AN EFFICIENT TOOL FOR ONLINE ASSESSMENT AT THE POLISH-JAPANESE INSTITUTE OF INFORMATION TECHNOLOGY

Paweł Lenkiewicz, Małgorzata Rzeźnik and Lech Banachowski
Polish-Japanese Institute of Information Technology (PJIIT), Warsaw 02-008, ul. Koszykowa 86, Poland

Keywords: e-Learning, Distance Learning, Online Learning, LMS, Platform EDU, Online Assessment.

Abstract: The purpose of the article is to present the development over the years of the tool for online assessment along with its usage in the online studies at the Polish-Japanese Institute of Information Technology (PJIIT) (Banachowski and Mrówka-Matejewska and Lenkiewicz, 2004; Banachowski and Nowacki, 2007).

1 INTRODUCTION

Polish-Japanese Institute of Information Technology (PJIIT) is a leading Polish university specializing in Computer Science. It was founded in 1994 as a result of an agreement between the governments of Poland and Japan. The Institute offers undergraduate, graduate and postgraduate courses in the main fields of Computer Science. Since 1994 the lectures at PJIIT have been gradually digitized. What is more, the Internet has become a powerful medium of communication between students and faculty and between PJIIT and partner universities in Japan. In the year 2001 we started teaching on-line courses on an experimental basis in cooperation with the University of North Carolina, Charlotte. The participants of the courses were graduate students from both the universities. In June 2000 the Senate of PJIIT took a decision to start preparation for online studies towards B.Sc. degree in Computer Science. The new studies commenced in September 2002 supported by specially designed and built dedicated LMS system Edu. In 2006 the studies were extended by the addition of the studies towards M.Sc. degree in Computer Science. Since 2008 we have offered also online postgraduate studies in Computer Science. These studies are directed towards IT specialists as well as specialists in other domains who want to use IT solutions supporting their everyday activities at work, specifically in conducting software projects. These studies constitute the first step to promote continuing education and life-long learning. This year we have started developing an on-line system for supporting processes of life-long learning of PJIIT's students, graduates and academic teachers (Banachowski and Nowacki, 2009).

Each online student has to come to the Institute for one-week stationary sessions two or three times a year. During these visits they take examinations and participate in laboratory courses requiring specialized equipment. The studies are based on the educational, multimedia materials available on-line.

The courses run either exclusively over the Internet or in the blended mode: lectures over the Internet and laboratory classes at the Institute's premises. Each course comprises 15 units treated as lectures. The content of one lecture is mastered by students during one week. At the end of the week the students send the assignments to the instructor and take tests, which are automatically checked and graded by the system. The grades are entered into the gradebook - each student can see only his or her own grades.

Besides home assignments and tests there are online office hours held once a week; seminars and live class discussions. Bulletin boards, timetables, discussion forum and FAQ lists are also available. It is important that during their studies the students have remote access to the PJIIT's resources such as software, applications, databases, an ftp server, an e-mail server. Partial grades obtained during the semester (coming mainly from home assignments and tests) contribute to the final grade for the on-line part. Besides this grade we have always the second grade resulting from the face-to-face examination

Lenkiewicz P., Rzenik M. and Banachowski L. (2010).
AN EFFICIENT TOOL FOR ONLINE ASSESSMENT AT THE POLISH-JAPANESE INSTITUTE OF INFORMATION TECHNOLOGY.
In *Proceedings of the 2nd International Conference on Computer Supported Education*, pages 362-366
Copyright © SciTePress

administered in the PJIIT building.

Over the years enrollment in online studies has been growing, as has the number of teachers using the PJIIT's e-learning platform to enhance the quality of their stationary courses. At one point there appeared the need for an effective and customizable testing system. The old system did not allow generation of multiple versions of a test, or ascribing weights, or difficulty levels to test questions. The present paper describes how the new system was built.

