Multi Level Knowledge in Modeling Qualitative Physics Learning

FILIPPO NERI

neri@mfu.unipmn.it

Università del Piemonte Orientale "A. Avogadro", Dipartimento di Scienze e Tecnologie Avanzate, Corso Borsalino 54, 15100 Alessandria Italy

Editor: Floriana Esposito, Ryszard Michalski, & Lorenza Saitta

Abstract. The goal of the reported research is the development of a computational approach that could help a cognitive scientist to interactively represent a learner's mental models, and to automatically validate their coherence with respect to the available experimental data. In a reported case-study, the student's mental models are inferred from questionnaires and interviews collected during a sequence of teaching sessions. These putative cognitive models are based on a theory of knowledge representation, derived from psychological results and educational studies, which accounts for the evolution of the student's knowledge over a learning period. The learning system WHY, able to handle (causal) domain knowledge, shows how to model the answers and the causal explanations given by the learner.

Keywords: functional cognitive modeling, causal reasoning, multistrategy learning

1. Introduction

The multiplicity of tasks humans routinely perform suggests that the knowledge a person acquires and exploits in his/her life might be richly articulated and diverse with respect to its nature, putative representation, and preferential domain of use. However, amazing performance achieved on even difficult tasks does not always match a corresponding ability to single out reasons or to *explain* the "whys" underlying observations and actions involved in problem-solving activity. On the other hand, abstract knowledge of fundamental laws does not automatically provide people with concrete tools to solve problems: in order to be of any use in practice, this knowledge must become *operational*. This dichotomy between performing and explaining corresponds to Ohlsson's distinction between "learning to do", requiring *practical* knowledge, and "learning to understand", involving declarative knowledge (Ohlsson, 1996).

Practical knowledge is performance-oriented, and its development has been widely investigated in the field of skill acquisition. A number of successful models, of remarkable degree of sophistication, have been proposed since long to describe this type of learning [Anzai & Simon, 1979; Anderson, 1985]. On the contrary, modeling declarative (in Ohlsson's sense) knowledge acquisition (*higher order* learning) is still very preliminary. In fact, if it is usually easy to define a sufficiently precise notion of "good performance" for a task, it is not the same for a definition of what must be considered a "good explanation", or what

does it mean to have “understood” an argument or a phenomenon. One of the difficulties encountered in investigating “learning to understand” is the unavoidable role played by the discourse, which is the expressive means almost exclusively used to explain and argue. In fact, a body of experiments supports the hypothesis that human argumentative reasoning lies largely outside the boundaries of formal logics. Then, even precisely framing the problem and interpreting the data (usually verbal transcripts) is an elusive task for the researcher.

Even though the distinction between problem solving and understanding can be often observed in humans’ everyday life, it is of utmost importance in teaching, from elementary schools up to universities. Orienting teaching/learning toward a deeper understanding, by the part of the student, of the taught notions and phenomena, would result not only in a more stable and comprehensive models of the world, but also in more powerful and versatile tools for increasing competence in problem solving.

“Learning to understand”, primarily from the perspective of Education Sciences, is the focus of our research. In particular, we are interested in teaching/learning Physics, domain in which the separation between problem solving and explanation ability is striking [Hestenes, 1987; Tiberghien, 1994; Duit, 1995].

Understanding a domain is usually achieved over time, through a series of updates of current ideas about the world, stimulated by new incoming information. A particular form of development has been called *conceptual change* which, intuitively, consists in the modification of a person’s conceptualization of the world [Carey, 1983; Tiberghien, 1985, 1989, 1994; Dykstra, 1992; Vosniadou & Brewer, 1992, 1994; Caravita & Halldén, 1994; Chi, Slotta, & de Leeuw, 1994; Vosniadou, 1994, 1995; Duit, 1995; Saitta et al., 1995].

Investigating students’ conceptualizations of the world is important and potentially fruitful under several respects. First of all, the results can shed light on basic human cognitive mechanisms taking part in learning, contributing so to advances in Cognitive Sciences. Second, it can help science education researchers to develop more realistic learning theories, useful to design innovative and principled teaching methodologies. But the teachers are those that can potentially benefit the most: they can receive well-founded suggestions to devise more effective teaching strategies, and be aided to identify better suited progressions in the presentation of the notions to teach and of the exercises to propose. Moreover, they can be offered concrete tools to design non-conventional didactic material (books, working notes, exercises, experimental work), and to assess the advances (if any) of the students toward the target knowledge. Particularly important, for teaching, is the possibility of developing accurate models of *individual* students, to support “personalized” tutoring.

Due to the complexity and inherent vagueness of the notion of “understanding”, we have circumscribed our investigation along three dimensions, namely domain of investigation, research target, and epistemological framework. The domain is limited to elementary Physics, in particularly, basic notions of *Calorimetry*. The research target is on *knowledge representation*, specifically the identification of interrelated bodies of knowledge, allowing the emergence of observed cognitive phenomena to be interpreted and (possibly) explained. Finally, concerning “understanding”, we rely upon a set of epistemological assumptions implicitly made in Tiberghien’s research in education science [Tiberghien, 1994], which we use as *working hypotheses*. These hypotheses, although biasing the approach, allow the researcher’s particular perspective to be taken into consideration when interpreting the results.

The first hypothesis states that conceptualizing a domain and changing one's current conceptions are both knowledge-intensive processes, which bring to bear a complex body of implicit and explicit pre-existing knowledge. The second hypothesis asserts that this knowledge is decomposable, both according to its nature and content, and according to its depth. In particular we consider two main layers, "shallow" and "deep" knowledge, challenged by new information with different ease. Deep knowledge is supposed to hold "unquestioned" assumptions, theories, epistemological concepts. Shallow knowledge usually contains facts and "rules of thumb" used when a specific problem has to be solved. A complex interplay may exist between these two layers. Articulation of knowledge structures has been proposed by several researchers [diSessa, 1993, 1996; Tiberghien, 1994; Vosniadou, 1995; Ohlsson, 1996]. The novelty of our approach resides in modeling precise *operational* links between the content and role of the different types of knowledge involved in conceptual change. The third hypothesis accepts Ohlsson's distinction between "learning to do" and "learning to understand" [Ohlsson, 1996]. From this starting point we go further, by proposing a formal definition of the bodies of knowledge involved in the two processes, and by relating them computationally. The last hypothesis assumes that a reasonable explanatory framework is offered by *causality*. More precisely, "understanding" a phenomenon will be interpreted as seeking a causal explanation for it. Even though the notion of cause has been severely challenged in modern philosophy, it seems nevertheless well rooted in everyday life, where it proves to be an effective tool for explaining and predicting events in the world [Salmon, 1989]. There is in fact little doubt that people routinely exploit causal relations, and this notion is spontaneously imposed on systems by people in order to understand how they work (see, for instance, [White & Frederiksen, 1988]). Again, in our approach causal knowledge can be represented, automatically manipulated, and functionally related to problem-solving oriented knowledge.

If "understanding" is the focus of the research, more complex types of knowledge need to be represented and manipulated than those sufficient to model performances. The main goal of this paper is to show how, from a corpus of student's verbal productions, the researcher can identify elements of knowledge, classify them and reconstruct knowledge structures which account for their utility in providing answers or explanations. To reach this goal, flexible and powerful languages must be used: the researcher could choose among the wide spectrum of formal languages Artificial Intelligence has to offer. However, if, in addition, automated treatment of the formalized knowledge is ultimately envisaged, a tool supporting both representation of complex knowledge and automated manipulation of it is needed. With this requirement in mind, we selected the Machine Learning system WHY, which acquires and revises knowledge by exploiting a causal model of the domain and a set of examples [Saitta, Botta, & Neri, 1993; Baroglio, Botta, & Saitta, 1994]. Even though the system WHY would allow for a complete simulation of the learning process, in this paper we only describe its representational facilities.

Finally, some remarks are in order. The first one concerns model complexity: in human learning, complexity is an issue that a model must handle, still remaining logically and computationally tractable. Hence, the model should be able to capture at least that level of complexity that allows it to be realistic, yet discarding complications. We believe that the "minimum" complexity level, which we cannot pass of [Tiberghien, 1994], can be

suitably modeled into the articulated knowledge representation framework presented here. The second remark concerns the paper's goal: we would like to stress that this paper is not aimed at supporting some cognitive learning mechanisms against some others. Its goal is, on the contrary, to provide a tool for describing and interpreting observed learning phenomena, as perceived by the educator, in a high-level graphical language that teachers found useful for representing their student's mental models. Finally, we notice that the system WHY is a symbolic system, which uses a logical language to represent knowledge. With this choice, we do not make any claim about the type of internal representation humans use. The logical language is simply a means to express findings in a comprehensible way *for the researcher*, and to eventually support automated simulations.

Consequently, in this paper we propose a functional model of the student's knowledge as far as learning qualitative Physics is concerned. Our framework complies with two cognitive hypotheses about knowledge organization. The first, the distinction between knowledge used to perform prediction and to interpret experimental situations; the second, the use of a naive explanatory framework, based on a simple notion of causality, which the student uses to "make sense" of what she observes or is taught of. Our model is grounded on an epistemological framework and previous experiments by Tiberghien (1989, 1994). Within this context, we have given a precise meaning to the relation between the layers of shallow and deep knowledge. The methodological improvement has been made possible by the availability of the system WHY [Neri & Saitta, 1993; Saitta, Botta, & Neri, 1993].

