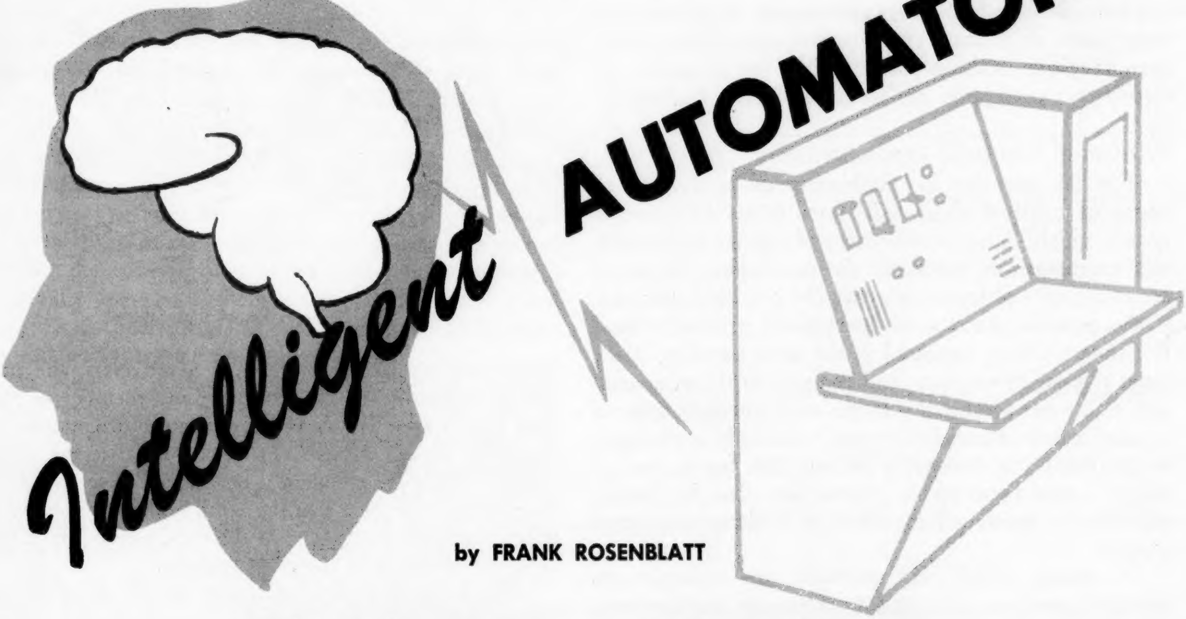
Vol. VI, No. 2, Summer 1958

FILE COPY

research trends

CORNELL AERONAUTICAL LABORATORY, INC., BUFFALO 21, NEW YORK

The Design of an *Intelligent* AUTOMATON



by FRANK ROSENBLATT

Introducing the perceptron — A machine which senses, recognizes, remembers, and responds like the human mind.

STORIES about the creation of machines having human qualities have long been a fascinating province in the realm of science fiction. Yet we are now about to witness the birth of such a machine — a machine capable of perceiving, recognizing, and identifying its surroundings without any human training or control.

Development of that machine has stemmed from a search for an understanding of the physical mechanisms which underlie human experience and intelligence. The question of the nature of these processes is at least as ancient as any other question in western science and philosophy, and, indeed, ranks as one of the greatest scientific challenges of our time.

Our understanding of this problem has gone perhaps as far as had the development of physics before Newton. We have some excellent descriptions of the phenomena to be explained, a number of interesting hypotheses, and a little detailed knowledge about events in the nervous system. But we lack agreement on any integrated set of principles by which the functioning of the nervous system can be understood.

We believe now that this ancient problem is about to yield to our theoretical investigation for three reasons:

First, in recent years our knowledge of the functioning of individual cells in the central nervous system has vastly increased.

Second, large numbers of engineers and mathematicians are, for the first time, undertaking serious study of the mathematical basis for thinking, perception, and the handling of information by the central nervous system, thus providing the hope that these problems may be within our intellectual grasp.

Third, recent developments in probability theory and in the mathematics of random processes provide new tools for the study of events in the nervous system, where only the gross statistical organization is known and the precise cell-by-cell "wiring diagram" may never be obtained.

