

Modelling Physics Knowledge Acquisition in Children with Machine Learning

Filippo Neri

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

Lorenza Saitta

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

Andrée Tiberghien

Equipe COAST de l'UMR GRIC
CNRS - Université Lyon 2
Ecole Normale Supérieure de Lyon
46, Allée d'Italie
69364 LYON Cedex 07 (France)
atibergh@cri.ens-lyon.fr

Abstract

A computational approach to the simulation of cognitive modelling of children learning elementary physics is presented. Goal of the simulation is to support the cognitive scientist's investigation of learning in humans. The Machine Learning system WHY, able to handle domain knowledge (including a causal model of the domain), has been chosen as tool for the simulation of the cognitive development. In this paper the focus will be on knowledge representation schemes, useful to support further modelling of conceptual change.

Introduction

Conceptual Change is a well known phenomenon in developmental psychology and educational science [Carey, 1983; Tiberghien, 1989, 1994; Vosniadou & Brewer, 1992, 1994; Smith et al., 1992; Caravita & Halldén, 1994; Chi et al., 1994; Vosniadou, 1994, 1995; Slotta & Chi, 1996]. Even though a quite large body of experimental findings has been collected over the years, still no single definition of conceptual change is universally accepted. In fact, a sound explication of this notion would presuppose a precise definition of "concepts", a plausible hypothesis about their internal representation, and an (at least approximate) understanding of human learning mechanisms. In addition, the strong interpersonal, intercultural and interdomain differentiations, emerged in all the above aspects, suggest a multifaced and complex network of underlying interrelated phenomena, difficult to capture into a general theory.

To model human learning, Machine Learning (ML) methods and systems are natural candidates to provide computational modeling tools. In recent years, they have been used so far in two contexts: either building student models in a ITS environment [Sleeman et al., 1990; Baffes & Mooney, 1996], or describing knowledge acquisition and evolution [Klahr & Siegler, 1978; Sage & Langley, 1983; Hardiman et al., 1984; Anderson et al., 1990; Shultz et al., 1994; Schmidt & Ling, 1996]. Works in the first group try to build up a picture of what a student knows on a specific subject at a given moment, whereas works in the second group take explicitly into account

conceptual change and/or human learning mechanisms. Further works that are of direct relevance concern Qualitative Physics [Forbus & Gentner, 1986].

Most models of human learning presented so far in the ML literature are based on excessively simplifying assumptions. Basically, learning is reduced to a simple classification task, performed on the basis of knowledge consisting in a set of rules or a neural net. In fact, in the first place, the heuristic knowledge that a person possesses in a specific domain (substantially the one modelled in the ML systems) is neither acquired nor used in isolation, but it is embedded into, and biased by, a pre-existing deeper knowledge structure [Murphy & Medin, 1985; Tiberghien, 1994; Vosniadou, 1994], which gives it its meaning. Several education and cognitive scientists have clearly pointed out that *misconceptions* and errors can be traced back to conflicts between taught concepts and this deeper layer. The deep knowledge, in fact, plays a substantial role in learning, specifically in the understanding of concepts. This layer is not necessarily modified by acquiring skill in solving problems. For this reason, we keep, in our model, the two layers separate, i.e., the heuristic knowledge and the explanatory framework; this last, ignored in most ML models of human learning, is the basis for supplying *explanations* of phenomena.

In this paper a new approach to model human conceptual change is presented. The modeling tool is the ML system WHY, which acquires and revises a First Order Logic theory by exploiting a causal model of the domain and a set of examples [Saitta et al., 1993; Baroglio et al., 1994]. The overall goal of the research is to model conceptual change occurring in young students, acquiring basic concept in Physics, specifically *Heat* and *Temperature* concepts [Tiberghien, 1989, 1994]. In this paper we will concentrate on the knowledge representation schemes, as details on modelling changes in the knowledge can be found elsewhere [Neri et al., 1996]. One of the main novelties, with respect to previous models, is the differentiation between the knowledge a student uses to answer questions and to interpret experimental results, and an *explanatory framework*, based on the notion of simple linear *causality*. The computational model is grounded on an

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

epistemological framework and previous experiments by Tiberghien [1989, 1994].

We would like to underline that in no way we advocate the presented logical knowledge representation scheme as being the actual one used by humans. The representation is only intended to be a description tool both understandable by the experimenters and implementable as a program.

