Semiotic modelling of biological processes: Semiosis as an emergent process

João Queiroz[a,b,c] & Charbel El-Hani[a,b,c]

a. Research Group in History, Philosophy, and Biology Teaching, Institute of Biology, Federal University of Bahia, Brazil.
b. Graduate Studies Program in History, Philosophy, and Science Teaching, Federal University of Bahia/State University of Feira de Santana, Brazil.
c. Graduate Studies Program in Ecology and Biomonitoring, Federal University of Bahia, Brazil.

General abstract: Here we introduce biosemiotics as a field of research that develops models of life processes focusing on their informational aspects. Peirce's general concept of semiosis can be used to analyze such processes, and provide a powerful basis for understanding the emergence of meaning in living systems, by contributing to the construction of a theory of biological information. Peirce's theory of sign action is introduced, and the relation between 'information processing' and sign processes is discussed, in fact, a semiotic definition of information is introduced. Three biosemiotic models of informational processes, at the behavioral and molecular levels, are developed, first, a model of genetic information processing in protein synthesis; second, a model of signal transduction in B-cell activation in the immune system; and, finally, a model of symbolic non-human primate communication. We also address some perspectives for the development of applied semiotic research in fields such as Artificial life, cognitive ethology, cognitive robotics, theoretical biology, and education.

In this lecture, we will summarize a systematic analysis of the variety of emergence theories and concepts developed by Stephan (1998,1999). This will lead us to pose fundamental questions that have to be answered (Lecture 3) in order to ascribe a precise meaning to the term "emergence" in the context of an analysis of biological semiosis.

1. Central characteristics of emergentism: what questions about semiosis do they raise?

Semiosis can be described as an "emergent" process in semiotic systems. But what do we exactly mean by this idea? To provide a clear answer to this question is particularly important, in view of the revitalization of the emergence debate in the 1990s (Kim 1998, 1999; Stephan 1999; Cunningham 2001; Pihlström 2002; El-Hani 2002), which made the term "emergence" and its derivatives much more popular than they had been throughout the whole 20th century. This is particularly true of research on computer models of non-linear dynamical systems, complex systems research, artificial life, cognitive sciences, etc. In the case of Alife, for instance, Langton (1989, p. 2) even states that the key concept in this field is that of "emergent behavior".

As the concept of emergence is increasingly used, it becomes more and more important to avoid employing it in vague and imprecise ways, inasmuch as this concept has carried for a long time a burdensome load of confusion about its metaphysical and epistemological aspects. In fact, the price of being careless about the use of this concept is already surfacing in the current debates about emergence, in the form of a certain perplexity about what is really meant by "emergence", "emergent properties", and so on.

The term "emergence" is often employed in an intuitive and ordinary way, referring to the idea of a "creation of new properties". This idea comes back to one of the original sources of the emergentist thinking, the works of the British psychologist Conwy Lloyd Morgan. As Emmeche and colleagues (1997) show, a discussion of the key concepts in this idea, "novelty", "property", and "creation", can result in an understanding of some of the main issues in emergentism. Nevertheless, this idea is not enough for grasping the concept of emergence, mainly because it is focused on characteristic claims of one type of emergentism, namely, "diachronic emergentism" (see below).

In a technical sense, "emergent" properties can be understood as *a certain class* of higher-level properties related *in a certain way* to the microstructure of *a class of systems*. The reason why such a broad definition, with open clauses, seems at first more adequate than a definition with more content and precision has to do with the fact that the concept of emergence and its derivatives are employed in the most diverse fields. Consequently, a more detailed definition is likely to apply to some fields but not to others. It is true, however, that a more concrete and operational definition is needed when one is dealing with particular cases of emergence, as it is the case when we discuss the emergent nature of semiotic process. One should not rest content, therefore, with the above definition, with its open clauses. But attempts to made it more precise should be dealt with case by case, considering specific theoretical and empirical constraints on the meaning of "emergence" in different research fields. It is part of the task of an emergence theory – applied to a particular research field – to fill in the open clauses in this definition (shown in italics). It should provide, among other things, an account of which systemic properties of a class of systems are to be regarded as "emergent" and offer an explanation of how they relate to the microstructure of such systems. Moreover, it should establish which systems exhibit a certain class of emergent properties. If we extend this definition to encompass processes[i], a first question to be answered in order to characterize semiosis as an emergent process concerns the demarcation of the class of systems which show semiosis. We can frame it as follows: (1) what is a semiotic system?

There is no unified emergence theory. Rather, emergence theories come in various shapes and flavors. Nevertheless, it is possible to recognize in the diversity of emergence theories a series of central characteristics (Stephan 1999, chapter 3; cf. also Stephan 1998). In the following sections, we will discuss in detail the fundamental tenets of emergentist philosophies, at least in their scientifically-compatible versions. Nevertheless, for some readers already familiar with emergentism, it may be unnecessary to go through all these details. Accordingly, we will present in the next paragraph a brief overview of these basic concepts.

