ISSN 0980-1472

Revue Internationale de systémique

Vol. 3, N° 3, 1989

afcet

Dunod

AFSCET

Revue Internationale de systémique

Revue Internationale de Sytémique

volume 03, numéro 3, pages 237 - 249, 1989

Noise, complexity, and meaning in cognitive systems

Henri Atlan

REVUE INTERNATIONALE DE SYSTEMIQUE
Vol. 3, N° 3, 1989, pp. 237 à 249

NOISE, COMPLEXITY AND MEANING IN COGNITIVE SYSTEMS *

H. ATLAN

Hadassah Medical Center [1]

Abstract

The classical computer model is criticized in spite of its heuristic value both in psychology (the brain as a computer) and in biology (the genetic program metaphor). Learning with language acquisition is presented as a self-organizing process whereby diversity and complexity increase at the expense of initial redundancy. According to the complexity from noise principle, literal meanings are not viewed as primary but rather as acquired, in a process of differentiation, starting from an undifferentiated state which appears polysemic and metaphorical.

Résumé

Le modèle d'ordinateur classique est critiqué malgré sa valeur heuristique aussi bien en psychologie (le cerveau-ordinateur) qu'en biologie (la métaphore du programme génétique). L'apprentissage avec acquisition du langage est présenté comme un processus auto-organisateur par lequel la diversité et la complexité augmentent aux dépens d'une redondance initiale. En application du principe de complexité par le bruit les significations littérales ne sont pas vues comme primaires mais plutôt comme acquises dans un processus de différenciation à partir d'un état indifférencié qui apparaît comme polysémique et métaphorique.

1. Department of Medical Biophysics, p.o.b. 12000, 91120 Jerusalem, Israël.

* This article is the transcription of a lecture given in 1984 as a contribution to a conference on Perspectives in Cognitive Neuropsychology (Bressanone, Italy). Since then, as suggested in the conclusion, the automata network paradigm which was then in its infancy (2, 12) has established itself under the name of "neo-connectionism". However, the criticism of the computer model presented here should be extended to the new models of parallel computing if they are to be taken literally as the "true" representation of the way in which our mind works.

Revue internationale de systémique. 0980-1472 Vol. 3/89/03/237 - 249/13/$ 3.30 © Gauthier-Villars

The subject of this paper has to do with cognition only in a limited sense since it concerns non human cognitive systems, namely biological systems exhibiting learning properties. More precisely the development of the immune system or the Central Nervous System in an individual, or else, the evolutionary process at the level of the species, can be viewed as non directed learning processes. Several basic questions are asked regarding their mechanisms, and these questions may be relevant, to some extent, to those encountered in the study of human cognition. Even though they would not be relevant, at least they would offer an alternative to the computer paradigm which seems very often the only one people can have in mind when they want to build mechanistic models of human information processing.

In this context, I shall try to answer, at least partly, the following questions :

– how complexity can grow out of a less complex system ?

– what is the meaning of information, if any, in non human information processing systems ?

– how is non directed learning possible at all, since it implies a kind of paradox : in order to learn something new one has to perceive it as a pattern to be learned, that is one has to recognize it. And if one is able to recognize it, it means that this pattern is not new to him, that he knows it already. Therefore it would seem that no real learning can ever take place ;

– another question linked to the previous one : is it possible to describe some irreversible process in non directed learning, whereby some quantity decreases while some other quantity – which will be called the knowledge – increases ?

All these questions have been asked in basic biology and the "dogmatic" answer, i.e. the one coming from the so-called dogma of molecular biology, has been given in the form of a metaphor, the metaphor of the genetic program : every individual is determined by a computer-like program written in the DNA sequences of its genes. His development, his growth, his reactions to the environment acting as a source of stimuli, are the results of the execution of this program. This program has been written during the evolutionary process by natural selection, such that the organism is adapted to its environment ; i.e. it reacts to its environment by sub-routines written in its genetic program, and called by given stimuli.

The problem with the genetic program is that it is a metaphor which should not be pushed too far because it leads to practical difficulties and

the only way to set around it is to modify the idea of a computer program to such a point that almost nothing is left from the metaphor.