With particular reference to learn Physics, an Epistemology of Physics is proposed by diSessa [1993]. The main claim is that humans gradually acquire knowledge elements, called *phenomenological primitives* (p-prims), which constitute a large and complex knowledge system, and for which diSessa advocates a connectionist representation.

Vosniadou and co-workers (see, for instance, [Vosniadou, 1994, 1995; Vosniadou & Brewer, 1992, 1994]) present a theoretical framework that hypothesizes that acquisition of knowledge about the physical world is biased by a set of fundamental constraints, called *entrenched presuppositions*. An attempt to build a computational model of the day/night cycle has been done in [Morik & Vosniadou, 1995].

An interesting theoretical framework for interpreting Physics learning is Forbus and Gentner's Qualitative Process and Structural Mapping theory [Forbus & Gentner, 1986]. The main claim of the proposal is about the centrality of the notion of *process*, as a fundamental representation construct.

In a number of papers, Chi and co-workers have tried to explain why certain conceptual changes in Physics are so difficult to induce whereas others are not (see, for instance, [Chi et al., 1994; Slotta & Chi, 1996]). They hypothesize that the learner has a cognitive model based on disjoint (ontological) category trees. Chi brings evidence that restructuring the interior of an ontology tree is more easily accomplished than moving a concept across trees.

Finally, in [Tiberghien, 1994] a theoretical framework for interpreting such difficulties has been proposed. The framework has its foundation both in pedagogical studies and in the epistemology of science and claims that, in experimental sciences, questions are strongly linked to three main factors: the theoretical background, the experimental facts considered, and the explanations produced. In the proposed theoretical framework, interpretation and prediction in Physics imply a modeling process articulated on three levels: "theory", "model" and "experimental field" of reference.

The Learning Context

The specific learning context considered in this paper is the following: in three classes of the first and second year of secondary school (12-13 year old students, 6-5th grades), Physics teaching took place under controlled conditions, in the sense that the teachers agreed to propose the same teaching materials, experimentations and questions. Content of the course were basic concepts and qualitative relations in the domain of *heat* and *temperature*, including *change of state* and *heat transfer* in everyday life situations.

Individual interviews before and after the set of teaching sessions and an experimental task have been performed with two students of each class. Written questionnaires, before and after teaching, have been filled by each student.

In this paper we want to describe how to model the process of learning in individual students. A basic assumption is that explanation of the observed phenomena consists in finding causes and causal chains. Causality has been acknowledged before as playing a crucial role in naïve Physics learning. [White & Frederiksen, 1987; Rozier, 1988]. Taking into account the age of the learner (12-13 years), Aristotelian causalities are used as reference. In particular, *material* causality (used when students, for instance, consider that wool heats "because it is wool") and *efficient* causality (involved when there is a change, for example, when a battery lights a bulb) are considered here.

The Learning System WHY

WHY is a system that learns and revises a knowledge base for classification problems using domain knowledge and examples [Saitta et al., 1993; Baroglio et al., 1994]. The domain knowledge consists of a *causal model* of the domain, and a body of *phenomenological theory*, describing the links between abstract concepts and their possible manifestations in the world. A complex inference engine, combining induction, deduction, abduction and prediction, is the core of the system.

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, and, more importantly, 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. All the knowledge structures share, in WHY, a First Order Logic based language, whose atomic predicates are partitioned into *operational* and *nonoperational*. Operational predicates are observable, whereas nonoperational predicates are only deducible.

The causal model C is represented as a directed, labelled graph, as the one reported in Figure 2. The phenomenological theory P is represented as a set of Horn clauses, and the examples are represented as ground logical formulas.

The goal of WHY is to acquire a knowledge base KB of heuristic rules, sufficient to solve a set of problems in the chosen domain. Moreover, the system gives causal explanations of its decisions. 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 answers/solutions to questions, as it is done in diagnostic systems working from first principles [Reiter, 1984]. 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 et al., 1993]) guarantees their validity and correctness (obviously with respect to that of C) and also allows 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 empirical inductive basis. In this case, KB will give unjustified classifications (correct or not), for which no explanations exist with respect to C. Exploiting these different types of relations between KB and C, all the findings emerged in the experimentation with children learning Physics can be modeled. 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 heuristic knowledge base KB corresponds to the compiled rules David normally uses to answer questions and solve problems, when no explanation of his answers is required; to give explanations, David shall exploit the causal model. The heuristic rules, then, can be considered as the predictors of the answers, when David is questioned about possible outcomes of experiments. The causal model C is used by David when an explanation is requested or when he does not have yet a heuristic rule to answer a question.

