ENERGY AND INFORMATION

Author(s): Myron Tribus and Edward C. McIrvine

Source: *Scientific American*, Vol. 225, No. 3 (September 1971), pp. 179-190
Published by: Scientific American, a division of Nature America, Inc.
Stable URL: https://www.jstor.org/stable/10.2307/24923125

ENERGY AND INFORMATION

The flow of energy in human societies is regulated by the tiny fraction of the energy that is used for the flow of information. Energy and information are also related at a much deeper level

by Myron Tribus and Edward C. McIrvine

Scientia potestas est—"Knowledge is power"—the Romans said, and 20 centuries later science has given an old phrase new meaning. The power to which the Romans referred was political, but that is a small detail. Science does not hesitate to give precise definition to everyday words such as "work," "power" and "information," and in the process to transform proverbial truths into scientific truths. Today we know that it takes energy to obtain knowledge and that it takes information to harness energy.

Research into the relation between energy and information goes back many years, but the era of precise yet general quantification of information began only with Claude E. Shannon's famous 1948 paper "The Mathematical Theory of Communication." It has long been known that any dynamic measuring instrument placed in a system must draw some power in order to actuate its mechanism. For example, a meter connected to an electrical circuit uses some power to cause the deflection of a pointer. The effect is reciprocal: whereas the theory of instrumentation shows why energy is needed to obtain information, recent advances in information theory show why information is needed for transformations of energy. In this article we shall pursue both lines of thought.

Ideas about energy are part of the education of every scientist. Since the concept of energy is discussed in detail elsewhere in this issue, we shall not repeat that discussion. The fundamentals of information theory are less well known. We shall therefore dwell at some length on the fundamental ideas of information before we take up the interaction of energy and information.

Ideas about probability play a central role in any theory of knowledge. In modern information theory probabilities are treated as a numerical encoding of a state of knowledge. One's knowledge about a particular question can be represented by the assignment of a certain probability (denoted p) to the various conceivable answers to the question. Complete knowledge about a question is the ability to assign a zero probability ($p = 0$) to all conceivable answers save one. A person who (correctly) assigns unit probability ($p = 1$) to a particular answer obviously has nothing left to learn about that question. By observing that knowledge can be thus encoded in a probability distribution (a set of probabilities assigned to the set of possibilities), we can define information as anything that causes an adjustment in a probability assignment. Numerous workers have demonstrated that Shannon's measure of uncertainty, which he called entropy, measures how much is expected to be learned about a question when all that is known is a set of probabilities.

Shannon's contribution to the theory of information was to show the existence of a measure of information that is independent of the means used to generate the information. The information content of a message is accordingly invariant to the form and does not depend on whether the message is sent by dots and dashes, by impressing a particular shape on a carrier wave or by some form of cryptography. Once this invariance is understood it becomes an engineering task to design a communications channel. One intriguing aspect of communication theory lies in the observation that the merit of a particular communications-channel design lies not in how well the actual message is sent but in how well the channel could have sent all the other messages it might have been asked to convey. A voltmeter that accurately indicates 1,000 volts when connected across a 1,000-volt potential drop is not of any value if it always reads 1,000 volts no matter what the actual potential is!

Although Shannon's measure of uncertainty was postulated for the purpose of designing better communications channels (and has served admirably for that purpose), it has much broader applicability. After all, any piece of physical instrumentation can be viewed as a communications system. Thus a probe (for example a thermocouple, a pressure transducer or an electrode) serves as a sender. Amplifiers, wires, dials and mechanisms serve as a communications channel. The human observer serves as a receiver. One can apply Shannon's ideas not only to the design of the apparatus but also to the code used for conveying the information. It is in the latter respect that the connection between information and energy is most interesting.

First we must define Shannon's measure. Suppose we have defined a question, denoted Q, and are uncertain of the answer. The statement "We have defined a question" needs to be made precise. We require that all possible answers be enumerated and that our confusion be over which of the possible answers is the correct one. If we ask something without knowing what the possible answers are, then we have not really posed a question; we have instead requested help in formulating a question. In order to define Shannon's measure we must deal with a well-defined question Q and have in mind a set of possible answers without necessarily knowing which answer is correct. (Suppose our question is: "Which number will turn up on this roulette wheel?" The possible answers consist of the numbers on the roulette wheel, and our uncertainty arises over which number to select.)

To make things compact we let the symbol X represent our knowledge about Q. (In our example X stands for all the things we know about the roulette wheel including our experience with the casino owner, the past history of the wheel and the actions of the disreputable-looking person standing near the table.) This knowledge, X, leads to an assignment of probabilities to the various possible answers. Assigning $p = 0$ to any one answer is the same as saying, "That answer is impossible." Assigning $p = 1$ to an answer is the same as saying, "That answer is certain." Unless X is of a very special nature we shall end up assigning intermediate values between 0 and 1 to all the possible results. Shannon's measure is represented symbolically by $S(Q \mid X)$ to emphasize that the uncertainty or entropy S depends on both the well-defined question Q and the knowledge X [*see upper illustration on this page*].

$$S(Q|X) = -K \sum p_i \ln p_i$$

ENTROPY IN COMMUNICATIONS was formulated mathematically by Claude E. Shannon in 1948. Shannon's entropy S is defined in terms of a well-defined question (Q) and knowledge (X) about Q. In Shannon's formula the symbol K represents an arbitrary scale factor, and the sign Σ means to "sum over," or simply add up, for each possible answer to the question Q the product of the probability (p_i) assigned to that answer and the "natural" logarithm of the probability ($\ln p_i$). Shannon went on to define the information (I) in a message as the difference between two entropies, or uncertainties: one that is associated with knowledge X before a message and the other that is associated with knowledge X' after a message; in symbols, $I = S(Q \mid X) - S(Q \mid X')$.

$$S' - S = \int_{X}^{X'} \frac{dQ_r}{T}$$

ENTROPY IN THERMODYNAMICS was defined by Rudolf Clausius in 1864 in terms of a "transformation" that always accompanies a conversion between thermal and mechanical energy. According to Clausius' formula, when a system changes from a state described by X to another state described by X', the entropy change ($S' - S$) is calculated by dividing each increment of reversible heat addition (dQ_r) by the absolute temperature (T) at which the heat addition occurs and adding the quotients over the change from state X to state X'; the integration sign ($\int$) symbolizes this mathematical operation. It can be shown that Shannon's function and Clausius' function are the same.

