Electronic testing system

Boyka Gradinarova and Ognjan Jelezov
Technical University of Varna, Bulgaria

Abstract: Information technology's fast development makes the education process more sophisticated, either in terms of course material presentation or testing and assessment realisation. In this paper we introduce a system for on-line electronic testing (ETS), developed at the Technical University of Varna, Computer Science and Engineering Department. Compared to similar programs such as WebCT, this system offers improvements such as input simplicity, and the enhanced possibility to implement html and other code including more pictures, video, tables and animations.

Key words: web technologies, testing system

1.   INTRODUCTION

This paper reports a 12 month development of concepts and software for computer-based testing and its implementation at the Technical University of Varna, Computer Science and Engineering Department. The original idea was to create a system that could help create the examinations for a group of students larger than 150 in number and allow for the concept of each student being allowed to apply for an examination each month. This concept would create a great deal of effort for the lecturer, as (s)he has to generate different tests for more groups of students, and use much more effort to check tests and assessments. We will present the ideas that motivated us to develop specialised software through ideas of tree category test generation, data entry and report generation.

2.   MOTIVATION

The test generation concepts were defined at the start of this project, as our primary goal was to help lecturers set examinations every month with different sets of questions for a relatively large number of students. Afterwards these ideas were expanded to a concept that allowed the computer to generate different tests each time the student applied for a test. Of course this was created not only for different questions, but had to be realised for the same type of questions that represented the same concept but had different data, the way the sentence was asked, the way the answer was answered, etc. Also the concept had to cover areas of questions where one question concerning a knowledge skill was asked. Test generation was made simple by being based on random generation of the line number of questions in the given area. However, the possibility of applying for an examination each month meant that students could learn by heart a set of questions and line numbers of the answers instead of learning the concepts. Therefore we created computer-based testing with random positioning of answers, so the position number could mislead students if they were learning the position numbers by heart. Thus, instead of memorising questions and the line numbers of the answer options, students were forced to learn the relations between questions and answers, i.e. concepts and other knowledge skills. Therefore the concept created in our program includes random generation of questions in a given area and random position of the answer options.

Ideas for data entry were based on complete interactive multimedia concepts in a windows-oriented operating system. Two individuals were responsible for data entry: the lecturer who filled the data base of questions, and the student who applied for an examination. The student data entry area used a classical interface with windows dialogs, radio buttons, multi-option selections, confirmations and action buttons. The visual display was standard, using conventional internet browsing techniques.

The concept of areas and sets of questions had to allow the lecturer to define the question relatively easily, without standard dialog windows which asked what the question was, what the given answer options were and which option was correct. Since the area consisted of a set of questions that had different values for some variables, but the question and text in the question were the same, we developed the concept with an interface that allowed very fast data entry and data base connection.

The lecturer had the opportunity to generate several types of reports. He was able to print the selected areas and questions from selected areas, to correct some of the questions or to add new questions, and to define test strategies. According to the tree organisation of the questions and areas, he

was able to delete from, add to, or recognise the structure of the areas and questions.

Test re-evaluation was enhanced, since the lecturer was capable of seeing the statistics concerning the examination, the number of students who applied for it, the number of students who were successful, the questions not answered, the average percentage of answered questions and given points, etc. The lecturer was also capable of examining a given student's test and helping him to make an assessment of the student's knowledge. All data collected could be used within a spreadsheet and structural packages.

3.   DEFINITION OF ETS

A lot of authors have tried to define the terms concerning computer-based testing or electronic testing systems. For example, Olsen (see web-site) tried to locate the development and administration of test and assessment applications. According to Olsen, computer-based tests are defined as tests or assessments that are administered by computer in either stand-alone or networked configuration or by other technology devices linked to the internet or the world wide web. A very useful definition of testing is given in Landon (see web-site) as a technology that includes facilities to assist in the making up of practice quizzes, tests, examinations, and other assignments.

If we agree on this definition, we conclude that computer-based testing is a method for testing by using computer hardware and software. As a follow-up we define electronic testing systems as dynamic computer systems that support computer-based testing. They include various methods for data base management and artificial intelligence to support decision-making processes for the generation of testing material, realised as web technologies.

4.   COMPARISON TO OTHER SOFTWARE
PRODUCTS

A really useful survey of the comparison of on-line course delivery software products can be found as a web tool (Landon, see web-site). We have used some of these data to develop computer-based testing as an efficient tool for different purposes like conventional education, not just distance learning or on-line learning software tools, job-hunting quizzes, etc. It is assumed that nearly all the products that are examined include multiple

choice questions, true/false questions, matching questions, short answer and quantitative questions.

