A semiotic analysis of the genetic information system*

CHARBEL NIÑO EL-HANI, JOÃO QUEIROZ,
and CLAUS EMMECHE

Abstract

Terms loaded with informational connotations are often employed to refer to genes and their dynamics. Indeed, genes are usually perceived by biologists as basically 'the carriers of hereditary information.' Nevertheless, a number of researchers consider such talk as inadequate and 'just metaphorical,' thus expressing a skepticism about the use of the term 'information' and its derivatives in biology as a natural science. First, because the meaning of that term in biology is not as precise as it is, for instance, in the mathematical theory of communication. Second, because it seems to refer to a purported semantic property of genes without theoretically clarifying if any genuinely intrinsic semantics is involved. Biosemiotics, a field that attempts to analyze biological systems as semiotic systems, makes it possible to advance in the understanding of the concept of information in biology. From the perspective of Peircean biosemiotics, we develop here an account of genes as signs, including a detailed analysis of two fundamental processes in the genetic information system (transcription and protein synthesis) that have not been made so far in this field of research. Furthermore, we propose here an account of information based on Peircean semiotics and apply it to our analysis of transcription and protein synthesis.

Keywords: *gene; information; process philosophy; semiosis; biosemiotics; C. S. Peirce.*

1. Introduction: The gene concept and its problems

The gene concept has certainly been one of the landmarks in the history of science in the twentieth century. Keller (2000), for instance, refers to the twentieth century as 'the century of the gene.' Grós (1989) claims that we live in a 'civilization of the gene.' The term 'gene' was introduced

Semiotica 160–1/4 (2006), 1–68
DOI 10.1515/SEM.2006.039
0037–1998/06/0160–0001

by the Danish geneticist W. L. Johannsen, who regarded it as a kind of accounting or calculating unit; a very handy term but with no material counterpart that could be related to it with any degree of confidence (Johannsen 1909; cf. Falk 1986). Indeed, in the beginnings of genetics, an instrumentalist view about the status of 'gene,' as a theoretical concept prevailed (Falk 1986). The 'gene' was often regarded as nothing but a useful abstract concept to express regularities in the transmission of phenotypic traits.

Nevertheless, a realist, material view about the status of 'gene' was also found in classical Mendelian genetics. Herman J. Muller, for example, advocated the idea that genes were material units in their own rights, even though they could only be recognized through their effects. As Falk (1986) convincingly argued, the tension between instrumentalist and realist attitudes towards the status of the gene concept resulted in a fertile dialectics, described by him as a development on the pattern of 'Russian dolls,' in which discoveries about the chemical nature of the gene led, in turn, to the elaboration of new functional definitions, which, in turn, led to the investigation of a deeper structural meaning, which, in turn, led to a still deeper level of functional meaning, and so on.

Genes were regarded in classical genetics as units of recombination, function, and mutation. However, as a result of the development of the understanding of the gene on the pattern of 'Russian dolls,' it became eventually clear that genes were not units of either recombination or mutation. In the end, the prevailing meaning of the term in the twentieth century was that of a gene as a 'unit of function.' But, after the proposal of the double helix model and the flourishing of molecular biology, the gene was redefined as a material entity, concretely existent in DNA, and it became widely accepted to think of the gene also as a structural unit. Finally, the introduction of an informational vocabulary in molecular biology and genetics resulted in the so called 'information talk,' and genes came to be often regarded also as informational units, leading to what has been called the informational conception of the gene, a very popular notion in textbooks, in the media, and in public opinion. What is meant by 'information' in this case is merely sequence information in DNA or proteins (Sarkar 1998), an idea we will challenge throughout this paper.

With the proposal of the double helix model of DNA by James Watson and Francis Crick in 1953, a realist view about the gene prevailed. DNA was established as the material basis of inheritance, and the road to the so-called classical molecular gene concept was paved. Indeed, the *classical molecular gene concept*, according to which *a gene is a sequence of DNA that encodes a functional product, a polypeptide or an RNA*, can be seen as an outgrowth of the advances of molecular biology in the 1950s and

1960s. Genes seemed to be reducible, then, to concrete entities at the molecular level, namely, strings of DNA, and the structural and functional definitions of the gene were focused on a single entity (Stotz et al. 2004), resulting in a model with remarkable heuristic power.

The classical molecular gene concept is closely connected with the 'central dogma of molecular biology,' conceived as a statement about the 'flow' of 'information' in a cell. In a manner that dramatically shows the strong reductionist tendency that marked molecular biology since its beginnings (although this science seems to be gradually adopting a less reductionist view in recent years), the very idea of the dogma was that DNA makes RNA, RNA makes proteins, and proteins make the organism (see Crick 1958). But Crick also expressed the dogma more carefully as follows: 'once [sequential] information has passed into protein, it cannot get out again' (Crick 1958: 152–153). This 'dogma' became one of the elements in the hard core of molecular biology as a research program. In this context, the problem that no clear conception of 'information' is available in biological thought becomes quite central to molecular biology.

Since the beginnings of molecular biology, 'information' was conflated or simply identified with a string of DNA constituting a 'gene.' When information is conceived as sequences of nucleotides in DNA, we find ourselves in a difficult position to identify other kinds of information in a cell or even in the organism as a whole. Even if we point out to other 'informational' molecules, such as RNAs and proteins, the 'information' they allegedly 'contain' or 'carry' can be directly traced down, through the central dogma of molecular biology, to DNA. When information is conceptualized this way, DNA becomes a sort of reservoir from where all 'information' in a cell flows and to which it must be ultimately reduced. Our understanding becomes, so as to say, seduced by this purported 'information reservoir' and we tend, then, to overplay the role of DNA in cell systems, turning it into a complete 'program for development' or an all-powerful 'controller' of cell metabolism. But, as we are enchanted by this quite controversial picture of the role of DNA,[1] we simply forget that DNA seems to play the role of a set of data rather than that of a program in cell systems (Atlan and Koppel 1990); or, to put it differently, that DNA is a source of materials for cells, playing a role that is *obviously important*, but cannot be correctly described as that of a sort of master agent (or master molecule) in cell processes (Nijhout 1990). It is not DNA that *does things to* the cell; rather, it is the cell that *does things with* DNA.

The widespread usage of the informational conception of the gene makes the consequences of the understanding of genetic information as just sequential information in DNA go far beyond conceptual issues in

genetics and molecular and cell biology. Oyama (2000 [1985]) identifies a connection between the typical way of rendering the notion of 'genetic information' and genetic determinism, which has important consequences for the public understanding of science and a whole series of social, economical, and political issues related to the knowledge and applications in the fields of genetics and molecular biology.

In sum, it is an important task to clarify the concept of information in biology. But to reach any worthy result in this task, we should employ appropriate conceptual and methodological tools. Biosemiotics (for introductions to biosemiotics, see, e.g., Hoffmeyer 1996; Kull 1999), still a somewhat neglected perspective in current debates about the gene concept, offers a theoretical 'toolbox' for dealing with the notion of information in biology that can help us reach a precise and coherent understanding of this central notion. We also believe biosemiotics makes it possible to formulate the notion of genetic information in a manner that does not lend support to genetic determinism.

As regards the gene concept, several discoveries in molecular biology, including transposons, split genes, alternative splicing, consensus sequences, overlapping and nested genes, mRNA editing, transplicing, etc., posed very difficult problems to the generic or consensus view of genes, much in line with the classical molecular gene concept. These discoveries led, in Falk's (1986: 164) words, to '... an age of anarchy in the instrumental formulation of genetic entities,' in which a great number of heterodox entities was admitted into the 'expanding zoo of genetic units.' It was realized that the gene is neither discrete (there are overlapping and nested genes), nor continuous (there are introns within genes). It does not necessarily have a constant location (there are transposons), and it is neither a unit of function (there are alternatively spliced genes and genes coding for multifunctional proteins), nor a unit of structure (there are many kinds of *cis*-acting sequences affecting transcriptions [promoters, enhancers, terminators, etc.], split genes, and so on [cf. Falk 1986: 169; Fogle 1990: 356–363]). In this scenario, the question 'What is a gene, after all?' became a topic of strong debate in the philosophy of biology (for reviews about these discoveries and the problems they bring to the gene concept, see, for example, Falk 1986; Portin 1993; Keller 2000; Fogle 1990, 2000).

But it is not only in the philosophy of biology where we find a growing recognition of the problems surrounding the gene concept. Doubts about the status of this concept are also found in empirical papers within molecular biology (possibly indicating a crisis in molecular biology as a 'normal science'). To quote just two recent examples, we find Wang et al. (2000), in a study of the origin of a particular gene and the complex

modular structure of its parental gene, claiming that this structure '... manifests the complexity of the gene concept, which should be considered in genomic research' (Wang et al. 2000: 1294), for example, when one tries to predict a gene from genome data (Wang et al. 2000: 1300). Kampa et al. (2004) considers that their findings in an in-depth analysis of the transcriptome (the set of all transcripts of a cell) of human chromosomes 21 and 22 '... strongly support the argument for a reevaluation of the total number of human genes and an *alternative term for* "*gene*" to encompass these growing, novel classes of RNA transcripts in the human genome' (Kampa et al. 2004: 331; emphasis added). Although they do not suggest that we should abandon the term 'gene' altogether (as, for instance, Keller 2000; see below), they comment that '... the use of the term "gene" to identify all the transcribed units in the genome may need reconsideration, given the fact that this is a term that was coined to denote a genetic concept and not necessarily a physical and measurable entity. With respect to the efforts to enumerate all functional transcribed units, it may be helpful to consider using the term "transcript(s)" in place of gene' (Kampa et al. 2004: 341).

In the last three decades, a realist, material view of the gene has been superseded by a pluralist view that was captured by Falk in the following statement: 'Today the gene is not *the* material unit or *the* instrumental unit of inheritance, but rather *a* unit, *a* segment that corresponds to *a* unit-function as defined by the individual experimentalist's needs' (Falk 1986: 169). Ambiguities have been, however, a feature of the gene concept throughout its whole history (Kitcher 1982; Falk 1986) and they even have been heuristically useful in the past. Even though Falk is uncertain as to whether or not the current ambiguities will also be helpful, he does not seem to consider the sort of attitude he describes in current scientists as a reason to lose our hope as regards the status of the gene concept (see Falk 1986, 2000, 2001). Other researchers, however, consider that the conceptual variation currently observed in the case of the gene can lead to confusion (e.g., Fogle 1990, 2000).

As the twentieth century came to a close and we entered what seems to be a whole new era in biological research, the future of the gene didn't look bright for some thinkers. Keller (2000), for instance, considered the gene a concept 'in trouble' and suggested that maybe the time was ripe to forge new words and leave that concept aside (see also Portin 1993; Gelbart 1998). Although some authors agreed with Keller's proposal (e.g., Rios 2004), it has not found wide acclaim; rather, it was rejected by many reviewers of her book, such as Coyne (2000), Magurran (2000), Maynard Smith (2000b), Hall (2001), and Wilkins (2002). Symptomatically, other philosophers of biology and also practicing scientists foresee

a brighter future for the gene concept. Falk, for instance, takes a more optimistic view: while admitting that the gene is a concept 'in tension' (Falk 2000), he seeks ways to 'save' it (Falk 2001). Waters is even more optimistic, considering that different definitions of the gene can be unified by a concept with a number of 'open' clauses, such as that of 'a gene for a linear sequence in a product at some stage of genetic expression' (Waters 1994: 178). Hall (2001) is also optimistic, arguing that, despite published obituaries (Gray 1992; Neumann-Held 1999; Keller 2000), the gene is not dead, but alive and well, even though 'orphaned,' 'homeless,' and seeking a haven from which to steer a course to its 'natural' home, the cell as a fundamental morphogenetic unit of morphological change in development and of evo-devo (the interface between evolution and development).

The attempts to save the gene also led to distinctions between different concepts, as, for instance, Griffiths and Neumann-Held's (1999) distinction between the 'molecular gene' and the 'evolutionary' gene, and Moss' (2001, 2003) distinction between gene-P (the gene as a determinant of phenotypes or phenotypic differences) and gene-D (the gene as a developmental resource).[2] Moss forcefully argues that genes can be productively conceived in these two different ways, '*albeit with nothing good resulting from the conflation of the two*' (Moss 2001: 85). Gene-P, on the one hand, is the '... expression of a kind of instrumental preformationism' (Moss 2001: 87), showing its usefulness due to the epistemic value of its predictive power and its role in some explanatory games of genetics and molecular biology. In these terms, Moss doesn't attack the much criticized construct of the 'gene for' one or another phenotypic trait, recognizing its value for some theoretical and empirical tasks. Rather, the focus of his criticism is on the tendency to conflate this first conceptualization of the gene with a second one, that of gene-D. A gene-D is conceived, in a more realist tone, as a developmental resource defined by a specific molecular sequence and functional template capacity, which plays an entirely different explanatory role, in comparison to that of gene-P. Gene-P and gene-D are, in short, distinct concepts with different conditions of satisfaction for what it means to be a gene.

We will be specifically interested in gene-D in the present paper. Our task here is to begin the construction of a theoretical framework for a semiotic analysis of the concepts of 'gene' and 'information,' on the grounds of a case study about protein-coding genes.[3] We should emphasize the originality of this approach, not only in the specific context of molecular biology, but also in the general context of biosemiotics. We think it is important to develop biosemiotics by providing new sets of modeling tools and some exemplars or case studies to understand the precise sense in which specific life processes can be conceived as involving the

action of signs, as generally claimed by biosemioticians. Furthermore, by applying the formal notion of semiosis to model some aspects of the genetic information system, we intend to produce a radically new explanation of 'genetic information' as a semiotic process. In this effort, we will move towards a reinterpretation of what is information in a cell that hopefully avoids a number of problems detected in information talk not only in biology but also in science as a whole.

2. Biosemiotics and information talk in biology

During the 1950s and 1960s, genetics and cell and molecular biology were swamped by terms borrowed from information theory. This 'information talk' still pervades these fields, including widely used terms such as 'genetic code,' 'messenger RNA,' 'transcription,' 'translation,' 'transduction,' 'genetic information,' 'chemical signals,' 'cell signaling' etc. As the concept of information and its plethora of associated notions were introduced in biology, so did several problems with which the tradition of biology was unprepared to cope. Instead of deepening the discussion about the problems involved in information talk, the trend in the biological sciences was one of treating 'information' as merely sequence information in DNA or proteins (Emmeche and Hoffmeyer 1991; Sarkar 1998).

As a result, 'information' turned into one of the most important but problematic concepts in biology (see Oyama 2000 [1985]; Stuart 1985; Sarkar 1996; Griffiths 2001; Jablonka 2002). The concept of information in biology has been recently a topic of substantial discussion (see, for example, Maynard Smith 2000a; Godfrey-Smith 2000; Sarkar 2000; Sterelny 2000; Wynnie 2000; Jablonka 2002; Adami 2004). Furthermore, the evolution of new kinds of information and information interpretation systems in living beings has received a great deal of attention recently (see, for example, Jablonka 1994; Jablonka and Szathmáry 1995; Maynard Smith and Szathmáry 1995, 1999; Jablonka, Lamb, and Avital 1998). The evolution of different ways of storing, transmitting, and interpreting 'information' can even be regarded as a major theme in the history of life (Maynard Smith and Szathmáry 1995, 1999; Jablonka 2002).

Shannon and Weaver's highly influential 1949 book *The Mathematical Theory of Communication* showed how one can define the amount of information as the measure of the probability of selection of a particular message among the set of all possible messages. The probabilistic measure of information provided by this theory is non-semantic, indifferent to meaning (Shannon and Weaver 1949: 31; Cover and Thomas 1999; Jablonka 2002). There is controversy about the prospects of such a

non-semantic understanding of information in biology. Jablonka (2002), for instance, argues that the meaning-free concept of information theory is not sufficient for the understanding of information in biology by pointing out that, for instance, a DNA sequence encoding a functional enzyme and a same-length sequence coding for a completely non-functional enzyme would contain, according to the above-mentioned measure, the same amount of information. In biology, a semantic and pragmatic concept of information is necessary. Nevertheless, while in the case of the gene, a number of definitions have been coined and discussed, semantic and pragmatic concepts of information have been rarely defined in biology (Jablonka 2002). Moreover, several authors, particularly Susan Oyama (2000 [1985]), argued that the usual way of applying the concept of information to biological systems raises a number of important problems (see above).

A number of researchers consider information-talk as inadequate and 'just metaphorical,' thus expressing a skepticism about the use of the term 'information' and its derivatives in biology as a natural science. First, exactly because the meaning of that term in biology is not as precise as it is, for instance, in the mathematical theory of communication. Second, because it seems to refer to a purported semantic property of genes without theoretically clarifying if any genuinely intrinsic semantics is involved. Stuart (1985) and Sarkar (1996), for instance, argued that information talk should be eliminated from biology, since 'information' is a foreign metaphor in this science and its use may lead to erroneous views of explanation in fields such as molecular biology.

By assuming a biosemiotic point of view, we disagree with this position, claiming instead that the notion of information and other related ideas grasp some fundamental features of biological systems and processes that might be otherwise neglected. The concepts of 'code,' 'information,' 'signals,' 'message,' 'signaling,' 'transduction' and so on can be seen as necessary to understand the organization of relations in living beings in such a way that makes it clear that what happens in such beings is much more than simple chemistry.

For instance, Bray, in a symposium about reductionism in 1997, argued that as 'about fifty percent of the genome of a multicellular organism may code for proteins involved in cell signaling, ... organisms can be viewed as complex information-processing systems, where molecular analysis alone may not be sufficient' (cited in Williams 1997: 476–477). Similarly, Nurse argues that 'there's a need to realize that information may be transmitted in ways that may be lost by studying molecules alone,' and, furthermore, that 'it may not be possible or even necessary to explain all cellular phenomena in terms of precise molecular interactions'

(cited in Williams 1997: 476–477). These statements indicate that a reason why we can say that more than just chemistry is taking place in living beings lies in the fact that these systems process 'information' in quite complex ways, as Signs are produced, communicated, interpreted, translated, etc. In other words, biological meaningfulness is emerging all the time in such systems.

It is not surprising, then, that biologists felt the need to talk about 'information' as they were delving more and more into the molecular microstructure of living systems. It was just the case that they needed a way of conveying the idea that, even though all cellular processes have physical-chemical properties, more than just physics and chemistry is going on there. In this context, it is quite difficult to see what would be the real advantage of stripping off biology of information talk, instead of making it more precise and exploring its consequences in more depth.

The concept of information and related notions in biology should not only be taken seriously, but also clarified by employing appropriate conceptual tools. The use of semiotic concepts and theories to interpret information talk can significantly contribute to a precise and coherent formulation of the notion of information in biology. A semiotic treatment of information talk in biology can significantly contribute to an understanding of the role of genes in biological systems which avoids the reference to notions much criticized such as genetic 'blueprints' and 'programs,' while preserving the concept of 'information,' albeit radically reinterpreted. As we will see below, such a treatment lends support to the now widely accepted idea that there is more to information in living systems than just genes (see, for example, Jablonka 2002).

As Griffiths (2001) sums up, 'genetic information' is a metaphor in search of a theory. In this paper, we intend to make a contribution for the construction of this theory, by developing an account of genes as Signs and a semiotic modeling of information in biological systems. Both steps are fundamental, in our view, to the construction of a theory of information in biology.

We will concentrate our efforts in this paper on genetic information simply for methodological reasons. Although there are several other information systems in living beings, the genetic information system offers a good starting point for a semiotic treatment of information in biology, given its central role in biological thinking. Nevertheless, we should not simply extrapolate the conclusions taken from an analysis of this peculiar system for all other types of information systems in living beings. As Jablonka (2002: 579) argues, the genetic system, despite its importance, is highly specific and unusual, and, therefore, should not be taken as a prototype for thinking about information in biology. Accordingly, we do

not intend here to take the genetic information system as a prototype for other biological information systems. Rather, we consider its analysis just a first step in a research program aiming at a general semiotic analysis of information systems in living beings. Our intention is to proceed in subsequent works with semiotic analyses of other biological information systems, using the theoretical framework built here, but adapted to each specific case under analysis.

3. Information, meaning, and semiosis

When Peircean semiotics is used as a theoretical framework for case studies of specific meaning processes in biology, one should remember that the notion of Sign in Peirce is not the same as a simple 'unit' of information or communication as these terms are often used in several fields of research. It is a notion related to formal attempts to describe inferential processes in general, and it is not equivalent to the dyadic concept of representation in linguistics.

It is our primary aim here to apply some central general notions of Peirce's semiotics to understand the nature of genetic information. Nevertheless, such an application necessarily involves interpretation and, thus, decisions about how to see, for example, the relationship between what molecular biologists and biophysicists call forms of information processing (i.e., production and interpretation of Signs) in a complex living system such as the cell and forms of causality in that system. The analysis of the genetic information system given below is obviously not the only way to apply Peircean semiotics to this particular case; and some might object to the particular way we addressed the problem. In any case, we think that we have been faithful both to the basic insights and concepts of semiotics and to the findings of molecular biology, and that the few changes we have made in specific semiotic conceptions (as we shall explicate below) are necessitated by the growth of scientific knowledge about the system analyzed.

Peirce's conception of Semiotics as the 'formal science of signs' has had a deep impact in philosophy, psychology, theoretical biology, and cognitive sciences (see Freadman 2004; Fetzer 2001; Hookway 2002; Violi 1999; Houser 1997; Deacon 1997; Brunning and Forster 1997; Hoffmeyer 1996; Tiercelin 1995; Colapietro 1989; Freeman 1983; Jakobson 1969). Peircean semiotics is based on a theory of categories, including a list of categories (Firstness, Secondness, Thirdness) which can be logically described as an exhaustive system of hierarchically organized classes of

relations (monadic, dyadic, triadic) (Houser 1997; Brunning 1997). This system is the formal foundation of his 'architectonic philosophy' (Parker 1998) and of his model of *semiosis* (Sign action) (Murphey 1993: 303–306).