The mathematical definition of Shannon's entropy has the interesting property that if one correctly assigns $p = 1$ to one of the answers and (therefore) $p = 0$ to all the others, S is 0. (If you know the right answer, you have no uncertainty.) On the other hand, if all the probabilities are equal, S is a maximum. (If your information is so slight that you must assign equal probabilities, you are as uncertain as possible about the answer.)

In the preceding discussion we used the knowledge X about a question to define the entropy S regarding the uncertainty of the answer. Conversely, we could have used S to define X by saying that any X that maximizes $S(Q \mid X)$ is a state of maximum ignorance about Q. A man who does not know one answer from another is as ignorant about Q as he can possibly be. The only state of greater ignorance is not to know Q. Hence we can use the Shannon formalism to describe X quantitatively. Otherwise X is a qualitative concept.

For a given question (Q constant) it is of course possible to have different states of knowledge. Shannon defined the information in a message in the following way: A message produces a new X. A new X leads to a new assignment of probabilities and thus a new value of S. To obtain a measure of the information Shannon proposed that the information (I) be defined by the difference between the two uncertainties: in symbols, $I = S(Q \mid X) - S(Q \mid X')$.

The information content of a message, then, is a measure of the change in the observer's knowledge (from knowledge X before the message to knowledge X' after the message). A message that tells you what you already know produces no change either in knowledge (X remains the same) or in probability assignment and therefore conveys no information.

Shannon's measure is an invention. It was designed to fill a specific need: to provide a useful measure of what is transmitted on a communications channel. It has also been shown to be the only function that satisfies certain basic requirements of information theory. In the 23 years since Shannon put forward his measure thousands of papers have been written on the subject and no one has found a replacement function, or even a need for one. On the contrary, many alternative derivations have been found. We conclude that the Shannon entropy measure is fundamental in information science, just as the Pythagorean theorem is fundamental in geometry. Accordingly Shannon's concept of entropy should be a useful starting point for reasoning about information processes in general.

The word "entropy" had of course been used before Shannon. In 1864 Rudolf Clausius introduced the term in his book *Abhandlungen über die mechanische Wärmetheorie* to represent a "transformation" that always accompanies a conversion between thermal and mechanical energy. If a physical system changes from a state described by X (a particular combination of pressure, temperature, composition and magnetic field, for example) to another state defined by X' (a different combination of pressure, temperature, composition and magnetic field), then according to the Clausius definition, the entropy change is calculated by dividing each increment of heat addition by the absolute temperature at which the heat addition occurs and adding the quotients [*see lower illustration on this page*]. Except for the fact that Shannon's entropy and Clausius' entropy are represented by the same symbol and the same name, there appears at first sight nothing to indicate that the two functions are in fact the same function.

What's in a name? In the case of Shannon's measure the naming was not accidental. In 1961 one of us (Tribus) asked Shannon what he had thought about when he had finally confirmed his famous measure. Shannon replied: "My greatest concern was what to call it. I thought of calling it 'information,' but the word was overly used, so I decided to call it 'uncertainty.' When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage.' "

