

# On evolution of thinking about semiosis: semiotics meets cognitive science

**Piotr Konderak**

Department of Logic and Cognitive Science  
Maria Curie-Skłodowska University in Lublin  
*kondorp@bacon.umcs.lublin.pl*

## Abstract

The aim of the paper is to sketch an idea—seen from the point of view of a cognitive scientist—of cognitive semiotics as a discipline. Consequently, the article presents aspects of the relationship between the two disciplines: semiotics and cognitive science. The main assumption of the argumentation is that at least some semiotic processes are also cognitive processes. At the methodological level, this claim allows for application of cognitive models as explanations of selected semiotic processes. In particular, the processes of embedded interpretation (in contrast to interpretability in principle) are considered: belief revision, dynamic organization of meaning and metaknowledge. The explanations are formulated in terms of artificial cognitive agents of the GLAIR/SNePS cognitive architecture. Finally, it is suggested that even if someone rejects the idea of artificial cognitive systems as simulations of semiotic processes, they may acknowledge the usefulness of cognitive modeling in analysis of semiotic processes in virtual, simulated worlds and in the area of “new media”.

**Keywords:** cognitive semiotics; cognitive science; computational modeling; semiosis; Peirce

## 1. Setting the scene

Cognitive Semiotics can be defined as an interdisciplinary matrix of disciplines and methods, focused on the multifaceted phenomenon of meaning or as an emerging field with the ambition of

...integrating methods and theories developed in the disciplines of cognitive science with methods and theories developed in semiotics and the humanities, with the ultimate aim of providing new insights into the realm of human signification and its manifestation in cultural practices. (Zlatev 2012)

As a newly developed discipline, cognitive semiotics is intended to be the study of meaning and meaning-making activity. The Lund School in cognitive semiotics (Zlatev 2012, Sonesson 2012) highlights the role of phenomenological studies and the role of enactive cognitive science. Here I will explore a somewhat different approach, suggesting that it is fruitful to apply methods of cognitive science (mainly cognitive modeling) in studies on sign-usage, meaning and meaning-making.<sup>21</sup>

### **1.1. Semiotics**

In the eyes of a cognitive scientist, semiotics is an inhomogeneous field of study. “There are many ways to think of meaning and the possibility and impossibility of studying it”. (Brandt 2004: 257) The proliferation of differing approaches to the phenomenon of meaning makes cooperation between researchers in semiotics and cognitive science somehow difficult. To disambiguate the notion of meaning and meaning-making, I will formulate claims with a particular semiotical theory in mind. In my opinion, the philosophical (mainly epistemological) consequences of Peircean theory of signs make the theory especially appropriate for an analysis of cognitive-semiotic phenomena. Therefore, I will rely on the Peircean version of semiotics, where “a sign or representamen is something that stands to somebody for something in some respect or capacity. [...] The sign stands for something, its object. It stands not in all respects, but in reference to a sort of idea which I have sometimes called the ground of the representamen” (CP 2.228). In consequence, the notions of a representamen (in particular: a *symbol*), its *interpretation*, and the *respect* in which the representamen stands for something else are of special importance. Meaning, in turn, arises in the process of interpretation of the sign. Instead of a dyadic relationship between a sign and its object, we have here a triadic relationship involving interpretation as the third element. Finally, the reference to “somebody” in the above quotation indicates that it is necessary to take the interpreter of the sign into account.<sup>22</sup>

In this context, I regard dynamicity of semiosis ('action of signs') as a crucial feature of semiotic activity: “By semiosis I mean ... 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”. (CP 5.484).

Even within the Peircean approach, the processes of creation, use and (re-)interpretation of signs are—from the point of view of cognitive science—underspecified. Semioticians formulate their explanations in very general

---

<sup>21</sup> In that sense, my proposal is in line with the vision of a marriage of cognitive science and semiotics, as suggested by U. Eco (1999).

<sup>22</sup> However, it is important not to identify and confuse an interpretant with an interpreter.

categories and in quite broad frameworks that it seems can be supplemented by results of research on cognitive systems.<sup>23</sup>

### **1.2. A perspective of cognitive science**

Cognitive science seeks an explanation in more specific, functional terms, i.e. in terms of plausible psychological, neurobiological or/and computational mechanisms responsible for a given cognitive ability. For instance, analyzing a process of interpretation of a sign considered as a cognitive process cognitive scientists may deliver functional explanations of some aspects of the process: when it occurs (e.g. in the case of contradiction), how it proceeds (e.g. backtracking), or why it runs this way. One method that allows for such explanations is cognitive modeling, a basic method in cognitive science. Cognitive models are in fact neural or psychological theories of human cognitive activity expressed in terms of computer programs and implemented as executable procedures (Anderson 1993, Konderak 2005). This method usually involves creation of artificial cognitive agents (henceforth ACAs), characterized by the ability to construct and process representations of objects, states of affairs, or situations. Nowadays, the construction of ACAs is facilitated by particular cognitive architectures such as ACT-R (Anderson), Soar (Laird, Newell, Rosenbloom) or GLAiR (Shapiro, Rapaport).

In sum, I want to indicate the possibility that studying some artificial systems (i.e. cognitive models treated as simulations; see below) may be fruitful for the purposes of semioticians. In other words, I argue that cognitive modeling may also be a useful method of discovering properties of semiotic systems or processes. Creation of ACAs as semiotic systems may deliver a better understanding of the semiotic processes as described by stoics, Saussure, Peirce etc. What is more, it is also possible that one cannot create a credible model of a cognitive system unless one considers it as a sign-using system.

The very idea of a functional approach to semiosis is a sign of an evolution of thinking about semiotics from methods typical of humanities and philosophy (e.g. phenomenological methods, Sonesson 2012), to experimental methods as are used within natural sciences, including cognitive science.

