A CAI SYSTEM TO IDENTIFY WEAK PARTS OF A LEARNER: ON THE NUMBERS OF CATEGORY LAYERS AND QUESTIONS IN A SET OF QUESTIONS

Fuminori Hirose¹, Kazuhiro Uenosono² and Seiichi Komiya¹
¹ Shibaura Institute of Technology, 3-9-14, shibaura, Minato-ku, Tokyo 108-8548, Japan
{skomiya, ma11134}@shibaura-it.ac.jp
² Aoyama Gakuin University, 5-10-1 Fuchinobe, chu-o-ku, Sagamihara-shi, Kanagawa, Japan
k-uenosono@aim.aoyama.ac.jp

ABSTRACT
A teaching method which a teacher gives a lecture to a number of students in a lump, that is, “reception learning” is an effective way to give many students much knowledge in short period of time. However, even if there are great differences among many students in the level of understanding, the teaching method can not fill the knowledge gap [1]. For this reason, the authors are proposing an individual learning system using a CAI (Computer Aided Instruction) system in order for the students to study more effectively to overcome their weak parts of study. However, since the existing CAI systems have no concept of “categories” to divide a learning field into several parts of study in the Student Model, the students may have to study the parts where they already have enough knowledge. Therefore, It is able to say that the students could not learn efficiently by using the past CAI systems.

This led us to propose a CAI system [18] that each student only has to learn only his/her weak parts by automatically identifying the weak parts. In our CAI system, if a learning field is decided, it is able to naturally determine the number of category layers of questions in the learning field. However, in the past it was not able to clarify the appropriate number of questions which compose a set of questions on each category layer. Therefore, we conducted a series of simulations to solve this issue.

This paper consists of the following sections: Section 2 describes the mechanism of our CAI system to identify weak parts of a learner. Section 3 describes the details of the simulations used to find out the appropriate number of questions which compose a set of questions on each category layer, and the results of these simulations. Section 4 introduces related works. Section 5 describes the conclusion.

2. Mechanism to Identify Weak Parts of Study

2.1 Definition of Target Study Areas and Weak Parts of Study

In order to enhance learning effects of each learner, we build such a CAI system that the system makes the learner do an intense study of his/her weak parts. That is to say, our goal is to build a new CAI system which each student only has to learn only his/her weak parts by identifying weak parts of a learner automatically.

To define weak parts of study, we introduce the concept of “category” which presents classification of a learning field, and divides the field into several learning areas (i.e. categories), and hierarchically and iteratively divides each area (i.e. category) into several areas (i.e. categories) based on whole-part relationship. We assume that the results from the hierarchical categorization for areas of...
study based on the whole-part relationship can always be represented in tree structures. In contrast, when the results from the hierarchical categorization based on the whole-part relationship cannot be described in tree structures, such areas are not considered as areas of study.

Figure 1 shows an example of a hierarchically-categorized learning subject called “Information Processing Technology” based on whole-part relationship by using “IT Common Body of Knowledge” and “Fundamental Information Technology Engineer Know-How Body of Knowledge and Core Body of Knowledge” for Skill Standards for IT Professionals [2] as references.

The following is our definition of an area where a student is weak at. Suppose that a learning subject has been categorized hierarchically based on the whole-part relationship. In this case, a certain number of easy questions (for example, five questions) selected from a category are prepared and logically integrated into one screen. A student is given this set of questions and if all the answers the student produced are incorrect, the lowest level of the category among the questions that the student answered incorrectly is defined as an area where the student is weak at. For example, suppose that a student is provided a certain number of questions selected from the “Binary” category in category layer 3 on Figure 1 and logically integrated into one screen. If the student fails to answer all questions shown on the screen, we regard the “Binary” category as a weak category (i.e. weak part).

This type of category is referred to as “a weak category.”

