ADALINE: Smarter than Sweet

Stanford engineers have given this maidenly nickname to their adaptive linear computer. She can forecast the weather, play blackjack, recognize voices, and solve all kinds of problems on her own.

by BERNARD WIDROW

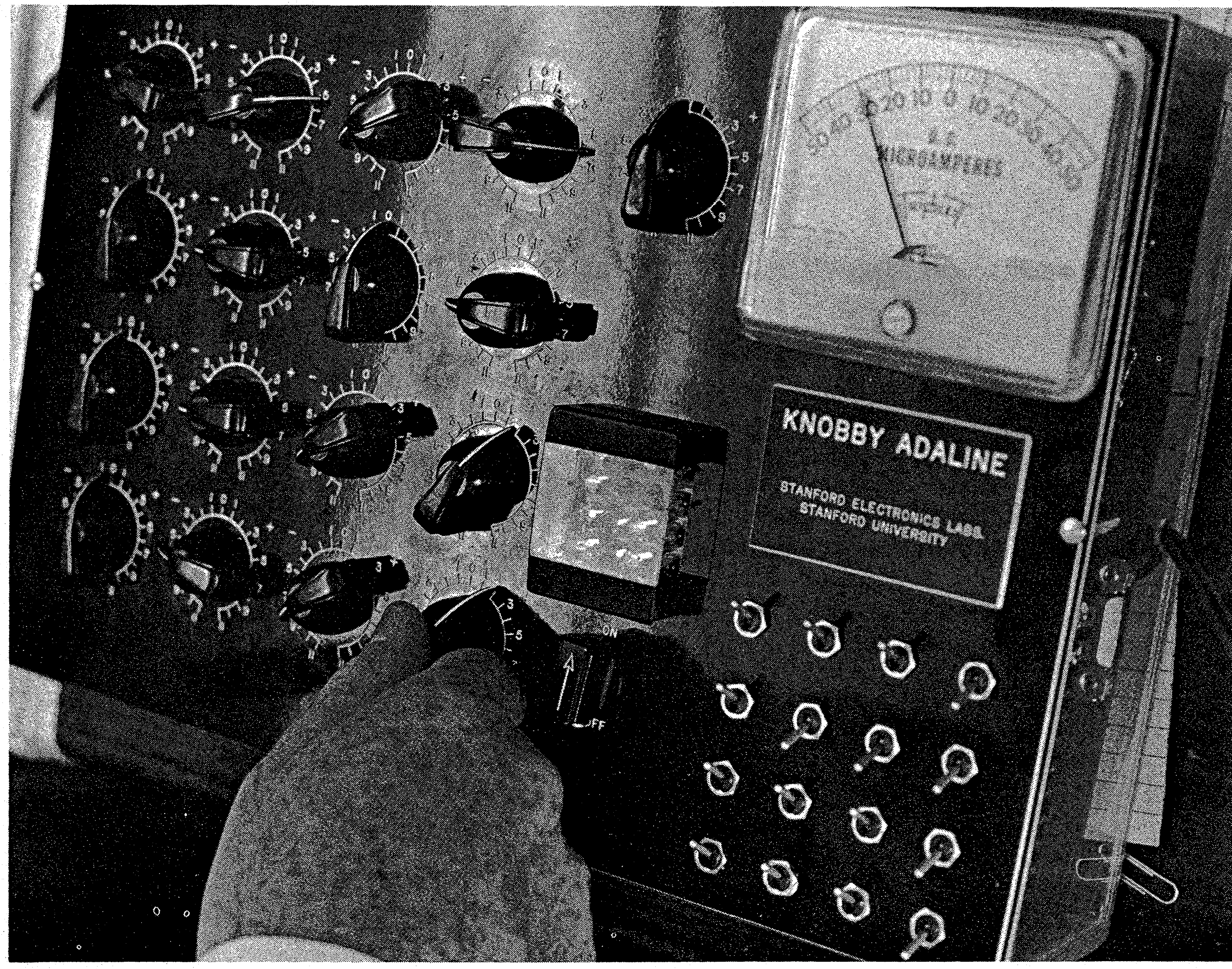
If we had been told ten years ago that a bunch of computers somewhere near Washington, D.C. would calculate our income tax and then check up on us, most of us would have been amazed and maybe even a little frightened. In those days it was common to refer to computers awesomely as "electronic brains." Today we accept them rather passively. We realize that these devices are really great big arithmetic machines which have to be instructed by human brains in every detail of their operation.

