6

CYBERNETICS

NORBERT WIENER • November 1948

Cybernetics is a word invented to define a new field in science. It combines under one heading the study of what in a human context is sometimes loosely described as thinking and in engineering is known as control and communication. In other words, cybernetics attempts to find the common elements in the functioning of automatic machines and of the human nervous system, and to develop a theory which will cover the entire field of control and communication in machines and in living organisms.

It is well known that between the most complex activities of the human brain and the operations of a simple adding machine there is a wide area where brain and machine overlap. In their more elaborate forms, modern computing machines are capable of memory, association, choice and many other brain functions. Indeed, the experts have gone so far in the elaboration of such machines that we can say the human brain behaves very much like the machines. The construction of more and more complex mechanisms actually is bringing us closer to an understanding of how the brain itself operates.

The word cybernetics is taken from the Greek *kybernetes,* meaning steersman. From the same Greek word, through the Latin corruption *gubernator,* came the term governor, which has been used for a long time to designate a certain type of control mechanism, and was the title of a brilliant study written by the Scottish physicist James Clerk Maxwell 80 years ago. The basic concept which both Maxwell and the investigators of cybernetics mean to describe by the choice of this term is that of a feedback mechanism, which is especially well represented by the steering engine of a ship. Its meaning is made clear by the following example.

Suppose that I pick up a pencil. To do this I have to move certain muscles. Only an expert anatomist knows what all these muscles are, and even an anatomist could hardly perform the act by a conscious exertion of the will to contract each muscle concerned in succession. Actually what we will is not to move individual muscles but to pick up the pencil. Once we have determined on this, the motion of the arm and hand proceeds in such a way that we may say that the amount by which the pencil is not yet picked up is decreased at each stage. This part of the action is not in full consciousness.

To perform an action in such a manner, there must be a report to the nervous system, conscious or unconscious, of the amount by which we have failed to pick up the pencil at each instant. The report may be visual, at least in part, but it is more generally kinesthetic, or to use a term now in vogue, proprioceptive. If the proprioceptive sensations are wanting, and we do not replace them by a visual or other substitute, we are unable to perform the act of picking up the pencil, and find ourselves in a state known as ataxia. On the other hand, an excessive feedback is likely to be just as serious a handicap. In the latter case the muscles overshoot the mark and go into an uncontrollable oscillation. This condition, often associated with injury to the cerebellum, is known as purpose tremor.

Here, then, is a significant parallel between the workings of the nervous system and of certain machines. The feedback principle introduces an important new idea in nerve physiology. The central nervous system no longer appears to be a self-contained organ receiving signals from the senses and discharging into the muscles. On the contrary, some of its most characteristic activities are explainable only as circular processes, traveling from the nervous system into the muscles and re-entering the nervous system through the sense organs. This finding seems to mark a step forward in the study of the nervous system as an integrated whole.

The new approach represented by cybernetics—an integration of studies which is not strictly biological or strictly physical, but a combination of the two—has already given evidence that it may help to solve many problems in engineering, in physiology and very likely in psychiatry.

This work represents the outcome of a program undertaken jointly several years ago by the writer and Arturo Rosenblueth, then of the Harvard Medical School and now of the National Institute of Cardiology of Mexico. Dr. Rosenblueth is a physiologist; I am a mathematician. For many years Dr. Rosenblueth and I had shared the conviction that the most fruitful areas for the growth of the sciences were those which had been neglected as no-man's lands between the various established fields. Dr. Rosenblueth always insisted that a proper exploration of these blank spaces on the map of science could be made only by a team of scientists, each a specialist but each possessing a thoroughly sound acquaintance with the fields of his fellows.

