An exercise of applied epistemology: Peirce’s Semiosis implemented in the Representation of Protein Molecules

DOI:10.36253/Substantia-2558

An Exercise of Applied Epistemology: Peirce's Semiosis Implemented in the Representation of Protein Molecules

Elena Ghibaudi*
Department of Chemistry, University of Torino, Via Giuria 7, 10125 Torino, Italy

**Contacts of corresponding author*
Via Giuria 7, I-10125 Torino (Italy)
Tel. ++39-(0)11-6707951
Fax ++39-(0)11-6707855
E-mail: elena.ghibaudi@unito.it

Received: Mar 06, 2024 **Revised**: May 27, 2024 **Just Accepted Online**: May 29, 2024 **Published**: Xxx

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record.
Please cite this article as:

E. Ghibaudi (2024) An Exercise of Applied Epistemology: Peirce's Semiosis Implemented in the Representation of Protein Molecules. **Substantia**. *Just Accepted*. DOI: 10.36253/Substantia-2558

Abstract
In their disciplinary communication chemists make a broad use of iconographic means. In this paper, some aspects of the iconographic communication are discussed, with specific reference to the representation of protein molecules in the light of Peirce's and Eco's semiotics. As far as Pierce's thought is concerned, I discuss two triads (*representamen*, *interpretant*, and *object)* and (*icon*, *index*, *symbol).* Eco's distinction between s-codes and codes is equally applied to the analysis of protein icons. The symbolic and iconic aspects of proteins' representations are discussed, in the light of various conventions that regulate the use of shapes, lines, shadows, colors in the building up of images. The iconic aspect turns out to be the most surprising, not just because it makes 'visible' what is inherently invisible, but also because of its heuristic potential. I argue that the construction of protein images and their use in research qualify their epistemic status as that of conjectures.

Keywords: semiosis, protein models, iconographic communication, Pierce, Eco

Introduction[1]
Scientific communication makes a broad use of iconographic means. Chemistry is no exception, having a long tradition of inventing and using iconographic means of expression. A renown image in the history of chemistry is the table published by John Dalton in 1808, whose caption was the following: «This plate contains the arbitrary marks or signs chosen to represent the several chemical elements or ultimate particles».[2] In fact, chemistry relies heavily on the symbolic dimension[3], that finds expression on both the linguistic level (chemical symbols, molecular formulas, chemical equations, etc.) and the iconographic level (diagrams, structure formulas, molecular models, iconic representations, etc.). The linguistic and iconographic aspects of chemical communication have been the object of a wealth of studies (see, for example, refs.[4,5,6,7,8]). Interestingly, iconography has become more and more relevant in chemistry: Weininger remarks that «the decreasing importance of linguistic signs such as names, compared to iconic signs such as structural formulas, accords with and reinforces the intensely visual character of chemistry».[9]
Over time, the use of graphical means has become increasingly widespread in chemistry, due to the need to process and make intelligible an increasing amount of structural data. Concurrently, the availability of large computing powers and the impressive development of visual softwares (diagrams, pictures, etc.) made sophisticated graphic representations available,

and scientists of various disciplines developed the most different ways of representing molecular properties. The representation of the molecular realm through these tools has opened up new perspectives for chemistry researchers; at the same time, the design and use of such representations raise considerable epistemic issues[8,10,11].

Chemical iconographical expressions are veritable working tools for chemists, as they are endowed with heuristic potential: according to Klein «Scientists often apply inscriptions or systems of signs not to present, illustrate and justify already existing knowledge, but as tools on paper for producing new representations and new knowledge».[12] Mestrallet goes even further, by stating that «molecular representations are tools for modifying reality»[13].

Weininger[9] underlines the strict relationship ("entanglement") between chemical theory, practice and representation. This view is shared by Hoffman and Lazslo who, referring to structural formulas, remark that «The writing of structures is not innocent. It is ideology-laden. It carries the modern reunification of the theoretical and the experimental».[14]

Moreover, in the educational literature, special attention has been put on the role played by molecular graphics in chemistry learning and understanding: Khanfour-Armalé and Le Maréchal remark that «the use of representations is essential for visualizing microscopic phenomena and thus helping students to solve problems»[15]. In fact, Lazlo reminds that 'to think chemically' implies the ability to absorb both textual and iconic information: «Chemists are endowed with what can be termed schizovision» [8].

Given the amplitude of the literature in this field, I have chosen to confine my analysis to the varied epistemic functions of iconic representations of proteins. In doing so, I will follow a philosophical approach that can be dubbed of 'applied epistemology'. The term is neither new nor recent. In 1989 Mark Battersby defined critical thinking as an 'applied epistemology' whose task would be to analyse the «epistemological claims that are not necessarily addressed in any discipline and deserve philosophical reflection»[16]. In 2006, Battersby wrote on *Informal Logic*: «Applied epistemology," [...] focuses the discipline towards the actual practice of how people come to and should come to justified beliefs. In an analogy with applied ethics, the study of people's actual epistemological practices can provide both information and challenges for the theoretician of reasoning»[17]. In the same year, Larry Laudan gave a good definition of 'applied epistemology': «Applied epistemology, in general, is the study of whether systems of investigation that purport to be seeking the truth are well engineered to lead to true beliefs about the world»[18]. My position towards applied epistemology is that philosophy, epistemology and semiotics have provided a wealth of sophisticated tools that can be applied to the study of theoretical and experimental scientific practices, in details. In the present exercise of applied epistemology, I mainly exploit the 'tools' developed by two authors who provided great contributions to both philosophy and semiotics: Charles Sanders Peirce and Umberto Eco.

The title of the present paper relates Peirce and semiosis; in fact, the American philosopher is counted among the founders of semiotics. Nevertheless, Peirce was more than an expert in semiotics, as he provided important contributions to the philosophy of science. These contributions are especially relevant as they come from a professional scientist, a man who knew science from inside: Peirce was a chemist by training, and from 1896 to 1902, he worked as a consulting chemical engineer for St. Lawrence River Power Co. in Massena (NY)[19]. Hence, Peirce was well-prepared to link the stringent materiality of empirical and technical facts with the abstract character of their interpretation.

Iconographic Communication: semiotic tools for an interpretation

Peirce's intellectual production is impressive. Most of it lies in the huge amount of manuscripts conserved in Harvard (more than 100,000 pages) that cover several decades. Therefore, studying Pierce is a challenge from both the quantitative and the qualitative standpoint, because the American philosopher changed his views on important points of his philosophical system, repeatedly. In the present study, I choose to refer to the last period of Pierce's thought; quotations are excerpts from writings that date between 1893 and 1907. Similarly, as regards to Umberto Eco, I choose to mention only one of his earlier works, the treatise *A Theory of Semiotics*, published in the United States in 1976, out of the wide production of this Italian philosopher.

