An Abductive Theory of Scientific Reasoning

Lorenzo Magnani

Abstract— **What I call theoretical abduction (sentential and model-based) certainly illustrates much of what is important in abductive reasoning, especially the objective of selecting and creating a set of hypotheses that are able to dispense good (preferred) explanations of data, but fails to account for many cases of explanations occurring in science or in everyday reasoning when the exploitation of the environment is crucial. The concept of manipulative abduction is devoted to capture the role of action in many interesting situations: action provides otherwise unavailable information that enables the agent to solve problems by starting and performing a suitable abductive process of generation or selection of hypotheses. Many external things, usually inert from the epistemological point of view, can be transformed into what I call epistemic mediators, which are illustrated in the last part of this paper, together with an analysis of the related notion of of "external representation".**

Keywords— **Abduction, Scientific Discovery, Scientific Reasoning, Epistemic Mediators, External Representations.**

I. Theoretical Abduction

What is abduction? Many reasoning conclusions that do not proceed in a deductive manner are *abductive*. For instance, if we see a broken horizontal glass on the floor we might explain this fact by postulating the effect of wind shortly before: this is certainly not a deductive consequence of the glass being broken (a cat may well have been responsible for it). Hence, *theoretical* abduction [1] 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: 1) abduction that only generates "plausible" hypotheses (*selective* or *creative*) and 2) abduction considered as *inference to the best explanation*, which also evaluates hypotheses (cf. Figure 1). To illustrate from the field of medical knowledge, the discovery of a new disease and the manifestations it causes can be considered as the result of a creative abductive inference. Therefore, creative abduction 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. All we can expect of our "selective" abduction, is that it tends to produce hypotheses for further examination that have some chance of turning out to be the best explanation. Selective abduction will always produce hypotheses that give at least a partial explanation and therefore have a small amount of initial plausibility.

L. Magnani is with the Department of Philosophy and Computational Philosophy Laboratory, University of Pavia, Pavia, Italy and with the Program in Philosophy, Science, and Technology, Georgia Institute of Technology, Atlanta, GA, USA. E-mail: lmagnani@cc.gatech.edu.

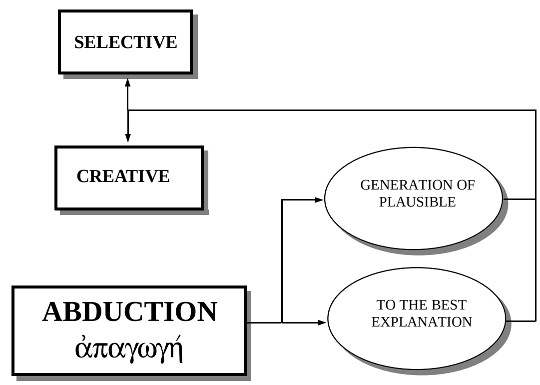


Fig. 1. Creative and selective abduction.

Finally, many attempts have been made to model abduction by developing some formal tools in order to illustrate its computational properties and the relationships with the different forms of deductive reasoning. This kind of sentential frameworks exclusively deals with selective abduction (diagnostic reasoning) and relates to the idea of preserving consistency. If we want to provide a suitable framework for analyzing the interesting cases of creative reasoning (in science too), we do not have to limit ourselves to the sentential view of theoretical abduction but we have to consider a broader inferential one which encompasses both sentential and what I call model-based elements of creative abduction.

II. The Logical Framework

An important way of modeling abduction resorts to the development of suitable logical systems, that in turn are computationally exploitable in the area of the so-called logic programming. Many attempts have been made to model abduction by developing some formal tools in order to illustrate its computational properties and the relationships with the different forms of deductive reasoning (see, for example, [2], [3], [4]). Some of the more recent formal models of abductive reasoning, for instance [5], are based on the theory of the epistemic state of an agent ([6], [7]), where the epistemic state of an individual is modeled as a consistent set of beliefs that can change by expansion and contraction (*belief revision framework*). I have discussed these deductive models in [1], where I have shown their importance and how they do not adequately account for some roles played by abduction in many forms of explanatory creative reasoning.

