Developing a System to Report Levels of Understanding from Educational Data

Takeshi Morishita
Institute of Education
Shinshu University
Japan
morisita@shinshu-u.ac.jp

Yoshinori Higashibara
Institute of Education
Shinshu University
Japan
higashi@shinshu-u.ac.jp

Abstract: We developed a system to automatically analyze educational data from individual elementary and junior high students and to graphically represent trends in student understanding for teacher use. With this tool, teachers could see the percentage of questions answered correctly and trends in errors by analyzing students' learning records. We found that this can enable teachers to provide appropriate learning support and support individual progress by providing appropriate adaptive learning.

Introduction

Technological advances have advanced the analysis of big data, and the application of data analysis to the social sciences has been recently investigated. Educational Data Mining (EDM) is one such type of big data analysis. EDM is defined by the Journal of Educational Data Mining (JEDM) as "*an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in*."

Bousbia & Belamri (2010) noted that "*EDM objective aim to improve several aspects of educational systems in general and CBLE (Computer Based Learning Environments) in particular. In this specific context, the learner modeling is a key point to accomplish several goals and tasks (tutoring, adaptation, personalization, etc.)*." Children require classroom education appropriate for their understanding and abilities. Therefore, a teacher must be able to measure individual ability precisely. CBLE is useful for this, and adaptive learning, an education method automatically optimized to offer personalized learning contents, can be used to address individuals' error tendencies.

Before the emergence of computers, teachers assessed students' academic skills the best they could, but it was not possible to provide an advice and the problem appropriate to it. Teachers also gave pre-tests and placed students in classes according to their academic abilities and provided the teaching materials most appropriate for every class and changed instruction methods as needed. However, for the requisite condition, including a class unit, and the topic unit of the particle, size was coarse and was far from the ideal that many school teachers had. The emergence of computers changed this situation completely, and education data mining began to attract interest.

Miyahara & Higashibara (2014) argued that there was a limit to the individualized instruction possible in a simultaneous class in the classroom. In elementary schools, a class can advance even if not all students understand the material. However, systems supporting individual learning can automatically provide instruction appropriate to each student's strengths and weaknesses. They can return students' learning records to the teacher, who analyzes them to shape his/her instruction and class structure. This improves students' academic skills and

motivation for learning, showing that using ICT (Information and Communication Technology) is an effective technique for realizing individual learning. However, the analysis of learning records requires statistical and technical knowledge and cannot be done easily by laypeople.

Thus, the purpose of this study is to develop a system to automatically analyze and graphically represent individual students' educational data. This system, for use by teachers in elementary and junior high schools, will show whole-class and individual tendencies in understanding.

System Requirements

Adaptive learning addresses how to adapt to students' needs. Kogo (2013) suggested the following possibilities on his weblog.

1. Content choice according to personal understanding
2. Presentation method according to personal learning preferences
3. Choice of learning method

Criterion 1 eliminates uniform content for all students and mandates subjects and tasks suitable for specific individuals. The criteria are expected to develop from criterion 1 to 2 to 3 in the future, but in this study, we focus on criterion 1. We thought that teacher would like to know the following after class or during class, on realizing criterion 1.

- What questions can students easily answer?
- What questions trigger the most wrong answers? Which require extra attention? What kind of error trends are there?
- What kind of support leads students to correct answers? What kind of support did students come to be able to lead a correct answer through?
- Which questions do students answer most consistently? Which students give the most correct answers?
- Which students have trouble with problems? Which students have an unidentified need for support? Which problems are hardest to answer correctly? Why do students make errors?
- How is the class as a whole progressing? What kind of characteristics define the whole class?

The next function we implemented was to provide the information necessary to allow a teacher to obtain the answers to these questions: grade reports, percentage of questions answered correctly and time needed, and erroneous answer data. This system displays these data for each student or for the whole class and has functions to graphically represent the data as a chart for teacher use.

The Individual Learning Support System

In this study, an elementary school teacher used an individual learning support system called "Interactive STUDY," and students learned individually in the class. Interactive STUDY is a system to support independent learning and the acquisition of underlying and basic knowledge on web (Higashibara *et al.* 2002). This system is designed to provide learning contents and supplementary materials that allow teachers to evaluate students' understanding and learning needs. The system employs computers in class to exploit interpersonal interactions between the student and teacher and avoid simple repetitive learning. The system has design features that encourage students to reach out to friends or teachers if they face any difficulties.