Peirce's semiotic analysis and semiosis

It is well known that Pearce's thought was largely based on triads. He designed several triads, with different nature and meanings. I will comment on two fundamental triads that refer to the nature and the classification of signs (Figure 1). Peirce gave many definitions of sign[20], but probably the best known (and clear) is the following:

«A sign, or *representamen*, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the *interpretant* of the first sign. The sign stands for something, its *object*. It stands for that object, not in all respects»[21] (italics in the text).

This definition states the fundamental relation between representamen, object and interpretant (not to be confused with the interpreter). Peirce's semiotics is entirely developed from this triad, taken as a starting point. The *representamen* is intended as «something which stands to somebody for something in some respect». This is one of the main features of a sign: it has no universal character and is always addressed to somebody (not to anybody). The *interpretant* is what is created in the mind of the observer as a result of the perception/reception of the sign. The *object* cannot be fully represented by the sign («not in all respects»). Peirce seems to suggest that the expressive effectiveness of a sign is tied to both what is present and what is absent in the sign. The expressions «in some respect» and «not in all respects» draw attention towards the viewpoint from which the sign is employed. They are not logically equivalent: in fact, only the latter guarantees that other viewpoints may be taken into consideration, in addition to the one that is proposed. What is absent in the first interpretation of the sign can be (partly) found in the *process of semiosis*.

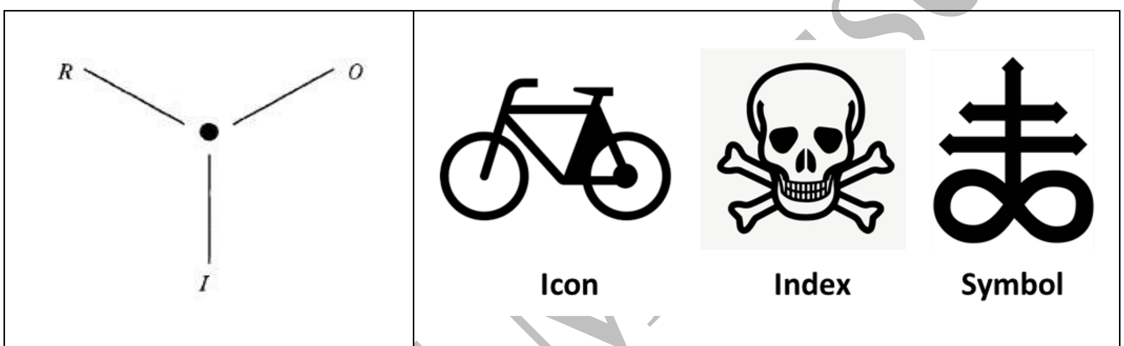


Figure 1. Left panel: nature of signs; R = representamen, O = object, I = interpretant[22]. Right panel: classification of signs. The sign associated with 'Icon' has a similarity relationship with its object. The sign associated with 'Index' is the toxicity mark, widely used by chemists to signal toxic substances. The sign associated with 'Symbol' is the Leviathan cross (sulphur in alchemy).

A sign is understood through a mental process that Pierce calls '*semiosis'***:**

«An action, or influence, which is, or involves, a cooperation of three subjects, such as a sign, its object, and its interpretant, this tri-relative influence not being in any way resolvable into actions between pairs»[23].

Semiosis has to be intended as a series of mental acts. The first one – according to Peirce – is an action that involves a cooperation of three subjects, sign, object and interpretant. But the interpretant, which in turn is a sign, may become a new representamen; so, a new triad may come out and foster a further semiotic act that leads to a second interpretant, and so on. Semiosis is fostered by the fact that the interpretant is in itself «a more developed sign». Hence, semiosis is a progressive process and, in principle, an endless process. As a matter of fact, at a personal and social level, the process reaches a stable state whenever a *habit* is established. A personal habit is acquired through two interlinked processes: a recurrent experiential relationship with the sign and the correlated semiosis. Nevertheless, a personal habit is never acquired in an isolated environment. A strong interaction with other actors (i.e. a social environment) may concur to the establishment of a habit, as is the case of a research team or a teaching class. Peirce remarks that «A habit is not an affection of consciousness; it is a general law of action»[24]. A habit corresponds to the establishment of a conventional behaviour: whenever the interpretant becomes a habit, his action is no longer limited to the cognitive sphere.
Here is a simple example of semiosis: in a laboratory, a flame test is performed and the appearance of a violet color is interpreted as a (macroscopic) sign of the presence of potassium in the sample. A well-trained student correlates the presence

of potassium with the occurrence of an emission process, the (microscopic) transition of an electron from a higher to a lower energy state in a potassium atom. So, the interpretant (potassium in the sample) becomes «a more developed sign», related with the excited states of potassium atoms. The process may even drift away from the scientific context and lead to the suggestion of violet as a fine color for clothes. The pragmatic aspect of semiosis is here very clear.

The effectiveness of a sign depends on the existence of a cooperation between the three elements of the triad, whose mutual relationship is dynamic. In his unpublished *A Survey of Pragmaticism*, Peirce wrote:

> «[B]y 'semiosis' I mean […] an action, or influence, which is, or involves, a coöperation of three subjects, such as a sign, its object, and its interpretant, this tri-relative influence not being in any way resolvable into actions between pairs. Σημείωσις in Greek of the Roman period, as early as Cicero's time, if I remember rightly, meant the action of almost any kind of sign; and my definition confers on anything that so acts the title of a 'sign'»[25].

Peirce stresses that the relationship is genuinely triadic and it does not result from the overlap of dual connections: «this tri-relative influence not being in any way resolvable into actions between pairs». In other words, meaning is the consequence of triadic relation of sign-object-interpretant (S-O-I) as a whole: «The triadic relation is *genuine*, that is, its three members are bound together by it in a way that does not consist in any complexus of dyadic relations».[26] It is noteworthy that the term |action| appears three times in this definition. In the same vein, Gérard Deledalle remarks that «The representamen, the object, and the interpretant stand for relations or functions»[27]. The term *relation* refers to the observer who perceives the sign; the term *function* refers to the making up of its meaning. This aspect is stressed by Floyd Merrell, who identifies the central point of Figure 1 (left panel) as a node: «a central 'node' [...] the fountain head of all sign *relata* and the locus of meaning in flux»[28]. Hence, the meaning comes out from a net of relations: «Meaning is in the interrelations, in the interaction, the interconnectedness»[29].

Peirce's Signs

Peirce's triad reported in Figure 1 (right panel) refers to the types of sign, and it is really crucial in the history of semiotics. It is worth reading Peirce's definitions with some comment. In a 1903 manuscript, Peirce wrote:

> «An *Icon* is a sign which refers to the Object that it denotes merely by virtue of characters of its own, and which it possesses, just the same, whether any such Object actually exists or not»[30] (italics in the text).

