*Semiotica,* 2006, Vol. 158 (1/4), pp. 233-254

# Life and Semiosis

## The real nature of information and meaning

**Marcello Barbieri**

Dipartimento di Morfologia ed Embriologia
Via Fossato di Mortara 64, 44100 Ferrara, Italy
brr@unife.it

ABSTRACT – Genes and proteins are molecular *artifacts* because they are *manufactured* by molecular machines that physically stick their subunits together in the order provided by external templates. This implies that all biological objects are artifacts, and therefore that *"life is artifact-making"*. Natural objects can be completely accounted for by physical quantities, whereas artifacts require additional entities like sequences and codes, or equivalent entities like information and meaning. Here it is shown that organic information and organic meaning are brought into existence by the molecular processes of copying and coding, which implies that, far from being metaphors, they are as real as the processes that produce them. It is also shown that they can be defined by operative procedures that make them as objective and reproducible as physical quantities. The result is that organic information and organic meaning are a new type of *fundamental* natural entities that here are referred to as *nominable* entities because they can be specified only by naming their components in their natural order. Any organic code is a correspondence between the objects of two independent worlds (genes and proteins) which is established by molecules that belong to a third world (RNAs). The elementary act of organic coding is therefore a relationship between three objects that can be referred to as "sign, meaning and adaptor", whereas the elementary act of cultural semiosis consists, according to Peirce, of "sign, meaning and interpretant". It is underlined that "organic semiosis" is implemented by codemakers and consists of *objective* organic reactions, whereas "mental semiosis" is performed by interpreters and is a *subjective* process. This means that organic semiosis does not require the existence of a mind at the molecular level, and the organic codes are natural processes that are based on objective, reproducible and fundamental natural entities.

KEYWORDS – *Information, meaning, organic codes, organic semiosis, mental semiosis, biosemiotics.*

### Introduction

Molecular biology is based on two key discoveries: one is that the linear sequence of nucleotides represents the *information* carried by a gene; the other is that the sequence of nucleotides in genes determines the sequence of amino acids in proteins. In both genes and proteins, therefore, *biological information* was identified with, and defined by, the specific sequence of their subunits.

This concept of information has solved the century old mystery of inheritance, and has turned the whole of biology from an energy-based into an information-based science. Despite all this, however, it has been pointed out that biological information is not a physical quantity, neither a fundamental nor a derived one, because a sequence cannot be measured. This is further underlined by the fact that there is another type of information which can be measured and which is a true physical quantity. It is the information of a message that Shannon defined in 1948 with an entropy-like formula, and which can be referred to as *physical information*.

The crucial point is that Shannon's information does not depend on the sequence of subunits, while biological information is defined precisely by that sequence. Physical information, in other words, has nothing to do with *specificity*, while biological information has everything to do with it. The two concepts are literally worlds apart, and this reinforces the conclusion that biological information is not, and cannot be, a physical quantity. So what is it?

According to an influential school of thought, biological information is a *metaphor*. It is a linguistic construction that we use in order to avoid long periphrases when we talk about living systems, but no more than that. It is like those computer programs that allow us to write our instructions in English, thus saving us the trouble to write them with the binary digits of the machine language. Ultimately, however, there are only binary digits in the machine language, and in the same way, it is argued, there are only physical quantities at the most fundamental level of Nature.

This conclusion, known as *"the physicalist thesis"* has been proposed in various ways by a number of scientists and philosophers (Chargaff, 1963; Sarkar, 1996; 2000; Mahner and Bunge, 1997; Griffiths and Knight, 1998; Griffith, 2001, Boniolo, 2003). It is probably one of the most deeply dividing issues of modern science. Many biologists are convinced that biological information is a real and fundamental component of life, but physicalists insist that it is only a metaphorical concept and that there is nothing fundamental about it because it is not a physical quantity. The purpose of this paper is to show that the physicalist thesis would be true if genes and proteins were made by spontaneous assemblies, like all inorganic molecules, but they are not. They are *molecular artifacts* because they are manufactured by molecular machines, and the physicalist thesis collapses because there is a fundamental difference between natural objects and artifacts. Natural objects can be completely accounted for by physical quantities whereas artifacts require additional *fundamental* entities that here are referred to as organic information and organic meaning.

## A new definition of life

Genes and proteins differ from inorganic molecules not because they have different structures and functions but because they are *produced* in a totally different way. All inorganic molecules are made by self-assembly and their structure is determined *from within*, i.e., by *internal* factors. Genes and proteins, instead, are produced by molecular machines which physically stick their subunits together in an order provided *from without*, by *external* templates. They are assembled by molecular robots on the basis of outside instructions, and this makes them as different from ordinary molecules as *artificial* objects are from *natural* ones. Indeed, if we agree that objects are natural when their structure is determined from within and artificial when it is determined from without, then we can truly say that genes and proteins are *artificial molecules,* that they are *artifacts made by molecular machines*. This in turn implies that all biological objects are artifacts, and we arrive at the general conclusion that the whole of life is artifact-making.

We may find it difficult to accept this idea, but let us not forget that it is based on the most basic experimental properties of genes and proteins. It is the direct consequence of the most surprising discovery of Molecular Biology, the discovery that there is a totally unexpected gulf between life and non-life. The great divide is not between organic and inorganic structures. It is between structures which are built from within and structures which are built from without, between molecules which are made by spontaneous assemblies and molecules which are manufactured by molecular machines.