Figure 1 represents the flow of individual learning using Interactive STUDY. This system is aimed at mastery learning. It is programmed to identify learning achievement situations and error trends with the aim of delivering appropriate support.

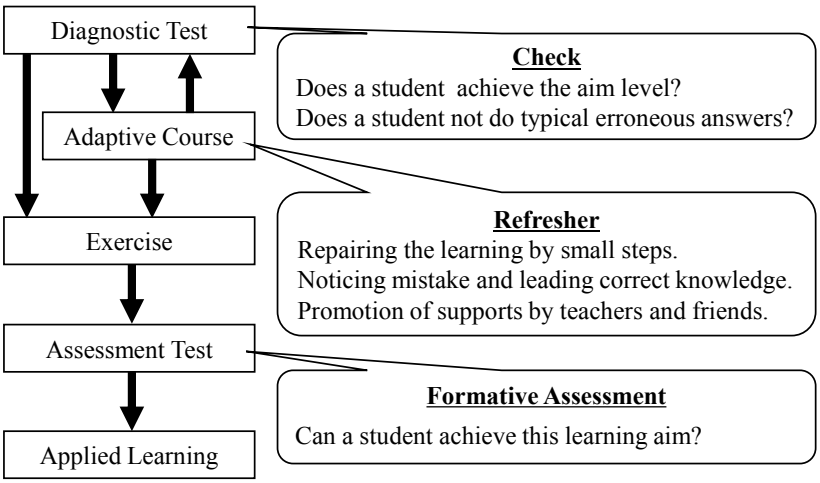

Figure 1: The Learning Flow in Interactive STUDY

First, students answer questions to examine their understanding in a diagnostic test. This test checks whether the student has met the learning goals and does not typically make mistakes. If a student fails to meet the goals or makes more than an acceptable number of errors, he/she moves to an adaptive course shaped to his or her learning patterns.

In adaptive courses, students work on supplementary materials to overcome deficiencies and meet the goals set forth in the diagnostic test. Each student learns by small steps and corrections and is able to arrive at the right answers, since the program notices erroneous answers. Students retake the diagnostic test if they complete the supplementary program, regardless of whether they are able to meet the goal level. They advance to the next exercise if they meet the goal but repeat the adaptive course as many times as necessary if unable to meet the goal. Students passing the diagnostic test advance to the next exercise. They attempt a number of questions depending on their learning goals. Students uncomfortable with the difficulty of the exercise can retake the treatment course before moving to the assessment test.

The student understanding gained in the exercise allows the student to advance to the assessment test. Those who pass the assessment test advance to applied learning and attempt more progressively challenging questions and gain increasing confidence.

This system also compiles a learning record of the types of questions answered by students, how many seconds they took to answer them, and the kinds of answers students provided. The teacher analyzes students' understanding or error tendencies and identifies students needing support. Thus, these learning records make it possible to provide individualized instruction. In this study, we developed a support system to support teachers more effectively using these learning records.

System Summary

The system automatically tallies and records what students learn in Interactive STUDY. Teachers can generate reports of both overall group learning and individual learning after downloading the recorded learning data and a table showing learning activity and student status from Interactive STUDY. The recorded learning data is a record of who answered which questions at which times in the learning activity. The latter student status table contains individual learner records such as a learning frame label on time, the understanding situation (aiming to understand situation, erroneous answer situation, and so on) and the score. It can be used in case of the divergence to the most suitable learning subject and enables teachers to evaluate learner understanding of the material more accurately by graphically representing the whole learning process.

In the following sections, we explain this system with real data.

A. Percentage of Questions Answered Correctly and Average Time

Figure 2 shows total number of those who answered the question ("Headcount"), percentage (%) of questions answered correctly ("Percentage"), and average time ("Avg. Time") in a table listing every question in the teaching materials. This system was originally created in Japanese, but has been translated into English for this paper. We gather some information for non-indication so that it can be easy to see the figure (the same as follows in all figures).

The "C" in the "Answer" category shows the number of correct answers and "Percentage" shows the percentage of questions answered correctly among those who answered the question. The cell is displayed in yellow if the "Percentage" of the question concerned is less than 80%, and in green if fewer than 60%.
The "E" shows the number of erroneous answers, and the "0" and "2" in the "Answer" category represent the number who answered with a previously predicted erroneous answer pattern. The teacher can click the number under "E" to identify students giving wrong answers (Figure 3).

