RICHARD T. ELY LECTURE

ECONOMICS OF INQUIRING, COMMUNICATING, DECIDING*

By JACOB MARSCHAK
University of California, Berkeley

We hear much of today's "informational revolution." We are also told of the rapid growth of the "knowledge industry." Informational revolution is exemplified by TV pictures of the moon surface and also by robotized stock market transactions and, hopefully, by computerized professors. Fritz Machlup defined the knowledge industry to include education and research as well as publishing and broadcasting. He estimated its share in the gross national product of 1958 at 23 percent to 29 percent, and its growth rate at about 10 percent, or twice that of the GNP. Projecting to the present, the share of the knowledge industry would then appear to straddle the 40 percent mark!

There is a suspicious overlap between these activities and those which Adam Smith and Karl Marx called "unproductive" and which include the work of kings and professors, none of whom add to the vendible and visible stocks of the nation. To be sure, recent analysis—for example, by T. W. Schultz and Carl v. Weizsaecker—found it both convenient and feasible to define human capital and thus to consider education as investment. But the notable fact remains that professors and kings or the modern equivalent of kings—managers, both public and private—are strongly involved in those trends: informational revolution and growing knowledge industry.

Professors and managers, but also computers and TV sets, are involved in still another trend relevant to my talk. A growing proportion of both manhours and machine-hours is not employed for using large amounts of energy (muscular or otherwise) to transform or transport matter. Instead, so-called "brains" (human or otherwise) are employed to manipulate symbols. A sequence or network of such symbol manipulators uses up a minute amount of energy to eventually release, trigger-like, large amounts of energy through the more brutal medium of generators, muscles, and machine tools. In a modern assembly or disassembly plant (sawmill, meat packing), a growing majority of people,

* Supported by a grant from the Ford Foundation to the Graduate School of Business Administration, University of California at Berkeley, and administered through the Center for Research in Management Science. Facilities of the Western Management Science Institute, Univ. of California at Los Angeles (supported by grants or contracts of the National Science and the Ford Foundations and of the Office of Naval Research) were used in preparing the present revised version.

wearing white collars, or blue denims as well, do the brain work of inspecting, deciding, reporting—shunting, pushing buttons—and not the muscular work of shaping or carrying material masses; and a growing proportion of machines, called control mechanisms, are also busy with inspecting, reporting, deciding, and not with transforming or transporting matter and energy.

My topic is the economics of what I shall call the services of inquiring, communicating, deciding. Data are gathered. They are communicated to the decision-maker. He, on the basis of the message received, decides upon the action to be taken. A higher-order decision must have been made previously. Someone representing the interests of the economic unit considered—its head, leader, organizer—must have chosen a particular combination of these three services from all those available in their respective markets. The maker of this higher-order decision (the "meta-decider," sometimes called "the organizer") may happen to be the same person who will decide upon acting. Or more generally, the organizer will hire the services of the decision-maker—who, in appropriate cases, may be just a robot.

I might also call my topic the economics of the instruments, or devices, human or otherwise, for inquiring, communicating, and deciding. For it is not relevant, for my purposes, to distinguish between purchased instruments and hired services, provided the length of the hire contract is specified. In any case, I shall be concerned with symbol manipulators, human or otherwise, rather than with processors or transporters of matter or energy.

Here is what I plan to do. I shall present, in turn, from the user's point of view, the successive links in the sequence of symbol-manipulating services: inquiry, or data gathering; communication of messages; and deciding upon actions on the basis of messages received. It will turn out, in fact, that the link called "communication" must be broken into two distinct services: on the one hand, the service of "encoding and decoding" which, at the present state of arts and in the most numerous and socially most important cases, is best supplied by men; and on the other hand, the service of "transmission" which is best supplied by inanimate equipment. As to the supply conditions of services of inquiry, or data production, and of decision making, I shall be able to submit nothing but crude illustrations, I am afraid. As to the demand side, economists will not be surprised that to make an economical—that is, optimal, efficient—choice the user must choose those links, or components, simultaneously (just as a manufacturer cannot choose between rail and road as means of bringing him fuel without making up his mind, at the same time, whether the fuel should be coal or oil). Hence, the jointness of demand for services of inquiry, communication, and decision.

To be sure, current engineering science finds it convenient to isolate a pure theory of communication—a theory of efficient coding and transmission alone, essentially created by Claude Shannon and streamlined by Jack Wolfowitz. At the other extreme, statistical decision theory culminating in the work of David Blackwell leaves out the communication component and only analyzes, from the point of view of a perfect decision-maker, the optimal choice of inquiry, or data producing, services, also called "experiments." I shall later state the implicit tacit assumptions made in each case. If they are not satisfied, the user guided by those subtheories will have suboptimized. This is not to say that we ought not to break up a complex problem into subproblems, assuming them independent as a provisional first approximation. Given our human limitations, this may even be the optimal research strategy. It just happens that the economist is aware of interdependencies: he calls them complementarity and substitutability of goods. He is also traditionally permitted—as is the philosopher—to attack complexities with ridiculously simple examples in order to get directly to the general and fundamental.