The definition of life as *artifact-making* is centered on a single feature and this appears to go against a long tradition according to which life cannot be reduced to any one property. In reality no such contrast exists, because the new definition does not account for *all* aspects of life. It singles out the one feature that divides inanimate matter from the living world, and in so doing it tells us why living systems are fundamentally different from non-living ones, but that does not give us a complete description of life. What it does give us, however, is a new beginning. More precisely, a new theoretical framework for the study of the origin and the evolution of life.

## The origin of genes

The discovery that genes and proteins are *manufactured* molecules has direct implications for the origin of life, because it tells us that primitive molecular machines came into existence long before the origin of the first cells. The simplest molecular machines we can think of are molecules that could join other molecules together by chemical bonds, and for this reason we may call them *bondmakers*. Some could form bonds between amino acids, some between nucleotides, others between sugars, and so on. It has been shown, for example, that short pieces of ribosomal RNA have the ability to form peptide bonds, so it is possible that the first bondmakers were RNA molecules of small or medium-size molecular weights. Among the various types of bondmakers, furthermore, some developed the ability to join nucleotides together in the order provided by a *template*. Those bondmakers started *making copies* of nucleic acids, so we can call them *copymakers*, and their appearance was a real turning point, because it set in motion an extraordinary sequence of events.

The copying of a template is the elementary act of gene duplication, the very first step toward the phenomenon of *heredity*. When a process of copying is repeated indefinitely, furthermore, another phenomenon comes into being. Copying mistakes become inevitable, and in a world of limited resources not all changes can be implemented, which means that a process of selection is bound to take place. Molecular copying, in short, leads to *heredity*, and the indefinite repetition of molecular copying leads to *natural selection*. That is how natural selection came into existence. Molecular copying started it and molecular copying has perpetuated it ever since.

In the history of life, molecular copying came into being when the first copymakers appeared on the primitive Earth and started making copies of nucleic acids. This implies that *natural* nucleic acids had already been formed by spontaneous reactions on our planet, but that was no guarantee of evolution. Only the copying of genes could ensure their survival and have long-term effects, so it was really the arrival of copymaking that set in motion the extraordinary chain of processes that we call evolution. The first Major Transition of the history of life is generally described as the origin of genes, but it seems more accurate to say that it was the origin of molecular *copying*, or the origin of *copymakers*, the first molecular machines that started multiplying nucleic acids by making copies of them.

## The origin of proteins

The transition from natural to manufactured molecules was relatively simple for genes but much more complex for proteins, because genes can be copied while proteins cannot. Manufactured genes could be made simply by copying natural genes, and all that was required to that purpose were molecules which had a polymerase-like activity. Manufactured proteins, instead, could not be made by copying, and yet the information to make them had to come from molecules that can be copied, because only those molecules can be inherited. The information for manufacturing proteins, therefore, had to come from genes, so it was necessary to bring together a carrier of genetic information (a messenger RNA), a peptide-bondmaker (a piece of ribosomal RNA) and molecules that could carry both nucleotides and amino acids (the transfer RNAs). The first *protein makers*, in short, had to bring together three different types of molecules (messenger, ribosomal and transfer RNAs), and were therefore much more complex than copymakers.

The outstanding feature of the protein-makers, however, was not the number of components. It was the ability to ensure a one-to-one correspondence between genes and proteins, because without it there would be no biological specificity and no heredity. If the links between genes and proteins could have been determined by *stereochemistry*, as one of the earliest models suggested, a one-to-one correspondence would have been automatically ensured. Protein synthesis would have been but a form of *indirect copying*, and the problem of explaining its specificity would be relatively simple. The stereochemical hypothesis, however, turned out to be false. There simply is no chemical necessity in the links between codons and amino acids, and a one-to-one correspondence between them could only be the result of conventional rules. In short, only a genetic code could guarantee biological specificity, and this means that the evolution of the translation apparatus had to go hand in hand with the evolution of the genetic code. Protein synthesis arose from the integration of two different processes, and the final machine was a *"code-and-template-dependent-peptide-maker"*, or, more simply, a *codemaker*.

The second Major Transition of the history of life is generally described as the origin of proteins, but again it would be more accurate to say that it was the origin of *codemaking*, or the origin of *codemakers*, the first molecular machines that discovered molecular coding and started populating the Earth with codified proteins.

## Organic information

Biological (or *organic*) information has been defined as *the specific sequence of a molecule*, but this is not entirely satisfactory because it gives the impression that information is a *property* of molecules, something that molecules have simply because they have a sequence. In reality, there are countless molecules which have a sequence but only in a few cases this becomes information. This happens only when copymakers use it as a guideline for copying. Even copymakers, however, do not account, by themselves, for information. Copymakers can stick subunits together and produce sequences, but without a template they would produce only *random* sequences, not specific ones. Sequences alone or copymakers alone, in other words, have nothing to do with information. It is only when a sequence provides a guideline to a copymaker that it becomes *information* for it. It is only during an act of copying, in other words, that information comes into existence.

This tells us that organic information is not just the specific sequence of a molecule, but *the specific sequence produced by a copying process*. This definition underlines the fact that information is not a *thing* or a *property*, but *the result of a process*. It is, more precisely, an "operative" definition, because information is defined by the process that brings it into existence. There simply is no difference between saying that molecule B is a copy of molecule A, and saying that molecule B carries the same information as molecule A.

We realize in this way that organic information is as real as the copying process that generates it, but we still do not know if we can reduce it to simpler entities, and in particular to physical quantities. Luckily, this problem has a straightforward solution because the sequences of genes and proteins have two very special characteristics. One is that *a change to a biological sequence may produce a sequence which has entirely new properties*. This means that although a biological sequence can be said to have "components", it is at the same time a single indivisible whole. The second outstanding feature is that *from the knowledge of n elements of a biological sequence it is impossible to predict the element (n+1)*. This is equivalent to saying that *a specific sequence cannot be described by anything simpler than the sequence itself*. Organic information, in short, cannot be reduced to anything else, and this makes of it an *irreducible* (or *fundamental*) entity of Nature. But what kind of entity is it? How does it fit into our schemes?

