

February 1996

# Cognition, Redundancy, and Learning in Organizations

Richard N. Langlois  
*University of Connecticut*

Pierre Garrouste  
*Universite Lumiere/Lyon*

Follow this and additional works at: [http://digitalcommons.uconn.edu/econ\\_wpapers](http://digitalcommons.uconn.edu/econ_wpapers)

---

## Recommended Citation

Langlois, Richard N. and Garrouste, Pierre, "Cognition, Redundancy, and Learning in Organizations" (1996). *Economics Working Papers*. 199610.  
[http://digitalcommons.uconn.edu/econ\\_wpapers/199610](http://digitalcommons.uconn.edu/econ_wpapers/199610)



University of  
Connecticut

*Department of Economics Working Paper Series*

**Cognition, Redundancy, and Learning in Organizations**

Richard Langlois  
University of Connecticut

Pierre Garrouste  
Universite Lumiere/Lyon

Working Paper 1996-10

February 1996

---

341 Mansfield Road, Unit 1063  
Storrs, CT 06269-1063  
Phone: (860) 486-3022  
Fax: (860) 486-4463  
<http://www.econ.uconn.edu/>

## **Introduction.**

There is a well-developed tradition in the literatures of business and economics that sees organizations as information-processing systems. It is perhaps quite natural, then, that students of organization have long been fascinated with cybernetics and the mathematical theory of information as a way to understand organizational learning and structure. Apart from providing concrete content to the notion of “information,” this theory also has the benefit of linking closely the ideas of information content and of “organization” itself through the formalism of what statistical thermodynamics calls entropy. However, the severe limitations of this approach for matters economic are equally well understood: the measure of information *per se* does not take into account the *economic* value of information (Langlois 1983; Arrow 1974). Furthermore, cybernetic models of organizations also typically take as fixed the information structure of a system. To the extent that organizations learn, they do so the way economic actors do in the neoclassical economics of information, that is, by receiving signals that update probability distributions over known and given contingencies. There is little in the literature about how organizations create categories of understanding in the first place, about how information builds a knowledge structure.

Nonetheless, there may well remain a place for the cybernetic theory of information within economics. Information theory is in many ways about structure and complexity. And present-day economics is concerning itself increasingly with such questions, in such areas as evolutionary economics (Nelson and Winter 1982) and the economics of institutions (Langlois 1986). This essay attempts to revisit the cybernetic theory of information and some related ideas in order to develop a notion of self-organization relevant to understanding organizational learning.

## **Knowledge and Structure.**

It is conventional to see the distinction between knowledge and information as a distinction between a stock and a flow. This is certainly unobjectionable, and maybe even useful, as long as we don't take the metaphor too seriously. Knowledge is not a stock in the same sense that oil in a tank is a stock, something modified in a purely quantitatively way by the inflow or outflow of info-fluid<sup>1</sup> (Langlois 1983, pp. 586-7). Knowledge is about structure. As the late Kenneth Boulding put it,

we cannot regard knowledge as simply the accumulation of information in a stockpile, even though all messages that are received by the brain may leave some sort of deposit there. Knowledge must itself be regarded as a structure, a very complex and frequently quite loose pattern, ... with its parts connected in various ways by ties of varying degrees of strength. Messages are continually shot into this structure; some of them pass right through its interstices ... without effecting any perceptible change in it. Sometimes messages “stick” to the structure and become part of it. ... Occasionally, however, a message which is inconsistent with the basic pattern of the mental structure, but which is of a nature that it cannot be disbelieved hits the structure, which is then forced to undergo a complete reorganization. (Boulding 1955, pp. 103-104, quoted in Machlup 1983, p. 643n).

In order for a message to “stick” to the structure — or, more importantly, for the message to modify the structure in a useful way — that message must be meaningful to the receiving system. The message must somehow “fit.” As Kenneth Arrow (1974, chapter 2) notes, individuals and organizations have information structures that are in the nature of message decoders. To understand messages in Chinese, for example, one needs to have learned Chinese. Choosing an information structure, like learning a language, thus involves an investment that is typically costly in both money and time. To put it another way, information structures develop or evolve slowly and cannot be recreated or “reengineered” quickly or costlessly.

---

<sup>1</sup>. On this point cf. also Hayek (1952, p. 105) on the “storage” theory of memory.

The association of knowledge with structure is intuitively appealing, if still rather vague. What makes a structure “knowledge”? At some level, a structure constitutes knowledge if that structure is ordered in a way that produces results.<sup>2</sup> Think of genetics. We can say that DNA is a knowledge structure because it is an orderly arrangement that “knows how” to do something, namely how to generate an organism. That organism in turn is also an ordered structure that does something, namely survive the evolutionary process. Thus knowledge is a pudding whose proof is in the eating, even if modern philosophers of science don't agree about how much the eating proves.