This is true, at least, of the widely-used digital and analog computers, both of which rely on stored information, or programs. But in our laboratories we are working on a third kind of data-processing system, a radical departure from the first two, which more nearly deserves the name of "electronic brain." It cannot think and it is not remotely intended to replace the human brain, but it definitely can learn, discriminate, and remember.

We have taught our Adaline (for adaptive linear neuron) and our Madaline (for multiple-Adalines) to balance a rod, forecast the weather, read electrocardiograms, and type out simple sentences that are spoken to it. Before long we expect them to repeat the sentences out loud, and larger Madaline systems could be taught to drive a car, fly a plane, or perform countless other jobs.

The human who deals with the adaptive computer, rather than being a programmer as in the case of the analog or digital machines, is more of a teacher. He gets the adaptive computer to solve problems by "showing" it examples of what he would like done. From a limited number of training examples the adaptive computer is able to generalize and to infer a way of solving new problems that are related to the training problems but are not identical.

The teacher has to repeat his message to the computer over and over and over again until the message "socks in," but once it's in, the computer can make sense out of a lot of related things. It is something like listening to a symphony. A basic theme is introduced and developed and embellished and repeated, and by the end of the movement you get the message. If you really want to get it, you listen to the symphony many



times, and every time you hear it it's more beautiful because you make more and more associations and connections. This is what happens in the adaptive computer.

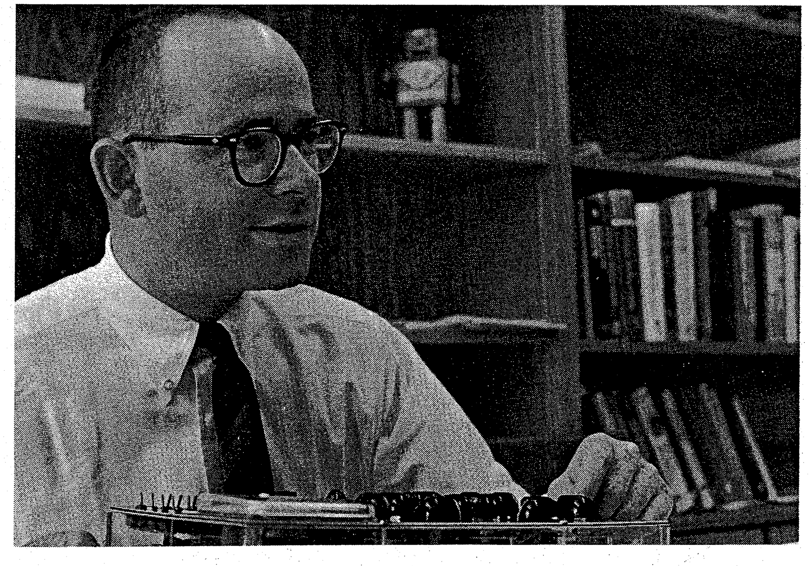
In a very recent experiment, we found that an Adaline can even learn without a teacher, at least in a very limited way. A digital computer played the part of dealer in a game of blackjack, or "twenty-one"; the Adaline was the player. The Adaline was given certain arbitrary information about card values but was not taught the strategy of the game. Every time she won a hand she was "rewarded" by having her responses reinforced. Whenever she lost she was "punished" by having her responses diminished. After several hundred hands, she developed a blackjack strategy quite close to a theoretical optimum. It is too soon to predict with any assurance, but apparently we are at the beginning of a completely new phase of adaptive systems research, that of autonomous learning.

We have also built systems which can learn many skills simultaneously. But, like the human, they tend to be more expert at the things they have seen most recently. If you don't repeat the old with the new, the old responses will wash out, but if you repeat them a little bit you can regenerate the old rather quickly and have the new in addition.

It seems quite clear that the systems we have evolved here, although we're not trying to make them so, are analogous in their action to certain of the functions performed by neurons, the key building blocks of living nervous systems, particularly in their ability to draw specific reactions from gross experience.

Science has no completely-founded explanation of how our brains remember things over long periods of time. We believe that our adaptive systems may provide some insight toward explaining this phenomenon.