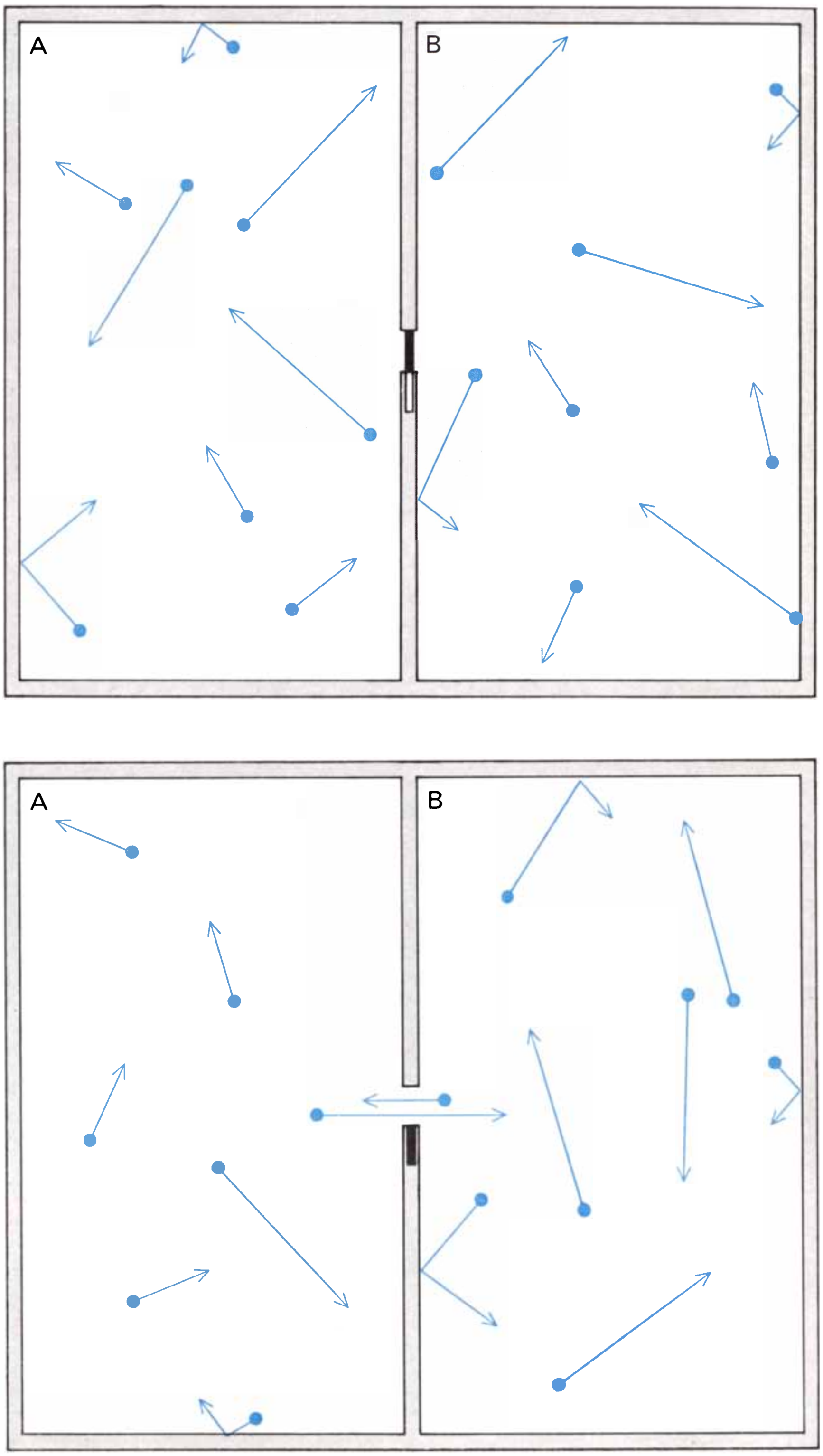
The point behind von Neumann's jest is serious. Clausius' definition of entropy has very little direct physical appeal. It can be derived with satisfactory mathematical rigor and can be shown to have interesting and useful properties, particularly in engineering, but in a direct aesthetic sense it has not been satisfactory for generations of students. Simple physical arguments lead one to believe in the correctness of most quantities in physics. Surrounding Clausius' entropy there has always been an extra mystery.

The appearance of Shannon's measure, with the same name and the same functional representation as the earlier measure in statistical thermodynamics, aroused great interest among theoretical physicists. One of the best-known contributors to the subsequent discussion was Leon Brillouin, who treated the two entropies as the same in a series of papers and in the book *Science and Information Theory*. The proof that they are indeed the same (and not merely analogues) has been dealt with extensively elsewhere and will not be treated here.

The unit of information is determined by the choice of the arbitrary scale factor K in Shannon's entropy formula. If K is made equal to the ratio $1/ln\ 2$ (where the expression $ln\ 2$ represents the "natural" logarithm of 2), then S is said to be measured in "bits" of information. A common thermodynamic choice for K is kN, where N is the number of molecules in the system considered and k is 1.38×10^{-23} joule per degree Kelvin, a quantity known as Boltzmann's constant. With that choice for K the entropy of statistical mechanics is expressed in units of joules per degree.

The simplest thermodynamic system to which we can apply Shannon's equation is a single molecule that has an equal probability of being in either of two states, for example an elementary magnet. In this case both p_1 and p_2 equal 1/2, and hence S equals $+k\ ln\ 2$. The removal of that much uncertainty corresponds to one bit of information. Therefore a bit is equal to $k\ ln\ 2$, or approximately 10^{-23} joule per degree K. This is an important figure, the smallest thermodynamic entropy change that can be associated with a measurement yielding one bit of information.

In classical thermodynamics it has long been known that the entropy of mixing, per molecular weight of mixture, is a function of the fractional composition. Obviously the fractional concentration of a particular molecular species represents the probability of picking out a molecule of that species in a random sampling of the mixture. What does the act of mixing signify if we use the entropy of mixing as a measure of information? Imagine that we mix half a molecular weight of each of two isotopes. The resulting entropy change would be $N_0 k\ ln\ 2$, where N_0 (Avogadro's number, 6×10^{-23}) is the number of molecules per molecular weight. Numerically this change is about six joules per degree K., or 6×10^{23} bits. The latter number represents the number of



MAXWELL'S DEMON, a hypothetical being invoked by James Clerk Maxwell in 1871 as a possible violator of the second law of thermodynamics, was assumed to operate a small trapdoor separating two vessels full of air at a uniform temperature (*top*). By opening and closing the trapdoor so as to allow only the swifter molecules to pass from *A* to *B* and only the slower ones to pass from *B* to *A*, the demon could, without expenditure of work, raise the temperature of *B* and lower that of *A* (*bottom*), in contradiction to the second law of thermodynamics. The demon was finally "exorcised" in 1951 by Leon Brillouin, who pointed out that if the demon were to identify the molecules, he would have to illuminate them in some way, causing an increase in entropy that would more than compensate for any decrease in entropy such a being could effect. Without the input of energy represented by the illumination, the demon lacks sufficient information to harness the energy of the molecules.

decisions that would have to be made if a person were to sort the isotopes one at a time.

Just as the entropy of information has meaning only in relation to a well-defined question, so the entropy of thermodynamic analysis has meaning only in relation to a well-defined system. In our present understanding of physical science that system is defined by quantum theory. The question is: "In what quantum state is this system?" The answer is: "In some statistical combination of states defined by the quantum-mechanical solutions of the wave equation." In fact, these solutions define the possibilities we alluded to in discussing information and uncertainty. The probabilities encode our knowledge about the occupancy of the possible quantum states (possible, that is, for a given state of knowledge).