2 WHY AUTOMATED ONLINE TESTING?

As we have mentioned in part 1, testing is closely connected to assessment and in most courses is used alongside other methods of evaluating students' progress. Not all tests are done for the purpose of assessment though (Bachmann, 1990) – some quizzes may only function as exercises providing training opportunities (or the fun part of a course). The type of test may also vary. While by the term *test* we usually mean a set of questions requiring short or multiple choice answers, some examinations may involve essay writing, and they will still be called tests (e.g. Test of English as a Foreign Language). The system discussed in the paper follows the standard meaning of the word *test*, i.e. it is a system which allows generation of sets of questions which would be characterized by parameters pre-defined by the teacher or the examiner. The system is simple enough to enable formative (ongoing) assessment using short tests and sophisticated enough for the purpose of summative (final) assessment which requires the development of large, complex tests.

Modern teaching methodology stresses the importance of incorporating technology into all types of educational endeavors. The rising popularity of e-learning and blended learning (the mixture of the traditional and virtual classroom environments) seems to prove that this is the right course of action. Whilst nowadays there is a trend towards more collaborative learning, involving higher order information processing, testing, preferably on a computer, is still a vital part of any well-designed course. In fact the scale of automated testing seems to be growing rapidly. There are several reasons (Bednarek and Lubina, 2008; Chapelle and Douglas, 2006; Paloff and Pratt, 2003):

1. Automated tests are checked quickly, and learners (or their teachers) get almost instant feedback. The importance of immediate feedback cannot be overestimated. It is one of those features which draw people to e-learning. Automated tests are ideal for formative assessment and course monitoring.
2. Automated tests are scalable – the argument crucial to educational institutions running examinations. They may also allow creating several different versions of the same test to prevent cheating.
3. Computerized testing may allow tests' self-adjustment to the level of a test taker. That is important in, for example, language examinations, as it shortens the time a test would take, while at the same time increasing its accuracy.
4. They free the faculty from tedious test checking.
5. The can be easily modified and updated.
6. They are reusable, especially for formative assessment. In the case of summative assessment this is only possible if there is a large enough repository of test questions to be chosen from.
7. They can be carried at a distance, or at the school's premises, depending on the purpose and type of assessment.
8. Some people are shy or afraid of losing face while being assessed by a human. Automation lifts this constraint.
9. They are good for self-study purposes, hence the proliferation of sites with quizzes, tests etc on the World Wide Web.

The biggest downside of tests is the fact that the widespread use of automatically checked tests has led to the appearance of databases with exam questions available from the Internet, and that cheating is easier when technology is used. On the other hand, there have appeared sophisticated technologies for the verification of test takers' identity. Moreover, an institution interested in security issues can always make sure that the tests are supervised by people - proctors, teachers or assistants - in order to minimize the risk of cheating.

Another critical opinion concerning testing is that it does not encourage high-level information processing, and promotes guesswork rather than learning. The quality of a test will always depend more on the content of test questions than the technology used. The technology itself is developing rapidly to allow for even more sophistication. It is now possible to automatically check quite complex answers (essays, texts). According to data from research, the agreement between scores awarded by one such a system for automated scoring of essays, called *e-rater* (the property of Education Testing

Services) and a human examiner is about 50+% (depending on the type of essay) and is comparable to the agreement between two human raters (Attali and Burstein, 2006). One can only hope that systems like *e-rater* become more refined and more easily available on a larger scale and at an affordable price soon.

Nevertheless, current practices show that even the simplest true/false tests play a valuable role in the process of education and, judging by the proliferation of self-study quizzes and tests on the World Wide Web, they are not likely to disappear.