THE BASIC CIRCUIT of the adaptive computer is an adaptive threshold logic block which can recognize and store patterns of information through repetitive conditioning. Photo cells, microphones, manually-operated measuring equipment, or a wide variety of other sensory



Bernard Widrow must be put in the group that has come to be known as "the new breed" or "the take-over generation." He himself would laugh off any such showboat labels, but the fact remains that at 34 (and he *looks* even younger) he is well known in this country and abroad among electronics engineers for his original contributions in adaptive computers. He is an intense, tireless worker, a rapid-fire talker, a transcontinental commuter, and in every way fits the picture. Yet to all of this he adds a good humor and a helpful, straightforward attitude toward students which make him an effective teacher at both the undergraduate and graduate levels.

A native of Norwich, Connecticut, he took all of his degrees at M.I.T.: S.B. in '51, S.M. in '53, and Sc.D. in '56. He had worked there on the staff of the Lincoln Laboratory and as assistant professor when he was appointed assistant professor at Stanford in 1959. Men usually stay in this grade a minimum of three years but at the end of two Stanford jumped him to his present rank of associate professor.

He and his wife and their two girls (three years old and five months old) live in a redwood-and-glass home in Pine Hill, the faculty subdivision on campus.

input devices allow the computer to "see," "feel," or "hear" its instructions. This is the training phase. When the computer is turned loose on a new problem, it identifies patterns in the problem which it associates with patterns in its stored experience.

Let's illustrate this by the way a doctor observes a multitude of symptoms, some precisely measured, such as temperature or blood pressure, and some more subtle, such as coloring, pain patterns, or demeanor. What the doctor does is almost subconsciously to attach a weight or a significance to each of the symptoms, based on his gross experience with many diseases and many patients, and he combines these effects to arrive at a diagnosis.

It is entirely possible to build an adaptive computer with a set of dials for pulse rate, temperature, X-ray readings, and other medical symptoms. The physician would feed information into the computer in the form of electrical signals whose strength would be proportional to the dial setting. The computer, on the basis of its training for the disease in question, knows how much weight to attach to each signal. These are finally combined in the summing device, and the sum either exceeds or falls short of a threshold value—either yes, the patient does have the disease, or no, he does not. The answer would probably come in the form of a meter reading which would show a degree of sickness or wellness.

The advantage of the computer is that it can be trained on the experiences of many doctors—the best men in each of the specialties. Every time the computer sees another case, it learns a little bit—in other words, the weights change—until at the end of the training period it comes up with an integrated experience. But I emphasize that the skill of the individual doctor is still the key to a successful diagnosis. Like the stethoscope, the computer simply extends and enhances his knowledge.

Applications of adaptive computers to medical and other fields are not our primary concern at Stanford. Our job is to develop the basic systems and components. We feel that we are already in a position of leadership, but we want to grow because we want to push ahead in this major new field of information processing. In ten or fifteen years the impact of adaptive computation on society should be comparable to that of the digital and analog computers today—and these machines, particularly the digital, are just in their infancy.

Among the other professors closely identified with this activity are William R. Rambo, director of the Stanford Electronics Laboratories and head of the Systems Techniques Laboratory; John G. Linvill, head of the Solid State Laboratory; Gene F. Franklin, head of the Systems Theory Laboratory; James B. Angell; and Richard L. Mattson. We have worked closely with the faculty in medicine, speech audiology, and statistics and with staff members of the Stanford Computation Center, Palo Alto Medical Research Foundation, and U.S. Weather Bureau at the San Francisco International Airport.

Graduate students are also an integral part of our research operation. We are blasting into a wide-open, completely new field that is a bottomless pit of Ph.D. thesis topics.

