

Chapter 2 - What is Information? Why is It Relativistic? and What is Its Relationship to Materiality, Meaning and Organization?

We have represented a discrete information source as a Markoff process. Can we define a quantity, which will measure, in some sense, how much information is 'produced' by such a process, or better, at what rate information is produced? – Shannon (1948)

To live effectively is to live with adequate information – Wiener (1950)

Information is a distinction that makes a difference – MacKay (1969)

Information is a difference that makes a difference – Bateson (1973)

Information... arises... as natural selection assembling the very constraints on the release of energy that then constitutes work and the propagation of organization – Kauffman, Logan, Este, Goebel, Hobill & Shmulevich (2007)

Juxtaposing these five modern definitions of information we begin to see the issues that we face in developing an understanding of information. We see that Shannon's definition of information is a purely mathematical notion totally devoid of meaning or context. With Wiener's cybernetic concept of information we see that information takes on a functional role. The MacKay and Bateson definitions address their principal critique of Shannon information, namely that it does not deal with meaning. Finally the last definition of Kauffman et al. (2007) deals with the materiality of biotic information and its relationship to organization, two features missing from Shannon information. These five quotes embrace the issues of information's materiality, meaning and relationship to organization that we will address in this chapter. We will show that the definition of information developed by Shannon that is commonly used in information theory only begins to scratch the surface of this complex phenomenon.

We begin by considering the historic development of the concept of information from its earliest usages in English to its formal formulation by the reputed father of information theory, Claude Shannon (1948), and the use of information in cybernetics by Norbert Wiener (1948 & 1950). We then study the critiques of Shannon's formulation of information by Doug MacKay (1969) and Gregory Bateson (1973) who insist that information is more than a number of bits but that it also entails meaning. We also examine the relationship of information, energy and entropy arguing, as have many physicists before us, that information and entropy are opposites and not parallel as suggested by Shannon.

We then examine the way that information, which had always been associated with the human mind was introduced into biology by those considering the way genetic information is transmitted from one generation to another and by those considering the transmission of signals in living organisms.

We then review the work of Kauffman et al. (2007) that demonstrated that Shannon information cannot describe the information contained of a living organism. This work led to the introduction of the notion of the relativity of information and the realization that what we consider to be information depends on the context of where and how it is being generated and used.

Next we will examine the relationship of information to meaning and materiality within information theory, cybernetics and systems biology. And finally we examine the link between information and organization showing that in biotic systems that information and organization are intimately linked. We will also identify a similar link between information and organization in the various aspects of human culture including language, technology, science, economics and governance.

We conclude the chapter by discussing to what extent living organisms including humans are just flesh and to what extent they are information.

Origins of the Concept of Information

We begin our historic survey of the development of the concept of information with its etymology. The English word information according to the Oxford English Dictionary (OED) first appears in the written record in 1386 by Chaucer: “Whanne Melibee hadde herd the grete skiles and resons of Dame Prudence, and hire wise informacions and techynges.” The word is derived from Latin through French by combining the word inform meaning giving a form to the mind with the ending “ation” denoting a noun of action. This earliest definition refers to an item of training or molding of the mind. The next notion of information, namely the communication of knowledge appears shortly thereafter in 1450. “Lydg. & Burgh *Secrees* 1695 Ferthere to geve the Enformacioun, Of mustard whyte the seed is profitable.”

The notion of information as a something capable of storage in or the transfer or communication to something inanimate and the notion of information as a mathematically defined quantity do not arise until the 20th century.

The OED cites two sources, which abstracted the concept of information as something that could be conveyed or stored to an inanimate object:

1937 *Discovery* Nov. 329/1 The whole difficulty resides in the amount of definition in the [television] picture, or, as the engineers put it, the amount of information to be transmitted in a given time.

1944 *Jrnl. Sci. Instrum.* XXI. 133/2 Information is conveyed to the machine by means of punched cards.

The OED cites the 1925 article of R. A. Fisher as the first instance of the mathematization of information:

What we have spoken of as the intrinsic accuracy of an error curve may equally be conceived as the amount of information in a single observation belonging to such a distribution... If p is the probability of an observation falling into any one class, the amount of information in the sample is $S\{(\partial m/\partial \theta)^2/m\}$ where $m = np$, is the expectation in any one class [and θ is the parameter] (Fisher 1925).

Another OED entry citing the early work of mathematizing information is that of R. V. L. Hartley (1928, p. 540) "What we have done then is to take as our practical measure of information the logarithm of the number of possible symbol sequences." It is interesting to note that the work of both Fisher and Hartley foreshadow Shannon's concept of information, which is nothing more than the probability of a particular string of symbols independent of their meaning.

Shannon and the Birth of Information Theory

Despite the early work of Fisher and Hartley cited above the beginning of the modern theoretical study of information is attributed to Claude Shannon (1948), who is recognized as the father of information theory. He defined information as a message sent by a sender to a receiver. Shannon worked at Bell Labs and wanted to solve the problem of how to best encode information that a sender wished to transmit to a receiver. Shannon gave information a numerical or mathematical value based on probability defined in terms of the concept of information entropy more commonly known as Shannon entropy. Information is defined as the measure of the decrease of uncertainty for a receiver. The amount of Shannon information is inversely proportional to the probability of the occurrence of that information, where the information is coded in some symbolic form as a string of 0s and 1s or in terms of some alphanumeric code. Shannon (1948, pp. 392-94) defined his measures as follows:

We have represented a discrete information source as a Markoff process. Can we define a quantity, which will measure, in some sense, how much information is 'produced' by such a process, or better, at what rate information is produced? Suppose we have a set of possible events whose probabilities of occurrence are p_1, p_2, \dots, p_n . These probabilities are known but that is all we know concerning which event will occur. Can we find a measure of how much 'choice' is involved in the selection of the event or of how uncertain we are of the outcome? If there is such a measure, say $H(p_1, p_2, \dots, p_n)$... we shall call $H = - p_i \log p_i$ the entropy of the set of probabilities p_1, \dots, p_n ... The quantity H has a number of interesting properties, which further substantiate it as a reasonable measure of choice or information.