The concept of an inherent connection between the entropy of Clausius and the intuitive notion of information preceded Shannon's work by many years. In fact, the information-theory approach to thermodynamics is almost as old as thermodynamics itself. Clausius' 1864 book represents the earliest complete formulation of classical (nonstatistical) thermodynamics. By 1871 James Clerk Maxwell had introduced the role of information by proposing his famous demon [see "Maxwell's Demon," by W. Ehrenberg; SCIENTIFIC AMERICAN, November, 1967]. He suggested that a demon of minute size ought to be able to operate a small trapdoor separating two vessels, permitting fast molecules to move in one direction and slow ones in the other, thereby creating a difference in temperature and pressure between the two vessels [*see illustration on preceding page*]. Maxwell's demon became an intellectual thorn in the side of thermodynamicists for almost a century. The challenge to the second law of thermodynamics was this: Is the principle of the increase of entropy in all spontaneous processes invalid where intelligence intervenes?

From Maxwell's time on many leading investigators pondered the relation between observation and information on the one hand and the second law of thermodynamics on the other. For example, in 1911 J. D. Van der Waals speculated on the relation between entropy change and the process of reasoning from cause to effect. In 1929 Leo Szilard commented on the intimate connection between entropy change and information. In 1930 G. N. Lewis wrote: "Gain in entropy means loss of information; nothing more." Until Shannon came on the scene, however, there was no measure of information, so that the discussions could not be quantitative.

What Shannon added was the recognition that information itself could be given a numerical measure. If any of the early thermodynamicists had chosen to do so, he could have defined information to be consistent with thermodynamic entropy. After all, the "entropy of mixing" was well known. Any one of those men could have chosen to define information as the number of decisions required to "unsort" a mixture. The Shannon measure would have followed.

Shannon, who had no direct interest in thermodynamics, independently developed a measure of information. For practical reasons he chose to require the measure to meet certain logical criteria of consistency and additivity. In retrospect a logician can show that with these criteria Shannon was bound to produce a measure that would be consistent with thermodynamic entropy. Once it is recognized that the two subjects derive from common considerations it is straightforward to derive one from the other.

The Shannon formulation is somewhat more general since it is entirely a mathematical theory and is applicable

ACTIVITY	ENERGY (JOULES)	INFORMATION CONTENT (BITS)	ENERGY PER INFORMATION (JOULES PER BIT)
CHARACTER RECORD ACTIVITIES:			
TYPE ONE PAGE (ELECTRIC TYPEWRITER)	30,000	21,000	1.4
TELECOPY ONE PAGE (TELEPHONE FACSIMILE)	20,000	21,000	1
READ ONE PAGE (ENERGY OF ILLUMINATION)	5,400	21,000	.3
COPY ONE PAGE (XEROGRAPHIC COPY)	1,500	21,000	.07
DIGITAL RECORD ACTIVITIES:			
KEYPUNCH 40 HOLLERITH CARDS	120,000	22,400	5
TRANSMIT 3,000 CHARACTERS OF DATA	14,000	21,000	.7
READ ONE PAGE COMPUTER OUTPUT (ENERGY OF ILLUMINATION)	13,000	50,400	.3
SORT 3,000-ENTRY BINARY FILE (COMPUTER SYSTEM)	2,000	31,000	.06
PRINT ONE PAGE OF COMPUTER OUTPUT (60 LINES x 120 CHARACTERS)	1,500	50,400	.03

RATIOS OF ENERGY TO INFORMATION for various information-preparation, information-processing and information-distribution activities involving symbols are presented in this table. The energy values used are those typically involved in powering the mechanisms employed for these activities and in most cases are accurate only to about an order of magnitude. For character records information content is assumed to be about seven "bits" per character. The energy/information ratio for information systems based on character records and digitally encoded records varies from a few joules per bit down to a few hundredths of a joule per bit.

to all kinds of uncertainty. The thermodynamic theory is less general since it is bound to our "real world" environment of atoms, molecules and energy. Brillouin tried to emphasize this distinction by speaking of "free information" for the abstract Shannon quantity and "bound information" for the quantity when it described physically real situations (and thus thermodynamics). As we shall see in considering practical information-processing activities, the only real distinction is that natural physical situations involve much larger amounts of information than we appear able to control in our human-oriented information systems. There is no conflict between abstract Shannon information and thermodynamic information, as long as the questions we ask are physically real questions.

An interesting application of information theory to thermodynamics is provided by Brillouin's "exorcising" of Maxwell's demon. As we mentioned above, the demon led to apparent thermodynamic paradoxes. Brillouin pointed out that if the demon were to see the molecules, he would have to illuminate them in some special way. Since the black-body radiation in a gas vessel is the same in all directions, without a torch the demon would have no way of distinguishing the location of individual gas molecules. It takes special information to harness the energy of the molecules, and this information is above and beyond the normal thermodynamic information that serves to distinguish the system itself from its surroundings. Without the departure from equilibrium represented by the torch, Maxwell's demon lacks the information on which to act.

Unlike the demon, we do not live inside a gas vessel in equilibrium at a uniform temperature. Suppose for a moment that we did. Imagine that the earth is contained in a totally absorbing "black box" at a uniform temperature of 290 degrees K. (63 degrees Fahrenheit), a reasonable estimate of the real earth's average surface temperature. We would be as helpless as Maxwell's demon without a torch. In spite of the large energy flux and the moderate average surface temperature, an earth at equilibrium in an ambient environment would be inhospitable to life. No information could be processed; no energy would be available in the thermodynamic sense.

The actual case is of course different. The earth is part of a "sun-earth-space" system that is quite out of equilibrium. The sun plays the same role for us that the torch did for Maxwell's demon. By providing a departure from equilibrium it becomes a source of information and useful energy.