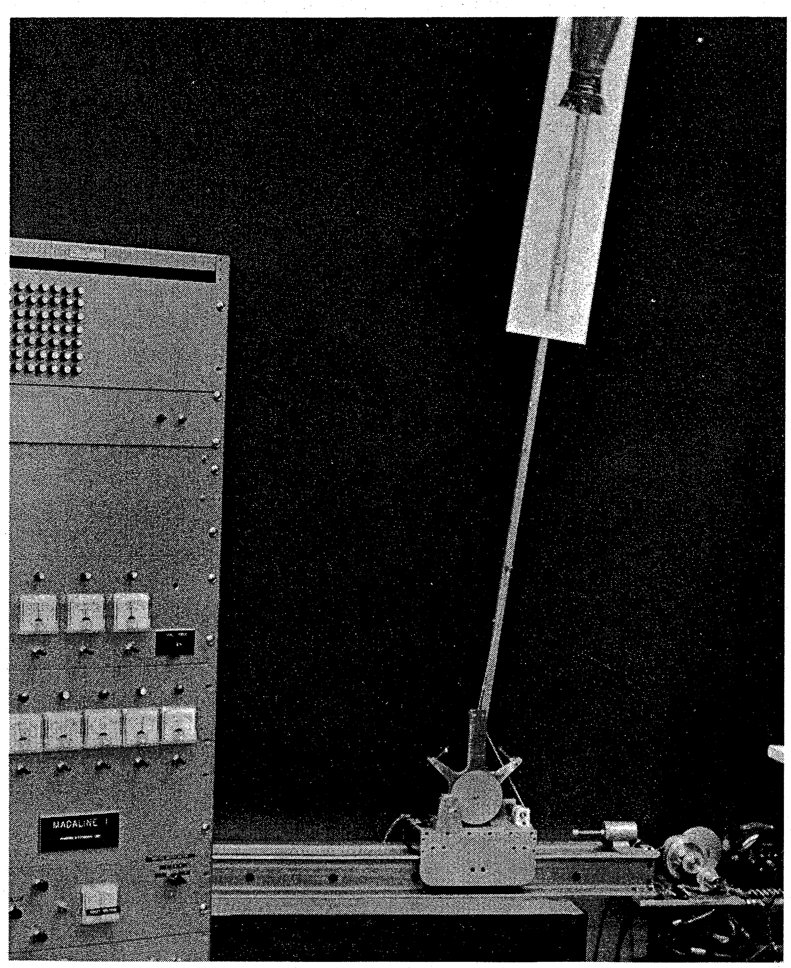
My first doctoral student when I came to Stanford four years ago was Marcian Hoff. Together we developed some of the original training procedures for the Adaline and the first successful adaptive component to provide the various weights and significances that must be given to the computer's stored experience. This component, called a memistor, is electro-chemical (we were inspired by the liquid state electronic system of the brain). A variable resistance effect is obtained in the memistor by varying the thickness of electroplated metallic films. In a given Adaline there will be many memistors, each one corresponding to a particular weight or significance factor.

Traditionally a theory is developed mathematically before anything is attempted experimentally. But the problems we deal with are so complex and so many approximations have to be made that we have found it profitable to go directly into the experimental stage. We get an idea of how something might be done and we go ahead and build it. Usually it doesn't work. We fiddle with it until it does. Then we sit down to find out why.

Michael J. C. Hu, a graduate student from Hong Kong, is conducting experiments based on weather forecasting. He fed into an Adaline the patterns of atmospheric pressure readings on selected days from over 500 million square miles extending from Alaska to northern Mexico and from Hawaii to the Rockies. Correlated in the training were the patterns of actual weather in the Bay Area on corresponding days. Then he asked Adaline to make predictions based on the pressure data alone for 18 days on which it had not been trained. Adaline scored 78 per cent right on today forecasts, 89 per cent on tonight forecasts, and 83 per cent on tomorrow forecasts. For the same 18 days, U.S. Weather Bureau forecasters scored 78, 89, and 67 per cent. Hu is now working on a much more comprehensive study.

One of our graduate students, Donald Specht, conceived the idea of using adaptive principles to forecast the stock market. This didn't work out very well so he switched to diagnosing electrocardiograms, with the advice and data of Drs. Von der Groeben and Toole of the School of Medicine. Information derived by these cardiologists was quantified in digital form and presented to the adaptive machine as a pattern. In one experiment the system was trained on 50 cases, half normal and half abnormal. The system was then presented with 50 cases it had never seen before and it discriminated normals from abnormals with a reliability of about 85 per cent.

Our work in voice recognition was begun two summers ago by a group of graduate students. They had just finished final exams and decided to blow off steam by devising an electronic system. They filtered sounds into groups of different frequency components and presented them as patterns to a Madaline. The teacher speaks single words into a microphone and lets the system know which pattern this word fits into. The Madaline can assimilate a vocabulary of 20 words in about an hour. If the person who did the training then speaks sentences into the microphone using this vocabulary, Madaline will type out the sentences almost letter perfect. When another person does the speaking, the difference in voice inflection may cause reliability to drop to about 90 per cent. When we had a group of NATO visitors, we asked Frenchmen and Germans to address the machine in English and in spite of the accents Madaline came through with 85 per cent



Balancing a broom with a computer is not idle fun. It's a complicated study of trainable control systems.

accuracy. More training would increase her reliability and vocabulary.