The remaining of the paper is organized as follows: in Section 2 related work is surveyed. In Section 3, the investigated learning context is described, together with the concrete experimental setting. Section 3 also presents the cognitive/education framework which we draw upon. Section 4 describes the system WHY, whereas Section 5 discusses how the learning problem is mapped onto WHY's epistemic components and operators through a worked example. Section 6 presents an overview of the experimental findings, and Section 7 presents some conclusions and hints for future work.

2. Related work

For investigating development of human conceptions, several models have been proposed, both in Cognitive Sciences (e.g., VanLehn, 1990, 1993; diSessa, 1993, 1996; Tiberghien, 1989; Vosniadou, 1995; Chi, Slotta, & de Leeuw, 1994; Anderson, 1985; Simon, Newell, & Klahr, 1991) and in Artificial Intelligence [Klahr & Siegler, 1978; Sage & Langley, 1983; Forbus, & Gentner, 1986; Laird, 1983; Rumelhart, 1989; Newell, 1990; Sleeman et al., 1990; Pazzani, 1991; Shultz, Mareschal, & Schmidt, 1994; Baffes & Mooney, 1996; Schmidt & Ling, 1996]. However, most computational models of human learning, even though interesting under many respects, are based on excessively simplifying assumptions, especially for what concerns representation of knowledge, which is often reduced to a set of simple "Condition $\rightarrow$ Action" rules. Moreover, and more importantly, most current models are performance-oriented, and they lack the ability to provide *explanations* that go beyond a trace of activated rules, or some pre-compiled natural language statements.

In order to keep the paper focused, we only review previous work related to conceptual change and to computational modeling, with special attention to knowledge representation

and articulation. Also, preference will be given to Physics learning: important approaches oriented to other disciplines, such as Mathematics (e.g., Weil-Barais & Vergnaud, 1990), will be left outside.

2.1. Education and cognitive sciences

In modeling human conceptualization, two main tendencies emerged. One supposes the existence of a coherent core of knowledge, whereas the other hypothesizes a set of knowledge pieces that may be incoherent. The first tendency encompasses several approaches. For Chi (1992), the basic structure of knowledge is composed by (ontological) category trees. Some among them are the “Matter”, the “Process” and the “Mental States” trees. A node of a tree is a concept. Any ontological tree is characterized by a set of ontological attributes that describe the concepts in a tree. For example, the attribute of “duration” applies to a process concept and not to a matter concept, whereas for the “color” attribute the opposite is true. Carey (1983) approached conceptual change in domain-specific knowledge using the notion of a “set of core principles that define the entities covered by the domain and support reasoning about those entities” [Carey & Spelke, 1994, p. 169]. In a similar perspective, Vosniadou (1994) proposed a more complex core, which consists of a *framework theory*, with ontological and epistemological presuppositions, and a *specific theory*. This notion of coherent core is also shared by several authors in science education [Dykstra, 1992; Caravita & Hallen, 1994; Samarapungavan & Wiers, 1997].

Contrarily to claims of the first tendency, diSessa (1993) supposes the learner’s initial knowledge as being a non coherent structure. Considering, more specifically, Physics learning, he proposes an epistemology of Physics based on simple and sparse knowledge pieces. These pieces, called *phenomenological primitives* (p-prims), constitute, as an ensemble, a large and complex knowledge system. Many p-prims seem to act by simple recognition of situations, but some of them can exhibit a self-explanatory nature. A system of priorities determines order in reasoning with p-prims. The knowledge sparseness view is shared by Hestenes (1987), who also claims that “. . . Physical intuition can be regarded as a system of unarticulated beliefs. . .”. These beliefs are only weakly interrelated and frequently inconsistent.

The hypothesized nature of knowledge strongly affects the types of learning supposedly used to acquire it. For example, for Chi and co-workers, conceptual change consists in restructuring an ontology tree [Chi, Slotta, & de Leeuw, 1994].

A central claim of diSessa’s phenomenology [diSessa, 1993] is that learning is changing the system of p-prim priorities, modifying the p-prim applicability context, and also adding new p-prims. Learning may also consist in changing the p-prim’s role: some are no more self-explanatory, but need to refer to deeper knowledge structures, such as physical laws. Physics learning can be interpreted to a great extent as moving from a phenomenological to a deeper knowledge layer by reorganizing and changing *priority* to existing p-prims. Furthermore, diSessa claims that humans gradually acquire a *sense of mechanism*, i.e., a qualitative and heuristic understanding of “how things work”, used as an explanatory framework.

In Vosniadou’s framework, conceptual change is seen as a gradual lifting of presuppositions and constraints, which can be added, removed, suspended or revised. Some mental

models, proved useful in the past, may be stored in long-term memory and retrieved when needed. It is argued that the mental models are the points where new information is incorporated into existing knowledge. Mental models fall into three categories: *initial*, *synthetic* and *scientific* ones. Initial models are those which people (children) form exclusively from everyday experience, before being exposed to any teaching. Attempts to resolve conflicts between presuppositions and taught information generate synthetic models, which usually require that some presupposition is removed or revised. Finally, scientific models encode the culturally accepted view [Vosniadou & Brewer, 1992, 1994].

Tiberghien (1994) does not hypothesize learning mechanisms of her own, but she exploits the classification, proposed by Rumelhart & Norman (1977), of learning phenomena into three broad classes: "Accretion", "Tuning", and "Restructuring". The categorization of learning into the classes proposed by Rumelhart and Norman received a wide consensus, even though the precise meaning of the terms may vary among researchers. These authors considered frame-like structures, called *schemata*, as basic knowledge units. Accretion, tuning, and restructuring are similar to Piaget's *Assimilation*, *Accommodation* and *Re-Equilibration*, respectively [Piaget, 1974].

The above categorization of learning types can be made more precise when some particular representation formalism is hypothesized. For knowledge represented as a set of rules, for example, learning can be described by means of a set of basic transformations, partitioned by Ohlsson (1996) into four groups. First of all, rules can be *created* by proceduralization [Anderson, 1985; Ohlsson, 1987], analogy [Anderson, 1993], or induction [VanLehn, 1990]. Rules can, in principle, be also deleted, but it is not always clear through which mechanism. Moreover, rules can be *modified*, by generalization [Anderson, 1985], or discrimination [Langley, 1987; Ohlsson, 1987], and *optimized* via composition [Lewis, 1987]. Finally, *priorities* associated to the rules can be changed, by strengthening or weakening [diSessa, 1993].

VanLehn & Jones [1993] proposed an approach/tool, named CASCADE, for developing a computational theory of human cognitive skill acquisition. CASCADE's activity started for investigating the self-explanation effect in student learning Newtonian Physics [Chi et al., 1989], that is how students use solved examples to deal with new Physics problems. CASCADE learns by doing the same sort of academic work that students do: analyzing solved examples, solving problems, getting feedback from a tutor. Knowledge is represented by mean of "Condition -> Action" rule that are activated according to a backward-chaining control strategy. Learning is, in CASCADE, "impasse"-driven: when the system cannot explain a line in a solved example, i.e. to provide a derivation of the line from the available knowledge, a learning section starts. This impasse-driven learning philosophy is also present in our approach. However, our approach essentially differs from VanLehn's work, because we are not concerned with the development of a cognitive theory of human learning; instead, we are interested in providing a teacher with an interactive tool for the functional modeling of students.

The above overview mainly concerns modeling "learning to do". On the other hand, most of the approaches trying to address "understanding" make use of a more or less formalized notion of "causality". For instance, White & Frederiksen (1987, 1988) exploit causal modeling to set up a learning environment that incorporates features of microworlds

and of intelligent tutoring systems. The authors suggest possible progressions between models, in the context of a theory of model evolution.

2.2. Artificial intelligence

Among the models of learning proposed in the field of Artificial Intelligence, the one which is closer to the approach taken in this paper is Forbus and Gentner's theoretical framework for learning in physical domains [Forbus & Gentner, 1986]. This framework is based on the Qualitative Process (QP) theory, which stresses the relevance, in studying Physics, of the notion of *process*. A canonical learning sequence is hypothesized to occur. At each stage, theories are generated by exploiting both observations and theories developed at previous stages. The first stage of learning is that of *Protohistories*, which are "... rich, contextually specific, highly perceptual representations of phenomena, capturing expectations about typical phenomenological patterns...", acting as summaries of experience. Protohistories are automatically available when a trigger condition is encountered.

The second stage develops a *Causal corpus*, which consists of a set of simple statements, inferred from event co-occurrence, positing mechanisms underlying causal connections among events; these statements are conjectured to be computed from protohistories. Even though the notion of process does not yet enter the casual corpus, the authors conjecture that people, at this stage, are able to make a distinction between a mechanistic/causal connection and a definitional one. The causal corpus can be used directly as a collection of heuristics to draw inferences. The next stage is that of *Naive Physics*; at this stage, the notion of *Process* becomes central and provides the mechanisms left unspecified in the causal corpus. The mechanisms are represented as "processes", which explain observed changes. Finally, *Expert models* are created: they are quantitative representations of phenomena and of their connections. The transition to expert models involves several kinds of learning, including developing new ontologies, and translating qualitative into quantitative statements. The expert models also include heuristic "rules of thumb", which may derive from learning new protohistories. At this stage there may be a substantial ontological shift, driven both by the desire to handle more complex systems, and by the emergence of useful abstractions. In Forbus and Gentner's model, the basic knowledge representation formalism consists of *frames*. Forbus and Gentner's model is not bound to any particular Physics domain.

