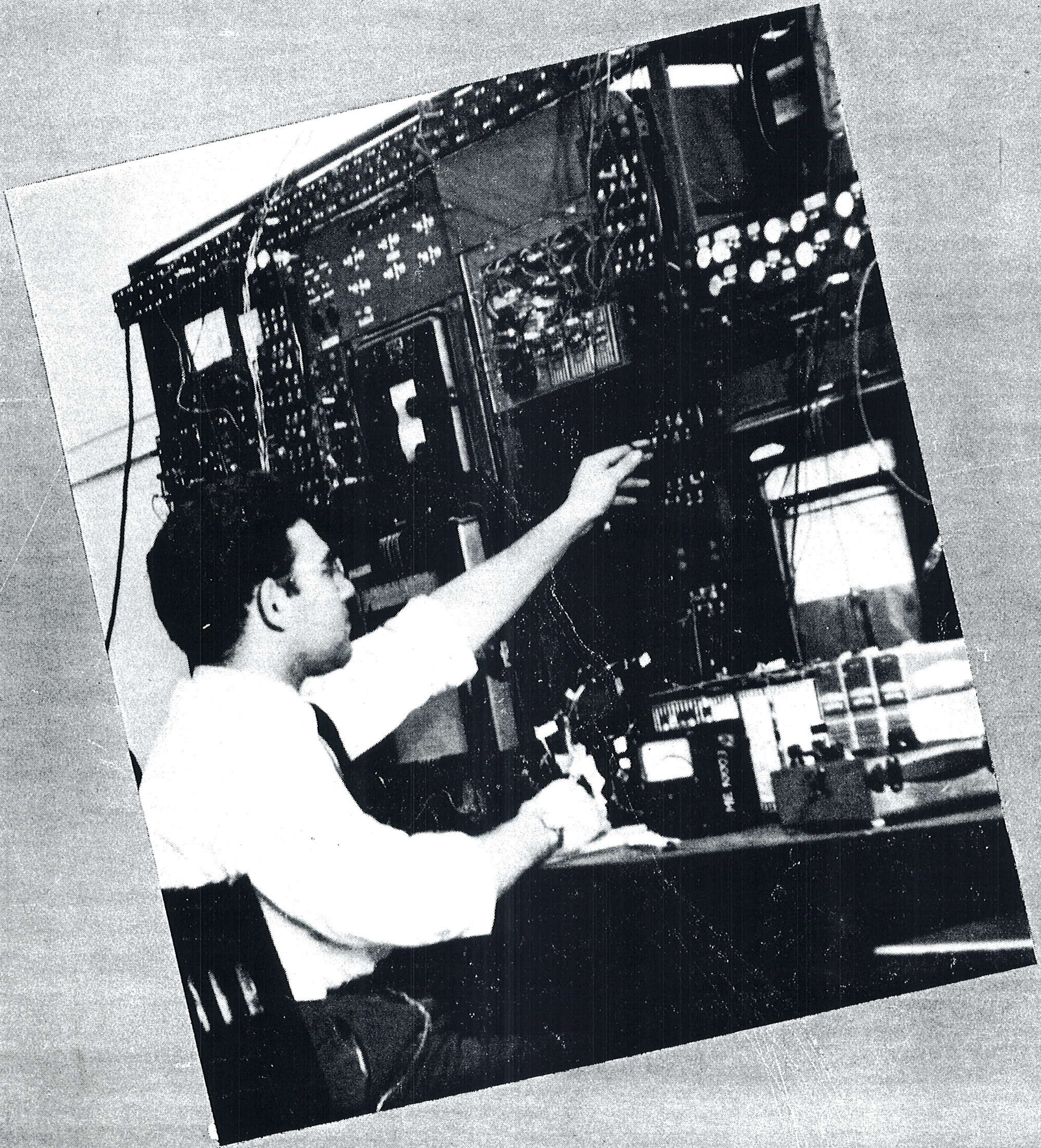
Engineering QUARTERLY



JANUARY 1950

JANUARY 1950

THINKING MACHINES

A New Field in Electrical Engineering

DR. LOFTI A. ZADEH
ELECTRICAL ENGINEERING DEPT.