In scientifically-compatible accounts, emergentism is a naturalistic and physicalistic position, according to which the evolution of physically constituted systems show, from time to time, critical turning points, in which new organizational patterns arise, and, thus, new classes of systems exhibiting novel properties and processes. Among these novel properties and processes, emphasis is given to emergent properties, a particular class of systemic properties (i.e., properties observed at the level of the whole, but not of the parts). Emergent properties are not treated, in a scientifically-compatible emergentist philosophy, as free-floating properties, but rather they are conceived as being grounded on the system's microstructure, by which they are synchronically determined. But, despite synchronic determination, emergentists also treat these properties as irreducible, basically on two different senses: (i) emergent properties can be irreducible because they cannot be analyzed in terms of the behavior of a system's parts (unanalyzability), or (ii) because they depend on the parts' behavior within a system of a given kind, and this behavior, in turn, does not follow from the parts' behavior in isolation or in other (simpler) kinds of system (non-deducibility). This latter concept of irreducibility is related to a rather important but quite controversial idea in emergentism, that of a downward determinative influence of the system as a whole on the behavior of its parts, from which it follows the non-deducibility of the latter behavior (downward determination). Finally, another fundamental tenet opposes reductionistic treatments of emergent processes and properties, and, consequently, of the systems exhibiting them, namely in principle theoretical unpredictability, i.e., the idea that emergent properties or processes are not only novel but also cannot be theoretically predicted before its first appearance. After this overview, we will now delve into a detailed discussion about each of these basic ideas.

First, emergentists should, in a scientific spirit, be committed to *naturalism*, claiming that only natural factors play a causal role in the evolution of the universe. Even though naturalism and materialism (or, for that matter, physicalism) philosophically do not coincide, it is the case that, in the current scientific picture, a naturalistically-minded emergentist should also stick to the idea that all entities consist of physical parts. This thesis can be labeled "*physical monism*": there are, and will always be, only physically constituted entities in the universe, and any emergent property or process is instantiated by systems that are exclusively physically constituted. Therefore, we can pose the following question: (2) are semiotic systems exclusively physically constituted?

A second characteristic mark of emergentism is the notion of *novelty*: new systems, structures, processes, entities, properties, and dispositions are formed in the course of evolution. This idea entails the following question: (3) do semiotic systems constitute a new class of systems, instantiating new structures, processes, properties, dispositions, etc.?

Emergence theories require, thirdly, a distinction between *systemic* and *non-systemic properties*. A property is systemic if and only if it is found at the level of the system as a whole, but not at the level of its parts. Conversely, a non-systemic property is also observed at the parts of the system. If we similarly propose a distinction between systemic and non-systemic processes, the next question can be raised: (4) can semiosis be described as a systemic process?

A fourth characteristic of emergence theories is the assumption of a *hierarchy of levels of existence*. Thus, it is also necessary to answer the following question to convincingly characterize semiosis as an emergent process: (5) how should we describe levels in semiotic systems and, moreover, how do these levels relate to the emergence of semiosis?

A fifth characteristic is the thesis of *synchronic determination*, a corollary of physical monism: a system's properties and behavioral dispositions depend on its microstructure, i.e., on its parts' properties and arrangement; there can be no difference in systemic properties without there being some difference in the properties of the system's parts and/or in their arrangement.[ii] The next question to be addressed, then, is the following: (6) in what sense can we say (and explain) that semiosis, as an emergent process in semiotic systems, is synchronically determined by the properties and arrangement of its parts?

Sixthly, although some emergentists (e.g., Popper in Popper & Eccles [1977]1986) have subscribed to indeterminism, one of the characteristics of emergentism (at least in the classical British tradition[iii]) is a belief in *diachronic determination*: the coming into existence of new structures would be a deterministic process governed by natural laws (Stephan 1999, p. 31). This is certainly one feature of classical emergence theories which is incompatible with Peirce's theoretical framework, as he rejected the belief in a deterministic universe (CP 6.201). But this does not preclude the treatment of emergence in connection to a Peircean account of semiosis, as there are also emergence theories committed to indeterminism. It is not necessary at all to be imprisoned in the old British tradition of emergentist thought.

Seventhly, emergentists are committed to the notion of the *irreducibility* of a systemic property designated as "emergent". An eighth important notion used is that of *unpredictability* (in principle). We should, then, pose two more questions: (7) in what sense can we say that semiosis, as observed in semiotic systems, is irreducible? (8) In what sense can we claim that the instantiation of semiosis in semiotic systems is unpredictable in principle?

> Finally, the ninth characteristic of emergentism is the idea of *downward causation*: novel structures or new kinds of states of "relatedness" of preexistent objects manifest downward causal efficacy, determining the behavior of a system's parts. Given this idea, yet another question should be raised: (9) is some sort of downward causation involved in semiosis? We will discuss these latter notions in a fine-grained manner in the next section.

2. Varieties of emergentism: what questions about semiosis do they raise?

Several different emergence theories have been proposed throughout the 20th century. The characteristic marks discussed above allow one to define several varieties of emergentism, significantly differing from one another in strength (see Stephan 1998; 1999, chapter 4).

For the sake of our arguments, we will consider just three basic varieties of emergentism – *weak*; *synchronic*; and *diachronic* emergentism.[iv] Weak emergentism assumes (1) physical monism, (2) a distinction between systemic and non-systemic properties, and (3) synchronic determination. This comprises the minimal conditions for a physicalist emergentist philosophy. Thus, weak emergentism is the common basis for all stronger physicalist emergence theories. However, this view in itself is weak enough to be compatible with reductive physicalism (Stephan 1998, p. 642; 1999, p. 67). Consequently, weak emergentism faces a fundamental problem as regards the basic motivations underlying the efforts of most emergence theorists, who typically take emergentism to be by definition an anti-reductionist stance.