Figure 1. An example of the category layers based on the whole-part relationship of learning subjects

### 2.2 Strategic Knowledge and Instruction Programs for Identifying a Weak category

#### 2.2.1 An example of screen prepared for executing the CAI in Drill & Practice mode and its processing flow

In this system, a depth-first search which prioritizes depth is adopted to identify one’s weak category. If no weak category is detected when searching through end nodes in a general depth-first search, searching will be continued towards the upper layer direction. In the process of identifying one’s weak category, however, if no weak category is detected by searching through end nodes, it is considered as a failure of detecting a weak category, and any further searching operation must be cancelled. It is because that the searching towards the upper layer direction after the searching through end nodes will not be based on the hierarchically-categorized manner between the whole area and part of areas of study if the relationship between the whole area and part of areas of study is accurately broken down into a hierarchy. For this reason, we apply a breadth-first search to pin down the location of a weak category before conducting a depth-first search. To detect a category (i.e., a basic route for searching) of the category layer 1 which is likely to contain a weak category in the category layer 1 or its lower layers, perform a breadth-first search only on the category layer 1. The breadth-first search is used only when we detect a basic route for searching. A question screen (Decision Frame) dedicated for the breadth-first search is referred to as Type I. Figure 2 shows a Decision Frame for Type I. A Decision Frame for Type I is a logically integrated screen which shows a set of questions. Each question is chosen from each category belonging to the category layer 1, and the status of each category does not violate the ordering restriction for questions.

After retrieving a basic route (a category of category layer 1) for search by using a Decision Frame for Type I, we assume that a weak category exists in the retrieved category or its subordinate category and perform a depth-first search on these categories. A Decision Frame used for this operation is referred to as Type II. Figure 3 shows a Decision Frame for Type II. A Decision Frame for Type II is a logically integrated screen which focuses on a category layer, and shows a certain number of questions (for example, five questions) chosen from the same category of that category layer. A Decision Frame for Type II can be used in the following two situations:

1. A situation where you need to check whether a category in a focused category layer is a weak category or not.
2. A situation where you need to check whether a weak category has been remediated or not.

When a category in a category layer is identified as a weak category, appropriate knowledge or information for remediation will be presented to the student since it is considered a lack of his or her knowledge or information required to work on questions in that category. Figure 4 shows an example screen for knowledge or information (Text Frame) provided for remediation.

On a Text Frame, a set of questions is provided to measure his or her comprehension of the weak category. Those questions are the ones that have never been on the test, and the difficulty is the same as those previously on the test. If the student got all the answers right, that category is considered fully comprehended.
In this system, two types of Decision Frames, Type I and Type II, are provided for students to identify their weak categories. Once a weak category is identified, a Text Frame is presented to provide supplemental knowledge required to answer questions belonging to the weak category. Since this system is a CAI system that works in Drill & Practice mode, Text Frames will be presented after some Decision Frames. On the other hand, for a CAI system in Tutorial mode, Decision Frames will be presented after some Text Frames. Therefore, in Drill & Practice mode, not only the indication order of Decision Frames and Text Frames is reversed, but also the indication order of material screens, main part of the system, is reversed.

2.2.2 Finding a category to be set as the starting point for weak category detection

The search method we adopted takes the following steps to find a basic route for identifying a weak category possibly included in a subordinate position of the category layer 1:

[1] Select one question on each category chosen from the category of category layer 1, and list these questions on a Decision Frame for Type I.

[2] A student answers all the questions provided. Based on the results of the test, either one of the following processes will be taken:

(2-1) If a student got all the answers right, it is safe to say that the category is considered fully comprehended by the student. In this case, you may back to (1).

(2-2) If the student fails to answer even one question in the test, we assume that he/she has a weak category in the category layer 1 where the failed question is included or its subordinate position and choose a category for the category layer 1 where the failed question belongs to. Then, we proceed to the process for narrowing down choices for a weak category.

In the process above, the breadth-first search is only applied to the category layer 1 described in Section 2.2.2.