Let me, then, go ahead with a simple example. I must decide this Thursday night whether to fly West next Saturday morning. Visibility and winds along the airplane's route the day after tomorrow will determine whether, if I do fly, I shall arrive in time for an important appointment or shall be killed in a crash. If I don't fly, I miss the appointment. But I cannot know what the weather will be. Instead, I may look tonight at the hotel barometer; or I may rely on the radio reports of other, more numerous and accurate barometer readings; or I may rely on the *Farmer's Almanac*. If the cost of these various services were equal, I would choose the one which gives data most closely reflecting (in some sense) the actual event I am interested in: the weather on Saturday. But perfection is costly, and I shall choose a service that is known not to mislead too grossly or too frequently, yet will be relatively cheap.

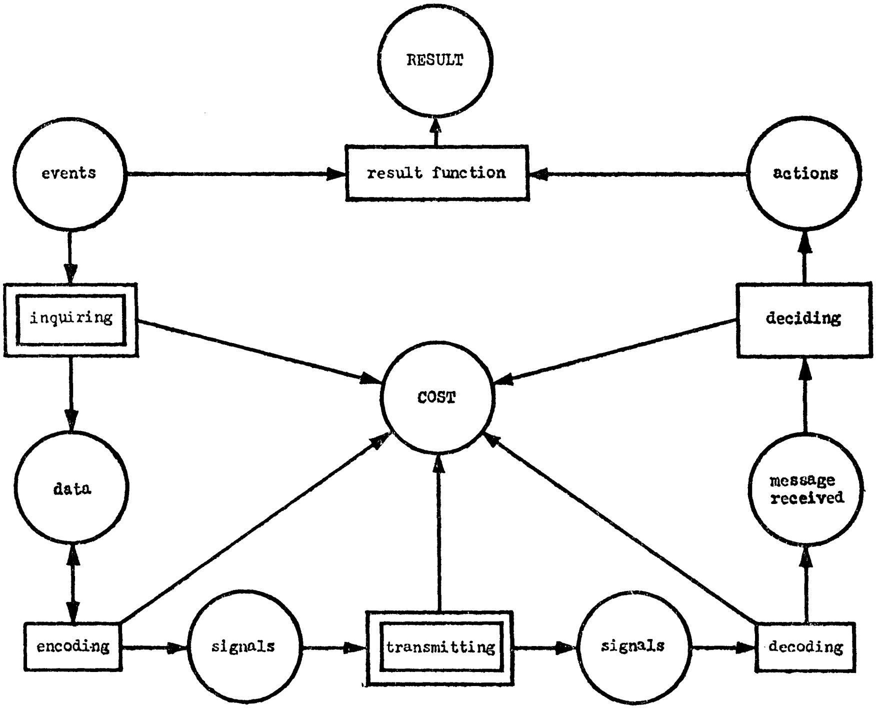
Take another example. A store's profit will be influenced by its inventory policy, given the actual future demand for its merchandise. Lacking the knowledge of this demand, the firm will have to choose between various services of market forecasters differing in precision and accuracy but also in the fees charged.

So much about services that inquire; i.e., produce data. These data are not identical with, yet do reflect in some sense the events that are relevant to, the result of a decision. Now, the decision-maker may or may not be able to obtain such data directly. Another service called "communication" will bring to him, not those data, but a message, possibly delayed or distorted, about those data. He must decide on the basis of such a message, which is now twice removed, in a sense, from the actual, result-relevant event: weather on Saturday, demand next month, and so on.

The inventory example illustrates also the nature of decision services.

Inventory policy is a rule stating whether and how much to reorder when the stock at hand is at a given level and you have some—usually imperfect—knowledge related to the prospective demand of your customers. One policy is similar to the one you use when you decide whether to refill your car's oil tank up to a certain upper level (this you will do whenever oil is below a certain lower level) or to leave it alone. Except that in the inventory case the two critical levels themselves are not fixed but should depend on what the store has learned—however imperfectly—about future demand; that is, on the message it has received about the data produced by a market forecast. Such a decision rule or strategy—a rule of how to respond to the message—may require the sophisticated services of a specialist or a computer. Contrast with this a simple routine rule: to refill the inventory every Monday to a constant level. This can be handled by an unskilled clerk or a cheap robot. The more sophisticated, flexible, nonroutine rule would be preferable if it were not for its higher cost.