Already from the beginning of molecular biology – and this can be found in every text book – the so-called genetic program was described as a very peculiar program since it needs the products of its reading and execution to be read and executed ; or else, that it functions as a self-programming program. Additional questions were raised about the language in which it is written and it was recognized very soon [1] that the genetic code certainly does not constitute a computer language, being at most a lexicon. Nevertheless the computer model is useful everytime one has to deal with the problem of "what is organization" because the computer represents today a kind of maximum, in the sophistication of an organized machine.

However, the computer model has a big disadvantage : one finds in it, at the end, what one wanted precisely to eliminate at the beginning, i.e. some aspect in the crude observed phenomena which was difficult to account for by a scientific explanation, so that the model was built in order to eliminate it. I am thinking for example about finalism in biology and subjective introspective data in psychology, that the computer models are supposed to eliminate.

In biology one wanted to eliminate the apparent purposes in the observed development and adaptative phenomena by sticking to mechanistic processes but one finds them back in the computer model since a purpose is always present in any computer program.

In cognitive sciences also it seems that one wants to eliminate the subjective introspective nature of human experience but one finds it back disguised within the semantic components of any computer-like model, where the semantic content must be fed from the spontaneous or natural human experience.

This is due to the fact that the computer is not a natural system, just offered by nature to be used as a model for other natural systems like the living organism or the human cognitive system. It is itself a production of a human cognitive system. As a matter of fact it is only a particular instance, a particular kind of production of a human cognitive system.

The fact that computers can do things that human brain cannot do must not be misleading : it has to do mainly with an increase by several orders of magnitude in the speed of information processing. This speeding up of the process is far from being trivial and it has far reaching consequences [2] but the logical structure involved in the functioning of a computer program is no more than the logical structure created by the programmer. As such it is a particular production of a human cognitive system and it should not be more legitimate to use it as a paradigm for how the human cognitive system functions than any other production of a human cognitive system, such as a metaphysical theory, or a philosophical theory, or any system of interpretations that a man can imagine and use to put some order into its experience in the world.

It is certainly an interesting problem in the history of science to understand why nevertheless the computer model has imposed itself not only in human sciences but also in basic biology, after a century where the central model was the steam machine, which itself had replaced the watch and the pendulum. As I said before the metaphor of a computer program in basic biology was meant to account for apparently finalistic genetic determinations. The word teleonomy was invented in the fifties by Pittendrish [3] to replace teleology, in order to get rid of finalism in biology. Apparent purposes are observed in adaptation and development and classical biology was always mixed up with teleological, i.e. finalist thinking. In order to get rid of it, the goal of research was stated as looking for mechanisms with teleonomy, the difference being that in one case (teleology) one deals with purposeful end-directed processes governed by an "intelligent, designing mind", while in teleonomy one is looking for non purposeful end-directed processes. The computer came perfectly to provide a model for such a non-purposeful end directed process : the computer executes a program, mechanistically, without purpose, although it works as an end-directed process. The only problem of course is, that in the metaphor of the genetic program the existence of a programmer and his purpose are forgotten ; of course the nature of the programmer who wrote the genetic program is left ouside the scope of the metaphor... unless one wants to say that natural selection works as a programmer writing a program. Then one is led back to the purpose always present in the work of a programmer.

It is to try to avoid these difficulties that I got interested in possible theories of self-organization, where the question was : which logical prerequisites must be met in order that a self-organizing system exist ? i.e. a system where no program would have to be fed from the ouside by a programmer, or a system working like a self programming program.

Now there was a famous demonstration by Ashby [4] showing that self-organization cannot exist.

The demonstration goes as follows : if, by organization we mean the laws that govern the behaviour and evolution of a system, a change in these laws must be produced by the functioning of the system itself in order for it to be self organizing. Now, if the functioning of the system can produce a change in its laws, then it is the law governing this change which must be looked at as its organization and not the laws themselves ; and the law by which the laws are changed, is, itself constant ; therefore the organization of the system is constant and there is no self-organization. A true change in organization must come from the ouside, i.e. the environment of the system and then, again, there is no self organization.