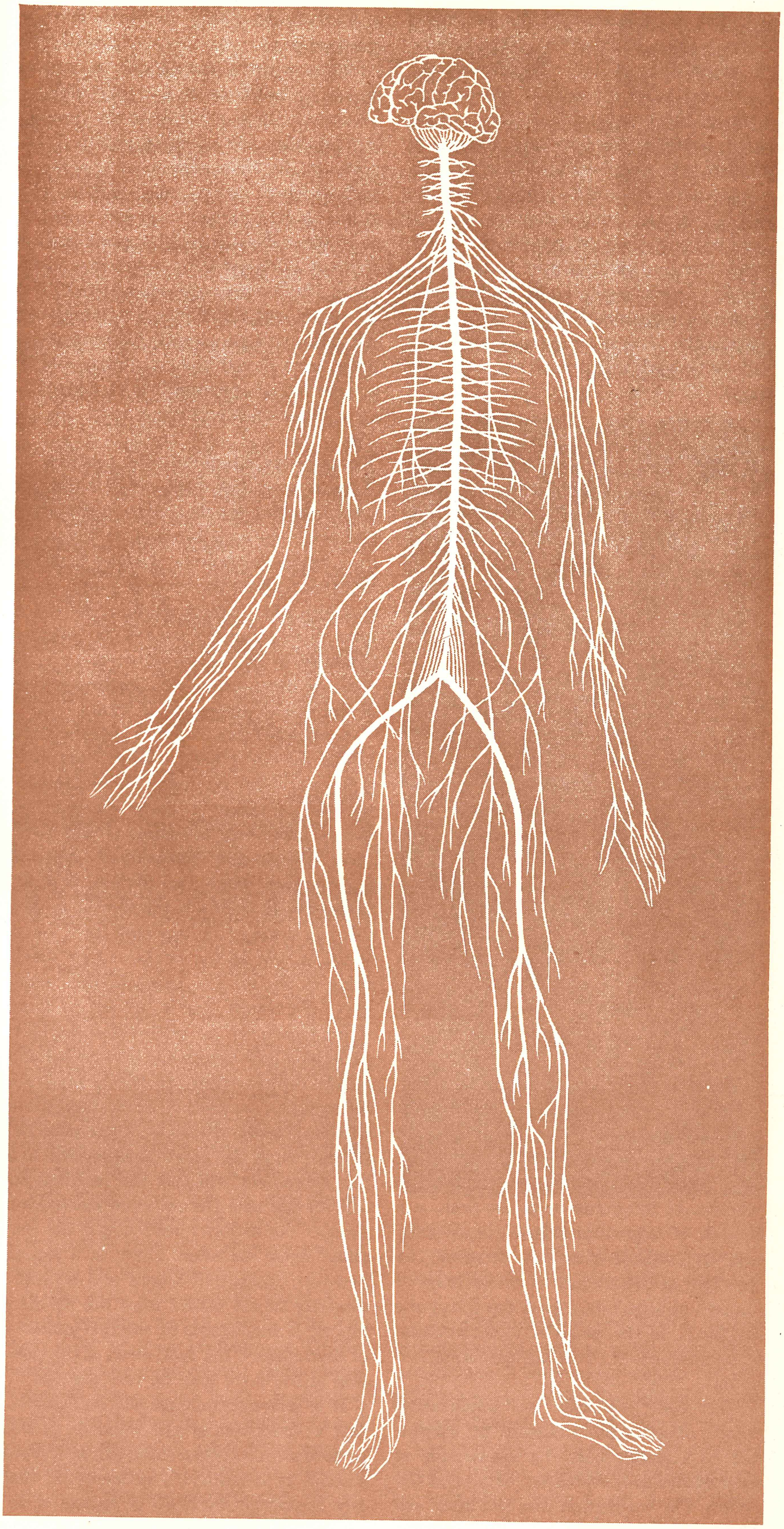
Our collaboration began as the result of a wartime project. I had been assigned, with a partner, Julian H. Bigelow, to the problem of working out a fire-control apparatus for anti-aircraft artillery which would be capable of tracking the curving course of a plane and predicting its future position. We soon came to the conclusion that any solution of the problem must depend heavily on the feedback principle, as it operated not only in the apparatus but in the human operators of the gun and of the plane. We approached

Dr. Rosenblueth with a specific question concerning oscillations in the nervous system, and his reply, which cited the phenomenon of purpose tremor, confirmed our hypothesis about the importance of feedback in voluntary activity.

The ideas suggested by this discussion led to several joint experiments, one of which was a study of feedback in the muscles of cats. The scope of our investigations steadily widened, and as it did so scientists from widely diverse fields joined our group. Among them were the mathematicians John von Neumann of the Institute for Advanced Study and Walter Pitts of Massachusetts Institute of Technology; the physiologists Warren McCulloch of the University of Pennsylvania and Lorente de No of the Rockefeller Institute; the late Kurt Lewin, psychologist, of M.I.T.; the anthropologists Gregory Bateson and Margaret Mead; the economist Oskar Morgenstern of the Institute for Advanced Study; and others in psychology, sociology, engineering, anatomy, neurophysiology, physics, and so on.