This kind of sentential frameworks exclusively deals with selective abduction (diagnostic reasoning) and relates to the idea of preserving *consistency*. Exclusively considering the sentential view of abduction does not enable us to say much about creative processes in science. It mainly refers to the *selective* (diagnostic) and *explanatory* aspects of reasoning and to the idea that abduction is mainly an inference to the best explanation: when used to express the creativity events it is either empty or replicates the well-known *Gestalt* model of radical innovation. It is empty because the sentential view stops any attempt to analyze the creative processes.

Already in the Peircian syllogistic and sentential initial conception of abduction - as the fallacy of affirming the consequent, we immediately see it is perfectly compatible with the *Gestalt* model of discovery. In both cases the event of creating something new (for example a new concept) is considered so radical and instantaneous that its irrationality is immediately involved. In this case the process is not considered as algorithmic: "the abductive suggestion comes to us like a flash. It is an act of insight, although of extremely fallible insight" ([8] 5.181). Moreover, Peirce considers abduction as "a capacity of guessing right", and a "mysterious guessing power" common to all scientific research ([8], 6.530).

Notwithstanding its non-algorithmic character it is well known that for Peirce abduction is an *inferential process* (cf. [1], chapter 2, section 2.2, for an explanation of the exact meaning of the word "inference" in Peircian thought).

The abductive inference includes all the operations whereby hypotheses and theories are constructed ([8] 5.590) (see also [9]). Hence abduction has to be considered as a kind of *ampliative* inference that is not logical and truth preserving (in the sense of deductive): indeed valid deduction does not yield any new information.

As previously said, if we want to provide a suitable framework for analyzing the most interesting cases of conceptual changes in science we do not have to limit ourselves to the sentential view of abduction but we have to consider a broader *inferential* one which encompasses both sentential and what I call *model-based* sides of creative abduction.

III. Model-Based Creative Abduction

A. Conceptual Change and Creative Reasoning in Science

I have analyzed in detail elsewhere ([10], [1]) some limitations of the sentential models of abduction (cf. the previous section) in accounting for some reasoning tasks; for example they do not capture 1. the role of statistical explanations, where what is explained follows only probabilistically and not deductively from the laws and other tools that do the explaining; 2. the sufficient conditions for explanation; 3. the fact that sometimes the explanations consist of the application of *schemas* that fit a phenomenon into a pattern without realizing a deductive inference; 4. the idea of the existence of high-level kinds of *creative* abductions; 5. the existence of model-based abductions; 6. the fact that explanations usually are not complete but only furnish *partial* accounts of the pertinent evidence ([11]); 7. the fact that one of the most important virtues of a new scientific hypothesis (or of a scientific theory) is its power of explaining *new*, previously unknown facts: "[...] these facts will be [...] unknown at the time of the abduction, and even more so must the auxiliary data which help to explain them be unknown. Hence these future, so far unknown explananda, cannot be among the premises of an abductive inference" ([9] p. 507), observations become real and explainable only by means of new hypotheses and theories, once discovered by abduction.

It is in terms of *model-based abductions* (and not in terms of sentential abductions) that we have to think for example of the case of a successful synthesis of two earlier theoretical frameworks which might even have seemed incompatible. The old epistemological view sees Einstein's theory as an attempt to "explain" certain anomalies and facts such as the Michelson-Morley experiment: "The most instructive way of looking at Einstein's discovery is to see it as a way of reconciling Maxwell's electromagnetic theory with Newtonian mechanics [...] it would be ridiculous to say that Einstein's theory 'explains' Maxwell's theory any more than it 'explains' Newton's laws of motion" ([9] p. 510). This kind of abductive movement does not have that immediate explanatory effect illustrated by the sentential models of abduction: the new framework usually does not "explain" the previous ones but provides a very radical new perspective.

If we want to deal with the nomological and most interesting creative aspects of abduction we are first of all compelled to consider the whole field of the growth of scientific knowledge cited above.

We may also see belief change from the point of view of *conceptual change*, considering concepts either cognitively, like mental structures analogous to data structures in computers, or, epistemologically, like abstractions or representations that presuppose questions of justification. Belief revision is able to represent cases of conceptual change such as adding a new instance, adding a new weak rule, adding a new strong rule ([12]), that is, cases of addition and deletion of beliefs, but fails to take into account cases such as adding a new part-relation, adding a new kind-relation, adding a new concept, collapsing part of a kind-hierarchy, reorganizing hierarchies by branch jumping and tree switching, in which there are reorganizations of concepts or redefinitions of the nature of a hierarchy.