Receives Navy Support

In July, 1957, Project PARA (Perceiving and Recognizing Automaton), an internal research program which had been in progress for over a year at Cornell Aeronautical Laboratory, received the support of the Office of Naval Research. The program had been concerned primarily with the application of probability theory to the problem of memory and perception. In undertaking this investigation, the author assumed at the outset that the organization of the sensory world of light, sound, temperature, pressure, etc., is learned, rather than being immediately self-evident to the perceiving system.

In other words, an organism fully equipped with visual apparatus, and exposed to an environment of, say, squares and circles, would not be able to tell these forms apart unless it has specifically learned to do so. This means in the fullest sense that the two kinds of forms would be indistinguishable at the outset, i.e., that two squares, chosen at random, would appear to be no more alike than a square and circle, similarly chosen at random. Inasmuch as people are unable to report their experiences as infants, experimental observations have been unable to establish a definite case for or against the theory that perception of "similarity" must be learned.

Problem of Perceptual Generalization

For the engineer or mathematician attempting to construct a system which will "learn to perceive" (i.e., a system which, in the environment of squares and circles, will spontaneously arrive at the conclusion that there are two classes of forms present), the principal difficulty is the problem known as "perceptual generalization." If a square always appeared in the same position, at the same angular orientation, and reduced to the same size, and if all other geometrical forms were similarly reduced to some standard transformation, it would be a relatively simple matter to distinguish among such forms, and to assign a new form to its appropriate class by simply matching it against all members of a library of stored images.

A system which will perform this reduction to standard position, size, etc., is extremely cumbersome, however, and still leaves the more baffling problem of how non-rigid forms, such as a man or an ocean wave, can be recognized. The problem of perceptual generalization is concerned with how, after exposure to a

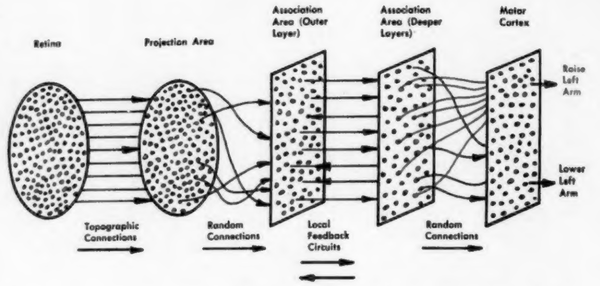


FIG. 1 — **Organization of a biological brain. (Red areas indicate active cells, responding to the letter X.)**

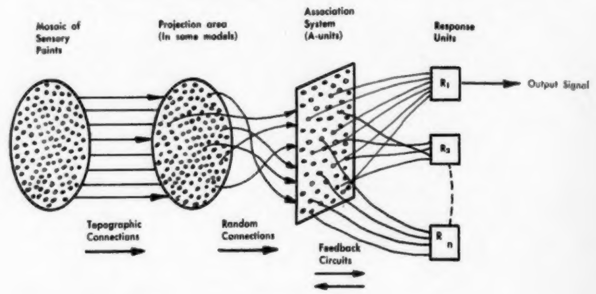


FIG. 2 — **Organization of a perceptron.**

limited sample of forms from a given class, the perceiving system is able to recognize any member of that class (e.g., a man in any posture, angular orientation, or costume), even if it has never seen the particular image before. While the problem is here stated in terms of the visual sense, it is clear that the same problem exists when other senses are used. One of the most interesting forms of this problem is in speech recognition.

The design of a physical system which can recognize "similarities" in our complex environment, where countless demands are made on all of our senses, and which tends, spontaneously, to form meaningful classifications of stimuli in such an environment, has been the main objective of Project PARA.

THE COVER

Resistance thermometers for measuring heat transfer rates in shock tunnels have been successfully developed by CAL in conjunction with its hypersonic research. Here, the skilled hands of a technician insert a glass button with a resistance thermometer mounted on it into a slender wedge. Five such wedges are used in a rake, or series of probes, to calibrate flow in a hypersonic shock tunnel. Each wedge has a resistance thermometer on both faces. Thus, by measuring the heat transfer rate on both sides of the wedge, it is possible to measure both flow angularity and flow Mach number. Resistance thermometers are fabricated in the Materials Department for the hypersonic research activities of the Aerodynamic Research Department.