However, there are two possible ways by which the environment can induce changes in the organization. One of them is by means of instructions, as in a program, and then, really, there is no reason whatsoever to talk about self organization, since the organization is imposed by a preexisting plan from the ouside. However, there is another way, namely when changes in the organization are produced by random perturbations coming from the environment.

Then, if random perturbations can produce a change which is not mere disorganization, there is some reason to talk about self organization, even though the triggering of the change comes from outside the system [5]. By the way, this holds for random perturbations coming from inside as well, as it is the case, for example, in thermal fluctuations.

From this comes the idea that randomness must play some role in self organization. In order to formalize this kind of thing one had also to define more precisely what is meant by organization. This can be done in a number of ways by different formalisms, such as automata theory, or information theory. When one looks in the literature to see what is meant by organization in general, in a kind of naive or self evident way, one finds two different, in fact opposite properties. One is repetition, order in the form of strong constraints between parts, or between states in time. This property in its most general form is called redundancy and it is a measure of how much knowledge about parts of the system allows to know something about other parts, due to known constraints between them. However, depending on the authors, one finds also variety or diversity as a property typical of the existence of an organization, especially when complex adaptable behaviour is exhibited. This property is in fact opposed to the first one, and it is a measure of the uncertainty or unexpectedness, or the so-called information content of a system in the formalism of information theory. As such it can be used to measure the complexity of a natural organization.

These two opposite properties appear both as obvious features of what is organization and therefore a good theory must take them both into account in such a way that an optimal organization will appear as a kind of compromise between maximum order or redundancy and maximum disorder or complexity [6].

From then on, it is possible to see how random perturbations can produce a change in organization by reducing the redundancy and increasing the complexity of a system, at least, up to a certain point, as long as there is enough redundancy to keep the system going. This is the so-called complexity from noise principle, at the root of the formal theory of self organization that I have worked out several years ago. [5, 6, 7, 8, 9]

This theory states something about necessary conditions for self organization :

– the initial redundancy must be high enough to start with, since it has the role of a potential for self organization to be reduced in the process ;

– some inertia must exist in the response of the system to random perturbations so that it is not going to be destroyed too quickly in order for the initial redundancy not to disappear right away. (A good example is that of liquid crystals and biological macromolecules where thermal noise can produce changes in conformation before it destroys the whole structure of the molecule, whereas the high repetitive order of crystals can exist only in one form and is completely destroyed almost at once when the "dose" of thermal noise – i.e. the temperature – reaches some threshold).

Of course these are not sufficient conditions because the sufficient conditions have to do with specific forms of functional organization, i.e. with specific functions performed by the organized system, and giving a meaning to its organization. I shall come back to this point later.

This theory can give some answers to the questions asked at the beginning :

– what increases in learning is the complexity of the learning system itself and what decreases is its redundancy ;

– the paradox of the impossibility of learning is solved by assuming at every stage a given pattern within the learning system to be projected into the environment and fit approximately with something in the environment. This fit is enough for this thing to be "recognized". However, the fit is not perfect and the errors trigger modifications in the pattern whereby a new pattern is generated and something is learned. Again the newness comes from the errors or the noise in the fit.

What about the meaning ?

This theory was stated in a formalism where the meaning of information is completely disregarded, namely the probabilistic Shannon information theory. But this had an advantage in pointing out clearly where the effect of something like the meaning of information had to be taken into account, even in non human, natural, information processing systems.

More precisely, on the one hand, it is possible to analyze channels of transmitted biological information in the purely probabilistic Shannon sense, at the molecular level of the organization of living cell : the linear structure of the DNA as sequences of nucleotides can be viewed as input to a channel ; the linear structure of the proteins as sequences of amino acids can be viewed as output ; and the correspondence between them as it takes place in protein synthesis, can be treated as a noisy channel in Shannon sense, without taking into consideration any element of meaningfullness of the transmitted genetic information.