Adding new part-relations occurs when in the part-hierarchy new parts are discovered: an example is given by the introduction of new molecules, atoms, and subatomic particles. Thomson's discovery that the "indivisible" atom contains electrons was very sensational.

Adding new kind-relations occurs when it is added a new superordinate kind that combines two or more things previously taken to be distinct. In the nineteenth century scientists recognized that electricity and magnetism were the same and constructed the new concept of electromagnetism. Another case is shown by differentiation, that is the making of a new distinction that generates two kinds of things (heat and temperature were considered the same until the Black's intervention).

The last three types of conceptual change can be illustrated by the following examples. The Newtonian abandon of the Aristotelian distinction between natural and unnatural motion exemplifies the collapse of part of the kind-hierarchy. Branch jumping occurred when the Copernican revolution involved the recategorization of the earth as a kind of planet, when previously it had been considered special, but also when Darwin reclassified humans as a kind of animal. Finally, we have to say that Darwin not only reclassified humans as animals, he modified the meaning of the classification itself. This is a case of hierarchical tree redefinition:

These last cases are the most evident changes occurring in many kinds of creative reasoning in science, when adopting a new conceptual system is more complex than mere belief revision. Related to some of these types of scientific conceptual change are different varieties of *model-based abductions*. In these cases the hypotheses "transcend" the vocabulary of the evidence language, as opposed to the cases of simple inductive generalizations, like for example in the well-known case of the abductive discovery of totally new physical concepts during the transition from classical mechanics to quantum mechanics.

B. Model-Based Abduction

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" ([8] 5.283), and, in Kantian words, all synthetic forms of cognition. That is, a considerable part of the thinking activity is *model-based*. Of course model-based reasoning acquires its peculiar creative relevance when embedded in abductive processes.

For Peirce ([13]) a Kantian keyword is synthesis, where the intellect constitutes in its forms and in a harmonic way all the material delivered by the senses. Surely Kant did not consider synthesis as a form of *inference* but, notwithstanding the obvious differences, I think synthesis can be related to the Peircian concept of inference, and, consequently, of abduction. After all, when describing the ways the intellect follows to unify and constitute phenomena through imagination Kant himself makes use of the term *rule* "Thus we think a triangle as an object, in that we are conscious of the combination of the straight lines according to a rule by which such an intuition can always be represented" ([14], A140, B179-180, p. 182), and also of the term *procedure* "This representation of a universal procedure of imagination in providing an image for a concept, I entitle the schema of this concept" ([14], A140-B179-180, p. 182). We know that rules and procedures represent the central features of the modern concept of inference. Moreover, according to Peirce, the central question of philosophy is "how synthetical reasoning is possible [...]. This is the lock upon the door of philosophy" ([8] 5.348), and the mind presents a tendency to unify the aspects which are exhibited by phenomena: "the function of conception is to reduce the manifold of sensuous impressions to unity" ([8] 1.545).

Most of these forms of constitution of phenomena are creative and, moreover, characterized in a model-based way. Let me show some examples of model-based inferences. It is well known the importance Peirce ascribed to diagrammatic thinking, as shown by his discovery of the powerful system of predicate logic based on diagrams or "existential graphs". As I have already stressed, Peirce considers inferential any cognitive activity whatever, not only conscious abstract thought; he also includes perceptual knowledge and subconscious cognitive activity ([15]). For instance in subconscious mental activities visual representations play an immediate role.

We should remember, as Peirce noted, that abduction plays a role even in relatively simple visual phenomena. *Visual abduction*, a special form of non verbal abduction, occurs when hypotheses are instantly derived from a stored series of previous similar experiences. It covers a mental procedure that tapers into a non-inferential one, and falls into the category called "perception". Philosophically,[1] *perception* is viewed by Peirce as a fast and uncontrolled knowledge-production procedure. Perception, in fact, is a vehicle for the instantaneous retrieval of knowledge that was previously structured in our mind through inferential processes. Peirce says: "Abductive inference shades into perceptual judgment without any sharp line of demarcation between them" ([19] p. 304). By perception, knowledge constructions are so instantly reorganized that they become habitual and diffuse and do not need any further testing: "[...] a fully accepted, simple, and interesting inference tends to obliterate all recognition of the uninteresting and complex premises from which it was derived" ([8] 7.37). Many visual stimuli - that can be considered the "premises" of the involved abduction - are ambiguous, yet people are adept at imposing order on them: "We readily form such hypotheses as that an obscurely seen face belongs to a friend of ours, because we can thereby explain what has been observed" ([20] p. 53). This kind of image-based hypothesis formation can be considered as a form of *visual* (or *iconic*) *abduction*. Of course such subconscious visual abductions of everyday cognitive behavior are not of particular importance but we know that in science they may be very significant and lead to interesting new discoveries ([21], [22]). If perceptions are abductions they are withdrawable, just like the scientific hypotheses abductively found. They are "hypotheses" about data we can accept (sometimes this happens spontaneously) or carefully evaluate.

