Customized Online Test Database Systems and Automatic Generation System for Testing Papers

[a]Zhi Liu, [b]Ronghua Liang, [c]Xiang Pan

[a, b, c] College of Computer Science and TechnologyZhejiang University of Technology Hangzhou 310023, China

Abstract

In the online test database system, a major issue is how to input the questions with formulas and pictures to the test database, while maintaining the format of test questions in order to support the printing layout. In this paper, a test database system supporting the formulas and pictures stored is designed and implemented. Based on this database system, an automatic generation system for test paper is achieved based on improved genetic algorithm. The J2EE technology is used to develop the online system. The system has friendly interface and it is easy to be operated. This system can support multidisciplinary test database and it can be customized according to different chapters and sections. The experiments show it is a useful online test database system and automatic generation system for testing papers.

Index Terms: test database system; automatic generation system for testing paper; online; improved genetic algorithm

1. Introduction

In recent years, computer technology and network technology are developing rapidly and their applications in education are more and more widely. In order to support the separation of examination and teaching, many test database system and automatic generation system for testing papers appears. But most of them are just designed for supporting the text questions stored in the database. In practical applications, many questions are expressed in formulas or pictures. How to edit them and input them into the database is a very important issue. In this paper, some methods are proposed to give an answer about it. In order to implement the automatic generation system, we proposed the improved genetic algorithm. Based on the traditional genetic algorithm, the encoding settings, the adaptive function, evolutionary operator are improved. The strategies for suppressing prematurity in the evolution process and population updating strategies are proposed in this paper. The automatic generation algorithm for test papers based on the improved genetic algorithm is realized. It will be convenient for teacher to input questions and generate the test papers. It is also useful for students to test themselves on this system. All test questions are supplied with the reference solutions.

Corresponding author:
E-mail address: [a]lzhi@zjut.edu.cn

2. System structure

In this test database system, there are main 6 modules: chapters management, sections management, test questions management, automatic generation system for test papers management, user permissions management and setup management. The system structure is shown in Fig. 1.

Figure 1.   The system structure of test database system

The main function modules are discussed below:

(1) The chapter management module

This system is designed to support test questions management for multi disciplines. In this module, user can create a new discipline and create the various chapters under this discipline. It is shown in Fig 2.

Figure 2.   The chapter management

(2) The section management module

In this module, user can create various sections according requirement under a chapter. It is shown in Fig 3.

Figure 3.   The section management

(3) The questions management module

In this module, user can upload the questions to the test database. The editor is supported in this system. User can edit the question, copy and paste the question. The formulas and pictures can be copied and pasted too. It is shown in Fig 3.
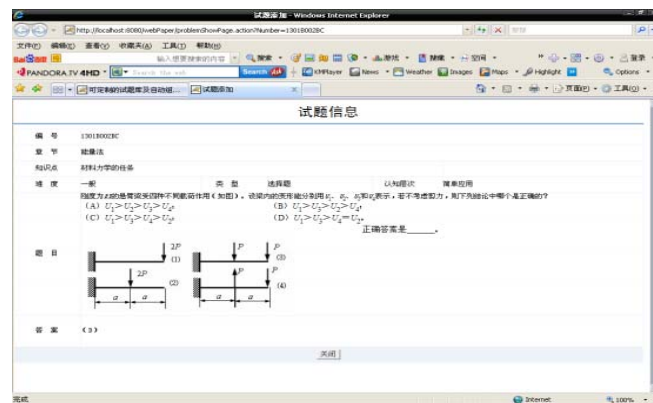

Figure 4.   The question management

In order to save the question with the formulas and pictures, the text information in question in the Word format will be saved as in the HTML format firstly. The formula saved as vector format will be converted into two picture files with suffix of .wmz and .gif. In order to prevent the files to be overwritten, the files uploaded to the service are renamed as the time when the files are uploaded. On the service, the source file in HTML format is gotten and the source code is analyzed. The picture source address stored in the field "<v:imagedata src="">" will be replaced to the relative path the picture is saved adding file name. Then the whole HTML source codes will be stored in database. When the question is shown on the webpage, these HTML source codes will be read and analyzed. The editor is shown in Fig 5.
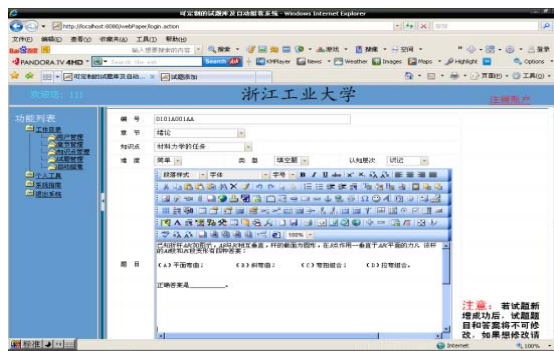
Figure 5.   The editor for test question