An Example of Model Construction

In this section we will go through an example of using WHY to model the knowledge of "David", a 12 year old student of 6th grade, exposed to the teaching course on heat and temperature mentioned in Section 3. At this grade, the specific content of teaching was mostly directed toward the notion of change of state in Physics.

The data available from David's history, used to build up the model, are the answers to two questionnaires and an interview both before and after teaching. Moreover, the answers to questions, the predictions of outcomes from practical manipulations, and the given explanations during each teaching session are available as well.

In order to use WHY to hypothesize David's mental models, the task of answering a question or predicting an experiment's outcome has to be mapped on a discrimination task, whose possible answers have been a-priori individuated by the teacher. Moreover, each experiment or question is represented as an example, consisting of two parts: a description of the experimental setting, from the point of view of the teacher, and a question. The correct answer labels the example. According to the teaching protocol, it is assumed that the students understand the example descriptions. For the sake of exemplification, let us consider a question, occurring in a questionnaire, reported in Figure 1.

The first step in the modelling process is to set up the vocabulary used in teaching. By analysing the whole questionnaires and interviews, all the words relevant to the specific Physics domain have been extracted and transformed into atomic predicates of the language. Some of

the predicates derived from the question in Figure 1 are reported:

amount(x,u,t), different(u,v), gas-stove(x),
inside(x,y), not-boiling(x,t), on(x,y),
person(z), same-features(x,y), water(x), temp(x,T), ...

The complete vocabulary contains 95 words. For what concerns the semantics of the predicates, 69 of them are operational, i.e., their evaluation can be made directly on the experimental setting. For instance, person(z) and temp(x,T) are operational. Other predicates (precisely 26) are non-operational, i.e. their truth value can be determined by deduction. For instance, the rule

$$\text{gas-stove}(x) \wedge \text{ignited}(x) \Rightarrow \text{FLAME}(x)$$

states that the predicate FLAME(x) can be asserted true on x if x is an ignited gas-stove.

Professor Tournesol makes the following experiment: He takes two saucepans A and B, pours water from a faucet into them, and he also puts a thermometer inside each of them.

The saucepans A and B are equal.

The thermometers are equal.

The two flames are equal and the saucepans are put on the gas stoves at the same time.

The quantities of water in A and B are different.

After 3 minutes, the water in A and B does not boil yet. Tintin reads the indication on the thermometers inside A: it shows 50°C.

(1) Does the thermometer in B show a reading:

- Greater than 50°C
- Equal to 50°C
- Less than 50°C

(2) Why ?

Figure 1 - Example of questions occurring in the questionnaires.

Notice that for non-operational predicates, the semantics of the same term for the teacher and for David can be different. For instance, the two following rules:

$$\text{TEMP}(x,T) \wedge \text{greater}(T,\theta_1) \Rightarrow \text{HOT}(x) \quad (\text{Teacher})$$

$$\text{feel-hot}(x) \Rightarrow \text{HOT}(x) \quad (\text{David})$$

show that the teacher evaluates the hotness of an object x according to its temperature, whereas David relies on his tactile perception.

The second step consists in transforming all the questions in examples for WHY. For instance, the question in Figure 1 is described as follows:

Example # 2 : Description

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

thermom-reading(h_a , 50, final)

Example # 2: Decisions

{GREATER-THERMOM-READING(h_b , h_a , final),
 SAME-THERMOM-READING(h_b , h_a , final),
 LOWER-THERMOM-READING(h_b , h_a , final)}

In the description of the example we may notice that the quantities have been rendered in qualitative form, such as “small” and “large” amounts of water, a “short” time period (for 3 minutes), “initial” and “final” for the beginning and ending times of the experimentation. Moreover, some background information, which David and the teacher do not need to say explicitly, are added for the system’s sake, such as the fact that the room temperature is 20°C. The predicate “same-feature(x , y)” denotes functional equality between x and y without object identity.

In the decision part of the example the “classes” are defined according to the alternative possible answers.