In this work, we intend to characterize semiosis as an emergent process in a stronger sense. Therefore, we have to analyze in more detail the concepts of "irreducibility" and "unpredictability", assumed in stronger forms of emergence theories, committed to synchronic and/or diachronic emergentism.

Synchronic and diachronic emergentism are closely related, being often interwoven in single emergence theories, but, for the sake of clarity, it is important to distinguish between them. Synchronic emergentism is primarily interested in the relationship between a system's properties and its microstructure. The central notion in synchronic emergentism is that of *irreducibility*. Diachronic emergentism, by its turn, is mainly interested in how emergent properties come to be instantiated in evolution, focusing its arguments on the notion of *unpredictability*.

Modes of irreducibility

By adding to the three tenets of weak emergentism the thesis of the irreducibility of systemic properties, synchronic emergentism yields a doctrine incompatible with reductive physicalism. Stephan (1998, pp. 642-643; 1999, p. 68) distinguishes between two kinds of irreducibility. The first notion is based on the behavioral unanalyzability of systemic properties:

> (I_1) [Irreducibility as unanalyzability] Systemic properties which cannot be analyzed in terms of the behavior of a system's parts are necessarily irreducible.

This notion plays an important role in the debates about qualia and is related to a first condition for reducibility, namely, that a property *P* will be reducible if it follows from the behavior of the system's parts that the system exhibits *P*. Conversely, a systemic property *P* of a system *S* will be irreducible if it does not follow, even in principle, from the behavior of the system's parts that *S* has property *P*.

A second notion of irreducibility is based on the non-deducibility of the behavior of the system's parts:

> (I_2) [Irreducibility of the behavior of the system's parts] A systemic property will be irreducible if it depends on the specific behavior the parts show within a system of a given kind, and this behavior, in turn, does not follow from the parts' behavior in isolation or in other (simpler) kinds of system (cf. Stephan 1998, p. 644).

It is here that the notion of downward causation (DC) enters the scene: there seems to be some downward causal influence of the system where a given emergent property *P* is observed on the behavior of its parts, as we are not able to deduce this behavior from the behaviors of those very same parts in isolation or as parts of different kinds of system. A second condition for reducibility is violated in this case. This condition demands "that the behavior the system's parts show when they are part of the system follows from the behavior they show in isolation or in simpler systems than the system in question" (Stephan 1998, p. 643). It follows from this condition that a systemic property *P* of a system *S* will be irreducible if it does not follow, even in principle, from the behavior of the system's parts in systems simpler than *S* how they will behave in *S*, realizing property *P*.

More recently, Stephan, along with other authors, grasped the notions of irreducibility as unanalyzability and as non-deducibility of the behavior of the system's parts in two conditions for emergence they call "vertical" and "horizontal" (Boogerd et al., 2005).

Taking Broad's works (1919, 1925) as a starting point, Boogerd et al. (2005) distinguish between two independent conditions for emergence that Broad himself did not explicitly differentiate (Figure 1). A systemic property P_R of a system $R(A,B,C)$ is emergent if either of these conditions is fulfilled. The vertical condition captures the situation in which a systemic property P_R is emergent because it is not explainable, even in principle, with reference to the properties of the parts, their relationships within the entire system $R(A,B,C)$, the relevant laws of nature and the required composition principles. The horizontal condition grasps the situation in which a systemic property P_R is emergent because the properties of the parts within the system $R(A,B,C)$ cannot be deduced from their properties in isolation or in other wholes, even in principle. Since these two conditions are independent, there are two different possibilities for the occurrence of emergent properties: (i) a systemic property P_R of a system S is emergent if it does not follow, even in principle, from the properties of the parts within S that S has property P_R; and (ii) a systemic property P_R of a system S is emergent, if it does not follow, even in principle, from the properties of the parts in systems different from S how they will behave in S, realizing P_R.

The vertical condition for emergence expresses in a different way the idea of unanalyzability. Even if we know (i) what properties and relations A, B, and C show within the system $R(A,B,C)$, (ii) the relevant laws of nature, and (iii) all necessary composition principles, yet we will not be able to deduce that the system has property P_R. This is a case in which the condition of analyzability is violated, as it does not follow, even in principle, from the behavior of the parts A, B, and C in system $R(A,B,C)$ that the system has P_R. This is basically the idea of emergence that appears in most metaphysical discussions, particularly in discussions about qualia (see, *e.g.*, Levine 1983, 1993; Kim 1999). As Boogerd et al. (2005) comment, if some phenomenon is emergent in this sense, it will be fundamental and irreducible, in the sense that it is neither predictable nor explainable in terms of the properties and relations of the system's own constituents.