An even more general model of learning has been proposed in the system SOAR [Laird, Rosenbloom, & Newell, 1986], which exploits a kind of compilation mechanism to transform sequences of successful actions into knowledge "chunks", available to subsequent reasoning. Forbus and Gentner's model addresses the problem of understanding and explaining, whereas SOAR's approach is performance-oriented. Also limited to model competence in problem-solving are a number of studies concerning the evolution of childrens' conceptions on the balance scale [Klahr & Siegler, 1978; Sage & Langley, 1983; Shultz, Mareschal, & Schmidt, 1994; Schmidt & Ling, 1996]. These models exploit different knowledge representations, namely decision trees, neural networks or production rules; even though they can accurately reproduce observed effects, such as the U-shaped performance curve, no notion of deep understanding of the domain, nor of explanation is introduced.

In Machine Learning, explanations in terms of causality have been considered in a few approaches. For instance, Nordhausen & Pazzani (1993) proposes a method to learn causal

relation from observations, and Saitta and co-workers developed the system WHY, which exploits a causal model to learn and revise a set of rules [Saitta, Botta, & Neri, 1993; Baroglio, Botta, & Saitta, 1994]. More oriented to the development of tutoring systems, Sleeman et al. (1990) present a model of a student learning high-school algebra. Problem solving in the domain is modeled as a search in a space, whose states are equations and operators are (possibly incorrect) manipulations of the equations. Operators are represented as rules, and an incorrect operator is called a *mal-rule*. The student model consists of a set of rules. Rules are also the basic knowledge components in Baffes and Mooney's system ASSERT (1996), which can be used to develop intelligent tutoring systems (ITSs) relaying on a refinement-based student model. The described domain of application is teaching concepts in C++ programming. As for Sleeman et al. (1990), the idea is to start with a correct knowledge base, and to modify it with the aim of covering erroneous answers given by the students. The revised theory is then assumed to model the student misconceptions. Rules in the correct theory are associated to a textual comment; thus, an "explanation" can be built up by chaining all the comments along the correct inferential chain. Sison, Numao, & Shimura (1999) also investigate the learning behaviour of novice programmers. They automatically acquire a bug library of the most relevant program versions for a given function. Understanding a novice erroneous behavior requires to classify her program into the current bug library and then provide an intensional definition for the observed errors.

Our approach differs from the ITS oriented ones, such as Baffes and Mooney's, Sleeman's et al., or Andes [Gertner, Conati, & VanLehn, 1998], because tutoring systems are student oriented, whereas our approach is teacher-oriented. In the former, the student model is built with the aim of selecting the next most effective hint during problem solving; instead, we help the teacher to describe her perceived evolution of the student's mental model, and we delegate to her experience the choice about the "next lesson" topic. Essentially we chose to operate at a different time scale: days instead of minutes or hours. In addition, we focalize on the effective teaching/learning of (complex) theories more than on providing hints to the student while she is solving a specific problem.

By summarizing, modelling higher order learning, i.e. learning necessary to gain a deep understanding of a domain, may bring substantial improvements both to the clarification of basic human cognitive activities, and to the design of more intelligent teaching/learning tools. Substantial advances along this line cannot avoid considering the detailed and accurate analysis of the types of knowledge involved in the process.

3. Domain of investigation: basic concepts and qualitative relationships in physics

The general aim of Physics is to interpret and predict events in the physical world, through the construction of an understanding of the world itself. Several studies in education research report on students' difficulties in learning Physics [Driver, Guesne, & Tiberghien, 1985; Halloun & Hestenes, 1987; Hestenes, 1987; Duit, 1995; diSessa, 1996]. We believe that the interpretation of these difficulties, through their modelization, may produce valuable insights into human learning mechanisms. One distinctive facet of teaching/learning Physics is that students are no *tabula rasa*: they have gathered a wealth of experience in everyday

life. This experience, mostly based on perception and simple observation of actions' effects, is often in conflict with the scientifically accepted view. Then, the students develop, during their life, alternative conceptions, which are well radicated and difficult to change.

Another component of the students' difficulties in learning Physics is that physical phenomena can be described at several levels of detail, using a broad spectrum of quantitative and qualitative descriptions, and the knowledge involved in reasoning about Physics is articulated and differentiated. The richness and diversity of the physical world that appears to our senses, on the one hand, and the relative stability and reproducibility of phenomena, on the other, clearly support the hypothesis that Physics epistemology needs to be organized at least on two levels. A "deep" level, with assumptions, presuppositions, abstract laws and principles, and a "shallow" one, with phenomenological information.¹ The deep layer is rather stable, has a unifying role, and provides explanations; the shallow one is endlessly different, perceptually rich and requires efficient performance to be dealt with. The interplay between the knowledge layers, the diffuse presence of alternative conceptions, the high level of conceptualization, the deep understanding reached by the discipline, the abundance of experiments and findings make Physics an ideal field in which to study learning mechanisms.

3.1. *The educational context*

Our domain of investigation is the educational context studied by Tiberghien [1994], who provided us with the psychological/educational data she had collected. In this context, a group of eleven students, attending the first and second year of secondary school (12–13 year-old), were exposed to a Physics course, outside normal teaching, consisting of 11 sessions, once a week, including experimentation, questions, discussions and explicit teaching. Interviews before and after the set of teaching sessions with each student have been performed, and one group of two students was video-recorded throughout every session. The students also filled two questionnaires, one at the beginning and one at the end of the course.

Content of the course were basic concepts and qualitative relations in the domain of *heat* and *temperature* in everyday life situations. In Physics, temperature describes an intensive aspect of heat, whereas energy and entropy extensive ones [Duit, 1995]. In everyday life, instead, "heat" and "temperature" are often not differentiated, and heat is associated with high temperatures only. Moreover, heat is often considered as a liquid-like substance [Duit, 1995; Erickson, 1979]. The teaching sequence considered here was focused on heat conduction and change of state. The knowledge to be taught required a first distinction between temperature and heat to be stimulated through suggested experiments. The teaching sequencing was based on two types of general learning hypotheses:

1. Relevance of students' activities—Manipulating experiments by themselves should lead the students to a better understanding of the notions of the taught knowledge involved in the experiments.
2. Relevance of "cognitive conflict"—Experiments are selected to show observable facts which are in contradiction with students' predictions and interpretations. Such situations were designed from previous research results on students' conceptions.

The order of presentation of the notions was chosen on the basis of possible contradictions, and also with the aim to reinforce and extend the students' knowledge of the validity of the notions already acquired in a previous experiment.

3.2. *From education to functional modeling of Physics learning*

Drawing upon pedagogical studies and epistemology of science, Tiberghien (1985) proposed that in experimental sciences questions are linked to three main factors: the theoretical background, the experimental facts considered, and the explanations produced. Tiberghien assumes that when interpreting (or predicting) material situations, the learner constructs a "model" of the situation (analogical and/or propositional), depending on her own point of view [Tiberghien, 1994]. During the modeling process, the learner is assumed to select the objects and events according to her own theory. This hypothesis agrees with the findings of psychological research [Brown, 1989; Carey, 1983; Vosniadou, 1989]. For example, Brown (1989) states that young children not only can transfer their knowledge on "deeper bases than mere-appearance matches", but also they can use relational information in which causality plays a fundamental role. The hypothesis that the learner's modeling is based on the learner's theory is confirmed by the observation of causal explanation even in young children. All the above considerations lead Tiberghien (1994) to hypothesize that a knowledge framework well suited to model learning to understand in Physics could be articulated on three levels: a *theory* level, a *model* level, and an *experimental field* of reference.

In this paper we add the further hypothesis that the "theory" level contains a causal model of the domain; thus, an explanation denotes *causal* attribution. In the learner's theory, causality plays a crucial role [White & Frederiksen, 1987]. Taking into account the age of the learner (12–13 years), Aristotelian causalities are used as reference [Kuhn, 1971]. In particular, we consider *material* causality (e.g., "wool heats because it is wool") and *efficient* causality (e.g., "a battery lights a bulb"). Efficient causality may put into play a theoretical construct, including hidden variables, which may not correspond to a direct perception.

In Tiberghien's original framework the "model" level was hypothesized to consist of qualitative and quantitative functional relations between physical quantities, in order to represent the selected aspects of a set of material situations. In this paper we consider only the qualitative aspect of this level. The experimental field of reference involves the experimental situations that belong to the domain of validity of the considered theoretical construction.

Concerning Physics learning, Tiberghien's position on student's conceptions (1994) can be summarized as follows:

- By student's "conception" it is intended a hypothetical set of propositions, procedures, or skills, which are attributed to the student by the researcher, aiming to account for the students' behaviour in a given set of situations.
- It is possible to differentiate between conceptions on concepts (for example, heat and temperature), and conceptions on experiments (conceptions on material situations).

The second point is very important for the differentiation between teaching and learning. It allows a certain autonomy between the analysis of students' knowledge and that of the knowledge to be taught. When the aim is to probe the students' understanding of the notions of heat and temperature, it is necessary to understand the knowledge which was taught. When the goal is understanding about a part of the material world, it is up to the student to use, or not, the terms and/or the notions which have been taught to him/her. As we accept the hypothesis that students have an explanatory system with principles behind their explanations [Carey, 1983], we can compare the students' explanatory system, reconstructed from their productions, to the explanatory system of the knowledge to be taught. This forces the knowledge to be taught to be reformulated in order to make explicit its underlying explanatory system.