A story is told that Shannon did not know what to call his measure and von Neumann advised him to call it entropy because nobody knows what it means and that it would therefore give Shannon an advantage in any debate (Campbell 1982, p. 32). This choice was criticized by Wicken (1987, p. 183) who argued that in science a term should have only one meaning. Schneider and Sagan (2005) referring to the use of the term entropy in both thermodynamics and information theory also suggests Shannon's use of the term is

confusing when they wrote: “There is no simple correspondence between the two theories.”

The Relationship of Information and Entropy

Understanding the efficiency of a steam engine through thermodynamics led Clausius to the idea of entropy as a measure of the mechanical unavailability of energy or the amount of heat energy that cannot be transformed into usable work. He referred to it in German as *Verwandlungsinhalt*, which may be translated roughly into English as “transformation content”. Clausius then coined the term entropy deriving the root *trope* from the Greek word *τροπή* meaning transformation. He added the prefix *en* because of the close association he felt that existed between energy and entropy. One can therefore roughly translate entropy from its etymology as energy transformation. Clausius felt the need to define entropy because the energy of the universe is conserved but its entropy is constantly increasing.

The relationship between entropy and probability is due to the work of Boltzmann from his consideration of statistical mechanics, which is an alternative way of looking at thermodynamics. He showed that the entropy of a gas is proportional to the logarithm of W where W is the number of microstates of the gas that yield identical values of the thermodynamic variables of pressure, temperature and volume. The formula he derived, namely, that $S = k \ln W$ where k is the Boltzmann constant is what inspired Shannon to call his expression for the measure of the information content of a message “information entropy” despite the difference in sign and the fact that the proportionality constant or Boltzmann constant has the physical dimensions of energy divided by temperature.

The relationship between entropy and information as developed by physicists arose from a consideration of Maxwell’s demon and is quite opposite to the one proposed by Shannon. Maxwell in 1867 postulated a gedanken experiment in which a demon standing in a doorway between two rooms filled with gas would allow only fast moving molecules to pass from one room to another so as to create a temperature difference in the two rooms from which usable work could be extracted in violation of the second law of thermodynamics. Leo Szilard in 1929 analyzing the problem that Maxwell’s Demon presented showed that to obtain the information he needed the demon caused an increase of entropy elsewhere such that the net entropy did not decrease. He suggested that the demon is only able to temporarily reduce entropy because it possesses information, which is purchased at the cost of an increase in entropy. There is no violation of the Second Law because acquisition of that information causes an increase of entropy greater than the decrease of entropy represented by the information. As a result of Szilard’s analysis one must conclude that entropy and information are opposite. He also pointed out that the net energy gained by the demon was not positive because of the energy cost in obtaining the information by which the demon selected the fast moving molecules and rejecting the slow moving ones. Since the information was purchased at the cost of an increase in entropy the information has an effective net negative entropy. Following Szilard, Gilbert N. Lewis (1930, p. 573) also saw an inverse relationship between information and

entropy. He wrote, "Gain in entropy always means loss of information, and nothing more."

Schrödinger (1944, pp. 71-72) first explicitly introduced the notion of negative entropy:

Every process, event, happening - call it what you will; in a word, everything that is going on in Nature means an increase of the entropy of the part of the world where it is going on. Thus a living organism continually increases its entropy - or, as you may say, produces positive entropy - and thus tends to approach the dangerous state of maximum entropy, which is death. It can only keep aloof from it, i.e. alive, by continually drawing from its environment negative entropy - which is something very positive as we shall immediately see. What an organism feeds upon is negative entropy. Or, to put it less paradoxically, the essential thing in metabolism is that the organism succeeds in freeing itself from all the entropy it cannot help producing while alive (Chapter 6).

Both Wiener (1950) and Brillouin (1951) both adopted Shannon's definition of information and its relation to entropy with the one exception of its sign, likely influenced by the arguments of Szilard (1929) and Schrödinger (1944).

Wiener wrote,

Information is "negative entropy"; it expresses purpose (1948).

Messages are themselves a form of pattern and organization. Indeed, it is possible to treat sets of messages as having entropy like sets of states in the external world. Just as entropy is a measure of disorganization, the information carried by a set of messages is a measure of organization. In fact, it is possible to interpret the information carried by a message as essentially the negative of its entropy, and the negative logarithm of its probability. That is, the more probable the message, the less information it gives (p.39)...This amount of information is a quantity which differs from entropy merely by its algebraic sign and a possible numerical factor (1950, p. 129).

Brillouin (1951) also argued that a living system exports entropy in order to maintain its own entropy at a low level. Brillouin used the term negentropy to describe information rather than negative entropy.

The reason that Wiener and Brillouin consider entropy and information as opposites or regard information as negative entropy follows from the tendency in nature for systems to move into states of greater disorder, i.e. states of increased entropy and hence states for which we have less information. Consider a system, which is in a state for which there is a certain finite number of possible configurations or microstates all of which are equivalent to the same macro state. The tendency of nature according to the second law of thermodynamics is for the number of microstates that are equivalent to the macrostate of the system to increase. Because there are more possible microstates as time increases

and we do not know which particular microstate the system is in, we know less about the system as the number of possible microstates increases. It therefore follows that as the entropy increases the amount of information we have about the system decreases and hence entropy is negative information or vice-versa information is the negative of entropy. In other words the second law of thermodynamics tell us that when system A evolves into system B that system B will have more possible redundant or equivalent micro states than system A and hence we know less about system B than system A since the uncertainty as to which state the system is in has increased.

Wiener and Brillouin relate information to entropy with a negative sign whereas Shannon uses a positive sign. Hayles (1999, p. 102) notes that although this difference is arbitrary it had a significant impact. Observing that Shannon used the positive sign she also noted that “identifying entropy with information can be seen as a crucial crossing point, for this allowed entropy to be reconceptualized as the thermodynamic motor driving systems to self-organization rather than as the heat engines driving the world to universal heat death.” For Wiener, on the other hand she wrote, “life is an island of negentropy amid a sea of disorder (ibid.)”.

