

Simulating Children Learning and Explaining Elementary Heat Transfer Phenomena: A Multistrategy System at Work

Filippo Neri

Università di Torino
Dipartimento di Informatica
Corso Svizzera 185
10149 Torino (Italy)
neri@di.unito.it

Keywords: cognitive modelling, multistrategy learning, conceptual change, causality

Abstract

The multistrategy learning system WHY is used as a testbed for investigating a computational cognitive model of conceptual change in children learning elementary physics¹. Goal of the simulation is to support the cognitive scientist's investigation of learning in humans.

The student's mental model is manually inferred by the cognitive scientist, and by interacting with WHY, from a sequence of interviews collected along a period of eleven teaching sessions. The hypothesized cognitive models are based on a theory of conceptual change, derived from psychology results and educational experiences, which accounts for the evolution of the student's knowledge over a learning period.

The multistrategy learning system WHY, able to handle domain knowledge (including a causal model of the domain), has been chosen as tool for the *interactive* simulation of the cognitive models evolution. The system is able to model both the answers and the causal explanations given by the children. An example of modelisation of an observed conceptual change is provided.

1 Introduction

People acquire, in their lifetime, models of the world that they use to interpret data, to explain phenomena and to make predictions. These models usually evolve when new information is gathered, and their evolution can be described as a particular aspect of learning, called *conceptual change* [Tiberghien, 1989, 1994; Vosniadou & Brewer, 1994; Caravita & Halldén, 1994; Chi et al., 1994; Vosniadou, 1994]. The issue of conceptual change has been addressed from a variety of perspectives, but, even though quite a large body of experimental findings has been collected, still no single definition of conceptual change is universally accepted [White, 1994]. Conceptual

¹ This work has been performed within the project "Learning in Humans and Machines", supported by the European Science Foundation

change has been mainly studied in the context of learning Mathematics or Physics [Forbus & Gentner, 1986; diSessa, 1993; Vosniadou, 1994; Chi et al., 1994].

Goal of our research is to help the cognitive scientist in developing a **computational model** of the student. In turn, this could also produce a deeper understanding of how and when conceptual change happens thus allowing to fully automatize the model evolution process. Instead, we are not currently concerned with the task of providing the cognitive scientist with some machine-generated student's model to be tested with time consuming psychological experiments.

Models of conceptual change proposed in Cognitive Science have a **descriptive** nature: they describe mental models or knowledge states, but do not provide an account for the actual mechanisms of transition from a knowledge state to another.

Rumelhart and Norman [1977] have categorized the type of transitions occurring as *Accretion*, *Tuning* and *Restructuration*, which are reminiscent of Piaget's *Assimilation*, *Accommodation*, and *Self-Regulation*. Accretion involves addition of new information to existing theories, and presents no problem when the new information does not contradict previous knowledge. When the new information is inconsistent with previous theories, tuning or restructuring may occur. However, when a contradiction emerges, also *failures in learning* may happen, taking the form of *inert knowledge* or *misconceptions*.

Computational model of human learning have been studied in [Sleeman et al., 1990; Baffes & Mooney, 1996; Sage & Langley, 1983; Newell, 1990; Schmidt & Ling, 1996; Shultz et al., 1994]. However, two aspects are overlooked in these models: the first is the strict *interconnection* between the *heuristic knowledge* in a specific domain (substantially the one modelled in the Machine Learning systems) and pre-existing *deeper knowledge structures* or theories [Murphy & Medin, 1985; Vosniadou, 1994, 1995; Tiberghien, 1994; Chi et al. 1994]. The second aspect is the importance of *explanation*. Human learning is, to a great extent, a search for explanations; then, any model of human learning should provide an explanatory framework, allowing not only answers to questions to be predicted, but also reasons put forward in support of those answers to be formulated. Consequently, along the paper, we consider student's learning in the sense of EBL [Mitchell, Keller & Kedar-Cabelli; 1986]. That is the capability to explain previously unknown phenomena and, also, the ability to change the explanation level for the observed phenomena.

