a system, if a stimulus pattern can occur with equal probability anywhere in the retinal space (and eliminating special boundary conditions, as in the toroidally connected model), the expected value of every A unit after a long series of stimulus exposures will be exactly zero. Such a system clearly would not learn at all, if stimuli were distributed uniformly in space. If the stimuli were *not* uniformly distributed, however, the values would tend to correlate with any bias existing in the input signals, and it was predicted that such a system should learn to discriminate. The second simulation program was originally set up to study linear systems of this type, both in forced learning and spontaneous learning experiments. The theory of such systems in spontaneous learning is considered elsewhere [7]. While linear systems have now been abandoned, a typical experiment will be considered presently, as it illustrates several points of interest.

The problem of spontaneous learning—the ability of a perceptron to form meaningful classifications of stimulus patterns without any assignment of “correct” responses by a human experimenter—has prompted an extensive series of experiments. The effect was originally demonstrated with the second simulation program, where two disjunct classes of stimuli were properly separated, in a number of experiments. More interesting results were obtained with the third program, which eventually pointed the way to the development of the “cross-coupled association system,” which promises to yield substantially improved performance in a large variety of problems [8]. In studying these spontaneous

learning effects, the first question was whether they could actually be obtained at all, and the second was how much experience would be required, a question for which no satisfactory theoretical answer had been found at the time the simulation experiments were undertaken. In this area, there has been particularly close feedback between simulation work and development of the theory, the simulation program frequently demonstrating the existence of special cases, involving particular parameters or particular stimulus forms, which had not been anticipated. More recent theoretical models owe a great deal to this period of empirical exploration.

RESULTS OF SIMULATION EXPERIMENTS

The first experiments which we shall consider are concerned with the discrimination of the letters “E” and “X” in a forced learning situation, and are illustrated in Fig. 5. The stimuli were constrained to a central portion of the field (as shown by the insert) partly to facilitate learning, and partly to prevent truncation at the boundaries, since the toroidal stimulus space was not used in this program. Fig. 5(a) shows the probability of correct generalization (P_g) as measured on a sample of 20 X's and 20 E's. The stimulus sequence consisted of ten X's followed by ten E's, followed by a test of performance; then ten more X's, ten more E's, and a second test, for a total of 100 training stimuli. The data points shown in the figure are means obtained from ten 100 A unit perceptrons, each of them having a different connection network, but exposed to the same sequence of stimuli. The curves in Fig. 5(b) show the performance of a larger (1000 A unit) perceptron, on a more difficult variation of the same problem. In the solid curve, we see the performance of the system for stimuli rotated by some integral number of degrees selected at random between 0 and 30 degrees. This rotation is combined with vertical and horizontal translations selected within the same limits as in the preceding case. For rotations up to 30 degrees, note that the system attains perfect performance after only ten stimuli of each type. The broken curve shows the performance of the same system for rotations up to 359 degrees, combined with translations as above. In this case, there is a definite decline in the perceptron's performance, although it has attained a P_g of better than 0.90 after 30 stimuli of each type.

The next experiment (Fig. 6) was designed to check the hypothesis that performance on outline figures should be better than on solid figures, since unlike figures represented by their contours would have a minimum intersection on the retina, while solid areas might still have a large intersection even though their shape was different. The figures used were squares (illustrated in the inset) and diamonds, which covered the same areas as the squares, rotated 45 degrees. As shown by the two curves, the outline figures did indeed yield a better performance than the solid figures, giving a per-

³ F. Rosenblatt, “The Perceptron: A Theory of Statistical Separability in Cognitive Systems,” Cornell Aero. Lab., Buffalo, N. Y., Rept. no. VG-1196-G-1; January, 1958. See p. 53.

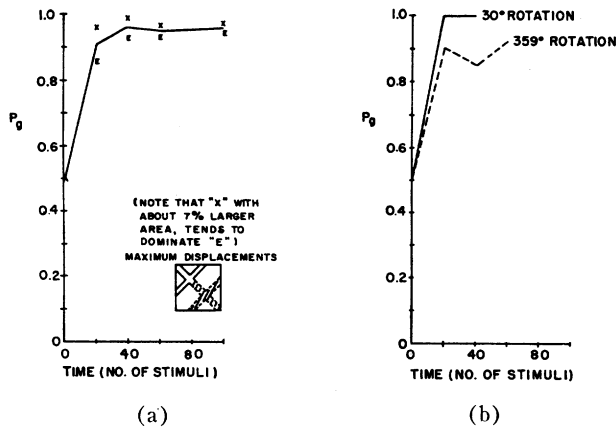


Fig. 5—(a) Experiment 10. "E" vs "X." No rotation. Centers placed in 13 by 13 field, in 72 by 72 retina. $N_A=100$, $\theta=2$, $x=5$, and $y=5$. (b) Experiments 20, 21. "E" vs "X" with shifting plus rotation. $N_A=1000$, $\theta=4$, $x=10$, and $y=0$.