Understanding the Perceptron

To understand the proposed machine — or perceptron — it is necessary first to understand something of the nature of the brain and how it works. Figure 1 represents the basic organization of the human brain, including the motor cortex, which controls physical responses. This organization has been well established through physiological and anatomical studies. The connections from the retina to the visual projection area provide a sort of map of the visual field in the brain. Beyond this point, however, connections appear to become increasingly random, so that in the association areas (which appear to be mainly responsible for learning and memory) it is no longer possible to relate a particular point to some specific location in the retina. The association cells of the brain are likely to respond to any one of a vast number of different stimuli from any of the five senses.

Inputs to the association area tend to arrive at the surface layers of cells, while outputs emanate from the deeper layers. Feedback circuits between these layers are so organized that a cell in a deeper layer is more likely to feed back to the same outer layer cells which caused its activity than to cells which take no part in its stimulation. When impulses arrive at the motor cortex, an intelligible order appears to have been restored. The motor cortex apparently contains a kind of map of the body surface, so that stimulation of a particular location will lead to a specific muscular response. Thus the confusion of connections through the association area has somehow "recognized" the visual stimulus, and developed an output signal which is constrained to particular, relevant channels.

Mystery Still Exists

The channels into and out of the central nervous system have been rather well mapped. We know what points in the projection area will respond, say, to a ray of light in the lower right quadrant of the visual field, and we know where in the motor cortex the signal which causes a man to raise his left arm originates. The big mystery is how the apparently unintelligible tangle of connections in the association area manages to record the fact that a beam of light (or a dog, or a landscape) is actually seen, and how the impulses from the visual stimulus are interpreted in such a manner as to select the single appropriate response channel.

In Figure 2 is shown the organization of a system whose "anatomy" is completely known — the perceptron. This system is capable of the same functions of sensing, recognition, retention, and response selection as its biological counterpart. Although the similarity of organization to the biological brain is clearly evident, certain differences and simplifications should be noted.

First, the projection area, which is found in all advanced biological systems, is not essential for the perceptron. In simplified models, the retinal points are assumed to be connected directly to randomly selected units (A-units) in the association system. In other words, each sensory point may be connected to one or more A-units chosen at random from all possible units in the system.

Second, the responses (R-units) of the perceptron are typically binary devices which are either on or off, or which may sometimes have a third "neutral" condition. Little attention has been given to responses which must vary in intensity, the R-units of the perceptron being used to signal the state of the system.

Third, the R-units of the perceptron actually combine the functions of the second association layer with those of the motor cortex. The R-units transmit feedback signals to the same A-units which are responsible for activating the unit in the first place.

Meaning of Responses

These response units of the perceptron are more like special association cells whose activity represents the brain's recognition response to various stimuli, rather than cells in the motor cortex which regulate speech or movement. The activation of a particular response for the perceptron might mean, for example, that a triangle is present, or that a man's voice is being heard. Each response is thus capable of representing a particular concept, or abstraction, in terms of which the environment is organized.

At the outset, when a perceptron is first exposed to stimuli, the responses which occur will be random, and no meaning can be assigned to them. As time goes on, however, changes which occur in the association system cause individual responses to become more and more specific to such particular, well-differentiated classes of forms as squares, triangles, clouds, trees, or people.

In order to clarify the foregoing process, it is necessary first to point out a fundamental feature of the perceptron — a feature whose biological counterpart has not yet been demonstrated. When an A-unit of the perceptron has been active, there is a persistent aftereffect which serves the function of a "memory trace." The assumed characteristic of this memory trace is a simple one: whenever a cell is active, it gains in "strength," so that its output signals (in response to a fixed stimulus) become stronger, or gain in frequency or probability. The strength of an A-unit is measured in units of "value" (**v**), a hypothetical quantity. All theoretical attempts to account for learning in the nervous system have ultimately been forced to assume some functional change which serves the same purpose as **v**.