The paper is structured as follows. Section 2 presents the idea of cognitive interpretation of semiosis (on the basis of Peircean theory of signs). Cognitive reading of semiosis suggests application of the methods of cognitive science to cognitive semiotics. Section 3 briefly presents various interpretations of the general relationship between computational models and semiosis, and the simulational approach is suggested as being relevant for purposes of cognitive

---

<sup>23</sup> As Allen Newell declared during his last lecture: "The answer must have the details. I have got to know how the gears clank and how the pistons go and all the rest of that detail. My question leads me down to worry about architecture". (Anderson 2007: 3-4)

semiotics. Section 4 focuses on the methodology of cognitive science; in particular, on the notions of cognitive architectures and cognitive models. Criteria allowing for evaluation of cognitive models are presented. Section 5 is a presentation of four features of artificial cognitive agents which are necessary in the context of semiotic activity: the ability to (actually) interpret signs, the dynamic character of semiosis, the ability to resolve contradictory interpretations, and metacognitive capacities. The section ends with a presentation of a computational model of vocabulary acquisition as an instantiation of cognitive-semiotic cognitive agents. Finally, the role of semiotic activity of artificial cognitive agents in virtual reality is mentioned.

## **2. Between semiotics, cognition and models**

As declared above, an interpretation of Peircean philosophy and semiotics is a promising starting point to show the connections between cognition and semiosis. Numerous quotations from Peircean writings indicate a relationship between the two: for example, “All thought is in signs” (CP 2.213) and in consequence, all thoughts are “thought-signs”; “we think only in signs” (CP 2.302). Leaving aside doubts concerning ambiguous (from the point of view of contemporary cognitive science) words like “thought” or “thinking”, we may conclude that at least some cognitive processes (called here “thinking”) are also semiotic activities; therefore, it seems that by studying cognition we can elicit some properties of semiotic structures and processes. In a similar vein, Eco (1999: 8) notices: “And it does not displease me that semiotics has come to be included in this confederation [of cognitive sciences - PK], independently of the question (still debated) whether semiotics is a cognitive science or cognitive sciences are a branch of semiotics. [...] I have stressed the link between semiotics and cognitive sciences [...]—its cognitive orientation should always be taken in serious consideration”.

There are two assumptions which underlie the idea of integrating semiotics and cognitive science:

1. cognitive modeling or creation of artificial cognitive agents is a fruitful method of investigating cognitive structures and processes. The assumption is a key methodological claim within cognitive science;
2. semiotic processes are in fact cognitive processes. For now, I leave open the question of whether all semiotic processes are cognitive processes, or only some of them. Even if only some of them are cognitive, we can still gain knowledge about the nature of semiosis by studying appropriate cognitive activities.

These two assumptions strongly suggest that cognitive modeling using some kind of cognitive architecture may be a successful methodology in the investigation of semiotic processes. This suggestion is in line with Zlatev's (2012: 14) declaration concerning the methodology of cognitive semiotics. This discipline should help to integrate humanities and science by way of first-person, second-person and third-person "triangulation of methods". First- and second-person methods are characteristic of humanities and are supplemented by the third-person, or objective methods typical of science (in particular for cognitive science). In other words, although "our knowledge of the world, including our scientific knowledge, arises from a first-person perspective" (Gallagher & Zahavi 2008: 89), it requires justification in the form of "objective" methods such as psychological experiments, brain imaging or cognitive modeling. In this context, Zlatev stresses the validity of all the methods in the area of studies on meaning and meaning-making. The paradigm—an often-quoted example of such triangulation—is neurophenomenology (Varela 1996), an approach that starts with phenomenological analyses of experience and supplements them with results of neuroimaging. I treat cognitive modeling as a form of third-person explanation of selected cognitive processes involving semiotic activity.

### 3. Between computers and semiosis

Certain contemporary artificial systems—called *symbolic machines* or *sign-vehicles*—seem to be an intriguing and promising object of semiotic research. In consequence, several approaches and a number of perspectives have emerged to account for the intersection between semiosis and computer models (Clarke 2001, Mehler 2003). These approaches reflect the evolution of thinking about artificial symbol processors as semiotic systems.

- *Computer Semiotics*, as Andersen (1990) presents it, is a framework for understanding and designing computer systems as sign systems - computers are "targets of interpretations". The focus of this approach is a "symbolic machine constructed and controlled by means of signs" and semiotics delivers useful tools of interpretation. Computer semiotics studies the special nature of computer-based signs and how they function in use.<sup>24</sup> Such an approach—although interesting—does not elicit the properties of the process of semiosis, i.e. the process of creation and interpretation of signs. In the case of Andersen's semiotic computer systems, semiotic activity is on the side of the external interpreter.

---

<sup>24</sup> In a similar vein Fetzer (1997) tries to combine semiotic research with Artificial Intelligence and cognitive science. He analyses "computers" as special kinds of signs, which may be interpreted as symbols, icons or indexes.

- *Computational Semiotics* is defined by Gudwin (1999) in the following way: “Computational Semiotics corresponds to the proposition of a set of methodologies that in some way try to use the concepts and terminology of semiotics”. Just as in the case of the engineering branch of Artificial Intelligence, the aim of the approach is the construction of autonomous artificial agents; therefore, the goal is mainly practical. As the author states: “We are looking at semiotics in order to build more intelligent systems”. I would call the approach *applied semiotics* (the use of semiotic notions within the domain of AI research). The approach is of minor interest for the semiotic as well as the cognitive semiotic community.
- Finally, the idea presented in the paper is more limited and more semiotics-oriented: it could be called *simulational study of meaning*. Here, the key problems are not constructional (as Gudwin suggests), but conceptual. I use functioning ACAs to discover the properties of dynamic sign systems in order to suggest some processes that lie behind semiotic activity and to clarify the meaning of such terms as representation, interpretation (and reinterpretation), synechism or fallibilism. Just like the COSINE approach (“computational semiotics in the narrow sense”), we “do not realize semiotic systems, but simulate their procedural organization”. (Mehler 2003: 76)<sup>25</sup> *Simulation* in this context (explaining the relationship between semiotic processes and ACAs in the sense of the third approach, above) may be defined as formal descriptions (models) of the processes or objects modeled. (Mehler 2003: 74). More precisely, simulations may be understood as dynamic, procedural models of selected functions of modeled entities. We experiment with a simulation in order to understand the system being modeled.

