5 ?ca 1955 ? The Impact of Communication Engineering on Philosophy N. WIENER MC 22

about 5000 words

THE IMPACT OF COMMUNICATION ENGINEERING ON PHILOSOPHY

by

Professor Norbert Wiener,
Massachusetts Institute of Technology
Cambridge, Massachusetts

Many of the main themes of philosophy in all ages have centered around the subject of communication, or the acquisition, use, and transfer of information. By communication I mean, more explicitly, the transmission or employment of one specified alternative among a collection of possible alternatives. This notion of the actual as immersed in a universe or ensemble of possible situations is one which has appeared in other physical sciences, and in particular in the science of statistical mechanics. In the pre-statistical Newtonian physics, the world is considered as a single world in which future states follow present states according to certain fixed physical laws. Even before the full impact of Newton's work was realized, Leibnitz introduced the notion of other possible universes. This notion is fundamental in the writings of Boltzman and Gibbs. These authors ask the question, not how the laws of Newton express themselves in the development of a single system, but how the same laws express themselves in the development of all the systems of a set in which there is a probability or distribution of the individual systems. In this there is no violation of the Newtonian laws, but a new idea not to be found in Newton. They are not concerned with what happens in the individual system so much as what happens in almost all systems of a given statistical distribution, where this notion of "almost all" has been identified by Borel with Lesbegue's notion of "except for a set of measure zero."

In precisely the same way a significant message obtains its significance just because it is only one of a number of messages which might have been sent. To illustrate this, let us point out that the significance of a message can be a very different thing from the number of words it seems to contain. We are all familiar with the Christmas and birthday messages sent by the telegraph companies. They read something like this: "It is a great privilege, dear friend, to convey to you our most sincere wishes for a Merry Christmas and Happy New Year. We are all well and thinking of you." Actually what the telephone company sends is a code number, say 73, by which the message can be identified at the other end from a book of standard messages. The effective degree of choice shown in such a message is not measured by the number of words which it contains but by the number of possible messages belonging to the same code system. If the code system contained only one such message, the only significance of the message would be that it is sent or not sent. Communication, to be effective, is then the transmission of an alternative among other alternatives and can only be studied by considering the transmission of all possible alternatives and the transmission of a given alternative of a choice among these.

It is perhaps as well to point out that we are conceiving the choices to be made in communication as independent choices, and between alternatives which we have no reason to favor, one over another. In evaluating choices actually made, it may be pointed out that we can reduce any finite set of choices to a finite equivalent set of independent choices. The weighting of each independent choice is to be made in the following way: if it selects from a set of alternatives with total _a priori_ probability A a sub-set with total a priori probability B, it amounts to $\log_2 A/B$ effective unitary choices. Thus if we have two equally probable alternatives, and select one, the choice is equivalent to

$$\log_2 \frac{1}{\frac{1}{2}} = 1$$

effective choices. I bring this connection with an *a priori* probability out because of its importance later, when we shall mention the notion of entropy.

To return to our definition of communication: Epistemology is the science of the validity of our communication with the external world and experience from the external world. Our experiences are regarded as significant just to the extent to which they convey to us one alternative when they might have conveyed others. Logic is the science of the internal or formal validity of communication, that is, the science of the identification of content in apparently different communications, so that we can really penetrate to the effective degree of multiplicity of our world of communication in a particular situation. Psychology is often supposed to be the study of the self, but experimental psychology at least has had much more to do with the study of our sense organs or of the means of communication from the external world to ourselves. Other branches of psychology are the study of association, or of the means by which communications previously unrelated become incorporated into the system; the study of the will, or of our means of communication proceeding outwards into the external world and of our ability to impose this communication on the external world; the study of memory, or of the storage of communications; the study of affective coloring or of that element of communication both from the outer world and from within which leaves us to accept and to seek for the renewal of certain parts of it and to reject others. Parallel with this intrinsic study of communication are Neurology or the study of the physiological channel of communication through the nervous system; Endocrinology, or the study of the humoral transmission of messages; and Myology, or the study of the muscular effectors by which communication is carried from within out. By effectors in a more general sense, I mean muscles and those glands and electric organs which are the channels by which the communication of an animal leave its body and impose themselves on the outer world.

The external transmission of communication leads to several different sciences. A group of these, the philological group, concerns language as a means of communication. Particularly important among the philological sciences is Semantics, or the study of language as an effective vehicle of meaning.