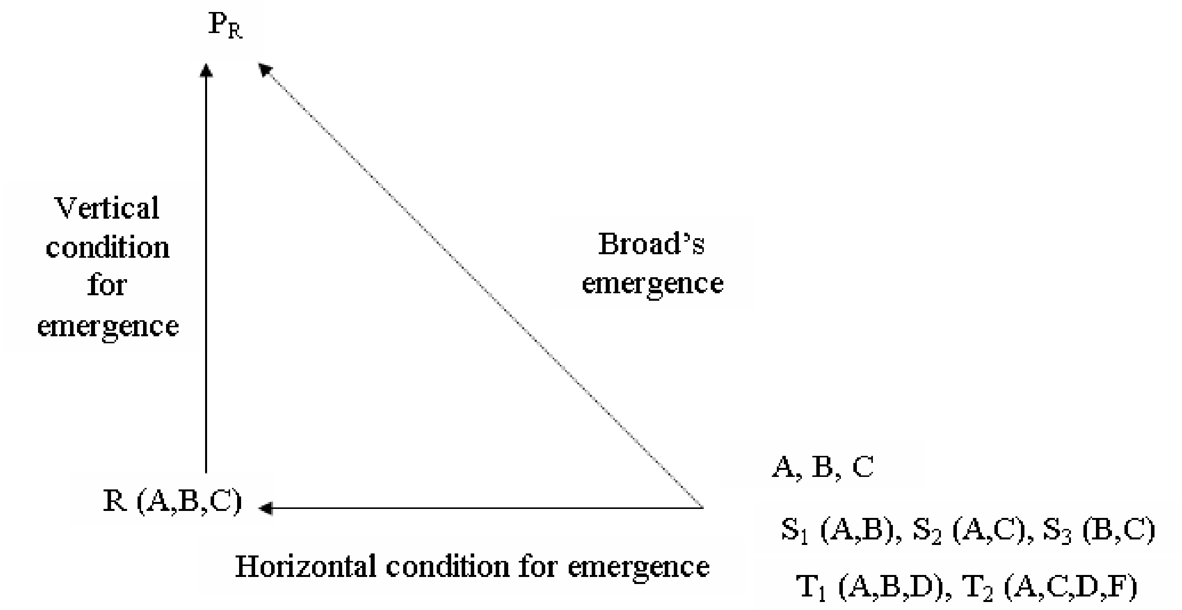

Figure 1: Vertical and horizontal conditions for emergence. *A*, *B*, and *C* are the parts making up the system $R(A,B,C)$, which shows P_R, a systemic property. $S_1(A,B)$, $S_2(A,C)$, and $S_3(B,C)$ are simpler systems including these parts. $T_1(A,B,D)$ is a system with the same number of parts, and $T_2(A,C,D,F)$ is a system with more parts than $R(A,B,C)$. The diagonal arrow represents Broad's idea of emergence. The horizontal and vertical arrows capture the two conditions implicit in Broad that Boogerd and colleagues made explicit. (From Boogerd et al., 2005).

The horizontal condition for emergence expresses in a different way the idea of irreducibility based on the non-deducibility of the behavior of the system's parts. In this case, if we know the structure of the system $R(A,B,C)$, we will be able to explain and predict the behavior of the parts within it, and, also, the instantiation of the systemic property P_R.

Boogerd et al. (2005) discuss the resources available for deducing the behavior of the parts within $R(A,B,C)$ from other kinds of systems, in order to establish what would be the proper basis for such a deduction. We may deduce the behavior of the parts in $R(A,B,C)$ from their behavior in systems of greater, equal or less complexity. As Figure 1 shows, the possible bases for deduction of the parts' behavior in $R(A,B,C)$ include: (i) more complex systems, such as $T_2(A,C,D,F)$; (ii) systems with the same degree of complexity, such as $T_1(A,B,D)$; (iii) simpler systems, such as $S_1(A,B)$, $S_2(A,C)$, and $S_3(B,C)$; and (iv) the parts A, B, and C in isolation.[v] Boogerd et al. convincingly argue that only (iii) is an interesting basis for deduction since (iv) trivializes emergence – as, in this case, each and every property of a system would seem to be 'emergent' –, and (i) and (ii) trivialize non-emergence – since, in this case, each and every property of a system would seem to be 'non-emergent'. They conclude that the key case for understanding the horizontal condition for emergence is (iii), in which we attempt to deduce the behavior of $R(A,B,C)$ or its parts on the basis of less complex systems.

A more fine-grained analysis of the irreducibility concept naturally leads to a reframing of the seventh question raised above: (7) Which interpretation of irreducibility is more adequate to understand Peirce's claims about the irreducibility of semiosis? Furthermore, the explanation of irreducibility as non-deducibility makes it evident that question 9, "is some sort of downward causation involved in semiosis?", should be posed in connection with this particular interpretation. This raises a number of difficult questions, as the problem of downward causation (DC) is the most debated in the contemporary literature on emergence (see, e.g., Schröder 1998; Stephan 1999; Andersen et al. 2000; El-Hani 2002; Hulswit, 2005). Therefore, we will not pursue this debate here in all its details. Rather, we will discuss some central ideas and controversies about DC, in order to subsequently consider them with regard to semiotic phenomena.

Downward causation

Emmeche and colleagues (2000) identified three versions of DC, each making use of a particular way of interpreting the causal mode (or modes) involved in the influence of a whole over its parts: strong, medium, and weak DC. Strong DC interprets the causal influence of a whole over its parts as a case of ordinary, efficient causation. Nevertheless, we need to postulate a sharp distinction between a higher and a lower level, each being constituted by different kinds of substances, if we want to claim that a higher level exerts an efficient causal influence over a lower one (Emmeche et al. 2000; Hulswit, 2005). Strong DC implies, thus, substance dualism, and this makes it, in turn, quite untenable in a current scientific understanding of emergence. Moreover, this notion faces a number of other important difficulties. If we consider the standard case in discussions about DC, i.e., “reflexive” and “synchronic” downward causation (Kim 1999), in which some activity or event involving a whole at a time *t* is a cause of, or has a causal influence on, the events involving its own micro-constituents at that same time *t*, then a strong account of DC looks like a bizarre metaphysical bootstrapping exercise (see Symons 2002).

