A Prediction Model Uses the Sequence of Attempts and Hints to Better Predict Knowledge: Better to Attempt the Problem First, Rather Than Ask for a Hint

Hien D. Duong Linglong Zhu Yutao Wang Neil T.Heffernan

Department of Computer Science

Worcester Polytechnic Institute

100 Institute Road, Worcester, MA

hdduong@wpi.edu lzhu@wpi.edu yutaowang@wpi.edu nth@wpi.edu

ABSTRACT

Intelligent Tutoring Systems (ITS) have been proven to be efficient in providing students assistance and assessing their performance when they do their homework. Many research projects have been done to analyze how students' knowledge grows and to predict their performance from within intelligent tutoring system. Most of them focus on using correctness of the previous question or the number of hints and attempts students need to predict their future performance, but ignore how they ask for hints and make attempts. In this paper, we build a Sequence of Actions (SOA) model taking advantage of the sequence of hints and attempts a student needed for previous question to predict students' performance. We used an ASSISTments dataset of 66 students answering a total of 34,973 problems generated from 5010 questions over the course of two years. The experimental results showed that the Sequence of Action model has reliable predictive accuracy than Knowledge Tracing.

Keywords

Knowledge Tracing, Educational Data Mining, Student Modeling, Sequence of Action model.

1. INTRODUCTION

Understanding student behavior is crucial for Intelligent Tutoring Systems (ITS) to improve and to provide better tutoring for students. For decades, researchers in ITS have been developing various methods of modeling student behavior using their performance as observations. One example is the Knowledge Tracing (KT) model (Corbett and Anderson, 1995), which uses a dynamic Bayesian network to model student learning. But KT focuses attention on students' performance of correctness, ignoring the process a student used to solve a problem. Many papers have shown the value of using the raw number of attempts and hints (Feng, Heffernan and Koedinger, 2009, Wang, Heffernan 2011). However, most EDM models we are aware of (with one notable exception of Ben Shih, et al. (2012)) have ignored the sequencing of action.

Consider a thought experiment. Suppose you know that Bob Smith asked for one of the three hints and makes one wrong answer before eventually getting the question correct. What if someone told you that Bob first made an attempt then had to ask for a hint compared to him first asking for a hint and then make a wrong attempt? Would this information add value to your ability to predict whether Bob will get the next question correct? We suspected that a student who first makes an attempt might be a better student.

In this work, we define a Sequence of Action (SOA) model that uses the information about the action sequence of attempts and hints for a student in previous question to better predict the correctness of next question. In SOA, students' sequences of actions are divided into five categories: One Attempt, All Attempts, All Hints, Alternative Attempt First and Alternative Hint First. The results of tabling methods indicate that it is better to attempt the problem first rather than ask for a hint. Another highlight of this paper is that we used the next question's percent correct from the tabling method as a continuous variable to fit a binary logistic regression model for SOA. The experimental results show that the SOA outperforms KT in all three metrics (MAE, RMSE, AUC).

2. Sequence of Action Model

2.1 Tabling Method

There are many different sequences of actions. Some students answered correctly only after one attempt and some students kept trying many times. Some students asked for hints and made attempts alternatively, which we believe that they were trying to learn by themselves. In the data, there are 217 different sequences of actions. We divided them into five bins: (1) One Attempt: the student correctly answered the question after one attempt; (2) All Attempts: the student made many attempts before finally get the question correct; (3) All Hints: the student only asked for hints without any attempts at all; (4) Alternative, Attempt First: the students asked for hints and made attempts alternatively and made an attempt at first; (5) Alternative, Hint First: the students asked for hint and made attempts alternatively and asked for a hint first.

We used 34,973 problem logs of sixty-six 12-14 year-old, 8th grade students participated in one class from ASSISTments, which is an online tutoring system giving tutorial assistance if a student makes a wrong attempt or asks for help. Questions in each problem set are generated randomly from several templates and there is no problem-selection algorithm used to choose the next question. Table 1 shows the sequence of action division and some examples in each category, and the correct percent of next question from tabling method (Wang, Pardos and Heffernan2011). Notice that each sequence ends with an attempt because in ASSISTments, a student cannot continue to next question unless he or she fills in the right answer of the current problem. In Table 1, 'a' stands for answer and 'h' stands for hint. For example, 'aha' indicates a student makes an attempt and then asks for a hint before finally types the right the answer.

