



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

C. R. Biologies 326 (2003) 133–140



Biological modelling / Biomodélisation

# What do we mean by biological complexity? Qu'entendons-nous par complexité biologique ?

Jacques Ricard

*Institut Jacques-Monod, CNRS, université Paris-7, 2, place Jussieu, 75251 Paris cedex 05, France*

Received 15 July 2002; accepted 9 December 2002

Presented by Jacques Ricard

---

## Abstract

The purpose of the present paper is to offer a precise definition of the concepts of integration, emergence and complexity in biological networks through the use of the information theory. If two distinct properties of a network are expressed by two discrete variables, the classical subadditivity principle of Shannon's information theory applies when all the nodes of the network are associated with these properties. If not, the subadditivity principle may not apply. This situation is often to be encountered with enzyme and metabolic networks, for some nodes may well not be associated with these two properties. This is precisely what is occurring with an enzyme that binds randomly its two substrates. This situation implies that an enzyme, or a metabolic network, may display a joint entropy equal, smaller, or larger than the corresponding sum of individual entropies of component sub-systems. In the first case, the collective properties of the network can be reduced to the individual properties of its components. Moreover, the network is devoid of any information. In the second case, the system displays integration effects, behaves as a coherent whole, and has positive information. But if the joint entropy of the network is smaller than the sum of the individual entropies of its components, then the system has emergent collective properties and can be considered complex. Moreover, under these conditions, its information is negative. The extent of negative information is enhanced if the enzyme, or the metabolic network, is far away from equilibrium. **To cite this article: J. Ricard, C. R. Biologies 326 (2003).**

© 2003 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

## Résumé

Le but de cet article est de présenter une définition précise des notions d'intégration, d'émergence et de complexité dans les réseaux biologiques, à l'aide de la théorie de l'information. Dans sa forme originelle, toutefois, la théorie de Shannon est une théorie de la communication qui s'établit entre la source et la destination de messages. Une telle théorie doit être modifiée et adaptée pour rendre compte des phénomènes d'intégration et d'émergence apparaissant dans les réseaux biologiques, notamment les réseaux métaboliques. Si deux variables discontinues décrivent deux propriétés d'un réseau à l'état de quasi-équilibre, le principe de sub-additivité, qui est central pour la théorie classique de l'information, s'applique seulement si tous les nœuds du réseau sont associés à ces variables. Or, ces conditions ne sont, en général, pas satisfaites dans le cas des réseaux biologiques, car certains nœuds peuvent n'être associés à aucune, ou à une seule des deux variables. C'est, par exemple, ce qui se passe pour une réaction enzymatique à deux substrats, où la protéine fixe de manière aléatoire chacun de ces deux substrats. Dans le milieu réactionnel coexistent alors des molécules d'enzyme qui ont fixé un substrat, deux substrats, ou aucun substrat. Dans le cas où le principe de sub-additivité n'est pas une condition nécessaire, l'entropie jointe du réseau peut être

---

*E-mail address:* [meunier@ijm.jussieu.fr](mailto:meunier@ijm.jussieu.fr) (J. Ricard).

égale, inférieure ou supérieure à la somme des entropies individuelles des composants du système global. Si l'on se trouve dans le premier de ces trois cas, les propriétés du réseau peuvent être déduites de celles de ses composants, et ce réseau ne possède aucune information. Si l'on se trouve, en revanche, dans le deuxième cas, le système considéré présente un phénomène d'intégration de ses composants en un tout cohérent. Il possède, de plus, une information dont le signe est positif. Si, enfin, le réseau possède une entropie jointe supérieure à la somme des entropies individuelles de ses composants, alors il peut être considéré comme complexe et présente des propriétés émergentes par rapport à celles de ses composants. L'information du réseau possède alors un signe négatif. De nombreuses réactions enzymatiques, ainsi que des réseaux métaboliques, entrent certainement dans ces deux dernières catégories. Si le réseau n'existe pas dans un état de quasi-équilibre, mais dans des conditions d'état stationnaire, c'est-à-dire s'il se comporte comme un système thermodynamique ouvert, alors le flux de matière qui parcourt le système crée des propriétés nouvelles. Si l'on fait évoluer le réseau, de son état initial de quasi-équilibre vers un nouvel état stationnaire, en augmentant la constante catalytique de la réaction, l'information du système devient de plus en plus fortement négative. Cette situation est observée même si, dans l'état initial de quasi-équilibre, le système ne possédait aucune information. L'évolution d'un système enzymatique vers un état de non-équilibre engendre donc la complexité. La notion de complexité est ainsi totalement distincte de celle de complication. Un système apparemment peu compliqué, car formé d'un petit nombre de composants, peut donc posséder un comportement complexe et, réciproquement, un système compliqué peut ne pas posséder un tel comportement. **Pour citer cet article : J. Ricard, C. R. Biologies 326 (2003).**

© 2003 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS. Tous droits réservés.

