Information in the Biological Sciences

Alfredo Marcos* and Robert Arp**

*Department of Philosophy, University of Valladolid, email: amarcos@fyl.uva.es

**Independent Scholar, email: robertarp320@gmail.com

Abstract Information has been a central concept for contemporary work in the biological sciences (and other sciences) especially after the publication of Claude Shannon and Warren Weaver's, *The Mathematical Theory of Communication*, in 1949. In fact, the pervasiveness of Shannon's information theory—as well as of the very terms themselves—becomes evident when one takes a moment to reflect upon just a few of the concepts that are standard in the biomedical sciences, such as genetic *code*, *messenger* RNA, ion *channel*, cell *signal*ing, intracellular *communication*, *signal* transduction, pathogen *transmission*, positive *feedback* loop, expressive *noise* minimization, and many others. In this chapter we first give a historical introduction concerning the concept and nature of information, with a special emphasis upon the biological sciences. Then, we provide a few important examples of information at work in the biological sciences. Next, we consider the debate regarding the reality and nature of bioinformation, arguing that bioinformation is best understood as a relationship between and/or among entities; for instance, DNA is informational only in relation to a given cellular context, and it is misguiding to locate information in a particular molecule. We then go on to show how bioinformation relates to other concepts such as entropy, order, organization, complexity, and knowledge. Finally, we approach education itself as an informational process in order to draw some consequences for the teaching of biology.

1 Introduction

> Life, too, is shaped by information. All living creatures are information-processing machines at some level…
>
> Charles Seife, *Decoding the Universe*

Why does information matter in the teaching of biology? How can the biology educator benefit from the philosophy of biology regarding information? These are the two basic questions that we explore in this chapter.

1.1 Information is Pervasive in Biology

Concerning the first question above, the concept of information is important in the teaching of biology simply because it is integral to the biological sciences themselves. Charles Seife (2007) is correct above in noting that “all living creatures are information-processing machines at some level,” and information has been a central concept for contemporary work in the biological sciences (and other sciences), especially since the publication of Claude Shannon and Warren Weaver’s, *The Mathematical Theory of Communication*, in 1949 and the discovery of the genetic code around the middle of the 20th century by Marshall W. Nirenberg and coworkers, who won the Nobel Prize in 1968 “for their interpretation of the genetic code and its function in protein synthesis” (NPO, 2012). In fact, the pervasiveness of Shannon’s information theory—as well as the very terms themselves—becomes evident when one takes a moment to reflect upon just a few of the concepts that are standard in the biomedical sciences, such as genetic *code*, *messenger* RNA, ion *channel*, cell *signaling*, intracellular *communication*, *signal transduction*, pathogen *transmission*, positive *feedback* loop, expressive *noise* minimization, and many others. Biology has developed what we might call an *informational paradigm*. This is a fact.

One may take a positive or a negative view regarding this fact, or even remain indifferent; indeed, all these positions are present in contemporary literature. But this leaves the fact unchanged. And this fact is important, for teaching biology cannot be achieved without a reflective and critical understanding of informational concepts.

There is, moreover, another reason why the concept of information should interest *every* teacher: the educational process itself may be considered an informational relationship existing between and among multiple minds engaged in communicating, processing, and learning.

1.2 Philosophy of Biology and Information

These considerations lead us to the second question posed above: How can the biology educator benefit from the philosophy of biology regarding information? Stated in another way: What does the philosophy of biology contribute regarding the concept of information and its relationship to the biological sciences?

The word *philosophy* comes from two Greek words: *philos* deriving from *philein*, “love,” and *sophos* meaning “wisdom.” Love here means something like an intense desire for something, while wisdom is arguably a kind of knowledge gained from experience, whether this is practical experience (gained from living life with all of its ups and downs) or theoretical experience (gained from understanding, evaluating, critiquing, and synthesizing ideas, positions, and concepts).

Ever the theoretician, the philosopher has always been the person who not only desires to look deeper into some claim, idea, argument, event, or state of affairs by questioning assumptions and challenging status quo thinking, but also attempts to explain and systematize aspects of reality as it is perceived. In Bertrand Russell's (1912/1999) words, which are appropriate given the nature of this book, "Philosophy, like all other studies, aims primarily at knowledge. The knowledge it aims at is the kind of knowledge which gives unity and system to the body of the sciences, and the kind which results from a critical examination of the grounds of our convictions, prejudices, and beliefs" (p. 9).

The word *biology* comes from two Greek words as well: *bios* meaning, "life" and *logos* meaning, "word" or "rational account." Thus, biology is the kind or type of rational account (or science) that studies life, which most of us already know. Whereas *biology* can be characterized as a set of sub-disciplines (the biological or life sciences) under science, the concern of which includes the description, classification, analysis, explanation, prediction, and ultimately control of living things, *philosophy of biology* can be characterized as a sub-discipline of philosophy, the concern of which is the meta-leveled attempt on the part of philosophers, biologists, and other thinkers to understand, evaluate, and critique the methods, foundations, history, and logical structure of biology in relation to other sciences, disciplines, and life endeavors so as to better clarify the nature and purpose of biological science and its practices (see Ayala & Arp, 2009; Rosenberg & Arp, 2009; Rosenberg & McShea, 2007; Ruse, 2008; Sober, 1993).

Now, the epistemological, computational, linguistic, and logical aspects of information have been dealt with extensively in the philosophical tradition. When the use of informational concepts was extended to biology, philosophers immediately began to react, reflect, ruminate, and even ridicule, so we can expect major contributions from the philosophy of biology.

Specifically, we expect this discipline to help us understand the meaning of the different versions of the concept of information—especially bioinformation—from historical as well as from contemporary perspectives. Philosophy of biology also contributes to clarifying the scope of the use of informational terms in biology, that is, whether they are used metaphorically, in a linguistically instrumental way, or in such a way as to capture the real, objective aspects of living things. If the philosophy of biology can offer no definitive answer to this issue of scope, it can at least make us aware of the problems and ensure that they are clearly posited.

Furthermore, the philosophy of biology also helps the educator understand the complex relationships existing between different concepts that have a great presence in the biological literature. We refer here to the concept of information itself and to others such as *form*, *correlation*, *order*, *organization*, *complexity*, *meaning*, *knowledge*, and *entropy*, to name just a few. The concept of entropy is now used standardly when discussing protein synthesis in cellular functions, for example, and since this kind of entropy—known as *Shannon entropy*—quantifies the expected value of the genetic information contained in the messages delivered between and among various mRNA molecules so that protein synthesis may occur, we can see how a clear understanding of the concepts of information and entropy, as well as their relationship to one another, is crucial for the biology educator if a robust explanation of protein synthesis is to be put forward (Ewens, 2010; Collier, 2003; Brooks & Wiley, 1988; Weber, Depew & Smith, 1988; Wicken, 1987).

Philosophers of biology also make contributions to the problem of the *location* of information. For example, we often wonder where hereditary information is to be found. Seemingly in the genes and the configuration of codons and switches (see Burian this volume); but there is no doubt that the epigenetic level is also important for the development of the organism (see Uller this volume), as is the cellular cytoplasm, the configuration of tissues, the organism itself as a whole and, in general, the environment. We sometimes speak of information as if it resided exclusively in the genes, but on other occasions we speak of it as if it were present everywhere. In short, all these questions come into play in the teaching of biology, and in all of them, the philosophy of biology can be of help, as we shall see.

Finally, the biologist—*qua* educator—may be interested in the informational aspects of the educational process itself. For this topic, valuable contributions may also be expected from the field of philosophy, especially the philosophy of education, as well as from communication theory, linguistics, psychology, sociology, anthropology, and other related disciplines.

1.3 Outline of the Chapter

The underlying viewpoint in this chapter is that teaching biology cannot be achieved without a reflective and critical understanding of informational concepts. So we begin in Section 2 by looking at the different ways in which the concept of information was historically understood up to the 1950s, when it began to make its presence felt in the life sciences. Next, in Section 3, we examine the influence that the informational paradigm has had in the different areas of life sciences, such as genetics, cell biology, neurobiology, and ecological studies. Here, we provide several standard examples of information processing in living systems.

Once the apparent pervasiveness of informational terms has been demonstrated through examples from different areas of the life sciences, in Section 4 we then examine some of the debates to which this pervasiveness has given rise. In the first place, there is an argument on the advisability of using informational concepts in biology, with some researchers maintaining that informational jargon should be kept out of the life sciences, while others argue that the informational perspective is indispensable for understanding biological phenomena. Second, those authors who accept the informational perspective as legitimate continue to debate about its possible interpretation: for some, informational concepts must be taken as metaphors in biology; for others, they have a merely instrumental use; while still others consider information to be a real and substantial aspect of living things. Finally, there is an argument concerning the very nature of bioinformation, which may be considered as a thing, a property, or a relationship. We think that bioinformation is best understood as a relationship between and/or among entities; for instance, DNA is informational only in relation to a given cellular context, and it is misguiding to locate information in a particular molecule.

In Section 5, we turn attention to the relationship between the concept of information and related concepts that are integral to the life sciences, such as entropy, organization, complexity, and knowledge, as well as the problem of the location of information in living systems. In Section 6, we offer some final thoughts concerning the philosophy of education in light of information existing as a relational and informing phenomenon. Our hope is that the information we provide *about* information in this chapter will be helpful for biology educators.

2 From Information to Bioinformation: A Historical Overview

The English word *information* derives from the Latin noun *informatio*, which can mean, “representation,” “idea,” or “explanation.” Also, the Latin verb *informo* can mean, “to sketch,” “to draw,” or “to represent” something as well as “to give shape or form” to something. In ancient times, the term was used in both everyday and learned discourse, as for instance, in the works of Virgil, Cicero, Tertullian, and Augustine of Hippo (Capurro, 1978; Floridi, 2003, 2011). It was used in different domains: ontological (“to shape something”), epistemological (“to become acquainted through the sensorial or intellectual reception of a form”), pedagogical, and moral (“to instruct,” “to form”). But it was not the object of any special philosophical reflection.

During the Middle Ages, the verb *informo* and its derivatives were incorporated into philosophical language from Scholastic discourse. Throughout this period, the verb retained its ontological, epistemological, didactic, and moral connotations as well as its active sense, whereby *informatio* was an action rather than a thing. It referred to the action of shaping and its result. Interestingly enough, the great medieval philosopher and theologian, Thomas Aquinas (1225-1274 CE), used information to refer to the act of shaping/forming and its result when he defined *per modum informationis*, a natural biological process whereby a living thing begins to exist. Also in connection with biological domains, we can refer to Marcus Terentius Varro (116-27 BCE), who describes the development of a fetus as a process of information, whereby it is “shaped” or “informed” (*informatur*) (Capurro & Hjørland, 2003).

During the 14th and 15th centuries, the use of the word *information* spread into European languages from French. At that point, *investigation*, *education*, and *intelligence* were added to its traditional meanings. However, and perhaps because of the rejection of Scholastic terminology, from then on *information* ceased to be a philosophical term, and others, such as *impression*, *idea*, and *representation* came into play, especially when discussing mental forms of information. Descartes, Locke, Hume, Bacon, Kant, and other modern philosophers did not think of their philosophy in terms of information, and in the few places where we find a word that derives from the term *information*, it came to be understood as an idea or a representation inside the subject's mind. In his dialogue, *Alciphron* (1732/1901) George Berkeley (1685-1753) has Euphranor claim "I love information upon all subjects that come in my way, and especially upon those that are most important" (Dialogue 1, Section 5). Modern *idea-ism*—that is, the preference for philosophizing about ideas rather than things—is clearly related to this change from the view of information as an action to an idea (Collins, 1956; Musgrave, 1993). Interestingly enough, Thomas Reid (1710-1796), one of the authors who most bitterly criticized the modern *idea-ism*—the "theory of ideas" in his own terms—was also one of the few who used the term *information* profusely. In Reid's *Inquiry into the Human Mind on the Principles of Common Sense* (1764), the term appears no less than fifteen times in different contexts and with varied meanings, frequently in connection with the term *knowledge*, and even sometimes in connection with terms such as *input*, *artificial language*, *sign*, *receive*, *perception*, and *channel* (Reid, 1764/2001, pp. 48, 53, 61, 64, 103, 117-123).

It was during the 19th century that the term *information* grew in importance, to the point of acquiring a crucial place in contemporary culture. It was bound up with the expansion of communication technologies, such as the telegraph, and with the use given to it in military intelligence service (Adriaans & van Benthem, 2007; also the papers in Davies & Gregersen, 2010). Thus, information acquired a great economic and political value. A 1902 issue of *The Economist*, for example, notes that the telegraph has "taken the place of the Ambassador" whose "business [...] undoubtedly is to collect information" (*The Economist*, 1902, p. 1881).

Since then, mathematical theories of communication have been developed that seek to facilitate the transmission of the greatest amount of information at the lowest possible cost, in the shortest possible time, and with the maximum security. After World War II, these developments accelerated thanks to the progress of information technology. The linking of communication and computation, and the growth of their social presence, has done the rest. As a result, the term *information* currently occupies a central place in ordinary speech and in almost all sciences and disciplines, from communications to computer science, statistics to systems theory, and criminology to cytology (Miller, 2005; Seife, 2007; Floridi, 2003, 2007, 2011; Gleick, 2011).

