Semiotic Brains and Artificial Minds
How Brains Make Up Material Cognitive Systems

Lorenzo Magnani

Department of Philosophy and Computational Philosophy Laboratory, University of Pavia, Pavia, Italy and *Department of Philosophy, Sun Yat-sen University, Guangzhou (Canton), P.R. China* lmagnani@unipv.it

Abstract: Our brains make up a series of signs and are engaged in making or manifesting or reacting to a series of signs: through this semiotic activity they are at the same time engaged in "being minds" and so in thinking intelligently. An important effect of this semiotic activity of brains is a continuous process of "externalization of the mind" that exhibits a new cognitive perspective on the mechanisms underling the semiotic emergence of abductive processes of meaning formation. In this perspective we can see that at the root of thinking abilities there is a process of externalization/disembodiment of mind that presents a new cognitive perspective on the role of external models and representations. To illustrate this process I will take advantage of Turing's comparison between unorganized brains and logical and practical machines and of the analysis of some aspects of the cognitive interplay between internal and external representations. I consider this interplay critical in analyzing the relation between meaningful semiotic internal resources and devices and their dynamical interactions with the externalized semiotic materiality suitably stocked in the environment. Hence, minds are material, "extended" and artificial in themselves. A considerable part of human abductive thinking is occurring through an activity consisting in a kind of reification in the external environment and a subsequent re-projection and reinterpretation through new configurations of neural networks and chemical processes. The last part of the paper will describe the concept of mimetic mind that I have introduced to shed new light on the role of computational modeling and on the decline of the so-called Cartesian computationalism.

1. INTRODUCTION

What I call *semiotic brains* are brains that make up a series of signs and that are engaged in making or manifesting or reacting to a series of signs: through this semiotic activity they are at the same time engaged in "being minds" and so in thinking intelligently. An important effect of this semiotic activity of brains is a continuous process of disembodiment of mind that exhibits a new cognitive perspective on the mechanisms underling the semiotic emergence of meaning processes.

To illustrate this process I will take advantage of some paleoanthropological results on the birth of material culture, that provide an evolutionary perspective on the origin of intelligent behaviors. Then I will describe the centrality to semiotic cognitive information processes of the *disembodiment of mind* from the point of view of the cognitive interplay between internal and external representations. I consider this interplay critical in analyzing the relation between meaningful semiotic internal resources and devices and their dynamical interactions with the externalized semiotic materiality already stocked in the environment. Hence, minds are "extended" and artificial in themselves.

Unorganized brains organize themselves through a semiotic activity that is reified in the external environment and then re-projected and reinterpreted through new configurations of neural networks and chemical processes. I also think the disembodiment of mind can nicely account for low-level semiotic processes of meaning creation, bringing up the question of how could higher-level processes be comprised and how would they interact with lower-level ones. To the aim of explaining these higher-level semiotic mechanisms I provide the analysis of model-based and manipulative abduction, and of external representations (Zhang, 1997) where many external things, usually inert from the semiotic point of view, can be transformed into what we can I call "epistemic mediators" (Magnani, 2001a) that give rise - for instance in the case of scientific reasoning - to new signs, new chances for interpretants, and new interpretations.

In the last part of the paper (section 7) the concept of *mimetic mind* is introduced to shed new cognitive and philosophical light on the role of computational modeling and on the decline of the so-called Cartesian computationalism and to emphasize the possible impact of the construction of new types of universal "practical" machines, available over there, in the environment, as new tools underlying the emergence of meaning processes.

2. THE CENTRALITY OF ABDUCTION

If we decide to increase knowledge on the semiotic character of cognition it is necessary to develop a cognitive model of creativity able to represent not only "novelty" and "unconventionality", but also some features commonly referred to as the entire creative process, such as the expert use of background knowledge and ontology (defining new concepts and their new meanings and searching heuristically among the old ones) and the modeling activity developed in the so called "incubation time" (generating and testing, transformations in the space of the hypotheses). The philosophical concept of *abduction* may be a candidate to solve this problem, and offers an approach to model creative processes of meaning generation in a completely explicit and formal way, which can fruitfully integrate the narrowness proper of a merely psychological approach, too experimentally human-oriented.

A hundred years ago, C. S. Peirce (*CP*, 1931-1958) coined the concept of abduction in order to illustrate that the process of scientific discovery is not irrational and that a methodology of discovery is possible. Peirce interpreted abduction essentially as an "inferential" *creative process* of generating a new hypothesis. Abduction has a logical form (fallacious, if we model abduction by using classical syllogistic logic)[1] distinct from deduction and induction. Reasoning which starts from reasons and looks for consequences is called *deduction*; that which starts from consequences and looks for reasons is called *abduction*.

Abduction – a distinct form of reasoning – is the process of *inferring* certain facts and/or laws and hypotheses that render some sentences plausible, that *explain* or *discover* some (eventually new) phenomenon or observation; it is the process of reasoning in which explanatory hypotheses are formed and evaluated. There are two main epistemological meanings of the word abduction (Magnani, 2001a): 1) abduction that only generates "plausible" hypotheses ("selective" or "creative") and 2) abduction considered as inference "to the best explanation", which also evaluates hypotheses. An illustration from the field of medical knowledge is represented by the discovery of a new disease and the manifestations it causes which can be considered as the result of a creative abductive inference. Therefore, "creative" abduction

[1] The abductive inference rule corresponds to the well-known fallacy called affirming the consequent (simplified to the propositional case)

$$\frac{\varphi \rightarrow \psi \qquad \psi}{\varphi}$$

deals with the whole field of the growth of scientific knowledge. This is irrelevant in medical *diagnosis* where instead the task is to "select" from an encyclopedia of pre-stored diagnostic entities. We can call both inferences ampliative, selective and creative, because in both cases the reasoning involved amplifies, or goes beyond, the information incorporated in the premises (Magnani, 1992).

Theoretical abduction certainly illustrates much of what is important in creative abductive reasoning, in humans and in computational programs, but fails to account for many cases of explanations occurring in science when the exploitation of environment is crucial. It fails to account for those cases in which there is a kind of "discovering through doing", cases in which new and still unexpressed information is codified by means of manipulations of some external objects (*epistemic mediators*). I have introduced this concept of theoretical abduction in Magnani (2001a, 2002): I maintain that there are two kinds of theoretical abduction, "sentential", related to logic and to verbal/symbolic inferences, and "model-based", related to the exploitation of internalized models of diagrams, pictures, etc., cf. below in this paper. The concept of *manipulative abduction*[2] captures a large part of scientific thinking where the role of action is central, and where the features of this action are implicit and hard to be elicited: action can provide otherwise unavailable information that enables the agent to solve problems by starting and by performing a suitable abductive process of generation or selection of hypotheses.

In section 6 I will describe how manipulative abduction can nicely account for the relationship between meaningful behavior and dynamical interactions with the environment. The following sections illustrate that at the roots of the creation of new meanings there is a process of *disembodiment of mind* that exhibits a new cognitive description of the mechanisms underling the emergence of meaning processes through semiotic delegations to the environment.

3. FROM THE PREHISTORIC BRAINS TO THE UNIVERSAL MACHINES

I have said that what I call *semiotic brains* are brains that make up a series of signs and that are engaged in making or manifesting or reacting to a series of signs: through this semiotic activity they are at the same time engaged in "being minds" and so in thinking intelligently. In this section I will

[2] Manipulative abduction and epistemic mediators are introduced and illustrated in Magnani (2001a).

illustrate the process of "disembodiment of mind" as an important aspect of this semiotic activity of brains.

Following Turing's point of view (Turing, 1949) a big cortex can provide an evolutionary advantage only in presence of a massive storage of meaningful information and knowledge on external supports that only an already developed small community of human beings can possess. Evidence from paleoanthropology seems to support this perspective. Some research (Mithen, 1996, 1999, Humphrey, 2002, Lewis-Williams, 2002) in cognitive paleoanthropology – even if rather speculative - teaches us that high level and reflective consciousness in terms of thoughts about our own thoughts and about our feelings (that is consciousness not merely considered as raw sensation) is intertwined with the development of *modern language* (speech) and *material culture*. After 250.000 years ago several hominid species had brains as large as ours today, but their behavior lacked any sign of art or symbolic behavior. If we consider high-level consciousness as related to a high-level organization – in Turing's sense – of human cortex, its origins can be related to the active role of environmental, social, linguistic, and cultural aspects.

Handaxes were made by Early Humans and firstly appeared 1,4 million years ago, still made by some of the Neanderthals in Europe just 50.000 years ago. The making of handaxes is strictly intertwined with the development of consciousness. Many needed capabilities constitute a part of an evolved psychology that appeared long before the first handaxed were manufactured. It seems humans were pre-adapted for some components required to make handaxes (Mithen, 1996, 1999)

1. imposition of *symmetry* (already evolved through predators escape and social interaction). It has been an unintentional by-product of the bifacial knapping technique but also deliberately imposed in other cases. Dennett (1991) hypothesizes that the attention to symmetry may have developed through social interaction and predator escape, as it may allow one to recognize that one is being directly stared at. It also seems that "Hominid handaxes makers may have been keying into this attraction to symmetry when producing tools to attract the attention of other hominids, especially those of the opposite sex" (Mithen, 1999, p. 287);
2. understanding *fracture dynamics* (for example evident from Oldowan tools and from nut cracking by chimpanzees today);
3. ability to *plan* ahead (modifying plans and reacting to contingencies, such unexpected flaws in the material and miss-hits), still evident in the minds of Oldowan tool makers and in chimpanzees;
4. high degree of *sensory-motor control*: "Nodules, pre-forms, and near finished artefacts must be struck at precisely the right angle with pre-

> cisely the right degree of force if the desired flake is to be detached" (Mithen, 1999, p. 285). The origin of this capability is usually tracked back to encephalization – the increased number of nerve tracts and of the integration between them allows for the firing of smaller muscle groups - and bipedalism – that requires a more complex integrated highly fractionated nervous system, which in turn presupposes a larger brain.

The combination of these four resources produced an important semiotic revolution: the birth of what Mithen calls technical intelligence of early human mind, that is consequently related to the construction of handaxes and their new semiotic values. Indeed they indicate high intelligence and good health. They cannot be compared to the artefacts made by animals, like honeycomb or spider web, deriving from the iteration of fixed actions which do not require consciousness and intelligence.

3.1 Private Speech and Fleeting Consciousness

Two central factors play a fundamental role in the combination of the four resources above:

- the exploitation of *private speech* (speaking to oneself) to trail between planning, fracture dynamic, motor control and symmetry (also in children there is a kind of private muttering which makes explicit what is implicit in the various abilities);
- a good degree of *fleeting consciousness* (thoughts about thoughts).

Of course they furnish a kind of blackboard where the four - previously distinct - resources can be exploited all together and in their dynamic interaction. In the meantime these two aspects obviously played a fundamental role in the development of consciousness and thought:

> So my argument is that when our ancestors made handaxes there were private mutterings accompanying the crack of stone against stone. Those private mutterings were instrumental in pulling the knowledge required for handaxes manufacture into an emergent consciousness. But what type of consciousness? I think probably one that was fleeting one: one that existed during the act of manufacture and that did not the endure. One quite unlike the consciousness about one's emotions, feelings, and desires that were associated with the social world and that probably were part of a completely separated cognitive domain, that of social intelligence, in the early human mind (p. 288).

This use of private speech can be certainly considered a semiotic internal "tool" for organizing brains and so for manipulating, expanding, and exploring minds, a tool that probably coevolved with another: talking to each other.[3] Both private and public language act as tools for thought and play a fundamental role in the evolution "opening up our minds to ourselves" and so in the emergence of new meaning processes.

3.2 Material Culture and Semiosis

Another semiotic tool appeared in the latter stages of human evolution, that played a great role in the evolutions of primitive minds, that is in the organization of human brains. Handaxes also are at the birth of *material culture*, so as new cognitive chances can co-evolve:

- the mind of some early humans, like the Neanderthals, were constituted by relatively isolated semiotic cognitive domains, Mithen (1999) calls *different intelligences*, probably endowed with different degrees of consciousness about the thoughts and knowledge within each domain (natural history intelligence, technical intelligence, social intelligence). These isolated cognitive domains became integrated also taking advantage of the role of public language;
- *degrees of high level consciousness* appear, human beings need thoughts about thoughts;
- social intelligence and public language arise.

It is extremely important to stress that *material culture* is not just the product of this massive cognitive chance but also cause of it. "The clever trick that humans learnt was to *disembody* their minds into the material world around them: a linguistic utterance might be considered as a disembodied thought. But such utterances last just for a few seconds. Material culture endures" (cit., p. 291).

In this perspective we acknowledge that material artefacts are tools for thoughts as is language: tools – as new "signs" - for exploring, expanding, and manipulating our own minds. In this regard the evolution of culture is inextricably linked with the evolution of consciousness and thought.

