Rodrigo, M. M. T. et al. (Eds.) (2021). Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Toward Educational Explainable Recommender System: Explanation Generation based on Bayesian Knowledge Tracing Parameters

Kyosuke TAKAMI*, **Brendan FLANAGAN, Yiling DAI & Hiroaki OGATA**
Academic Center for Computing and Media Studies, Kyoto University, Japan
*takami.kyosuke.2z@kyoto-u.ac.jp

Abstract: Explainable recommendations, which provides explanations about why an item is recommended, helps to improve the transparency, persuasiveness, and trustworthiness. However, there are few research in educational technology that utilize explainable recommendation. Previous research has identified that learner motivation can deteriorate while using educational recommender systems, and that providing additional forms of feedback can improve performance and increase trust. In this paper, we propose an explanation generator using the following parameters from Bayesian knowledge tracing models: guess (giving a correct answer despite not knowing the skill) and slip (knowing a skill, but giving a wrong answer) for a quiz recommended by the system. Recommended quizzes were categorized into different feature types according to the value of the model parameter and explanation texts are generated based on these feature types.

Keywords: Explainable recommendation, Bayesian knowledge tracing, model interpretation

1. Introduction

Explainable recommendation, which provides explanation about why an item is recommended, can improve transparency, persuasiveness and trustworthiness of users (Zhang, Y., & Chen, X., 2020). In education, explanations from a learning system may also have the additional benefits for students learning from the explanations given by the system. Previous research on intelligent tutoring systems has shown that student motivation in system-based self-regulated learning can be improved by prompting and feedback mechanisms, leading to higher achievement (Duffy & Azevedo, 2015).

Recommendation explanations can be generated from different sources and presented in different display styles (Tintarev and Masthoff, 2015), for example a relevant user or item, a sentence, an image, or a set of reasoning rules. Explainable recommendation studies are categorized either model-intrinsic or model-agnostic approach (Lipton, 2018; Molnar, 2019). The model-intrinsic approach develops interpretable models, whose decision process is transparent, and thus, the explanations can be naturally provided. The model-agnostic approach allows the decision process to be a blackbox. Instead, an explanation generator provides explanations for the recommendation model after the recommendations have been provided, preparing several pre-defined possible explanation templates based on data mining methods. However, previously these studies were mainly conducted in the context of e-commerce recommender systems such as Netflix or Amazon, and much research did not attempt to implement explainable recommendation in education. To address this problem, we propose an Explanation Generator from Bayesian Knowledge Tracing (BKT) parameters. First, we tried to interpret the decision process of the BKT model from the features of the parameters and classified them into different feature types. Finally, we developed an explanation text generator based on these feature types.

2. Method

2.1 Bayesian Knowledge Tracing Model

Bayesian Knowledge Tracing (BKT) model (Corbett & Anderson, 1995) has been used to model student knowledge in various educational systems, including tutors for reading skill (Beck & Chang, 2007), computer programming (Corbett & Anderson, 1995) and mathematics (Koedinger, 2002). This model calculates the probability that a student knows a skill at a given point in time by combining data on the student's performance up to that point with model parameters. Main two parameters are guessing (giving a correct answer despite not knowing the skill) and slipping (knowing a skill, but giving a wrong answer). Students who do not know the skill will generally give incorrect answers, but there is a constant probability (called the guess parameter) that the student will give the correct answer. Correspondingly, a student who knows the skill will generally give the correct answer, but there is a certain probability (called the slip parameter) that the student will give the wrong answer. See (Baker, R.S., Corbett, A.T. & Aleven, 2008) more theoretical BKT background detail. We applied a BKT model for our quiz recommender system and propose an explanation generator from the analysis of the guess and slip BKT parameters.