2.2.3 Strategic knowledge for identifying a weak category

Once we have selected a category as the starting point, we give the student a certain number of questions chosen from that category and narrow down choices of categories to pin down a weak category by analyzing the answers produced from the student. When giving a student a certain number of questions chosen from a category, the results of the test are divided into following three cases of [S1], [S2] and [S3]. We have said as follows in [22] about strategic knowledge ([1], [2] and [3]).

[S1] All the answers are incorrect. In this case, the strategic knowledge indicates that this category is a weak category.

[S2] All the answers are correct. In this case, the strategic knowledge indicates that this category is not a weak category but is a category understood by the student.

[S3] Some answers are correct and the others are not. In this case, the current category is still too broad to figure out a weak category. We need to go deeper to consider a more specific category layer. When going one category layer down from the starting point, set the categories including questions which the student gave incorrect answers as B1, B2, ⋯, Bm, and set the categories including questions which the student gave correct answers as G1, G2, ⋯, Gm. Then the strategic knowledge of [S3] can be expressed as follows. It is possible to say that the student’s weak category may be included in categories of B1, B2, ⋯, Bm or in their subordinate categories. Furthermore, categories of G1, G2, ⋯, Gm and their subordinate categories may not include any weak categories. In case of [S3], the top level of a category
which may include a weak category for the student can be expressed by using the arithmetic mean of set theory “∪” which means “or” in mathematical field. With “∪”, the category is shown as $B_1∪B_2∪...∪B_m = G_1∪G_2∪...∪G_n$. Therefore, the strategic knowledge of $[S3-1]$ and $[S3-2]$ can be delivered from the strategic knowledge of $[S3]$. $[S3-1]$ When $B_1∪B_2∪...∪B_m = G_1∪G_2∪...∪G_n$ is not an empty set, if $B_1∪B_2∪...∪B_m = G_1∪G_2∪...∪G_n = \{C_i, C_2, \cdots, C_p\}$, we assume that weak categories may exist in subordinate categories under each of $C_1, C_2, \cdots, C_p$. $[S3-2]$ When $\{B_1∪B_2∪...∪B_m\} ∩ \{G_1∪G_2∪...∪G_n\}$ is not an empty set, if $\{B_1∪B_2∪...∪B_m\} ∩ \{G_1∪G_2∪...∪G_n\} = \{D_1, D_2, \cdots, D_q\}$, we assume that weak categories and other categories may co-exist in subordinate categories under each of $D_1, D_2, \cdots, D_q$. Therefore, we need to analyze each of $D_1, D_2, \cdots, D_q$ by going one category layer down from the starting point. The strategic knowledge $[S3-2]$ is the knowledge about subordinate categories of $[S3-1]$. Hence, $[S3-2]$ will be applied after applying the strategic knowledge of $[S3-1]$.

### 2.3 Difficulty Level of Questions

Before actually adopting a group of questions to the CAI system, these questions need to be evaluated to determine if they are appropriate or not. To do so, we conduct a paper-based test to a group of students who are prospective users of the CAI system or to a group of students similar to those prospective users. The difficulty level of a question is computed by multiplying the incorrect rate for an appropriate question for the CAI system (in unit of “%”) by a numerical value of “100.” That is, the higher the incorrect rate for a question, the higher the difficulty level.

### 2.4 A Set of Questions

To create a set of questions, classify questions into ten (10) ranks based on their difficulty, and collect those questions with the same level and put them into a group. In this phase, the following two conditions must be met simultaneously:

(1) Minimize the variability of difficulty levels between sets of questions.

(2) Minimize the variability of difficulty levels for questions belonging to a set of questions.

Above two conditions (1) and (2) must be met by a set of questions. The reasons are as follows:

(1) The reason why the condition (C1) is required to be met

We need to prepare a set of questions chosen from a group of questions belonging to the same rank of the category in order to minimize the variability of difficulty levels. It is because we need to use an appropriate set of questions to prevent an inappropriate category to be chosen as a weak category which can be identified as a weak category at one point and not at another point. For example, there are two students, A and B, whose comprehension levels for a category are the same. For those two students, the CAI system provides some difficult questions to Student A and some easy questions to Student B. As a result, although these two students may have a good understanding of the category, that category is identified as a weak category for Student A and not for Student B. Therefore, we need to prepare questions that can minimize the variability of difficulty levels.