The main novelty of our approach is in the *differentiation* between the *pragmatic knowledge* a student uses to answer questions and/or to interpret experimental results, and an *explanatory framework*, which the student uses to "make sense" of what he/she observes or is taught. A central hypothesis of the approach is that explanation corresponds to *causal* attribution. This hypothesis derives from a number of previous studies (for instance, [diSessa, 1993; Tiberghien, 1994]), and from the direct observation that children, even young ones, spontaneously use verbal constructs suggesting causality. The learning/teaching context we want to model envisages the task of acquiring basic concept in Physics, specifically *Heat* and *Temperature* concepts by middle school students. The specific learning context is: a group of secondary school students (12-13 years old, 6-5th grades), were exposed to a Physics course, outside normal teaching, consisting of 11 sessions, once a week, including experimentation, questions, discussions and explicit teaching. Content of the course were basic concepts and qualitative relations in the domain of *heat transfer* in everyday life situations.

As computational modelling tool it has been selected the multistrategy learning system WHY [Saitta, Botta & Neri, 1993; Giordana et al., 1997], which learns and revises a knowledge base for classification problems using domain knowledge and examples. The domain knowledge consists of a *causal model C* of the domain, stating the relationships among basic phenomena, and a body of *phenomenological theory*, describing the links between abstract concepts and their possible manifestations in the world. The causal model provides explanations in terms of causal chains among events, originating from “first” causes.

The evolution of our modelling approach together with a critical discussion can be found in [Saitta, Neri and al., 1995; Neri, Saitta and Tiberghien, 1997a; Saitta, Neri and Tiberghien, 1997].

For sake of completeness, we note that WHY can automatically make changes from one model to the following by using automatic induction as in [Sleeman et al., 1990; Baffes & Mooney, 1996]. But, this is not the primary goal of our research. In fact, we want to help the cognitive scientist in developing what she believe to be the evolution of the student's model. The long term goal of our research, instead, may be the (fully) automatization of the learning process. In order to accomplish this long term goal, we discuss some temptative computational definitions of the different kinds of conceptual change that have been observed during the psychological experiments.

The last important point to be kept present while reading this work, is that we are going to study a model that capture the “functionality” of the student's understanding and answering to problems. Instead, we are not making any claim about the (unlikely) cognitive validity of the described model.

2 The Modelling Tool WHY

In this section a brief description of the functionalities of the learning system WHY is given. WHY learns and revises a knowledge base for classification problems using domain knowledge and examples. The domain knowledge consists of a *causal model C* of the domain, stating the relationships among basic phenomena², and a body of *phenomenological theory*, describing the links between abstract concepts and their possible manifestations in the world.

The causal model *C* provides explanations in terms of causal chains among events, originating from “first” causes. The phenomenological theory *P* contains the semantics of the vocabulary terms, structural information about the objects in the domain, ontologies, taxonomies, domain-independent background knowledge (such as symmetry, spatial and temporal relations); finally, *P* contains a set of rules aimed at describing the manifestations of abstractly defined concepts in terms of properties, objects and events in the specific domain of application.

The causal model *C* is represented as a directed, labelled graph. Three kinds of nodes occur in the graphs: *causal nodes*, corresponding to processes or states related by cause-effect relations, *constraint nodes*, attached to edges and representing conditions which must be verified in order to instantiate the corresponding cause-

² In general, *C* may contain any “deep” model of the domain, not necessarily a causal one.

effect relation, and *context* nodes, associated to causal nodes, representing contextual conditions to be added to the cause in order to obtain the effect.

The goal of WHY is to build up or revise a knowledge base **KB** of heuristic classification rules. A causal explanation (justification) of any **KB** revision is automatically provided.

It is important to clarify the relations between the causal model **C** and the heuristic knowledge base **KB**. The causal model could be used directly to obtain classifications. However, causal reasoning is slow, and the rules in **KB** act as shortcuts, compiled from **C**. On the other hand, the fact that the rules are justified by **C** (being derived from it according to the method described in [Saitta, Botta & Neri, 1993]) guarantees their validity and correctness (with respect to that of **C**) and also allows for explanations of the given classification in terms of the deep knowledge. On the other hand, **KB** and **C** may not be related at all, for instance in the case that **KB** is not derived from **C** but is directly “taught” by a teacher or acquired by the learner on a pure inductive basis. In this case, **KB** will give unjustified classifications (correct or not), for which no explanation exists with respect to **C**. Exploiting these different types of relations between **KB** and **C**, all the learning models emerged in the experimentation can be modelled. In the interplay between **KB** and **C**, the knowledge in **P** supplies the links between the general principles stated in **C** and the concrete experiments. The content of **P** contributes, as well, to enrich the modelling of students’ misconceptions and conceptual change. Actually, it is in **P**, for instance, that ontological shift occurs.