In considering the human use of energy and information, we must take into account the radiation balance of the earth's surface. The earth receives 1.6×10^{15} megawatt-hours of energy from the sun each year in the form of solar electromagnetic radiation, and it reradiates this energy principally as black-body radiation. Thus the earth approximately balances its energy budget. Man's use of energy on the earth's surface actually constitutes internal transactions with energy fluxes that are thermodynamically available, that is, usable before the energy is thermally degraded to the average surface temperature or chemically degraded by diffusion to the environment. Taking commonly accepted average values for the temperatures of the sun and the earth, the 1.6×10^{15} megawatt-hours of energy radiated to outer space carries with it the capability for an entropy decrease, or "negentropy flux," of 3.2×10^{22} joules per degree K. per year, or 10^{38} bits per second.

Of course, this negentropy flux derives from the energy flux from the sun to the earth to deep space. By storing the energy and negentropy in various systems (fossil fuels, lakes, clouds, green plants and so forth), the earth creates subsystems that are out of equilibrium with the general environment. In addition to the solar flux, then, we have the stored energy and negentropy of the earth's resources. Only in the case of the potential use of deuterium in fusion reactors does this stored energy exist in amounts significantly greater than one year's solar-energy flux. For practical purposes we may consider that both the energy flux and the negentropy flux at the earth's surface are due to solar processes. On the surface of the earth, therefore, the maximum steady-state rate at which information can be used to affect physical processes is of the order of 10^{38} bits per second. A great deal of this information is "used" in meteorological processes (cloud formation, thunderstorms, the establishment of high-altitude lakes and watersheds and so on). A large additional amount is "used" for the life processes of plants and animals. A comparatively small quantity is under the control of man, yet this quantity is responsible for man's technological reshaping of his environment.

The limitation of 10^{38} bits per second is not a stringent one. Consider the information rate of a television broadcast. Television stations broadcast 30 frames per second, each frame containing 525 scan lines. The resolution along each scan line allows about 630 bits of information to be encoded. The resulting information rate is $30 \times 525 \times 630$, or 10^7 bits per second.

Suppose we now consider a totally nonredundant television broadcast: one in which each dot is uncorrelated with the other dots on the same frame and each frame is unrelated to other frames. No human being could possibly absorb information from a television tube at such a rate. Even if the material being broadcast were a typical printed page (and on such a page there is a great deal of correlation between dots), it would take a person of reasonable skills about 60 seconds to absorb the information from one frame. Thus we can estimate the probable bit rate required to engage a human being in intellectual attention as being less than 10^4 bits per second. With a world population of less than five billion, the entire human race could have its information channels individually serviced and saturated with a bit rate of 5×10^{13}, very much less than the 10^{38} bits per second available.

Many old maxims point out that talk is cheaper than action. A comparison of the entropy balance and the energy balance on the surface of the earth indicates that the maxims are indeed a reflection of our experience. The amplification of information is easier than the amplification of power.

The total muscle-power output of the human race is estimated to be about 3×10^9 megawatt-hours per year (somewhat less than one megawatt-hour per person per year). Worldwide power usage under human direction is of the order of 7×10^{10} megawatt-hours per year, so that the energy-amplification ratio currently is about 25 to one. In the U.S. the amplification is much larger: approximately 250 to one. If all the thermodynamically available solar energy were used, an amplification of 500,000 to one is theoretically possible.

The possible amplification of information-processing activities is much greater. By eliminating the human operator from the chain of data-processing and machine control over physical systems, any direct dependence on the natural rate of human data-handling is removed. A human operator, working with a fixed set of questions, can use a modern digital computer to amplify his abilities by a factor well in excess of 10^6, perhaps by a factor of 10^{12}. The weak limitation on information-processing rates due to radiation to space (10^{38} bits per second) would indicate the theoretical maximum amplification to be in excess of 10^{24}.

It is worth observing that this great gap between the achieved and the achievable gives information technology a character different from that of materials technology or energy technology. In materials technology and energy technology scientists are accustomed to studying fundamental limitations and natural structures and engineers are accustomed to designing within a few orders of magnitude of these limitations. In information technology scientists find fundamental theorems not at all restrictive, and entrepreneurs discover that the freedom from constraints makes possible the construction of an almost totally new environment of information. Hence the advent of television programming, automatic telephone solicitation, computer-generated junk mail and other artifacts of an "information overload" culture. In the case of material and energy nature often cries "Halt!" to the changes wrought by technology. In the case of information man himself must issue the directives to ensure that technology is used for human betterment.

We have noted Brillouin's exorcising of Maxwell's demon. Beginning with this act Brillouin was led to investigate the relation between the entropy of an observation and the thermodynamic entropy, and he concluded that one bit of information requires $k \ln 2$ thermal entropy units. As Dennis Gabor once put it: "You cannot get something for nothing, not even an observation."

This comment has a special meaning for those of us who are engaged in the design of xerographic copying equipment. Among other reasons for concern about information and energy processes, we are interested in the minimum energy requirement for making a copy. The actual energy used at present is inconveniently high (mainly because the fixing of the final image takes about 90 percent of the power and involves thermally fusing the black toner to the paper), and so we shall not discuss the subject further. We discovered, however, that reading our copies often takes much more energy than making them. (A typical reading requirement might be 5,400 joules from a 90-second exposure of a 60-watt lamp.) Just as Maxwell's demon could not see molecules without a torch, so we cannot see images without illumination. Some luminous flux must be present involving electromagnetic radiation from a temperature greater than that of the environment. In the process of thermally degrading that energy a signal is generated. The distribution of reflectivity over the surface of the paper modulates the negentropy flux from the illuminator. Both the paper and the illuminator are required for the retrieval of information.