To state more precisely the problem facing the user of the data producing, communication, and decision services, it is convenient to represent each service as a transformer of inputs into outputs. (Transformer, transformation, and function mean essentially the same thing.) A data producing service such as a barometer is a transformer whose input is the result-relevant event (weather next Saturday) and whose output is an observed value, a datum (the barometer reading tonight). We say that the data service is both precise and reliable if to each result-relevant event corresponds one observation or datum, and conversely. But this perfection is almost never attained. Generally, each event may be reflected in various alternative observed values, with some alternatives more likely than others. We have here the case of an "unreliable" (probabilistic, stochastic, noisy) transformer. For example, suppose that if Saturday's weather is going to be good, the chance that the barometer shows high pressure tonight is 80 percent. We say that the likelihood of the observation "high pressure," given the event "good weather," is 80 percent. Suppose the likelihood of low pressure if the weather is going to be dangerous is also 80 percent. Suppose that on a second barometer both these likelihoods are lower: 60 percent and 60 percent, say. If you have access to both barometers at the same cost or effort, you will prefer to be guided by the first one. For, in an obvious sense it is more reliable (more informative, in Blackwell's terminology). Indeed, in the case of perfect reliability the two likelihoods would be 100 percent and 100 percent; and clearly our first barometer (with 80 percent, 80 percent) comes closer to this perfection than the second (with 60 percent, 60 percent). In fact, the second comes closer than the first to the other extreme: likelihoods 50 percent, 50 percent, in which case the barometer would be useless.



○ Circles = variables (generally random).
□ Boxes = transformers (noisy if double-framed). User maximizes average result minus cost (assuming additive utilities).

NOTE: In statistical theory, the lower row encoding → · · · → decoding is omitted so that messages received = data. In communication theory, data = events; actions = messages received; and result = "bad" or "good" according as message received is or is not identical with datum.

Or consider a consumer survey conducted by a government agency. The success of the government decision to undertake one or another administrative action depends on the attitudes of all consumers. But only a certain number are sampled. The larger the sample the more reliable is the estimate of people's attitude. But also the more expensive. It is not different with research into the laws of physics or biology. What I called "result-relevant events" the statisticians call "hypotheses." The data producing service (e.g., a sampling) they call "experiment," and the data are called "observations."

Unfortunately, it is not always possible to compare two data producing services on the basis of the likelihoods only. How does our first barometer, with the 80 percent likelihood of high pressure given good weather, and 80 percent likelihood of low pressure given dangerous weather compare with the following one: if the weather is going to be good, the barometer will show high pressure for sure, i.e., with likelihood

100 percent; but given dangerous weather, it will show high or low pressure, not with 80:20 but with 50:50 chances. Thus, whenever the new barometer shows low pressure, it gives you absolute certainty that the weather cannot be good. But when it shows low pressure you are left guessing, and you might be better off with the original barometer. Which barometer should guide you?

Here I must remind you that, just as economists are Keynesians or non-Keynesians, the statisticians are Bayesians or non-Bayesians. The Bayesians, having given serious thought to our problem, tell me to consult two further items: (1) the approximate, so-called "prior," probabilities that I would assign to dangerous versus good weather for next Saturday—in the absence of any barometer, e.g., on the basis of my previous experience with December weather; (2) the utilities that I assign to various results of my actions and whose actuarial value I would like to be as high as possible. In our case, the best result is: surviving and making the appointment. The second best is: surviving but missing the appointment. The worst is death. What matters is the ratio of the disadvantage of missing the appointment (but staying alive) to the disadvantage of death. Now, unless a barometer is useless, I would fly if it shows high pressure and not fly otherwise (or vice versa). The probabilities of these pressure readings depend on my prior probability of the two states of weather and on the likelihoods characterizing each of the two barometers. Therefore, those pressures will be read on the two barometers with different probabilities. Hence the average utility of the results of actions dictated by each barometer and weighted by the probabilities of its readings will differ as between the two barometers. I prefer the one whose dictation will yield, on the average, the higher utility.

Suppose, then, the prior probability of bad (i.e., deadly dangerous) weather is, in my judgment, about 10 percent. Should I use the "old" barometer (whose likelihoods are 80 percent and 80 percent) or the "new" barometer (whose likelihoods are 100 percent and 50 percent), still assuming that the costs are equal. It turns out that I should stick to the old barometer as long as I judge the disadvantage of missing the appointment (while staying alive) to be less than one-seventh of the disadvantage of death. This is surely my case!

I believe this business of assigning prior probabilities to events and of utilities to results is a headache familiar to cost-benefit analysts, certainly present in this very room! It surely requires some soul-searching to appraise and reappraise the subjective probabilities ("beliefs") and utilities ("goals," "values," "tastes") of your government agency. You will presumably try out various plausible assumptions and see whether your boss likes the decisions derived from them; and whether under repeated trials of this kind his choices reveal a consistency of his beliefs and of his tastes. Or perhaps such trials will gradually train him toward consistency—toward learning what he wants and believes.