2.1 The Shannon Model of Information

The unavoidable locus for the theory of information is the classical work by Claude E. Shannon and Warren Weaver (1949). However, the term *information* does not even appear in its title, *The Mathematical Theory of Communication*. The expression, *theory of information*, probably comes from an article by Ralph Hartley (1928) entitled, “Theory of Information Transmission.” Although Shannon focuses attention on communication, we should understand that his theory deals specifically with the communication *of information*. The explanation of this concept given by Warren Weaver is still very useful:

> Information must not be confused with meaning… To be sure, this word *information* in communication theory relates not so much to what you *do* say, as to what you *could* say. That is, information is a measure of one’s freedom of choice when one selects a message. If one is confronted with a very elementary situation where he has to choose one of two alternative messages, then it is arbitrarily said that the information, associated with this situation, is unity […] The amount of information is defined, in the simplest cases, to be measured with the logarithm of the number of available choices. (Shannon & Weaver, 1949, pp. 8-9)

More specifically, the transmission of information concerns the reduction of statistical uncertainty in the communication between transmitter and receiver (Cover & Thomas, 2006; Yeung, 2006; West & Turner, 2006). In this way, the information of a message is measured by a probabilistic function, $I(m_i) = -\log P(m_i)$, where $I(m_i)$ is the information attributed to a message m_i. In consequence, the amount of information generated by a source of messages is measured by this formula: $H(M) = -\Sigma_i P(m_i) \bullet \log P(m_i)$. This magnitude is also called the *entropy* of a source. The name “entropy” was chosen by Shannon in attention to the formal similarity between this formula and Boltzmann’s formula for thermodynamic entropy. We shall return below to this point and its conceptual implications. Another way to think about this is that a message is informative insofar as it reduces the receiver’s uncertainty about some state of affairs and communicates something new to the receiver. So, if someone learning English for the first time did not know that the word *psychology* begins with an “s” sound, rather than a “p” sound, then conveying that message to that person would be informative.

Shannon identifies the elements that comprise the communication of information processes. He represents them by means of the diagram in figure 1, which is our rendition of it (Shannon & Weaver, 1949, p. 34). Shannon’s objective was to apply his theory to technical systems of communication, such as a telephone or a telegraph system. For this reason, his diagram includes a transmitter and a receiver. The function of the transmitter is to transform the original message—or instance, a sequence of letters—into a signal suitable for transmission over the channel. Shannon defines a channel as a “pair of wires, a coaxial cable, a band of radio frequencies, a beam of light, etc” (p. 34). For its part, the receiver performs the inverse operation of that which is performed by the transmitter. But, we could devise diagrams with more boxes, depending on the nature of the problems to which we are applying the theory (see, for example, Moles, 1972). In Shannon’s diagram, the functions of encoding and decoding the message are performed by the transmitter and the receiver, respectively, but we could design new boxes for an encoder and a decoder.

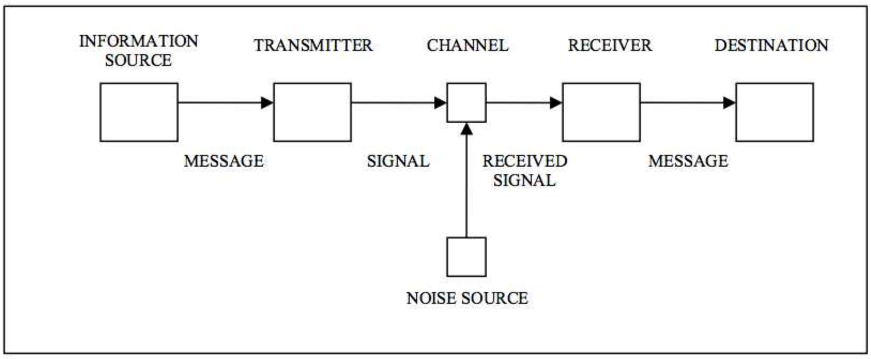

Figure 1: A Rendition of Shannon’s Diagram

It is possible to construct simpler diagrams with no more than three elements: a source or emitter, a channel, and a receiver. And we can even adopt an abstract interpretation of Shannon’s theory free from spatiotemporal connotations. In this regard, Abramson (1963) interprets an information channel as a simple mathematical relationship between the probabilities of two sets of symbols. A channel of information consists only of an incoming alphabet, an outgoing alphabet, and a set of conditional probabilities. For instance, $P(b_j|a_i)$ is the probability of receiving the symbol b_j, if a_i were emitted. Here, a source of information is no longer imagined as a dimensional *box*. It is an abstract entity comprising a set of symbols and their corresponding probabilities (Cover & Thomas, 2006; Yeung, 2006; West & Turner, 2006).

2.2 Problems to Understand and Overcome

As Shannon himself warns, there are more problems regarding the concept of information than those that his theory deals with. In order to organize the many informational problems, we can follow the threefold classification suggested by Weaver (Shannon & Weaver, 1949, p. 31).

First, there are *technical problems* concerning the maximum amount of information a message can convey. These concern the statistical regularities of the source, such as the internal structure and constraints of the messages, together with the conditions of noise and equivocation of the channel itself. Given these conditions, we ask: “What is the best possible configuration of the message?” That is, which configuration optimizes the balance between length and reliability of the message. Thus, we have problems at a *syntactic* level, of the type dealt with by Claude Shannon’s mathematical theory of communication. Let us add that the measure of complexity proposed by Andrey Kolmogorov (1903-1987)—namely, the measure of the computational resources necessary to specify an object, or what has come to be known as the *Kolmogorov complexity*—remains also at the syntactic level (Kolmogorov, 1965; Solomonoff, 2003; Li & Vitányi, 1997; Grünwald & Vitányi, 2003).

Second, there exist *semantic problems* that concern the meaning and theoretical truth of the messages, and the correlation between the message and some other thing. Weaver makes it clear that Shannon's theory does not seek to explain problems at this level or at the next one. In the last few decades, several theories have appeared that do deal with semantic aspects of information (Barwise & Seligman, 1997).

Finally, there are *pragmatic problems* concerning the efficiency of the message to modify the receiver's behavior. Weaver says that, "the *effectiveness problems* are concerned with the success with which the meaning conveyed to the receiver leads to the desired conduct on his part" (p. 5). In biological terms, we find here the functional aspects of information, its ability to affect the receiver's behavior in a functional or adaptive sense.

More recently, Luciano Floridi (2007, 2011) distinguishes between information *as* reality, information *about* reality, and information *for* reality, and it is tempting to correlate these categories with Weaver's levels. On the syntactic level, what we study is information as reality, that is, the properties of the message itself. On the semantic level, we deal with information about reality, or what a message tells us about another part of reality. On the pragmatic level, we observe the capacity of a message to alter reality. This is like saying that we observe the message as information for (making or modifying) reality. A variety of approaches have arisen to address the syntactic, semantic, and pragmatic levels of information (Shannon, 1993; Landauer, 1996; Winder, 2012). However, our main interest here is bioinformation, and so our concern is mainly with pragmatic or functional problems.

3 The Many Faces of Bioinformation

One of the earliest links between information and biology in the 20th century occurs in August Weismann's 1904 book, *The Evolution Theory* (Weismann, 1904; Maynard Smith, 2000). In an important paper over one hundred years later, Artmann (2008) affirms the central role of bioinformation, and his ideas are worth quoting at length:

> The most remarkable property of living systems is their enormous degree of functional organization. Since the middle of the twentieth century, scientists and philosophers who study living complexity have introduced a new concept in the service of explaining biological functionality: the concept of information […] Let us adduce some of the highly controversial theses that the proponents of biological information theory claim to be true: In molecular *genetics*, a set of rules for transmitting the instructions for the development of any organism has been discovered that is most appropriately described as a genetic code. The main research problem of *developmental biology* is how the decoding of these ontogenetic instructions depends upon changing biochemical contexts. *Neurobiology*

> cannot make decisive progress before neural codes that are needed for storing, activating, and processing simple features of complex cognitive representations are discovered. *Ethology* is a science of communication since it studies the astonishing variety of information-bearing signals, whose transmission can be observed, for example, in social insects, birds, and primates. Information-theoretical considerations are also of great importance to *evolutionary biology*: macroevolutionary transitions—from co-operative self-replication of macromolecules, to sexual reproduction, to human language—established more and more complex forms of natural information processing. If all these claims prove true, the following answer must be given to the old problem of defining life: *life is matter plus information*. (pp. 22-23, italics added)

Consistent with Artmann's claims, since the 1950s the notion of information has become increasingly important in most fields of biology (see Paton, 1992). It has even been used to define life itself (see Tipler, 1995, pp. 124-127; Küppers, 1990, 2000, p. 40; Ruiz-Mirazo & Moreno, 2011). The biological sciences have adopted a theoretical stance derived from information theory. This perspective holds that all biological processes involve the transfer, processing, or storage of information, and has been referred to as *bioinformational equivalence* in a famous paper by C.I.J.M. Stuart (1985; Burian & Grene, 1992, p. 6).

A glance at the current bibliography will suffice to show that, since Stuart's 1985 paper, the use of the concept of information in biology has become widespread (for a historical perspective, see Kull, 1999; also Queiroz, Emmeche, & El-Hani, 2007; Jablonka, 2002; Artmann, 2008; Collier, 2007). In molecular biology, biomolecules are considered to contain information and are the result of informational processes (Holzmüller, 1984). In genetics especially, biological thinking is shaped by the idea of information transfer (Brandt, 2005; Kjosavik, 2007), while in developmental biology and aging, much is said about the expression of information and phenotypic information (Waddington, 1968; Oyama, 2000; Atlan, 1972, p. 96; Peil, 1986). In cell biology, tissue biology, zoology, and botany, we study different ways of communicating information with chemical, neuronal, or linguistic bases (Albrecht-Buehler, 1990; Marijuan, 1991; Stegmann, 2005; Pfeifer, 2006). In ecology, the concepts of complexity and biodiversity are closely bound up with information through notions of entropy and order (Margalef, 1968).

In neurophysiology and endocrinology, the study of communication, storage, and processing of information is central, as are the various electric and chemical codes (Baddeley, Hancock, & Földiák, 2000). The immune system is also researched in terms of knowledge understood as information, both acquired and accumulated (Forrest & Hofmeyr, 2000). Evolution, from the origin of life onward, is thought of as the accumulation of information in macromolecules (Elsasser, 1975; Küppers, 1990; MacLaurin, 1998; Moreno & Ruiz-Mirazo, 2002; for information and the origin of life, see Yockey, 1977, 1981, 2005). The latest research into the human genome, and the genomes of other organisms, has required the application of powerful methods of computation, classification, and querying of data and information, and this coming together of disciplines has given rise to what is known as *bioinformatics* (see Arp, Smith, & Spear, in preparation; Nishikawa, 2002).

The concept of information, however, is also central to epistemology and the cognitive sciences and, as several research programmes are attempting to link the cognitive phenomenon with its biological basis, it would be desirable to have one general concept of information that could be applicable to both cognitive and biological contexts. Examples of such programs include evolutionary epistemology along the lines of Lorentz and Wuketits (1983) or Popper (1990), Piagetian epistemology (Piaget, 1970), psychobiology (Bond & Siddle, 1989), evolutionary psychology (Horan, 1992), cognitive ethology (Allen, 1992), neural Darwinism (Edelman, 1987) and, in general, a widespread current tendency to naturalize epistemology (Giere, 1988). An analogy could be drawn between the programs of artificial life, computational science, and the social sciences, where the confluence with biology is evident and the need for a common concept of information is urgent.

Below are a few more-detailed examples of information at work in the biological sciences at various levels. As we hope to demonstrate, many basic life processes—from the molecular foundations of inheritance to the behavior of higher organisms in relation to their environments—are self-organizing processes of storing, replicating, varying, transmitting, receiving, and interpreting information.

3.1 Genetic Information

In general, biologists and other researchers who describe biological phenomena are aligned with Mayr (1996) in his description of organisms as “hierarchically organized systems, operating on the basis of historically acquired programs of information” (Yockey, 2005; Terzis & Arp, 2011; Gould, 2002; Bogdan, 1994; Boi, 2011).

The “programs of information” part of Mayr’s description of organisms above is what is significant for us here. But what does this mean? As most people know, a *gene* is a functional segment of deoxyribonucleic acid (DNA) located at a particular site on a chromosome in the nucleus of all cells. DNA and ribonucleic acid (RNA) are composed of nucleotides that specify the amino-acid sequences of all the proteins needed to make up the physical characteristics of an organism, much like a cryptogram or code. This genetic code consists of specific sequences of nucleotides that are composed of a sugar (deoxyribose in DNA, ribose in RNA), a phosphate group, and one of four different nitrogen-containing bases, namely, adenine, guanine, cytosine, and thymine in DNA (uracil replaces thymine in RNA). These four bases are like a four-letter alphabet, and triplets of bases form three-letter words or *codons* that comprise the “program of information” which identifies an amino acid or signals a function.

DNA is the template from which RNA copies are made that transmits genetic information concerning an organism’s physical and behavioral traits (phenotypic traits) to synthesis sites in the cytoplasm of the cell. mRNA takes this information to ribosomes in a cell where amino acids, and then proteins, are formed according to that information. The proteins are the so-called *building blocks* of life, since they ultimately determine the physical characteristics of organisms (Boi, 2011; Carroll, 2005).