Simple Memory Hypothesis

The variable **v** appears to be the simplest, and in some ways the most plausible, memory hypothesis advanced to date. Further, the perceptron is the first system proven to be workable with so simple a memory mechanism. An examination of the behavior of A-unit values in more advanced models of the perceptron makes it clear that such a variable would be exceedingly difficult to detect, physiologically. The values of the A-units tend to a terminal equilibrium condition, from which they may fluctuate slightly, either positively or negatively. It is not surprising, therefore, that such an erratic variable has escaped detection in physiological experiments. Nonetheless, such slight fluctuations exert a mass statistical effect which can be demonstrated to enable the perceptron to form new associations, to "store" memories, and to select appropriate responses.

A simple perceptron is shown in detail in Fig. 3. The circles in this figure represent sets of units, and there

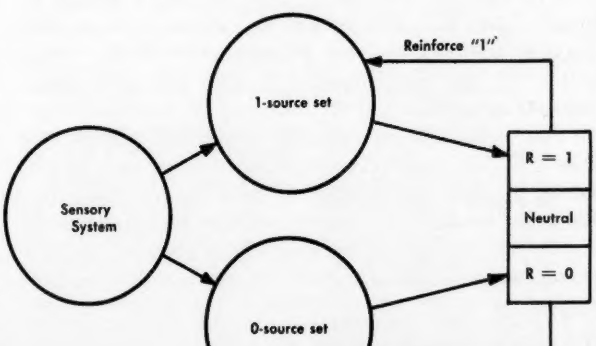


FIG. 3 — Detailed organization of a single perceptron.

might be hundreds or thousands of units in each set. This perceptron has only a single response unit, which has three possible states: "1", "0", or "neutral".* In the absence of any signal, the response unit is in a neutral state, and delivers no output. In the presence of a signal it tends to oscillate between the neutral state and the "1" or "0" condition.

The association system is divided into two subjects, or "source sets", one of which tends to activate a 1-response, while the other tends to activate a 0-response. If the total output signal from the 1-source set is greater than the total output signal from the 0-source set, the response R = 1 tends to occur. If the total signal from the "0" set is greater, the response R = 0 tends to occur.

EDITOR'S NOTE

Because of the unusual significance of Dr. Rosenblatt's article, Research Trends is proud to devote this entire issue to it.

In the Fall issue, we will return to our policy of presenting two articles highlighting trends in the Laboratory's research.

In a more elaborate perceptron there may be a large number of such responses, and the source sets for these responses will typically cross-cut one another, so that the same A-unit may be in the source sets of a number of responses. It can be shown that such multiple functioning of the A-units does not interfere excessively with their performance.

Feedback Signals

The R response causes a feedback signal to be sent back to the members of its own source set. These feedback signals have the effect of multiplying the rate of activity of the A-units which receive them. Thus if the response should happen to be "1," each unit in the 1-source set might have its rate of activity doubled, while the members of the 0-source set remain unaffected. This not only increases the slight original tendency to maintain the response R = 1, it also means that the A-units in the 1-source set will gain in value at a **greater rate** than the units in the 0-source set. The increase in value is referred to as a "reinforcement." An example series of pertinent experiments is shown in Fig. 4.

It should be emphasized that the condition shown in Part D (Fig. 4 Example) could easily be reversed, if the perceptron were "trained" with only a single square and circle, prior to testing. In order to make this performance reliable, the perceptron must first see a **sample** of squares (say, 100 - 200 squares in various positions and angular orientations) and a sample of circles, being forced by the experimenter to give the appropriate response to each.

The predicted performance of typical perceptrons with 100, 200, and 500 association cells in each source set, in learning to discriminate two figures which are about as "similar" as a square and a circle, is shown in Fig. 5. The broken curves show the probability that the perceptron will give the correct response if it is shown, as a test figure, one of the identical stimuli used during the training period, as in Part B of the example.

The solid curves show the "probability of correct generalization." This is the probability that the perceptron will give the appropriate response for **any** member of the stimulus class picked at random, as in Part D of the example.

