DEVELOPMENT OF AN ON-LINE ASSESSMENT SYSTEM TO TRACK THE PERFORMANCE OF STUDENTS

Raquel Trillo, Sergio Ilarri
IIS Department, Univ. of Zaragoza, María de Luna 3, 50018, Zaragoza, Spain
raqueltl@unizar.es, silarri@unizar.es

Juan R. López, Nieves R. Brisaboa
Database Laboratory, Univ. of A Coruña, Facultade de Informática, A Coruña, Spain
mon@udc.es, brisaboa@udc.es

Keywords: On-line tests and assessment, e-learning.

Abstract: An important feature of an e-learning environment is the ability to continuously assess the progress of the students. With an adequate tracking of the students' knowledge and performance, it is possible to re-orient the teaching strategies in time to ensure the success of the learning process. Suitable computer tools are required to help the teachers to achieve such difficult goal.
In this paper, we present a tool for the definition, execution and evaluation of on-line tests, which can be easily integrated in an existing e-learning system. Prior to the development of the tool, we performed an extensive study of other existing alternatives, both commercial and free, that led our design. Thus, the developed tool presents the desirable features of those alternatives and others that we consider interesting.

1 INTRODUCTION

Web technologies have opened up many opportunities to develop new educational systems, since they allow an anywhere and anytime interaction between students and teachers. Thus, e-learning systems have emerged as a complement of traditional face-to-face classes and even in some cases as a replacement for them.

At the beginning, in this context, the web was used mainly as a medium to disseminate learning materials to students, and the only existing interaction was that of students downloading reading materials by following links in static HTML pages. These web pages were not part of an active learning system, and therefore lacked some functionalities that we demand today, such as the possibility to track and analyze the performance of the students. Due to this, different e-learning environments appeared, according to a certain teaching theory. In any of them, it is very important to be able to detect as soon as possible any deviation from the expected results (learning goals), both in individual students and in the whole group. One of the elements that can be used to check the progress of students is to test their performance in solving quizzes, exercises and problems. This is not only very useful for the teacher, who can adapt his/her methodology and react to the needs detected; it is also very valuable to the students, as it provides them with a reference on how they are improving their knowledge and skills and the topics/areas where they need to seek clarification or invest more effort. Therefore, the traditional value of the evaluation (*summative assessment*, used for grading the students) gives way to a formative value (*formative assessment*, which promotes learning), much more important from the point of view of the learning process, as it provides both students and teachers with orientation about areas that they need to improve.

The *Facultade Virtual* (`http://fv.udc.es`) is an e-learning system used at the University of *A Coruña* (Spain) which offers the classical functionalities for supporting, accessing and managing information for Higher Education, both face-to-face and online; thus, it offers general information about the University, an schedule of events, and course-specific information (syllabus, interesting links, bibliography, list of students, and electronic materials such as readings, slides, videos or audio). However, it did not provide any facility for the assessment of students: the only possibility was to offer links to documents containing the exercises and ask the student to send

the solutions by email to their teacher. Unfortunately, with this approach the possibilities that an online environment should offer would be wasted; for example, if a student could receive immediate feedback after submitting a test on-line, he/she could direct his/her own learning in a better and more efficient way. The motivation for this work was to develop an assessment system that would allow to define and correct tests via web. Although our work was conceived in the context of the *Facultade Virtual*, the developed assessment system can be easily integrated in any other e-learning system or even operate on its own.

The structure of the rest of the paper is as follows. In Section 2, we compare different systems available for evaluating students, assessing the convenience of developing a new one. In Section 3, and as a result of the previous comparison, we identify the functionalities that should be provided in the new assessment system. In Section 4, we describe the system from a functional and technological point of view. Finally, in Section 5, we draw some conclusions and set some lines for future work.