As our tests show, and as others conclude, only some software products can generate a question data base, support reporting features and on-line testing and grading. Not all software products supply self-test tutorial questions, timed quizzes or generate random sets of questions. The following is a list of software products that supply on-line testing and is ordered by the number of web-based hits (Landon, see web-site): WebCT, Black Board, Learning Space, Top Class, IntraLearn, eCollege.com, Web Course in a Box. Landon also generated and tested benchmark tests for several software products, grading the following topics in the field of computer-based testing: can easily create course rosters; easily check that data is represented correctly in web browser; easily create a data file using minimal coding beyond basic word processing; has a record-keeping function; student can access tests and grades at all times; does comparative statistics on grades; customisable grade report; can easily administer web-based quizzes; instructors can create quizzes within the program; can create a ten-question quiz within an hour; can easily utilise a variety of question types; can easily allow the student to view the answer key; students can check/change responses before submission; supports instructor's on-line grading of quizzes; provides scoring data in an easy-to-read/use format.

So far there were no capabilities identified to generate randomised questions, a randomised order of answer options, selecting from a question data base, generating a test by defining the strategy to choose questions from selected areas, etc. That was our real motivation when producing this type of software.

5.   PRACTICAL IMPLEMENTATION

The use of the internet as a base for developing technologies is a subject reviewed by many authors. A useful survey of the internet for business infrastructure can be found in Jelezov (2001). In our environment, in the version of the software we produced, we use Microsoft IIS Web Server and a Windows operating system, ASP as a scripting language, Microsoft SQL server as a data base server and internet web-browsing technology.

Each test measures verbal, quantitative and analytical skills related to a specific field of course study. A different time constraint and score mark is associated with each question. We distinguish three classes of questions: verbal, quantitative and analytical questions, similar to those at the web site www.gre.com/cbttest.html.

A verbal question defines a specific concept of definition and usually 2 points are able to be gained in 60 seconds. The verbal measure tests the ability to analyse and evaluate written material and to synthesise information obtained from it, to analyse relationships among component parts of sentences, and to recognise relationships between words and concepts.

In the case of quantitative questions where a specific quantitative measure is expected to be calculated, or a more sophisticated concept to be explained, usually 4 or 6 points are able to be gained in between 120 and 240 seconds. The quantitative measure tests basic skills and understanding of elementary concepts, as well as the ability to reason quantitatively and to solve problems in a quantitative setting or to explain more sophisticated concepts.

For an analytical question, where a more complicated formula is used to calculate the expected answer, or a rather difficult concept has to be explained, then up to 10 points are given within a time constraint of up to 360 to 480 seconds. The analytical measure tests the ability to understand structured sets of relationships, deduce new information from sets of relationships, analyse and evaluate arguments, identify central issues and hypotheses, draw sound inferences, and identify plausible causal explanations. Questions in the analytical section usually measure reasoning skills developed in virtually all fields of the study course.

6.   ETS CHARACTERISTICS

The program consists of a fully web internet operational module. After authorisation of the user as user or lecturer, the appropriate menu shell is opened. Students are given typical options to choose from a self-testing option, examination or report tool. The lecturers can choose from several options such as data entry tool, the tool for question data base editing and organisation, test and examination definitions, and report tool.

The self-testing procedure allows the students to become familiar not only with the multimedia interactive internet web technology to realise the test, but to become familiar with the course material and concepts to be learned. These sessions can be carried out in an informal atmosphere and students are allowed to talk freely or refer to notes. The self-testing procedure is a realisation of web-based testing, that offers 24-hour access to testing, immediate scoring, and to a reduced need for test administrators, leading to convenient, cost-effective and efficient testing.

Although a lot of the questions available for self-testing were the same, there still were questions in the data base which were solely for the

examination. Also many modifications to the original questions for self-testing were made.

We implemented non-adaptive tests since these tests can be used efficiently for computer-based assessment. In these tests the student is allowed to skip a question and to go onto the next question. Afterwards the student is allowed to come back to the questions not answered and to spend more time thinking about the correct answer. The availability to confirm the answer before it is sent to the program is also another benefit that makes the student more comfortable with the test. In the case of multi-choice questions, the student can answer one or all possibilities of the given answer options. He must understand how the test is scored, since wrong answers bring negative points. It is better not to answer the question instead of clicking by guessing. Each correct answer scores the given percentage of the points, whilst each incorrect answer scores negative points equal to half the percentage of the points obtained as a correct answer. This methodology is used in addition to the strategy that no question has all correct answer options and there must be at least one incorrect answer option. If the student tries to cheat and clicks on all answer possibilities, he scores negative points.

The data entry tool allows an easier definition of questions in all segments. Although there are conventional dialog windows when asking for entry, the question itself, the possible answer option scoring points, time allowed and other details were found to be more attractive in a macro definition and we used conventional MS Word and spreadsheet processing programs to define questions and then activated a special macro to transfer the questions into the data base. Not only is this method more attractive, but it is the fastest way to define a bulk of questions, especially if we define similar questions with the same answer options.