Despite the difference in the sign of information entropy assigned by Shannon and Wiener, Shannon was heavily influenced by Wiener’s work as indicated by the way Shannon (1948) credits Wiener for his contribution to his thinking in his acknowledgement: “Credit should also be given to Professor N. Wiener, whose elegant solution of the problems of filtering and prediction of stationary ensembles has considerably influenced the writer’s thinking in this field.” Shannon also acknowledge his debt to Wiener in footnote 4 of Part III:

Communication theory is heavily indebted to Wiener for much of its basic philosophy and theory. His classic NDRC report, *The Interpolation, Extrapolation and Smoothing of Stationary Time Series* (Wiley, 1949), contains the first clear-cut formulation of communication theory as a statistical problem, the study of operations on time series. This work, although chiefly concerned with the linear prediction and filtering problem, is an important collateral reference in connection with the present paper. We may also refer here to Wiener’s *Cybernetics* (Wiley, 1948), dealing with the general problems of communication and control.

MacKay’s Counter Revolution: Where is the Meaning in Shannon Information?

According to Claude Shannon (1948, p. 379) his definition of information is not connected to its meaning. Weaver concurred in his introduction to Shannon’s *A Mathematical Theory of Communication* when he wrote: “Information has ‘nothing to do with meaning’ although it does describe a ‘pattern’.” Shannon also suggested that information in the form of a message often contains meaning but that meaning is not a necessary condition for defining information. So it is possible to have information without meaning, whatever that means.

Not all of the members of the information science community were happy with Shannon's definition of information. Three years after Shannon proposed his definition of information Donald Mackay (1951) at the 8th Macy Conference argued for another approach to understanding the nature of information. The highly influential Macy Conferences on cybernetics, systems theory, information and communications were held from 1946 to 1953 during which Norbert Wiener's newly minted cybernetic theory and Shannon's information theory were discussed and debated with a fascinating interdisciplinary team of scholars which also included Warren McCulloch, Walter Pitts, Gregory Bateson, Margaret Mead, Heinz von Foerster, Kurt Lewin and John von Neumann. MacKay argued that he did not see "too close a connection between the notion of information as we use it in communications engineering and what [we] are doing here... the problem here is not so much finding the best encoding of symbols...but, rather, the determination of the semantic question of what to send and to whom to send it." He suggested that information should be defined as "the change in a receiver's mind-set, and thus with meaning" and not just the sender's signal (Hayles 1999b, p. 74). The notion of information independent of its meaning or context is like looking at a figure isolated from its ground. As the ground changes so too does the meaning of the figure.

Shannon whose position eventually prevailed defined information as the pattern or the signal and not the meaning. The problem with MacKay's definition was that meaning could not be measured or quantified and as a result the Shannon definition won out and changed the development of information science. The advantage that Shannon enjoyed over MacKay by defining information as the signal rather than meaning was his ability to mathematize information and prove general theorems that held independent of the medium that carried the information. The theorizing that Shannon conducted through his combination of electrical engineering and mathematics came to be known as information theory. It is ironic that the OED cites the first use of the term "information theory" as that of MacKay's who used the term in a heading in an article he published in the March 1950 issue of the Philosophical Magazine.

Shannon's motivation for his definition of information was to create a tool to analyze how to increase the ratio of signal to noise within telecommunications. People that shared MacKay's position complained that Shannon's definition of information did not fully describe communication. Shannon did not disagree—he "frequently cautioned that the theory was meant to apply only to certain technical situations, not to communication in general (ibid., p. 74)." He acknowledged that his definition of information was quite independent of meaning, however, he conceded that the information that was transmitted over the telecommunication lines he studied often had meaning as the following quote from his original paper written at the Bell Labs indicates:

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one **selected** from a set of possible messages.

The system must be designed to operate for each possible **selection**, not just the one that will actually be chosen since this is unknown at the time of design. If the number of messages in the set is **finite** then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. (Shannon 1948 – bolding is mine)

I ask the reader to note that Shannon requires the number of possible messages to be finite as this will be a critical concern when we examine biotic information. I admire Shannon's frankness about his definition of information, which he devised to handle the engineering problems he faced. He was quite clear that his definition was not the unique definition of information but merely one definition of information suited for his engineering requirements. In the abstract to his paper, *The Lattice Theory of Information* Shannon (1953) wrote,

The word "information" has been given many different meanings by various writers in the general field of information theory. It is likely that at least a number of these will prove sufficiently useful in certain applications to deserve further study and permanent recognition. It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field. The present note outlines a new approach to information theory, which is aimed specifically at the analysis of certain communication problems in which there exist a number of information sources simultaneously in operation.

What I find extraordinary is that his definition of information limited in scope by his own admission became the standard by which almost all forms of information were gauged. There have been some slight variations of Shannon information like Kolomogorov information used to measure the shortest string of 0s and 1s to achieve a programming result or represent a text on a computer or a Turing machine. But despite these small variations Shannon information has been accepted as the canonical definition of information by all except for a small band of critics.

I have purposely bolded the term selected and selection in the above quote of Shannon to highlight the fact that Shannon's definition of information had to do with selection from a pre-determined set of data that did not necessarily have any meaning. MacKay used this selective element of Shannon information to distinguish it from his own definition of information, which, unlike Shannon, incorporates meaning explicitly. He also defended his definition from the attack that it was subjective.

Mackay's first move was to rescue information that affected the receiver's mindset from the 'subjective' label. He proposed that both Shannon and Bavelas were concerned with what he called 'selective information,' that is information calculated by considering the selection of message elements from a set. But selective information alone is not enough; also required is another kind of information that he called 'structural.' Structural information indicates how

selective information is to be understood; it is a message about how to interpret a message—that is, it is a metacommunication (Hayles 1999a, pp. 54-55).

Structural information must involve semantics and meaning if it is to succeed in its role of interpreting selective or Shannon information. Structural information is concerned with the effect and impact of the information on the mind of the receiver and hence is reflexive. Structural information has a relationship to pragmatics as well as semantics where pragmatics tries to bridge the explanatory gap between the literal meaning of a sentence and the meaning that the speaker or writer intended. Shannon information has no particular relation to either semantics or pragmatics. It is only concerned with the text of a message and not the intentions of the sender or the possible interpretations of the receiver.

