An Item Selection Strategy Based on Association Rules and Genetic Algorithms

Ming-Hsiung Ying, Shao-Hsuan Huang, Luen-Ruei Wu
Chung Hua University, Taiwan
E-mail: {mhying, M09610023, M09610028}@chu.edu.tw

Abstract—The online learning and testing have been as important topics of information education. The main purpose of academic testing is to improve learning. Students could take online test to evaluate their achievements to learning goals. Many online test systems randomly generate test papers from an item bank. A high-quality test paper must to consider the following questions. Is the depth and breadth of test items appropriate? Can test items examine student ability at different cognitive levels? Can test items avoid relationships among test items? Can a test identify student ability and provide learning suggestions appropriate? Therefore, it is the important issue to solve above problems by using information technology. This study applies a novel item selection strategy implemented by computer and is based on assessment theory, association rule, genetic algorithms and a revised Bloom taxonomy. The proposed strategy ensures that test is high quality.

Index Terms—item selection strategy, association rule, genetic algorithms, revision of Bloom’s taxonomy, assessment theory

I. INTRODUCTION

The online learning and testing have been as important topics of information education. Test is a kind of tool to evaluate students’ achievement, and it can also obtain large amounts of data related to learning progress and are a fair way to evaluate student learning status because subjective factors do not affect test results. Therefore, testing can enhance learning performance, and it is the best way to evaluate student learning status [4].

With the rapid development of information technology in recent years, there are numerous benefits for computer-based testing (CBT) and online testing (OLT) delivery, including data-rich test results, immediate test feedback, convenient test times and places, reduction in the time-consuming job of scoring tests and preparing items, test items can be created in ‘banks’ and delivered at random [14][8][13]. Therefore OLT and CBT have become important issues in information education.

In 1956, Bloom identified a learning taxonomy. The Bloom’s taxonomy has been used by instructional designers and teachers at all levels of education. Anderson and Krathwohl revised Bloom’s original taxonomy by combining both the cognitive process and knowledge dimensions in 2001. This new expanded taxonomy can help instructional designers and teachers to set meaningful learning objectives, and it provided the measurement tool for thinking [3].

In recent years, many researchers applying genetic algorithms (GA) to test paper generation, GA can improve test paper generation operational efficiency [15].

This study uses assessment theory, association rules, GA and a revised Bloom taxonomy to ensure that test paper is high-quality. To determine whether test items have appropriate difficulty and extensively knowledge content, one must ensure the correlation coefficient between items is minimized, and evaluate different identification knowledge level items.

The proposed item selection strategy considers various factors that can affects item appropriateness and subject relevance of parameters such as difficulty, discrimination, distribution rate of the revised of Bloom taxonomy, and the content of items.

The purposes of this study summarized as follows.

1. Use association rules, assessment theory, genetic algorithm, and revised Bloom taxonomy to ensure that tests paper is high quality.
2. Consider some item-related factors that may affect the test quality, such as difficulty, discrimination, revised Bloom taxonomy distribution rate, content of items, and ensure the correlation coefficient between items is minimized, and evaluate different identification knowledge level items.
3. Achieve efficient management of tests and maintain a bank of items for measuring student learning performance.
4. A preliminary study of the item selection strategy confirmed that the computer automatically generated high-quality test papers.

II. RELATED WORK

The purpose of education is to lead to cause students to change their behaviors, and teachers have to use tests well to assess their change condition. The main purpose of tests is to improve the learning performance [5]. Test theory is primarily divided into classical test theory and modern test theory. Classical test theory is based on the true score model [12][19], whereas modern test theory is based on item response theory[20]. Therefore, via a “self-study” of online learning environments, learners can use online tests to evaluate their performance. Although students can get a test score, the most important factor is to combine the items with educational goals. To achieve this objective, an online testing system is included in the information and knowledge level data for each item.
Yu argues that item analysis focuses on statistical characteristics with quantitative evaluations, and these characteristics include both discrimination and difficulty [24]. Via analysis teachers can gain information for various items to develop high-quality test papers, indirectly enhancing test reliability and validity, and the selection of test items from an item bank. Teachers can use the test results to implement item analysis and retain excellent items for future reference. To achieve time and energy efficiency, teachers can combine test and evaluation functions to improve their teaching.

Association rules are typically utilized in data mining. Data mining is a technique used to find information in hidden knowledge. For instance, a retail store manager determined that customer will buys soft drinks and snacks together. A strategy will place the two products near each other. Therefore, sales of both products can be increased.

Holland's GA is based on Darwin's evolution theory. The goal is to find optimal solution algorithms [16]. This algorithm is based on the evolutionary concept that animals and plants suited to an environment will outlive those unsuited to the same environment. Through the designed fitness function, the offsprings will continue to mate and produce offsprings that are relatively more suited to the environment [7].

