The Research and Application of General Item Bank Automatic Test Paper Generation Based on Improved Genetic Algorithms

Zhen-hua Jia, Chun-e Zhang, Hao-shuai Fang

Department of Computer Science and Engineering
North China Institute of Astronautic Engineering
Langfang, China
jiazhenhualf@126.com, zcxe76@sohu.com, huagongfang@163.com

***Abstract*—Automatic test paper generation is a core stage of computer aided test (CA), is one of the indispensable important technology of computer network test system. This article aim to the goal request of test paper generation, adopt different weight coefficient to different target according to its important, define the fitness function of automatic test paper generation, and use the improved genetic algorithm. The experimental results show that the algorithm can solve the problem of automatic test paper generation compared to other algorithm in item bank very good, and has good performance, high efficiency and quality. At the same time, provide a new method and thought for solving similar multi-objective constraint problems and not neighboring combination problems.**

Keywords: Automatic Test Paper Generation; Genetic Algorithm; Item Bank

I. INTRODUCTION

With the wide application of computer assistant testing (CAI), Using computers to establish item bank, and realize the computer automatic select test and composing test paper is an important part of the CAT. Use computer to automatic test paper generation, the computer according to the test paper attributes appointed by test paper setting people can extract test questions which satisfy the above constraint condition to component test paper. The algorithm m of automatic test paper generation influences the examination paper quality directly. Now commonly used automatic test paper generation algorithms usually have 3 kinds: random choice method, backtracking heuristics and genetic algorithm. Random choice method is mainly using the random numbers generated by computer to select test questions. This method can not consider all the constraint conditions. Backtracking heuristics has certain advantages in considering hardness constraint condition of test questions. For the item bank system which state types and choosing questions quantity are smaller, this method has good composing efficiency. But the performance often decreases sharply when the search space of item bank increase, however the search problem feasible solution space often spend large time. Genetic algorithm is a random optimization algorithm which analog natural selection and natural genetic mechanism, its main characteristic is group search strategy and information exchange between individuals in group. It can produce different progenies according to different environment, has dynamic and adaptability, thus satisfy the changing needs of the item bank.

II. IMPORTANT ATTRIBUTE INDEX OF TEST QUESTIONS

The attribute index of test questions describes the intrinsic nature, external feature and its roll in exam of each test question. It is the foundation of select questions and composing test paper by computer. You need to make sure the attributes of each test questions before composing test paper. The important attribute state as follow:

(1) Question type (qt), refers to the test questions type included in test papers, namely composing test papers by what type of test questions.

(2) Question difficulty coefficient (dfct), test question difficulty is an important of test questions quality indicators. The difficulty is usually refers to the difficulty degree when complete questions or project task.

(3) Distinguishing of test questions (dfrc). The distinguishing of test questions is measured the index of literacy level of a test questions for different levels examination people, it is the measurement for distinguish the ability of testee.

(4) Answer-question time (ta), refers to the estimate time of complete a certain questions, general with minutes for unit.

(5) Test questions score (ps), namely the score of each test question. Consider to set different scores by various users to different test questions in this model, this index will influenced by test questions difficulty and knowledge point set.

(6) Test questions mastery degree (ms), divide the degree of examinee need to master into 5 levels according from easy to difficult, include know, understand, master, application and comprehensive application.

(7) Knowledge point covering set (ks), namely the knowledge point covering of test questions and its proportion.

III. MEASUREMENT INDEX OF TEST QUESTIONS QUALITY

Test questions index is a key factor of composing test paper process, the users can put forward relevant requirements to entire test paper.

(1) Reliability and validity of test paper

978-1-4244-9600-6/11/$26.00 ©2011 IEEE

In educational measurement, the basic index of measure test paper quality is the reliability and validity of test paper.

The reliability of test paper is the index of credible degree when test paper as measurement tools.

The validity of test paper refers to a test or scale can measure the characteristic degree of need measure, namely refers to a single measure how effective.

(2) Difficulty and distinguishing of test paper

In view of the reliability and validity of test paper shall be evaluate only after exam implement, cannot as the basis of judgment generating test paper quality when composing test paper. But has closely connection with the difficulty and distinguishing of test paper.