4. The system WHY

In this section a description of the functionalities of the learning system WHY is given [Saitta, Botta, & Neri, 1993]. WHY learns and revises a knowledge base of classification rules using domain knowledge and examples. The domain knowledge is differentiated and stratified. WHY differs from most current Machine Learning systems, because it is provided, beside an inductive component, also with a sophisticated reasoner, which searches for explanations of phenomena by abduction on a network of causal relations. Abduction regresses, through chains of cause-effect relations, from observations to a set of *first causes*, which are the ultimate hypotheses we are disposed to accept without further explanations. In this section we will consider WHY only from a computational perspective. In the next section its suitability to model human learning processes will be discussed.

4.1. Knowledge representation framework

A First Order Logic representation framework underlies all types of knowledge handled by WHY. Indeed, WHY's representation framework is constituted by this logical language plus a graph-oriented description of causal relations. The *representation framework* is then defined as the pair $(L, C(S, \rightarrow))$, where L is a function-free Horn clause predicate logic language, and C is an oriented graph, called *causal graph* or *causal model*. Predicates in L are called *operational* or *observable* if their truth values can be evaluated directly on the example data. Instead, predicates of L that occur in the head of some Horn clause are called *not operational* or *not observable*, meaning that their truth values can only be determined indirectly on the basis of some other operational predicates [Mitchell, Keller, & Kedar-Cabelli, 1986]. Operational predicates are denoted by lower case letters (ex: water(x)), non operational predicates are denoted by upper case letters (ex: HEAT-SOURCE(x)).

4.1.1. The causal model. The second components of WHY's representation framework, the causal graph C , is defined by means of a set S of primitive graphical symbols and a special oriented link, $\rightarrow$, named *causal link*, which denotes a *causal relationship*. The symbols belonging to the set S are reported in figure 1(a).

A typical graphical unit occurring in C is represented in figure 2(b). Such units are the building blocks of the graph. Causal links connecting primary nodes form *causal paths*.

Figure 1. (a) Graphical elements occurring in a causal graph C . Primary nodes represent *processes* or *states*, involved in causal relations. Secondary nodes represent either *contexts* or (*necessary*) constraints. (b) Basic building unit of a causal graph. Nodes A and B are primary nodes, connected by causal links. The letter inside a node denotes a logical formula, whose operationalization set (if any) is specified by the triangle connected to the node itself. X, Y, Z and V denote sets of variables, not necessarily disjoint. Contexts (denoted by clouds) can only be associated to primary nodes, whereas constraints (denoted by rectangles) can only be associated to causal links.

Figure 2. Example of a causal model describing a qualitative understanding of everyday life situations concerning heat transfer. T_{Boil} and T_{Melt} are the boiling and melting temperature of substance x at the atmospheric pressure.

Primary nodes may or may not be observable. The primary nodes denote abstract situations (processes or states), whose instantiation produce changes, perceived as causally related pairs. Context nodes specify contextual conditions that must be true when the situation denoted by its related primary node is instantiated. Context nodes allow the common perception that a cause, to be effective, must find enabling conditions in the environment, to be modeled. Given their transient nature, contexts usually specify environment properties that may change from an experiment to another. Quite differently, constraint nodes specify time-independent, intrinsic properties of objects. Constraints must necessarily be true for the cause-effect relation to be instantiated. The difference between contexts and

constraints is similar to that proposed by Hestenes (1987) between “state variables” and “object variables”.

The *casual link* between the two primary nodes A and B, denoted by $A \rightarrow B$, postulates the existence of a *mechanism* which actually produces (“causes”) B whenever some conditions specified by A (and possibly some additional conditions in the context) become true. According to this definition, an implicit temporal relation is established between A and B, as B cannot precede A in time.

The causal graph C admits an easy-to-understand visual representation when WHY interacts with the user. Instead, for its internal use, the system utilizes a *frame-oriented* representation, containing all the information in a format more suitable to automatic processing.

4.2. Bodies of knowledge involved in learning

We will now enter into more details about the nature of the knowledge handled by WHY, and its role in the learning and problem solving processes. The knowledge handled by WHY can be roughly divided into four bodies: deep knowledge (the causal theory C), phenomenological knowledge, **Ph**, heuristic knowledge base, **KB**, and examples.

4.2.1. Deep knowledge. The deep knowledge constitutes the core of WHY’s explanatory framework. In figure 2, a simple causal model of heating is reported. This model describes the causal mechanism that underlies the observable behavior of a heated object as it may be perceived and understood by people in everyday life.

The depicted knowledge says that an observed substance that is being heated may melt, boil or just increase its temperature depending on its initial temperature value. The same causal knowledge can be used to explain several observable phenomena in terms of abstract and possibly not observable concepts. For instance, the causal model explains that heating some water with a heat source may result in the water boiling and becoming hot.

A set of phenomena occurring in the causal model deserves special attention: the ones represented by elliptic node without incoming arrows. These phenomena are called *first causes* as they are actually causing all the relevant domain phenomena. When looking for an explanation of an observed phenomenon, it is important to trace it back to the smallest set of first causes that potentially caused it.

4.2.2. Phenomenological knowledge. The phenomenological knowledge **Ph** contains the information needed to precisely define the domain of investigation, and to allow a concrete specification of any significant experimental setting. The phenomenological knowledge is represented as a set of Horn clauses of the language **L**. **Ph** contains different type of knowledge. In particular, **Ph** includes the domain vocabulary, which is organized into a set of *a-kind-of* taxonomies and is represented by rules such as:

$$\begin{array}{ll} \text{aluminum}(x) \Rightarrow \text{MATERIAL}(x) & \{\text{aluminum is a kind of material}\} \\ \text{thermometer}(x) \Rightarrow \text{OBJ}(x) & \{a \text{ thermometer is a kind of object}\} \end{array}$$

The most important information contained in **Ph** is a set of “phenomenological” rules that describe the alternative concrete manifestations of abstract concepts in terms of properties,

objects and events in the considered domain. This set of rules allows one to separate the part of knowledge that holds across several domains (mostly contained in **C**) from the more specific one, whose scope is limited to the context of experimentation. For instance:

$$\begin{array}{ll}
 \text{water}(x) \wedge \text{BOILING}(x) \Rightarrow \text{HOT}(x) & \{\text{boiling water is hot}\} \\
 \text{ice}(x) \rightarrow \text{COLD}(x) & \{\text{ice is cold}\} \\
 \text{to-make-bubbles}(x) \Rightarrow \text{BOILING}(x) & \{\text{if something makes bubbles then} \\
 & \text{it is boiling}\}
 \end{array}$$

Notice that even though the concepts of “hotness” and “coldness” are transversal to different domains, they are linked, by the above rules, to a concrete experimental setting involving substances and situations typical of everyday life. When changing the context, only the phenomenological part of the knowledge has to be replaced, but what has been learned at the deep level shall still hold. For example, when melting a piece of metal, the graph of figure 2 is still valid, but a different rule for assessing hotness may be used:

$$\begin{array}{l}
 \text{METAL}(x) \wedge \text{incandescent}(x) \Rightarrow \text{HOT}(x) \\
 \{a \text{ piece of metal appearing incandescent is hot}\}
 \end{array}$$

If composite apparata are considered in the domain, their structural description belongs to **Ph**. For instance the following rule describes a setting for heating liquids:

$$\begin{array}{l}
 \text{CONTAINER}(x) \wedge \text{PLATFORM}(y) \wedge \text{on}(x,y) \wedge \text{camping-gas}(z) \wedge \text{on}(y,z) \wedge \\
 \wedge \text{part-of}(x,k) \wedge \text{part-of}(y,k) \wedge \text{part-of}(z,k) \Rightarrow \\
 \Rightarrow \text{LIQUID-HEATING-APPARATUS}(k) \\
 \{a \text{ container placed on a platform which is on a camping-gas, builds up a} \\
 \text{heating apparatus}\}
 \end{array}$$

The last type of knowledge encoded in **Ph** is domain-independent background knowledge. This includes, for instance, the notions of symmetry, transitivity, spatial and temporal relations. Some examples are:

$$\begin{array}{ll}
 \text{full-of}(x,y) \Rightarrow \text{INSIDE}(y,x) & \{\text{if } x \text{ is full of } y \text{ then } y \text{ is inside } x\} \\
 \text{INSIDE}(x,y) \Rightarrow \text{CONTACT}(x,y) & \{\text{if } x \text{ is inside } y \text{ then } x \text{ is in} \\
 & \text{contact with } y\} \\
 \text{CONTACT}(x,y) \Rightarrow \text{CONTACT}(y,x) & \{\text{symmetry of the contact relation}\} \\
 \text{greater}(x,y) \wedge \text{greater}(y,z) \Rightarrow \text{GREATER}(x,z) & \{\text{transitivity of the} \\
 & \text{“greater-than” relation}\}
 \end{array}$$

4.2.3. Heuristic knowledge base. The heuristic knowledge base **KB** contains the knowledge used to perform classification over a set **H** of predefined classes. It consists of a set of rules in the language **L**. **WHY** can learn the knowledge base from scratch, using a set of examples; or, it can refine an already available knowledge base in the light of new incoming information: new examples or new phenomenological information. In the case of refinement, the previous **KB** can either have been learned by **WHY** itself, or it can have been provided by an expert of the domain (teacher).

Figure 3. Internal structure of the heuristic rules. To each of them, a set of alternative operational rules is associated by means of AND-OR forest.

In WHY the performance module is fully integrated with the learning module. The knowledge base **KB** is kept in an internal format which mirrors the way decision rules are acquired, and allows a fast classification process to be performed. Generally, rules in **KB** contains non-operational predicates thus allowing the number of rules to be kept low, and their structure as simple and as understandable as possible. An example of a heuristic rule with its internal structure is given in figure 3.

