REVUE INTERNATIONALE DE SYSTÉMIQUE
Vol. 9, N° 3, 1995, pp. 347 à 371

CHARACTERIZING EMERGENT PHENOMENA (2): A CONCEPTUAL FRAMEWORK

Eric BONABEAU [1], Jean-Louis DESSALLES [2], Alain GRUMBACH [2]

Abstract
The lack of a unifying conceptual framework for representing, characterizing, and dealing with emergence and emergent phenomena, led us to study the building of such a framework, based on the notions of levels of organization and of levels of detection. Information theory, and concepts related to theories of complexity, will help understand the nature of emergent phenomena.

Résumé
L'absence d'un cadre de pensée clair et unificateur qui permettrait la représentation, la caractérisation et le traitement de l'émergence et des phénomènes émergents, nous a conduit à proposer un cadre d'étude englobant, reposant sur les notions de niveaux d'organisation et de niveaux de détection. La théorie de l'information, combinée à des concepts récents liés aux théories de la complexité, permettent une meilleure compréhension de la nature des phénomènes émergents.

I. INTRODUCTION

The term "emergence", whose numerous definitions have been explored from the existing literature in our previous paper (Bonabeau *et al.*, 1995), obviously applies to a wide spectrum of phenomena, with a variety of different meanings, strengths, and consequences. We believe it both possible and worthwhile, however, to understand all these apparently irreconciliable conceptions of emergence within a common conceptual framework, which, to the best of our knowledge, is still lacking. Such a tentative framework, built

1. CNET Lannion B-RIO/TNT, 22307 Lannion Cedex.
2. Département Informatique, Télécom Paris, 46, rue Barrault, 75013 Paris.

Rev. intern. systémique. 0980-1472 Vol. 9/95/03/ $ 4.00/© Afcet Gauthier-Villars

upon levels of organization, levels of detection, and theories of complexity which make a unifying integration feasible, is developed in the present paper. We first introduce levels of organization (section 2) and levels of detectors (section 3), to eventually integrate them into a common view of emergence and emergent phenomena (section 4).

II. LEVELS OF ORGANIZATION

In this section, we make explicit some characteristic features which are common to some of the examples presented in our previous paper (Bonabeau *et al.*, 1995). We introduce the notion of level of organization which is first studied from an intuitive point of view, and then from a more formal point of view based on category theory.

II.1. Characteristic features

Let us consider for instance the following three examples (see Bonabeau *et al.*, 1995, figures 1, 2 and 3): the picture of a dog, an ant bridge and economic systems.

These three examples share certain features, corresponding to some important characteristics of emergence. In particular, we note the presence of:

(i) actors: interacting agents with **local perception** and the ability to **act locally**.

(ii) spectators: one or several entities sensitive to the emergent phenomenon, and possessing **global perception**.

(iii) a process including:

– an initial state, based on an organization level N

– a sequence of events leading to

– a final state corresponding to an upper organization level N'

(iv) a time scale for this evolution; this time scale is compatible with actor and spectator time scales.

To make these aspects clearer, let us consider ants:

(i) the agents are able to:

– perceive their local environment

– act in this environment.

(ii) the notion of an entity that is sensitive to an emergent phenomenon is not easy to define. In effect, the emergent aspect of a phenomenon is related to the point of view of an observer of this phenomenon: it is not intrinsic to the phenomenon, but related to the global system (phenomenon + observer). A first kind of observer may just look at the formation of the emergent phenomenon (for instance, the man who sees the ant bridge) but this observer is neither active in the situation, nor directly concerned by the phenomenon. We suggest to introduce a new kind of entity sensitive to the phenomenon, for which the absence of the phenomenon is critical, which we call the "involved being". For instance in the ant bridge example, the involved being is the ant colony. Many ant bridges exist in nature, without human spectator; but in each of them the ant colony is concerned, the bridge increasing the adaptive capacity of the colony.

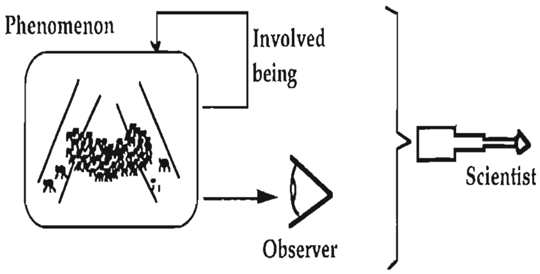


Figure 1.

(iii) The two levels of organization are:

– the set of ants

– the bridge

(iv) Finally, the time scale is around one minute.

We can study other examples very similarly. A summary of the characteristic features concerning each example is given in the following table which includes: levels of organization, time scales, agents and the different functions (local perception, local action, global perception). The first three lines describe the examples. The last three lines describe counter-examples which are different from examples through one feature only (near-misses).

The dog picture example is not as obvious as it seems: for this example, it is not the environment which is dynamical, but the observer's perceptual process. As regards the (artificial) bridge counter-example, the time scale is too long: the involved being is the walker or the driver, but he could not see the building of the bridge since this operation takes several years. Lastly the neural network counter-example concerns those networks which evolve towards a stable state, such as Hopfield nets. These models take as

feature / example	n1 level	n2 level	time scale	local perception	local action	global perception
dalmatian dog	point	concept	second	sensor	neuron	observer
ant bridge	ant	bridge	minute	ant	ant	colony
economic phenom.	purchase/sale	crash	hour	agent	agent	agent
dog	point	concept	*millisec.*	sensor	neuron	observer
bridge	brick	bridge	*year*	pontoneer	pontoneer	traveller
neural network	pixel	pattern	second	formal neuron	formal neuron	

Figure 2.

input a noisy pattern and give as output the corresponding prototype (not the corresponding synthetic information associated with it [the symbol]). If they are considered as one link of an automatic processing chain, the output will not be usable for further processing. Thus this artificial agent lacks global perception.