"PSYCHOLOGISTS Report Memory is Electrical," "Electric Brain Able to Translate Foreign Languages is Being Built," "Electronic Brain Does Research," "Scientists Confer on Electronic Brain,"—these are some of the headlines that were carried in newspapers throughout the nation during the past year. What is behind these headlines? How will "electronic brains" or "thinking machines" affect our way of living? What is the role played by electrical engineers in the design of these devices? These are some of the questions that we shall try to answer in this article.

WHAT IS A THINKING MACHINE*

In the past five years and particularly since the publication of Wiener's "Cybernetics" (1948), an increasing number of scientists and laymen have become aware of a scientific development which promises to revolutionize not only many fields of science and engineering, but also our whole way of living. Briefly stated, what has happened is that on one side an increasing number of psychologists, neurologists, physiologists, and others, began to recognize that in many respects a human being behaves like a large-scale digital computer controlling a system of so-called "servomechanisms." On the other side, and at about the same time, mathematicians and electrical engineers have found that the body of knowledge which grew around the techniques used in automatic telephony and large-scale digital computers, had advanced to a point where it has become possible to design and construct electronic devices which function in much the same manner as a human being. These two converging developments have given rise to a new field of science, cybernetics, which despite its infancy has already profoundly influenced the trend of scientific thought in many time-honored fields of science.

Broadly speaking, cybernetics deals with the study of animate and inanimate systems from the point of view of control and communication. It has basically two branches. The first branch is devoted primarily to mathematization of the processes of control and communication in a human being. This branch of cybernetics is essentially macroscopic in nature and is not concerned directly either with the chemical processes going on inside the human body or with the mysteries of life.

The second branch of cybernetics is devoted to the study of the problems of control and communication in man-made devices, and the development of devices which can, in certain respects, approximate the higher forms of human activity. In this article we are interested in what amounts to a division of the second branch of cybernetics, namely, the principles and organization of machines which behave like a human brain. Such machines are now variously referred to as "thinking machines," "electronic brains," "thinking robots," and similar names.* The distinguishing characteristic of thinking machines is the ability to make logical decisions and to follow these, if necessary, by executive action.

A convenient introduction to the subject of thinking machines and their relation to other branches of cybernetics can be obtained by inspecting some of the recent work in this field. Typical of the work done in the first branch of cybernetics is the following abstract of a paper presented by A. H. Copeland before a meeting of the American Mathematical Society.

". . . A linear graph (neural network) is constructed which is capable of storing information, recognizing similarities between new and old experiences, reminiscing, and reasoning. The graph's memory is by association and it is shown that this association gives rise to a Boolean algebra which forms the basis of the reasoning mechanism. The Boolean algebra of reasoning is only indirectly related to the binary number system although this system plays an important role in the construction of the graph. In this model, inhibition is not produced by special inhibitory impulses, but rather by differences in firing thresholds and by timing. The assumption concerning differential thresholds has some experimental justification and its consistency is established by the model. Incidentally, with a moderate alteration the graph becomes the wiring diagram of an ultra rapid electronic device to do the work of an I.B.M. sorting machine."

The type of work done in the second branch of cybernetics can not well be illustrated, since most of the advanced research in this field is carried out for the Armed Forces and hence is classified. However, an idea of the principles involved in a thinking machine can be obtained from the description of a Tit-Tat-Toe playing device which was recently demonstrated by Robert Haufe of Caltech before a meeting of the American Institute of Electrical Engineers. In Haufe's words:

". . . The game consists of two players alternately placing markers (usually *x* and *o*) in a field of nine squares, formed by two vertical and two horizontal lines. The game is won by a player if he is able to get three of his markers in a straight line.

"The purpose of this machine is to alternate with a human player in filling the squares of an indicating device in the form of a Tit-Tat-Toe board. The machine consists of a num-

*The descriptive material in this and the following section lays no claim to technical rigor and accuracy. A thorough and precise discussion of the subject is beyond the scope of the present article.

*The same names are frequently ascribed to devices which are not "thinking machines" in the sense used in this article.

ber of relays and their associated equipment, among whose functions is to keep a record of which squares are filled by each party; to cause the machine to make a rational move each time it has its turn; and to determine the result of the game when it is finished.