Symptomatically, Emmeche and colleagues (2000) emphasize that there are only two viable candidates for a scientifically-compatible account of DC, both related to an interpretation of DC as a case of synchronic formal causation: medium and weak DC. We can summarize the key points in Emmeche and colleagues' arguments for medium DC as follows: (i) a higher-level entity comes into being through the realization of one amongst several possible lower-level states. (ii) In this process, the previous states of the higher level operate as a "factor of selection" (p. 24) for the lower-level states. (iii) The idea of a factor of selection can be made more precise by employing the concept of "boundary conditions", introduced by Polanyi (1968) in the context of biology, particularly in the sense that higher-level entities are boundary conditions for the activity of lower levels, *constraining* which higher-level phenomenon will result from a given lower-level state. (iv) Constraints can be interpreted in terms of the characterization of a higher level by "organizational principles" – law-like regularities – that have a downward effect on the distribution of lower-level events and substances. (iv) Medium DC is committed to the thesis of "constitutive irreductionism" (p. 16), namely, the idea that even though higher-level systems are ontologically constituted by lower-level entities, the higher level cannot be reduced to the form or organization of the constituents. (v) Rather, the higher level must be said to "constitute its own *substance* and not merely to consist of its lower-level constituents" (p. 16, emphasis in the original), or, else, a higher-level entity should be regarded as a "real substantial phenomenon in its own right" (p. 23). (vi) This interpretation of DC may assume either a thesis they call "formal realism of levels" (p. 16), stating that the structure, organization or form of an entity is an objectively existent feature of it, which is irreducible to lower-level forms or substances, or a thesis designated as "substantial realism of levels" (p. 16), claiming that a higher-level entity is defined by a "substantial difference" from lower-level entities. Thus, an important difference between medium and strong DC seems to lie in the necessary commitment of the latter to the thesis of a "substantial realism of levels". Another difference highlighted by Emmeche and colleagues (2000, p. 25) is that "medium DC does not involve the idea of a strict 'efficient' temporal causality from an independent higher level to a lower one".

In turn, Emmeche and colleagues' (2000) treatment of weak DC can be summarized in terms of the following arguments: (i) in the weak version, DC is interpreted in terms of a "formal realism of levels", as explained above, and "constitutive reductionism" (p. 16), the idea that a higher-level entity ontologically consists of lower-level entities organized in a certain way. (ii) Higher-level forms or organization are irreducible to the lower level, but the higher-level is not a "real substantial phenomenon", i.e., it does not add any substance to the entities at the lower level. (iii) In contrast to the medium version, weak DC does not admit the interpretation of boundary conditions as constraints. (iv) If we employ phase-space terminology, we will be able to explain weak DC as the conception of higher-level entities as attractors for the dynamics of lower levels. Accordingly, the higher level is thought of as being characterized by formal causes of the self-organization of constituents at a lower level. (v) The relative stability of an attractor is taken to be identical to the downward "governing" of lower-level entities, i.e., the attractor functions as a "whole" at a higher level affecting the processes that constitute it (p. 28). (vii) The attractor also functions as a whole in another sense of the word, given that it is a general *type*, of which the single phase-space points in its basin are *tokens* (p. 29).

Even though Emmeche and colleagues' contribution to the debates about DC has a lot of merit, particularly because it stressed a diversity of DC accounts that has been often neglected and, moreover, tried to make advances in organizing the variety of such accounts, their typology faces a number of problems. But this is not an exclusive feature of their work; rather, many attempts to explain DC available in the literature are confronted with important difficulties (see Hulswit, 2005).[VI] Particularly, the distinctions between strong, medium, and weak DC should be further clarified. For instance, it seems necessary to explain in more detail in what sense strong and medium DC differ as regards the idea that a higher-level entity is a "substantial" phenomenon, or, else, how one would differentiate medium versions committed to the thesis of a "substantial realism of levels" from strong DC. For the sake of our arguments, we will simply work below with an interpretation which comes close to medium DC by interpreting boundary conditions as constraints, but, at the same time, departs from it, by resolutely rejecting "constitutive irreductionism". It also comes close, thus, to weak DC. We will not try, however, to classify our account in terms of Emmeche and colleagues' typology. We will rather concentrate on explaining how we will conceive here the relationship between DC and constraints.

In order to do so, we will begin by considering that, when lower-level entities are composing a higher-level system, the set of possible relations among them is constrained, as the system causes its components to have a much more ordered distribution in spacetime than they would have in its absence. This is true in the case of both entities and processes, since processes also make the elements involved in them assume a particular distribution in spacetime. We can take a first step, then, towards explaining why the same lower-level entity can show different behaviors depending on the higher-level system it is part of.[VII] The parts are, so as to say, "enslaved" by a particular pattern of constraints on their relations which is characteristic of systems of a given kind.

The “causes” in DC can be treated, in these terms, as higher-level general organizational principles which constrain particular lower-level processes (the “effect”), given that the particular relations the parts of a system of a given kind can be engaged in depend on how the system’s structures and processes are organized. In this framework, DC can be interpreted as a “formal cause” by recasting the notion of higher-level “constraints” (or “constraining conditions”), much discussed in works about the nature of complex systems (e.g. Salthe 1985), in terms of Aristotle’s set of causal concepts (see Emmeche et al. 2000, El-Hani & Pereira 2000, El-Hani & Emmeche 2000, El-Hani & Videira 2001).