*Keywords:* complex networks; emergence; information

*Mots-clés :* réseaux complexes ; émergence ; information

## 1. Introduction

Classical Science, that is the scientific activities that have sprung up since the days of Descartes, Newton and Leibniz, has always been in search for simplicity. More explicitly, it is based upon four 'principles' expressed by Descartes in the *Discours de la Méthode* [1]. These 'principles' are expressed below.

*Le premier était de ne recevoir jamais aucune chose pour vraie, que je ne la connusse évidemment être telle : c'est-à-dire d'éviter soigneusement la précipitation et la prévention ; et de ne comprendre rien de plus en mes jugements, que ce qui se présenterait si clairement et si distinctement à mon esprit, que je n'eusse aucune occasion de le mettre en doute.*

*Le second, de diviser chacune des difficultés que j'examinerais, en autant de parcelles qu'il se pourrait, et qu'il serait requis pour les mieux résoudre.*

*Le troisième, de conduire par ordre mes pensées, en commençant par les objets les plus simples et les plus aisés à connaître, pour monter peu à peu, comme par degrés, jusques à la connaissance des plus composés ; et supposant même de l'ordre entre ceux qui ne se précèdent point naturellement les uns les autres.*

*Et le dernier, de faire partout des dénombrements si entiers, et des revues si générales, que je fusse assuré de ne rien omettre.*

The second 'principle' is no doubt reductionist in its essence and, for this reason, adverse to present studies on complex systems. In spite of the fact that these analytic and reductionist approaches have led to important discoveries, there is little doubt that the World is complex in its essence and cannot be correctly understood through the classical reductionist approach. This is precisely what has recently become evident in different fields of Science ranging from fundamental Physics to Social Sciences [2–8]. It is clear that the living systems on Earth are complex systems and that it is hopeless to try to understand "what is life" through the sole study of the macromolecules that make up the substance of these systems.

The aim of the present contribution is threefold:

- first, to present a tentative definition of reduction, integration, emergence and complexity;
- second, to give a brief overview of the main features of complex systems;
- third, to discuss more thoroughly the most important of these features, the one which is at the very basis of the concept of complexity, namely information.

Although this discussion will be general, it is more specifically oriented towards dynamic biological systems.

## 2. A tentative definition of reduction, integration, emergence and complexity

From a philosophical viewpoint, the term of reduction can, at least, have two different meanings. It may refer to the mental process of deriving a scientific theory from another one, more general and embracing [9,10]. In this perspective, a biological theory, for instance, could be reduced to a more general physical one. If such a reduction could be pursued ad infinitum, this would imply unity of Science.

But there is a second type of reduction that is directly related to the problem of emergence and complexity. Let us consider a system made up, as any system, of a number of sub-systems. One can expect three types of situations to occur. The overall system and the set of the isolated component sub-systems have the same degrees of freedom, or the same entropy, or the same properties. Then, the overall system is not a real one, but the simple union of the two component sub-sets. The second possible situation occurs when the overall system has less degrees of freedom, or smaller entropy, than the set of the component sub-systems. It therefore behaves as a real system, for it displays some sort of integration of its elements as a coherent whole. The integrated system therefore displays collective properties distinct from those of the isolated component sub-systems. The last situation is observed if the overall system has more degrees of freedom, or more entropy, than the set of the component sub-systems. One may expect this system to have more wealthy collective properties than the integrated system considered above. It can be defined as an emergent, or complex, system. It results from this definition that neither the properties of an integrated system, nor those of a complex system can be reduced to the properties of their component sub-systems.

As the concepts of integration and emergence are related to that of information, it is thus clear that the analysis and discussion of this concept should be central in any discussion on complexity.