3 The Modelling Methodology

As WHY learns knowledge for classification tasks, and provide explanations thereof, the task considered in this paper has to be mapped accordingly. We have chosen to follow the individual evolution of three students over two years (6th and 5th grades), because our goal was not to verify general hypotheses about learning, but rather to show how conceptual change in an individual can actually be modelled with WHY. In this paper we will show how the model can be built up and used through a specific example: the knowledge evolution of the student “David” with respect to learning in the 6th grade.

Before outlining the methodology used, we will briefly describe the material available from the David’s history:

T_0 = Before teaching

Answers to Questionnaires Q_1 and Q_2 and their explanations. Answers to the initial Interview I_0 and their explanations. (An example of the type of questions is reported in Appendix A)

T_i ($1 \leq i \leq 11$) = During teaching

Answers to questions, predictions of outcomes from practical manipulations, and given explanations during the i -th teaching session.

T_f = After teaching

New answers to Questionnaires Q_1 and Q_2 and their explanations. Answers to the final Interview I_f and their explanations.

T_t = Post test

Answers to a test Interview I_t and their explanations, six months after the end of teaching.

In order to use WHY to model David, each practical experiment is represented as an example, consisting of two parts: a description of the experimental setting and a question. The experimental setting corresponds to the description of the example, whereas the possible answers to the question 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. An examples of this mapping is given in Appendix A.

In the current experimentation, the various knowledge bodies **C**, **P** 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 that can be triggered when one of WHY's explanations (answers) does not match the student's ones. This provides useful information to the educational/cognitive scientist in discovering where her hypothesized student's model is incorrect.

4 An Instance of Observed Conceptual Change

Due to space bound we cannot fully described our modelling methodology. So we only give an idea of the formalism used to represent the evolution of part of the knowledge of David (a student). Moreover, we also sketch how David's conceptual change can be pointed out by comparing its knowledge before and after teaching. In this section, we will compare two knowledge state of David represented as two WHY's causal models. A causal model consists in an oriented graph where causal relationships among the domain's abstract phenomena are represented.

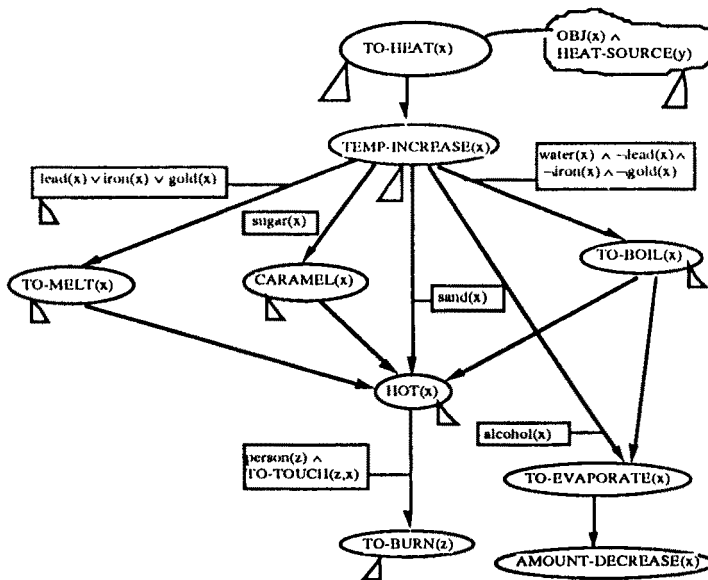


Figure 1 – Part of the causal model C_0 , hypothesized to represent David's knowledge before teaching. Elliptic nodes contains the domain relevant phenomena. Arrows represent causal relationships among them. Rectangle and clouds represent accessory conditions.

