CONTROL INFORMATION THEORY:
THE “MISSING LINK” IN THE SCIENCE OF CYBERNETICS

Peter A. Corning, Ph.D.
Institute for the Study of Complex Systems
3501 Beaverton Valley Road
Friday Harbor WA 98250 USA
E-mail: pacorning@complexsystems.org

(Systems Research and Behavioral Science, 24: 297-311, 2007)

ABSTRACT

Norbert Wiener's cybernetic paradigm represents one of the seminal ideas of the 20th century. It has provided a general framework for analyzing communications and control processes in “purposeful” systems, from genomes to empires. Especially notable are the many important applications in control engineering. Nevertheless, its full potential has yet to be realized. For instance, cybernetics is relatively little used as an analytical tool in the social sciences. One reason, it is argued here, is that Wiener's framework lacked a crucial element -- a functional definition of information. The functional (content and meaning) role of information in cybernetic processes cannot be directly measured with Claude Shannon's statistical approach, which Wiener also adopted. Although so-called Shannon information has made many valuable contributions and has many important uses, it is blind to the functional properties of information. Recently, we proposed a radically different approach to information theory. After briefly critiquing the literature in information theory, this new kind of cybernetic information will be described. We call it "control information." Control information is not a thing or a mechanism but an attribute of the relationships between things. It is defined as: the capacity (know how) to control the acquisition, disposition and utilization of matter/energy in “purposive” (cybernetic) processes. We will briefly elucidate the concept, and we will describe a proposed formalization in terms of a common unit of measurement, namely the quantity of available energy that can be controlled by a given unit of information in a given context. However, other metrics are also feasible, from money to allocations of labor (time and energy). Some illustrations will be provided and we will also briefly discuss some of the implications.

Keywords: Information theory, cybernetics, second-order cybernetics, semiotics, communications theory
Introduction

Norbert Wiener's *Cybernetics: Or Control and Communication in the Animal and the Machine* (1948), can truly be called one of the seminal scientific contributions of the 20th Century. Thanks to Wiener's inspired vision, cybernetic control processes are now routinely described and analyzed at virtually every level of living systems, inclusive of social, political and technological systems.¹ Cybernetic processes, including especially feedback processes, are observable in morphogenesis (the translation of genetic instructions into a mature organism), in cellular activity, in plants (see Gilroy and Trewavas 2001), in the workings of multicellular organisms with differentiated organ systems, in the behavioral dynamics of socially-organized species (such as *Apis mellifera*, the true honey bee), in the operation of household thermostats, in robotics, in aerospace engineering, and much more. Cybernetics has given us a framework for understanding one of the most fundamental and distinctive aspects of living systems -- their dynamic “purposiveness”, or goal-directedness. (Biologists refer to this property as “teleonomy” – an evolved internal teleology.) Much productive research has flowed from this paradigm, in fields as disparate as control engineering, molecular biology, plant physiology, neurobiology, psychology and economics.

And yet, cybernetics is still far from realizing its full potential. For instance, it has been relatively little-utilized as a rigorous analytical tool by social scientists, despite the efforts of such theorists as Karl Deutsch (1963), David Easton (1965), William Powers (1973), James Grier Miller (1978) and the present author (1983), among others. One reason for this shortfall, we believe, is that an important element is missing from Wiener's paradigm, and this omission has diminished its utility as an analytical tool.

Actually, Wiener's oversight involved more than an omission. To be precise, Wiener pointed his followers down a false trail, and this has had unfortunate consequences over the years, not only for the development of cybernetics but also for the related fields of semiotics, information theory, and communications theory in sociology. The problem, in essence, has to do with how information is defined and measured. Wiener failed to develop a functional definition of information, which is essential to an understanding of the role and dynamics of "communication and control" in cybernetic systems. Instead, he adopted an engineering approach which was similar to that of his colleague Claude Shannon, the "father" of information theory.²

Information Theory

In his classic 1948 article and his 1949 book with Warren Weaver, Shannon confined his formulation of "communications theory" (as he initially called it) to the problem of measuring uncertainty/predictability in the transmission of "messages" between a sender and a receiver. As Shannon and his co-author wrote: "The fundamental problem of communication is that of reproducing at one point either exactly or approximately the message selected at another point. Frequently the messages have meaning...[But] these semantic aspects of communication are irrelevant to the engineering problem" (p. 3).
Accordingly, in Shannon's usage information refers to the capacity to reduce statistical uncertainty. If one were to utilize the binary bit as a unit of measurement, the degree of informational uncertainty would be a function of the number of bits required to eliminate it. Shannon also adopted the thermodynamics term "entropy" (at the suggestion of mathematician John von Neumann) to characterize the degree of statistical uncertainty in a given communications context before the fact. More formally, Shannon's information can be represented by the equation:

$$I_x = \log_2 \frac{1}{P_x}$$  \hspace{1cm} (1)

where the information content $I$ of an event $x$ in bits is the logarithm to the base 2 of the reciprocal of its probability. Shannon's expression for entropy, then, was:

$$H = K \sum P_i \log_2 P_i$$  \hspace{1cm} (2)

where $K$ refers to Ludwig Boltzmann's famous constant ($1.38 \times 10^{-16} \text{ erg/}^\circ \text{C}$) and $P_i$ refers to the number of equiprobable states.

