# Dynamic Evaluation of both Students' and Questions' Levels for Generating Appropriate Exercises to Students Automatically

Suganuma, Akira
Department of Intelligent Systems, Kyushu University

Mine, Tsunenori
Department of Intelligent Systems, Kyushu University

Shoudai, Takayoshi
Department of Intelligent Systems, Kyushu University



出版情報：Proceedings of Knowledge-Based Software Engineering, pp.325-328, 2002-09
バージョン：
権利関係：

# Dynamic Evaluation of both Students' and Questions' Levels for Generating Appropriate Exercises to Students Automatically

Akira SUGANUMA, Tsunenori MINE, and Takayoshi SHOUDAI
*Graduate School of Information Science and Electrical Engineering*
Kyushu University, Kasuga 816-8580, Japan
{suga@is, mine@is, shoudai@i}.kyushu-u.ac.jp

**Abstract.** This paper discusses AEGIS(Automatic Exercise Generator based on the Intelligence of Students) that generates exercises of various difficulty levels according to each student's achievement level, marks his/her answers and returns the marked results to him/her. Guessing the achievement level of each student from his/her trial history, AEGIS selects the most suitable exercise for him/her according to his/her achievement level. It is necessary for AEGIS to evaluate dynamically not only the level but also the difficulty level of the exercises in order that it gives each student a suitable exercise. In this paper, we describe the method to re-estimate them.

## 1 Introduction

As the Internet has come into wide use, WWW environments provide lots of opportunities to various fields. In the educational domain, many Internet technologies enable us to hold lectures using Web contents as a teaching material and even develop new lecture methods using the technologies. Web data are, therefore, being expanded rapidly as useful materials.

We are devoting ourself to develop a Web-based self-teaching system and to build the tools for helping students understand their subjects [4, 6]. Through our experiences teaching in classes and developing such systems, we recognize the necessity of both a method evaluating students' achievement levels and generating exercises suitable for the students automatically. The well-considered exercises are useful not only to measure the achievement level of the students, but also to improve their performance. Unfortunately, it is not an easy task for any teacher to make exercises with the difficulties suited to their achievement levels. Besides, it is very important to mark their answers to the exercises and return the marked results to them for keeping their learning enthusiasm. These tasks become harder in proportion to the number of the students in a class.

In this paper, we present an automatic student's achievement level evaluator that generates exercises in three question-types from one tagged document, presents them to students and marks their answers automatically. We call the system AEGIS(Automatic Exercise Generator based on the Intelligence of Students)[5]. It is necessary for AEGIS to evaluate dynamically both student's achievement level and the difficulty level of exercise. Although many CAI systems have been proposed[1, 2, 3], our system is different from them in the points of reusability of pre-existing electronic materials and re-estimation of both the levels.

## 2 Re-estimation of Achievement Level and Difficulty Level

### *2.1 Achievement Level of Students*

It is very important for AEGIS to estimate an achievement level of each student. AEGIS measures it whenever the student answers a question because such a student level fluctuates constantly. The achievement level of student $i$ at time $t$ is calculated with the following formula:

$$s_{i,t} = \begin{cases} s_{i,t-1} + \dfrac{\sum_{j \in Q} (q_{j,t} - s_{i,t-1}) \delta_{i,j}}{\sum_{j \in Q} \delta_{i,j}} & \text{if } \sum_{j \in Q} \delta_{i,j} \neq 0 \\ s_{i,t-1} & \text{otherwise} \end{cases}$$

where $Q$ is a set of questions that he/she answered in the recent 30 trials and $q_{j,t}$ stands for the difficulty level of question $j$ at the time when the achievement level $s_{i,t}$ is calculated. The value, $\delta_{i,j}$, stands for 1 if student $i$ correctly answered question $j$ whose difficulty level is more than his/her achievement level $s_{i,t-1}$ or he/she incorrectly answered question $j$ whose level is less than $s_{i,t-1}$, or 0 otherwise. The achievement level of student $i$ is initialized to 1 when he/she tries a question at first time.

It is quite natural that a student correctly answers the question whose difficulty level is less than his/her achievement level and vice versa. AEGIS, therefore, neglects these trials. AEGIS increases his/her achievement level when he/she can correctly answer the more difficult question and decreases it when he/she cannot correctly answer the easier question.