After building up the dictionary and describing the examples, we have encoded, for reference, the teacher’s phenomenological and causal theories. These bodies of knowledge are not meant to describe all the knowledge the teacher has in the field, but only the part that he/she decides is relevant for teaching. Some rules belonging to the teacher’s phenomenological theory P^* are the following (the complete theory contains 121 rules):

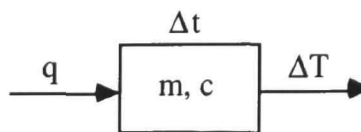
- aluminum(x) \Rightarrow METAL(x)
- METAL(x) \Rightarrow MATERIAL(x)
- TEMP(x , T) \wedge greater(T , θ_1) \Rightarrow HOT(x)
- gas-stove(x) \wedge ignited(x) \Rightarrow FLAME(x)
- electric-plate(x) \wedge turned-on(x) \Rightarrow HEAT-SOURCE(x)
- OBJ(x) \wedge HEAT-SOURCE(y) \wedge CONTACT(x , y) \Rightarrow
 \Rightarrow TO-HEAT(x , y)
- SAME-TEMP(x , y) \Rightarrow THERMAL-EQUILIBRIUM(x , y)
- full-of(x , y) \Rightarrow INSIDE(y , x)
- CONTACT(x , y) \Leftrightarrow CONTACT(y , x)

A part of David’s causal model is reported in Figure 2. The graph explains that the temperature of an object increases if it is heated and its initial temperature is below its boiling threshold. As we may notice, the model may be criticised under many respects with respect to a complete theory of heat transfer. However, it represents what the teacher wants David to understand in this preliminary course.

The *Matter* ontology of the teacher is reported in Figure 3. We notice that the teacher knows that an object has a temperature value associated to it, and that a material has characteristic temperatures associated to changes of state.

The teacher also uses a heuristic knowledge base KB^* , containing 23 rules, one of which is the following:
 MATERIAL(x) \wedge SOLID(x) \wedge TO-HEAT(x) \Rightarrow
 \Rightarrow MELTING(x)

Moreover, the teacher knows the relations among all the quantities appearing in the following heating process:



The teacher’s knowledge remains the same along the whole teaching course. It is useful to model it both as a reference for representing the teaching goal, and as a mean to evaluate David’s progress toward a correct understanding of the phenomena to which he is exposed.

The teacher’s knowledge can be considered “correct”, from his/her point of view; in fact, the teacher him/herself can validate it. On the contrary, David is unable to articulate his models of the world, and his knowledge can then be only guessed from his answer and explanations. For this reason, it needs to be validated experimentally. From the data available before teaching, an initial content can be attributed to David’s C_0 , P_0 and KB_0 .

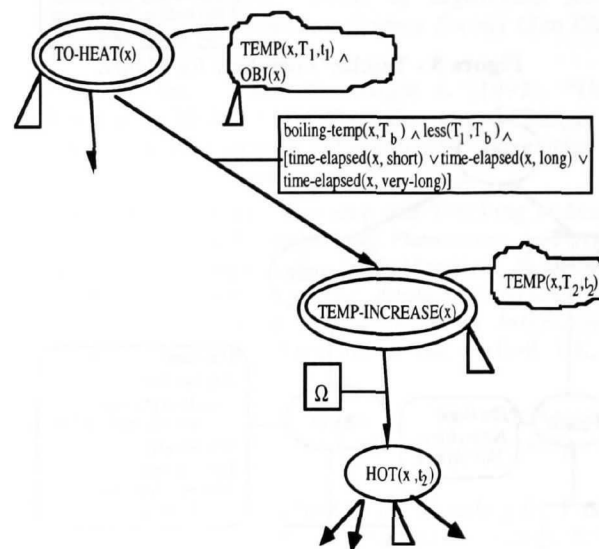


Figure 2 - Part of the teacher’s causal model. Three kinds of nodes occur in the graphs: *causal* nodes, corresponding to processes or states related by cause-effect relations, *constraint* nodes, attached to edges and representing physical or structural properties of objects, and *context* nodes, associated to causal nodes, representing contextual conditions (concomitant causes) referring to the environment.

The phenomenological theory P_0 contains 94 rules, and is not very different from the teacher’s one, except for the semantics of some predicates, based on direct perception instead of an objective measurement. For the sake of comparison, in Figure 4 David’s *Matter* ontology tree is reported. As we can see, David attributes to the objects the properties of being cold, warm or hot, but he does not relate them, at the beginning, with the object’s temperature.