According to a long tradition, natural entities are divided into *quantities* and *qualities*. Quantities can be measured and are objective, while qualities are subjective and cannot be measured. In the case of organic information, however, this scheme breaks down. Organic information is not a quantity because a specific sequence cannot be measured. But it is not a quality either, because linear specificity is a feature that we find in organic molecules, and is therefore an *objective* feature of the world, not a subjective one. A scheme based on quantities and qualities alone, in short, is not enough to describe the world. In addition to quantities (*objective and measurable*) and qualities (*subjective non-measurable)* we must recognize the existence in Nature of a third type of entities (*objective but not measurable*). Information is one of them, and we can also give it a suitable name. Since it can be described only by *naming* its sequence, we can say that organic information is a *nominable* entity, or that it belongs to the class of the nominable entities of Nature (Barbieri, 2004).

## Organic meaning

A code is a set of rules which establish a correspondence between the objects of two independent worlds. The Morse code, for example, is a correspondence between combinations of dots and dashes with the letters of the alphabet, and in the same way the genetic code is a correspondence between combinations of nucleotides and amino acids. Let us notice now that establishing a correspondence between, say, object 1 and object 2, is equivalent to saying that object 2 is the *meaning* of object 1. In the Morse code, for example, the rule that "dot-dash" corresponds to letter "A", is equivalent to saying that letter "A" is the meaning of "dot-dash". In the code of the English language, the mental object of the *word* "apple" is associated to the mental object of the *fruit* 'apple', and this is equivalent to saying that that fruit is the meaning of that word. By the same token, the rule of the genetic code that a codon corresponds to an amino acid is equivalent to saying that that amino acid is the organic meaning of that codon. Anywhere there is a code, be it in the mental or in the organic world, there is meaning. We can say, therefore, *that meaning is an object which is related to another object by a code*, and that organic meaning exists wherever an organic code exists (Barbieri, 2003).

The existence of meaning in the organic world may seem strange, at first, but in reality it is not stranger than the existence of codes because they are the two sides of the same coin. Saying that a code establishes a correspondence between two objects is equivalent to saying that one object is the meaning of the other, so we cannot have codes without meaning or meaning without codes. All we need to keep in mind is that *meaning is a mental entity when the code is between mental objects, but it is an organic entity when the code is between organic molecules*.

Modern biology and modern physics have readily accepted the concept of information but have carefully avoided the concept of meaning, and yet we have seen that organic information and organic meaning are both the result of natural processes. Just as it is an act of *copying* that creates organic information, so it is an act of *coding* that creates organic meaning. Copying and coding are the processes; copymakers and codemakers are their agents; organic information and organic meaning are their results, or the kind of natural entity that they belong to. But the parallel goes even further than that. We have seen that organic information *cannot be measured*, and the same is true for organic meaning. We have seen that organic information is an *objective* entity, because it is defined by the same sequence for any number of observers, and that is also true for organic meaning, which is defined by coding rules that are the same for all observers. Finally, we have seen that organic information is an *irreducible* entity, because it cannot be described by anything simpler than its sequence, and the same is true for organic meaning, which cannot be defined by anything simpler than its coding rules.

Organic information and organic meaning, in short, belong to the same class of entities because they have the same general characteristics: they both are *objective-but-non-measurable* entities, they both are *irreducible* (or *fundamental*) entities of Nature, and since we can describe them only by naming their components, they both are *nominable* entities (Barbieri, 2004). Finally, let us underline that they both are the pillars of life because organic information comes from the copying process that produces genes, while organic meaning comes from the coding process that generates proteins.

## Operative definitions

Physical quantities have three fundamental properties: (1) they are *objective*, (2) they are *reproducible*, and (3) they are *defined by operative procedures*. This last property is particularly important because it has provided the solution to one of the most controversial issues of physics. The critical point was the theoretical possibility that the entity which is measured may not be the same entity which has been defined. This led to the idea that there should be no difference between what is measured and what is defined, i.e., to the concept of *operative* (or *operational*) definition: *a physical quantity is defined by the operations that are carried out in order to measure it.*

It was this operational approach that solved the definition problem in physics, and it is worth noticing that we can easily generalize it. Instead of saying that *a natural entity is defined by the operations that measure it*, we can say that a *natural entity is defined by the operations that evaluate it in an objective and reproducible way*. The advantage of this generalized approach is that it applies to all *objective entities*, so it can be used not only in physics, but in biology as well. To this purpose, we only need to notice that *a measurement* is an objective and reproducible description of a physical quantity, just as *the naming of a specific sequence* is an objective and reproducible description of organic information and just as *the naming of a coded object* is an objective and reproducible description of organic meaning.

While the physical quantities are evaluated *by measuring*, in other words, our biological entities are evaluated *by naming their components*, but in both cases the entities in question *are defined by the operations that evaluate them*, and this is the essence of the operative approach. Alternatively, we can say that organic information and organic meaning are defined by the processes of copying and coding that bring them into existence, and this too amounts to an operative definition (Barbieri, 2004).

We conclude that organic information and organic meaning can be defined by generalized operative definitions which are as reliable as the operative definitions of physics. This should ensure that they are no longer at the mercy of endless debates on terminology as they have been in the past. The operative definitions are scientific tools which are justified by their own prescriptions, so there is no point in asking if they are right or wrong. All we can ask of them is whether they contribute or not to our description and to our understanding of Nature.