## 3. Main features of complex systems

Before discussing more thoroughly the logical foundations of the concepts of information, integration and emergence, it is of interest to present briefly the main features of complex systems [7]:

- a complex system should possess information; this important matter will be discussed later on;
- a complex system exhibits a certain degree of order; it is neither strictly ordered, nor fully disordered;
- complex systems should display collective properties, that are properties different from those of the component sub-systems;
- a complex system is usually not in thermodynamic equilibrium; it is thermodynamically open and displays non-linear effects;
- a complex system is in a dynamic state and has a history; this means that the present behaviour of the system is in part determined by its past behaviour;
- a complex system has emergent collective properties.

As far as we know, it seems that properties 1, 2 and 6 are mandatory for the definition of a complex system.

## 4. Information, integration and emergence in biological complex networks

Total information of the living cell is often identified to genetic information. This is too restrictive a view of the concept of information. As a matter of fact, one may expect most biological networks to store information for they associate in space and time, according to a strict order, many chemical reagents. Classical Shannon's communication theory [11–14], however, is not ideally suited for describing what network information is. It has therefore to be revisited and altered in order to allow this study. Moreover, many papers have recently been devoted to networks, in particular to metabolic networks [15–19]. Before studying integration and emergence, one has to present a brief overview of what a biological network is.

#### 4.1. Biological networks

A network is a set of nodes connected according to a certain topology. In a metabolic network, for instance, the nodes are the various metabolites, or the enzyme–metabolite(s) complexes, and the edges that connect the nodes the corresponding chemical reaction steps. Networks belong to different types known as random, regular, small-world and scale-free [15]. In random graphs, the nodes are mutually connected in a random manner according to a Poisson distribution, whereas in regular networks, node connection is effected through a fixed topological rule. Small-world networks display a fuzzy topology, midway between pure randomness and strict regularity. Sale-free graphs display both poorly and highly connected nodes.

A network is characterized by a parameter called diameter. If we consider all the possible pairs of nodes in a graph and the shortest distance, that is the smallest number of steps within a pair, the network diameter is the mean of the shortest distances. The immediate consequence of this definition is the increase of the diameter of a random graph as the number of nodes increases.

Metabolic networks possess an interesting property that is worth discussing briefly here [18]. Sequencing genomes of very different organisms, ranging from mycoplasmas and bacteria to man, have allowed us to know most of the enzymes present in these organisms, and therefore most of the reactions they catalyse. As a consequence, this allows us to construct, in principle, the metabolic networks of these living systems. Depending on they are primitive or evolved, the number of nodes of their metabolic graphs are very different. Still, the graph diameter remains nearly constant along the phylogenetic tree. The reason for this constancy is that metabolic networks are not random but possess a fuzzy organization. They are of the small-world type, and more and more connected as the organism is more evolved [18,19].

#### 4.2. Subadditivity principle, reduction and integration

Let us consider a composite dynamic network made up of four classes of nodes. A first class, X, collects nodes associated with a discrete variable  $x_i$ . A second class, Y, describes a different property associated with

a different discontinuous variable  $y_j$ . A third one, XY, has both properties  $x_i$  and  $y_j$ , and a last one has none of them. Each node of X and Y has a certain probability of occurrence,  $p(x_i)$  and  $p(y_j)$ , respectively. Similarly, each node of the class XY has a probability  $p(x_i, y_j)$ . If the events  $x_i$  and  $y_j$  are independent, then:

$$p(x_i, y_j) = p(x_i)p(y_j) \quad (1)$$

But if the events  $x_i$  and  $y_j$  are correlated, then:

$$p(x_i, y_j) = p(x_i)p(y_j/x_i) = p(y_j)p(x_i/y_j) \quad (2)$$

which is the classical Bayes' relationship. Here  $p(y_j/x_i)$  and  $p(x_i/y_j)$  are conditional probabilities, that is the probability of  $y_j$  given the value  $x_i$  and the probability of  $x_i$  given the value  $y_j$ , respectively.