It can be seen from Fig. 5 that a perceptron with 100 A-units in each source set, which has been trained with 100 squares and 100 circles, should have a probability of 0.92 of giving the correct response if it is shown one of the squares (or circles) that it has seen before. It should also have a probability of 0.85 of giving the correct response to a completely random square or circle which it may never have seen before. If learning experience is continued indefinitely, both probabilities converge to the same limit, in this case 0.887. Thus, in the limit, it makes no difference whether the perceptron has seen the particular stimulus before or not; it does equally well in either case.

The mathematical proof of the foregoing statement constitutes a proof of this machine's ability to form perceptual generalizations.

As the number of association units in the perceptron is increased, the probabilities of correct performance approach unity. From Fig. 5 it is clear that with an amazingly small number of units — in contrast with the human brain's 10^{10} nerve cells — the perceptron is capable of highly sophisticated activity.

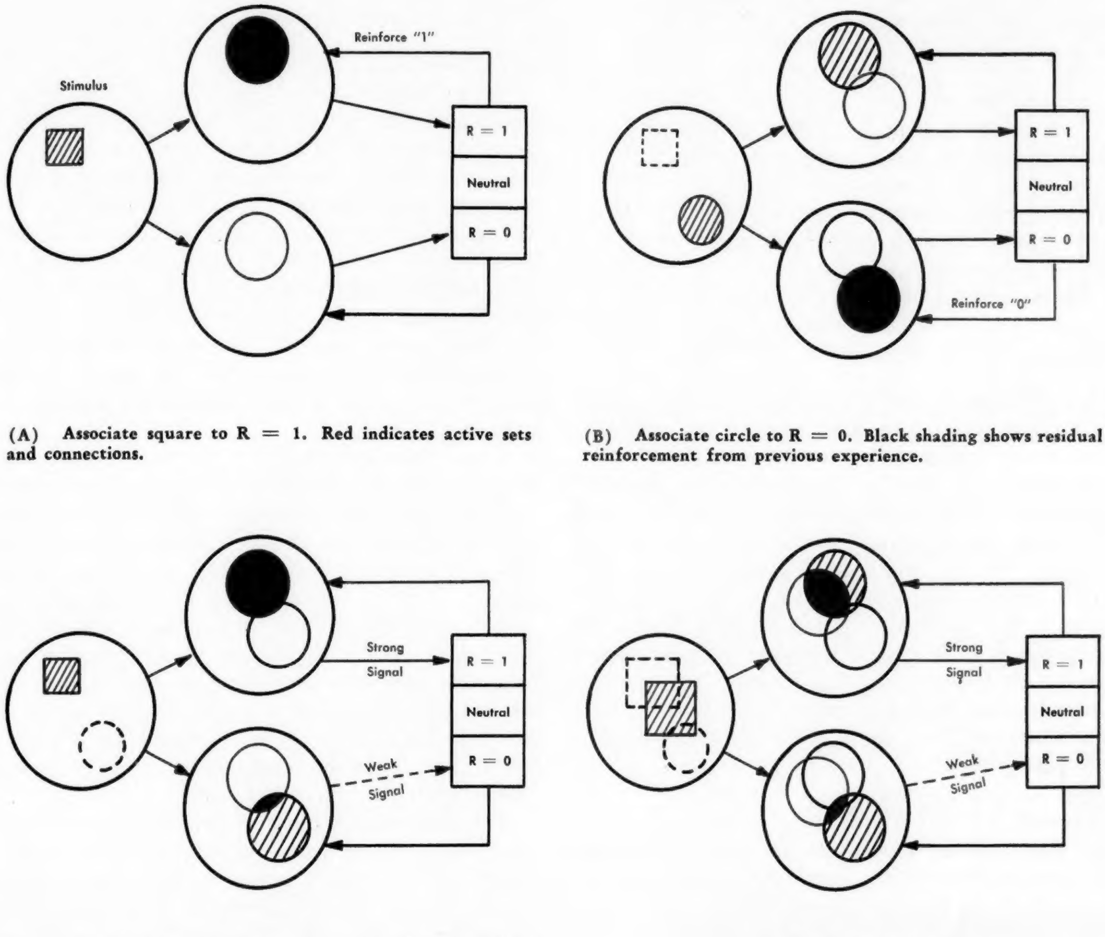
Can Recognize Patterns

It is important to recognize that the mode of operation of this system does not limit it to such simple, rigid forms as geometrical figures. Any classes of forms, which meet certain conditions of similarity, can be distinguished by the perceptron, including such diverse patterns as the letters of the alphabet, human profiles, or types of aircraft. With some very slight modifications, it can be shown that the perceptron should be capable of recognizing patterns in time (such as speech and movement) as well as patterns in space. A large increase in the "vocabulary" of the perceptron can be obtained with a relatively slight increase in the number of binary response units.

These results were well established, theoretically, by the Fall of 1957 (Ref. 3). At this time a simulation program was started, using the IBM 704 computer, to determine how well the theory would hold up in practice. While no digital computer can approach the perceptron in speed and flexibility of performance, such a computer can examine each connection and A-unit of the system in turn, can then compute the appropriate signals which would be transmitted in a physical network, and can next calculate the performance of a perceptron in response to a series of visual forms. Many such simulation experiments are now complete, and all main predictions of the theory are substantiated.

*The "neutral" state shown in these R-units is introduced in order to avoid excessive complexities in the discussion. In the proposed perceptron, the R-units will actually be simple binary devices with no intermediate neutral condition.

FIG. 4 — EXAMPLE EXPERIMENT



(A) Associate square to R = 1. Red indicates active sets and connections.

(B) Associate circle to R = 0. Black shading shows residual reinforcement from previous experience.