As Emmeche and colleagues (2000) argue, the notion of “boundary conditions” can be used for characterizing these higher-level constraints (see also Van Gulick 1993). Polanyi (1968) argued that a living system, as a naturally designed entity, works under the control of two principles: The higher one is the principle of design or organization of the system, and this harnesses the lower one, which consists in the physical-chemical processes on which the system relies. As the physical-chemical processes at the lower level are harnessed, the components come to perform functions contributing to the maintenance of the dynamical stability of the system as a whole. The (higher-level) constraining conditions are related to the higher-level organizational principles, which restrain the activity of the components at the lower level, selecting among the set of states that could be realized by the lower level that one which will be actually realized at a given time t.

Hulswit (2005) recently argued that most of the discussions about DC do not really refer to causation, but rather to downward *explanation* or *determination*. He correctly pointed out that the meanings usually ascribed to the supposedly causal influence of the higher on the lower level are not clearly related to our intuitive use of the verb “to cause” (in the sense of “bringing about”). This can be seen as a result of an impoverishment of the meaning of the term “cause” in modern science, due to the fact that classical physics critically appraised, and, ultimately, denied a number of theses related to Aristotelian philosophy, many of them concerned with the principle of causality (El-Hani & Videira 2001). Ultimately, only two of the four Aristotelian causal modes, efficient and final causes, ended up being taken into account in the meaning of the term “cause” in most modern languages.

Symptomatically, the Greek word translated as "cause" in Aristotle's works does not mean "cause" in the modern sense (Ross [1923]1995, p. 75; Lear 1988, p. 15). For Aristotle, a "cause" was not only an antecedent event sufficient to produce an effect or the goal of a given action, but *the basis or ground of something*. He stated that we understand something when we know why it is what it is, and the primary cause provides the grounds for our understanding of the 'why' of things being what they are (*Physics* II.3, 194b17-20. Aristotle 1995:332). And, moreover, he identified the why or the primary cause of a thing with its form. In his view, the form provided us with the best understanding of what a thing most truly is and why it is the way it is (Lear 1988:27).

It is in this sense that Aristotle claimed that form (and also matter) could be treated as having the aspect of causes – in terms of his formal and material causal modes. It is not surprising, then, that, if we stick to our currently intuitive ideas about causation, as Hulswit does, Aristotle's causal modes and, therefore, interpretations of DC which appeal to ideas such as that of formal causes seem more similar to modes of explanation than to modes of causation.

We will not use this line of reasoning, however, as a basis for counteracting Hulswit's arguments. We will rather explore his remarks to the effect that, although verbs usually related to the causing activity of a higher level in DC, such as "to restrain", "to select", "to organize", "to structure", "to determine", etc., may be understood as being related to "causing", they are not equivalent to "causing", in the modern sense. If we accept this line of reasoning, it will be an important task to try to understand what is the relationship between such "activities" ascribed to the higher level and "causing", so as to illuminate a pathway to a reinterpretation of DC.

It seems to us that the important relationship in this case lies in the fact that in considering either DC or our intuitive ideas about causation, we are dealing with some kind of *determination*. As Hulswit (2005) stresses, the main difference between "determining" and "causing" is that the former primarily involves necessitation (in the sense of "it could not be otherwise") while the latter primarily involves the idea of "bringing about".

We suggest here, even though in a preliminary way, that we should move from a notion of "downward causation" to one of *downward (formal) determination*. Instead of proposing that an understanding of the influence of wholes over parts demands causal categories other than efficient causation, we will rather claim that such an understanding requires other kinds of determination than just causation. For the sake of our arguments, consider, first, that most of the debates about DC are already about determination or explanation rather than causation. Second, that a similar move has been made in the case of another determinative but mereological relation, namely, physical realization (and, thus, supervenience), that cannot be properly accounted for as "causal" (see Kim 1993). Thus, it is largely accepted in other current philosophical debates, such as those about supervenience, the introduction of non-causal determinative relations.

Anyway, as much as in the case of DC, a proper explanation of downward determination will demand a clear theory about the relata at stake and the connections between this kind of determination and other basic categories, such as “law” and “cause” itself. El-Hani and Queiroz (2005) advances an account of downward determination (DD) in which a higher-level organizational pattern, interpreted as a general principle or disposition, is treated as the determiner, while lower-level particular processes are treated as the determined. Based on Ransdell’s (1983, p. 23) argument that the notion of determination, in the context of Peirce’s philosophy, carries a logical and a causal sense, they focus on the logical sense of determination, related to material implication (if p, then q). In this sense, ‘determination’ should be understood as a constraining rather than a causally deterministic process. El-Hani and Queiroz treat this determinative influence, then, as a propensity relation: if some lower-level entities a,b,c,…,n are under the influence of a general organization principle, W, they will show a tendency to behave in certain specific ways, and, thus, to instantiate a set of specific processes. The determining influence in this case is from a higher-level general organization principle on particular lower-level processes, and can be framed as follows: if a,b,c,…,n are under the influence of W, then they will show a tendency, a disposition, to instantiate process p. In other terms, we treat here the relation of implication, $p \rightarrow q$, as a ‘would be tendency’, as a relation leading to a higher likelihood that a given process happens. The difference between causal and logical determination is striking: while causal determination is a productive event, which brings about effects, logical determination is rather a subtractive event, which constrains the possibilities of behaviors of the components of a system. By eliminating possible behaviors of the parts, and, thus, coordinating them, it makes it possible that the system as a whole show an enhanced set of capabilities.