2.2 Overview of Data Collecting System

The explainable recommender system proposed in this paper is built on the LEAF framework, which was developed to support the distribution of learning materials, collection and automated analysis of learning behavior logs in an open and standards-based approach (Flanagan & Ogata, 2018). The main components of the framework are: Moodle LMS which acts as a hub for accessing various courses; the BookRoll reading system for learning material and quiz exercise distribution; an LRS for collecting learning behavior logs from all of the components; and the LAView learning analytics dashboard to provide feedback to students, teachers and school administrators. This framework enables us to collect and analyze learning behaviors in real time and provide feedback to stakeholders. Quiz books used in the mathematics classes were uploaded to BookRoll and multiple-choice quiz questions were created to enable the collection of answers in learning log data.

2.3 Used data for BKT Recommender System and Explanation Generator

We collected the log data from one unit of quizzes in junior high school mathematics classes over the period of April-May, 2021. This log data included 120 students answered data right or wrong to 53 quizzes. The recommendation of learning paths in mathematics has been popularized by the application of BKT in intelligent tutoring systems (Fischer et al., 2020), and is employed to model the degree of mastery of each mathematical skill (quiz) in the recommended system based on the analysis of answers in learning log data. In this process BKT parameters guess and slip were computed in each skill (quiz). Quiz recommendations are made based on the probability that the student will correctly answer a question as determined by the BKT model, with extremely high or low probability of correct answer quizzes having less weight in the recommendation.

2.4 Feature of Guess and Slip Parameter

Since the BKT model computes parameters from learners' sequential quiz answers in free order, the relationship between quizzes and their parameters is a black box. To interpret this black box we investigated the features of guess and slip parameter in each quiz. Figure 1 (left) shows plots of a series of quiz numbers with their guess (top) and slip (bottom) values. Figure1 (right) is part of the list of quiz titles for one unit. A single unit is made up of several components, for example: quizzes numbered 1 to 5 relate to the same component about 'Monomial and polynomial'. Comparing the quiz numbers 1, 18 and 40 marked in blue on the guess plot with the list on the right, the first of the components tended to have a higher guess value. On the contrary, Quiz 27 and 37 marked in green in the list, which are advanced exercise quizzes called STEP B that depend on previously learnt components' knowledge to be solved, representing lower guess value in the guess plot. These results show that the first quiz in a component tends to be higher guess value because students haven't acquired the skills yet, and the quiz of acquiring a skill and then using that skill tends to be lower guess value, using an acquired skill, and therefore they are not guessing. It is interesting to note that STEP B quizzes do not necessarily have a lower guess value implying the guess value indicates the quiz characteristics based on the learning history of the students, which is different from the difficulty level set by the quiz author.

Next we checked the slip parameters. If students were knowing a skill, but giving a wrong answer, slip values were estimated to be higher as shown at Quiz 10 and 37 marked with yellow in Figure 1 (slip plots and right list). Quiz 37 is lower guess value and high slip value, then we marked with green and yellow in the right list.