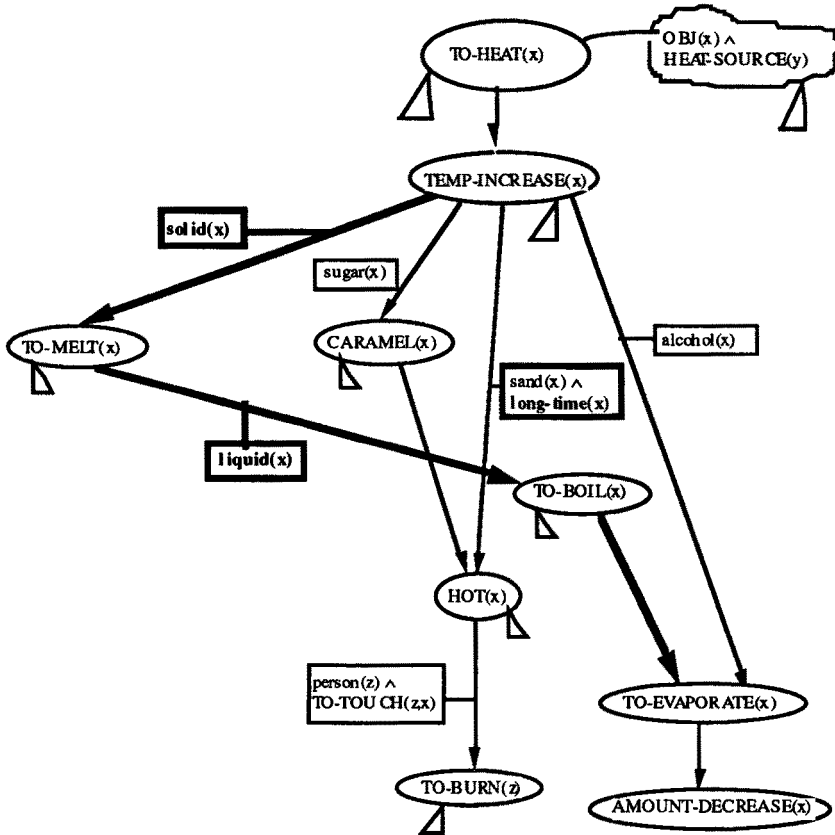


Figure 2 – Part of David’s hypothesized knowledge after teaching. Conceptual change is represented by the bold causal path.

David’s knowledge before teaching has been manually inferred on the basis of his initial answers to questionnaires and interviews. Part of David’s initial knowledge, represented as a WHY’s causal model, appears in Figure 1. By analysing the causal model (network), we may notice that David uses a notion of material causality linked to the “substance” of a body; in fact, what happens to the 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 shows evidence to believe that “boiling” and “melting” are alternative (and mutually exclusive) behaviours, exhibited by different substance. In fact, he say that iron, gold and lead shall not boil, *because they melt*.

By linking the behaviour of the bodies, with respect to heating, to specific substances, David is then unable to answer questions about materials that he does not about: for instance, he answers “I don’t know” to questions about the possibility of diamond, salt and aluminium to become liquid or gaseous.

After the teaching course, during which David has seen several other experiments, involving different materials, he is able to fill the final questionnaires and the final interview in such a way that we may infer that his deep knowledge of the world is changed, under various respect. In the following, we will just illustrate one of the changes, which can be considered as a **conceptual change**. In Figure 2, David’s hypothesized knowledge after the cycle of lesson is reported. Two changes can be noticed. The first one is that David is able to consider the importance of time in determining the final state of a material. At the beginning, in fact, he simply said that the sand would become hot, when heated. Now, he is able to understand that the effect of heating takes time to happens, suggesting the idea of a “process”. This finding is confirmed by the fact that now David says that, in order for the water to start boiling, “at least a quarter of an hour is necessary”.

However, the most important change, with respect to the goal of the teaching course, is that David seems less committed to a material causality for determining behaviours. In fact, he is able to generalize, from “iron”, “lead”, “gold” and “ice”, that any “solid” may become liquid is sufficiently heated. Moreover, “to boil” and “to melt” are no more mutually exclusive behaviours, but they are possibly in sequence, as it should be. David’s 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 actual transition from the two states of knowledge could be sketched as follows. During the lessons, David performs manipulation in which a number of different substances undergo a change of state: he observes ice transformed into water, water into ice and water into vapor, as well as lead and salt to melt. Then, he keeps adding to his causal network additional links, corresponding to the new observed substance. Observing commonalities (also helped by the teacher) he notices that all the substance that melt are solid and that vapors derive from liquids. Technically, the constraints on the edges of the causal net are generalized from specific substances to “solid” or “liquid”, respectively and, then, all these edges collapse into a single one.