2 COMPARISON OF ASSESSMENT SYSTEMS

Nowadays, there are many assessment systems available. Therefore, first of all we needed to analyze the most relevant ones in order to decide whether one of them could be used in our context or, on the contrary, it was convenient to design and implement a new one. So, we analyzed 21 assessment systems that we consider significant. In this section, we briefly show the conclusions of our analysis. We evaluated to which degree they supported the following features:

1. *Functionalities offered to the different roles involved in the learning-teaching process*. We consider the existence of three types of users: administrators, teachers and students. Each of them require to access and use the system in a different way. Therefore, the assessment system should take into account the needs of all of them.
2. *Features of the graphical user interface*. We consider the usability of the interface (i.e., whether it can be easily used by people not familiarized with computers) and whether the interface is available in several languages. Moreover, we take into account if knowledge about some computer language (e.g., HTML) is needed in order to manage the system.
3. *Features of the tests that can be generated*. We evaluate the available test presentation formats (web pages, plain text, proprietary formats, etc.) and whether modifications can be easily performed with the goal of adapting them to other environments. We also check if it is possible to structure the tests in sections including different types of questions, if we can set a maximum number of attempts and the maximum amount of time allowed for the test, and whether the exercises and questions can be generated choosing randomly among several alternatives. Finally, we also consider if the system supports the inclusion of multimedia materials (images, video, audio, etc.).
4. *Features of the questions/exercises allowed*. We consider which types of questions are supported, specially if it is possible to include multiple/single choice and free-text questions (e.g., essays), as these are the most common types of questions in Higher Education e-learning environments. We also consider interesting to check whether it is possible to define clues that can help and guide the students when they find difficulties. Finally, it is also interesting to be able to classify the questions in different topics and according to their difficulty.
5. *Features of correction*. We consider if the system has the ability to automatically or semi-automatically correct some types of tests (e.g., tests not including free-text questions). We also analyze the quality of the information presented to the student when he/she submits his/her answers/exercises (e.g., grade obtained, advice on which topics should revise, correct choices, sample correct answers, etc.).
6. *Support to track the performance of students*. This is a key feature, as otherwise teachers would not be able to monitor the learning process and adapt themselves to the needs and the unexpected situations detected.
7. *Security*. We consider whether measures are taken to keep the privacy and integrity of the information stored, and whether there are mechanisms to try to prevent cheating when performing exams online.
8. *Features concerning the technologies used to implement the system*. We consider whether a proprietary or open technology has been used, its scalability, and whether it is easily extensible to include new modules/functionalities.

We present a summary of the comparison in Table 1, where we use the following symbols[1]:

[1]The complete survey is available (in Galician) at `http://webdiis.unizar.es/~raqueltl/Archivos/Ficheros/Memoria.doc.gz`.

Symbol	Meaning
χ	No
$\surd$	Yes
$\sim$	Partially
H	HTML
C	Configurable
T	Text
w	WebCT
X	XML
L	Latex
CP	CGI and Perl
D	Delphi
L	LAMP
A	ASP

Looking at Table 1, we can see how some features are supported by most of the systems (e.g., single-choice tests). However, others appear less frequently (e.g., role management, tracking of students, or providing clues to the students). The realization that any of the evaluated systems supported all the features that we required in our context, led us to develop a new assessment system.

3 REQUIREMENTS OF AN ASSESSMENT SYSTEM

As we have explained before, the development of the new system was motivated by: 1) the absence of an assessment system in the e-learning platform of the *Facultade Virtual* at the University of *A Coruña*, and 2) the fact that existing systems lack some interesting features.

The in-depth study of the existing (commercial and free) systems (that we summarized in Section 2) and several meetings and interviews with the teams involved in the development and maintenance of the *Facultade Virtual* (teachers, educators, and Computer Science engineers) led us to conclude the following requirements:

- *Web-enabled*. We consider it important that the system can be accessed using a standard web browser, avoiding the need to install any software on the user's computer (therefore, allowing seamlessly access from any computer), following the current trend *from desktop to webtop* (Shubin and Perkins, 1998). Moreover, if the system can be accessed through the Internet, the inconvenience of fixed schedules and meeting places can be avoided. In this way, the system can support both face-to-face and distance learning.
- *Intuitive graphical user interface*. The system must be easy to use by students, teachers and administrators, even if they are not used to computers. The interface must be also internationalizable, that is, it should be easy to show it in different languages and translate it to others.
- *Support of different roles*. The system should consider the different types of users involved in the process of learning-teaching in an e-learning environment: teachers, students, and administrators. Different functionalities should be offered to each of them.
- *Support of different types of questions*. Students must be allowed to perform tests, exercises and problems on-line. It is specially important to support single-choice and multiple-choice tests, along with free-text questions (e.g., essays). We consider that these are the types of questions present in most of the evaluation systems and the ones with a greater utility. The software architecture of the system should facilitate the integration of new types of questions with a minimum effort.
- *Support of different features for the tests*. For example, it must be possible to define timed tests and whether there is a maximum number of attempts allowed. It is also important that tests can be structured in sections covering different topics.
- *Support of repositories of questions*. Repositories facilitate reusing existing questions, and they should be managed by the teacher responsible for each course. It must be possible to classify questions according to the topics they cover (e.g., by associating keywords or selecting from a predefined set of topics) and level of difficulty. In this way, it is possible to adapt the tests generated to the student's knowledge level, making them suitable to his/her learning pace.
- *Support of random generation of tests and questions*. The system must support the automatic generation of tests by selecting questions randomly, according to some template with indications regarding the topics and levels of difficulty. For example, several sections could be defined for a test, each with a certain number of questions covering a different topic.
- *Support of automatic corrections*. The system must support automatic grading of the answers provided by the students whenever it is possible (e.g., for single and multiple-choice questions). As we are facing a learning-teaching context that increasingly demands more work and responsibility to the teachers, the system should release them from the burden of performing tasks that can be automated; in this way, they can invest their efforts in other activities that have a greater impact on the learning of students. Supporting automatic corrections is also very valuable to the student, as he/she can receive immediate feedback (see below).

Table 1: Comparison of assessment systems.

	MkLesson	Tutorial Gateway	OLAA	EasyQuiz	EasyTestCreator	WebQuiz	HotPotatoes	Quirex	QuizTest	SFESurvey	RandomQuiz	NueQuiz	MojoQuiz	QuickQuiz	QuizMaster	Extropia/WebExam	WebAssign	Quia	AulaEscolar	MicroCampus	PHPTest
Presentation format	H	H	H	H	H	H	T,w,X	H	H	H	H	H	H	H	H	H	H	H	H	H	H,X
Presentation changes	χ	χ	√	√	√	√	√	√	√	√	√	χ	√	√	√	√	√	√	χ	χ	√
Structure in tests	χ	χ	χ	χ	√	χ	χ	√	χ	√	χ	χ	χ	χ	χ	χ	χ	χ	χ	χ	χ
1 question type/exam	√	√	√	√	χ	χ	√	χ	χ	√	√	√	χ	√	√	√	√	χ	√	√	√
Several attempts	√	√	√	√	√	√	√	C	C	√	C	√	√	√	√	√	√	C	χ	C	√
Random tests	χ	χ	χ	χ	χ	∼	∼	√	χ	χ	√	χ	χ	χ	χ	χ	χ	∼	∼	χ	∼
Multimedia contents	χ	χ	χ	χ	χ	∼	∼H	H	H	H	H	H	H	H	H	H	H	H,L	H	H	H
Single-choice tests	√	√	√	χ	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Multiple-choice tests	χ	χ	χ	√	√	√	√	√	χ	χ	χ	χ	χ	χ	χ	√	√	√	√	χ	χ
Free-text questions	χ	χ	√	√	√	√	√	√	√	χ	χ	√	χ	χ	χ	χ	√	√	χ	χ	√
Clues to students	χ	√	χ	χ	χ	χ	√	χ	χ	χ	χ	χ	χ	χ	χ	χ	√	∼	∼	χ	χ
Several topics	χ	χ	√	χ	χ	χ	χ	√	χ	χ	χ	χ	√	χ	χ	√	√	χ	χ	χ	χ
Several difficulties	χ	χ	√	χ	χ	χ	χ	√	χ	χ	χ	∼	√	χ	χ	√	√	χ	χ	χ	χ
Automatic correction	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Feedback on correct.	√	√	χ	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Role management	χ	χ	χ	χ	χ	χ	∼	∼	√	χ	√	χ	√	χ	χ	χ	√	∼	√	√	√
Tracking students	χ	χ	χ	χ	χ	∼	χ	χ	χ	χ	χ	χ	√	χ	√	χ	χ	√	√	∼	√
Answers visible	χ	χ	χ	χ	√	∼	√	χ	χ	χ	χ	χ	√	χ	√	χ	χ	χ	χ	χ	χ
Access to private data	N/A	N/A	N/A	N/A	N/A	N/A	χ	∼	√	N/A	√	N/A	χ	N/A	√	χ	√	χ	χ	χ	χ
Internationalizable	χ	χ	χ	√	χ	χ	√	√	χ	χ	χ	χ	√	χ	χ	χ	χ	√	χ	√	χ
HTML required	√	√	√	χ	χ	χ	χ	χ	∼	∼	∼	∼	∼	√	χ	√	√	∼	∼	χ	χ
Technology	CP	CP	CP	CP	D	CP	CP	CG,H	CP	CP	CP	CP	CP	CP	CP	CP	CP	A	A	A	L