II.2. Distinctive features

Let us now try to list some distinctive features. One interesting distinctive feature is related to the following three functions: local perception, local action, global perception. For the dog picture, the three agents performing these three functions are all different; in the ant bridge, both local functions are achieved by the same agent; as regards the economy, the three functions are achieved by the same agent.

II.3. Levels of organization

Emergence, as characterized by many, depends on the notion of level of organization. Its scope being wider than that of emergence, we study it from a more general point of view.

Examples

– A paper such as this one, is organized into chapters, sections, sentences, phrases, words, characters, pixels, and so on;

– A human body is organized into molecules, cells, organs, systems;

– A processing unit such as an engine, a television set, a computer, even a human being;

– A set, a society, a colony of entities...

Let us notice that among these examples, some entities are static (paper), others are dynamic (cells, processing unit, society), some are closed, i.e. without interactions with the environment (paper), others do interact (cells, processing unit, society).

Characteristic features

Let us now extract common features which characterize levels of organization:

– multiplicity of levels;

– a composition law defines upper level entities starting from lower level ones;

– relations between levels (analogy, hierarchy)

– irreducibility: it is not possible to suppress one level, to connect level n-1 directly to level n+1;

– autonomy of a level: existence of properties, relations, behavioral laws concerning entities at a given level, independently from other levels.

Synthesis

These characteristic features may be synthetized as illustrated by the following schema:

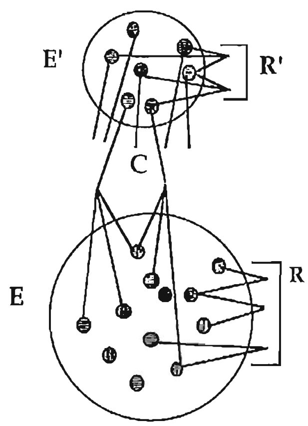


Figure 3.

E is a set containing elements e. C is an external composition law from finite sequences of E elements to E′ which is the target set. R is a potential relation on E^n. R′ is a potential relation on $E'^{n'}$.

$\langle E', R' \rangle$ is a level of organization if there exists a R′ relation on $E'^{n'}$ which cannot be deduced from R relations on E^n

For instance, E may be an alphabet, E′ the set of all words on that alphabet, C letter composition giving words, R′ a syntactic or semantic relation between words; R′ cannot be deduced from relations between letters. This definition allows to find most characteristic features which were mentioned above:

– multiplicity: sets E and E′
– composition and relation between levels: C composition law
– autonomy: relations on each level (R and R′).

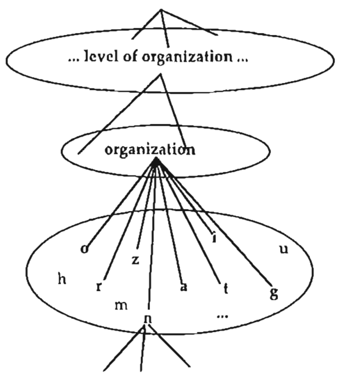


Figure 4.

As a counter-example, let E be the set of integers from 1 to 10, C the addition of 10, E′ be the set of integers from 11 to 20. If R′ is the order relation on E′2, R′ can be expressed through a R relation on E2: x′ R′ y′ if and only if exist x and y in E such that x′ = C(x,10) = x + 10, y′ = C(y,10) = y + 10 and x R y, R being the order relation on E2.

Of course, one major problem is to understand "... R′ cannot be deduced from R...". Gödel's theorem can serve as an analogy, to which we shall return in the last part of the paper. The analogy is the following:

– E set and E′/N integer set
– R relation/provable theorems on N
– R′ relation/true but not provable theorems.

Let us call N and N′ the organization levels ⟨E,R⟩ et ⟨E′,R′⟩, and e and e′ the corresponding elements. We are now able to come back to emergence and identify the organization levels corresponding to each example:

– dalmatian dog picture: e: point or patch (location, color), e′: dog (concept or image)

– ant bridge: e: ant (location, size), e′: bridge

– economy: e: transaction (purchase or sale, moment), e′: economic evolution, aggregate quantity.

II.4. Category theory

Ehresmann and Vanbremeersch (Ehresmann & Vanbremeersch, 1989) suggested an interesting formalism based on category theory, which they apply to a model of neural network, and can be mapped easily to our problem. Let us consider a set of objects with links between them ("morphisms"), and let us call it a "pattern". A "global link" of a pattern P to an object O consists of:

– the set of links from the objects of P to O;

– each f_{ij} link from O_i of P to O_j of P which is on a path from O_i to O, i.e. so that there exist links f_i and f_j satisfying: $f_i = f_{ij} \circ f_j$.

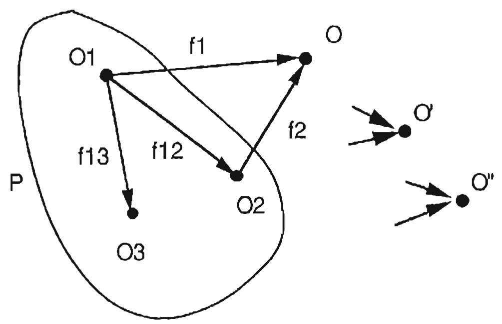


Figure 5.

On the previous figure, the global link is the set; $\{f_1, f_2, f_{12}\}$ (not f_{13} because there is no link between O_3 and O). An object O is a limit of a pattern P, if the links from O to any object O′ bijectively correspond to global links from P to O′. Thus, a limit is an object such that each operation on the initial pattern can be replaced by an operation on the limit. The limit is an object whose elements are the pattern objects, whose internal organization is composed of the pattern links. Besides, it may be given new links, which are not mediated by its elements. This latter aspect makes the connection with the major idea we mentioned about levels of organization: a relation (link) which cannot be expressed through basic element links. On the word-letter example, the letters are the elements of a pattern, the word is the limit pattern; the word may have links (syntactic or semantic relations) with other words.