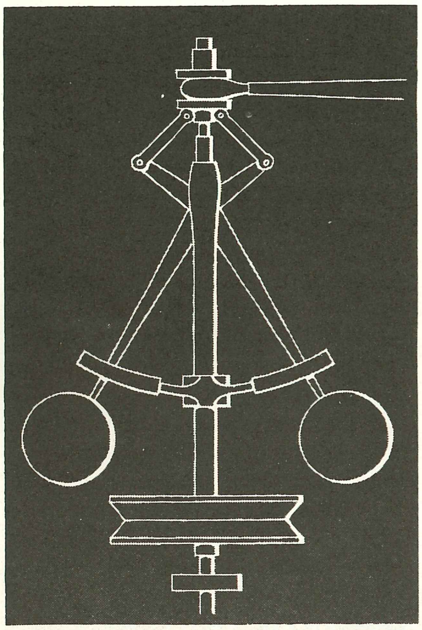
The study of cybernetics is likely to have fruitful applications in many fields, from the design of control mechanisms for artificial limbs to the almost complete mechanization of industry. But in our view it encompasses much wider horizons. If the 17th and early 18th centuries were the age of clocks, and the latter 18th and 19th centuries the age of steam engines, the present time is the age of communication and control. There is in electrical engineering a division which is known as the split between the technique of strong currents and the technique of weak currents; it is this split which separates the age just passed from that in which we are living. What distinguishes communication engineering from power engineering is that the main interest of the former is not the economy of energy but the accurate reproduction of a signal.

At every stage of technique since Daedalus, the ability of the artificer to produce a working simulacrum of a living organism has always intrigued people. In the days of magic, there was the bizarre and sinister concept of the Golem, that figure of clay into which the rabbi of Prague breathed life. In Isaac Newton's time the automaton became the clockwork music box. In the 19th century, the automaton was a glorified heat engine, burning a combustible fuel instead of the glycogen of human muscles. The automaton of our day opens doors by means of photocells, or points guns to the place at which a radar beam picks up a hostile airplane, or computes the solution of a differential equation.



THE NERVOUS SYSTEM, in the cybernetic view, is more than a self-contained apparatus for receiving and transmitting signals. It is a circuit in which a feedback principle operates as certain impulses enter muscles and re-enter the nervous system through the sense organs.

Under the influence of the prevailing view in the science of the 19th century, the engineering of the body was naturally considered to be a branch of power engineering. Even today this is the predominant point of view among classically minded, conservative physiologists. But we are now coming to realize that the body is very far from a conservative system, and that the power available to it is much less limited than was formerly believed. We are beginning to see that such important elements as the neurones—the units of the nervous complex of our bodies—do their work under much the same conditions as vacuum tubes, their relatively small power being supplied from outside by the body's circulation, and that the bookkeeping which is most essential to describe their function is not one of energy.



GOVERNOR of a steam engine is an example of feedback, one of the most important fundamental concepts in cybernetics.

In short, the newer study of automata, whether in the metal or in the flesh, is a branch of communications engineering, and its cardinal ideas are those of the message, of the amount of disturbance or "noise" (a term taken from the telephone engineer), of the quantity of information to be transmitted, of coding technique, and so on.

This view obviously has implications which affect many branches of science. Let us consider here the application of cybernetics to the problem of mental disorders. The realization that the brain and computing machines have much in common may suggest new and valid approaches to psychopathology, and even to psychiatry.

These begin with perhaps the simplest question of all: how the brain avoids gross blunders or gross miscarriages of activity due to the malfunction of individual parts. Similar questions referring to the computing machine are of great practical importance, for here a chain of operations, each of which covers only a fraction of a millionth of a second, may last a matter of hours or days. It is quite possible for a chain of computational operations to involve a billion separate steps. Under these circumstances, the chance that at least one operation will go amiss is far from negligible, even though the reliability of modern electronic apparatus has exceeded the most sanguine expectations.