So far, we have seen how the consistent user should choose between available inquiry services when their costs are equal. This has required, for some though not all pairs of such services, to take account of the user's utilities and prior probabilities in order to compute the average utility attainable on the basis of an inquiry. If costs are not equal, the knowledge of prior probabilities and utilities becomes necessary in all cases. For simplicity, a tacit assumption is made. One assumes, in terms of our example, that it is possible to represent the utility of, say, "having made the appointment and having my wealth diminished by the dollar cost of a particular inquiry" as a sum of two utilities. In other words, utility is assumed to be additive and to be commensurable with dollars. Under this assumption, one may call the average utility of results of actions guided by an inquiry simply its dollar value to the user, his demand price. He compares it with the dollar cost, which is the supply price. The additivity assumption is implicitly made by statisticians. They assume in effect that the disutility of a result of action based on a sampling survey is measured by the size of the error of an estimate, and then vaguely compare its average with the dollar cost of the sampling survey. Engineers indulge in similar practices, as we shall see. Not so the economists. Sophisticated in matters of substitution, income elasticity, and risk aversion—all of which question the assumption of additive utility—they raise a warning finger. They do so at least when they talk theory and are not themselves engaged in the practical pursuits variously called "management science," "operations research," "system analysis," "cost-benefit analysis."

My barometer case—with just two alternative events, two possible observations, and two actions—is probably the simplest nontrivial problem in statistical decision theory. As I said before, this theory neglects the communication link. Result-relevant events, each having some prior probability, are transformed into data by a transformer called an "inquiry," or an "experiment" (like a barometer, or a sampling survey). Data flow directly into the transformer called "decision-maker," who applies a decision rule (e.g., "fly if barometric pressure high") and puts out actions. Finally, events and decisions are joint inputs of a transformer which may be called "result function": its output is a result. Assuming additive utility, a dollar amount can be attached to each result. And the probability of each result is determined by the prior probability of each event, by the array (characteristic of the data producing service) of the likelihoods of the data, given each event, and by the decision rule (characteristic of the given decision service) transforming data into actions. The probabilities of the results thus derived serve as weights to obtain the average of their utilities. This average may be called the "gross value," to a given user, of the given pair of data producing and decision-making services. It is the maximum demand price offered by their user. Their combined cost asked by the suppliers is the minimum supply price. The difference may be called "the net combined value of these two services to the user." He will choose the combination with highest net value.

This net value depends, then, on the one hand, on the choice of services made by their user. On the other hand, it depends on conditions outside of his control: viz., his utilities and prior probabilities and the costs of available services. His problem is to maximize the net value of the data producing and decision-making services, given those noncontrolled conditions.

Those familiar with what has been called "information theory" in the last two decades will have noticed that, so far, we have not used the concept of amount of information, measured in units called "bits." My uncertainty about a set of alternative events is the same as the amount of information that I would receive if that uncertainty were completely removed; that is, if I would know exactly which particular event does happen. Roughly speaking, this uncertainty is measured by the smallest average number of successive yes-and-no answers needed to completely remove uncertainty. This number depends roughly on the prior probabilities of events. Suppose, for example, that the following four events have equal probabilities (one-quarter each): the bride at the neighborhood church next Saturday will or won't wear a mini-dress, and her dress will or won't be white. To learn the color of the dress I need one yes-or-no question; so my uncertainty about color measures one bit. For the same reason, my uncertainty about both color and style is two bits since I need two yes-and-no answers to answer it. Thus the uncertainty, measured in bits, about those two mutually independent sets of events is the sum $(1+1=2)$ of the uncertainties about each of them. The number of bits is, in this sense, additive: a property that we require of every genuine measure, such as that of time, distance, volume, energy, dollar income, and dollar wealth.

If the four bridal events were not equally probable—for example, if the odds for a maxi-dress were not 1:1 but 9:1 (while a dress of each style were still as likely as not to be white)—the average necessary number of yes-and-no questions and thus the number of information bits would be smaller in the long run, i.e., over a long sequence of such events: for we can then profitably identify the more probable sequences (i.e., those mostly consisting of maxi-dresses) by asking a few questions only—as any skilled player of the "20 questions game" knows. As before, the count of bits agrees remarkably with the intuitive use of the

English word uncertainty: for when the odds are 9:1 I am almost certain, and with odds 1:1, I am fully ignorant, am I not?

Now suppose you have the choice between learning both the style and the color of the bride's dress and learning, with equal speed and for the same fee, the future price of your stocks. Suppose the price is as likely to rise as to fall. Depending on your selling or purchasing now, you may lose or double your fortune. A service that will tell you correctly whether the stock price will rise or fall conveys only one bit of information; whereas the service telling you correctly both the style and the color of the dress provides two bits. Yet you will prefer to learn about the stocks, not about the dress. There is, thus, no relation between the number of bits conveyed and the gross value of the data producing service. Nor does there seem to be a relation between the number of bits and the cost of a data producing service. For example, the cost of a sampling survey depends on its size, and this is not clearly related to the number of bits involved.