IV. THE BUILDING OF TEST QUESTIONS AUTOMATIC TEST PAPER GENERATION

Composing test paper process is abstract amount of test questions from item bank generating one or more test papers satisfy certain constraint condition. For test paper, consider the following constraint conditions:

The test paper can be regard as composing of amount of test questions selected from item bank, the evaluation matrix of one test paper include m test questions can be said by a vector $T = \left[EI_i \right] (\text{which i=1..m})$.

(1) Test paper difficulty coefficient, namely the entire difficulty of test paper. Calculation method is $Tdfct = (\sum_{i=1}^{m} dfct_{\text{i}} \times ps_{\text{i}}) / Tps$, which $dfct_{\text{i}}$ is the question difficulty coefficient of EI_i, ps_{i} is test questions scores, Tps is test paper total scores.

(2) Distinguishing index of test paper, namely the entire difficulty of test paper. Calculation method is $Tdfrc = (\sum_{i=1}^{m} dfrc_{\text{i}} \times ps_{\text{i}}) / Tps$, which $dfrc_{\text{i}}$ is the question difficulty coefficient of EI_i, ps_{i} is test questions scores, Tps is test paper total scores.

(3) Answer-question time, the total time used by examinee complete test paper. For a certain test paper, overall time is the total time to finish each test questions. $Tta = \sum_{i=1}^{m} ta_{\text{i}}$, Which ta_{i} is the answer-question time in EI_i.

(4) Knowledge point set, namely test paper contains what knowledge point and the score percentage of this knowledge point is a set. If select u knowledge point, which the first $j(j \le u)$ element is $Tk_j =< k_j, kp_j >$, which k_j is the certain selected knowledge point code, kp_j is the percentage of this knowledge point in test paper scores.

(5) The setting of question type score, namely what question type of test paper and the quantity and total score of test questions include question type, is a set. Which the first j element is $Tqt_j =< qt_j, qn_j, qtps_j >$, qt_j is the code of certain selected question type, qn_j is the total quantity of this question type.

(6) Mater degree setting, namely the different percentage of different master degree test questions in test paper, is a set. Which the first $j(j \le 5)$ element is $Tm_j =< m_j, mp_j >$, m_j is the master degree code, mp_j is the score percentage of master degree, obviously $\sum_{i=1}^{5} mp_i = 1$.

V. THE SOLUTION BASED ON IMPROVED GENETIC ALGORITHM

A. The Building of Multi-Objective Optimization Model

For the constraint conditions in composing test paper model, which see the soft constraint as the target of optimization algorithm, see the hardness constraint as the constraint condition of feasible solution, change the composing test paper model to multi-objective optimization model as follow:

Target (1), test paper difficulty error $f1 = \frac{\left| TdfctS - Tdfct \right|}{TdfctS}$

Target (2), test paper distinguishing error $f2 = \frac{\left| TdfrcS - Tdfrc \right|}{TdfrcS}$

Target (3), answer-question time error $f3 = \frac{\left| TtaS - Tta \right|}{TtaS}$

Target (4), knowledge point percentage total error $f4 = \sum_{i=1}^{u} \frac{\left| kpS_i - kp_i \right|}{kpS_i}$

Target (5), master degree percentage total error $f5 = \sum_{i=1}^{5} \frac{\left| mpS_i - mp_i \right|}{mpS_i}$

Constraint (1), test questions knowledge point constraint $\left[k_i \right] (\text{which i=1..u'}) \subseteq \left[k_i \right] (\text{which i=1..u})$

Constraint (2), test questions type score allocation constraint $TqtS_j =< qt_j, qnS_j, qtpsS_j >$.

B. The Implement of Solution Based on Genetic Algorithm

The genetic algorithm is the compute model of simulate the biological evolution process which simulate the genetic selection and natural selection of Darwin. The genetic algorithm processing of solution composing test paper model designed by this article as show in figure 1.