(C) Test with original square. Solid red areas show effect of previous reinforcement.

(D) Test with random square.

In Part A the perceptron illustrated in Fig. 3 is responding with R = 1 to the image of a square, in the upper part of the visual field. The red connections are active. Note that an equal subset of A-units tends to respond to the stimulus in each of the two source sets (small red circles). It is assumed that either by chance, or because of "forcing" by the experimenter, the response unit goes to the state R = 1. Consequently, the 1-source set is reinforced at a rapid rate, relative to the 0-source set. The reinforcement is indicated by the solid red area, in the set of A-units responding to the square. At this point, it is clear that if the same square were to be repeated, the signal from the 1-source set would be stronger than the signal from the 0-source set, so that R = 1 would almost certainly be repeated.

In Part B a second stimulus (circle) is shown to the perceptron, which still carries the residual effects of its previous experience. It is assumed that the response R = 0 occurs, either spontaneously, or because of forcing by the experimenter. The added reinforcement is shown by the solid red area in the 0-source set. The question now is: will the perceptron still give the "appropriate" response (R = 1) if it is again shown the original square?

Part C shows that while some fraction of the reinforcement due to the circle is expected to carry over to the square, because of the intersections of the responding sets of A-units, the total reinforcement picked up in the 1-source set is greater than the total reinforcement picked up in the 0-source set, so that the appropriate response is expected to occur.

But the foregoing experiment is, in a sense, a trivial one, since we clearly cannot pre-train the perceptron on every possible square and circle, so as to guarantee that it will give the proper response in a particular case. The critical question is asked in Part D. What happens when we show the perceptron, which was trained in parts A and B, a new square, picked at random, which may or may not coincide in size and position with the square which was previously seen? Will the perceptron still show any tendency to prefer the correct response, or will its choice of response be entirely random?

Now, if a new square is picked at random, it can be demonstrated that it is likely to activate a set of A-units in each source set which has more members in common with the sets of A-units responding to other squares, than with the sets of A-units responding to circles. Consequently, the condition shown in Part D is most likely to result. Under these circumstances, while some reinforcement is expected to be picked up from both the previous square-circle reinforcement, the total reinforcement (solid red area) in the 1-source set remains greater, and the appropriate response should occur.