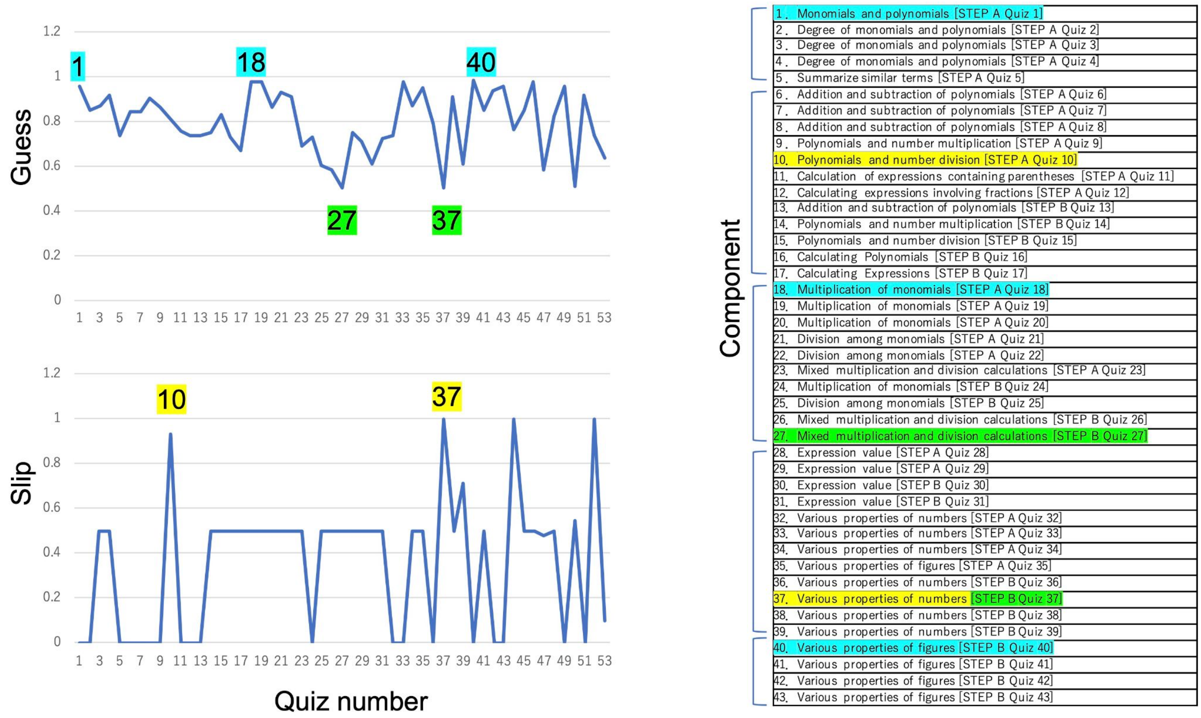

Figure 1. Guess and slip plot (left), part of the list of quiz titles for a single unit (right).

2.5 Four Different Types of Quizzes Based on BKT Parameters

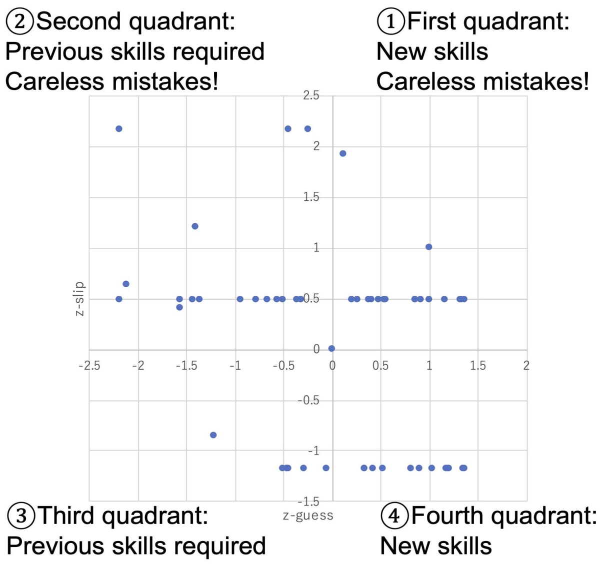

Figure 2. Four different types of quizzes based on BKT parameters: z-scaled guess-slip plot.

Using BKT guess and slip parameters, we can divide the quiz features into four categories as shown in Figure 2 (guess and slip values are z-scaled, i.e., mean of guess = 0, SD = 1). ①New skills and careless mistakes!: high guess and high slip quizzes which are the first part of the component requiring new skills and easy to make mistakes even if students know the skill. ②Previous skills and careless mistakes!: low guess and high slip quizzes which require previous components' knowledge and easy to make mistakes. ③Previous skills required: low guess and low slip quizzes requiring previous knowledge. ④New skills: high guess and low slip quizzes requiring new skills. Based on the four types of quizzes characterized by the parameters of BKT, we propose a method for generating reason explanations for the recommended quizzes.