However, at the level of the cell itself, this meaning exists and is manifested as the function of the cell. The transmission of information in the protein synthesis channel will result in correct or incorrect proteins and this will have different consequences at the level of the cell metabolism i.e., of the functioning of the cell. In other words the meaning of the genetic information is to be found in the effects of this information on its receiver, the receiver being the cell itself. Or else, the meaning of the genetic information is to be found in the phenotype. Thus, one can ascribe some meaning to information transmitted in a non human communication channel, and this meaning can be described as the effect of the information on the state of the receiver.

(Let us notice that this definition is certainly poor as compared with what we know or do not know about meaning in human communication, but it does not seem to be wrong. At least a part of what the meaning in human language appears to be can fit this definition : it seems to me that the effect of a message on the state of the receiver when the receiver is a human brain is certainly a part of what makes the meaning of a linguistic message).

Now, the basic observation which was used to formalize the theory of self-organization by positive effects of noise, had to do with a change in the level of observation [6, 7].

To recall it briefly, the effect of noise acting on a channel of communication between x and y is to reduce the information transmitted in the channel by an amount called the ambiguity, thus counted netatively : for example, one of the Shannon formulae for the transmitted information from x to y is

$$T(x ; y) = H(y) - H(y/x)$$

where H(y) is the information content or complexity of the output y and H(y/x) is the ambiguity-function, produced by the noise.

However, if one is interested in looking at the system, not at the level of the channel output but at that of the whole system containing x and y as related parts, then the same ambiguity function H(y/x) will be counted positively, since

$$H(x,y) = H(x) + H(y/x)$$

In other words, the contribution of x and y to the overall information content or complexity of the whole system increases when the ambiguity between x and y increases. Thus, the sign of the ambiguity changes depending upon the level at which the observation is made. Moreover, this change in the level of the observation is also a change in the level of the organization itself.

In the example of the cell and the protein synthesis channel, what is seen as detrimental by the cell, i.e. errors in protein synthesis, may be seen as beneficial, i.e. an increase in variety and diversity of cells at the level of the organism.

This means that we have also to consider transmission of information from one level to the other in a multi-level organization ; and this information from one level to the other must not be taken in the Shannon sense, but in its complete sense, i.e. including its meaning, defined as its effect on the receiver.

Thus, the fact that a multilevel system is able to utilize random perturbations means that it is able to create new meaning of information transmitted from one level to the other. This creation of a new meaning is what reorganization is about. We, as external observers do not have access to these meanings, since we only see the end products and the changes in the structure and performances of the whole system. That is why the complexity from noise principle could be stated in a formalism where the meaning is not taken into account and the appearance of a paradox (noise creating organization) because it worked as a double negation : destruction of a kind of information where the meaning is absent is a way for us as external observers to describe the creation of meaning within a natural system, when we do not have enough control on it to have access to the information transmitted from one level to the other.

Now if one wants to come back to human cognitive systems and try to apply the theory of self organization by complexity from noise, then, we have seen that non directed learning can be thought of as a process of complexification with decrease in the initial redundancy and increase in diversity. By this process, new meanings are created by and within the cognitive system itself. When we apply this idea to the acquisition of language, seen as a self organizing system, the older stages or first steps in the process appear as states of undifferentiation, with high redundancy in the sense that many different signs and symbols are equivalent because they have vague, overlapping ambiguous meanings. It is only at later stages that the language has become more diversified, more differentiated, as a result of a reduction in redundancy such that every sign becomes efficient in a different specific way.

Therefore a semantic unambiguous content is not to be found at the beginning of the process as some property to be added to, or part of, built in syntactic structures, but rather at the end of the process ; starting from a non specific, polysemic and to some extent presyntactic structure, ending as a strictly defined syntax, with completely non ambiguous meanings, as in artificial logicomathematical language.

From this point of view the evolution of the semantic component of language is not seen as an acquisition of polysemy and metaphorical meaning starting from well defined literal, non ambiguous meaning, but just the opposite.