On the other hand, the number of yes-and-no symbols involved is clearly relevant to the performance and the cost of the transmission service regardless of whether these symbols refer to the length of the bridal skirt or to the trend of prices of your stock. To the economist, the contrast between production and transmission of data is strikingly analogous to the contrast between production and transportation of goods. A gallon of whiskey is more valuable than a gallon of gasoline: their costs to the producer and their values to the buyer are quite different. Yet to transport one gallon over one mile costs about the same for any liquid.

When, some twenty years ago, those elusive things labeled by the vague English words "uncertainty" and its negative, "information," were harnessed, subjected to genuine measurements (as was energy some hundred years ago, and mass and force much earlier), it was easy to understand the enthusiasm of people in elusive fields such as psychology. But also, to some extent, in statistics and in mathematics, where it was partly due to deep and beautiful theorems developed in this context. It is remarkable that C. Shannon who first proposed these theorems clearly limited their application to communications.

Statistical decision theory deals only with the choice of experiments and of decision rules; that is, with the choice of data producing and of deciding services. It omits the lower row of the chart reproduced above: encoding of data into signals, transmitting signals through a "communication channel," and decoding them back into messages that the decider would understand. In other words, for the statistician the decision is taken on the basis of a message which is simply the same thing as the data produced by the inquiry or experiment.

Not so with the communication engineer. His responsibility is to construct channels for the fast and reliable transmission of signals. (It all started in the telephone industry!) He is therefore also interested in devising appropriate codes which translate ordinary English into signals and signals back into English. But, to concentrate his mind on pure communication economics he makes, in effect, the following simplifying assumptions: First, events and data are identical, for he is not interested in the imperfections of the data producing service. Second, deciding is the same thing as decoding; so that action is simply the same thing as the message received. Third, as we have observed for the case of non-equiprobable events (which is, of course the usual case), the count of bits presupposes, in general, long sequences of events; and, as we shall see, such long sequences are also essential to make the crucial concept of "channel capacity" useful. Fourth, in most though not all[1] of their work, communication engineers assume an extremely simple result function. There are only two results: bad (say, minus one), when the decoded, received message is not identical with the datum sent; good (say, zero), when the two are identical. That is, all errors are equally important, have the same disutility, whether an inch is taken for a mile or merely a colon is taken for a semicolon. Finally, utility is assumed to be additive; i.e., it is conceived as the sum of certain measurable advantages and disadvantages, appropriately converted into dollars. We have seen that statisticans make the same assumption when they compare the sampling error with the dollar cost of the sample. The economist who detects and warns against this assumption is somewhat of a purist. The assumption is surely convenient for practical purposes and its removal is perhaps not that urgent.

Indeed this last assumption permits the engineer to ask the following economic question on behalf of the user of transmission channels and of coding services. Given the dollar costs of available channels, what is the best combination of the following evils: the probability of error (any error); the cost of the channel; the average time delay, which depends both on the length of signal sequences transmitted at a time and on the size of the code's vocabulary. That is, disutility is thought of, in effect, as a sum of dollars that buy a given channel; plus the dollar-equivalent of an error (any error); plus the dollar-equivalent, to a given user, of each time delay arising in the coding and transmission of a given datum. The user's problem is to choose that combination of channel and code which will minimize the sum of the averages of these amounts, weighted by appropriate probabilities. What do these probabilities depend on? On the uncertainty about data (=events); on the likelihood array characterizing the channel's reliability (the array of conditional probabilities

[1] Not, in particular, in Claude E. Shannon's work on a "fidelity criterion," which does correspond to a general result function.

of output symbols given an input symbol); and on the coding and decoding procedures.

Clearly, an appropriately redundant code can almost overcome the lack of reliability of the channel; that is, it can almost eliminate the occurrence of errors. For example, the encoder just lets every "yes" or "no" to be repeated many times, and the decoder takes "no" for the answer if he has received more "no"'s than "yes"'s. "Don't!—repeat, don't!—repeat, don't shoot!" If I have heard two don'ts and only one do, I shan't shoot. However, we may need great redundancy of the code if the channel is very unreliable; and this will cause long delays if the channel is slow. But a channel that is fast and reliable is expensive.

If the user can afford to wait for a long sequence of data to flow in before they are encoded, the problem of choosing between channels is simplified, for their variety is reduced as follows. Instead of a whole array of likelihoods (of channel output symbols, given each input symbol) it becomes sufficient to use a single reliability measure (in bits per input symbol) which, multiplied by the channel's speed of transmission (in input symbols per time unit), gives the channel's "capacity," in bits per time unit.[2] Provided this capacity is larger than the number of bits per time unit that characterizes the uncertainty and speed of the flow of data, it has been shown that the user can achieve any desired probability of errors, however small, by using an appropriate, though redundant, code. Assuming that such codes have indeed been constructed (quite a difficult problem, solved only to a small part), it would be for the user to weigh against each other the disadvantages of errors, of time delays and of the high costs of high-capacity channels.