Peirce defined semiosis as an irreducible triadic relation between Sign-Object-Interpretant (S-O-I) (*EP* 2 2.171, *CP* 2.274; see Savan 1987–1988; Hookway 1992: 121). That is, according to Peirce, any description of semiosis involves a relation constituted by three irreducibly connected terms, which are its minimal constitutive elements (*MS* 318:81):

> My definition of a sign is: A Sign is a Cognizable that, on the one hand, is so determined (i.e., specialized, *bestimmt*) by something *other than itself*, called its Object, while, on the other hand, it so determines some actual or potential Mind, the determination whereof I term the Interpretant created by the Sign, that that Interpreting Mind is therein determined mediately by the Object. (*CP* 8.177)

Peirce conceives a 'Sign'[4] or 'Representamen' as a 'First' which stands *in such a genuine triadic relation* to a 'Second,' called its 'Object,' so as to be capable of 'determining a Third,' called its 'Interpretant,' to assume the same triadic relation to its Object in which it stands itself to the same Object (*CP* 2.274. See also *CP* 2.303, 2.92, 1.541). This triadic relation was regarded by Peirce as *irreducible*, in the sense that it is not decomposable into any simpler relation:

> ... by 'semiosis' I mean ... an action, or influence, which is, or involves, a co-operation of three subjects, such as a sign, its object, and its interpretant, this tri-relative influence not being in any way resolvable into actions between pairs. (*CP* 5.484)

One of the most remarkable characteristics of Peirce's theory of Signs is its dynamical nature. According to Merrell (1995: 78), 'Peirce's emphasis rests not on content, essence, or substance, but, more properly, on dynamics relations. Events, not things, are highlighted.' The complex S-O-I is the focal factor of a dynamical process (Hausman 1993: 72). Peirce was a truly process thinker (see Rescher 1996).

Sign processes are relationally extended within the spatiotemporal dimension, so that something physical has to instantiate or realize them. This means that Signs cannot *act* unless they are spatiotemporally realized (see Emmeche 2003; Ransdell 2003). If a Sign is to have any active mode of being, it must be materially embodied.

It is also important to avoid losing sight of the distinction between the interpreter, which is the system which interprets the Sign, and the Interpretant. The interpreter is described by Peirce as a 'Quasi-mind' (*CP*

4.536), a description which demands, for its proper interpretation, a clear recognition of Peirce's broad concept of 'mind' (Ransdell 1977; Santaella-Braga 1994). It is far from being the case that only conscious beings can be interpreters in a Peircean framework. Rather, a transcription machinery synthesizing RNA from a string of DNA or a membrane receptor recognizing a given hormone can be regarded as an interpreter in such a framework. A basic idea in a semiotic understanding of living systems is that these systems are interpreters of Signs; that is, that they are constantly responding to selected signs in their surroundings. The interpreter does not have to be a conscious being, not even an organism, as it may be some part or subsystem within an organism, or a humanly-designed product.[5]

We also need to consider here Peirce's distinctions regarding the nature of Objects and Interpretants (For a review of these topics, see Savan 1987–1988; Liszka 1990; Short 1996). He distinguishes between the Immediate and the Dynamical Objects of a Sign as follows:

> We must distinguish between the Immediate Object — i.e., the Object as represented in the sign — and ... the Dynamical Object, which, from the nature of things, the Sign *cannot* express, which it can only *indicate* and leave the interpreter to find out by *collateral experience*. (*CP* 8.314)

Or else:

> ... we have to distinguish the Immediate Object, which is the Object as the Sign itself represents it, and whose Being is thus dependent upon the Representation of it in the Sign, from the Dynamical Object, which is the Reality which by some means contrives to determine the Sign to its Representation. (*CP* 4.536)

And we should also take into account his distinction between the following two kinds of interpretants:[6]

> The *Immediate Interpretant* is the immediate pertinent possible effect in its unanalyzed primitive entirety ... The *Dynamical Interpretant* is the actual effect produced upon a given interpreter on a given occasion in a given stage of his consideration of the Sign. (*MS* 339d: 546–547).

Let us first consider Peirce's distinction between the Immediate and the Dynamic Objects of a sign. The Immediate Object of a Sign is the Object as it is immediately given to the Sign, the Dynamical Object in its semiotically available form. The Dynamical Object is something in reality that determines the Sign to its representation, and which the Sign can only

indicate, something that the interpreter should find out by collateral experience (*EP* 2 2.498).

In turn, Peirce defines the Dynamical Interpretant as the actual effect of a Sign, while the Immediate Interpretant is its 'range of interpretability' — the range of possible effects that a sign is able to produce (see Johansen 1993: 166–167). The Dynamical Interpretant is the instantiation of one of the possible effects established in the Immediate Interpretant. As the effect of the Sign upon the interpreter, the Dynamical Interpretant can be treated as being essentially equal to the significance of the Sign when seen in a dynamic and process-oriented perspective.

Peirce (see Fitzgerald 1966: 84; Bergman 2000) defined *meaning* as connected to the triadic relation as a whole (*EP* 2 2:429), as well as to different correlates of a triad — e.g., Object (*MS* 11, *EP* 2 2:274), Interpretant (*EP* 2 2:496, *EP* 2 2:499, *CP* 4:536). The notions of 'meaning,' 'information,' and 'semiosis' intersect and overlap in different ways (see Johansen 1993). For Debrock (1996), Peirce defined 'information' at least ordinarily (*CP* 2.418), metaphysically (*CP* 2.418), as a connection between form and matter, and logically (*W* 1.276), as the product of extension and intension of a concept. In this paper, we systematically refer to *information* as the communication of a *form* from O to I through S. The communication of a form amounts to the transference of a habit embodied in the Object to the Interpretant, so as to constrain (in general) the Interpretant as a Sign or (in biological systems) the interpreter's behavior. It should also be clear at this point that by 'communication' we mean more than mere transmission of a form.

Or, to put it in more detailed terms, the production of an effect of the Sign on the interpreter results from the communication of the form of the Object (as a regularity), via Sign, to the Interpretant. The Interpretant then becomes itself a Sign which refers to the Object in the same manner in which the original Sign refers to it (i.e., there is an invariance in the reconstruction of the form of the Object by the interpreter).

According to this approach, 'information' can be strongly associated with the concepts of 'meaning' and 'semiosis.' Peirce spoke of Signs as 'conveyers,' as a 'medium' (*MS* 793), as 'embodying meaning.' In short, the function of the Sign is to convey the form (*EP* 2 2:391):

> ... a Sign may be defined as a Medium for the communication of a Form ... As a medium, the Sign is essentially in a triadic relation, to its Object which determines it, and to its Interpretant which it determines ... That which is communicated from the Object through the Sign to the Interpretant is a Form; that is to say, it is nothing like an existent, but is a power, is the fact that something would happen under certain conditions. (*MS* 793:1–3. See *EP* 2:544, note 22, for a slightly different version)

What is a Form? There is a movement in Peirce's writings from 'form as firstness' to 'form as thirdness.' Form is defined as having the 'being of predicate' (*EP* 2 2.544) and it is also pragmatically formulated as a 'conditional proposition' stating that certain things would happen under specific circumstances (*EP* 2 2.388). It is nothing like a 'thing' (De Tienne 2003), but something that is embodied in the object (*EP* 2 2.544, note 22) as a habit, a 'rule of action' (*CP* 5.397), a 'disposition' (*CP* 2.170), a 'real potential' (*EP* 2 2.388) or, simply, a 'permanence of some relation' (*CP* 1.415).

We can say that Peirce follows a *via media* in which 'form' has the characteristics of both firstness and thirdness. This is in accordance with Bergman's (2000: 236) proposal of communicated *form* as a *First* of a *Third*. He based his proposal on the modalities associated with Firstness (possibility), Secondness (existence), Thirdness (habit, law), and on the principle of the interdependence of categories (see Potter 1997).

Peirce defines a Sign, in the passage quoted above, both as 'a Medium for the communication of a Form' and as 'a triadic relation, to its Object which determines it, and to its Interpretant which it determines.' If we consider both definitions of a Sign, we can then say that semiosis is a triadic process of communication of a *form* from the Object to the Interpretant by the Sign mediation (figure 1). Therefore, in this framework, we can say that semiosis is information, if we define this latter concept as above.[7] And, as *meaning* is also defined by Peirce as something communicated in semiosis (*NEM* 4: 309), it is plausible to also explain it as being associated with the *interpretant*, which, after all, will embody the reconstructed form of the Object.

Peirce (*CP* 8.177) writes that a Sign determines an Interpretant in some 'actual' or 'potential' Mind (in other passages, a 'quasi-mind.' See

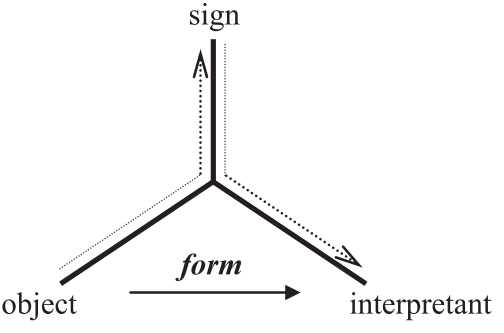

Figure 1. *Semiosis as the communication of a form from the Object to the Interpretant through the mediation of the sign*

CP 4.536). It is indeed possible to differentiate between 'potential' and 'effective' semiosis. Potential semiosis is defined as a triadically-structured process which is not taking place, which is only in potency. Effective semiosis, in turn, is a Sign in effective action, i.e., a Sign that, by being actualized, has an actual effect on the interpreter.

Following the distinction between potential and effective semiosis, we can define potential information as a process of communicating a form that could be realized in a given moment, while effective information is the communication of a form from an Object to an Interpretant through the Sign, i.e., a *Sign in effective action*.

The notion of information as form communicated from O to I through the mediation of S allows us to conceive it in a processual way, as a constraining factor of possible patterns of interpretative behavior. When applying this general semiotic approach to biological systems, information will most often be an *interpreter-dependent objective* process. It cannot be dissociated from the notion of a situated (and actively distributed) communicational agent (potential or effective). It is interpreter-dependent in the sense that information *triadically* connects representation (*Sign*), *Object*, and an effect (*Interpretant*) on the interpreter (which can be an organism or a part of an organism). The form — as a regularity in the Object — acts as a constraint on the interpreter's behavior, but the interpreter always reconstructs the form of the Object when interpreting a Sign. Nevertheless, the interpreter does so in such a manner that an invariance is retained, which makes possible, in fact, the very act of interpretation.

In sum, information in a biological system depends on both the interpreter and the Object (in which the *form* communicated in information is embodied as a constraining factor of the interpretative process). According to our interpretation of Peirce's remarks quoted above, information has a processual nature: information is the process of communicating a form from the Object to the Interpretant through the Sign.

As a way of stressing the difference between this account and more usual explanations about what is information, consider, for instance, Maynard Smith and Szathmáry's (1999: 9–10) argument that information is 'that something' which is conserved throughout a series of changes in the material medium underlying a communication process. We see this as resulting from a tendency to substantialize information. According to the account developed above, 'that conserved something' is not information, but rather an invariance in the reconstructed form. Information is rather the process by which a form is communicated through several different media (Signs) in such a way that an invariance is conserved throughout the process, even though the Object's form is continually reconstructed.

A framework for thinking about information as a process can be built in Peircean terms by employing the following definitions:

> [Information ≈ semiosis] A triadic-dependent process through which a form embodied in the Object in a regular way is communicated to an Interpretant through the mediation of a Sign.
>
> [Potential information ≈ potential semiosis] A process of communicating a form from an Object to an Interpretant through the mediation of a Sign that could take place in a given moment, changing the state of the interpreter.
>
> [Effective information ≈ effective semiosis] The process by which a Sign effectively exerts an effect (Interpretant) on some system (an interpreter) by making the Interpretant stands in a similar relation to something else (the Object of the Sign) as that to which the Sign stands, thus mediating the relation between Object and Interpretant. The Sign effectively communicates, thus, a form from the Object to the Interpretant, changing the state of the interpreter.

To formulate the above definitions in a sufficiently clear way, we should define what we mean by 'process.' 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: 38).

These definitions certainly raise several questions and face a number of difficulties when they are seen against the background of information theory. We shall leave to a subsequent paper, however, a discussion about such questions and difficulties.

4. Some basic notions about the genetic information system

It suffices for the analysis we perform here to present some very general notions about transcription, mRNA splicing, and protein synthesis. We will deliberately avoid introducing a large number of details, which can be easily found in any molecular and cell biology textbook (e.g., Griffiths et al. 1999; Lodish et al. 2003; Alberts et al. 2002; Lewin 2004).[8] This also means that we will keep our analysis simple in the present paper for methodological reasons.

Let us consider first a very simple model of the process of gene expression. During the synthesis of pre-mRNA, the four-base language of DNA (as a sequence of nucleotides including the bases adenine, A, guanine, G, cytosine, C, and thymine, T) is copied or 'transcribed' into the four-base

language of RNA (with uracil, U, replacing T). Transcription results in functional mRNAs (messenger RNA), rRNAs (ribosomal RNA), tRNAs (transfer RNA), snRNAs (small nuclear RNA), and scRNAs (small cytoplasmic RNA), but we will focus here on the synthesis of mRNA. Other functional RNAs that play important roles in various steps in DNA processing will be mentioned in passing.

During transcription, one DNA strand acts as a 'template,' determining by base pairing the order in which monomers (ribonucleoside triphosphates) are assembled to form a complementary RNA polymer, by a polymerization reaction catalyzed by the enzyme RNA polymerase.

The effects of a protein-coding gene on a given cell or organism are regulated mainly by control of gene expression at the level of transcription initiation. The transcription of a gene can be either *repressed*, when the corresponding mRNA and encoded protein or proteins are synthesized at low rates or not synthesized at all, or *activated*, when both the mRNA and encoded protein or proteins are, *ceteris paribus*, produced at much higher rates. Through the control of gene expression, only a subset of all genes present in any cell type in a multicellular organism is really expressed. Thus, from all the potential protein products a given cell type might have, only a specific number and variety will be present. This is the fundamental basis for cell differentiation in multicellular organisms.

In the end of the 1970s, it was found that eukaryotic genes are split into pieces of coding sequence, named 'exons,' separated by non-coding segments, named 'introns' (after Gilbert 1978). The discovery of split genes was one of the challenging discoveries that eventually led to the current debates about the gene concept. Now, it is well known that introns are common in multicellular eukaryotes, uncommon in many unicellular eukaryotes, and extremely rare in eubacteria and archaea. The vast majority of genes in multicellular eukaryotes contain multiple introns and the presence of such introns allows the expression of multiple related proteins from a single stretch of DNA by means of a process known as 'alternative splicing' (see below), which poses yet another challenge to the gene concept.

In eukaryotic protein-coding genes, introns are excised from a long 'primary transcript' (precursor mRNA or pre-mRNA), i.e., the RNA copy of an entire DNA sequence containing both exons and introns, in a process known as RNA 'processing,' which includes other events not described here. After the introns are excised, the coding exons are joined back together into a functional mRNA, which will be transported to the cytoplasm of the eukaryotic cell, where protein synthesis will take place.

Alternative splicing is rather common in mammalian genomes. Recent genome-wide analyses indicate that thirty-five to fifty-nine percent of

human genes produce alternatively spliced forms (Modrek and Lee 2002). Even though a significant portion of the predicted splicing variants are not functional, resulting from aberrant rather than regulated splicing, and, therefore, the frequencies of alternatively spliced gene products mentioned above are probably overestimated (Sorek et al. 2004), it is still the case that alternative splicing should be regarded as one of the most significant components of the functional complexity of the genome of our and many other species (Modrek and Lee 2002).

Alternative RNA splicing requires that the conceptualizations of genes move far beyond the simple scheme captured in formulas such as 'one gene-one protein or polypeptide.' One might argue, however, that such a challenge to the gene concept can be easily assimilated by simply replacing this formula by a new one, for instance, 'one gene-many proteins or polypeptides.' However, the situation is not so simple. As Keller (2000) argued, the situation is such that it does not allow us to be clear about where is the gene after all. For instance, should we call a 'gene' that piece of sequence in the DNA that can generate dozens of different proteins? Or should we apply this concept to each individual spliced mRNA by formulating such an idea as that of one mature mRNA-one protein? If we opt for the second alternative, a number of other problems will follow. For instance, the mRNA molecule itself can be further modified (RNA editing) and the final transcript can be assembled from exons derived from different pre-mRNAs (trans-splicing). More importantly, mRNAs are structures much more transient than quite basic (and, arguably, correct) intuitions about genes and their stability through generations require.

Alternative RNA splicing is an important mechanism for the production of different forms of proteins (isoforms) by different cell types. The fibronectin (FN) gene, for instance, generates more than twenty different FN isoforms. The FN gene has approximately 75,000 nucleotides (75 kb) and contains numerous exons. After the FN pre-mRNA is transcribed from DNA, it undergoes cell type-, development- and age-specific splicing. Each FN isoform is encoded by a differently, alternatively spliced mRNA, and, therefore, each isoform results from a unique combination of exons found in the FN gene (see figure 2).

The combinations of exons in each isoform change its causal dispositions. This can be clearly seen in the case of the splicing of FN pre-mRNA in fibroblasts and hepatocytes. In fibroblasts, splicing of the FN pre-mRNA results in mRNAs containing exons EIIIA and EIIIB. The fibroblast FN isoform contains amino acid sequences that bind tightly to proteins in the plasma membrane, ascribing it specific causal dispositions. This specific FN isoform contributes to the adhesion of fibroblasts to the

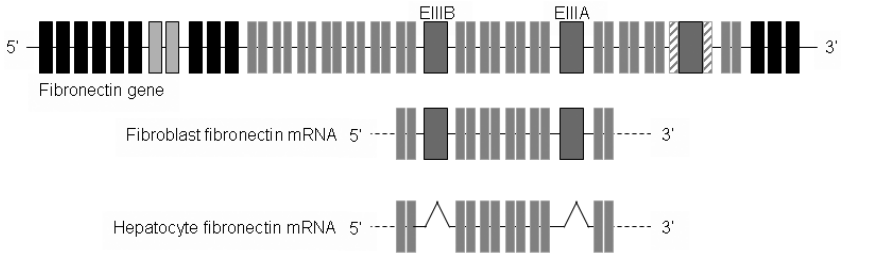

Figure 2. *Cell type-specific splicing of fibronectin pre-mRNA in fibroblasts and hepatocytes. The 75-kb FN gene (top) contains multiple exons. Introns are shown in the diagram as thin lines and are not drawn to scale. Most of the introns are much longer than any of the exons. The FN mRNA produced in fibroblasts includes the EIIIA and EIIIB exons, whereas these exons are spliced out of FN mRNA in hepatocytes (Redrawn from Lodish et al. 2003).*

extracellular matrix. In hepatocytes, the major cell type in the liver, cell-type specific splicing results in functional FN mRNAs lacking exons EIIIA and EIIIB. As in the case of fibroblasts, we have here a FN isoform with specific causal dispositions. First, it does not show the causal dispositions the fibroblast isoform shows: FN secreted by hepatocytes does not adhere tightly to fibroblasts or most other cell types. The lack of such causal dispositions is very important to the functionality of this FN isoform, as it allows it to freely circulate in the blood stream. Nevertheless, when the wall of a vase is ruptured, hepatocyte FN plays a fundamental role in the formation of blood clots, showing its specific causal disposition, which result from the presence in the protein of fibrin-binding domains, amino acid sequences that bind to fibrin, one of the main constituents of blood clots. When hepatocyte FN is bound to fibrin, it shows yet another causal disposition, interacting with integrins, cell-adhesion protein molecules found in the membranes of activated platelets. As a result, the blood clot is expanded through the addition of platelets.

The effects of genes on the functioning of a cell or organism can also be regulated by means of alternative pre-mRNA splicing, so as to produce different gene products from the same pre-mRNA. Particularly remarkable examples of genetic regulation at the level of RNA splicing are found, for instance, in the sex determination pathway of *Drosophila* (for a review, see, for example, Black 2003).

Finally, translation is an essential part of protein synthesis, consisting in the process by which the nucleotide sequence of an mRNA serves as a template for the synthesis of a polypeptide chain, i.e., for a series of events in which amino acids are ordered and joined to form the primary structure of a protein. Three types of RNA molecules are involved in

translation, performing different but cooperative functions. mRNAs are the 'vehicles' of the genetic information transcribed from DNA. The 'message' at stake is 'written' in the form of a series of three-nucleotide sequences, called 'codons,' each of which specifying a particular amino acid. tRNAs play a fundamental role in the process of deciphering the codons in mRNA. Each type of amino acid has its own subset of tRNAs. They act as transporters, binding amino acids and carrying them to the growing end of a polypeptide chain in response to specific codons in the mRNA. The reason why the correct tRNA with its attached amino acid is selected at each step in protein synthesis lies in the fact that each specific tRNA molecule contains a three-nucleotide sequence, called an 'anticodon,' that base-pairs with its complementary codon in the mRNA. In this manner, for each specific codon in mRNA a specific amino acid, carried by a specific tRNA, is included in a polypeptide chain, according to the rules expressed in the almost universal 'genetic code.' Along with 100 different proteins, several types of rRNA are components of ribosomes, the complex and large macromolecular structures that act, so as to say, as guides to coordinate the assembly of the amino acid chain of a protein. In fact, an rRNA (a ribozyme), and not a protein, is probably the catalyst involved in the formation of peptide bonds in protein synthesis.