One more example is given by the fact that the perception of tone arises from the activity of the mind only after having noted the rapidity of the vibrations of the sound waves, but the possibility of individuating a tone happens only after having heard several of the sound impulses and after having judged their frequency. Consequently the sensation of pitch is made possible by previous experiences

[1] In philosophical tradition perception was viewed very often like a kind of inference ([14], [16], [17], [18], [4])

and cognitions stored in memory, so that one oscillation of the air would not produce a tone.

To conclude, for Peirce all knowing is *inferring* and inferring is not instantaneous, it happens in a process that needs an activity of comparisons involving many kinds of models in a more or less considerable lapse of time.[2] This is not in contradiction with the fact that for Peirce the inferential and abductive character of creativity is based on the instinct (the mind is "in tune with nature") but does not have anything to do with irrationality and blind guessing. Hanson ([24], pp. 85-92) perfectly recognizes the model-based side of abductive reasoning, when he relates (and reduces) it to the activity of "interpretation" ("pattern of discovery") resorting to the well-known example of reversible perspective figures of *Gestalt* psychology. Unfortunately, this kind of analysis inhibits the possibility of gaining further knowledge about model-based reasoning. I think Hanson is inclined to consider the abductive event as instantaneous and not susceptible to further cognitive and epistemological examination.

All sensations or perceptions participate in the nature of a unifying hypothesis, that is, in abduction, in the case of emotions too:

> Thus the various sounds made by the instruments of the orchestra strike upon the ear, and the result is a peculiar musical emotion, quite distinct from the sounds themselves. This emotion is essentially the same thing as a hypothetic inference, and every hypothetic inference involved the formation of such an emotion ([8] 2.643).

Following Nersessian ([25], [26]), I use the term "model-based reasoning" to indicate the construction and manipulation of various kinds of representations, not necessarily sentential and/or formal. She proposes the so-called cognitive history and philosophy of science approach, which affords a reframing of the problem of conceptual formation and change in science that not only provides philosophical insights but also pays attention to the practices employed by real human agents in constructing, communicating and replacing representation of a domain. Common examples of model-based reasoning are constructing and manipulating visual representations, thought experiment, analogical reasoning, but also the so-called "tunnel effect" ([27]), occurring when models are built at the intersection of some operational interpretation domain - with its interpretation capabilities - and a new ill-known domain.

We have to remember that visual and analogical reasoning are productive in scientific concept formation too, where the role they play in model-based abductive reasoning is very evident; scientific concepts do not pop out of heads, but are elaborated in a problem-solving process that involves the application of various procedures: this process is a *reasoned process*. Visual abduction, but also many kinds of abductions involving analogies, diagrams, thought experimenting, visual imagery, etc. in scientific discovery processes, can be just called *model-based*. Additional considerations about the intersections between abduction and model-based reasoning (especially in experiment and thought experiment) are illustrated by Gooding ([28]): the ability to integrate information from various sources is crucial to scientific inference and typical of all kinds of model-based reasoning also when models and representations are "external", like verbal accounts, drawings, various artifacts, narratives, etc.

[2] This corresponds to Peirce's "philosophical" point of view, which delineates a very particular meaning of the word "inference", as illustrated above. We have to say that recently very many philosophers and others have accepted that perceptual knowledge is both non inferential and instantaneous (although the latter may be debatable) ([23]).