So far we have concentrated on the information needed to harness energy. Now it is time to examine the practical aspects of the energy requirements of information systems. The world of information technology includes both digital and analogue representations and the uses we make of them. The dramatic rise of the utility of digital computers sometimes leads us to overlook the other common representations of information: images and audio signals. For all three species of information representation we can consider the following distinguishable activities: the preparation of records, the storage of records, the processing of records and the distribution of records.

Information storage is a passive activity, and it does not intrinsically require the continuous input of energy (although in fact some forms of storage, such as a semiconducting digital-computer memory, do have power requirements). Information preparation, information processing and information distribution all require energy. The table on page 182 lists the information content and the energy in certain practical equipment configurations used for a number of activities involving character records and digitally encoded character records. Since the characters are chosen from a limited set, the information content is about seven bits per character. (In the case of nondigital characters we neglect

ACTIVITY	ENERGY (JOULES)	INFORMATION CONTENT (BITS)	ENERGY PER INFORMATION (JOULES PER BIT)
AUDIO RECORD ACTIVITIES:			
TELEPHONE CONVERSATION (ONE MINUTE)	2,400	288,000	.008
HIGH FIDELITY AUDIO RECORD PLAYBACK (ONE MINUTE)	3,000	2,400,000	.001
AM RADIO BROADCAST (ONE MINUTE)	600	1,200,000	.0005
PICTORIAL RECORD ACTIVITIES:			
TELECOPY ONE PAGE (TELEPHONE FACSIMILE)	20,000	576,000	.03
PROJECTION OF 35 MM SLIDE (ONE MINUTE)	30,000	2,000,000	.02
COPY ONE PAGE (XEROGRAPHIC COPY)	1,500	1,000,000	.002
PRINT ONE HIGH QUALITY OPAQUE PHOTOGRAPHIC PRINT (5" x 7")	10,000	50,000,000	.0002
PROJECT ONE TELEVISION FRAME (1/30 SECOND)	6	300,000	.00002

LOWER ENERGY/INFORMATION RATIOS generally exist for information activities involving audio and pictorial representations. This table overstates information content since the full bandwidth of available frequencies is never used in audio activities and the typical pictorial record contains a great deal of redundancy. As a result the energy/information values are in most cases accurate to less than an order of magnitude, but they are suggestive. It is clear, for example, that the information system that uses the smallest amount of energy per unit information is a hypothetical nonredundant television broadcast. Even this comparatively efficient information system uses for the purpose of communication only a tiny fraction of the thermodynamic information it requires to operate.

Basic Research at Honeywell
Research Center
Hopkins, Minnesota

Optimal Control of Oil Refineries

A Unique Application of Optimal Control Technology May Improve the Yield and Profit of the Catalytic Cracking Process.

Optimal control synthesis techniques offer a basis for more efficient control of many complex, industrial processes where several variables must be carefully controlled. Problems of this kind are common in the hydrocarbon industry. Although the efficiencies of petroleum companies have increased steadily, there appears to be a strong need for more precise control of their operations, especially among process functions and their associated systems control interfaces. Typically, control functions in the hydrocarbon industry are carried out semi-automatically under control of an operator who monitors a large control board. In the catalytic cracking process, as well as in many other processes, hazards to safety and equipment are always present. For example, reactor and regenerator temperatures can reach dangerous levels very quickly. In addition, a small percentage increase in yield or quality of yield can amount to sizeable dollar amounts.

Quadratic optimal control theory, and computer techniques for rapid solution of the optimizing Ricatti equations, have both been available to the control engineer for some time, but they have not been used extensively in the petroleum refining industry. Petroleum companies have been reluctant to change over to on-line computer installations because the benefits of continuous operation far outweigh any possibility of a shutdown. Early computers were unreliable and expensive and, thus, computers in refineries have been used almost exclusively in an off-line mode. Another great deterrent is the fact that the optimal solution obtained is a result of feedback of the state of the system whose variables are not easily measured.

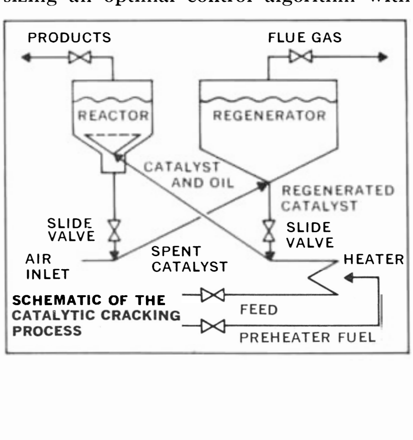
Two means of coping with this problem are: (1) estimation of the full state of the process from the measurable quantities available and (2) simplification of the model or controller or both to include only measurable quantities. State estimation has usually led to a more costly control than seems warranted for the improved performance obtained. However, model simplification introduces more uncertainty about the correspondence of model and performance. A number of recent investigations have focused on this problem of optimal control.

Honeywell scientists and engineers, in attempting to improve the safety and yield and profitability of hydrocarbon processing, have developed this latter control method to the point where it is useful for synthesizing linear controllers for complex multivariable processes when a specified sub-set of the state variables is measurable.

Honeywell chose the catalytic cracking process for its initial study because of its importance to the hydrocarbon industry. This process converts a high-boiling fraction of distilled crude oil into lower-boiling materials such as gasoline.

Several variables in this process can be adjusted by the operator or automatically: rate of air delivery to the regenerator, reactor catalyst level, catalyst flow rate, fuel gas to feed pre-heater, and input feed rate. The objective is to maintain, by proper settings and manipulation of these variables, the conditions for required product yields as well as safe stable process operation. It is particularly important to maintain the regenerator temperature and flue gas oxygen below certain specified values. The principal disturbance affecting the operation of the process is the fluctuation of feed properties.