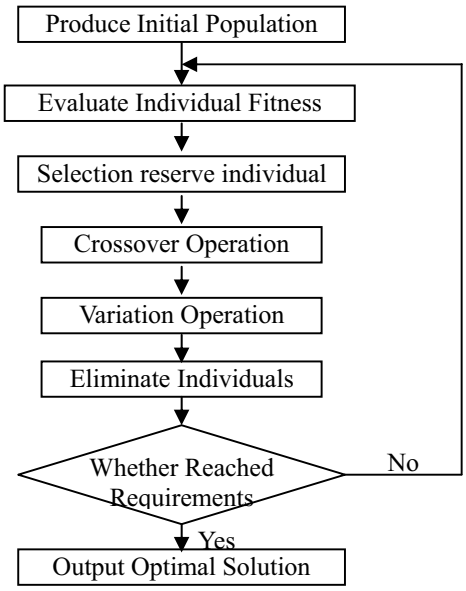

Figure 1. The basic frame of genetic algorithm

For each part of the genetic algorithm realization as follows:

(1) Chromosome coding method and initializes population, new individual creation methods

The new generation individual must satisfy the constraints of (1) and (2). For the constraint (1), just query the question type of including set knowledge point when select item bank. For the constrain (2), it can be seen as equality constraints, the traditional solution method is adopt penalty function, change it to inequality constraint. But in composing test paper system, this condition is must be satisfied hardness constraint, so adopt penalty function easily leading the composing does not satisfy constraint (2) and leading to the failure of composing test paper. According to the features of constraint (2), this algorithm adopt resolution numeric algorithm to generate new individual, for certain question type qt_i, the test question select method of satisfy the constraint of $TqtS_j =< qt_j, qnS_j, qtpsS_j >$ as follow:

① Split the question type score allocation for this question type, produce split combination, namely split $qtpsS_j$ to qnS_j number. Set can produce q kinds of combinations, notes for $cps_i (i \le q)$.

②Screen the combinations, there will generate available combinations after filter. Namely query item bank based on question type, knowledge point and scores, if the test questions number does not satisfy this combination, the combination is not available combination. Set available combinations as p kinds, denote $cps_i (i \le p)$, if not exists a combination of scores, then the test questions in item bank does not satisfy the requirements of composing test paper.

③For this question type, randomly select an available combination, randomly choose not repeat questions according to the score value in combination.

We all use the above method carry on extract for each question type in test paper can produce a new individual, initial population need circulated generate a new set of individuals.

For test questions qt_i the selected test questions is $Qqt_i =< cps_j, [tq_k] >$, which cps_j is the selected split combination code, $[tq_k]$ is selected test questions set. So the code for individual is $Q = [Qqt_i](i \le w)$, which w is the setting kinds of quantity of test questions.

(2) The calculation of individual fitness

According to important degree, target (1) – (6) coefficient choice for 1.5, 1.5, 1.0, 1.5, 1.0, the determination of fitness function is $f = 1.5f1 + 1.5f2 + f3 + 1.5f4 + f5$, obviously, the numerical smaller the test paper closer to expectation for each test paper, and the greater the fitness.

(3) Selection reserve individual (elite strategy mechanism)

According to the Markov's convergence model of chain evolutionary algorithm, genetic algorithm should adopt elite strategy mechanism. In the composing test paper model algorithm, according to the fitness, divide individual which in population into optimal, better, general and poorer, which the optimal and the poorer at least, the better and the general individual majority.

(4) Crossover operation

Choose from population with better classification individuals to a certain probability crossover operation, specific crossover operation as follow:

① Choose a question type according with certain probability, exchange all test questions of this question type.

②For the same question type, if the combination code is same, exchange the corresponding sequence test questions, if existing the same test questions after exchange, then select new test questions from item bank; if the combination code is different, for both of them carry on re-election test questions operation.

(5) Variation operation

Specific variation operation as follow:

① Choose a question type according with certain probability, choose combination again and choose test questions according to the new combination.

②After choosing a question type according with certain probability, the combination dose not change, and choose test questions again.

(6) Eliminated individuals

Make the poorer individuals in population as eliminated individuals, eliminated directly, produce the new individuals to complement the eliminated individuals, namely eliminated inferior individuals, increase the diversity of population at the same time.

(7) The choice of algorithm termination conditions

Because the information uncertainty in item bank, so the fitness function value may produce great fluctuation for different test questions, in order to enhance the generality of algorithm, adopt fixed evolution algebra to control the results of algorithm.

VI. THE ALGORITHM ANALYSIS AND TESTING

The item bank use "The Principles and Applications of Database" course item bank. The subjects were divided into 12 chapters (knowledge point), belonging to four modules. The specific conditions of test questions show in table 1, the master degree information of test questions show in table 2.