Given the arguments presented above, we can reformulate the ninth question as follows: (9) can we describe any sort of downward determinative relation in semiosis?

Unpredictability

We can now turn to diachronic emergentism, which can be treated swiftly here. This variety of emergentism is concerned with the doctrine of "emergent evolution". All diachronic theories of emergence are ultimately grounded on the thesis that "novelties" occur in evolution, opposing any sort of preformationist position. But merely the addition of the thesis of novelty does not turn a weak emergence theory into a strong one. Strong forms of diachronic emergentism demand the thesis of the "in principle theoretical unpredictability" of novel properties or structures.[VIII] The notion of "genuine novelty" then enters the scene, as one claims that a given property or structure is not only novel but also could not be theoretically predicted before its first appearance.

A systemic property can be unpredictable in this sense for two different reasons (Stephan 1998, p. 645): (i) because the microstructure of the system exemplifying it for the first time in evolution is unpredictable; (ii) because it is irreducible, and, in this case, it does not matter if the system's microstructure is predictable or not. As the second case does not offer any additional gains beyond those obtained in the treatment of irreducibility, we will focus our discussion on the unpredictability of the structures of semiotic systems and processes. We can reformulate, then, the eighth question raised in the previous section as follows: (8) is the structure of semiotic systems or processes in principle theoretically unpredictable?

Now, we should turn to our tentative answers to the questions raised along the discussion of emergentism and its varieties. Nevertheless, to do so, we should first present a general model for explaining the emergence of semiosis in semiotic systems developed by taking as a starting point Salthe's hierarchical structuralism (Queiroz & El-Hani 2006). We will devote the next lecture to this task.

Acknowledgement: João Queiroz and Charbel El-Hani are indebted to the Brazilian National Research Council (CNPq) and the State of Bahia Foundation for Research Support (FAPESB).

Next lecture (Multi-level model of emergent semiosis)> In the next lecture we will employ Salthe's (1985) hierarchical structuralism as a basis for developing a model for explaining the emergence of semiosis in systems which produce, process, and interpret signs.

REFERENCES

Andersen, P. B.; Emmeche, C.; Finnemann, N. O.; Christiansen, P. V. (eds.). 2000. *Downward Causation: Minds, Bodies and Matter*. Aarhus: Aarhus University Press.

Aristotle. 1995. *The Complete Works of Aristotle. The Revised Oxford Translation*. Barnes, J. (Ed.). 2 Vols. Princeton: Princeton University Press

Boogerd, F.C.; F.J.Bruggeman; R.C.Richardson; A.Stephan, H.V.Westerhoff 2005: "Emergence and Its Place in Nature: A Case Study of Biochemical Networks". *Synthese* 145 (1): 131 - 164.

Broad, C. D. 1919. "Mechanical explanation and its alternatives". *Proceedings of the Aristotelian Society* 19:86-124.

__. 1925. *The Mind and its Place in Nature*. London: Kegan Paul, Trench, Trubner & Co. Available in http://www.ditext.com/broad/mpn/mpn.html

Cunningham, B. 2001. "The Reemergence of 'Emergence'". *Philosophy of Science* 68 (PSA 2000 Proceedings): S62-S75.

El-Hani, C.N. 2002. "On the Reality of Emergents". *Principia*, Special Issue on Emergence and Downward Causation (El-Hani, C. N. Ed.) 6:51-87.

El-Hani, C. N. & Emmeche, C. 2000. "On some theoretical grounds for an organism-centered biology: Property emergence, supervenience, and downward causation. *Theory in Biosciences* 119 (3/4): 234-275.

El-Hani, C. N. & Pereira, A. M. 2000. Higher-Level descriptions: Why should we preserve them? In: P. B. Andersen, C. Emmeche, N. O. Finnemann & P. V. Christiansen (eds.), *Downward Causation. Minds, Bodies and Matter*. Aarhus: Aarhus University Press. (pp. 118-142).

El-Hani, C. N. & Queiroz, J. 2005. "Downward Determination". *Abstracta* 1(2): 162 – 192.

El-Hani, C. N. & Videira, A. A. P. 2001. "Causação Descendente, Emergência de Propriedades e Modos Causais Aristotélicos". *Theoria* 16:301-329.

Emmeche, C.; Køppe, S. & Stjernfelt, F. 1997. "Explaining Emergence: Towards an Ontology of Levels". *Journal for General Philosophy of Science* 28: 83-119.

__. 2000. "Levels, Emergence and Three Versions of Downward Causation", in P. B. Andersen, C. Emmeche, N. O. Finnemann & P. V. Christiansen (eds.). *Downward Causation: Minds, Bodies and Matter*. Aarhus: Aarhus University Press. (pp. 13-34).

Hulswit, M. 2001. "Semeiotic and the Cement of the Universe: a Peircean Process Approach to Causation", *Transactions of the Charles S. Peirce Society: A Quarterly Journal in American Philosophy,* Summer, XXXVII (3) 339-363.

____. 2005. "How causal is downward causation?" *Journal for General Philosophy of Science* 36(2): 1-26.

Kim, J. 1993. *Supervenience and Mind*. New York: Cambridge University Press.

__. 1998. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, Mass., and London: MIT Press.

__. 1999. "Making Sense of Emergence", *Philosophical Studies* 95: 3-36.

Langton, C. 1989. "Artificial life" in C. Langton, C. (ed.). *Artificial Life, Santa Fe Institute Studies in the Sciences of Complexity*. Vol. 5. (Proceedings of the First Conference on Artificial Life, Los Alamos, September, 1987). Reading, Mass.: Addison-Wesley.