The concept of information is related to that of uncertainty. The larger the uncertainty of a message and the larger is its information content. Put in other words, the information of a system is its ability to perform a difficult, and therefore improbable, task. The higher this difficulty, and the larger is the information required to perform the task. Therefore, the essence of information should be related to that of uncertainty. Any mathematical function aimed at measuring the degree of uncertainty should meet two axioms: that of monotonicity and that of additivity. Monotonicity means that the uncertainty should increase regularly with the number of states  $x_i$  or  $y_j$ . Additivity expresses the view that if the discrete variables X and Y are independent, the uncertainty of the couples XY should be equal to the sum of the uncertainties of X and Y. The only simple function of a probability  $p_i$  that meets these two requirements is:

$$f_i = \log\left(\frac{1}{p_i}\right) \quad (3)$$

The corresponding mean uncertainty function  $H$ , of probability  $p_i$ , is therefore:

$$H = \sum_i p_i f_i = - \sum_i p_i \log p_i \quad (4)$$

One can thus define mean uncertainty functions (or entropies expressed in bits) of X and Y as

$$\begin{aligned} H(X) &= - \sum_i p(x_i) \log_2 p(x_i) \\ H(Y) &= - \sum_j p(y_j) \log_2 p(y_j) \end{aligned} \quad (5)$$

Similarly, one can also derive the expression of the mean uncertainty function of the pairs XY (or mean joint entropy), namely

$$H(X, Y) = - \sum_i \sum_j p(x_i, y_j) \log_2 p(x_i, y_j) \quad (6)$$

Simple inspection of equations (5) and (6) shows that if the variables X and Y are independent, then:

$$H(X, Y) = H(X) + H(Y) \quad (7)$$

This equation shows that the entropies, or the degrees of freedom, of the system XY and of the sum of the sub-systems X and Y are the same. In line with what has been expressed in Section 1 of the present paper, equation (7) therefore implies that the system XY can be reduced to its component sub-systems X and Y.

Now, if the discrete variables X and Y are correlated and if the sum of the probabilities,  $p(x_i)$ ,  $p(y_j)$ , and  $p(x_i, y_j)$ , are all equal to unity, then one can demonstrate that the mean value of the conditional probabilities is larger than the mean of the corresponding probabilities. This implies that:

$$\langle \log_2 p(x/y) \rangle > \langle \log_2 p(x) \rangle \quad (8)$$

and

$$\langle \log_2 p(y/x) \rangle > \langle \log_2 p(y) \rangle \quad (9)$$

and this leads to:

$$H(X, Y) < H(X) + H(Y) \quad (10)$$

Expressions (7) and (10) define the so-called subadditivity principle, which read:

$$H(X, Y) \leq H(X) + H(Y) \quad (11)$$

Subadditivity principle is at the very basis of classical communication theory. When applied to networks, however, this principle is not a necessary condition, for the corresponding probability space is different from that of a simple communication process. In the latter, the elements of the probability space are  $p(x_i, y_j)$  with  $i, j = 1, \dots, n$ , whereas in the former two additional elements are present, namely  $p(x_0)$  and  $p(y_0)$ . One can demonstrate that, under these conditions, the sign of expression (10) can be reversed. Therefore, broadly speaking, the concept of information is more general than that of communication and can be defined as the ability of a system to associate, in a specific

manner, some of its elements as to generate a function, namely communication between a source and a destination, signalling, catalysis etc. The implications of the lack of subadditivity will be discussed in the next section. If, however, this principle applies, the difference,  $I(X:Y)$ , between the two members of expression (11) is defined as the information of the network. Therefore, one has:

$$H(X, Y) + I(X:Y) = H(X) + H(Y) \quad (12)$$

Information is therefore the piece of entropy that should be added to the joint entropy  $H(X, Y)$  in order to obtain the sum of individual entropies  $H(X)$  and  $H(Y)$ . If the subadditivity principle applies, information is positive or nil, but if it does not, that is if

$$H(X, Y) > H(X) + H(Y) \quad (13)$$

then information, as defined by Eq. (12), has to be negative. In that case, information is the piece of entropy that has to be deduced from the joint entropy in order to obtain the sum of individual entropies. If condition (10) applies, information measures the degree of integration of the network. Alternatively, if it is condition (13), information expresses the degree of emergence of the system.

#### 4.3. Integration and emergence in uncorrelated networks

The above reasoning, and the subadditivity principle which is at the heart of conventional Shannon's communication theory, rely upon the view that variables are correlated. If, however, these variables are not statistically correlated, but represent events that physically interact, the concept of information can still be applied to this situation. To illustrate this idea, let us consider a simple enzyme-catalysed chemical reaction  $AX + B \rightarrow A + XB$ . If the reaction is catalysed by enzyme E, this may imply that AX and B can form binary and ternary complexes with the enzyme E, as shown in Fig. 1. Catalysis takes place on the ternary E·AX·B complex, which decomposes, thus leading to the regeneration of the free enzyme and to the appearance of the reaction products A and XB.