Donald MacKay thinks of a system's structure as defining “conditional states of readiness” on which a signal operates. It is the overall configuration that determines the meaning — and the meaningfulness — of a message. “It isn't until we consider the range of other states of readiness, that *might have been considered but weren't*, that the notion of meaning comes into its own. A change in meaning implies a different selection from the range of states of readiness. A meaningless message is one that makes no selection from the range. An ambiguous message is one that could make more than one selection” (MacKay, 1969, p. 24, emphasis original). MacKay offers the metaphor of a railroad switching yard in which the configuration of tracks and switches stands ready to direct the trains passing through it. By sending the right electronic signal (or, in older yards, by inserting the correct key in a switch-box) one can rearrange the configuration of tracks. The meaningfulness of a message thus depends on its form — on the shape of the key.

---

2. Indeed, cybernetic information theory has tended to think of knowledge and information in behaviorist terms. A stimulus is information to the extent that it elicits some response from the structure it stimulates. As MacKay (1969) notes, however, such a Skinnerian conception is as naive in this as it is in other matters. A signal may change a knowledge structure in a way that is meaningful — i.e., it may modify the future or potential behavior of the system — without that change resulting in any directly observable response. In fact, as Machlup points out, “[a]ny kind of experience — accidental impressions, observations, and even ‘inner experience’ not induced by stimuli received from the environment — may initiate cognitive processes leading to changes in a person's knowledge. Thus, *new knowledge can be acquired without new information being received*” (Machlup 1983, p. 644, emphasis original).

And that meaning consists in the change the message effects in the arrangement of the yard, the selection it makes from the set of all possible configurations.

But where does the structure of knowledge — the railroad switching yard — come from? How does it form, and how is it modified by experience? In a work only now being appreciated by cognitive psychologists (Weimer 1982; Edelman 1987), F. A. Hayek (1952) put forward a rich and sophisticated theory of mind as structure. In this theory, “that which we call knowledge is primarily a system of rules of action assisted and modified by rules indicating equivalences or differences of various combinations of stimuli” (Hayek 1978, p. 41).

To survive, an organism must respond appropriately to the stimuli — the information — provided by its environment. Both phylogenetically and ontogenetically, organisms, in Hayek's view, use the pattern of stimuli to which they are subjected to create complex interpretive or classificatory systems that help them take appropriate action in response to future stimuli. The neural system of the brain (and, more generally, the nervous system as a whole) creates, with experience, a semipermanent structure or “map” that guides action — not only in response to new stimuli but also through processes of internal reclassification and recombination that lead to innovation.

In short, learning — whether in the organism or in the organization — is a matter of self-organization, that is, of the creation of structure. How is such self-organization possible? Can we begin to understand the process of self-organization in a way that is relevant to understanding the cognitive processes of learning in the brain or in the economy? In what follows, we pick up some neglected strands in the cybernetic theory of information and use them to put forward a picture of, and potentially a framework for analyzing, self-organization.

### **The theory of information, redundancy, and learning.**

The theory of information defines the quantity of information contained in a message by Shannon's well-known formula. Let  $x \in \{x_i\}$  be some elementary event, e.g., the occurrence of one particular symbol out of a set of collectively exhaustive and mutually exclusive possible symbols. Then the quantity of information — or, alternatively, the entropy<sup>3</sup> — of the system  $x$  is

$$H(x) = -\sum_i p(i) \log_2 p(i),$$

where  $p(i)$  is the probability that the symbol  $x_i$  will occur. It is well known that the entropy of a system is greatest when all its microstates are independent and equiprobable. Saviotti (1991, p. 180) provides a relevant example. In an organization in which all members have the same skills, functions, and power (that is, an organization with no division of labor), one could randomly reassign workers to tasks without degrading performance. We could say that the microstates of such an organization are equiprobable. Once labor is divided and skill differences emerge, however, such a random rearrangement would degrade performance. All microstates are not equiprobable. The organization with a division of labor is thus a lower-entropy or more “ordered” system than the organization with undivided labor and undifferentiated skills.