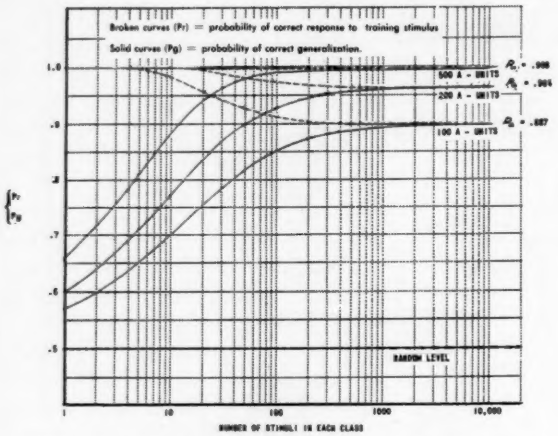


FIG. 5 — **Learning curves for three typical perceptrons.**

Now although the perceptron developed to that point could be shown to have an impressive capacity for learning and remembering those concepts imposed on it by an experimenter, it soon became clear that it could not spontaneously form meaningful classes. In fact, such a perceptron, turned loose in an environment with no intervention on the part of the experimenter, tends toward a terminal condition in which it gives either the response $R = 1$ universally, to everything it sees, or the response $R = 0$, equally universally, without any discrimination between stimulus classes. The responses of such a perceptron clearly give no information about the environment. Such perceptrons are referred to as Class C perceptrons.

Proof Found

Recently, a proof was found for a second theorem which indicates that with a seemingly trivial modification, a perceptron having strikingly different properties results. The modification required is simply that the values of the A-units should be allowed to decay at a rate proportional to their present magnitude. The resulting exponential decay is characteristic of practically all biological quantities which require the continued application of energy for their persistence; and it is probable that a similar rule must hold for biological memory traces.

The effect of introducing a decay component is that, instead of growing indefinitely, the values of the A-units tend toward a terminal equilibrium condition which depends on the input signals and activity of the unit. Perceptrons organized in this way are members of the class C'. The characteristics of this class are stated in an "existence theorem," which is of such fundamental importance that it seems worth stating here in simplified form:

> A class C' perceptron can be expected to divide the stimuli of any arbitrary environment into classes, without any assistance or training by a human operator. The system will form its own concepts and these concepts will tend to be meaningful; that is, they represent an organization of the environment on the basis of similarity and dissimilarity.

In brief, a Class C' perceptron is the first non-biological system known to be capable of classifying, conceptualizing, and symbolizing its environment — particularly a completely new and unanticipated environment — in the absence of any human training or control.

Perceptron Being Built

A working model of a Class C' perceptron is scheduled for completion within the next year. Although the economical design of such a system presents several difficult problems, considerable headway has already been made in the design of suitable components. Meanwhile, the predictions as to the terminal states of a Class C perceptron have already been tested, using the 704 as a simulator, and a program to investigate the Class C' perceptrons in a similar fashion is in operation.

There have, of course, been many theoretical brain models before the perceptron. We might profitably summarize the main points which set the perceptron off from other attempts.

(1) The perceptron is the first system which appears to be economical, in the sense that it can operate successfully on non-trivial problems, with a smaller number of units than are present in the human nervous system. All previous system designs, which are in any way comparable, are of a completely prohibitive size and cost.

(2) The perceptron is not built to rigid logical specifications, in which the failure of a particular unit is likely to cause a breakdown of operation. The design of the system is based on a small number of statistical parameters and some general logical constraints, but within these limits the actual connections can be drawn from a table of random numbers.

(3) The perceptron does not recognize forms by matching them against a stored inventory of similar images, or by performing a mathematical analysis of characteristics. The recognition is direct and essentially instantaneous, since the "memory" is in the form of new pathways through the system, rather than a coded representation of the original stimuli. There is no way of reconstructing the original stimuli from the memory with any absolute certainty. Nonetheless, the probability of obtaining an appropriate recognition response, or "naming response," can be made virtually perfect.

(4) As a model for the biological brain, the perceptron does not violate any known information about the central nervous system. Its size, the logic of its connections, the degree of reliability required of individual units, the permissible random variation in its "wiring diagram," and the kinds of signals employed are all consistent with current anatomical and physiological data or the latest assumptions of these characteristics. The differences from the nervous system are generally in the direction of simplification, rather than complication, since it is often possible to achieve effects in an electronic model which would require many cells and connections in a biological system. At only one point — the assumed "value" of the A-units — is there an assumption which does not have a clearly identifiable counterpart in the biological brain, and this appears to be due to difficulties of measurement, rather than incompatibility of the concept.