- *Feedback to the student*. Providing a valuable feedback to the students is a requirement of the learner-centered paradigm (Iahad et al., 2004). Thus, an appropriate feedback can have a very positive impact on the student's learning (Dalziel, 2001; Juwah et al., 2004). To this end, the assessment system should support:
  - The definition of *clues* that provide the students with orientations that may help them to solve questions that they may find difficult. This is particularly important in self-assessment tests: thanks to this extra help, the student may be encouraged to dig deeper or revise his/her knowledge on related helpful topics before giving up on a question.
  - When an automatic correction is not possible (such as in the case of free-text questions), feedback can also be provided through the inclusion of graded *sample answers*.
  - Along the same lines, the student can be presented with *additional information related to his/her performance in the test*, in order to clarify some doubts that can be inferred from the answers he/she provided and advice him/her what to do next.
  - Finally, the teacher can provide the students with comments about exercises he/she has corrected.
- *Support of multimedia*. Multimedia materials have been considered useful not only for the design of contents but also exercises (Fasli and Michalakopoulos, 2005). Teachers must be offered the opportunity to easily (e.g., without requiring knowledge on programming languages or HTML) include multimedia material that may facilitate the transmission of information to the students.
- *Facilities for exporting/importing questions and tests to/from different formats*. In this way, the interoperability of the system with other tools (e.g., *Microsoft Word* or structured text) would be supported. Similarly, it must be easy to add new modules that offer the capability of exporting/importing considering new formats that may be required in the future.
- *Safe*. It is necessary to consider some issues related to the safety of the system, taking into account the different types of information it manages (questions, answers, scores, feedback provided to the students, etc.) and the different roles of the users (e.g., only teachers must be allowed to define questions for their courses and for other courses for which they have been explicitly granted permission). On the other hand, it is also important to avoid cheating when the students perform exams: some safety measures that can be applied when an assessment system is used for grading are described in (Marais et al., 2006).
- *Efficient tracking of the learning process*. It must be possible to track the progress of the students and to obtain information about their performance on different types of questions. The teacher must be able to obtain graphical representations of this

information (e.g., by selecting statistics corresponding to specific time intervals).

- *Flexible and extensible*. The system must be easily maintained and escalable, that is, new types of questions, users and other functionalities must be added without much complication, and not having to modify the existing software structure.

These requirements have been taken into account in the development of our assessment system. In the following section, we indicate some technical details about the system developed.

4 DEVELOPMENT OF THE ASSESSMENT SYSTEM

The assessment system has been developed as a client-server three-tier architecture, considering the requirements outlined in the previous section, and using architectural and design patterns and the formal methodology *Unified Software Development Process* and the *Web Applications Extension* for *UML* - WAE- (Conallen, 2000). Traditional technologies such as *Java*, *JSP*, *XML*, and *XSLT* have also been used (Rockwell, 2001). The developed system is composed of three main parts, that correspond to the different roles of the users who take part in the learning-teaching process:

1. *For teachers*, three main functionalities are offered:
   - *Definition and evaluation of tests* (see Figure 1). These tests can be structured in different sections. Moreover, we consider the inclusion of multimedia material in multiple formats (text, video, audio, etc.), leveraging the multimedia and graphical capabilities of the existing web technologies to enhance user interaction.
   - *Management of information regarding students and courses*. Each teacher can manage information about the students that have signed up for his/her courses (see Figure 2), for example to add comments or grades, and the information on such courses.
   - *Obtention of statistical information*. The assessment system can present statistics of the performance of a student (see Figure 3) or the whole class. In this way, the teachers can analyze the progress of the students and detect needs as soon as they arise, which allows them to adapt their methodology accordingly. The statistics generated can also concern only particular topics or tests.