Consider now a system that can transmit two different kinds of messages,  $a \in \{x_i\}$  and  $b \in \{x_j\}$ . These might, for example, be two different characteristics of a product, such as its technical characteristics and its demand characteristics (Saviotti 1991, p. 199). Each of the two characteristics can take on a range of possible states. We can calculate the information content of a message about characteristic  $b$  conditional on having received a message about characteristic  $a$  as:

$$H(b|a) = -\sum_{i,j} p(i) p(j|i) \log_2 p(j|i),$$

---

<sup>3</sup> Entropy is, however, in different units, as the thermodynamics variant of this formula is multiplied by Boltzmann's constant.

where  $p(i)$  and  $p(j)$  are the probabilities that  $a = x_i$  and  $b = x_j$ , and the  $p(j/i)$  are the conditional probabilities that  $b = x_j$  given  $a = x_i$ . (If we change the interpretation so that  $p(j/i)$  is the probability that the system is in state  $j$  at time  $t+1$  given that it was in state  $i$  at time  $t$ , then this formula is identical to that of a first-order Markov chain.) In effect,  $H(b/a)$  measures the residual uncertainty about characteristic  $b$  left after having already received information about  $a$ . The expression  $H(b/a)$  is usually referred to as *equivocation* or *ambiguity*. For example, if information about one of the product's characteristics given the other reduces the overall uncertainty about the product, then technical characteristics and service characteristics are not fully independent as far as the information emitted by the product is concerned. At the same time, however, if knowledge of one characteristic does not convey complete information about the other, there is ambiguity in this sense.

The problem of organization becomes more interesting when we consider the perturbations to the system from the outside environment. In communications theory, such perturbations always enter as the phenomenon of noise. If we consider  $a$  and  $b$  in this context to be signals input into and output from a communications channel, respectively, then  $H(a) \neq H(b)$  implies the presence of noise. In communications theory, then, ambiguity arising from noise is always destructive, in the sense that the quantity of information transmitted along the channel is diminished. To correct for this loss of information, one can make the message redundant, that is, one can repeat certain symbols or otherwise use extra symbols for error-checking (e.g., by including the sum of a series of digits as a supplemental digit). This will have the effect of reducing the average information content of the message transmitted, since with redundancy some symbols have only an error-checking function and not a message-conveying function, which means that it takes more symbols to convey the same message. Another way to reduce the deleterious effects of noise is to make the system more reliable by increasing the number of parallel channels through which the same signal is transmitted (von

Neumann 1956). That is, one can introduce redundancy into the system itself rather than into the message transmitted.

The redundancy of a system is defined as

$$R = 1 - \frac{H_R}{H_{\max}}, \quad (1)$$

where  $H_{\max}$  is the quantity of information of the message with no redundancy (with no extra symbols or redundant channels) and  $H_R$  is the quantity of information with redundancy. Moreover, since redundancy essentially means that the symbols transmitted are not independent of one another (that is, the point of redundant information is to convey information about the original message), we can write the information content of the message with redundancy in a way analogous to our equation for ambiguity. If we assume the relationship among messages to be that of a first-order Markov chain, we have

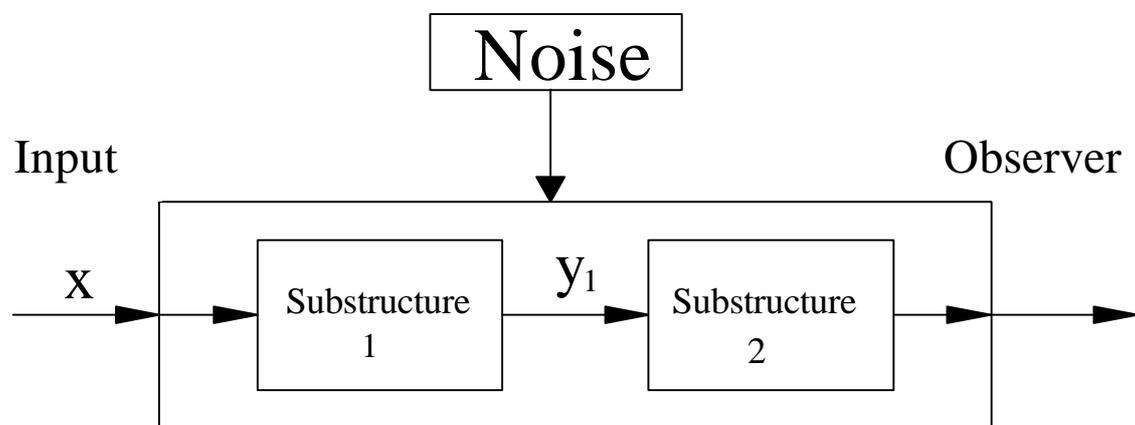
$$H_R = -\sum_{i,j} p(i)p(j|i)\log_2 p(j|i),$$