Early human brain becomes a kind of universal "intelligent" machine, extremely flexible so that we did no longer need different "separated" intelligent machines doing different jobs. A single one will suffice. As the engineering problem of producing various machines for various jobs is replaced by the office work of "programming" the universal machine to do these jobs,

[3] On languages as cognitive artefacts cf. Carruthers (2002), Clark (1998, 2003, 2005), Norman (1993), and Clowes & Morse (2005).

so the different intelligences become integrated in a new universal device endowed with a high-level type of consciousness.[4]

From this perspective the semiotic expansion of the minds is in the meantime a continuous process of *disembodiment* of the minds themselves into the *material world* around them. In this regard the evolution of the mind is inextricably linked with the evolution of large, integrated, material cognitive semiotic systems. In the following sections I will illustrate this extraordinary interplay between human brains and the cognitive systems they make.

3.3 Semiotic Delegations through the Disembodiment of Mind

A wonderful example of meaning creation through disembodiment of mind is the carving of what most likely is the mythical being from the last ice age, 30.000 years ago, a half human/half lion figure carved from mammoth ivory found at Hohlenstein Stadel, Germany.

> An evolved mind is unlikely to have a *natural home* for this being, as such entities do not exist in the natural world: so whereas evolved minds could think about humans by exploiting modules shaped by natural selection, and about lions by deploying content rich mental modules moulded by natural selection and about other lions by using other content rich modules from the natural history cognitive domain, how could one think about entities that were part human and part animal? Such entities had no home in the mind (cit., p. 291).

[4] On the relationship between material culture and the evolution of consciousness cf. (Donald, 1998 and 2001; Dennett, 2003).

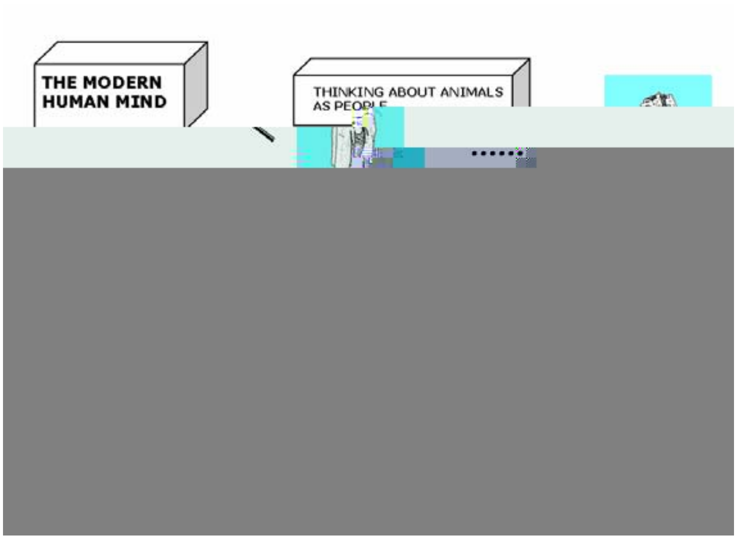

Figure1. In Mithen, 1999.

A mind consisting of different separated intelligences cannot come up with such entity (Figure 1). The only way is *to extend* the mind into the *material word*, exploiting in a semiotic way rocks, blackboards, paper, ivory, and writing, painting, and carving: “artefacts such as this figure play the role of anchors for ideas and have no *natural home* within the mind; for ideas that take us beyond those that natural selection could enable us to possess” (cit., p. 291).

In the case of our figure we face with an anthropomorphic thinking created by the material representation serving to semiotically anchor the cognitive representation of supernatural being. In this case the material culture disembodies thoughts, that otherwise will soon disappear, without being transmitted to other human beings, and realizes a systematic semiotic delegation to the external environment. The early human mind possessed two separated intelligences for thinking about animals and people. Through the mediation of the material culture the modern human mind can arrive to *internally* think about the new concept of animal and people at the same time. But the new meaning occurred over there, in the external material world where the mind picked up it.

Artefacts as *external semiotic objects* allowed humans to loosen and cut those chains on our unorganized brains imposed by our evolutionary past. Chains that always limited the brains of other human beings, such as the Neandertals. Loosing chains and securing ideas to external objects was also a way to creatively re-organize brains as universal machines for thinking.

In the remaining part of the paper I will describe the centrality to semiotic cognitive information processes of the disembodiment of mind from the point of view of the cognitive interplay between internal and external representations. I consider this interplay critical in analyzing the relation between meaningful semiotic internal resources and devices and their dynamical interactions with the externalized semiotic materiality already stocked in the environment. Hence, minds are "extended" and artificial in themselves.

4. MIMETIC AND CREATIVE REPRESENTATIONS

We have seen that unorganized brains organize themselves through a semiotic activity that is reified in the external environment and then reprojected and reinterpreted through new configurations of neural networks and chemical processes. I also think the disembodiment of mind can nicely account for low-level semiotic processes of meaning creation, bringing up the question of how could higher-level processes be comprised and how would they interact with lower-level ones.

4.1 External and Internal Representations

We have said that through the mediation of the material culture the modern human mind can arrive to *internally* think the new meaning of animals and people at the same time. We can account for this process of disembodiment from an impressive cognitive point of view.

I maintain that representations are external and internal. We can say that

- *external representations* are formed by external materials that express (through reification) concepts and problems already stored in the brain or that do not have a *natural home* in it;
- *internalized representations* are internal re-projections, a kind of recapitulations, (learning) of external representations in terms of neural patterns of activation in the brain. They can sometimes be "internally" manipulated like external objects and can originate new internal reconstructed representations through the neural activity of *transformation* and *integration*.

This process explains why human beings seem to perform both computations of a *connectionist* type[5] such as the ones involving representations as

[5] Here the reference to the word "connectionism" is used on the plausible assumption that all mental representations are brain structures: verbal and the full range of sensory representations are neural structures endowed with their chemical functioning (neurotransmitters and hormones) and electrical activity (neurons fire and provide electrical inputs to other neurons). In this sense we can reconceptualize cognition neurologically: for example the solution of a problem can be seen as a process in which one neural structure representing an

- (I Level) *patterns of neural activation* that arise as the result of the interaction between body and environment (and suitably shaped by the evolution and the individual history): pattern completion or image recognition,

and computations that use representations as

- (II Level) *derived combinatorial syntax and semantics* dynamically shaped by the various external representations and reasoning devices found or constructed in the environment (for example geometrical diagrams); they are neurologically represented contingently as pattern of neural activations that "sometimes" tend to become stabilized structures and to fix and so *to permanently belong to the I Level* above.

The I Level originates those *sensations* (they constitute a kind of "face" we think the world has), that provide room for the II Level to reflect the structure of the environment, and, most important, that can follow the computations suggested by these external structures. It is clear we can now conclude that the growth of the brain and especially the synaptic and dendritic growth are profoundly determined by the environment.

When the fixation is reached the patterns of neural activation no longer need a direct stimulus from the environment for their construction. In a certain sense they can be viewed as *fixed internal records* of *external structures* that *can exist* also in the absence of such external structures. These patterns of neural activation that constitute the I Level Representations always keep record of the experience that generated them and, thus, always carry the II Level Representation associated to them, even if in a different form, the form of *memory* and not the form of a vivid sensorial experience. Now, the human agent, via neural mechanisms, can retrieve these II Level Representations and use them as *internal* representations or use parts of them to construct new internal representations very different from the ones stored in memory (cf. also Gatti & Magnani, 2005).[6]

In the following section I will illustrate some fundamental aspects of the interplay above in the light of basic semiotic aspects of abductive reasoning.

---

explanatory target generates another neural structure that constitutes a hypothesis for the solution.

[6] The role of external representations has already been stressed in some central traditions of cognitive science and artificial intelligence, from the area of distributed and embodied cognition and of robotics (cf. Brooks, 1991, Clark, 2003, Zhang, 1997) to the area of active vision and perception (Gibson, 1979, Thomas, 1999).

5. MODEL-BASED ABDUCTION AND SEMIOSIS BEYOND PEIRCE

I think there are two basic kinds of external representations active in this process of externalization of the mind: *creative* and *mimetic*. Mimetic external representations mirror concepts and problems that are already represented in the brain and need to be enhanced, solved, further complicated, etc so they sometimes can also creatively give rise to new concepts and meanings. In the examples I will illustrate in the following sections it will be clear how for instance a mimetic geometric representation can become creative and give rise to new meanings and ideas in the hybrid interplay between brains and suitable cognitive environments.

What exactly is model-based abduction from a philosophical point of view? Peirce stated that all thinking is in signs, and signs can be icons, indices, or symbols. Moreover, all *inference* is a form of sign activity, where the word sign includes "feeling, image, conception, and other representation" (Peirce, *CP*, 5.283), and, in Kantian words, all synthetic forms of cognition. That is, a considerable part of the creative meaning processes is *model-based*. Moreover, a considerable part of the meaningful behaviour (not only in science) occurs in the middle of a relationship between brains and external objects and tools that have received cognitive and/or epistemological delegations (cf. the previous and the following subsection).

Following this Peircian perspective about inference I think it is extremely useful from a cognitive point of view to consider the concept of reasoning in a very broad way (cf. also Brent, 2000, p. 8). We have three cases:

1) reasoning can be fully conscious and typical of high-level worked-out ways of inferring, like in the case of scientists' and professionals' performances;

2) reasoning can be "acritical" (Peirce, *CP*, 5.108), which includes every day inferences in conversation and in various ordinary patterns of thinking;

3) reasoning can resort to "operations of the mind which are logically analogous to inference excepting only that they are unconscious and therefore uncontrollable and therefore not subject to logical criticism" (Peirce, *CP*, 5.108).

Immediately Peirce adds a note to the third case "But that makes all the difference in the world; for inference is essentially deliberate, and self-controlled. Any operation which cannot be controlled, any conclusion which is not abandoned, not merely as soon as criticism has pronounced against it, but in the very act of pronouncing that decree, is not of the nature of rational inference – is not reasoning" (*ibid*.).

As Colapietro clearly states (2000, p. 140), it seems that for Peirce human beings semiotically involve unwitting trials and unconscious processes.

Moreover, it seems clear that unconscious thought can be in some sense considered “inference”, even if not rational; indeed, Peirce says, it is not reasoning. Peirce further indicates that there are in human beings multiple trains of thought at once but only a small fraction of them is conscious, nevertheless the prominence in consciousness of one train of thought is not to be interpreted an interruption of other ones.

In this Peircian perspective, which I adopt in this essay, where inferential aspects of thinking dominate, there is no intuition, in an anti-Cartesian way. We know all important facts about ourselves in an *inferential* abductive way:

> […] we first form a definite idea of ourselves as a hypothesis to provide a place in which our errors and other people’s perceptions of us can happen. Furthermore, this hypothesis is constructed from our knowledge of “outward” physical facts, such things as the sounds we speak and the bodily movements we make, that Peirce calls signs (Brent, 2000,, p. 10).

Recognizing in a series of *material*, physical events, that they make up a series of signs, is to know the existence of a *mind* (or of a group of minds) and to be absorbed in making, manifesting, or reacting to a series of signs is to be absorbed in “being a mind”. “[…] all thinking is dialogic in form” (Peirce, *CP*, 6.338)**,** both at the intrasubjective[7] and intersubjective level, so that we see ourselves exactly as others see us, or see them exactly as they see themselves, and we see ourselves through our own speech and other interpretable behaviors, just others see us and themselves in the same way, in the commonality of the whole process (Brent, 2000, p. 10).

As we will better explain later on in the following sections, in this perspective minds are material like brains, in so far as they consist in intertwined internal and external semiotic processes: “[…] the psychologists undertake to locate various mental powers in the brain; and above all consider it as quite certain that the faculty of language resides in a certain lobe; but I believe it comes decidedly nearer the truth (though not really true) that language resides in the tongue. In my opinion it is much more true that the thoughts of a living writer are in any printed copy of his book than they are in his brain” (Peirce, *CP*, 7.364).

[7] “One’s thoughts are what he is ‘saying to himself’, that is saying to that other self that is just coming to life in the flow of time. When one reasons, it that critical self that one is trying to persuade: and all thought whatsoever is a sign, and is mostly in the nature of language” (Peirce, *CP*, 5.421).

5.1 Man is an External Sign

Peirce's semiotic motto "man is an external sign" is very clear about the materiality of mind and about the fact that the conscious self[8] is a cluster actively embodied of flowing intelligible signs:

> It is sufficient to say that there is no element whatever of man's consciousness which has not something corresponding to it in the word; and the reason is obvious. It is that the word or sign which man uses *is* the man himself. For, as the fact that every thought is a sign, taken in conjunction with the fact that life is a train of thoughts, proves that man is a sign; so, that every thought is an *external sign*, proves that man is an external sign. That is to say, the man and the *external sign* are identical, in the same *sense* in which the words *homo* and *man* are identical. Thus my language is the sum total of myself; for the man is the thought (Peirce, *CP*, 5.314).

