A New Computer-based Test System: An innovative Approach in E-learning

Maha E. K. AlSadoon and Rasha S. AbdulWahhab
Ahlia University
Manama, Bahrain
malsadoon@ahliauniversity.edu.bh & rasha_sh_abdulwahab@ahliauniversity.edu.bh

Abstract: - Computer-based system forms a rapidly changing in both technical and educational fields. Computer-based test system will help instructors to evaluate the learning status of students. This paper addresses the problem for generating a proper tests exam sheet by proposing new items test's selection strategy. The proposed system is composing from four units whereas the automatic test generator (ATG) is considered the main unit. To optimize the difficulty degree of the generated test sheet in ATG, an effective coding method and new heuristic genetic modification were used. Experimental results have shown that the proposed system can provide efficient test sheet. To evaluate the performance of the proposed system, two tests sheet with different difficulty degree has been tested on four different groups. The final results validated the feasibility of the proposed system for composing test sheet automatically.

Key-Words: - Genetic Algorithm, Test Generation, E-Learning.

1 Introduction

Many researchers have recently been concentrating how automated computer systems could help to compose test sheet automatically in order to measure the performance status of students[12][9]. In the context of education, test is considered the most popular measurement tool of learner's knowledge [13]. A well mine test supports instructors who want to verify whether students digest relevant knowledge and skills [9]. Moreover, several researchers have been debating that test influences evaluation of student's ability. For this reason, most of the investigation was to focus on the way of construction a well punctual test sheet [10][9].

In most of the existing algorithms to compose test sheet, the test's items are selected either manually or randomly [3][10]. These algorithms are unable to diagnosis students learning problems and therefore they will not reflect actual student learning results [3][9]. Hwang [10] is one of the leading researchers in the area of generating test sheet. He described the importance of selecting an appropriate test's items during exam construction. He also used an organized algorithm for composing good and efficient exams. The study of automatic computer assisted testing systems is examined by several researchers. Chenxia [1] pointed out that in order to generate good test sheets, as required in educational institutions; an efficient algorithm should be used for selecting appropriate items quickly according to the constraints proposed by the instructors. To cope this limitations, a number of heuristic approaches, such as genetic algorithm [17], immune [9], Tabu [7], particle swarm [16][14], have been proposed to generate optimal exams in a reasonable time. Lee [8] proposed a new system, which utilizes the characteristics of bidirectional concept beside GA to produce effective test sheets. The results of his proposal show that it is possible to produce an optimal test sheet that fit the learning target. Another proposal is made by [15]. He proposed an educational tool for helping instructors to select appropriate items from item bank with multiple restricted condition item strategy. In his approach an improved evolutionary technique of Simple Genetic Algorithm (SGA) has been applied. While in Hwang [6] was suggested as an optimization techniques to maximize (or minimize) some objective function by combining several measures such as degree of difficulty, discrimination and number of test items. He illustrated that the required time to set up an optimal test sheet grows exponentially especially when the number of test items in an item bank are large.

This paper examines two issues about developing an intelligent test generating system. We propose an automated testing system in which an appropriate item selection strategy is used. In addition, a novel genetic algorithm was used to support the proposed automated computer systems testing system. The novelty of GAs was achieved in this paper by proposing a new representation method and a new crossover operator. The proposed crossover operator is used as a heuristic method to enhance tests and convert them into feasible one. Thus we should expect some improvement in performance to solve test generating problem from the utilization of this new outline of GA.

The structure of the paper is as follows. The research objectives are illustrated in section 2. The main units of the proposed testing system are presented in section 3. Section 4 provides the experimental results of the proposed system and discusses those experiments. Finally, section 5 includes the overall conclusions and the directions of future work for extending the proposed system.

2 Research Objectives

The aim of this study is to develop an efficient computer-based test system that is used to generate a valid, reliable and high quality tests sheet automatically. In order to manage the requirements identified by the instructor, an appropriate items' selection strategy should be used. In such system, the interaction between system's interface and instructors is required. In order establishing these objectives, a novel GA to improve the efficiency of composing near optimal test sheet is examined.