Number	Answer 0	2	C	E	Headcount	Percentage	Avg. Time	Question
1q1102 診断第1問			1		1	100.0	154.0	＜11－80
1q1108 診断第2問			1		1	100.0	7.0	＜11－80
1q1202 診断第1問				2	2	0.0	246.5	まんなかの
1s0102 診断第1問			23	4	27	85.2	66.7	まんなかの
1s0108 診断第2問			22	1	23	95.7	40.7	まんなかの
1s0202 診断第1問	1		25	2	28	89.3	49.8	まんなかの
1s0208 診断第2問			25		25	100.0	29.2	こうえんに
1s0302 診断第1問		1	19	5	25	76.0	48.2	まんなかの
1s0308 診断第2問			15	5	20	75.0	46.3	まんなかの
1s0402 診断第1問			18		18	100.0	36.1	まんなかの
1s0408 診断第2問			16		16	100.0	20.2	まんなかの
(snip)								
Average						90.7	25.5	
Total	1	1	343	29	374	91.7	39.0	

Figure 2: A Sample Grade Report

Number	ID	Name	Answer
1q1202 診断第1問	10	2310	4
	22	2322	4
	22	2322	9
	22	2322	5
	22	2322	10
1s0102 診断第1問	4	2304	58
	4	2304	7
	10	2310	15
	15	2315	12
1s0108 診断第2問	22	2322	14
1s0202 診断第1問	6	2306	225

Figure 3: A List of Incorrect Answer

B. The Individual Learning Process

Thus, teachers can track the exact learning process of each student. Figure 4 graphically represents a student's learning process. In the figure, the check marks "✓" show questions answered correctly on the first attempt and the dagger marks "†" show those answered correctly on the second attempt. The minus signs "－" show those the student never answered correctly, and the asterisk marks "*" show when the student moved to the Adaptive Course. In addition, the forming up in ranks expresses temporal axes.

For example, in Figure 4, the student was able to correctly answer questions No. 1s0102 and No. 1s0108 immediately, but missed No. 1s0202. The contents of the questions change as the fourth column changes, and teachers know the specific places where students make mistakes. Therefore, where the asterisk marks "*" appears, the system moved the student to the Adaptive Course to definitively correct any misunderstanding of the material. The student then answered the next question, No. 1s0302, correctly the first time after the Adaptive Course, but missed No. 1s0308 and moved to the Adaptive Course again. The student then answered No. 1s0608 correctly within the class period.

Figure 4 also shows the learning process in chronological order, allowing the teacher to see how many times the student retook the Adaptive Course. Figure 5 shows one such case. The table on the left side shows the

Number	Question										
1s0102	診まんなかの	✓									
1s0108	診まんなかの		✓								
1s0202	診まんなかの			－	*						
1s0302	診まんなかの					✓					
1s0308	診まんなかの						－	*			
1s0402	診まんなかの								✓		
1s0602	診こどもが な									✓	
1s0608	診バスていにて										†

Figure 4: An Example of a Student's Learning Process

The Student of Figure 4

	Next	Easy	Adaptive
A	3		
B			
C			1
D			
E			1

Another Student

	Next	Easy	Adaptive
A	2		1
B			
C			
D		1	
E		1	

Figure 5: Examples of Learning Process Lists

learning process of the student in Figure 4, allowing the teacher to see immediately that the student moved to the Adaptive Course twice. The right side shows an example of a different student and shows the teacher that the student attempted the easier question twice in the learning process for review and became more confident. Thus, if a student needs attention, the teacher can understand in detail his/her learning process over time as well as the general pattern of learning processes for all students.

C. Individual Students' Answers

Teachers can see answer patterns for every student in this system. Figure 6 shows the answer pattern for the student in Figure 4. It allows teachers to see the time (sec.) required for each answer, true or false (T/F) answers, and which questions the student missed.

This also allows the teacher to compare individual students' performance with the answer patterns of the whole class. Thus, the teacher can see what materials are appropriate for the whole class or for any particular student by reviewing the percentage of questions answered correctly by the whole class. In addition, "Time Required in Class" shows the ratio of students' answer times to the average time for the whole class. When this ratio is under 1.0, the teacher can see that the student answered the question in less than the class average. The teacher can also identify students in need of support from the time they need to answer questions.