In figure 3, the rule states that the quantity of heat supplied by two heating sources that appear to be the same is actually the same. Note that in figure 3, the predicates 'HEAT-SOURCE' and 'SAME-HEAT-SUPPLIED' are non-operational, whereas the predicate 'same-features' is operational. In the picture, the AND-OR trees describe the types of heat sources (camping gas or gas stove or candle) observed in the experimentation. Note that allowing the occurrence of non operational predicates in the rule body, the structure reported in figure 3 accounts for 9 different heuristic rules containing only operational predicates in their bodies.

4.2.4. Examples. The last body of knowledge used by WHY consists of a set of labeled examples. Examples are described as ground first order logic formulas, and are represented extensionally in tables of a relational database, from which corresponding ground logical formulas can be derived. Examples may be complex scenes that contains several objects.

4.3. Reasoning mechanisms

WHY is provided with a sophisticated reasoner, able to perform induction, deduction, prediction and abduction. Induction is performed by invoking an inductive module that generates a knowledge base **KB** starting from a set of labelled examples. In the purely inductive learning mode, the deep theory is not used, and the acquired knowledge base may be totally unrelated to **C**.

Deduction is performed on the phenomenological theory **Ph** in the same way as in Explanation Based Learning EBL [Mitchell, Keller, & Kedar-Cabelli, 1986]. Goal of deduction is to assert the truth of non observable predicates, starting from the truth state of observable ones.

Prediction and abduction are oriented to manipulate causal knowledge. In fact, the causal graph **C** can be used both to make predictions about the occurrence of phenomena, and to perform abduction in search for first causes of observed phenomena. In both cases, chains of cause-effect relations are to be activated, from causes to effects in the former, and from

effects to causes in the latter. Both processes must be triggered by the activation of some node in $\mathbf{C}$. A node in the causal graph $\mathbf{C}$ is said *activated* if any of the operationalizations associated to the node is true.

If a primary node (i.e., an elliptic one) has been activated, all its effects become activated in turn. Then, a primary node in $\mathbf{C}$ can become activated either because it has been activated by some observable predicates, or because it has been predicted by a cause-effect chain. As the causal reasoning is the core of WHY's capabilities, we want to describe it more precisely. Let us consider the unit in figure 1(b), and the formula:

$$\Psi_p(\mathbf{X}, \mathbf{Y}, \mathbf{V}) = A(\mathbf{X}) \wedge C(\mathbf{Y}) \wedge N(\mathbf{V})$$

In *prediction*, the truth of $B(\mathbf{Z})$ can be asserted when $A(\mathbf{X})$ and $N(\mathbf{V})$ are true, whereas $C(\mathbf{Y})$ may be either true or unknown (but not false). On the opposite, if some phenomenon is observed, the chains of cause-effect relations can be followed starting from the observed effect(s), trying to arrive at some first cause. When using *abduction*, we consider the formula:

$$\Psi_a(\mathbf{Z}, \mathbf{V}) = B(\mathbf{Z}) \wedge N(\mathbf{V})$$

Then, from the truth of $\Psi_a(\mathbf{Z}, \mathbf{V})$ we assert the truth of both $A(\mathbf{X})$ and $C(\mathbf{Y})$. However, we have to be careful in this type of reasoning. In fact, abduction is not a truth-preserving mechanism and, hence, the truth of the cause (and of its associated context) is actually only *hypothesized* to have occurred.

Before describing the relationships existing among the bodies of knowledge, and the roles they play, we want to clarify the notion of explanation in WHY. An *example explanation* denotes a causal chain that “identifies the causal mechanism” underlying a phenomenon observable in the example. To be more precise, let $\mathbf{T} = \mathbf{C} \cup \mathbf{Ph}$ be the domain theory. Let, moreover, Σ be the set of first causes in the causal network $\mathbf{C}$. An example $q = \langle d, h \rangle$ is a pair consisting of a description d (a ground formula of $\mathbf{L}$) and a classification h corresponding to an observed property. Formally:

Definition 1 Given a theory $\mathbf{T}$ and an example $q = \langle d, h \rangle$, we call *explanation* $E(q)$ of q , with respect to $\mathbf{T}$, a first cause $\sigma \in \Sigma$, such that:

$$\mathbf{T} \cup \sigma(q) \cup d(q) \vdash h(q)$$

Intuitively, Definition 1 states that, given an example q , there exists a first cause σ in $\mathbf{C}$ from which a causal path, explaining the phenomenon h , starts.

4.4. An instance of example explanation in WHY

The relationships between the bodies of knowledge of WHY can be described by showing a case-study. In the following an instance of example explanation and derivation of a heuristic rule from it is described. Let us consider a specific heating experiment: some water (a), contained into a pot (b), is heated by means of a camping gas (c); the water starts boiling and it becomes hot. The observed heating phenomenon is that “some water is hot”.

Figure 4. Explanation of the observed phenomenon: "Water is hot". The relationships between non operational predicates of *C* and phenomenological rules describing their operationalizations are shown by using AND-OR forests. Notice that "contact" denotes thermal contact. The explanation in WHY is actually ground.

Figure 4 shows the *example explanation* for the phenomenon "Water is hot" observed in the considered experimental situation. The causal path on the left of figure 4 has been automatically determined by WHY through the successive applications of abductive reasoning steps starting from the node labeled 'Hot(x)'. The node 'Hot(x)' is activated by the given experimental situation as it corresponds to the phenomenon directly observed in the experimental situation. The resulting causal chain explains that a liquid (in this case, the water) is hot because of the heat transfer produced by an heat source (the camping gas). The AND-OR trees starting from some non operational predicates of the causal path represent those phenomenological rules that 'fired' when the instance description was supplied to the analysis of the system. Notice that the explanation in WHY is ground; that is, WHY explains a particular phenomena.

After finding an *example explanation*, WHY memorizes the performed reasoning activity by building a heuristic rule, i.e. a concise piece of knowledge, to be stored into its knowledge base. The heuristic rule to be built is almost a compilation of the discovered causal path, made by turning the explanation constants into variables as it is done in EBL [Mitchell, Keller, & Kedar-Cabelli, 1986]. Thus the rule is justified by the original explanation. For this case-study, WHY learns this heuristic rule:

$$\begin{aligned} & \text{HEAT-SOURCE}(y) \wedge \text{CONTACT}(x,y) \wedge \text{LIQUID}(x) \\ & \wedge \text{temp}(x, T_{\text{Boil}}) \wedge \text{temp-higher-than}(x, T_{\text{Hot}}) \Rightarrow \text{HOT}(x) \end{aligned}$$

WHY builds the heuristic rule as follows. First, it defines a class HOT(x) according to the observed phenomenon, then it collects all the operational and non operational predicates encountered along the causal path, with the exception of those inside the elliptic nodes. In fact elliptic nodes represent abstract processes or states that may not be measurable. In the

Figure 5. Knowledge structure for the heuristic rule build from the phenomena: "Hot water". The relationships between the 'abstract' form of the heuristic rule and the predicate operationalization are shown.

end, a heuristic rule only contains a set of sufficient conditions to assert a classification, but it is not complete compilation of a deep reasoning activity.

Figure 5 shows the relationships between a heuristic rule and the phenomenological theory. The non operational predicates occurring in the rule body have to be verified on a future example by using some phenomenological rules in **Ph**. It is important to note that only those rules activated during the construction of the original causal path are actually linked with the rule's non-operational predicates.

In case the same causal path is activated for another example with a different experimental setting (for instance, the heat source is a gas stove), possibly new operationalizations of a same predicate are OR-ed with the old ones. It is easy to imagine the considerable compactness obtained in the **KB**, due to the summarization of a possibly exploding number of rules into a few abstract ones, much easier to memorize and retrieve. A small price to pay for this economy in memory and computation is a possibly increased time in rule matching; however, this also can be alleviated by a matching strategy that exploits a suitable predicate ordering inside a rule body.

As concerns the conceptual links between the heuristic knowledge base **KB** and the deep knowledge, we observe that the rules are compiled from the causal graph. These compiled rules, on the one hand, are effective short-cuts to speed up deep reasoning, which is usually slow, and, on the other hand, they provide answers to question which are backed-up by the theory, and hence, justified. Learning to "understand" consists, in this framework, in moving from a heuristic knowledge base unrelated with the causal model to one in which each rule is justified. This process corresponds to a shift from novice to expert.

4.5. Learning modes in WHY

WHY can perform either one-step learning or incremental learning. In the scope of this work, we exploited both modalities: one-step learning allowed the acquisition of a knowledge base to be compared with the performances of the student we were modeling; incremental

Figure 6. Scheme of WHY's one-step learning mode. WHY's output is a heuristic knowledge base of justified rules and a set of explanations of the training examples.

learning allowed the investigation of the limits of the current model and the refinement of the student's model as long as new information was acquired by the student during the teaching period. A concise description of WHY learning modes is provided here and a more detailed description can be found in [Saitta, Botta, & Neri, 1993]. Figure 6 shows the scheme of one-step learning mode.

In one-step learning, WHY only acquires totally justified rules, starting from a set of labelled examples, and is able to explain all the examples. In this phase, the domain theory T is supposed to be sufficiently correct and complete to explain the given examples. If it is not so, the teacher, in an interaction phase, makes to T the necessary additions and/or corrections. WHY's refinement mode is described in figure 7.

When a classification error is uncovered, WHY can suggest possible theory revisions by means of the analysis of the content of C and Ph . In the current version of the system, WHY only points at the error, and the modifications to the prior knowledge have to be done manually by the teacher. The human expert can benefit from WHY's theory analysis in order to guide the system toward more promising rule specializations or generalizations, telling explicitly which predicates try to remove or add.