The distinctions listed above demonstrate that the relationship between semiotics and research on “computer” models (and, indirectly, cognitive science) has evolved from Fetzer's (1997) and Andersen's (1990) notion of a computer as a sign or target of interpretation, via Godwin's approach, to the notion of an artificial cognitive agent as a simulation of a semiotic system, i.e. a system that is capable of interpreting, creating and using signs.

---

<sup>25</sup> The approaches sketched above coincide somehow with the distinction between the three processes: realization (Andersen's computer semiotics), emulation (Gudwin computational semiotics) and simulation.

#### **4. On methodology: architectures and models**

Cognitive semiotics as a field of research rests on integration (in Zlatev's words: triangulation, see section 2) of methods. Although cognitive semiotics gives priority to first-person methods (for example conceptual analyses, phenomenological reduction) and second-person methods (empathy, imaginative projection), these methods should be also supplemented by third-person methods (like psychological experiments, brain imaging or cognitive modeling) (cf. Zlatev 2009: 178). In this context, the role of cognitive modeling is to provide third-person methods that are applied to data obtained by other first-, second-, and third-person data. Semiotic analyses may start with phenomenological analyses of the experience of a sign-user (cf. Sokolowski 2000), for example analysis of differentiation between expression and content from the standpoint of a subject (Sonesson 2009, Piaget 1975). The models may work using third-person data, or as various linguistic or semiotic corpora.

##### ***4.1. Cognitive architectures and cognitive models***

In contemporary "cognitivist" cognitive science, researchers often design their models within so-called cognitive architectures (Anderson 1983). A cognitive architecture (the term was introduced into cognitive science by Allan Newell) is an answer to the question of how a structure achieves certain functions (cf. Anderson 2007: 5). In Anderson's words, a cognitive architecture is "a theory of the basic principles of operation built into the cognitive system" (Anderson 1983: ix, 2007: 5). A cognitive architecture specifies a number of mental modules and the interconnections and principles of communication between them (i.e. structure). More precisely, cognitive architectures offer representations and cognitive mechanisms responsible for our cognitive activity. In this sense, a cognitive architecture may be considered a theory of cognition: it implements or realizes basic assumptions on cognition as formulated within cognitive psychology and/or cognitive neuroscience.

Cognitive architecture allows simulations of cognition to be designed (cf. section 3). As stated above, such simulations take the form of cognitive models. In other words, a cognitive architecture is a general framework for particular models simulating concrete cognitive abilities or capacities.

A single cognitive architecture may support several cognitive models of a given cognitive phenomenon. In this context, an important question about assessing the validity of cognitive models arises. Researchers also face the problem of relative evaluation of competing cognitive models of the same cognitive phenomenon.

The most obvious criterion is based on the similarity of a model's behavior in comparison to the cognitive agent being modeled (human being).<sup>26</sup> However, as suggested above, one may provide several different models producing the same (or similar) behavior. Consequently, the scientific practice within the field of cognitive modeling suggests the following, additional criteria (cf. Taatgen & Anderson 2008):

- a model should not just describe behavior, but should also predict the future behavior of an agent;
- a model should be able to acquire (learn) specific knowledge;
- researchers should limit arbitrariness in their models; they should introduce as few free parameters (parameters that can be associated with arbitrary values) as possible.

#### **4.2. A short overview of cognitive architectures**

Currently, the following three types of cognitive architectures have been developed: purely symbolical architectures (Soar, SNePS/GLAIR); connectionist architectures, based on distributed representations; and hybrid architectures involving both symbolic and distributed representations (ACT-R, Clarion).

The Soar architecture is designed to model problem-solving activities in terms of a goal state, an end state, as well as problem spaces. The architecture assumes a single long-term memory, a very restricted learning mechanism (chunking) and symbolic representations. The architecture reflects the early interests of cognitive scientists, who mostly deal with well-defined problems which can be solved by the consequent application of a limited number of rules<sup>27</sup>.

ACT-R cognitive architecture (Anderson 1990) is, in turn, based on the following three elements: modularity assumption (the mind is modeled as a set of interconnected, specialized modules), procedural and declarative memory (the former consists of "production rules", the latter of encodings of facts and utility values) and the mechanism of rational analysis. ACT-R is a hybrid architecture, as activation of symbols depends on subsymbolic processes. The ACT-R architecture was initially applied in research on human memory (Anderson & Bower 1973), human problem solving and, more recently, cognitive tutoring.

---

<sup>26</sup> The comparison is often dependent on first-person methods, as Zlatev suggests (e.g. on subjects' self-reports).

<sup>27</sup> Simon and Newell's book: Human Problem Solving clearly presents the class of problems that can be addressed with a help of Soar architecture.

*Clarion* (Sun 2002) is also a hybrid architecture (using both symbolical and connectionist representations), but it uses distributed representations more widely than ACT-R. *Clarion* is designed to model the division between implicit cognition and explicit cognition. Behavior is here a result of interaction of implicit (bottom-up) learning and explicit (top-down) learning. *Clarion* was designed to model motivational processes prior to cognition, as a basis of action and cognition. It is also intended to model learning in the presence or absence of explicit domain knowledge. Finally, it is supposed to model social interaction, in which individuals can appreciate other agents' motivations. (cf. Sun 2006)

What is important about the aforementioned architectures is their relatively limited areas of application. All architectures strive to cover a broad area of cognitive phenomena, but their application is limited only to certain class(es) of cognitive phenomena.<sup>28</sup>

It is also worth noting that such architectures—due to their structural and representational features—constrain the set of possible models implemented within them. In particular, the form of representation (symbolic or/and distributed), types of memories (declarative, procedural), limitations imposed on memory, assumed learning mechanisms: all these features decide which architecture may be used as a basis of a model of a particular cognitive process.<sup>29</sup>

### **4.3. SNePS/GLAIR as a cognitive architecture**

Before proceeding with the examples of—as I claim—semiotic activity of artificial cognitive systems, it is worth noting that I do not intend to analyze all artificial systems as possible semiotic systems. Only specialized systems based on artificial cognitive architectures and developed within these architectures' ACAs may function as sign users or interpreters. Among others, ACAs have the ability to construct and process representations of objects and states of affairs according to some theory of cognition. In particular, the architectures may support knowledge storage, reasoning, belief revision, planning or natural language understanding and generating.