The justification for calling this quantity entropy came from its similarity to Boltzmann's and Willard Gibbs' statistical equations for thermodynamic entropy. However, this conflation of terms and meanings served only to exacerbate an already serious muddle. The problem first arose when physicists -- beginning with Boltzmann and notably including Erwin Schrödinger in his legendary book, *What is Life* (1945) -- began to blur the distinction between thermodynamic (energetic) entropy (or its converse, which Schrödinger called "negative entropy") and physical (structural) order/disorder. The former usage refers to the availability of energy to do work, whereas the latter usage may be quite unrelated to any work potential. (More on this matter below.) Shannon was careful to differentiate between informational entropy and thermodynamic entropy, but other information theorists have not been so punctilious. Some of Shannon's followers have even suggested that there is an isomorphy, or equivalence, between statistical, energetic and physical order/disorder. However, this is not correct.\(^3\)

One consequence of this conceptual and theoretical conflation was that Shannon's form of information came to be viewed by many theorists as having more potency as an instrumentality for creating order/organization in the natural world than any purely statistical measure can properly support. It imputes causal efficacy to the statistical properties of the messages themselves without regard to their content. Unfortunately, Wiener followed the same approach. In his landmark book, published in the same year that Shannon's classic article appeared, Wiener did discuss the functional aspect of information in various places (e.g., Chapter VII on "Information, Language and Society"), but his formal definition and mathematical treatment involved what he called "a statistical theory of the amount of information" (p. 10). Thus, "the transmission of information is impossible save as a transmission of alternatives....Just as the amount of information in a system is a measure of its degree of organization, so the entropy of a system is a measure of its disorganization" (pp. 10,11). Later on Wiener described enzymes, animals and other cybernetic processes as "metastable Maxwell's Demons, decreasing entropy....Information represents a negative entropy" (p. 58).\(^4\) (In fact, Wiener did not provide
an explicit formalization in his long, discursive, and mathematically challenging chapter on the subject; instead, he focused on how to measure the "amount" of information.)

The suggestion that information is somehow equivalent to negative entropy (i.e., Schrödinger's neologism for available energy, or statistical/structural order, depending upon which version of the term entropy is being referenced) has also encouraged a tendency to reify the concept of information. Biologist Tom Stonier (1990) is perhaps the most emphatic proponent of this view. He argues that information is "real". He writes: "Information exists. It does not need to be perceived to exist. It requires no intelligence to interpret it. It does not have to have meaning to exist. It exists [his emphasis]" (p. 21). It is an embedded property of all physical order, he says. Indeed, physicist Stephen Hawking (1988) asserted that information is swallowed up and destroyed inside black holes, though he has never explained exactly what information is and how to measure it.

What we refer to as "statistical" and "structural" (i.e., order-related) formulations of information theory have made many important contributions to communications technology, computer science and related fields. However, these approaches cannot lead to a unifying theory of information for the simple reason that they are blind to the functional (teleonomic) basis of information in living (and human) systems, as Shannon acknowledged. Indeed, objections to various overclaims for information theory began almost immediately after Shannon published his path-breaking formulation. As early as 1956, Anatol Rapoport published an important rebuttal article entitled "The Promise and Pitfalls of Information Theory." Rapoport noted that "it is misleading in a crucial way to view ,information" as something that can be poured into an empty vessel, like a fluid or even energy." In what might in retrospect be considered a major understatement, Rapoport commented that "the transition from the concept of information in the technical (communication engineering sense) to the semantic (theory of meaning) sense" will be "difficult."

In a similar vein, Heinz von Foerster (1966, 1980, 1995, inter alia) stressed the functional importance of information for living systems. The nonsense sentences "Socrates is identical" or "4+4 = purple" differ profoundly from sentences that have meaning. Likewise, the aggregate number of light photons that might be processed by the retina of a human eye is less relevant from a functional point of view than the analytical and interpretative processes that go on in the brain (the uses that are made of those photons). As von Foerster noted, "Information" is a relational concept that assumes meaning only when related to the cognitive structure of the observer.

MacKay (1961/1968) also pointed out that Shannon's information, and similar formulations, are crucially dependent upon the existence of a sender and a receiver; otherwise, one is only describing a physical process -- a flow of electrons, photons, and the like. For instance, a television screen may display $10^7$ bits of statistical information per second. If one were to transmit an entirely new pattern once each second, the number of bits involved (the “amount” of information) would soon become astronomical, but it would have absolutely no meaning to a viewer. (Similar arguments can be found in Ackoff 1957-58; von Bertalanffy 1968; Bateson 1972, 1979; Cherry 1978; Krippendorff 1979; Maturana and Varela 1980, 1998; Eco 1986; Brier 1992; and Qvortrup 1993, among many others).
Nevertheless, the literature associated with statistical and structural information theory has continued to grow over the years, while the problem of "meaning" and, more broadly, the functional aspect of information has been ignored, skirted, or acknowledged but largely passed over by the workers in information theory, with some recent exceptions. Other theorists have finessed the problem by working within the framework of a particular information coding system, whether it is DNA codons or phonemes. Yet the fundamental theoretical problem remains unresolved. If information is said by some to do work, how can it be differentiated from energy? If information is equated with thermodynamic order, how does it differ from available energy, or physical order (depending upon which version of the term is being referenced)?