Two significant processes utilized by researchers that have contributed to, and continue to contribute to, our understanding of the genetic code are *genetic sequencing* and *genetic annotation*. Genetic sequencing refers to the methods and technologies used since the early 1970s to determine the specific order of the bases in a molecule of RNA (adenine, guanine, cytosine, and uracil) or DNA (adenine, guanine, cytosine, and thymine). Walter Fiers and colleagues (1976) published ground-breaking work in RNA sequencing in their *Nature* paper titled, “Complete Nucleotide Sequence of Bacteriophage MS2 RNA: Primary and Secondary Structure of the Replicase Gene.” A sequencing-by-separation technique was developed by Frederick Sanger and Alan Coulson (1975) for DNA in 1975, and this “plus and minus” method still acts as the basis for a lot of gene sequencing performed today. Various genetic sequencing methods have been utilized for RNA and DNA since the 1970s, including what is known as *high-throughput sequencing* that can produce millions of sequences at once (Shendure, Mitra, Varma, & Church, 2004).

Understanding the particular configurations of As, Gs, Cs, and Ts (or Us) in the genetic code is one thing; understanding what processes are initiated, amino acids are identified, or functions are signaled by virtue of these particular configurations is another. Genome annotation—*annotation* read here as “commentary” or “explanation”—refers to the methods and technologies used to identify the locations of genes (as well as the coding regions in a genome) and determine specifically what those genes do. “What are all these genes doing, how do their functions interact, and how may we take advantage of the sequences to advance understanding and cure human disease” (FlyBase, 2001). This is the question posed at the beginning of one of the earliest white papers produced by members of FlyBase, a consortium of researchers devoted to annotating the genetic makeup of *Drosophila melanogaster*, a fruit fly. In fact, many organ systems in mammals have well-conserved homologues in *drosophila*, and this species of fruit fly not only was utilized by Thomas Hunt Morgan and his researchers in the early 1900s—in the now famously dubbed “Fly Room”—so as to understand genetic functioning generally (Morgan, Sturtevant, Muller, & Bridges, 1915), but it was also utilized by various groups attempting to annotate the human genome through the Human Genome Project, which was completed in 2003 (HGPI, 2012). It is estimated that some 66% of human disease genes having a clear cognate in *drosophila* (Stein, 2001; Reiter, Potocki, Chien, Gribskov, & Bier, 2001; Tweedle et al., 2009).

3.2 ATP, Euglenas, and Information

Cells use energy, and one of the primary functions of the mitochondrion of an animal cell is by using the energy released during the oxidation of sugars to produce a nucleic acid called adenosine triphosphate (ATP). However, this can happen only if there is a line of communication between other organelles of the cell and the mitochondria themselves. ATP acts as the material catalyst of information communicated between mitochondrion and other organelles. When there are low levels of ATP, the mitochondria receive this information and oxidize more sugars; conversely, when sugars are oxidized (this activity, among other activities), the other organelles receive this information and cellular homeostasis can be maintained.

Euglena gracilis is an abundant one-celled microorganism that is a member of the protist kingdom found in freshwater environments; in colloquial terms, it is known as a kind of algae. It is about 10 micrometers in length and looks like a sperm cell with a more elongated body. It is equipped with a flagellum, eyespot, vacuoles, chloroplasts, mitochondria, plastids, and a cell nucleus. Each one of these components has a function: the flagellum is a whip-like tail that enables the euglena to move around; the eyespot is light/dark sensitive so that the euglena can move toward sunlight, its food source; vacuoles allow for wastes to be disposed; chloroplasts and mitochondria work together to transform sunlight energy to food through ATP; plastids store the food; the cell nucleus contains a nucleolus that synthesizes and encodes ribosomal RNA, which is important for euglena structure and reproduction (Buetow, 1982).

For an organism like the euglena to function effectively in some external environment—basically, live its life in its microbial world—it is necessary that information be exchanged between and among the various subsystems of this system. Food storage in the euglena can be viewed as a subsystem activity, which itself is made up of processes dealing with electron transport and oxygen exchange in photosynthesis. Concerning these processes in the euglena, electrons are transferred from a donor molecule (such as nicotinamide adenine dinucleotide) to an acceptor molecule (such as O_2) across a membrane, with the resulting H^+ ions used to generate energy in the form of ATP. The information must be exchanged in these processes; otherwise, there would be no storage of food. At the same time, this subsystem works with the subsystems concerning food acquisition and mobility. If information were not being exchanged between the eyespot and the flagellum, then there would be no movement toward sunlight; in turn, there would be no photosynthesis, and then no food storage.

3.3 Action Potentials, Reflex Arcs, and Information

When a neuron produces an action potential (colloquially, when it *fires*), information associated with spiking signals is communicated between that neuron and at least one other neuron. In the language utilized by Shannon (see figure 1 above), the axon of one neuron A acts as a *transmitter* and the dendrites of another neuron B, to which the axon of neuron A is connected, acts as a *receiver*. Protein synthesis in neurotransmitter release is the information that is communicated between neurons. Depending on the amount and intensity of the neurotransmitter emitted from the transmitter neuron, the receiver neuron may become excitatory, making it more likely to produce its own action potential. Networks of neurons can fire more quickly when they are used more frequently, as if the information associated with the particular network's firing has been stored. The complex workings of trillions of these connections throughout an animal with a complex nervous system enable it to fight, flee, forage, feast, and the like (Kandel, Schwartz, & Jessell, 2000).

A clear illustration of the communication of neuronal information in a systemic fashion is a mammal's muscle coordination in a *reflex arc*. In this activity, information is communicated to and from the spinal cord and a particular muscle group of the body (Kandel et al., 2000; Pelligrino, Fadiga, Fogassi, Galleste, & Rizzolatti, 1996). Consider a situation where a very curious cat decides to jump atop a very hot stove. The intense motion of the molecules from the stovetop is impressed upon the pads of the cat's paws. That motion affects the sensory neurons in the cat's skin, causing them to fire. The sensory neurons send a message to the spinal cord. These *messages* consist of billions of action potentials and neurotransmitter releases, affecting cell after cell that is along the pathway of this particular reflex arc. In an instant, the spinal cord then sends a message back to the muscle groups associated with the cat's legs, diaphragm, and back. In a flash, the cat jumps off the stove, screaming while arching its back.

However, now the cat must coordinate its fall to the ground. This time, information is sent from the visual system to the brain, and then back through the spinal cord to other muscles in the cat's body. All of this information must be integrated by the brain, and motor responses must be orchestrated by the combined effort of brain-body communication of information. The cat narrowly avoids falling into the garbage can placed next to the stove.

3.4 Visual Perception and Information

In their textbook devoted to the principles of neuroscience, Keith Kandel et al. (2000) describe the processes associated with perception in the cerebral cortex using a hierarchical model: "Sensory information is first received and interpreted by the primary sensory areas, then sent to unimodal association areas, and finally to the multimodal sensory areas. At each successive stage of this stream more complex analysis is achieved, culminating eventually, as with vision, for example, in object and pattern recognition in the inferotemporal cortex" (p. 353).

Kandel et al. actually divvy up the hierarchy of sensory systems into four parts, viz., (a) the primary sensory areas, (b) the unimodal areas, (c) the unimodal association areas, and (d) the multimodal association areas.

The primary sensory areas act as the base level, and they refer to the way in which information initially is communicated to the spinal cord and/or brain through one of the five sensory modalities, viz., touch, hearing, taste, smell, and vision. For example, in the visual system the primary sensory area is comprised of the eye, lateral geniculate nucleus, and the primary visual cortex located in the occipital lobe of the brain.

The unimodal areas build upon the data received from some prior particular primary sensory area, and refer to a higher-leveled integration of the data received from one of the primary sensory areas processed in a part of the brain different from that of the primary sensory area. In the visual system, there are two primary unimodal areas that process information concerning *where* an object is and *what* an object is, located along trajectories between the occipital lobe and parietal and temporal regions, respectively.

The unimodal association areas, in turn, refer to an even higher-leveled integration of the data received from two or more unimodal areas. In the visual system, the unimodal association area integrates data about the color, motion, and form of objects, and is located in the occiptotemporal area of the brain.

Finally, the multimodal association areas build upon the data received from the unimodal association areas and, depending upon the sensory modality, process this information in the parietotemporal, parietal, temporal, and/or frontal areas of the brain (also see van Essen, Anderson, & Felleman, 1992).

The result is this: information is exchanged at the various levels of the visual system and between the visual system and the central nervous system and, because of these exchanges, an animal is able to form a coherent picture of an object in its visual field, a visual perception (Crick & Koch, 2003; Baddeley, Vincent, & Attewell, 2011; Gray, 1999; Singer, 1999; Bullot, 2011; Arp, 2008).

3.5 Environments and Information

Organisms interact with external environments. However, because organisms are hierarchically organized living systems composed of subsystems, processes, and components engaged in various operations, they have their own internal environments as well. An *environment* can be defined as any pressure or force that interacts with, or affects somehow, the organism and its components. We can draw a distinction between the information that is exchanged *within* the organism's environment and the information that is exchanged *between* the external environment and the organism. So, there are really two types of environments, namely, environments that are *internal to* an organism and environments that are *external to* an organism.

Concerning internal environments, for example, the other organelles, nucleus, ATP, water, and various organic molecules act as the environment for a mitochondrion in the eukaryotic cell; other eukaryotic cells, cancerous cells, water, and all kinds of organic molecules and chemical elements act as the environment for a typical eukaryotic cell; a myriad of molecules including hydrogen, carbon, nitrogen, and oxygen surround and exert influence upon organs in a multi-cellular organism's body; a piece of food taken in from the environment external to the organism becomes part of the environment within the organism and, depending on the content, may be digested or expelled.

At the same time, the organism itself is interacting with external environments that are exerting pressures upon, as well as exchanging and communicating information with, the organism. Concerning external environments, we see that organisms are members of species that live in populations. These populations usually co-exist with other populations in communities. Many communities living with their non-living surroundings comprise an ecosystem, and the sum of all ecosystems make up the biosphere of the earth. Other members of a species, different species, and the non-living surroundings of an organism all are considered as parts of the external environment for an organism. The organism constantly experiences environmental pressures, and these pressures can be described in terms of information that is exchanged between the environment and the organism (Brandon, 1984, 1992). This kind of information exchange can be witnessed as a result of research accrued and experiments performed by biologists and other thinkers.

It is common knowledge that an organism's survival is dependent upon both genetic and environmental factors. For example, if there is an alteration in a rodent's genetic makeup causing it to have a malformed foot, then it is more likely to be eaten by a hawk out on the open range. However, if the same handicapped rodent lives in a forested area where it can hide under rocks and bushes, it is less likely to become a predator's victim. Also, if an environment happens to be made up of trees having fruit high up on its branches, and it just so happens that a fruit-eating animal's genes coded it to have a neck long enough to reach the fruit, then such an animal likely will survive. Conversely, if your animal genes coded you to have a short neck, it is unlikely you would survive in such an environment (that is, if the fruit high up in the trees was your only food source). In the words of Tim Berra (1990): "The environment is the selecting agent, and because the environment changes over time and from one region to another, different variants will be selected under different environmental conditions" (p. 8).

Another famous example that illustrates the informational transfer between the environment and an organism has to do with the finches that Darwin (1859/2009) described on the Galapagos Islands during his voyage on *The Beagle*. These finches clearly exhibit *adaptive radiation*, i.e., in the words of Berra (1990): "the evolutionary divergence of members of a single phylogenetic lineage into a variety of ecological roles usually resulting, in a short period of time, in the appearance of several or many new species" (p. 163). Darwin noted several different beak shapes and sizes that apparently were modified in the finches, depending upon the ecological niche the particular bird inhabited. Some finches had massive beaks ideal for crushing their seed food source, others had thinner pointed beaks ideal for probing flowers, still others had curved beaks ideal for picking food out of woody holes. In this situation, the environments in which the various finches inhabited were all different, and the finches with beaks most fit for a particular environment survived to reproduce.

Phenotypic traits are the physiological characteristics or behaviors of organisms that are under genetic control. The genetic information determines what a particular member of a species will look like, how fast it will run, what coloration it will have, how successful it will be at mating, etc. In the finch example, the different beaks represent the variety of phenotypic characteristics under genetic influence. If it just so happened that a certain beak style was effective in gathering food in an environment, then that finch would survive and pass its genes onto its offspring. Soon, that particular niche would be dominated by the beak style that was most advantageous for that environment.

Research has been conducted on animals to determine how the external environment affects the functioning of various systems of the body. One experiment has to do with occluding or removing the eyes of cats, rats, and birds at various stages of development to see if the neural connections of the brain necessary to the visual system either would develop abnormally, or cease to function altogether. These studies indicated that when occluding or removing the eyes, certain neural connections in the brains of these animals would not be made. This resulted in the cessation of certain visual processes, causing the overall subsystem to be underdeveloped in relation to other animals that have not had their eyes occluded or removed (Shatz, 1992; Clayton & Krebs, 1994). This example illustrates what happens when information *is not* exchanged between environment and organism.

A final example that demonstrates the information exchange between an organism and its environment has to do with the artificially controlled speciation of the fruit fly, *drosophila*. Experimenters are able to take out, move around, or add genetic sequences in the DNA of the fruit fly, causing radical phenotypic alterations in it to occur such as the deletion of some organ, legs growing where antennae should be, and antennae growing where legs should be. The experimenter's adjustments to the genetic material of the fruit fly are analogous to the radioactive material and other kinds of natural external forces of mutation that alter the genetic codes of fruit fly populations. We find similar monstrosities in fruit flies when we study them in their natural habitats (Duncan, Burgess, & Duncan, 1998). Just as researchers tap into and alter the genetic codes of fruit flies in controlled experiments, so too, external forces "tap into" and alter the genetic makeup of fruit fly populations in nature. These fruit fly abnormalities are another example of the property of environmental-organismic information exchange found in organisms.

