

# Personalizing Knowledge Tracing: Should We Individualize Slip, Guess, Prior or Learn Rate?

Junjie Gu, Yutao Wang, and Neil T. Heffernan

Department of Computer Science, Worcester Polytechnic Institute  
Worcester, MA, USA  
{jgu2, yutaowang, nth}@wpi.edu

**Abstract.** The intelligent tutoring system field is concerned with ways of personalizing to the student. Wang and Heffernan introduced the Student Skill model and showed that it was reliably better than the Knowledge Tracing (KT) model in predictive accuracies. One limitation of their work is that they only investigated one particular way of personalizing, which individualizes all four KT parameters simultaneously. But it may be better if we just use some of the parameters to personalize the model. More generally, we want to address the research question: What are the most important features to personalize? In this work, we systematically explored all 16 possible ways of incorporating student features into the model. We found that prior and slip are the two most important features to individualize, and the best model is the one with all four parameters individualized. Additionally, the one parameter that can be dropped without any hurt to performance is guess.

**Keywords:** Knowledge Tracing, Bayesian Networks, prediction, personalization, Intelligent Tutoring System.

## 1 Introduction

The traditional way of modeling student knowledge is Corbett and Anderson's Knowledge Tracing (KT) model [1]. Wang and Heffernan introduced the Student Skill (SS) model [5] and showed that it was reliably better than the KT model in predictive accuracies. The goal of our experiment is to search for the best structures of the SS model by trying all 16 possible ways of incorporating student features. The dataset we used came from the 2009-2010 school year of ASSISTments, containing 1775 distinct students, 123 distinct skills and 695,732 data points. The code and data used in the experiments are available online [6].

## 2 Methodology and Discussion

In this paper, we investigated the research question: which of the four features: slip, guess, prior, or learn rate of student are most important to individualize in a Bayesian Knowledge Tracing framework. We extended Wang and Heffernan's work by exploring more structures of the SS model and searched for the best combination of individualization features.

Two major observations were made from the experiments. First, the results showed that if we individualize only one feature for student, the most valuable feature would be slip or prior. It is not surprising that prior is an important feature to individualize since students' prior knowledge differs greatly. Since slip represents the probability of a wrong answer given the student knows the skill, the fact that individualizing slip makes the greatest difference suggests teachers or tutoring systems may need to pay attention to the students with large slip rates to check if they lose interest after mastering a skill or if they are still confused with some aspects of the skill while already mastered the major part of it, and take different actions accordingly.

Second, the single best model is the one with all four parameters personalized for student, but is not reliably different than the one without student guess. This result indicates that if we don't want to individualize all four parameters due to efficiency or data amount, guess rate could be the first feature to consider removing.

This paper investigated a new research question. No one in the ITS field has looked at what parameters to best individualize but this opens up a whole new idea. Our finding that prior and slip are more important to individualize is a novel contribution. But we did not answer the question, why is this so? What is it about prior and slip that gives this extra boost in precision? This raises a new question about what might be better ways to individualize.

**Acknowledgements.** We acknowledge funding from NSF (#1316736, 1252297, 1109483, 1031398, 0742503 ), ONR's 'STEM Grand Challenges' and IES (# R305A120125 & R305C100024).

## References

1. Corbett, A., Anderson, J.: Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. *User Modeling and User-Adapted Interaction* 4, 253–278 (1995)
2. Gong, Y., Beck, J.E., Heffernan, N.T.: Comparing Knowledge Tracing and Performance Factor Analysis by Using Multiple Model Fitting. In: Alevan, V., Kay, J., Mostow, J. (eds.) *ITS 2010, Part I. LNCS*, vol. 6094, pp. 35–44. Springer, Heidelberg (2010)
3. Murphy, K.P.: *The Bayes Net Toolbox for Matlab*, Computing Science and Statistics (2007), <http://www.cs.ubc.ca/~murphyk/Software/BNT/bnt.html>
4. Pardos, Z.A., Heffernan, N.T.: Modeling Individualization in a Bayesian Networks Implementation of Knowledge Tracing. In: De Bra, P., Kobsa, A., Chin, D. (eds.) *UMAP 2010. LNCS*, vol. 6075, pp. 255–266. Springer, Heidelberg (2010)
5. Wang, Y., Heffernan, N.T.: The Student Skill Model. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) *ITS 2012. LNCS*, vol. 7315, pp. 399–404. Springer, Heidelberg (2012)
6. Personalizing Knowledge Tracing, <http://tinyurl.com/ohofyrg>