IV. Action-Based Reasoning and Manipulative Abduction

Manipulative abduction happens when we are *thinking through doing* and not only, in a pragmatic sense, about doing. For instance, when we are creating geometry constructing and manipulating an external triangle, like in the case given by Kant in the "Transcendental Doctrine of Method". In the case of natural sciences the idea of manipulative abduction goes beyond the well-known role of experiments as capable of forming new scientific laws by means of the results (the nature's answers to the investigator's question) they present, or of merely playing a predictive role (in confirmation and in falsification). Manipulative abduction refers to an extra-theoretical behavior that seeks to create communicable accounts of new experiences to integrate them into previously existing systems of experimental and linguistic (theoretical) practices.

The existence of this kind of extra-theoretical cognitive behavior is also evidenced by the many everyday situations in which humans are perfectly able to perform very efficacious (and habitual) tasks without the immediate possibility of realizing their conceptual explanation. In some cases the conceptual account for doing these things was at one point present in the memory, but now has deteriorated, and it is necessary to reproduce it, in other cases the account has to be constructed for the first time, like in creative settings of manipulative abduction in science. Hutchins [29] illustrates the case of a navigation instructor that for 3 years performed an automatized task involving a complicated set of plotting manipulations and procedures. The insight concerning the conceptual relationships between relative and geographic motion came to him suddenly "as lay in his bunk one night". This example explains that many forms of learning can be represented as the result of the capability of giving conceptual and theoretical details to already automatized manipulative executions. The instructor does not discover anything new from the point of view of the objective knowledge about the in-volved skill, however, we can say that his conceptual awareness is new from the local perspective of his individuality.

In this kind of *action-based* abduction the suggested hypotheses are inherently ambiguous until articulated into configurations of real or imagined entities (images, models or concrete apparatus and instruments). In these cases only by experimenting, we can discriminate between possibilities: they are articulated behaviorally and concretely by manipulations and then, in-creasingly, by words and pictures. Gooding [28] 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 involving visual and tactile performances that are often apparatus and machine based. They belong to the pre-verbal context of ostensive operations, that are practical, situational, and often made with help of words, visualizations, or concrete artifacts. The embodied expertise deals of course with an expert manipulation of objects in a highly constrained experimental environment, and is directed by abductive movements that imply the strategic application of old and new (non-conceptual) templates of behavior mainly connected with extra-theoretical components.

Peirce gives an interesting example of model-based abduction related to sense activity: "A man can distinguish different textures of cloth by feeling: but not immediately, for he requires to move fingers over the cloth, which shows that he is obliged to compare sensations of one instant with those of another" ([8], *CP* 5.221). This idea surely suggests that abductive movements also have interesting extra-theoretical characteristics and that there is a role in abductive reasoning for various kinds of manipulations of external objects. When manipulative aspects of external models prevail, like in the case of manipulating diagrams in the blackboard, we face what I call *manipulative* abduction (or action-based abduction).

Some common features of the tacit templates of manipulative abduction (cf. Figure 2) that enable us to manipulate things and experiments in science are related to: 1. sensibility to 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; 2. preliminary sensibility to the *dynamical* character of the phenomenon, and not to entities and their properties, common aim of manipulations is to practically reorder the dynamic sequence of events into 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 possible stable and repeatable sources of information about hidden knowledge and constraints (Davy well-known setup in term of an artifactual tower of needles showed that magnetization was related to orientation and does not require physical contact); 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).

The whole activity of manipulation is devoted to building various external *epistemic mediators* that function as 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).

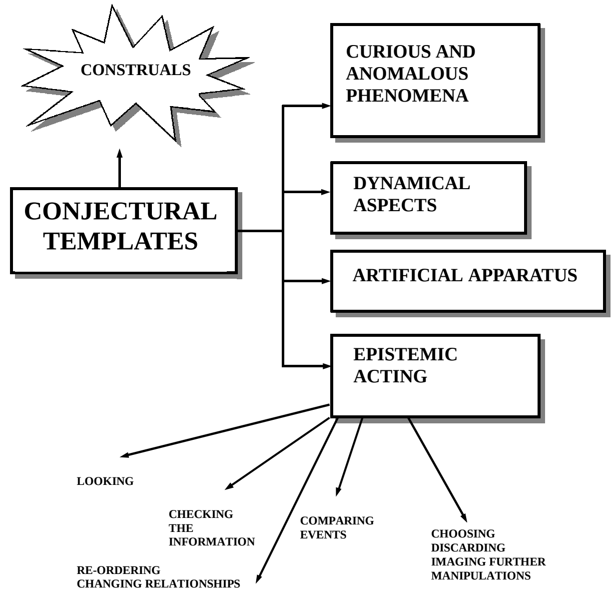


Fig. 2. Conjectural templates.