At this point, we can summarize all the above concepts with the following statements:

(1) *Organic information is the specific sequence produced by a copying process*.

(2) *Organic meaning is the object which is related to another object by a code*.

(3) *Organic information and organic meaning are neither* quantities *nor* qualities. *They are a new kind of natural entities which are referred to as* nominable *entities.*

(4) *Organic information and organic meaning have the same scientific status as physical quantities because they are* objective *and* reproducible *entities which can be defined by operative procedures.*

(5) *Organic information and organic meaning have the same scientific status as* fundamental *physical quantities because they cannot be reduced to, or derived from, simpler entities*.

## How many organic codes?

According to modern biology, the genetic code is the only organic code that exists in the living world, whereas the mental world is populated by a virtually unlimited number of cultural codes. We know, furthermore, that the genetic code came into being with the origin of life, whereas the cultural codes arrived almost four billion years later, with cultural evolution. This appears to suggest that organic evolution went on for four billion years, almost the entire history of life on Earth, without producing any other organic code after the first one. At a closer inspection, however, we realize that that is far from the truth.

To this purpose, let us underline the difference that exists between the *copying* of genes and the *coding* of proteins, i.e. between *transcription* and *translation*. In transcription, an RNA sequence is assembled from the linear information of a DNA sequence, and in this case a normal biological catalyst (an RNA polymerase) is sufficient, because each step requires a single recognition process. In translation, instead, two independent recognition processes must be performed at each step, and the system that performs the reactions (the ribosome) needs special molecules, first called *adaptors* and then *transfer RNAs*, in order to associate codons to amino acids according to the rules of the genetic code. Without a code, in fact, a codon could be associated with different amino acids and *biological specificity*, the most precious of life’s properties, would be lost.

These concepts can easily be generalized. We are used to think that all biochemical processes are *catalyzed* reactions, but in reality we must distinguish between *catalyzed* and *codified* reactions. The catalyzed reactions are processes (like transcription) that require only one recognition process at each step. The codified reactions, instead, require two independent recognition processes at each step and a set of coding rules. The catalyzed reactions, in other words, require *catalysts*, whereas the codified reactions require *adaptors*, i.e. catalysts plus a code.

Any organic code is a set of rules that establish a correspondence between two independent worlds, and this necessarily requires molecular structures that act like *adaptors*, i.e. that perform two independent recognition processes. The adaptors are required because there is no necessary link between the two worlds, and a set of rules is required in order to guarantee the specificity of the correspondence. The adaptors, in other words, are necessary in all organic codes. They are the molecular *fingerprints* of the codes, and their presence in a biological process is a sure sign that that process is based on a code. The discovery of adaptors has been in fact the objective criterion which has shown that the genetic code is only the first of a large family, and that many other organic codes exists in the living world.

## The signal transduction codes

Cells react to a wide variety of physical and chemical signals from the environment with the expression of specific genes but the external signals (known as *first messengers*) never reach the genes. They are invariably transformed into a different world of internal signals (called *second messengers*) by a process known as *signal transduction* (Sutherland, 1972). The surprising thing is that there are hundreds of first messengers (hormones, growth factors, neurotransmitters, etc.) whereas the known second messengers are only four (cyclic AMP, calcium ions, inositol trisphosphate and diacylglycerol) (Alberts *et al*., 1994).

First and second messengers, in short, belong to two different worlds, and this suggests that signal transduction may be based on organic codes. The proof that this is the case comes, as we have seen, from the presence of adaptors, and the molecules of signal transduction have indeed the characteristics of true adaptors. The transduction system consists of at least three types of molecules: a *receptor* for the first messengers, an *amplifier* for the second messengers and a *mediator* in between (Berridge, 1985). The system performs two independent recognition processes, one for the first and the other for the second messenger, and the two steps are connected by the bridge of the mediator. The connection however could be implemented in countless different ways because any first messenger can be coupled with any second messenger, and this makes it imperative to have a code in order to guarantee biological specificity. In signal transduction, in short, we find all the three characteristics of the codes:

(1) a correspondence between two independent worlds,

(2) a system of adaptors which give meanings to molecular structures,
(3) a collective set of rules that guarantee biological specificity.

The effects that the external signals have on cells, in conclusion, do not depend on the energy or the information that they carry, but only on the *meaning* that cells give them with rules that have been referred to as *signal transduction codes* (Barbieri, 1998; 2003).

## The splicing codes

One of the greatest surprises of molecular biology came from the discovery that the primary transcripts of the genes are often transformed into messenger RNAs by removing some RNA strings (called *introns*) and by joining together the remaining pieces (the *exons*). The result is a true assembly, because exons are assembled into messengers, and we need therefore to find out if it is a *catalyzed* assembly (like transcription) or a *codified* assembly (like translation). In the first case the cutting-and-sealing operations, collectively known as *splicing*, would require only a *catalyst* (comparable to RNA-polymerase), whereas in the second case they would need a catalyst and a set of *adaptors* (comparable to ribosome and tRNAs).