Part of the resistance to MacKay information was that its definition involved subjectivity, which orthodox scientists could not abide in their theories. Rather than deal with the fact that the exchange of information among humans involves a certain amount of subjectivity proponents of Shannon information theory chose to ignore this essential element of information and communications. Taken to its logical conclusion this attitude would limit science to study those areas that do not involve subjectivity, which would forever condemn linguistics and the other social sciences to non-scientific analysis. Rule out subjectivity in science or social studies and social science becomes a contradiction in terms.

This raises the question of whether subjectivity can be studied scientifically. I would suggest that an approach that parallels quantum physics is needed. Just as the measurement of sub-atomic particles changes their behavior and requires a quantum mechanic representation that includes the Heisenberg Uncertainty principle, something similar is required for a science of the subjective - something I would call quantum rhetoric. What is the study of communications and media ecology after all but the study of how one set of subjective humans communicates with another set of subjective humans. Shannon successfully exorcised the subjectivity from communications, which was fine for his engineering objectives. I totally respect Shannon because he always warned that his definition was not intended to be a theory of communications. My problem is with those that misuse his work and over extend it.

Information: The Difference that Makes a Difference

Although Shannon's notion of information divorced from meaning became the central theme of information theory MacKay's counter-revolution was not without some effect and resulted in a slight shift in the way information was regarded. No doubt the reader is familiar with Gregory Bateson (1973) famous definition of information as **“the difference that makes a difference.”** Buried in this one-liner is the notion that it is the meaning of the information that makes the difference. Although Bateson gets credit for this idea it is likely that he was influenced by MacKay (1969) who in his book *Information, Mechanism and Meaning* published four years before the appearance of Bateson's one-liner wrote: **“information is a distinction that makes a difference.”**

Bateson, MacKay and Shannon were all participants in the Macy conferences so Bateson was quite familiar with MacKay's ideas. The use of the term "distinction" in MacKay's one-liner is more closely tied to the idea of "meaning" than the term "difference". It is ironic that MacKay who pointed out the shortcomings of Shannon information, was the first to use the term "information theory" and was the first to point out that the importance of information is its meaning and the fact that it makes a difference. MacKay is certainly a scholar who made a difference and he deserves more credit and attribution than he usually receives.

Another one line definition of information that incorporates the notion of its meaning is this one by Ed Fredkin which I would put in a league with Mackay and Bateson's one-liners. "The meaning of information is given by the processes that interpret it." This is a very interesting definition because it explicitly incorporates the notion that information depends on context.

If information is the distinction (McKay) or the difference (Bateson) that makes a difference then if there is no distinction or no difference then there can be no information. This would mean chaos or random numbers contain no information because there is no difference or distinction in one part of the stream of numbers as opposed to another part of the stream because of a lack of organization. This is opposite to the conclusion of Shannon who claims that a stream of random numbers contains the maximum information possible. While it is true each element is different from the next and is a complete surprise it is also true that the overall pattern of chaos and randomness is the same and hence there is no distinction nor is there any difference in the stream of random numbers. A gas, which remains uniformly at the same temperature, pressure and volume, is constantly changing but one cannot make a distinction between the gas at one moment and the gas at another moment. There is no difference in the way the gas behaves at these different moments. The only information one can discern about the gas is its volume, pressure and temperature, which is unchanging. No work can be done by this gas. If, however, in this volume of gas there is a temperature differential then work can be extracted from the gas and there is information in the gas by virtue of the way in which the temperature differential is organized. This raises the question of whether or not organization is information, a point we will return to later in this chapter once we have dealt with the nature of information in biotic systems.

Information in Biotic Systems

We have seen that as early as 1925 the notion of information as an abstraction was first introduced by Fisher (1925) and formalized by Shannon (1948). It was not long after this development that biologists also began to talk about information. The OED cites the first uses of the term in biology in 1953:

1953 J. C. ECCLES *Neurophysiol. Basis Mind* i. 1 We may say that all 'information' is conveyed in the nervous system in the form of coded arrangements of nerve impulses.

1953 WATSON & CRICK in *Nature* 30 May 965/2 In a long molecule many different permutations are possible, and it therefore seems likely that the precise sequence of the bases is the code which carries the genetical information.

The use of information in this context was not the mathematization of information as was done by Fisher and Shannon but rather information was thought of qualitatively as something capable of being transferred or communicated to or through a living organism or stored in a living organism in the form of nucleic acid.

Life as Propagating Organization

Stuart Kauffman (2000) defined an autonomous agent (or living organism) acting on its own behalf and propagating its organization as an autocatalytic system carrying out at least one thermodynamic work cycle. The relationship of the information found in living organisms to the kind of information treated in Shannon information theory was not clear even though a lot of attention has been given in recent times to the notion of information in biotic systems by those pursuing systems biology and bioinformatics. It was to examine this relationship that a group of us undertook a study to understand the nature and flow of information in biotic systems. This led to an article entitled Propagating Organization: An Enquiry (POE) authored by Kauffman, Logan, Este, Goebel, Hobill and Shmulevich (2007) in which we demonstrated that Shannon information could not be used to describe information contained in a biotic system. We also showed that information is not an invariant independent of its frame of reference.

In POE we argued that Shannon's (1948) classical definition of information as the measure of the decrease of uncertainty was not valid for a biotic system that propagates its organization. The core argument of POE was that Shannon information "does not apply to the evolution of the biosphere" because Darwinian preadaptations cannot be predicted and as a consequence "the ensemble of possibilities and their entropy cannot be calculated (Kauffman et al. 2007)." Therefore a definition of information as reducing uncertainty does not make sense since no matter how much one learns from the information in a biotic system the uncertainty remains infinite because the number of possibilities of what can evolve is infinitely non-denumerable. I remind the reader that in making his definition that Shannon specified that the number of possible messages needed to be finite.

