Reductionist perspectives and the notion of information

CrossMark

Nicole Perret [a, *], Giuseppe Longo [a, b]

[a] *Centre Cavaillès, République des Savoirs, CNRS USR 3608, Ecole Normale Supérieure, Paris, France*
[b] *Department of Integrative Physiology and Pathobiology, Tufts University School of Medicine, Boston, MA, USA*

ARTICLE INFO

ABSTRACT

Reductionism is the dominant stance of biology. According to this perspective, biological phenomena have to fit with physical explanations. Some biologists thought that the introduction of the idea of program was a sound way to overcome both physicalism and reductionism. We argue instead that the introduction of information theory into biology did not liberate biology from reductionism. We argue that the adoption of information in biology is an erroneous transposition from a specific mathematical domain to one where it does not belong. Indeed, the mathematical framework of the information theory is too rigid and discrete to fit with biological phenomena. Therefore, information in biology represents an inappropriate metaphor. Then, we make explicit the use of metaphors and the choice of explanation mode. We argue that the choice of explanation is not neutral. Furthermore, the use of metaphors in science becomes dangerous when they take the place of theories and they lose their paradoxical content.

© 2016 Elsevier Ltd. All rights reserved.

Contents

"To think that the genome completely determines the organism is almost as absurd as thinking that the pipes in a large cathedral organ determine what the organist plays."

Noble D. (2012). A theory of biological relativity Interface Focus.

1. Introduction

For at least two centuries two distinct and even antagonistic stances co-existed in biology. In current terminology, we refer to them as reductionism and organicism. The former is going to be addressed in this article while the latter is addressed in Mossio et al, 2016. The history of physics shows that a new phenomenon has always engendered new observables and principles. For instance, Galileo proposed momentum and its principle of conservation; thermodynamics studies trajectories in a relevant phase space: pressure, volume, temperature. A new observable, entropy, has greatly enriched physics by providing a principle that can be applied to any form of energy transformation: the second principle of thermodynamics.[1] Does biology operate similarly regarding the choice its observables and invariants? At the beginning of the 20th century, the central goal was to find observables and principles to understand the phenomenon of reproduction resulting in the

---

* Corresponding author.
   E-mail addresses: nicole.perret@ens.fr (N. Perret), Giuseppe.Longo@ens.fr (G. Longo).

---

[1] See Longo and Soto, 2016; see also Bailly and Longo, 2011; Chibbaro et al., 2015.

hereditary transmission of phenotypes. In this context, the search for a proper observable specific to biology headed toward the notions of encoding and of program that are at the core of the theories of information.

In this paper, we argue that information is problematic for biology for at least three of the reasons that we will analyze here. First of all, we claim that the transposition of the mathematical theory of information into the biological field is scientifically erroneous. In the first section, we analyze the incompatibility between the information sciences and the biological object.

The second reason is related to a general problem of reductionism. According to the distinguished biologist Ernst Mayr, the information field provides an anti-reductionist framework for biology. Despite this viewpoint, we argue that applying the theory of information to biology belongs to a reductionist attitude. In the second section, we analyze this reductionist approach, and we point out the relationship with determinism. We show that the reductionist stance hides the general idea according to which classical determinism is the regular form of scientific knowledge and that this is also true when information theory is applied to biology. For the reasons analyzed in the first section, it will be clear that the deterministic theoretical framework is inappropriate for the biological object.

Faced with this kind of criticism, oftentimes biologists defend the use of the idea of information, as well as the concepts of signal and program, as *just* useful metaphors or fruitful ways with which to approach a phenomenon by using an image borrowed from common sense. That is why, in the last and conclusive section, we analyze the general use of metaphors and common sense in science, and we show their dangerous consequences. This is particularly the case of the idea of information and the genetic program in biology because, as we will explain, here we face a dead metaphor.