"The nature of the circuits is such that each time the machine has its turn, all of the squares, whether they have been filled or not, are classified in order of decreasing strategic desirability; another circuit then investigates these classes successively until it finds a square which is empty. The machine is then caused to take this square. Auxiliary circuits perform other functions such as preventing the human player from taking more than one square at each turn, preventing further play after the game has been won, and restoring the relays to their initial state at the beginning of each new game.

"The strategy incorporated into the machine is such that under any ordinary circumstances the machine will either win or force the human player to a draw. In order to make the game more interesting, however, a switch has been built into the machine which, when thrown, permits a skillful player to win."

Haufe's machine is much simpler than an average thinking machine. A better indication of the capabilities of such machines can be gained from the fact that it is now possible to design a thinking machine that will play a reasonably good game of chess against a human adversary; and that—what is perhaps more striking—the machine can be made to improve its game automatically through self-analysis of its experience.

Despite its simplicity, Haufe's machine is typical in that it possesses a means for arriving at a logical decision based on evaluation of a number of alternatives. More generally, it can be said that a thinking machine is a device which arrives at a certain decision or answer through the process of evaluation and selection. On the basis of this definition it follows that many devices that have been in use for years should be classified as thinking machines. For example, an ordinary sorter is a very elementary thinking machine. A telephone exchange is a complicated, though not very "brilliant" thinking machine. On the other hand, the mere ability to perform difficult computations does not so qualify a device. Thus, M.I.T.'s differential analyser is not a thinking machine, for it can not make any decisions, except trivial ones, on its own initiative. However, the recently built large-scale digital computers such as UNIVAC and BINAC are endowed with the ability to make certain non-trivial decisions and hence can be classified as thinking machines. In UNIVAC, BINAC, and most other decision-making computers, the ability to think is very limited. It is used almost wholly for computational purposes and does not constitute an end by itself. Thus, despite their stupendous computational abilities, the modern large-scale digital computers are relatively narrow thinkers. In fact, at the present time most of the "brilliant" thinking machines are being designed not for the purpose of computing, but for the direction of operations involving selection, direction, and distribution.

It should be emphasized at this point that all thinking machines are special purpose devices. Thus a machine designed to play chess could not be used for any significantly different purpose without undergoing extensive alteration. The contrasting versatility of the human brain can not but fill a designer of thinking machines with a feeling of admiration and humility.

HOW A THINKING MACHINE WORKS

The basic elements of a typical thinking machine are shown in Fig. 1. The roles played by these elements and their interconnections in performing the process of "thinking" are as follows:

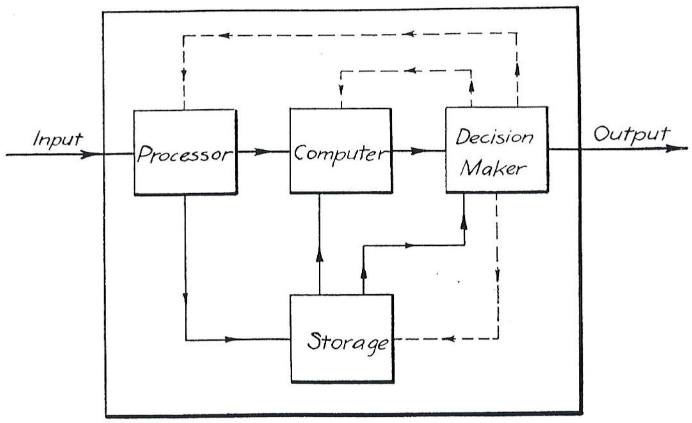


Figure 1—A schematic diagram illustrating how the basic elements of a thinking machine are arranged.

The input data are accepted in the box labeled *Processor*, where they are sorted and assimilated. A part of the processed data is sent to the box labeled *Storage*, where it is stored for future or delayed use. This element may have the form of a file of punched cards, a tape recorder, a delay line (for short-term storage), a cathode-ray tube such as the Selectron, etc., and it has the same function as memory in a human brain.

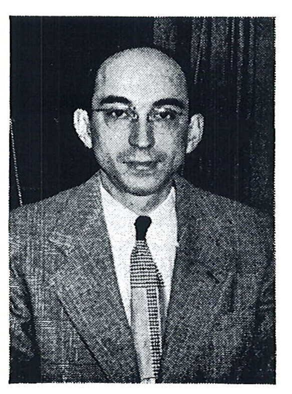
Another part of the processed data, as well as some data from *Storage* are fed into the box labeled *Computer*, where necessary computations are performed. The *Computer* is not an essential part of a thinking machine except where computation is either the end result, or where it is needed in the course of arriving at the desired decision.

