What is More Important for Student Modeling: Domain Structure or Response Times?

Jiří Řihák and Radek Pelánek

Masaryk University Brno

Abstract. Many features can be considered when designing a student model for an adaptive educational system. What is the relative importance of different modeling aspects? Where should we focus our attention in developing models for real word applications? We report comparison of two aspects: the choice of a domain model and the utilization of response times. The case study (an adaptive system for practice of basic arithmetic) suggests that response times deserve more attention in student modeling.

1 Introduction

A student model is a key part of an adaptive educational system. There is wide range of student modeling approaches and many features which can be included. In this work we compared impacts of two selected aspects of student modeling. The first one is the modeling of the domain structure [1], i.e., definition of skills and a mapping between skills and items. The second one is the utilization of response times [2], which is an additional information to the correctness of answers. As a case study we use a real adaptive educational system in its early stage of application, where the choice of a student modeling approach is a real, pressing development issue. We explore a range of domain models and response time uses, discuss their relations and comparison, and study parameter stability.

2 Setting

To explore the issue of model selection we utilize data from an adaptive practice application MatMat (`matmat.cz`), which covers the area of basic arithmetic. The system is available freely online and its behaviour and default student model are described in [4]. The currently available data comprise 150000 answers to 2000 items, which are divided into 5 high level concepts (counting, addition, subtraction, multiplication, division).

The used models are extensions of the Elo rating system [3], which can be seen as a heuristic for parameter estimation of the Rasch model. For comparison we used three domain models: the model with a single global skill; the model with skill parameters for each of the 5 main concepts; and the most complex model where skills are described in tree-like structure [4]. To incorporate response times we combine them with correctness of answer into a single performance measure r.

For correct answers we transform the value 1 into an interval $[0, 1]$ by one of the following approaches: no use of time: $r = 1$; the discrete decrease: $r = 1$ for fast responses ($< 7s$), $r = 0.5$ for slow responses; the exponential decrease: $r = 1$ for fast responses, $r = e^{1-t/7}$ otherwise [4]; and the linear decrease: $r = max(0, 1 - t/14)$.

3 Results

For comparison of predictive accuracy of models we use RMSE and AUC with student stratified train/test set division. With respect to domain modeling, the results show that more complex models are able to improve predictions, although increasing complexity of models brings only diminishing improvements. The evaluation of models which consider timing information is more difficult because different models are trained to predict different absolute values. Thus only AUC (which consider only relative order of predictions) seems to be meaningful and according to this metric the best results are achieved using the linear decrease.

To get insight into differences between models we analyze correlations between item difficulty parameters, which have clear interpretation. Unsurprisingly, we found a large gap between the baseline model and other more sophisticated models. The impact of domain modeling is nontrivial, but not pronounced. Different utilization of time, however, brings considerably different parameters. The degree of change is proportional to the intensity of time utilization. Results also suggest that domain modeling and time modeling are almost independent aspects and provide change (and possible improvement) in different directions.

We also studied how many answers are necessary to stabilize these difficulty parameters and found higher increase in stability for models utilizing response times. This increase in stability is probably mainly due to the use of more "bits of information" per each answer.

For the studied case study, the main conclusion is that differences in modeling of response times have larger impact than differences in domain modeling. This result is interesting, since much more research has been devoted to domain modeling than to response times modeling.

References

1. Michel C Desmarais and Ryan SJ d Baker. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, 22(1-2):9–38, 2012.
2. S Klinkenberg, M Straatemeier, and HLJ Van der Maas. Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. *Computers & Education*, 57(2):1813–1824, 2011.
3. Radek Pelánek. Application of time decay functions and Elo system in student modeling. In *Educational Data Mining (EDM)*, pages 21–27, 2014.
4. Jiří Řihák. Use of time information in models behind adaptive system for building fluency in mathematics. In *Educational Data Mining, Doctoral Consortium*, 2015.