4 A Few Debates Concerning Bioinformation and Bioinformation as a Relation

Despite its application to a broad range of disciplines—including the aforementioned examples in the biological sciences—appealing to the notion of information as an explanatory feature of living systems is a matter of much dispute, which in recent years has arisen over its need and usefulness.

4.1 Bioinformation: Metaphor or Reality?

Some authors consider information a distinctively linguistic phenomenon, so that its application in other fields is purely metaphorical and instrumentally useful. For example, when we were discussing action potentials of neurons above, we noted that protein synthesis in neurotransmitter release is the information that is communicated between neurons. It is possible to render information in the above description as a purely linguistic tool utilized for explanatory purposes—here, one might say, "protein synthesis in neurotransmitter release is not *really* information," and, in fact, we can skip the informational part and go directly to the *real* explanation associated with action potentials, namely, *protein synthesis in neurotransmitter release*. The notion of information is just that—a metaphorical notion—and does some explanatory work in explaining action potentials. But the real entities and processes doing all of the work consist of proteins, neurotransmitters, and the like physico-chemical phenomena.

This last point strikes a reductionist tone, and there are many reductionists who argue that the use of information concepts is redundant in fields like biology, which are subject to general laws of matter and energy. Such researchers think that biological phenomena should be explained in mechanical, electromagnetic, chemical, and thermodynamic terms, thus rendering informational conceptions—as well as other conceptions, for example, *function*—superfluous. According to this reductionist perspective, to speak of information in biology would just be an odd way of speaking of correlation and causation (Stuart, 1985; Griffiths, 2001; Sarkar, 1996, 2000; Janich, 1992; Kitcher, 2001).

Many researchers, however, think that the informational perspective sheds considerable light on biological phenomena, allowing us to understand living things in a way that would be otherwise impossible (Terzis & Arp, 2011; Maynard Smith, 2000, 2000a; Queiroz, Emmeche, & El-Hani, 2007; Godfrey-Smith, 2000; Griffiths, 2001; Roederer, 2005; Avery, 2003; Yockey, 2005). Proponents of biological information theory argue that many basic life processes include the storing, replicating, varying, transmitting, receiving, and interpreting of *real pieces of information* of various types; and these processes are perhaps irreducible to physical and chemical terms. Stated simply, such researchers are convinced that "there is more to informational talk in biology than mere metaphor" (Sarkar & Plutynski, 2008, p. xxi; also Sarkar, 1996, 2000, 2005; Griffiths, 2001).

There is no doubt that the presence of metaphors in biological texts is ubiquitous, and it is not just a question of informational metaphors (Keller, 1995, 2002). Darwin himself was called a "master of metaphor" by Stephen Jay Gould (Gould, 1989). From "natural selection" to the "immune self" (Tauber, 1994), all branches of biology constantly use very diverse metaphors. And all this is not incompatible with a realist reading of biological texts, for metaphors themselves may be interpreted in a realist way (Marcos, 1995, 1997, 2010, chapter 10).

The use of information theory as an instrument is very common in biology. As biological systems—from macromolecules to organisms—are very complex, we can use information theory to measure their structural complexity. In John Collier's (2007) words:

> I will compare the use of information as a technology of measurement, which does not imply that there is anything present that might be called 'information' with a stronger usage of information in biology that attributes information to biological systems in a non-instrumental way. This distinction between instrumental and substantive uses of information in biological studies often turns on the notion of information used, so it is important in each case to be clear what is at stake […] The instrumental usefulness of information technologies does not in itself imply the existence of substantive information. (p. 763)

But this instrumental application of the theories of information is also found outside biology. Any structure, living or otherwise, may be studied from this point of view. Following Collier (2007): "Some of the applications, however, present interesting issues for the philosophy of biology, especially concerning whether the instrumental use of information is sufficient to explain the use of the idea of information by biologists" (p. 767).

In other words, an instrumental interpretation is possible if we do not consider the purely biological, that is, if we consider living beings as mere physico-chemical structures. But, then, what sets living beings apart? "Arguably," Collier (2007) affirms, "to be alive requires this sort of separation of function and the requisite dynamical decoupling between metabolism and replication" (p. 770). So, the mutual reference between metabolism and replication must surely have an informational and functional character (also see Brooks & Wiley, 1988; Maynard Smith & Szathmáry, 1995). We can shed light on the structure of a gene only by showing its informational connection with a protein. We can say then that the function of a given fragment of DNA is to encode a protein. In an analogous way, we can explain the structure of a protein only by its reference to a vital function. And there are vital functions only when there exists an individual living being. So, living beings distinctively include information. This is the best possible explanation of the usefulness of informational concepts in biology. So, in the opinion of many authors a substantive explanation of bioinformation is required as part of the broader explanation of genetics.

4.2 Bioinformation as a Triadic Relationship

It is probably better to use the term *realist* rather than *substantialist* here. This is because when we speak of a substantialist interpretation of bioinformation, it would seem that we take for granted that bioinformation is a *substance*. From our point of view, bioinformation is a real entity, but not necessarily a substance. This observation leads us to another debate: If we accept that bioinformation is a real entity, what kind of entity is it precisely?

Some authors have viewed information as a thing, third *substance*, or primitive element. Wiener (1961), for example, thinks that information straightforwardly is “information, not matter or energy” (p. 132). Also, information has been seen as a *property* of a thing in terms of form, order, organization, negative entropy (Brillouin, 1962), complexity (Kolmogorov, 1965), or diversity (Margalef, 1980). Information as a property raises the problem of its location, which is a recurrent difficulty and, as such, one of the major arguments against the bioinformational paradigm. Actually, the problem of information location will be unsolvable unless we abandon this view of information as a simple property. Further, we find information conceptualized as a dyadic (semantic) and a triadic (pragmatic or functional) *relation*, as we hinted at in Section 2 above. As Barwise (1986) notes: “But is information relational? Surely so. The basic intuition about the information content C_S of a situation *s* is that it is information *about* something besides *s*”(p. 326; also Dennett 1987; Mackay, 1969; Küppers, 1990; Queiroz, Emmeche,& El-Hani, 2007).

On the other hand, information as a thing or basic substance should be the last hypothesis to explore, for the principle of ontological economy implies that, all things being equal, if some other hypothesis works, it is clearly preferable. The other three possibilities could be equated with the three parts of Weaver’s classical distinction (1949), which we explored briefly in Section 2.

The *technical problems*, which Weaver places at level A, are studied by considering the formal and statistical properties of messages. At this level, we are dealing with information as a property. The *semantic problems*, or level B problems, are concerned with the dyadic relationship between the message and its meaning. The *effectiveness problems*, or problems of level C, imply three elements. Weaver (1949) suggests that they are the message, its meaning, and a change in the receiver’s behavior caused by the reception of the message (p. 5). Therefore, problems of level C have a pragmatic aspect, which in biological contexts could be construed as a function. For instance, the change in cell behavior caused by the reception of a genetic message may consist in the accomplishment of a given function such as the synthesis of a determined protein.

In light of the above distinctions, we argue that bioinformation should be conceived as a triadic relationship, i.e., a relation involving three entities. The pragmatic or functional concept of information as a triadic relationship is the concept that best adapts to biological contexts, where functional explanations are very common (Cummins, 2002; Millikan, 2002; Perlman, 2004; Arp, 2006). We consider an explanation for the existence of an organ or a molecule satisfactory only if it includes reference, not only to its structure and material composition, but also to its *function* in the organism. Thinkers cannot seem to get around Trivers' (1985) claim that "even the humblest creature, say, a virus, appears organized to *do* something; it acts as if it is trying to achieve some purpose" (p. 5), or Arnhart's (1998) observation that "although the evolutionary process does not serve goals, the organisms emerging from that process do. Darwin's biology does not deny—rather, it reaffirms—the immanent teleology displayed in the striving of each living being to fulfill its specific ends [...] Reproduction, growth, feeding, healing, courtship, parental care for the young—these and many other activities of organisms are goal-directed" (p. 245). And what has been communicated in this paragraph above comports with the thinking of many biologists and philosophers of biology, including Collier (2007): "The relevant level for specifically biological information is the functional level" (p. 771).

Having said all of this, we can construe information—including bioinformation—in the following triadic, relational way. Information implies a relationship between:

1. a message, m, which may be any event, linguistic, or otherwise;
2. a system of reference, S, which the message informs the receiver about; and
3. a receiver, R.

Let's consider a fragment of mRNA as a concrete example of a message, while its system of reference could be a fragment of a protein. The receiver is a formal scheme residing in a concrete subject (a human being, another living system, a part of a living system, an ecosystem, a computer, or the cytoplasm of a cell for the preceding example). A concrete subject could, of course, use more than one receiver and use them alternately (playing with different "hypotheses") or successively (owing to an evolutionary or individual process of learning). We can also see the receiver as an internal (that is, residing in a concrete subject) predictive model of S, along the lines suggested by Rosen (1985), who characterizes living beings as "anticipatory systems."

A system of alternative messages in one relation can be a system of reference in another relation, and vice versa—and the process could be iterated. A segment of DNA can be a message informing the appropriate part of the cell about the mRNA to be synthesized. The same mRNA, initially part of a system of reference, may later become a message informing the cytoplasm about the synthesis of a certain protein, and so on. As Queiroz, Emmeche, & El-Hani (2007) state using Peircean vocabulary, "semiosis entails the installation of chains of triads" (p. 60). This is why a metaphor like "the flow of information" is sometimes useful.

Comparing this triadic view to the classical Shannon one (see figure 1), it may seem surprising that the emitter or source is not even mentioned. However, as Millikan (1989) rightly notes: we should "focus on representation *consumption*, rather than representation production" (pp. 283-284). Furthermore, there is often no specific emitter in non-linguistic contexts, like some biological ones, so a concept of bioinformation should not demand the presence of an emitter.

The matter of the channel is more complex because, usually, we have a dimensional image of it. However, it is possible to construe a channel in a more abstract way, as a set of conditional probabilities, along the lines suggested by Abramson (1963). In the same spirit, Barwise and Seligman (1997) note that a channel could be understood, basically, as an objective correlation of any degree between two domains.

Most of the conceptual problems concerning information actually stem from the ellipsis of some element of the informational relation. We often speak about the information of a message with no reference to a receiver or a referential system, although both of them exist implicitly. Bioinformation is always functional, transitive, and pragmatic. The message is always referred to something by a receiver; otherwise it is not a message, just an event (Millikan, 1989, p. 286). If messages were not referred to something by a receiver, Griffiths (2001) would be perfectly right to say that “most information talk in biology is a picturesque way to talk about correlation and causation” (p. 400).

However, factors conditioning information are often mistaken for information itself. Such is the case regarding the formal characteristics of the system of reference, and either those of the message or the system to which it belongs. The correlation between the messages and the system the information is about also affects the amount of information involved, but neither this *correlation* nor *form* constitutes the information itself.

The relationship among the three above-mentioned elements (m, R, and S) is informative when it changes the receiver’s knowledge of the system of reference. By *knowledge*, we mean the distributions of probabilities of the possible states of the system of reference in the receiver. Knowledge, therefore, should be understood here along the lines suggested by Karl Popper (1990) in a very general way: “Can only animals know? Why not plants? Obviously, in the biological and evolutionary sense in which I speak of knowledge, not only animals and men have expectations and therefore (unconscious) knowledge, but also plants; and, indeed, all organisms[…] Flowering plants know that warmer days are about to arrive […] according to sensed changes in radiation” (pp. 9, 10, 35).This understanding of knowledge does not necessarily imply consciousness, so the notion is applicable to human as well as to non-human living systems. Consider Rosen’s (1985) claim as well: “I cast about for possible biological instances of control of behavior through the utilization of predictive models. To my astonishment I found them everywhere[…] the tree possesses a model, which *anticipates* low temperature on the basis of shortening days” (p. 7).

We can describe information (I) as a relationship between a message (m), a receiver (R), and a system of reference (S). In this relationship can be found the triad formed by a message, receiver, and system of reference where the message alters the receiver’s previous knowledge of the system of reference (Dretske, 1981, 2007). Moreover, the more probable an alternative is to a receiver, the more information will be received when a message says that a different one has occurred, unless it is a simple contradiction. So, for example, the introduction of a certain genetic message into the cytoplasm increases the probability of the cell carrying out a certain function, for the probabilities of alternative behavior decrease. Now, we can say that the receiver knows—or knows better—how to do something. Again, this understanding of knowledge does not necessarily imply consciousness.

The informational relation, in accordance with our realistic interpretation of bioinformation, may be perfectly objective (see Barwise, 1986; Fodor, 1986; Denbigh & Denbigh, 1985). For example, it is quite objective that a genetic message informs the cytoplasm about synthesizing proteins. But, this does not necessarily mean that the information has been in the world since the beginning, preceding any subject capable of using it, as Dretske (1981) notes. Without cellular machinery there is no connection between DNA and protein. As Moreno and Ruiz-Mirazo (2002) state, the genetic message is, in principle, “decoupled from the dynamical organization of the system” (p. 73).