3 The Proposed System

In this section, intelligent computer-based Test Sheet system using a novel Genetic Algorithm (ITNGA) is presented. ITNGA is aimed to optimize the difficulty degree of the generated test sheets with a fixed number of test items. The developed system consists of four main units outlined as shown in Fig 1. Each unit in ITNGA is responsible to feed the required information to generate an optimal tests sheet. The main unit of ITNGA system is the Automatic Test Generator (ATG) which is responsible for generating an automatic tests sheet. These units are explained in the following:

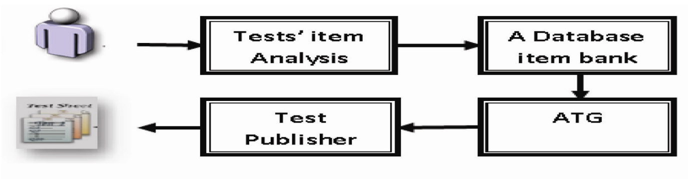

Fig1: ITNGA system units.

3.1 Test's Item Analysis

The difficulty degree of the test items in the item bank are set after analyzing the progress of the course by the lecturer referencing to the result obtained by other experts. In defining the difficulty degree, D_i refers to difficulty degree of the ith test item. In proposed system the difficulty degree is defined to range from 0.0 to 1.0 comprising five different grades which are: D_1: Very easy [0.8 - 1.0], D_2: Easy [0.6 - 0.8], D_3: Moderate [0.4 - 0.6], D_4: Hard [0.2 - 0.4], D_5: Very hard [0.0 - 0.2].

3.2 Database of ITNGA

A database item bank must be created for providing ITNGA system with the required information. This Database will be used to generate an automatic test sheet. The basic structure of this database is made of a set of tables where information about a particular entity (Course,Concept,Question,ect.) is represented in columns and rows. The item pool adopted in ITNGA system is the test pool of "Introduction to Programming" for information technology taught as an undergraduate course at Ahlia University. This item pool consists of six different concepts indicated as chapters. Each test for a particular concept includes two different types of items including multiple choice and true/false items.

3.3 Automatic Test Generator (ATG)

ATG in ITNGA system is responsible for selecting the items that will be used to generate the optimal test sheet for students. The generation process will be guided by instructor test specifications.

3.3.1 Retrieving the Potential Items Test

To generate genes of tests sheet, an item pool from the database should be retrieved. ITNGA system will send query to the database model for retrieving the required items based on instructor predefined parameters. The retrieved items are saved in a virtual table called TempQTable. However, to reduce the search space and the time required for retrieving items form database, the engine of SQL is integrated with ATG. Thus a value to the performance of proposed system is added from the utilization of such routine.

3.3.2 ATGs' Initial Tests

Hwang employs binary code to represent the test sheet as either 1 or 0 indicating that the test item is currently selected or not [5]. However, extra time is needed with 1 or 0 coding method for determining the number of test items. Therefore, transferring 0 to 1 will increased the time. To avoid this limitation and increase the speed of test items selection with the large database, an integer number coding is adopted in ATG to construct the test sheet. This type of representation will reduce the complicated of searching and data matching.

Let *pop* denotes the set of initially generated test sheets while *ps* represent the size of the population of *pop*. Each TS_k in *pop* is consisting of set of genes *g* indicating the test item which is relevant to one or more of specified course concepts C_x. Each item is generated as a random number and distributed uniformly over the range [1..n], where n is represent the predefined size of item bank. Each TS can be expressed as [g_1,g_2,... , g_{qnum}], where $g_i \neq g_j$ for $1 \leq i \neq j \leq n, TS \in pop$, and q_{num} represents the number of items in the TS. As mentioned earlier, all tests sheet in the initial population have a fixed length. The tests sheet included in the initial population can be created by using a random procedure to achieve a certain level of diversity.

3.3.3 Fitness Function

Choosing an appropriate evaluation function is an essential step for successful application of GAs to any problem domain. Roughly speaking, one can say that fitness value of each test sheet in the current population reflects its merit to survive in the next generation. For this reason a fitness value for each test sheet must be calculated. The fitness function in the test generator evaluates the quality of each test sheet. To evaluate the fitness of test sheets, the following indicator E_D should be calculated:

$$E_D = \sum_{j=0}^{q_{num}-1} |D_j - D| \qquad 1$$

Where E_D is the summation of the differences between the selected test items and the objective difficulty value. Whereas D_j represents the difficulty degree of the selected item of test sheet while D is the expected difficulty weight of ith test sheet that is determined by the instructor. The fitness function used in ATG is shown in 2.

$$F(TS) = \frac{1}{1 + E_D} \qquad 2$$

The objective function value of test generator is rescaled to a fitness value of range between 0.0 and 1.0.