2.6 Generating Explanation from Guess and Slip Parameters

In order to generate more appropriate explanations, a teacher with experience in teaching mathematics considered what the best explanation would be for each parameter of the quiz by reviewing each quiz. As a result of that reviewing, the area of guess-slip was divided into more detailed areas shown in n Figure 3. We set the range of values for guesses in three regions as shown by the dotted lines. As shown by the blue dotted line, we set the range at $\pm 3\sigma$ sigma around 0. The explanation type labels were attached to these regions in order from right to left on the guess axis, from Exp-1 to Exp-3. As you move from right to left on the guess axis, the quiz moves from basic skill quizzes to more applied quizzes that require the use of basic skills. In the light of these quizzes' characteristics, we can prepare three types of explanations according to guess values indicated in Table 1. The text of these explanations was created with reference to Persuasive technologies, and it shows that some persuasive messages are effective in the field of education (Hamari, J., Koicisto, J.,& Pakkanen, T., 2014). One of the four texts in each explanation type is randomly selected according to quiz' BKT parameters. In addition to guess oriented explanations, if the slip value is greater than the pink dash-dotted line showing the slip value from 0 to 3σ plus, Exp-4 explanation text is appended after the Exp-1, Exp-2 or Exp-3. For example, if $guess = -1.4$, $slip = 1.2$, an explanation will be generated choosing one of the texts in Exp-3 and one of the texts Exp-4 such as 'Now it's time to challenge applied problems! Make full use of the knowledge you have gained so far! Watch out for careless mistakes!'

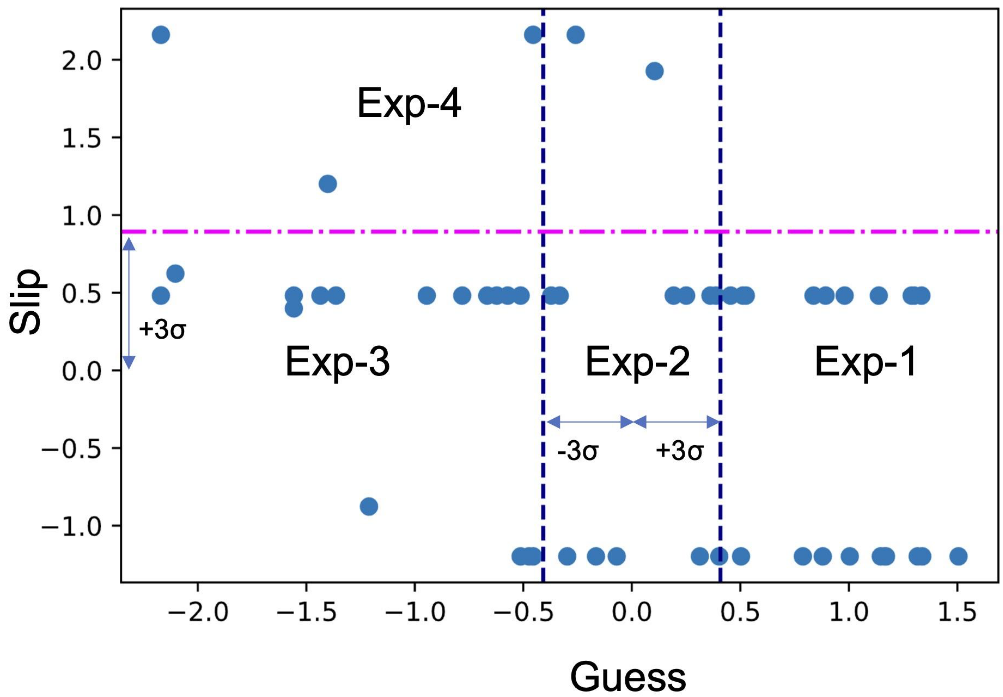

Figure 3. Generating explanation from Guess and Slip parameters: Z-scaled guess-slip plot.