Natural language is of particular interest for semioticians as it is considered a primary and paradigmatical sign system. The primacy of language can be seen not only in traditional semiotics (cf. the notion of *semiology*, Saussure 1983), but is also important in the case of ACAs: natural language understand-

---

<sup>28</sup> However, one has to notice that researchers attempt to extend their architectures to explain new aspects of cognition. Soar architecture, initially designed for symbolic problem-solving activities has been recently extended to take into account effects of emotions on cognition (Marinier & Laird 2004)

<sup>29</sup> The detailed analysis of different cognitive architectures and their applicability to certain tasks can be found in: Taatgen&Anderson (2008).

ing is considered an “AI-complete” domain of research (Shapiro 1992), meaning that to implement a cognitive system we need to implement natural language abilities. In other words, whatever else artificial cognitive agents are, they also process natural language symbols (or “marks”).

SNePS (<http://www.cse.buffalo.edu/sneps>) has been designed as a system that is able to use natural language and represent knowledge in the form of semantic networks (Quillian 1968). As an architecture, SNePS/GLAIR is based on symbolic representations, a feature that makes it particularly relevant for modeling such “representation-hungry problems” (Clark&Toribio 1994) as semiotic and linguistic activity. Knowledge in the system is represented in the form of a semantic network, which is particularly convenient for representing semantic (i.e. meaning-related) relationships between elements. Nowadays SNePS is a part of GLAIR, a multi-layered cognitive architecture for “embodied agents operating in real, virtual, or simulated environments containing other agents” (Shapiro, Bona 2010: 307). Interaction with an environment allows modeling of grounding of signs and language expressions in interactions with an agent’s surroundings.<sup>30</sup>

The architecture consists of three main layers:

- a *Knowledge Layer* (KL) in the form of a semantic network: (propositional) knowledge is represented as a network of *nodes* (representing propositions) connected by *links* (representing relationships between propositions). It is said that a KL contains the beliefs of a cognitive system. There are two subsystems that operate on semantic networks: SNeRE, a reasoning system, and SNeBR, i.e. SNePS Belief Revision system;
- *sensori-actuator layer* (SAL) is responsible for “perception”, i.e. internalizing information from the world outside a cognitive agent, and for acting in that world. To realize these tasks, the SAL layer contains controllers of sensors and effectors;
- *perceptuo-motor layer* (PML) is an intermediate level between sensors, effectors and the knowledge layer. More precisely, it is responsible for grounding of semantic network symbols in perceptual structures and actions; in addition, it contains various registers for providing the agent’s sense of situatedness in the environment.

The GLAIR/SNePS architecture is a framework for creating artificial cognitive agents embedded in the world (due to the SAL layer), representing knowledge, reasoning and “understanding” natural language utterances, as well as gener-

---

<sup>30</sup> The inclusion of sensors and effectors in the architecture is a (indirect, I suppose) result of criticism of purely internalist approaches (see Rowlands 2010) formulated within 4e cognitive science.

ating them. It is important to stress that artificial cognitive agents are not meant to realize or implement semiotic systems, but rather to simulate their functional organization.

## **5. Cognitive modeling as simulational semiotics**

As mentioned in section 2, cognitive semiotics should integrate methods and theories developed in humanities (semiotics in particular) and in cognitive science. A widely-accepted method within cognitive science is the use of artificial cognitive models to find answers to some cognition-related questions. I argue that it may be also a useful method for discovering properties of semiotic systems. As stated at the beginning of the paper, the key assumption is that processes of semiosis (e.g. in the sense of Peircean theory) could be also considered as cognitive processes.

The main idea of cognitive modeling is to look for answers by investigation of functioning cognitive models. A model of a GLAIR-based system that uses natural language allows one to address and (to some extent) answer the following semiosis-related questions: how to distinguish between interpretability and some actual interpretation; how to model the dynamic character of processes of semiosis; how to resolve the problem of contradictory beliefs; and finally, the question of metaknowledge.

### **5.1. Interpretability and actual interpretation**

Artificial cognitive systems are sometimes described as formal systems that manipulate symbols on the basis of these symbols' shapes. Therefore, operations on symbols (creating, modifying, removing) are considered to be "syntactic" operations. According to Harnad (1994), these operations must be globally semantically interpretable. The criterion of interpretability ("cryptographers' constraint", Harnad 1994) is a *precondition* for being a cognitive system. In other words, such a system may be interpreted as recognizing objects, discriminating shapes, drawing conclusions, planning activities etc. However, it seems that the condition of *being interpretable* (in principle) is too weak in the context of considerations on semiotic systems. The condition of interpretability allows for—so to say—external interpretations of a behavior of a system. Let us assume that a driver stops her car at a red traffic light. An external observer (a pedestrian, let us say) may interpret the behavior as respecting the obligation to stop at the red light. The driver, if asked why she stopped, may explain: *I just gave way to the car approaching the junction from the right*. The driver's behavior is interpretable as if she had respected the red light, whereas the red light had no influence on her behavior. The red light **is not** a sign for this driver (despite external interpretation). Obviously, the con-

sistency of the driver's behavior is a constraint on possible interpretations, but we may imagine a situation in which a driver stopping is consistent with both the "red-light" and the "giving way" interpretations. The parallel argument is formulated in reference to computers (or calculators). Their "behavior" may be interpretable as adding, dividing etc. Such interpretability does not force us to admit that numerals on an LCD display are signs for a computer/calculator. Consequently, the requirement of interpretability is only an initial requirement. Any artificial semiotic system must *actually* interpret symbols (e.g. natural language utterances) *itself* in the context of its knowledge, abilities and environment<sup>31</sup>

They are thoughts about something, they are meaningful, and they are not about what they are about merely because they are systematically interpretable by you as being about what they are about. They are *about them autonomously and directly*, without any mediation. (Harnad 1994, my emphasis).