2. Information sciences and biology

The use of the concept of information in biology appeared in the middle of the 20th century, but it is related to what happened at the beginning of that century when the possibility of isolating chromosomes was coupled with the new concept of the mendelian gene as a functional unit of recombination. Mendel's writings did not directly imply this concept, which appeared after 1900 when Hugo de Vries, Erich von Taschermark, and Carl Correns "rediscovered" Mendel (see Pichot, 1999). Johannsen (1911), then, replaced the term *mendelian factor* with the term *gene* and suggested the consequent distinction between genotype and phenotype (see Moss, 2004). Therefore, the idea of associating a phenotype with a segment of these chromosomes appeared. Schrödinger (1944) then, realized that this association was not well founded in a law: "It seems neither adequate nor possible to dissect into discrete 'properties' the pattern of an organism which is essentially a unity, a 'whole'. Now, […] a pair of ancestors are different in a certain well-defined respect […] we locate in the chromosome the seat of this difference. Difference of properties, to my view, is really the fundamental concept rather than the property itself" (Schrödinger, 1944, p. 10). This great physicist understood that differential analysis does not allow for the deduction of a *law* in the physical sense. In order to obtain a law in the proper sense, it would be necessary to propose a direct causal link, between the *wild gene* and the normal phenotype. He introduced the notion of encoding, borrowed from the new sciences of coding, in order to provide a theoretical framework and establish this hypothetical correlation. In other words, the fact that a mutation modifies the structure of an enzyme does not allow for the deduction of a direct *one gene - one enzyme* correlation[2] (following Beadle and Tatum, 1941). The notion of information was introduced as a theoretical framework providing this direct causal link. However, it is scientifically inexcusable to adopt this notion in biology without clarifying its usage with respect to at least two of the fields which make rigorous use of it. These fields are information elaboration (Turing-Kolmogorov) (Turing, 1936), or algorithmic theory of information (see Calude, 1994; Davis, 1958), and information transmission (Shannon-Brillouin) (Brillouin, 1962; Shannon, 1948).

The Central Dogma of molecular biology (Crick and Watson, 1953; Crick, 1970) suggested that the description of the chemical structure of the DNA molecule represents well the core of the informational/algorithmic view of biological phenomena. The idea here is that the expression from nucleic acid to protein is a unidirectional flow of information. Which information theory is involved here? Despite the different scientific implications of these theories, there is a significant confusion in biology which is rarely clarified. Maynard Smith (2000) explicitly refers to information elaboration (Turing-Kolmogorov) and to information transmission (Shannon-Brillouin) while emphasizing the relevance of the *latter* in biology. However, in the same text, he explains how molecular encoding can work as a short "recipe" (his wording) for generating complex, but organized (ordered) objects. The analogy is then the recipe for describing a circle by the three parameters which determine it. This recipe is less complex and contains less information than a point by point description of the circle. On the contrary, a totally disordered set can only be described point by point, as it does not obey as a rule. Now, this notion of a short (compressed) program for an organized object and of (maximal) informational complexity of disorder is that of Kolmogorov, and it is *covariant* with entropy (total disorder has maximal entropy). Note instead that, according to Shannon and Brillouin, complexity, as covariant with the quantity of information, is *contra-variant* with entropy and is in fact negentropy. This is also how physicists describe it, for sound reasons internal to the theory of "transmission of information" which thus differs greatly from Kolmogorov's one, a theory of "elaboration of information" (Longo et al., 2012). Another important difficulty lies in the fact that information in the two senses explored above deals with the realm of the discrete. Now, in this discrete framework, that is a precise concept in mathematics, only the dynamics of the discrete parts are relevant for the explanation of the entire system. In biology these discrete parts are molecules, so molecules alone forcibly and fully retain the researchers' attention. In this context it would be very hard to integrate, as a *positive* contribution to the expression of information, others events such as torsion, pressure (see for example Lesne and Victor, 2006; Farge, 2003), the dynamics of contact, geometries and relative distances, which all causally contribute to gene expression. A computer (Turing) or a cable (Shannon) does not receive a positive contribution from these observables which are better understood using continuous mathematics.[3] In fact, in both the mathematical theories of information, such dynamics can only cause an increase in entropy as noise. On the contrary, in biology we need a theory that includes such dynamics in a causal sense because they have a specific role in the gene expression and in the morphological constitution of the organism.

Furthermore, in both theories of information, the flow of

---

[2] Which we now know to be an erroneous bijective correspondence.

[3] It would be absolutely possible and interesting to provide a theory of bio-information by integrating the theory of continua. Control theory, for instance, deals with information by differential equations in a continuum, not to mention the new productive area of Information Geometry (Barbaresco and Djafari, 2015), entirely ignored in (molecular) biology.