5. The modeling methodology and an instance of conceptual change

As WHY learns knowledge for classification tasks, and provide explanations thereof, the student's modelization has to be mapped accordingly. From a practical point of view, this requires to make a decision about which kind of knowledge available to the student has to be described as a causal model of the domain, or as phenomenological knowledge or as heuristic knowledge. In Table 1 a summary of the chosen correspondence between the knowledge organization in WHY and in the student is reported.

The knowledge decomposition of Table 1 takes into account primarily the epistemological framework of knowledge organization proposed by Tiberghien (1994), discussed in Section 3.2, and, secondarily, the peculiar capabilities of representing and exploiting knowledge of WHY. We believe the resulting mapping to be representative and respectful of Tiberghien's proposal.

On the basis of this knowledge organization and using the experimental data provided by Tiberghien (1994), we will show how the model can be built up and used through a specific

Table 1. Correspondence between knowledge decomposition in WHY, in the human teacher/learner, and in Tiberghien's framework. The epistemic constructs and the phenomenological knowledge are transversal to all other types of knowledge.

WHY	Student and teacher	Educational framework
Causal theory	Knowledge with explanatory power (Material and efficient causality)	Theory
Heuristic knowledge base	Domain-specific knowledge (Knowledge with prediction power)	Model
Examples	Questions experiments	Experimental field of reference

Figure 7. Scheme of WHY's knowledge refinement learning mode. WHY receives in input an initial heuristic knowledge base **KB** and a new example, q , and outputs either a new knowledge base, **KB'**, with corresponding example explanations, or reports a situation of error ("impasse").

example: the knowledge evolution of the 12-year-old student “David”. The material available from David’s history contained answers to questionnaires and their motivations, before and after the teaching period; instances of these data can be found in figure 8 and figure 10.

In order to model David’s knowledge with WHY, each practical experiment is represented as an example, consisting of two parts: a description of the experimental setting and a

Questionnaires (before teaching)			
What happens ...	when water is heated but before it starts boiling?	When the water is boiling?	Why?
The water makes bubbles	False	True	Because when it boils, it produces bubbles
The water becomes steam	False	False	Because heating need to be very strong to let the water evaporate
The water temperature does not change	False	False	Because when you heat water, it becomes hot
The water temperature changes	True	True	Because when you heat water, it becomes hot

What happens to ... when heated?	Does it become a Liquid?	Does it become a Gas?	Why?
Salt	I do not know	I do not know	
Lead	Yes	False	When the lead reaches a given temperature, it melts

Interview (before teaching)
Teacher: “You put some water into a pot and you heat it with a camping gas. You can heat the water as long as you want. What do you think will happen to its temperature?”
David: “It will keep raising. I think that it will keep raising all the time.”
Teacher: “You put some sand into a pot and you heat it with a camping gas. You can heat the sand as long as you want. What do you think will happen to its temperature?”
David: “It will keep raising. I think that it will keep raising all the time”

Analysis done by the Educational Scientist
Causal schemata are observable in David’s reasoning:
“X heats Y CAUSES Y becomes hot”
“The more X heats Y CAUSES the hotter Y becomes”
Observation about David’s conceptions:
a) The phenomenon of phase transition does not appear to be known.
b) The kind of phenomena occurring during heating are strongly dependent on the substance involved: water will boil, lead will melt, etc.
c) The temperature of a substance is going to increase as long as the substance is heated (i.e., David does not appear to know that during phase transition the temperature stays constant)

Figure 8. A selection of David’s answers provided at the beginning of the cycle of lessons. The part of the questionnaire filled in by David is reported in plain text.

question. The experimental setting corresponds to the description of the example, whereas the possible answers to the questions are considered as alternative classes. Then, the process of predicting the outcome of an experiment is mapped onto the problem of predicting the correct answer.

In the current experimentation, the various knowledge bodies **C**, **Ph** and **KB** have been manually constructed and encoded by the experimenters. WHY relies on a sophisticated algorithm for uncovering errors or incompleteness in its knowledge; it can be triggered when one of WHY's explanations (answers) does not match the student's one. This provides useful information to the educational/cognitive scientist in discovering where his/her hypothesized student's model is incorrect.

In the following, we give an idea of the data and of the formalism used to represent the student's knowledge in WHY. Then, we will compare two knowledge states of David, represented as two WHY's causal models. Finally, we sketch how, by comparing David's knowledge before and after teaching, a *conceptual change*, actually observed in David during the teaching period, appears in our formalism.

David's knowledge before teaching has been inferred on the basis of his initial answers to two questionnaires and interview; a selection of them appears in figure 8.

By analyzing the questionnaires and the interview, we could infer that David used a notion of material causality linked to the "substance" of a body; in fact, what happens to a body depends on what "it is". For instance, water will eventually boil, if heated, whereas lead or iron or gold will melt; for this reason, they become hot. Similarly, sugar becomes "caramel" and, again, it becomes hot. Questioned on the subject, David showed evidence to believe that "boiling" and "melting" are alternative (and mutually exclusive) behaviors, exhibited by different substances. In fact, he says that iron, gold and lead shall not boil, *because they melt*. Because of linking the behavior of the bodies under heating to specific substances, David is unable to answer questions about materials that he does not know about: for instance, he answers "I don't know" to questions about the possibility of diamond, salt and aluminum to become liquid or gaseous.

After analyzing all the available data, we were able to come up with a formal representation in WHY's language of David's knowledge relevant for the task. David's knowledge has been represented as heuristic rules, phenomenological rules, or causal relationships, depending on the peculiar role played by each knowledge part. However, in the rest of the paper we focalize our discussion on David's theory (the causal model of the domain), it is at the level of the theory that the most interesting and deep changes occur during learning. In figure 9, part of David's theory is represented in WHY's formalism. In the picture, it is easy to recognize David's conception as reported in figure 8 along with those that have been found through the complete analysis of the available data. For instance, the conception that "while a substance is heated, its temperature is going to steadily increase" is represented by the causal relation between "TO-HEAT" and "TEMP-INCREASE".

Note also that we can represent "don't know" situations, such as in the case of heating salt, by not describing the phenomena in the causal model. Reading it in the reverse way, any knowledge not explicitly appearing in the causal model does correspond to "don't know" answers by David. After David's theory is represented as a causal model, WHY can be used to provide both explanations and predictions about situations unseen by David. In turn,

Figure 9. Causal model hypothesized to represent part of David's theory before teaching. Material causality dominates the model.

when the answers provided by the system show an “incorrect” behavior with respect to the target “official” knowledge of the course, David can be confronted on the same situation and eventually a better-focalized lesson could be planned.

An example of use of the causal model by WHY has been given in Section 4.5.

After the teaching period David has seen several experiments involving different materials. From the analysis of his final questionnaires and his final interview, we infer that his deep knowledge of the world is changed, under various respects. Figure 10 shows a selection of David's answers and explanations provided at the end of the cycle of lessons.

By comparing David's knowledge at the beginning and at the end of the cycle of lessons, we were able to analyze what David has learned. In the following, we will illustrate just one of the most important changes to David's theory, which can be considered as a conceptual change. The most important change, with respect to the goal of the teaching course, is that David seems less committed to material causality for determining behaviors. In fact, he is able to generalize, from “iron”, “lead”, “gold” and “ice”, that any “solid” may become liquid if sufficiently heated. Moreover, “to boil” and “to melt” are no more mutually exclusive behaviors, but they are possibly in sequence, as it should be. David's principle of causality shows a shift of the “cause” from the “substance” to some underlying process, which, on the other hand, he is not yet capable of pinning down. The educational scientist has described the dynamic of David's conceptual change as a two-phase activity:

- 1) During the lessons, David performed manipulations in which a number of different substances undergo a change of state: he observes ice transformed into water, water into

Questionnaires (after teaching)			
What happens ...	when the water is heated but before it is boiling?	when the water is boiling?	Why?
The water temperature does not change	False	True	
The water temperature changes	True	False	

What happens to ... when heated?	Does it become a Liquid?	Does it become a Gas?	Why?
Salt	Yes		
Lead	Yes	Yes	Because I did the experiments

Interview (after teaching)

Teacher: "You put some water into a pot and you put some salt into a different pot. Then you heat both with two camping gas. You can heat them as long as you want. What do you think will happen to their temperature?"

David: "The water will reach about 100 °C and then it will become steam, (the water) stays at about 100 °C until all the water has evaporated."

The experience is done in practice, and while the water evaporates the salt does not change.

Teacher: "Now, what do you think has happened to their temperature?"

David: "The water reached about 100 °C and then it became steam, instead the salt could reach till 120 °C also 130 °C..."

There (the water) started as liquid, there (the salt) started as solid. I think that all the substance, if we take the water or the lead, I think the lead is solid ..., so below 0 °C but non a 0 °C it (the water) is ice, well the lead has another temperature, it is solid lead, a solid substance, after over 0 °C it becomes water till 100 °C; the lead is about 327 °C, I believe, it will melt, it will become liquid, and afterwards (the water) becomes a gas and also (the lead) becomes a gas...

Then, I thought that this (the salt) became liquid at a given temperature as the water does solid, liquid, gas, may be, I think that it (the salt) also should becomes solid, liquid, gas at a given temperature. "

Analysis done by the Educational Scientist

Observations about David's conceptions are:

- The phenomenon of phase transition appears in David's explanation.
- The phenomena occurring during heating are not dependent on the substance involved.
- The temperature of a substance is going to increase as long as the substance is heated, unless it is changing phase.
- Each substance has characteristic temperature for melting and evaporating.

