



Modeling Student Learning Behaviors in ALEKS: A Two-Layer Hidden Markov Modeling Approach

Guoyi Wang¹, Yun Tang¹ , Junyi Li¹, and Xiangen Hu^{1,2}

¹ Central China Normal University, Wuhan 430079, Hubei, China
tangyun@mail.ccnu.edu.cn

² University of Memphis, Memphis, TN 38152, USA

Abstract. This study analyzed 25,783 log data entries of student learning activities on a self-paced online intelligent tutoring system. The behavior patterns between the high- and low-achievement students and of different mathematical topics were compared using two-layer hidden Markov model. The results showed that high-achievement students exhibited more effective learning behaviors, such as asking for explanation and practicing after making an error. In contrast, low-achievement students tended to make consecutive errors without seeking help. Moreover, students' learning behaviors tended to be more effective when learning simple topics. Our findings implied that intelligent tutoring systems could track the behavior patterns of students and detect ineffective learning states, so as to provide learning support accordingly.

Keywords: Intelligent tutoring system · Two-layer hidden Markov model
Behavior pattern

1 Introduction

It is commonly believed that human tutors are more effective than computer tutors when teaching the same content. Human tutors were found to be better at scaffolding learners, as well as giving feedback that encourages learners to engage in interactive and constructive behaviors [1]. Another key property of effective tutoring was the granularity of the interaction [1]. For example, Intelligence tutoring systems (ITS) can be classified into answer based, sub step-based, and step-based tutor according to the grain size. The step-based ITSs—which have the finest granularity—were nearly as effective as human tutoring [1–3]. To determine the granularity of interaction and deliver necessary support in ITSs, it is critical to unveil the behavior patterns through student modeling. Because ITSs automatically record student activities in the log data, educational data mining techniques can be applied for understanding the behavior patterns and further facilitate the improvement of learning process [4].

In the present study, we propose to use two-layer hidden Markov model (TL-HMM) to model the sequential patterns in ITS log data. TL-HMM is a method developed on the basis of Hidden Markov Models (HMM) [5]. It decomposes the process into two nested layers of models. The first layer is treated using a non-hidden Markov model. Each micro

activity is treated as an observable symbol and the state is produced from the transition between each activity upon the distribution condition. The second layer treats the state from the first layer as a latent state in a hidden Markov model. The transitions between latent states are constructed by the relations and dependencies of these micro activities. TL-HMM can be used to decode the latent state sequences of students. Instructors may use the representation of student behavior patterns and behavior transitions to identify the hidden knowledge and understand the dynamic change of the latent behaviors in massive interaction datasets [5].

The present study attempted to discover specific behavior patterns from ITS log data using TL-HMM. The investigation into behavioral patterns may help us to understand how students learn in the system and predict their performance level.

2 Methods

The data in the present study were obtained from Assessment and Learning in Knowledge Spaces (ALEKS) [6], a web-based intelligent tutoring system. In ALEKS, practice problems are organized into topics, while each problem is attached with an explanation of concepts and procedures in the form of worked examples. Students may request for explanations at any time, or ALEKS would prompt students with explanations after two consecutive errors. After a certain number of practice, students would be graded as mastering or failing the topic according to the scoring rules, of which the artificial intelligence is based in the Knowledge Space Theory [7].

The data we used consisted of 25,783 entries from six high schools in the United States from 2010 to 2014. Table 1 lists the four topics. The unique ID of schools and students were omitted for anonymity. Each entry represents a completed sequence of actions ($M = 5.81$, $SD = 4.32$) for practicing a specific topic. There were four kinds of actions: L (receiving an explanation prompted by ALEKS), E (requesting an explanation), C (answering correct) and W (answering wrong). The duration of each action was also contained in the data. Due to the lack of individual information, we treated each entry as from an independent student and categorized the students who mastered the topic as high-achievement, while those who failed as low-achievement.

Table 1. Contents of the four mathematical topics in the data set.

Topic	Content
Algebra 208	Solving a two-step equation with signed fractions
Algebra 209	Solving a linear equation with several occurrences of the variable
Algebra 703	Solving a word problem using a quadratic equation with rational roots
Algebra 224	Solving a word problem using a system of linear equations

The sequences of actions were analyzed using TL-HMM implemented in the MeTA toolkit [8]. Two patterns were derived to characterize student behaviors: the latent state representations and the latent state transitions. We then used python-igraph [9] to visualize these behavior patterns.

3 Results and Discussion

In this section, we compared the behavior patterns of high- and low-achievement students. All data entries were classified into failure and mastery according to the end results and were respectively fitted. TL-HMM uses latent states to model these behavior patterns, while the transitions between latent states captured the progress of learning over time. [5] suggested that the number of states should be empirically determined and based on the goal of analysis. Accordingly, we fit models of 2-10 states and plotted the corresponding latent state representations. We then examined these representations and picked the ones with meaningful patterns. Likelihood-based or predictive accuracy-based methods, such as BIC or cross validation, may also be used to assist the model selection process.

Figure 1 shows the latent state representations of a 3-state TL-HMM that best fits the sequence of actions of low-achievement students (i.e., those who failed an item). Figure 2 shows the latent states of a 4-state TL-HMM that best fits the actions of high-achievement students (i.e., those who mastered an item). In each representation, the nodes stand for different behaviors. The size of every node is set to be proportional to its probability of being visited during a random walk. The arrows indicate the direction of transfer between two nodes. The thickness of the curves reflects the probability of the nodes transfer.