(5) The perceptron is the first system which has proved capable of spontaneous organization and symbolization of its environment, along lines which bear some definite relationship to the human concept of similarity. While statistical schemes for the correlation and differentiation of patterns have been proposed previously, and might be implemented by a digital computer, the perceptron appears to be the only system which inherently operates in this fashion, as a property of its organization, rather than through the execution of a logical program.

Applications

The ultimate applications of a system such as the perceptron, if such a system can indeed be built economically, open possibilities which still seem difficult to imagine. In principle, the perceptron can not only read print and script, but can respond to verbal commands as well.

One stage beyond the level which now seems attainable by the perceptron, is the possibility of an automatic translator which can receive spoken inputs in one language and produce written or verbal outputs in another language. And it is possible that ultimately the coupling of a perceptron with a conventional digital computer may carry us over the remaining obstacles of grammar and syntax.

The application of such a system to library research and data gathering, for scientific purposes, is a definite possibility. In this application, the perceptron might be expected to digest and prepare abstracts of relevant material, as well as to locate references.

In the more distant future, automatic navigation and landing systems, automatic pilots, and recognition systems of every variety might make use of the perceptron. Finally, coming at a time when the scientific exploration of outer space is just getting started, the possibility of a robot passenger, capable of describing and classifying new environments, may make possible the completion of many useful explorations under difficult environmental conditions.

Extend Theory

Such speculation, however, cannot really be evaluated at this time. We must first extend the basic theory of the perceptron, which is still in its infancy. We must lower the cost of an A-unit to a few hundredths the cost of units which can now be built with conventional components. We must study the behavior of laboratory models in environments ranging from the simple mixtures of geometrical forms, which are simulated in our current programs, to such complex problems as the discrimination of speech and human faces. We must develop sensing devices suitable for providing visual and auditory inputs to the system.

This program is a major undertaking, and we cannot expect practical applications in the immediate future. Nonetheless, whether the next stage takes two years or ten years, it now seems clear that with the perceptron, a new field of research, both for engineering and for the theory of intelligent systems, has come of age.

REFERENCES

1. HAYEK, F. A. *The Sensory Order.* Univ. of Chicago Press, Chicago, 1952.
2. HEBB, D. O. *The Organization of Behavior.* John Wiley & Sons, New York, 1949.
3. ROSENBLATT, F. *The Perceptron: A Theory of Statistical Separability in Cognitive Systems.* CAL Report No. VG-1196-G-1, January, 1958.

ABOUT THE AUTHOR

FRANK ROSENBLATT, author of "The Design of an Intelligent Automaton," became interested in the problems of measurement and data analysis which appeared to be fundamental to scientific progress in psychopathology six or seven years ago. He was at that time a Fellow of the U.S. Public Health Service and was engaged in research on schizophrenia. Subsequently, his doctoral thesis dealt with the application of an analysis technique to problems of personality measurement. At the same time, however, Dr. Rosenblatt nurtured a growing conviction that the main content of psychology will become amenable to sound theoretical treatment only after a more secure basis is established, through an improved understanding of the biophysical processes in memory and cognition. With this thought in mind, Dr. Rosenblatt had already increased his emphasis on physiological problems and mathematical brain models.

His training in electronics and computer design stems largely from the construction of a special digital computer, the EPAC, which he built as an aid to the data analysis required for his thesis. Employed for the last three years as a Research Psychologist at Cornell Aeronautical Laboratory, Dr. Rosenblatt has made contributions to the design of information processing and weapon control systems, in addition to his work as Project Engineer responsible for the perceptron program.

Dr. Rosenblatt was born in New Rochelle, N. Y., in 1928. He attended the Bronx (N. Y.) High School of Science and graduated from Cornell University, where he majored in social psychology, in 1950. He received his PhD in Psychology, also from Cornell, in 1956.