(2) The reason why the condition (C2) is required to be met

When the variability of difficulty levels for questions belonging to a set of questions gets bigger, that set will be more vulnerable to inappropriate questions of which difficulties vary. Consequently, the probability that a student gets all the answers correct or incorrect will be lower, because difficulty levels in a set of questions are not equivalent. It leads to a situation where the process to identify a weak category or the process to check the progress of remediation for a weak category might be repeated indefinitely. Therefore, we need to minimize the variability of difficulty levels for questions composing a set of questions. In terms of the magnitude of the variability of difficulty levels for the conditions C1 and C2, the variability in C2 has to be smaller than that in C1. It is because if the condition C2 is not met, the process might be repeated indefinitely. In case of failing to meet the condition C1, it is still possible to complete the process. Therefore, it is required to minimize the variability of difficulty levels between questions belonging to a set of questions.

### 3. Relationship between the Number of Questions in a Set of Questions and the Number of Required Category Layers

The probability that a student answers all the questions correctly or incorrectly can be computed by the number of questions composing a set of questions. For example, if a set of questions consists of three questions of which difficulty levels are 50, the probability that a student answers all the questions correctly or incorrectly will be 12.5% calculated by $(1/2)^3$. For a set of questions composed of four questions, the probability will be 6.25% by calculating $(1/2)^4$. Likewise, for a set of questions composed of five questions, the probability will be 3.125% by calculating $(1/2)^5$.

On the other hand, the number of category layers that areas of study can be broken down is determined by areas themselves. Since the number of category layers is determined by areas of study, it is necessary to clarify the number of questions to compose a set of questions in order to identify a weak category by searching through end nodes. In other words, the number of category layers required to identify a weak category by searching through end nodes needs to be clarified based on the number of questions composing a set of questions. This is because the number of category layers required to identify a weak
category is proportional to the number of questions composing a set of questions. The more questions are included in a set of questions, the more hierarchical category layers are required. We conducted a series of simulations to reveal the relationship between the number of questions composing a set of questions and the number of category layers required to identify a weak category. In the simulations, several different levels of difficulties were applied to questions.

Note that the process to check the remediation for a weak category is not the process to find a category, and has nothing to do with the relationship between the number of questions composing a set of questions and the number of category layers.

Table 1 shows the probability of identifying a weak category when the difficulty levels of all questions are set to 50.

<table>
<thead>
<tr>
<th>Number of questions</th>
<th>Probability of identifying a weak category</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>12.5% ((1/2)^3)</td>
</tr>
<tr>
<td>4</td>
<td>6.25% ((1/2)^4)</td>
</tr>
<tr>
<td>5</td>
<td>3.125% ((1/2)^5)</td>
</tr>
</tbody>
</table>

3.1 Details of Simulations

Table 2 shows the details of our simulations. In our simulations, the number of category layers prepared for learning subjects was set to unlimited. The purpose of these simulations is to study the number of category layers required to identify a weak category when the numbers of questions composing sets of questions are set to 3, 4, and 5, and the difficulty levels for those questions are set to 50, 60, and 70, respectively. The number of students being tested was supposed to be one million. Figure 6 shows the process flow of the simulations. First, we selected a difficulty level for questions and the number of questions composing a set of questions. In this procedure, the correctness of answers from students was supposed to be determined by the rates of correct answers, and we generated random numbers that represent the answers from students. For example, when the difficulty level was set to 70, the answers were judged as incorrect if a generated random number was less than 70, and judged as correct if its value was greater than or equal to 70. We generated random numbers as many as the number of questions composing a set of questions, and judged if each of the random numbers was correct or not based on its value. If the value was greater than or equal to 70, it was correct. Otherwise it was incorrect.