This suggests immediately that splicing is a codified process because it is implemented by structures that are very similar to those of protein synthesis. The splicing systems, known as *spliceosomes*, are huge molecular machines like ribosomes, and employ small molecular structures, known as *snRNAs* or *snurps*, which are very much comparable to tRNAs. The similarity, however, goes much deeper than that, because the snRNAs have properties that fully qualify them as *adaptors*. They bring together, in a single molecule, two independent recognition processes, one for the beginning and one for the end of each intron, thus creating a specific correspondence between the world of the primary transcripts and the world of messengers. The two recognition steps are independent because the first step could be associated with different types of the second one, as demonstrated by the cases of *alternative splicing*. The choice of the beginning and of the end of an intron, furthermore, is the operation that actually defines the introns and gives them a *meaning*. Without a complete set of such operations, primary transcripts could be transformed arbitrarily into messenger RNAs, and there would be no biological specificity whatsoever. In RNA splicing, in short, we find the three basic characteristics of the codes:
(1) a correspondence between two independent worlds,
(2) the presence of molecular adaptors,
(3) a set of rules that guarantee biological specificity.

The processing of RNA transcripts into messengers is truly a codified process based on adaptors, and takes place with rules that have been given the name of *splicing codes* (Barbieri, 1998; 2003).

## A stream of codes

Molecular adaptors has been found in a wide variety of biological processes. The transport of proteins to different cell compartments, for example, is implemented by molecules that function like molecular addresses and perform two independent recognition steps. The crucial point is that there is *no necessary correspondence* between the proteins and their geographical destinations. The molecular signals are purely conventional labels, like the names that we give to streets, to cities, to airports and to holiday resorts. The existence of eukaryotic compartments, in other words, is based on natural conventions which have been called *compartment codes* (Barbieri, 2003).

Another process that is based on organic codes is the mechanism of *dynamic instability* that allows the cytoskeleton to build a potentially unlimited number of cellular structures. The best proof of this enormous versatility is the fact that the cytoskeleton was invented by unicellular eukaryotes but was later exploited by metazoa to build completely new structures such as the axons of neurons, the myofibrils of muscles, the mobile mouths of macrophages, the tentacles of killer lymphocytes and countless other specializations. The versatility of the cytoskeleton is due to the fact that its microtubules can be attached to an endless number of cell structures by a few anchoring molecules. But this is possible only because there is *no necessary relationship* between the common structure of the cytoskeleton and the cellular structures that the cytoskeleton is working on. The anchoring molecules (or accessory proteins) are true *adaptors* that perform two independent recognition processes: microtubules on one side and different cellular structures on the other side. The resulting correspondence is based therefore on arbitrary rules, on true natural conventions that have been referred to as *the cytoskeleton codes* (Barbieri, 2003).

Many other organic codes have been discovered with a wide variety of theoretical and experimental criteria. Among them, the *Sequence Codes* (Trifonov, 1989; 1996; 1999), the *Adhesive Code* (Redies and Takeichi, 1996; Shapiro and Colman, 1999), the *Sugar Code* (Gabius, 2000; Gabius et at., 2002), and the *Histone Code* (Strahl and Allis, 2000; Jenuwein and Allis, 2001; Turner, 2000; 2002; Gamble and Freedman, 2002; Richards and Elgin, 2002). Even if the evaluation criteria and the definitions of code have been somewhat different from case to case, all these findings tell us that the living world is teeming with organic codes and we simply cannot ignore their existence any longer.

## The concept of organic semiosis

The discovery that there are organic codes and cultural codes in the world gives us the problem of understanding the relationship that exists between them, and more in general between life and culture. The cultural codes are the result of "semiosis", and semiosis is any event which is based on "signs". The traditional compact definition of sign is "something that stands for something else", and a general definition of semiosis has been given in the 4-volume treatise *Semiotik/Semiotics* (1997-2003) edited by Roland Posner, Klaus Robering and Thomas Sebeok.

*"We stipulate that the following is a necessary and sufficient condition for something to be a semiosis: A interprets B as representing C. In this relational characterization of semiosis, A is the Interpretant, B is some object, property, relation, event, or state of affairs, and C is the meaning that A assigns to B."* (Posner et al., 1997).

This is a modern formulation of the classical definition of semiosis proposed by Charles Peirce where A is the *Interpretant* that some *Interpreter* uses to relate B, the *Sign* (or *Representamen*), to C, the *Object* (or *Meaning*). According to Peirce, in short, the elementary act of semiosis is a triadic relationship between Sign, Object and Interpretant which is implemented by an Interpreter, as represented in Figure 1.

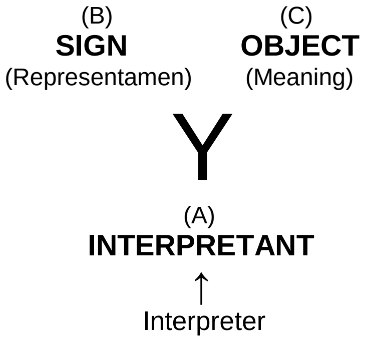

*Figure 1 – According to Charles Peirce, the elementary act of semiosis is a triadic relationship between Interpretant (A), Sign, or Representamen, (B) and Object, or Meaning, (C) implemented by an Interpreter*

The genetic code, as we have seen, is a relationship between codons and amino acids that is implemented with transfer-RNAs by the apparatus of protein synthesis, and can be illustrated by the scheme of Figure 2. More in general, any organic code can be described as a correspondence implemented by a codemaker with adaptors (A) between the objects of two independent worlds (B) and (C). In these cases, the molecules of the first world (B) represent, or "stand for" other molecules, and function therefore as organic signs, whereas the molecules of the second world (C) are their organic meanings. The elementary act of coding, in short, is a triadic relationship between organic Sign, organic Meaning and Adaptor which is implemented by a Codemaker (Figure 3).