and

$$H_{\max} = -\sum_i p(i)\log_2 p(i).$$

Atlan (1972) uses the ideas of ambiguity and redundancy to generate a theory of the development and decay of complexity, which he views as a theory of self-organization. In this theory, the presence of noise plays a role that is not solely destructive but can in fact be creative, in that, by increasing ambiguity, noise can increase the information content of a system in a way that is equivalent to increasing variety and complexity. This approach rests on a crucial distinction, which we hinted at above, between internal communication among the

substructures of the system and external communication with an observer. Instead of two different product characteristics, consider now two subsystems of a larger system — substructures of an organization, for example — each composed of elements  $y$ . See Figure 1. The substructures communicate with one another internally, and the system communicates with an outside observer. Noise from outside the system impinges on both transmissions, the external signal  $x$  processed through the system and the internal signal  $y_1$  transmitted from substructure 1 to substructure 2.



**Figure 1**

If there were no noise, that is, if all signals were transmitted without ambiguity, then the information content of the internal system would be  $H = H(y_1) = H(y_2)$ . In effect, the information in substructure 2 would be a perfect replica of the information in substructure 1. In such a case, Atlan would say that the two substructures exert complete *constraints* on one another. If, by contrast,  $H = H(y_1) + H(y_2)$ , the two substructures would be completely independent, and information about substructure 2 would convey no information about, and would therefore not constrain, substructure 1.

If we think in terms of the linear system of figure 1, however, these two extremes are similar, in that they reflect structures that do not have as high an

information content as possible and thus, to Atlan, are not as highly complex as possible. (We return below to the interpretation of organization as complexity.) Atlan (1972, p. 258, translation ours) puts it this way. “The important point is that these two limiting cases, total absence of constraint and total constraint between the substructures, both correspond to the absence of organization in the system: in the first case, all we have is a juxtaposition of completely independent structures one on the other; and, in the second, all we have is the same structure replicated  $N$  times.” Consider the books in a library. All books in actual libraries contain references to one another (footnotes, allusions, etc.): these connections are constraints in Atlan’s sense. If all books were completely independent, in the sense that no book in the library ever referred to any other in any way, we would consider the society that produced the library not to have had a culture, as that term is normally understood, and we would consider the library to be uncomplex, or at any rate unorganized, in an important sense. At the same time, however, if, at the other extreme, all books were perfect replicas of one another, we would also call the library a non-complex structure. The implication: organizational complexity requires ambiguity.

In communications theory, as we saw, ambiguity can reflect only a deterioration in the quality of the signal. And, in Figure 1, this is also the case in the transmission of the external signal  $x$  through the system. To the extent that  $y_1$  does not perfectly replicate  $x$ , information has been lost. But within the system, ambiguity can mean an increase in information content. For if substructure 2 is somewhat independent of substructure 1, that is, it is not an exact copy, then it must contain some information not contained in substructure 1. Atlan calls the first kind of ambiguity, that involving signal loss, *destructive ambiguity*; and he calls the second kind, that arising from the increased independence of the second substructure, *autonomy ambiguity*. Formally, the information content of the system is

$$H = H(y_1) + H(y_2|y_1) - H(y_1|x).$$

The term  $H(y_2|y_1)$  reflects autonomy ambiguity, and the term  $H(y_1|x)$  reflects the destructive ambiguity of imperfect replication of information in communication with the outside world.