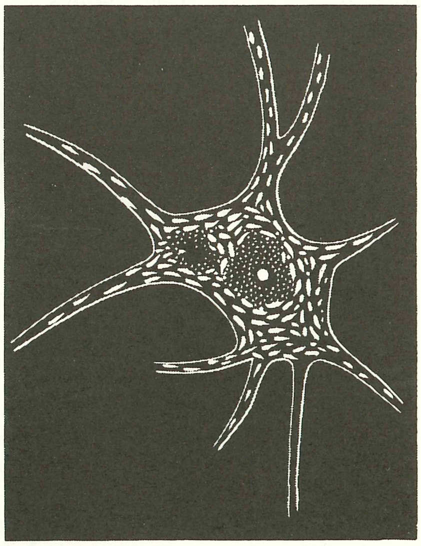
In ordinary computational practice by hand or by desk machines, it is the custom to check every step of the computation and, when an error is found, to localize it by a backward process starting from the first point where the error is noted. To do this with a high-speed machine, the check must proceed at the pace of the original machine, or the whole effective order of speed of the machine will conform to that of the slower process of checking.

A much better method of checking, and in fact the one generally used in practice, is to refer every operation simultaneously to two or three separate mechanisms. When two such mechanisms are used, their answers are automatically collated against each other; and if there is a discrepancy, all data are transferred to permanent storage, the machine stops and a signal is sent to the operator that something is wrong. The operator then compares the results, and is guided by them in his search for the malfunctioning part, perhaps a tube which has burned out and needs replacement. If three separate mechanisms are used for each stage, there will practically always be agreement between two of the three mechanisms, and this agreement will give the required result. In this case the collation mechanism accepts the majority report, and the machine need not stop. There is a signal, however, indicating where and how the minority report differs from the majority report. If this occurs at the first moment of discrepancy, the indication of the position of the error may be very precise.

It is conceivable, and not implausible, that at least two of the elements of this process are also represented in the nervous system. It is hardly to be expected that any important message is entrusted for transmission to a single neurone, or that an important operation is entrusted to a single neuronal mechanism. Like the computing machine, the brain probably works on a variant of the famous principle expounded by Lewis Carroll in *The Hunting of the Snark:* "What I tell you three times is true."

It is also improbable that the various channels available for the transfer of information generally go from one end of their course to the other without connecting with one another. It is much more probable that when a message reaches a certain level of the nervous system, it may leave that point and proceed to the next by one or more alternative routes. There may be parts of the nervous system, especially in the cortex, where this interchangeability is much limited or abolished. Still, the principle holds, and it probably holds most clearly for the relatively unspecialized cortical areas which serve the purpose of association and of what we call the higher mental functions.

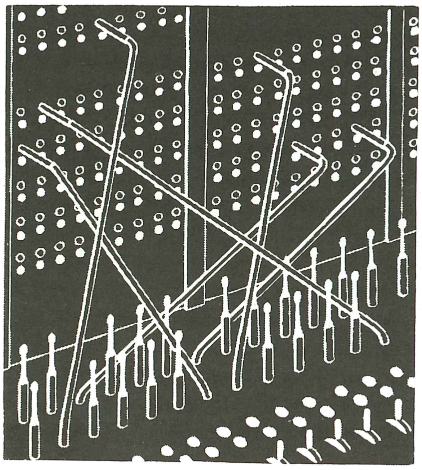
So far we have been considering errors in performance that are normal, and pathological only in an extended sense. Let us now turn to those that are much more clearly pathological. Psychopathology has been rather a disappointment to the instinctive materialism of the doctors, who have taken the view that every disorder must be accompanied by actual lesions of some specific tissue involved. It is true that specific brain lesions, such as injuries, tumors, clots and the like, may be accompanied by psychic symptoms, and that certain mental diseases, such as paresis, are the sequelae of general bodily disease and show a pathological condition of the brain tissue. But there is no way of identifying the brain of a schizophrenic of one of the strict Kraepelin types, nor of a manic-depressive patient, nor of a para-



NERVE CELL performs its functions under much the same conditions as a vacuum tube, obtaining its power from outside.

noiac. These we call functional disorders.

This distinction between functional and organic disorders is illuminated by the consideration of the computing machine. It is not the empty physical structure of the computing machine that corresponds to the brain—to the adult brain, at least—but the combination of this structure with the instructions given it at the beginning of a chain of operations and with all the additional information stored and gained from outside in the course of its operation.