Ehresmann and Vanbremeersch use this formalism to model neural networks. Objects are neurons; links are synaptic connections. A pattern is the set of neurons which are activated by a stimulus. A limit is associated with the pattern, through the following property: each (downstream) neuron will receive the same action potential from the limit as from the neurons of the pattern. Hofstadter (Hofstadter, 1979) also suggests an illustration of the notion of limit, in the algorithmic domain. He introduces a chunking process which consists in replacing part of a graph (the pattern) by a node (the limit) so that links with other elements of the graph are kept.

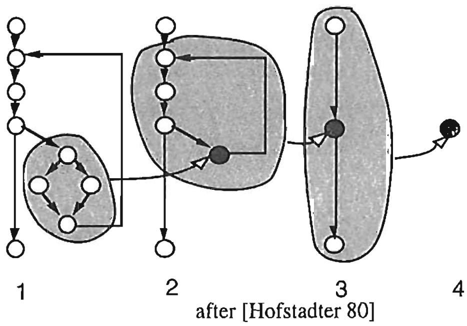


Figure 6.

On that example, graph 1 contains 9 nodes; a first chunking replaces a pattern of 4 nodes by one new node, while keeping relations with other nodes. And so on.

II.5. Section's conclusion

This study could be extended by including aspects such as intentionality: on the local level, agents act following personal goals. On the global level however, it is not easy to introduce intentionality,... except from an ecological or teleological point of view.

feature / examples	n1 level	n2 level	time scale	local perception	local action	global perception
ant bridge	ant	bridge	minute	ant	ant	colony
?	**human being**	?	?	**human being**	**human being**	**?**

Figure 7.

Finally, we have seen that a level of organization is defined by a set of elements and relations R between them (note that the composition law, that allow to generate higher-level elements, is at least compatible with, if not defined by R): these relations R depend on the observer's abilities to dectect them. Therefore, we have to understand how such relations (or almost equivalently higher-level objects) are detected. This is the topic of the next section.

III. DETECTING EMERGENCE

We have already emphasized the importance of the observer. In this section we propose a model of some cognitive processes that certainly take place in the observer's mind when she "feels" that something is emerging.

III.1. Changing levels of description

A comparison between some of the examples given in (Bonabeau *et al.*, 1995) shows that the emerging object is qualitatively different from its components, expressed in section 2 by the idea that R' is not deducible or derivable from R. Indeed we need to make a fundamental distinction between

example	**emerging object**	**component**	**new property**
spiral	4 arms spiral	straight segments	rounded curve
traffic jam	vehicle gathering	car	backward motion
temperature	group of molecules	molecule	heat
H_2S	molecule	atom	smell
particle	quantum particle	plane wave	localization
word	English word	phonem	reference
run for 20	complex rule	elementary rule	generality

Figure 8. Parallel between examples of emergence given in text. Last column gives a property of the emerging object that is not present at the component level.

two levels of description, since properties or operations that are defined at the constituent level cannot be applied as such to sets of constituents. Emergence corresponds to an activation of the higher level of description.

Levels of description appear as "situated": they depend on the observer, what she is currently looking at, and her current point of view. Emergence is also "situated". But let us first describe more precisely what a description level is.

Biological point of view		**Sociological point of view**	
level of description	*example*	*level of description*	*example*
...		...	
species	Pan troglodytes	ecosystem	the arboricultural system
individual	a given chimpanzee	society	groups of chimps frequently meeting
organ	the liver	group	a given group
cell	a liver cell	bound individuals	a mother with some of her offsprings
		individual	a given chimpanzee
macromolecules	hemoglobin	...	
simple molecules	a given aminoacid		
atoms	carbon		
...	...		

Figure 9. Two different description hierarchies to describe the same reality (living beings), but from two different points of view: biological and sociological. Many other hierarchies could be used (e.g. for a genetician, the "good" hierarchy would be base-pair/codon/gene/individual/kin/population/species).

III.2. Description levels as membership hierarchies

Description levels (not to be confused with organization levels discussed in the previous section) are modeled here as levels of a membership hierarchy: one element of level n + 1 is a set of elements of level n. For instance in figure 9, an individual is a set of organs from the biological point of view. Elements of a chimpanzee group, from the sociological point of view, are for instance a mother-offsprings clan, a group of cooperating males. Let us stress two points:

(i) Description levels are used when describing perceived reality. Describing english written words (in general) as made of 26 characters makes use of organization levels: a lower level with 26 elements, an upper level with 100 000 elements. A word like "rational" appears as a combination of seven different characters. But when reading the word "rational", the observer sees it at the word description level as a set of eight different actual characters. The two "a" are indeed two independenly perceived objects.

(ii) Membership is not transitive. In the above examples, a biologist would not consider a given species as a set of many organs; a sociologist would refuse to see a given group of chimpanzees as a mere set of individuals; for a linguist it would be a nonsense to consider a given sentence as a set of characters or a set of phonems.

The membership relation which links two adjacent description levels is fundamental, since it is a distinctive feature of a description hierarchy, as opposed to inclusion hierarchies (like those given by iterative clustering techniques), inheritance hierarchies (as in an object-oriented computer program), functional hierarchies (x is eaten by y which is eaten by z...), etc. The model presented here claims that emergence is a phenomenon which takes place between two description levels, i.e. two levels linked by a membership relation.

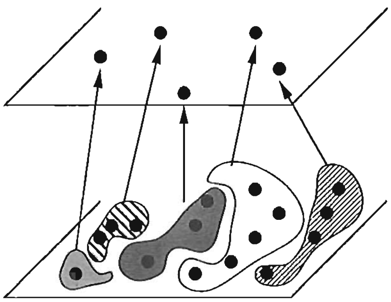


Figure 10. Description levels as defining a membership hierarchy.