Translation involves three stages: initiation, when ribosomal units assemble near the translation start site in the mRNA with the tRNA carrying the amino acid methionine base-paired with the start codon, most commonly AUG; chain elongation, in which a four-step cycle is repeated, involving the binding of a tRNA carrying an amino acid, the release of the tRNA involved in the previous step in the elongation, transfer of the growing polypeptide to the incoming amino acid catalyzed by one of the rRNAs, and translocation of the ribosome to the next codon in the mRNA; and termination, in response to stop codons UAA, UGA, and UAG.

Recognition of a codon in mRNA specifying a given amino acid by a particular tRNA is, in fact, the second step in 'decoding' the genetic 'message.' The first step is the attachment of the appropriate amino acid to a tRNA in a reaction catalyzed by a specific aminoacyl-tRNA synthetase. The specificity of the attachment between amino acids and tRNAs results from the capacity of each one of these enzymes of recognizing *one* amino acid and *all* its compatible, or 'cognate,' tRNAs. Therefore, the rules captured in the genetic code ultimately depend on the recognition activity of aminoacyl-tRNA synthetases.

Although the terms 'translation' and 'protein synthesis' are usually employed interchangeably, this is not correct, since, although translation is obviously an essential step in protein synthesis, this process involves further steps. Polypeptide chains undergo post-translational folding and

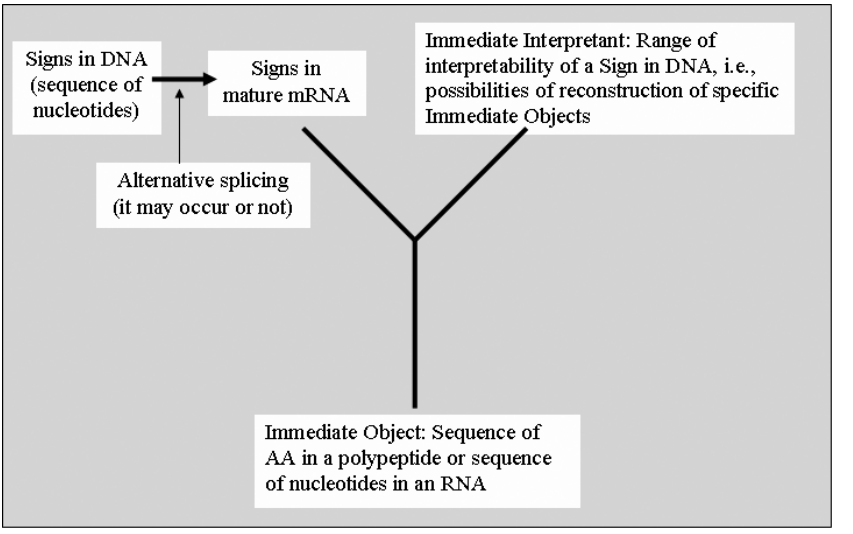

Figure 3. *A general semiotic analysis of the gene as a Sign*

often other changes, as, for instance, chemical modifications and association with other polypeptide chains, which are required for production of functional proteins. All these steps in protein synthesis can undergo regulation.

5. A semiotic analysis of genes and genetic information: First take

If we take Peirce's concepts of Sign and semiosis as bases for analyzing what is a gene, it will be the case that the action of a gene as a Sign will have to be understood as a relationship between three elements (figure 3). Given the definition of information proposed in section 3, genetic information can be described as a semiotic process. In these terms, we should conclude that there's more to genetic information than just the sequence of nucleotides in a piece of DNA.

In this picture, a string of DNA is a Sign. In this sense, the FN gene can be treated as a Sign. As a protein-coding gene, it stands — in a triadic-dependent relation — for a specific sequence of amino acids (Immediate Object) — one of the FN isoforms, translated out of a mature mRNA after alternative splicing (which, as figure 3 shows, can take place or not, depending on the string of DNA we are analyzing)[9] — through a process of reconstruction of a specific form (Interpretant).[10]

A Sign, after all, is the mediating element in a semiotic process through which a form is communicated from an Object to an Interpretant. This is the reason why we consider the Interpretant here as the reconstruction of a form (habit) which was embodied in an Object. To be more explicit, we defined the information above as the communication of a form from the Object to the Interpretant, and we also argued that such a communication constrains the behavior of the interpreter. What we mean by 'reconstruction' here is a process by which the form of a protein in a cell generation is communicated through Signs in DNA (in potency) to the form of a protein in the next cell generation, and the latter constrains the behavior of the cell as an interpreter. Thus, a regularity obtains (with obvious evolutionary consequences) in the three-dimensional structure and the function of proteins over generations.

We will introduce the qualifiers 'Composite' and 'Simple' to incorporate a part-whole relationship in the semiotic analysis of genes developed here, referring to a stretch of DNA or mature RNA as a whole as a Composite Sign, formed by clusters of Simple Signs, codons. We can now turn to a first refinement in our analysis, introducing the distinction between Immediate and Dynamical Object, and Immediate and Dynamical Interpretant in a systematic way.

The Dynamical Object of a gene is a functional, folded, and chemically modified protein, which is often not entirely specified in the sequences of nucleotides or amino acids, but it is rather indicated by such sequences. Functional proteins are not always simply translated out of nucleotide sequences by a cell, but they are rather found out through resources the cell acquire by collateral experience, i.e., by habits that a cell acquire in its development towards the states characteristic of a given cell type, and can be traced back to evolutionary processes.[11] A functional FN isoform, for instance, is a Dynamical Object.

The Composite Immediate Object of a protein-coding gene is the sequence of amino acids of a polypeptide, as this is the object represented in the gene's vehicle, a string of DNA.[12] Each amino acid, in turn, is a Simple Immediate Object. If we consider the sequence of amino acids of a specific FN isoform, we will say, in the terms of our analysis, that such a sequence is an Immediate Object of the FN gene. It is important to bear in mind, however, that it is *an* Immediate Object, not *the* Immediate Object. After all, the FN gene codes more than twenty different FN isoforms, all of them being possible Immediate Objects of the FN gene as a Sign in DNA.

The sequence of amino acids, the Composite Immediate Object, is the Dynamical Object in its semiotically available form. The sequence of

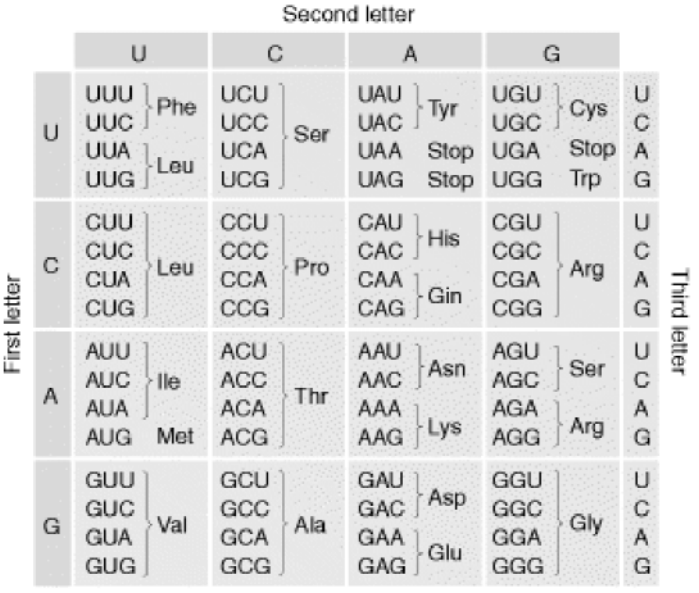

Figure 4. *The genetic code. Sets of three nucleotides (codons) in an mRNA molecule are translated into amino acids during protein synthesis according to the rules shown in the table above (from Griffiths et al. 1999).*

amino acids of each FN isoform amounts to a specific protein coded — in its semiotically available form — in a mature RNA which results, after splicing, from a pre-mRNA transcribed from the FN gene.

The Immediate Object, a sequence of amino acids, can indicate a range of possible functional proteins, Dynamical Objects, as a single amino acid sequence can be folded in different ways in different cellular contexts. But we should not lose from sight, however, that such an indication by the Immediate Object plays a fundamental role in the reconstruction of the Dynamical Object, since it is not the case that any three-dimensional protein can be produced from a given amino acid sequence.[13]

The Immediate Interpretant of a codon as a Simple Sign is the range of interpretability established by the rules of base pairing by which specific nucleotides in DNA determine specific nucleotides in mRNA, or the range of interpretability of three-nucleotide sequences in mature mRNA as established in the genetic code, a set of rules by means of which nucleotide sequences determine the addition of specific amino acids to a growing polypeptide chain (Figure 4). Symptomatically, 'coding' can be defined as a system of constraints which establishes a range of possible effects of a

Sign (see Nöth 1995: 210–211). The Dynamical Interpretant of a codon as a Simple Sign amounts, then, to the realization of one of the rules of base pairing or of the genetic code.

A Composite Sign in DNA determines a range of possible Composite Immediate Objects. It is true that there are cases in which a stretch of DNA codes for only one protein product. In this case, the Composite Sign in DNA determines only one Immediate Object. Nevertheless, in eukaryotic cells at least, most stretches of DNA codes for several distinct proteins, as in the case of the FN gene. Therefore, we can define the Immediate Interpretant of a Composite Sign as the range of interpretability of that Sign in DNA, i.e., as the possible Immediate Objects, the possible sequences of amino acids, that can be produced from that Sign in DNA. Alternative RNA splicing is understood, in these terms, as one of the processes that enrich the range of interpretability, the Immediate Interpretant, of a stretch of DNA. In the case of the FN gene, its Immediate Interpretant comprises more than twenty possible Composite Immediate Objects.

This analysis is in accordance with the definition of a Sign as medium for communicating the form of an Object to an Interpretant. The Interpretant can be seen, thus, as a reconstruction of the form of an Object. It follows that the Immediate Interpretant of a stretch of DNA or mRNA as a Composite Sign, i.e., its range of interpretability, amounts to the diversity of possibilities of reconstruction of the form of the Composite Immediate Object, the sequence of amino acids in a polypeptide.

The Dynamical Interpretant of a stretch of DNA or mRNA as a Composite Sign corresponds to the effective reconstruction of a sequence of amino acids. In an alternatively spliced gene, such as the FN gene, this realization involves the instantiation of a specific splicing pattern in a given cell type, at a given developmental stage. Thus, one of the possibilities established in the range of interpretability of a stretch of DNA, in its Immediate Interpretant, is actualized. In a fibroblast, for instance, when a specific Immediate Object is synthesized, the fibroblast-specific FN isoform, this means that, from the range of possible sequences of amino acids that might be made out of the FN gene — its Immediate Interpretant — a specific sequence was reconstructed — its Dynamical Interpretant.

After it is actualized, an Immediate Object indicates a particular Dynamical Object — say, a specific FN isoform —, which the cell finds out through habits acquired in evolution and development. It is the Dynamical Object, then, that has an effect on the cell as a global interpreter. We can define, then, a Dynamical Interpretant of the Dynamical Object, a particular effect on a cell, among a range of possible effects — the Imme-

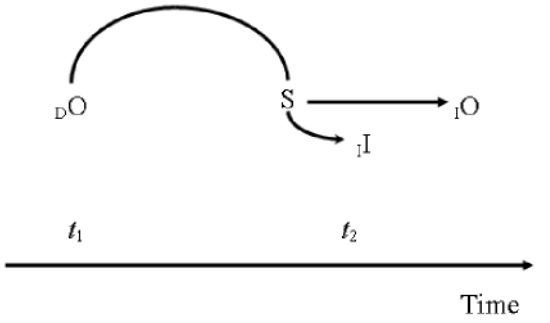

Figure 5. *The Dynamical Object (functional protein) as the primary constraining factor of semiosis in the genetic information system. S, Sign; $_DO$, Dynamical Object, $_IO$, Immediate Object; $_II$, Immediate Interpretant; t, generation time.*

diate Interpretant of the Dynamical Object. This Dynamical Interpretant is the actualization of one of the possible effects that a Composite Sign might have on the interpreter. Its range of interpretability is the Immediate Interpretant of the Composite Sign.

The analysis presented in this section faces the potential problem that it seems to treat the Sign as the primary constraining factor in semiosis, while this role is reserved for the Dynamical Object in Peirce's theory of Signs.[14] After all, we are describing here how S (a sequence of nucleotides in DNA) determines O (a sequence of amino acids in a polypeptide) through I (a range of possibilities of reconstruction of sequences of amino acids).[15] We accommodate this description to a Peircean framework by examining the constraining action of the Object in evolutionary terms (see figure 5). Consider two different generations of a population, in times t_1 and t_2, and a protein (Dynamical Object) in t_1 that increases the likelihood of successful, adaptive experiences of organisms possessing it. Therefore, that protein increases the likelihood that a gene (Sign) encoding it will be present in high frequencies in the next generation, in t_2. Indeed, the sequence of a gene is determined, by past natural selection, because of the effects it produces (Maynard Smith 2000: 177). This gene, in turn, will bring to the next generation the potency to produce that protein, as a Dynamical Object, by indicating it through its semiotically available form, its Immediate Object. This means that that gene, as a Sign, exerts a determining influence on the range of possibilities of reconstructing sequences of amino acids in the next generation. If we follow this set of ideas, we will be able to see how, in evolutionary terms, O determines I through S, in conformity with Peirce's account of semiosis.

Nevertheless, the role of O as the primary constraining factor of semiosis depends, in the genetic information system, on the role of S, in a given generation, in determining O through I. We can say, in short, that the fact that S determines O through I in a given population in t_2 is itself determined by the fact that O determined I by increasing the likelihood of S being present in a high frequency in t_2, by means of its involvement in successful experiences in t_1.

The relationship between Signs in DNA and the sequence of amino acids of a protein (the Composite Immediate Object) is established by a complex mechanism of interpretation, involving transcription, RNA processing and translation. Thus, to interpret a string of DNA, more than one interpretative system is required, including, for instance, RNA polymerases, involved in the transcription of DNA into RNA, and ribosomes, involved in the translation of mRNA into proteins. These interpretative systems are parts or subsystems of a cell as a global interpreter, and their actions are subordinated to the latter.

The idea that the cell can be seen as a global interpreter to which a series of interpretative subsystems in the genetic information system are subordinated is dramatically reinforced by recent analyses of the functional organization of proteomes. For instance, Gavin et al. (2002) showed that the vast majority of the protein complexes in yeast are associated with one another, directly or indirectly, through common proteins. As a researcher told Sampedro (2004: 61), it is as if 'the whole cell was a single machine.' More than half of the protein complexes analyzed by Gavin et al. are involved in the genetic information system: transcription/DNA maintenance/chromatin structure (twenty-four percent); RNA metabolism (twelve percent); protein synthesis/turnover (fourteen percent); signaling (nine percent); and protein/RNA transport (five percent). Even more interestingly, the multi-component cellular systems involved in transcription, RNA processing, and RNA transport do not form a simple linear assembly line, but a complex and extensively coupled network (Maniatis and Reed 2002). It is this network structure which makes it possible the coordination of the interpretative subsystems in the genetic information system by the cell. It is clear, then, that we cannot easily move from claims at the cell level to claims at the molecular level while pondering about which system is interpreting genes as signs. We think that these recent studies clearly show that, when a gene is interpreted, the interpretation process is indeed taking place at the cellular level, albeit multi-component molecular subsystems are necessary to this endeavor.

The idea that ultimately the whole cell participates in the network necessary for the interpretation that is demanded for the effect of a gene

product to take place (cf. Emmeche and Hoffmeyer 1991) is further supported by the role of an impressive array of signaling pathways regulating the interpretation of Signs in DNA. As Fogle (2000: 19) sums up, 'DNA action and function become meaningful in the context of a cellular system. Coding information in the DNA is necessary but insufficient for the operation of living systems.'

A Peircean approach to the gene concept entails that genetic structures should not be seen in isolation from the larger system by which they are interpreted. From this perspective, the meaning of a gene to its interpreter, the cell, or, to put it differently, the biological meaningfulness of a gene, is found not only in the entity that is normally identified with it, namely, a stretch of DNA. After all, there is more to genetic information than just a sequence of nucleotides in DNA. We will have to include the effect of the gene-as-a-Sign on the cell or organism, and, in fact, the very role of cellular subsystems as interpreters of strings of DNA, in such a way that they relate Signs to specific Dynamical Objects, proteins which play a function inside the cellular system and have an effect on it or on the organism of which the cell is a part.

The identification of genetic information with sequential information in DNA molecules makes it impossible to understand it as a triadic-dependent, semiotic process, as we propose here. In other words, in the classical molecular gene concept, information was often considered to be simply reduced to its vehicle, DNA, isolated from all the other elements in what we analyze as a triadic process that comprises the *action* of a gene as a Sign. We propose here that we should regard information as precisely this action of a gene as a Sign, understanding it as a process including more elements than just Signs in DNA.

In our view, this first-take semiotic analysis of the genetic information system leads to the following conclusions:

- Genes should be regarded as Signs in DNA, which can only have any effect on a cell through a triadic-dependent process (semiosis);
- This process *is* genetic information and involves more than just genes as Signs in DNA but also Objects and Interpretants;
- Genetic information is the process by means of which a form in a Dynamical Object (a functional protein) is communicated to an Interpretant (the reconstruction of a specific sequence of amino acids in a cell) by means of Signs in DNA.

In the next section, we will turn to a more detailed analysis of some processes in the genetic information system. The conclusions presented above will be both substantiated and extended in significant ways by this more fine-grained analysis.

6. Refining the semiotic analysis of genes and genetic information

Our strategy to refine the semiotic analysis presented above will consist in elaborating a case study of some processes involved in the genetic information system. As we intend to build a basic framework to be subsequently used in the analysis of a variety of sign processes in living systems, we will deliberately keep this case study simple at this first step, avoiding many biological details. It is for methodological reasons, then, that we will focus our attention here mainly on transcription and translation.

Let us begin by posing some questions concerning the semiotic analysis developed in this paper. First, does this analysis lead to new predictions about genetic and cellular systems? This is not the case for the moment being, but the use of the semiotic concepts and tools employed here to analyze other features of the genetic information system as well as other, less well-known information systems in living beings, such as the epigenetic cellular and organismic information systems (see Jablonka 2001, 2002), may eventually result in new predictions.

Second, does the semiotic analysis developed here lead to new insights into genetic and cellular systems? We believe the answer is in the affirmative, as this analysis allows us to explain more precisely what is 'information' in the genetic information system. This conclusion will become clearer after we refine the semiotic analysis of genes and genetic information.

6.1. *Levels of semiosis: A general model*

The semiotic analysis of the genetic information system can be further refined by considering that semiosis in cellular (and other kinds of) systems involves relationships at several levels. Here, we will model semiosis at three levels at a time, on the grounds of Salthe's (1985) 'basic triadic system,' clearly influenced by Peirce (see also Queiroz and El-Hani 2004, in press). The basic triadic system plays a fundamental role in Salthe's 'hierarchical structuralism,' conceived by him as a coherent and heuristically powerful way of representing natural entities. This role follows from the prospect of discovering by means of this system general rules and principles of constraint within which the laws of nature must operate.

According to the basic triadic system, to describe the fundamental interactions of a given entity or process in a hierarchy, we need (i) to consider it at the level where we actually observe it ('focal level'); (ii) to investigate it in terms of its relations with the parts described at the next lower level; and (iii) to take into account entities or processes at the next higher

level, in which the entities or processes observed at the focal level are embedded. Both the lower and the higher levels have constraining influences over the dynamics of the entities and/or processes at the focal level. These constraints allow us to explain the emergence of entities or processes (e.g., semiosis) at the focal level.

At the lower level, the constraining conditions amount to the 'possibilities' or 'initiating conditions' for the emergent process, while constraints at the higher level are related to the role of a (selective) environment played by the entities at this level, establishing the boundary conditions that coordinate or regulate the dynamics at the focal level.[16]

In this model, an emergent process at the focal level is explained as the product of an interaction between processes taking place at the next lower and higher levels.[17] The phenomena observed at the focal level should be '... among the possibilities engendered by permutations of possible initiating conditions established at the next lower level' (Salthe 1985: 101). Nevertheless, processes at the focal level are embedded in a higher-level environment that plays a role as important as that of the lower level and its initiating conditions. Through the temporal evolution of the systems at the focal level, this environment or context selects among the states potentially engendered by the components at the lower level those that will be effectively actualized. As Salthe (1985: 101) puts it, 'what actually will emerge will be guided by combinations of boundary conditions imposed by the next higher level.' Figure 6 shows a scheme of the determinative relationships in Salthe's basic triadic system.