Two points should be emphasized. The first is that the icon refers to the object through the specific, graphical, visible and proper characters of the icon itself. Peirce often speaks of diagrams as typical examples of icons. The second point is that an icon may refer to a non-existing object: this means that the icon opens towards possible worlds. This highlights the 'arbitrary' and creative nature of the making of an icon, as we will argue further on, while discussing Eco's s-codes.

The image of the bycicle reported in Figure 1 is an Icon in that it has a similarity relationship with its object ("a sign by Firstness"[31]). Concerning the index, Peirce wrote:

> «An *Index* is a sign which refers to the Object that it denotes by virtue of being really affected by that Object. In so far as the Index is affected by the Object, it necessarily has some Quality in common with the Object and it is in respect to these that it refers to the Object»[32] (italics in the text)

In the laboratory, chemists create experimental situations in which precipitates of various colours and behaviours are produced, so that a chemist may deduce, for instance, the presence of Ag^{+} or Pb^{2+} ions in solution. In these cases, perceptible colours and behaviours are indexes that refer to specific microscopic events and particles. The image of the skull reported in Figure 1 is an Index in that chemists use it as a mark of toxicity for poisonous substances.

The last fundamental kind of sign is the following:

> «A *Symbol* is a sign which refers to the Object that it denotes by virtue of a law, usually an association of general ideas, which operates to cause the Symbol to be interpreted as referring to that Object»[33] (italics in the text)

The term ‘law’ remarks the conventional, inter-subjective, cultural nature of symbols. Namely, the number 9 and the term Country are symbols by virtue of very different conventions; nonetheless, both are very structured, and induce a specific behaviour. The Leviathan cross reported in Figure 1 is a Symbol as it used to stand for the element Sulphur in alchemy.
The three categories proposed by Peirce do not define disjoint sets of signs, otherwise many images found in scientific texts could not be interpreted according to Peirce's semiotics. It is evident that an image such as Figure 2 cannot be catalogued under just one out of the three fundamental categories of signs. Peirce was well aware of the question, and stated:

> «One sign frequently involves all three modes of representation; and if the iconic element is altogether predominant in a sign, it will answer most purposes to call it an icon»[34]

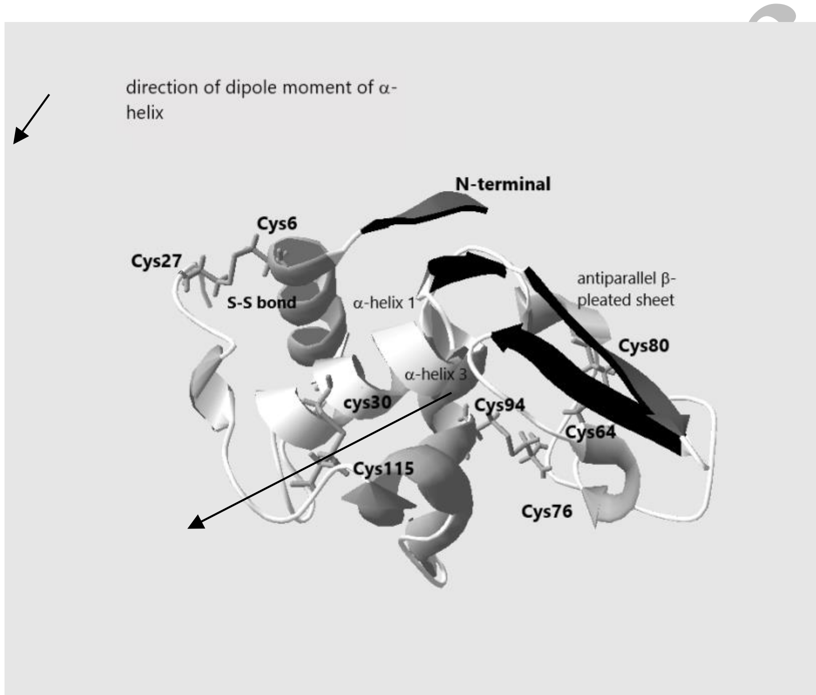


Figure 2. Representation of the 3D structure secondary and tertiary interactions of the enzyme lysozyme (PDB entry: 2VB1; image generated by Swiss PDB Viewer 4.1.0)

Figure 2 is just an example of the use of mixed signs. The different types of arrow mean distinct properties: the direction of the helix dipolar moment, the direction of beta-strands, etc. All of them have distinct implications in terms of interpretation of the behaviour of the object. The symbolic character is made evident by the presence of letters or writings that describe the nature of specific sites within the protein. The indexical character is evident in the arrows that are related with the dipole moment of helix, that is actually found in the object. Finally, the iconic character is given by the ensemble of this representation of a protein.
The two triads that we have analysed display a very different character. The first triad (representamen, interpretant, object) has a philosophical nature: it defines the sign and helps interpreting its function. Therefore, it is the basis for semiosis.
The second triad (icon, index, symbol) has systematic purposes: it helps classifying signs. Like any classification, it may be further detailed, depending on the features that need to be highlighted. Peirce himself proposed further classifications of signs and triadic diagrams: a first division in 10 classes (exposed in the Syllabus diagram, 1903) that subsequently engendered a system of 66 classes (Welby diagram)[35]. According to Romanini: «The 66 classes of signs arranged in […] triangular shape

show regular periods revealing the increase of complexity of semiosis as it reaches communication, as well as phases that describe the whole process of inquiry»[36]. In our context, we refer just to the first level of such classification.
Semiosis can be understood as the dynamic process of producing subsequent interpretations by the observer. The signs as such also possess an 'extrinsic' dynamics:

> «Symbols grow. They come into being by development out of other signs, particularly from icons or from mixed signs partaking of the nature of icons and symbols. We think only in signs. These mental signs are of mixed nature; the symbol-parts of them are called concepts. So it is only out of symbols that a new symbol can grow"»[37]

In addition, signs are not fixed: they evolve with time. The causes of such evolution are varied. Figure 3 shows two renown models, that marked key-steps in protein crystallography. The model in Figure 3 (left panel) was built when the group led by John Kendrew obtained data of myoglobin's structure at 6 Å-resolution. At the time, Kendrew's problem was to get myoglobin's structure at atomic resolution; in 1960 he was finally able to build up the model reported in Figure 3 (right panel), where each single atom of myoglobin (apart hydrogen atoms) has its 'place' in the 3D space.