(*Continued on page* 30)

THE AUTHOR

Dr. Lotfi A. Zadeh received the degree of Bachelor of Science in electrical engineering from the University of Teheran, Teheran, Iran, in 1942. In 1944, after working for a year as a technical contractor with the United States Army Forces in Iran, he came to the United States and for a brief period was associated with International Electronics Laboratories, New York City. Resuming his studies in 1945, he received the degree of Master of Science from Massachusetts Institute of Technology in 1946, and the Ph.D. degree from Columbia University in 1949. Since 1946 he has been an instructor in electrical engineering at Columbia University.

Dr. Zadeh is an associate member of the American Institute of Electrical Engineers and the Institute of Radio Engineers. He is a member of the American Physical Society, the American Mathematical Society, and Sigma Xi.



THINKING MACHINES—L. A. Zadeh

(Continued from page 13)

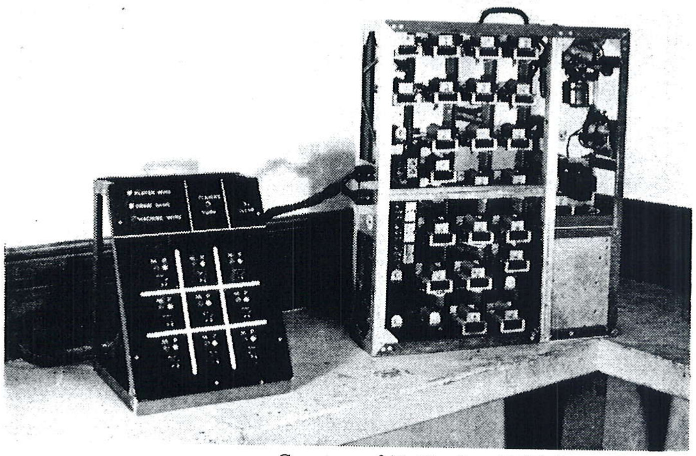
The box labeled *Decision Maker* is the most important part of a thinking machine. The function of this element is to make logical decisions and also to serve as the clearing-house for the rest of the system. Here all relevant information supplied by *Computer* and *Storage* is examined and weighted in accordance with the instructions and criteria set into the machine, and on the basis of these the final answer or decision is produced as the output. The manner in which the *Decision Maker* operates will be described in greater detail at a later point.

The broken lines leading from the *Decision Maker* to the other three elements of the machine represent the so-called feedback connections. The function of these is to make the operation of the other three elements dependent, if necessary, upon the results obtained in the *Decision Maker.* For example, in a certain application it might be required that the standards used in processing the input data be governed by the nature of the decision arrived at by the *Decision Maker.* Such a situation would be indicated by a feedback connection between the *Decision Maker* and the *Processor.*

To provide an illustration of how a machine of the type described above would operate in practice, let us take a look into the future and imagine the following scene. (The example should not be taken seriously except as it serves to illustrate the operation of the machine).

It is the year 1965. Three years ago for reasons of economy and efficiency the trustees of Columbia University have decided to disband the Office of University Admissions and to install in its place a thinking machine to be called the Electronic Director of Admissions. The design and construction of the Director was entrusted to the Department of Electrical Engineering, and the installation was completed in the spring of 1964. Since then the Director has been functioning perfectly and has won the unanimous acclaim from administration, faculty and student body alike.

The mode of operation of the Electronic Director of Admission is quite simple. The applicant's records are first processed by a human receptionist who expresses their content in the form of n numbers $a_1, a_2, a_3, \ldots, a_n$, each evaluating an item in the record. For example, a_1 might stand for the I.Q. of the applicant, a_2 for the evaluation of a personality trait, and so forth.



Courtesy of R. Haufe—Calif. Inst. of Tech.

Figure 2—The two units of the Tit-Tat-Toe machine.