(4) The automatic generation paper

This module is the key part in the whole system. Here, an improved genetic algorithm is used to generate test paper according to the requirements from users. The generation paper is shown in Fig 6.
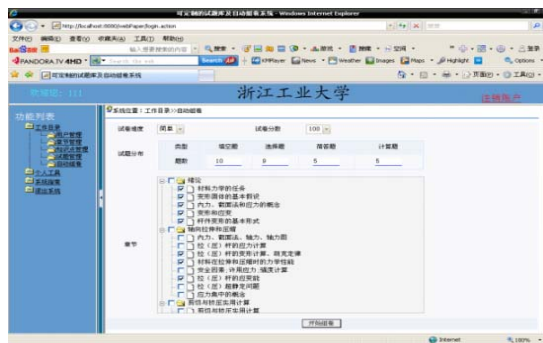

Figure 6.   The generation paper management

Users can specify the testing scope, including the testing chapter and testing sections. They can also specify the question forms, such as the filling in the blank, multiple Choice, Q&A questions, problem-solving, and so on. Users can also specify the test difficulty level and test time. After these specifying, the test paper can be generated automatically. The generated paper is shown in Fig 7.
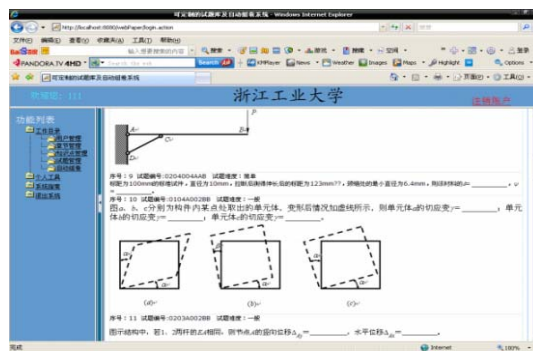

Figure 7.   The generated paper

In order to analyze the test paper, users can check the paper and find the various test question ratio, such the difficulty ration, chapter ratio, and test question form ratio. It is shown in Fig. 8.

Figure 8. The test question ratio

If the generated paper can not meet the user's need, user can also change the test questions by manual method. They can choose new item from the test database directly in special chapters or sections and replace the old test question. They can move the idea item from the right window to the left window easily. It is shown in Fig 9.
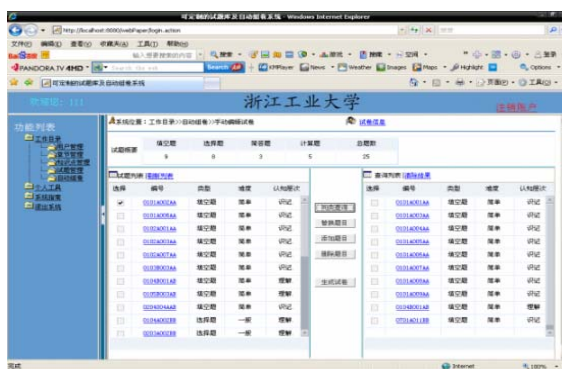

Figure 9. Manual method of modifying the test questions

The final test paper can be downloaded and saved as a file with word format and can be printed directly.

3. Improved genetic algorithm

In this system, the algorithm of automatic generation paper is a very important issue. An improved genetic algorithm is used in this system. Here, several key problems are discussed.

(1) Test paper structure: The test paper is generated according to different constraints, such as total score, test form ratio, test difficulty, use frequency of test question, and so on. We define these factors as below:

①Total score: $\sum_{i=1}^{n} Mark_i = P$. The P is the total score of the test paper. $Mark_i$ is the score of item i .

②Score of question with different forms: $\sum_{i=1}^{n} Style_{i,m} * Mark_i = P_m$

③Difficulty of test paper: $\sum_{i=1}^{n} Dif_i = D$. Dif_i is the difficulty value of item i.

④ Use frequency of test question: $(\sum_{i=1}^{n} Ordr_i)/n = e$

(2) Mathematical Model of generation paper: in this paper, the algorithm of generation paper is based on genetic algorithm. The generation paper problem is a problem to solve approximate optimal solution under multi constraints conditions. So we define four vectors here.

$A = \{a_1, a_2, a_3 \cdots\cdots a_n\}$ is the 为 attribute set, such as difficulty, question form, and so on.

$V = \{v_1, v_2, v_3 \cdots\cdots v_n\}$ is the value set. It includes error tolerance ranges of each attribute variable.

$R = \{r_1, r_2, r_3 \cdots\cdots r_n\}$ is the constraints set. It includes the constraint relationships between each property variables and values.