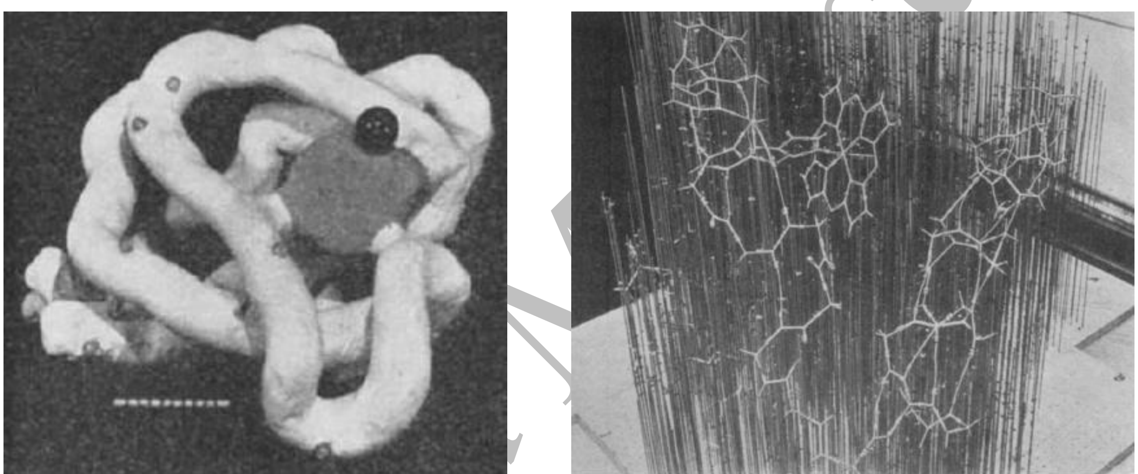


Figure 3. Kendrew's models of myoglobin's structure. Left panel: the model was made up with plasticine in 1957, at 6 Å-resolution. Right panel: the model of myoglobin at 2 Å-resolution, dubbed 'forest of rods', was built in 1960[38,39] (Reprinted by permission from Macmillan Publishers Ltd).

The models in Figure 3 are typical of cognitive situations in which the scientist is faced to different sets of data and builds up a sequence of models whose 'completeness' depends on the quantity and quality of data that the model aims to interpret. The two pictures in Figure 3 emphasize the 'materiality' of Kendrew's models, a materiality that does not prevent models from being perceived as signs. In a 1908 manuscript, Peirce refers to the monuments that are found in North American towns and villages, which portray a Union soldier of the Secession War. Peirce states: «That statue is one piece of granite, […] yet it is what we call a 'General' sign»[40]. A similar position on the iconic character of material objects was expressed in 1938 by Charles Morris in an important contribution to the theory of signs: «A photograph, a star chart, a model, a chemical diagram are icons, while the word 'photograph', the names of the stars and of chemical elements are symbols»[41]. Even though the information to be represented is the same, the mode of representation may be different and is subjected to evolution.
Figure 4 still refers to myoglobin and reports its 3D structure as a 'ribbon diagram' (or 'Richardson diagram'). Aim of this kind of model is to highlight structural features (secondary structures, the position of prosthetic group, etc.) that may help understanding the protein function.
In spite of referring to a same object (myoglobin), Figure 3 and 4 convey distinct messages and originates from different purposes.

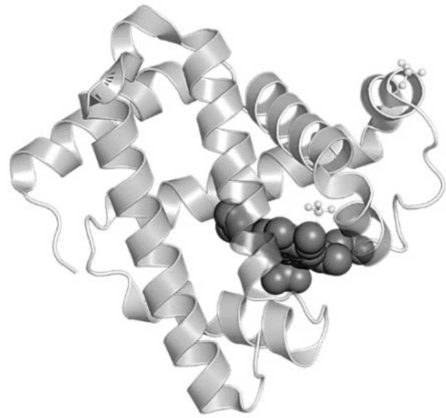


Figure 4. The 3D structure of myoglobin represented as a 'ribbon diagram' (PDB entry 2OH8; image generated by Swiss PDB Viewer 4.1.0)

s-codes vs. codes: Eco's lesson on signs

It is now worth recalling a classification proposed by semiologist Umberto Eco in 1976. It refers to the distinction between ‘s-codes’ and ‘codes’, that turns out to find a direct validation in the images that we are analysing. Eco defines three types of s-codes (where the ‘s’ stays for system): syntactic, semantic and behavioural. He also differentiates them from codes. In his own words:

(a) A syntactic system is «A set of *signals* [signs] ruled by internal combinatory laws»; «they could convey different notions about things and they could elicit a different set of responses»[42] (italics in the text).

(b) A semantic system is «A set of *notions* [...] which can become [...] a set of possible communicative contents»; «they could be conveyed by any [...] type of signal, such as flags, smoke, words, whistles, drums and so on»[43] (italics in the text).

(c) A behavioural system is «A set of possible *behavioural responses* on the part of the destination»; «these responses are independent of the (b) system». «They can also be elicited by another (a) system»[27] (italics in the text).

(d) A code: «A *rule* coupling some items from the (a) system with some from the (b) or the (c) system». «This rule establishes [...] that both the syntactic and the semantic units, once coupled, may correspond to a given response». «Only this complex form of rule may properly be called a *'code'* »[27] (italics in the text).

Systems (a), (b) and (c) share the property of being completely arbitrary and independent from each other. These features are relevant to the communication process because they leave the greatest possible freedom to those who have the task or the intention to communicate. A code assures a great freedom of choice over systems (a), (b) and (c), too.

Eco points out that the confusion between codes and s-codes can lead to «considerable theoretical damage». Thus, he introduces a sharp terminological distinction:

> «I shall therefore call a system of elements such as the syntactic, semantic and behavioral ones outlined in (a), (b) and (c) an *s-code* (or code as system); whereas a rule coupling the items of one s-code with the items of another or several others-codes, as outlined in (d), will simply be called a *code*»[44] (italics in the text).

In a nutshell, s-codes may be understood as |sets| and codes as |functions| that connect different sets. Or, with a less restricted definition, one can speak of s-codes as |rules of the game| and speak of codes as |game's styles|.

An example of the application of semantic, syntactic and behavioral s-codes and codes to the representation of protein structures may be found in Figure 5. The use of different colors or color shades, as well as continuous or dotted lines, pertain to the syntactic aspects that define the ‘pictorial alphabet’ on which the representation is based. The semantic s-codes disclose the interpretation-keys of the figure: e.g. dashed lines are H-bonds, black lines stay for aminoacids interacting with the heme group, etc. The behavioural s-code follows closely, as it defines how the picture must be read. It influences the reader's reaction in that the observer may formulate working hypothesis based on this picture and plan experiments aimed at verifying them. Another example of interpretation of chemical contents in terms of codes and s-codes is offered by chemical equations that chemists use to describe chemical transformations. In that case, the syntactic s-code is represented by the use of different symbols and combinations of symbols and numbers, to give chemical formulas and their combination (according to Jacob, «It is possible to distinguish between a chemical orthography and a chemical grammar»[7]). The semantic s-code lies in the interpretation rules of such formulas. The behavioural s-code allows reading that combinations of alpha-numerical signs in term of specific substances reacting on a molar bases. Coupling these s-codes result in the possibility, for the reader, to foresee the outcome of a chemical process carried out in specific conditions, or to design another process based on the premises posed by this one. This goes along with Wightman's analysis of chemical formulas, as he maintains that combining letters (chemical symbols) with numbers «changed the status of 'symbols' in the restricted sense from mere abbreviations […] into the elements of an 'algebra' and later a 'geometry' or 'topology'».[5] Other examples could be proposed: resorting on Eco's codes and s-codes is useful in that it allows highlighting the multifaceted character of the symbolic language through which chemistry describes its mental and practical operations.