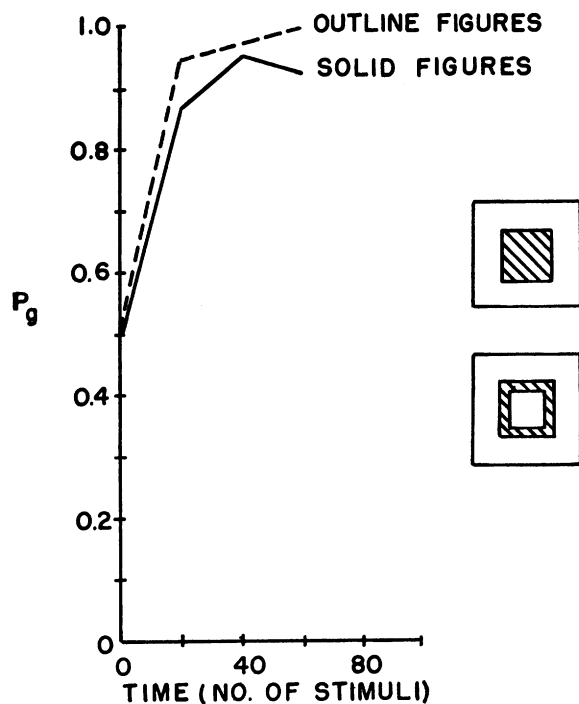


Fig. 6—Experiments 16, 17. Square-diamond discrimination. $N_A=1000$, $x=10$, $y=0$, and $\theta=4$. Centers placed in 13 by 13 field.

fect response record after 60 training stimuli (30 of each class). In this experiment, of course, rotation was eliminated to avoid confusion of squares and diamonds, and the figures were merely displaced in the same manner as the E's and X's in the preceding experiment.

Fig. 7 shows two experiments concerned with part-whole discrimination, which was discussed in the preceding section. In Experiment 18, illustrated in Fig. 7(a) a system with only excitatory connections to the A units was simulated. The stimulus is shifted at random in the central portion of the field, as before. In this case, the letter "E" was correctly learned, but the system was unable to learn to give the opposite response to the letter "F." In Experiment 22, shown in Fig. 7(b), we see that

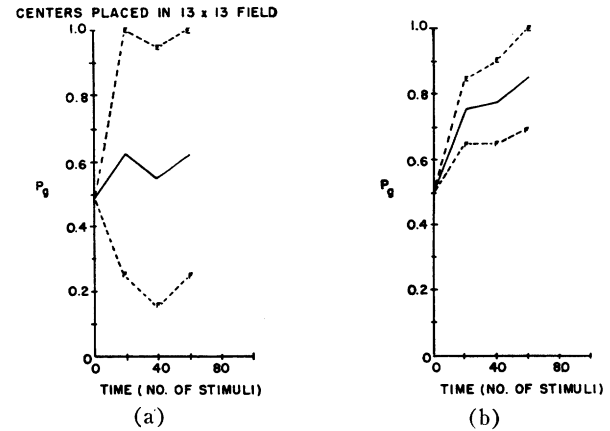


Fig. 7—(a) Experiment 18. "E" vs "F." $N_A=100$, $x=10$, $\theta=4$, and $y=0$. Centers placed in 13 by 13 field. (b) Experiment 22. "E" vs "F." $N_A=1000$, $x=5$, $\theta=3$, and $y=5$.

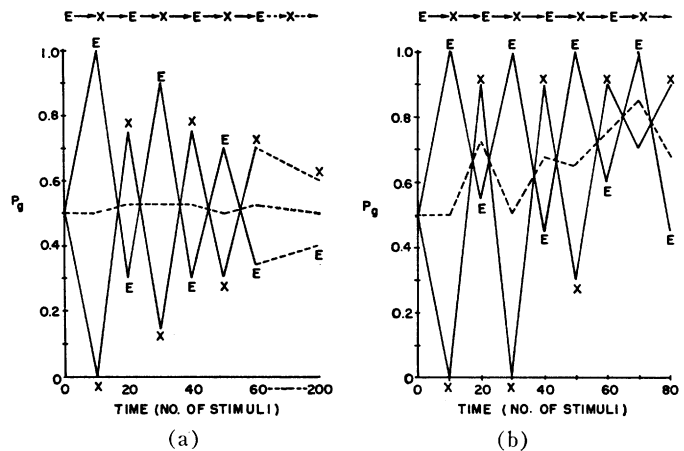


Fig. 8—Linear system experiments ("E" vs "X"). (a) Experiment 4-14, 15. $N_A=500$, $x=4$, $y=4$. Centers placed in 13 by 13 field. (b) Experiment 4-16. $N_A=500$, $x=4$, and $y=4$. Centers placed in 5 by 5 field.