To avoid errors in our mutual understanding, let me be redundant, mindful of my low transmitting capacity and of your limited memory. I said a short time ago that engineers have isolated the pure communication problem by not concerning themselves with the services that produce data and that decide on acting; and also by usually refusing to distinguish between important and unimportant errors. I also pleaded, a longer time ago, on behalf of economists who emphasize that the demand for all services, all the transformers on my chart, is a joint one. Indeed, the user can improve the reliability of messages on which decisions are based by improving the communication service, but also by improving the data producing service which he is free to choose. Similarly, the user (the "meta-decider") is free to choose the deciding service; for example, he may prefer not to burden the unskilled but inexpensive decider with messages written in a vocabulary that is too rich and fine.

[2] For example, two channels with equal transmission speeds and each characterized by the same array of likelihoods as, respectively, the old and the new barometer of our previous illustration have approximately equal capacities, in bits per second.

Moreover, depending on the user's result function, he may fear some errors of communication but be indifferent to others. He may be indifferent to the music of the voice at the other end of the telephone; so he does not really need a high-fidelity telephone.

On the other hand, statisticians have isolated their problem, also essentially an economic one, by omitting the communication components. As I said before, this may be a good research strategy. I am told that in the early space vehicles rectangular pieces of equipment were used although the vehicles had circular cross-section. That is, the problem of building a good battery (say) was solved separately from, not jointly with, the problem of building a fast vehicle. Our problem-solving (decision-making) capacity is limited to only a few good solutions per manhour. To take up all problems at once is desirable but not cheap and perhaps not feasible. However, as time goes on the joint approach should be tried. Hence this economist's appeal to both statisticians and engineers.

I have just said that the limitation of the research capacity of all of us explains and possibly justifies the fact that engineers and statisticians have broken up the economics of symbol manipulation into separate sections, neglecting the essential complementarity of the several services from the point of view of the demand by the user.

However, this separation seems to be partly justified also by the economics of the services themselves; viz., by the supply side. I mean in particular the conditions for the production, and therefore for the supply, of inanimate transmission channels, such as telephones, the broadcasting apparatus, perhaps even the products of the old-fashioned printing press.

To be sure, you may not be anxious to learn about the bridal dress and be very much interested in the stock market. Yet your morning newspaper will bring you both a society page (which you will throw away) and a stock market page. Any page costs as much to print as any other page. The cost depends on the number of symbols on the page, and this corresponds to the number of bits transmitted by the printed messages. And the cost per bit turns out to be smaller if every subscriber receives both the social page and the stock market report and the sports page and the political news, regardless of his special tastes. Similarly, I am forced to subscribe to a high-fidelity telephone service although I am not interested in the music of the other person's voice. I suppose this is due to the economies of mass production. It is cheaper to produce instruments that will minimize the probability of transmission error—any error, however unimportant to me personally—than to custom-make instruments which would suit people's individual preferences. Remember that, in this country at least, with its large total demand for clothing and for food, consumers do find it advantageous to buy ready-made suits and standardized groceries. To go to a Bond Street tailor or to buy fancy foods is slightly more satisfactory but so much more expensive!

The problem is familiar to operations researchers as that of optimal assortment. It is also known to social and economic statisticians, editors of census volumes, and makers of production indices. They call it "optimal aggregation." What indeed is the most economical way to break down a collection of items into groups, each to be treated as if it were homogeneous, when every detail suppressed involves some sacrifice, yet also saves cost?

Thus, it is just possible that, for the purposes of large markets (but not, I would think, for the purpose of building a particular satellite!) the isolated theory of transmission channels that minimize the probability of error—any error—is exactly what one needs. Yet, to be sure of this, we ought to have at least an approximate idea as to whether the services immediately complementing the transmission, that is, the services of coding, also exhibit advantages of mass production; and that the imperfections of available data producing and decision-making services are indeed negligible as to their economic effects.

Inanimate transmission channels do display the advantages of mass production. This makes it useful, when studying their supply conditions, to apply the pure theory of communication and to derive economically significant results from measuring information in bits. But what about other symbol-manipulating services: inquiry, coding, deciding? What can we say, in particular, about those supplied not by machines but by humans?

Before commenting on this most fascinating question, let us remind ourselves of the principles of the analysis of demand, supply, and the markets, and apply them, in particular, to the markets of symbol-manipulating services.

The demands of individual users are aggregated into the total demands for various data-providing services: total demands for weather forecasters and market prophets; for the output of research laboratories, for services of spies and detectives—given the prices of each of these services. Similarly with the total demands for various communication services—television, telephones, post office, newspapers, but also schools! —given, again, the prices of each. And so also with the demand for deciders—inventory clerks and vice-presidents for finance, and humble robots. Some of these services are substitutes for one another: for example, TV, radio and newspapers; telephone and mail. Some are mutual complements: the demand for weather data and for radio sets boosts each other.