Instead of Shannon information we defined a new form of information, which we called instructional or biotic information,

not with Shannon, but with constraints or boundary conditions. The amount of information will be related to the diversity of constraints and the diversity of processes that they can partially cause to occur. By taking this step, we embed the concept of information in the ongoing processes of the biosphere, for they are causally relevant to that which happens in the unfolding of the biosphere. We therefore conclude that constraints are information and ... information is constraints... We use the term "**instructional information**" because of the

instructional function this information performs and we sometimes call it “**biotic information**” because this is the domain it acts in, as opposed to human telecommunication or computer information systems where Shannon information operates (ibid.).

A living organism is an open system, which von Bertalanffy (1968, p. 149) “defined as a system in exchange of matter with its environment, presenting import and export, building-up and breaking-down of its material components.” Instructional or biotic information may therefore be defined as the organization of that exchange of energy and matter. The fact that a biotic system is an open system can be used to argue against the association of instructional or biotic information with cybernetics because cybernetics focuses strictly on the flow of information and does not deal with the flow of energy and matter.

In POE we argued that constraints acting as instructional information are essential to the operation of a cell and the propagation of its organization.

The working of a cell is, in part, a complex web of constraints, or boundary conditions, which partially direct or cause the events which happen. Importantly, the propagating organization in the cell is the structural union of constraints as instructional information, the constrained release of energy as work, the use of work in the construction of copies of information, the use of work in the construction of other structures, and the construction of further constraints as instructional information. This instructional information further constrains the further release of energy *in diverse specific ways*, all of which propagates organization of process that completes a closure of tasks whereby the cell reproduces (ibid.).

In POE we associated biotic or instructional information with the organization that a biotic agent is able to propagate. This contradicts Shannon’s definition of information and the notion that a random set or soup of organic chemicals has more Shannon information than a structured and organized set of organic chemicals found in a living organism.

The biotic agent has more meaning than the soup, however. The living organism with more structure and more organization has less Shannon information. This is counterintuitive to a biologist’s understanding of a living organism. We therefore conclude that the use of Shannon information to describe a biotic system would not be valid. Shannon information for a biotic system is simply a category error. A living organism has meaning because it is an autonomous agent acting on its own behalf. A random soup of organic chemicals has no meaning and no organization (ibid.).

The key point that was uncovered in the POE analysis was the fact that Shannon information could be defined independent of meaning whereas biotic or instructional

was intimately connected to the meaning of the organism's information, namely the propagation of its organization. Thus we see organization within a system as a form of information, which is a much more dynamic notion of information than Shannon information which is merely a string of symbols or bits.

According to Shannon's definition of information a set of random numbers transmitted over a telephone line would have more information than the set of even numbers transmitted over the same line. Once 2, 4, 6, 8, 10, 12 was received the receiver, who is assumed to be a clever person, would be able to correctly guess that the rest of the numbers to follow the sequence would be the set of even numbers. The random numbers have no organization but the even numbers are organized so the mystery of the relevance of Shannon information deepens as one must counter-intuitively conclude that information and organization can be at cross-purposes in Shannon's scheme of things.

This argument completely contradicts the notion of information of a system biologist who would argue that a biological organism contains information. It is by virtue of this propagating organization that an organism is able to grow and replicate, as pointed out by Kauffman (2000) in *Investigations*. From the contradiction between Shannon and biotic information we already have a hint that there is possibly more than one type of information and that information is not an invariant like the speed of light in relativity theory, which is independent of its frame of reference. We also see that perhaps Shannon's definition of information might have limitations and might not represent an universal notion of information. After all Shannon formulated his concept of information as information entropy to solve a specific problem namely increasing the efficiency or the signal to noise ratio in the transmission of signals over telecommunication lines.

The Relativity of Information

Robert M. Losee (1997) in an article entitled A Discipline Independent Definition of Information published in the Journal of the American Society for Information Science defines information as follows:

Information may be defined as the characteristics of the output of a process, these being informative about the process and the input. This discipline independent definition may be applied to all domains, from physics to epistemology.

The term information, as the above definition seems to suggest, is generally regarded as some uniform quantity or quality, which is the same for all the domains and phenomena it describes. In other words information is an invariant like the speed of light, the same in all frames of reference. The origin of the term information or the actual meaning of the concept is all taken for granted. If ever pressed on the issue most contemporary IT experts or philosophers will revert back to Shannon's definition of information. Some might also come up with Bateson definition that information is the difference that makes a difference. Most would not be aware that the Shannon and Bateson definitions of information are at odds with each other. Shannon information does not make a difference because it has nothing to do with meaning; it is merely a string of symbols or bits. On the

other hand, Bateson information, which as we discovered should more accurately be called MacKay information, is all about meaning. And thus we arrive at our second surprise, namely the relativity of information. Information is not an invariant like the speed of light, but depends on the frame of reference or context in which it is used.

We discovered in our review of POE that Shannon information and biotic or instructional information are quite different. Information is not an absolute but depends on the context in which it is being used. So Shannon information is a perfectly useful tool for telecommunication channel engineering. Kolomogorov (Shiryayev 1993) information, defined as the minimum computational resources needed to describe a program or a text and is related to Shannon information, is useful for the study of information compression with respect to Turing machines. Biotic or instructional information, on the other hand, is not equivalent to Shannon or Kolomogorov information and as has been shown in POE is the only way to describe the interaction and evolution of biological systems and the propagation of their organization.

Information is a tool and as such it comes in different forms just as screwdrivers are not all the same. They come in different forms, slot, square, and Philips---depending in what screw environment they are to operate. The same may be said of information. MacKay identified two main categories of information: selective information not necessarily linked to meaning and structural information specifically linked to meaning. Shannon information was formulated to deal with the signal to noise ratio in telecommunications and Kolomogorov information was intended to measure information content as the complexity of an algorithm on a Turing Machine. Shannon and Kologomorov information are what MacKay termed selective information. Biotic or instructional information, on the other hand, is a form of structural information. The information of DNA is not fixed like Shannon selective information but depends on context like MacKay structural information so that identical genotypes can give rise to different phenotypes depending on the environment or context.