But more important, from a functional perspective information is not equivalent either to thermodynamic entropy or "negative entropy" (order). If it were, why confuse matters by using different terms for the same thing? In fact, this conflation of different phenomena involves a fundamental dimensional error. Information (properly defined) has no dimensions, while thermodynamic entropy has the dimensions of energy divided by temperature. It is comparable to equating voltage with length, or mass with velocity. Indeed, physicist Rolf Landauer (1996) has devised a thought experiment which supports his argument that there is no minimum energy expenditure that is necessarily associated with information flows; in theory, the information flow could be made reversible (see also Bennett 1988).

Also, information (unlike energy) can be endlessly reused; there is no law of informational entropy. Nor is information "conserved" in accordance with the first law of thermodynamics; it can be multiplied indefinitely (we will provide an example below). It has also been observed that, in some communications systems, information may flow in the opposite direction from the energy flow (for example, the old-fashioned Morse Code telegraph). Also, highly organized biological systems tend to be relatively more efficient users of energy; they use information to economize on energy consumption and, in so doing, validate the distinctions between information, energy and biological organization.

A further objection is that information by itself cannot do anything; it cannot control a thermodynamic process without the presence of a user that can do purposeful work. In other words, information must be distinguished functionally from the process of exercising control, yet many theoreticians simply take this operation for granted, as F. Clerk Maxwell did with his demon (and as many other physicists have done since). It is this overlooked aspect -- this free ride -- that has allowed physical scientists to theorize about informational processes without acknowledging the necessary role of cybernetic control processes. Indeed, cybernetic processes cannot even be described by the laws of physics (see Corning and Kline 1998a).

Another theoretical problem with traditional information theory concerns the contexts in which information does not have a statistical aspect. This can be illustrated by embellishing an example used by Wicken (1987) to show how Shannon information depends upon the existence of alternatives. Flipping a coin repeatedly is said to produce information -- a unique sequence among many possible alternatives. But if the coin is two-headed, the outcome of each flip is predetermined, and so no statistical information is generated. Now suppose that there are two bettors, one of whom does not know that the coin is two-headed (at least initially).
Consequently, some money might change hands, even though no statistical information is produced. Furthermore, after a few flips of the coin the "sucker" might get suspicious and challenge the process, precisely because of the absence of statistical properties. Clearly, some other kind of information -- what we call "control information" -- was also involved in this situation.

Defining information as a manifestation, or embedded property of physical order (e.g., Tribus, Riedl, Brooks and Wiley, Stonier, Wicken and others) presents similar difficulties. First, there is the problem of defining order in any empirically-consistent, measurable way. We do not gain anything by conflating certain properties of the physical-biological world with a concept that has an inescapably functional connotation for living systems. To the contrary, we obscure the many properties of information which cannot be associated with physical order per se, such as the feedback in cybernetic processes that can even produce disordering effects. (Feedback is highly sensitive to phase relationships in periodic systems; in a poorly "tuned" system, feedback can produce all manner of destructive consequences.) Indeed, to an information theorist, feedback is an incomprehensible concept.

In fact, whole categories of information in living systems are excluded altogether by equating information with order, or with binary bits for that matter. For many organisms, physical phenomena of various kinds (gravity, the earth's magnetic field, thermal or chemical gradients, moisture, even the ambient flow of solar photons) provide vitally important information. Living organisms are constantly sensing, filtering, storing and deleting data on a real-time basis, but only some of it is used. This information is not so much ordered as sensed or detected and then utilized in purposeful ways -- only a portion of which can be said to be order-creating. One example is the role of facial expressions in shaping the interactions among humans (and other animals), as Paul Ekman (1973, 1982) has demonstrated, following Darwin's lead in *The Expression of the Emotions in Man and Animals* (1873/1965). Facial expressions, with or without intent, can convey important information, but only to another animal that can properly interpret their meaning.

But perhaps most important, definitions of information that equate physical/statistical order with functional organization involve a fundamental typological error. Biological organization has properties that are not reducible to physical order. (On this point, see Corning and Kline 1998a, b.) In fact, cybernetic processes have the perverse property of being relational in nature -- they are always dependent upon the relationship between a given system (inclusive of its goals) and its specific environment -- a fact that is frequently stressed by the proponents of second-order cybernetics (see below).

**Control Information**

Accordingly, we have proposed that a categorical distinction should be made between what we have called statistical and structural definitions of information (which have their uses) and control information -- which we have designated "$I_C$", and which we formalize below. We define control information as: **The capacity (know how) to control the acquisition, disposition and utilization of matter/energy in purposive (teleonomic) processes**.
Control information has a number of distinctive properties. First and foremost, it does not have any independent existence. It is not a concrete thing, or a mechanism. It is defined (and specified) by the relationship between a particular cybernetic system (a user) and his/her/its environment(s) – external and internal. In this paradigm, the environment contains latent or potential control information (which we designate “I”), but this potential does not differ in any way from the physical properties of the environment. This is a crucial point; there are no discrete embedded properties out there. Moreover, this potential is only actualized when a purposeful (cybernetic) system makes use of it. In other words, the very existence and functional effects produced by control information are always context-dependent and user-specific. A few examples may help to clarify this seemingly paradoxical, even counter-intuitive notion:

First, imagine a traffic intersection with a stoplight that has just turned red. The information conveyed by the photons of light that are emitted by the stoplight and the behavioral consequences that ensue will depend completely upon the circumstances. A motorist who does not see the light may drive right through it. Another motorist, in a hurry late at night, might observe the light and then deliberately decide to ignore it. A third will obey the law and stop. However, to the inhabitant of a remote, hunter-gatherer society -- say a Yanomamö tribesman -- the red stoplight may represent only a puzzling apparition, while it may only be a bright colored light to an infant. Thus, the user and the informational source together determine the informational value and the degree of behavioral control that results.