Once a communication has left us in the form of language or gesture or some other similar means, its further transmission becomes the object of a branch of engineering, namely Communication Engineering. Closely related to this and, indeed, indistinguishable from communication engineering in its main principles is Control engineering, which deals with the employment of information whether of human origin or not, in remote mechanical or electrical effectors of an artificial nature. In control engineering it is important, indeed, not only to cause messages proceeding to the effectors, but messages arising in the effectors, and enabling the machine to act more effectively by taking account not only of its expected performance but of its actual performance. This is, indeed, a parallel to the branch of physiology dealing with the kinaesthetic sense. When we wish to perform such an act as that of picking up a lead-pencil, we do not consciously wish to contract certain muscles, and indeed we do not usually know what muscles we contract. What we wish to do is to diminish the amount by which we have not yet picked the body up, and this is conveyed to us by our eyes and our kinaesthetic organs or proprioceptors. This is specifically a feed-back mechanism. Its breakdown leads to ataxia, and is surprisingly similar in its manifestations to the wider and wider oscillations which we observe in the breakdown of a servo-mechanism. It is a notable symptom of cerebellar disease or injury.

It is a well-known fact that in more complicated servo-mechanisms, more than one channel of feed-back is necessary for stability. Similarly, in the body, voluntary feed-back must themselves be stabilized by what we may call postural feed-backs. It is these latter which break down in Parkinsonianism, or shaking palsy. The centers disordered in this disease are known, and lie outside the cerebellum. In postural tremors like Parkinsonianism, the tremor occurs in a resting state, and is alleviated by voluntary motion, while in the break-down of voluntary activity, the tremor only exists when the subject tries to perform some action.

To return to the basic notions of communication engineering: the fundamental idea is that of a decision, i.e. of a yes or no. This represents the smallest amount of definite information which can be transmitted. To describe a more complicated message it should be broken up into yeses and no's. It is perfectly clear, or course, that these yeses and no's do not receive communication value until they are interpreted by the recipient as yeses or no's answering a concrete question, but in their transmission from individual to individual they assume the form of yeses or no's, pure and simple, which form a code which is interpreted according to rules, natural or arbitrary, by the recipient. Suppose now that we transmit a quantity. How much information does a statement of this quantity contain? If the quantity is specified with complete precision, the answer is an infinite amount of information. On the other hand, we are manifestly unable to transmit and use an infinite amount of information. Just how does this limitation make itself apparent?

Let me put on an electrical circuit a voltage of 32 volts or less, suppose from generators which I do not know, there is additively superimposed upon this a random voltage of plus or minus one volt. How many decisions can be effectively transmitted? The simplest answer is to represent the voltage less than 32 volts on the binary scale wither the total voltage is between 16 and 32 or it is less than 16. This is one significant decision. We represent it by a digit 1 or 0. If the first digit is one, the voltage may lie between 24 and 32 or between 16 and 24. If the first digit is 0, the voltage may lie between 8 and 16 or it may lie between 0 and 8. Whatever the first digit may be, we have a binary choice of the second which we shall take to be 1 where the larger possibility is taken and 0 when the smaller possibility is taken. In the same way, we may split our range of decision into ranges of 4 and of 2, which means that we had two more choices of digit to our possibilities. On the other hand, the next choice falls afoul of thefact that we have an uncertainty of plus or minus one in the measured value of voltage. We thus have four affective decisions of yes or no. It will be seen that if we represent the number we are trying to transmit by A and the uncertainty of this number by B, then $\log_2$ of A/B represents the number of determinable digits which we are trying to convey, i.e. the number of decisions and hence the amount of information transmitted is $\log_2$ of A/B.

This is a very significant remark. As we have already stated, A/B is a ratio of probabilities, i.e. that the message lies within the whole range allowable and the probability that it lies within a certain range which cannot be resolved because of the error level or, as we call it in communication engineering, the noise level. Thus, the amount of information is measured by the logarithm of a probability or a probability ratio. To those of us who are familiar with thermodynamics, the notion of the logarithm of a probability ratio is familiar. The logarithm of the probability of a certain set of states of a dynamical system as compared with a set of all possible states is called minus the entropy, in other words amount of information is a quantity which has the character of a negative entropy. This is in accordance with the fact that on the one hand, the entropy of a physical system can increase but never decrease spontaneously, but on the other hand, in the transmission of information, we may lose it but never gain it.