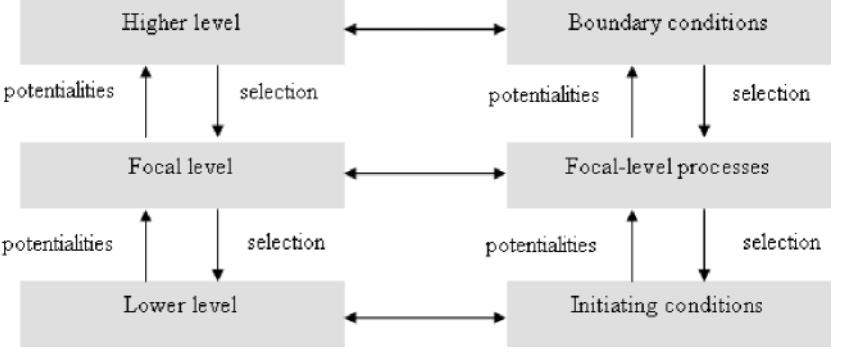

Figure 6. *A scheme of the determinative relationships in Salthe's basic triadic system. The focal level is not only constrained by boundary conditions established by the higher level, but also establishes the potentialities for constituting the latter. In turn, when the focal level is constituted from potentialities established by the lower level, a selection process is taking place, since among these potentialities some will be selected in order to constitute a given focal-level process.*

The term 'emergence' is often employed in an intuitive and ordinary way, referring to the idea of '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 et al. (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' (Stephan 1998, 1999). Here, we employ the concept of emergence and its derivatives in a rather technical sense, according to which 'emergent' properties or processes should be understood as a certain class of higher-level properties or processes related in a certain way to the microstructure of a class of systems. Salthe's triadic system entails that these higher-level properties or processes result from an interaction between constraints (initiating conditions) established by a lower level, on which the emergent properties or processes are grounded, and another set of constraints (boundary conditions), established by a higher level in which the focal-level emergent properties or processes are embedded.[18]

For the sake of the argument, let us begin by taking as the 'focal level' that level in which a given semiotic process is observed. Semiotic processes at the focal level are described here as *chains of triads*. We can treat, then, the interaction between semiotic processes at the focal level, potential determinative relations between elements at the immediately lower level ('micro-semiotic level'), and semiotic processes at the immediately higher level ('macro-semiotic level'). In the latter, *networks of chains of triads* which embed the semiotic process at the focal level are described. The micro-semiotic level concerns the relations of determination that *may* take place within each triad S-O-I. The relations of determination provide the way the elements in a triad are arranged in semiosis. According to Peirce, the Interpretant is determined by the Object through the mediation of the Sign (I is determined by O through S) (Peirce *MS* 318: 81). This is a result from two determinative relations: the determination of the Sign by the Object relatively to the Interpretant (O determines S relatively to I), and the determination of the Interpretant by the Sign relatively to the Object (S determines I relatively to O) (De Tienne 1992).

In the micro-semiotic level, we consider that, given the relative positions of S, O, and I, a triad $t_i = (S_i, O_i, I_i)$ can only be defined as such in the context of a chain of triads $T = \{\ldots, t_{i-1}, t_i, t_{i+1}, \ldots\}$ (see Gomes et al. 2003; Queiroz and El-Hani 2004). Semiosis, as a Sign in action, entails the instantiation of chains of triads. As Savan (1986: 134) argues, an Interpretant is both the third term of a given triadic relation and the first

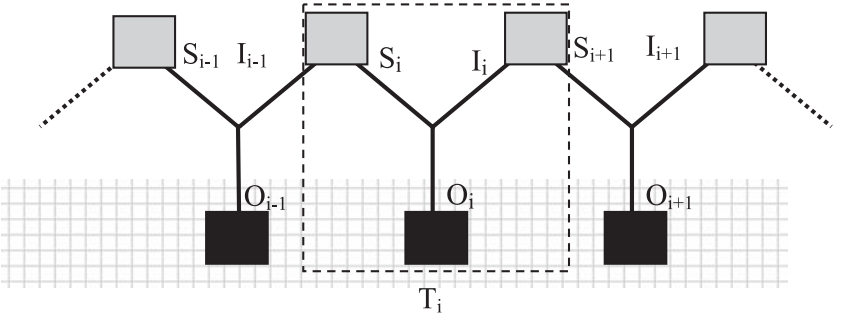

Figure 7. *Scheme showing that a triad can only be defined within a chain of triads. The grid at the bottom part of the figure shows that O_{i-1}, O_i, and O_{i+1} are Immediate Objects of the same Dynamical Object.*

term (Sign) of a subsequent triadic relation. This is the reason why semiosis cannot be defined as an isolated triad; it necessarily involves chains of triads (see Merrell 1995) (see figure 7).

In short, given the framework of Salthe's hierarchical structuralism, we should analyze semiosis by considering three levels at a time: Each chain of triads will be located at a focal level, and, correspondingly, we will talk about focal-level semiotic processes. Micro-level semiotic processes will involve the relations of determination within each triad. Macro-level semiotic processes will involve networks of chains of triads, in which each individual chain is embedded. Focal-level semiosis will emerge as a process through the interaction between micro- and macro-semiotic processes, i.e., between the relations of determination within each triad and the embedment of each individual chain in a whole network of sign processes.

Following Salthe's explanation of constraints, micro-semiosis establishes the initiating conditions for focal-level semiotic processes. Each chain of triads always indicates the same Dynamical Object, through a series of Immediate Objects, as represented in each triad (see figure 7). The potentialities of indicating a Dynamical Object are constrained by the relations of determination within each triad. That is, the way O determines S relatively to I, and S determines I relatively to O, and then how I is determined by O through S leads to a number of potential ways in which a Dynamical Object may be indicated in focal-level semiosis, i.e., to a set of potential triadic relations between Immediate Objects, Signs, and Interpretants.

We need to consider, then, the distinction between *potentiality* and *actuality* in the context of our analysis. For this purpose, we introduce

the definitions of *potential* Signs, Objects, and Interpretants. A 'potential Sign' is something that *may* be a Sign of an Object to an Interpretant, i.e., it may stand for that Object to an Interpretant. A 'potential Object,' in turn, is something that *may* be the Object of a Sign to an Interpretant. And, finally, a 'potential Interpretant' is something that *may* be the Interpretant of a Sign, i.e., it may be an effect of that Sign. The micro-semiotic level is the domain of potential Signs, Objects, and Interpretants.

We should consider, then, a whole set W of possible determinative relations between these three elements, which can generate, in turn, a set of possible triads. These triads cannot be fixed, however, by the micro-semiotic level, since it establishes only the initiating conditions for chains of triads at the focal level. To fix a chain of triads, and, consequently, the individual triads which are defined within that chain, boundary conditions established by the macro-semiotic level should also play their selective role. That is, networks of chains of triads constitute a semiotic environment or context that plays a fundamental selective role for the actualization of potential chains of triads. Chains of triads are actualized at the focal level by a selection of those triads that will be effectively actualized amongst those potentially engendered at the micro-semiotic level. After all, as we saw above, a triad $t_i = (S_i, O_i, I_i)$ cannot be defined atomistically, in isolation, but only when embedded within higher-level structures and/or processes, including both chains of triads $T = \{\ldots, t_{i-1}, t_i, t_{i+1}, \ldots\}$ and networks of chains of triads $ST = \{T_1, T_2, T_3, \ldots, T_n\}$. In short, these structures and/or processes provide the context for the actualization of potential determinative relations within each chain.

Considering the dynamics of semiotic processes at the focal level, we can say that the temporal evolution of such processes is determined by events of actualization of potential chains of triads and potential triads. Triads are actualized, realizing a specific chain at the focal level, through the operation of two constraints. First, potential determinative relations (*initiating conditions*) at the micro-semiotic level constrain the universe of potential chains of triads, given that the whole set W of possible determinative relations between potential Signs, Objects, and Interpretants is always smaller than the universe U of all potentially existent triads. That is, given the initiating conditions established at the micro-semiotic level, a given chain of triads realized at time *t* will be among the elements of a set $W = U - x$ of potential chains of triads that might be actualized at *t*. Then, a second kind of constraint acts on the set W, namely, the boundary conditions established by the macro-semiotic level, in the context of which a given chain of triads will be effectively realized. The boundary conditions will select, among all the potential chains of triads which could

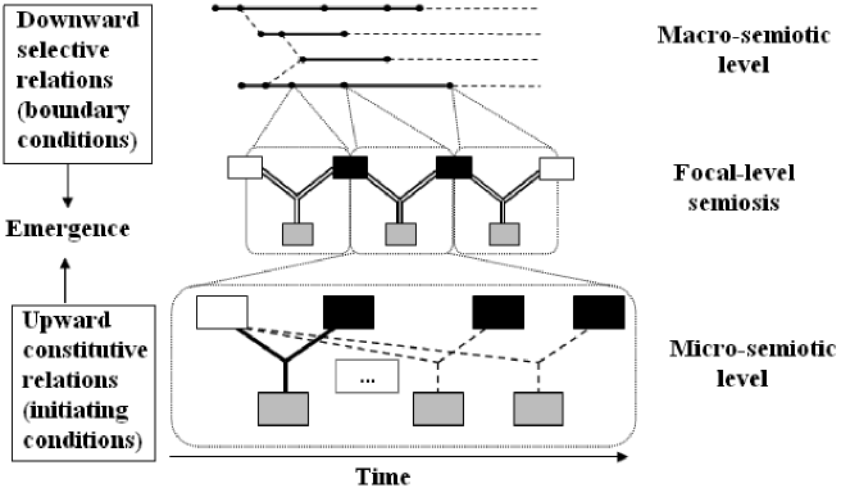

Figure 8. *A model of semiosis in three levels. The upward arrow shows the constitutive relation from individual triads to chains of triads, corresponding to Salthe's initiating conditions. The downward arrow shows selective relations from networks of chains of triads to chains of triads, corresponding to Salthe's boundary conditions.*

be realized from the set W of potential determinative relations S-O-I, a specific chain $T_i = \{\ldots, t_{i-1}, t_i, t_{i+1}, \ldots\}$ to be actualized.[19]

It is in this sense that the emergence of semiotic processes at the focal level, in which chains of triads are actualized, is explained in this model as resulting from an interaction between the potentialities established by the micro-semiotic level and the selective, regulatory influence of the macro-semiotic level. The general ideas involved in this model of semiosis in three levels are shown in figure 8.

6.2. *Levels of semiosis in the genetic information system*

6.2.1. *The micro-semiotic level: Strings of DNA as potential Signs.* A set of three nucleotides (codon) in an open reading frame (ORF) of a coding string of DNA is treated, in our analysis, as a *potential* Simple Sign, i.e., as a Simple Sign which is not involved at a particular time *t* in an effective triadic process involving Objects and Interpretants, which is not partaking in effective semiosis, but potentially can do so. In a similar way, we will refer here to a *potential* Composite Sign.[20]

Let us consider the string of DNA corresponding to the FN gene (with all its exons and introns) in a given cell (figure 2). In this case, each codon

in the FN gene is a potential Simple Sign standing for one specific amino acid as its potential Simple Immediate Object, given the rules of the genetic code. The FN gene is a potential Composite Sign in DNA, which can have a range of effects on cells as interpreters. In a given cell, one (or, sometimes more than one) of these effects will be actualized, i.e., a Dynamical Interpretant of a Dynamical Object, a FN isoform encoded by the FN gene. When the FN gene, as a potential Composite Sign which can undergo alternative splicing, is actualized, a particular splicing pattern will be selected among all possible patterns that might be selected. Thus, a particular sequence of amino acids (Composite Immediate Object) will be selected and reconstructed (Dynamical Interpretant) among all possible sequences of amino acids that might be synthesized in that cell type and developmental state (the range of interpretability, the Immediate Interpretant of the Composite Sign). That Immediate Object will, then, indicate a particular Dynamical Object, say, the fibroblast FN isoform.

To understand the idea of 'potentiality' in this explanation, consider, for the sake of our argument, a cell in a given state at time t in which the FN gene is not being transcribed, i.e., the codons in the FN gene are merely potential Simple Signs and the FN gene as a whole, a potential Composite Sign. This is a situation in which we can talk only about a micro-semiotic level as a set of initiating conditions for effective semiotic processes which are not instantiated at that time t. After all, in a string of DNA which is not transcribed in mRNA, or in a string of mRNA that is not translated into a polypeptide, codons are not effectively *acting* as Signs. In these circumstances, codons can potentially be Signs of an Object for an Interpretant. Free amino acids, by their turn, can potentially be Objects of that Sign for an Interpretant. Finally, the potential Interpretant amounts to the potentiality of a specific sequence of amino acids (Composite Immediate Object) being reconstructed from a Composite Sign in DNA, by means of the processes of transcription, RNA processing, translation etc.

To inquire further into the idea that a non-transcribed reading frame in DNA is nothing but a potential Composite Sign, consider a hypothetical situation in which the FN gene remain non-transcribed in all states of a given cell, in any given time t. In this case, the FN gene and the codons composing it will never effectively act as Signs; rather, they will remain potential. The string of DNA containing these codons will always remain as a silent structure that might — potentially — engage in the process of becoming effective information.

As we interpret what is a gene from a semiotic standpoint, an exciting conclusion is suggested from the very beginning, namely, that informa-

tion in a gene is not an entity, but a process. DNA will not 'harbor' or 'carry' information as sequences of nucleotides, but only the potentiality of engaging in processes by means of which the form of an Object can be communicated to an Interpretant, i.e., what we call here 'potential information.' We are moving towards a reinterpretation of what is information in a cell that hopefully avoids a problem detected by Oyama ([1985] 2000), namely, that genetic determinism is implied by the way we represent genes as they carried information for the development (or functioning) of an organism. We will come back to this point later.

The idea that a string of DNA encoding the sequence of amino acids of a protein is just a potential Sign when it is not being transcribed can benefit from Emmeche's notion of experience as '... traces of particular significant interactions between a system and its surroundings that for some period are represented within the system' (Emmeche 2003: 325). Experiences are, thus, '... fossilized signs ... or quasi-stable forms of movement that organize the system's past forms of movement in such a way as to have significant consequences for the system's future movement' (Emmeche 2003: 327). Similarly, DNA sequences can be regarded as 'fossilized' Signs that represent within the system past interactions with its surroundings in such a way as to have significant, i.e., adaptive consequences for the system's future dynamics. As Emmeche (Emmeche 2003: 328) writes, 'in biological systems like the cell, experiences are, among other things, the genetic "fossils" in DNA witnessing the specific proteins that were functionally participating in earlier ancestor cell lines to maintain the metabolic form of movement.' DNA sequences are just physical carriers of past experience, i.e., 'fossils,' potential Signs. When they are put into effective action in a cell (rather than act on their own), they become part of an effective triadic process, by which they can have an effect on a cell by irreducibly involving also Objects and Interpretants.

If we go on with the analogy, we will be able to see that an unexpressed gene in a cell is a potential Sign as much as an undiscovered fossil deep down in a mountain. We can postulate that a hypothetical fossil buried in a rock but never seen before is a Sign on the grounds of our previous experiences of the Sign action of fossils: we have already a habit of interpreting patterns of rock as Signs of an ancient fish or dinosaur. Similarly, we have enough knowledge of genes to postulate that particular tokens of genes may be potential Signs, *i.e.*, they may be readable by the cell as Signs for the process of synthesis of a specific protein, in response to some necessity. Pragmatically, also, a potential Sign is known by its effects, these being as hypothetical as the very Sign; yet we judge these future effects as real based upon an inference that relate past tokens of similar type of Signs to their Objects and Interpretants. The unseen gene

is as silent as the unseen fossil. Collateral evidence about expressed genes or neighboring disclosed fossils supports our claim about the possible existence of unseen, silent, potential Signs. A potential Sign is information that does not — yet — have an effect on the interpreter, but has the power to do so in the future, in a given interpretative context.

In living systems, experience became so intensified in semiotic terms that it can reach forward in time. This is true not only of experience in quite complex but also in much simpler living systems, where experience can take, for instance, the form of a genetic memory (compare the term 'form,' as defined in section 3), which, given the stability of DNA, can represent traces of significant interactions between a living being and its surroundings for quite long periods, reaching the future not only in the restricted time scale of somatic life, but also in the much more expanded time scale of evolutionary processes. The representation of experience as a quasi-stable dynamical pattern, as a 'fossil,' renders the system anticipatory, endowing it with the capacity of operating with models of possible future states.

It is not the case at all that we are claiming that the genetic information system might be prescient in some sense or another. It is just that, if the selective regimen for a given lineage remains stable in the relevant variables, the past selective events —, i.e., the past events of differential survival and reproduction — endow the future generation with 'fossils,' potential Signs in DNA, which are traces of adaptive interactions between a system and its surroundings, and are likely (but not surely) to create conditions for successful future interactions.

6.2.2. *Transcription, RNA processing, and protein synthesis as processes of gene actualization.* Transcription, RNA processing, and protein synthesis can be understood, in semiotic terms, as processes of actualization of potential signs in protein-coding genes. Consider, for instance, a given hepatocyte *h*, in which the FN gene is transcribed and the corresponding mRNA, after cell type-specific splicing, is translated into the hepatocyte-specific FN isoform. These processes actualize potential Signs in a string of DNA, turning them into actual Sign processes, Signs in effective action in an organism. When put into action, the nucleotide sequences in that string of DNA become part of effective semiosis, a triadic-dependent process by means of which the FN gene as a Composite Sign indicates, through a process involving the actualization of each Simple Sign composing it, the functional hepatocyte FN isoform as a Dynamical Object. This FN isoform has in turn an effect on the organism in which it is expressed (its Dynamical Interpretant), participating in its adaptive

interactions with its surroundings, and, thus, contributing to the presence of the FN potential gene in the next generation in a high frequency.

The actualization of a potential gene in a string of DNA depends on boundary conditions established by a higher-level semiotic network, a network of signaling processes, involving many chains of triads, which will regulate or coordinate gene expression, ultimately determining the likelihood of transcription of a given gene, or splicing of a given pre-mRNA according to a particular pattern, or chemical modification of a given protein in a manner that modulates its function in a particular way (e.g., a phosphorylation), and so on. A variety of regulatory mechanisms studied in cellular and molecular biology can be thus interpreted as a macro-semiotic environment establishing boundary conditions which will downwardly determine which potential genes in a string of DNA will be turned into actual genes, into genes in effective action in a cell. These mechanisms determine which sequence of amino acids will be actually reconstructed (Dynamical Interpretant) among all those that might be reconstructed (Immediate Interpretant, the range of interpretability of a Sign) out of a string of DNA (Composite Sign).

This shows how several complexities involved in the gene concept and in gene expression itself can be introduced in our analysis: boundary conditions established by this macro-semiotic environment will determine, for instance, which stretch of DNA will be read (e.g., allowing for an analysis of transcription of overlapped or nested genes), which pattern of RNA splicing or RNA editing will be instantiated in order to produce a particular mature mRNA (allowing for the subtleties of alternative RNA splicing or RNA editing to be taken into account), which functional protein will be effectively constructed by the cell (allowing for chemical and/or structural modifications suffered by the primary amino acid sequence of a protein to be considered), and so on. However, we will avoid introducing a great deal of complex details now; rather, we will concentrate on an analysis of transcription and translation, as our goal here is to establish a set of concepts, tools, and procedures for the analysis of information systems in living beings, not to provide an exhaustive analysis of the host of processes involved in these systems, not even at the cellular level.

6.2.3. *Semiotic analysis of transcription.* The first step in the actualization of potential Signs in a string of DNA is transcription. This process turns the potential Signs in DNA into potential Signs in pre-mRNA. It is easy to see that Signs in pre-mRNA are still potential, since they will become part of actual triads only if they are effectively translated.[21] It can be the case, for instance, that a given codon in pre-mRNA is located

in an alternatively spliced exon that is eliminated from the final transcript in a given cell type, developmental stage or age. In this case, the actualization process is not completed and that codon remains in the condition of a potential Simple Sign. Consider, for instance, Exons EIIIA and EIIIB in the FN gene. As they are spliced out of FN mRNA in hepatocytes, the codons in those exons are never actualized, remaining as potential Signs in this cell type. In fibroblasts, however, these potential Signs will be indeed actualized.

Transcriptional control is the major mechanism for regulating the production of a protein encoded by a given stretch of DNA, involving both repression and activation of specific genes in response to signals originating from the cell itself and, more often, from the extracellular environment. In terms of the general model presented above, this means that, as a first step in the actualization of potential Signs at the micro-semiotic level, transcription is constrained by boundary conditions established by networks of chains of triads (macro-semiotic level) which ultimately determines the likelihood of transcription of a given potential Sign in DNA. Transcriptional regulation amounts to the selection (by the macro-semiotic environment) of a specific chain of triads to be actualized, among many potential chains that might be actualized in a given moment. Furthermore, transcriptional regulation is not at all a case of random selection, but rather the result of mechanisms selected in the course of the evolution of an organism, due to the differential fitness of varying responses of the cellular regulatory systems (as a cellular macro-semiotic environment) to boundary conditions or selective regimens established by the environment outside the cell, and outside the organism as a whole.

Let us now analyze in more detail transcription as a semiotic process. We will consider here two views of the processes at stake, the 'horizontal' and the 'vertical' perspectives. If we take a 'horizontal' view of transcription, we will see a mechanistic process in which RNA polymerase moves along a string of potential Simple Signs in DNA, triggering subsequent semiotic events, in which those potential Signs become part of triads including Objects and Interpretants. Let us focus first on a Simple Sign in DNA, i.e., a set of three nucleotides in a coding region (figure 9). The Simple Immediate Object, by its turn, is a set of three nucleotides in mRNA. In our example, a codon in the FN gene is a potential Simple Sign that is actualized when that gene is transcribed.[22] As we argued above, the Immediate Interpretant of a Simple Sign is the range of interpretability established by the rules of base pairing, and its Dynamical Interpretant is the realization of a particular rule by means of which specific nucleotides in DNA determine specific nucleotides in mRNA. When a triad in transcription is actualized, the interpretative subsystem of the

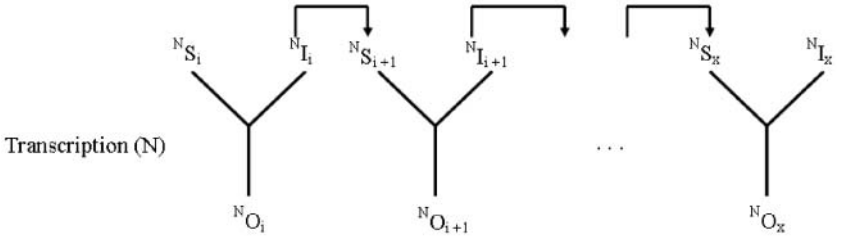

Figure 9. *A 'horizontal' view of transcription. The letter 'N' is used along this paper to identify the elements of a triad at the level of transcription.* ^{N}*S: Simple Signs in DNA (codons);* ^{N}*O: Simple Immediate Objects in pre-mRNA (codons);* ^{N}*I: Immediate Interpretants in transcription, the range of interpretability established by the rules of base pairing by which specific nucleotides in DNA determine specific nucleotides in mRNA. RNA polymerase is the interpretative system performing the transcription. The arrows represent the movement of the interpretative subsystem, RNA polymerase, to the next Simple Sign, when a triad is actualized in transcription.*

cell (as a global interpreter), RNA polymerase, moves to the next codon in the string of DNA, i.e., to the next Simple Sign.