TELEPHONE EXCHANGE, when it is overloaded, has breakdowns rather similar to the kind that occur in human beings.

This information is stored in some physical form—in the form of memory. But part of it is in the form of circulating memories, with a physical basis that vanishes when the machine is shut down or the brain dies, and part is in the form of long-time memories, which are stored in a way at which we can only guess, but probably also in a form with a physical basis that vanishes at death.

There is therefore nothing surprising in considering the functional mental disorders fundamentally as diseases of memory, of the circulating information kept by the brain in active state and of the long-time permeability of synapses. Even the grosser disorders such as paresis may produce a large part of their effects not so much by the destruction of tissue which they involve and the alteration of synaptic thresholds as by the secondary disturbances of traffic, the overload of what remains of the nervous system and the rerouting of messages which must follow such primary injuries.

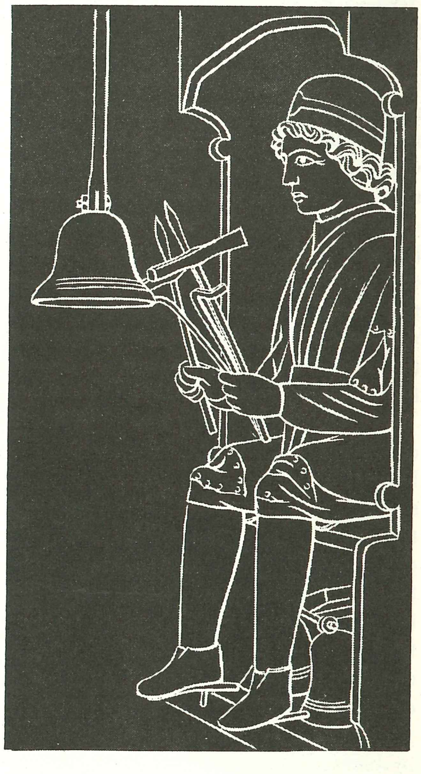
In a system containing a large number of neurones, circular processes can hardly be stable for long periods of time. Either they run their course, dissipate themselves and die out, as in the case of memories belonging to the specious present, or they embrace more and more neurones in their system, until they occupy an inordinate part of the neurone pool. This is what we should expect to be the case in the malignant worry that accompanies anxiety neuroses. In such a case, it is possible that the patient simply does not have the room—*i.e.,* a sufficient number of neurones—to carry out his normal processes of thought. Under such conditions, there may be less going on in the brain to occupy the neurones not yet affected, so that they are all the more readily involved in the expanding process. Furthermore, the permanent memory becomes more and more deeply involved, and the pathological process which began at the level of the circulating memories may repeat itself in a more intractable form at the level of the permanent memories. Thus what started as a relatively trivial and accidental disturbance of stability may build itself up into a process totally destructive to the normal mental life.

Pathological processes of a somewhat similar nature are not unknown in the case of mechanical or electrical computing machines. A tooth of a wheel may slip under such conditions that no tooth with which it engages can pull it back into its normal relations, or a high-speed electrical computing machine may go into a circular process that seems impossible to stop.

How do we deal with these accidents in the case of the machine? We first try to clear the machine of all information, in the hope that when it starts again with different data the difficulty will not recur. If this fails and the difficulty is inaccessible to the clearing mechanism, we shake the machine or, if it is electrical, subject it to an abnormally large electrical impulse in the hope that we may jolt the inaccessible part into a position where the false cycle of its activities will be interrupted. If even this fails, we may disconnect an erring part of the apparatus, for it is possible that what remains may be adequate for our purpose.

In the case of the brain, there is no normal process, except death, that can clear it of all past impressions. Of the normal non-fatal processes, sleep comes closest to clearing the brain. How often we find that the best way to handle a complicated worry or an intellectual muddle is to sleep on it! Sleep, however, does not clear away the deeper memories, nor indeed is a malignant state of worry compatible with adequate sleep.