This relation between information and entropy becomes less surprising furthermore, when we consider the negative entropy as a measure of the organization of a system and that organization and information are ideas that are thoroughly germane to one another. However, I am sadly afraid that many communication engineers and even more physiologists and psychologists have not grasped this entropic character of information. Many times have I heard psychologists discourse in terms of energies and I have even heard that word neurophysiology as if it represented a fundamental idea concerning the functioning of the nervous system. Many times, too, have I heard the communication engineers of the older school talk in terms of resistances, capacities and inductances as if passive networks consisting only of these and not containing any external sources of energy were in some ways more natural material for communication engineering than systems containing vacuum tubes and other intercurrent sources of energy. All this is based on a misapprehension. It can be shown that a theory of communication engineering receives no essential complication if we are allowed to feed in energy by amplifying and other similar devices, nor is neurology made more complicated if we realize that the energy supply to the nervous system occurs at every point and is primarily metabolic, so that in an energy balance the nerves are of trifling importance compared with the cerebral blood vessels. Until this is clearly realized, no important progress is possible.

The parallelism between statistical mechanics and communication theory is further emphasized by the fact that in the practice of each of them, there is a certain important principle which is always utilized in a negative, prohibitive way. In statistical mechanics, this is the second law of thermodynamics, which says that a system may lose order, but never gain order as a whole. In communication engineering, it is the fact that the past may influence the future, but not the future the past. It is not going too far afield to interpret this statement as asserting that we may lose information, but never gain it, beyond the level at which it exists initially.

Besides the notion of information and amount of information, an extremely important concept is that of the rate of transmission of information. The connection of the two ideas is obvious. The rate of transmission of information is measured not in decisions but in decisions per second. It is obvious, furthermore, that the notion of rate of information presupposes something of the nature of what statisticians call a time series, i.e., a sequence, continuous or discrete, of quantities or other vehicles of information arranged one after another in time. For example, day-by-day prices of General Electric stock in the stock market form what we could call a simple, discrete, numerical time series. On the other hand, the variable voltage at a given point in the line during telephone conversation also furnishes a time series, simple and numerical, but not discrete. The theories of time series in statistics and of varying quantities in communication engineering have a history which almost completely separates their techniques, but I wish to emphasize that this separation is accidental and marked out by the professional difference between the statistician and the electrical engineer and does not correspond to any real difference whatever in the nature of the sciences.

One of the most important problems in the statistical mechanics of messages is that of the prediction of the future of a message when the past is given. This can only be done if we know the statistics of the ensemble of messages. It is fundamentally the question, not of the exact future of the message, but of the relative distribution of the ensemble of possible futures. It may be that we wish to know all of the statistical parameters of this future, or only some of these. In exceptional cases, such as those which occur in astronomy, it might be a sensible question to ask for the precise future of a system,--except for two things, that we do not know an infinite range of data covering the past of the system, nor do we know those data available to us to more than a finite degree of precision.

The problem of working with data not known to us to more than a finite degree of precision--that is, of messages corrupted with noise,-- and of separating out the valid information, is essentially the problem of the design of wave-filters in telephone engineering. The author has given valid mathematical methods for the treatment of this problem as well as the prediction problem, both on the assumption that all methods of operation, linear and non-linear, are available, and on the assumption that only linear methods are available. The application of these methods to short runs of data involves the same sort of problems of working with insufficient data that are universal in statistics, and much remains to be done.

Let it be noted that the idea of the statistical design of engineering apparatus such as wave filters is a new one. The idea is that such pieces of apparatus, at any rate in the case of communication engineering, have to deal with a certain ensemble or distribution of messages, and that, given this ensemble, they have to perform a certain task in an optimum manner. Very often this task cannot be performed in a strictly perfect manner. Thus, the problem of design is reduced to what mathematicians call a minimization problem. It is interesting to see that there is thus the possibility of a mathematical theory of invention.

A mode of communication which is perhaps as remote as possible from the method of linear communication is that method which depends on elements which can only exist in a finite number of distinguishable states and in particular in two states. The wheels of the standard desk computing machine have only a finite number, say ten, stable, distinguishable positions. The determination of the positions of what we may call the earlier wheels of a system carries with it the determination of the later wheels. The computing machine depending on wheels is held down, however, to certain rather narrow restrictions. Wheels have friction. They have a large amount of inertia and due to these causes they cannot be operated at an excessive speed. In order to satisfy the needs of modern computational work, it is desirable to have elements which have a very low inertia and very low frictional forces, and which can consequently be operated at vastly greater speeds. The electron beam is a member which well lends itself to such operation. The electron beam, however, while it can be operated very simply so as to have two alternative states of equilibrium, requires a great deal more apparatus if it is to be operated in such a way as to have a large number of intelligible states of equilibrium. This means that the easiest way of representing, let us say ten choices, by means of an electron beam, is to build an apparatus with four consecutive stages and conceivably capable of sixteen different positions of equilibrium, from which we discard six. In other words, if a computing machine is to be worked at highest efficiency, it is appropriate under very general conditions to have it make nothing but binary choices.