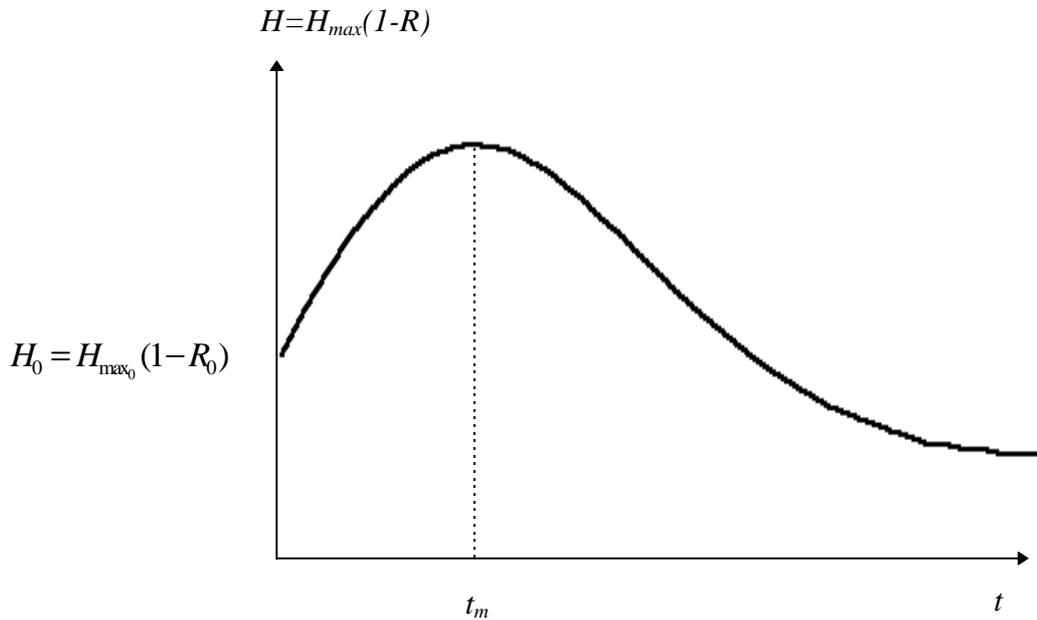
A return to the library might help clarify. Imagine that the library is in a monastery, perhaps of the kind depicted by Eco (1983). This monastery takes in various newly discovered ancient texts ( $x$ ), reads them ( $y_1$ ), copies them ( $y_2$ ), and then distributes them to other libraries. To the extent that the monks make mistakes in reading the texts (because of “noise”), then they will copy them imperfectly, and there will be destructive ambiguity — information loss. If monks subsequently read and copy imperfect copies, which are in turn read and copied subject to noise, *ad infinitum*, we would expect the signal (the books sent to other libraries) to deteriorate progressively over time. (This is, of course, the Markov-chain interpretation of the process.) But suppose that “noise” also effects the connection between reading and copying the manuscript; that is, because the monk reads the manuscript imperfectly, he purposely makes an imperfect copy — an interpretation or gloss rather than (or in addition to) a literal copy. Such a process would also introduce ambiguity, since  $y_2$  would come to diverge from  $y_1$ . But such ambiguity would add to the information in the system, since the monk would bring to the interpretation knowledge not contained in the original text. Indeed, more noise might in this context mean more information content, to the extent that more difficulty in reading the original forces the monk to apply greater originality in the commentary. The information the monk adds could, in fact, more than compensate for the signal loss, and would constitute what Atlan views as self-organization. By generalizing the two-substructure case to the case of many interconnected substructures — many scholars reading and commenting on many texts — we can envisage what von Neumann (1966) called an “extremely highly complicated

system.” In the language of Ashby (1956), increasing autonomy ambiguity can mean an increase in *variety*.

In Atlan’s formulation, this notion of ambiguity autonomy has implications for the life-cycle of organizations, both biological and social. Recall that redundancy is the key to a system’s success in overcoming noise. And the information content of the system, again, is  $H = H_{\max}(1-R)$ . If we think in terms of the system moving through time, and of the cumulative effect of noise on the system, we can differentiate this equation with respect to time, yielding:

$$\frac{dH}{dt} = H_{\max} \left( \frac{-dR}{dt} \right) + (1 - R) \frac{dH_{\max}}{dt}.$$

Since both redundancy and  $H_{\max}$  ought to decrease over time under the cumulative effects of noise, the first term on the right-hand side is positive and the second negative. The first term reflects the effects of autonomy ambiguity: an increase in information content as, in effect, redundancy is transformed into complexity by the effects of noise on the internal communications within the system. The second term reflects the destructive effect of noise on communication with the outside world. Integrating the differential equation yields a time path of  $H$ , which, can, of course, take many shapes, depending on the parameters and functional forms one assumes.



**Figure 2**

One possible pattern, however, would be one (as suggested in Figure 2) in which there is self-organization. That is,  $H$  may *increase* until some critical time  $t_m$  before decreasing. The increasing phase of  $H(t)$  reflects the effects of increasing autonomy ambiguity, which for a period overcompensates for the destructive effects of noise. During this self-organizing phase, which Atlan likens to a period of non-directed learning, complexity comes through a reduction in the system's redundancy. In effect, redundancy is transformed into complexity. After  $t_m$ , however, the system, with increasingly less redundancy at its disposal, succumbs to the destructive effects of noise as  $H_{max}$  progressively declines.