When both correct and incorrect answers co-existed in a category, we went one category layer down from that point. When all the answers are judged as incorrect in a category, the evaluated category is identified as a weak category, and when all the answers are judged as correct in a category, the evaluated category is identified as a remediated category.

When the difficulty level is set to 50, an answer is judged as incorrect if the value of random number is less than 60, and judged as correct if it is equal to or greater than 60. Since the number of students being tested was supposed to be one million, we conducted a series of simulations one million times. The program used in those simulations was written in java, and the number of steps reached 57.

3.2 Results of Simulations

The results of the simulations are shown in Tables 3 through 5 for the difficulty levels of 50, 60 and 70, respectively. We calculated the average number of category layers, and the maximum and minimum depths of categories required to identify a weak category or a strong category (a student got all the answers correct) by using the results of simulations. The minimum number of category required to identify a weak category was one \((1)\) in any conditions.

When the difficulty level was set to 50 for a set of three questions, the average number of category layers checked for identification was 4.505 and the maximum depth was 46. For a set of 4 questions, the average number was 8.505 and the maximum depth was 102. For a set of five questions, the average number was 16.480 and the maximum depth was 220.

When the difficulty level was set to 60 for a set of three questions, the average number was 3.911 and the maximum depth was 43. For a set of four questions, the average number was 6.711 and the maximum depth was 94. For a set of five questions, the average number was 11.512 and the maximum depth was 129.

When the difficulty level was set to 70 for a set of three questions, the average number was 2.814 and the maximum depth was 29. For a set of four questions, the average number was 4.505 and the maximum depth was 46. For a set of five questions, the average number was 8.504 and the maximum depth was 102.

Table 3 shows the number of questions in a set of questions and the appropriate number of category layers.

<table>
<thead>
<tr>
<th>Number of questions</th>
<th>Average number of category layers</th>
<th>Maximum number of category layers</th>
<th>Minimum number of category layers</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4.505098</td>
<td>46</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>8.504702</td>
<td>102</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>16.479518</td>
<td>220</td>
<td>1</td>
</tr>
</tbody>
</table>
questions are assigned to a certain number of students, score table [3] or the item response theory [4].

examinations of a series of simulations above. three questions with the difficulty set to 50 from a weak category is 4.5 when a set of questions contains by (1/2)^3. Therefore, the probability that a student gets all three questions, the probability will be 12.50% calculated by (1/2) 5, and when a set of questions contains five questions, the probability that a student gets all the answers incorrect will be 3.125% questions contains. For example, when a set of test with more questions. For example, when a set of may have less chance to get all the answers incorrect on a weak category. As a result, the higher the difficulty of identifying a weak category will decrease since a student needs to composed by three questions since a student may have less chance to get all the answers incorrect on a test with more questions. For example, when a set of questions contains five questions, the probability that a student gets all the answers incorrect will increase in a set which contains fewer questions.

It already has been revealed that a category layer can be broken down into five layers when we use Fundamental Information Technology Engineer Examination as an area of study. Therefore, a set of questions presented to students needs to be composed by three questions since the average number of category layers needed to identify a weak category is 4.5 when a set of questions contains three questions with the difficulty set to 50 from examinations of a series of simulations above.

### 3.3 Considerations on the Simulation Results

This section describes considerations on the simulation results. Based on the results shown in Tables 4 through 6, it can be observed that the higher the difficulty levels of questions, the lower the average number and the maximum number of category layers required to identify a weak category. As a result, the higher the difficulty of questions, the more depthless category layer where a weak category can be identified. This is because when the number of questions in a set increases, the probability of identifying a weak category will decrease since a student may have less chance to get all the answers incorrect on a test with more questions. For example, when a set of questions contains five questions, the probability that a student gets all the answers incorrect will be 3.125% calculated by (1/2)^5, and when a set of questions contains three questions, the probability will be 12.50% calculated by (1/2)^3. Therefore, the probability that a student gets all the answers incorrect will increase in a set which contains fewer questions.