It is by way of signs that we ourselves *are* semiotic processes – for example a more or less coherent cluster of narratives. If all thinking is in signs it is not true that thoughts are in us because we are in thoughts.

I think it is at this point clearer what I meant when I said in the previous section, when I explained the concept of model-based abduction, that all thinking is in signs, and signs can be icons, indices, or symbols and that, moreover, all *inference* is a form of sign activity, where the word sign includes feeling, image, conception, and other representation. The model-based aspects of human cognition are central, given the central role played for example by signs like images and feeling in the inferential activity "[…] man is a sign developing according to the laws of inference. […] the entire phenomenal manifestation of mind is a sign resulting from inference" (Peirce, *CP*, 5.312 and 5.313).

Moreover, the "person-sign" is future-conditional, that is not fully formed in the present but depending on the future destiny of the concrete semiotic activity (future thoughts and experience of the community) in which he or she will be involved. If Peirce maintains that when we think we appear as a sign (Peirce, *CP*, 5.283) and, moreover, that everything is present to us is a phenomenal manifestation of ourselves, feelings, images, diagrams, conceptions, schemata, and other representations are phenomenal manifestations that become available for interpretations and thus are guiding our actions in a positive or negative way. They become *signs* when we think and interpret them. It is well-known that for Peirce all semiotic experience – and

[8] Consciousness arises as "a sort of public spirit among the nerve cells" (Peirce, *CP*, 1.354).

thus abduction - is also providing a guide for action. Indeed the whole function of thought is to produce habits of action.[9]

Let us summarize some basic semiotic ideas that will be of help in the further clarification of the cognitive and computational features of model-based and manipulative abduction. One of the central property of signs is their reinterpretability. This occurs in a social process where signs are referred to *material objects*.

As it is well-known for Peirce *iconic* signs are based on similarity alone, the psychoanalytic patient who thought he was masturbating when piloting the plane interpreted the cloche as an extension of his body, and an iconic sign of the penis; an ape may serve as an icon of a human. *Indexical* signs are based on contiguity and dynamic relation to the object, a sign which refers to an object that it denotes by virtue of being "really affected" by that object: a certain grimace indicates the presence of pain, the rise of the column of mercury in a thermometer is a sign of a rise in temperature, indexical signs are also the footprints in the sand or a rap on the door. Consequently we can say indexical signs "point". A *symbol* refers to an artificial or conventional ("by virtue of a law") interpretation of a sign, the sign ∞ used by mathematicians would be an example of Peirce's notion of symbol, almost all words in language, except for occasional onomatopoeic qualities, are symbols in this sense, associated with referents in a wholly arbitrary manner.

We have to immediately note that from the semiotic point of view *feelings*, too, are signs that are subject to semiotic interpretations at different levels of complexity. Peirce considered feelings elementary phenomena of mind, comprising all that is immediately present, such as pain, sadness, cheerfulness. He believes that a feeling is a state of mind possessing its own living qualities independent of any other state of the mind. Neither icon, index, nor symbol actually functions as a sign until it is interpreted and recognized in a semiotic activity and code. To make an example, it is the evolutionary kinship that makes the ape an icon of the man, in itself the similarity of two animals does not mean anything.

Where cognition is merely possible, sign action, or *semiosis*, is working. Knowledge is surely inferential as well as abduction, that like any inference requires three elements: a *sign*, the *object* signified, and the *interpretant*. Everywhere "A signifies B to C".

There is a continuous activity of interpretation and some of this activity – as we will see - is abductive. The Peircian notion of interpretant plays the role of explaining the activity of interpretation that is occurring in semiosis. The interpretant does not necessarily refer to an actual person or mind, an

[9] Cf. for example the contributions contained in recent special issue of the journal *Semiotica* devoted to abduction (Queiroz and Merrell, 2005).

actual interpreter. For instance the communication to be found in a beehive[10] where the bees are able to communicate with the others by means of signs is an example of a kind of “mindless” triadic semiosis: indeed we recognize that a sign has been interpreted not because we have observed a mental action but by observing another material sign. To make another example, the person recognizing the thermometer as a thermometer is an interpretant, as she generates in her brain a thought. In this case the process is conscious, but also unconscious or emotional interpretants are widespread. Again, a person points (index) up at the sky and his companion looks up (interpretant) to see the object of the sign. Someone else might call out “What do you see up there?” that is also another interpretant of the original sign. As noted by Brent “For Peirce, any appropriate response to a sign is acting as another sign of the object originally signified. A sunflower following the sun across the sky with its face is also an interpretant. Peirce uses the word interpretant to stand for any such development of a given sign” (Brent, 2000, p. 12).

Finally, an interpretant may be the thought of another person, but may as well be simply the further thought of the first person, for example in a soliloquy the succeeding thought is the interpretant of the preceding thought so that an interpretant is both the interpretant of the thought that precedes it and the object of the interpretant thought that succeeds it. In soliloquy sign, object, and interpretant are all present in the single train of thought.

Interpretants, mediating between signs and their objects have three distinct levels in hierarchy: feelings, actions, and concepts or habits (that is various generalities as responses to a sign). They are the effect of a sign process. The interpretant produced by the sign can lead to a feeling (*emotional* interpretant), or to a muscular or mental effort, that is to a kind of action - *energetic* interpretant (not only outward, bodily action, but also purely inward exertions like those “mental soliloquies strutting and fretting on the stage of imagination” - Colapietro, 2000, p. 142). Finally, when it is related to the abstract meaning of the sign, the interpretant is called *logical*, as a generalization requiring the use of verbal symbols. It is a further development of semiosis in the hierarchy of iconic, enactive, and symbolic communication: in short, it is “an interpreting thought”, related for instance not only to the intellectual activity but also to initiate the ethical action in so far as a “modification of a person’s tendencies toward action” (Peirce, *CP*, 5.476).

The logical interpretants are able to translate percepts, emotions, unconscious needs, and experience needs, and so to mediate their meanings to arrive to provisional stabilities. They can lead to relatively stable cognitive or intellectual habits and belief changes as self-controlled achievements like many abductive conceptual results, that Peirce considers the most advanced forms of semiosis and the ultimate outcome of a sign. Indeed abduction –

[10] This kind of communication is studied in Monekosso, Remagnino, & Ferri.

hypothesis - is the first step toward the formation of cognitive habits: "every concept, every general proposition of the great edifice of science, first came to us as a conjecture. These ideas are the *first logical interpretants* of the phenomena that suggested them, and which, as suggesting them, are signs" (Peirce, *CP*, 5.480).

Ortogonal to the classification of interpretants as emotional, energetic, and logical is the alternate classification given by Peirce: interpretants can also be immediate, dynamic, and normal. Some interpreters consider this classification a different way of expressing the first one. It is sufficient to note this classification can be useful in studying the formation of a subclass of debilitating and facilitating psychic habits (Colapietro, 2000, pp. 144-146). Colapietro proposes the concept of quasi-final interpretants – as related to the Peircian normal interpretants - as "effective in the minimal sense that they allow the conflict-ridden organism to escape being paralyzed agent: they permit the body-ego to continue its ongoing negotiations with these conflicting demands, even if only in a precarious and even debilitating manner. In brief, they permit the body-ego to go on" (p. 146). For instance there are some sedimented unsconscious reactions of this type in immediate puzzling environments – later on useless and stultifying in wider settings - but there also is the recurrent reflective and – provisionally - productive use of fallacious ways of reasoning like hasty generalizations and other arguments (Woods, 2004).

In the following sections I will describe how the interplay of signs, objects, and interpretants is working in important aspects of abductive reasoning. Of course model-based cognition acquires its peculiar creative relevance when embedded in abductive processes. I will show some examples of model-based inferences. It is well known the importance Peirce ascribed to diagrammatic thinking (a kind of iconic thinking), as shown by his discovery of the powerful system of predicate logic based on diagrams or "existential graphs". As we have already stressed, Peirce considers inferential any cognitive activity whatever, not only conscious abstract thought; he also includes perceptual knowledge and subconscious cognitive activity. For instance in subconscious mental activities visual representations play an immediate role.[11]

Many commentators always criticized the Peircian ambiguity in treating abduction in the same time as inference and perception. It is important to clarify this problem, because perception and imagery are kinds of that model-based cognition which we are exploiting to explain abduction: in (Magnani, 2006) I conclude we can render consistent the two views, beyond Peirce, but perhaps also within the Peircian texts, taking advantage of the

[11] Cf. Queiroz and Merrell, 2005.

concept of *multimodal abduction*, which depicts hybrid aspects of abductive reasoning (Magnani, 2006)..

6. CONSTRUCTING MEANING THROUGH MIMETIC AND CREATIVE EXTERNAL OBJECTS

6.1 Constructing Meaning through Manipulative Abduction

Manipulative abduction occurs when many external things, usually inert from the semiotic point of view, can be transformed into what I have called “epistemic mediators” (Magnani, 2001a) that give rise - for instance in the case of scientific reasoning - to new signs, new chances for interpretants, and new interpretations.

We can cognitively account for the process of disembodiment of mind we have seen in the perspective of paleoanthropology taking advantage of the concept pf *manipulative* abduction. It happens when we are thinking *through* doing and not only, in a pragmatic sense, about doing. It happens, for instance, when we are creating geometry constructing and manipulating an external, suitably realized, icon like a triangle looking for new meaningful features of it, like in the case given by Kant in the “Transcendental Doctrine of Method” (cf. Magnani, 2001b; cf. also the following subsection). It refers to an extra-theoretical behavior that aims at creating communicable accounts of new experiences to integrate them into previously existing systems of experimental and linguistic (semantic) practices.

Gooding (1990) refers to this kind of concrete manipulative reasoning when he illustrates the role in science of the so-called “construals” that embody tacit inferences in procedures that are often apparatus and machine based. The embodiment is of course an expert manipulation of meaningful semiotic objects in a highly constrained experimental environment, and is directed by abductive movements that imply the strategic application of old and new *templates* of behavior mainly connected with extra-rational components, for instance emotional, esthetical, ethical, and economic.

The hypothetical character of construals is clear: they can be developed to examine or discard further chances, they are provisional creative organization of experience and some of them become in their turn hypothetical *interpretations* of experience, that is more theory-oriented, their reference/meaning is gradually stabilized in terms of established observational practices. Step by step the new interpretation - that at the beginning is completely "practice-laden" - relates to more "theoretical" modes of understanding (narrative, visual, diagrammatic, symbolic, conceptual, simulative), closer to the constructive effects of theoretical abduction. When the reference/meaning is stabilized the effects of incommensurability with other established observations can become evident. But it is just the construal of certain phenomena that can be shared by the sustainers of rival theories. Gooding (1990) shows how Davy and Faraday could see the same attractive and repulsive actions at work in the phenomena they respectively produced; their discourse and practice as to the role of their construals of phenomena clearly demonstrate they did not inhabit different, incommensurable worlds in some cases. Moreover, the experience is constructed, reconstructed, and distributed across a social network of negotiations among the different scientists by means of construals.

It is difficult to establish a list of invariant behaviors that are able to describe manipulative abduction in science. As illustrated above, certainly the expert manipulation of objects in a highly semiotically constrained experimental environment implies the application of old and new *templates* of behavior that exhibit some regularities. The activity of building construals is highly conjectural and not immediately explanatory: these templates are hypotheses of behavior (creative or already cognitively present in the scientist's mind-body system, and sometimes already applied) that abductively enable a kind of epistemic "doing". Hence, some templates of action and manipulation can be *selected* in the set of the ones available and pre-stored, others have to be *created* for the first time to perform the most interesting creative cognitive accomplishments of manipulative abduction.

Moreover, I think that a better understanding of manipulative abduction at the level of scientific experiment could improve our knowledge of induction, and its distinction from abduction: manipulative abduction could be considered as a kind of basis for further meaningful inductive generalizations. Different generated construals can give rise to different inductive generalizations.

Some common features of these tacit templates that enable us to manipulate things and experiments in science to favour meaning formation are related to: 1. sensibility towards the aspects of the phenomenon which can be regarded as *curious* or *anomalous*; manipulations have to be able to introduce potential inconsistencies in the received knowledge (Oersted's report of his well-known experiment about electromagnetism is devoted to describe some anomalous aspects that did not depend on any particular theory of the nature of electricity and magnetism; Ampère's construal of experiment on electromagnetism - exploiting an artifactual apparatus to produce a static equilibrium of a suspended helix that clearly shows the role of the "unexpected"); 2. preliminary sensibility towards the *dynamic* character of the phenomenon, and not to entities and their properties, common aim of manipulations is to practically reorder the dynamic sequence of events in a static spatial one that should promote a subsequent bird's-eye view (narrative or visual-diagrammatic); 3. referral to experimental manipulations that exploit *artificial apparatus* to free new possibly stable and repeatable sources of information about hidden knowledge and constraints (Davy well-known set-up in terms of an artifactual tower of needles showed that magnetization was related to orientation and does not require physical contact). Of course this information is not artificially made by us: the fact that phenomena are made and manipulated does not render them to be idealistically and subjectively determined; 4. various contingent ways of epistemic acting: *looking* from different perspectives, *checking* the different information available, *comparing* subsequent events, *choosing*, *discarding*, *imaging* further manipulations, *re-ordering* and *changing relationships* in the world by implicitly *evaluating* the usefulness of a new order (for instance, to help memory).