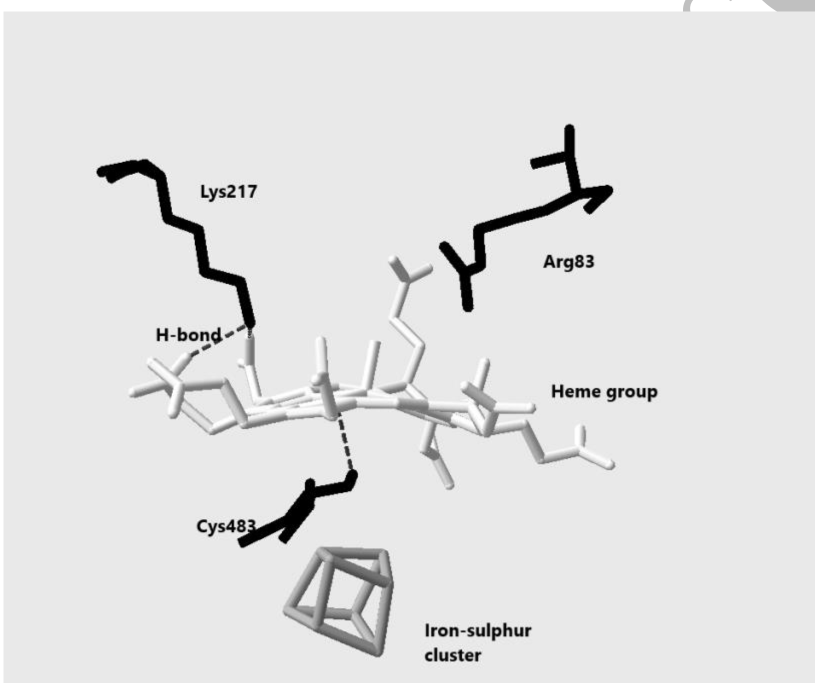


Figure 5. Representation of the active site of enzyme sulfite reductase
(PDB entry 1AOP; image generated by Swiss PDB Viewer 4.1.0)

In the perspective of Peirce's thought, protein icons and their related codes and s-codes foster abduction processes, as they play a role in the process of construction of scientific knowledge. The concept of abduction is probably the most important contribution given by Peirce to the philosophy of science. The distinction between abduction and induction or deduction is clearly explained:

> «Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea; for induction does nothing but determine a value, and deduction merely evolves the necessary consequences of a pure hypothesis»[45]

The relationship between abduction and icon is also well expressed by Peirce:

«An [...] Abduction, is an argument which presents facts in its Premiss which present a similarity to the fact stated in the Conclusion, [...] so that we are not led to assert the Conclusion positively but are only inclined toward admitting it as representing a fact of which the facts of the Premiss constitute an Icon»[46]

More concisely, a 1903 manuscript reports: «Abduction, or the suggestion of an explanatory theory, is inference through an Icon»[47]. Scientific research practices confirm the rightness of Peirce's thought, because the inspection of protein icons fosters the formulation of hypothesis that could not be disclosed by the mere examination of big sets of numerical data.

The Representation of Protein Molecules

An astonishing aspect of protein icons is their ability to represent what is not representable: they are an attempt to visualize microscopic objects and chemical-physical properties. One could say that these pictures *make abstractions visible*. Let's take the example of potential surfaces: they do not materially exist and are the transposition, in pictorial terms, of sets of numbers related with a physical property. Should these numbers be arranged in a table, the reader would not be able to place them within a horizon of meaning. The switch to iconic representations allows providing numbers with an operational meaning, that fosters the formulation of further hypothesis and the planning of investigations. The switch from numbers to icons conceals the use of analogy as interpreting tool: iconic language allows representing channels, grooves, surfaces, shapes, etc... These terms, borrowed from natural language, clearly refer to the macroscopic realm and, taken as such, could not apply to the microscopic world. Nevertheless, resorting to analogies discloses the impressive heuristic potential of icons. In fact, based on these images the reader may formulate hypothesis about the possibility of hosting a molecule inside a 'protein pocket' or to make substrate-protein contacts mediated by a specific charge distribution on the 'protein surface'. In other words, analogies inherent to iconic representations of properties and features of the microscopic world allow handling an otherwise inaccessible realm and interpreting its behavior.

The application of s-codes to the making of a pictorial representation responds to both inter-personal communication and interpretation purposes. The inter-personal communication is achieved when distinct people in distinct places read the picture in a same way. This implies the definition of conventional s-codes, that are employed in the making of the figure and are approved by a disciplinary community. Interpretation purposes are directly related with behavioral s-codes: the icon makes evident what numbers would conceal, e.g. the possibility to accommodate a substrate within a pocket, to track an internal electron-transfer pathway, to make adducts with other molecules, etc..

The analysis of signs according to Eco's s-codes emphasizes the conventional character of signs, and therefore, according to Peirce, their symbolic aspect. A relevant feature of symbols is that their meaning is multifarious. Peirce remarks that:

«The entire intellectual purport of any symbol consists in the total of all general modes of rational conduct which, conditionally upon all the possible different circumstances and desires, would ensue upon the acceptance of the symbol»[48]

A similar viewpoint is echoed by a passage of Roland Barthes's writings:

«It follows that the meaning of a text lies not in this or that interpretation but in the diagrammatic totality of its readings, in their plural system», «The meaning of a text can be nothing but the plurality of its systems, its infinite (circular) 'transcribability'»[49]

The reference to 'the total of all general modes of rational conduct' in Peirce's excerpt, just as Barthes' «diagrammatic totality of its readings», underlines that a symbol has a multiplicity of interpretations that are in dialogic relationship between each other. In the language of semiotics, one can say that there are multiple codes and s-code for each symbol. Such multiplicity discloses different aspects of the symbol's object. Hence, «Different systems of expression are often of the greatest advantage»[50], but the actual nature of the object to which the symbol refers remains somehow unattainable.