Although we introduced the notion of the relativity of information in POE we were unaware at the time of the formulation of a similar idea long ago by Nicholas Tzannes (1968). He “wanted to define information so that its meaning varied with context... [and] pointed out that whereas Shannon and Wiener define information in terms of what it is, MacKay defines it in terms of what it does (Hayles 1999a, p. 56).” Both Shannon and Wiener’s form of information is a noun or a thing and MacKay’s form of information is a verb or process. We associate instructional or biotic information with MacKay as it is a process and not with Shannon because DNA, RNA and proteins are not informational “things” as such but rather they catalyze “processes” and actions that give rise to the propagation of organization and hence the transmission of information—information with meaning at that. Put simply instructional information is structural information as the root of the word instructional reveals.

Another distinction between Shannon information and biotic or instructional information as defined in POE is that with Shannon there is no explanation as to where information comes from and how it came into being. Information in Shannon’s theory arrives *deus ex*

machina, whereas biotic information as described in POE arises from the constraints that allow a living organism to harness free energy and turn it into work so that it can carry out its metabolism and replicate its organization. Kauffman (2000) has described how this organization emerges through autocatalysis as an emergent phenomenon with properties that cannot be derived from, predicted from or reduced to the properties of the biomolecules of which the living organism is composed and hence provides an explanation of where biotic information comes from.

Information and Its Relationship to Materiality and Meaning

O, that this too too solid flesh would melt – Shakespeare’s Hamlet (Act 1, Scene 2)

Where is the wisdom we have lost in knowledge?

Where is the knowledge we have lost in information? – TS Eliot

Where is the meaning we have lost in information? – RKL

To drive home the point that information is not an invariant but rather a quantity that is relative to the environment in which it operates we will now examine the relationship of information to materiality and meaning drawing on the work and insights of Katherine Hayles (1999a & b). She points out that although information is used to describe material things and furthermore is instantiated in material things information is not itself material. “Shannon’s theory defines information as a probability function with no dimension, no materiality, and no necessary connection with meaning. It is a pattern not a presence (Hayles 1999a, p. 18).”

The lack of a necessary connection to meaning of Shannon information is what distinguishes it from biotic information. Biotic information obviously has meaning, which is the propagation of the organism’s organization. Information is an abstraction we use to describe the behavior of material things and often is sometimes thought of as something that controls, in the cybernetic sense, material things.

Hayles (1999a) traces the origin of information theory to cyberneticians like Wiener, von Forester and von Bertalanffy and telecommunication engineers like Shannon and Weaver. She points out that they regarded information as having a more primal existence than matter. Referring to the information theory they developed she wrote: “It (information theory) constructs information as the site of mastery and control over the material world.”

She further claims, and I concur, that Shannon and cybernetic information is treated as separate from the material base in which it is instantiated. Wiener (1961, p. 132), for example, wrote in his book *Cybernetics, or Control and Communication in the Animal and the Machine* that “information is information, not matter or energy” (1961, 132). The question that arises is whether or not there is something intrinsic about information or is it merely a description of or a metaphor for the complex patterns of behavior of material things. Does information really control matter or is information purely a mental construct

based on the notion of human communication through symbolic language, which in turn is a product of conceptual thought as described in Logan (2006 & 2007)?

While it is true that the notion of information as used by the cyberneticians like Wiener, von Forester and von Bertalanffy and that used by Shannon and Weaver influenced each other and in the minds of many were the same they are actually quite different from each other. The notion of information as the master or controller of the material world is the view of the cyberneticians beginning with Wiener (1950): "To live effectively is to live with adequate information. Thus, communication and control belong to the essence of man's inner life, even as they belong to his life in society."

For communication engineers information is just a string of symbols that must be accurately transmitted from one location, the sender, to another location, the receiver. Their only concern is the accuracy of the transmission with the relationship to the meaning of the information being meaningless to their concerns. If we consider the relationship of information and meaning for the moment then there is a sense in which the cybernetician's notion of information has meaning as a controller of the material realm whereas Shannon information has no relationship as such to meaning. In fact one can question if Shannon's used the correct term "information" when he described $H = - \sum p_i \log p_i$ as the measure of "information". The quantity H he defined is clearly a useful measure for engineering in that it is related to the probability of the transmission of a signal — a signal that might or might not contain meaning. It is my contention that a signal without meaning is not truly information. I agree with MacKay and Bateson that to qualify as information the signal must make a difference, as is also the case with the way Wiener defines information in the context of cybernetics. Sveiby reports that Shannon himself had some second thoughts about the accuracy of his use of the term 'information':

Shannon is said to have been unhappy with the word "information" in his theory. He was advised to use the word "entropy" instead, but entropy was a concept too difficult to communicate so he remained with the word. Since his theory concerns only transmission of signals, Langefors (1968) suggested that a better term for Shannon's information theory would therefore perhaps be "signal transmission theory" (from the following Web site visited on 9/9/07:

<http://sveiby.com/portals/0/articles/Information.html#Cybernetics>).

I find myself in agreement with Langefors that what Shannon is analyzing in his so-called information theory is the transmission of signals or data. It is consistent with some of my earlier work in the field of knowledge management and collaboration theory, in part inspired by the work of Karl Erik Sveiby, where Louis Stokes and I developed the following definitions of data, information, knowledge and wisdom:

- Data are the pure and simple facts without any particular structure or organization, the basic atoms of information,
- Information is structured data, which adds meaning to the data and gives it context and significance,

- Knowledge is the ability to use information strategically to achieve one's objectives, and
- Wisdom is the capacity to choose objectives consistent with one's values and within a larger social context (Logan and Stokes 2004, pp. 38-39).

I also found the following description of the relationship of data and information that I accessed on Wikipedia on September 12, 2007 particularly illuminating:

Even though information and data are often used interchangeably, they are actually very different. Data is a set of unrelated information, and as such is of no use until it is properly evaluated. Upon evaluation, once there is some significant relation between data, and they show some relevance, then they are converted into information. Now this same data can be used for different purposes. Thus, till the data convey some information, they are not useful.