The initial stages are characterized by a lack of specificity, i.e. a wealth of non differentiated meanings, acting as a potential for future literal meanings. And the latters appear at the end as a result of a reduction in the initial redundancy and a creation of a one to one correspondence between a symbol and its effect or meaning. (There seems to be a contradiction with what I have said before since this process appears to be a reduction in ambiguity whereas we have seen that complexification is produced by an increase in ambiguity between the parts. Again this is a consequence of the change in the level of observation. The meaning of language we are talking about now, is in fact the value of each part of the system, i.e. each sign or symbol, viewed from outside, by us as external observers of the language considered as a self organizing system. The ambiguity between parts increases in the sense that every one of them becomes independent in a semantic "space", by acquiring a specific value or effect – what we call its meaning – independant from that of the others ; whereas at the beginning, in the undifferentiated state, with polysemy and apparently metaphorical meanings, many different signs are equivalent, sharing common meanings, ill defined to our eyes which implies that the constraints between them are strong, i.e. the ambiguity in the communication between them is minimum).

It is interesting to see what happens when the process is led to its extreme, in the form of an artificial logico mathematical language. There is a one to one correspondence between the symbol and its meaning but at the same time, as a formal language, it can exist as a structure of symbols without them having any meaning ; in other words, whenever the language reaches the state of having non ambiguous meanings, it can be formalized and exist without any meaning. This would imply that what makes really the meaning of things and words in non formal language is precisely their polysemic nature rather than their possible unambiguous theoretical semantic content.

This will lead me to a second remark regarding the relationship between linguistic theories and the acquisition of natural language by so-called native speakers. The first remark had to do with the application of the complexity from noise principle according to which specificity and complexity, appearing as literal meaning in this case, come at the end as the result of a progressive restriction of initial non differentiated, apparently metaphorical, meaning. (This does not imply necessarily that a truly metaphorical meaning is developed in the early stages of language acquisition, before the understanding of literal meaning is acquired by the child. This non differentiated meaning appears to be "metaphorical" by analogy with our adult experience of metaphors which is based of course on our ability to recognize literal meanings and distinguish them from metaphorical ones. Therefore, it may be the case that a second stage with (adult) metaphorical meanings appears, or is retained, after the acquisition of literal meaning took place. It has been suggested [8] that such a recharging in redundancy might be one of the functions of the paradoxical sleep. Electrical cortical activity seems to be widespread to the whole cortex indicating less inhibitory associative activity ; simultaneously, differentiation effects of unambiguous identification and definitions seem to be removed and (metaphorical) associations occur in dreams while being otherwise inhibited.)

The second remark has to do with the fact that grammar and linguistic theories have taken the form of deterministic procedures written in formal languages, which makes them irrelevant, to some extent, to the problem of natural language acquisition. Not only native speakers do not seem to have any kind of innate rational grammar but such grammar must be imposed on them by directed learning as logics and mathematics. Quine's [10] distinction between fitting and guiding rules is relevant here. The rules discovered by the grammarian are aimed at fitting the actual production of native speakers' well formed strings. Even though they would succeed perfectly, it would not mean that they are identical with the unconscious rules which guide this production. In the same way as an electron is not supposed to know, even unconsciously, the rules deduced from Shroedinger's equation which fit the observations of its behavior. It is true that there is always a temptation to believe that the electron – and the native speaker – is guided by the rules. This may be more justified for the electron because it is no more than a solution of the Shroedinger equation in the sense that it does not manifest itself outside the scope of experimental and theoretical physics. Can we say that the native speaker's language is no more than the grammar in the sense that it would not manifest itself outside the scope of linguistic theories ? In his above mentioned article (concluded "with a plea against absolutism"), Quine showed how the purpose of logical analysis of sentences is different from that of grammar, and that both purposes are operational. They paraphrase sentences of ordinary language with technical symbols in order to gain some clarity and efficiency having in mind a given, specific, operation to be performed (logical analysis to remove paradoxes, or algorithmic production of well formed strings). The tools used in the paraphrases such as logical symbols and quantifiers, or algorithmic trees, are no more implicit in the ordinary language than other specific technicalities such as the fourth dimension in Physics to deal with time or the binary code for computer programming. Again, natural scientists are tempted to believe that these specific language tools are implicit in the natural phenomena they succeed to account for or to master. It seems to me that it would be even less justified to believe that logical or linguistic structures are implicit in the ordinary language viewed as a natural phenomenon.

