Proposal of Quiz-Classification Method in Adaptive Learning System

Kodai ABE[1]*, Haruhiko MITSUNAGA[2],
Hiroto YAMAKAWA[3], and Hiroshi KOMATSUGAWA[1]
[1] Graduate School of Photonics Science, Chitose Institute of Science and Technology, Japan
[2] Graduate School of Education and Human Development, Nagoya University, Japan
[3] Faculty of Science and Technology, Chitose Institute of Science and Technology, Japan * Fax: 81-123-40-2308, e-mail: abe215@kklab.spub.chitose.ac.jp

We propose a quiz-classification method using statistical approach based on LRT considering missing values. Then, we apply LRT to learners' learning results with missing values that are replaced with dummy learning data based on the difficulty degree of the quiz.
Key words: Quiz-Classification, Adaptive Learning System, Latent Rank Theory

1. INTRODUCTION

It is important for learners to acquire knowledge in academic disciplines that require advanced expertise, and necessary to confirm the degree of their knowledge in various situations of lectures. In our previous study (Ueno et al., 2017), we proposed a learning model for a flipped classroom using an adaptive learning system that provided learning quizzes corresponding to learners' understanding degree of knowledge. The system mainly provides two functions, that is, (i) an adaptive test function (ATE) and (ii) an adaptive training function (ATR). The results showed that a learning model using an adaptive learning system positively contributed to the improvement in the learners' degrees of knowledge. In the study, difficulty of the learning quizzes used in the adaptive learning system was not determined through the statistical approach but by the experts. Then, the difficulty was not always consistent with that calculated from learners' actual answers using ATE and ATR. Furthermore, the answer data includes missing values because the learners do not solve some learning quizzes. In this study, we propose a quizclassification method using statistical approach based on Latent Rank Theory (LRT) (Shojima, 2007) considering missing values of learners' answers that are replaced to dummy learning data based on the difficulty of the quiz.

2. LEARNING QUIZZES

Difficulty of the learning quizzes is defined at 7 levels categorized by the expert.
Understanding knowledge and concept is set to learning objective at level 1 and 2. Knowledge utilization is set to level 3, 4 and 5. Application of knowledge is set to level 6 and 7.

3. QUIZ-CLASSIFICATION METHOD

LRT is a clustering algorithm using learners' learning results and can estimate learners' ability and the difficulty of the learning quizzes using SOM-based clustering algorithm. Self-organizing maps (SOM) is an unsupervised machine learning method for non-linear classification problems. However, using this method, we meet problems for lacking experts' instructional policy in which they want to make structured materials for their lesson. To avoid the problem, we introduce experts' policy in the LRT algorithm. Then, we apply LRT to learners' learning results with missing values that are replaced with dummy learning data based on the difficulty degree of the quiz.

3.1. Procedure

The procedure of our proposed classification method is as follows.

I. We prepared correct or incorrect information based on learners' actual answers of each knowledge which is stored in the e-learning system.

II. Initially, missing values have no data. So, Missing data are replaced to dummy data based on the experts' policy. Supplementing these dummy data, we can arrange the complete data set for calculating LRT. (see section 3.2).

III. The difficulty of each learning quizzes using LRT can be also estimated.

IV. Using LRT, the difficulty of each learning quizzes can be determined and selected as a new level.

3.2. Supplement of missing values

I explain the detail how to supplement missing values using figure 1. We generate a probability-based correct rate using logistic model described in Eq.1. This is the image table of learners' answer situation. 0 means the incorrect state. 1 means the correct state. This striped area is the dummy data that was initially no data and replace for missing value. Dummy data was set 0 or 1 based on value of probability determined by this equation. Note that coefficient b is the difficulty degree of quiz determined by experts, constant value θ is the ability value of student. So, the probability-based correct or incorrect state is determined and corresponding to this value.

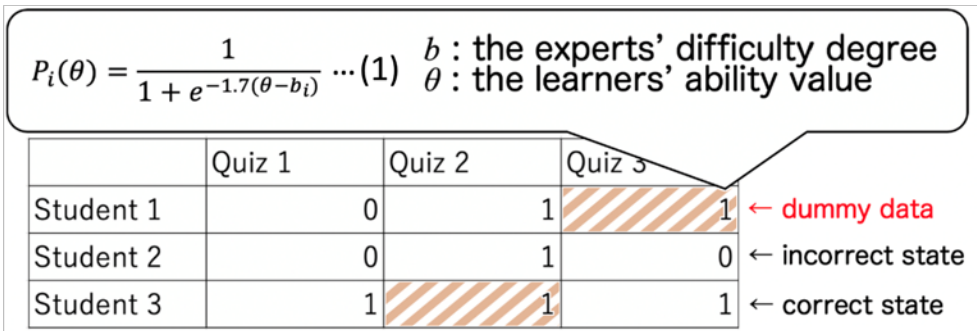

Fig.1. Logistic model and the image table of leaners' answer situation.