Moreover, David attributes the constance of temperature during a change of state to a maximum allowed temperature for the substance.

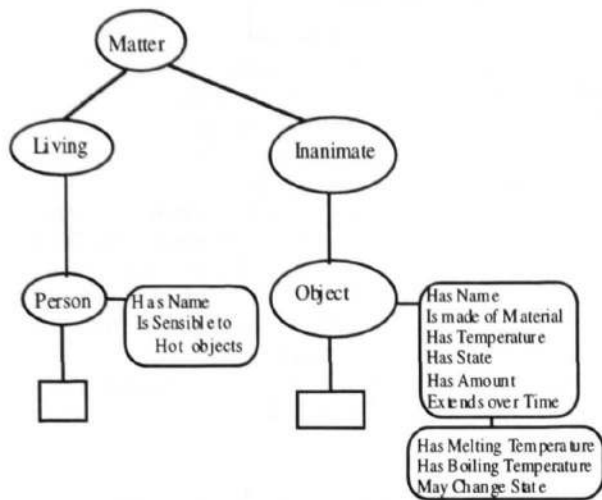


Figure 3 - Teacher's ontology for *Matter*.

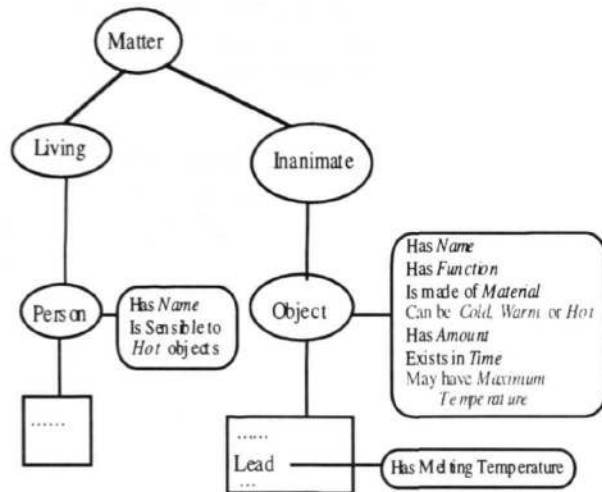


Figure 4 - David's ontology tree for *Matter*

In Figure 5, David's causal model C_0 is reported. The model is substantially different from the teacher's one, because it is mostly oriented to deducing effects on the basis of the properties of the involved substances. Then, material causation is underlying David's explanations.

Also the heuristic knowledge base KB_0 is rather different, because David, in order to answer the questions, seems to apply rules that can be paraphrased as follows:

- “What is hot heats”
- “What is cold cools”
- “A greater cause has a greater effect”

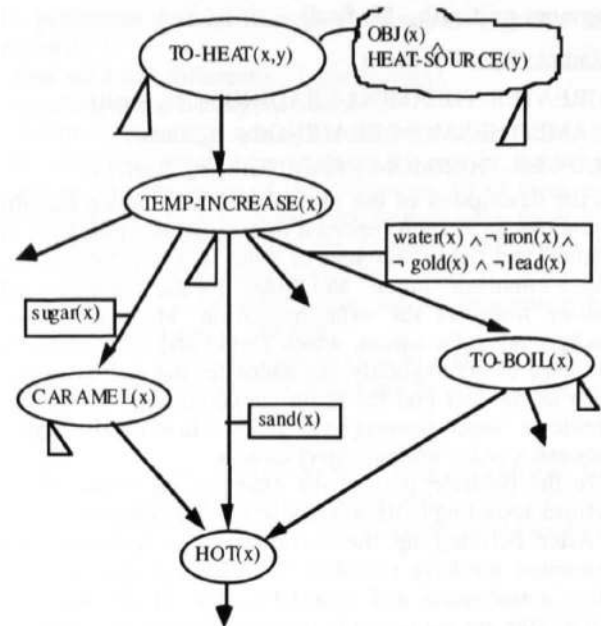


Figure 5 – David's initial causal model.