### *2.2 Difficulty Level of Hidden Region*

The difficulty level of a question is carefully configured because AEGIS uses it to estimate the students' achievement level as described in section 2.1 and refers it to generate a suitable question for a student. Since the teachers set it up with the attribute LEVEL of DEL tag, which is one of tags we defined to embed the information to generate exercises in teaching documents. However, a question evaluated by the teachers as an easy one may not always be answered correctly by lots of students, and vice versa. AEGIS, therefore, utilizes the value of the attribute LEVEL as an initial value of the difficulty level, and re-estimates the level dynamically at regular intervals with the following formula:

$$q_{j,t} = \begin{cases} q_{j,t-1} + \dfrac{\sum_{i \in S} (s_{i,\tau} - q_{j,t-1}) \xi_{i,j}}{\sum_{i \in S} \xi_{i,j}} & \text{if } \sum_{i \in S} \xi_{i,j} \neq 0 \\ q_{j,t-1} & \text{otherwise} \end{cases}$$

where $S$ is a set of students who answered the question $j$ between time $(t-1)$ and $t$, $s_{i,\tau}$ is a student's achievement level at time $\tau\,(t-1 \leq \tau \leq t)$. The value, $\xi_{i,j}$, stands for 1 if students whose achievement level is more than the difficulty level $q_{j,t-1}$ answered it wrongly or students whose level is less than $q_{j,t-1}$ answered it correctly, or 0 otherwise. The initial difficulty level of question $j$, $q_{j,0}$, is given with the attribute LEVEL of DEL tag by teachers.

Figure 1 illustrates the concept of re-estimating question's difficulty level based on students' achievement level. Point $q_{j,t-1}$ is the difficulty level of question $j$ at time $t-1$. The achievement level of students AEGIS gave question $j$ is distributed in the neighborhood of $q_{j,t-1}$. The distribution seems to be represented by the curve shown in Fig. 1. If a student whose achievement level is less (greater) than $q_{j,t-1}$ answers it correctly (incorrectly), he/she falls into the shaded (dotted) area. Assuming that the difficulty level of the question is overestimated (underestimated) if the area is not empty, AEGIS decreases (increases) the level. An intersection point of the horizontal axis and the line connecting $(M_c, N_w)$ and $(M_w, -N_c)$ finally becomes a new difficulty level of the question, where $M_c$ ($M_w$) is the mean of the achievement level of students who are contained in the shaded (dotted) area, and $N_c$ ($N_w$) is the number of the students.

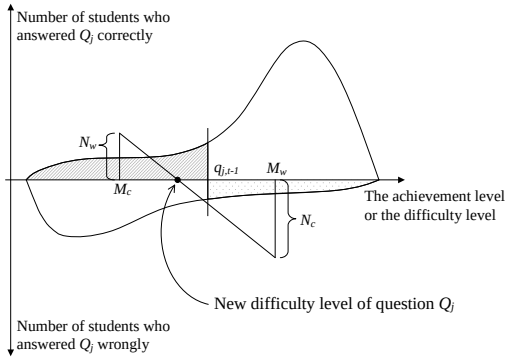



Figure 1: Renewing the difficulty level of a question based on students' achievement level

## 3 Evaluation with Simulator

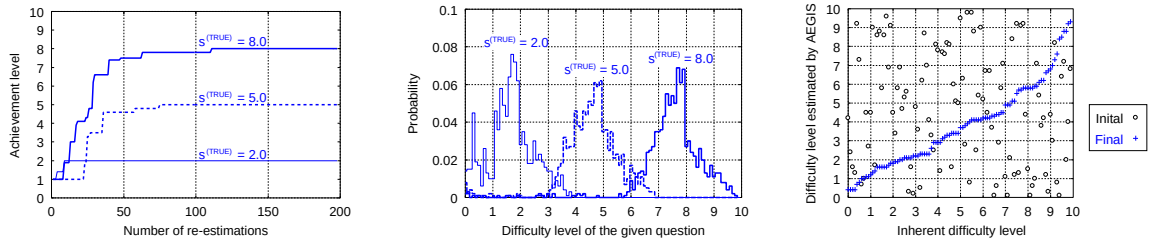
AEGIS estimates dynamically both the achievement level and the difficulty level with the equations defined in Section 2. In order to examine their validities, we experimented with a simulator. We assumed that a student correctly (incorrectly) answered questions whose inherent difficulty level($q^{(TRUE)}$) was less (more) than his/her inherent achievement level($s^{(TRUE)}$) and he/she could correctly answer the question with 50% probability if $q^{(TRUE)}$ was equal to $s^{(TRUE)}$. We prepared 100 questions whose inherent difficulty levels were distributed at the equal interval from 0 to 10. Each difficulty level ($q^{(AEGIS)}$) is initialized by $q^{(TRUE)}$. We also prepared 100 students whose inherent levels are distributed between 0 and 10 at the equal intervals. We investigated the following three things:

(1) How does AEGIS estimate the achievement level of a student?