I would interpret the signals transmitted between Shannon's sender and receiver as data. Consistent with MacKay and Bateson's position information makes a difference when it is contextualized and significant. Knowledge and wisdom represent higher order applications of information beyond the scope of this chapter. We will return, however, to the topic of knowledge and science in Chapter 8. The contextualization of data so that it has meaning and significance and hence operates as information is an emergent phenomenon. The communication of information cannot be explained solely in terms of the components of the Shannon system consisting of the sender, the receiver and the signal or message. It is a much more complex process than the simplified system that Shannon considered for the purposes of mathematizing and engineering the transmission of signals. First of all it entails the knowledge of the sender and the receiver, the intentions or objectives of the sender and the receiver in participating in the process and finally the effects of the channel of communication itself independent of its content as in McLuhan's observation that "the medium is the message". The knowledge and intention of the sender and the receiver as well as the effects of the channel all affect the meaning of the message that is transmitted by the signal in addition to its content.

The Meaning of Information in Biotic Systems

Biotic or instructional information, defined in POE as the constraints that allow an autonomous agent, i.e. a living organism, to convert free energy into work so that the living organism is able to propagate its organization through growth and replication, is intimately connected with meaning. "For Shannon the semantics or meaning of the message does not matter, whereas in biology the opposite is true. Biotic agents have purpose and hence meaning (Kauffman et al. 2007)." One can therefore argue that since the meaning of instructional information is propagating organization that we finally understand the meaning of life – the "meaning of life" is propagating organization. This remark is not meant to trivialize the great philosophical quest for the meaning of life from a human perspective but there is a sense in which the meaning of life including human

life is indeed the propagation of organization. The purpose of life is the creation or propagation of more life.

In addition to the fact that Shannon information does not necessarily entail meaning whereas biotic or instructional information always entails meaning there is one other essential difference between the two. Shannon information is defined independent of the medium of its instantiation whereas biotic information is very much tied to its material instantiation in the nucleic acids and proteins of which it is composed. The independence of Shannon and cybernetic information from the medium of its instantiation is what gives rise to the notion of strong artificial intelligence and claims like those of Moravic, Minsky and to a certain extent Wiener that human intelligence and the human mind can somehow be transferred to a silicon-based computer and does not require the wet computer of the human brain. Shannon and cybernetic information can be transferred from one material environment to another, from one computer to another or in the case of Shannon information from one telephone to another or from a computer to a hard copy of ink on paper. This is not the case with living organisms in the biosphere where information is stored in DNA, RNA and proteins.

Shannon information whether on paper, a computer, a DVD or a telecommunication device, because it is symbolic, can slide from one medium or technology to another and not really change, McLuhan's "the medium is the message" aside. This is not true of living things. Identical genotypes can produce very different phenotypes depending on the physical and chemical environment in which they operate. Consider the fact that identical twins are not "identical". The reason identical twins are not "identical" is that the environment in which the biochemical interactions between biomolecules takes place alters the outcome.

The Materiality of Information in Biotic Systems

Information is information, not matter or energy. No materialism which does not admit this can survive at the present day. - Norbert Wiener (1948)

Shannon's theory defines information as a probability function with no dimension, no materiality, and no necessary connection with meaning. It is a pattern not a presence. - Hayles (1999a, p. 18)

Shannon information cannot be, nor was it meant to be, naively applied to complete living organisms, because the information in a biotic system like DNA is more than a pattern—it is also a presence. A receptor for food or toxins is not just a pattern—it is also a presence. A biological system is both an information pattern and a material object or more accurately information patterns instantiated in a material presence. Schrödinger (1944, 21) long ago before the discovery of DNA described this dual aspect of chromosomal material metaphorically. "The chromosome structures are at the same time instrumental in bringing about the development they foreshadow. They are law-code and executive power - or, to use another simile, they are architect's plan and builder's craft -

in one.” It is the dynamic of the interaction between the patterns of information and the material composition of the biotic agents that determines their behavior.

As previously discussed, the issue hinges on the degree to which one can regard a biotic agent as a fully physical computational system. It is clear that a biotic system cannot be described only by Shannon information for which the information is abstracted from its material instantiation and is independent of the medium. The same argument can be made for the inappropriateness of Kolomogorov information for biotic systems. Kolomogorov information, which is defined with respect to Turing machines, is another case where the information pattern is separated from its material instantiation. Biology is about material things not just mathematical patterns. As Kubie once warned at one of the Macy conferences, “we are constantly in danger of oversimplifying the problem so as to scale it down for mathematical treatment (Hayles 1999, p. 70).” As noted above the physical environment changes the meaning of the information embedded in the DNA of the genome.

Another way to distinguish the difference between biotic or instructional information and either Shannon or Kolomogorov information is that the latter are symbolic which is not the case for biotic or instructional information. The information coded in the chemical alphabet of biomolecules that make up living organisms acts through the chemical interactions of those biomolecules. “DNA is a molecule interacting with other molecules through a complex set of mechanisms. DNA is not just some text to be interpreted, and to regard it as such is an inaccurate simplification (Sarkar 1996, p. 860).” It is not the symbolic nature of DNA that gives rise to messenger RNA and it is not the symbolic nature of RNA that gives rise to proteins but rather the chemical properties of DNA that produce or catalyze the production of RNA and the chemical properties of RNA that produce or catalyze proteins and the chemical properties of proteins that carry out the protein’s various functions such as:

1. serving as enzymes to catalyze biochemical reactions vital to metabolism,
2. providing structural or mechanical functions, such as building the cell’s cytoskeleton,
3. playing a role in cell signaling, immune responses, cell adhesion and the cell cycle.

DNA, RNA and proteins are both the medium and the content, the message and the messenger. Not so for Shannon and Kolomogorov information where one can distinguish between the medium and the message, the message and the messenger. The message is the information, which operates independent of the medium in which it is instantiated, McLuhan aside. For biotic information, on the other hand, the medium and the message are the same – they cannot be separated. For biotic information the medium is the message in the McLuhan sense and it is also the content. For human symbolic information described by Shannon information, the information or content and the medium are quite separate. For biotic systems not only is the medium the message in the McLuhan sense that a medium has an effect independent of its content but the medium is also the content because it is the chemical properties of the medium that affects the organism. In fact the medium is the message because it is literally the content and the

content of the message is unique to that medium and is instantiated in it and it cannot be transferred to another medium. To repeat it is not possible to transfer the content or the message of the medium to another medium. There is an isomorphism between the medium and its content. The medium is the content and hence also the message. The medium is both the message and the content for a biotic system because information in a biological system is not symbolic but rather chemical. It is for this reason that the notion of transferring the contents of the human brain to a computer is pure nonsense.