a system, in which half of the connections to the A units are inhibitory, is able to learn the correct response to both classes of stimuli, although the F response is considerably less consistent than the E response. Experiments 18 and 22 are, unfortunately, not fully comparable, as the perceptron in the second case was a thousand-unit system, while in Experiment 18 only a hundred A units were used. The character of the curves in these experiments, however, is definitely not a function of the size of the systems, but rather of the stimulus relationships, as shown by supporting evidence from many other cases. These results are in closer agreement with the theoretical predictions referred to earlier.

The next experiment (Fig. 8) was performed with the second simulation program, and represents the learning which is possible with a purely linear model, if the stimuli are constrained to one region of the retinal field. In this experiment, instead of testing the perceptron after every twenty stimuli, as in previous experiments, it was tested after every ten stimuli, which yields the characteristic pattern of converging oscillations shown in the figure. The first ten stimuli were all E's, and after these

ten exposures, we find that the system has learned the "E" perfectly, but always gives the wrong response to stimuli of the opposite class (the letter "X"). The perceptron was then shown ten X's, to which the opposite response was forced, and we find at time 20 that it has now learned to give the desired response to the X, but has almost completely forgotten the proper response to the letter E. The amplitudes of such oscillations are apt to be increased by a large decay rate for the values of the A units (which makes more recent reinforcement more effective than earlier experience), but in the experiment illustrated here the decay rate was zero. Note that in Experiment 4-16 [illustrated Fig. 8(b)] the mean learning curve, shown by the broken line, climbs towards a high probability level as experience with both stimuli increases. At the same time, the swings in performance become considerably less pronounced, as each series of ten stimuli represents a progressively diminishing portion of the total experience of the system. The important conclusion from this experiment is that discrimination learning is possible for a linear system, provided the stimuli are sufficiently constrained in location. The retinal field in this case was 20 by 20 units, and the centers of the stimuli were constrained to a 5 by 5 region in the center of the retina. In Experiment 4-14 [shown in Fig. 8(a)], where the stimuli were distributed more freely over the retina (with the centers in a 13 by 13 field), no learning was demonstrated even after 200 stimuli. As a methodological experiment, these results indicate the importance of making sure that the stimulus distribution employed does not include "location cues" which are sufficient to indicate which stimulus is present, if we wish to test the ability of the perceptron to discriminate pattern characteristics exclusive of location. This can be fully guaranteed, in general, only by a uniform stimulus distribution over the entire field, with the elimination of special boundary effects by assuming a closed space, or an infinite space, as with the Born-von Kármán boundary conditions referred to in the Introduction.

Experiment 4-36, shown in Fig. 9, was again carried out with the second simulation program, this time with a more conventional perceptron. The threshold of zero, employed here, is sufficient to make the system fundamentally nonlinear, by eliminating the output of A units in the presence of negative input signals. The experiment was designed to show the performance of the system in the presence of a high degree of randomness, or noise, in the initial values of the A units. The stimuli for this experiment were vertical and horizontal bars, 4 units in width and 20 units long. A 5 per cent decay rate was introduced for the values of the A units. Note that in spite of the high decay rate and high initial noise level, the system achieved perfect performance on both classes of stimuli after a total of only 50 stimuli. This should be compared with the performance of very large (or infinite) perceptrons, in a spontaneous learning experiment with the same types of stimuli, which is illustrated in Fig. 10.