We will consider that emergence occurs when an object is suddently perceived at the n+1 level when only objects of level n were perceived just before. But we must first explain how some of the n-objects perceived are grouped together to form a perceived (n+1)-object. Even if we restrict ourselves to partitions (in most examples, emerging objects are disjoint sets of lower level objects), we must explain how a given partition is chosen to form a perceived (n+1)-object among the B_k possible partitions, where B_k is Bell's number (see Favaron, 1990) given by:

$$B_k = \frac{1}{e} \sum_{i \geq 0} \frac{i^k}{i!}$$

This number rapidly reaches large values (there are 1.4 billion ways to partition 15 objects). Of course when a similarity relation is available at level n, then clustering techniques (similarity based categorization, cf. [de Amorim, 1990]) allow to find a partition which maximizes the sum of similarities of pairs of objects belonging to the same cluster (finding this partition is NP-hard).

Unfortunately, if there are N levels, a clustering technique requires $N-1$ similarity relations. Merely estimating similarities at level $n+1$ using similarities at level n (e.g. average similarity between sets) yields indeed inclusion hierarchies, not membership hierarchies. Moreover, it is not obvious that $(n+1)$-objects are perceived as similar because their constituents are similar, for any conceivable similarity relation, at level n.

III.3. Emergence and detection

We propose here that the structuration of a given perception into description levels is a direct consequence of the existence of a detection hierarchy. The observer must have accordingly many detectors organized into a detection hierarchy in order to be able to perceive reality in a hierarchical way. A detector of level n reports the occurrence of a n-object, so that

the sudden activity of a (n+1)-detector triggered by a set of activated n-detectors

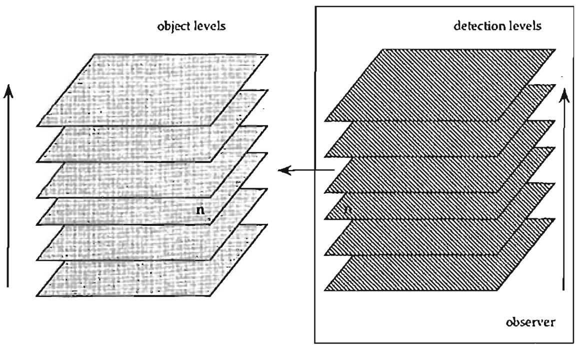


Figure 11. Description hierarchy as a consequence of the hierarchical structuration of detectors.

The concept of detector is used here with the meaning it has in communication theory. The presence of a detector is what makes the difference between analog and digital communications. In analog communication, the important thing is to make sure that the receiver gets a very accurate copy of the message sent by the source. Digital communication works on totally different grounds. The emitter and the receiver must first agree on a set of possible symbols and their physical representation (e.g. electric pulse), and any message then consists of a sequence of such symbols. Under certain hypotheses, the optimal detector computes correlations between the received physical pulses and an exemplar of each expected pulse, and then takes a decision (threshold comparison).

III.4. Examples

Neuronal detection hierarchy

The first example of detection hierarchy comes from situations where detectors are "hard-wired". One can imagine that a n-detector detects the presence or absence of the object to which it is tuned in the input space, each input being the output of a (n − 1)-detector.

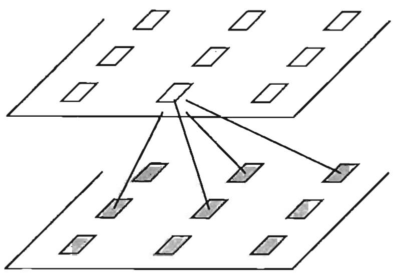


Figure 12. Hard-wired detector hierarchy: each detector receives inputs from the outputs of detectors belonging to the level just below. One can imagine that detectors are "AND" gates, or artificial neurons, or groups of such neurons.

The type of object O_n detected by a n-detector D_n can be characterized according to objects $\{O_{n-1}\}$ which are detected in the current environment by the detectors $\{D_{n-1}\}$ which send outputs towards D_n. D_n detects O_n if and only if some of the O_{n-1} are also detected by $(n-1)$-detectors. One can thus associate O_n to a subset of the present (i.e. detected) $\{O_{n-1}\}$. The hierarchy among objects perceived in the current environment is, as expected, a membership hierarchy. For instance if one thinks of neurons, thresholded weighted sums are not associative, and the link $\{O_{n-1}\} \rightarrow O_{n+1}$ that one could consider according to the network topology is meaningless (as expected with a membership hierarchy, and contrary to the transitivity in inclusion hierarchies where one level can be omitted without destroying the hierarchy). O_n is an object since it is detected (as soon as D_n is active), and it appears as composed of other objects which are present (= detected) in the current environment.

First levels of the visual system are perhaps organized according to a hard-wired detection hierarchy (detectors in this case are groups of neurons): pixel-detectors/on-off region detector/line detector... Admittedly higher levels of detection in the visual system are unlikely to be hard-wired this way. Gestalt theory taught us that we have "good shape detectors", e.g. those which are able to detect a traffic jam on an aerial photography or the logo spiral (cf. examples given above). But we can imagine that such shape-detectors are also connected to lower level hard-wired detectors in the way illustrated in the previous figure, even if these connections are not materialized in some simple axonal circuitry.

Language detection hierarchy

The linguistic capabilities of human beings, as they are modeled by several linguistic theories, offer another example of hierarchical detection. If we look at the so-called "double articulation" (Martinet, 1967): phonem/word and word/sentence, we come again to three levels: phonems, words, sentences. Each competent speaker has a set of word-detectors for the words of his/her language. One may think that some (phonemic or whatever) rules help him/her build these word-detectors, in real-time or during learning. The competent speaker has also grammar rules that help him/her elaborate sentence-detectors that detect well-formed sentences. A rare word may thus suddenly emerge from its detected phonemic constituents, or a complicated sentence may sometimes emerge from its detected words.

One may imagine that word-detectors are memorized and exist before a linguistic message is heard, and that they are "fed" by phonem-detectors. Our competence allows us also to detect some unusual or even unknown words, and to associate a syntactical category (e.g. "gromic" can be heard as an adjective, while "frogrorlo" will be refused as an English word).

Sentence detectors, on the other hand, seem to be built up in real time. In the generative grammar model, such a detector is instantiated by application of transformational rules which are compatible with present (i.e. detected by word-detectors) words. Word-detectors create a partition among detected phonems, and sentence-detectors create a partition among detected words. One can imagine that this membership hierarchy includes further levels, for instance the argumentation level where each argument can be recognized by an argument-detector which works on detected constituents (that we can represent using logical propositions) (Dessalles, 1990).