The premise at Honeywell was that if a mathematical model and a computer program could be formulated for the cracking process, an on-line computer could monitor and automatically set the controls on the control panel. Honeywell engineers and scientists pioneered in applying these operational techniques to the catalytic cracking process by formulating a mathematical model of a catalytic cracker and synthesizing an optimal control algorithm with eight multi-variable control laws of varying degrees of complexity.



SCHEMATIC OF THE CATALYTIC CRACKING PROCESS

The model describes the process in terms of seven non-linear heat and material balance equations, plus a group of kinetic equations relating conversion, catalyst fouling, and carbon burning to the process parameters. The model also provides performance indices in terms of product yields and maintenance of safe operating conditions. The most complex control law postulates five control variables operating on information fed back from seven process variables. The simplest uses two control variables and two process measurements. The investigating team wrote a computer program which simulates the catalytic cracking unit based on the mathematical description and the control laws which were derived. The initial results based on data from an operating unit were promising. These data were analyzed by scientists and engineers from Honeywell's Corporate Research Center, from the Aerospace and Defense Group's Systems and Research Center, and by the process control specialists in Honeywell's Industrial Division.

Honeywell will soon begin a joint venture with a petroleum company in an effort to verify these results in an actual installation. Sensors and data recorders in the refinery will monitor and record temperature, fluid flow and pressure fluctuations. This information will then be fed into a computer which is wired into the control boards of the refinery. The input of information will then automatically alter the settings on the control board. This control system will not eliminate the job of the operator who monitors the control board but it will increase efficiency and decrease the risk of operational error or explosion.

If you are working in the area of quadratic optimal control, applied to the petroleum refining industry, and want to know more about Honeywell's investigations, please contact Mr. Jack Post, Honeywell Corporate Research Center, Hopkins, Minnesota. Honeywell carries out basic research in all of the sciences pertinent to its business at its Solid State Electronics Center, Plymouth, Minnesota, and Information Sciences Center, Cambridge, Massachusetts, as well as its Corporate Research Center, Hopkins, Minnesota, under the direction of Dr. John Dempsey, Vice President of Science and Engineering.

Honeywell

The Automation Company

the possibility of added information in the form of changes of font, boldface, italics and other additions to sets of characters.) The energy is the energy involved in powering the mechanisms typically employed in these activities and in illuminating pages for reading. Most of the values are accurate to within about an order of magnitude. In the table we calculated the illumination from a 60-watt lamp. Obviously one can also read under a high-power arc lamp or by the light of a candle. The energy per information unit typically varies from a few joules per bit down to a few hundredths of a joule per bit.

Similar information is contained in the table on page 184 for practical equipment configurations used in activities involving audio representations and pictorial-image representations. We have used typical values for slide-projector illumination, radio and television receiver power and telephone central-office power. Clearly the values vary in individual instances. Additional uncertainties arise from our estimates of information content. Audio-intensity modulation, pictorial gray scale and color constitute multilevel coding techniques. For multilevel coding the information capacity of a channel is related to the signal-to-noise characteristics as well as to the frequency bandwidth. From a practical point of view, however, the table overstates information content; the full bandwidth is never used in audio activities and the typical pictorial record has a great deal of redundancy. The values are thus accurate to less than an order of magnitude, but they are nonetheless suggestive.

An extreme case of redundancy attends a 35-millimeter pictorial image of a page of print. Assuming a resolution of 100 line pairs per millimeter, about two million bits are used to represent approximately 3,000 characters. Most of the information is redundant and is used to convey the white spaces, the details of the character font and other material that may be of no importance to the message. Even a four-letter word could take up a million bits in a high-resolution photograph. At the other extreme a simple, nonredundant, two-level Baudot code can be used to represent the same four-letter word in only 24 bits.

Information technology has as one of its present concerns the best use of energy and physical structures to convey information as needed for human purposes, without undue redundancy and yet with veracity, style and taste. As a part of that concern the joint processes of energy flow and information flow are of special interest. We are led back to the consideration of the thermodynamic functions of a physical system that is involved in information processing. In the discussion that follows we shall make use of ideas recently developed by Robert B. Evans, now at the Georgia Institute of Technology, who has devoted a decade to the unraveling of the question.

The information-theory treatment of thermodynamics clarifies the concept of equilibrium. A few moments' thought should serve to convince one that the concepts "Distinguishable from the environment" and "Out of equilibrium" are the same. Our ability to recognize a system depends on the fact that it differs from its environment. "Thermodynamic information" is conceptually the same as "Degree of departure from equilibrium." If each of these quantities is measured in such a way as to satisfy the elementary properties of additivity, consistency and monotonic increase with the system's size, then apart from units of measure each will be the same mathematical expression, since they really refer to the same thing.

Thermodynamic information is defined as the difference between two entropies: $I = S_o - S$. S refers to the entropy of a system of given energy, volume and composition. S_o is the entropy of the same system of energy, volume and composition when it is diffused into (indistinguishable in) a referenced environment. It measures the loss of information in not being able to distinguish the system from its surroundings (as when an iceberg melts in the open sea).