Figure 10. A selection of David's answers provided at the end of the cycle of lessons. The part of the questionnaire filled in by David is reported in plain text.

ice and water into vapor, as well as lead and salt to melt. Then, he keeps adding to his theory additional information, corresponding to the new observed substances. This phase involves also learning of basic terms or notions such as that "aluminum will behave as the gold, when heated" or that "lead, gold, iron and aluminum are metals".

- 2) Then, observing commonalities (also helped by the teacher) he notices that all the substances that melt are solid and that vapor derives from liquids: he generalizes the heating behavior of specific substances by exploiting the concepts of "solid" or "liquid".

Figure 11. Part of David's hypothesized knowledge after teaching. Conceptual change is represented by the bold causal path.

In figure 11, part of David's theory at the end of the cycle of lesson is represented. As before it is easy to recognize David's conceptions corresponding to the excerpt of the answers reported in figure 12. Also the comparison of the graphical representation of David's knowledge before and after teaching (compare figure 9 and figure 11) does consistently match the learning dynamics reported by the educational scientist.

However, the graphical comparison also tells us something that was missed during the manual analysis of the data by the educational scientist. It appears that even after teaching, heating some substances such as alcohol or sugar can still be considered by David's as subject to a different kind of phenomenon. This observation is consistent with the experimental data collected by the educational scientist. In fact, she was favorably impressed by this result and recognized that questions about these specific substances should have been asked to David in order to test his degree of generalization.

6. Future work: modeling the dynamics of learning

The next research step consists in studying and possibly defining, inside our framework, the basic knowledge refinement mechanisms that underlay learning. Efforts in this direction

have been accomplished but work is still needed to obtain a sufficiently detailed definition of such knowledge refinement mechanisms. In the following we report the work in progress.

As a reference for knowledge transitions during learning, we consider Rumelhart and Norman's work (1977). They categorized knowledge transitions into three kinds: *Accretion*, *Tuning* and *Restructuring*. Accretion involves addition of new information that does not contradict previous knowledge to existing theories. When the new information is inconsistent with previous theories, tuning or restructuring may occur. Tiberghien's epistemological framework (1994) is consistent and, actually, she implicitly relies on these mechanisms to explain student's learning. However, neither Rumelhart and Norman nor Tiberghien advanced any hypothesis about overall learning strategies, i.e. when and in which order refinements are performed, in such a way to be sufficiently detailed to be implemented in a computer program.

As said above, within the scope of this research, WHY has been essentially used as a smart knowledge revision system that helped the teacher in developing her model of the students by providing useful hints during the interactive sessions. Given the current state of the art, we can only reserve for the future the proposal of a control strategy for applying the knowledge revision operators in such a way to completely mimic the evolution of the student's knowledge.

Notwithstanding this limit, WHY proved to be an interesting computational framework for analyzing more sophisticated learning strategies. In the following, we show that WHY can represent types of knowledge transitions such as Rumelhart and Norman's ones. Then, we discuss how accretion, tuning, and restructuring can be mapped to the application of one or more WHY's revision operators to a specific portion of the domain knowledge.

Accretion: Learning by accretion *increases coverage*, in the sense that more experimental situations can be handled, independently of their correct interpretation/prediction. A typical situation in which accretion occurs is when a student memorizes, without explanation, a piece of information taught by the teacher. This information can be added, as a rule, to the student's knowledge, without checking for compatibility with previous knowledge. The rule can be accessed, on the basis of recency, for giving a correct answer for a while, until it is forgotten, or other pieces of knowledge, incorrect but supported by the student's deep beliefs, gain higher priority.

As in WHY, a "rule" is applicable to a situation when all the conditions specified by the rule's antecedent are defined, i.e. they have a "true" or "false" value, we could associate to accretion the following *operational* properties:

- 1) Accretion may affect both the *phenomenological theory* and the *heuristic knowledge base*, by adding to any of them a new rule.
- 2) Accretion can be implemented with WHY's "adding rule" operator.

Tuning: Tuning aims at *increasing the number of correct predictions or explanations*, but does not modify the student's deep explanatory framework. A tuning situation is, for instance, learning that "aluminum will behave as the gold when heated". In WHY terms, tuning has the following operational properties:

- 1) Tuning may affect the *heuristic knowledge base* or the *phenomenological theory*, by changing the preconditions in some rule or by adding new rules.
- 2) Tuning may also affect the *causal model*, but only by adding/deleting conditions in the constraint and context nodes (i.e., causal nodes and links between them cannot be changed).
- 3) Tuning can be implemented with WHY's operators that generalize/specialize rule antecedents, or add and delete rules.

Restructuring: Restructuring corresponds to a deep change in the way phenomena are understood. Then, restructuring affects the explanatory framework. In the case of David, a restructuring situation is learning to interpret heating phenomena in terms of phase transitions. In WHY's terms, restructuring has the following operational properties:

- 1) Restructuring affects the *causal model*, via addition/deletion of causal nodes, or modification of causal links.
- 2) Restructuring can be implemented with WHY's operators that add/delete nodes or links in the causal model.

As a general observation, we believe that this mapping straightforwardly matches WHY's knowledge representation, and it is respectful of the cognitive meaning of the learning mechanisms. Moreover, the knowledge stored in different bodies can only undergo changes that are significant for that body of knowledge. For instance, restructuring does apply to the causal model, but, as a side effect, changes in other knowledge bodies may be stimulated. On the contrary, it does not make sense to restructure the heuristic knowledge leaving unchanged the causal model.

In the light of the above information, we may reconsider the case-study of the previous section. The teacher individuates two important steps in David's learning:

- 1) the addition of pieces of information involving different behaviors for different substances. This can be thought of as a sequence of Accretion/Tuning steps;
- 2) the generalization of several pieces of information from specific substances to the concept of "solid" or "liquid".

From an operational perspective, the first step results in the addition of constraints, representing alternative substances, on the links of the causal net. The second step corresponds to a generalization of them from specific substances to "solid" or "liquid", and, eventually, in a collapse of the causal links into a single one. This collapse affects the causal structure and can be considered restructuring.

7. Conclusion

We discussed a way of interpreting learning in relation with teaching in the domain of Physics. Our analysis is based on knowledge processing as a relevant frame of reference

from the points of view of perceived personal knowledge of a learner. Our framework allows us to establish independence between analysis of the learner's acquisition and of Physics knowledge. Thus, it is possible to take into account the *coherence of the learner per se*, even if it is incompatible with Physics. The structuring of the analysis in terms of causal theory, phenomenological knowledge, and field of applicability allows the learner's knowledge at different moments to be compared. This independence and this comparison are essential to the characterization of different types of learning. WHY's articulated knowledge representation allows most of phenomena observed in children learning elementary Physics to be modeled, notably their explanation in terms of simple causality, and the interdependence of "surface" heuristic knowledge and deep beliefs.

In the future, this work could be extended along two directions. The first is the study of specific learning strategies, allowing a complete simulation of a student's evolution to be performed, by exploiting the elementary operators available in WHY. The second direction is to extend the experimentation. We plan to collect more in-field data about the conception of high school students on mechanics and to test the use of WHY by a part of a group of prospective teachers in such a way to evaluate its impact in the selection of teaching sequences or strategies.

We feel that such ambitious tasks can be achieved by means of a strong interdisciplinary cooperation and by exploiting automatic systems able to deal with knowledge representation that tries to closely match the one used by the expert of the domain. Computational systems that match this requirement are more easily accepted by the domain expert as she/he can think in familiar terms about the application and then exploit at best her/his experience. We have shown how multi-strategy learning systems, such as WHY, can provide a significant contribution to the application of Machine Learning to this new challenging domains.

Note

1. In Physics, several layers of explanatory knowledge may be envisaged. It is up to the teacher to select the ultimate explanations that she considers congruent with the goal of teaching.

References

- Anderson, J. R. (1985). *The Architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Anzai, Y. & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review*, 86, 124–140.
- Baffes, P. T. & Mooney, R. J. (1996). A novel application of theory refinement to student modelling. *Proc. of Thirteenth National Conference on Artificial Intelligence* (pp. 403–408) Portland, OR.
- Baroglio, C., Botta, M., & Saitta, L. (1994). WHY: A system that learns from a causal model and a set of examples. In R. Michalski & G. Tecuci (Eds.), *Machine learning: A multistrategy approach, Vol IV* (pp. 319–348) Los Altos, CA: Morgan Kaufmann.
- Brown, A. L. (1989). Analogical learning and transfer: What develops? In S. Vosnadiou & A. Ortony (Eds.), *Similarity and analogical reasoning*, Cambridge, MA: Cambridge University Press. pp. 369–412.
- Caravita, S. & Halldén, O. (1994). Re-framing the Problem of Conceptual Change. *Learning and instruction*, 4, 89–111.
- Carey, S. (1983). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