It is Fetzer who also stresses the difference between interpretability and actual interpretation:

[...] [T]here is a crucial difference between sign-using systems and other systems that [...] do not qualify as semiotic systems. Ordinary digital computers are systems of this [latter] kind, because the marks by means of which they operate are meaningless *for systems* of that kind. While they may be meaningful *for users of those systems*, they are not meaningful for those systems themselves. (Fetzer 2001: 118; my emphasis)

The above quotation may be considered as an invitation to introduce the idea of "embeddedness". As semiotic systems, artificial cognitive systems need to be embedded in some world, i.e. they should interact with the world, create new (arbitrary) symbols in their knowledge bases, modify and remove them as a response to external stimuli and, finally, they should act in the world (on the basis of their knowledge). It turns out that to be a semiotic system an artificial cognitive agent needs sensors and effectors that assure the connection between "mind" and "world".<sup>32</sup> An actual interpretation requires an explicit connection between a symbol (a sign) and an object, as detected by sensors or an element of a system's internal knowledge base.

SNePS architecture forces actual interpretation. Data received from the environment must be actually interpreted in terms of the existing knowledge base. Interpretation is a necessary condition for inclusion in the knowledge base. In addition, connections between elements of knowledge base and SAL/PML layers provide (at least partial) grounding for symbols.

---

<sup>31</sup> The environment is understood here widely: as physical, cultural or social.

<sup>32</sup> Harnad (1994) would treat such sensors and effectors as a partial solution of the problem of "grounding of symbols".

## 5.2. The dynamic character of meaning and interpretation

Any ACA must have some means of adding new and changing or removing existing information. It is often assumed that modeling semiotic activity involves a model's ability to learn. It must be emphasized, however, that learning could not be based on association of a sign and its fixed meaning. As it is a consequence of Peircean fallibilism, the dynamicity of the process of interpretation precludes (pre-) established semantics for a set of signs. An artificial semiotic system has to “discover” the meanings itself (using sensory/acting modules or various reasoning mechanisms). As these “discoveries” are fallible, the knowledge base should be dynamic i.e. susceptible to constant reinterpretation. As I have argued elsewhere (Konderak 2015), an interpretation may be understood as placing a sign (representing a belief) within a network of interconnected signs (beliefs). The full meaning of a sign (at any given moment) is determined by a whole network of signs (beliefs). In consequence, the meaning of a sign depends on its location in the entire network representing an agent's knowledge and history of word acquisition (connections are always made in the context of learning).<sup>33</sup>

Restating the above in terms of a SNePS/GLAIR cognitive agent, the full meaning of a node of its semantic network is in principle determined by the whole network representing the agent's knowledge. It should be emphasized that no single node has a meaning (as with a single sign in Peircean theory). The *actual* meaning of the node is some part of the surrounding network. Therefore, in practice the meaning of a node depends on the whole network and its structure, the place of the node in the network and the *scope* of the part of the network taken into account (i.e. used in reasoning). A change in any part of the network results in a change in the meaning of a node. Such an approach constitutes an example of the *holistic theory of meaning* (Quine 1951). As Quillian (1968: 238) states, “a word's full concept [i.e. node - PK is defined to be all the nodes that can be reached by an exhaustive tracing process, originating at its initial, patriarchal type node”. In consequence, the meaning of any node (and any word) instantly changes along with the cognitive development of an agent and the dynamic character of interpretation is built into the knowledge representation formalism.

---

<sup>33</sup> In addition, the set of nodes existing in a network will also depend on the order of presentation (“order dependency”). An agent interpreting the statement: *Jan wished to know whether Charles Dodgson was the author of Alice in Wonderland* would create two nodes: representing Dodgson and an author of *Alice in Wonderland*. An agent believing that Dodgson is the author of *Alice in Wonderland* would create only one node associated with two different language labels.

### 5.3. *Contradictory interpretations*

The process of the interpretation of signs is not unrestricted. Even in the context of Peircean theory, every act of interpretation is determined by some Dynamical Object. It seems that one of the basic features of a process of interpretation of signs is avoidance of contradictions. On the one hand, however, sign users—treating the world they live in as consistent (at both the metaphysical and ontological levels)—usually avoid or eliminate inconsistencies in their belief systems. This is also somehow justified psychologically: when realized, inconsistencies cause discomfort and elicit trials to eliminate the inconsistencies. Psychological argument is supported by a logical argument: according to Duns Scotus' law ( $p \wedge \sim p \rightarrow q$ ), anything is derivable from an inconsistency.

The avoidance or elimination of contradictions, however, does not always take place. The reason is that cognitive agents (either natural or artificial) are not omniscient: they do not know all the possible consequences of their beliefs. They also are not aware of all their beliefs at any given moment of time. In consequence, cognitive agents, in fact, often maintain contradictory beliefs without realizing it. To deal with this situation, we need some formal tools for representing contradictory beliefs in the knowledge base of a cognitive agent. To create an artificial cognitive agent using natural language, we need some kind of logic that is weaker than classical logic and would allow contradictions to be dealt with. There have been several approaches called *paraconsistent logics* (like a relevance logic) that allow for inconsistencies within a system (e.g. a system of one's beliefs) without acceptance of consequences of the Duns Scotus law. To put it differently, our cognitive agent should be able to believe contradictory interpretations of a sign without believing everything. (cf. Pańniczek 1986) In addition, it is a fact that we usually somehow interpret expressions denoting contradictory objects (such as a square circle)—trying, for instance, to prove their non-existence (cf. Meinong 1981).

An unrealized contradiction within a belief system must be distinguished from conscious acceptance of contradictory beliefs. The latter case is omitted here, as interpreters usually try to maintain consistency of their beliefs (sign interpretations). Artificial cognitive agents may suggest the course of the processes of re-interpretations of signs in the context of conflicting (e.g. contradictory) interpretation of signs.

SNePS/GLAIR cognitive architecture delivers a kind of explanation of the process. It is the SNeBR (SNePS Belief Revision system) module that detects contradictions and takes relevant action. When an agent realizes a contradiction, it has to retract one of the base beliefs that underlies the contradiction. The retracted belief is retained in the knowledge base as an unasserted belief. There may be a situation in which a system has to decide which of the two or more beliefs has to be retracted. The system solves the problem of distinguishing its beliefs by taking into account, for example, their importance (con-

sistency maintenance procedures will force the removal of “less important” beliefs); in order to accomplish this, ACAs mark credibility according to some criteria (source, history of acquisition, category).

