Modelling Conceptual Change: An Interdisciplinary Approach

Filippo Neri*, Lorenza Saitta*, Andrée Tiberghien+

* Università di Torino
  Dipartimento di Informatica
  Corso Svizzera 185
  10149 Torino (Italy)
  {saitta, neri}@di.unito.it

+ 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

Content Areas: cognitive modeling, machine learning, causality

Abstract

A computational approach to the simulation of cognitive models of conceptual change in children learning elementary physics is presented. The student’s mental model is inferred from a sequence of interviews collected along a period of eleven teaching sessions. Goal of the simulation is to support the cognitive scientist's investigation of learning in humans. 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. A Machine Learning (ML) system, able to handle domain knowledge (including a causal model of the domain), has been chosen as tool for the simulation of the cognitive models evolution. The system performs knowledge revision and provides causal explanation for its conclusions.

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, 1992, 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].

In order to make a step ahead toward a deeper understanding of conceptual change, one possibility is to build models, which offer, at the very least, a mean to obtain predictions from tentative hypotheses. For what concerns conceptual change, models 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.

1 This work has been performed within the project “Learning in Humans and Machines”, supported by the European Science Foundation.
Rumelhart and Norman [1977] have categorized the type of transitions occurring as Accretion, Tuning or Restructuration, which are reminiscent of Piaget's Assimilation, Accomodation, 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 restructuration may occur. However, when a contradiction emerges, also failures in learning may happen, taking the form of inert knowledge or misconceptions.

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

Starting from a descriptive model, a first step toward building a computational one consists in adding an operational definition of the mechanisms involved in conceptual change. One possibility is to simulate these mechanisms with a model sufficiently precise to be run as a program on a computer. The idea is to let the model (program) run in a set of situations comparable to a specific experimental setting, and to test the predictions from the model with the actual behaviour of human learners. In the Machine Learning literature, two approaches have been followed in modeling human learning: in the first one, a static snapshot of what a student knows at a given instant is inferred from his/her answers to a set of problems [Sleeman et al., 1990; Baffes & Mooney, 1996]. In the second approach the learning process is modeled as a sequence of knowledge states [Anderson et al., 1990; Klahr & Siegler, 1978; Sage & Langley, 1983; Newell, 1990; Schmidt & Ling, 1996; Shultz et al., 1994; Shultz et al., 1995]. An attempt to build up a computational model of the day/night cycle has been done in [Morik & Vosniadou, 1995], using the Machine Learning system MOBAL [Morik et al., 1993].

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]. 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, and suggest that education should be mainly directed toward this deeper level and not primarily toward the domain-specific one [Hestenes, 1987; Vosniadou & Brewer, 1992, 1994; White, 1994; Caravita & Halldén, 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.

In this paper, a new approach to modeling human conceptual change, which is intended to extend current modeling practice and to overcome the said limitations, is presented. One of the main novelties is 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, 1996; Tiberghien, 1994]), and from the direct