Information can be measured from the magnitude of its effects, that is, by the changes to the receiver’s knowledge of the system of reference (for a measure of this kind, see Marcos, 2011). This is a traditional and standard way of measuring different physical magnitudes. Measuring information—like measuring anything else—requires a subject to do it, and this subject acts according to theoretical grounds. To assess the quantity of information given by a genetic message to the cytoplasm, we need extensive biochemical knowledge, but this does not make the informational relation any less objective.

5 Bioinformation and Related Concepts

The concept of information is usually presented in connection with others that seem to form a constellation. The relationships between them, however, do not usually appear with sufficient clarity, which may make educational tasks difficult. At this point, a philosophical approach may be helpful, one that introduces clarity to the concepts and their mutual relationships. Throughout the chapter, such concepts as *correlation* and *form* have been appearing, and in this section others will make their appearance, such as *entropy*, *order*, *organization*, *complexity*, and *knowledge*. All of them are closely related to the notion of information, but none of them simply identifies itself with it.

5.1 Bioinformation and Thermodynamic Entropy, Order, and Organization

Above, we mentioned that August Weismann correlated information with the biological sciences in his 1904 book, *The Evolution Theory*. However, information first appeared in biology in connection with the concept of physical entropy and its different measures (thermodynamic or statistic) through physicists Rudolf Clausius (1822-1888) and Ludwig Boltzmann (1844-1906), who formulated the measures of entropy. Clausius was the first to introduce the term *entropy* to thermodynamics in 1876, while Boltzmann gave a statistical interpretation to the term. Boltzmann considered that a macrostate of a given system is more entropic in the same measure as it is compatible with a greater number of microstates.

The classic example is that of a gas-filled box. The box has two compartments, right (R) and left (L), connected by a door. The system can be in a macrostate A, in which the temperature in one compartment is relevantly higher than in the other one, or in a macrostate B, with equal temperature in both. According to the kinetic theory of heat, the temperature of each compartment varies depending on the kinetic energy of the particles in it. Thus, if the temperature in R is higher than the temperature in L, this is because the particles in R are on average faster than the particles in L. If there were the same temperature in both compartments, this would be due to a uniform distribution of the fast and the slow particles along the box. Let's call "a microstate" to a concrete distribution of the particles. So, the macrostate A has obviously a lower statistical probability than B because it is compatible with fewer microstates than B, and so B has higher entropy than A.

Boltzmann proposed the following formula for measuring thermodynamic entropy: $S = K \,\mathrm{Ln}\, W$, where S is the entropy of a given macrostate of a system, K is Boltzmann's constant, and W is the number of microstates compatible with this given macrostate. As can be easily observed, this equation is similar to the Shannonian formula for informational entropy, $H(S) = -K \sum_i P(s_i) \cdot \log P\,(s_i)$.

James Maxwell (1831-1879) took the next step with a thought experiment. If we place inside the box a demiurgic being (Maxwell's Demon) that allows the faster particles to pass to one compartment and the slower ones to the other, then the system evolves toward a less entropic state. Apparently, this situation is incompatible with the second law of thermodynamics.

Leo Szilard (1898-1964) found a sound answer to Maxwell's paradox. Maxwell's Demon overcomes the universal tendency to entropy thanks to the information he obtains about the speed of the particles. However, he had to measure the speed by means of whatever physical process, and any measurement process must involve some transaction of energy and increase of entropy. It seems, therefore, that an (inverse) link exists between entropy and information.

Taking inspiration from this idea, Léon Brillouin (1889-1969) developed the concept of *negentropy*, or negative entropy, as equivalent to information (Brillouin, 1962). Thinkers such as Tribus, Shannon, and Evans (1966) and Layzer

(1990) have attempted to equate information with a positive magnitude, the distance from thermodynamic equilibrium (also see Brooks & Wiley, 1988; Weber, Depew, & Smith, 1988; Marcos, 1991).

The last step prior to the *solidification* of the concept of information in biology was Erwin Schrödinger’s (1944) classic, *What is Life? The Physical Aspect of the Living Cell*, where he claims that living things overcome the universal tendency to entropy by exporting entropy to their environment, as Maxwell’s Demon does. Thus, a connection was made between thermodynamic order and biological complexity. Schrödinger contributes to the link between biological phenomena and physical entropy, and physical entropy had already been connected with information, so the stage was set for the encounter of information and biology. A slogan for this approach applied to biology could be, “A gain in (physical) entropy means a loss of (biological) information” or even Schrödinger’s (1944) own claim that, “life feeds on negative entropy” (p. 70).

From the point of view proposed here, thermodynamic entropy conditions the information that the macrostate of a system can offer about its possible microstates to a receiver equipped with the right physical laws. If the particles of the system act together, the system as a whole is more dynamic. Correspondingly, the movement of the system offers a great deal of information about its elements. If entropy increases, the system is less dynamic and reflects less efficiently the positions and moments of its components. Thus, thermodynamic entropy is linked *specifically* with the information that a macrostate can give about a system’s currently accessible microstates. So, the basis for a general measure of information could not be entropy, negentropy, or distance from equilibrium (Marcos, 1991).

Physical entropy is currently linked with (structural) order and (functional) organization, but order and organization are, respectively, relative to a structure and a function. Several types of order or organization may be identified even within the same system. Organization is also relative to a receiver connecting the message and the system of reference. A fragment of DNA is organized for the synthesis of a certain protein only if one knows how the cellular apparatus works. Physical entropy, therefore, should not be considered a general measure of organization; rather, it is a correct approach to *one* type of organization able to render work (Denbigh & Denbigh, 1985; Nauta, 1972). In biology, organization is always established with regard to a certain function. It is not just a question of structural regularity.

This is why Schrödinger (1944) conjectured, before the discovery of the double helix, that genetic information must be contained in some kind of *aperiodic* crystal. As it is well known, each crystal is formed by the periodic repetition of the same module. Biological macromolecules, such as proteins and DNA, are also modular compounds, but they are not formed by a periodic repetition of a singular module. In this sense, one can speak of them as aperiodic crystals. Schrödinger’s book, with its concept of aperiodic crystal, exerted a great influence on physicists and paved the way for many physicists to move into biological studies. One of the most prominent was Francis Crick, who noted that the book was “extremely well written" and made the subject seem "exciting" (Crick, 1965). Schrödinger's idea favored also the use of radiographic methods for the study of biological molecules; methods that were first developed for the study of the structure of the crystals. James D. Watson and Francis Crick, the two co-discoverers of the structure of DNA in 1953, used X-ray diffraction data collected by the British crystallographer, Rosalind Franklin (Ceccarelli, 2001; Ridley, 2006).

5.2 Bioinformation and Shannon's Entropy

Some remarks are in order here about the relation between Shannon's entropy and bioinformation. On the one hand, the structure of the system the message belongs to affects the information, but in the opposite way to that of the system of reference. When we try to pass information, we do not want the system to which the message belongs to impose any structural limitations on our communication, or at least we want them kept to a minimum. This is what Shannon calls *entropy* (freedom of choice within a source), and is recommended for a system acting as a symbolic one. A symbolic system is a system whose elements have a symbolic function, such as for instance the genetic system and the language. The word "table" symbolizes a table, as the codon UCG given the correct context symbolizes Serine. This is why in some parts of biological systems—for example, in neuronal, genetic, immunological, and linguistic domains—unities like can be combined in many different ways, for they must be flexible when representing other parts of the systems or external realities.

On the other hand, a higher level of structure or regularity in the system of reference brings about the possibility of transmitting more information about it with a given number of symbols, in line with common sense and philosophical tradition (Eco, 1962; Moles, 1972; Kolmogorov, 1965). For instance, a few words could be enough for describing the full molecular structure of a crystal, but the same number of linguistic symbols could say almost nothing on the molecular structure of a volume of gas.

Consequently, this matter is sometimes shrouded in confusion. It could be seen as a paradox that some authors correlate information with freedom of choice or low structural constraints, as Shannon does; while others, like Eco and Moles, correlate information with structural order, constraints, or regularity. But it is not paradoxical at all, but rather expresses two aspects of information. One aspect is the relative order of the system of reference, while another aspect is that of the symbolic system. Shannon's entropy of the symbolic system correlates positively with information: while regarding the object that the system informs about, it is the case that the greater the order and organization, the greater the amount of information that can be produced by a given sequence of symbols. Finally, another factor limiting the amount of information is the correlation between the structure of the message and that of the referential system. If it is perfect, a maximum amount of information can be transmitted. No greater correlation exists than between a system and itself. In this regard, Shannon's measure is an absolute limit on the amount of information: no more information can be given about a system than is given by the system itself. Therefore, Shannon's measure is often referred to as a measure of possible information.

5.3 Bioinformation and Complexity

Another approach to information has appeared more recently, based on the work of Andrey Kolmogorov (1965) and Ray Solomonoff (2003): algorithmic or computational theory. Here, information is viewed as a special kind of complexity. Any sequence describing a text, image, musical composition, etc., may be generated by means of a program and a suitable computer. If the sequence shows any regularity, symmetry, or redundancy, the program could be shorter than the sequence itself. If the sequence is more *complex*, or even random, it will be less susceptible to compression, so the greater the complexity, and the lesser the compressibility.Thus, for instance, under this approach, the "aperiodic crystals" to which Schrödinger refers are more complex than a periodic standard crystal.

But, it must be remembered that information, unlike complexity, is not a property of a single thing, but a relation between at least three entities (as we have mentioned already), so some remarks may be made on the relationship between Kolmogorov's complexity and bioinformation. First, the relationship between information and the complexity of a sequence is not a direct one, that is, complexity cannot be simply equated with information. The need for a long program to generate a sequence does not translate directly into that sequence "having" a great deal of information. It would be counterintuitive, for random sequences would be the most informational ones. Kolmogorov's measure of complexity can distinguish between a crystal and a protein, but a relevant concept of bioinformation must also distinguish between a functional protein and a random peptidic compound.

Second, Kolmogorov's notion of complexity has also been used to calculate the informational content of an individual object as a direct function of the length of the shortest program describing or producing it. Here, we must remember the difference between things and words. When complexity is assessed from the compressibility of a description encoded in a binary sequence, it could normally be referred to a universal Turing machine. The input into such a computer is a binary sequence, as is the output, so the computer cannot relate the description to the object itself. Therefore, a measure of the complexity of sequences is available, but this does not mean that we can calculate the complexity of the object described, because the information that a description gives about an object is always referred to a certain receiver in a concrete subject. For example, a DNA sequence is a good description of the three-dimensional structure of a protein *to certain cytoplasmic* *machinery*, but it would not make sense to say that it is generally, or for a Turing machine (see Rosen, 1985).

Third, there are doubts as to whether natural selection *alone* can explain the increase in complexity throughout evolution (Marcos, 1991a, 1992). After all, organisms exist that are very simple but seem perfectly adapted, a classic objection to Darwinism (Bertalanffy, 1968). The connection we have established between complexity and information may clarify the issue. Later variants in evolutionary succession may "take into account" those already existing, but not vice versa (Rosen, 1985). Once an organism A is settled into its environment, any other organism B will adapt to this environment more effectively if it is equipped to relate informationally with A. This informational asymmetry means that both the environment and organisms become more and more complex, and so maintain their adaptational dynamics throughout evolution. In this regard, complex biological organisms could be indicative of a complex environment, for more information is required to adapt to a complex environment than to exist in a simple one. The existence of living beings that adapt to an environment in which others already exist may ensure the survival of the latter, rather than threaten it, since the environment to which the new system adapts is also that on which they depend. Humanity's acceptance of this idea is not unconnected with the increase in ecological awareness. This remark, of course, has a direct link to the teaching of biology, which should promote ecological awareness.

5.4 Bioinformation and Knowledge

Information is also related to knowledge, which we hinted at in our discussion of mammalian visual perception in Section 3. In his highly influential work, *Knowledge and the Flow of Information*, Fred Dretske (1981) defines information as "a commodity that, given the right recipient, is capable of yielding knowledge" (p. 47). So a triadic relation is also needed here: we have the message, the circumstances it informs about, and the "right recipient." Information is therefore related to knowledge in a dual way: it depends on the receiver's previous knowledge, while knowledge is an effect produced by information. So, knowledge itself can be viewed as the property of a subject (edification), or as a dyadic relation between subject and object (correspondence, correlation). It is easy to connect the first notion of knowledge with biology: bioinformation contributes to the construction of living beings themselves. It is more difficult to apply the demand for truth. Nevertheless, we think that even in biological contexts information somehow requires truth (see Devlin, 1991).

5.5 Location of Bioinformation

Where is bioinformation? In our opinion, conceiving information as a relation could avoid the (pseudo-)problem of finding the location of bioinformation. It could be (dis)solved by considering information, not as being already present somewhere (in the genes, cytoplasm, proteins, environment, ecosystem, brain, or wherever) but as being established by the *interactive relations* between and among the parts and processes of living systems. This is a very important point for biology educators to consider and may help to clarify many misunderstandings.

The typical textbook presentation of DNA as "encoding" or "including information" make people think of it as the Holy Grail of biology. However, we must remember that DNA on its own codes for nothing; it is informational only in the cellular context. So, in this sense, genetic information is not *wholly* genetic. Further than this—as we have tried to show primarily in Section 3—biological information is not an exclusive property of the genes, but exists as a relationship between biological entities of different levels. Especially since the research that has resulted from the Human Genome Project, we have witnessed the development of various *omic* sciences, such as transcriptomics (the study of the set of all RNA molecules produced in one or more cells), metabolomics (the study of metabolites and other products associated with metabolism), and proteomics (the study of the structure and function of proteins). This is another indicator that bioinformation is not a simple property of the genes, but a complex relationship between different biological entities.