Thus we are often forced to resort to more violent types of intervention in the memory cycle. The most violent of these involve surgery on the brain, leaving behind permanent damage, mutilation and the abridgement of the powers of the victim, for the mammalian central nervous system seems to possess no power of regeneration. The principal type of surgical intervention that has been practiced is known as prefrontal lobotomy, or leucotomy. It consists in the removal or isolation of a portion of the prefrontal lobe of the cortex. It is currently having a certain vogue, probably not unconnected with the fact that it makes the custodial care of many patients easier. (Let me remark in passing that killing them makes their custodial care still easier.) Prefrontal lobotomy does seem to have a genuine effect on malignant worry, not by bringing the patient nearer to a solution of his problem, but by damaging or destroying the capac-



AUTOMATON of the 15th century was one of a long series of attempts to produce a working simulacrum of a living organism.

ity for maintained worry, known in the terminology of another profession as the conscience. It appears to impair the circulating memory, *i.e.,* the ability to keep in mind a situation not actually presented.

The various forms of shock treatment—electric, insulin, metrazol—are less drastic methods of doing a very similar thing. They do not destroy brain tissue, or at least are not intended to destroy it, but they do have a decidedly damaging effect on the memory. In so far as the shock treatment affects recent disordered memories, which are probably scarcely worth preserving anyhow, it has something to recommend it as against lobotomy, but it is sometimes followed by deleterious effects on the permanent memory and the personality. As it is used at present, it is another violent, imperfectly understood, imperfectly controlled method to interrupt a mental vicious circle.

In long-established cases of mental disorder, the permanent memory is as badly deranged as the circulating memory. We do not seem to possess any purely pharmaceutical or surgical weapon for intervening selectively in the permanent memory. This is where psychoanalysis and the other psychotherapeutic measures come in.

Whether psychoanalysis is taken in the orthodox Freudian sense or in the modified senses of Jung and of Adler, or whether the psychotherapy is not strictly psychoanalytic at all, the treatment is clearly based on the concept that the stored information of the mind lies on many levels of accessibility. The effect and accessibility of this stored information are vitally conditioned by affective experiences that we cannot always uncover by introspection. The technique of the psychoanalyst consists in a series of means to discover and interpret these hidden memories, to make the patient accept them for what they are, and thus to modify, if not their content, at least the affective tone they carry, and make them less harmful.

All this is perfectly consistent with the cybernetic point of view. Our theory perhaps explains, too, why there are circumstances in which a joint use of shock treatment and psychotherapy is indicated, combining a physical or pharmacological therapy for the malignant reverberations in the nervous system and a psychological therapy for the damaging long-time memories which might re-establish the vicious circle broken up by the shock treatments.

We have already mentioned the traffic problem of the nervous system. It has been noted by many writers that each form of organization has an upper limit of size beyond which it will not function. Thus insect organization is limited by the length of tubing over which the spiracle method of bringing air by diffusion directly to the breathing tissues will function; a land animal cannot be so big that the legs or other portions in contact with the ground will be crushed by its weight, and so on. The same sort of thing is observed in engineering structures. Skyscrapers are limited in size by the fact that when they exceed a certain height, the elevator space needed for the upper stories consumes an excessive part of the cross section of the lower floors. Beyond a certain span, the best possible suspension bridge will collapse under its own weight. Similarly, the size of a single telephone exchange is limited.

In a telephone system, the important limiting factor is the fraction of the time during which a subscriber will find it impossible to put a call through. A 90 per cent chance of completing calls is probably good enough to permit business to be carried on with reasonable facility. A success of 75 per cent is annoying but will permit business to be carried on after a fashion; if half the calls are not completed, subscribers will begin to ask to have their telephones taken out. Now, these represent all-over figures. If the calls go through a number of distinct stages of switching, and the probability of failure is independent and equal for each stage, in order to get a high probability of final success the probability of success at each stage must be higher than the final one. Thus to obtain a 75 per cent chance for the completion of the call after five stages, we must have about 95 per cent chance of success at each stage. The more stages there are, the more rapidly the service becomes extremely bad when a critical level of failure for the individual call is exceeded, and extremely good when this critical level of failure is not quite reached. Thus a switching service involving many stages and designed for a certain level of failure shows no obvious signs of failure until the traffic comes up to the edge of the critical point, when it goes completely to pieces and we have a catastrophic traffic jam.