3 DEVELOPMENT OF THE NEW TESTING SYSTEM

In the initial versions of the Edu system a simple testing module was created. It allowed generation of simple questionnaires and tests. All students got the same set of questions and answers. The module did not offer too many configuration options. Very soon it became not adequate to our needs. After many discussions with the lecturers it became obvious that the ideas about an ideal system of testing vary widely. According to many requirements stated by teachers, the module should be very flexible, should offer diversified scoring systems, ways of creating and selecting questions, with the possibility of manual, as well as automatic, evaluation. Before we started the implementation we tried to answer the question: What should the ideal system of testing for a technical university look like? We made efforts to gather as many ideas from our lecturers as possible, in order to create a system catering to diverse needs. After a few months an initial version of the system was developed. The system was created in ASP technology with Microsoft SQL Server database. Most of the data operations were moved to database layer. It gave us very good system speed and response time. Below we present the most important options and features of the current version.

4 THE TESTING SYSTEM FEATURES

4.1 Creating Questions

One of the basic requirements was that questions should be grouped into sets. When a lecturer creates a test, he or she may define how many questions from each set should be selected randomly. We added the possibility of easy set management, transferring and copying questions between sets etc.

A test author may define an unlimited number of answers to each question, state whether a number of answers will be defined on the level of the test or a question. Each question has a level of difficulty. Basing on that, the author of a test can decide how many questions of at a given difficulty level will be randomly selected to the test. Every answer can be evaluated using simple correct / incorrect information, as well as a number of points.

One of the key requirements was the option of different formatting of questions and answers, as well as supplementing them by graphics. We implemented a web editor with the basic possibilities of text formatting. The system of tags makes it possible to include graphics and mathematical expressions inside a text. We decide to use LaTeX language for mathematical expressions.

4.2 Creating Tests

After the creation of the questions sets, the author is able to define the test. The test's basic parameters are: title, dates of beginning and finishing, time limit (optionally). A lecturer can also protect the test using a password which will be known to some students. The system has 3 ways of test evaluation:

- "big points" – the student obtains 1 point for one question, when all answers are correctly selected or not selected.
- "small points" – the student obtains 1 point for every answer properly selected or not selected;
- "weights" – every answer has a weight between -100 and 100.

One of the most important parameters of the test is random questions and answers selection. Apart from the number of questions and answers for every question (it can be also defined on the level of a question – in this case different questions may have different numbers of answers), the author may decide that each question contains at least one correct answer. The lecturer may also define the number of questions for each level of difficulty. The author may include all parameters or only some of them, for example: the test contains 20 questions, 5 of them should be easy, 5 should be difficult, and the rest random. The author may also define that all or only some of the questions will be single choice questions. The system may show the test in two ways. First, in which the students see the entire test on the screen, and the second where only one question is visible at a given moment. The lecturer may decide whether the students can see the results

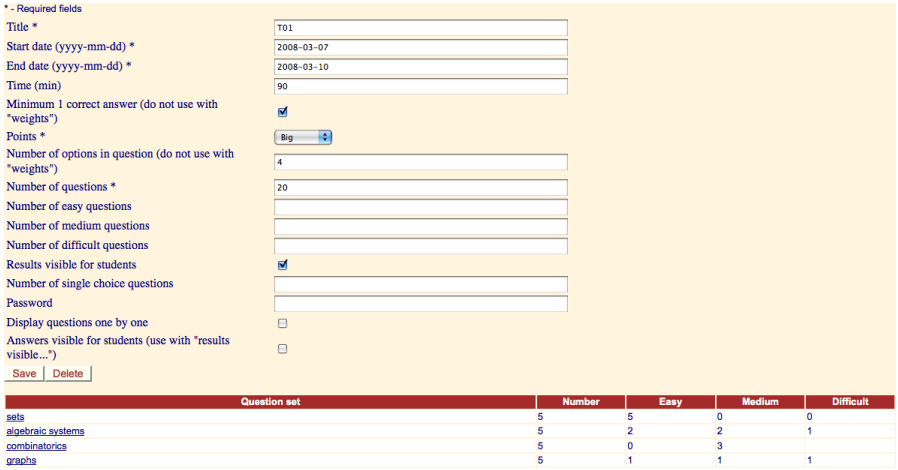

Figure 1: Parameters of a test.