From the general point of view of everyday situations manipulative abductive reasoning exhibits other very interesting templates: 5. action elaborates a *simplification* of the reasoning task and a redistribution of effort across time when we "need to manipulate concrete things in order to understand structures which are otherwise too abstract" (Piaget, 1974), or when we are in presence of *redundant* and unmanageable information; 6. action can be useful in presence of *incomplete* or *inconsistent* information - not only from the "perceptual" point of view - or of a diminished capacity to act upon the world: it is used to get more data to restore coherence and to improve deficient knowledge; 7. action as a *control of sense data* illustrates how we can change the position of our body (and/or of the external objects) and how to exploit various kinds of prostheses (Galileo's telescope, technological instruments and interfaces) to get various new kinds of stimulation: action provides some tactile and visual information (e. g., in surgery), otherwise unavailable; 8. action enables us to build *external artifactual models* of task mechanisms instead of the corresponding internal ones, that are adequate to adapt the environment to the agent's needs: experimental manipulations exploit *artificial apparatus* to free new possible stable and repeatable sources of information about hidden knowledge and constraints.[12]

[12] The problem of manipulative abduction and of its tacit features is strongly related to the whole area of recent research on embodied reasoning (Anderson, 2002), but also relates to studies on external representations and situated robotics (Clancey, 2002, Agree & Chapman, 1990, Brooks & Stein, 1994). I have illustrated its role in ethical reasoning in Magnani (2007).

The whole activity of manipulation is devoted to build various external *epistemic mediators*[13] that function as versatile semiotic tools able to provide an enormous new source of information and knowledge. Therefore, manipulative abduction represents a kind of redistribution of the epistemic and cognitive effort to manage objects and information that cannot be immediately represented or found internally (for example exploiting the resources of visual imagery).[14]

If we see scientific discovery like a kind of opportunistic ability of integrating information from many kinds of simultaneous constraints to produce explanatory hypotheses that account for them all, then manipulative abduction will play the role of eliciting possible hidden constraints by building external suitable experimental structures.

6.2 Manipulating Meanings through External Semiotic Anchors

If the structures of the environment play such an important role in shaping our semiotic representations and, hence, our cognitive processes, we can expect that physical manipulations of the environment receive a cognitive relevance.

Several authors have pointed out the role that physical actions can have at a cognitive level. In this sense Kirsh & Maglio (1994) distinguish actions into two categories, namely *pragmatic actions* and *epistemic actions*. Pragmatic actions are the actions that an agent performs in the environment in order to bring itself physically closer to a goal. In this case the action modifies the environment so that the latter acquires a configuration that helps the agent to reach a goal which is understood as physical, that is, as a desired state of affairs. Epistemic actions are the actions that an agent performs in a semiotic environment in order to discharge the mind of a cognitive load or to extract information that is hidden or that would be very hard to obtain only by internal computation.

In this section I want to focus specifically on the relationship that can exist between manipulations of the environment and representations. In particular, I want to examine whether external manipulations can be considered as means to construct external representations.

[13] I derive this expression from the cognitive anthropologist Hutchins, that coins the expression “mediating structure” to refer to various external tools that can be built to cognitively help the activity of navigating in modern but also in “primitive” settings (Hutchins, 1995 and 1999).

[14] It is difficult to preserve precise spatial relationships using mental imagery, especially when one set of them has to be moved relative to another.

If a manipulative action performed upon the environment is devoted to create a configuration of signs that carries relevant information, that action will well be able to be considered as a cognitive semiotic process and the configuration of elements it creates will well be able to be considered an external representation. In this case, we can really speak of an embodied cognitive process in which an action constructs an external representation by means of manipulation. We define *cognitive manipulating* as any manipulation of the environment devoted to construct external configurations that can count as representations.

An example of cognitive manipulating is the diagrammatic demonstration illustrated in Figure 2, taken from the field of geometry. In this case a simple manipulation of the triangle in Figure 2(a) gives rise to an external configuration - Figure 2(b) - that carries relevant semiotic information about the internal angles of a triangle "anchoring" new meanings.

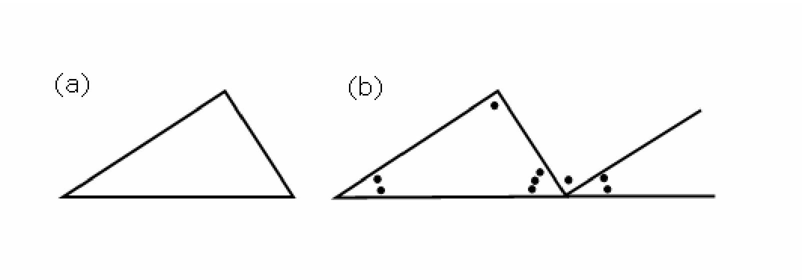

Figure 2. Diagrammatic demonstration that the sum of the internal angles of any triangle is 180°. (a) Triangle. (b) Diagrammatic manipulations.

The entire process through which an agent arrives at a physical action that can count as cognitive manipulating can be understood by means of the concept of manipulative abduction (Magnani, 2001a). Manipulative abduction is a specific case of cognitive manipulating in which an agent, when faced with an external situation from which it is hard or impossible to extract new meaningful features of an object, selects or creates an action that structures the environment in such a way that it gives information which would be otherwise unavailable and which is used specifically to infer explanatory hypotheses.

In this way the semiotic result is achieved on *external* representations used in lieu of the internal ones. Here action performs an *epistemic* and not a merely performatory role, for example relevant to abductive reasoning.

6.3 Geometrical Construction is a Kind of Manipulative Abduction

Let's quote Peirce's passage about mathematical constructions. Peirce says that mathematical and geometrical reasoning "consists in constructing a diagram according to a general precept, in observing certain relations between parts of that diagram not explicitly required by the precept, showing that these relations will hold for all such diagrams, and in formulating this conclusion in general terms. All valid necessary reasoning is in fact thus diagrammatic" (Peirce, *CP*, 1.54). This passage clearly refers to a situation like the one I have illustrated in the previous section. This kind of reasoning is also called by Peirce "theorematic" and it is a kind of "deduction" necessary to derive significant theorems: "[…] is one which, having represented the conditions of the conclusion in a diagram, performs an ingenious experiment upon the diagram, and by observation of the diagram, so modified, ascertains the truth of the conclusion" (Peirce, *CP*, 2.267). The experiment is performed with the help of "imagination upon the image of the premiss in order from the result of such experiment to make corollarial deductions to the truth of the conclusion" (Peirce, 1976, IV, p. 38). The "corollarial" reasoning is mechanical (Peirce thinks it can be performed by a "logical machine") and not creative, "A Corollarial Deduction is one which represents the condition of the conclusion in a diagram and finds from the observation of this diagram, as it is, the truth of the conclusion" (Peirce, *CP*, 2.267, cf. also Hoffmann, 1999).

In summary, the point of theorematic reasoning is the transformation of the problem by establishing an unnoticed point of view to get interesting – and possibly new – insights. The demonstrations of theorems in mathematics are examples of theorematic deduction.

Not dissimilarly Kant says that in geometrical construction of external diagrams "[…] I must not restrict my attention to what I am actually thinking in my concept of a triangle (this is nothing more than the mere definition); I must pass beyond it to properties which are not contained in this concept, but yet belong to it" (Kant, 1929, A718-B746, p. 580).

We have seen that manipulative abduction is a kind of abduction, usually model-based, that exploits external models endowed with delegated (and often implicit) cognitive and semiotic roles and attributes.

1. The model (diagram) is *external* and the strategy that organizes the manipulations is unknown *a priori*.
2. The result achieved is *new* (if we, for instance, refer to the constructions of the first creators of geometry), and adds properties not contained before in the concept (the Kantian to "pass beyond" or "ad-

> vance beyond" the given concept, Kant, 1929, A154-B193/194, p. 192).[15]

Iconicity in theorematic reasoning is central. Peirce, analogously to Kant, maintains that "philosophical reasoning is reasoning with words; while theorematic reasoning, or mathematical reasoning is reasoning with specially constructed schemata" (Peirce, *CP*, 4.233); moreover, he uses diagrammatic and schematic as synonyms, thus relating his considerations to the Kantian tradition where schemata mediate between intellect and phenomena.[16] The following is the famous passage in the *Critique of Pure Reason* ("Transcendental Doctrine of Method"):

> Suppose a philosopher be given the concept of a triangle and he be left to find out, in his own way, what relation the sum of its angles bears to a right angle. He has nothing but the concept of a figure enclosed by three straight lines, and possessing three angles. However long he meditates on this concept, he will never produce anything new. He can analyse and clarify the concept of a straight line or of an angle or of the number three, but he can never arrive at any properties not already contained in these concepts. Now let the geometrician take up these questions. He at once begins by constructing a triangle. Since he knows that the sum of two right angles is exactly equal to the sum of all the adjacent angles which can be constructed from a single point on a straight line, he prolongs one side of his triangle and obtains two adjacent angles, which together are equal to two right angles. He then divides the external angle by drawing a line parallel to the opposite side of the triangle, and observes that he has thus obtained an external adjacent angle which is equal to an internal angle – and so on.[17] In this fashion, through a chain of inferences guided throughout by intuition, he arrives at a fully evident and universally valid solution of the problem (Kant, 1929, A716-B744, pp. 578-579).

As we have already said for Peirce the whole mathematics consists in building diagrams that are "[…] (continuous in geometry and arrays of repeated signs/letters in algebra) according to general precepts and then [in] observing in the parts of these diagrams relations not explicitly required in the precepts" (Peirce, *CP*, 1.54). Peirce contends that this diagrammatic na-

[15] Of course in the case we are using diagrams to demonstrate already known theorems (for instance in didactic settings), the strategy of manipulations is not necessary unknown and the result is not new.

[16] Schematism, a fruit of the imagination is, according to Kant, "[…] an art concealed in the depths of the human soul, whose real modes of activity nature is hardly likely ever to allow us to discover, and to have open to our gaze" (Kant, 1929, A141-B181, p. 183).

[17] It is Euclid's Proposition XXXII, Book I, cf. above Figure 2.

ture is not clear if we only consider syllogistic reasoning “which may be produced by a machine” but becomes extremely clear in the case of the “logic of relatives, where any premise whatever will yield an endless series of conclusions, and attention has to be directed to the particular kind of conclusion desired” (Peirce, 1986, pp. 11-23).

In ordinary geometrical proofs auxiliary constructions are present in terms of “conveniently chosen” figures and diagrams where strategic moves are important aspects of deduction. The system of reasoning exhibits a dual character: deductive and “hypothetical”. Also in other – for example logical - deductive frameworks there is room for strategical moves which play a fundamental role in the generations of proofs. These strategical moves correspond to particular forms of abductive reasoning.

We know that the kind of reasoned inference that is involved in creative abduction goes beyond the mere relationship that there is between premises and conclusions in valid deductions, where the truth of the premises guarantees the truth of the conclusions, but also beyond the relationship that there is in probabilistic reasoning, which renders the conclusion just more or less probable. On the contrary, we have to see creative abduction as formed by the application of *heuristic procedures* that involve all kinds of good and bad inferential actions, and not only the mechanical application of rules. It is only by means of these heuristic procedures that the acquisition of *new* truths is guaranteed. Also Peirce’s mature view illustrated above on creative abduction as a kind of inference seems to stress the strategic component of reasoning.

Many researchers in the field of philosophy, logic, and cognitive science have sustained that deductive reasoning also consists in the employment of logical rules in a heuristic manner, even maintaining the truth preserving character: the application of the rules is organized in a way that is able to recommend a particular course of actions instead of another one. Moreover, very often the heuristic procedures of deductive reasoning are performed by means of a model-based abduction where iconicity is central. We have seen that the most common example of creative abduction is the usual experience people have of solving problems in geometry in a *model-based* way trying to devise proofs using diagrams and illustrations: of course the attribute of creativity we give to abduction in this case does not mean that it has never been performed before by anyone or that it is original in the history of some knowledge.

