Expanding Knowledge Tracing to Prediction of Gaming Behaviors

Sarah E Schultz
Worcester Polytechnic Institute
100 Institute Rd
Worcester, MA
seschultz@wpi.edu

Ivon Arroyo
Worcester Polytechnic Institute
100 Institute Rd
Worcester, MA
iarroyo@wpi.edu

ABSTRACT

Knowledge tracing has been used to predict students' knowledge and performance for almost twenty years. Recently, researchers have become interested in looking at students' behaviors, especially those considered gaming behaviors. In this work, we attempt to leverage a variation of knowledge tracing to predict gaming behaviors without damaging the prediction of performance. We compare the predictions of this model to those of knowledge tracing and a separate engagement tracing model.

Keywords

Knowledge tracing, affect, engagement, gaming, behavior

1. INTRODUCTION

When Corbett and Anderson first published the knowledge tracing model in 1995, they claimed that their goal was "to implement a simple student modeling process that would allow the tutor to […] tailor the sequence of practice exercises to the student's needs" [1]. While knowledge tracing is generally able to predict students' performance "quite well," it does not take into account the possibility of disengagement. Traditionally, knowledge tracing is used with the probability of transition from a learned to an unlearned state set at 0, so students who become disengaged are not presumed to be forgetting the skill. When the forgetting transition is allowed, models such as knowledge tracing can become confounded, mistaking disengagement for unlearning, as illustrated in Figure 1.

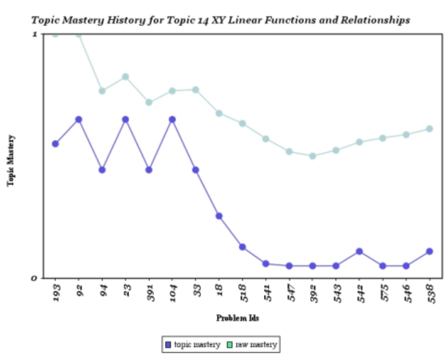

Figure 1- Bayesian Knowledge Estimation of a student on one skill (bottom line)

Figure 1 suggests that this student was un-learning, while after looking at the logs in detail, it was clear that, after the 7th problem, the student was just clicking through all the available multiple-choice answers without attempting to answer correctly.This type of behavior is defined by Baker et al as "gaming the system" [2] and is considered to be an indicator of disengagement or negative affect. Some work has been done in modeling engagement and affect in Intelligent Tutoring Systems [3], but relatively little research has focused on combining these methods with ways of tracking knowledge, such as knowledge tracing, in order to create a model that can predict both student performance and disengaged behavior and intervene appropriately.

2. PREVIOUS WORK

2.1 Bayesian Knowledge Tracing

Corbett and Anderson's Bayesian Knowledge Tracing (BKT) [1] (Figure 2) is a hidden Markov model. At each time step there is a latent node, knowledge, and an observed node, performance. The parameters for this model are $P(L_0)$, the probability that a student already knows the skill; P(T), the probability of learning the skill from one time-step to the next; P(G), the probability that a student who does not know the skill but correctly guesses; and P(S), the probability that a student who does know the skill slips and gets the answer incorrect. As mentioned in the introduction, P(F), forgetting, is traditionally set at 0, however for this work we allow forgetting in order to see if looking at behavior affects the amount of forgetting that students appear to do.

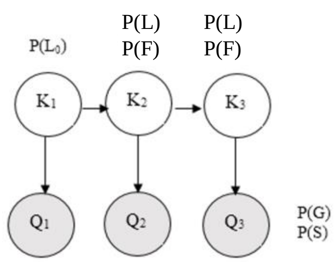

Figure 2- Bayesian Knowledge Tracing

2.2 HMM-IRT

In 2006, Johns and Woolf proposed the Dynamic Mixture Model (DMM) for predicting student knowledge and engagement [4]. They used a hidden Markov model like BKT for tracing engagement, but paired it with an Item Response Theory-like model for predicting knowledge. Rather than predicting knowledge at each time step, there is a single knowledge node for every skill and students' performance relies on that in addition to their engagement state. This allowed more accurate knowledge predictions than IRT alone, as disengagement, indicated by gaming behaviors, could explain away some incorrect attempts, rather than attributing those to knowledge.