5 Tentative Computational Definition of Conceptual Change

In order to make possible the future automatization of the learning process, we associate, as described in the following, each type of observed change to the application of one or more WHY’s revision operators.

Accretion

Learning by accretion *increases coverage*, in the sense that more experimental situations can be handled (independently of their correct interpretation/prediction). 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. Accretion, then, affects both the *phenomenological theory* or the *heuristic knowledge base*, by adding to any of them a new rule. Accretion may also consists of addition of a property to an

ontological node. In our approach, accretion is not considered conceptual change in a strict sense. Accretion can be implemented with the “adding rules” operator.

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.

Tuning

Tuning *increases the number of correct predictions or explanations*, but does not modify the deep explanatory framework. It may affect the *heuristic knowledge base* or the *phenomenological theory*, by changing the preconditions in some rule or by adding new rules. Tuning may also involve the *causal model*, but only with addition/deletion in the constraint and context nodes (causal nodes and links between them cannot be changed). Tuning can be implemented with the operators that generalize or specialize rule antecedents, or add and delete rules. Also tuning is not considered here a conceptual change in a strict sense.

Restructuration

Restructuration affects the explanatory framework. It involves the *causal model*, via addition/deletion of causal nodes, or modification of causal links, and the *ontologies* of the domain, via addition/deletion of nodes or changing a node from one ontology to another. Restructuration is considered conceptual change.

As it can be seen, we have given a stricter definition of conceptual change than in some approaches (for instance, [Vosniadou, 1994, 1995]), but larger than in others (for instance, [Chi et al., 1994]).

6 Conclusion

We discussed a way of interpreting learning in relation with teaching in the domain of physics. Our analysis is based on a type of knowledge processing which can be a relevant frame of reference from the points of view of both personal knowledge of a learner and "official" knowledge, such as that of Physics. Our framework allows us to establish an 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 and Physics knowledge to be compared. This independence and this comparison are essential to the characterisation of different types of learning.

Moreover, WHY's articulated knowledge representation allows most of phenomena observed in children learning elementary physics to be modelled, notably their explanation in terms of simple causality, and the interdependence of "surface" heuristic knowledge and deep beliefs.

Appendix A. Representing Questionnaires in WHY

In this appendix we will show, through an example, how questions are translated into a first order logic ground formula to become training instances for WHY. The questions in the interview are expressed in natural language and describe experiments, whose outcome the student is supposed to explain, guided by the

teacher. The experimental material is actually provided during the interview, and the student is allowed to freely manipulate it. The student is supposed to answer verbally.

Question: “Two saucepans A and B, with thermometers inside, are put on two gas stoves g_a and g_b , which are equal (denoted by the predicate “same-feature”). A contains a smaller amount (a) of water than B does (b). The initial temperature of the system is 20 °C. After a short time interval, the thermometer in A indicates 50 °C.

1) Will the thermometer h_b in B indicate the same, a greater or a lower temperature than that h_a in A?

2) Please, explain your answer”

To answer the question is modelled as a classification problem, in which there are three classes

{GREATER-THERMOM-READING(h_b, h_a, t_2),

SAME-THERMOM-READING(h_b, h_a, t_2),

LOWER-THERMOM-READING(h_b, h_a, t_2)},

among which David has to choose. Then, he is confronted with the actual outcome of the manipulation (In this case he answered correctly).

person(David) \wedge saucepan(A) \wedge saucepan(B) \wedge
 \wedge same-features(A,B) \wedge water(a) \wedge water(b) \wedge thermometer(h_a) \wedge thermometer(h_b) \wedge
 same-features(h_a, h_b) \wedge gas-stove(g_a) \wedge
 \wedge gas-stove(g_b) \wedge same-features(g_a, g_b) \wedge ignited(g_a) \wedge
 \wedge ignited(g_b) \wedge on(A, g_a) \wedge on(B, g_b) \wedge person(Tournesol) \wedge
 \wedge person(Tintin) \wedge person(Haddock)
 \wedge to-put-inside(Tournesol,a,A) \wedge amount(a,small) \wedge
 \wedge to-put-inside(Tournesol,b,B) \wedge to-put-inside(Tournesol, h_a ,A) \wedge
 \wedge to-put-inside(Tournesol, h_b ,B) \wedge amount(b,large) \wedge
 \wedge temp(a, 20,initial) \wedge temp(b, 20,initial) \wedge not-boiling(a,initial) \wedge
 \wedge not-boiling(b,initial) \wedge time-elapsed(a,short) \wedge
 \wedge time-elapsed(b,short) \wedge time-elapsed(g_a ,short) \wedge
 \wedge time-elapsed(g_b ,short) \wedge thermom-reading(h_a ,50,final)