So man, with the best developed nervous system of all the animals, probably involving the longest chains of effectively operated neurones, is likely to perform a complicated type of behavior efficiently very close to the edge of an overload, when he will give way in a serious and catastrophic manner. This overload may take place in several ways: by an excess in the amount of traffic to be carried; by a physical removal of channels for the carrying of traffic; or by the excessive occupation of such channels by undesirable systems of traffic, such as circulating memories that have accumulated to the extent of becoming pathological worries. In all these cases, a point is reached—quite suddenly—when the normal traffic does not have space enough allotted to it, and we have a form of mental breakdown, very possibly amounting to insanity.

This will first affect the faculties or operations involving the longest chains of neurones. There is appreciable evidence, of various kinds, that these are precisely the processes recognized as the highest in our ordinary scale of valuation.

If we compare the human brain with that of a lower mammal, we find that it is much more convoluted. The relative thickness of the gray matter is much the same, but it is spread over a far more involved system of grooves and ridges. The effect of this is to increase the amount of gray matter at the expense of the amount of white matter. Within a ridge, this decrease of the white matter is largely a decrease in length rather than in number of fibers, as the opposing folds are nearer together than the same areas would be on a smooth-surfaced brain of the same size. On the other hand, when it comes to the connectors between different ridges, the distance they have to run is increased by the convolution of the brain.

Thus the human brain would seem to be fairly efficient in the matter of the short-distance connectors, but defective in the matter of long-distance trunk lines. This means that in the case of a traffic jam, the processes involving parts of the brain quite remote from one another should suffer first. That is, processes involving several centers, a number of different motor processes and a considerable number of association areas should be among the least stable in cases of insanity. These are precisely the processes which we should normally class as higher, thereby confirming our theory, as experience does also, that the higher processes deteriorate first in insanity.

The phenomena of handedness and of hemispheric dominance suggest other interesting speculations. Right-handedness, as is well known, is generally associated with left-brainedness, and left-handedness with right-brainedness. The dominant hemisphere has the lion's share of the higher cerebral functions. In the adult, the effect of an extensive injury in the secondary hemisphere is far less serious than the effect of a similar injury in the dominant hemisphere. At a relatively early stage in his career, Louis Pasteur suffered a cerebral hemorrhage on the right side which left him with a moderate degree of one-sided paralysis. When he died, his brain was examined and the damage to its right side was found to be so extensive that it has been said that after his injury "he had only half a brain." Nevertheless, after his injury he did some of his best work. A similar injury to the left side of the brain in a right-handed adult would almost certainly have been fatal; at the least it would have reduced the patient to an animal condition.

In the first six months of life, an extensive injury to the dominant hemisphere may compel the normally secondary hemi-