Conceptual hierarchies

Conceptual emergence may occur because concepts are organized locally in hierarchies. Let us give first an example. In the game "run for 20", each player can add 1 or 2 to the number obtained by the opponent. The winner reaches 20 first. A beginner soon discovers that reaching 18 is very bad, or that reaching 17 is a very good thing (whatever the opponent plays (adding 1 or 2), it is always possible to reach 20). Then (s)he will notice that reaching 14 is the best way to be sure to obtain 17. From such elementary rules a complex and powerful rule may suddenly emerge: "if I can reach a multiple of 3 minus 1, I'm able to win". Such a rule replaces a whole set of elementary rules. In a problem-solving situation like "run for 20", a conceptual object is relevant if it can be used as constituent in a solution. During an argumentation, a conceptual object is relevant if it can be included in an argument. Conceptual emergence occurs when a conceptual object of higher generality is suddenly recognized as relevant. Concepts can be locally organized according to their generality. Figure 13 shows how detected objets are accordingly layered in a membership hierarchy. In this schema, $\{Y_i\}$ should not be confused with the extension of the concept X_1. This extension would be a set of "basic" objects that are supposed to be given by perception. This point of view is often adopted in psychological theories of concept formation and typicality (Rosch, 1978). But generalization creates an inclusion hierarchy among extensions, not a membership hierarchy.

To explain conceptual emergence, we have to notice that conceptual objects may be locally organized into a membership hierarchy. Here the condition "By reaching 14" belongs to the condition "By reaching a multiple of 3 minus 1", in the sense given in figure 4. Emergence results from a jump towards the next higher level where the relevance of a more general conceptual object is suddenly perceived.

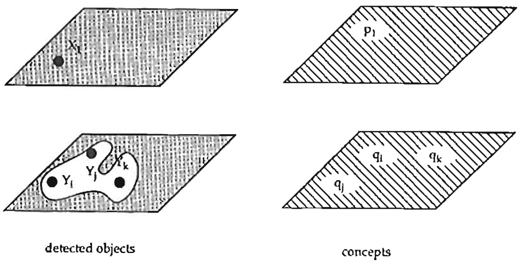


Figure 13. Concepts as detectors hierarchically organized according to their generality. Objects are defined by the concepts which detect them. The object Y_i, detected by qi, belongs to the object X_l detected par p_l, if q_i is equivalent to p_l & e_i, where e_i is an "elementary" concept, i.e. e_i cannot be split by conjonction in the current context.

IV. A FRAMEWORK FOR CHARACTERIZING EMERGENT PHENOMENA

IV.1. Introduction

In this section we propose to build a tentative framework for describing emergent phenomena. The central notion is that emergence is a dynamic process through which some quantity [complexity] is rapidly/dramatically varying with respect to the time constant/the spatial granularity, or more generally to the model or the level of description used by the observer. All definitions have to be conditional, relative to some set of tools available. These tools are mostly detectors, that enable the observer to detect "regularities", or special shapes we shall call pregnant. Emergence can then be defined relative to an equivalence class of observers having certain detecting abilities in common. The generic properties characterizing the equivalence class of observers may either be well formalized (symmetry groups, universal Turing machines, languages from Chomsky's hierarchy,...) or less well-defined (perceptual and cognitive capacities).

IV.2. Complexity and organization

The notion of complexity is essential in the process of characterizing emergence, because emergence can be considered as an increase in the complexity of a system (in fact, whether it corresponds to an increase or to a decrease somewhat depends on the point of view adopted).

IV.2.1. Intuitions of complexity

The first idea that comes to mind when speaking of complexity is that of interrelatedness: many interacting elements give rise to unpredictable but structured behavioral patterns. In that sense complexity has something to do with organization. But this is not necessarily the only idea. Purely random processes may also look very complex if one tries to describe them with a great degree of "deterministic" precision, while they turn out to be quite simple in probabilistic terms. One important conclusion to draw from these simple considerations is that complexity is a relative concept: it is to be defined with respect to a set of tools (feature detecting tools and modeling tools). Some tools are not always appropriate: for example, probabilities are of interest when it is harmless to smooth individual features, while they are not adequate when one wants to account for strong individual deviations (such deviations from statistical laws can be crucial for the understanding of the phenomena under observation) (Kampis, 1991). If one accepts the fact that (seemingly) purely random processes are easily modeled (as easily as completely ordered processes), then the classical concept of complexity holds: complexity lies somewhere in the middle of the entropic scale.

IV.2.2. Formal measures of complexity

Here we review very briefly some formal measures of complexity (Bennett, 1990; Chaitin, 1979; Crutchfield & Young, 1990; Grassberger, 1989; Kampis, 1991; Langton, 1990). One candidate is Gibbs-Boltzmann-Shannon (GBS) entropy, which measures the lack of information of an observer about a system. It more or less measures the disorder associated with the system (or rather with a model of it). In Kolmogorov-Chaitin-Solomonoff (KCS) information theory, the complexity K(D) of a string D of symbols (or of some pattern) is defined as the size of its minimal Universal Turing Machine (UTM) representation. Thus, a (purely random) string which is its own best representation is of maximal K(D), while a highly structured string (111111111 for instance) has minimal K(D). KCS Complexity can be expressed in a more general form: $C(D)=INF_{(S)}\ C(D/S)$, S is a symmetry, and C(D/S) designates the conditional complexity of D, which is the amount of information in equivalence classes induced by the symmetry S in D, plus the amount of data that is unexplained by S (Crutchfield, 1990). Using this notation, Kolmogorov complexity is defined by $K(D)=C(D/\{UTM\})$, where {UTM} denotes the set of all symmetries computable by a UTM. One can as well define C(D)=C(D/{BTM}), where {BTM} designates the set of all symmetries computable by a BTM (a Bernouilli Turing Machine, which is a Turing machine with a random register, so that both purely random and fully predictable processes are easily described) (Crutchfield & Young, 1990).