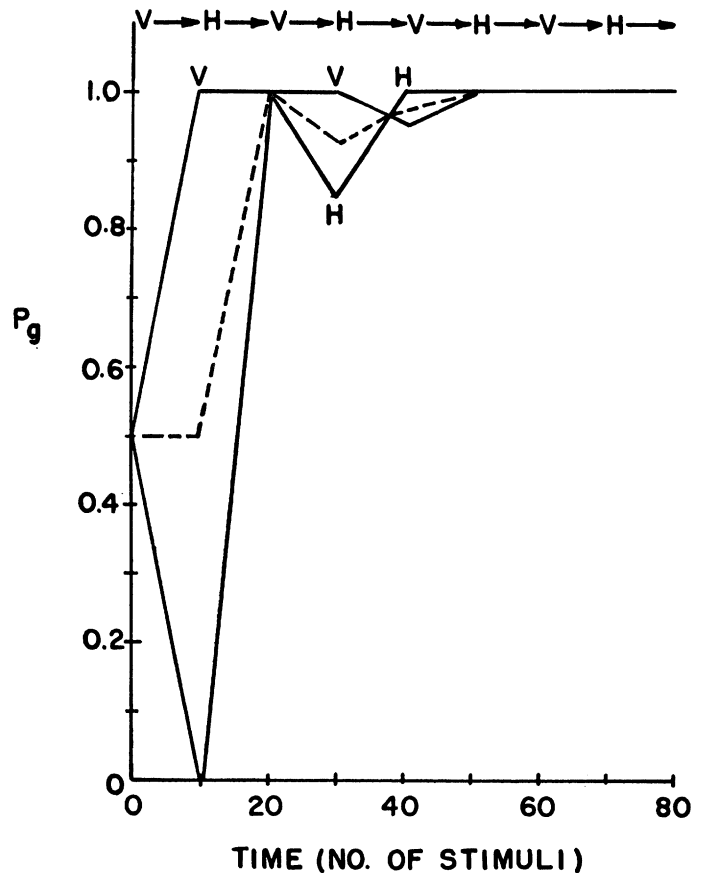


Fig. 9—Experiment 4-36. Forced learning experiment with vertical and horizontal bars. 500 A units. $\delta = 0.05$, $\theta = 0$, $x = 4$, $y = 4$, and V_0 between +500 and -500. Centers in 5 by 5 field, in 20 by 20 retina.

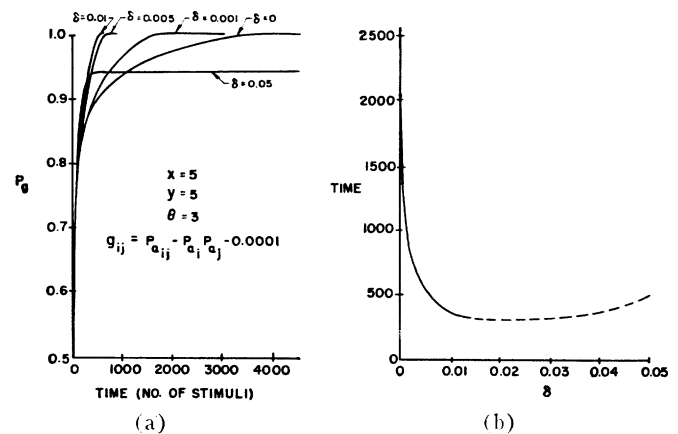


Fig. 10—Experiment 5-4. (a) Spontaneous organization of infinite perceptron in environment of 4 by 20 vertical and horizontal bars. (b) Expected waiting time to perfect performance, as a function of decay rate (means of 10 runs).

In the experiment shown in Fig. 10, stimuli were placed with equal probability at any position in a 20 by 20 retinal field, with Born-von Kármán boundary conditions. The stimuli were 4 by 20 horizontal and vertical bars, as in the previous case. The perceptron used in this experiment is one in which the A units are reinforced for the response $R = 1$, but are left unaltered if the response $R = 0$ occurs. Unlike all of the previously illustrated experiments, this is a spontaneous learning experiment, in which no attempt is made to control the

response during the learning procedure, reinforcement being applied for whichever response is elicited by a given stimulus. The perceptron here was assumed to have an infinite number of A units, and the calculations were done with the third program, which was specifically designed to handle these conditions. The family of curves in Fig. 10(a) shows the performance as a function of the decay rate, δ . We find that for a zero decay rate, the system eventually learns to dichotomize the bars correctly 100 per cent of the time, *i.e.*, it learns to assign one response to all horizontal bars, and the opposite response to all vertical bars. However, this takes upwards of 3000 stimuli in most cases.⁴ As the decay rate increases, performance improves progressively, until a decay rate is reached (0.05 in this case) for which the system is unstable, and never attains perfect performance. The effect of the decay short of the instability level appears to be to keep previous reinforcements from accumulating to such a degree that they are difficult or impossible to undo, as the system settles into a more satisfactory terminal state; in other words, the decay keeps the system flexible, by making it possible to reverse the effects of previous learning more readily. At the instability level, previous reinforcements are reversed so readily that they are unable to maintain their effect at all, and associations are likely to be lost and reformed continually. The curve in Fig. 10(b) which shows expected waiting time to perfect performance, for the same series of runs, indicates the same phenomenon. We find that there is a clear optimum in performance as a function of the decay rate, for $\delta \approx$ approximately 0.01. Beyond this point, instability begins to occur, as indicated by the broken curve in the figure.