Hence we have to say that theoretical model-based abductions – as so iconicity - also operate in deductive reasoning. Following Hintikka and Remes’s analysis (1974) proofs of general implication in first order logic need the use of instantiation rules by which “new” individuals are introduced, so they are “ampliative”. In ordinary geometrical proofs auxiliary constructions are present in term of "conveniently chosen" figures and diagrams. In Beth's method of semantic tableaux the strategic "ability" to construct impossible configurations is undeniable (Hintikka, 1998; Niiniluoto, 1999).[18]

This means that also in many forms of deductive reasoning there are not only trivial and mechanical methods of making inferences but we have to use *models* and *heuristic procedures* that refer to a whole set of strategic principles. All the more reason that Bringsjord (1998) stresses his attention on the role played by a kind of "model based deduction" that is "part and parcel" of our establishing Gödel's first incompleteness theorem, showing the model-based character of this great abductive achievement of formal thought.[19]

I think the previous considerations also hold for Peircian theorematic reasoning: indeed Peirce further distinguished a "corollarial" and a "theoric" part within "theorematic reasoning", and connects theoric aspects to abduction (Hoffmann, 1999, p. 293). Of course, as already stressed, we have to remember this abductive aspect of mathematical reasoning is not in itself creative. It can be performed both in creative (to find new theorems and mathematical hypotheses) and non creative (merely "selective") ways, for example in the case we are using diagrams to demonstrate already known theorems (for instance in didactic settings), where selecting the strategy of manipulations is among chances not necessarily unknown and the result is not new. With respect to abduction in empirical sciences abduction in mathematics aims at hypothesizing ideal objects, which later we can possibly insert in a deductive apodictic and truth preserving framework.

The example of diagrams in geometry furnishes a semiotic and epistemological example of the nature of the cognitive interplay between internal neuronal representations (and embodied "cognitive" kinesthetic abilities) and external representations I have illustrated above: also for Peirce, more than a century before the new ideas derived from the field of distributed reasoning, the two aspects are intertwined in the pragmatic and semiotic view, going beyond the rigidity of the Kantian approach in terms of schematism. Diagrams are icons that take material and semiotic form in an external environment endowed with

[18] Also Aliseda (1997, 2006) provides interesting use of the semantic tableaux as a constructive representation of theories, where abductive expansions and revisions, derived from the belief revision framework, operate over them. The tableaux are so viewed as a kind of reasoning where the effect of "deduction" is performed by means of abductive strategies.

[19] Many interesting relationships between model-based reasoning in creative reasoning and its possible deductive models are analyzed in Batens (2006), Meheus (1999), Meheus et-al. (2002), Meheus and Batens (2006), also related to the formal treatment of inconsistencies.

- constraints depending on the specific cognitive delegation performed by human beings and
- the particular intrinsic constraints of the materiality at play.

Concrete manipulations on them can be done for instance to get new data and cognitive information and/or to simplify the problem at issue (cf. the epistemic templates illustrated above in section 6.1)

6.4 The Semiosis of Re-Embodiment

Some interesting semiotic aspects of the above illustrated process can be nicely analyzed. Imagine that a suitable *fixed internal record* exists – deriving from the cognitive exploitation of the previous suitable interplay with *external structures* - at the level of neural activation and that for instance it embeds an abstract concept endowed with all its features, for example the concept of triangle. Now, the human agent, via neural mechanisms and bodily actions, can "re-embody" that concept by making an external perceivable *sign*, for instance available to the attention of other human or animal senses and brains. For instance that human agent can use what in semiotics is called a *symbol* (with its conventional character: *ABC*, for example), but also an *icon* of relations (a suitable diagram of a triangle), or a *hybrid representation* that will take advantage of both. In Peircian terms:

> A representation of an idea is nothing but a sign that calls up another idea. When one mind desires to communicate an idea to another, he embodies his idea by making an outward perceptible image which directly calls up a like idea; and another mind perceiving that image gets a like idea. Two persons may agree upon a conventional sign which shall call up to them an idea it would not call up to anybody else. But in framing the convention they must have resorted to the primitive diagrammatic method of embodying the idea in an outward form, a picture. Remembering what likeness consists in, namely, in the natural attraction of ideas apart from habitual outward associations, I call those signs which stand for their likeness to them *icons*.
>
> Accordingly, I say that the only way of directly communicating an idea is by mean of an icon; and every indirect method of communicating an idea must depend for its establishment upon the use of an icon (Peirce, *MS* 787 *CSP* 26-28). [20]

[20] We have to note that for Peirce an idea "[…] is not properly a conception, because a *conception* is not an idea at all, but a *habit*. But the repeated occurrence of a general idea and the experience of its *utility*, results in the formation or strengthening of that habit which is the conception" (Peirce, *CP*, 7.498).

It is well-known that for Peirce every picture is a icon and thus every diagram, even if it lacks a sensuous similarity with the object, but just exhibits an analogy between the relations of the part of it and of the object:

> Particularly deserving of notice are icons in which the likeness is aided by conventional rules. Thus, an algebraic formula is an icon, rendered such by the rules of commutation, association, and distribution of the symbols; that it might as well, or better, be regarded as a compound conventional sign. It may seem at first glance that it is an arbitrary classification to call an algebraic expression an icon; that it might as well, or better, be regarded as a compound of conventional sign. But it is not so. For a great distinguishing property of the icon is that by direct observation of it other truths concerning its object can de discovered than those which suffice to determine its construction. Thus, by means of two photographs a map can be drawn, etc. Given a conventional or other general sign of an object, to deduce any other truth than which it explicitly signifies, it is necessary, in all cases, to replace that sign by an icon. This capacity of revealing unexpected truth is precisely that wherein the utility of algebraic formulae consists, so that the icon in character is the prevailing one (Peirce, *MS* 787 *CSP* 26-28).

Stressing the role of iconic dimensions of semiosis[21] in the meantime celebrates the virtues of analogy, as a kind of "association by resemblance", as contrasted to "association by contiguity".

Human beings delegate cognitive features to external representations through semiotic attributions because for example in many problem solving situations the internal computation would be impossible or it would involve a very great effort because of human mind's limited capacity. First a kind of "alienation" is performed, second a recapitulation is accomplished at the neuronal level by re-representing internally that which was "discovered" outside. Consequently only later on we perform cognitive operations on the structure of data that synaptic patterns have "picked up" in an analogical way from the environment. We can maintain that internal representations used in cognitive processes like many events of *meaning creation* have a deep origin in the experience lived in the semiotic environment.

I have already indicated that I think there are two kinds of artefacts that play the role of *external objects* (representations) active in this process of disembodiment of the mind: *creative* and *mimetic*. Mimetic external representations mirror concepts and problems that are already represented in the

[21] We have to remember that in this perspective any proposition is a diagram as well, because it represents a certain relation of symbols and indices.

brain and need to be enhanced, solved, further complicated, etc. so they sometimes can creatively give rise to new concepts and meanings.

Following my perspective it is at this point evident that the “mind” transcends the boundary of the individual and includes parts of that individual’s environment. It is in this sense that the mind is semiotic and artificial.

6.5 External Diagrammatization and Iconic Brain Co-Evolution

Following our previous considerations it would seem that diagrams can be fruitfully seen from a semiotic perspective as external representations expressed through icons and symbols, aimed at simply “mimicking” various humans’ internal images. However, they can also play the role of creative representations human beings externalize and manipulate not just to mirror the internal ways of thinking of human agents but to find room for concepts and new ways of inferring which cannot – at a certain time – be found internally “in the mind”.

In summary, we can say that

- diagrams as external iconic (often enriched by symbols) representations are formed by external materials that either mimic (through reification) concepts and problems already internally present in the brain or creatively express concepts and problems that do not have a semiotic “natural home” in the brain;
- subsequent internalized diagrammatic representations are internal re-projections, a kind of recapitulations, (learning) in terms of neural patterns of activation in the brain (“thoughts”, in Peircian sense), of external diagrammatic representations. In some simple cases complex diagrammatic transformations – can be “internally” manipulated *like* external objects and can further originate new internal reconstructed representations through the neural activity of transformation and integration.

We have already stressed that this process explains – from a cognitive point of view – why human agents seem to perform both computations of a connectionist type such as the ones involving representations as

- (I Level) patterns of neural activation that arise as the result of the interaction (also presemiotic) between body and environment (and suitably shaped by the evolution and the individual history): pattern completion or image recognition, and computations that use representations as

- (II Level) derived combinatorial syntax and semantics dynamically shaped by the various artificial external representations and reasoning devices found or constructed in the semiotic environment (for example iconic representations); they are – more or less completely – neurologically represented contingently as patterns of neural activations that “sometimes” tend to become stabilized meaning structures and to fix and so to permanently belong to the I Level above.

It is in this sense we can say the “System of Diagrammatization”, in Peircian words, allows for a self-controlled process of thought in the fixation of originally vague beliefs: as a system of learning, it is a process that leads from “absolutely undefined and unlimited possibility” (Peirce, *CP*, 6.217) to a fixation of belief and “by means of which any course of thought can be represented with exactitude” (Peirce, *CP*, 4.530). Moreover, it is a system which could also improve other areas of science, beyond mathematics, like logic, it “greatly facilitates the solution of problems of Logic. [...] If logicians would only embrace this method, we should no longer see attempts to base their science on the fragile foundations of metaphysics or a psychology not based on logical theory” (Peirce, *CP*, 4.571).

As already stressed the I Level originates those sensations (they constitute a kind of “face” we think the world has), that provide room for the II Level to reflect the structure of the environment, and, most important, that can follow the computations suggested by the iconic external structures available. It is clear that in this case we can conclude that the growth of the brain and especially the synaptic and dendritic growth are profoundly determined by the environment. Consequently we can hypothesize a form of co-evolution between what we can call the *iconic brain* and the development of the external representational systems. Brains build iconic signs as diagrams in the external environment learning from them new meanings through interpretation (both at the spatial and sentential level) after having manipulated them.

When the fixation is reached – imagine for instance the example above, that fixes the sum of the internal angles of the triangle – the pattern of neural activation no longer needs a direct stimulus from the external spatial representation in the environment for its construction and can activate a “final logical interpretant”, in Peircian terms. It can be neurologically viewed as a fixed internal record of an external structure (a fixed belief in Peircian terms) that can exist also in the absence of such external structure. The pattern of neural activation that constitutes the I Level Representation has kept record of the experience that generated it and, thus, carries the II Level Representation associated to it, even if in a different form, the form of *semiotic memory* and not the form of the vivid *semiotic sensorial experience* for example of the triangular construction drawn externally, over there, for instance in a blackboard. Now, the human agent, via neural mechanisms, can retrieve that II Level Representation and use it as an internal representation (and can use it to construct new internal representations less complicated than the ones previously available and stored in memory).

At this point we can easily understand the particular *mimetic* and *creative* role played by external diagrammatic representations in mathematics:

1. some concepts, meanings, and "ways of inferring" performed by the biological human agents appear hidden and tacit and can be rendered explicit by building external diagrammatic mimetic models and structures; later on the agent will be able to pick up and use what was suggested by the constraints and features intrinsic and immanent to their external semiotic materiality and the relative established conventionality: artificial languages, proofs, examples, etc.;
2. some concepts, meanings, and "new ways of inferring" can be discovered only through a problem solving process occurring in a distributed interplay between brains and external representations. I have called this process disembodiment of the mind: the representations are mediators of results obtained and allow human beings

    (a) to re-represent in their brains new concepts, meanings, and reasoning devices picked up outside, externally, previously absent at the internal level and thus impossible: first, a kind of alienation is performed, second, a recapitulation is accomplished at the neuronal level by re-representing internally that which has been "discovered" outside. We perform cognitive geometric operations on the structure of data that synaptic patterns have "picked up" in an analogical way from the explicit iconic representations in the environment;

    (b) to re-represent in their brains portions of concepts, meanings, and reasoning devices which, insofar as explicit, can facilitate inferences that previously involved a very great effort because of human brain's limited capacity. In this case the thinking performance is not completely processed internally but in a hybrid interplay between internal (both tacit and explicit) and external iconic representations. In some cases this interaction is between the internal level and a computational tool which in turn can exploit iconic representations to perform inferences (cf. above section 6.1).

An evolved mind is unlikely to have a natural home for complicated concepts like the ones geometry introduced, as such concepts do not exist in a definite way in the natural (not artificially manipulated) world: so whereas evolved minds could construct spatial frameworks and perform some trivial spatial inferences in a more or less tacit way by exploiting modules shaped by natural selection, how could one think exploiting explicit sophisticated geometrical concepts without having picked them up outside, after having produced them?

A mind consisting of different separated implicit templates of thinking and modes of inferences exemplified in various exemplars expressed through natural language cannot come up with certain mathematical and geometrical entities without the help of the external representations. The only way is to extend the mind into the material world, exploiting paper, blackboards, symbols, artificial languages, and other various semiotic tools, to provide *semiotic anchors* for finding ways of inferring that have no natural home within the mind, that is for finding ways of inferring that take us beyond those that natural selection and cultural training could enable us to possess at a certain moment.