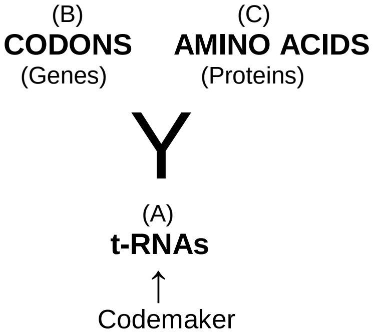

*Figure 2 – The elementary act of coding in the genetic code is a triadic relationship between t-RNAs (A), codons (B) and amino acids (C) implemented by the translation apparatus (Codemaker)*

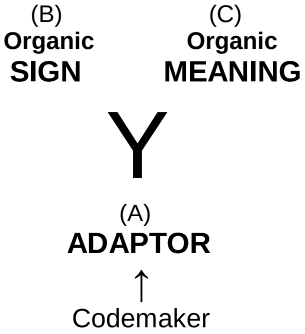

*Figure 3 – The elementary act of coding in any organic code (organic semiosis), is a triadic relationship between Adaptor (A), Organic Sign, (B) and Organic Meaning, (C) implemented by a Codemaker*

The fact that some molecules are used as signs means that a process of organic coding is automatically also a process of *organic semiosis*, and the mere existence of the genetic code implies therefore that organic semiosis is a reality. One could be tempted to conclude that all forms of semiosis are equivalent, and in particular that organic semiosis is equivalent to cultural semiosis, but this is far from being the case. If we compare the scheme of Figure 1 with that of Figure 3, we realize that the terms *Interpretant* and *Interpreter* are replaced by *Adaptor* and *Codemaker*, and this substitution makes an enormous difference. In cultural semiosis the *Interpreter* is always the human mind, and the concepts of Interpretant, Sign and Meaning, are bound to be "subjective" or "mentalistic" entities. In organic semiosis, instead, the agent is a *Codemaker*, a system made of organic molecules (for example the translation apparatus) and the operations that take place are described by "objective" and "reproducible" entities.

The concept of mind, however, is not necessarily restricted to the human mind. Charles Peirce was convinced that there are many types of semiosis in the universe, but in all cases he linked semiosis to interpretation and interpretation to mind, and concluded therefore that mind is everywhere in the Universe. *"Mind has its universal mode of action, namely final causation. The microscopist looks to see whether the motion of a little creature show any purpose. If so, there is a mind there. But the being governed by a purpose or other final cause is the very essence of the psychical phenomenon in general"* (Peirce, 1902).

The Peircean scheme of Figure 1, in conclusion, represents not only cultural, or human, semiosis but any type of *mental semiosis*, i.e. any semiosis that depends upon interpretation and therefore upon some kind of interpreting mind. The scheme of Figure 3 shows that *organic semiosis* is substantially different from all types of mental semiosis precisely because interpretation is not required and there is no need to postulate the existence of a mind.

## The making of Biosemiotics

In the 1960s Thomas Sebeok started a lifelong campaign in favor of the idea that language must have biological roots, and in 1963 he introduced the term *Zoosemiotics* to designate the new science of animal semiotics. According to Sebeok, any semiosis is “defined” by the triadic relationship proposed by Peirce, and interpretation is its *sine qua non* condition. He insisted that *“there can be no semiosis without interpretability”* (Sebeok, 2001) and that interpretations is *“a necessary and sufficient condition for something to be a semiosis”*. This deep link between semiosis and interpretation explains why Sebeok came to the conclusion that Zoosemiotics as a science had been in existence since the 1920s and that its foundations had been laid by Jacob von Uexküll in *Theoretische Biologie* (1928).

Von Uexküll’s main concept was the idea that animals live in a subjective world, or *Umwelt*, and that they are *interpreters* and not just *receivers* of signals. Animals have the ability to turn *signals* into *signs* and are therefore true semiotic agents. This led Sebeok to conclude, in the 1970s, that von Uexküll’s biology was the natural research field of Zoosemiotics. Soon, however, applications were found in many other fields and the experimental horizon became much wider.

In 1981 Martin Krampen showed that plants too engage in semiosis (phytosemiotics), and in 1988 Sorin Sonea proposed that semiosis goes on even in bacteria. In 1977, Giorgio Prodi suggested that there is natural semiosis in any cell, and in 1988 he spoke of *signs and codes in immunology*. The word *Zoosemiotics* became increasingly inadequate and Sebeok decided to replace it officially with *Biosemiotics*, a term proposed by Juri Stepanov in 1971, but which appeared for the first time (with a restricted meaning) in 1961, when Friedrich Rothschild used it to indicate a semiotic approach to psychology.

The making of biosemiotics gathered further momentum in the 1990s, with the joining in of a younger generation and the increasing participation of biologists. Among these, Jesper Hoffmeyer and Claus Emmeche, who founded a biosemiotic Group in Copenhagen, Kalevi Kull, who directed the Jakob von Uexküll Center in Tartu, Estonia, and Anton Markos who promoted a school of hermeneutic biology in Prague. A significant step forward was achieved in 2001, when the first Gathering *exclusively* dedicated to biosemiotics took place in Copenhagen and the official directory listed 61 biosemioticians from all countries of the world. In a relatively short time, biosemiotics became much more than the study of the biological roots of semiosis. It became also the study of *the semiotic roots of life*, and this had far-reaching implications for the whole of biology.

Eventually even the origin of life, and in particular the origin of the genetic code, came under scrutiny, and this brought to the surface all the hidden theoretical implications of the basic concept of semiosis. The process of genetic coding is certainly a process of semiosis, but the model that describes it is not the classical scheme of Peirce based on interpretant and interpreter, as illustrated in Figure 1. It is the scheme of organic semiosis based on adaptors and codemakers, as represented in Figures 2 and 3.