- Carey, S. & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind* (pp. 169–200). Cambridge, Cambridge University Press.
- Chi, M. T. H. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models of science*, (pp. 129–160) Minnesota Studies in the Philosophy of Science, Minneapolis, MN: University of Minnesota Press.
- Chi, M. T. H., Bassok, M., Lewis M. L., Reiman, P., & Glaser, R. (1989). Self-explanation: how students study an use examples in learning to solve problems. *Cognitive Science*, 13, 145–182.
- Chi, M. T. H., Slotta, J. D., & de Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and instruction*, 4, 27–43.
- diSessa, A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10, 105–225.
- diSessa, A. (1996). What do “just plain folk” know about physics. In D. Olson & N. Terrance (Eds.), *Handbook of education and human development*, (pp. 709–730) Blackwell Publ..
- Driver, R., Guesne, E., & Tiberghien, A. (1985). *Children's ideas in science*. Milton Keynes, UK. Open University Press.
- Duit, R. (1995). “Constraints on knowledge acquisition and conceptual change the case of physics. *Proc. of Symposium on Constraints on Knowledge Construction and Conceptual Change: A Look across Content Domains* at 6th European Conference for Research on Learning and Instruction, Nijmegen, The Netherlands.
- Dykstra, D. (1992). Studying conceptual change: Constructing new understandings. In R. Duit, F. Goldberg, & H. Niedderer (Eds.), *Research in physics learning: Theoretical Issues and Empirical Studies* (pp. 40–57) Kiel, Germany: Institut für die Pädagogik der Naturwissenschaften.
- Erickson, G. L. (1979). Children's conceptions of heat and temperature. *Science Education*, 63, 221–230.
- Forbus, K. D. & Gentner, D. (1986). Learning physical domains: Toward a theoretical framework. In R. Michalski, J. Carbonell, & T. Mitchell (Eds.), *Machine learning: An artificial intelligence approach, Vol. II* (pp. 311–348) Los Altos, CA: Morgan Kaufmann.
- Gertner, A. S., Conati, C., & VanLehn, K. (1998). Procedural help in Andes: Generating hints using a bayesian network student model, *Proc. of National Conference on Artificial Intelligence (AAAI'98)*, Madison (WI), pp. 106–111.
- Halloun, I. A. & Hestenes, D. (1987). Modeling instruction in mechanics. *American Journal of Physics*, 55, 455–462.
- Hestenes, D. (1987). Toward a modelling theory of physics instruction. *American Journal of Physics*, 55, 440–454.
- Klahr, D. & Siegler, R. S. (1978). The representation of children's knowledge. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (pp. 61–116) Academic Press, New York, NY.
- Kuhn, T. S. (1971). Les notions de causalité dans le développement de la physique. In M. Bunge, F. Halbwachs, J. Piaget, & L. Rosenfeld (Eds.), *Les théories de la causalité* (pp. 7–18) P.U.F., Paris, France.
- Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*, Cambridge, MA: Harvard University Press.
- Laird, J. E., Rosenbloom, P. S., & Newell, A. (1986): Chunking in SOAR: The anatomy of a general learning mechanism, *Machine Learning*, 1, 11–46.
- Langley, P. (1987). A general theory of discrimination learning. In D. Klahr, P. Langley, & R. Neches (Eds.), *Production system models of learning and development*, Cambridge, MA: MIT Press.
- Lewis, C. (1987). Composition of productions. In D. Klahr, P. Langley, & R. Neches (Eds.), *Production system models of learning and development*, Cambridge, MA: MIT Press.
- Mitchell, T., Keller, R. M., & Kedar-Cabelli, S. (1986). Explanation-based generalization: A unifying view. *Machine Learning*, 1, 47–80.
- Murphy, L. G. & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289–316.
- Neri, F. (1998). Simulating children learning and explaining elementary heat transfer phenomena: a multistrategy system at work. *Proc. of European Conference on Machine Learning (ECML98)*, Lecture Notes in Artificial Intelligence series, vol. 1398 (pp. 67–76).
- Neri, F. (1999). Computer aided tracing of childrens physics learning: a teacher oriented view. *Proc. of International Joint Conference on Artificial Intelligence 99 (IJCAI99)*, Stockholm, Sweden. In press.
- Neri, F. & Saïta, L. (1993). Exploiting sample selection and ordering to speed-up learning. *Proc. AAAI Symposium on Training Issues in Incremental Learning* (Stanford, CA), pp. 54–69.

- Newell, A. (1990). *Unified theories of cognition*, Cambridge, MA: Harvard University Press.
- Nordhausen, B. & Pazzani, P. (1993). An integrated framework for empirical discovery. *Machine Learning*, 12, 17–47.
- Ohlsson, S. (1987). Truth versus appropriateness: Relating declarative to procedural knowledge. In D. Klahr, P. Langley & R. Neches (Eds.), *Production system models of learning and development*, Cambridge, MA: MIT Press.
- Ohlsson, S. (1996). Learning to do and learning to understand: A lesson and a challenge for cognitive modeling. In P. Reiman & H. Spada (Eds.), *Learning in Humans and Machine* (pp. 37–62) Pergamon Press, Oxford.
- Pazzani, M. (1991). “Influence of prior knowledge in concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 416–432.
- Piaget, J. (1974). *Understanding causality*, New York, NY: Norton & Co.
- Rumelhart, D. E. (1989). Toward a microstructural account of human reasoning, In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 299–312) New York, NY: Cambridge University Press.
- Rumelhart, D. E. & Norman, D. A. (1977). Accretion, tuning and restructuring: Three modes of learning. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic factors in cognition*, Hillsdale, NJ: Erlbaum.
- Sage, S. & Langley, P. (1983). Modeling cognitive development on the balance scale task. *Proc. 8th Int. Joint Conf. on Artificial Intelligence* (pp. 94–96) Karlsruhe, Germany.
- Saitta, L. et al. (1995). Knowledge representation changes in humans and machines. In P. Reimann & H. Spada (Eds.), *Learning in Humans and Machines: Towards an Interdisciplinary learning science*, (pp. 109–129) Oxford, UK: Elsevier Publ. Co.
- Saitta, L., Botta, M., & Neri, F. (1993). Multistrategy learning and theory revision. *Machine learning*, 11, 153–172.
- Salmon, W. (1989). *Four decades of scientific explanation*, Minneapolis, MN: University of Minnesota Press.
- Samarapungavan, A. & Wiers, R. W. (1997). Children’s thoughts on the origin of species: A study of explanatory coherence. *Cognitive Science*, 21, 147–177.
- Schmidt, W. C. & Ling, C. X. (1996). A decision-tree model of balance scale development. *Machine Learning*, 24, 203–230.
- Shultz, T. R., Mareschal, D., & Schmidt, W. (1994). Modeling cognitive development on balance scale phenomena. *Machine Learning*, 16, 57–86.
- Simon, T., Newell, A., & Klahr, D. (1991). A computational account of children’s learning about number conservation, In D. H. Fisher, M. J. Pazzani, & P. Langley (Eds.), *Concept formation: Knowledge and experience in unsupervised learning* (pp. 423–462) San Francisco, CA: Morgan Kaufmann.
- Sison, R., Numao, M., & Shimura, M. (1999). Multistrategy detection and discovery of novice programmer errors. *Machine Learning*, this issue.
- Sleeman, D., Hirsh, H., Ellery, I., & Kim, I. (1990). Extending domain theories: two case studies in student modeling. *Machine Learning*, 5, 11–37.
- Tiberghien, A. (1985). “Heat and temperature”, The development of ideas with teaching. In R. Driver, E. Guesne, & A. Tiberghien (Eds.), *Children’s ideas in science* (pp. 67–84) Milton Keynes, UK: Open university press.
- Tiberghien, A. (1989). Learning and teaching at middle school level of concepts and phenomena in physics. The case of temperature. In H. Mandl, E. de Corte, N. Bennett, & H. F. Friedrich (Eds.), *Learning and instruction. European research in an international context, Volume 2.1*, Oxford, UK: Pergamon Press, pp. 631–648.
- Tiberghien, A. (1994). Modelling as a basis for analysing teaching-learning situations. *Learning and Instruction*, 4, 71–87.
- VanLehn, K. (1990). *Mind Bugs: The origins of procedural misconceptions*. Cambridge, MA: MIT Press.
- VanLehn, K. & Jones, R. (1993a). Learning by explaining examples to oneself: A Computational model. In S. Chipman & A. L. Meyrowitz (Eds.), *Foundations of knowledge acquisition: cognitive models of complex learning*, Boston, MA: Kluwer.
- VanLehn, K. & Jones, R. (1993b). Learning by explaining examples to Oneself: A computational model. In S. Chipman & A. L. Meyrowitz (Eds.), *Foundations of knowledge acquisition: cognitive models of complex learning*, Boston, MA: Kluwer.
- Vosniadou, S. (1989). Analogical reasoning in knowledge acquisition. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 413–437) Cambridge, MA: Cambridge University Press.
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and instruction*, 4, 45–69.

- Vosniadou, S. (1995). A cognitive psychological approach to learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines: Towards an interdisciplinary learning science* (pp. 23–36) Oxford, UK: Elsevier Publ. Co.,
- Vosniadou, S. & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, 535–585.
- Vosniadou, S. & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science*, 18, 123–183.
- Weil-Barais, A. & Vergnaud, G. (1990). Student's conceptions in physics and mathematics: biases and helps. In J. P. Caverni, J. M. Fabre, & M. Gonzalez (Eds.), *Cognitive biases* (pp. 69–84) North Holland, Elsevier Science Publ.
- White, B. Y. & Frederiksen, J. R. (1987). Causal model progressions as a foundation for intelligent learning environments. Techn. Report No. 6686, BBN Laboratories, Cambridge, MA.
- White, B. Y. & Frederiksen J. R. (1988). Explorations in understanding how physical systems work. *Proc. of 10th Annual Meeting of the Cognitive Science Society*, Lawrence Erlbaum, Hillsdale, NJ: Lawrence Erlbaum.

Received December 22, 1998

Accepted June 18, 1999

Final manuscript June 18, 1999