In the second example, imagine that a large boulder straddles a hiking trail in a mountainous area. The physical properties of the boulder are invariant, but the information "extracted" by four different hikers, and the functional consequences, may vary considerably. One hiker may see the boulder merely as an obstacle and will take action to walk around it. A second one, very tired, may see it as a place to sit down and rest. A third hiker may recognize it as the landmark for a diverging trail that he/she was instructed to take. Now imagine a fourth hiker who is a gold prospector. Observing a small vein of gold, he/she proceeds to demolish the boulder to remove the gold and, in the process, destroys forever the boulder's informational potential. Again, the informational process involves an interaction -- a specific system-environment relationship.

A final example involves the properties of language. Linguists have long insisted that the functional properties of language (or meaning) cannot be reduced to an invariant, quantitative unit, like a binary bit. Thus, the letters in "RAT" "TAR" "ART" and "TRA" have energetic and statistical properties that are equivalent. Yet the meaning (if any) depends upon the configuration -- the gestalt. Moreover, for a small child or an adult who does not know English, none of the words have any meaning at all. In fact, written language involves an essentially arbitrary relationship between configurations of two-dimensional physical patterns and the associations that are produced, if any, in the specific reader's mind. This explains why the same configuration of letters can have very different meanings in different languages. An example is the word "gift". In English it means a present; in German it means poison.

The key point here is that control information causes purposeful work to be done in or by cybernetic systems. If energy, in accordance with the classical definition, is "the capacity to do work," control information is "the capacity to control the capacity to do work." Virtually everything in the universe might, potentially, have informational value (i.e., be used by
cybernetic systems for some purpose), but control information is not located in the physical objects alone. Again, it is not a discrete embedded property. It is defined by the precise relationship between a given object and a given observer/user. Indeed, biological systems vary tremendously in their ability even to detect different aspects of the external world. Thus, the pheromone "signals" that control the behavior of army ants will go unnoticed and ignored by humans. Elephants can detect and respond to very low sound frequencies and dogs can detect very high frequencies that humans cannot even hear. And hawks have some eight times the number of photoreceptors per millimeter of retina as do humans; there is a definite physical basis for the old expression about being "hawk-eyed."

As the foregoing indicates, control information has a number of distinctive properties. First, control information is always relational and context-dependent and has no independent material existence; it cannot be identified or measured independently of a specific cybernetic process. However, it can be measured (see below). Moreover, there may or may not be a sender, or a formal communications channel, or a message for that matter, but there must always be a user -- a living system or a human-designed system. For instance, if you disassemble an automobile into its 15,000 or so component parts, it will no longer be able to utilize cybernetic control instructions from a driver.

Second, control information does not exist until it is actually used. An unread book, an unread genome, or an undetected animal pheromone represent only latent or potential control information (I_p). Accordingly, the various mechanisms which exist in nature and human societies for coding, storing and transmitting potential control information are reducible to their underlying physical processes; their informational properties arise only from the variety of ways in which these physical media may actually be utilized for informational purposes. Moreover, potential control information is equally prevalent in the state properties of physical objects -- temperature, mass, velocity, viscosity, etc. There is no fundamental physical distinction between the two types of latent information; there is only a functional distinction. To be sure, one can always make estimates or predictions about it, but control information cannot actually be measured except in vivo and in situ.

This distinction is important to bear in mind. Potential control information is very often embodied in various specialized information-storage and transmission media. But its seductive concreteness may bear little relationship to its utility -- its functional potency as an influence in a given cybernetic process. The various kinds of information vehicles have only the potential to exercise cybernetic control, and the vehicle must not be confused with the driver. In fact, much time and energy in the real world are devoted to establishing and manipulating relationships between “I_p” and “I_c”.

Accordingly, control information has no fixed structure or value. It is not equivalent to any specific quantity of energy, or order, or entropy, or the like. To illustrate, a single binary bit may (in theory) control an energy flow as small as a single electron or as vast as the signal for a nuclear war; its power can vary tremendously, depending upon the context. (Another way of stating it is that all bits are not created equal.) Control information is analogous to money, whose value is not intrinsic but can be defined only in terms of specific, real-world transactions.
Very often control information has synergistic properties; it emerges from an "ensemble" of informational "components" or "fragments" that may be combined in many different ways. Language provides an obvious example. A change in the arrangement of an identical set of letters converts the declaration "I shall go" into the question "shall I go?" Similar informational synergies are commonplace also with physical phenomena. Thus, the sight of a swarm of bees coming at you conveys an aggregate informational effect that is lacking if only a single bee is doing so.

By the same token, much of the information used by (and within) organisms involves processes that might be characterized as inferential -- that is, they derive from the weight of the evidence rather than from a deterministic message. To illustrate: you may hear a fire alarm; you smell smoke; you see people running out of your building; you assess the context and your experiential “data base” and may infer that there is a fire and that it would be advisable to vacate the building. In a similar vein, it could be said that the testimony presented at a trial consists of informational components, but only the verdict represents control information (i.e., produces definitive action).

Lies, myths, misinformation or disinformation of various kinds may also serve as control information insofar as they affect a user's behavior. It is not the veracity which counts in the control information paradigm but the functional effects that are produced. (Recall the two-headed coin example above.) There is, in fact, a large literature in biology on the evolution and use of deception as a strategy for achieving various functional outcomes.