#### 5.4. Metaknowledge

*Metacognition* can be understood as any cognitive process that controls or monitors any aspect of cognition. (Moses, Baird 1999) In turn, *Metaknowledge* is defined as “knowledge about knowledge” and embraces, in the context of GLAIR cognitive agents, beliefs about beliefs (metabeliefs).

Fetzer (2001) claims that being a semiotic system requires more than using symbols (e.g. using natural language expressions). The system must realize (be aware) that it *uses signs as signs*, i.e. the system needs both meta-knowledge embracing the usage of signs and some metaprocesses that control the interpretation of signs. Taking semiotic activity as an aspect of cognition, we may (as Fetzer does) expect semiotic systems to demonstrate such metacognition and metaknowledge.

In general, it is assumed that any system that has knowledge and reasons about it needs a model of itself. There are two aspects of self-modeling:

- beliefs about itself (i.e. beliefs with the self-term such as I as an argument) and
- a sense of embodiedness and situatedness in the world (Shapiro, Rapaport *et al.* 2007: 21).

SNePS/GLAIR agents have some features that facilitate metacognition and metaknowledge. The first aspect mentioned above is possibly due to the fact that propositions are represented as *terms* rather than as *sentences*, so that propositions can occur as arguments of propositions (I believe that *he will be a philosopher*) without the necessity of using second-order logic.

The second aspect is realized via a set of PML deictic registers (including the registers *I*, *you*, as well as *now*) and modality registers containing the terms representing currently performed acts or current state of an agent. Thanks to the above registers, an agent is able to formulate statements in the first-person, in the past or future tense (*I realized that he was sad*). “Affective and effective actions implemented in the PML are the source of first-person privileged knowledge about what the agent is sensing or doing”. (Shapiro, Rapaport *et al.* 2007: 28)

Such an approach turns formulas of traditional, propositional logic that represent sentences with well-defined truth values such as  $p \rightarrow q$  into meta-formulas such as “ $P \rightarrow Q$ ”, stating that if an agent interprets a sign as meaning P, the agent is also justified in interpreting some other sign as meaning Q.

The notion of metaknowledge is used in several areas of research (Shapiro, Rapaport *et al* 2007): self-awareness (Cox 2005), action in situations where there is a lack of sufficient knowledge (McCarthy 1977), math-capable agents (Cohors-Fresenborg & Kaune 2001), or contextual vocabulary acquisition (Rapaport 2005).

### **5.5. An example implementation: vocabulary acquisition**

As an anonymous reviewer of the paper noted, “a good case study is preferable to several promissory notes”.

The presentation so far resembles just wandering in some direction without presenting a particular, working solution. As far as I am aware, there is no cognitive model of (general) semiotic activity so far. Instead, researchers attempt to model various particular aspects of the process of meaning-making or semiosis. In particular, researchers focus on one particular sign system, namely natural language *as a symbolic system of signs*. The emphasized statement is quite important in the context of this paper: we are less interested in language as a means of communication, or as a means of influence, or in its cognitive functions (cf. Carruthers 2002), but rather we are interested in language as a *semiotic system*, a system of utterances which stand for something else in one respect or another on the basis of conventions. Accordingly, not all models of language should be taken into account, but only models instantiating important features of the semiotic systems presented above: actual, embedded interpretation of utterances as signs, dynamicity of interpretation, the ability to cope with contradictory interpretations and finally metaknowledge or metasemiotic aspect.

The short presentation of the particular computational model of vocabulary acquisition (Rapaport & Ehrlich 2000) should be understood as a kind of justification of the choice of SNePS/GLAiR as an architecture relevant for modeling certain aspects of semiotic activity.

The vocabulary acquisition model implemented within SNePS architecture is characterized as “a computational theory of how natural-language-understanding systems can automatically acquire new vocabulary by determining from context the meaning of words that are unknown, misunderstood, or used in a new sense”. (Rapaport & Ehrlich 2000)

In their paper, Rapaport and Ehrlich (2000: 349) present a model which tries to understand (i.e. interpret) a previously unknown word, which in the presented case is the word “brachet”. The model has some (limited) knowledge on meanings of words that is represented in the form of a semantic network. The

second clue used by the system is context<sup>34</sup> (e.g. textual: the word appears in Mallory's *Morte Darthur*).

The model may go through the three stages of vocabulary acquisition; namely, constructing a new definition of an unknown word, correcting a definition of a misunderstood word and expanding the definition of a word being used in a new sense. The first stage is the actual interpretation of the expression in terms of the semantic network of the model: it builds a preliminary definition of the word, an actual hypothesis about its meaning. Importantly, the system does it autonomously, without external (e.g. human) intervention. The system is not told, for instance, that the initial definition is incorrect. The definition changes as a result of interaction of the model with the text. Initial characterization does not match all the uses of the word in the text, so the model adjusts the definition accordingly. The process (performed partially by the belief revision system) reflects the dynamicity of the interpretation. As Rapaport and Ehrlich report, the agent of the SNePS system builds the following four definitions:

- A brachet is a physical object that may be white.
- A brachet is an animal that may bite and can be small and white.
- A brachet is an animal that may bite and can be small, valuable, and white.
- A brachet is a hunting dog that can bay and bite.

Although this particular case does not display operations on contradictory characteristics, the system is ready for such a situation: “when SNeBR is invoked by a derived contradiction, the belief(s) with least priority (i.e., held with least certainty) in the conflict set will be selected for revision”. (Rapaport & Ehrlich 2000). In other words, the system chooses the most uncertain statement as a candidate for revision. Finally, as Shapiro *et al.* (2007: 24-25) explicitly state, vocabulary acquisition is an example of cognitive activity involving metacognition. The system must display the ability to represent its own beliefs, to reason about them, to compare them, to apply them in various contexts and to evaluate the application; all of them are clearly metacognitive abilities and may be treated as mechanisms responsible for Peircean re-interpretation of signs.