4. EVALUATION

We applied our proposed method to the case of a C programming class in 2018 and 2019. The learning quizzes were re-categorized based on answers in 2018 and updated and used in 2019. Updated knowledges are “Variables”, “If”, “Loops” and “Arrays” of C programming. We evaluated effectiveness of our proposed method through two metrics. One is correspondence rate. The other is rate of correct number. We determined the corresponding rate using a ratio of a “priori” level categorized by experts and a “calculated” level obtained from our model. The results of 2018 and 2019 are shown in Table 1.

Table 1. Correspondence rate using ratio of a “priori” level categorized by experts and a “calculated” level obtained from our model

Knowledge	Correspondence rate in 2018	Correspondence rate in 2019
Variables	0.76	0.95
If	0.59	0.92
Loops	0.64	0.89
Arrays	0.91	0.93

As for the next metric, we determined rate of correct numbers of quizzes Based on learners' actual answers. The results of 2018 and 2019 are shown in Figure 2. The vertical line shows the rate of correct numbers of quizzes and the horizontal one shows the level obtained from our model. The black parts represent the rate of correct numbers. The white parts represent the rate of incorrect numbers.

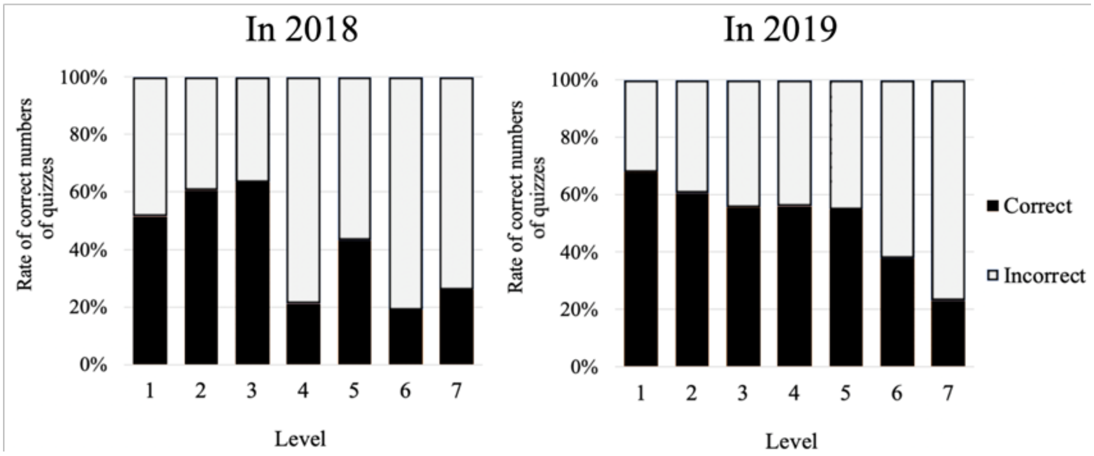

Fig.2. Rate of correct numbers of quizzes based on learners' actual answers.

5. CONCLUSION

At the first metrics, we found that the number of the rate of correspondence of “Variables” increased from 76% to 95%, “If” increased from 59% to 92%, “Loops” increased from 64% to 89% and “Arrays” increased from 91% to 93%. At the second metrics, we found the improvement of the rate of correct of learners’ actual answers between 2018 and 2019. In 2019, we observed the correct answer rate tends to decrease as the difficulty of increasing. However, in 2018, we did not observe same tendency. These results indicate that our proposed method contributes to the improvement of a quiz classification with missing values of learners’ answers.

REFERENCES

Ueno H., Kato T., Yoshida F., Tsukada N., Tateno H., Fukamachi K., Yamakawa H., & Komatsugawa H. (2017). *A Model of Flipped Classroom Using an Adaptive Learning System*. Work-in-Progress Poster, The 25th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education 1-3.

Shojima K. (2007). *Latent rank theory: Estimation of item reference profile by marginal maximum likelihood method with EM algorithm*, DNC Research Note, 07-12.