Table 4. The number of questions in a set of questions and the appropriate number of category layers (When the difficulty level is set to 60)

<table>
<thead>
<tr>
<th>Number of questions</th>
<th>Average number of category layers</th>
<th>Maximum number of category layers</th>
<th>Minimum number of category layers</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.910600</td>
<td>43</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>6.710848</td>
<td>94</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>11.511690</td>
<td>129</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5. The number of questions in a set of questions and the appropriate number of category layers (When the difficulty level is set to 70)

<table>
<thead>
<tr>
<th>Number of questions</th>
<th>Average number of category layers</th>
<th>Maximum number of category layers</th>
<th>Minimum number of category layers</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2.81393</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4.07765</td>
<td>38</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>5.89213</td>
<td>72</td>
<td>1</td>
</tr>
</tbody>
</table>

### 4. Related Studies

#### 4.1 Studies on Evaluation Methods of Difficulty Levels for Students

Several analysis methods for test results have been proposed to evaluate difficulty levels for students, including the methods using the Student-Problem (S-P) score table [3] or the item response theory [4].

In the S-P table, a certain number of “true or false” questions are assigned to a certain number of students, and results are listed in the S-P table, with the entries are one (1) for correct answers and zero (0) for incorrect answers. After calculating the scores, those students and questions are sorted in the score table in the descending order of score and rate of correct answers. Based on the data in this table, we can obtain two separate curves. One is a curve reflecting students’ cumulative distribution of total scores, called S-curve. The other is a curve representing questions’ cumulative distribution of the number of students who answer questions correctly, called P-curve. In addition, an area surrounded by those two curves (amount of differences), called D, can be obtained. When the S-curve and the P-curve are overlapped, the amount of variance will be minimized. This indicates that questions are homogeneous and their difficulty levels are low. On the other hand, when the amount of variance increases, questions become heterogeneous, and their difficulty levels get higher.

Kuwabara [5] proposed a method to assess the difficulty levels of study materials by measuring the following three elements: the amount of time students spent to complete assigned issues, the number of questions students asked in the course of the learning process, and students’ deductive inference capability. However, other than Object-C, a study material of computer languages for beginners, it has not clarified yet any other types of study materials that Kuwabara’s method can be applied effectively.

#### 4.2 Systems to Automatically Generate Questions (with the concept of difficulty for questions)

According to Yamashita [6], the research for automatically generating a set of questions originated at P. Suppes [20]. In his research, questions can be automatically generated with 14 different levels of difficulties by using a structure called “Stand structure.” However, there is no concept or mechanism of the student model in his CAI system so the system cannot automatically figure out the comprehension level of each student. In his system, answer records of students are analyzed by a human researcher, and the comprehension level of each student is sorted into 14 levels to automatically generate appropriate questions for each student.

In addition, researchers including Beulen [7] and C. Brown[12] have been studying automatic generation of questions with the concept of difficulty level for questions.

#### 4.3 Systems to Automatically Generate Questions (without any concept of difficulty level for questions)

A system to automatically generate questions for students by using a computer is called generative courseware. The mechanisms for automatically generating questions without any concept of difficulty for questions are proposed by the following studies: Kunichika [11], Hoshino [8], Cai [9], Edmond [10], Lazcorreta [13].
4.4 Systems to Automatically Generate Questions in Accordance with Students’ Comprehension Level