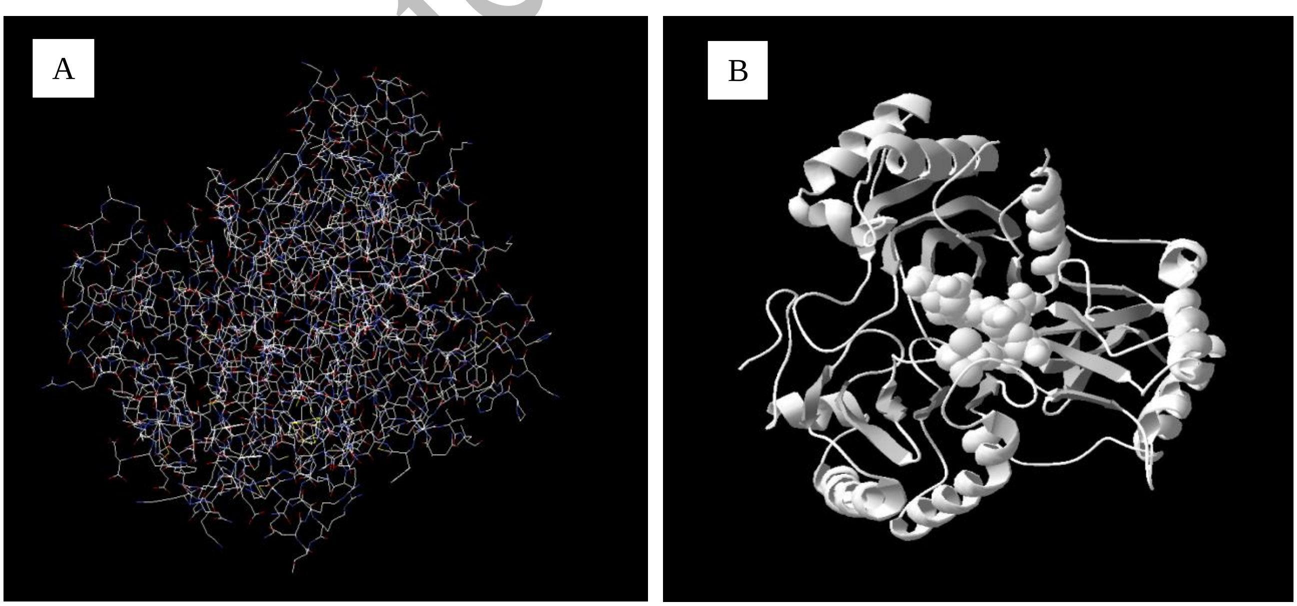
The set of iconic representations in Figure 6 are a good exemplification of Peirce's and Barthes's statements: all these images refer to a same object, a same protein molecule. Each of them provides a different view of the system, depending on the properties that the designer wishes to highlight. Some of them provide structural hints, by highlighting secondary and tertiary structures; some others suggest functional or relational behaviors, by highlighting regions of distinct charge density or the presence of acid or basic sites. Interestingly, the less useful representation is the one that contains the higher amount of the structural information (panel A), as it reports the position of all atoms in the structure and implies no data selection. The comparison between distinct representations shows quite clearly that data selection is essential to the assignment of meaning. Namely, secondary structures become evident only by discarding information about lateral chains of aminoacids; otherwise, any possibility of spotting features like crucial interactions within the structure or stable domains, is missed.

Figure 6 is also representative of distinct syntactic and semantic s-codes applied to a same microscopic object. Here it becomes clear that s-codes are the base for representation, whereas codes are the base for interpretation and depend strictly on the purpose of the modeler. So, this figure discloses very clearly the role of the modeler's intentionality: depending on the modeler's aim, a different set of s-code is used and a different code is generated. Depending on one's interest in structural stability aspects or in understanding the way a protein may interact with another protein, one may choose to represent the system in distinct ways.

The peculiar epistemic value of icons is well expressed by this excerpt of Peirce's writings:

> «[A] great distinguishing property of the icon is that by the direct observation of it other truths concerning its object can be discovered than those which suffice to determine its construction. [A] capacity of revealing unexpected truth»[51]

This is strongly related with the previously mentioned heuristic potential of icons. According to Peirce, the mere observation of an icon has the power of disclosing truths about the object, truths that escaped even to the icon's maker. This statement is amazing in several ways. As we have seen, Peirce clearly assigns to icons an important role in abduction processes: they are tools in the construction of (scientific) knowledge. We already remarked that the inspection of an icon representing a protein model allows disclosing the presence of binding sites and regions of structural stiffness or plasticity, the ability to establish specific interactions with other molecules through superficial contacts, etc…Interestingly, these aspects are all enclosed in the set of numerical data that are needed for building up the pictures. But the heuristic value belongs to the pictorial representation and not to the raw sets of data: it does not belong to numbers, it belongs to the iconic representations of subset of numbers, chosen by the modeler's intention. Numbers have clearly no inherent meaning: they need a hermeneutic action, capable of putting them inside a specific frame that provides them with a meaning. The cognitive potential does not lie in numbers; it rather belongs to their various representations that, in turn, are the result of specific choices of the icon's designer.[52]



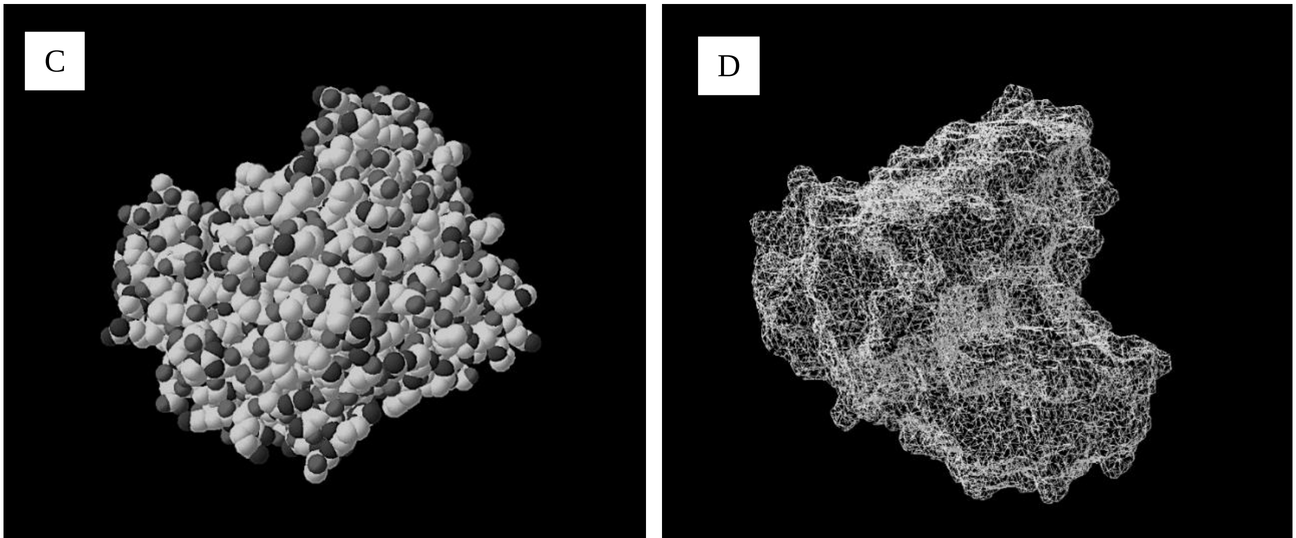


Figure 6. Representations of the 3D structure of enzyme Sulfite reductase: A) CPK representation; B) ribbons representation; C) Van der Waals radii representation; d) protein surface representation (PDB entry 1AOP; image generated by Swiss PDB Viewer 4.1.0)