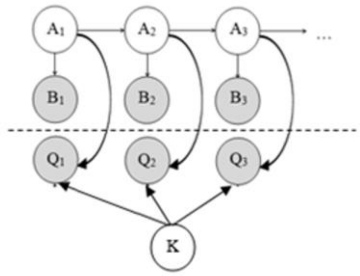

Figure 3- Dynamic Mixture Model

2.3 The KAT Model

In our previous work [5], we proposed the knowledge and affect tracing (KAT) model (Figure 5), which combines two hidden Markov models, BKT and the engagement tracing piece of DMM. As in DMM, affect influences performance. This model was able to predict both performance and behavior better than the dynamic mixture model, but did not predict performance as well as standard BKT, perhaps due to over-parameterization [5].

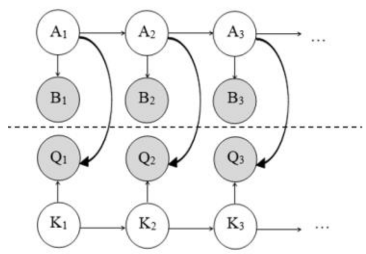

Figure 4- The KAT Model

3. THE KTB MODEL

We propose the "Knowledge Tracing with Behavior" (KTB) model. This model has only one latent node, which we call "knowledge"-- although in reality is a combination of both knowledge and engagement-- and two observables, performance and gaming behaviors. This model is shown in Figure 5.

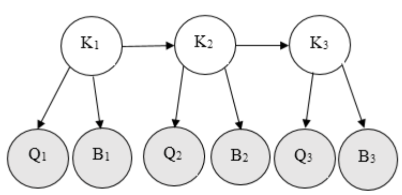

Figure 5- KTB Model

This model has fewer parameters than the dynamic mixture model or KAT model, but still can predict both performance and disengaged behavior of the students.

The variable called Gaming Behavior (B) is defined as either gaming or normal. See our definition for "gaming" in this context in our previous work [5].

4. BAYESIAN ENGAGEMENT TRACING

Since the performance prediction of the KTB model can be compared to that of Bayesian Knowledge Tracing, it is necessary to have a model of engagement tracing to compare the behavior predictions. To that end, we include a model of "Bayesian Engagement Tracing" (BET) in this work, which is the same as the HMM part of Johns and Woolf's model or the engagement piece of the KAT model, but not connected to any other model (top part of figure 4).

5. DATASETS AND METHODS

The data and methods used in this work was the same as that used in [5]. The data came from two tutors for middle and high school mathematics, ASSISTments and Wayang Outpost. For details, please see [5] in the main conference proceedings.

6. RESULTS AND ANALYSIS

While KT and KTB both outperform KAT and DMM in all predictions, in seven of the nine knowledge components, KTB was better able to predict performance than standard knowledge tracing, although the only significant difference between the two was in the ASSISTments skill "Circle Graph" ($p=0.03$). Interestingly, the Bayesian engagement tracing model was better able to predict students' behavior than KTB in eight of the nine knowledge components, although the differences are again not significant, except in two cases, "Box and Whisker," and "Triangles" ($p=0.02$).

7. DISCUSSION

We have proposed a new model, knowledge tracing with behavior, which can predict both student performance and behavior, and have shown that it can do so at least as well as BKT and a separate Bayesian engagement tracing, at predicting future behaviors (correctness at responding math problems and gaming behaviors). KTB seems to stop the false forgetting effect that is recorded by KT when forgetting is not allowed to be zero.

ACKNOWLEDGEMENTS

This research is supported by the Office of Naval Research, STEM Challenge Award, # N0001413C0127US. We also acknowledge funding from NSF (#1316736, 1252297, 1109483, 1031398, 0742503), and IES (# R305A120125 & R305C100024). Any opinions or conclusions expressed are those of the authors, not necessarily of the funders.

REFERENCES

[1] Corbett, A.T., Anderson, J.R., "Knowledge tracing: Modeling the acquisition of procedural knowledge." *User Modeling and User-Adapted Interaction*, 1995, 4, p.253-278.

[2] Baker, R.S., Corbett, A.T., Koedinger, K.R., Wagner, A.Z. (2004) Off-Task Behavior in the Cognitive Tutor Classroom: When Students "Game The System". In *Proceedings of ACM CHI 2004: Computer-Human Interaction*, 383-390.

[3] Beck, J.E. "Engagement tracing: using response times to model student disengagement." *Proceedings of AIED conference*, 2005. p. 88-95. IOS Press

[4] Johns, J. and Woolf, B.P. "A Dynamic Mixture Model to Detect Student Motivation and Proficiency." *Proceedings of AAAI Conference*, 2006, 1, p. 163-168.

[5] Schultz, S. and Arroyo, I. "Tracing Knowledge and Engagement in Parallel in an Intelligent Tutoring System." To appear in *Proceedings of the 7th Annual International Conference on Educational Data Mining*, 2014