From the point of view of economic organization, however, this model may not be wholly appropriate. For one thing, in what sense does information content ( $H$ ) measure *complexity*? Is such complexity what we want to mean by organization? As we suggested early on, high entropy implies maximum *disorder* in a system. Saviotti (1991, p. 183) provides another library example. The

information content of the words arranged in the books in a library is lower than what the information content of the library would be if we cut up all the letters in the books and dumped them in a pile on the floor. The pile of letters is the  $H_{max}$  of the library, which occurs when all the letters are equiprobable. The  $H$  of the actual library is less than  $H_{max}$  because, in effect, there is redundancy: sequences of letters are repeated in ordered ways. As Saviotti suggests, the *stored information* of the actual library is greater than its *potential information* ( $H_{max}$ ). Clearly, the actual library is more ordered, more organized, and perhaps even more complex than the  $H_{max}$  library.

Indeed, von Foerster (1960, p. 37) has suggested that the appropriate measure of organization — he uses the term “order” — is the redundancy measure  $R$ . This has the nice property that when  $H = H_{max}$ , that is, when the system’s information content is at its maximum,  $R = 0$ , implying complete disorder. When, on the other hand, the elements of the system are arranged so that information about any one element conveys complete information about all others, then  $R = 1$ , and the system is completely ordered or constrained.

One result of this definition is that von Foerster’s conception of self-organization is rather the opposite of Atlan’s. For von Foerster, self-organization occurs when  $R$  *increases*, not when  $R$  is converted into increasing  $H$ . Thus, a system is self-organizing (is becoming more orderly) when  $\frac{dR}{dt} > 0$ . Differentiating

the expression for  $R$  (equation 1), we have

$$\frac{dR}{dt} = -\frac{H_{max} \left(\frac{dH}{dt}\right) - H \left(\frac{dH_{max}}{dt}\right)}{H_{max}^2}.$$

So long as we do not start out with a system of zero maximum entropy, the denominator is always positive, and the condition for self-organization becomes:

$$H \frac{dH_{max}}{dt} > H_{max} \frac{dH}{dt}. \quad (2)$$

If we consider the important case in which  $H_{max}$  is a (positive) constant — as in the case of the pile of letters on the library floor — then the condition for self-organization reduces to

$$\frac{dH}{dt} < 0.$$

Increasing order means lower entropy, which is not, of course, surprising. Notice again that this is the opposite of Altan's model of self-organization: here complexity (or information content, at any rate) is turned into redundancy instead of the other way around. In order for information content to decrease, holding  $H_{max}$  constant, there must be a change in the conditional probabilities in the direction of lower ambiguity. There must be a monk in the library rearranging the letters in the pile (into books!) so that the probability of seeing the letter  $u$  conditional on having just seen the letter  $q$  is no longer the same as, say, the probability of seeing an  $x$  conditional on having just seen a  $q$ . In von Foerster's terms, the monk is an *internal demon*.

Suppose that, instead of holding  $H_{max}$  constant, we hold  $H$  constant. In that case, the condition for self-organization becomes:

$$\frac{dH_{max}}{dt} > 0.$$

That, is the *maximum* entropy of the system has to increase — without changing the *observed* entropy. If maximum entropy is governed by the possible letters in our pile, then this condition means that we would have to add new characters to the alphabet, but in such a way that the relations among the letters, and the resulting conditional probabilities, do not change. A more relevant and sensible image, however, would be to return to the monastery library and to think in terms of new ideas entering by way of recently unearthed Greek classics. The library would be self-organizing — would be becoming more orderly — if, as the new ideas entered, they could be reconciled with the existing body of ideas in a way that kept  $H$  constant. In von Foerster's terms, there must be an *external demon* (this time more

closely akin to Maxwell's famous spirit) who filters and arranges new elements of the system so as to prevent entropy from increasing. This sort of process is probably not far from what actually went on in monastery libraries — or, indeed, from what frequently goes on in organizations in general. (We will return to this theme below.)

It is possible, of course, for  $H$  and  $H_{\max}$  to vary simultaneously. To see what this would mean, we can reinterpret equation (2) in terms of the behavior of the internal and external demons. The total differential of  $H$  is

$$dH = \frac{\mathcal{J}H}{\mathcal{J}H_{\max}} dH_{\max} + \frac{\mathcal{J}H}{\mathcal{J}t} dt$$

and

$$\frac{dH}{dt} = \frac{\mathcal{J}H}{\mathcal{J}H_{\max}} \frac{dH_{\max}}{dt} + \frac{\mathcal{J}H}{\mathcal{J}t}.$$