Now, to explain the "given prices" in the markets and the kind and volume of transactions that actually come about, we need to know also the supply conditions. What does it cost to produce a market survey; to print a mass-circulation paper or a professional periodical; or to run a school? And to rear and train a vice-president or to build an automatic pilot? Again, the supply conditions are interrelated, although perhaps not as closely as the demand conditions. An automatic pilot combines the services of inquiring and of deciding, and it might be more costly to produce these services separately.

At any rate, the supply of a given service or a bundle of services—at given prices!—will depend on the costs of producing various kinds and amounts of them. Under competition the price will, then, equate demand and supply.

Is this not classroom economics? Yes indeed. But it should include the more advanced parts of it which allow for oligopoly, uncertainty, and other such things, mildly called "imperfections." Particularly important are the facts of indivisibility, or more precisely, the lack of homogeneity, of standardization of many of the symbol manipulators. There exist almost unique, irreplaceable research workers, teachers, administrators; just as there exist unique choice locations for plants and harbors. The problem of unique or imperfectly standardized goods is not peculiar to the economics of inquiring, communicating, and deciding. But it has been indeed neglected in the textbooks.

Let us return to the comparison of services supplied by men and by machines. The subject has seriously worried the most creatively imaginative pessimists of science fiction—from Karel Čapek to Ray Bradbury. It has also fascinated, and has led to some serious work of, psychologists and computer scientists. The results of this work, however tentative, are of great interest to us economists.

To begin with, humans are very poor transmission channels. "Indeed," says George Miller, a leading psycholinguist, "it is an act of charity to call a man a channel at all. Compared to telephone or television channels, man is better characterized as a bottleneck. Under optimal conditions it is possible for a skilled typist or piano player to transmit 25 bits per second. . . . We shall have to regard 25 bits per second as near the upper limit." More usually, the transmission capacity of an average person in our culture is only 10 bits or less, that is, we are unable to identify a stimulus without error when the stimulus is selected from more than 2^{10}, i.e., about a thousand equiprobable alternatives (that is, when the identification logically requires at least ten yes-or-no questions). As to the so-called "short memory," an important accessory of many transmission instruments, George Miller says that "no self-respecting photographic plate would even bother to sneer at us."

But what about the other symbol-manipulating services? Take coding. The lady who is typing the almost illegible manuscript of this lecture has an uncanny gift of recognizing the intended meaning of letters and words. I think she does this mostly by looking at the context of a whole sentence, or even of the whole paper itself. This we can interpret either by saying that she has the ability of encoding almost without errors the data presented in longhand into the symbols of the typewriter alphabet; or that she decodes the longhand symbols given to her, into messages, and these into actions; viz., into pressing the keys, mostly in such a way that no error occurs. As you know, the computer industry has just begun to construct machines that may one day match the human ability to recognize simple visual patterns such as hand-written individual letters (not whole sentences!). But some people believe it will take a very long time (generations or centuries, Y. Bar-Hillel thinks) until a machine can conduct "intelligent conversation" with a man or with another machine. The key words are "heuristics" and "intuition." They are as vague as "pattern recognition," "Gestalt," and "context," and all these words are perhaps intended to have the same meaning. It is remarkable, in fact, in this very context, that you and I vaguely understand each other as to what the word "context" is intended to mean. We understand each other not letter by letter, not even word by word, but by grouping symbols into large chunks—letters (or, rather, phonemes) into words, and words into sentences, and even into larger entities, each including all sentences with the same so-called "meaning." The chunks forming the vocabulary of a human coder or decoder are of course much less numerous than the ensemble of all possible combinations of a given number of letters, say. The use of chunks diminishes therefore the flow of signals through the channel; it is more efficient, more economical. Consider the three letters C, O, W. They can be combined in six possible ways. But only one of the six combinations occurs in your vocabulary: cow. And, remarkably, it invokes not just the few properties listed by taxonomists who define a cow, but a whole image of shapes and sounds and colors and the tail waving the flies away. A most efficient, economical code—the living human language!

Most important, when you, a man, talk to a fellow human you adjust your code to the receptor, and keep readjusting it, sensitive to his feedback responses. Is this not what characterizes a good teacher? James Boswell, young and smug, wrote in his diary:

> Health, youth, and gold sufficient I possess;
> Where one has more, five hundred sure have less.

He could as well have said: my wealth is at the uppermost half percentile. This would be economical if you would address income statisticians: you would utilize a "subroutine" that has been educated into them! But when you address other people, better recite a poem.

Even the talking to computers is better done by men, at least today. The encoding, or programming, of a difficult problem for a computer is said to be an art, not a science. People who say this probably mean precisely this: the activity of programming cannot be delegated by men to machines, at least not in serious cases and not in the present state of technology. Hence the very large proportion (one-half, I understand) that the human activity of programming contributes to the value added by computing organizations.