Lear, J. 1988. *Aristotle: The Desire to Understand*. New York: Cambridge University Press.
Levine, J. 1983. "Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly* 64: 354-361.
Peirce, C. S. (EP1, 1992; EP2, 1998.). *The Essential Peirce. Selected Philosophical Writings*. (Vol. 1 ed. by N. Houser & C. Kloesel; Vol 2 ed. by the Peirce Edition Project). Bloomington and Indianapolis: Indiana University Press. (quoted as EP, followed by volume and paragraph).
Peirce, C.S. 1931–1935. *The Collected Papers of Charles Sanders Peirce*. Electronic edition reproducing Vols. I-VI [C. Hartshorne & P. Weiss (eds.), Cambridge: Harvard University Press, 1931–1935]; Vols. VII–VIII [A. W. Burks (ed.), same publisher, 1958]. Charlottesville: Intelex Corporation. (quoted as CP, followed by volume and paragraph.)
Peirce, C.S. 1967. *Annotated Catalogue the Papers of Charles S. Peirce*. (ed.) R.S. Robin. Massachusetts: The University of Massachusetts Press. [quoted as MS, followed by the number of the manuscript].
Pihlström, S. 2002. "The Re-Emergence of the Emergence Debate", *Principia*, Special Issue on Emergence and Downward Causation (El-Hani, C. N. Ed.) 6: 133-181.
Polanyi, M. 1968. "Life's Irreducible Structure", *Science* 160: 1308-1312.
Popper, K.R. & Eccles, J.C. [1977] 1986. *The Self and Its Brain*. London: Routledge and Kegan Paul.
Queiroz, J. & El-Hani, C. N. 2006. "Towards a multi-level approach to the emergence of meaning processes in living systems". *Acta Biotheoretica* 54 (3): 174-206.
Ransdell, J. (1983). *Peircean semiotic*. Manuscript.
Rescher, N. 1996. *Process Metaphysics: An Introduction to Process Philosophy*. New York: SUNY Press.
Ross, D. [1923] 1995. *Aristotle*. Routledge: London.
Salthe, S. N. 1985. *Evolving Hierarchical Systems: Their Structure and Representation*. New York: Columbia University Press.
Schröder, J. 1998. "Emergence: Non-deducibility or Downwards Causation?". *Philosophical Quarterly* 48:433-452.
Stephan, A. 1998. "Varieties of Emergence in Artificial and Natural Systems", *Zeitschrift für Naturforschung* 53c:639-656.
__. 1999. *Emergenz: Von der Unvorhersagbarkeit zur Selbstorganisation*. Dresden and München: Dresden University Press.
Symons, J. 2002. "Emergence and Reflexive Downward Causation". *Principia*, Special Issue on Emergence and Downward Causation (El- Hani, C. N. Ed.) 6: 183-202
Van Gulick, R. 1993. "Who is in charge here? And who's doing all the work?" in J. Heil & A. Mele (eds.). *Mental Causation*. Oxford: Oxford University Press. pp. 233-256.

[i]**Notes**

It is true that one may treat a process as a property exhibited by a given system, but, in order to emphasize the dynamic nature of semiosis, we will systematically describe it as an emergent “process”, rather than as a “property”. We follow here Rescher in his definition of a process as “... a coordinated group of changes in the complexion of reality, an organized family of occurrences that are systematically linked to one another either causally or functionally” (Rescher 1996, p.38).

[ii] Previously, this idea has been typically stated in a weaker version, as the thesis of ‘mereological supervenience’. The reason why this latter thesis is to be regarded as weaker than synchronic determination lies in the issue that it does not entail, as it is usually thought, the determination and dependence of the system’s properties on its microstructure. Supervenience simply states a pattern of covariance between two sets of properties, and we cannot directly derive a metaphysical relation of dependence/determination from property covariance (for treatments of this issue, see, e.g., Kim 1993, 1998; Heil 1998; Bailey 1999). In this respect, it seems better to rely on a stronger statement, directly addressing determination, instead of relying upon the notion of supervenience.

[iii] On British emergentism, see Blitz (1992), McLaughlin (1992), Stephan (1999).

[iv] Stephan begins his systematic analysis of varieties of emergentism by discussing these three basic forms, but later expands his typology to include six different emergentist positions. We will not deal with these six positions in the scope of this lecture. For more details, see the original works.

[v] Boogerd et al. are obviously aware that complexity does not depend only on the number of components, but also on the system’s structure and mutual interactions of the parts. They only indicate differences of complexity by the number of parts for the sake of the argument.

[vi] Indeed, the previous contributions to the emergence debate by one of the authors (C. N. El-Hani) face many problems identified by Hulswit (2005). See, for instance, El-Hani & Pereira (2000), El-Hani & Emmeche (2000), El-Hani & Videira (2001).

[vii] As we argued above, this is the basis for irreducibility as the non-deducibility of the behavior of a system’s parts.

[viii] Notice that theoretically-unpredictable structures or properties can be inductively-predictable (Kim 1999), given that, once a structure or property appears for the first time, it is possible that further occurrences of that structure or property are adequately predicted, given the thesis of synchronic determination. Moreover, “in principle” unpredictability is introduced in opposition to “practical” unpredictability, which is dependent on our cognitive limitations and state of knowledge.