We pulled out the achievement level ($s^{(AEGIS)}$) of three students whose inherent achievement levels are high ($s^{(TRUE)} = 8.0$), middle ($s^{(TRUE)} = 5.0$) and low ($s^{(TRUE)} = 2.0$). Figure 2-(a) shows the variance of $s^{(AEGIS)}$ in a simulation. They were gradually separated each other because they answered questions correctly or incorrectly based on their inherent achievement level. As was expected, each curve of $s^{(AEGIS)}$ in our simulation approximates closely the value of $s^{(TRUE)}$ after about 50 re-estimations.

(2) Can AEGIS provide only questions suitable for a student?

We applied 1,000 questions of various difficulty levels to these students so as to confirm that AEGIS generates good questions suitable for the students' achievement level. Figure 2-(b) shows the distributions of the difficulty levels of questions which each student tried to solve. A student of high (resp. middle, low) achievement level tried a lot of questions of high (resp. middle, low) difficulty level. Let $(x, y)$ be a pair of the mean value $x$ and the standard deviation $y$ of each distribution. $(x, y)$ of each curve is (1.7, 0.82), (4.6, 1.02) and (7.5, 1.19), respectively. This result shows that AEGIS generates questions suitable for the students' achievement level.




(a) The achievement level

(b) The distribution of the given questions

(c) The distribution of the difficulty level estimated by AEGIS

Figure 2: Results of the evaluation of AEGIS

(3) How does AEGIS estimate the difficulty level of a question?

We made AEGIS calculate the difficulty level ($q^{(AEGIS)}$) which are initialized by a random number. The distribution of $q^{(AEGIS)}$ is shown in Figure 2-(c). The value of $q^{(AEGIS)}$ of the question whose $q^{(TURE)}$ is high (low) increases (decreases) every re-estimation. The points of $q^{(AEGIS)}$ finally place in order of $q^{(TRUE)}$. We conclude that our method can well estimate the inherent difficulty level of each question.

## 4 Conclusion and Further work

AEGIS is consequently utilized as not only a system generating exercises but also a tool classifying questions because the re-estimated level keeps close to their real difficulty level. Experimental results with the simulator showed the effectiveness of the algorithm estimating both the achievement level of a student and the difficulty level of a question as were expected. We have a plan to evaluate this system by applying it to the real courses of Computer Literacy, which are taken by more than 2300 students at Kyushu University. We hope it will work fine as an educational tool for every student and help him/her to understand his/her subjects.

This research was partly supported by Kyushu University, the Grant for Special Academic Research P&P, Type C and supported by the Grants-in-Aid for Scientific Research, Priority Areas Research.

## References

[1] P. Browning, “TUTORIAL MARKUP LANGUAGE - A CBA SYSTEM,” 1998, http://www.soton.ac.uk/ ukgec/workshop/5-cba/minutes.htm#TUTORIAL.

[2] A. Carbone and P. Schendzielorz, “A Web-Based Quiz Generator for Use in Tutorials and Assessment,” Global J. of Engng. Educ., vol.I, no.3, 1997, http://www.eng.monash.edu.au/usicee/gjee/vol1no3/paper20.htm.

[3] ClassBuilder GradeBook And Exam Creation Software, http://www.classbuilder.com.

[4] T. Mine, D. Nagano, K. Baba, T. Shoudai, and S. Hirokawa, “On-web visualizing a mechanism of a single chip computer for computer literacy courses,” Proc. Int. Conf. on Computers in Education, ICCE’98, vol. 2, pp.496–499, 1998.

[5] T. Mine, A. Suganuma, and T. Shoudai, “The Design and Implementation of Automatic Exercise Generator with Tagged Documents based on the Intelligence of Students: AEGIS,” Proc. Int. Conf. on Computers in Education, pp.651-658, 2000.

[6] A. Suganuma, R. Fujimoto, and Y. Tsutsumi, “An WWW-based Supporting System Realizing Cooperative Environment for Classroom Teaching,” Proc. World Conf. on the WWW and Internet, pp.830-831, 2000.