If we take a 'vertical' view, we will consider the relationship between semiotic processes in transcription and translation. Then, we will see a dynamical process in which the Simple Immediate Object of each triad actualized in each step of transcription, *i.e.*, a three-nucleotide sequence in mRNA, *becomes* a potential Simple Sign in the next Sign process in the actualization of a gene, translation (figure 10).[23]

6.2.4. *Semiotic analysis of protein synthesis.* We can now go on to analyze protein synthesis in semiotic terms, considering both the recognition of codons in mRNA by particular tRNAs and the attachment of appropriate amino acids to specific tRNAs.

Let us consider, first, the attachment of amino acids to tRNAs. In this case, the Sign in a given triad is the three-dimensional structure of a particular tRNA, which is recognized by the interpretative system in this process, a specific aminoacyl-tRNA synthetase. The Simple Immediate Object in each triad is a specific amino acid, which is also recognized by aminoacyl-tRNA synthetase on the grounds of its three-dimensional structure. The aminoacyl-tRNA synthetase establishes a relationship between one amino acid (the Simple Immediate Object) and all its cognate tRNAs (with specific three-dimensional structures as Signs) due to its capacity of specific recognition. This enzyme actualizes, thus, one of the rules expressed in the genetic code, the Immediate Interpretant of a Simple Sign. This actualization is the Dynamical Interpretant.

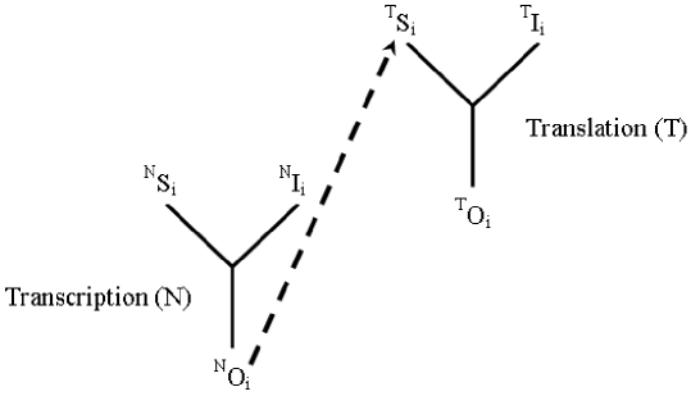

Figure 10. *A 'vertical' view of the relationship between Sign processes in transcription and translation. The letter 'T' is used along this paper to identify the elements of a triad at the level of translation. The letter 'N' and the symbols ^{N}S, ^{N}O, and ^{N}I are used as indicated in the caption for figure 9. ^{T}S: Simple Signs in mRNA (codons); ^{T}O: Simple Immediate Objects (particular tRNAs with specific anticodons); ^{T}I: Immediate Interpretants in translation, the range of interpretability established by the rules of base pairing by which specific nucleotides in mRNA are paired with specific nucleotides in tRNA. For more details on the semiotic analysis of translation, see next section. The arrow indicates that three-nucleotide sequences in mRNA, the Simple Immediate Objects in transcription, become potential Simple Signs in translation.*

As regards the recognition of codons in mRNA by particular tRNAs, the Simple Signs are three-nucleotides sequences in mRNA (codons), the Simple Immediate Objects are particular tRNAs with specific anticodons, and the Immediate Interpretant of the Simple Sign is the range of interpretability established by the rules of base pairing by which specific nucleotides in mRNA are paired with specific nucleotides in tRNA, with the caveat that nonstandard base pairing often occurs between codons and anticodons. The Dynamical Interpretant is the actualization of a specific base pairing. When a triad in this step of translation is actualized, the ribosome, as an interpretative subsystem, moves to the next codon in the string of mRNA, i.e., to the next Simple Sign.

As several processes in protein synthesis, including translation, protein folding, association of different polypeptide chains, and post-translational chemical modifications are often regulated, this step in the actualization of potential Signs in DNA is also constrained by boundary conditions established by networks of chains of triads in a macro-semiotic level, which select determinative relations between S, O, and I at the micro-semiotic level.

Again, we can see these processes in a 'horizontal' or a 'vertical' view. Consider, first, the semiotic processes involved in the recognition of

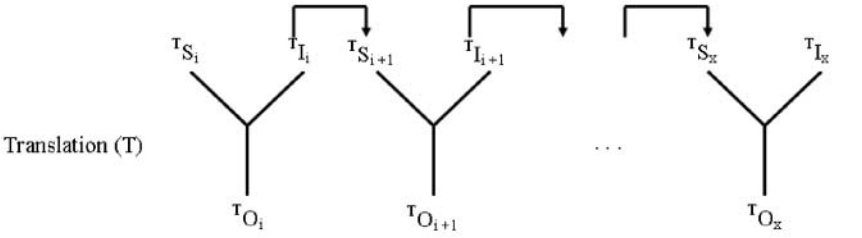

Figure 11. *A 'horizontal' view of translation events in the ribosome. The letter 'T' and the symbols ^{T}S, ^{T}O, and ^{T}I are used as explained in the caption for figure 10. The ribosome is the interpretative system performing this step in protein synthesis. The arrows represent the movement of the ribosome to the next Simple Sign, when a triad is actualized in translation.*

codons in mRNA by tRNAs. If we take a 'horizontal' view of this process, we will see, as in transcription, a mechanistic process consisting in the triggering of a sequence of Sign events by the movement of ribosomes along strands of mRNA (figure 11).

If we take a 'vertical' view of the relationship between semiotic processes in translation and in aminoacyl-tRNA synthesis, we will see a dynamical process in which the Simple Immediate Object of a triad, a tRNA with an anticodon that matches a codon in mRNAs, is also a potential Simple Sign[24] in the semiotic process in which a specific aminoacyl-tRNA is synthesized (figure 12).

Other steps in protein synthesis can also be analyzed semiotically. For instance, protein folding, at least when it involves molecular chaperones, a special class of proteins that help guide the folding of many proteins, is regulated by processes involving signaling pathways. Nevertheless, we will not develop a semiotic analysis of this process in the context of this work.

Finally, it is worth discussing start and stop codons in the context of the semiotic analysis developed here. Translation is always initiated by the recognition of a start codon in mRNA, usually AUG, by a tRNA carrying the amino acid methionine. Translation is, therefore, a semiotic process with a peculiar characteristic: it typically begins with the same Simple Sign (AUG) and always with the same Simple Immediate Object (methionine). This Immediate Object, however, is in most cases subsequently eliminated from the sequence of amino acids which indicates the Dynamical Object of the Composite Sign, and, therefore, we have here an Immediate Object which is not really related to the semiotic availability of the functional protein indicated by a gene. In this case, the Dynamical Interpretant is the actualization of a rule of the genetic code, by which AUG usually encodes methionine, and the Dynamical Object is the instruction that translation should be initiated.

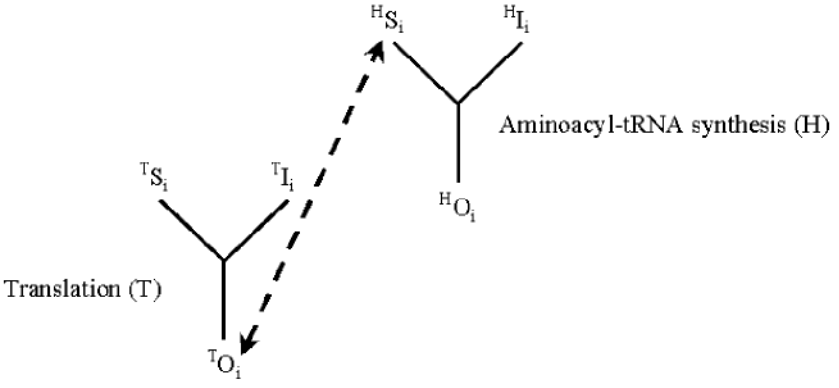

Figure 12. *A 'vertical' view of the relationship between semiotic processes in translation and in aminoacyl-tRNA synthesis. The letter 'T' and the symbols* ^{T}S, ^{T}O, *and* ^{T}I *are used as explained in the caption for figure 10. The letter 'H' is used to identify the elements of a triad at the level of aminoacyl-tRNA synthesis.* ^{H}S*: Signs: three-dimensional structure of tRNAs;* ^{H}O*: Simple Immediate Objects: specific amino acids;* ^{H}I*: Immediate Interpretant, the range of interpretability of each codon as a Simple Sign, established by the rules of the genetic code. The arrow indicates that the same element, a tRNA with a specific anticodon that matches a codon in mRNAs, plays the different functional roles of Simple Immediate Object and Simple Sign in different triads.*

Stop codons (UAA, UAG, and UGA), in turn, are usually involved in the termination of the semiotic process of translation. None of the stop codons is recognized by a tRNA. Rather, they are recognized by proteins called 'release factors,' which act at the ribosomal site occupied, in the case of other codons, by an aminoacyl-tRNA. When a release factor binds this site, a molecule of water, instead of an amino acid, is added to the growing polypeptide chain, resulting in its release from the last tRNA. In this case, the Dynamical Object of the semiotic process is the instruction that the process should be interrupted, and this Dynamical Object is made semiotically available by the fact that a molecule of water, rather than an amino acid, is the Immediate Object of the Simple Sign at stake.

6.2.5. *A global picture.* It is time, then, to look at the processes discussed above from a global perspective, which will allow us to use the detailed analysis we carried out to reach, as an overall conclusion, the initial semiotic analysis presented in section 5. Figure 13 shows a complete view of all the semiotic processes involved in the actualization of potential Signs in DNA that were explained in the previous sections. It is important

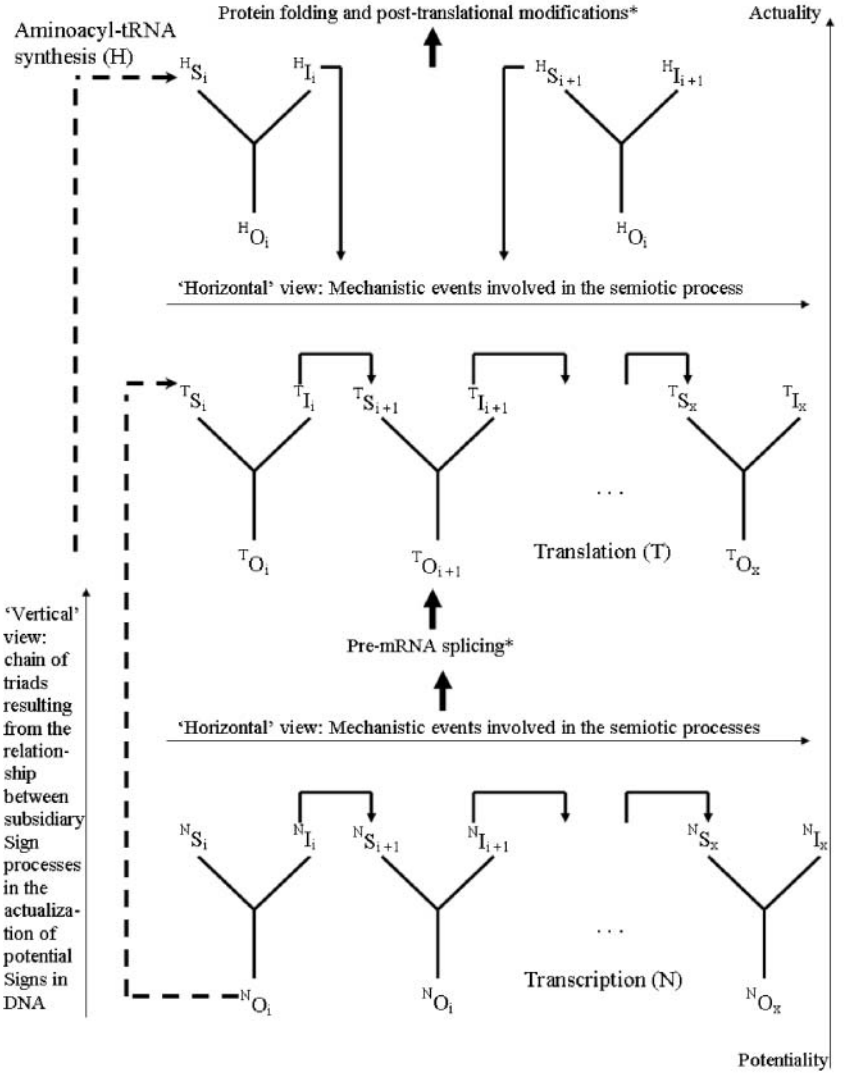

Figure 13. *Whole view of the semiotic processes involved in the actualization of potential Signs in DNA. Letters 'N,' 'T,' 'H,' and symbols* ${}^{N}S$, ${}^{N}O$, ${}^{N}I$, ${}^{T}S$, ${}^{T}O$, ${}^{T}I$, ${}^{H}S$, ${}^{H}O$, ${}^{H}I$ *are used as explained in figures 9, 10, and 12. Dashed arrows represent relationships between transcription and translation, and translation and aminoacyl-tRNA synthesis, as explained in figures 10 and 12. Continuous lines indicate the horizontal and vertical views explained in the text. Asterisks indicate signaling processes that can be analyzed semiotically, but were not addressed here for reasons of space.*

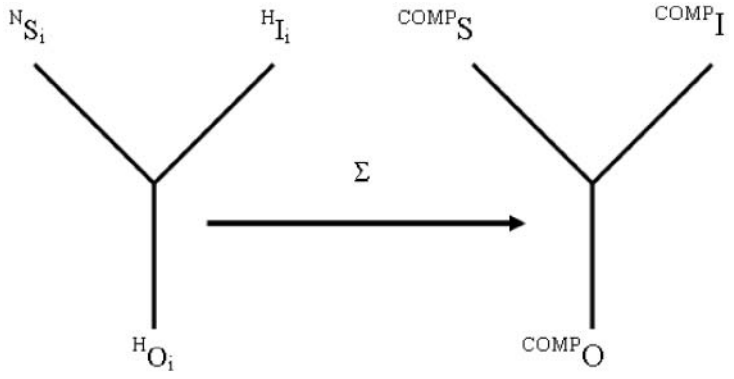

Figure 14. *A global picture of genetic information as a triadic process. The letters 'N' and 'H' indicate transcription and aminoacyl-tRNA synthesis, respectively. The symbols ^{N}S and ^{H}O indicate a Simple Sign in DNA (a codon) and a Simple Immediate Object in the synthesis of aminoacyl-tRNA, that is, a specific amino acid. By means of the specificity of recognition of amino acids and tRNAs by aminoacyl-tRNA synthetases, the rules connecting sixty-one codons and twenty amino acids in the system of constraints expressed in the genetic code (the Immediate Interpretant of the Simple Sign) obtain. In the picture, it is indicated by ^{H}I. By the summation of individual chains of triads, we obtain a global picture of genetic information as a semiotic process involving a gene as a Composite Sign, ^{COMP}S, i.e., a string of DNA; ^{COMP}O, the Composite Immediate Object of the gene, i.e., the linear sequence of amino acids in a protein (or polypetide), in the case of a protein-coding gene; and the Immediate Interpretant of the Composite Sign (^{COMP}I), i.e., its range of interpretability, the possibilities of reconstruction of sequences of amino acids from that Sign in DNA. The expression 'COMP' stands here for 'Composite.' This global model is equivalent to the picture shown in figure 3.*

to notice that the chains of triads are, as shown in the figure, arranged vertically, going from transcription to translation and aminoacyl-tRNA synthesis, each codon a time. The actualization of potential Signs in DNA is, indeed, a complex process including an impressive array of subsidiary semiotic processes, described in a stepwise manner above and expressed as a whole in figure 13. This is in accordance with the general idea that in living nature there are different levels of handling of information, i.e., generation, translation, coding, re-coding, and interpretation of Signs.

From the view of the semiotic processes involved in the actualization of potential Signs in DNA shown in figure 13, we can obtain a global picture (figure 14) that corresponds to the semiotic analysis presented in section 5. As we argued above, if genes are treated as Signs, they can only have an effect on a cell through a triadic process which *is* genetic information, and involves an irreducible relationship between three elements: the Composite Sign, which is a string of DNA, and can be transcribed into

RNA, processed, and, in the case of protein-coding genes, translated into a protein (or a polypeptide), by means of the semiotic processes analyzed above and shown in figure 13; the Composite Immediate Object, which is, in the case of protein-coding genes, a linear sequence of amino acids, and, in the case of RNA genes, a linear sequence of ribonucleotides; and the Immediate Interpretant of a Composite Sign, which is its range of interpretability, i.e., the possibilities of reconstruction of sequences of amino acids (Immediate Objects) from that Sign in DNA. The Dynamical Interpretant of a Composite Sign corresponds, in turn, to the effective reconstruction of a sequence of amino acids from a Sign in DNA.[25]

As the global picture in figure 14 illustrates, a model of genetic information interpreted as a Sign process can be obtained by the generative emergence resulting from the summation of lower-level semiotic processes, involving triads of which the codons in DNA (Simple Signs) are the first correlates. In this sense, the Composite Sign (a stretch of DNA) and the Composite Immediate Object (a linear sequence of amino acids or ribonucleotides) can be treated as resulting from an accumulative process of interpreting Simple Signs (codons in DNA) of Simple Immediate Objects (amino acids or ribonucleotides).

The Dynamical Object of a Composite Sign in DNA is a functional, folded, and chemically modified protein, which can exert a particular effect (the Dynamical Interpretant of the Dynamical Object) on a cell or organism of which the cell is part, among a range of possible effects (the Immediate Interpretant of the Dynamical Object). It is only then that a potential Sign, a potential gene in DNA, turns into an actual Sign, a gene effectively involved in the Sign process we interpret here as genetic information.

To put it differently, the full actualization of a string of DNA, which is only a potential Sign, demands the ultimate indication of a Dynamical Object, a functional protein, by the Composite Immediate Object, a polypeptide chain (in the case of a protein-coding gene). Only then the path from potential to effective information is completed in the genetic system. The actualization of potential Signs in DNA requires a series of interpretative subsystems, such as RNA polymerases, ribosomes, aminoacyl-tRNA synthetases, etc. The regulatory influence of the macro-semiotic level, as a network of signaling processes, on interpretative subsystems, and, thus, on transcription, splicing, translation, shows that, in the end, we have to consider, the whole cell as ultimately participating in the network necessary for the actualization of potential genes in DNA (see section 5). The cellular network of chains of triads is, in turn, highly responsive to environmental factors, given the semi-open nature of living systems.

Finally, we should consider how the process of actualization of a potential gene in DNA can be embedded into the model of semiosis in three levels shown in figure 8.[26] As figure 15 shows, potential Signs, potential genes in DNA, are actualized in response to regulatory dispositions arising from a network of signaling pathways that elicit cellular specific responses to signals, i.e., to Signs arising from the extra- or intracellular environment.[27] A controlled, regulated answer by a cell is impossible without signaling. When a particular gene product is needed, a signal from the environment activates the expression of a given gene by means of signaling mechanisms. The cell, as an interpreter, answers to an environmental cue by means of a specific alteration of its internal states, triggered by a whole network of signal transduction culminating in a change at some level of gene regulation (for a biosemiotic analysis of signal transduction systems, see Bruni 2003). These relations cannot be understood only in terms of the molecular interactions taking place in the network of signal transduction, because the latter crucially involves triadic-dependent interpretative processes, as the widespread usage of information talk in modeling and explaining signaling pathways clearly suggests. Through signaling pathways, cells are able to interpret Signs from the extra- and intracellular environment as meaning something, that is, being interpretable by the cell as signifying something beyond the chemical carrier of the Sign itself. Thus, the presence of an antigen bound to a membrane receptor may mean, for instance, that the organism is under the threat of a pathogen. In response to such an interpretation of the environment, a signal-transduction cascade can be activated in B-cells, leading to the actualization of genes associated with B-cell activation, and, thus, to specific effects on the B-cell, making it, for instance, engage in a process of presenting peptides derived from the antigen in its cell surface, where they can be recognized by T-cells, leading to T-cell activation. As molecules come to mean something else than just being molecules (in our example, the threat by a pathogen) and cells use them as parts of a process of interpreting its circumstances, something more than chemistry is going on there. Needless to say, there is nothing supernatural or at any rate mystical going on; it is just the case that semantic and pragmatic information plays a fundamental role in the lives of organisms, and information can be interpreted as a semiotic, triadic-dependent process, as we have argued throughout this paper.