3.3.4 Selection Operator

The tournament selection is adopted in ATG. In a single iteration a tournament operates by selecting two parents (tests sheet) usually called tournament size from the population chosen at random. In a tournament of size two, a best fit test sheet wins the tournament. A copy of the winner after applied genetic operators such as crossover will replace the loser or replace a randomly chosen test sheet of the population [2]. The other test sheet is selected in the same way.

3.3.5 Heuristic Modification Operator

Reproduction directs the search toward the best existing individuals (test sheet) but does not create any new individuals. The new tests sheet are principally generated by crossover. Crossover operator allows the tests sheet in the mating pool to exchange their genetic martial such that better combination of genes can be preserved in their offsprings. In some implementations of the GA, simple crossover operator presents no problems. For example, if every possible individual permutation is valid within binary representation individuals, crossover can be allowed to proceed without interference problem. However, the crossover operator of standard GA cannot be directly applied for generating tests sheet, since integer representations of individuals have to be used and these operators may produce illegal offspring i.e. duplicated genomes (items) on the same test sheet especially when the numbers of items related to a specific concept (search space) are small. In ATG a newly generated test sheet may becomes a member of the *pop(t)* where t in [1..*ps*], given that the following conditions are satisfied:

1. The length (number of items) of the new offspring (test sheet) must be exactly the same as its parent length.
2. No duplication (i.e. same item) is allowed to occur within each offspring.

In order to cope with the above constraints, an enhanced crossover operator is proposed. This new operator follows the principle of one point crossover as well as the ability to clean up and preventing the duplication of items in the newly generated trail tests sheet. Tow parents TS_1 and TS_2 of q_{num} length are aligned with each other and one crossover locus cp is chosen at random over the range$[1..q_{num}]$. The selected crossover locus will divide TS_1 and TS_2 into two parts. The proposed operator will produce two new test sheets $TS_{\bar{1}}$, and $TS_{\bar{2}}$. The first part $TS_{\bar{1}}$'s items $[g_{0\,..}\, g_{cp}]$ are copied from the original parent TS_1. The tail of the second parent TS_2 from the onward point are switched to create the remaining items of TS_1. Through the job of coping the items from TS_2, one must consider the uniqueness of the test items in $TS_{\bar{1}}$, i.e., each test item can occur only once. This is implemented by using the cleanup operator. The cleanup operator is called to check the uniqueness of g_i in $TS_{\bar{1}}$, if it is already in $TS_{\bar{1}}$, one item from $tempTS_1$ is chosen. $tempTS_1$ is obtained by ordering the items of TS_1 in an ascending manner. The same process is repeated until q_{num} items are generated. The same operation in enhanced crossover are repeated for generating items of $TS_{\bar{2}}$. This operator guarantees the correctness of the final production of $TS_{\bar{1}}$, and $TS_{\bar{2}}$.

3.3.6 Stopping Criteria

The evolution of natural organism dose not stops unless it extinguished. However, for the simulation of test generator program, a stopping criterion must be setup. There are three different criteria used to decide when test generator is terminated, these include:

1. An upper limit on the number of generations is reached, or
2. An upper limit on the number of evaluations of the fitness function is reached which is equal to 1.0, or
3. The chance of achieving changes in the next generations is extremely low.

3.4 Publisher Unit

In this section, the communication between the item bank database and other components in ITNGA system to publish the test sheet is explained. The publishing means the processes of formatting and packaging tests for delivery within the selected testing requirements. This process of publishing is demonstrated through two different tasks, in which an embedded SQL is used. As mentioned earlier, ATG's best solution is a list of integers which when input a suitable generator causes that generator to output the final test style. The mapping from the integer solution to exam test sheet is called OntoTest mapping. Fig 2 outlines the diagram of OntoTest in ITNGA system.

Fig 2: The diagram of OntoTest mapping.

OntoTest mapping has two different tasks. The first task is to map integer best solution produced by ATG into a dynamic table. Algorithm 2 presents the pseudo code of the first task of OntoTest generator.

Algorithm 2: Pseudo code for implementing the first task of OntoTest mapping.

$TS_{\bar{b}}$ indicate the best test sheet after completing GA execution
For $i \leftarrow 0$ **To** $q_{num} - 1$
$Exam \leftarrow \sigma_{QID,\ Q,\ QType}\ (\pi_{\ QID\ =\ TS_{\overline{\bar{b}\ (i)}}})\ (Question)$
End For

To link the possible items to their correct answers a second task will be called. Algorithm 3 presents the pseudo code of the second task provided by OntoTest mapping.