This leads to discussing Peirce's view on the process of building up of scientific knowledge. Pierce's analytical mind established a classification of epistemic shades, based on the subtle distinction between presumption, guess, conjecture, surmise: «An increase of information by induction, hypothesis, or analogy, is a *presumption*. [...] A very weak presumption is a *guess*. A presumption opposed to direct testimony is a *conjecture*, or, if weak, a *surmise*»[53]. Based on his own words, scientific knowledge aims at ascertaining the truth, but this truth is not the mere result of strictly logical operations. A good deal of creativity lies in what Pierce calls abduction:

> «it must be admitted that the only method of ascertaining the truth is to repeat this trio of operations: conjecture *[the abductive hypothesis];* deductions of predictions from the conjecture; testing the predictions by experimentation (not necessarily what is technically so called, but essentially the same thing -- trial)»[54]

Two important aspects of Peirce's thought emerge from this two excerpts from Peirce's writings: i) the definition of conjecture, intended as a presumption that does not rely on direct testimony; ii) the role of conjectures in the abduction process. According to Peirce's thought, any microscopic object, an atom, a molecule, a subatomic particle, that cannot be the object of direct testimony is a conjecture. So, the basic conceptual tools of chemistry, taken as a science that explains macroscopic phenomena through the conceptualization of a microscopic level, are conjectures. This is true for proteins' structural models as well. Conjectures are put at the centre of the abduction process. In the construction of scientific knowledge, abduction comes first: it foregoes deduction or induction, which are way of relating a hypothesis with an experimental evidence. Instead, abduction deals with the way hypothesis takes shape. Peirce closely links the epistemological process of abduction to the physical one of visual perception: «abductive inference shades into perceptual judgment without any sharp line of demarcation between them»[55]. He remarks that «Abduction furnishes all our ideas concerning real things, beyond what are given in perception, but is mere conjecture, without probative force»[56]. This implies that researchers represent molecular systems projecting shapes and entities belonging to their own macroscopic ontological level (balls, ribbons, channels, pockets, etc.) to the atomic-molecular level.

The main consequence of this shift in focus is to emphasize that the basis of scientific knowledge is not pure rationality, but rather creativity. Investigators of a given system or phenomenon need *to imagine* relationships and behaviors, whose plausibility will subsequently be proved or disproved through properly built experiments. This resorting to something that does not belong exclusively to the rational sphere is found in both experimental and theoretical practices: «The role of the imaginary and the fictional in chemistry […] exceeds that found in any other science».[9]

Based on the premises of Peirce’s thought about conjectures and Eco’s thought about s-codes and codes, I believe that proteins’ icons can be taken as conjectures in a double sense:

i) they are conjectures on the representability of the microscopic realm, that is not accessible to direct testimony. They are attempts to make visible the invisible (or the abstract). In proteins’ icons, we find visual representation of a number of physical properties and features that have a purely abstract character and can be expressed by numbers.
ii) they are conjectures in a more strictly epistemic sense, in that they are tools of the abductive process. Thanks to iconic representations, one can formulate hypothesis on the stability, function, behaviour and/or mechanism of the microscopic object that is represented.

Two further aspects of conjectures that may be related with Eco’s thought on codes and s-codes are formulations and effectiveness. Formulation of a conjecture implies the use of a set of s-codes, that define ‘the rule of the game’. S-codes are tools, whereas the code employed in handling the conjecture is the final aim. It is the code that leads to the formulation of hypothesis.
In summary, I tried to use some elements of the semiotic theory by Peirce and Eco’s theory of codes as tools for carrying out an epistemic analysis of visual representations of proteins and the way these icons are employed by researchers.
My conclusion is that protein icons, and more generally, icons referring to microscopic objects that are part of the chemical explanation of phenomena, are conjectures (in Peirce’s acceptation).
Other relevant aspects of the use of icons of the molecular world are the following: i) the assignment of meaning requires selection of data: big dataset do not possess an inherent meaning; ii) iconic representations make visible what cannot be visualized, either because microscopic or abstract. In doing so, they contribute to the construction of reality: “If we think of chemical signs as nothing more than ‘fingers’ pointing to a predetermined reality, we slight their unique and invaluable creative functions”[9]; iii) icons possess a heuristic potential, that make them instruments of the abduction process. As regards this latter aspect, I’d like quoting the motivation provided by the Royal Swedish Academy of Sciences to the assignment of the 2017 Nobel prize for Chemistry to the reserchers who developed cryo-electron microscopy: «A picture is a key to understanding. Scientific breakthroughs often build upon the successful visualisation of objects invisible to the human eye»[57]. It is doubtless that the impressive advances in the understanding of the behaviour and structure-function relationships of biomolecules occurred in the last decades owe something to iconic representations.
In conclusion, I believe this reflection on the role of iconic representations in scientific investigations, in the light of Peirce’s and Eco’s thought, provide some hints on the mental processes that chemists operate when they propose explanations at the microscopic level for phenomena observed in the macroscopic realm.

Acknowledgments
The author acknowledges support from the Project CH4.0 under the MUR program “Dipartimenti di Eccellenza 2023-2027” (CUP: D13C22003520001)

Bibliography

1 The present paper is dedicated to the memory of Luigi Cerruti, with whom the author shared long discussions concerning the issues addressed in this work; part of this work was presented by Cerruti and myself at the 21^{st} conference of the International Society for the Philosophy of Chemistry (Paris 2017).

2 J. Dalton, *A New System of Chemical Philosophy*, Bickerstaff, Manchester (UK), **1808**, opposite p. 219. The plate can be seen at URL [https://library.si.edu/digital-library/book/newsystemofchemi11dalt] Last accessed on March 5^{th} 2024.

3 A. H. Johnstone, *Why is Science Difficult to Learn? Things Are Seldom What They Seem., J. Comp. Assist. Learn*. **1991**, 7, 75-83.

4 M.P. Crosland, *Historical Studies in the Language of Chemistry* **1962**, Harvard University Press, Cambridge (MS).

5 W.P.D Wightman, *Essay Review: The Language of Chemistry, Annals of Science* **1963**, 19, 259-267.

[6] B. Bensaude-Vincent (**2002**). Languages in chemistry. In M. J. Nye (Ed.), *The Cambridge history of science* (Vol. 5, The modern physical and mathematical sciences) (pp. 174-190).

[7] C. Jacob, *Analysis and Synthesis: Interdependent Operations in Chemical Language and Practice, HYLE - Int. J. Phil. Chem.* **2001**, 7, 31-50.