Signaling pathways in a cellular system play the role of establishing boundary conditions to processes at the focal- and micro-semiotic levels, downwardly selecting particular strings of DNA, potential genes, to be actualized, among all the potential Signs at the micro-semiotic level that might be actualized at a given time t. It is the actualization of a specific

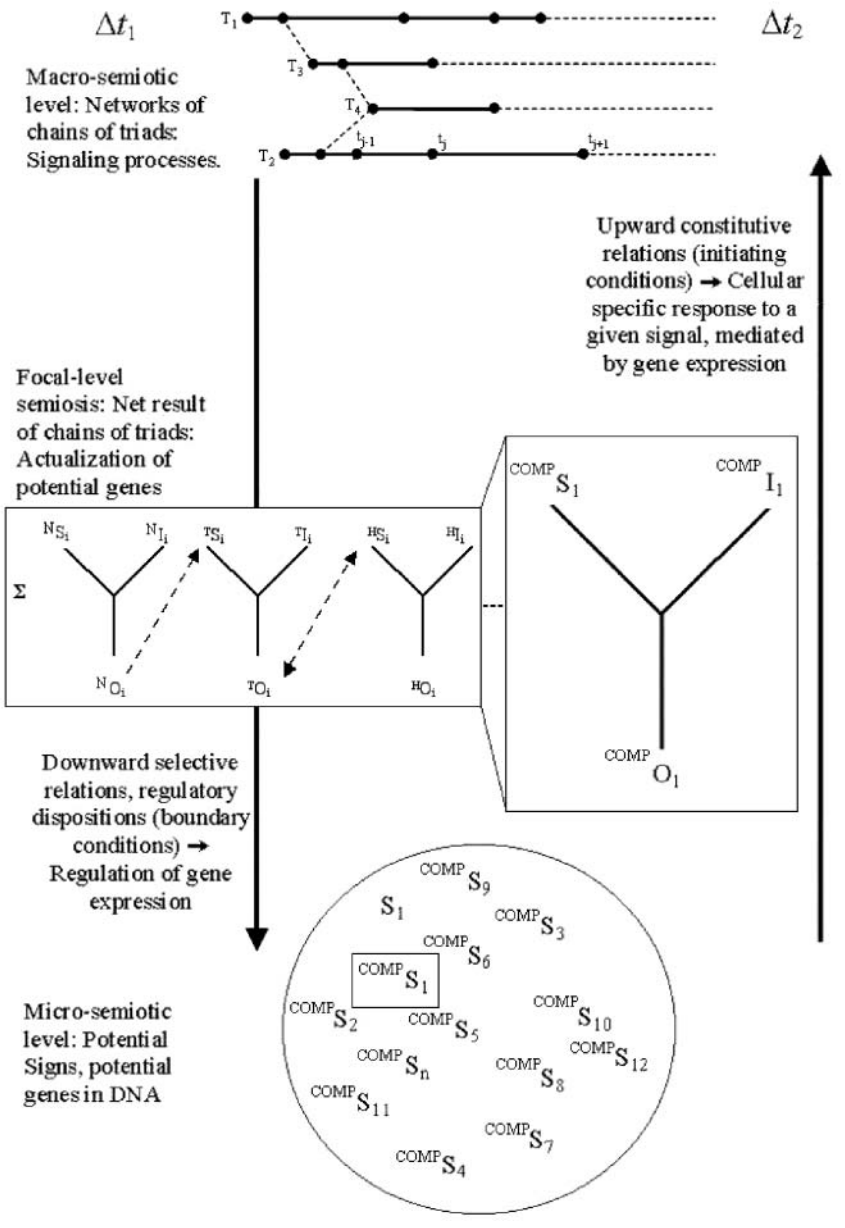

Figure 15. *The process of actualization of a potential gene in DNA embedded into the model of semiosis in three levels shown in figure 8.* ${}^{COMP}S$, ${}^{COMP}O$, ${}^{COMP}I$, ${}^{N}S$, ${}^{N}O$, ${}^{N}I$, ${}^{T}S$, ${}^{T}O$, ${}^{T}I$, ${}^{H}S$, ${}^{H}O$, *and* ${}^{H}I$ *are used as explained in the previous figures.* ${}^{COMP}S_{1-n}$ *stands for potential Signs at the micro-semiotic level.* Δt_1 *and* Δt_2 *in the upper part of the figure indicate the diachronic nature of the semiotic processes involved. At the focal-level of semiosis, we show a chain of triads at the focal level and the global picture shown in figure 14.*

set of potential genes (which can — but not necessarily should — include only one element, as shown in the hypothetical case in figure 15) that allows the cell to answer to a given signal in a specific way, by means of Dynamical Objects and their Dynamical Interpretants.

In the whole process of gene actualization, it is the cell and, as Hall (2001) emphasizes, its immediately adjacent peri- and extracellular matrices that carry out the responses to environmental changes. That is, although potential genes in DNA are being actualized in the process of responding to a given intra- or extracellular signal, it is not the case at all that genes are in the command; rather, they are commanded by the interpretive mechanisms of the cell to supply the materials needed for such a response.

7. What is genetic information?

In this section, we will come back to the claims we put forward in the end of section 5, namely, that genes can be regarded as Signs in DNA, which can only have an effect on a cell through a triadic-dependent process, which, in turn, *is* genetic information and involves more than just genes as Signs in DNA but also Objects and Interpretants. And, moreover, that information is the process by means of which a form in a Dynamical Object (a functional protein) is communicated to an Interpretant (the reconstruction of a specific sequence of amino acids in a cell) by means of Signs in DNA.

At the focal level of semiosis, the actualization of a potential gene is an emergent process, depending on two sets of constraints (see figure 15): First, as a given organism, as a product of a historical process, does not and cannot contain each and every possible coding string of DNA but rather only a specific set of them, its response to a given extra- or intracellular signal is constrained by the very fact that it contains a restricted set of available potential genes in its genome. According to the model developed in section 6.1, this set of potential genes establishes initiating conditions for the organism's response.

Secondly, a network of signaling pathways at the macro-semiotic level also constrains the organism's response, with the result that the response turns out to be quite specific to a given class of signals. In the general model presented in section 6.1, this amounts to the establishment of boundary conditions by the macro-semiotic level, *i.e.*, to a set of higher-level constraining conditions that result in the selective choice of a set (which can contain only one member) of potential genes available in that organism's DNA to be actualized.

But, then, what is, precisely speaking, *information* in this emergent process? Putting the concepts of information discussed in section 3 to work, we can say that *the actualization of a potential gene triggers a triadic-dependent process by means of which that gene has an effect on the cell. This process is effective information*. A gene has an effect on the interpreter because it mediates, as a Sign, a process by which the form of a Dynamical Object (a functional protein) makes a difference to that interpreter. It is clear, then, that effective information — as defined here — is not contained in DNA, but, rather, is *a semiotic process which is irreducibly triadic, involving a gene in action, a dynamical Sign that has an effect on its interpreter by determining an Interpretant (the effect of the Sign as a difference) to stand in a similar relation to something else (the Object of the Sign) as that to which the Sign stands, thus mediating the relation between Object and Interpretant*.[28]

In the context of our analysis, we can say that, when a cell, as an interpreter, responds to an environmental cue, by means of a set of signaling pathways that ends up altering its pattern of gene expression, triggering the actualization of a set of potential genes, what happens is that the interpretation systems of a cell are acting to create differences inside the cell in correspondence to differences in the external environment interpreted by them. In response to a state of the system plus its environment, a difference is established between two or more classes of gene expression patterns, in which different sets of potential genes are actualized, that is, *become elements in effective semiosis*, i.e., in *effective information*, as defined in this paper. These differences in patterns of gene expression have an effect on the cell by altering its internal states, through the action of Dynamical Objects indicated by different sets of potential genes that are getting actualized.

The semiotic analysis developed in this paper suggests that potential genes can be regarded as a kind of 'tacit' representational patterns having strings of DNA as their vehicles. A 'potential Sign' is something that *may* be a Sign of an Object to an Interpretant. A potential Sign, therefore, is a Sign which is not involved in effective semiosis (i.e., effective information), in a given time t. Potential information is defined here as a process of communicating a form embodied in an Object to an Interpretant through the mediation of a Sign that could take place in a given moment (see section 3). A potential gene, as a potential Sign, is just one element in semiosis. This means that potential genes, as patterns in DNA, are *not* and, also, do not *carry* information. Rather, they are only the first correlates of a triadic-dependent process that we define here as information. Potential genes and, therefore, DNA, *carry, harbor, convey* only the potentiality of a process we call 'information.'[29]

Even if potential genes are treated as patterns in DNA, it will still be the case that to have any effect on the cell as an interpreter they must be subordinated to information as a process, an idea which is consistent with the general claim, in process philosophy, that substances are conceptually and ontologically subordinated to processes (Rescher 1996). It is within such a process perspective that we treat potential genes as potential Signs, a kind of disposition or pattern showing propensities for having certain effects on interpreters, i.e., as potential (semiotic) processes, or tendencies. That is, the framework developed here gives privilege to processes, not substances, and treats propensities or general tendencies as real. In these terms, it can be made consistent with the general picture about genetic information we find in current molecular biology and genetics, provided that this picture is reinterpreted within a general process philosophical stance.[30]

Effective information, in turn, is not carried by and cannot be identified with entities in DNA, but is, as defined here, the very triadic-dependent, semiotic process by means of which a gene can have any effect on a cell. As such a process, it irreducibly involves Signs, Objects, and Interpretants in a dynamical relationship. Notice, moreover, that it is the interpreter that coordinates the semiotic processes at stake. Biologically speaking, the genetic material does not do things to the cell, but, rather, it is the cell, as an interpreter, that does something with the genetic material.

The semiotic analysis we developed also allows us to offer an interesting account of the 'transmission' of information. It is not effective information that is being communicated when we observe, for instance, 'vertical transmission,' say, from parent to offspring.[31] From the perspective of our results, what is being communicated is only *potential information*, i.e., the potentiality of the process we call information, which can be said, as explained above, to be carried by stretches of DNA. Signs in DNA will only become elements in effective information when interpreted by the cell. Effective information itself cannot be carried from one system to another, but only potential information can be 'carried' by the first correlates of triads, Signs (the vehicles of which, in biological systems, are typically physicochemical entities).[32]

In this connection, we think Jablonka points in the right direction when she uses the term 'transmission' (taking into account Oyama's criticism of the typical usage of informational terms in biology) 'not for the handing over of preexisting entities, but to denote any process that results in an organization pattern from one entity being reproduced in another. Thus, when talking about heredity, an entity is related to others by special processes that lead to the reconstruction of its organization in those other entities' (Jablonka 2002: 588–589). Considering the genetic inheritance

system, this can be taken to mean, in our terms, that the pre-existent entities that are transmitted from one generation to another carry only potential information. They are 'fossils' in DNA, which will become, only when actualized, Signs in action, developmental resources (among others) in a set of processes that will end up reproducing the form or organization of an entity in other entities.

If we consider the communication of forms by genes from an evolutionary perspective, we will be in a position to claim that in this case forms are communicated from Dynamical Objects (functional proteins) to interpreters through genes as Signs, this being the reason why the Dynamical Object is the primary constraining factor in semiosis (see figure 5). To clarify the matter, suppose, for the sake of the argument, that a stretch of DNA which codes for a sequence of amino acids that does not indicate any functional protein suffers a mutation in time t_1 that turns it into a potentially functional gene, *i.e.*, that, after that mutation, if that potential gene is actualized, it will indicate a functional protein. This protein, in turn, plays an adaptive role in a given lineage, affecting the likelihood of survival and successful reproduction of the organisms carrying it in t_1. That potential gene, that fossilized Sign in DNA, will tend to be preserved by natural selection in the future generations of that lineage, in times $t_2, t_3, \ldots, t_n$, if the selective regimen remains the same in the significant variables affecting the survival and reproduction of the organisms at stake. The form of the Dynamical Object in t_1 increases the chance of the Sign indicating it being present in the next generation of interpreters, in t_2, in high frequency. The form of the Dynamical Object is communicated to the interpreters in the future generations through the mediation of the Sign. It is in this sense that we can say that form is communicated, from an evolutionary perspective, from a functional protein, as a Dynamical Object, to a gene, as a Sign. Notice that we are not postulating any inversion of the central dogma (as if sequences of amino acids in proteins might determine sequences of nucleotides in DNA). We are referring, rather, to the effect of functional proteins on the likelihood of certain Signs, certain genes, being present in future generations.

In somatic time scale, in a given generation, the form — as a type[33] — which was evolutionarily communicated from a Dynamical Object at t_1 — as a token of that type — to the interpreter by the mediation of the Sign is then communicated from the Sign to the Dynamical Object — as a token of the type which was communicated — in t_2, through the mediation of the Interpretant. Thus, the interpreter will be able to produce through habits acquired in evolution and development a new token of the Dynamical Object. Examining the process, therefore, from a proximal, rather than a distant (evolutionary) perspective, we can say that

effective information is a process by means of which a form is communicated from ^{COMP}S (as a fossilized Sign) to ^{COMP}O through ^{COMP}I, indicating a Dynamical Object, the final form of which will depend on constraints established by both ^{COMP}S (at the micro-semiotic level) and a series of habits, regularities, at work at the macro-semiotic level. When a Dynamical Object, a functional protein, is finally put into action in a cell, its actual effect on the cell, its Dynamical Interpretant, takes place in t_2. The presence of that functional Dynamical Object in a given cell at t_2 is mediated by the communication of the fossilized form as a Sign, which constrains the future semiotic processes in daughter cells. This constraint on the future semiotic processes increases the likelihood of repeating the successful, adaptive past interactions with the circumstances.

A final argument should be offered to support the claim that information should be identified with a process by which a sign has an effect on an interpreter, and not with any single element of the triadic-dependent process itself. To build this argument, let us consider in turn each element in a triadic-dependent process that might be regarded as information. Consider, first, that the presence of a given Sign S in DNA cannot be information in itself, since S is present in each and every cell of an organism, even in those in which it is not expressed and, therefore, has no actual effect on the cell. Secondly, consider that the presence of a Dynamical Object in a given cell do have an effect on it, as its Dynamical Interpretant. Nevertheless, the Dynamical Object has an effect on the cell by means of the communication of its form to an Interpretant by the mediation of the Sign. Information lies in the process of communicating a form, of in-*form*-ing the interpreter, and not in the form itself of the Dynamical Object. By the same token, we should not identify information with the Immediate Object, which simply indicates the Dynamical Object. Thirdly, the Immediate Interpretant is the range of interpretability of a Sign, and, thus, it doesn't have by itself an effect on the interpreter, but is rather a set of habits that allows something to mean something else and, thus, have an effect on the interpreter. Finally, we should consider the possibility that an environmental cue *E* to which a given cell responds is information. Surely, there are reasonable grounds for claiming that an environmental cue is 'informative,' and, no doubt, when we focus on the cell as a whole, a cue E to which the cell responds is involved in a process by which it has an effect on the interpreter (the cell). On then grounds of the framework developed here, we will call this process 'information.' However, by exploring this intuition further, and focusing on the relation between E and changes in gene expression as a subset of possible responses, we can see that when a cell answers to *E* by changing its pattern of gene expression, it is as if a 'fossilized' Sign in its DNA 'comes alive,' allowing

the cell to answer to *E*. That is, the cell can retrieve, given the genetic mechanisms of transmission of potential Signs, past successful, adaptive interactions with environmental cues that have the same character as *E* in evolutionary events that happened to the lineages from which a given organism descends. But this means, then, that the cellular system operates to answer to a difference in its environment by *changing its internal states*. We should still look, in this case, inside the cellular system in order to find the change which takes place when it answers to an environmental cue. We are back, then, at the focal level of our analysis, and, at this level, we already discarded Signs in DNA, Immediate Objects, Immediate Interpretants, Dynamical Objects, and Dynamical Interpretants as possible single 'bearers' or 'units' of information. We reach, as a conclusion of this argument, the same idea we have been advocating throughout this paper, namely, that genetic information is a triadic-dependent, semiotic process.

8. Concluding remarks

In this paper, we developed an analysis of the genetic information system which is, in our view, in full accordance with Peirce's theory of Signs and his general process approach to philosophy. Consider, for instance, the claim that potential genes carry only the potentiality of information interpreted as a process rather than an entity, and, accordingly, that information, not even in a potential sense, corresponds to sequences of nucleotides in stretches of DNA. This idea can be straightforwardly related to the claim that a sequence of nucleotides has no intrinsic meaning in the absence of a cell to interpret it. Furthermore, DNA becomes effective information only when it is used as 'data' (Atlan and Koppel 1990) (or, as we prefer, Signs) by an active and complex system of interpretation in the cell, i.e., when potential genes are actualized in response to intra- and/or extracellular signals. A nucleotide sequence means nothing apart from the dynamics of the cell. This is exquisitely consistent with Peirce's claim that '... it is impossible to deal with a triad without being forced to recognize a triad of which one member is positive but ineffective, another is the opponent of that, a third, intermediate between these two, is allpotent' (*CP* 4.317). Signs in DNA, potential genes, can be understood, if we borrow Peirce's terms, as being 'positive but ineffective.' Indeed, DNA, in a cell system, cannot do anything by itself, while the cell, in turn, can do things with DNA, by actualizing potential Signs fossilized in it, so as to indicate some useful molecule, say, a functional protein (an 'opponent'), a Dynamical Object, which determines a third, a Dynamical

Interpretant, which is 'all-potent,' playing a given function or resulting in a given dysfunction, depending on the specific case at stake.

Furthermore, the account developed here is also in accordance with the general picture of genes and how they are expressed in molecular biology, with the fundamental difference that information is differently conceptualized. Surely, this partly stems from the fact that we opted to present here a more conservative interpretation of our results, leaving a bolder interpretation — roughly similar to Neumann-Held's 'molecular process gene concept' (see Neumann-Held 1999, 2001; Griffiths and Neumann-Held 1999) — for further discussion in future works. One can consider, however, the compatibility of this conservative interpretation with molecular biology as a strength rather than a weakness of the account, since it can be taken to mean that it is possible to develop in this way a more consistent understanding of 'information' which can be smoothly integrated into established knowledge in molecular biology. It is just a question, then, of evaluating the pros and cons of building a semantic/pragmatic concept of information inside current paradigms in molecular biology, or of striving for promoting a conceptual revolution in this science and other fields from the standpoint of biosemiotics. One or the other project will seem attractive for different groups of researchers.

The arguments developed in this work hopefully show how the conceptual and methodological tools offered by biosemiotics can help us make it more precise what is information in biological systems. Thus, we may go beyond the unfortunate situation of information talk in biology as a loose bunch of metaphors with no clear meaning, to such an extent that some philosophers have suggested that the best thing to do would be to eliminate it. We hope our arguments have shown how biosemiotics can contribute to the project of building a theory about information in biology, including both semantic and pragmatic dimensions of information.

Rather than eliminating information talk from biology, the biosemiotic point of view regards it as essential to a way of conceiving biological systems that grasp their fundamental difference from standard chemical and physical systems. It is just the case of employing adequate tools, such as those offered by Peirce's theory of signs, to clarify the nature of information in a living system. Consider, for instance, Oyama's (2000 [1985]) argument that, while genes are represented as if they contained information about how an organism will develop, they will continue to be treated as determining causes, lending support to genetic determinism. The notion of information arising from the biosemiotic analysis presented above suggests that the problem is not really in the notion of information in itself, but rather in the way information has been typically conceived in genetics

and molecular biology, namely, in such a way that it was reduced to merely sequential information in a string of DNA. If we consider, as above, that strings of DNA only contain potential genes; that potential information, although arguably carried by stretches of DNA, should be treated as the potentiality of a process; and effective information is a triadic-dependent process including not only Signs in DNA but also Objects and Interpretants, we will be in a much better position to picture information as being fundamentally dynamical and distributed, being related not only to genes but to any structure that can act as a Sign. Maybe, we can even follow Keller (2000: 146) in her suggestion of a cellular program that is not limited to DNA, but is rather a shared program in which all cell components function alternatively as 'instructions' and 'data' (or, more precisely, Signs).

Williams (1997: 476–477), in his summary of a symposium promoted by Ciba Foundation to discuss the future of the reductionist approach in biology, states that the biggest challenge to reductionism comes from the concept of information. Some participants of that meeting, he reports '... felt that a deeper understanding of the role of information may yet throw a spanner in the grand reductionist scheme.' In a sense, he argues, information in biological systems is 'fully consistent with' reductionist principles of physics and chemistry, because it is 'carried and received by molecules.' In terms of a biosemiotic analysis, these molecules that carry and receive 'information' are regarded as just potential Signs. These potential Signs, however, to be effectively 'informational' should be part of triadic-dependent, semiotic processes, involving more than just these potential entities, as we argued above. The conclusion that information is a process rather than an entity shows how a careful analysis of what is information in biological systems, based on a coherent framework such as that of semiotics, can indeed overcome a one-sided reliance on reductionist approaches to biology.

According to the picture presented in this paper, the meaning of a gene is highly context-sensitive. After all, information is highly context-sensitive, and genes can only mean something by being Signs within a triadic-dependent process defined here as information. A Peircean approach to the concepts of gene and information entails that both should be seen in the contexts in which information is handled by an interpreter, a conclusion which is in accordance with ideas stressed by a number of authors involved in the debates about these concepts (e.g., Nijhout 1990; Keller 2000; Hall 2001; Jablonka 2002) and highlights the pragmatic dimension of genetic information. The meaning of a gene is not contained in the sequence of nucleotides in a string of DNA, but rather emerges as a process involving the larger system by which genes are interpreted.

As Hall (2001) emphasizes, however, one of the major unresolved problems in biology is how to place genes *in context*. His answer is in agreement with our conclusions in this paper: 'Simply, the gene's home, context, and locus of operation is the cell.'[34] This may seem at first sight too obvious to be of any relevance. Nevertheless, as Hall stresses, we have been slow to recognize that the cell is not only the place where genes reside, but 'the cell enables the gene, allowing it to play its role(s) in development and evolution' (Hall 2001: 226).