**Formalizing Control Information**

The term "control information" may be novel, but the concept itself is not idiosyncratic or alien. Many other theorists over the years have articulated similar ideas. To cite a few examples: Raymond (1950) pointed out that information controls the expenditure of energy. Rapoport (1956) characterized information as a means for resisting the Second Law and reducing entropy. MacKay (1961/1968) noted that information "does logical work" -- it has "an organizing function" (well, some of the time at least). Biologist Paul Weiss (1971) insisted that information and biological functions are inseparable. Wicken (1987) differentiated between statistical information and what he called "functional information," which he associated with the creation of biological "structures". Similarly, Küppers (1990), following Manfred Eigen, took the argument to the level of nucleic acids and the very origins of life and spoke of the functional role of template-based information in creating living structures.

The problem, of course, is how to convert this perspective into an analytical framework. Specifically, the question is, how can you measure something that does not exist as a concrete physical entity? Our view, in essence, is that it can be measured in relation to what it does -- in relation to its "power" to control and utilize available energy and matter in or by a purposeful system. One can measure its qualitative effects, or its “meaning”, in terms of the results that are produced -- the cybernetic work that is accomplished. (As an aside, this formulation does not exhaust the meaning of “meaning”; it is confined here to cybernetic control functions.) Potentially, there are many different ways of measuring these results. However, we have chosen to confine our measuring-rod (initially) to the "thermoeconomic" realm -- that is, the capacity to
control purposeful work. Accordingly, our basic formalization utilizes available energy. Our
definition is as follows:

$$I_{cf} = \ln A_u - \ln A_i$$

(3)

where $A$ = available energy as defined by Keenan (1941, 1951), or the energy available to do
work net of the entropy of a system and its surroundings, namely,

$$A = E + P_o V - T_o S_C$$

(4)

where $E$ is the total stored energy, $V$ is the volume, $S_C$ is the (Clausius) entropy of the system, $P_o$
is the pressure and $T_o$ the absolute temperature of the surroundings. Accordingly, in our
formalization, $A_u =$ the total quantity of available energy potentially accessible for cybernetic
control in a given situation by a given cybernetic system, $A_i =$ the total available energy cost
associated with bringing the available energy under control and exercising control over its use,
inclusive of the cost of reducing/eliminating Shannon entropy ($S_S$) or the cost of Shannon
information ($I_S$), and $f$ represents a multiplier for the quantity of a given type of informational
unit that may be present in a given context. Use of the $\ln$ form allows one to handle a large
range of numbers while expressing both the magnitude and efficacy (or power) of a given unit or
ensemble of information. Also, if we take the exponential we get the amplification ratio, a
measure of the relative efficiency of a given informational unit/ensemble. Thus,

$$\exp I_{cf} = \left[\frac{A_u}{A_i}\right]$$

(5)

This formalization, it should be noted, deals only in the currency of energy. Yet cybernetic
processes utilize many different kinds of currencies -- from electron flows to biochemical
interactions, animal and human behavior, manufacturing processes, even monetary transactions.
We believe that the utility of our formalization can be broadened by making appropriate
conversions from these units into energetic equivalencies -- a well-established technique in
energetic analyses dating back to various efforts to develop energy theories of economic value in
the 1930s. Therefore, we propose the use of energetic units (initially at least) as a common
currency for measuring control information. A similar approach can be found in the efforts of
Howard Odum (1988) to develop an energetic measuring-rod for the cost of various kinds of
embodied information in human societies. Odum used specifically an energy-scaling factor
(solar emjoules per joule) of energy inputs, which he calls "energy." However, we use the more
conventional available energy measure, and we focus instead on the benefits (or outputs) that are
produced.

Some Illustrations

We can illustrate this formulation by revisiting the examples provided above. In the red
stoplight example, the signal produces a clearly observable change in the behavior of any
motorist who responds by stopping, and this can readily be converted to a quantity of purposeful
work output. (A proper accounting should also include the work performed by the automobile.)
But what about the motorist who "runs" the stoplight? Here the analysis becomes more subtle
and difficult. The potential information very likely would result in a change in the driver's
degree of alertness, heart rate, blood pressure, etc., and may also result in a slowing down, speeding up, or both, of the automobile. The energetic consequences would be much smaller, but they would still be significant; the information would exercise some influence over the behavior of the driver (and the car). Conversely, in accordance with our definition, no control information would exist for the motorist who did not see the light, or for the Yanomamo tribesman, or the infant, and there would be no measurable energetic consequences.

Similar energetic analyses could be done for the hiker example. In each of the four hypothetical cases described above, the boulder generated different quantities of control information by virtue of its influence on the behavior of each hiker. Likewise, in the language example, it is axiomatic that words have the power to influence human behavior. A time-honored example is the proscription against shouting "fire" in a crowded theater. This venerable legal dictum illustrates both the potential power and the context-dependent nature of control information. Indeed, advertisers and their agencies spend untold billions of dollars/pounds each year trying to find just the right words, and images.