From the point of view of general everyday situations manipulative abductive reasoning exhibits very interesting features (for example, we can find the first three in geometrical constructions) (cf. Figure 3): 1. action elaborates a *simplification* of the reasoning task and a redistribution of effort across time [29], when we need to manipulate concrete things in order to understand structures which are otherwise too abstract, or when we are in presence of *redundant* and unmanageable information; 2. 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; 3. 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 agent's needs. 4. 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.

Also natural phenomena can play the role of external artifactual models: under Micronesians' manipulations of their images, the stars acquire a structure that "becomes one of the most important structured representational media of the Micronesian system" ([29], p. 172).

A. External Representations

Certainly a big portion of the complex environment of a thinking agent is internal, and consists of the proper soft-

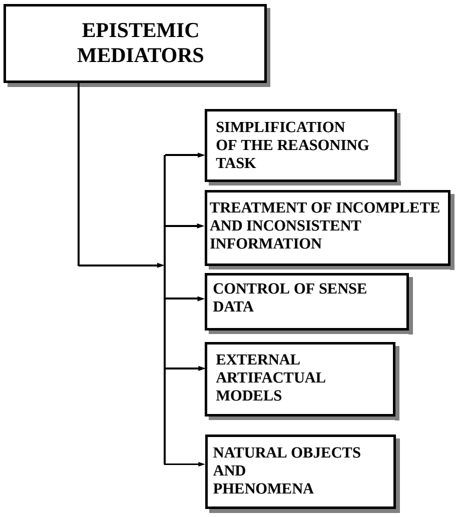


Fig. 3. Epistemic mediators.

ware composed of the knowledge base and of the inferential expertise of that individual. Nevertheless, any cognitive system is composed by a "distributed cognition" among people and "external" technical artifacts ([29], [30]).

In the case of science, specific experiments serve as states and the implied operators are the manipulations and observations that transform one state into another. The scientific 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. There is a kind of an epistemic negotiation between the sensory framework of the scientist and the external reality of the object used. This process involves an external representation consisting of artificial objects (for example written symbols and figures and artifacts) that are manipulated "by hand".

The cognitive system is not merely the mind-brain of the person performing the scientific task, but the system consisting of the whole body (cognition is *embodied*) of the person plus the external physical representation. In scientific discovery the whole activity of cognition is located in the system consisting of a human together with the external objects.

An external representation can modify the kind of computation that a human agent uses to reason about a problem: the Roman numeration system eliminates, by means of the external signs, some of the hardest parts of the addition, whereas the Arabic system does the same in the case of the difficult computations in multiplication [31]. All external representations, if not too complex, can be transformed in internal representations by memorization. But this is not always necessary if the external representations are easily available. Internal representations can be transformed in external representations by externalization, that can be productive "if the benefit of using external representations can offset the cost associated with the externalization process" (*ibid.*, p. 181).

Hence, contrarily to the received view in cognitive science, not all cognitive processes happen in an internal model of the external environment. The information present in the external world can be directly picked out without the mediation of memory, deliberation, etc. Moreover, various different external devices can determine different internal ways of reasoning and cognitively solve the problems, as it is well-known. Even a simple arithmetic task can completely change in presence of an external tool and representation.

It is indeed interesting to note that also in mathematics model-based and manipulative abductions are present. For example, it is clear that in geometrical construction all these requirements are fulfilled. Geometrical constructions present situations that are curious and "at the limit". These 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 some of the features indicated below, typical of all abductive epistemic mediators, not just of the ones which play a scientific role: simplification of the task and the capacity to get visual information otherwise unavailable.

In general 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 an *optical* role - microscopes (that look at the infinitesimally small details), telescopes (that look at infinity), windows (that look at a particular 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).[3] I describe them as *epistemic mediators* [1] 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.).

V. Conclusions

The main thesis of this paper is that abduction is a significant kind of scientific reasoning, helpful in delineating the first principles of a new theory of science. The interdisciplinary character of abduction is central and its fertility in various areas of research evident. The study of the high-level methods of abductive reasoning is situated at the crossroads of philosophy, epistemology, artificial intelligence, cognitive psychology, and logic. The various aspects of abduction I have described provide a better understanding of the processes of explanation and discovery in science. Their analysis certainly increases knowledge about creative and expert inferences that complete the epistemological and cognitive examination of important features of scientific reasoning.

