Iyer, S. et al. (Eds.) (2022). Proceedings of the 30th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Automated Test Set Quiz Maker Optimizing Solving Time and Parameters of Bayesian Knowledge Tracing Model Extracted from Learning Log

Kyosuke TAKAMI[a*], Gou MIYABE[b], Brendan FLANAGAN[a] & Hiroaki OGATA[a]
[a]*Academic Center for Computing and Media Studies, Kyoto University, Japan*
[b]*Saikyo Junior High School Attached to Saikyo High School, Kyoto, Japan*
*takami.kyosuke.2z@kyoto-u.ac.jp

Abstract: Creating a set of quizzes for the students' test is almost an irreplaceable task to teachers. In practice, a teacher could use a learning analytics dashboard while creating a test to control the quiz difficulty and the amount of time it takes to solve. This paper draws inspiration from this practical example, and we propose an automated test set quiz maker by optimizing the time and learning parameters that have been estimated from the analysis of learning logs. First, we estimate the Bayesian Knowledge Tracing (BKT) model parameters: in particular the guess and slip probability from the quiz answer log history. The system automatically generates a test set of quizzes if the user inputs the desired amount of time it should take to solve the test by optimizing the selection of quizzes based on BKT parameters and estimated solving time. This function is expected to reduce the burden of preparing examination questions for teachers, and it can be used as a trial test before the exam for students.

Keywords: Automated test set quiz generation, Bayesian Knowledge Tracing, Mathematical Optimization, Knapsack Problem

1. Introduction

Learning Analytics (LA) is an emergent field and its aims are better understanding of students and providing intelligence to students, teachers, and administrators using learning log data (Law & Liang, 2020). In school practice, teachers use learning log data to monitor students' learning status, including their study time and percentage of correct answers through a dashboard (Majumdar et al., 2020). In particular, Mr. Miyabe, an experienced junior high school math teacher, practices creating exam questions by controlling answer time and difficulty level based on the percentage of correct answers to questions for the entire class obtained from the dashboard. He said ‘Until now, when creating test questions, average scores and answer times were estimated "based on teacher experience". By using the analysis dashboard tool, even inexperienced teachers can easily do this and reduce the time required to create test questions’. Inspired by this practical example, we propose an automated test set quiz maker by optimizing time and learning parameters from learning log. The system is expected to be used at schools to create exam questions based on students' learning history (students’ comprehension and time spent answering quizzes) reducing the burden on teachers if the correct answer rate and answer time can be controlled and the optimal set of quizzes can be presented automatically. Students can also use the same feature to practice before an upcoming exam.

2. Related work

2.1 Learning Analytics

Learning Analytics (LA) is an emergent field which aims at a better understanding of students and providing intelligence to learners, teachers, and administrators using learning log data (Law & Liang, 2020). LA has been used in Higher Education for improving the services and students' retention rate (Bienkowski et al., 2012). Although most LA studies were conducted in higher education (Li et al., 2015), LA research has recently been spreading to K12 as well (Aguerrebere et al., 2022).

2.2 Testing effects

The testing effect also known as retrieval practice, active recall, practice testing, or testenhanced learning (Dunlosky et al., 2013; Roediger & Butler, 2011) suggests devoting part of the learning period to retrieving information from memory increases long-term memory (Bruce Goldstein, 2010). For example, Roediger and Karpicke (Roediger & Karpicke, 2006) asked participants to study two sentences, with one sentence studied twice and the other studied once and tested once. One week after a test, the tested sentence was better recalled than the restudied one. The testing effect has been demonstrated in numerous studies using different educational materials in both laboratories and classrooms (Roediger & Karpicke, 2006; Rowland, 2014). Testing is an effective way to learn, but it requires a great deal of teacher knowledge and effort to create a set of quizzes for the test.

In this study to reduce the teachers' burden of creating tests, we propose an automated test set quiz maker by optimizing time and learning parameters from learning log. This research will develop a function by not creating a new quiz one by one but by developing a function that uses mathematical optimization to select the optimal quiz set from already created quizzes. Namely, this function automatically generates the test set of quizzes if the user inputs the desired amount of time it should take to solve the test by optimizing the selection of quizzes based on BKT parameters and estimated solving time.

3. Methods

3.1 System overview

The automated test set quiz maker proposed in this paper was plugged into the recommender system on the LEAF framework. LEAF framework is an infrastructure system 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. The BKT recommender (Takami et al., 2021, 2022) was developed on the LEAF framework and automated test set quiz maker was plugged into the recommender as one of function recommendations.

3.2 Bayesian Knowledge Tracing algorithm

The Bayesian Knowledge Tracing (BKT) algorithm functions as a Hidden Markov Model (HMM) in its traditional form and assumes a student's knowledge as a binary variable showing whether or not a student has mastered a skill (Corbett & Anderson, 1994). This model calculates the probability of a skill at a given point in time by combining data on the student's performance up to that point with model parameters. The guess (giving a correct answer despite not knowing the skill) and slip (knowing a skill, but giving a wrong answer) parameters of BKT for each quiz are calculated from the data of correct and incorrect answers for all quizzes for all students in the relevant course using Python Library of Bayesian Knowledge Tracing Models (Badrinath et al., 2021).