The tool for question data base editing and organisation enables the lecturer or the tutor to organise the tree structure, to establish new areas, to rename names, to move questions from one to another area, to delete a question or an area, to define sub and super areas, and to change question or answer option body text. Also the test administrator is allowed to enable or disable each question for the self-testing procedure for the examination.

The tool for test and examination definitions states the definition of the strategy for the examination and testing. It can be done by looking at the previous strategies or by defining a new strategy, etc. Test administrators can efficiently use the report tool to see student progress, the student's test and to check his knowledge or absence of specific skills. The lecturers are allowed to see each student log file of activities and the test itself.

7.   RESULTS

The area of global testing is developing and is further supported by the International Test Commission (2000). Computer-based testing, including PC-based, web-based and computer adaptive testing, has seen a continuous increase since the 1980s (Lyman, 1998). The Association of Test Publishers (2000) has developed computer-based testing standards. A useful survey of mega-trends towards personal testing, whether used for job-hunting, interviews, scoring, etc. is given in Jones and Higgins (2001).

Many authors have stated that advances in technology significantly increase the acceptance of computer-based testing as an effective means of delivering assessments for a wide range of testers. As Bokij and Greasley show in their article (2000) computer-based testing gives higher scores in comparison to conventional methods. One of the reasons for these results is the fact that computer-based testing tends to be completed more quickly than conventional assessments. As the time taken to complete the assessment decreases, performance increases. Some other factors concerned with improvements from our research reported that students felt more comfortable and relaxed when seated in a laboratory as opposed to an examination hall: it was quicker to work through than a conventional multiple-choice test. Some typical comments were: "Not as much panic involved – to sit, read and click seems to take away that element of stress"; "Removal to write down the answers allows more rapid test competition"; "The computer allowed you to think rather than wasting time trying to write things down"; "Because the computer checks you have decided on your answer, so you always re-read it to make sure your answer is the right one".

Moreover, students feel that assessment by computer-based testing is more objective, since the same criteria are used for everyone, so the computer-based testing is less threatening than conventional examinations, is more accurate and tests a wider range of skills than a conventional examination. However, it should also be clear that computer-based testing must not be seen as a 'quick-fix' for problems such as rising student numbers. If one adopts a considered and methodological approach to computer-based testing, positive benefits might include both increased efficiency and quality (Bocu, 2000).

The practical experience of carrying out the exams for more than 150 students by computer-based testing was gained in three computer labs under formal examination conditions. In order to deal with the relatively large number of students that took part in the examination, three computer rooms were used simultaneously and students were given a time schedule for the examination. Students were assigned to a specific computer room and required to attend at a specific time and place. Security measures were taken

to prevent cheating during the test. These included staff supervision of assessments, password protection, randomised questions asked, randomised answer options given and encryption of the results files.

The students were allowed to use a self-testing procedure. In this case they became familiar with the test questions, they were prepared, and unexpected surprise was less for them than in conventional testing. Once the students knew what to expect on the test, they practiced and prepared, so they put more concentration on the subject instead of psychological stress and fear from the unexpected.

It was interesting for us to see a large number of students using self-testing. This process led to more motivation, since while they were learning, they tested their knowledge and accomplished the process of learning more deeply. The process of assessment of the knowledge and correction of the learned concepts during the test evaluation brought very good results from the students. Compared to the results obtained by applying the same procedure on different courses, or by applying the same computer-based tests and conventional testing, we concluded that the motivated students ended up with better marks. The electronic testing system enabled the students to learn more skills in a shorter time and more efficiently.

REFERENCES

Association of Test Publishers (2000) Guidelines for Computer-Based Testing. Washington, D.C.: Association of Test Publishers

Bocu, P. and Graesley, A. (2000) Can Computer-Based Testing Achieve Quality and Efficiency in Assessment? International Journal of Educational Technologies, 1, 1

Graduate Record Examination. Description of General Test. Available at: www.gre.com/cbttest.html

International Test Commission (2000) International Guidelines for Test Use. Netherlands: ITC

Jelezov, O. (forthcoming) Infrastructure for Electronic Business on the Internet.

Jones, J.W. and Higgins, K.D. (2001) Megatrends in Personnel Testing: A Practitioners Perspective. Association of Test Publishers. Available at: http://www.testpublishers.org/journal03.htm, Jan. 2001

Landon, B. Online Educational Delivery Applications: a web tool for comparative analysis. Available at: http://www.c2t2.ca/landonline/index.html.

Lyman, H.B. (1998) Test Scores and What they Mean. Sixth edition. Boston: Allyn and Bacon

Olsen, J.B. Guidelines for Computer-Based Testing. Association of Test Publishers (ATP). Available at: http://www.isoc.org/oti/articles/0500/olsen.html

BIOGRAPHY

Boyka Gradinarova is a lecturer in the Department of Computer Systems and Technologies at Varna Technical University.