Let us also consider a comparative cost-benefit example -- operating an automobile versus pedaling a bicycle. The costs in monetary terms for operating a given automobile in a given setting are already quite well known and could be converted to energetic equivalents. However, we must be careful to separate the costs associated with actually performing the work from the control costs for the process. From this perspective, the control information costs \( A_i \) turn out to be relatively low compared with the work that an automobile can perform \( A_u \). To simplify the analysis, the control information cost \( A_i \) could be equated with the labor (time/energy) consumed by the controller -- the driver. So, the quantity (power) of the control information associated with driving a car could be calculated in terms of the available energy consumed by the car in doing work, minus the labor cost for the operator \( A_u - A_i \). Now compare this with pedaling a bicycle. The control costs \( A_i \) are approximately the same, while the available energy that can be controlled \( A_u \) is reduced to the muscle work performed by the rider/controller in propelling the bike. Obviously, driving a car greatly amplifies the power of a given quantity of neuronal activity (control information). This illustrates again the context-specific nature of control information; similar quantities of neuronal activity may control very different quantities of cybernetic work.

The economic aspect of our approach should also be mentioned. As noted above, our basic equation for control information is designed to measure not the total available energy involved in a particular context but the "profits", net of entropy and the informational costs associated with the exercise of control. This approach, we maintain, brings our equation out of the realm of theory and locates it in the real-world of economic analyses, where the relationship between costs and benefits plays an important, even decisive, role in determining whether or not potential control information becomes actualized. If the efficiency (benefit-cost ratio) is very low, the likelihood that a given form of potential control information may actually be utilized to exercise control will be reduced commensurately. It is likely to remain in the realm of latency. Indeed, our equation (5) above expresses precisely the reason why we will never see a real-world Maxwell's demon, even if it were technically feasible. There is no way that we know of for the demon to achieve an energetic profit. Maxwell's demon has unwittingly identified a law-like principle of control information theory; if the energetic costs of a particular type of control
An obvious illustration of this economic aspect can be seen in the world of commercial advertising, where the objective is to induce a desired behavioral response from the reader, listener or viewer, and where the relative costs and benefits are decisively important in determining which advertisements (potential control information) are utilized and in which contexts. Thus, a one-minute TV commercial for the 2001 Superbowl cost an advertiser $2.3 million. However, the audience included some 130 million domestic viewers alone.

What about the relationship between control information and organization (biological structures)? Many theorists have pointed to the key role of information in building and maintaining biological systems. It is also a truism that much biological information is encoded, stored and transmitted in various ways. Indeed, information is an integral part of all biological processes. To some theorists, therefore, it has seemed logical to seek a concrete informational measuring-rod for biological organization. We believe that no such structural measuring-rod will be found. We believe that is it important to maintain a clear distinction between the properties of the various physical media that may serve informational purposes and their precise functional dynamics. By insisting that structural information, like any other kind, is only latent control information (like an unread book) and of no direct functional significance until it is actually used in some way, we do not then have to explain such paradoxes as the fact a significant portion of the DNA in the genome of any given species may not code for anything -- i.e., may not have any functional value. (The question of why so much so-called "junk" DNA exists is another matter.) In our scheme, potential control information \( I_p \) becomes control information \( I_c \) if and when it is utilized, and its power is a function of its organizing ability -- the organizing work that it can do with the available energy at hand in relation to a given system.

It should also be noted that we have made no provision in our paradigm for developmental or capital costs -- say the energetic investment in designing and building a demon, or an automobile. Aside from the formidable analytical challenges, and the problem of infinite regress (how far back do you go with the bookkeeping process?), this would be likely to produce some highly skewed results. A more logical approach is to follow the lead of economists and accountants, who utilize various cost-allocation and amortization procedures to apportion the developmental costs for various economic processes. Thus, in our automobile-versus-bicycle example above, the (external) information costs associated with learning to drive, or to ride a bicycle, as well as the cost of providing traffic control systems (stoplights, road-signs, etc.), if allocated over the number of uses and users, might add a very small increment to the total information costs. In any case, much productive research could be done in this vein, including cost-benefit analyses of the cost to acquire and utilize potential control information (knowledge and experience) compared with the benefits.

**Control Information and Feedback**

One other concern relates to how control information is related to feedback. Feedback is a fundamental, and quintessential, aspect of any cybernetic process, needless to say. In fact, it is the most reliable way of documenting the autonomy and “purposiveness” of any dynamic
system. Feedback is routinely observed and measured in practice, and in many different media, yet it remains a problematical concept in formal information theory simply because it is not the statistical properties of information that are important in feedback processes but their functional effects; it’s not feedback if nothing happens. Indeed, feedback is precisely the means by which cybernetic systems, especially living systems, cope with uncertainty and, equally important, any discrepancies between the subjective experience of the system and the objective properties of the world outside. Feedback enables the system to align the subjective and objective realms. Feedback cannot be comprehended and measured in conventional information theory because it is indifferent to the functional consequences of information. In contrast, the control information paradigm is fully able to measure feedback effects independently of any particular information medium, inclusive of the many forms of perceptual data that do not have the properties of being signals or messages.

**Control Information and Semiotics**

To anyone who is familiar with the large and productive field of semiotics (the doctrine of "signs"), the concept of control information may seem to be quite similar. In fact, these two formulations are convergent but have different purposes and foci. As articulated by Thomas A. Sebeok (1986), one of the leading figures in modern semiotics, the doctrine of signs and their meanings traces its roots to ancient Greece (see also Nöth 1990). Indeed, it has been an important theme in the entire tradition of philosophical discourse, from Plato and Aristotle to St. Augustine, Leibniz, Locke, Berkeley and Charles Sanders Peirce.