References

- Baffes P. T. and Mooney R. J. (1996). “A Novel Application of Theory Refinement to Student Modelling”. *Proc. of Thirteenth National Conference on Artificial Intelligence* (Portland, OR), pp. 403-408.
- Caravita S. and Halldén O. (1994). “Re-framing the Problem of Conceptual Change”. *Learning and Instruction*, 4, 89-111.
- Chi M. T. H., Slotta J. D. and 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.
- Forbus K.D. and 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*, Morgan Kaufmann, Los Altos, CA, pp. 311-348.
- Giordana A. and Neri F., Saitta L. and Botta M. (1997). "Integrating Multiple Learning Strategies in First Order Logics". *Machine Learning*, 27, 209-240.
- Mitchell T., Keller R., Kedar-Cabelli S. (1986). "Explanation Based Generalization". *Machine Learning*, 1, 47-80.
- Murphy L.G. and Medin D.L. (1985). "The Role of Theories in Conceptual Coherence". *Psychological Review*, 92, 289-316.
- Neri F., Saitta L. and Tiberghien A. (1997a). "Modelling Physical Knowledge Acquisition in Children with Machine Learning". Proc. of *19th Annual Conference of the Cognitive Science Society*, Stanford (CA), Morgan Kaufmann, pp. 566-571.
- Newell A. (1990). *Unified Theories of Cognition*, Harvard University Press, Cambridge, MA.
- Rumelhart D. E. and Norman D. A. (1977). "Accretion, Tuning and Restructuring: Three modes of Learning", in Cotton J. W. and Klatzky R. L. (Eds.), *Semantic Factors in Cognition*, Erlbaum (Hillsdale, NJ).
- Sage S. and Langley P. (1983). "Modeling Cognitive Development on the Balance Scale Task". Proc. *8th Int. Joint Conf. on Artificial Intelligence* (Karlsruhe, Germany), pp. 94-96.
- Saitta L., Botta M., Neri F. (1993). "Multistrategy Learning and Theory Revision". *Machine Learning*, 11, 153-172.
- Saitta L., Neri F. and al. (1995). "Knowledge Representation Changes in Humans and Machines". In P. Reimann and H. Spada (Eds.), *Learning in Humans and Machines: Towards an Interdisciplinary Learning Science*, Elsevier (Oxford), pp. 109-128.
- Saitta L., Neri F. and Tiberghien A. (1997). "World Model Construction in Children during Physics Learning". Proc. of *International Symposium on Methodologies for Intelligent Systems '97 (ISMIS 97)*, Lecture Notes in Artificial Intelligence series, Springer Verlag (Berlin, Germany), in press.
- Schmidt W.C. and Ling C.X. (1996). "A Decision-Tree Model of Balance Scale Development". *Machine Learning*, 24, 203-230.
- Shultz T.R., Mareschal D. and Schmidt W. (1994). "Modeling Cognitive Development on Balance Scale Phenomena". *Machine Learning*, 16, 57-86.
- Sleeman D., Hirsh H., Ellery I. and Kim I. (1990). "Extending Domain Theories: two case Studies in Student Modeling". *Machine Learning*, 5, 11-37.
- 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 and H.F. Friedrich (Eds.), *Learning and Instruction. European Research in an International Context, Volume 2.1*, Pergamon Press, Oxford, UK, pp. 631-648.
- Tiberghien A. (1994). "Modelling as a Basis for Analysing Teaching-Learning Situations". *Learning and Instruction*, 4, 71-87.
- Vosniadou S. (1994). "Capturing and Modeling the Process of Conceptual Change". *Learning and Instruction*, 4, 45-69.
- Vosniadou S. and Brewer W.F. (1994). "Mental Models of the Day/Night Cycle". *Cognitive Science*, 18, 123-183.
- White R. T. (1994). "Commentary Conceptual and Conceptional Change". *Learning and Instruction*, 4, 117-121.