Let us assume for the moment that the overall system is close to thermodynamic equilibrium, which implies that the rate constant of decomposition of the ternary complex E·AX·B is small relative to the

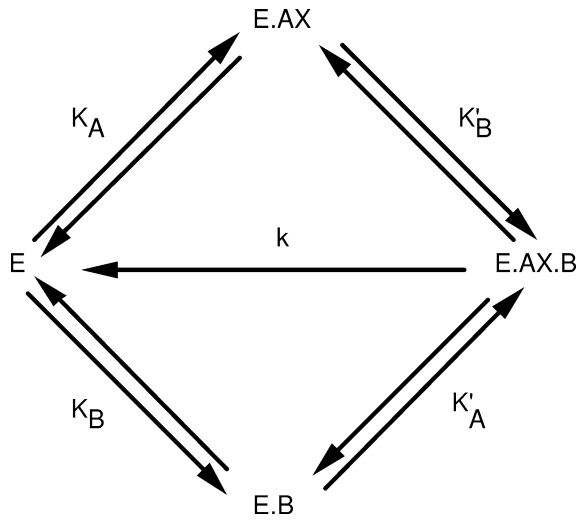


Fig. 1. Random addition of substrates, AX and B, to an enzyme E.  $K_A$ ,  $K'_A$ ,  $K_B$ , and  $K'_B$  are the apparent binding constants of the substrates to the enzyme.  $k$  is the catalytic constant.

other rate constants. Then, one can easily derive the expression of the probabilities for the enzyme to bind AX,  $p(AX)$ , or to bind B,  $p(B)$ , or to bind both AX and B,  $p(AX, B)$ . Similarly, one can derive the expression of conditional probabilities,  $p(AX/B)$  and  $p(B/AX)$ . If the concentrations of AX and B are constant in the reaction mixture, the probabilities  $p(AX)$  and  $p(B)$  are not correlated. As a matter of fact, each of them assumes one fixed value only. But one can still extend the concepts of entropy and information to the present situation by setting:

$$\begin{aligned}
 H(AX) &= -\log_2 p(AX) \\
 H(B) &= -\log_2 p(B) \\
 H(AX, B) &= -\log_2 p(AX, B)
 \end{aligned}
 \tag{14}$$

These expressions are indeed identical to equations (5) and (6), if X, Y and XY each assume one value only. Expressions of the probabilities that appear in equations (14) are dependent on the concentrations of the reagents, AX and B, as well as upon the affinity constants  $K_A$ ,  $K'_A$ ,  $K_B$  and  $K'_B$ . One can demonstrate that if  $K_A = K'_A$ , which implies from thermodynamics that  $K_B = K'_B$ , then

$$\begin{aligned}
 p(AX/B) &= p(AX) \\
 p(B/AX) &= p(B)
 \end{aligned}
 \tag{15}$$

Under these conditions:

$$H(AX, B) = H(AX) + H(B)
 \tag{16}$$

and the enzyme network does not contain any information. The corresponding system, of joint entropy  $H(AX, B)$ , can be reduced to its sub-systems, of entropies  $H(AX)$  and  $H(B)$ .

Alternatively, if  $K_A < K'_A$  and  $K_B < K'_B$ , then

$$\begin{aligned}
 p(AX/B) &> p(AX) \\
 p(B/AX) &> p(B)
 \end{aligned}
 \tag{17}$$

and

$$H(AX, B) < H(AX) + H(B)
 \tag{18}$$

The system behaves as an integrated coherent entity. Its entropy,  $H(AX, B)$ , is smaller than the sum of the entropies,  $H(AX)$  and  $H(B)$ , of the two component sub-systems. The information of the network,  $I(AX:B)$ , is thus positive and equal to the extent of the integration process.

The last case is even more interesting. If  $K_A > K'_A$  and  $K_B > K'_B$ , then

$$\begin{aligned}
 p(AX) &> p(AX/B) \\
 p(B) &> p(B/AX)
 \end{aligned}
 \tag{19}$$

and

$$H(AX, B) > H(AX) + H(B)
 \tag{20}$$

Under these conditions, the system behaves as an emergent coherent entity. Its entropy,  $H(AX, B)$ , is now larger than the sum of the entropies,  $H(AX)$  and  $H(B)$ , of the two component sub-systems. In line with previous definitions, it can therefore be considered a complex system. Here, information,  $I(AX:B)$ , defined as a measure of the extent of the emergent process, has a negative value. It is the piece of entropy that should be deduced from the joint entropy,  $H(AX, B)$ , in order to obtain the sum of individual entropies  $H(AX)$  and  $H(B)$ .