Hence, we can hypothesize – for example – that many valid spatial reasoning habits which in human agents are performed internally have a deep origin in the past experience lived in the interplay with iconic systems at first represented in the environment. As I have just illustrated other recorded thinking habits only partially occur internally because they are hybridized with the exploitation of already available or suitably constructed external diagrammatic artefacts.

6.6 Delegated and Intrinsic Constraints in External Agents

We have said that through the cognitive interplay with external representations the human agent is able to pick up and use what suggested by the constraints and features intrinsic to their external materiality and to their relative established conventionality: artificial languages, proofs, examples, etc. Let us consider the example above (section 6.2) of the sum of the internal angles of a triangle. At the beginning the human agent – that is an interpretant in Peircian sense - embodies a sign in the external world that is in this case an icon endowed with “intentional” delegated cognitive conventional and public features – meanings - that resort to some already known properties of the Euclidean geometry: a certain language and a certain notation, the definition of a triangle, the properties of parallel lines that also hold in case of new elements and “auxiliary” constructions obtained through manipulation, etc. Then he looks, through diagram manipulations, for possible necessary consequences that occur over there, in the diagram/icon and that obey both

- the conventional *delegated* properties and
- the properties *intrinsic* to the materiality of the model.

This external model is a kind of autonomous cognitive *agent* offered to new interpretants of the problem/object in question. The model can be picked up later and acknowledged by the human agent through fixation of a new neural configuration – a new "thought" (in the case the new result concerning the sum of the internal angles).

The distinction above between delegated and intrinsic and immanent properties is also clear if we adopt the Peircian semiotic perspective. Peirce – speaking about the case of syllogistic logic, and not of geometry or algebra - deals with this problem by making an important distinction between what is going on in the brain of the logical human agent and the autonomous power of the chosen external system of representation or diagrammatization (Hoffmann, 2003). The presence of this "autonomous power" explains why I attribute to the system of representation a status of cognitive agency similar to the one of a human person, even if of course lacking aspects like direct intention and responsibility. Any diagram, Peirce says, makes use

> [...] of a particular system of symbols - a perfectly regular and very limited kind of language. It may be a part of a logician's duty to show how ordinary ways of speaking and of thinking are to be translated into that symbolism of formal logic; but it is no part of syllogistic itself. Logical principles of inference are merely rules for the illative transformation of the symbols of the particular system employed. If the system is essentially changed, they will be quite different (Peirce, *CP*, 2.599).

Of course the argumentation above also holds for our case of iconic geometric representation. This distinction integrates the one I have introduced above in the two levels of representations, and in some sense blurs it by showing how the *hybrid* character of the system composed by the two levels themselves, where the whole package of sensorial and kinesthetic abilities are involved.

The construction of the diagram also depends on those delegated semiotic properties that are embedded in what Peirce calls "precept": he says in the passage we have already quoted above that mathematical reasoning "[…] consists in constructing a diagram according to a general precept" (Peirce, *CP*, 1.54) and not only on the constraints expressed by the materiality of the model itself.[22]

Pickering (1995) depicts the role of some externalities (representations, artefacts, tools, etc.) in terms of a kind of non-human agency that interactively stabilizes with human agency in a dialectic of resistance and accommodation (p. 17 and p. 22). The two agencies, for example in scientific reasoning, originate a co-production of cognition the results of which cannot be presented and identified in advance: the outcome of the co-production is intrinsically "unpredictable". Latour's notions of the de-humanizing effect of technologies are based on the so-called "actor network theory",[23] which also stresses the semiotic role of externalities like the so-called non human agents. The actor network theory basically maintains that we should think of science, technology, and society as a field of human and non-human (material) agency. Human and non-human agents are associated with one another in networks, and they evolve together within these networks. Because the two aspects are equally important, neither can be reduced to the other: "An actor network is simultaneously an actor whose activity is networking heterogeneous elements and a network that is able to redefine and transform what is it made of […]. The actor network is reducible neither to an actor alone nor to a network" (Callon, 1997, p.93).

The operation on a diagram has reduced complexity enabling concentration on essential relations and has revealed new data. Moreover, through manipulations of the diagram new perspectives are offered to the observation, or interesting anomalies with respect the internal expectations are discovered. In the case of mathematicians, Peirce maintains, the diagram "puts before him an icon by the observation of which he detects relations between parts of the diagram other than those which were used in its construction" (Peirce, 1976, III, p. 749): "unnoticed and hidden relations among the parts" are discovered (Peirce, *CP*, 3.363). This activity is a kind of "thinking through doing": "In geometry, subsidiary lines are drawn. In algebra permissible transformations are made. Thereupon, the faculty of observation is called into play. […] Theorematic reasoning invariably depends upon experimentation with individual schemata" (Peirce, *CP*, 4.233).

We have said that firstly the human agent embodies a sign in the external world that is in this geometrical case an icon endowed with "intentional" delegated cognitive conventional and public features – meanings - that resort to some already known properties of the Euclidean geometry: these features

[22] It is worth noting that this process is obviously completely related to the Peircian idea of pragmatism (Hoffmann, 2004), that he simply considers "the experimental method" which is the procedure of all science.

[23] This theory has been proposed by Michel Callon, Latour himself, and John Law (cf. Callon, 1994, 1997, Latour, 1987, 1988, Callon and Latour, 1992, and Law, 1993).

can be considered a kind of immanent rationality and regularity (Hoffmann, 2004) that establishes a disciplinary field to envisage conclusions.[24] The system remains relative to the chosen conventional framework. They are real as long as there is no serious doubt in their adequacy: "The 'real,' for Peirce, is part of an evolutionary process and while 'pragmatic belief' and unconscious habits might be doubled from a scientific point a view, such a science might also formulate serious doubts in its own representational systems" (cit., p. 295).

Let us imagine we choose a different representational system still exploiting material and external diagrams. Through the manipulation of the new symbols and diagrams we expect very different conclusions. An example is the one of the non-Euclidean discoveries. In Euclidean geometry, by adopting the postulate of parallels we necessarily arrive to the ineluctable conclusion that the sum of internal angles of a triangle is 180°, but this does not occur in the case of the non-Euclidean geometry, that I will illustrate in the following section.

6.7 Mirror and Unveiling Diagrams

It is well-known that in the whole history of geometry many researchers used internal mental imagery and mental representations of diagrams, but also self-generated diagrams (external) to help their thinking (Otte and Panza, 1997). For example, it is clear that in geometrical construction many of the requirements indicated by the manipulative templates (cf. above section 6.1) are fulfilled. Indeed iconic geometrical constructions present situations that are curious and "at the limit". Because of their iconicity, they are constitutively dynamic, artificial, and offer various contingent ways of epistemic acting, like looking from different perspectives, comparing subsequent appearances, discarding, choosing, re-ordering, and evaluating. Moreover, they present the features typical of manipulative reasoning illustrated above, such as the simplification of the task and the capacity to get visual information otherwise unavailable.

We have seen that manipulative abduction is a kind of abduction, usually model-based and so intrinsically "iconic", that exploits external models endowed with delegated (and often implicit) cognitive and semiotic roles and attributes. In section 6.3. above we have said that 1) the model (diagram) is

[24] Paavola, Hakkarainen, & Sintonen (2006) consider the interplay between internal and external aspects of abductive reasoning in the framework of the interrogative model of the so-called "explanation-seeking why-questions". They emphasize the interaction with the "environment" and show the importance of the heuristic strategies and of their trialogic nature (inquirer and fellow inquirers, object of inquiry, mediating artefacts and processes), also taking advantage of Davidson's ideas concerning triangulation.

external and the strategy that organizes the manipulations is unknown *a priori*; 2) the result achieved is *new* (if we, for instance, refer to the constructions of the first creators of geometry), and adds properties not contained before in the concept (the Kantian to "pass beyond" or "advance beyond" the given concept, Kant, 1929, A154-B193/194, p. 192).

Hence, in the construction of mathematical concepts many external representations are exploited, both in terms of diagrams and of symbols. I am interested in my research in the diagrams which play various iconic roles: an *optical* role - microscopes (that look at the infinitesimally small details), telescopes (that look at infinity), windows (that look at particularly situation), a *mirror* role (to externalize rough mental models), and an *unveiling* role (to help to create new and interesting mathematical concepts, theories, and structures). I also describe them as the *epistemic mediators* (cf. above) able to perform various abductive tasks (discovery of new properties or new propositions/hypotheses, provision of suitable sequences of models able to convincingly verifying theorems, etc.). Elsewhere I have presented some details concerning the role of optical diagrams in the calculus (Magnani and Dossena, 2003).

We have seen that diagrams serve an important role in abduction mainly because they can be manipulated. In mathematics diagrams play various roles in a typical abductive way. Two of them are central:

- they provide an intuitive *explanation* able to help the understanding of concepts difficult to grasp or that appear obscure and/or epistemologically unjustified,[25]
- they help to *create* new previous unknown concepts, as illustrated in the case of the non-Euclidean geometry.

In the case of the construction and examination of diagrams in geometrical reasoning, specific experiments serve as states and the implied operators are the manipulations and observations that transform one state into another. The geometrical outcome is dependent upon practices and specific sensory-motor activities performed on a non symbolic object, which acts as a dedicated external representational medium supporting the various operators at work. We have illustrated in the previous sections that there is a kind of an epistemic and semiotic negotiation between the sensory framework of the geometer and the external reality of the diagram. This process involves an *external representation* consisting of written symbols and figures that are manipulated "by hand". The cognitive system is not merely the mind-brain

[25] Some new optical diagrams (microscopes within microscopes), which provide new mental representations of the concept of tangent line at the infinitesimally small regions, are introduced in the already cited Magnani & Dossena (2003).

of the person performing the geometrical task, but the system consisting of the whole body (cognition is *embodied*) of the person plus the external physical representation. In geometrical discovery the whole activity of cognition is located in the system consisting of a human together with diagrams.

We stated above that in mathematics mirror and unveiling diagrams play various roles in a typical abductive way. Now we can add that:

- they are *epistemic mediators* able to perform various abductive tasks in so far as
- they are *external representations* which, in the cases we will present in the following sections, are devoted to provide explanatory abductive results.

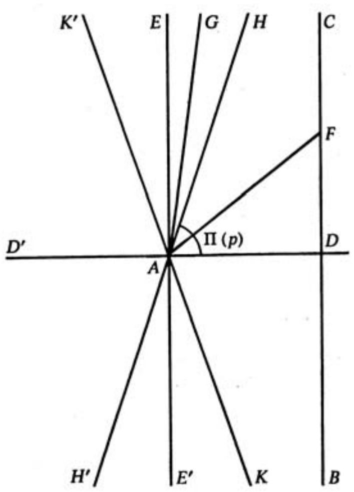

Figure 3 .Parallel lines.

Let us consider some aspects of the role of mirror and unveiling diagrams in the Lobachevskyan discovery of the elementary non-Euclidean geometry. The example indicates that the use of a different background – different delegated cognitive conventional and public geometrical features, non-Euclidean – (cf. previous section), new icons endowed with the same materiality (and related constraints) of the ones exploited in the Euclidean case, can lead to different new results.

A *mirror diagram* (for example the diagram of the drawn parallel lines - cf. Figure 3 - Lobachevsky, 1891) is a kind of external *analogous* both of the mental image we depict in the mental visual buffer and of the symbolic-propositional level of the postulate definition (the fifth postulate). In general this diagram mirrors the internal imagery (thus it is a sign of a "thought", in Peircian terms) and provides the possibility of detecting anomalies. The external representation of geometrical structures often activates direct perceptual operations (for example, identify the parallels and search for the limits) to elicit consistency or inconsistency routines. Sometimes the mirror diagram biases are inconsistent with the task and so they can make the task more difficult by misguiding actions away from the goal. If consistent, they can make the task easier by guiding actions toward the goal. In certain cases the mirror diagrams biases are irrelevant, they should have no effects on the decision of abductive actions, and play lower cognitive roles. In the case of Figure 3 the diagram of the parallel lines was used in the history of geometry to make both consistent and inconsistent the fifth Euclidean postulate and the new non-Euclidean perspective (more details on this epistemological situation are given in Magnani, 2002).

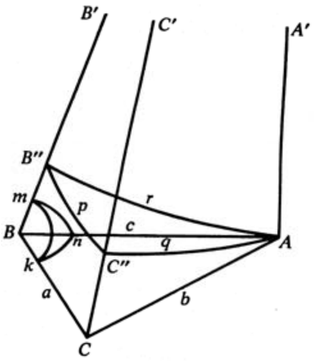

Figure 4. Euclidean/non-Euclidean model.