In the testing of the algorithm of composing test paper, parameter setting evolution generation for 250, population size for 100. Which evolutionary selection elitist strategy number set to 10, cross number to 60, variation number to 15, out quantity to the remaining for 15.

Compared to the algorithm, test backtracking algorithm and traditional genetic algorithm at the same time. Traditional genetic algorithm across the variability parameters setting the same with this algorithm, using the penalty function carry on process for constraint condition (2). Test paper composing select all the knowledge point of 12 chapter, select 2 different sets of expectations parameters, each parameter of various algorithms are run 50 times, find the average value of desired value error and the failure times of composing test paper.

TABLE I. THE BASIC INFORMATION IN ITEM BANK

Question type / Parameters	Choice	Blanks	Judgment.	Short answer	Applied
The average number of each chapter	60	50	40	20	20
Average difficulty	0.403	0.419	0.396	0.488	0.564
Average distinguishing	0.367	0.378	0.361	0.375	0.389
Test questions total quantity of multi knowledge point	125	110	92	130	125

TABLE II. THE BASIC INFORMATION OF MASTER DEGREE

Question type / Master degree	Choice	Blanks	Judgment	Short answer	Applied
Know	205	201	194	52	0
understand	234	194	162	61	27
Master	191	163	98	67	51
Application	29	51	11	69	90
Comprehensive application.	4	3	0	41	74
Total number	663	612	405	290	252

TABLE III. THE TEST RESULTS OF THREE ALGORITHMS

Value / Target	This algorithm	Traditional genetic algorithm	Backtracking algorithm
Difficulty error	0.060	0.119	0.321
Distinguishing error	0.058	0.127	0.346
Time error	0.084	0.166	0.229
Knowledge point error sum	0.869	1.827	2.567
Master degree error sum	0.381	1.115	1.625
Composing failure frequency	0	63	26
Average time-consuming(seconds)	16.8	14.9	37.5

It can be seen from the test data, the backtracking algorithm takes most time, achieve the effect is also the worst. This improved algorithm of time consuming minor longer than the traditional algorithm, but due to consider the hardness constraint conditions when produce feasible solution, so only the test questions number in item bank can satisfy, the composing test paper will not appear to fail.

VII. CONCLUSION

This article from the actual characteristics of "The Principle and Application of Database" ,etc course, establish the test questions evaluation standard system and the automatic test paper generation mathematical model based on genetic algorithm, propose the solutions combining specific questions for solving hardness constraint problems. Another introducing of elite strategy and improved crossover and variation operators, According to the Markov chain model, the algorithm in evolutionary algebra can satisfy constraint conditions, the algorithm test also reflects its effectiveness. This improved algorithm has implement in generic item bank developed based on J2EE technology. The system is running well, after several tests and a lower probability of generating repeat and similar test paper, the algorithm can satisfy the requirements of automatic test paper generation.

ACKNOWLEDGEMENTS

This paper is supported by the school teaching research project of North China Institute of Astronautic Engineering. The project number is JY-2008-015-Y.

REFERENCES

[1] Wang Wen-fa, Ma Yan, etc. The Application of Backtracking Algorithm in Multi Constraint Allocation[J]. Micro Computer Application. 2008,29(11):87-90.

[2] Lu Jing. The Study of Intelligent Composing Test Paper Based on Improved Genetic Algorithm[D]. Changsha. Central South University. 2007

[3] Ma De-liang, Lu Chang-hui, etc. The Intelligent Composing Test Paper Generation Based on Improved Genetic Algorithm [J]. Computer Application. 2009, 29(7): 1884-1886

[4] Yu Hong-chao, Fu Yun-qing. The Study of Automatic Test Paper Generation Algorithm Adapt to Network Education. [J]. Computer and Modernization. 2009,(1):66-68.

[5] Ding Wei-ping, etc. The Application Study of Intelligent Composing Test Paper Generation Based on Genetic Algorithm [J]. 2009,(1):66-68.

[6] Jia Jun, Kang Li-shan, etc. The Convergence Definition and Example Study of Multiple Objective Genetic Algorithm. Computer Application and Software. 2..6. 23(1): 1-3.