The functioning of any living system (or part of a living system) depends on various factors. For example, the three-dimensional structure of a protein depends on DNA, but also on the very machinery of the cell. What the message is and what the receiver is are chosen conventionally but not arbitrarily. A message is usually defined as a small factor of great specificity in relation to a given function and displaying a high potential for variability. The DNA codifying a certain protein possesses these characteristics in relation to the function of synthesizing the protein in question, and the protein in relation to its biological function. In other words, the slightest alteration of the DNA could destroy the structure of the protein, and the slightest change in a protein could destroy its function, as it happens in widely known genetic diseases such as sickle-cell anemia.

Such an effect is unlikely to be the result of a similar change in an environmental factor. But this does not force us to identify the information with a property of the message. The information in a fragment of DNA about a protein obviously depends on its specificity, but only regarding a given receiver (Sattler, 1986). Actually, the probability of any given protein arising in a prebiotic environment (Yockey, 1977, 1981), even in the presence of a specific DNA, is much smaller than in a cytoplasmic environment. Therefore, information is located neither before nor after the triadic relation. Kampis and Csányi (1991) state: "we have to give up the idea of a complete localization of information" (p. 23; also see Kampis, 1990).

On the other hand, any one fragment of DNA may, of course, produce information on more than one function, and not necessarily in the same quantity. For example, attempts could be made to calculate the amount of information in a fragment of DNA in relation to the transportation of oxygen, which is different from the function of producing a particular protein. The difference lies in the fact that the same function can be performed by different proteins or variants of a protein.

Finally, let us deal very briefly with the location of information in living systems according to different hierarchical levels (Collier, 2003). An organism can be conceived of as a hierarchically organized living system made up of components that are engaged in processes constituting coordinated subsystems, with the product of these processes and subsystems being a *particularized* homeostasis relative to their operations that contributes to the overall *generalized* homeostasis of the organism.

For all intents and purposes, in the absence of connecting principles, the amount of information obtained by an external observer—for example, a scientist—on a living system at different levels should be considered as amounts of information about different systems. Otherwise, more information would supposedly be derived about a living being from the knowledge of, for example, its atomic state than of its genetic makeup (see Atlan, 1972). Information concerning the atomic state is not about the living being *per se*, unless we have theoretical principles connecting atomic states with some functional characteristics. Developing principles of connection between levels is like developing a receiver that allows us to obtain information about one level from another, acting as a message. We know that, given certain principles of connection, one biological level can inform us about another, but we also know that a complete reduction is not viable, for any concrete informational relation is subject to imperfections.

6 Information and Education

6.1 Bioinformatics

Humans seem to be the only species that can produce *information* about information, biological or otherwise. In the last fifty years, the development of new technologies and the massive increase in the use of computers in all areas of human activity have led to a veritable explosion in the amount of data and information (about information) that is produced, used, and in need of management worldwide, constituting a veritable sea of extraordinary depth and breadth. This is especially true in the biological sciences, medical research, and medical practice. In these disciplines, thousands of scientists and clinicians are contributing daily to the accumulation of a massive body of biomedical knowledge and information, which we have hinted at already in our discussions of genome annotation and the newer omic sciences above.

Bioinformatics is now the word used for the categorizing, cataloguing, and coding of this biomedical information with the help of computers. The 11th edition of the *Merriam-Webster Dictionary* (2004) defines *bioinformatics* as "the storage, classification and analysis of biological information using computers" (p. 71), while Baxevanis and Ouellette (2005) define it simply as the "storage, organization, and indexing of biomedical information in computers" (p. 77). The challenge nowadays definitely concerns the ability to collect, categorize, manage, store, process, retrieve, disseminate, mine, and query all of this biomedical data and information appropriately and efficiently by computational means (see Arp, Smith, & Spear, in preparation; Mathura & Kangueane, 2009; Nishikawa, 2002; and the journals *Bioinformatics* and *Bioinformation*). In fact, every science, organization, and business has its own informatics, teeming with data (Beynon-Davies, 2003; Taylor & Joudry, 2008).

Further, when the biomedical data and information possessed by experts in the various subfields of biology and medicine is organized and stored in interconnected, calibrated, interoperable computer repositories, it is accessible to anyone anywhere in the world, in real time, and could be continuously updated in light of new scientific and medical discoveries. Also, the information contained in these databases could, in principle, be used as the basis for certain kinds of automated reasoning that would independently assist in furthering the goals of scientific research and clinical practice. And one can imagine the ways in which this is immediately beneficial for biomedical research, the curing of diseases, the treatments of patients, the construction of new technologies, the annotating of data, and the general welfare of humankind. Think of a doctor with immediate access to the most current information about all known diseases at the click of a mouse. Or, imagine a single, calibrated, integrated biomedical knowledge base—a kind of Great Bioinformatics Encyclopedia—comprehensive of all biomedical knowledge within one system. The authors of a 2007 *Scientific American* article concerning bioinformatics and the World Wide Web share a similar dream of a database that, when queried, would "give us a single, customized answer to a particular question without our having to search for information or pore through results" (Feigenbaum, Herman, Hongsermeier, Neumann, & Stephens, 2007).

6.2 Human Communication

The organism can be conceptualized as a hierarchical organization whereby levels of operation, in the forms of subsystems and processes, function interdependently with one another in this unified system. In order for all of the functions to take place in this living system, informational relationships must exist on different levels, from the genetic to the social. Genes communicate information, cells communicate information, subsystems and processes communicate information, the environment surrounding the organism communicates information, and we humans are unique in communicating *conceptual* information precisely about these various forms of bioinformation (Boeckx & Uriagereka, 2011). Conceptual information that exists in the social sphere of human communication and interaction is of particular importance to the biology educator for research reasons, as well as for teaching reasons having to do with conveying concepts concerning biological research, ideas, and principles in books, journals, the classroom, the lab, or online.

Some years back, John Tyler Bonner (1980) described culture itself as being rooted in informational terms. According to Bonner, culture is understood as the transfer of information through behavior and especially by virtue of the process of teaching and learning (also see Hintikka, 1973; Hintikka & Suppes, 1970; Baddeley, Hancock, & Földiák, 2000). Bonner did not limit this process of teaching and learning to human cultures; rather, he extended the concept of culture to other species. We can also speak of cultural information in different living systems, not only in humans (Laland & Galef, 2009). For example, one may speak of cultural learning in some aspects of birdsong, and in other forms of animal communication (see Oller & Griebel, 2008). We especially find transmission of information in primates, too (Goodall, 2000).

In 1976 Richard Dawkins described units of cultural information analogous to genes that he termed *memes*, and such an idea virtually single-handedly spawned an area of study known as *memetics* (Dawkins, 1976), that has become significant for biologists, psychologists, sociologists, anthropologists, philosophers, and many other researchers (Blackmore, 1999). Like genes, memes can replicate, mutate, compete, and even go extinct. Examples given by Dawkins include fashion, catch phrases, melodies, and various forms of technology. Of course, concepts expressed as theories, hypotheses, ideas, data, arguments, principles, and the like that one would find in a standard discipline like biology exist as straightforward examples of memes, too.

Now, it seems clear that genetic evolution and cultural evolution inform one another in mutual ways. It could be argued that cultural evolution gives continuity to genetic evolution, as would seem to be the case with memes mirroring genes. Within this framework of mutual informing, one could understand the human educational process as a type of memetic informational relationship that prolongs biological evolution, interacts with it, and maintains certain analogies with it. For this reason, when we speak of the transmission of conceptual information between teacher and student in the educational process, we are not talking about something absolutely distinct from the kinds of bioinformation we have described already, such as genetic information, cellular information, visual information, and organismal/environmental information exchange.

Although, as may seem obvious, in the context of education, informational relationships also have their own distinctive features, of which we should like to point out the following. First, as we have argued, the conceptual informational relationship is centered on the receiver. This is never more certain than in the educative process, the locus of which must obviously be the student. Here, lectures from teachers, books, articles, and other educational media carry on the function of messages generators, whose mission is to propitiate changes in the student's knowledge. In our case, the system of reference will be the world of living systems.

This could appear to be an excessively passive view of education, where the student is characterized simply as a receiver; however, this is not correct. Indeed, the cognitive changes are produced *in the* student and *by* the student, by means of the construction and management of different possible receivers. If we see the educational process as an informational relationship, we realize that it depends on messages received, generated by the lectures from teachers and other educational media. It also depends on the activity of the system of reference, that is, in this case, of objective and dynamic biological reality. If nature were not active—in the field, in the laboratory, and in the classroom—we could learn or teach but little biology. The upshot is that education in biology critically depends on the activity of the student, who cannot *learn* without *doing* (Marcos, 2011a).

Also, if education itself is considered in informational terms, we underscore one of the historical meanings of information that we described in Section 2. Information here is a *formative* relationship—in the moral sense of the term—that forms and informs students, and the teacher as well. The educational process is, indissolubly, a process of information and of formation. But the old idea that it is simply the teacher who forms the disciple is erroneous and incomplete. Formation is the result of an informational relationship in which the teacher and the student take an active part.

When all is said and done, then, our hope is that we not only continue to be students of the biological sciences and philosophy of biology in our own research, but also that in this chapter we have played a bit of the role of teacher for you, the reader, concerning the concept of bioinformation.

References

Abramson, N. (1963). *Information theory and coding*. New York: McGraw-Hill.

Adriaans, P., & van Benthem, J. (2007). Introduction: Information is what information does. In P. Adriaans &J. van Benthem (Eds.), *Philosophy of information* (pp. 3-28). London: Elsevier.

Alberts, B. (Ed.). (2002). *Molecular Biology of the cell*. New York: Garland.

Albrecht-Buehler, G. (1990). In defense of "nonmolecular" cell biology. *International Review of Cytology, 120*, 191-241.

Allen, C. (1992). Mental content and evolutionary explanation. *Biology and Philosophy, 7*, 1-12.

Arnhart, L. (1998). *Darwinian natural right*. Binghamton, NY: SUNY Press.

Arp, R. (2006). Evolution and two popular proposals for the definition of function. *Journal for General Philosophy of Science, 37*,2-12.

Arp, R. (2008). *Scenario visualization: An evolutionary account of creative problem solving*. Cambridge, MA: MIT Press.

Arp, R., Smith, B., & Spear, A. (in preparation). *Building ontologies with Basic Formal Ontology*. Cambridge, MA: MIT Press.

Artmann, S. (2008). Biological information. In S. Sarkar and A. Plutynski (Eds.), *A companion to the philosophy of biology* (pp. 22-39). Oxford: Blackwell.

Atlan, H. (1972). *L'Organisation biologique et la theorie de l'information*. Paris: Hermann.

Avery, J. (2003). *Information theory and evolution*. London: World Scientific Publisher.

Ayala, F., & Arp, R. (Eds.).(2009). *Contemporary debates in philosophy of biology*. Malden, MA: Wiley-Blackwell.

Baddeley, R., Hancock, P., & Földiák, P. (Eds.). (2000). *Information theory and the brain*. Cambridge: Cambridge University Press. In G. Terzis and R. Arp (Eds.), *Information and living systems* (pp. 289-308). Cambridge, MA: MIT Press.

Baddeley, R., Vincent, B., & Attewell, D. (2011). Information theory and perception: The role of constraints, and what do we maximize information about?

Barwise, J. (1986). Information and circumstance. *Notre Dame Journal of Formal Logic*, *27*, 324-338.

Barwise, J., & Seligman, J. (1997). *Information flow: The logic of distributed systems*. Cambridge: Cambridge University Press.

Baxevanis, A., & Ouellette, B. (2005). *Bioinformatics: A practical guide to the analysis of genes and proteins*. Hoboken, NJ: John Wiley & Sons.

Berkeley, G. (1735/1901). *Alciphron, or The minute philosopher*. In A. Fraser (Ed.), *The works of George Berkeley*. Oxford: Clarendon Press.

Berra, T. (1990). *Evolution and the myth of creationism*. Stanford: Stanford University Press.

Bertalanffy, L. (1968). *General systems theory: Foundations, development, applications*. New York: George Brazilller.

Beynon-Davies, P. (2003). Information systems: An introduction to informatics in organizations. New York: Palgrave.

Blackmore, S. (1999). *The meme machine*. Oxford: Oxford University Pres.

Boeckx, C., & Uriagereka, J. (2011). In G. Terzis and R. Arp (Eds.), *Information and living systems* (pp. 353-370). Cambridge, MA: MIT Press.

Bogdan, R. (1994). *Grounds for cognition: How goal-guided behavior shapes the mind*. Hillsdale, CA: LEA Publishers.

Boi, L. (2011). Plasticity and complexity in biology: Topological organization, regulatory protein networks, and mechanisms of genetic expression. In G. Terzis and R. Arp (Eds.), *Information and living systems*(pp. 205-250). Cambridge, MA: MIT Press.

Bond, N., & Siddle, D. (Eds.). (1989). *PsychoBiology: Issues and applications*. Amsterdam: Elsevier.

Bonner, J. T. (1980). *The evolution of culture in animals*. Princeton: Princeton University Press.

Brandon, R. (1984). On the concept of environment in evolutionary biology. *Journal of Philosophy*, *81*, 613-665.

Brandon, R. (1992). Environment. In E. Keller & E. Lloyd (Eds.), *Keywords in evolutionary biology* (pp. 81-86). Cambridge, MA: Harvard University Press.