Mutual information quantifies the information processes going on between two elements A and B: intuitively, there is mutual information between A and B if A and B are able to affect one another's behavior. Mutual information is defined as MI(A,B)=H(A)+H(B)−H(A,B), where H(A) is the entropy of A, and H(A,B) is the entropy of the joint process (A,B). Mutual information can be measured between two elements, two sets of elements, or on a given element at different time steps.

We could continue forever this review of complexity. All definitions share common features and some differences, but what we must remember is that complexity is obviously a relative concept, which depends both on the task at hand, and on the tools available to achieve this task. Conditional complexity is a powerful tool for defining a general notion, that of relative complexity: C(D/S) can be interpreted as "the difficulty of decomposing D when S is used as a structuring element ($\Leftrightarrow$ what is explainable by S) + what is not decomposable under this structuring element ($\Leftrightarrow$ what is not explainable by S and remains to be explained otherwise)". Given a set of tools considered as structuring elements, there are some aspects of the object that can be explained (or compressed) through these tools, and there are some aspects of the object that cannot be understood using these tools. In order to explain the latter aspects, other tools may be required. We will call these tools detectors. For instance, a UTM represents an enormous set of detectors: the set of all detectors that are capable of recognizing a computable symmetry. In this context, we can see that the complexity of an object is roughly the number of classes of features it contains (given a set of observables).

IV.2.3. Levels of organization

Mutual algorithmic complexity, or mutual complexity in a more general form, constitutes a good means of defining levels of organization. This is largely based on a work by Chaitin (Chaitin, 1979), who proposed a characterization of organized systems by introducing K_d-complexity.

$$K_d(S) = \mathrm{Min}_\beta\,[KCS(\mu) + \Sigma_{i \leq n(\beta)} KCS(S_i)]$$

where β is a partition of the system S into parts $S_1, \ldots, S_{n(\beta)}$ so that for any i, $D(Si) \leq d$, D(X) being the diameter of X. μ is a procedure of reconstruction of the systems out of its parts, and KCS(μ) is the algorithmic cost of that reconstruction: $K(\mu)=K(S \mid \Pi S_i)+K[n(\beta)]$, if one assumes that $|\mu|=\min\{|p|:U(p[S_o*...*S_{n(\beta)}])=S\}$.

What does k_d-complexity represent? For $d \geq D(S)$, one obviously has $Kd(S)=KCS(S)$: in effect, the best partition – the one that minimizes the $\Sigma K(Si)$ part of Kd-complexity – consists in taking only one part, the whole system. It comes directly from the fact that $H(\langle X,Y\rangle)\leq H(X)+H(Y)+O(1)$. Thus $\{S\}$ is the optimal partition. When the diameter of "observation" d is decreased, $K_d(S)$ may increase dramatically if S is structured. Each time one encounters a new "critical diameter" under which it is no longer possible to find a "good partition" because some partitions are no longer accepted, this corresponds to the appearance of a structure. When the diameter is decreased, one thus needs to "break" some structures into several parts sharing a lot of mutual information. Of course, critical diameters need not be localized, since a given structure can cover a certain spectrum of diameters.

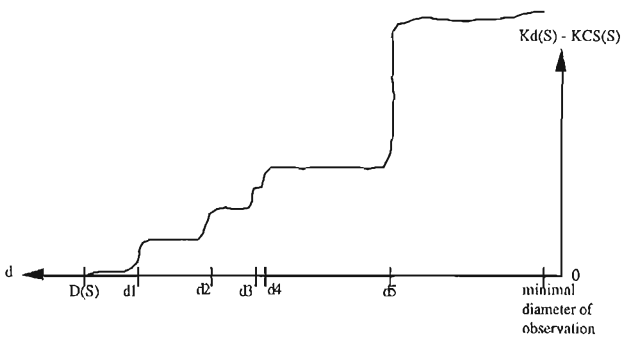


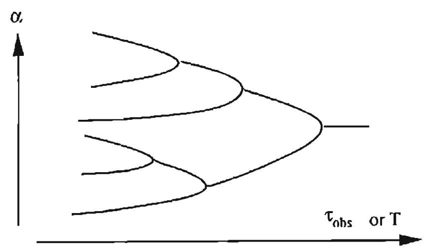
Figure 14. In this picture, d_1, d_2, d_3 and d_4 are critical diameters, but they are very close – one can imagine that this corresponds to a loosely localized type of structure –, while d_5 is a highly critical diameter, clearly defining one type of structure.

It has been shown by Chaitin that the degree of algorithmic independence between n variables $X_1, ..., X_n$, was well estimated by $\Sigma K(Xi)-K(\langle \Pi Xi\rangle)$, that is, if one neglects the cost of the procedure of reconstruction μ, $K_d(S)$ represents the mutual information between $S_1, ..., S_n$: the optimal partition is the one that allows for minimal mutual information between its parts. One can also modify K_d-complexity so as to make it iterable, that is all partitions for a diameter d are chosen among the under-partitions of the optimal $(d+\Delta d)$-partition. It is interesting to remark that the fact that $K_d(S)$ diminishes when the diameter is increased does not mean that the average mutual complexity per partition $K_d(S)/|\beta|$ also diminishes.

IV.3. Emergence

IV.3.1. Emergence and phase space reduction

An emergent phenomenon implies that a given indistinguishable set of microstates transforms into several distinguishable macrostates: macroscopic entropy increases, while microscopic entropy decreases, since microstates that were previously associated with the same microstate are now distinguished. There has also been a reduction of the space of possible microstates associated with macrostates, that is not all microstates are permitted for one particular macrostate. This reduction is detected by the observer for whom there has been a *broken ergodicity.* Broken ergodicity is more general than broken symmetry, where a system fails to show the symmetry of its underlying hamiltonian (Palmer, 1989). It occurs when a system undergoes a transition which prevents it from exploring all the states of its phase space (Reichl, 1980; Palmer, 1989). In an emergent phenomenon, the observer changes some parameter of observation, such as the time scale or the spatial granularity, and becomes aware of more and more details of the structures of the system. That is, there is an apparent broken ergodicity, leading to the appearance of components which were not observed at another scale. This is so if the scale is finer and finer.