Consecutive binary choices are adequate for a description of the entire number system of Arithmetic and the entire system of Logic. If we introduce binary choices which are conditioned by the results of previous binary choices, we can by means of them represent every arithmetical and every logical relation. This is, in fact, what is done in modern ultra-rapid computing machines like the EDVAC. It is also, strikingly enough, what is done by our nervous system. The elements of our nervous system are neurons, and these neurons can exist physiologically only in two states, that of activity and that of passivity. On each neuron there terminate fibers from other neurons joined to it by organs called synapses. If the proper combination of synaptic connections fires, which may mean quite as well that some of them do not fire as that others do fire, the outgoing neuron will fire. For example, the synaptic connections of a given neuron may be N in number, and if M, or more of them, fire within a certain brief interval of time, the outgoing neuron may fire, and not otherwise. Organizationally, such a system is highly analogous to the ultra-rapid computing machine, although the mechanisms which determine the firing or non-firing may be totally different.

Both the nervous system and the computing machine must contain organs of memory or retentiveness, and may contain organs of association. In both cases, the demands on memory are likely to be such as to require several different sorts of memory; namely, at least to a short-time memory which can be impressed at high speed, read at high speed, and erased at high speed, and a long-time memory in which speed is sacrificed to permanence. The short-time memory of man is that which constitutes what we know as the specious present, and may very well consist in impulses travelling in circular paths along chains of neurons. This device of the circulating path is also used in the computing machine, where an impulse is impressed on a rotating steel wire or another mechanism of the sort, is read off that wire before it becomes blurred to a critical extent, and is re-impressed on that or another wire by an instrument(which we call a telegraph type repeater) in which older impulses, instead of reproducing themselves directly as newer impulses, are used to trigger new impulses of the same sharpness as they originally possessed. The longer storage in the computing machine is likely to depend on semi-permanent changes of elements of apparatus, such as the charging or discharging of a condenser, or the flopping over of a mechanical relay or its equivalent electric circuit. How the longer time memory is carried in the brain we do not know. It may very well be a change in the actual number or in the permeability of synaptic connections. It is highly unlikely that the human brain or that of the higher vertebrates ever makes new system connections during its adult lifetime, but it is quite possible that memory may be resident in the channeling of nervous impulses by the destruction of already existing synapses. We may, indeed, be leading an existence suggestive of Balzac's story Peau de Chagrin where all of our success and accomplishment is directly due to the destruction of the very margin of our existence. At any rate, it is an established anatomical fact that the adult has less neurons and probably less synapses than the child.

There are two important problems concerning the nervous system and representing lines of construction which can be followed in the computing machine. These are: (1) the problem of the recognition of what is called "Gestalt" ~~in German~~, or, in other words, of a not necessarily visual form, and (2) the problem of association. As to (1) it is perfectly clear that we can recognize a square as a square whatever its size or orientation is. It would not be impossible to make a machine which would recognize a square by comparing it with the entire set, say of projective transformations of a given square, in such a way that if the coincidence of outlines were more perfect than a certain amount, the recognition would be performed. This could be done conceivably by a scanning process over the entire set of admissible transformations analogous to the scanning process by which in television we view one after the other of the points of a plane area. It is most certainly not done in precisely this way in the mind. Consecutive processes in time are very much better done by appropriate machinery than by the nervous system, whereas the nervous system contains a complexity of simultaneous parts far greater than we can hope to achieve in a man-made machine of even comparable size. We must make up by our advantage in time our disadvantage in space, if we are to make satisfactory automata. For example, television is much the same phenomenon as we should have if, instead of a retina with many rods and cones, we had one single element moving over the retina with such a velocity that it would seem to us to everywhere at once.

It is relatively easy to give an account of a mechanism by which we could recognize any specific item of recognizable form. It is much more difficult to give an apparatus for recognizable form in general. It is perhaps worth remarking, however, that recognizable form in general is probably incapable of a philosophical definition and only signifies such forms as are actually recognized by the human mechanism, however disjoint this assemblage may be. A great deal can be done, it is true, in individual cases, by showing the marks we do actually recognize. The Bell Telephone Company, by its work on the Vocoder, a mechanical device for reproducing speech, has shown what a small part of the entire acoustical pattern is necessary for such recognition. A few code quantities such as high pitch, low pitch, abrupt start, abrupt termination, smooth start, smooth termination, and the like, will give us the greater part of the intelligibility given by the entire gamut of the voice. This question then of the relation between coding and Gestalt deserves much more study.