Tsumori [14] narrowed the area of study for the CAI system down into IT Passport Test to focus on vocabulary used in that area, and implemented the automatic generation system of appropriate questions for a student whose comprehension level was figured out during the learning process by introducing the concept of distance between vocabularies based on their similarity. Suganuma [15] has developed a system to automatically generate questions in three different forms, which are the alternative form, fill-in-the-blank form and error-correcting form, based on the XML documents. According to Suganuma, the difficulty levels of questions vary if the forms of questions are switched between those three forms although the contents of questions still remain the same. The alternative form which gives multiple choices to a student was determined as the easiest form. In the fill-in-the-blank form, students must answer questions without any alternative choice or clue for answers. Therefore, questions in this form are harder than questions in the alternative form. In the error-correcting form, students need to find wrong parts in problem statements, and correct the wrong parts without any clue. Therefore, questions in the error-correcting form are considered as the most difficult type of questions. On the other hand, Ogura [16] has conducted additional studies on Suganuma’s study and clarified that transforming a problem from the sentence or calculation form to the error-correcting form might lower the difficulty level of the problem since the error-correcting form might give the student some hints of answer.

Questions are automatically generated by controlling the forms of questions in accordance with each student’s comprehension level figured out during the learning process. A tag is also prepared to categorize a set of practice questions tailored by its contents, and control the order of these questions based on each category. However, any efficient method to categorize a set of practice questions by using the tag has not been proposed yet. Christian [21] implemented a measurement algorithm that can be used to manage the comprehension levels of the students based on their response times and percentages of correct answers. Then, the measurement results were used to control the questions presented to the students subsequently.

The study we conducted is the first attempt to categorize the learning target into some areas, and figure out the comprehension level of each student in each area (specifically, identify each student’s weak category). Therefore, this paper is the first study that identifies a weak part of a student and promotes the student’s learning focused on the weak part.

4.5 Comparison with the Author’s Previous Studies

According to the proposals by Yamamoto [17], they assumed that a weak category can be extracted by inductively inferring a category (a weak category) which is common to incorrect answers from a student. We decided not to use this method since it used inductive inference and required a lot of questions to figure out a weak category.

There was a situation in which a category may be identified as a weak category at one point and not at another point depending on the questions selected from an identical category. Uenosono [18] introduced the concept of difficulty level in order to solve this issue. Proposals from Uenosono [19] had no strategic knowledge [S3-2] until then. Before introduction of [S3-2], there was a problem in which a weak category could not be narrowed down further, if an identical category contained question answered correctly and those answered incorrectly at the same time. Introduction of [S3-2] enabled the author to narrow down a weak category in such a case.

However, the relationship between the number of questions which compose a set of questions and the number of category layers has not been clarified in this proposal. We have successfully resolved this problem from the results of simulations we conducted, and proposed this improvement in this paper.

5. Conclusion and Future Direction

The CAI system based on usual Overlay model[23] has no concept of different areas of study in the Student Model, and students may have to study the areas where they have understood enough. Therefore, we can safely say that this system is not providing an ideal learning method. To solve this problem, we need a system to identify a weak part of study where a student has not comprehend enough, and apply instruction programs which force the student to only focus on his or her weak part. We focused on students’ weak parts as the target field. Our goal here is to build a CAI system which forces students to only focus on their weak parts by automatically identifying the weak parts of students. We introduced the concept of a category meaning a classification, and tried to identify a category classification (category) where a part of study belongs to when we categorized those areas of study based on the whole-part relationship.

We proposed the student model and instruction programs for a CAI system in Uenosono [18] to materialize the concept above, and discussed its effectiveness. However, although our traditional versions of the CAI system can be used to determine the number of category layers based on the areas of study, the system cannot clarify the appropriate number of questions which compose a set of questions on each category layer. Therefore, we conducted a series of simulations to solve this problem. As a result, we figured out the appropriate number of questions that compose a set of questions on each difficulty level.

In this system, Decision Frames are provided first to identify a weak category of a student. Once a weak category is identified, Text Frames are presented to
provide supplemental knowledge required to produce right answers for questions chosen from the weak category. In short, our CAI system presents some Decision Frames before presenting Text Frames. On the other hand, traditional CAI systems present some Text Frames before presenting Decision Frames. In addition to the order to present Text Frames and Decision Frames is reversed, our CAI system mainly focuses on Decision Frames while the traditional CAI systems mainly focused on Text Frames. We need further study on the features of Text Frames in our system.

References