[8] P. Laszlo, *Chemistry, Knowledge Through Actions?, HYLE - Int. J. Phil. Chem.* **2014**, 20, 93-116.

[9] S. Weininger, *Contemplating The Finger: Visuality and the Semiotics of Chemistry, HYLE - Int. J. Phil. Chem* **1998,** 4, 3-27.

[10] E. Francoeur, *The Forgotten Tool: The Design and Use of Molecular Models, Social Studies of Science* **1997**, 27, 7-40.

[11] E. Francoeur, *Cyrus Levinthal, the Kluge and the origins of interactive molecular graphics, Endeavour* **2002,** 26, 127-131.

[12] U. Klein *, Paper Tools in Experimental Cultures, Stud. Hist. Phil. Sci.* **2001**, 32, 265-302.

[13] R. Mestrallet, *Communication, Linguistique et Sémiologie. Contribution à l'étude de la sémiologie. Études sémiologique des systèmes de signes de la chimie*, **1981,** Bellaterra, Barcelona, p.21.

[14] R. Hoffmann, P. Laszlo, *Representation in Chemistry, Angew. Chem. Int. Ed* **1991**, 30, 1-16.

[15] R. Khanfour-Armalé, J.-F.Le Maréchal, *Représentations moléculaires et systèmes sémiotiques, Aster*, **2009,** 48, pp. 63-88.

[16] M. Battersby, *Critical Thinking as Applied Epistemology: Relocating Critical Thinking in the Philosophical Landscape, Informal Logic* **1989**, 11, 91-100.

[17] M. Battersby, *Applied Epistemology and Argumentation in Epidemiology, Informal Logic* **2006**, 26, 41-62.

[18] L. Laudan, *Truth, Error, and Criminal Law, An Essay in Legal Epistemology*, Cambridge University Press, Cambridge **2006**, p. 2.

[19] C.S. Peirce, *Writings of Charles S. Peirce: A Chronological Edition*, Vol.8 (1890–1892), Indiana University Press, Bloomington, **2010**, p.21.

[20] Robert Marty, University of Perpignan, collected and analyzed 76 different definitions of sign found in Peirce's texts, from 1865 to 1911: R. Marty, *76 Definitions of The Sign by C.S.Peirce*, **2012**, [http://perso.numericable.fr/robert.marty/semiotique/76defeng] Accessed March 5th 2024.

[21] CP 2.228 **1897**; the notation CP [volume#.paragraph#] refers to *Collected Papers of Charles Sanders Peirce*, vol. 1-6 (Eds.: C. Hartshorne, P. Weiss); vol. 7-8 (Ed.: A. Burks), Harvard University Press, Cambridge (MS), **1931-35, 1958**. Chronology added, following: *Schriften von Charles Sanders Peirce*, [https://de.wikipedia.org/wiki/Schriften_von_Charles_Sanders_Peirce]. Accessed March 5th 2024.

[22] F. Merrell, *Peirce, Signs, and Meaning*, University of Toronto Press, Toronto, **1997**, p.13.

[23] CP 5.484 **1907**

[24] CP 2.148 **1902**

[25] CP 5.484 **1907**

[26] *The essential Peirce, Selected philosophical writings,* Vol. 2, (N. Houser, J.R. Eller, A.C. Lewis, A. De Tienne, C.L. Clark, D.B. Davis, eds), Indiana University Press, Bloomington, **1998**, p.274.

[27] G. Deledalle, *Charles S. Peirce's Philosophy of Signs: Essays in Comparative Semiotics*, Indiana University Press, Bloomington, **2000**, p.18.

[28] Merrell **1997**, op. cit., p. 26.

[29] Merrell **1997**, op. cit., p. xii.

[30] CP 2.247 **1903**

[31] CP 2.276 **1902**

[32] CP 2.248 **1903**

[33] CP 2.249 **1903**

[34] MS 491 **1903**, quoted from D.R. Anderson, *Creativity and the Philosophy of C.S. Peirce*, Nijhoff, Dordrecht, **1987**, p. 69.

[35] Discussion of these classifications is beyond the scope of this article. For further insight, see P. Farias, J. Queiroz, *Notes for a dynamic diagram of Charles Peirce's classifications of signs, Semiotica* **2000**, 131, 19-44; P. Farias, J. Queiroz, *On diagrams for Peirce's 10, 28, and sixty-six classes of signs, Semiotica* **2003**, 147, 165–184; P. Borges, *79. Peirce's System of 66 Classes of Signs. Charles Sanders Peirce in His Own Words: 100 Years of Semiotics, Communication and Cognition* (T.Thellefsen and B. Sorensen, eds.), De Gruyter Mouton, Berlin-Boston, **2014**, pp. 507-512.

[36] V. Romanini, *Minute semeiotic: The Periodic Table of Classes of Signs*, July **2014**, DOI: 10.13140/2.1.4753.6327

[37] CP 2.302 **1893**

[38] J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, *A three-dimensional model of the myglobin molecule obtained by X-ray analysis*, *Nature* **1958**, 181, 662-666.

[39] J.C. Kendrew, R.E. Dickerson, B.E. Strandberg, R.G. Hart, D.R. Davies, *Structure of myoglobin. A three-dimensional Fourier synthesis at 2Å resolution*, *Nature* **1960**, 185, 422-427.

[40] CP 8.357 **1908**

[41] C.W. Morris, *Foundations of the Theory of Signs*, in *International Encyclopedia of Unified Sciences*, Vol. 1(2) (Ed.: O. Neurath), University of Chicago Press, Chicago, **1938**, p. 24.

[42] U. Eco, *A Theory of Semiotics*, Indiana University Press, Bloomington, **1976**, p. 36.

[43] Eco **1976**, op. cit., p. 37.

[44] Eco **1976**, op. cit., p. 38.

[45] CP 5.171 **1903**

[46] CP 2.96 **1902**

[47] C.S. Peirce, *Pragmatism as a Principle and Method of Right Thinking. The 1903 Harvard Lectures on Pragmatism*, (Ed.: P.A. Turrisi), State University of New York Press, Albany (NY), **1997**, p. 176.

[48] CP 5.438 **1905**

[49] R. Barthes, *S/Z*, Blackwell Publishing, New York, **1974**, p. 120.

[50] CP 2.222 **1903**

[51] CP 2.279 **1897**

[52] Incidentally, these thoughts have clear implications regarding the current debate on Big Data and the ability of Artificial Intelligence to disclose meanings. AI spans huge amounts of data, but cannot disclose the symbolic content of an icon. Even further, AI is not capable of abductive reasoning.

[53] CP 2.430 **1893**

[54] CP 7.672 **1903**

[55] CP 5.181 **1903**

[56] CP 8.209, **1905**

[57] Royal Swedish Academy of Sciences, *Press release: The Nobel Prize in Chemistry 2017*, **2017** [https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2017/press.html] Accessed March 5th 2024.