3.3 Problem formalization as a Knapsack Problem

The knapsack problem is a problem in combinatorial optimization. It derives its name from the problem faced by someone who is constrained by a fixed-size knapsack and must fill it with the most valuable items. The knapsack problem has various kinds of practical applications i.e. portfolio optimization, cutting stock problems, scheduling problems (Martello & Toth, 1990) and cryptography (Chor & Rivest, 1988; Goodman & McAuley, 1985; Laih et al., 1989). In a given set of n items, each with a weight w_j and a value p_j, determine the number of each item to include in a collection so that the total weight is less than or equal to a given limit and the total value is as large as possible. the problem may be mathematically modeled as follows:

$$\text{Maximize } \sum_{j=1}^{n} x_j p_j$$

$$\text{Subject to } \sum_{j=1}^{n} w_j x_j \leq C$$

$$x_j \in \{0,1\},\ j = 1,...,n.$$

where x_j takes values either 0 or 1 which represents the selection or rejection of the ith item. We put BKT model parameter or correct rate into a value p_j and quiz solving time into weight w_j in a given set of n quizzes constrained desired time to solve as C. We used python package *pulp* for optimizing.

3.4 Maximizing Guess and Slip value of BKT calculated from all of the course students

First, the guess (giving a correct answer despite not knowing the skill) and slip (knowing a skill, but giving a wrong answer) parameters of BKT for each quiz are calculated from the data of correct and incorrect answers for all quizzes for all students in the relevant course using Python Library of Bayesian Knowledge Tracing Models (Badrinath et al., 2021). The parameters of guess and slip are estimated using all logs of the course every hour and updates them every hour. Therefore, the parameters will continue to be updated until the actual day of the test, and quiz can be selected according to the latest state of understanding just prior to the test.

For guess value, we reported the first quiz in a component in a quiz book 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 (Takami et al., 2021). From this study, we can consider lower guess value quizzes are difficult to be solved requiring basic skills. Therefore, in order to select difficult quizzes, we preprocessed guess value as -log (guess + δ) for a problem with a lower guess value to be constrained by time to be optimized (note: We set δ = 0.001 to avoid an error with a true log of 0 when guess = 0). For slip value, we input the raw slip value to optimize the more careless and error-prone quizzes that are constrained by time.

3.5 Optimization from individual student logs

For the convenience of students, we also introduced a method of optimization that constrains the percentage of correct answers and time for each individual student. One student's learning log was used based on whether he or she solved quizzes or not. We prepared two types of optimization. First, time constraints on his or her solved quizzes with low percentage of correct answers from his or her own learning log. Second, time constraints on his or her unsolved problems with high percentage of correct answers from other students' logs. The former type is expected to be used to effectively review areas of his or her poor performance before an exam, while the latter is expected to be used by students who have not studied much to gain confidence by solving easy problems before an exam.