Table 1. *Generated Texts for Explanation*

Explanation type	Texts of explanations
Exp-1	’You're not getting the basic skills. Let's go over the basics with this quiz!’ ‘Let's carefully go over some basic skills with this problem!’ ‘You are not getting the basic skills. Let's work on the basic problems!’ ‘You don't seem to have the basics down, so with this problem, let's get the basics down!’
Exp-2	‘If you can solve this quiz, you can try the applied exercise quiz!’ ‘This is the skill you need to solve the applied quiz!’ ‘If you can't solve this problem, you can't solve the applied quizzes!’ ‘If you can solve this quiz, you can improve your skills in applied quiz!’
Exp-3	‘It's an applied quiz that requires some skills. It's great if you can solve it!’ ‘Now it's time to challenge applied problems! Make full use of the knowledge you have gained so far!’ ‘This quiz is an applied exercise that requires the knowledge you've acquired so far. Let’s work on it.’ ‘Let's try this quiz! This is a quiz that you can solve by using your learned skills.’
Exp-4	‘This quiz is easy for everyone to make mistakes on, so be careful!’ ‘This quiz is so easy to miss!’ ‘Watch out for careless mistakes!’ ‘People often make careless mistakes on this question, so be careful!’

2.7 Generator Implemented in Quiz Recommender System

We have implemented this explanation generation algorithm into our recommendation system. Figure 5 shows a screenshot of the user interface of the implemented recommendation system. The reason for the recommendation will be displayed under the title of the recommended question. Students who see these explanations are expected to be convinced of the reason why the quiz was recommended, and to be persuaded to solve the quizzes, resulting in improved academic performance by actually solving the quiz.

Figure 4. Screenshot of recommendation UI.

3. Conclusion and Future Work

Explainable recommendation, which provides explanation about why an item is recommended, can improve transparency, persuasiveness and trustworthiness of users. In this paper, we propose an explanation generator using BKT parameters guess and slip for the quiz BKT recommendation system. Since the decision process of the BKT was not transparent, we tried to interpret it from the relationship between quizzes and parameters. Recommended quizzes were categorized into different feature types according to the value of the parameter and explanation texts are generated based on these feature types.

In future work, the effectiveness of generated explanations needs to be evaluated in the real world classroom whether students can solve the recommended quiz convincing themselves as to why the quiz was recommended by the explanation. Second, the method proposed in this paper is limited to only four types of explanations. By preparing a variety of explanations, we need to consider the best explanation for each individual's character or personality.

Acknowledgements

This work was partly supported by JSPS Grant-in-Aid for Scientific Research (B) 20H01722, JSPS Grant-in-Aid for Scientific Research (Exploratory) 21K19824, JSPS Grant-in-Aid for Scientific Research (S) 16H06304 and NEDO JPNP20006 and JPNP18013.

References

Baker, D., & RS, C. AT, and Aleven, V.(2008). More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing. In *Intelligent Tutoring Systems* (pp. 406-415).

Beck, J. E., & Chang, K. M. (2007, July). Identifiability: A fundamental problem of student modeling. In *International Conference on User Modeling* (pp. 137-146). Springer, Berlin, Heidelberg.

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, *4*(4), 253-278.

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338-348.

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education*, *44*(1), 130-160

Flanagan, B., Ogata, H. (2018). Learning analytics platform in higher education in Japan, *Knowledge Management & E-Learning: An International Journal, 10(4)*, 469-484.

Hamari, J., Koivisto, J., & Pakkanen, T. (2014, May). Do persuasive technologies persuade?-a review of empirical studies. In *International conference on persuasive technology* (pp. 118-136). Springer, Cham.

Koedinger, K. R. (2002). Toward Evidence for Instructional Design Principles: Examples from Cognitive Tutor Math 6.

Lipton, Z. C. (2018). The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31-57.

Molnar, C. Interpretable Machine Learning. Leanpub, 2019.

Tintarev N., Masthoff J. (2015) Explaining Recommendations: Design and Evaluation. In: Ricci F., Rokach L., Shapira B. (eds) Recommender Systems Handbook. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7637-6_10

Yongfeng Zhang and Xu Chen (2020), "Explainable Recommendation: A Survey and New Perspectives", Foundations and Trends® in Information Retrieval: Vol. 14: No. 1, pp 1-101. http://dx.doi.org/10.1561/1500000066