Brandt, C. (2005). Genetic code, text, and scripture: Metaphors and narration in German molecular Biology. *Science in Context*, *18*, 629-648.

Brillouin, L. (1962). *Science and information theory*. London: Academic Press.

Brooks, D., & Wiley, E. (1988). *Evolution as entropy*. Chicago: University of Chicago Press.

Buetow, D. (Ed.). (1982). *The biology of euglena*. Academic Press, New York.

Bullot, N. (2011). Attention, information, and epistemic perception. In G. Terzis and R. Arp (Eds.), *Information and living systems* (pp. 309-352). Cambridge, MA: MIT Press.

Burian, R., & Grene, M. (1992). Editorial introduction. *Synthese*, *91*, 1-19.

Capurro, R. (1978). *Information: Ein Beitrag zur etymologischen und ideengeschichtlichen Begründung des Informationsbegriffs*. Munich: Saur.

Capurro, R., & Hjørland, B. (2003). The concept of information. *Annual Review of Information Science and Technology, 37*, 343-411.

Carroll, S. (2005). *Endless forms most beautiful: The new science of EvoDevo and the making of the animal kingdom*. New York: W.W. Norton & Company.

Ceccarelli, L. (2001). *Shaping science with rhetoric: The cases of Dobzhansky, Schrödinger, and Wilson*. Chicago: University of Chicago Press.

Clayton, N., & Krebs, J. (1994). Hippocampal growth and attrition in birds affected by experience. *Proceedings of the National Academy of Science*, *91*, 7410-7414.

Collier, J. (2003). Hierarchical dynamical information systems with a focus on biology. *Entropy*, *5*, 100-124.

Collier, J. (2007). Information in biological systems. In P. Adriaans and J. van Benthem (Eds.), *Philosophy of information* (pp. 763-787). London: Elsevier.

Collins, J. (1956). *A history of Modern European philosophy*. New York: Bruce Publishing.

Cover, T., & Thomas, J. (2006). *Elements of information theory*. New York: Wiley.

Crick, F. (1965). Recent research in molecular biology: Introduction. *British Medical Bulletin, 21*, 183-186.

Crick, F., & Koch, C. (2003). A new framework for consciousness. *Nature Reviews Neuroscience, 6*, 119-126.

Cummins, R. (2002). Neo-teleology. In A. Ariew, R. Cummins, & M. Perlman (Eds.), *Functions: New essays in the philosophy of psychology and biology* (pp. 157-172). Oxford: Oxford University Press.

Darwin, C. (1859/2009). *The origin of species by natural selection: Or, the preservation of favored races in the struggle for life*. New York: Bantam Books.

Davies, P., & Gregersen, N.H. (2010). *Information and the nature of reality: From physics to metaphysics*. Cambridge, UK: Cambridge University Press.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Denbigh, K., & Denbigh, J. (1985). *Entropy in relation to incomplete knowledge.* Cambridge: Cambridge University Press.

Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Devlin, K. (1991). *Logic and information.* Cambridge: Cambridge University Press.

Dretske, F. (1981). *Knowledge and the flow of information*. Oxford: Blackwell.

Dretske, F. (2007). Epistemology and Information. In P. Adriaans & J. van Benthem (Eds.), *Philosophy of information* (pp. 29-48). London: Elsevier.

Duncan, D., Burgess, E., & Duncan, I. (1998). Control of distal antennal identity and tarsal development in *Drosophila* by *Spineless-aristapedia*: A homolog of the mammalian dioxin receptor. *Genes & Development, 12*, 1290-1303.

Eco, U. (1962). *Opera aperta.* Milano: Bompiani.

The Economist (1902). Journalists and diplomats. December 6.

Edelman, G. (1987). *Neural Darwinism: The theory of neural group selection*. New York: Basic Books.

Elsasser, W. (1975). *Chief abstractions in biology*. New York: American Elsevier.

Ewens, W. (Ed.). (2010). Special issue: Entropy in genetics and computational biology. *Entropy, 12*, 5-7.

Feigenbaum, L., Herman, I., Hongsermeier, T., Neumann, E., & Stephens, S. (2007). The Semantic Web in action. *Scientific American, 297*, 90-97.

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J… Ysebaert, M. (1976). Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature, 260*, 500-507.

Floridi, L. (Ed.). (2003). *The Blackwell guide to the philosophy of computing and information*. London: Blackwell.

Floridi, L. (2007). Trends in the philosophy of information. In P. Adriaans & J. van Benthem(Eds.), *Philosophy of information* (pp. 113-132). London: Elsevier.

Floridi, L. (2011). *The philosophy of information*. Oxford: Oxford University Press.

FlyBase. (2001). Drosophila white paper. Retrieved from: http://flybase.org/static_pages/news/whitepapers/DrosBoardWP2001.pdf.

Fodor, J. (1986). Information and association. *Notre Dame Journal of Formal Logic*, *27*, 307-323.

Forrest, S., & Hofmeyr, S. (2000). Immunology as information processing. In L. Segal & I. Cohen (Eds.), *Design principles for immune systems and other distributed autonomous systems* (pp. 361-388). New York: Oxford University Press.

Gatlin, L. (1972). *Information theory and the living systems*. New York: Columbia University Press.

Giere, R. (1988). *Explaining science*. Chicago: University of Chicago Press.

Gilbert, S., & Sarkar, S. (2000). Embracing complexity: Organicism for the 21st century. *Developmental Dynamics, 219*, 1-9.

Gleick, J. (2011). *The information: A history, a theory, a flood*. London: Harper Collins Publishers.

Godfrey-Smith, P. (2000). Information, arbitrariness, and selection: Comments on Maynard Smith. *Philosophy of Science,67*, 202-207.

Goodall, J. (2000). *Through the window: My thirty years with the chimpanzees of Gombe*. Boston: Mariner Books.

Gould, S.J. (1989). The wheel of fortune and the wedge of progress. *Natural History, 89, 3,* 14-21.

Gould, S.J. (2002). *The structure of evolutionary theory*. Cambridge, MA: The Belknap Press.

Gray, C. (1999). The temporal correlation hypothesis of visual feature integration: Still alive and well. *Neuron*, *24*, 31-47.

Griffiths, P. (2001). Genetic information: A metaphor in search of a theory. *Philosophy of Science*, *68*, 394-403.

Grünwald, P., &Vitányi, P. (2003). Kolmogorov complexity and information theory. *Journal of Logic, Language and Information.12*, 497-529.

Hanson, P. (Ed.), (1990). *Information, language, and cognition*. Vancouver: University of British Columbia Press.

Hartley, R. (1928). Transmission of information. *Bell System Technical Journal*, *7*, 535-563.

HGPI (Human Genome Project Information). (2012). Home page. Retrieved from: http://www.ornl.gov/sci/techresources/Human_Genome/elsi/patents.shtml.

Hintikka, J. (1973). *Logic, language games and information*. Oxford: Clarendon Press.

Hintikka, J., & Suppes, P. (1970). *Information and inference*. Dordrecht: Reidel.

Holzmüller, W. (1984). *Information in biological systems: The role of macromolecules*. Cambridge: Cambridge University Press.

Horan, B. (1992). What price optimality? *Biology and Philosophy, 7,* 89-109.

Jablonka, E. (2002). Information: Its interpretation, its inheritance, and its sharing. *Philosophy of Science*, *69*, 578-605.

Janich, P. (1992). *Grenzen der Naturwissenschaft*. Munich: Beck.

Kampis, G. (1990). *Self-modifying systems in biology and cognitive science*. Oxford: Pergamon Press.

Kampis, G., & Csányi, V. (1991). Life, self-reproduction and information: Beyond the machine metaphor. *Journal of Theoretical Biology*, *148*, 17-32.

Kandel, E., Schwartz, J. & Jessell, T. (2000). *Principles of neural science*. New York: McGraw-Hill.

Keller, E.F. (1995). *Refiguring life: Metaphors in twentieth-century biology*. New York: Columbia University Press.

Keller, E.F. (2002). *Making sense of life: Explaining biological development with models, metaphors, and machines*. Cambridge, MA: Harvard University Press.

Kitcher, P. (2001). Battling the undead: How (and how not) to resist genetic determinism. In R. Singh, K. Krimbas, D. Paul, & J. Beatty (Eds.), *Thinking about evolution: Historical, philosophical and political perspectives (Festchrifft for Richard Lewontin*) (pp. 396-414). Cambridge: Cambridge University Press.

Kjosavik, F. (2007). From symbolism to information? Decoding the gene code. *Biology and Philosophy, 22, 3,*333-349.

Kolmogorov, A. (1965). Three approaches to the quantitative definition of information. *Problems in Information Transmission,1*, 1-7.

Kull, K. (1999). Biosemiotics in the twentieth century: A view from Biology. *Semiotica, 127*, 385-414.

Küppers, B-O. (1990). *Information and the origin of life*. Cambridge, MA: MIT Press.

Küppers, B-O. (2000). The world of biological complexity: origin and evolution of life. In S. J. Dick (ed.), *Many worlds: the new universe, extraterrestrial life, and the theological implications* (pp. 32–43). Radnor, PA: Templeton Foundation Press.

Laland, K., & Galef, B. (Eds.). (2009). *The question of animal culture*. Cambridge, MA: Harvard University Press.

Landauer, R. (1996). The physical nature of information. *Physics Letters A, 217,*188-193.

Layzer, D. (1990). *Cosmogenesis: The growth of order in the universe*. New York: Oxford University Press.

Li, M., & Vitányi, P. (1997).*An introduction to Kolmogorov complexity and its applications*.New York: Springer-Verlag.

Lorenz, K., & Wuketits, F. (1983). *Die evolution des denkens*. Munich: Piper.

Mackay, D. (1969). *Information, mechanism and meaning*. Cambridge, MA: MIT Press.

MacLaurin, J. (1998). Reinventing molecular Weismannism: Information in evolution. *Biology and Philosophy*, *13*, 37-59.

Marcos, A. (1991). Información y entropía [Information and entropy]. *Arbor*, *549*, 111-135.

Marcos, A. (1991a). Información y evolución [Information and evolution]. *Contextos*, *9*, 197-214.

Marcos, A. (1992). *El Papel de la información en Biología* [*The role of information in Biology*]. Barcelona: University of Barcelona (Ph.D. Thesis).

Marcos, A. (1995). Biología, realismo y metáfora [Biology, realism and metaphor]. *Ágora, 14, 1,* 77-97.

Marcos, A. (1997). The tension between Aristotle's theories and uses of metaphors. *Studies in History and Philosophy of Science, 28,*123-139.

Marcos, A. (2010). *Ciencia y acción* [*Science and Action*]. Mexico DF: F.C.E.

Marcos, A. (2011). Bioinformation as a triadic relation. In G. Terzis and R. Arp (Eds.), *Information and living systems* (pp. 55-90).Cambridge, MA: MIT Press.

Marcos, A. (2011a). Learning by doing: *Paideia* and *Phronesis* in Aristotle. *Educação* (Brazil), *34,* 13-24.

Margalef, R. (1968). *Perspectives in ecological theory*. Chicago: Chicago University Press.

Margalef, R. (1980). *La Biosfera, entre la termodinámica y el juego*. Barcelona: Blume Ediciones.

Marijuan, P. (1991). Enzymes and theoretical biology.*BioSystems*, *25*, 259-273.

Mathura, V. & Kangueane, P. (Eds.). (2009). *Bioinformatics. A concept-based introduction*. New York: Springer.

Maynard Smith, J. (2000). The concept of information in Biology. *Philosophy of Science*, *67*, 177-194.

Maynard Smith, J. (2000a). Reply to commentaries. *Philosophy of Science*, *67*, 214-218.

Maynard Smith, J., & Szathmáry, E. (1995). *The major transitions in evolution*. Oxford: W.H. Freeman Spektrum.

Mayr, E. (1982). *The growth of biological thought*. Cambridge, MA: Harvard University Press.

Mayr, E. (1996). The autonomy of biology: The position of biology among the sciences. *Quarterly Review of Biology*, *71*, 98-106.

Miller, K.(2005). *Communication theories: Perspectives, processes, and contexts*. New York: McGraw-Hill, 2005.

Millikan, R. (1989). Biosemantics. *The Journal of Philosophy*, *86*, 281-297.

Millikan, R. (2002). Biofunctions: Two paradigms. In A. Ariew, R. Cummins, & M. Perlman (Eds.), *Functions: New essays in the philosophy of psychology and biology* (pp. 113-43). Oxford: Oxford UniversityPress.

Moles, A. (1972). *Théorie de l'information et la perception esthétique*. París: ÉditionsDenoël.

Moreno, A., & Ruiz-Mirazo, K. (2002). Key issues regarding the origin, nature and evolution of complexity in nature: Information as a central concept to understand biological organization. *Emergence, 4,* 63-76.

Morgan, T.H., Sturtevant, A., Muller, H., & Bridges, C. (1915). The mechanism of Mendelian heredity. New York: Henry Holt.

Musgrave, A. (1993). *Common sense, science and scepticism*. Cambridge: Cambridge University Press.

Nauta, D. (1972). *The meaning of information*. The Hague: Mouton.

Nelson, D., & Cox, M. (2004). *Lehninger principles of biochemistry*. New York: Freeman.

Nishikawa, K. (2002). Information concept in biology. *Bioinformatics*, *18*, 649-651.