The genetic code, in short, tells us that life is based on organic semiosis not on mental semiosis, and the triadic model of Peirce should be replaced by a more realistic type of trinity. As a matter fact, such a scheme had already been proposed in 1981 with the theory the cell is a trinity of *genotype, phenotype* and *ribotype* (Barbieri, 1981, 1985), because these are precisely the three categories of organic semiosis that are represented in Figures 2 and 3. Up until now, however, biosemiotics continues to be promoted as a science that is largely based on Peirce’s model. The transition from mental to organic semiosis has turned out to be extremely difficult, as any true change of paradigm, and still belongs to the future of biosemiotics (Barbieri, 2002).

## The problem of interpretation

Biologists have quickly accepted the idea of molecular information but not the idea of molecular meaning. Oddly enough, one of the most popular arguments against the existence of meaning at the *molecular* level has come from the study of animal communication, the very science that argued for the existence of language and meaning at the *animal* level. A signal that reaches an animal may set in motion an automatic response, and what we observe in these cases is the unfolding of a deterministic sequence of reactions. In other cases, however, there is an intermediate phase between signal and response, a phase where the signal is *interpreted* and the response changes accordingly. These are the cases where it is said that meaning appears, because interpretation is assumed to be the process that gives a meaning to a signal. Hence the conclusion that *meaning is always the result of an interpretation process*. No interpretation, no meaning (Markos, 2002).

In the case of protein synthesis, we know that codons are translated into amino acids according to the rules of the genetic code, but we also know that these reactions take place in a totally automatic way, with no room for interpretation. Hence the conclusion that there is no addition of meaning here, because protein synthesis is a deterministic chain of biochemical reactions. This argument is still popular and yet it is flawed because the determinism that we observe in protein synthesis is also observed in many cultural processes where we know that meaning does exist. The very act of speaking, for example, is based on the automatic application of prefixed rules. Language itself would not be possible if the meaning of its words had to be negotiated every time they are spoken. Once the basic rules have been fixed in the initial stage of learning, they are no longer changed and the code of a language becomes as deterministic as the genetic code.

The interpretation process that we observe in many animals can easily be understood as an evolution of their signal processing systems. It is likely that the most primitive reactions were heavily determined by genes, but the number of hard-wired responses could not increase indefinitely, and animals became more and more dependent on processes of learning in order to increase their behavioral repertoire. And learning how to respond to a signal means learning *how to interpret* that signal. Rather than memorizing an unlimited number of incoming stimuli, it was far more convenient to learn a few basic rules and let an interpretation phase decide the meaning of countless combinations of signals.

A process of learning, in other words, amounts to the construction of a behavioral code whose rules are *context-dependent*, and therefore *interpretation-dependent*. This gives the *impression* that the generation of meaning comes from the interpretation process, but the truth is rather different. The interpretation process is necessary to make a choice among a plurality of rules, not to create the meaning of those rules. Meaning, in short, is always the result of a code, but sometimes the code is context-dependent, and in these cases it is associated with a process of interpretation. We conclude therefore that the definition of meaning as *an object which is related to another object by a code* is valid in all cases, with or without interpretation.

## Conclusions

One of the most extraordinary results of molecular biology is the discovery that genes and proteins are *molecular artifacts*, i.e., molecules that are manufactured by molecular machines. Genes are manufactured by *copymakers*, machines that make copies of nucleotide sequences. Proteins are made by *codemakers*, machines that translate nucleotide sequences into amino acid sequences according to the rules of a code. When a copymaker is copying, the nucleotide sequence of the template becomes information for it, so it is an act of copying that brings organic information into existence. In a similar way, when a codemaker is assembling a protein, amino acids are associated to codons and become the *meaning* of those codons, so it is an act of coding that brings organic meaning into existence. Organic information and organic meaning, in short, are the inevitable results of the processes of copying and coding that produce genes and proteins.

Modern biology has readily accepted the concept of information but not the concept of meaning, and this cannot be right, because copying and coding, or genes and proteins, are equally important to life. It is a fact however that meaning is still excluded from biology, and this is due to the convergent opposition of two very different ideas.

One is physicalism, the idea that sequences and codes, or information and meaning, are useful metaphors but not fundamental entities because they cannot be expressed in terms of physical quantities. Here it is shown that only spontaneous molecules can be completely described by physical quantities whereas molecular artifacts like genes and proteins necessarily require new fundamental entities. It is shown furthermore that organic information and organic meaning have the same status as physical quantities because they can be defined by *operative definitions*, and are as objective and reproducible as physical quantities.

The second reason that effectively keeps meaning out of biology is the idea that meaning is always the result of a process of interpretation. If this were true, the genetic code would require the existence of a mind at the molecular level, and to most biologists this is unacceptable. Here it is shown that organic meaning is the result of processes of coding that require only adaptors and codemakers, and is completely accounted for by objective and reproducible entities. There is a substantial distinction between organic semiosis and mental semiosis, and that is all we need to realize that organic meaning is an objective natural relity.

We conclude that organic information and organic meaning are both fundamental entities of life and are as real as the processes of copying and coding that bring them into existence. Their real nature has so far eluded us because they are *nominable* entities, i.e. objective, reproducible and fundamental entities of nature that can be described only by naming their components in their natural order. They have also eluded us because nominable entities are necessary only for the production of *artifacts* and we have not yet come to terms with the idea that life is *artifact making*.

## References

Alberts, Bruce; Bray, Dennis; Lewis, Julian; Raff, Martin; Roberts, Keith and Watson, James D. (1994). *Molecular Biology of the Cell.3rd edition*. New York: Garland Publishing.

Barbieri, Marcello (1981). The Ribotype Theory on the Origin of Life. *Journal of Theoretical Biology*, 91, 545-601.