Next, the applicant's record, as expressed by n numbers a_1, a_2, etc., is suitably coded and then fed into the *Processor.* There it is processed and routed in part to *Computer* and in part to *Storage.* The function of the *Computer* is to calculate, on the basis of the applicant's and the University's records, the probability of various events, such as for example, the probability of failure after the first year of study, the probability of getting elected to Phi Beta Kappa, and so forth. This data, together with the data from *Storage* are fed into the *Decision Maker* where the final decision regarding the admission of the applicant is made. A sample of the criteria used in judging admissibility is given below:

1. Accept, if the probability of becoming a Phi Beta Kappa is greater than 60%.

2. Reject, if the probability that the applicant will not survive his first year in College is greater than 20%.

3. Accept, if certain conditions (formulated in terms of the a_1, a_2, etc.,) are satisfied and if the estimated registration is less than a prescribed number N.

And so forth.

The type of machine sketched above can be made as crude or as elaborate as it may be desired. It is not as fantastic as it may appear. In fact, such machines may be commonplace in anywhere from ten to twenty years hence. Furthermore, it is absolutely certain that thinking machines will play a major role in any armed conflict that may arise in the future.

As was mentioned previously, the part of the machine that is most essential for the ability to "think," is the decision making element. The *Decision Maker* is essentially a device that arrives at its conclusions in accordance with the rules of logic. In the case of the Electronic Director of Admissions, the *Decision Maker* had to decide whether a set of numbers representing the applicant's record and the probabilities of certain events, did satisfy certain prescribed criteria for admission. More generally, the function of the *Decision Maker* would be to deal with statements, classes and relations, The well known Boolean algebra, or the calculus of propositions, deals with such statements, classes and relations in an algebraic manner. Thus, in Boolean algebra, if A and B are certain statements, then the composite statement $A \cdot B$ means A and B; $A > B$ means A implies B; $A \vee B$ stands for A and/or B, and so forth. Every statement is either true or false, or in other words, it must have one of the two truth values—True and False. The rules of Boolean algebra permit the determination by a purely algebraic process whether a composite statement such as $(A \cdot B) > B$ (this is a very simple example) is true or false when the truth values of the component statements are known. Thus, through the use of Boolean algebra it is possible to reduce a wide class of logical processes to a series of algebraic operations.

In a decision-making device, the algebraic operations of Boolean algebra are simulated by means of electrical circuits consisting of relays and switches. The analogy existing between the basic operations in Boolean algebra and the behavior of relay circuits, is illustrated in the following table.

Relay Circuit Element	Symbolic Logic Interpretation
Circuit A	Statement A
Closed circuit	A is false
Open circuit	A is true
Series connection of A and B	A and/or B ($A \vee B$)
Parallel connection of A and B	A and B ($A \cdot B$)

In a thinking machine the decision-making element is a part, albeit an essential part, of the whole machine. There is, however, considerable interest, particularly on the part of logicians, in machines which would serve only for the purpose of simulating logical processes. Such machines are sometimes referred to as logical machines. An example of such a machine is a device constructed at the Harvard Computation Laboratory by Kalin and Burkhart.

Thinking machines are essentially electrical devices. But, unlike most other electrical devices, they are the brain children of mathematicians and not of electrical engineers. Even at the present time most of the advanced work on thinking machines is being done by mathematicians. This situation will last until electrical engineers become more proficient in those fields of mathematics which form the theoretical basis for the design of thinking machines. The most important of these fields is that of symbolic logic.

It is true that most of the fundamental principles on which thinking machines are based, have been contributed by mathematicians. But it is equally true that these principles could not be evolved into a physically realizable thinking machine, were it not for the ability of electrical engineers to supply the techniques that make possible the storage devices, processors, computors, decision makers, and other less important elements of thinking machines. Through their association with mathematicians, the electrical engineers working on thinking machines have become familiar with such hitherto remote subjects as Boolean algebra, multivalued logic, and so forth. And it seems that the time is not far distant when taking a course in mathematical logic will be just as essential to a graduate student in electrical engineering as taking a course in complex variable is at the present time. Time marches on.

The Department of Electrical Engineering at Columbia is doing its share in the development of thinking machines. Apart from the work done for the Armed Forces on a consulting basis, basic research has been initiated on the problem of synthesis of so-called fixed memory and variable memory systems. The latter systems could lead to thinking machines in which the number of possible decisions would not necessarily be finite, as is the case with ordinary decision-making devices. It is to be hoped that this and other projects will aid the Department to maintain an advanced position in the new and rapidly developing field of the thinking machine.