An example of *unveiling diagram* is the one illustrated by the Figure 4 (Lobachevsky, 1891). It is more abstract than the previous one and exploits "audacious" representations in the perspective of three dimensional geometrical shapes. The construction given in the figure aims at iconically "representing" a stereometric non-Euclidean form built on a rectilinear right angled triangle *ABC* to which theorems previously proved (for example, the one stating that the parallels *AA'*, *BB'*, *CC'*, which lie on the three planes are parallels in non-Euclidean sense) can be applied. In this way Lobachevsky is able to further apply symbolic identifications and to arrive to new equations which consistently (and at the same time) connect Euclidean and non-Euclidean perspectives. This kind of diagram strongly guides the geometer's selections of moves by eliciting what I call the *Euclidean-inside non-Euclidean* "model matching strategy". This maneuver also constitutes an important step in the affirmation of the modern "scientific" concept of model. This unveiling diagram constitutes a kind of gateway to the unexpected meanings of new imaginary entities.

In general we have to note that some perceptions activated by the diagram are of course disregarded as irrelevant to the task, as it usually happens when exploiting external diagrammatic representations in reasoning processes. Because not everything in external representations is always relevant to a task, high level cognitive mechanisms need to use task knowledge (usually supplied by task instructions, geometrical, in our case) to direct attention and perceptual processes to the relevant features of external representations. This external representation in terms of an unveiling diagram activates a perceptual reorientation in the external construction (that is identifies possible further constructions and enhancements); in the meantime the generated internal representation of the external elements activates directly retrievable information (numerical values) that elicits the strategy of building further non-Euclidean structures together with their *analytic* ("symbolic", in semiotic terms) counterpart (the non-Euclidean trigonometry equations).

The different selected representational system – that still uses Euclidean icons - determines in this case quite different possibilities of constructions, and thus different results from iconic experimenting. New results are derived in diagrammatic reasoning through modifying the representational systems, adding new meaning to them, or in reconstructing their systematic order.

Many commentators (and myself in Magnani, 2001b) contend that Kant did not imagine that non-Euclidean concepts could in some way be constructed in *intuition* (a Kantian expression which indicated our iconic external representation), through the mediation of a model, that is preparing and constructing a Euclidean model of a specific non-Euclidean concept (or group of concepts). Yet Kant also wrote that ‘‘the use of geometry in natural philosophy would be insecure, unless the notion of space is originally given by the nature of the mind (so that if anyone tries to frame in his mind any relations different from those prescribed by space, he will labor in vain, for he will be compelled to use that very notion in support of his figment)" (Kant, 1770, Section 15E). Torretti (2003, p .160) observes:

> I find it impossible to make sense of the passage in parentheses unless it refers precisely to the activity of constructing Euclidean models of non-Euclidean geometries (in a broad sense). We now know that one such model (which we ought rather to call quasi-Euclidean, for it would represent plane Lobachevskian geometry on a sphere with radius $\sqrt{-1}$ is mentioned in the *Theorie der Parallellinien* that Kant's fellow Königsbergian Johann Heinrich Lambert (1786) wrote about 1766. There is no evidence that Kant ever saw this tract and the few extant pieces of his correspondence with Lambert do not contain any reference to the subject, but, in

> the light of the passage I have quoted, it is not unlikely that Kant did hear about it, either from Lambert himself, or from a shared acquaintance, and raised the said objection.

I agree with Torretti, Kant had a very wide perspective about the resources of “intuition”, anticipating that a geometer would have been “compelled” to use the notion of space “given by nature”, that is the one that is at the origins of our external representation, “in support of his figment”, for instance the non-Euclidean Lobachevskyan abstract structures we have treated above in Figure 4. that exhibits the non-Euclidean through the Euclidean. Nevertheless, while Kant conceives the a priori forms and categories as absolute and unchangeable conditions of possible experience and “Erkenntnis,” the central idea of Peirce’s evolutionary philosophy is expressed in his claim that laws, general rules, and forms in themselves are the results of evolution. Peirce’s generally evolutive orientation is the main feature that distinguishes him from Kant.

7. MIMETIC MINDS AS SEMIOTIC MINDS

I contend that there are external representations that are representations of other external representations. In some cases they carry new scientific knowledge. To make an example, Hilbert’s *Grundlagen der Geometrie* is a “formal” representation of the geometrical problem solving through diagrams: in Hilbertian systems solutions of problems become proofs of theorems in terms of an axiomatic model. In turn a calculator is able to re-represent (through an artifact) (and to perform) those geometrical proofs with diagrams already performed by human beings with pencil and paper. In this case we have representations that *mimic* particular cognitive performances that we usually attribute to our *minds*.

We have seen that our brains delegate cognitive (and epistemic) roles to externalities and then tend to “adopt” and recapitulate what they have checked occurring outside, over there, after having manipulated – often with creative results – the external invented structured model. A simple example: it is relatively neurologically easy to perform an addition of numbers by depicting in our *mind* – thanks to that brain device that is called visual buffer – the images of that addition *thought* as it occurs concretely, with paper and pencil, taking advantage of external materials. We have said that mind representations are also over there, in the environment, where mind has objectified itself in various semiotic structures that *mimic* and *enhance* its internal representations.

Turing adds a new structure to this list of external objectified devices: an abstract tool, the (Universal) Logical Computing Machine (LCM), endowed with powerful mimetic properties. We have concluded the section 3.3 remarking that the creative "mind" is in itself extended and, so to say, both internal and external: the mind is *semiotic* because transcends the boundary of the individual and includes parts of that individual's environment, and thus constitutively artificial. Turing's LCM, which is an externalized device, is able to mimic human cognitive operations that occur in that interplay between the internal mind and the external one. Indeed Turing already in 1950 maintains that, taking advantage of the existence of the LCM, "Digital computers […] can be constructed, and indeed have been constructed, and […] they can in fact mimic the actions of a human computer very closely"(Turing, 1950, p. 435).

In the light of my perspective both (Universal) Logical Computing Machine (LCM) (the theoretical artifact) and (Universal) Practical Computing Machine (PCM) (the practical artifact) are *mimetic minds* because they are able to mimic the mind in a kind of universal way (wonderfully continuing the activity of disembodiment of minds and of semiotic delegations to the external materiality our ancestors rudimentary started). LCM and PCM are able to re-represent and perform in a very powerful way plenty of cognitive skills of human beings.

Universal Turing machines are discrete-state machines, DMS, "with a Laplacian behavior" (Longo, 2002; Lassègue, 1998, 2002): "it is always possible to predict all future states") and they are equivalent to all formalisms for computability (what is thinkable is calculable and mechanizable), and because universal they are able to simulate – that is to *mimic* – any human cognitive function, that is what is usually called mind.

A natural consequence of this perspective is that Universal Turing machines do not represent (against classical AI and modern cognitivist computationalism) a "knowledge" of the mind and of human intelligence. Turing is perfectly aware of the fact that brain is not a DSM, but as he says, a "continuous" system, where instead a mathematical modeling can guarantee a satisfactory scientific intelligibility (cf. his studies on morphogenesis).

We have seen that our brains delegate meaningful semiotic (and of course cognitive and epistemic) roles to externalities and then tend to "adopt" what they have checked occurring outside, over there, in the external invented structured and model. And a large part of meaning formation takes advantage of the exploitation of external representations and mediators.

Our view about the disembodiment of mind certainly involves that the Mind/Body dualist view is less credible as well as Cartesian computationalism. Also the view that Mind is Computational independently of the physical (functionalism) is jeopardized. In my perspective on human cognition in terms of mimetic minds we no longer need Descartes dualism: we only have *semiotic brains* that make up large, integrated, material cognitive systems like for example LCMs and PCMs. These are new independent semiotic agencies that constitute real artificial minds aiming at "universally" imitating human cognition. In this perspective what we usually call mind simply consists in the union of both the changing neural configurations of brains together with those large, integrated, and material cognitive systems the brains themselves are continuously building in an infinite semiotic process.

Minds are material like brains, in so far as they take advantage of intertwined internal and external semiotic processes. It seems to me at this point we can better and more deeply understand Peirce's semiotic motto "man is an external sign" in the passage we have completely quoted above in section 5.1: "[…] as the fact that every thought is a sign, taken in conjunction with the fact that life is a train of thoughts, proves that man is a sign; so, that every thought is an *external sign*, proves that man is an external sign" (Peirce, *CP*, 5.324).

The only problem seems "How meat knows": we can reverse the Cartesian motto and say "sum ergo cogito". In this perspective what we usually call mind simply consists in the union of both the changing neural configurations of brains together with those large, integrated, and material cognitive systems the brains themselves are continuously building.

8. CONCLUSION AND FUTURE TRENDS

The main thesis of this paper is that the disembodiment of mind is a significant cognitive perspective able to unveil some basic features of creative thinking and its computational problems. Its fertility in explaining the semiotic interplay between internal and external levels of cognition is evident. I maintain that various aspects of creative meaning formation could take advantage of the research on this interplay: for instance study on external mediators can provide a better understanding of the processes of explanation and discovery in science and in some areas of artificial intelligence related to mechanizing discovery processes.[26]

From the paleoanthropological perspective we have learnt that an evolved mind is unlikely to have a *natural home* for new concepts and meanings, as such concepts and meanings do not exist in the already known artificial and natural world: the cognitive referral to the central role of the relation between meaningful behaviour and dynamical interactions with the environment becomes critical to the problem of meaning formation. Finally, I

[26] On the recent achievements in the area of the machine discovery simulations of model-based creative tasks cf. (Magnani, Nersessian, & Pizzi, 2002).

think the cognitive role of what I call “mimetic minds” can be further studied also taking advantage of the research on hypercomputation. The imminent construction of new types of universal “abstract” and “practical” machines will constitute important and interesting new “mimetic minds” externalized and available over there, in the environment, as sources of mechanisms underlying the emergence of new meaning processes. They will provide new tools for creating meaning formation in classical areas like analogical, visual, and spatial inferences, both in science and everyday situations, so that this can extend the epistemological and the psychological theory.

The perspectives above, resorting to the exploitation of a very interdisciplinary interplay will further shed light on how concrete manipulations of external objects influence the generation of hypotheses and so on the characters of what I call manipulative abduction showing how we can find methods of constructivity – and their computational counterparts – in scientific and everyday reasoning based on external models and “epistemic mediators” (Magnani, 2004).

Another interesting application is given in the area of chance discovery (cf. Magnani, Piazza & Dossena, 2002): concrete manipulations of the external world constitute a fundamental passage in chance discovery: by a process of manipulative abduction it is possible to build semiotic prostheses that furnish a kind of embodied and unexpressed knowledge that holds a key role in the subsequent processes of scientific comprehension and discovery but also in ethical thinking and in moral deliberation. For example I have viewed moral reasoning as a form of “possible worlds” anticipation, a way of getting chances to shape the human world and act in it (Magnani, 2003). It could be of help to prefigure risks, possibilities, and effects of human acting, and to promote or prevent a broad variety of guidelines. Creating ethics means creating the world and its new semiotic directions, when facing different (real or abstract) situations and problems. In this way events and situations can be reinvented either as an opportunity or as a risk for new moral directions. I have also described some “templates” of manipulative behavior which account for the most common cognitive and moral acting related to chance discovery and chance production. I maintain this kind of research could be furthermore specifically addressed to the analysis of the construction of new meaning processes by chance.

REFERENCES

Agre, P. & Chapman, D. (1990). What are plans for? In P. Maes (Ed.) *Designing Autonomous Agents*. Cambridge, MA: MIT Press (pp. 17-34).

Aliseda, A. (1997). Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence. PhD Thesis. Amsterdam: Institute for Logic, Language and Computation.

Aliseda, A. (2006). *Abductive Reasoning. Logical Investigations into Discovery and Explanation*. Springer: Berlin.

Anderson, M.L. (2003). Embodied cognition: A field guide, *Artificial Intelligence* 149(1):91-130.

Batens, D. (2006), A diagrammatic proof search procedure as part of a formal approach to problem solving. In L. Magnani (Ed.). Model-Based Reasoning in Science and Engineering (pp. 265-284). London: College Publications.

Brent, J. (2000). A brief introduction to the life and thought of Charles Sanders Peirce. In J. Muller & J. Brent (Eds.), *Peirce, Semiosis, and Psychoanalysis*. Baltimore and London: John Hopkins (pp. 1-14).

Bringsjord, S. (2000). Is (Gödelian) model-based deductive reasoning computational? In Special Issue *Abduction and Scientific Discovery*, L. Magnani, N.J. Nersessian, & P. Thagard. *Philosophica*, 61, 51-76.

Brooks, R.A. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139-159.

Brooks, R.A. & Stein, L. (1994) Building brains for bodies. *Autonomous Robots* 1:7-25.

Callon, M. (1994). Four models for the dynamics of science. In: S. Jasanoff, G.E. Markle, J.C. Petersen, & T.J. Pinch (Eds.), *Handbook of Science and Technology Studies* (pp. 29-63). Los Angeles: Sage.