This experiment is the best demonstration to date of the "self-organizing" capability of a perceptron. Nonetheless, it can be demonstrated that minor changes in the stimulus environment will make it impossible for the same perceptron to achieve a satisfactory dichotomy. For example, if the 4 by 20 horizontal bars are replaced by double bars, composed of two 2 by 20 vertical bars separated by a space of 3 units, the perceptron will never spontaneously learn to distinguish the double bars from the single bars. Other classes of stimuli can be set up which are equally difficult, or impossible, for the system to learn spontaneously, although in each of these cases the problem would present no difficulty in a forced learning situation. Moreover, the curves in Fig. 10 are convex, indicating increasing difficulty in correctly associating the last few stimuli after most of the class has been learned. In a human subject faced with this task we would expect concave curves instead. These considerations indicate that the spontaneous learning capability of this perceptron, while interesting, is not sufficient to provide a basis for a biological theory of perceptual organization. This problem is considered in further detail elsewhere [8].

⁴ Individual runs differ from one another due to differences in stimulus sequence, even though the perceptrons are infinite; the curves shown are means of ten different runs.

CONCLUSIONS

The simulation experiments described above have gone a long way toward demonstrating the feasibility of a perceptron as a pattern-recognizing device. Both forced learning and spontaneous learning performances have been investigated, and some insight has been gained into conditions under which different systems break down, or deviate from typical biological learning phenomena. Although digital simulation is apt to be time-consuming and expensive, particularly for large networks, improved programming methods have cut down the running time considerably, so that for early investigations of all systems proposed up to this time, digital simulation is still competitive with the construction of actual hardware models. As the number of connections in the network increases, however, the burden on a conventional digital computer soon becomes excessive, and it is anticipated that some of the models now under consideration [8] may require actual construction before their capabilities can be fully explored.

Digital programs undertaken to date have been concerned exclusively with the logical properties of the network, rather than with any particular hardware embodiment; that is, there has been no attempt to introduce simulation of electronic noise, component variation, or other factors which might affect the performance of an actual system. The results of these programs, therefore, should be interpreted as indicating performances which might be expected from an "ideal," or perfectly functioning system, and not necessarily as representative of any particular engineering design. A Mark I perceptron, recently completed at the Cornell Aeronautical Laboratory, is expected to provide data on the performance of an actual physical system, which should be useful for comparative study.

A new program is currently being employed to simulate the "cross-coupled perceptron" described elsewhere [8]. The results of this study will be reported separately when they are available.

BIBLIOGRAPHY

- [1] W. A. Clark and B. G. Farley, "Generalization of Pattern Recognition in a Self-Organizing System," *Proc. W. J. C. C.*, pp. 86-91; 1955.
- [2] B. G. Farley and W. A. Clark, "Simulation of self-organizing systems by digital computer," *IRE TRANS. ON INFORMATION THEORY*, vol. IT-4, pp. 76-84; September, 1954.
- [3] D. O. Hebb, "The Organization of Behavior," John Wiley and Sons, Inc., New York, N. Y.; 1959.
- [4] N. Rochester, J. H. Holland, L. H. Haibt, and W. L. Duda, "Tests on a cell assembly theory of the action of the brain, using a large digital computer," *IRE TRANS. ON INFORMATION THEORY*, vol. IT-6, pp. 80-93; September, 1956.
- [5] F. Rosenblatt, "The Perceptron: A Theory of Statistical Separability in Cognitive Systems," Cornell Aero. Lab., Buffalo, N. Y., Rept. no. VG-1196-G-1; January, 1958.
- [6] F. Rosenblatt, "The perceptron, a probabilistic model for information storage and organization in the brain," *Psych. Rev.*, vol. 65, pp. 386-408; 1958.
- [7] F. Rosenblatt, "Two Theorems of Statistical Separability in the Perceptron," *Proc. of Symposium on the Mechanization of Thought Processes*; 1958. H. M. Stationery Office, London, Eng. Also printed without appendix by Cornell Aero. Lab., Buffalo, N. Y., Rept. no. VG-1196-G-2; September, 1958.
- [8] F. Rosenblatt, "Perceptual Generalization Over Transformation Groups," *Proc. of Interdisciplinary Conf. on Self-Organizing Systems*; 1959. Awaiting publication by Pergamon Press, New York, N. Y.