A key element of the semiotics paradigm in its contemporary form is the requirement for a source, or a producer of messages that are communicated via some channel to a receiver, or a destination. In other words, it envisions a highly structured process rather like the basic paradigm in information theory. However, semiotics embraces all the elements of that process. Equally important, semiotics focuses on the functional properties and meanings of the messages. It is concerned with the content, not the physical or statistical properties per se, as in traditional information theory. Although the semiotics paradigm rather obviously applies to human language and communications systems, it has also been applied by semioticians to communications processes in other living systems. There is even a nascent new inter-discipline called biosemiotics (Hoffmeyer 1997).

The control information paradigm is distinctive in three ways. First, it does not presuppose a discreet source of messages or structured channels. To repeat, in our paradigm every aspect of the phenomenal world represents latent information that may be detected and used in a myriad of different ways in cybernetic processes, and its role may be entirely passive. Indeed, even the absence of something may be of informational significance to a cybernetic system.

Second, our focus is on the user -- a cybernetic system and his/her/its goals and capabilities. Control information is always defined in terms of the functional relationship between the source and the user. But most important, our paradigm provides a way of measuring the meaning of various signs in terms of one or more quantitative metrics. We have proposed a way of measuring the relative power and efficacy of semiotic processes in cybernetic systems. We believe that semiotics as a science can benefit from the use of our control information concept.
Control Information and Second-Order Cybernetics

Control information also has profound implications for the claims of some information scientists, on the one hand, that information is an objective element of the natural world that is equivalent in importance to matter and energy (e.g., Vickery and Vickery 1988; Stonier 1990) and, on the other hand, the assertion of second-order cybernetics that information can only be defined in terms of the cognitive abilities and subjective experience of the user (see especially Maturana and Varela 1980, 1998; Brier 1992; von Foerster 1995; Franscùs 1997; Heylighen and Joslyn 2001). The control information paradigm represents a bridge – or a third way – between objective/universalistic and subjective/particularistic approaches. Control information is, indeed, defined by the user as an autonomous agency, as we have stressed. And yet, its effects are also objectively measurable in such a way that they can be fully comprehended and documented by an outside observer (in theory, at least). Thus, in the control information paradigm we do not need to erect a wall between the user and the observer. Second-order cybernetics can be accommodated within mainstream cybernetics as a unified discipline. To repeat, this is not to say that control information can encompass all of the many forms of subjectivity, perception, and meaning but only that a certain class of external (cybernetic) consequences can be fully understood and measured. Indeed, I believe it puts a scientific floor under second-order cybernetics.

Sociological Theories of Communication

Finally, reference should be made the various sociological theories (and theorists) of communication. In general, these tend to be highly abstract intellectual constructs and models that, by and large, have eluded operationalization as frameworks for concrete empirical research (though they have generated extensive academic debate and discipleship). First, there is the “theory of communicative action” of philosopher/sociologist Jürgen Habermas (1981). The focus of this theory is human minds in interaction, and the consequences that flow from this. Communication is thus of central importance in creating social systems. Information, language, rational discourse and an interpretation of social evolution as a learning process are all featured in this theory. But Habermas also has a normative side; he critiques existing societies and seeks greater enlightenment via improved, more rational communication.

While Habermas”s focus is useful and his aspirations commendable, his work does not rise to the level of science. He does not, to my knowledge, give us an operational definition of information. He does not address the fundamental cybernetic question of how information is related to the general problem of goal-directed cybernetic control and feedback in humankind, either among individuals or in social systems. Nor does he deal with the vast domain of non-verbal communications and control processes, not only in humans but in machines, and robots, and (not least) the rest of the natural world. His theory encompasses at best a very small subset of the vast universe of cybernetic processes, and it does not advance information science per se.

A second major communications theorist is Niklas Luhmann (1984/1995). Luhmann
focuses on communications in “minds” and “social systems”, both of which he reifies and characterizes as “autopoietic” (or self-organizing and self-maintaining), following Maturana and Varela’s (1980) unique vision. In Luhmann’s framework, “meaning” becomes a centrally important concept. It is the selective screen that extracts relevant aspects from the communication flows (the information) that define the system. Furthermore, Luhmann asserts, only human consciousness and human interactions can be provided with meaning (which, of course, makes it a very parochial concept). Though Luhmann uses the term information freely he does not define it except by allusion to Gregory Bateson’s (1972) cryptic characterization as “a difference that makes a difference.” However, he did speak of information as being “an event that selects system states” (quoted in Leydesdorff 2000). The problem with this formulation is how do you operationalize Luhmann’s abstract “systems”, or their “states”. Luhmann’s construct may have some descriptive value for sociologists, but it cannot be specified empirically. Luhmann is also reputed to have shared with his mentor, Talcott Parsons, an interest in cybernetics, but he failed to incorporate this empirically-grounded science into his theory.

There are a number of other variations on this theme in sociology and communications theory. Loet Leydesdorff (2000), for instance, differentiates between “meanings” and “functions” at various levels, and he has developed some formal communications models based on Luhmann’s insistence that functional codes of communication are binary (following the expansive claims of mathematician G. Spencer-Brown). Of course, this limits the definition of communications to a highly specialized, if not unique context. Even the genetic “code” is not binary. Leydesdorff also explores the problems of uncertainty and ambiguity in communications processes, and, following Giddens (1979, 1984), he recognizes the potential for interactions and co-evolution in multi-level systems and between a system and various outside observers. Conceptually, this seems quite useful, but it does not move us any closer to a definition of information that can be related to the problems of cybernetic control in concrete, real-world systems.