We discovered the inherent ability of adaptive computers to ignore their own defects while we were rushing through construction of a system called Madaline I for presentation at a technical meeting. The machine was finished late the night before the meeting and the next day we showed some very complex pattern discriminations. Later we discovered that about a fourth of the circuitry was defective. Things were connected backward, there were short circuits, and poor solder joints. We were pretty unhappy until it dawned on us that this system has the ability to adapt around its own internal flaws. The capacity of the system is diminished but it does not fail. This ability is becoming more and more important with the increasing use of microelectronic systems which are difficult to repair.

Balancing a rod may sound like a rather childish game but actually it is a complicated investigation of decision-making in control systems. The rod is mounted with a swivel joint on a small cart. The cart is on a track and is controlled by a motor that is always either full forward or full reverse. A Madaline keeps the rod upright by switching the cart back and forth, making decisions which depend on the state of affairs in the system at a given instant of time. It also must reverse the cart and rebalance the rod before the cart crashes into a stop at the end of the track—in other words, sequential decisions are required.

We are working on a more sophisticated version of the trainable balancer in which a retina of photocells observes an illuminated rod. It will balance the rod, after training, from the sequence of patterns observed by its "artificial eyeball." Pattern-recognizing control systems appear to be extremely flexible and should make possible economical and reliable automation and control of highly complex processes—including some whose complexities defy mathematical description and analysis.

The potential applications seem almost endless. We have cited a few examples. The possibilities in machine and process industries, management, teaching machines and programmed learning, information retrieval—these and many more leap to the mind. It is evident too that applications to military systems are legion, and we are indebted to a number of agencies in the Army, Air Force, and Navy, whose interest and financial support have made possible much of our basic research in the adaptive computer field.

Language translation with a conventional digital computer has been only a limited success thus far. The big problem is that a single word in, say, French might have ten different meanings in English. The number of English sentences that you could get from a ten-word French sentence would be ten to the tenth power—ten billion sentences, most of which would not make sense. The English equivalent you choose has to be based on context, and this is a matter of memory and association, both of which the adaptive computer is capable of.

To gain widespread application of adaptive computers, it will be necessary to produce replicas of prototype systems. We can picture, for instance, some day providing physicians with a little cracker box about the size of a transistor radio with a bunch of small dials and a meter on the front. The pre-determined weights and pattern significances for a certain diagnostic problem would be permanently stored in the box by means of fixed resistors. Such a device might even be attached to a seriously ill patient in the hospital for continuous recording. A nurse could read it and see whether or not she should call the doctor.

We're pumping something really new into the automation picture with the adaptive computer, and its potential should be viewed as a blessing. People are not going to be idled; they're going to be doing different things. If we are clever enough to automate things that are repetitive, we're clever enough to think of new industries. Already a lot of people, instead of shoveling coal, are making semiconductor devices, building television sets, or doing other jobs that are more healthful and rewarding. And if you want to think really far out, think of the new travel industry that is coming up. We'll be taking weekend trips to the moon and climb in the craters.

It follows that I do not believe the adaptive computer will undermine the human intelligence, rendering it obsolete or weak from disuse. The fact is that as we go deeper and deeper into making machines that are more powerful and flexible, the more we appreciate the human intellectual capability. It is inconceivable at this point that these machines will replace man in any real sense.

Stanford Today

AUTUMN 1963

THE "NEW MATHEMATICS" REALLY ISN'T

Mathematics moves in the most sophisticated circles of science, engineering, and behavioral studies. Yet most pre-college math teaching is chained to trivial business applications. The Stanford SMSG program is raising sights from kindergarten on.

TEACHING MACHINES AND CREATIVITY

Can technically-aided learning and free-wheeling "discovery thinking" prosper side by side? Ernest R. Hilgard believes they can.

ADALINE: SMARTER THAN SWEET

This is about a space-age computer nicknamed for an old-fashioned girl. Stanford engineers can teach her to solve all manner of problems on her own, without the programming other computers need. Bernard Widrow is the author.

STANFORD TODAY

Second class postage paid at Palo Alto, California. Issued four times a year, in March, June, September, and December.

SERIES 1, No. 6 SEPTEMBER 1963

STANFORD TODAY is produced quarterly by the Publications Service of Stanford University. *Editor:* Peter C. Allen, '36 • *Assistant Editor:* Ralph P. Handy, '51 • *Design:* Walter C. Peterson, '38 • *Photography:* José Mercado

Please address any correspondence regarding this publication to STANFORD TODAY, STANFORD, CALIFORNIA. When notifying of address change, please clip and enclose the address panel from back cover.