—— (1985). *The Semantic Theory of Evolution*. New York: Harwood Academic Publishers.

—— (1998). The Organic Codes. The basic mechanism of macroevolution. *Rivista di Biologia-Biology Forum*, 91, 481-514.

—— (2002). Has Biosemiotics come of age? *Semiotica*, 139, 1/4, 283-295

—— (2003). *The Organic Codes. An Introduction to Semantic Biology*. Cambridge UK: Cambridge University Press.

—— (2004). The Definitions of Information and Meaning. Two possible boundaries between physics and biology. *Rivista di Biologia-Biology Forum*, 97, 91-110.

Berridge, Michael J. (1985). The molecular basis of communication within the cell. *Scientific American*, 253, 142-152.

Boniolo, Giovanni (2003). Biology without Information. *History and Philosophy of the Life Sciences*, 25, 255-273.

Chargaff, Erwin (1963). *Essays on Nucleic Acids*. Amsterdam: Elsevier.

Gabius, Hans-Joachim (2000). Biological Information Transfer Beyond the Genetic Code: The Sugar Code. *Naturwissenschaften*, 87, 108-121.

Gabius, Hans-Joachim; André Sabine; Kaltner Herbert and Siebert Hans-Christian (2002). The sugar code: functional lectinomics. *Biochimica et Biophysica Acta*, 1572, 165-177.

Gamble, Matthew J. and Freedman Leonard P. (2002). A coactivator code for transcription. *Trends in Biochemical Sciences*, 27 (4), 165-167.

Griffith, Paul E. (2001). Genetic Information: A Metaphor in Search of a Theory. *Philosophy of Science*, 68, 394-412.

Griffith, Paul E. and Knight, Robin D. (1998). What is the developmental challenge? *Philosophy of Science*, 65, 276-288.

Jenuwein, Thomas and Allis David (2001). Translating the Histone Code. *Science*, 293, 1074-1080.

Johannsen, Wilhelm (1909). *Elemente der exacten* Erblichkeitslehre. Jena: Gustav Fischer.

Krampen, Martin (1981). Phytosemiotics. *Semiotica*, 36 (3/4), 187-209.

Mahner, Martin and Bunge, Mario (1997). *Foundations of Biophilosophy*. Berlin: Springer Verlag.

Markos, Anton (2002). *Readers of the Book of Life*. *Contextualizing Developmental Evolutionary Biology*. New York: Oxford University Press.

Peirce, Charles S. (1931-1935) *Collected papers of C. S. Peirce*. Cambridge Massachusetts: Harvard University Press, 1269.

Posner, Roland; Robering Klaus and Sebeok Thomas (1997). *Semiotik/Semiotics: A Handbook on the Sign-Theoretical Foundations of Nature and Culture.* Volume 1. Berlin: Walter de Gruyter, 4.

Prodi, Giorgio (1977). *Le basi materiali della significazione*. Milano: Bompiani.

—— (1988). Material bases of signification. *Semiotica*, 69 (3/4), 191-241.

Readies, Christoph and Takeichi, Masatoshi (1996). Cadherine in the developing central nervous system: an adhesive code for segmental and functional subdivisions. *Developmental Biology*, 180, 413-423.

Richards, Eric J. and Elgin Sarah C.R. (2002). Epigenetic Codes for Heterochromatin Formation and Silencing: Rounding up the Usual Suspects. *Cell*. 108, 489-500.

Rothschild, Friedrich S. (1961). The Psycology of the Self. Conference held at the New York Academy of Sciences. —— (1962) Laws of symbolic mediation in the dynamics of self and personality. *Annals of the New York Academy of Sciences*, 96, 774-784.

Sarkar, Sahotra (1996). Biological Information. A Skeptical Look at some Central Dogmas of Molecular Biology. In S. Sarkar (ed.) *The Philosophy and History of Biology*. Dordrecht : Kluwer Academic Publishers, 187-231.

—— (2000). Information in Genetics and Developmental Biology. *Philosophy of Science*, 67, 208-213.

Sebeok, Thomas A. (1963). Communication among social bees; porpoises and sonar; man and dolphin. *Language*, 39, 448-466.

—— (2001). Biosemiotics: Its roots, proliferation, and prospects. *Semiotica*, 134, ¼, 68.

Shannon, Claude (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, 27, 379-424, 623-656.

Shapiro Lawrence and Colman David R. (1999). The Diversity of Cadherins and Implications for a Synaptic Adhesive Code in the CNS. *Neuron*, 23, 427-430.

Sonea, Sorin (1988). The global organism: A new view of bacteria. *The Sciences*, 28 (4), 38-45.

Stepanov, Juri S. (1971). *Semiotika*. Moscow: Nauka.

Strahl Brian D. and Allis David (2000). The language of covalent histone modifications. *Nature*, 403, 41-45.

Sutherland Earl (1972). Studies on the mechanism of hormone action. *Science*, 177,401-408.

Trifonov, Edward N. (1989). The multiple codes of nucleotide sequences. *Bulletin of Mathematical Biology*, 51, 417-432.

—— (1996). Interfering contexts of regulatory sequence elements. *Cabios*, 12, 423-429.

—— (1999). Elucidating Sequence Codes: Three Codes for Evolution. *Annals of the New York Academy of Sciences*, 870, 330-338.

Turner, Brian M. (2000). Histone acetylation and an epigenetic code. *BioEssay*, 22, 836-845.

—— (2002). Cellular memory and the Histone Code. *Cell*, 111, 285-291.

Uexküll, Jakob von (1928). *Theoretische Biologie*. Berlin: Julius Springer.