To evaluate the vocabulary-acquiring model as a cognitive model, one has to refer to the criteria of assessment (see section 4). The presented model is informed by research on expanding human vocabularies (see Rapaport&Ehrlich 2000, section 4). However, the model should not be understood as just explanation of previously acquired data: in an important sense, the model predicts

---

<sup>34</sup> Context includes surrounding text, grammatical information, and background knowledge.

the behavior of human subjects. As Rapaport and Ehrlich show, the system's behavior is very similar to those of human subjects.<sup>35</sup> In addition, human subjects and the artificial system converged on a very similar definition. In consequence, the model proved its validity from a behavioral point of view.

It must be also noted that the researchers did not equip the system with detailed, domain-specific knowledge. The definitions are formulated on the basis of text and general background knowledge. In other words, the system is not "adjusted" to the text it should interpret: it uses general knowledge.

## **6. Virtual semiotic agents in virtual worlds?**

One may object that artificial cognitive agents are not appropriate models of human sign-interpreting beings as they neither live in nor ground their symbols in the same world. The requirements of the actual interpretation mentioned above indicate the strong dependence between the semiotic activity of an agent and the world (in fact worlds) in which the agent lives. It seems that signs are dependent on a world in at least two ways: Firstness and Secondness are parts of a world in which we function as semiotic systems. We derive both signs and their references from the worlds (either "real", or fictional) we are acquainted with. Secondly, our cognitive apparatus—as shaped by evolution—seems to be especially tuned to the world we function in.

In consequence, ACAs may turn out to be appropriate for explaining the processes of semiosis in virtual or simulated worlds. However, our means of communication and changes in our human sources of information are accompanied by an evolution of signs: sign systems are becoming increasingly grounded in virtual reality (virtual Facebook friends, the Bitcoin currency, Second Life). The Internet and media reality are becoming the new, virtual domain of the interpretation (or grounding) of the signs we use. However, grounding is only one aspect of the process of "virtualization" of signs and semiosis: the other one is a delimitation of possible interpretations by the structure of sources of information, e.g. the World Wide Web. Links placed within a text direct the flow of interpretation. As Eco (1999) writes: "In a continual deferral from text to text I had lost the opportunity to produce the habit which would have allowed me to pick up what I originally wanted".

The somehow speculative conclusion of this section is that to understand the process of human interpretation of the whole new domain of "virtual" signs, we need some extension of traditional semiotic theories. We may need artificial agents acting in virtual or simulated worlds and interpreting signs, thus grounding them in these worlds.

---

<sup>35</sup> Rapaport and Ehrlich compared the performance of the model with reports of two human subjects coping with the same task – providing a definition of the word „bracket“.

## **7. Conclusions**

The suggested approach is without doubt reductive. I do not dare to claim that cognitive modeling can explain semiotic processes *in toto*. I appreciate the phenomenological approaches to meaning and semiosis (Sonesson 2012) as well as those based on the notion of life (Zlatev 2003).

The key notion of the paper is that of an *artificial cognitive system* (or artificial cognitive agent) as a possible *simulation* of a cognitive-semiotic system. Accepting the links between cognitive science and semiotics, we should also consequently consider the adaptation of methods of cognitive science to the investigation of sign systems and sign processes.

One of the advantages of the suggested approach is that it can fill the gaps in traditional semiotic theories—in particular, in Peircean triadic conception—suggesting specific computational mechanisms responsible for representation and (re-)interpretations of signs. The examples which are shown above include mechanisms of dynamic reinterpretation of signs, resolving inconsistencies in interpretations and acquisition of new symbols from context. Moreover, we can observe not only the mechanisms, but also their development in time.

The evolution of thinking about semiosis can be seen on at least three levels: methodological, conceptual and ontological. First of all, the status of artificial cognitive systems has changed. Artificial systems can be viewed not only as “sign-vehicles” (Andersen), symbol-processors or signs themselves (Fetzer), but we can go a step further and treat them as simulations of semiotic processes. Cognitive science treats such systems as models of, for example, perception, reasoning or problem-solving activity. Similarly, it would be worth treating them seriously as models of semiotic processes.

Secondly, it seems that the range of possible signs and semiotic systems has changed. Increasingly, we are dealing with signs taken from and functioning in virtual and simulated worlds that have their own structure, restricted ontology and specific principles of organization and functioning. It is possible that to investigate our semiotic activity in these worlds we need new tools: artificial semiotic systems that are tuned to these worlds.

## **References**

- Andersen, P.B. 1990. *A Theory of Computer Semiotics: Semiotic Approaches to Construction and Assessment of Computer Systems*. Cambridge: Cambridge University Press.
- Anderson, J. 1993. *Rules of the Mind*. Hillsdale, New Jersey: Erlbaum.
- Anderson, J. 2007. *How Can the Human Mind Occur in the Physical Universe?* New York: Oxford University Press.

- Anderson, J., et al. 2004. An integrated theory of the mind. *Psychological Review*, 111(4): 1036–1060.
- Anderson, J.R., Bower, G.H. 1973. *Human associative memory*. Washington, DC: Winston and Sons.
- Brandt, P.Å. 2004. *Spaces, Domains and Meanings. Essays in Cognitive Semiotics*. Bern: Peter Lang.
- Carruthers, P. 2002. The cognitive functions of language. *Behavioral and Brain Sciences*, 25: 657–726.
- Chalupsky, H., Shapiro, S.C. 1994. SL: A subjective, intensional logic of belief. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Hillsdale, New Jersey: Lawrence Erlbaum: 165–170.
- Clark, A., Toribio, J. 1994. Doing without representing. *Synthese*, 101: 401–431.
- Clarke, R. 2001. Studies in Organisational Semiotics: an Introduction. *Australian Journal of Information Systems*, 8(2): 75–82.
- Cohors-Fresenborg, E., Kaune, C. 2001. Mechanisms of the Taking Effect of Metacognition in Understanding Processes in Mathematics Teaching. *Developments in Mathematics Education in German-Speaking Countries, Selected Papers from the Annual Conference on Didactics of Mathematics*: 29–38. Ludwigsburg, Germany: Goettingen University Press.
- Cox, M.T. 2005. Metacognition in computation: A selected research review. *Artificial Intelligence*, 169(2): 104–141.
- Eco, U. 1999. Semiotics in the Next Millennium, <http://www.umbertoeco.it/CV/Semiotics%20in%20the%20next%20millennium.pdf>
- Fetzer, J.H. 1997. Thinking and Computing: Computers as Special Kinds of Signs. *Minds and Machines*, 7(3): 345–364.
- Fetzer, J.H. 2001. *Computers and Cognition: Why Minds Are Not Machines*. Dordrecht: Kluwer Academic Publishers.
- Fodor, J., Lepore, E. 1992. *Holism: A Shopper's Guide*. Cambridge, MA: Basil Blackwell.
- Gudwin, R.R. 1999. From Semiotics to Computational Semiotics, *Proceedings of the 9th International Congress of the German Society for Semiotic Studies, 7th International Congress of the International Association for Semiotic Studies (IASS/AIS)*. Dresden.
- Harnad, S. 1994. Computation Is Just Interpretable Symbol Manipulation: Cognition Isn't. *Minds and Machines*, 4(4).
- Hintikka, J. 1962. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press.
- Konderak, P. 2005. Model kognitywny zdolności językowych [Cognitive Model of Language Faculty] (Unpublished doctoral dissertation). Maria Curie-Skłodowska University, Lublin.
- Konderak, P. 2015. On a cognitive model of semiosis. *Studies in Logic, Grammar and Rhetoric*, 40 (53),129-144.