The very fact that the mainstream representation of genes is such that all information in a cell ended up being deposited in DNA shows how slow we have been to recognize that DNA is enabled by the cell to perform the roles it performs. As a consequence of the semiotic analysis offered in this paper, the interpretation of what is information in a cell system and, in particular, of how potential Signs in DNA can be actualized so as to be part of effective information, clearly ascribes to the cell, as an interpreter of Signs in DNA, the capacity of *enabling* genes, much in the sense proposed by Hall (2001). Cells enable DNA to perform its roles by harnessing the behavior of this macromolecule so as to make it operate in a particular way that is demanded by a given environmental situation a cell faces at certain locations and times. The mechanisms that allow cells to constrain the operation of DNA to their own needs involve the establishment of boundary conditions by a macrosemiotic level of signaling pathways, as discussed above. They show that DNA molecules are governed by the cell, rather than command the cell in a dictatorial way, as the metaphors of genetic 'programs' and 'controllers' suggest. Biological systems function by means of a 'democratic' rather than a 'dictatorial' control structure, *i.e.*, there is neither genomic nor metabolic supremacy over other cellular processes (Bruggeman et al. 2002).

The recognition that the cell is the context of genes unravels new and difficult challenges, for instance, that of identifying and understanding the spatial and temporal contexts (often quite complex and multifaceted) in which genes operate, or, as Hall (2001: 228) puts it, 'to unravel the epigenetic code underlying developmental evolution,' which is far more complicated than the genetic code of four 'letters' arranged in groups of three, encompassing a whole array of genetic and non-genetic factors that ultimately lead from genotype to phenotype by means of development. We should leave, however, the analysis of other biological features relevant to the project of a semiotic analysis of the genetic information system for subsequent works. The arguments developed in this work are, in our view, sufficient to show both the relevance and the far-reaching consequences of such an analysis.

Notes

* Charbel Niño El-Hani thanks the Brazilian National Research Council (CNPq) for research grant number 302495/02–9, post-doctoral studies grant number 200402/03–0, and funding of the project number 402708/2003–2. João Queiroz thanks the State of São Paulo Research Foundation (FAPESP) for the post-doctoral studies grant number 02/09763–2. Claus Emmeche thanks the Faculty of Science, University of Copenhagen. We are indebted to Jesper Hoffmeyer, Kalevi Kull, Luis Bruni, Mia Trolle Borup, Peter Harries-Jones, Soren Brier, Alexei Sharov, Don Favareau, and Lenny Moss for discussions in the presentation of a previous version of this paper at Gatherings in Biosemiotics IV in Prague in 2004. We are also thankful to Wolfgang Hofkirchner and Toshiyuki Nakajima for their comments about the paper while one of us (C.N.E.H.) were visiting the Center for Advanced Studies in Information and Communication Technologies and Society (ICT and S) at the University of Salzburg.

1. For critiques of this way of representing DNA and, consequently, genes (if they are regarded, as usual, as strings of DNA), see, among several other works, Oyama 2000 [1985]; Nijhout 1990; Moss 1992; Smith 1994; Sarkar 1996; El-Hani 1997; Keller 2000).
2. It is important to avoid losing from sight that the distinction between gene-P and gene-D is not identical to the distinction between classical and molecular genes. Molecular entities can be treated as genes-P. For details, see the original works.
3. This is only a methodological decision, made for the sake of simplicity. We intend to build a theoretical framework for future analyses on the grounds of the more well established case, subsequently applying it to more difficult and less well-known cases.
4. 'Sign' was used by Peirce to designate the irreducible triadic relation between S, O and I as well as to refer to the first term of the triad (sometimes 'Representamen'). Some commentators proposed, then, that we should distinguish between 'Sign in a broad sense' and 'Sign in strict sense' (e.g., Johansen 1993: 62). Charles Morris (see Nöth 1995: 80) proposed the use of the expression 'Sign vehicle' in the place of 'Sign in strict sense.' We will systematically use the term 'Sign' in this paper to refer to the first term of the triad, and 'semiosis,' to refer to the whole triad. We will not use the expression 'Sign vehicle' often but sometimes we will employ it due to its interesting metaphorical connotations in the case of biological systems, in which the first term of a triadic-dependent process is typically a physical entity, such as a molecule or set of molecules, as, e.g., DNA. In these cases, we apply the notion of 'Sign vehicle' especially to emphasize the 'material quality' of a Sign (*CP* 5.287).
5. When a part or subsystem of a system is the interpreter, its actions as an interpreter will be typically subordinated, i.e., regulated by the system as a whole (that we will call, in this case, a 'global' interpreter). We can call, as Jablonka (2002), the subordinated interpreters 'interpretative systems' within a global interpreter. It can happen that a system loses its control over one or more of its included interpreters. In this case, dysfunctional states may result from the interpretation of Signs in that system. It is possible to analyze in these terms, for instance, events in carcinogenesis in which stretches of DNA are transcribed in a place and/or time in which they were not supposed to be transcribed. These would be misinterpretation events. By 'misinterpretation,' we mean the interpretation of a Sign that does not lead to a successful coping with a system's circumstances, i.e., that does not contribute to the maintenance of the dynamic stability of a system in a given circumstance.
6. In the context of our analysis, we will not employ the concept of Final Interpretant. It will not play an important role in our current arguments, and, thus, we consider we can leave it to subsequent works.

7. Notice that our argument is grounded on a passage found in a late work by Peirce, in which the idea of a Sign as a medium of communication of a form is prominent (*EP* 2 2.544, note 22; also *EP* 2 2:329, *EP* 2 2:389–2.391).
8. These books are also the basic sources in this section, unless otherwise noticed. As this section only summarizes some elementary ideas in cell and molecular biology, any reader who doesn't feel any need of perusing these notions can simply skip it.
9. If alternative splicing does not occur, it will be the case that Signs in DNA and Signs in mature mRNA will be equivalent.
10. In the case of genes, the Objects at stake are entities, as described above. Nevertheless, it is important to bear in mind that, in Peirce's framework, it is not the case that the Object of a Sign *should* necessarily be an entity, a thing, or even an existent. Consider, for instance, the following passage: 'The Objects — for a Sign may have any number of them — may each be a single known existing thing or thing believed formerly to have existed or expected to exist, or a collection of such things, or a known quality or relation or fact, which single Object may be a collection, or whole of parts, or it may have some other mode of being, such as some act permitted whose being does not prevent its negation from being equally permitted, or something of a general nature desired, required, or invariably found under certain general circumstances' (*CP* 2.232).
11. Symptomatically, Godfrey-Smith (1999) and Griffiths (2001) argue that developmental information is not stored in the genetic code, because the formal coding relation between codons in DNA and amino acids in polypeptides specifies only the primary structure of proteins. To be more precise, we should consider here proteins that acquire their mature conformation spontaneously. These proteins show the property of self-assembly. In this case, the three-dimensional structure of a protein simply follows from its primary sequence of amino acids, and, therefore, the Immediate Object directly determines the Dynamical Object. (Here we find yet another peculiar feature of the genetic information system, when compared to the standard Peircean framework). There are a number of proteins, however, that cannot self-assemble and should be assisted by proteins called chaperones in order to acquire their proper structures. In this case, the sequence of amino acids, the Immediate Object, only indicates the functional protein, the Dynamic Object. In the text, we are dealing particularly with this case, which fits Peirce's understanding of the relationship between Immediate and Dynamical Objects. Chaperones can be treated, in these terms, as part of the habits cells acquired in evolution.
12. In our analysis, we are dealing with genes-D, as defined by Moss (2001, 2003), but by introducing the idea that the sequence of amino acids of a protein is the Dynamical Object represented in a Sign in DNA in its semiotically available form, we can be accused of introducing in our analysis a risky conflation between gene-P and gene-D. Even though we agree with Moss' diagnosis that genetic determinism is supported by the confusion between genes as determinants of phenotypes (gene-P) and genes as developmental resources (gene-D), we think there is no problem with regarding genes as determinants of phenotypes at the level of the primary structure of a protein. Or, to put it differently, as regards the relationship between sequences of nucleotides and sequences of amino acids at the level of translation, we believe no serious problem follows from understanding the primary structure of proteins as being represented in DNA. We don't see how genetic determinism, the main threat Moss has in view when vigorously criticizing the conflation of gene-P and gene-D, might follow from the claim that components of the primary structure of proteins are semiotically available in DNA in the form of nucleotide sequences. This problem only appears, in our view, when one considers that phenotypic levels higher than that of the primary structure of proteins

are determined by genes. Genetic determinism is also avoided in the context of our analysis if we take into account that, yet at the cell level, functional proteins as Dynamical Objects often are not determined by DNA sequences, but only indicated by them and found by cells through habits established through evolution and development. Thus, often it is not even the case that the three-dimensional structure of a protein is completely determined by, or even represented in, DNA in our picture. Furthermore, our analysis keeps genes on the same plane as other biomolecules involved in development, giving them no causal privileging. We work here with genes as developmental resources which represent a range of possible proteins (their Immediate Interpretant), which will in turn be embedded in complex causal structures, in which many other molecules play important causal roles. In fact, we don't see a possible conflation of gene-P and gene-D as a sufficient basis for arguing that one should abandon such a basic idea in molecular biology as that of genetic coding. We consider that there is a fundamental difference between talking about genes-D as representing amino acid sequences and talking about genes-D as determinants of organismic phenotypes. This latter way of talking should be avoided, as Moss rightly argues, since it involves a commitment to a preformationist, determinist view of the relationship between genotype and phenotype. The former, however, is in fact implied by an idea that appears in Moss' works themselves, namely, that a gene-D is a specific nucleic acid template. We think our analysis is compatible with the idea that genes-D contain 'molecular template resources' involved in the synthesis of 'gene products,' as Moss argues (see, e.g., Moss 2001: 88).

13. In a Peircean framework, the Immediate Object can be understood as the characteristics selected in the Sign as a means of indicating the Dynamical Object. It is not the case, in this framework, that the Immediate Object is a condition of possibility to the Dynamical Object. Nevertheless, in the case we are analyzing here the interpreter creates a Dynamical Object of a given class (showing a given habit) on the grounds of indications present in the Sign. A cell uses Signs in DNA as a basis for synthesizing a Dynamical Object sufficiently resembling a past Dynamical Object that does not exist anymore but resulted in successful, adaptive experiences. This is the reason why we claim that, in this case, the Immediate Object establishes conditions of possibility to the Dynamical Object.
14. The irreducibility of the triadic relation S-O-I is a logical property. Therefore, while it makes no sense to sort out a primary constraining factor in such a logical relation, dynamically it makes sense to sort out the Dynamical Object as the primary constraining factor of semiosis (for a detailed discussion about this issue, see Short 1998: 31).
15. In this picture, it is important to take in due account that we are not claiming that DNA causes or brings about the protein as an Object, since DNA is a set of data (or, as we prefer, signs) rather than a program, a source of materials rather than a master agent in the cell. It is the DNA processing system that produces the proteins. We are not claiming, therefore, that the Sign causes the Object.
16. The regulation of a focal-level process by higher-level boundary conditions is interpreted here as a kind of selective process. Suppose that the causal relation between a given element of a system, *A*, and another element of the same system, *B*, is regulated. This is understood, in this framework, as the selection of *B* as the most probable effect of *A*, among other possible effects, by boundary conditions established by a level higher to the level where the causal relation at stake is taking place. This is related to ideas found in Polanyi (1968), who introduced the term 'boundary conditions' in the biological sciences, and Campbell (1974), who introduced the expression 'downward causation,' commonly employed in discussions about emergence. For Polanyi (1968), boundary conditions are higher-level general principles that control or delimit

lower-level processes. Campbell (1974: 180), in turn, claims that 'all processes at the lower level of a hierarchy are restrained by and act in conformity to the laws of the higher levels.' These higher-level laws act as 'selective systems' for lower-level processes. Van Gulick (1993) also understands downward causation as a selective restraint due to boundary conditions, which he regards basically as real and causally potent 'patterns of organization.'

17. The choice of a focal level depends on the purpose of a given research. Therefore, a researcher can choose as a focal level in her investigation a level in which other researcher, guided by another purpose, locates the boundary conditions in the triadic system she is studying. We will see below that a higher level constraining focal-level semiotic processes can include itself semiotic processes. The latter would turn into focal-level processes for a research aiming specifically at studying them.
18. There are several issues to be dealt with in order to apply in a consistent way the idea of emergence to the understanding of semiotic processes. Nevertheless, it is not in the scope of this paper to address all these issues, even though some of them are considered in the arguments put forward below. A thorough treatment of the conditions which should be fulfilled for semiosis to be characterized as an emergent process is found in Queiroz and El-Hani (2004, in press).
19. Even though we will not pursue this issue in this paper, we should emphasize that there is a clear correspondence between the hierarchical structure proposed by Salthe and Peirce's categories. The micro-semiotic level — at which processes relating S, O, and I are initiated — gives Sign processes an inevitable character of indeterminacy. It is straightforward, then, to associate the micro-semiotic level with firstness. Salthe himself stresses that this level exhibits a fundamentally stochastic behavior. At the focal level, specific, particular processes are spatiotemporally instantiated, as *tokens*, which are cases of secondness. The macro-semiotic level, in turn, gives Sign processes their *generality* and *temporality*, making them historical and context-dependent. We can say, thus, that macro-semiotic levels show the nature of thirdness. The stochastic behavior at the micro-semiotic level establishes *potentialities* for the particular Sign processes that are instantiated at the focal level. These potentialities are not the same as mere *possibilities*. For the sake of our arguments, consider Peirce's treatment of Quality as a 'mere abstract potentiality' (*CP* 1.422). Quality has the nature of firstness, being essentially indeterminate and vague. But we can also talk about a generality of Quality. In this case, we are beyond the realm of pure firstness, as generality refers to some law-like tendency, and thus to the nature of thirdness. Peirce works, in this case, with a merging of firstness and thirdness. It is in this latter sense that we understand potentialities at the micro-semiotic level here, as a particular set of potential Signs, Objects, and Interpretants which have been established due to the fact that the micro-semiotic level is embedded in a hierarchical system which includes levels showing the nature of secondness and thirdness (focal and macro-semiotic levels, respectively). These potentialities show, thus, the nature of a generality, being closer to a merging of firstness and thirdness than to pure firstness. Such a treatment seems to be compatible with Peirce's categoreal scheme, since, as Potter (1997: 94) stresses, 'the categoreal structure which Peirce uses is ... highly subtle and complex, admitting of various combinations.'
20. It is in this sense that we will talk about 'potential genes' in the following sections. Cf. Jablonka's (2002: 586–587) hypothetical example of proto-cells in which DNA was not used as an 'informational resource' but as a high-energy storage polymer. If, in such proto-cells, a given stretch of a storage DNA polymer had, by chance, the precise sequence coding for, say, fibronectin, this would have '... of course ... no special consequences for the proto-cell, since there is no cellular system that can interpret this

sequence in a specific way' (Jablonka 2002: 586–587). That is, that coding stretch of DNA would always remain a potential Composite Sign, a potential gene, and, as it would never get actualized, it would mean nothing to a cell, it would be no effective information at all, precisely for the lack of a Dynamical Interpretant.

21. For the sake of clarity, it is important to emphasize that this claim applies to the event in which nucleotide sequences in mRNA become part of actual triads *qua* Signs, since, if they were transcribed, they were already part of actual triads, but *qua* Immediate Objects.
22. But notice that, even though an actualization of potential triads indeed takes place in transcription, this is just a first step in the process of actualization of a potential gene in DNA, which will involve several other steps. The actualization of a potential Sign in DNA results in another potential Sign, now in pre-mRNA.
23. Just for the sake of the argument, we are skipping RNA splicing here. It is as if we were analyzing a string of DNA resulting in a pre-mRNA with only one mature mRNA as the result of its processing. This doesn't mean, however, that we don't take RNA splicing into account in our analysis, as the previous comments about the FN gene in this very section show. If we consider that splicing patterns are cell type-, development-, and age-regulated by mechanisms involving signaling pathways, we will see that semiotic processes play a role also in pre-mRNA splicing. For reasons of space, we will not explore this avenue in the scope of this paper, leaving it for subsequent works.
24. We should still talk about potential Signs in this case because the process of actualization of a potential gene can still be interrupted, as it depends on the availability of specific amino acids. Consider, for instance, the case of a starving animal that can lack some amino acids necessary for protein synthesis.
25. This analysis should be made somewhat more complex to accommodate mRNA editing (Hanson 1996; Lewin 2004), a process in which individual bases are added to or deleted from mRNA during processing. In this case, Simple and Composite Signs in mRNA are changed in such a manner that the Composite Immediate Object has a different sequence, in the end, from that which is semiotically available in DNA. As Lewin (2004: 742, emphasis added) puts it, 'RNA editing is a process in which *information changes at the level of mRNA*. It is revealed by situations in which the coding sequence in an RNA differs from the sequence in DNA from which it is transcribed.' Nevertheless, mRNA editing is a rare phenomenon, and, thus, we can say that it doesn't affect our analysis to a great extent, even though we should leave room to accommodate it.
26. In the following arguments, we will not focus on genes which are constitutively expressed, such as housekeeping genes, but rather on genes that can be turned on or off depending on the context in which a cell is embedded.
27. We use the expression 'extra- or intracellular environment' mostly for the sake of simplicity. There is, in fact, a hierarchy of 'contexts,' 'environments,' or, in our own terms, semiotic levels that can direct gene expression (*i.e.*, establish boundary conditions for the selection of potential genes in DNA), ranging from systems of gene-gene interactions to organisms, and passing through nucleus, cytoplasm, cell, cell surface, extracellular matrix, morphogenetic fields, collective condensations of cells (blastemas), organs, etc. (see, for example, Hall 2001).
28. This shows how the metaphor that information 'flows' in a cell is inadequate. First, it is redundant to say that information flows, since information is itself a process. Second, it leads to the misinterpretation that information is a sort of entity going from one place to another. See Hoffmeyer (2002).

29. In this connection, it is interesting to note that the genetic information system and the human symbolic culture are regarded by Jablonka (2002) as the only information systems capable of transmitting 'latent' ('potential,' 'non-expressed,' 'non-actualized') forms of information. Some kinds of epigenetic cellular information systems also can, in her view, sometimes transmit latent information. Given the framework developed in the present paper, this restriction of the capacity of transmitting potential information to only a subset of biological information systems demands careful appraisal. Nevertheless, we will limit, for the moment, our conclusions to the genetic information system, leaving this issue to be dealt with elsewhere.
30. In fact, the semiotic analysis we developed can be interpreted in a more conservative way, in which 'information' is treated as a process, but 'genes' are still treated in a manner which is close to viewing them as 'entities' in DNA. This interpretation is the one presented here. It is in line with a concept Griffiths and Neumann-Held (1999) named 'the contemporary molecular gene concept.' A bolder interpretation, in which 'genes' themselves are conceived as processes, particularly when they are expressed, can be also proposed, but we opted for leaving this interpretation to be presented elsewhere. It shows remarkable similarities with the 'molecular process gene concept' (Neumann-Held 1999, 2001; Griffiths and Neumann-Held 1999), and, as that concept, faces some difficulties as regards the individuation of genes, as Moss (2001) argues. As we don't have space here to address this controversy, we think it is better to leave this more daring interpretation to be dealt with elsewhere.
31. Two modes of 'information transmission' are usually recognized: 'horizontal transmission,' between individuals belonging to the same generation, and 'vertical transmission,' from one generation to another.
32. A more colloquial example may help us show the correctness of an understanding of what is usually taken as 'information transmission' as the communication of potential information. Suppose you send an e-mail to somebody else but, ultimately, the e-mail is lost in the intricacies of the worldwide web, so that the receiver never really reads the message. The message will obviously have no effect at all on that receiver. That is, what you sent through the web was not effective, but rather potential information, that only when interpreted, turned into an effective triadic-dependent process, i.e., effective information. As the message in this case never reached the recipient, it remained as just potential information. The idea that, without an interpretation that actualizes it, a Sign, when transmitted, is only potential information also help us understand why information in living systems is irreducible to the physicochemical carriers of its potentiality. Considering the above example, when you send an e-mail to a receiver that is never read, you indeed produced a change in physical states in a number of computers throughout the world, but, as interpretation never happened, no effective information was ever produced. Therefore, there is something more to effective information, to Signs in effective action, than simply physicochemical carriers, and this 'something more' is a triadic-dependent process by means of which Signs, Objects, and Interpretants are dynamically interrelated.
33. To understand what we mean by 'type' here, consider that if we have a protein-token, say, a particular molecule of fibronectin, it is a token of a type, fibronectin. It is because of the communication of a general form, which define 'fibronectin,' from one generation to another by Signs in DNA that a particular, a fibronectin-token, is reconstructed in a given generation.
34. In fact, as Hall (2001: 228) also recognizes, genes have multilevel homes or contexts, given the nested structure of living systems. Cells are given preeminence for their widely acknowledged role as fundamental units of organic structure and function,

and, also, for the fact that there are an enormous number of (unicellular) organisms which have no level higher than the cell.

References

Adami, Christoph (2004). Information theory in molecular biology. *Physics of Life Reviews* 1, 3–22.

Alberts, Bruce, Johnson, Alexander, Lewis, Julian, Raff, Martin, Roberts, Keith, and Walter, Peter (2002). *Molecular Biology of the Cell*, 4th ed. New York: Garland Science.

