Inferring the Differential Student Model in a Probabilistic Domain using Abduction Inference in Bayesian Networks

Nabila Khodeir[1], Nayer Wanas[2], Nevin Darwish[3], Nadia Hegazy[1]
nabilakhodeir@yahoo.com, nayerw@microsoft.com, ndarwish@ieee.org, nhegazy@mcit.gov.eg
[1]Informatics Dep., Electronics Research Institute, Tahrir St.,Giza, Egypt
[2]Cairo Microsoft Innovation Lab, 306 Chorniche El-Nile, Maadi, Cairo, Egypt
[3]Dept. of Computer Engineering, Faculty of Engineering, Cairo University, Giza, Egypt

Abstract. In this paper we aim to estimate the differential student knowledge model in a probabilistic domain within an intelligent tutoring system. The suggested algorithm aims to estimate the actual student model through the student answers to questions requiring diagnosing skills. Updating and verification of the model are conducted based on the matching between the student and model answers. Two different approaches to updating namely coarse and refined model are suggested. Results suggest that the refined model, although takes more computational resources, provides a slightly better approximation of the student model.

1 Introduction

The student model is a core component in any intelligent or adaptive tutoring system that represents many of the student features such as knowledge and individual traits [1]. Differential model is one from several models that have been suggested for the student knowledge modeling [5]. The differential model represents both the student knowledge and the differences between student and expert knowledge, which represents the gap in the student's knowledge [2].

We propose a method to build a differential student model in a probabilistic domain based. A Bayesian network is used to represent both the domain and the student model. Based on the domain structure we generate problems that require diagnostic skills to be solved. The discrepancies between the answers generated by the student model and the answers provided by the student are used to update the student model.

2 Estimating of the Student knowledge Model

A diagnostic question is generated and presented to the student. Subsequently, the answer provided by the student to the question is compared to that generated by the initial student model using an abduction inference through the Bayesian network [3,4]. If the answers match the student model doesn't require regulation. On the other hand, if there is a discrepancy between the answers the student model is adjusted. Since the answer is a ranked list of hypotheses, the difference between the two answers can be either (i) missing or (i) adding or (iii) incorrect order hypotheses or (iv) any combinations of these differences. The addition or absence of a hypothesis is related to difference in the relations between domain items, On the other hand, the difference between hypotheses order is referred to the differences in the weight of theses relations. Upon identifying the type of difference, the student model should be updated accordingly. Two different approaches are used to update the student model, namely (i) coarse model update and (ii) refined model update. The coarse model update is conducted by adding or removing relations of the differences in the hypotheses and swapping between relations weights for the differences in the hypotheses order. Refined model update, on the other hand, is performed using successive increase or decrease in the weights of the different hypotheses according to the nature of the differences. Model verification is evaluated to verify that the new model generates answer which matches the answer provided by the student.

Simulated students and different groups of questions are used to evaluate the algorithm. The questions are evaluated against the initial student model, generating the original match. The performance of the algorithm is evaluated after each question. Moreover, the set of questions are all tested against the new adjusted model for verification. The performance of the algorithm is evaluated using the accuracy of estimating the student answers to the questions. Coarse and refined updating model approaches are successful in estimating the student model with an accuracy of over 84% for individual match, and 70% for verification match.

3 Discussion and Conclusion

An experimental evaluation of the approaches has been conducted. The results suggested that the refined model updating exhibits similar performance with respect to accuracy compared to the coarse updating. However, it estimates models that are closer to the actual model by at least 8% compared to the coarse updating model. Applying the questions in any order doesn't have a significant effect on the overall performance of the algorithm. On the other hand, the proximity of the initial model selected to that of the student, or to the knowledge model improves the performance of this approach significantly.

References

[1] Brusilovsky, P., Millan, E. User Models for Adaptive Hypermedia and Adaptive Educational Systems, *In Brusilovsky, P., Kobsa, A., Nejdl, W.(eds.) The Adaptive Web, LNCS 4321*, 2007, pp. 3-53.

[2] Burton, R, Brown, J A tutoring and student modelling paradigm for gaming environments, *SIGCSE-SIGCUE*, 1976, pp. 236-246

[3] Elvira Project, http://www.ia.uned.es/~elvira/index-en.html

[4] Nilsson, D. An efficient algorithm for finding the M most probable configurations in probabilistic expert systems, *Stat. and Comp*. 1998, 8(2), pp. 159-173

[5] Pahl, C., Kennyand, C. Interactive Correction and Recommendation for Computer Language Learning and Training, *IEEE Trans. KDE 2009*, 21(6), pp.854-866.