- Marinier, L.P., Laird, J.E. 2004. Toward a comprehensive computational model of emotions and feelings. *Proceedings of the Sixth International Conference on Cognitive Modeling*. Mahwah, New Jersey: Erlbaum: 172–177.
- McCarthy, J. 1977. Epistemological Problems of Artificial Intelligence. *Proceedings of the 5th International Joint Conference on Artificial Intelligence (IJCAI-77)*, vol. 2: 1038–1044. Los Altos, CA: William Kaufmann, Inc.
- Mehler, A. 2003. Methodological Aspects of Computational Semiotics. *S.E.E.D. Journal (Semiotics, Evolution, Energy, and Development)*, 3(3): 71–80.
- Meinong, A. 1981. The Theory of Objects. *Realism and the Background of Phenomenology*. Chisholm, R. (ed.). Atascadero, CA: Ridgeview: 76–117.
- Moses, L.J., Baird, J.A. 1999. Metacognition., *The MIT Encyclopedia of the Cognitive Sciences*. Wilson, R. A., Keil, F. (eds.). Cambridge, MA: The MIT Press: 533–535.
- Paśniczek, J. 1986. Czy sprzeczność może być racjonalna? [Can Contradiction Be Rational?]. *Czy sprzeczność może być racjonalna?* Jodkowski, K. (ed.), Lublin: Wydawnictwo UMCS: 193–208.
- Peirce, Ch.S. 1931-58. *Collected Papers of Charles Sanders Peirce*. Cambridge: Harvard University Press.
- Peirce, Ch.S. 1982, *The Writings of Charles S. Peirce: A Chronological Edition*. vol. 1–6, 8. Peirce Edition Project. Bloomington I.N: Indiana University Press.
- Piaget, J. 1945. *La formation du symbole chez l'enfant*. Paris: Delachaux et Niestlé.
- Quillian, R.M. 1968. Semantic Memory. *Semantic Information Processing*. Minsky, M. (ed.). Cambridge: The MIT Press: 227–270.
- Quine, W.V.O. 1951. Two Dogmas of Empiricism. *The Philosophical Review*, 60: 20–43.
- Rapaport, W. 2005. In Defense of Contextual Vocabulary Acquisition: How to Do Things with Words in Context. *Proceedings of the 5th International and Interdisciplinary Conference on Modeling and Using Context (Context 2005): Lecture Notes in Artificial Intelligence 3554*, Berlin: Springer Publishing: 396–409.
- Rapaport, W. 2012. Semiotic Systems, Computers and the Mind: How Cognition Could Be Computing, *International Journal of Signs and Semiotic Systems*, 2(1): 32–71.
- Rapaport, W., Ehlich, K. 2000. A Computational Theory of Vocabulary Acquisition. *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language*. Iwańska, Ł., Shapiro, S. (eds.). Menlo Park, CA: AAAI Press/MIT Press: 347–375.
- Saussure, F. de ([1916] 1983). *Course in General Linguistics*. Harris, R. (trans.). London: Duckworth.
- Shapiro, S., Bona, J. 2009. The GLAIR Cognitive Architecture. *Biologically Inspired Cognitive Architectures II*. Menlo Park, CA: AAAI Press: 141–152.
- Shapiro, S., Rapaport, W., Kandefer, M., Johnson, F.J., Goldfain, A. 2007. Metacognition in SNePS. *AI Magazine*, 28(1): 17–29.

- Sokolowski, R. 2000. *Introduction to Phenomenology*. Cambridge, MA: Cambridge University Press.
- Sonesson, G. 2009. The view from Husserl's lectern: Considerations on the role of phenomenology in cognitive semiotics. *Cybernetics and Human Knowing*, 16(3-4): 107-148.
- Sonesson, G. 2012. The Phenomenological Road to Cognitive Semiotics, *Proceedings of the 10th World Congress of the International Association for Semiotic Studies (IASS/AIS)*. Universidade da Coruña: 855-866.
- Sun, R. 2002. *Duality of the Mind: A Bottom-up Approach Toward Cognition*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Sun, R. 2006. The CLARION cognitive architecture: Extending cognitive modeling to social simulation. *Cognition and Multi-Agent Interaction*, Sun, R. (ed.). Cambridge University Press, New York.
- Taatgen, N.A., Anderson, J.R. 2008. Constraints in cognitive architectures. *The Cambridge Handbook of Computational Psychology*. Sun, R. (ed.). Cambridge University Press: 170-185.
- Zlatev, J. 2003. Meaning = Life (+ Culture). An outline of a unified biocultural theory of meaning. *Evolution of Communication*, 4(2): 253-296.
- Zlatev, J. 2012. Cognitive Semiotics: An emerging field for the transdisciplinary study of meaning. *The Public Journal of Semiotics*, 4(1): 2-24.