The first term on the right-hand side represents the work of the external demon, who controls both the rate of increase of maximum entropy and the way in which changes in maximum entropy affect system entropy ( $H$ ). Alternatively, we can think of  $\frac{dH_{\max}}{dt}$  as reflecting the insertion of variety from outside the system. (In

our monastic library example, this would be the flow of rediscovered classical ideas.) In that case,  $\frac{\mathcal{J}H}{\mathcal{J}H_{\max}}$  represents the effect of the external demon as

gatekeeper. If this partial is 1, then all the injected variety is turned into entropy, and the external demon has had no effect; if the partial is 0, then the demon is a perfect gatekeeper, and variety enters the system without increasing  $H$ .  $\frac{\mathcal{J}H}{\mathcal{J}t}$

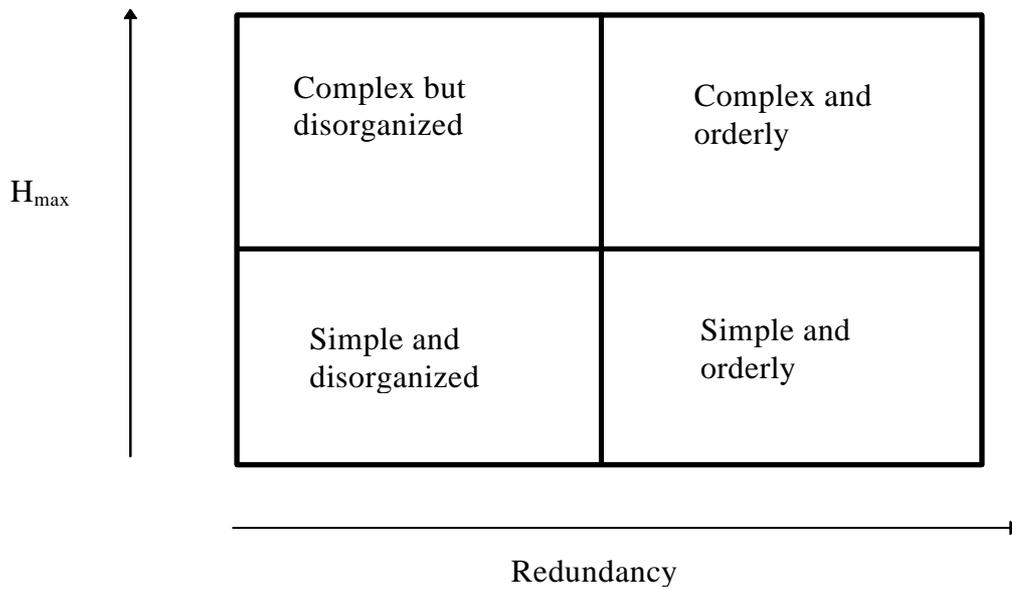
represents the work of the internal demon, the organizer who arranges the elements already in the system. Plugging  $\frac{dH}{dt}$  into equation (2) gives:

$$\left[ H_{\max} \frac{\mathcal{J}H}{\mathcal{J}H_{\max}} - H \right] \frac{dH_{\max}}{dt} < -H_{\max} \frac{\mathcal{J}H}{\mathcal{J}t}. \quad (3)$$

Notice that, when  $H_{max}$  is large, learning is relatively easy, in the sense that the efforts of the internal demon in keeping the right-hand side greater than the left-hand side (and thereby keeping the time derivative of  $R$  positive) are multiplied by a large number. As the internal demon works to lower  $H$ , however, the bracketed term on the left-hand side will grow larger. This means that the internal demon will have to work harder (make  $\frac{dH}{dt}$  more negative), or, more typically, the rate of growth of  $R$  will slow as  $R$  increases. Of course, as  $H_{max}$  increases, the efforts of the external demon will be multiplied by a larger number, and, in general, the external demon could help the internal demon maintain the rate of growth of  $R$  for a while by decreasing  $\frac{dH}{dH_{max}}$ . Eventually, however, unless  $H_{max}$  increases fast enough, the rate of growth of  $R$  will slow.

### **Self-organization: a schema.**

Where does this leave us in the theory of self-organization? We suggested that Atlan's notion of self-organization is unsatisfactory in that it equates self-organization with an increase in information content. By contrast, von Foerster's choice of redundancy as a measure of orderliness has some appeal. Nonetheless, in the end redundancy also provides an inadequate account of self-organization. The reason is that redundancy is a *relative* measure of order.  $R$  approaches 1 when  $H$  is small relative to  $H_{max}$ . This means, however, that this measure would count as equally organized a system with only two possible states and a system with a million possible states, so long as their observed entropies were the same fraction of their maximum entropies. Clearly, however, our intuitive notion of "highly organized" encompasses both Atlan's notion of complexity as high  $H$  — or, more appropriately, high  $H_{max}$  — and von Foerster's notion of orderliness as redundancy. In other words, self-organization means *both* increased complexity (or variety) and increased redundancy. Figure 3 summarizes this idea.