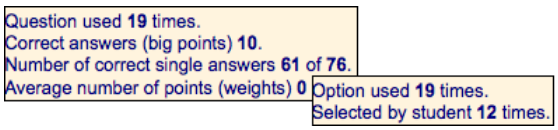

Figure 2: Questions' statistics.

Figure 3: Example of a test.

after test completion, and if they are able to see correct answers.

The last, optional stage is defining which sets of questions will be taken into account in the test generation. The author can define the sets and levels of difficulty for all questions, or leave some of the parameters empty. Then the system will select some of the questions randomly.

4.3 Statistics

From the point of view of the person using the same questions many times, it is very important to have the possibility of analyzing the statistics of questions usage. Therefore the author can observe each question and see how many times it has been used, as well as how many times each student has answered this question correctly or not.

4.4 Results

The lecturer has full access to the numbers of points obtained by students, as well as additional information, for example IP addresses, or how much time a given student spent on the test. The lecturer is able to see the test details, for example for manual evaluation or looking for errors in questions. From this level it becomes possible to allow the student to

repeat the test (with the same set of questions or the new, randomly generated test).

4.5 Algorithm of Random Test Generation

Due to the large number of options for the selection of questions and answers to the test, it was very important to create a good algorithm of random test generation, as well as validation of settings. The lecturer will be notified if there are not enough questions or answers for a given criterion. It was also very difficult to implement an algorithm in which some of the questions will not be preferred and distribution of usage of questions will be regular. The mentioned algorithm was implemented on the database server in stored procedure. So far it has allowed us to achieve acceptable performance, despite the rapidly growing numbers of students taking the e-tests, and the expanding base of questions.

5 CONCLUSIONS

The implementation of the new testing module has given our lecturers a very useful tool for evaluation of the students' progress. Thanks to active participation of our lecturers during the process of defining the requirements, the system became adequate to the technical university needs. The system is used very often, not only for Internet based studies but also for stationary studies. A lot of tests have been carried on using the system, including small exercises checking knowledge of single classes, as well as large tests of the exam rank. A very important feature is the possibility of analysis of questions usage statistics. Perhaps, the most important advantage is that a uniform, systematic, constantly developing database of test questions and answers of different topics has been created as a result of the development of the new testing tool. It is now an essential part of the knowledge repository for an on-line system supporting modern education.

REFERENCES

Attali Y. and Burstein J., 2006. "Automated essay scoring with e-rater v.2.0" (PDF). Journal of Technology, Learning, and Assessment, 4(3).

Bachmann L., 1990, *Fundamental Considerations in Language Testing*, Oxford: Oxford University Press

Banachowski L., Mrówka-Matejewska E., P. Lenkiewicz, 2004. "Teaching computer science on-line at the Polish-Japanese Institute of Information Technology", *Proceedings of IMTCI, Warsaw September 13-14 2004*, PJIIT Publishing Unit.

Banachowski L. and Nowacki J.P., 2009. "Application of e-learning methods in the curricula of the faculty of computer science, 2007 WSEAS International Conferences, Cairo, Egypt, December 29-31 2007", *Advances in Numerical Methods, Lecture Notes in Electrical Engineering, vol. 11*, ed. Nikos Mastorakis, John Sakellaris, Springer, 161-171.

Banachowski L. and Nowacki J.P., 2009. "How to organize continuing studies at an academic institution?", *Proceedings WSEAS International Conferences, Istanbul, Turkey, May 30 - June 1*.

Bednarek J., Lubina E., 2008. *Kształcenie na odległość. Podstawy dydaktyki*, Warszawa: Wydawnictwo Naukowe PWN SA.

Chapelle C.A., Douglas D., 2006. *Assessing Language through Computer Technology*, series ed. Alderson & Bachman, Cambridge: Cambridge University Press.

Paloff, R.M. & Pratt, K., 2003. *The Virtual Student*, San Francisco: Jossey-Bass.