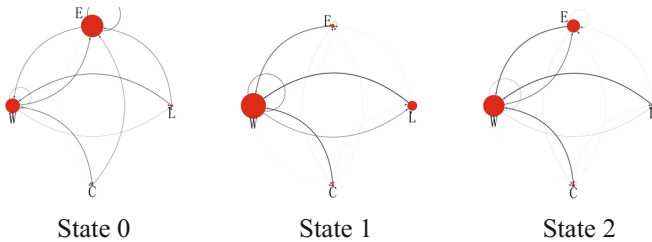


Fig. 1. A 3-state TL-HMM fit to actions preceding failure. The thickness of the paths indicate the relative size of transition probability.

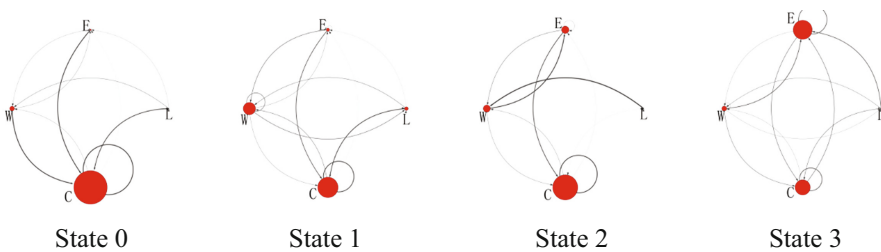


Fig. 2. A 4-state TL-HMM fit to actions preceding mastery.

Figure 3 presents the latent state transition diagrams. The nodes in the transition diagram stand for different behavior patterns reflected by the latent states in Figs. 1 and 2. For low-achievement students, there is a very high probability of staying in State 1 (constantly answer wrong without calling for explanation) while State 0 (help-seeking) has the lowest probability. On the other hand, for high-achievement students, there is a relatively high probability of State 0 (always answer correct) while State 3 (ineffectively overusing help) has the lowest probability. In either category, the probabilities of transferring between latent states are very small. It is reasonable because each learning sequence has an average length of about 12.5 min. Hence we did not expect a frequent transfer between latent states.

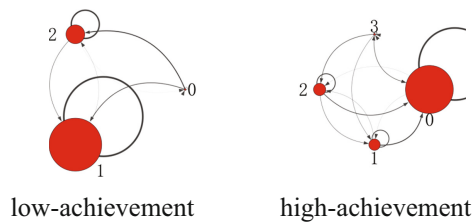


Fig. 3. The latent state transition diagrams for the 3-state TL-HMM fit to the low-achievement students (left panel) and the 4-state TL-HMM fit to the high-achievement students (right panel).

We further classified the data entries of failure and mastery into the four topics (see Table 1). The 3-states model learned on the “failure” students was retrofitted to train the data set of “failure” students on each topic. The model parameters associated with their Markov model representations were forced to be fixed, so the latent state meanings would not drift. Similar analysis was also performed to the mastery data set of four topics. The general patterns of latent state transitions were held in each topic, although there are some topic-specific differences of interest. On one hand, when solving complex topics (i.e., Algebra 224 and Algebra 703), the high-achievement students mastered the item without much help from explanations (State 0), while the low-achievement students requested more explanations (State 2) along with practice. It may be inferred that explanation alone was not as effective as the system may expect. On the other hand, when solving simpler topics (i.e., Algebra 208 and Algebra 209), the high-achievement students benefitted more from repeated practice (State 1) or help-seeking (State 2), while the low-achievement students were trapped in ineffective practice (State 1) but fail to learn from explanations.

4 Conclusion

The present study explored the patterns of learning behaviors in ALEKS using a TL-HMM approach. The results demonstrated that this approach could facilitate the interpretation of learning process in ALEKS or similar ITSs. Specifically, low-achievement students rarely asked for explanation. Upon detecting ineffective learning

states, ITSs may consider providing additional assistance or even bringing outside support (e.g., a human tutor) to help the students learn.

Acknowledgement. This study was partially supported by Self-determined Research Funds of CCNU from the Colleges' Basic Research and Operation of Ministry of Education, China (No. CCNU15A05049; No. CCNU16JYKX38).

References

1. VanLehn, K.: The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educ. Psychol.* **46**(4), 197–221 (2011). <https://doi.org/10.1080/00461520.2011.611369>
2. Hu, X., Craig, S.D., Bargagliotti, A.E., et al.: The effects of a traditional and technology-based after-school program on 6th grade student's mathematics skills. *J. Comput. Math. Sci. Teach.* **31**(1), 17–38 (2012)
3. Ma, W., Adesope, O.O., Nesbit, J.C., et al.: Intelligent tutoring systems and learning outcomes: a meta-analysis. *J. Educ. Psychol.* **106**(4), 901–918 (2014)
4. Koedinger, K.R., Brunskill, E., Baker, R.S.J.D., et al.: New potentials for data-driven intelligent tutoring system development and optimization. *AI Mag.* **34**(3), 27–41 (2013)
5. Geigle, C., Zhai, C.: Modeling student behavior with two-layer hidden Markov models. *J. Educ. Data Min.* **9**(1), 1–24 (2017)
6. Craig, S.D., Hu, X., Graesser, A.C., et al.: The impact of a technology-based mathematics after-school program using ALEKS on student's knowledge and behaviors. *Comput. Educ.* **68**, 495–504 (2013)
7. Huang, X., Craig, S.D., Xie, J., et al.: Intelligent tutoring systems work as a math gap reducer in 6th grade after-school program. *Learn. Individ. Differ.* **47**, 258–265 (2016)
8. Massung, S., Geigle, C., Zhai, C.: MeTA: a unified toolkit for text retrieval and analysis. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics—System Demonstrations*, pp. 91–96. Association for Computational Linguistics, Berlin (2016). <https://doi.org/10.18653/v1/p16-4016>
9. Csardi, G., Nepusz, T.: The igraph software package for complex network research. *InterJ. Complex Syst.* **1695**(5), 1–9 (2006)