Algorithm **Error! No text of specified style in document.** : Pseudo code for retrieving the answer ID from the database.

For $i \leftarrow 0$ **To** $q_{num} - 1$
$\sigma_{QID,\ Ans}\ (\pi_{\ Question.QID\ =\ Exam.QID}\ )\ (Question \cup\ Exam)$
End For

4 Experimental Results

The experimental work for evaluating and analyzing ITNGA system is presented in this section whereas the evaluation is conducted in terms of the following criteria:

1. Show that ITNGA is applicable at Ahlia University domain.
2. To get a good solution for generating test sheet.
3. See how an ITNGA scales-up with population size (*ps*).
4. Find best setup for parameter of test generator, i.e., *ps* and maximum numbers of generation t_{max}.

4.1 Success Rate of ITNGA System

In this section an empirical study were conducted to efficiently determine the performance and the usefulness of the proposed ITNGA system for constructing desirable tests sheet. In this experiment, the success rate is computed, i.e. the success rate represents the percentage run's terminating with success. The success rate is computed by testing ITNGA system with 100 independent runs with different values of ps and t_{max}. The criteria for sufficient solution quality is given in this experiment by neglecting all solution that are far from that optimality, i.e. their fitness value varies between [0.7 ... 1.0] are accepted to be a solution. Table 1 illustrates the success rate of ITNGA with different values of ps, q_{num} and t_{max}.

max_g	*ps*											
	25			50			75			100		
	20	40	60	20	40	60	20	40	60	20	40	60
q_{num}=5	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
q_{num}=10	74%	75%	86%	99%	100%	100%	100%	100%	100%	100%	100%	100%
q_{num}=15	6%	14%	16%	53%	85%	85%	91%	99%	97%	98%	100%	100%
q_{num}=20	0%	2%	2%	4%	40%	38%	15%	88%	89%	25%	98%	100%

Table1: The probability of successes of ITNGA with different values of q_{num}, ps and max_g

The realized conclusion by this experiment, large values of ps and t_{max} must be set when the requested number of test items is high. In this circumstance, ITNGA system is able to construct an optimal tests sheet quickly or easily. Once the ps value is increased, the probability to find optimal tests sheet in the early generations are increased. ITNGA is able to increase the probability of success for producing optimal test sheets if the appropriate value of ps is large.

4.2 The Applicability of ITNGA

The objective of this section is to show the applicability of ITNGA system at Ahlia University domain. During the Winter Semester of 2010, four groups are chosen to examine them by the test sheets generated by ITNGA system. The four groups met at similar times i.e. after cover "Chapter 2" and "Chapter 3". The total number of students in the first and second group was made up of 60 students. All students in both groups are in the same college in addition they have similar knowledge about the programming. Two tests were conducting each test consisted of twenty items related to the concept of "Chapter 2" and "Chapter 3" only. The total score of test is between 0-10. The difficulty degree for the first and the second test is equal to 0.8 and 0.3 respectively. Table 2 presents the success percentage of four groups.

Test	**Chapters**	**Item Type**	**Difficulty Degree *D***	**Success %**
First	2 and 3	Multiple choice and true/false	0.8	100 %
Second			0.3	66 %

Table 2: Students reselts.

Depending on the students' grades, the following points are recorded:

1. The percentage of success for the first test was 100 % which mean that all students are able to pass the test because of the difficulty level for the first test was equals to 0.8.
2. The percentage of success for the second test was decreases to 66 % that is because of the difficulty level for this test was equals to 0.3.

5 Conclusion

In this paper, ITNGA system is proposed to cope with an automatic test generating problem. ITNGA system consists from four different units whereas ATG is considered the main unit. ATG is used to generate tests sheet. To evaluate the performance of ITNGA system a series of experiments were conducted on 2000 item bank (those experiment based on 100 independent runs). The most perspective that can be concluded by implementing ITNGA system are as follows:

1. It is noticed that the performance of the proposed system is affected by population size ps.
2. By employing ITNGA system, the experimental results show that an optimal test sheet can be generated in acceptable time.