A hierarchy of components developing as temperature or observation time is lowered

Figure 15.

Let us for instance consider a traffic-jam: the emergent structure comes from the fact that cars cannot drive freely wherever they want due to constraints imposed by other cars and by the environment (roads). There is in this case an "observational broken ergodicity": all behaviors are possible a priori, while observed behaviors are much more specific.

IV.3.2. Detectors

From the previous paragraphs, we understand that elements of a system can be grouped so as to facilitate the explanation of the system's behavior at the higher level. This also corresponds to a decrease in the relative complexity C(Syst/{Det}) where {Det} is the set of available detectors. At time t

$$C_t(S/D_1, \ldots, D_n) = Ct(S/D_1, \ldots, D_{n-1})$$

and D_n is activated at time $t+\Delta t$

$$C_{t+\Delta t}(S/D_1, \ldots, D_n) < C_{t+\Delta t}(S/D_1, \ldots, D_{n-1}) = C_t$$

The transition occurs within Δt because detectors are usually "treshold devices". For instance, if an object contains long range symmetries, it might take a long time to detect these symmetries, but once a symmetry is about to be detected, the transition is sharp.

Detectors are sensitive to certain relationships present between elements of the system under observation. The passage from level N to level N+1 corresponds to a reduction of the space of possible behaviors due to mutual constraints between the elements: it is these remaining behaviors (or the associated constraints) that are detected, and thus they are "physically pregnant" in Thom's terminology (Thom, 1980). They become biologically pregnant when they evoke shapes which are especially important for the survival of the (biological) system: "*Mind and world in short have evolved together, and in consequence are something of a mutual fit. (...) That is to say that our various ways of feeling and thinking have grown to be what they are because of their utility in shaping our reactions on the outer world*" (James, 1879).

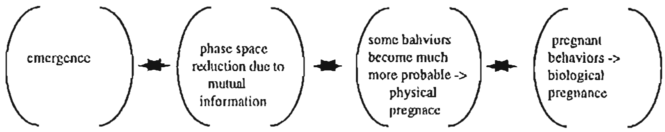


Figure 16.

IV.3.3. "Non-derivable" properties

It is worth returning to the notion of "derivation" mentioned in the first part of this paper: what does it mean that one cannot "derive" the properties of a system at a particular level given a model of how a lower level behaves?

There are several possible (and not incompatible) answers to this question. For the sake of simplicity, we describe these answers in a formal setting: all processes we are talking about are computational. We shall see that the concept of detector is of importance.

– We can understand the notion of derivation in the context of Gödel's theorem. In this case, derivation is something very formal, it designates a way of proving new theorems from a set of axioms. Not being able to derive observed higher-level properties from the set of lower-level properties means not being able to prove theorems about higher-level entities while we know that these theorems are true. We know since Gödel that such a situation is possible. The only way of knowing the properties of the higher-level entities is to *detect* them, to be *sensitive* to them, since it is impossible, even "in principle" to deduce them.

– Another cause for the impossibility to derive the higher-level properties from lower-level considerations is the time – and possibly the "memory" space – it would take to examine all the possible properties of the higher-level and determine those that are of interest. Time compression is allowed by computers, and in that sense we may speak of computational emergence since computers enable us to observe higher-level properties, indeed contained in the specification of the system from the beginning, but which are out of reach for a human being, because the space of possible properties is too huge. Computers might even not be sufficient, if the number of dimensions of the space of possibilities is "immense" in Elsasser's sense (Elsasser, 1981), i.e. a number that is "not tractable and cannot be acted upon with present-day computers" (Kampis, 1991). These properties may be logically deep: it might take a very, very long time even with the best program ever – not to speak of the difficulty of finding this program – to compute these properties. In this case, the only way of discovering a new "deep" property is to have a detector which is at least as deep as the property, and having a high value of mutual information with that property due to a common history.

– Putnam made a distinction between "to deduce" and "to explain" (Putnam, 1973). Being able to deduce the properties of a phenomenon from a set of causes is not equivalent to explaining this phenomenon, because only a few among the many possible causes may be relevant, "*certain systems can have behaviors to which their microstructure is largely irrelevant*" (Putnam, 1973). Explaining the phenomenon amounts to determining what the relevant causes are. Second, explanation is, according to him, not transitive, because explanations at one level are not of the same nature as explanations at another level. But worse, it may simply be impossible to deduce the properties of a phenomenon from a set of causes originating from one single discipline: this is so because "*the laws of the higher-level discipline are deducible from the laws of the lower-level discipline together with "auxiliary hypotheses" which are accidental from the point of view of the lower-level discipline*" (Putnam, 1973). The laws of the higher-level discipline therefore depend on both the laws of the lower-level discipline and "boundary conditions" which are "*accidental from the point of view of physics but essential to the description of the higher level*". It is through the huge space of possibilities allowed by physics that higher-level phenomena are somewhat autonomous relative to the laws of physics. What is essential is that explaining a phenomenon is not always possible, and the only way of having an idea about this phenomenon is not by explanation but by detection.