Some research in the area of artificial intelligence has shown that methods for discovery could be found that are

[3] The epistemic and cognitive role of optical diagrams in "perceiving the infinite and the infinitesimal world" in the calculus is illustrated in [32].

computationally adequate for abductively rediscovering - or discovering for the first time - empirical or theoretical laws and theorems[4]. Moreover, the study of diagnostic, visual, spatial, analogical, and temporal reasoning has demonstrated that there are many ways of performing intelligent and creative tasks that cannot be described with only the help of classical logic. However, non-standard logic has shown how we can provide rigorous formal models of many kinds of abductive reasoning such as the ones involved in defeasible, heterogeneous, and uncertain inferences [33]. Abduction is also useful in describing the different roles played by the various kinds of medical reasoning, from the point of view both of human agents and of computational programs that perform medical tasks such as diagnosis ([1], chapter 4). Finally, concrete manipulations of external objects influence the generation of hypotheses: what I call manipulative abduction shows how we can find methods of constructivity in scientific and everyday reasoning based on external models and "epistemic mediators".

I think the cognitive activity of abduction can be further studied in many areas of model-based reasoning [34], such as the ones involving creative, analogical, spatial inferences, and the exploitation of external representations and mediators, both in science and everyday situations, so that this can extend the epistemological and the psychological theory.

References

[1] L. Magnani, *Abduction, Reason, and Science. Processes of Discovery and Explanation*, Kluwer Acedemic/Plenum Publishers, New York, 2001.

[2] T. Bylander, D. Allemang, M.C. Tanner, and J.R. Josephson, "The computational complexity of abduction," *Artificial Intelligence*, vol. 49, pp. 25–60, 1991.

[3] A. Kakas, R.A. Kowalski, and F. Toni, "Abductive logic programming," *Journal of Logic and Computation*, vol. 2(6), pp. 719–770, 1992.

[4] J.R. Josephson and S.G. Josephson, Eds., *Abductive Inference. Computation, Philosophy, Technology*, Cambridge University Press, Cambridge, 1994.

[5] C. Boutilier and V. Becher, "Abduction as belief revision," *Artificial Intelligence*, vol. 77, pp. 43–94, 1995.

[6] C. Alchourrón, P. Gärdenfors, and P. Makison, "On the theory of logic change: partial meet functions for contractions and revision," *Journal of Symbolic Logic*, vol. 50, pp. 510–530, 1985.

[7] P. Gärdenfors, *Knowledge in Flux*, MIT Press, Cambridge, 1988.

[8] C.S. Peirce, *Collected Papers*, edited by C. Hartshorne and P. Weiss (vols. I-VI), and by A.W. Burks (vols. VII-VIII). Harvard University Press, Cambridge, MA, 1931–1958.

[9] J. Hintikka, "What is abduction? The fundamental problem of contemporary epistemology," *Transactions of the Charles S. Peirce Society*, vol. 34, pp. 503–533, 1998.

[10] L. Magnani, "Model-based creative abduction," in *Model-Based Reasoning in Scientific Discovery*, L. Magnani, N.J. Nersessian, and P. Thagard, Eds., New York, 1999, pp. 219–238, Kluwer Academic/Plenum Publishers.

[11] P. Thagard and C.P. Shelley, "Abductive reasoning: logic, visual thinking, and coherence," in *Logic and Scientific Method*, M.L. Dalla Chiara, K. Doets, D. Mundici, and J. van Benthem, Eds., Dordrecht, 1997, pp. 413–427, Kluwer Academic Publisher.

[12] P. Thagard, *Conceptual Revolutions*, Princeton University Press, Princeton, NJ, 1992.

[13] D.R. Anderson, "The evolution of Peirce's concept of abduction," *Transactions of the Charles S. Peirce Society*, vol. 22(2), pp. 45–58, 1986.

[4]On the recent achievements in the area of the machine discovery simulations of abductive creative tasks cf. [33].

[14] I. Kant, *Critique of Pure Reason*, Reprint 1998 - originally published 1787. MacMillan, London, 1992.