4. Implementation of prototype into the learning system

4.1 User Interface screenshot

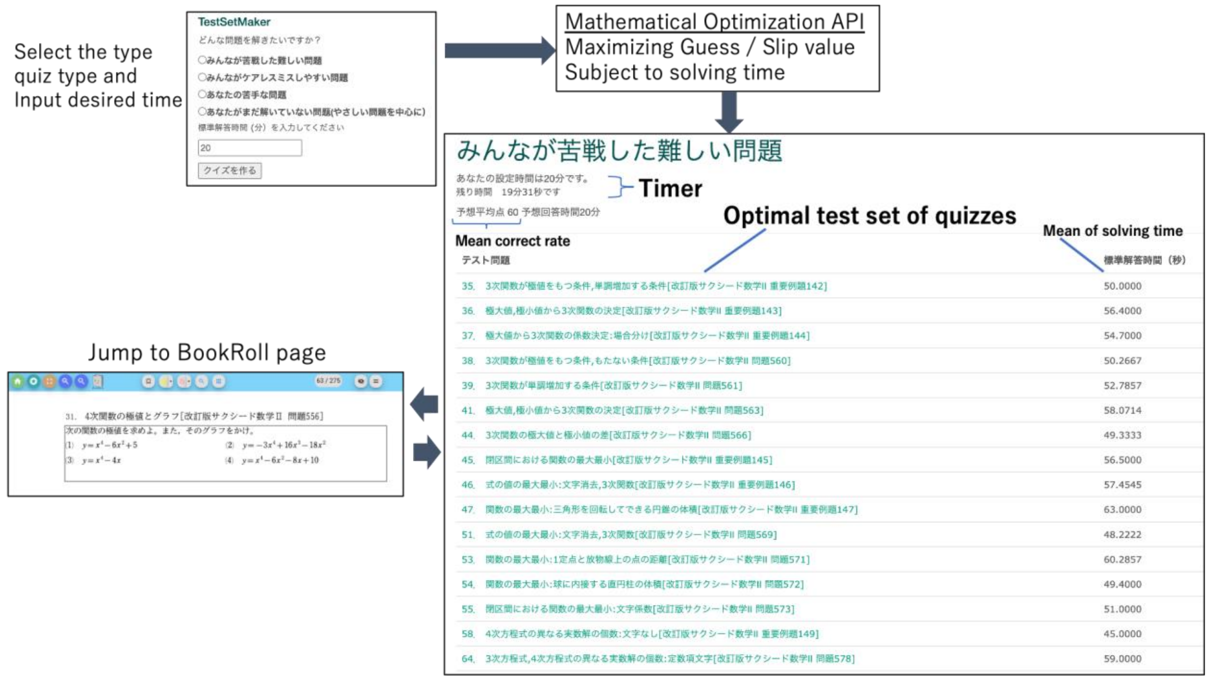

Figure 1. We implemented this mathematical Optimization API to the recommender system which we developed on LEAF framework

We implemented this mathematical Optimization API to the recommender system on the LEAF framework. The student selects the type of quiz to be optimized (BKT parameters or individual correct rate) and enters the time to be constrained, and this information is sent to the optimization API. The API performs the optimization process by constraining guess, slip, or the percentage of correct answers with the solution time, and outputs an optimal test set of quizzes that are optimal for the input time. As shown in Figure 1, optimized outputs test set of quizzes displayed on the page. At the top of the list of quizzes, a countdown timer and mean correct rate of quiz set are displayed. The right column of the optimal test set of quizzes, means of solving time are displayed. If a student clicks the title of the quiz, he or she can jump to the BookRoll page. After solving the quiz on the BookRoll page the student is supposed to return the optimal set of quizzes list page. When the timer reaches 0, a "Time up!" pop-up appears.

5. Conclusion and Future Work

In this study, inspired by a teachers' practical example using an e-learning dashboard for creating exam questions, we proposed an automated test set quiz maker optimizing BKT model parameters and solving time. Future study needed to evaluate whether reducing the burden of creating exams on teachers and the effectiveness on students using this function as a trial test before exam.

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, and NEDO JPNP20006 and JPNP18013.

References

Aguerrebere, C., He, H., Kwet, M., Laakso, M. J., Lang, C., Price-Dennis, C. M. D., & Zhang, H. (2022). Global Perspectives on Learning Analytics in K12 Education. by Charles Lang, Alyssa Friend Wise, Agathe Merceron, Dragan Gaševic, and George Siemens. 2nd ed. Vancouver, Canada: SOLAR.

Badrinath, A., Wang, F., & Pardos, Z. (2021). pyBKT: An Accessible Python Library of Bayesian Knowledge Tracing Models. 468–474.

Bienkowski, Feng, & Means. (2012). Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics: An Issue Brief. Office of Educational Technology, US. https://eric.ed.gov/?id=ED611199

Bruce Goldstein, E. (2010). Cognitive Psychology: Connecting Mind, Research and Everyday Experience. Cengage Learning.

Chor, B., & Rivest, R. L. (1988). A knapsack-type public key cryptosystem based on arithmetic in finite fields. IEEE Transactions on Information Theory / Professional Technical Group on Information Theory, 34(5), 901–909.

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.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. Psychological Science in the Public Interest: A Journal of the American Psychological Society, 14(1), 4–58.

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

Goodman, R. M. F., & McAuley, A. J. (1985). A New Trapdoor Knapsack Public Key Cryptosystem. Advances in Cryptology, 150–158.

Laih, C.-S., Lee, J.-Y., Harn, L., & Su, Y.-K. (1989). Linearly shift knapsack public-key cryptosystem. IEEE Journal on Selected Areas in Communications, 7(4), 534–539.

Law, N., & Liang, L. (2020). A Multilevel Framework and Method for Learning Analytics Integrated Learning Design. Journal of Learning Analytics, 7(3), 98–117.

Li, K. C., Lam, H. K., & Lam, S. S. Y. (2015). A Review of Learning Analytics in Educational Research. Technology in Education. Technology-Mediated Proactive Learning, 173–184.

Majumdar, R., Hiroyuki, K., Komura, K., Flanagan, B., & Ogata, H. (2020, March). LA Platform in Junior High School: Trends of Usage and Student Performance. Companion Proceedings 10th International Conference on Learning Analytics & Knowledge (LAK20).

Martello, S., & Toth, P. (1990). Knapsack problems: algorithms and computer implementations. John Wiley & Sons, Inc.

Roediger, H. L., 3rd, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. Trends in Cognitive Sciences, 15(1), 20–27.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves longterm retention. Psychological Science, 17(3), 249–255.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. Psychological Bulletin, 140(6), 1432–1463.

Takami, K., Dai, Y., Flanagan, B., & Ogata, H. (2022). Educational Explainable Recommender Usage and its Effectiveness in High School Summer Vacation Assignment. LAK22: 12th International. https://dl.acm.org/doi/abs/10.1145/3506860.3506882

Takami, K., Flanagan, B., Dai, Y., & Ogata, H. (2021). Toward Educational Explainable Recommender System: Explanation Generation based on Bayesian Knowledge Tracing Parameters. In 29th International Conference on Computers in Education Conference Proceedings (Vol. 2, pp. 532-537).