Based on the results of the empirical test were conducting on four groups of ITCS 122 students, the ITNGA system can be suitably applied to evaluate the students performance, and to practice in actual evaluation system in the future work. The study also provides quite for future researcher which will enable instructors to enhance student's ability to acquire skills and apply knowledge. This in turn will not only save time, but also improves the quality of the system to generate new methods in teaching. Assess to modern technologies allow greater transfer of knowledge as well as sharing information. Instructors are to benefit from these methods by communicating with students more frequently. The proposed system needs to be incorporated and integrated in future research in order to enhance the capability of educational system in Bahrain to achieve excellence. Such system will enable instructors to construct an effective and qualified tests sheet. This in turn will increase instructor efficiency to conduct their tests and generate test items. To do so, the following steps can be addressed in future research:

1. Adding online features.
2. Add student portfolio that will record student's performance and upgrade the items statistical such as difficulty degree and discrimination degree according to their records with some constraints.
3. Provide the students with their marks in addition to the feedback to improve their knowledge.

References:

[1] Chenxia, S.; Bo, L. Z.~J. and Chao, W., An adaptive algorithm for generating question paper wit multi-factors. In: International Conference on Computer Science an Education, pp 1048--1051,2009.

[2] Eiben, A. E.;~Smith, J.~E., Introduction to Evolutionary Computing, Springer,2007.

[3] Fengrui~W., Wenhong~W., Q.~P. and Z., F., A novel online test-sheet composition approach using genetic algorithm. In: Proceedings of the 4th International Symposium on Advances in Computation and Intelligence, Springer-Verlag Berlin Heidelberg, pp 1--10,2009.

[4] Hwang, G.J.;~Lin, B. M. T. L.~L., An effective approach for test sheet composition with large scale item banks, Computers and Education,v46, pp 122 - 139,2006.

[5] Hwang, G. J.;~Lin, B. M. T. T. H.~H. and Tsung, L.~L., On the development of a computer-assisted testing system with genetic test sheet-generating approach, IEEE Transactions on Systems, Man, and Cybernetics, v35, pp 590 - 594,2005.

[6] Hwang, G.~J., Development of an intelligent testing and diagnostic system on computer networks, In: Proceedings of the National Science Council of ROC. 9(1),pp 1-9,1999.

[7] Hwang G. J., Yin P. Y. and Yeh S. H., A Tabu Search Approach to Generating Test Sheets for Multiple Assessment Criteria, Journal, IEEE Transactions on Education, n1, v49, Feb,2006.

[8] Lee, C. L.;~Huang, C. H.~L. and J., C., Combination of Bidirectional Conceptual Map and Genetic Algorithm for E-Learning Evaluation System application, volume 6. Springer Berlin / Heidelberg,2007.

[9] Lee, C. L.;~Huang, C. H.~L. and J., C., Test-sheet composition using immune algorithm for e-learning application. Lecture Notes in Computer Science, v4570, pp 823-833, Springer Berlin / Heidelberg,2007.

[10] Lin, H. G. J. C. H. C. Y. P.~Y. and Y., J., An innovative parallel test sheet composition approach to meet multiple assessment criteria for national tests, Computers and Education, v51,2008.

[11] Lira, P.;~Bronfman, M. E.~J., Multitest II: A program for the generation, correction, and analysis Journals, Education, IEEE Transactions, v33,1990.

[12] Liu, P. G. L.-L. and Yao, Q., An intelligent test sheet composition system. In: Proceedings of 2008 IEEE International Symposium on IT in Medicine and Education , pp 503 -- 507,2008.

[13] Meng, A., Y.-L. R.~D. and Padilla, P., Genetic algorithm based multi-agent system applied to test generation, Journal Computers & Education, v49, n4, pp1205--1223 Dec,2007.

[14] Ming, L.; Jin-gang, M. and Jing, Z., Automatic generating test paper strategy based on improved particle swarm optimization, in ITIME 09, IEEE International Symposium, pp 711--715. IT in Medicine and Education,2009.

[15] Sun Koun-Tem, Lin Y., Huang Y., An efficient genetic algorithm for item selection strategy. Technical report, ICS, 2006.

[16] Yin P. Y., A Particle Swarm Optimization Approach to Composing Serial Test Sheets for Multiple Assessment Criteria, Journal, Educational Technology Society, n3 : V9,3 - 15,2006.

[17] Zhou Y., Li Y. and Feng C., Research on Intelligent Test Paper Auto-Generating Algorithm Based on Improved GA, In: Proceedings of the 21st annual international conference on Chinese control and decision conference,3938 -- 3942. IEEE Press Piscataway, NJ, USA,2009.