Table 1. Sequence of Action Category and Examples

Sequence of Action Bin	Examples	Next Question Correct Percent
One Attempt (a)	a	0.8339
All Attempts (a+)	aa, aaa, …, aaaaaaaaaaaa	0.7655
All Hints (h+)	ha, hha,…, hhhhhhha	0.4723
Alternative, Attempt First (a-mix)	aha, aahaaha,…, aahhhhaaa	0.6343
Alternative, Hint First (h-mix)	haa, haha,…, hhhhaha	0.4615

From the tabling results, shown in Table 1, we can see that the percent of next-question-correct is highest among students only using one attempt since they master the skill the best. They can correctly answer the next question with the same skill. For students in All Attempts category, they are more self-learning oriented, they try to learn the skill by making attempts over and over again. So they get the second highest next-question-correct percent. But for students in the All Hints category, they do the homework only relying on the hints. It is reasonable that they don't master the skill well or they don't even want to learn, so their next-question-correct percent is very low. The alternative sequence of action reflects students' learning process. Intuitively, these students have positive attitude for study. They want to get some information from the hint based on which they try to solve the problem. But the results for the two alternative categories are very interesting. Though students in these two categories alternatively ask for hints and make attempts, the first action somewhat decided their learning altitude and final results. For students who make an attempt first, if they get the question wrong, they try to learn it by asking for hints. But for students who ask for a hint first, they seem to have less confidence in their knowledge. Although they also make some attempts, from the statistics of action sequence, they tend to ask for more hints than making attempts. The shortage of knowledge or the negative study attitude makes their performance as bad as the students asking exclusively for hints first.

2.2 SOA Binary Logistic Regression Model

In this section we build a logistic regression model based on sequence of action to better predict students' performance. In this model, we want to use students' current sequence of action to predict their performance on next question in same skill. The dependent variable is students' actual performance on a question, correct or incorrect, and the independent variables are categorical factor Skill_ID and continuous factor Next_Question_Correct_Percent from Table 1, which indicates the sequence of action of current question. For example, if sequence of action of current problem is "hhhhaha", we use 0.4615 its value. We equally split 66 students into six groups, 11 students in each, to run 6-fold cross validation. The SOA and KT model are trained on the data from every five groups and are tested on the sixth group.

Table2 shows experimental result of three metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Area Under ROC Curve (AUC). Lower values for MAE and RMSE indicate better model fit while higher values for AUC reflect a better fit. The values are calculated by averaging corresponding numbers obtained in each experiment of the 6-fold cross validation. The raw data and results for the six groups is available at this website:

(http://users.wpi.edu/~lzhu/SOA/DataSet_and_Results.rar).

Table 2. Prediction accuracy of KT, SOA and Ensemble

	MAE	RMSE	AUC
KT	0.3032	0.3921	0.6817
SOA	0.2900	0.3813	0.6841
t-test p value	0.0000	0.0000	0.5286

Although most numbers seem very close, SOA outperforms KT in all three metrics. To examine whether the difference were statistically reliable, we did a 2-tailed paired t-test based on the result from the cross validation. The last row in Table 2 shows that the differences are significant in both MAE and RMSE.

3. CONTRIBUTIONS

In this work, we presented a Sequence of Actions (SOA) model, in which students' action of asking for hints and making attempts are divided into five categories shown in Table 1. The result of a tabling method shows that students who make an attempt first did better on next question with the same skill than those who ask for a hint first. The result from logistic regression shows that paying attention to the sequence of action increases prediction accuracy of students' performance.

4. ACKNOWLEDGMENTS

This work is funding in part by IES and NSF grants to Professor Neil Heffernan. The first three authors have all been funded by the Computer Science Department at WPI.

5. REFERENCES

[1] Wang Q.Y., Pardos, Z. A., & Heffernan, N. T. (2011). *Tabling Method–A simple and practical complement to Knowledge Tracing*. KDD.

[2] Feng, M., Heffernan, N.T., &Koedinger, K.R. (2009). *Addressing the assessment challenge in an Online System that tutors as it assesses*. UMUAI: The Journal of Personalization Research19(3), 243-266.

[3] Corbett, A. T., & Anderson, J. R. (1995). *Knowledge tracing: modeling the acquisition of procedural knowledge*. User Modeling and User-Adapted Interaction, 4, 253–278.

[4] Shih, B., Koedinger, K.R., &Scheines, R. (2010). *Discovery of Learning Tactics using Hidden Markov Model Clustering*. Proceedings of the 3rd International Conference on EDM.