To turn from coding to inquiring services. A biochemist (J. Lederberg) and a computer scientist (E. Feigenbaum) have told a computer how to generate the graphs of all imaginable organic compounds of a certain class, and also how to match them with certain empirically observed spectra. This was essentially a job of mathematical routine. But now comes the heuristics! The biochemist had accumulated enough experience and a flair to eliminate as unrealistic all but a few patterns from the thousands that the computer had omitted. Yet the human being, a Nobel laureate, was not able to articulate his flair, although he did learn in this direction from the cooperation with the computer. Given the abilities and the technologies, there is some optimal way of allocating the tasks between men and machines—as economists have known long ago. And we must not be too hard on the computer: its hardware is certainly much less complex than the man's genetic endowment, and the computer's short babyhood is not rich in experience.

Finally, the service of decision making or problem solving. How to allocate tasks of this nature among executives and machines? A delicate problem! It involves all echelons of a corporation, up to, but of course not including, the president, who cannot fire himself. I had better skirt this subject!

But let me remind you of the distinction I have made earlier, between decision making and the higher-order activity of choosing who or what should provide a given service of decision making or of inquiring or of coding or transmitting. The man in charge was called the "leader" or 'organizer." It is his judgment of prior probabilities and utilities, his "beliefs" and his "tastes" (or "values," in other words), that are used among the "givens" of the organizer's problem. He cannot delegate them, either to men or machines.

His problem may be, in fact, much, much larger than my chart suggests at first glance. The economic problem of organization is that of allocating numerous kinds of tasks, symbol manipulating as well as physical, to numerous transformers, arranged in a complex yet efficient network. And further complications, of a different kind, arise when a single organizer is replaced by several. Their beliefs and utilities are not the same. They engage in a nonconstant sum game. The economist's problem is then shifted from the search for optimality to the search for stability: he tries to explain, as does the biologist or anthropologist, why certain arrangements, certain allocations of tasks and incentives (rewards) have a greater chance to survive over a given period than other arrangements, and under what conditions.

The criterion of survival, viability, stability guides the social scientist who describes, and tries to explain, the existing institutions. Yet not everything that is stable is desirable. Some wicked dictatorships have been quite stable. Along with the stability criterion, the economist uses a weak collective optimality criterion, a modest common denominator on which people might agree in spite of their divergent utilities and beliefs: an arrangement of tasks and incentives is optimal in this modest sense if there is no feasible arrangement that would be better or at least not worse for all members of the organization.

What, then, if we consider the whole society as an organization? How should incentives and tasks be allocated in a way that is stable or is collectively optimal, or, if possible, both? Further, some of us cannot help but smuggle in our own values, in particular a high valuation of liberty and equity. I suppose "public policy," "public good," in our tradition, mean somehow to reconcile the criteria of stability and of collective optimality with those of liberty and equity. Though the economic theorist prefers to hide behind the technical term "welfare economics," he means not just Secretary Gardner's Department of Health, Education and Welfare, but much more, the whole public policy. Nor is our special concern only education, even if taken in the broad sense of the communication of what my chart calls "data," to the whole or some parts of the public. For research, inquiry has been also our concern here. Public policy problems in the field of symbol manipulation are crudely exemplified by questions such as, "When, if at all, should the government subsidize or protect research and teaching and the dissemination of news?"

As far as I know, welfare economics of symbol manipulation is at its beginning. Special problems, such as the theory of patents and of public versus private broadcasting and, most importantly, of the economics of education, have been studied and the names of Silberston, Coase, Gary Becker come to mind.

On the more abstract level, a basic distinction exists between the information about external facts and the information conveyed to a member of society about the doings of others. A preliminary analysis of economic policy of information about external facts has been made by my colleague Hirshleifer. If correct, his conclusions on teaching and research are quite relevant to the California battle of tuition fees, although Hirshleifer's analysis had to be based on some extreme, simplifying assumptions. To analyze the economics of information of people about other people is even harder. Game theorists have provided some building blocks. Ożga has worked on "imperfect markets through lack of knowledge" and Stigler on the information in the labor market. It is just one year ago that Leijonhufvud told this Association that Keynesian unemployment may be mostly due to lack of information. We know very little about the technology of such information; for example, about the optimal language. Indeed, many believe that the run on gold is dammed, not by verbal announcements in English or even in French, but by actually selling gold to all comers. And Radner has penetratingly pointed to the setup cost of information which makes for increasing returns to scale and makes it difficult to apply the classical theory of free markets, which reconciles optimality and stability.

All this discussion, mostly by young members of our Association, is very recent, very exciting, and, I believe, very important. The informational revolution is upon us, and the manipulation of symbols dominates our lives more and more. I do hope we shall soon understand how to harness and benefit from those trends in our culture.