In fact, even the fit is not perfect. The criterion of well formed strings is a criterion for grammarians which not only is not enough to account for the meaning, but contradicts the native speaker's criterion of meaningfulness. He must learn the grammarian's criterion as an artificial one, just like logics and mathematics, and in so doing he is forced to eliminate meaningful strings as not well formed or to include well formed strings which are spontaneously meaningless to him. He learns progressively to combine the two criteria together just as one learns being logical in life by combining logical criteria of decision with spontaneous ones whereas the set of spontaneous criteria was not logical to start with. Thus, the existence of an innate competence to form infinite meaningful strings does not imply that it has the rational and logical structure of a grammar. The most prominent feature of rationality and logics is the use of negation, with the derived law of non contradiction, and the even more derived use of demarcation by social consensus about what is meaningful and what is not. Generative grammars (like other grammars) are nothing else than systematizations of this directed learning process by setting written rules for such demarcations. As it was the case regarding more classical grammars, Wittgenstein's remark is still pertinent : "Grammar does not tell us how language must be constructed in order to fulfil its purpose, in order to have such-and-such an effect on human beings. It only describes and in no way explains the use of signs" [11]. What differentiate's generative grammars from more classical may be the logical level of its rules : while classical grammars were satisfied with rules of demarcation, generative grammars are looking for rules of production. As such they may be very useful for artifical languages to be produced in a logical way by algorithmic deterministic procedures whereas the old grammars would not be of any help in this task ; but nothing more as far as natural languages are concerned. Instead, it is possible that the use of probabilistic algorithms and heuristics would bring closer to natural language, since the use of some degree of indetermination and randomness seems to be a necessary ingredient to account for self organizing properties. Recent works on random networks [2] and networks of probabilistic automata [12] may provide some tools for progress in this direction.

References

[1] B. SHANON, "The Genetic Code and Human Language", *Synthese*, v. 39, 401-415, 1978, See also D.R. HOFSTADTER, Godel, Esher, Basic books, NewYork, 1979.

[2] H. ATLAN, F. FOGELMAN-SOULIE, J. SALOMON, J. WEISBUCH "Random Boolean Networks," *Cybernetics and Systems*, 12, 103-121 (1981).

[3] C. S. PITTENDRIGH, "Adaptation, Natural Selection, and Behavior" *Behavior and Evolution*, A. ROE & G. G. SIMPSON eds., Yale Univ. Press, New Haven Connect., 1958, pp. 390-416.

[4] W. R. ASHBY, "Principles of the self-organization system", In *Principles of Self Organization*. Eds. H. Von FOERSTER & G.W. ZOPF, Pergamon Press, N. Y., 1962, pp. 255-278.

[5] H. ATLAN, *L'organisation biologique et la Théorie de l'information*, Hermann, Paris, 1972.

[6] H. ATLAN, "On a formal definition of organization", *J. Theoret. Biol.* 45, pp. 295-304, 1974.

[7] H. ATLAN, "Application of information theory to the study of stimulating effects of ionizing radiation, thermal energy, and other environmental factors. Preliminary ideas for a theory of organization", *J. Theoret. Biol.* 21 : 1, pp. 45-70, 1968.

[8] H. ATLAN, *Entre le cristal et la fumée*, Paris, Seuil, 1979.

[9] H. ATLAN, "Hierachical Self Organization in living Systems, Noise and Meaning", in *Autopoiesis : A theory of the Living Organization*, Ed. M. ZELENY, North Holland, N. Y., 1981, pp. 185-208.

[10] W. V. QUINE, "Methodological Reflections on Current linguistic Theories", in *Semantics of Natural Language*. Eds. DAVIDSON and HARMAN, D. Reidel Publ. Co., Dordrecht, Holland, 1972, pp. 442-454.

[11] L. WITTGENSTEIN, *Philosophical Investigations*. English translation. GEM Anscombe. Basic Blackwell, Oxford, 3rd Edition, 1968, p. 138, nº 496.

[12] M. MILGRAM & H. ATLAN, "Probabilistic Automata as a Model for Epigenesis of Cellular Networks.", *J. Theoret. Biol.*, 103, pp. 523-547, 1983.