

A Probabilistic Model for Student Knowledge Diagnosis in Learning Environments

Jaime GÁLVEZ ^{a,1}

^a*Dpto. de Lenguajes y Ciencias de la Computación, University of Málaga, Spain*
e-mail: jgalvez@lcc.uma.es

1. Introduction

Nowadays, the way learners acquire and improve their knowledge has changed compared to several years ago. The use of new technologies, in conjunction with Artificial Intelligence (AI) and psychological research in the field of education, has made it possible to obtain a new generation of learning environments, which are more effective than traditional methods[1]. These intelligent tutoring environments are computer assisted tools capable of both representing knowledge and directing a learning strategy. Thus they are able to mimic the behavior of an expert, teaching the students how to apply knowledge and estimating their knowledge from how they interact with the system. As a result, learners can be redirected to the most appropriate action. To achieve these goals, the maintenance and use of a model that reflects the student knowledge is vital. Such a model allows the student weaknesses and strengths to be ascertained and a pedagogical strategy for appropriate instruction to be developed.

We have previous experience in the development of educational tools [2]. Nevertheless, this system is focused on assessing declarative knowledge using tests. We want to extend our previous work to cover both declarative and procedural knowledge assessment. For this reason, the main goal of our current research is to study existing approaches for building student models in problem solving environments, and to develop a novel methodology or framework to facilitate the elicitation of diagnosis modules for such environments. In this line, we have explored existing student modeling techniques and have tried to identify how we could improve them to create a versatile student model (in the context of the application domain), with accurate estimations of the student knowledge.

2. Research lines

Our initial and current line of research involves the Constraint Based Modeling approach (CBM) [3], whose effectiveness has been proven with the wide range of tutors built

¹Corresponding Author. E-mail address: jgalvez@lcc.uma.es

[4]. It is based on Ohlsson's theory of learning from performance errors [5]. According to it, a CBM-based tutor teaches the domain by detecting the misconceptions students have when trying to apply their knowledge to problem solving, and feedback is provided on any faulty knowledge. The main element of this approach is the "constraint", which represents a principle in the domain that cannot be violated by any solution for a given problem. The short-term student model is the set of constraints the student has violated during the resolution of problems. Using this set, the system can determine the main student weaknesses and adapt the next problem according to each individual's needs.

Unlike other CBM-based tutors, we do not focus on the tutorial use of CBM. Rather, we use it for diagnosis purposes. Accordingly, we have developed a diagnosis model which combines a CBM-based model with the Item Response Theory (IRT) [6]. The goal of our proposal is to extend the model of violated constraints with the information needed to supply well-founded and long-term student knowledge diagnoses. We aim to improve the heuristics commonly used in CBM tutors for this purpose, such as the proportion of constraints the student knows [7], by an accurate and invariant technique. The inclusion of IRT in our model may also lead to adaptive and CBM-based diagnosis systems, where the most appropriate problem to be solved by students at each step is selected dynamically. Additionally, the length of a problem solving session could be decided dynamically depending on the student level and the required accuracy; the problem difficulty could be inferred using data-driven methods.

We have already developed two CBM-based diagnosis tools: the OOPS system, for the fundamentals of object oriented programming; and another for the lineal optimization problem domain. The elaboration of both tools is part of our bottom-up strategy where first, we are going to build diagnosis tools for several domains, and then, common characteristics and patterns will be extracted to define a framework for building such systems. Our initial experiments with real students suggest that our proposal is viable and the diagnosis results are valid.

3. The Future Work

As future work we will continue to extend our model with some of the features described above and will also explore how our model could be used in learning systems.

References

- [1] Bloom, B., The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring, *Educational Researcher*, vol. 13, pp. 4–16, 1984.
- [2] Guzmán, E., Conejo, R., and J.L. Pérez-de-la-Cruz, Adaptive Testing for Hierarchical Student Models, *User Modeling and User-Adapted Interaction*, vol. 17, pp. 119–157, 2007.
- [3] Ohlsson, S. and Mitrovic, A., Constraint-based knowledge representation for individualized instruction, *Computer Science and Information Systems*, vol. 3, pp. 1–22, 2006
- [4] Mitrovic, A.; Martin, B. and Suraweera, P., Intelligent Tutors for All: The Constraint-Based Approach IEEE Intelligent Systems, *IEEE Educational Activities Department*, vol. 22, pp. 38–45, 2007
- [5] Ohlsson, S., Constraint-based student modeling, *Student Modelling: the Key to Individualized Knowledge-based Instruction*, pp. 167–189, 1994.
- [6] Thurstone, L. L., A method of scaling psychological and educational tests, *Journal of Educational Psychology*, vol. 16, pp. 433–451, 1925.
- [7] Mitrovic, A. and Martin, B., Evaluating Adaptive Problem Selection, *AH 2004*, pp. 185–194, 2004.