For what concerns validation of the models, we have to specify, first of all, that our aim is not that of making claims about average or most common behaviours among childrens that age, but to model the *individual* evolution of a single student. That means that we do not need to apply the same model to a statistically significant set of students, but, rather we have to apply the same *modelling process* (possibly yealding different models) to various students. In fact, with respect to this modelling procedure, a model is “good” if it is able to predict the answers of the modelled student to the questions he/she is presented with. Then, if this match is verified for a number of cases, there is suggested evidence that the hypotheses underlying the modelling methodology may be adequate for tracing individual knowledge evolution, in this specific Physics' area. In the analysed cases, satisfying results have been obtained.

Conclusion

We have introduced a new way of interpreting learning from the point of view of the learner's knowledge acquisition in relation with teaching, in the domain of Physics. The framework of our analysis is a specific type of knowledge processing that we term “modelling”. It is a relevant framework for analysis with respect to both the learner's personal knowledge of a learner and the content of Physics teaching. This approach allows two main types of knowledge to be distinguished, in particular, the pragmatic knowledge (heuristic), put into play during problem solving, and a deeper explanatory knowledge structure (causal), used to give explanations and, in general, to get a better grasp of the phenomena.

The eventual goal of the investigation is to come up with a method for designing more effective teaching strategies.

References

- Anderson J.R., Boyle C.F., Corbett A.T. and Lewis M.W. (1990). "Cognitive Modelling and Intelligent Tutoring", *Artificial Intelligence*, 42, 7-49.
- Baffes P. T. and Mooney R. J. (1996). "A Novel Application of Theory Refinement to Student Modeling". *Proc. of Thirteenth National Conference on Artificial Intelligence* (Portland, OR), pp. 403-408.
- 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*, Morgan Kaufmann, Los Altos, CA, pp. 319-348.
- Caravita S. and Halldén O. (1994). "Re-framing the Problem of Conceptual Change". *Learning and Instruction*, 4, 89-111.
- Carey S. (1983). *Conceptual Change in Childhood*. MIT Press, Cambridge, MA.
- 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*
- 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.
- Hardiman J., Well P. T. and Pollatsek A. D. (1984). "The Usefulness of the Balance Model in Understanding the Mean". *Journal of Educational Psychology*, 76, 792-801.
- Klahr D. and Siegler R.S. (1978). "The Representation of Children's Knowledge". In H.W. Reese and L.P. Lipsitt (Eds.), *Advances in Child Development and Behavior*, Academic Press, New York, NY, pp. 61-116.
- Morik K. and Vosniadou S. (1995). "A Developmental Case Study on Sequential Learning: the Day-Night Cycle". In Reimann P. and Spada H. *Learning in Humans and Machines: Towards an Interdisciplinary Learning Science*, pp. 212-227.
- Murphy L.G. and Medin D.L. (1985). "The Role of Theories in Conceptual Coherence". *Psychological Review*, 92, 289-316.
- Neri F., Saitta L., Tiberghien A., Zucker J-D. (1996). "Modelling Conceptual Change with a Machine Learning System". Technical Report, Univ. Torino, Dipt. Informatica.
- Rozier S. (1988). *Le Raisonnement Linéaire Causal en Thermodynamique Élémentaire*, Thèse de Physique. Université Paris 7, Paris, France.
- 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.
- 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.
- Slotta J. D. and Chi M. T. H. (1996). "Understanding Constraint-Based Processes: a Precursor to Conceptual Change in Physics". *Proc. of Eighteenth Annual Conference of the Cognitive Science Society* (San Diego, CA), pp. 306-311.
- Smith C., Snir J. and Grosslight L (1992). "Using Conceptual Models to Facilitate Conceptual Change: The Case of Weight-Density Differentiation". *Cognition and Instruction*, 9, 221-283.
- 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. (1995). "A Cognitive Psychological Approach to Learning". In P. Reimann and H. Spada (Eds.), *Learning in Humans and Machines: Towards an Interdisciplinary Learning Science*, Elsevier Publ. Co., Oxford, UK, pp. 23-36.
- Vosniadou S. and Brewer W.F. (1992). "Mental Models of the Earth: A Study of Conceptual Change in Childhood". *Cognitive Psychology*, 24, 535-585.
- Vosniadou S. and Brewer W.F. (1994). "Mental Models of the Day/Night Cycle". *Cognitive Science*, 18, 123-183.
- White B. Y. and Frederiksen J. R. (1987). "Causal Model Progressions as a Foundation for Intelligent Learning Environments". *Techn. Report No. 6686*, BBN Laboratories, Cambridge, MA.