As the subadditivity principle is at the root of standard communication theory, the last conclusion above is at variance with this theory. It is therefore of interest to make clear the reasons for this disagreement. Classical Shannon's communication theory is based on two different implicit postulates: the existence of a statistical correlation between discrete variables and

the assumption that the sums of the probabilities of occurrence of the corresponding states are all equal to unity. In the case of a simple enzyme network, as the one shown in Fig. 1, none of these conditions is fulfilled. Therefore, for an enzyme reaction, subadditivity may, or may not, apply and one can also reach the same conclusion for metabolic networks. As a matter of fact, for the enzyme process of Fig. 1, if  $K'_A > K_A$  and  $K'_B > K_B$ , the binding of B to the enzyme facilitates that of AX, and conversely. Subadditivity should apply to this situation. But if  $K_A > K'_A$  and  $K_B > K'_B$ , then the binding of B hinders that of AX, and conversely. In this case, subadditivity does not apply anymore. Therefore, in the present case, positive or negative information does not originate from statistical correlation between discrete variables, but from the physical interaction between two binding processes.

It is of interest to stress that, if subadditivity does not apply, this means that the probability of occurrence of the E·AX·B complex is smaller than the product of the corresponding probabilities of occurrence of the binary complexes E·AX and E·B. This implies in turn that the ternary complex E·AX·B is destabilised relative to the binary complexes E·AX and E·B. One can see that, under these conditions, the Michaelis constants for the two substrates AX and B as well as the catalytic constant are increased. Therefore, the intrinsic properties of the enzyme network are altered owing to the high information content of the E·AX·B complex. The situation that is expected to occur is just the opposite if there is less information in the E·AX·B complex than in the set of the E·AX and E·B binary complexes.

Thus far, the simple enzyme network was assumed to occur under quasi-equilibrium conditions. If, however, the system departs from this state, that is if the catalytic rate constant,  $k$ , assumes rather high values, the system can possibly reach a steady state. Under these conditions, the probabilities for the enzyme to bind AX, B, or both AX and B depend also upon the  $k$  value. As a matter of fact, the probabilities  $p(\text{AX})$ ,  $p(\text{B})$  and  $p(\text{AX}, \text{B})$  decrease as  $k$  is increased. However, the decrease of  $p(\text{AX}, \text{B})$  is steeper than that of  $p(\text{AX})$  and of  $p(\text{B})$ , in such a way that  $H(\text{AX}, \text{B})$  becomes larger than the sum of  $H(\text{AX})$  and of  $H(\text{B})$ . This means that, as the system departs from its initial pseudo-equilibrium conditions and becomes more 'open', it tends to generate negative information.

Moreover, this situation takes place even if  $K_A = K'_A$  and  $K_B = K'_B$ , that is even if the binding of the two substrates to the enzyme do not physically interact.

## 5. General conclusions

Classical Shannon's theory is aimed at deciphering how information is conveyed within a channel from a source to a destination. The basic principles and the mathematical formulation of this communication theory, however, have to be revisited and altered as to offer a clear-cut and mathematically sound definition of the concepts of reduction, integration, emergence and complexity in biological systems. The reason why Shannon's communication model is not ideally suited for studying biological networks relies upon the fact it is solely based on the idea of a statistical correlation between discrete variables and on the so-called subadditivity principle. As a matter of fact, if two discontinuous variables are correlated, subadditivity principle should apply only if all the nodes of the network are associated with these variables. It appears very likely that this condition is not fulfilled for biological networks. This is precisely what is expressed in the statement made previously that the probability spaces are different for a communication process and a biological network. This means that three possible situations can be expected to occur. The joint entropy of the dynamic process is equal to the sum of the individual entropies of its components. Then, this process does not contain any information and its properties can be reduced to those of its components. If, alternatively, the joint entropy of the network is smaller than the sum of individual entropies of the component sub-systems, the network behaves as an integrated entity whose properties cannot be reduced to those of its elements. The information content of the system is then a measure of its degree of integration. Last but not least, if the joint entropy of the network is larger than the sum of the individual entropies of its components, the system is complex and displays emergent effects associated with negative information.