A GA consists of five components, and these five components are [9]:
1. A method for encoding potential solutions into chromosomes.
2. A means of creating the initial population.
3. An evaluation function that can evaluate the fitness of chromosomes.
4. Genetic operators that can create the next generation population.
5. A way to set up control parameters, e.g., the population size, the probability of applying a genetic operator, and so on.

This computation process generates an initial population, chromosomes, and genetic coding and decoding. To calculate the value of a fitness function, there are two operations in the traditional genetic algorithm that we adopt in this research:
2. Evolution operation: selection.

As shown in Figure 1, crossover operator swaps the left potions of two chromosomes. Mutation operator replaces bits on a chromosome with randomly generated bits.

![Figure 1. Genetic operators of genetic algorithm](image)

GA are probabilistic search techniques, it has been widely used in engineering and science for solving optimization problems, as well for commercial and financial projections.

Even though the GA based approaches prove to be robust and efficient, another important factor to be considered in making GA efficient is the tuning of control parameter values. According to Grefenstette(1986), the parameters considered in GA include the population size (N), the crossover rate (C), the mutation rate (M), the generation gap (G), the scaling window (W), and the selection strategy (S). In his research, Grefenstette denoted a GA with some specific parameter values as GA (N, C, M, G, W, S)[11].

In designing test items, the educational objectives should be considered. Bloom's Taxonomy divides educational objectives into three "domains:" Affective, Psychomotor, and Cognitive. Within the taxonomy learning at the higher levels is dependent on having attained prerequisite knowledge and skills at lower levels[6][21]. Cognitive psychologists recognize three distinct types of knowledge : declarative, procedural, and conditional. Declarative knowledge is knowing "that". Procedural knowledge is knowing "how" to execute a skill or apply concepts and principles to specific situations. Conditional knowledge is "knowing when and why" to utilize declarative or procedural knowledge [1][10][22][18].

### Table I.

<table>
<thead>
<tr>
<th>Knowledge Dimensions</th>
<th>Cognitive Process Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Remember</td>
</tr>
<tr>
<td>Factual</td>
<td></td>
</tr>
<tr>
<td>Conceptual</td>
<td></td>
</tr>
<tr>
<td>Procedural</td>
<td></td>
</tr>
<tr>
<td>Meta-cognitive</td>
<td></td>
</tr>
</tbody>
</table>

In recent years, the taxonomy of educational objectives by Bloom et al. has been widely used. Bloom identified six levels within the cognitive domain, including knowledge, comprehension, application, analysis, synthesis and evaluation [3]. Anderson and Krathwohl revised the original taxonomy of Bloom by combining both the cognitive process and knowledge dimensions [3]. Table I describes the revised Bloom's taxonomy, the taxonomy comprises a two-dimensional table. One dimension identifies the knowledge (the kind
of knowledge to be learned), while the other identifies the cognitive process (the process used to learn). The knowledge dimension comprises four levels: factual, conceptual, procedural, and meta-cognitive. The cognitive process dimension comprises six levels: remember, understand, apply, analyze, evaluate, and create. This new expanded taxonomy can help instructional designers and teachers set meaningful learning objectives, and provide the measurement tool for thinking[2][17][23].

III. RESEARCH STRUCTURE

A. Research method

In the marketing field, researchers attempt to find association rules based on the correlation coefficient between commodities and present highly relevant merchandise in the same field to enhance sales. In the educational assessment field, teachers would like the test papers that can use the selection strategy to implement an online test system. Information technology is combined with test theory, such that an online test system can use the selection strategy to implement an online test system.

B. Research steps

The first step in this study is to build an item bank of “Enterprise Resource Planning” course. The item bank contains 313 items that lists four types of information about each item: difficulty, discrimination, knowledge level of the revised Bloom's taxonomy, and cognitive process level of the revised Bloom's taxonomy. This experiment adopts three chapters with 69 items because the item bank is too large. The system generates a test paper for students, according to test results, calculates item difficulty and discrimination. The formulas for item difficulty and item discrimination are as follows:

1. Item difficulty: is derived by DIF=R/N, where DIF is item difficulty, N is the total number of students who answered the item correctly, and R is the number of students who answered the item correctly.

2. Item discrimination: is derived by DIS=P_H-P_L, where DIS is item discrimination, P_H is the percentage of students who had the right answer and in the high score group, and P_L is the percentage of students who had right answer and in the low score group. Test scores are separated into three groups - the high score group, middle score group and low score group; 25-33% of the total number of students are chosen for each group.