[15] W.H. Davis, *Peirce's Epistemology*, Nijhoff, The Hague, 1972.

[16] J. Fodor, *The Modularity of the Mind*, MIT Press, Cambridge, 1983.

[17] W.V.O. Quine, Ed., *The Logic of Perception*, MIT Press, Cambridge,MA, 1983.

[18] R.L. Gregory, "Perception as hypothesis," in *The Oxford Companion to the Mind*, R.L.Gregory, Ed., New York, 1987, pp. 608–611, Oxford University Press.

[19] C.S. Peirce, *Visual cognition and cognitive modeling*, pp. 302–305, J. Buchler. Dover, New York, 1955.

[20] P. Thagard, *Computational Philosophy of Science*, MIT Press, Cambridge, MA, 1988.

[21] L. Magnani, S. Civita, and G. Previde Massara, "Visual cognition and cognitive modeling," in *Human and Machine Vision: Analogies and Divergences*, V. Cantoni, Ed., New York, 1994, pp. 229–243, Plenum Publishers.

[22] C. Shelley, "Visual abductive reasoning in archaeology," *Philosophy of Science*, vol. 63(2), pp. 278–301, 1996.

[23] K. Akins, *Perception*, Oxford University Press, Oxford, 1996.

[24] N.R. Hanson, *Patterns of Discovery. An Inquiry into the Conceptual Foundations of Science*, Cambridge University Press, Cambridge, 1958.

[25] N.J. Nersessian, "Opening the black box: cognitive science and history of science," *Osiris*, vol. 10, pp. 194–211, 1995.

[26] N.J. Nersessian, "Should physicists preach what they practice? Constructive modeling in doing and learning physics," *Science and Education*, vol. 4, pp. 203–226, 1995.

[27] A. Cornuéjols, A. Tiberghien, and G. Collet, "Decomposing the scientific discovery process using multiple interpretations of "notions"," *Foundations of Science*, vol. 5(2), pp. 129–155, 2000.

[28] D. Gooding, *The Experiment and the Making of Meaning*, Kluwer Academic Press, Dordrecht, 1990.

[29] E. Hutchins, *Cognition in the Wild*, MIT Press, Cambridge, MA, 1995.

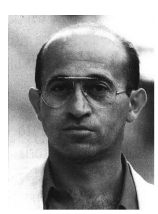
[30] D.A. Norman, *Things that Make Us Smart. Defending Human Attributes in the Age of the Machine*, Addison-Wesley, Reading, MA, 1993.

[31] J. Zhang, "The nature of external representations in problem solving," *Cognitive Science*, vol. 21(2), pp. 179–217, 1997.

[32] L. Magnani and R. Dossena, "Perceiving the infinite and the infinitesimal world: unveiling and optical diagrams and the construction of mathematical concepts," *Fourdations of Science*, forthcoming.

[33] L. Magnani, N.J. Nersessian, and C. Pizzi, Eds., *Logical and Computational Aspects of Model Based Reasoning*, Dordrecht, 2002. Kluwer Academic.

[34] L. Magnani and N.J. Nersessian, Eds., *Model-Based Reasoning. Science, Technology, Values*, New York, 2002. Kluwer Academic/Plenum Publishers.



Lorenzo Magnani is professor of Logic and of Epistemology in the Department of Philosophy at the University of Pavia where he directs the Computational Philosophy Laboratory. He is also Visiting Professor at Georgia Institute of Technology, Atlanta. His primary research interests are in the area of philosophy of science, logic, and artificial intelligence. Currently he is studying the processes of conceptual innovation and change in science. Visiting researcher at the Universities of Pittsburgh (CMU), Montreal (McGill), Waterloo, and Atlanta (Georgia Tech). Author of several books and articles. Recent books are *Abduction, Reason, and Science. Processes of Discovery and Explanation* (Kluwer Academic/Plenum Publishers, New York, 2001), and *Philosophy and Geometry. Theoretical and Historical Issues* (Kluwer, Dordrecht, 2001). Recent chair of the International Conference MBR'98 (Model-Based Reasoning) with N.J. Nersessian and P. Thagard, and MBR'01 (with N.J. Nersessian and K. Knoespel), and editor of the related Proceedings (Kluwer Academic/Plenum Publishers, New York, 1999 and 2002, and Kluwer, Dordrecht, APLS series, 2002).