Callon, M. (1997). Society in the making: the study of technology as a tool for sociological analysis. In W.E. Bjiker, T.P. Hughes, & T. Pinch (Eds.), *The Social Construction of Technological Systems.* Cambridge, MA: MIT Press (pp. 83-106).

Callon M. & Latour, B. (1992). Don’t throw the baby out with the bath school! A reply to Collins and Yearley. In A. Pickering (Ed.), *Science as Practice and Culture*. Chicago and London: The University of Chicago Press, (pp. 343-368).

Carruthers, P. (2002). The cognitive function of language. *Behavioral and Brain Sciences*, 25(6):657-74.

Clancey, W.J. (2002). Simulating activities: Relating motives, deliberation, and attentive coordination. *Cognitive Systems Research* 3(1-4):471-500.

Clark, A. (1998). Magic words, in P. Carruthers & J. Boucher (Eds.), *Language and Thought. Interdisciplinary Themes.* Oxford University Press: Oxford (pp. 162-183).

Clark, A., 2003, *Natural-Born Cyborgs. Minds, Technologies, and the Future of Human Intelligence.* Oxford: Oxford University Press.

Clark, A. (2005). Material symbols: from translation to co-ordination in the constitution of , thought and reason. In B. Bara, L. Barsalou & M. Bucciarelli (Eds.), *CogSci 2005, XXVII Annual Conference of the Cognitive Science Society*, CD. Stresa, Italy.

Clowes, R. W. & A. Morse (2005). Scaffolding Cognition with Words. Accepted draft for the 5th International Workshop on Epigenetic Robotics.

Colapietro, V. (2000). Futher consequences of a singular capacity. In J. Muller & J. Brent (Eds.) *Peirce, Semiosis, and Psychoanalysis*. Baltimore and London: John Hopkins (pp. 136-58).

Dennett, D. (1991). *Consciousness Explained*. New York: Little, Brown, and Company.

Dennett, D. (2003). *Freedom Evolves.* New York: Viking.

Donald, M. (1998). Hominid enculturation and cognitive evolution. In C. Renfrew, P. Mellars, & C. Scarre (Eds.), *Cognition and Material Culture: The Archaeology of External Symbolic Storage*. Cambridge: The McDonald Institute for Archaeological Research, (pp. 7-17).

Donald, M. (2001). *A Mind so Rare. The evolution of Human Consciousness.* New York: W.W Norton & Company.

Gatti, A. & Magnani, L. (2005). On the representational role of the environment and on the cognitive nature of manipulations. In L. Magnani (Ed.), *Computing, Philosophy, and Cognition*, Proceedings of the European Conference of Computing and Philosophy, Pavia, Italy, 3-4 June 2004, forthcoming.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception.* New York: Houghton Mifflin, New York

Gooding, D. (1990). *Experiment and the Making of Meaning*. Dordrecht: Kluwer.

Hameroff, A.R., Kaszniak, A.W. & Chalmers, D.J. (Eds) (1999). *Toward a Science of Consciousness III. The Third Tucson Discussions and Debates*. Cambridge, MA: MIT Press, Cambridge.

Hintikka, J. (1998). What is abduction? The fundamental problem of contemporary epistemology. *Transactions of the Charles S. Peirce Society* 34, 503-533.

Hintikka, J. and Remes, U. (1974). *The Method of Analysis. Its Geometrical Origin and Its General Significance*. Dordrecht: Reidel.

Hoffmann, M.H.G. (1999). Problems with Peirce's Concept of Abduction, *Foundations of Science*, 4(3), 271–305.

Hoffmann, M.H.G. (2003), Peirce's 'Diagrammatic reasoning as a solution of the learning paradox. In G. Debrock (Ed.), *Process Pragmatism: Essays on a Quiet Philosophical Revolution*. Amsterdam: Rodopi Press (pp. 121-143).

Hoffmann, M.H.G. (2004), How to get it. Diagrammatic reasoning as a tool for knowledge development and its pragmatic dimension. *Foundations of Science*, 9, 285-305.

Humphrey, N. (2002). *The Mind Made Flesh*. Oxford: Oxford University Press.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Hutchins, E. (1999) Cognitive artifacts, in R.A.Wilson & F.C. Keil, *Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press (pp. 126-7).

Kant, I. (1929). *Critique of Pure Reason*, trans. by N. Kemp Smith. London: MacMillan. Reprint 1998; originally published 1787.

Kant, I., (1968). *Inaugural Dissertation on the Form and Principles of the Sensible and Intelligible World* (1770). In I. Kant, *Kant. Selected Pre-Critical Writings*, edited and translated by G.B. Kerferd & D.E. Walford. Manchester: Manchester University Press (pp. 45-92). Also translated by J. Handyside, *Dissertation on the Form and Principles of the Sensible and Intelligible World*. In I. Kant, *Kant's Inaugural Dissertation and Early Writings on Space*. Chicago: Open Court (1929) (pp. 35-85).

Kirsh, D. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18, 513-549.

Lambert, J. H. (1786). Theorie der Parallellinien'. *Magazin für die reine und angewandte Mathematik*, 2, 137–164; 3, 325–358. (Written about 1766; posthumously published by J. Bernoulli.)

Lassègue, J. (1998). *Turing.* Paris: Les Belles Lettres.

Lassègue, J. (2002). Turing entre formel et forme; remarque sur la convergence des perspectives morphologiques. *Intellectica*, 35(2), 185-198.

Latour, B. (1987). *Science in Action: How to follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press.

Latour, B. (1988). *The Pasteurization of France.* Cambridge, MA: Harvard University Press.

Latour, B. (1999). *Pandora's Hope*. Cambridge, MA: Harvard University Press.

Law, J. (1993). *Modernity, Myth, and Materialism*. Oxford: Balckwell.

Lewis-Williams, D. (2002). *The mind in the cave*, London: Thames and Hudson.

Lobachevsky, N.J. (1891). *Geometrical Researches on the Theory of Paralleles* [1840]. (Trans. by G.B. Halsted). Austin: University of Texas.

Longo, G. (2002). Laplace, Turing, et la géométrie impossible du "jeu de l'imitation": aléas, determinisme e programmes dans le test de Turing. *Intellectica*, 35(2), 131-161.

Magnani, L. (1992). Abductive reasoning: philosophical and educational perspectives in medicine. In Advanced Models of Cognition for Medical Training and Practice, David A. Evans & Vimla L. Patel (Eds.) (pp 21-41). Berlin: Springer.

Magnani, L. (2001a). *Abduction, Reason, and Science. Processes of Discovery and Explanation*. New York: Kluwer Academic/Plenum Publishers.

Magnani, L. (2001b). *Philosophy and Geometry. Theoretical and Historical Issues*. Dordrecht: Kluwer Academic.

Magnani, L. (2002). Epistemic mediators and model-based discovery in science. In L. Magnani & N.J. Nersessian (Eds.) (pp. 325-329). New York, NY: Kluwer Academic/Plenum Publishers.

Magnani, L. (2003). Moral mediators. Prefiguring ethical chances in a human world. In H. Shoji & Y. Matsuo (Eds.), *Proceedings of the 3rd International Workshop on Chance Discovery* (pp. 1-20). HCI International Conference, Greece.

Magnani, L. (2004). Conjectures and manipulations. Computational modeling and the extra-theoretical dimension of scientific discovery. *Minds and Machines* 14, 507-537.

Magnani, L. (2006). Multimodal abduction. External semiotic anchors and hybrid representations. *Logic Journal of IGPS*, 14(1), 107-136.

Magnani, L. (2007). *Knowledge as a Duty. Distributed Morality in a Technological World*. Cambridge: Cambridge University Press, in press.

Magnani, L., N.J. Nersessian, & Thagard, P. (Eds.) (1999). *Model-Based Reasoning in Scientific Discovery*. New York: Kluwer Academic/Plenum Publishers.

Magnani, L. & Dossena, R. (2003). Perceiving the infinite and the infinitesimal world: unveiling and optical diagrams and the construction of mathematical concepts. In *Proceedings of CogSci2003*. Boston, MA: CD-ROM produced by X-CD Technologies.

Magnani, L., Nersessian, N.J., & Pizzi, C. (2002). *Logical and Computational Aspects of Model-Based Reasoning.* Dordrecht: Kluwer Academic.

Magnani, L., Piazza, M. & Dossena, R. (2002). Epistemic mediators and chance morphodynamics. In A. Abe (Ed.), *Proceedings of PRICAI-02 Conference, Working Notes of the 2nd International Workshop on Chance Discovery* (pp. 38-46). Tokyo.

Meheus, J. (1999). Model-based reasoning in creative processes. In L. Magnani, N.J. Nersessian, & P. Thagard (Eds.) (1999) (pp. 199-217).

Meheus, J., Verhoeven, L., Van Dyck, M. and Provijn, D. (2002). Ampliative adaptive logics and the foundation of logic-based approaches to abduction. In .L. Magnani, N.J., Nersessian, and C. Pizzi (Eds.) (2002) (pp. 39-71).

Meheus, J. & Batens, D. (2006). A formal logic for abductive reasoning. In *Logic Journal of IGPL*, 14(1), 221-236.

Mithen, S (1996). *The Prehistory of the Mind*. *A Search for the Origins of Art, Religion, and Science*. London: Thames and Hudson.

Mithen, S. (1999). Handaxes and ice age carvings: hard evidence for the evolution of consciousness. In Hameroff, et al. (Eds.) (1999) (pp. 281-296).

Monekosso, N., P. Remagnino, & F. J. Ferri (2004). Learning machines for chance discovery. In A. Abe & R. Oehlmann (Eds.) (2004) *Workshop 4: The 1st European Workshop on Chance Discovery* (pp. 84-93), Valencia, Spain, August 2004.

Niiniluoto, I. (1999). Abduction and geometrical analysis. Notes on Charles S. Peirce and Edgar Allan Poe. In: L. Magnani, N.J. Nersessian, and P. Thagard (Eds.) (1999) (pp. 239-254).

Norman, D.A. (1993). *Things that Make Us Smart. Defending Human Attributes in the Age of the Machine*. Reading. MA: Addison-Wesley.

Otte, M. & Panza, M. (Eds.) (1997). *Analysis and Synthesis in Mathematics*. Dorcrecht: Kluwer Academic.

Paavola, S., Hakkarainen, K., & Sintonen, M. (2006). Abduction with dialogical and trialogical means. *Logic Journal of IGPS*, 14(1), 137-150.

Peirce, C.S. (1931-1958) (*CP*). *Collected Papers,* 8 vols. C. Hartshorne & P. Weiss (vols. I-VI), (Eds.), & A.W. Burks (vols. VII-VIII) (Ed.), Cambridge, MA: Harvard University Press.

Peirce, C.S. (1976). *The New Elements of Mathematics by Charles Sanders Peirce*. C. Eisele (Ed.) (vols. I-IV). The Hague-Paris/Atlantic Highlands, NJ: Mouton/Humanities Press.

Peirce, C.S. (1986). *Historical Perspectives on Peirce's Logic of Science: A History of Science* C. Eisele (Ed.) (vols. I-II). Berlin: Mouton.

Peirce, C.S. (*MS*) *The Charles S. Peirce Papers: Manuscript Collection in the Houghton Library*, Harvard University (numbered according to Richard S. Robin, Annotated *Catalogue of the Papers of Charles S. Peirce*. Worcester, Mass., 1967: The University of Massachusetts Press). Available in the Peirce Microfilm edition. Pagination: CSP = Peirce / ISP = Institute for Studies in Pragmaticism.

Piaget, J. (1974). *Adaptation and Intelligence.* Chicago, IL: University of Chicago Press.

Pickering, A. (1995). *The Mangle of Practice. Time, Agency, and Science*. Chicago and London: The University of Chicago Press.

Queiroz, J. & Merrell, F. (2005). Abduction: between subjectivity and objectivity. Special issue of *Semiotica*, 153.

Thomas, N.J. (1999). Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science*, 23 (2), 207-245.

Torretti, R. (2003). Review of L. Magnani, *Philosophy and Geometry:Theoretical and Historical Issues,* Dordrecht: Kluwer Academic Publishers (2001), *Studies in History and Philosophy of Modern Physics*, 34B(1), 158-160.

Turing, A.M. (1950). Computing machinery and intelligence. *Mind*, 49, 433-460. Also in A.M. Turing (1992) (pp. 133-160).

Turing, A.M. (1969). Intelligent machinery [1948]. In B. Meltzer & D. Michie (Eds.), *Machine Intelligence* 5, 3-23. Also in A.M. Turing (1992), pages 3-23.

Turing, A.M. (1992). *Collected Works of Alan Turing, Mechanical Intelligence*. Ed. by D.C. Ince. Amsterdam: Elsevier.

Woods, J. (2004). *The Death of Argument*. Dordrecht: Kluwer.

Zhang, J. (1997). The nature of external representations in problem-solving. *Cognitive Science*, 21, 179–217