These three possible situations are precisely found to take place with simple enzyme-catalysed reactions occurring under quasi-equilibrium. In these uncomplicated networks, two substrate-binding processes are obviously not correlated, but they may physically interact, that is the binding of a substrate may facilitate

or hinder that of the other. Although there is no correlation between discrete variables, there is an information transfer between two processes. Information theory can be extended to express quantitatively this situation, and one can demonstrate that, if the binding of a substrate facilitates that of the other, the overall system is integrated and its information is positive. Alternatively, if the binding of a substrate hinders that of the other, the system is complex and its information is negative.

Thus, in the case of an enzyme reaction whose joint entropy is smaller than the sum of the corresponding individual entropies, the probability of occurrence of the  $E \cdot AX \cdot B$  state is smaller than the product of the probabilities of occurrence of  $E \cdot AX$  and  $E \cdot B$ . This implies in turn that the fundamental energy level of  $E \cdot AX \cdot B$  is low, leading to a high-energy barrier between this state and the corresponding transition state of the chemical reaction. If, conversely, the joint entropy is larger than the sum of individual entropies, the probability of occurrence of the ternary  $E \cdot AX \cdot B$  enzyme–substrates complex is destabilized relative to the binary  $E \cdot AX$  and  $E \cdot B$  complexes. In the former case above, one should expect the catalytic rate constant of the reaction to be small, whereas in the latter the catalytic constant should be high.

If the network does not occur under quasi-equilibrium but under steady-state conditions, which implies that it is thermodynamically open, the departure from quasi-equilibrium alone generates emergence of information. This situation may take place even in the absence of any physical interaction between the substrate binding processes. Departure from equilibrium may thus be sufficient to generate complexity. It therefore appears that, within the frame of the definitions given above, apparently uncomplicated biochemical systems may be considered complex.

## References

- [1] R. Descartes, *Discours de la Méthode*, Flammarion, Paris, 1992.
- [2] R.C. Strohman, The coming Kuhnian revolution in biology, *Nat. Biotech.* 15 (1997) 194–200.
- [3] R. Gallagher, T. Appenzeller, Beyond reductionism, *Science* 284 (1999) 79.
- [4] N. Goldenfeld, L.P. Kadanoff, Simple lessons from complexity, *Science* 284 (1999) 87–89.
- [5] G.M. Whitesides, R.F. Ismagilov, Complexity in chemistry, *Science* 284 (1999) 89–92.
- [6] L.H. Hatwell, J.J. Hopfield, S. Leibler, A.W. Murray, From molecular to modular cell biology, *Nature* 402 (Suppl.) (1999) C47–C52.
- [7] J.P. Setgna, K.A. Dahmen, C.R. Myers, Crackling noise, *Nature* 410 (2001) 241–250.
- [8] J. Ricard, *Biological Complexity and the Dynamics of Life Processes*, Elsevier, Amsterdam, 1999.
- [9] C.G. Hempel, *Philosophy of Natural Science*, Englewood Cliffs, New Jersey, Prentice Hall, 1966.
- [10] E. Nagel, *The Structure of Science*, Harcourt, Brace and World Inc., New York, 1961.
- [11] C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* 27 (1948) 379–423, 623–656.
- [12] C.E. Shannon, *The Mathematical Theory of Communication*, University of Illinois Press, USA.
- [13] H.P. Yockey, *Information Theory and Molecular Biology*, Cambridge University Press, 1992.
- [14] C. Adami, *Introduction to Artificial Life*, Springer Verlag, 1998.
- [15] S.H. Strogatz, Exploring complex networks, *Nature* 410 (2001) 268–276.
- [16] D.J. Watts, S.H. Strogatz, Collective dynamics of ‘small-world’ networks, *Nature* 393 (1998) 440–442.
- [17] R. Albert, H. Jeong, A.L. Barabasi, Error and attack tolerance of complex networks, *Nature* 406 (2000) 378–381.
- [18] H. Jeong, B. Tombor, R. Albert, Z.N. Oltavi, A.L. Barabasi, The large-scale organization of metabolic networks, *Nature* 407 (2000) 651–654.
- [19] D.E. Fell, A. Wagner, The small-world of metabolism, *Nat. Biotech.* 18 (2000) 1121–1122.