IV.3.4. Emergence relative to a model

It is now perfectly possible to understand "emergence relative to a model" (ERM) within the present framework. ERM corresponds to a shift in point of view. At a given level of description, ERM is associated with an increase of C(Syst/{Tools}), where Syst is the system under observation and {Tools} a set of tools available to the observer. The system apparently increases the dimension of its phase space and its behavior apparently becomes "random": less of the behavior of the system is understood, more of it becomes unexplained. The meaning of {Tools} can be twofold:

- it can represent a set of detectors, in which case ERM $\Leftrightarrow$ (increase of C(Syst/{Delt}),
- or it can represent a model of the system, and then ERM $\Leftrightarrow$ (increase of C(Syst/[P/{Det}])), where P/{Det} is a minimal program generating D with a particular input, using only allowed symmetries (detectable with {Det}). More clearly, in this case, P is the program that explains the system's behavior, and therefore if ERM occurs one will have to change the input for a longer one, and it might be sufficient to modify [P/{Det}] $\rightarrow$ [P′/{Det}] in order to lower C(Syst/[P/{Det}]); changing the model without changing the set of detectors (and thus the set of observables) is sufficient. This is similar to Cariani's syntactic emergence. If {Tools} = {Det}, then the initial level of complexity will never be reached again if no detector is added: ERM reflects the need for a new detector to be used in order to build more efficient models.

From these considerations, it is easy to understand how ERM is different from other conceptions of emergence, and why it corresponds to a shift in point of view: in effect, ERM sees emergence in cases where it becomes more difficult to model the system, while most other conceptions of emergence state that emergence leads on the contrary to a reduction in the modeling difficulty. To understand the difference it may help to see that such conceptions simply apply to the observer (assuming the system does not modify its behavior), while ERM applies to the system (whose behavior's modification requires a new model).

V. CONCLUSION

In conclusion, we have presented a framework, based on levels of organization, levels of detection and theories of complexity, to integrate most conceptions of emergence. We have emphasized the importance of the observer, and shown how, depending on the point of view, emergence corresponds to an increase or a decrease of complexity in a system's model. This framework is far from sufficient, since it is mostly descriptive, and gives only few prescriptions to achieve emergence: but we believe it constitutes a first step towards this goal.

References

H. ATLAN, *L'organisation biologique et la théorie de l'information*, Paris, Ed. Hermann, 1972.

H. ATLAN, *Self-Creation of Meaning*, *Physica Scripta*, 1987, *36*, pp. 563-576.

C.H. BENNETT, *How to Define Complexity in Physics, and Why*, in *Complexity, Entropy and the Physics of Information*, ed. W.H. Zurek, SFI, Addison-Wesley, 1990.

E. BONABEAU, J.-L. DESSALLES, A. GRUMBACH, Characterizing emergent phenomena (1): A critical review, *Rev. Int. Syst.*, 1995, *9*, 3, pp. - .

J.-P. BRONCKART, *Les théories du langage*, P. Mardaga ed., Dessart, Brussels 1977.

P. CARIANI, *On the design of devices with emergent semantic functions*, Ph.D. dissertation, State University of New York at Binghamton, 1989.

G.J. CHAITIN, *Toward a mathematical definition of life*, in *The maximum entropy formalism*, MIT Press, 1979.

C. CHAPPELIER, R. MUNOS, A. GRUMBACH, *Réseau connexionniste à structure aléatoire*, Mémoire d'option, ENST, 1991.

J. CRUTCHFIELD, K. YOUNG, *Computation at the onset of chaos*, *Complexity, Entropy and the physics of Information, SFI studies in the Sciences of Complexity*, vol. III, ed. W. Zurek, Addison-Wesley, 1990.

S.G. DE AMORIM, J.-P. BARTHÉLÉMY, C.C. RIBEIRO, *Clustering and Clique Partitioning: Simulated Annealing and Tabu Search Approaches*, *Les cahiers du Gerad*, 1990, G-90-35.

J.-L. DESSALLES, *The simulation of conversations*, in KOHONEN & FOGELMAN-SOULIÉ, *Cognitiva 90*, Elsevier, pp. 303-310, 1990.

J.-L. DESSALLES, *La perception de l'émergence*, Proceedings of the Workshop on emergence in cognitive science, ENST, Paris 1992.

A. EHRESMANN, J.P. VANBREMEERSCH, *Systèmes hiérarchiques évolutifs à mémoire auto-régulée*, Seminar on Synergie et Cohérence dans les systèmes biologiques, Paris, Jussieu, 1989.

W. ELSASSER, *Principles of a new biological theory: A summary*, *J. Theor. Biol.*, 1981, *89*, pp. 131-150.

O. FAVARON, M. MAHEO, I. FOURNIER, *Combinatoire et Algorithmique*, Orsay Plus, UPS, 1990.

P. GRASSBERGER, *Problems in Quantifying Self-Organized Complexity*, *Helvetica Physica Acta*, 1989, *62*, pp. 498-511.

D. HOFSTADTER, *Gödel, Escher, Bach: An eternal Golden Braid*, Basic Books, 1979.

W. JAMES, *Psychology. The briefer course*, Gordon Allport, ed., University of Notre-Dame Press, Notre Dame, Indiana, 1985.

G. KAMPIS, *Self-Modifying Systems in Biology and Cognitive Science*, Oxford, Pergamon Press, 1991.

B.-O. KÜPPERS, *Information and the origin of life*, MIT Press, 1991.

R. LAWLER, *Extending a Powerful Idea*, A.I. Memo 590, MIT AI Lab., 1980.

P.H. LINDSAY, D.A. NORMAN, *Traitement de l'information et comportement humain*, Vigot, 1980.

D. MEDIN, *Concepts and Conceptual Structure*, *Am. Psychologist*, 1989, *44*, 12, pp. 1469-1481.

G. NICOLIS, I. PRIGOGINE, *Exploring Complexity: An Introduction*, R. Piper, GmbH & Co. KG Verlag, 1989.

R. PALMER, *Broken ergodicity*, in *Complex Systems, SFI Studies in the Sciences of Complexity*, Ed. D. Stein, Addison-Wesley Logman Publishing Group Ltd., 1989.

H. PUTNAM, *Reductionism and the nature of psychology*, *Cognition*, 1973, *2*, pp. 131-146.

L.E. REICHL, *A modern course in statistical physics*, University of Texas Press, Austin, 1980.

E. ROSCH, *Principles of Categorization, in Rosch & Loyd, Cognition and Categorization*, Hillsdale, 1978.

R. THOM, *Modèles mathématiques de la morphogenèse*, Christian Bourgois, 1980.