As to association, we do not know how it is done, but there is certainly no contradiction in supposing a machine which gives two settings, which have often and consistently occurred together, a common code symbol, so that the exciting of the one will excite the other. Apart from this, we are not much further on than Locke. Locke, however, in his theory of association, supposes the mind to be a blank slate without any active powers of modifying the material in it. He combines a mechanical theory of association with an explicit denial of any machinery by which association can act. For him, the mind is devoid of inborn impressions and there is no place in his philosophy for an inborn set of the mind. Perhaps we can do him the most justice by saying for him the mind is merely a playing-field, on which the ideas act out their intrinsic attractions and repulsions. Nevertheless, they are supposed to be simple, and devoid of an internal structure to explain their behavior.

We shall not be far wrong if we associate the Lockean psychology with the faint forces and action at a distance of the classical Newtonian physics, while the more modern psychology is the direct heir to the philosophy of Leibnitz, and concerns itself with notions closer to the modern physical field theories. We are more accustomed than the authors of the enlightenment to expect complexity of structure to be a concomitant of complexity of behavior, and less likely to be content with a purely classificatory theory, like the Linnaean botany and zoology, which without analysing a complicated behavior, or a complex structure obviously bound up with a complex function, is content to describe it with a complex neo-Latin name, and to refer its object to its appropriate place in a hortus siccus or a collection of stuffed skins.

It is definitely worth while to correlate the point of view of this paper with those assumed in two other sciences: Mathematics (together with logic), and sociology. First as to Mathematics: the notion of the Machina ratiocinatrix is but the mechanization and logical extension of the calculus ratiocinator of Leibnitz, and has been made an effective tool in metamathematics by Turing, who is in fact engaged in the construction of such machines. The mechanization of the mind here considered is regarded as justifiable by many mathematical logicians; though some, like Church, regard the mind as having a flexibility of choice exceeding that of any single machine. Some of this apparent flexibility may well be specious, being due to the facts: (1), that at no specific time is the information available to any brain-machine given in its completeness; (2), that we do not, in fact, possess an explicit knowledge of the total set of running rules of any brain machine; (3), that we have, in fact, a multiplicity of brain machines, not all alike.

It is perhaps interesting to observe that the Russellian paradoxes and the difficulties of the problem of decision receive a definite interpretation in the performance of the machine. A logical machine comes to a definite solution of a problem when it settles down to a final state, stops, and turns off the power. If faced with a problem such as that of whether the barber who shaves those, and only those, who do not shave themselves, does or does not shave~~s~~ himself, ~~or not,~~ it hunts over an unending chain of "yes," "no", "yes," "no" answers. In the case of ~~the~~ Fermat's last theorem, it may run through an unending set of special processes not subsumable in finite terms. There is a certain analogy between the role played by time in these processes and by the type number in Russell's hierarchy.

One last group of remarks concerns the role of communication in sociology. We are ourselves beings with an internal communication system, with the aid of our voice, our writing, and our gestures, we are also members of a larger communication system embracing many individuals. Is there a way of distinguishing these two fields of communication and separating the order or organization of the individual from that of the race or community?

There is, in my opinion, such a way. The complexity of the experience of the individual is determined by the number of yeses or no's or decisions effective internally in his conduct. The organizational complexity of a society is measured by the number of acts of recognition of its constituent individuals which can be imparted to society as a whole. In an animal like the cat, which is thoroughly non-gregarious, this social organization is extraordinarily small when compared with the very high organization of the individual. In an animal like the ant, all that has so far been pointed out with any security as to the degree of social recognition is that by some sort of chemical sense the individual can recognize the nest or colony to which another individual belongs without recognizing him in any way as this or that individual. This is a poverty of communication comparable with a poverty of that part of our internal communication which takes place by the chemical transmission of hormones. Much as the organization of the individual ant is below our own, it is still probably considerably higher than that of the ant communication. On the other hand, in the human community, the direct personal means of recognition in communication are supplemented by the extremely complicated means known as spoken and written language. There is very little doubt that the storage capacity of a human community for significant communication vastly exceeds the internal memory capacity of any individual member. Difficult as it may be, these matters are subject to an ultimate numerical investigation, but the time is probably not ripe to carry such an investigation out.