$I = \{i_1, i_2, i_3 \cdots\cdots i_n\}$ is the test questions set, in which a test question meets the constraint relationship.

The automatic generation for test paper is designed to select a set of test questions from the test database randomly. All property variable set A in these questions will satisfy the constraints of R in the range of V value. So the problem o f generation for test paper will be changed to the problem of solving approximate optimal solution under multi constraints conditions and the approximate optimal solution is not the unique.

(3) Improved genetic algorithm: The traditional genetic algorithm has some difficult to be applied in this problem, such as encoding problem and premature problem.

① Encoding problem

The segment real is used to be encoded. The question number is used as gene when the paper is initialized. The genes with same question form will be set on one vector and all kinds of gene encoding in vectors with various test forms is combined to form a chromosome. The benefit of this method is not only to shorten the code length, and in this encoding method, the number of test questions and the forms of test questions are reflected in the coding. It avoids generating the test paper that does not satisfy the constraints.

②Design of fitness function

For the above four constraints, the error function of chromosome which meets the goal of generation paper is constructed:

$$E = \sum_{i=1}^{4} e_i * a_i$$

In which, $e_i \in \{0,1\}$ is the relative error and a_i is error weighting coefficient of relative error. Here, $\sum_{i=1}^{4} a_i = 1$. According to expert opinion, the relative error of each constraint should be limited in the range of 1% to 5%.

Fitness function is inversely proportional to error function. That means that error between individual and target is smaller, the fitness of individual is greater. According to the principle, the fitness function of satisfying goal of generation paper is constructed as:

$$Fitness = (E + \mu)^{-1}$$

In which, E is the value of error function and μ can be set the value as 0.01 according to the experience.

③Design evolution operator

In this paper, the roulette method is used to select operator. The segment single point crossover method is used. That means that when cross operation is running, every question with the same question form will be crossed with each other. It can be regarded as multi-point crossover. Then single point mutation method is used. That means mutation operation can turn out only once in order to avoid changes in the number of test questions in the kinds of questions.

④ Strategy of suppressing prematurity

In order to solve the prematurity problem in the genetic algorithm, the fitness of the population is changed with the linear transform. In the early stage of evolution, the fitness gap between the population individual fitness is reduced. It will prevent the competitive prominent individuals to control selection process. In the later stage of evolution, the fitness gap between population individuals is enlarged. It will increase the diversity of the population individuals in order to prevent the local convergence of the evolutionary process.

⑤Population update strategy

The optimal reservations policy is used in this paper. , that means that the best individual for each generation will be directly copied into the next generation. At the same time, in order to maintain the diversity of population individuals, the best individual will also participate in the mutation and crossover operation.

(4) Experiment analysis:

Under the demand in the default, in which the difficult level is medium and total score is 100, two typical generation paper is gotten in this system. In the experiment, the crossover and mutation probability is 0.4 and 0.2 respectively. They are shown Table 1 and Table 2.

TABLE I.   THE COMPARISON OF DIFFICULT LEVEL

Difficulty level	Easy	Rather easy	Medium	Difficult	Very Difficult
Level value (paper 1)	11	25	37	20	4
Level value (paper 2)	2	23	58	10	4

TABLE II.   THE COMPARISON OF QUESTIONS DISTRIBUTION

Chapter	1	2	3	4	5	7	8	9	10	13	14	15
Score paper 1	3	23	9	0	9	11	17	2	4	8	3	8
Score paper 2	3	17	2	17	6	11	11	0	0	5	9	16

From the view of above tables, we can find that the medium difficult questions are majority and the score error is about 3%. The distribution of chapters is proportional to the amount of test database roughly. So the improved genetic algorithm can satisfy the constraints in the generation for test paper among the difficult, score, test item distribution, and so on.

Conclusions

In order to generate an appropriate test paper, the test database system must support full test questions. If the test questions are not enough, the generated test paper will be looked like similar and it will influence the effect of test paper. Another important issue needs to be solved in this system is the uniqueness of test question. Two similar test questions can be represented in rather different ways and generation system will select them as two different questions in the same paper. At present, this problem can only be solved by manual modified.

References

[1] Wang You-ren, Zhang Zhai, Gui Jiang, Yao Rui, and Chu Jian-bo. Mathematical Model and Algorithm of Intelligent Test Paper Auto-generation System from Item Pool[J], System Engineering Theory and Practic, 2004(9), P85-89(In Chinese).

[2] Yang Jun, Application of Improved Genetic Algorithm in Automatic Test Paper Generation[J], Computer Applications and Software, 2009(12), Vol. 26, P225-227.(In Chinese).

[3]   Yang Luming, Chen Daxin, Application of an Improved Genetic Algorithm in Composing a Test Paper[J], Computer and Digital Engineering, 2004(5), Vol. 32, P76-79. (In Chinese).