Atlan, Henry and Koppel, Moshe (1990). The cellular computer DNA: Program or data? *Bulletin of Mathematical Biology* 52 (3), 335–348.

Bergman, Mats (2000). Reflections on the role of the communicative sign in semeiotic. *Transactions of the Charles S. Peirce Society* 36 (2), 225–254.

Black, Douglas L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry* 72, 291–336.

Bruggeman, Frank J., Westerhoff, Hans V., and Boogerd, Fred C. (2002). Biocomplexity: A Pluralist research strategy is necessary for a mechanistic explanation of the 'live' state. *Philosophical Psychology* 15 (4), 411–440.

Bruni, Luis Emilio (2003). A sign-theoretic approach to biotechnology. Ph.D. dissertation, Institute of Molecular Biology, University of Copenhagen.

Brunning, Jackeline (1997). Genuine triads and teridentity. In *Studies in the Logic of Charles Sanders Peirce*, Nathan Houser, D. Roberts, and J. Evra (eds.), 252–270. Bloomington: Indiana University Press.

Brunning, Jackeline and Forster, Paul (eds.) (1997). *The Rule of Reason: The Philosophy of Charles Sanders Peirce*. Toronto: University of Toronto Press.

Campbell, Donald T. (1974). 'Downward causation' in hierarchically organized biological systems. In *Studies in the Philosophy of Biology: Reduction and Related Problems*, Francisco Ayala and Theodosius Dobzhansky (eds.), 179–186. Berkeley: University of California Press.

Colapietro, Vincent (1989). *Peirce's Approach to the Self: A Semiotic Perspective on Human Subjectivity*. Albany: State University of New York Press.

Cover, Thomas M. and Thomas, Joy A. (1999). Information theory. In *The MIT Encyclopedia of the Cognitive Sciences*, Robert A. Wilson and Frank C. Keil (eds.), 404–406. Cambridge MA: MIT Press.

Coyne, Jerry A. (2000). The gene is dead: Long live the gene. *Nature* 408, 26–27.

Crick, Francis H. (1958). On protein synthesis. *Symposium of the Society of Experimental Biology* 12, 138–163.

Deacon, Terrence (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W.W. Norton.

Debrock, Guy (1996). Information and the metaphysical status of the sign. In *Peirce's Doctrine of Signs — Theory, Applications, and Connections*, V. Colapietro and T. Olshewsky (eds.), 80–89. Berlin/New York: Mouton de Gruyter.

De Tienne, André (1992). Peirce's semiotic monism. In *Signs of Humanity/L'homme et ses signes. Proceedings of the Fourth International Congress/Actes du Quatrième Congrès Mondial* (= Approaches to Semiotics 107), International Association for Semiotic Studies/Association Internationale de Sémiotique, G. Deledalle, M. Balat, and J. Deledalle Rhodes (eds.), 1291–1303. Berlin: Mouton de Gruyter.

—(2003). Learning *qua* semiosis. *S.E.E.D.* (3), 37–53.

El-Hani, Charbel Niño (1997). Explicações causais do desenvolvimento: são os genes suficientes? *Cadernos de História e Filosofia da Ciência*, série 3, 7 (1), 121–167.

Emmeche, Claus (2003). Causal processes, semiosis, and consciousness. In *Process Theories: Crossdisciplinary Studies in Dynamic Categories*, Johanna Seibt (ed.), 313–336. Dordrecht: Kluwer.

Emmeche, Claus and Hoffmeyer, Jesper (1991). From language to nature — the semiotic metaphor in biology. *Semiotica* 84 (1/2), 1–42.

Emmeche, Claus, Køppe, Simo and Stjernfelt, Frederik (1997). Explaining emergence: Towards an ontology of levels. *Journal for General Philosophy of Science* 28, 83–119.

Falk, Raphael (1986). What is a gene? *Studies in the History and Philosophy of Science* 17, 133–173.

—(2000). The gene — A concept in tension. In *The Concept of the Gene in Development and Evolution*, Peter J. Beurton, Raphael Falk, and Hans-Jörg Rheinberger (eds.), 317–348. Cambridge: Cambridge University Press.

—(2001). Can the norm of reaction save the gene concept? In *Thinking about Evolution: Historical, Philosophical and Political Perspectives*, Rama S. Singh, Costas B. Krimbas, Diane B. Paul, and John Beatty (eds.), 119–140. New York: Cambridge University Press.

Fetzer, James (2001). *Computers and Cognition: Why Minds are not Machines.* Dordrecht: Kluwer.

Fitzgerald, John (1966). *Peirce's Theory of Signs as Foundation for Pragmatism*. University of Notre Dame: Mouton.

Fogle, Thomas (1990). Are genes units of inheritance? *Biology and Philosophy* 5, 349–371.

—(2000). The dissolution of protein coding genes. In *The Concept of the Gene in Development and Evolution*, Peter J. Beurton, Raphael Falk, and Hans-Jörg Rheinberger (eds.), 3–25. Cambridge: Cambridge University Press.

Freadman, Anne (2004). *The Machinery of Talk: Charles Peirce and the Sign Hypothesis.* Stanford: Stanford University Press.

Freeman, Eugene (1983). *The Relevance of Charles Peirce.* La Salle IL: Monist Library of Philosophy.

Gavin, Anne-Claude, Bösche, Markus, Krause, Roland, Grandi, Paola, Marzioch, Martina, Bauer, Andreas, Schultz, Jörg, et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141–147.

Gelbart, William (1998). Data bases in genomic research. *Science* 282, 660.

Gilbert, Walter (1978). Why genes in pieces? *Nature* 271, 501.

Godfrey-Smith, Peter (1999). Genes and codes: Lessons from the philosophy of mind? In *Biology Meets Psychology: Philosophical Essays*, Valerie Gray Hardcastle (ed.), 305–331. Cambridge: MIT Press.

—(2000). Information, arbitrariness, and selection: Comments on Maynard Smith. *Philosophy of Science* 67 (2), 202–207.

Gomes, Antonio, Gudwin, Ricardo, and Queiroz, João (2003). On a computational model of Peircean semiosis. In *Proceedings of the International Conference on Integration of Knowledge Intensive Multi-Agent Systems*, KIMAS'03, H. Hexmoor (ed.), 703–708. Los Alamitos CA: IEEE.

Gray, Russell D. (1992). Death of the gene: Developmental systems fight back. In *Trees of Life: Essays in the Philosophy of Biology*, Paul E. Griffiths (ed.), 165–209. Dordrecht: Kluwer.

Griffiths, Anthony J. F., Gelbart, William M., Miller, Jeffrey H., and Lewontin, Richard C. (1999). *Modern Genetic Analysis*. New York: W. H. Freeman. Available online at http://www.ncbi.nlm.nih.gov/PubMed/.

Griffiths, Paul E. (2001). Genetic information: A metaphor in search of a theory. *Philosophy of Science* 68 (3), 394–403.
Griffiths, Paul E. and Neumann-Held, Eva (1999). The many faces of the gene. *BioScience* 49 (8), 656–662.
Grós, François (1989). *Civilização do Gene*. Lisbon: Terramar.
Hall, Brian K. (2001). The gene is not dead, merely orphaned and seeking a home. *Evolution and Development* 3 (4), 225–228.
Hanson, Maureen R. (1996). Protein products of incompletely edited transcripts are detected in plant mitochondria. *The Plant Cell* 8 (1), 1–3.
Hausman, C. (1993). *Charles Sanders Peirce's Evolutionary Philosophy*. Cambridge: Cambridge University Press.
Hoffmeyer, Jesper (1996). *Signs of Meaning in the Universe*. Bloomington: Indiana University Press.
—(2002). The central dogma: A joke that became real. *Semiotica* 138 (1/4), 1–13.
Hookway, Christopher (1992). *Peirce*. London/New York: Routledge and Kegan Paul.
—(2002). *Truth, Rationality, and Pragmatism: Themes from Peirce*. Oxford: Oxford University Press.
Houser, Nathan (1997). Introduction: Peirce as a logician. In *Studies in the Logic of Charles Sanders Peirce*, Nathan Houser, D. Roberts, and J. Evra (eds.), 1–22. Bloomington: Indiana University Press.
Jablonka, Eva (1994). Inheritance systems and the evolution of new levels of individuality. *Journal of Theoretical Biology* 170, 301–309.
—(2001). The systems of inheritance. In *Cycles of Contingency: Developmental Systems and Evolution*, Susan Oyama, Paul E. Griffiths, and Russell D. Gray (eds.), 99–116. Cambridge: MIT Press.
—(2002). Information: Its interpretation, its inheritance, and its sharing. *Philosophy of Science* 69, 578–605.
Jablonka, Eva and Szathmáry, Eörs (1995). The evolution of information storage and heredity. *Trends in Ecology and Evolution* 10, 206–211.
Jablonka, Eva, Lamb, Marion J., and Avital, Eytan (1998). 'Lamarckian' mechanisms in Darwinian evolution. *Trends in Ecology and Evolution* 13, 206–210.
Jakobson, Roman (1969). *Linguistica e Comunicação*. São Paulo: Cultrix.
Johannsen, Wilhelm (1909). *Elemente der Exakten Erblichkeitslehre*. Jena: Gustav Fischer.
Johansen, Jorgen Dines (1993). *Dialogic Semiosis*. Bloomington: Indiana University Press.
Kampa, Dione, Cheng, Jill, Kapranov, Philipp, Yamanaka, Mark, Brubaker, Shane, Cawley, Simon, Drenkow, Jorg, et al. (2004). Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Research* 14, 331–342.
Keller, Evelyn Fox (2000). *The Century of the Gene*. Cambridge: Harvard University Press.
Kitcher, Philip (1982). Genes. *British Journal for the Philosophy of Science* 33, 337–359.
Kull, Kalevi (1999). Biosemiotics in the twentieth century: A view from biology. *Semiotica* 127 (1/4), 385–414.
Lewin, Benjamin (2004). *Genes VIII*. Upper Saddle River NJ: Pearson Education.
Lizska, James J. (1990). Peirce's interpretant. *Transactions of the Charles S. Peirce Society*. 26 (1), 17–61.
Lodish, Harvey, Berk, Arnold, Matsudaira, Paul, Kaiser, Chris A., Krieger, Monty, Scott, Matthew P., Zipursky, S. Lawrence, and Darnell, James (2003). *Molecular Cell Biology*, 5th ed. New York: W. H. Freeman. Available online at http://www.ncbi.nlm.nih.gov/PubMed/.

Magurran, Anne (2000). Backseat drivers. Review of *The Century of the Gene* by E. F. Keller. *New York Times Book Reviews*, December 10, 26.
Maniatis, Tom and Reed, Robin (2002). An extensive network of coupling among gene expression machines. *Nature* 416, 499–506.
Maynard Smith, John (2000a). The concept of information in biology, *Philosophy of Science* 67 (2), 177–194.
—(2000b). The cheshire cat's DNA. *New York Review of Books* 47, 43–46.
Maynard Smith, John and Szathmáry, Eörs (1995). *The Major Transitions in Evolution*. Oxford: W. H. Freeman.
—(1999). *The Origins of Life: From the Birth of Life to the Origins of Language*. Oxford: Oxford University Press.
Merrell, Floyd (1995). *Peirce's Semiotics Now*. Toronto: Canadian Scholar's Press.
Modrek, Barmak and Lee, Christopher (2002). A genomic view of alternative splicing. *Nature Genetics* 30, 13–19.
Moss, Lenny (1992). A kernel of truth? On the reality of a genetic program. *PSA* 1, 335–348.
—(2001). Deconstructing the gene and reconstructing molecular developmental systems. In *Cycles of Contingency: Developmental Systems and Evolution*, Susan Oyama, Paul E. Griffiths, and Russell D. Gray (eds.), 85–97. Cambridge: MIT Press.
—(2003). *What Genes Can't Do*. Cambridge: MIT Press.
Murphey, Murray (1993). *The Development of Peirce's Philosophy*. Cambridge: Harvard University Press.
Neumann-Held, Eva (1999). The gene is dead — Long live the gene: Conceptualizing genes the constructionist way. In *Sociobiology and Bioeconomics: The Theory of Evolution in Biological and Economic Thinking*, Peter Koslowski (ed.), 105–137. Berlin: Springer.
—(2001). Let's talk about genes: The process molecular gene concept and its context. In *Cycles of Contingency: Developmental Systems and Evolution*, Susan Oyama, Paul E. Griffiths, and Russell D. Gray (eds.), 69–84. Cambridge: MIT Press.
Nijhout, Frederik H. (1990). Metaphors and the role of genes in development. *Bioessays* 12 (9), 441–446.
Nöth, Winfried (1995). *Handbook of Semiotics*. Bloomington: Indiana University Press.
Oyama, Susan (2000 [1985]). *The Ontogeny of Information: Developmental Systems and Evolution*, 2nd ed. Cambridge: Cambridge University Press.
Parker, Kelly (1998). *The Continuity of Peirce's Thought*. Nashville: Vanderbilt University Press.
Peirce, Charles S. (1931–1966). *The Collected Papers of Charles S. Peirce*, 8 vols., C. Hartshorne, P. Weiss, and A. W. Burks (eds.). Cambridge: Harvard University Press. [Reference to Peirce's papers will be designated *CP* followed by volume and paragraph number.]
—(1967). Manuscripts in the Houghton Library of Harvard University, as identified by Richard Robin, *Annotated Catalogue of the Papers of Charles S. Pierce.* Amherst: University of Massachusetts Press. [Reference to Peirce's manuscripts will be designated *MS*.]
—(1976). *The New Elements of Mathematics*, vol. 4, C. Eisele (ed.). Berlin: Mouton; Atlantic Highlands, NJ: Humanities Press. [Reference to volume 4 of Peirce's *New Elements* will be designated *NEM* 4.]
—(1982–). *Writings of Charles S. Peirce*, 5 vols., M. Fisch, E. Moore, and C. Kloesel (eds.). Bloomington: Indiana University Press. [Reference to Peirce's writings will be designated *W* followed by volume and page number.]
—(1998). *Essential Peirce. Selected Philosophial Writings*, vol. 2 (1893–1913). Peirce Edition Project (eds.). Bloomington: Indiana University Press. [Reference to volume 2 of *Essential Peirce* will be designated *EP* 2.]

Polanyi, Michael (1968). Life's irreducible structure. *Science* 160, 1308–1312.
Portin, Petter (1993). The concept of the gene: short history and present status. *Quarterly Review of Biology* 56, 173–223.
Potter, Vincent (1997). *Charles S. Peirce: On Norms and Ideals*. Amherst: University of Massachusetts Press.
Queiroz, João and El-Hani, Charbel Niño (2004) Towards a multi-level approach to the emergence of semiosis. *Tech Report DCA-FEEC-UNICAMP* 07–04, 1–21. ftp://ftp.dca.fee.unicamp.br/pub/docs/techrep/2004/RT-DCA-FEEC-07-2004.pdf
—(in press). Semiosis as an emergent process. *Transactions of the Charles S. Peirce Society*.
Ransdell, Joseph (1977). Some leadings ideas of Peirce's semiotic. *Semiotica* 19 (3/4), 157–178.
—(2003). The relevance of Peircean semiotic to computational intelligence augmentation. *S.E.E.D.* 3 (3), 5–36.
Rescher, Nicholas (1996). *Process Metaphysics: An Introduction to Process Philosophy*. New York: State University of New York Press.
Rios, Ricardo Iglesias (2004). O início do fim do gene, resenha de Keller, E. F. O Século do Gene. *Ciência Hoje* 34 (204), 72–73.
Salthe, Stanley N. (1985). *Evolving Hierarchical Systems: Their Structure and Representation*. New York: Columbia University Press.
Sampedro, Javier (2004). *Deconstruyendo a Darwin: Los Enigmas de la Evolución a la Luz de la Nueva Genética*. Barcelona: Crítica.
Santaella-Braga, Lucia (1994). Peirce's broad concept of mind. *European Journal for Semiotic Studies* 6 (3/4), 399–411.
Sarkar, Sahotra (1996). Biological information: A skeptical look at some central dogmas of molecular biology. In *The Philosophy and History of Molecular Biology: New Perspectives*, Sahotra Sarkar (ed.), 187–231. Dordrecht: Kluwer.
—(1998). Forty years under the central dogma. *Trends in Biochemical Sciences* 23, 312–316.
—(2000). Information in genetics and developmental biology: Comments on Maynard Smith. *Philosophy of Science* 67 (2), 208–213.
Savan, David (1986). Response to T. L. Short. *Transactions of the Charles S. Peirce Society* 22 (2), 125–143.
—(1987–1988). *An Introduction to C. S. Peirce's Full System of Semeiotic* (= Monograph Series of the Toronto Semiotic Circle 1). Toronto: Toronto Semiotic Circle.
Shannon, Claude E. and Weaver, Warren (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
Short, Thomas (1996). Interpreting Peirce's interpretant: A response to Lalor, Liszka, and Meyers. *Transactions of the Charles S. Peirce Society* 32 (4), 488–541.
—(1998). What's the use? *Semiotica* 122 (1/2), 1–68.
Smith, Kelly C. (1994). *The Emperor's New Genes: The Role of the Genome in Development and Evolution*. Ph.D. dissertation, Duke University.
Sorek, Roten, Shamir, Ron, and Ast, Gil (2004). How prevalent is functional alternative splicing in the human genome? *Trends in Genetics* 20 (2), 68–71.
Stephan, Achim (1998). Varieties of emergence in artificial and natural systems. *Zeitschrift für Naturforschung* 53c, 639–656.
—(1999). *Emergenz: Von der Unvorhersagbarkeit zur Selbstorganisation*. Dresden/Munich: Dresden University Press.
Sterelny, Kim (2000). The 'genetic program' program: A commentary on Maynard Smith on information in Biology. *Philosophy of Science* 67 (2), 195–201.
Stotz, Karola, Griffiths, Paul E., and Knight, Rob (2004). How biologists conceptualize genes: An empirical study. *Studies in the History and Philosophy of Biological and Biomedical Sciences* 35, 647–673.

Stuart, C. I. J. M. (1985). Bio-informational equivalence. *Journal of Theoretical Biology* 113, 611–636.

Tiercelin, Claudine (1995). The relevance of Peirce's semiotic for contemporary issues in cognitive science. In *Mind and Cognition: Philosophical Perspectives on Cognitive Science and Artificial Intelligence* (= Acta Philosophica Fennica 58), L. Haaparanta and S. Heinamaa (eds.), 37–74. Helsinki: Philosophical Society of Finland.

Van Gulick, Robert (1993). Who is in charge here? And who's doing all the work? In *Mental Causation*, John Heil and Alfred Mele (eds.), 233–256. Oxford: Oxford University Press.

Violi, Patrizia (1999). Semiotics and cognition. In *The MIT Encyclopedia of the Cognitive Sciences*, Robert A. Wilson and Frank C. Keil (eds.), 744–745. Cambridge: MIT Press.

Wang, Wen, Zhang, Jianming, Alvarez, Carlos, Llopart, Ana, and Long, Manyuan (2000). The origin of the *Jingwei* gene and the complex modular structure of its parental gene, *Yellow Emperor*, in *Drosophila melanogaster*. *Molecular Biology and Evolution* 17 (9), 1294–1301.

Waters, Kenneth (1994). Genes made molecular. *Philosophy of Science* 61, 163–185.

Wilkins, Adams S. (2002). Grappling with developmental complexity. *BioEssays* 24, 1193–1195.

Williams, Nigel (1997). Biologists cut reductionist approach down to size. *Science* 277 (5325), 476–477.

Wynnie, J. A. (2000). Information and structure in molecular biology: Comments on Maynard Smith. *Philosophy of Science* 67 (3), 517–526.

Charbel Niño El-Hani (b. 1968) is Associate Professor at the Instituto de Biologia, Universidade Federal da Bahia, Brazil ⟨charbel@ufba.br⟩. His research interests are philosophy of biology, biosemiotics, science education research, and animal behavior. His publications include 'Causação descendente, emergência de propriedades e modos causais aristotélicos' [Downward causation, property emergence, and Aristotelian causal modes] (with A. A. A. P. Videira, 2001); 'A pragmatic realist view of emergence' (with S. Pihlström, 2002); 'On the reality of emergents' (2002); and 'Modos de irredutibilidade das propriedades emergentes' [Modes of irreducibility of emergent properties] (with J. Queiroz, 2005).

João Queiroz (b. 1963) is a postdoc researcher at the State University of Campinas (UNICAMP) and an invited researcher at the Instituto de Biologia, Universidade Federal da Bahia (UFBA), Brazil ⟨queirozj@gmail.com⟩. His research interests are C. S. Peirce's pragmatism, biosemiotics, and cognitive science. His publications include *Semiose segundo C. S. Peirce* (2004); '10 cubes and 3N3: Using interactive diagrams to investigate Charles Peirce's classifications of signs' (with Priscila Farias, 2004); 'Abduction — between subjectivity and objectivity' (with Floyd Merrell, 2005); and *Artificial Cognition Systems* (with A. Loula and R. Gudwin, 2006).

Claus Emmeche (b. 1956) is Associate Professor and Center Director at the Center for the Philosophy of Nature and Science Studies, University of Copenhagen ⟨emmeche@nbi.dk⟩. His research interests are biosemiotics, philosophy of biology, and theoretical biology. His publications include 'On some theoretical grounds for an organism-centered biology: Property emergence, supervenience, and downward causation' (with Charbel N. El-Hani, 2000); *Reading Hoffmeyer, Rethinking Biology* (with Kalevi Kull et al., 2002); 'The chicken and the Orphean egg: On the function of meaning and the meaning of function' (2002); and 'Causal processes, semiosis, and consciousness' (2003).