3. Correlation coefficient between items: is derived by R_ab=(CT_ab+IT_ab)/ST_ab, where R_ab is the correlation coefficient between item a (I_a) and item b (I_b), CT_ab is the frequency of I_a and I_b that random chosen into the different student’s test, ST_ab is the frequency of students who answered the I_a and I_b simultaneously correctly, and IT_ab is the frequency of students who answered the I_a and I_b simultaneously incorrectly. S={I_1,I_2,…,I_n} is a combination of item bank, n is the total numbers of items in the item bank, For example, 50 students participate in a test, there are 50 different test papers generated. The item 7 and item 9 chosen in a test paper totally 25 times, 12 students answered the I_7 and I_9 simultaneously correctly, and 8 students answered the I_7 and I_9 simultaneously incorrectly. The correlation coefficient between item 7 and item 9 is R_{7,9}=(12+8)/25=0.8.

C. Chromosome structure and fitness function design

The fitness function design in this study is based on item difficulty and item discrimination of test theory, and a fitness ratio of Bloom’s taxonomy. The following brief outline of the GA illustrates how the GA functions, where notation S(t) is the population in the tth generation; si(t) is the ith member in S(t); f(si(t)) is the fitness value of si(t); and TOTFIT(t) is the sum of f(si(t)) for all si(t). We calculate the quality score (satisfaction degree) using the fitness function described as follows:

\[ f(si(t)) = \text{Wdif} \times \text{Adif}(si(t)) + \text{Wdis} \times \text{Adis}(si(t)) + \text{Wac} \times \text{Ac}(si(t)) + \text{Wrb} \times \text{RB}(si(t)) \]  

where Adif is the si(t) chromosome’s functional score for item difficulty, Adis is the si(t) chromosome’s functional score for item discrimination, AC is the correlation between two items, RB is fitness ratio of Bloom’s taxonomy, and Wdi, Wdis, Wac, Wrb are the weight for each factor. The factors of fitness function are described as follows:

1. Average fitness score of item difficulty-Adif and average fitness score of item discrimination-Adis: Before calculating the fitness score, Adif and Adis should be derived; the derivational formula is:

\[ \text{Adif}(s_i(t)) = \frac{1}{y} \sum_{j=1}^{x} \left( \frac{c_j \times PD_j(s_i(t))}{y - 0.5} \right) \]  

where x is the total number of items in the item bank, and the number of genes in all chromosomes si(t), y is the number of items needed in the test paper, PD_j(s_i(t)) is the difficulty of the jth gene in s_i(t), and C_j=0.1.
\[ AD(x) = \left( \sum_{i=1}^{n} |C_i \cdot D_i(x) + \sum_{i=1}^{n} \frac{D_i}{x} \right) - \sum_{i=1}^{n} D_i(x) \\ \text{where } D_i(s_i) \text{ is the item discrimination of the } j \text{th gene in } s_i(t), \text{ and } C_i = 1. \]

2. Average correlation coefficient-AC(s_i(t)):
For each bit in a chromosome where C_j=1, identify the item id, and calculate the correlation coefficient between any two items, the formula as follows:
\[
AC(s_i(t)) = \frac{\sum_{i=1}^{n} AC_i}{\sum_{i=1}^{n-n_{AC}}} - \frac{\sum_{i=1}^{n} AC_i}{c_i^2} 
\]

3. Fitness ratio of Bloom’s taxonomy- RB(s_i(t)):
This study use Bloom’s taxonomy theory which has two dimensions. First is knowledge dimension, another is cognitive process dimension. Table II describes item distribution for the item bank, those items are classified according to Bloom’s taxonomy:

<table>
<thead>
<tr>
<th>Cognitive Process Dimension</th>
<th>Knowledge Dimension</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Factual</td>
<td>Remember</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B Conceptual</td>
<td>Understand</td>
<td>10</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C Procedural</td>
<td>Apply</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

where x is the total number of items in the item bank. Thus, we expect that fitness value is as large as possible, to ensure that test quality is good and the test can evaluate student learning performance.

III. IMPLEMENTATION OF THE GENETIC ALGORITHM

The implementation steps for the GA are as follows:

**Step 1**
Generate the initial population S(t) in Table III, where t=0. The GA parameter is GA (30, 0.975, 0.075, 1.0, 1, E), it means that every generation has 30 chromosomes (POPSIZE=30); the crossover rate is CR=0.975, mutation rate is 0.075, and the number of generations is GENER=300. The weight of each factor is Wdif=Wdis=Wac=Wr=25.

We suppose that C_2, C_4, C_5, and C_9 are a selected combination of one chromosome (Fig. 1), thus, R_{2,4}, R_{2,4}, R_{2,4}, R_{2,4}, R_{2,4}, R_{2,4}, R_{2,4} are six R_{ab} C_4=6 kinds combination R_{ab}.