NPO (Nobel Prize Organization). (2012). The Nobel Prize in physiology or medicine 1968: Robert W. Holley, H. Gobind Khorana, Marshall W. Nirenberg. *Nobelprize.org*.

Oller, D., & Griebel, U. (Eds.). (2008). *Evolution of communicative flexibility: Complexity, creativity, and adaptability in human and animal communication*. Cambridge, MA: MIT Press.

Oyama, S. (2000). *The ontogeny of information: Developmental systems and evolution*. Durham: Duke University Press.

Paton, R. (1992). Towards a metaphorical biology. *Biology and Philosophy*, *7*, 279-294.

Peil, J. (1986). Information theoretical aspects of the aging process. *Archives of Gerontology and Geriatrics, 10,* 97-105.

Peirce, C. (1932-1935). *Collected papers*. Cambridge, MA: Harvard University Press.

Pelligrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1996). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176-180.

Perlman, M. (2004). The modern philosophical resurrection of teleology. *The Monist*, *87*, 3-51.

Pfeifer, J. (2006). The use of information theory in Biology: Lessons from social insects. *Biological Theory*, *1*, 317-330.

Piaget, J. (1970). *L'épistemologie génétique*. Paris: PUF.

Popper, K. (1990). Towards an evolutionary theory of knowledge. In K. Popper, *A world of propensities* (pp. 27-51). Bristol: Thoemmes.

Queiroz, J., Emmeche, C., & El-Hani, C. (2007). Information and semiosis in living systems: A semiotic approach. Retrieved from: http://www.library.utoronto.ca/see/SEED/Vol5-1/Queiroz_Emmeche_El-Hani_abstract.htm.

Reid, T. (1764/2001).In D. Brookes (Ed.), *An inquiry into the human mind on the principles of common sense*. State College, PA: Pennsylvania State University Press.

Reiter, L., Potocki, L., Chien, S., Gribskov, M., Bier, E. (2001). A systematic analysis of human disease-associated gene sequences in drosophila melanogaster. *Genome Research, 11*, 1114-1125.

Ridley, M. (2006). *Francis Crick: Discoverer of the genetic code*. New York: Harper.

Roederer, J. (2005). *Information and its role in nature*. Berlin: Springer-Verlag.

Rosen, R. (1985). *Anticipatory systems*. Oxford: Pergamon Press.

Rosenberg, A., & Arp, R. (Eds.). (2009). *Philosophy of biology: An anthology*. Malden, MA: Wiley-Blackwell.

Rosenberg, A., & McShea, D. (2007). *Philosophy of biology: A contemporary introduction*. London: Routledge.

Ruiz-Mirazo, K., & Moreno, A. (2011). The need for a universal definition of life in twenty-first-century biology.In G. Terzis and R. Arp (Eds.), *Information and living systems* (pp. 55-90). Cambridge, MA: MIT Press.

Ruse, M. (Ed.). (2008). *The Oxford handbook of philosophy of biology*. Oxford: Oxford University Press.

Russell, B. (1912/1999). *The problems of philosophy*. In J. Perry & M. Bratman (Eds.), *Introduction to philosophy: Classical and contemporary readings* (pp. 9-12). Oxford: Oxford University Press.

Salemi, M. & Vandamme, A. (Eds.). (2003). *The phylogenetic handbook: A practical approach to DNA and protein phylogeny*. Cambridge: Cambridge University Press.

Salthe, S., &Matsuno, K. (1995). Self-organization in hierarchical systems. *Journal of Social and Evolutionary Systems*, *18*, 327-338.

Sanger, F., & Coulson, A. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology, 94,* 441-448.

Sarkar, S. (1996). Biological information: A skeptical look at some central dogmas of molecular Biology. In S. Sarkar (Ed.), *The philosophy and history of molecular biology: New perspectives* (pp. 187-232). Dordrecht: Kluwer.

Sarkar, S. (2000). Information in genetics and developmental biology: Comments on Maynard Smith. *Philosophy of Science*, *67*, 208-213.

Sarkar, S. (2005). *Molecular models of life: Philosophical papers on molecular biology*. Cambridge, MA: MIT Press.

Sarkar, S. and Plutynski, A. (Eds.). (2008). *A companion to the philosophy of biology*. Oxford: Blackwell.

Sattler, R. (1986). *Biophilosophy*. Berlin: Springer-Verlag.

Schrödinger, E. (1944). *What is life? The physical aspect of the living cell.* Cambridge: Cambridge University Press.

Seife, C. (2007). Charles Seife, *Decoding the universe: How the new science of information is explaining everything in the cosmos, from our brains to black holes*. New York: Penguin Books.

Shannon, C.(1993).*Collected papers* (edited by N.J.A. Sloane and A.D. Wyner). New York: IEEE Press.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Shatz, C. (1992). The developing brain. *Scientific American*, *267*, 60-67.

Shendure, J., Mitra, R., Varma, C., & Church, G. (2004). Advanced sequencing technologies: Methods and goals. *Nature Reviews Genetics, 5*, 335-344.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron, 24*, 49-65.

Sober, E. (1993). *Philosophy of biology*. Boulder, CO: Westview Press.

Solomonoff, R. (2003). The universal distribution and machine learning. *Computer Journal*, *46*, 598-601.

Stein, L. (2001). Genome annotation: From sequence to biology. *Nature Reviews Genetics*, *2*, 493-503.

Stegmann, U. (2005). John Maynard Smith’s notion of animal signals. *Biology and Philosophy, 20*, 1011-1025.

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

Taylor, A., & Joudry, D. (2008). *The organization of information*. Santa Barbara, CA: Linworth / Libraries Unlimited.

Tauber, A. (1994). *The immune self: Theory or metaphor?* Cambridge: Cambridge University Press.

Terzis, G. and Arp, R. (Eds.). (2011). *Information and living systems*. Cambridge, MA: MIT Press.

Tipler, F. (1995). *The physics of immortality*. New York: Anchor Books.

Tribus, M., Shannon, P., & Evans, R. (1966). Why thermodynamics is a logical consequence of information theory. *Journal of the American Institute of Chemical Engineering, 12,* 244-248.

Trivers, R. (1985). *Social evolution*. Menlo Park, CA: Benjamin-Cummings.

Tweedle, S., Ashburner, M., Falls, K., Leyland, P., McQuilton, P., Marygold, S... The FlyBase Consortium. (2009). FlyBase: enhancing drosophila gene ontology annotations. *Nucleic Acids Research, 37,* D555-D559.

van Essen, D., Anderson, C., & Felleman, D. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science, 255,* 419-23.

Villanueva, E. (Ed.). (1990). *Information, semantics & epistemology*. Oxford: Blackwell.

Waddington, C. (Ed.). (1968). *Towards a theoretical biology*. Edinburgh: Edinburgh University Press.

Weber, B., Depew, D., & Smith, J. (Eds.). (1988).*Entropy, information and evolution*.Cambridge, MA: MIT Press.

Weismann, A. (1904).*The evolution theory* (trans. J. Thomson & M. Thomson). London: Edward Arnold.

West, R., & Turner, L. (2006). *Introducing communication theory: Analysis and application*. New York: McGraw-Hill.

Wicken, J. (1987). *Evolution, information and thermodynamics: Extending the Darwinian program*. Oxford: Oxford University Press.

Wiener, N. (1961). *Cybernetics, or control and communication in animal and machine*. Cambridge, MA: MIT Press.

Winder, R. (2012). *Philosophical aspects of information systems*. London: Taylor and Francis.

Yeung, R. (2006). *A first course in information theory*. New York: Springer.

Yockey, H. (1977). A calculation of the probability of spontaneus biogenesis by information theory. *Journal of Theoretical Biology*, *67*, 377-398.

Yockey, H. (1981). Self-organization origin of life scenarios and information theory. *Journal of Theoretical Biology*, *91*, 13-31.

Yockey, H. (2005). *Information theory, evolution, and the origin of life*. Cambridge: Cambridge University Press.

Glossary

Bioinformatics: The science concerned with collecting, categorizing, managing, storing, processing, retrieving, disseminating, mining, and querying biomedical data and information appropriately and efficiently by computational means.

Bioinformation (Biological Information): Information in the biological realm. Information implies a relationship between: (1) a message, m, which may be any event, linguistic, or otherwise; (2) a system of reference, S, which the message informs the receiver about; and (3) a receiver, R. Bioinforamtion occurs when there are biological entities—nucleic acids, cell cytoplasm, proteins, antibodies, neurons, sensory organs, organisms, or even ecosystems—involved as such in the informational relationship.

DNA Information: Since the middle of last century, DNA has been often identified as the informational molecule par excellence. It has become commonplace to say that the DNA "encodes," "contains," or "stores" information; even that it "transmits" or "conveys" hereditary information from one generation to another. What these expressions really mean is that DNA plays an important role in certain bioinformational relationships, such as reproduction and molecular synthesis. In these relationships, the DNA usually plays the role of a message. A message is usually defined as a small factor of great specificity in relation to a given function and displaying a high potential for variability. DNA possesses precisely these characteristics in relation to reproduction and metabolism. However, this does not force us to identify the bioinformation simply with a property of the DNA. We should see bioinformation, instead, as a complex relation in which the DNA has an important role.

Genetic Annotation: Reading *annotation* here as "commentary" or "explanation," the methods and technologies used to identify the locations of genes (as well as the coding regions in a genome) and determine specifically what those genes do.

Genetic Code: The specific sequences of nucleotides that are composed of a sugar (deoxyribose in DNA, ribose in RNA), a phosphate group, and one of four different nitrogen-containing bases, namely, adenine, guanine, cytosine, and thymine in DNA (uracil replaces thymine in RNA). These four bases are like a four-letter alphabet, and triplets of bases form three-letter words or *codons* that comprise the "information" which identifies an amino acid or signals a function. DNA is the template from which RNA copies are made that transmits genetic information concerning an organism's physical and behavioral traits (phenotypic traits) to synthesis sites in the cytoplasm of the cell. mRNA takes this information to ribosomes in a cell where amino acids, and then proteins, are formed according to that information. The proteins are the so-called *building blocks* of life, since they ultimately determine the physical characteristics of organisms.

Genetic Sequencing: The methods and technologies used since the early 1970s to determine the specific order of the bases in a molecule of RNA (adenine, guanine, cytosine, and uracil) or DNA (adenine, guanine, cytosine, and thymine).

Information: The English word *information* derives from the Latin noun *informatio*, which can mean, "representation," "idea," or "explanation." Also, the Latin verb *informo* can mean, "to sketch," "to draw," or "to represent" something as well as "to give shape or form" to something.

- **DNA**, see **DNA Information.**
- **Entropic**, see **Information Entropy.**
- **Molecule**, see **Molecule Information.**
- **Shannon**, see **Information Entropy.**
- **Triadic View of**, see **Triadic View of Information.**

Information Entropy (Shannon Entropy): Deriving from Claude Shannon's ideas and arguments in *The Mathematical Theory of Communication* (1949), a measure of uncertainty, usually expressed in bits, whose mathematical formula is $H(S) = -K \sum_i P(s_i) \cdot \log P(s_i)$, where H is the entropy of a source, S is a source (that is, a discrete random variable), s_i is one of the possible values of S, $P(s_i)$ is the probability of s_i, and K is a positive constant. In more intuitive terms, information entropy enables us to estimate the amount of uncertainty reduced on average by each symbol produced by a given source.

Informational Molecule: Any molecule capable of participating in a bioinformational relationship, either as a message, receiver, or a reference system such as a fragment of DNA, a neurotransmitter, an antigen, and a protein. When speaking specifically about genetic or hereditary information, it is very usual to ascribe the role of message to the DNA or RNA, the role of reference to the proteins, and the role of receiver to the molecules of the cytoplasm, such as those that make part of a ribosome. This ascription is not arbitrary, but it is worth noting that the same molecule may be involved in different informational relationships with different roles. For instance, an mRNA molecule may be seen as the reference of a DNA fragment, but it can be also seen as a message regarding a protein.

Philosophy: The word *philosophy* comes from two Greek words: *philos* deriving from *philein* "love," and *sophos* meaning "wisdom." Love here means something like an intense desire for something, while wisdom is arguably a kind of knowledge gained from experience, whether this is practical experience (gained from living life with all of its ups and downs) or theoretical experience (gained from understanding, evaluating, critiquing, and synthesizing ideas, positions, and concepts). Ever the theoretician, the philosopher has always been the person who not only desires to look deeper into some claim, idea, argument, event, or state of affairs by questioning assumptions and challenging status quo thinking, but also attempts to explain and systematize aspects of reality as it is perceived. In Bertrand Russell's (1872-1970) words, which are appropriate given the nature of this book, "Philosophy, like all other studies, aims primarily at knowledge. The knowledge it aims at is the kind of knowledge which gives unity and system to the body of the sciences, and the kind which results from a critical examination of the grounds of our convictions, prejudices, and beliefs."

Philosophy of Biology: A sub-discipline of philosophy, the concern of which is the meta-leveled attempt on the part of philosophers, biologists, and other thinkers to understand, evaluate, and critique the methods, foundations, history, and logical structure of biology in relation to other sciences, disciplines, and life endeavors so as to better clarify the nature and purpose of biological science and its practices.

Triadic View of Information: A concept of information put forward by thinkers such as Alfredo Marcos (see this volume) that implies a relationship between: (1) a message, m, which may be any event, linguistic, or otherwise; (2) a system of reference, S, which the message informs the receiver about; and (3) a receiver, R.