Conclusion

We believe that the concept of control information provides a new tool for analyzing cybernetic processes, including feedback processes, both in nature and in human systems. It provides both a qualitative and quantitative measure of information in terms of the functional consequences that are produced by a given informational unit in a given context. Moreover, it has many practical applications; indeed, it is already used implicitly as a measuring-rod in many different fields, from advertising to politics and education. As we noted above, it also lends itself well to economic analyses. In sum, we believe that control information enriches Wiener's original vision by providing a new and more fruitful way of measuring the relationship between communication processes and control functions. We believe that control information provides a missing element in Norbert Wiener's cybernetic paradigm.
Acknowledgments

This paper is the outgrowth of a close collaboration between this author and the late Stephen Jay Kline, Woodard Professor of Science, Technology and Society, and of Mechanical Engineering, Emeritus, at Stanford University. Two jointly-authored papers, "Thermodynamics, Information and Life Revisited," (Part I and Part II) appeared in the journal *Systems Research and Behavioral Science* (Corning and Kline 1998a,b). However, the core concept of control information is one of the present author's contributions to this collaboration. This paper elaborates on the concept and relates it specifically to cybernetics and to Norbert Wiener's theoretical framework. The author also wishes to thank the Collegium Budapest (Institute for Advanced Study) in Hungary for a fellowship that was of great assistance in completing this work, as well as Patrick Tower and Connie Sutton for their diligent and capable research support and Kitty Chiu for her varied contributions to the production of the final result. Two anonymous reviewers for this journal were also most helpful. An earlier version of this paper was the winner of the U.K. Cybernetics Society’s 30th Anniversary Prize Competition in 2000 and was subsequently published in the journal *Kybernetes* (2001).

Footnotes

1. Actually, the use of feedback mechanisms in technological systems dates back to antiquity (see O. Mayr 1970). However, Wiener provided a broader framework for understanding feedback processes in relation to goal-directed behaviors of all kinds. Contemporary theorists often distinguish between evolved, internal purposiveness (teleonomy) and an externally imposed purpose, or teleology.

2. The other leading figure among the pioneers in cybernetics, H. Ross Ashby, was even less helpful. In his much-cited classic, *Design for a Brain* (1952/1960), Ashby barely mentioned communications, and the term "information" was not even referenced in his index. Even the all-important concept of feedback merited only two index references. There are occasional allusions to information, however. Thus, in one place Ashby describes trial-and-error learning as a valuable part of "information gathering" for an animal, which he notes is essential to adaptation (p. 83). However, there is no explicit treatment of information in Ashby's book, much less the problem of measuring it.

3. A simple thought experiment can be used to illustrate. Imagine two alternative paradigms. In one case, there is a delicately-structured, heated crystal inside an isolated system with Gibbsian constraints (no gravity or other extraneous influences). It is in a highly ordered state and also has a certain heat content and available energy. Now imagine a second isolated system containing an identical crystal with the same available energy but in the form of a pile of disordered shards. Is there any difference in the ability of the two crystals to do work? Conversely, consider the case of an elaborate, highly ordered crystal floating in space at near-absolute zero degrees. It would be richly endowed with negative entropy in a structural sense but would possess no available energy since there would be no temperature gradient between the crystal and its surrounding environment.
4. Maxwell's demon refers to a famous 19th century "thought experiment," since recounted in innumerable discussions of thermodynamics. Physicist James Clerk Maxwell proposed a means by which, supposedly, the Second Law might be violated. Maxwell conjured up a fanciful creature that would be stationed at a wall between two enclosed volumes of gases at equal temperatures. (The term "demon" was actually coined by a contemporary colleague, William Thomson.) The demon would then selectively open and close a microscopic trap door in the wall in such a way as to be able to sort out the mixture of fast and slow gas molecules between the two chambers. In this manner, Maxwell suggested, a temperature differential would be created that could be used to do work, thereby reversing the otherwise irreversible thermodynamic entropy. The fundamental problem with this paradigm was that it would be impossible to build and operate a real-world equivalent of a demon at a profit.

5. Another problem with defining information as equivalent to physical order is that it entails the same kind of semantic pettifoggery that is associated with the concept of negative entropy. In fact, the term negative entropy is really a convoluted synonym for thermodynamic order. It means, literally, an absence of an absence of order. If information is equivalent to order/negentropy, then it is inextricably tied to available energy, or physical order of all kinds, or both, depending upon how the term negentropy is defined. If so, information is highly inflammable; it is consumed every time irreversible work is performed and every time entropy increases, for whatever reason.

6. Stuart Umpleby (2004), in a recent essay, also suggests the use of Gregory Bateson’s definition of information: “the difference that makes a difference.” Umpleby is seeking to give information, in the narrow sense of data or signals, a non-statistical definition. He wishes to clarify its functional significance and enhance its status as a fundamental property of nature, like matter and energy. He suggests that making a difference is a more “elementary” concept than information. However, there are also some problems with Bateson’s definition. One problem is semantic. Information and “differences” have different meanings in everyday language that are well understood, albeit imprecise. It would be a very hard sell to get people to talk about differences that make a difference when they mean information. But more to the point, the term lacks functional specificity. Matter and energy too can be characterized as differences that make a difference, at a fundamental level. In other words, how does one differentiate between informational differences and all of the other kinds of differences in the natural world. What is it exactly that makes information a different difference? One must conclude that Bateson’s definition does little to clarify the status of information as an existential phenomenon.

References