Thus, we expect that fitness value is as large as possible, to ensure that test quality is good and the test can evaluate student learning performance.
Step 2
Calculate the fitness value for each member, \( f(s_i(t)) \). The fitness value of each number (see Table III), e.g. the twenty-sixth chromosome has the optimal fitness value, and the satisfaction degree is 73.272%.

Step 3
Calculate the selection probability for each \( s_i(t) \). The selection probability is defined as
\[
P(s_i(t)) = \frac{f(s_i(t)) - \text{Min}(f(s_i(0)))}{\text{TOTFIT}}
\]
The 26th chromosome has the maximum fitness value, and has the maximum probability to be selected as a parent; the probability value is 0.04670 (Table III). The 8th chromosome has the minimum fitness value, and its selection probability is 0.

Step 4
Select a pair of members (parents) that compare the accumulation probability with a random number (ranging from 0 to 1), and reproduce the chromosome into the new population.

Step 5
Apply the genetic operators (crossover, mutation, and inversion) to parents. Replace the parents with the resulting offspring to form a new population, \( S(t+1) \), for the generation \( t+1 \). If the size of the new population is equal to 30, then go to step 6, else go to step 4. In our case, every generation including 30 chromosomes, we use the higher crossover rate of 0.975 that can generate more newly structure, and prevent dropping into the local optimal solution. Our mutation rate is 0.075. If the result falls in the optimal local area, we will consider it in the computation of the global optimal area to be a possible final number.

Step 6
If the current generation, \( t + 1 = 300 \), then stop; otherwise, go to step 2.

Finally, we got the optimal fitness value is 82.0916 (Table IV). Its chromosomal structure was \{0000010100 000001001100000000010001010100 01000000011100 1100 12.5\% 0 12.5\% 25\% 37.5\% 50\% 62.5\% 75\% 87.5\% 100\% 01000000011100 \}. The 6th, 8th, 16th, 19th, 20th, 22th, 24th, 29th, 30th, 33th, 37th, 46th, 48th, 51th, 53th, 54th, 58th, 66th, 67th, 68th test items were selected into the test paper. It means that this combination of items has the appropriate difficulty and knowledge levels, item discrimination and a variety of cognitive levels.

Figure 3 shows the stable maximum fitness value after 177 generations. The average fitness value increased rapidly after the initial stage reached a stably state. The minimal fitness descended gracefully with the chromosome mutation. The mutant chromosome had the minimal fitness value. Thus, we believe a high-quality test paper was generated.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Fitness function value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>73.2723</td>
</tr>
<tr>
<td>2</td>
<td>73.5823</td>
</tr>
<tr>
<td>3</td>
<td>73.7523</td>
</tr>
<tr>
<td>4~18</td>
<td>78.1932</td>
</tr>
<tr>
<td>19~49</td>
<td>80.4305</td>
</tr>
<tr>
<td>50~58</td>
<td>80.6444</td>
</tr>
<tr>
<td>59~165</td>
<td>81.6135</td>
</tr>
<tr>
<td>166~176</td>
<td>81.9647</td>
</tr>
<tr>
<td>177~300</td>
<td>82.0916</td>
</tr>
</tbody>
</table>

Figure 3. Fitness value of each generation

V. CONCLUSIONS
This study applies a novel item selection strategy executed by a computer. The strategy is based on assessment theory, association rule, GA and revised Bloom taxonomy, thereby ensuring that a high-quality test paper is generated.

Through the GA, the item selection strategy generates test papers that meet the needs of most teachers. Additionally, the strategy can easily deal with large amounts of data in the item bank. Furthermore, the item combinations generated by the item selection strategy ensure that the test has appropriate difficulty, high degree of discrimination, a maximum independence parameter between items and an appropriate knowledge level.

ACKNOWLEDGMENT
This research is supported by National Science Council of Taiwan, R.O.C. (Grant No: NSC 98-2410-H-216-006-MY2) and Chung-Hua University (Grant No: CHU-NSC 98-2410-H-216-006-MY2).

REFERENCES


Ming-Hsiung Ying: Dr. Ying is an assistant professor of information management at the Chung Hua University, HsinChu, Taiwan. He received the PhD in Management Information System from National Chengchi University, Taipei, Taiwan in 2006. His research interests center on E-learning, knowledge management, psychology of Internet, database system, and enterprise resource planning.

Shao-Hsuan Huang: He is a Master student at the Department of Information Management, Chung Hua University, Taiwan. His major is information management and E-learning.

Luen-Ruei Wu: He is a Master student at the Department of Information Management, Chung Hua University, Taiwan. His major is information management and E-learning.