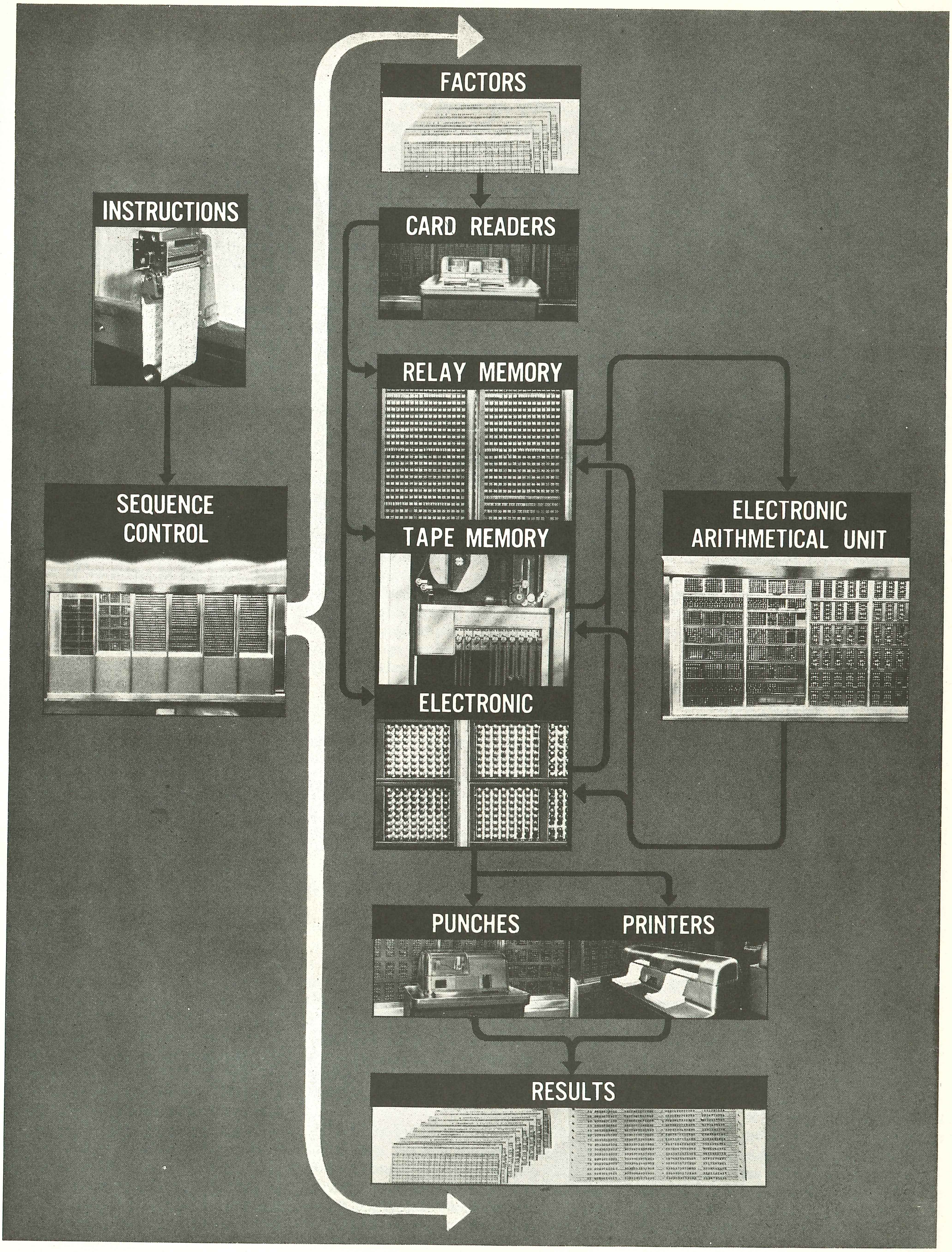


DIAGRAM of the Selective Sequence Electronic Calculator built by the International Business Machines Corporation, provides another cybernetic comparison. Physical structure of the machine is not analogous to the brain. The structure plus instructions and stored memories is analogous. The machine has electronic and relay circuits for temporary memory, punched cards for permanent memory.

sphere to take its place, so that the patient appears far more nearly normal than he would have been had the injury occurred at a later stage. This is quite in accordance with the great flexibility shown by the nervous system in the early weeks of life. It is possible that, short of very serious injuries, handedness is reasonably flexible in the very young child. Long before the child is of school age, however, the natural handedness and cerebral dominance are established for life. Many people have changed the handedness of their children by education, though of course they could not change its physiological basis in hemispheric dominance. These hemispheric changelings often become stutterers and develop other defects of speech, reading and writing.

We now see at least one possible explanation for this phenomenon. With the education of the secondary hand, there has been a partial education of that part of the secondary hemisphere which deals with skilled motions such as writing. Since these motions are carried out in the closest possible association with reading, and with speech and other activities which are inseparably connected with the dominant hemisphere, the neurone chains involved in these processes must cross over from hemisphere to hemisphere, and in any complex activity they must do this again and again. But the direct connectors between the hemispheres in a brain as large as that of man are so few in number that they are of very little help. Consequently the interhemispheric traffic must go by roundabout routes through the brain stem. We know little about these routes, but they are certainly long, scanty and subject to interruption. As a consequence, the processes associated with speech and writing are very likely to be involved in a traffic jam, and stuttering is the most natural thing in the world.

The human brain is probably too large already to use in an efficient manner all the facilities which seem to be present. In a cat, the destruction of the dominant hemisphere seems to produce relatively less damage than in man, while the destruction of the secondary hemisphere probably produces more damage. At any rate, the apportionment of function in the two hemispheres is more nearly equal. In man, the gain achieved by the increase in the size and complexity of the brain is partly nullified by the fact that less of the organ can be used effectively at one time.

It is interesting to reflect that we may be facing one of those limitations of nature in which highly specialized organs reach a level of declining efficiency and ultimately lead to the extinction of the species. The human brain may be as far along on its road to destructive specialization as the great nose horns of the last of the titanotheres.