To conclude we have argued that information is not an invariant independent of the frame of reference in which it operates. In the biotic frame of reference information is always associated with meaning, which is not necessarily the case with Shannon or Kolomogorov information. In the biotic frame information cannot be separated from the medium of its instantiation as is the case in the Shannon and Kolomogorov reference frames. In other words the information in DNA, RNA and proteins are embodied. They differ from human symbolic information, which can be disembodied and moved from one medium to another. Each generation makes a god of their latest technological or scientific achievement or breakthrough. For the Hebrews it was the written word and the law “written with the finger of God”. For the Greeks it was their deductive logic and rational thought disembodied from practical experience and empirical evidence of the physical world. For the Enlightenment it was Newtonian mechanics and God, the clock maker, where things were explained in terms of mechanical models. In the Information Age the god is disembodied information, information without context where everything is explained in terms of the transfer of information, and some times it is information without meaning.

Organization as Information

What is the relationship of organization and information? What we discovered in POE was that the autocatalysis of biomolecules led to the organization of a biological living organism whose organization of constraints allowed it to convert free energy into work that sustained growth and permitted replication. We identified the constraints as instructional or biotic information, which loops back into the organization of the organism. This model of information holds for biotic systems where autocatalysis is the organization and the components are the individual biomolecules.

The argument seems circular only because a living organism represents a self-organizing system. This is still another way that biotic information differs from Shannon information which is defined independent of meaning or organization. In fact organized information has less Shannon information because it does not reduce as much uncertainty as disorganized information. It is also the case as we mention above that this model provides a mechanism for the creation of information which in not the case with the Shannon model of information.

I believe that Hayles (1999a, p. 11) has come to a similar conclusion regarding the relationship of information and organization when she wrote about the paradigm of autopoiesis or self-organization:

Information does not exist in this paradigm or that it has sunk so deeply into the system as to become indistinguishable from the organizational properties defining the system as such.

It is the latter half of her statement that is congruent with our notion that the set of constraints or organization that give rise to an autonomous self-organizing system is a form of information.

Wiener like Shannon related information to entropy but, unlike Shannon, Wiener (1948, p. 18) saw a connection between organization and information, “The notion of the amount of information attaches itself very naturally to a classical notion in statistical mechanics: that of entropy. Just as the amount of information in a system is a measure of its degree of organisation, so the entropy of a system is a measure of its degree of disorganisation.”

We (Kauffman et al. 2007) make a similar claim in POE when we assert that the constraints that allow the propagation of organization in a living organism represents the information content of that organism. In other words the propagating organization of a living organism is its information content. Our position in a certain sense recapitulates similar sentiments expressed by Norbert Wiener (1954, p. 96) when he wrote “We are not stuff that abides but patterns that perpetuate themselves.”

However where I differ from Wiener is that while we are patterns that abide I also believe that we are patterns that are uniquely instantiated in flesh. I therefore believe that human intelligence cannot be transferred from a human brain onto a silicon-based computer as is claimed by some advocates of strong AI. The point that I would make is that the pattern cannot be separated from the medium in which it is instantiated as was argued above. The medium of flesh and its organization are what is critical. It is the pattern instantiated in the flesh and not just the pattern by itself that makes life. The information in a biological system is not symbolic but rather chemical. As we have already asserted the medium of the flesh is both the message and the content of a biotic system. We will return to the question of the relationship of information and organization once we have introduced some ideas about the origin of language and culture and their relationship to information. But before addressing this issue let us continue our discussion of whether or not living organisms are information or flesh.

Who Are We? What Are We, Information or Flesh?

Information in the form of words or language is symbolic. The word cat is a symbol that represents a class of living breathing creatures made of flesh. An actual cat is not a symbol of something else but an organization of organic chemicals that can propagate its organization through its metabolism and its ability to replicate.

The organic chemicals of which we are composed are continually replaced so that after seven years there is a completely new set of molecules. So we are not flesh or a particular

set of molecules but the organization of the molecules of which we are composed or more accurately we are a process and not a thing that can be duplicated.

One cannot make a replica of a person. Even twins that originated from the same fertilized egg are never exactly the same. But a text can be replicated or duplicated exactly. A text can also be transmitted and reformatted from one medium to another, for example from a computer file to a text printed on paper or from a live performance to a podcast.

I believe that the proponents of strong artificial intelligence (AI) and strong artificial life (AL) make the mistake of considering intelligence or life as merely reified information. They do not take into account that it is the interaction or organization of flesh-based matter that makes intelligence and life. The pattern of that interaction or organization that we identify as information cannot be abstracted away from the physical medium in which it is instantiated and remain unchanged or, even more importantly, continue as the process that gave rise to that intelligence or life in the first place.

A feature of both intelligence and life is that it is autonomous. A living organism is an autonomous agent that has the capacity to exploit free energy from its environment and use that energy in the form of work to carry out its metabolism, to replicate and to make use of its intelligence. The proponents of strong AI and AL overlook this important factor when they claim that intelligence and life is nothing more than information or a pattern that is independent of its physical instantiation. At best artificial life forms may be regarded as obligate symbionts with humans but not as independent living organisms as they are not autonomous.

We attempted to answer the question who are we, what are we: information or flesh? Our conclusion is that we are both but in order to more fully address the question we need to deal with the issue of language. In the next chapter we will examine the role of language in defining who and what we are. We will discover that language is a critical element of determining who and what we are because of the way that language extends the brain into the human mind and creates the conditions for the emergence of culture another unique feature of humankind that defines us.

Acknowledgement: This chapter draws heavily on two sources other than my earlier work, namely the paper *Propagating Organization: An Enquiry* (Kauffman et al. 2007) and the book *How We Became Posthuman* (Hayles 1999a). In a certain sense this paper is a remix of these two sources with help from the cited references.