The idea of using thermodynamic information as a generalized measure of the "availability" of energy was first put forward tentatively by Evans in 1965. (Although heat energy, mechanical energy and chemical energy can be converted into one another, they are not equally "available" to do work. What we call Carnot efficiency and Gibbs free energy were invented to deal with the availability of energy.) By 1969 Evans submitted his doctoral dissertation containing an entirely classical proof that a new quantity, obtained by multiplying his formula for thermodynamic information by an appropriate reference temperature, has most unusual properties. Evans has called this new function "essergy," for the essential aspect of energy [*see illustration at left*]. He has demonstrated that essergy is a unique measure of "potential work." Moreover, it incor-

$$S_o = \frac{E + P_oV - \sum \mu_{io} N_i}{T_o}$$

$$I = \frac{E + P_oV - T_oS - \sum \mu_{io} N_i}{T_o}$$

$$T_o I = E + P_oV - T_oS - \sum \mu_{io} N_i$$

THREE EQUATIONS show the derivation of the concept of thermodynamic information. The top equation is based on classical thermodynamics: S_o is the uncertainty when energy (E), volume (V) and the number of moles of various chemical species (N_i) are unrecognizable because they are distributed in an environment at a temperature T_o, a pressure P_o and chemical potentials μ_{io}. The middle equation is derived from the top one on the basis of the relation $I = S_o - S$, where I is information and S is the uncertainty about the system formed with energy E, volume V and composition N_i, and the system is now discernible from the environment. These equations were derived by Robert B. Evans, now at the Georgia Institute of Technology, in 1969. He showed that a new quantity obtained by multiplying the middle equation by T_o is the most general measure of disequilibrium or "potential work." Evans has named this new quantity "essergy" (for the essential aspect of energy).

porates as special cases all previously known measures of departure from equilibrium (such as Gibbs free energy, Helmholtz free energy, the function used in Germany under the name "exergy," the Keenan availability function and so on). One can also think of thermodynamic information as a fundamental quantity representing the "signal-to-noise ratio" for a system in an environment T_0. Recall that kT_0 is the thermal-noise energy per degree of freedom, and that Evans essergy is $T_0 I$. When I is of the order of $100kT_0$ or less, one is dealing with information processing. When I is larger, one is dealing with work and power. For example, in the flux of electromagnetic essergy the flux of essergy works out to be the same as what is called the Poynting vector. Close to a radar antenna the flux of essergy is large enough to cook a man; far away it becomes a weak signal.

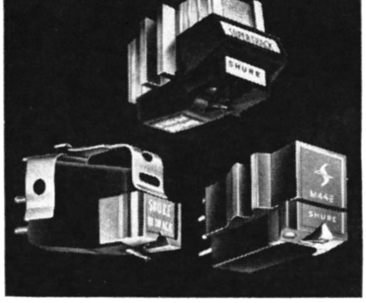
Evans' essergy function has proved useful in analyzing power cycles, particularly for economic optimization. After all, energy can be neither created nor destroyed, so that an energy balance for a steam power plant is not very enlightening. An essergy balance, however, enables one to track down specific plant inefficiencies and see what they cost. Looking at the earth as a whole, that solar energy flux of 1.6×10^{15} megawatt-hours per year would be useless if its reradiation to outer space were not accompanied by an entropy flux of 3.2×10^{22} joules per degree K. (The corresponding change in essergy is -2.6×10^{15} megawatt-hours per year.) Ideally an essergy analysis could be performed on the natural processes of the earth's surface. This would form a foundation for ecologically sound planning of energy utilization by man, since it would provide an indication of the disturbance caused by proposed large-scale changes. The usefulness of the essergy function in practical matters has already been demonstrated in a comparative analysis of methods of making fresh water from seawater and a generalized study of power plants.

If we return to the examples of information-systems activities listed in the tables on pages 182 and 184, it is now possible to compare energy and information fluxes by comparing the contributions to essergy change. The example that used the smallest amount of energy per unit information was the nonredundant television broadcast, with 300,000 bits of information at a cost of six joules. Yet that information change represents a contribution of only 1.3×10^{-15} joule of essergy. Indeed, in terms of energy requirements the information-processing aspect of a television broadcast is a sidelight. What one is mostly doing with the energy is heating the room. A similar analysis would show that each example of what we have called an information-systems activity is in fact an energy activity that carries with it a small amount of information. Not until information technology reaches the state of handling information at one bit per molecule will we be addressing real information processes in the thermodynamic sense. Perhaps it is this wide gap between the amounts of entropy that are measured in practical information systems and the entropy of physical systems that has led to a reluctance on the part of classical thermodynamicists to adopt the information-theory view of the foundations of thermodynamics.

It has been a decade since the appearance of the first textbook to develop thermodynamics on the basis of information theory. The interval has not seen a conversion from the classical tradition. The classical treatment seems more entrenched than ever, to judge from recent conferences of teachers of thermodynamics. In view of the ever increasing importance of information technology in the scientific world, this is unfortunate. Perhaps the reason lies in the comments of the Russian worker A. I. Veinik, who wrote in 1961 apropos of his own attempt to restructure thermodynamics: "The traditional separation [into the] branches of physics is due to historical reasons rather than to the nature of the subject matter. A logical development of science must lead inevitably to the unification of these fields, and such a unification has not yet come about only because the century-old traditional outer trappings adorning the queen of the sciences, classical thermodynamics, have been defended with particular zeal. Apparently a large part of this defense has been based on the authority of the geniuses who brought this queen of sciences into existence."

Whatever the short-term outcome in educational circles, it is certain that the conceptual connection between information and the second law of thermodynamics is now firmly established. The use of "bound information" in the Brillouin sense of necessity involves energy. The use of energy, based on considerations of thermodynamic availability, of necessity involves information. Thus information and energy are inextricably interwoven. We may well ponder the wisdom of the observation *Scientia potestas est.*

To each his own.



Not *everybody* needs a concert grand piano, nor does everybody *need* the best cartridge Shure makes to enjoy his kind of music on his kind of hi-fi system. Eventually, you'll want the renowned *V-15 Type II Improved,* the peerless cartridge for advanced systems and ample budgets. But, if your exchequer is a little tight, consider the *M91E,* widely acclaimed as the *second best* cartridge in the world. With a sharply circumscribed budget, all is *far* from lost. Choose any of the four models in the *M44 Series,* built for optimum performance in the easy-to-take $18-25 price range. Write for a complete catalog:

Shure Brothers Inc.,
222 Hartrey Ave., Evanston, Illinois 60204.

SHURE