**Figure 3**

Thus both Altan and von Foerster are necessary for a complete account of self-organization. Clearly, self-organization implies a movement from the southwest to the northeast corner of Figure 3. But that movement needn't be along a straight line. We can imagine an organization (for example, a firm or network of firms) that first moves directly north up the diagram and then gradually moves east. Such an organization thus has an early phase of non-directed learning — a phase of what, following Piaget, Atlan (1972, p. 267) calls “assimilation.” In this phase, growth in  $H_{max}$  may outstrip the powers of both the internal and the external demons to hold the line on  $H$ , and  $R$  may actually increase (the organization may veer temporarily to the west). Self-organization also requires, however, that eventually the redundancy of the system increase as potential is transformed into stored information.

Without wanting to make too much of the point, we should notice that such a pattern is congruent with life-cycle models of innovation and product development (Utterback 1979). In those models, the early stages of the life of a new product are

marked by fluidity, rapid change in characteristics, and a diversity of approaches. Eventually, however, a dominant design or paradigm emerges. At that point, product innovation gives way to process innovation, and learning — often rapid learning, as measured by declines in product price — takes place within a relatively fixed structure.

## References.

- Arrow, Kenneth J. 1974. *The Limits of Organization*. New York: Norton.
- Ashby, W. Ross. 1956. *An Introduction to Cybernetics*. London: Chapman and Hall.
- Atlan, Henri. 1972. *L'Organisation Biologique et la Théorie de l'Information*. Paris: Hermann.
- Boulding, Kenneth E. 1955. "Notes on the Information Concept," *Explorations* [Toronto] 6: 103-112.
- Eco, Umberto. 1983. *The Name of the Rose*. Trans. William Weaver. San Diego: Harcourt, Brace, Jovanovich.
- Edelman, Gerald M. 1987. *Neural Darwinism: the Theory of Neuronal Group Selection*. New York: Basic Books.
- Hayek, F. A. 1952. *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*. Chicago: The University of Chicago Press.
- Hayek, F. A. 1978. *New Studies in Philosophy, Politics, Economics, and the History of Ideas*. Chicago: The University of Chicago Press.
- Langlois, Richard N. 1983. "Systems Theory, Knowledge, and the Social Sciences," in Fritz Machlup and Una Mansfield, eds., *The Study of Information: Interdisciplinary Messages*, New York: John Wiley, pp. 581-600.
- Langlois, Richard N. 1986. "Coherence and Flexibility: Social Institutions in a World of Radical Uncertainty," in Israel Kirzner, ed., *Subjectivism, Intelligibility, and Economic Understanding: Essays in Honor of the Eightieth Birthday of Ludwig Lachmann*. New York: New York University Press, pp. 171-191.
- MacKay, Donald M. 1969. *Information, Mechanism, and Meaning*. Cambridge: MIT Press.
- Machlup, Fritz. 1983. "Semantic Quirks in Studies of Information," in Fritz Machlup and Una Mansfield, eds., *The Study of Information: Interdisciplinary Messages*, New York: John Wiley, pp. 641-671.
- Nelson, Richard R., and Sidney G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge: Harvard University Press.
- Saviotti, P. Paolo. 1991 "The Role of Variety in Economic and Technological Development," in Saviotti and J. Stanley Metcalfe, eds., *Evolutionary Theories of Economic and Technological Change*. Chur, Switzerland: Harwood Academic Publishers.

- Utterback, James M. 1979. "The Dynamics of Product and Process Innovation," in C. T. Hill and J. M. Utterback, eds., *Technological Innovation for a Dynamic Economy*. New York: Pergamon Press, pp. 40-65.
- von Foerster, H. 1960. "On Self-organizing Systems and their Environments," in Marshall C. Yovits and Scott Cameron, eds., *Self-organizing Systems*. New York: Pergamon Press.
- von Neumann, John. 1956. "Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components," in C. E. Shannon and J. McCarthy, eds., *Automata Studies*. Princeton: Princeton University Press.
- von Neumann, John. 1966. *Theory of Self Reproducing Automata*. Ed. A. W. Burks. Urbana: University of Illinois Press.
- Weimer, W. B. 1982. "Hayek's Approach to the Problems of Complex Phenomena: An Introduction to the Psychology of the Sensory Order," in W. B. Weimer and D. Palermo, eds., *Cognition and the Symbolic Processes*. Volume II. Hillsdale: Lawrence Erlbaum.