Number	Question	Student Answer	T/F	Correct Answer	% of Class	Time (sec.)	Time Required in Class
1s0102	真まんなかの	13	T		85.2	31.0	0.5
1s0108	真まんなかの	13	T		95.7	13.0	0.3
1s0202	真まんなかの	5	F	4	89.3	91.0	1.8
1s0302	真まんなかの	8	T		76.0	62.0	1.3
1s0308	真まんなかの	7	F	8	75.0	83.0	1.8
1s0402	真まんなかの	16	T		100.0	14.0	0.4
1s0602	真こどもが な	15	T		100.0	46.0	1.6
1s0608	真バスていに	15	F	17	91.7	31.0	1.5

Figure 6: An Example of a Student's Answer Pattern

Trial Practice

We used this system to analyze arithmetic learning in an elementary school. Twenty-five second graders in a class learned to add and subtract on Interactive STUDY, and we analyzed the percentage of correctly answered questions, the learning time, and the learning process for each student.

Figure 2 shows the percentage of questions answered correctly and the average time needed for each student in the class. The students started from question No. 1s0102 in the teaching materials. Questions No. 1q1102, No. 1q1108, and No. 1q1202 were given to the students for independent review of previous learning. No student answered question No. 1q1202 correctly, and student 2322 attempted it four times, but failed to answer it each time, as shown in Figure 3. Thus, we found it necessary for the teacher to provide careful, face-to-face individual guidance to this student.

Next, we focused on question No. 1s0302, which was answered correctly less than 80% of the time, as seen in Figure 2. Question No. 1s0302 was "Twelve people got on a bus. Seven people left at a stop and three people got on. How many people are on the bus now?" We can click on the number of incorrect answers to see the number and kind of errors made. Figure 7 shows the results. Here, one student belongs to Answer Category 2, which indicates that the teacher had predicted incorrect answers for all subtraction problems.

Other erroneous answers fell into Answer Category E, those not predicted by the teacher. Students 16 and 24 were wrong repeated by the same question. Figure 8 shows the learning process for Student 16, which shows he/she reviewed the previous question and learned in the Adaptive Course, but was still unable to answer correctly. We see from this data that he/she answered simple addition and subtraction problems correctly, so the problem was that he/she was unable to handle mixed addition and subtraction problems. This revealed that the teacher will need to offer appropriate individual instruction to this student in the future. Alternately, if the teacher had noticed the problem in class, he or she could have the four students who missed this question work together on the answer.

Number	Answer Category	Answer	Headcount	Student ID
1s0302 診断第1問	C	8	19	
	2	2	1	2
	E	15	2	24, 24
	E	10	1	16
	E	5	1	16
	E	6	1	25

Figure 7: The Breakdown of Answers for Question No. 1s0302

Number	Question									
1s0102	影5、2桁－1桁	✓								
1s0108	影5、2桁－1桁		✓							
1s0202	影5、2桁－1桁			✓			✓			
1s0208	影6、2量の合計				✓			✓	*	
1s0302	影5、2桁－1桁					†				−

Figure 8: The Learning Process for Student 16

Conclusion and Implications

We developed a system to automatically analyze individual learners' educational data and to graphically represent response patterns for both whole classes and individuals. From this trial system, we were able to interpret response patterns from students and give teachers the ability to identify students struggling with the material so that individual adaptive instruction can be provided.

Miyahara & Higashibara (op. cit.) improved students' learning results by providing learning support based on the analysis of learning records. The goal of this current project is to improve their system for easier teacher use, allowing instant visual comprehension of the data.

Acknowledgements

We would like to thank TEK Software Co., Ltd for providing the technical support to realize our study and Hirayama elementary school in Hino-city, Tokyo, for trial practice. This study was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant-in-Aid for Young Scientists (B) 15K21249.

References and Relevant URLs

Bousbia, N. & Belamri, I. (2010). Which Contribution Does EDM Provide to Computer-Based Learning Environments? In Pena-Ayala, A. (Ed.), *Educational Data Mining, Studies in Computational Intelligence 524* (pp.3-28). Switzerland: Springer International Publishing.

Miyahara, A. & Higashibara, Y. (2014). ICT Assisted Academic Achievement–Using Individual Assisted Learning System for Math–. *Studies on Educational Practices*. No.15 (pp.103-110), in Japanese.

Kogo, C. (2013). Adaptive Learning. Kogo Lab. https://kogolab.wordpress.com/2013/05/19/%E3%82%A2%E3%83%80%E3%83%97%E3%83%86%E3%82%A3%E3%83%96%E3%83%A9%E3%83%BC%E3%83%8B%E3%83%B3%E3%82%B0/ (visited on March 24, 2016), in Japanese.

Higashibara, Y., Igarashi, K. & Nishio, Y. (2002). The Development of Web Based CAI System 'Interactive STUDY'. *The Bulletin of Center of Educational Research and Teaching*. No.3 (pp. 41-50), in Japanese.