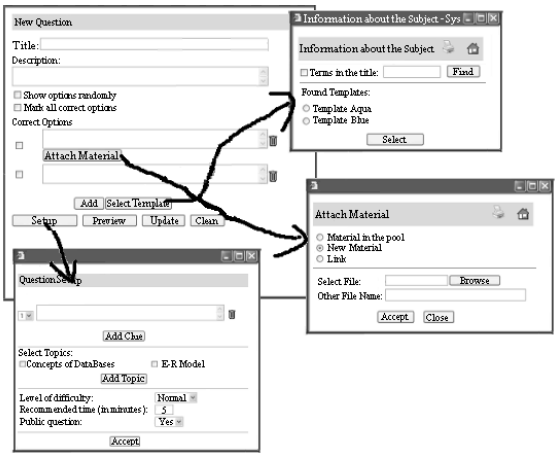

Figure 1: Definition of tests.

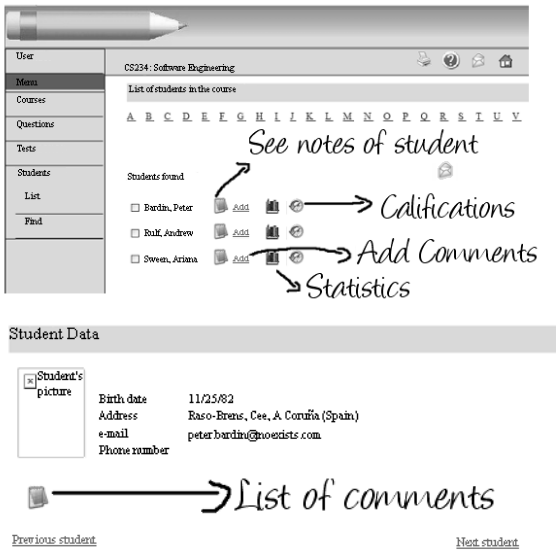

Figure 2: Access to information on students.

2. *For students*, it offers the possibility of performing exams and self-assessment tests (see Figure 4). Tests can be elaborated by teachers, but they can also be randomly generated from repositories of questions and depending on the performance of the student in previous tests: thus, the tests can be automatically adapted to the student's needs. Moreover, the students are provided with automatic test correction (when it is possible) and the correct/sample answers, giving him/her immediate feedback. Finally, we would like to indicate that the student can analyze his/her performance relative to the whole class, as peers can be a valuable reference for his/her own learning.

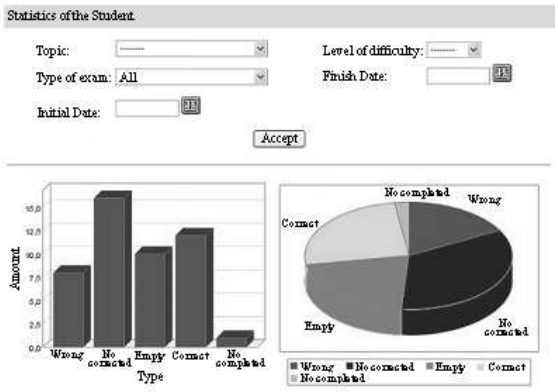

Figure 3: Statistics regarding the performance of students.

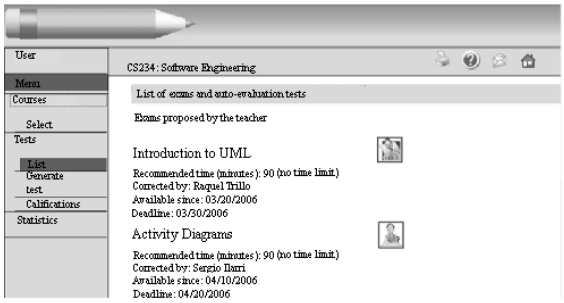

Figure 4: Access to tests by students.

3. *For administrators*, it provides them with the functionalities required to manage information about courses (e.g., allowing the definition of new courses), sign up new students, define the teachers assigned to the courses, and perform other tasks concerning the maintenance of the system.

5 CONCLUSIONS AND FUTURE WORK

In this paper, we have described an assessment system that we have developed for the definition, correction and tracking of students' performance. The system can be used in isolation or integrated in an e-learning platform. Previous to the development of the system, we have performed an extensive study of other existing alternatives and, taking into account their features, we have presented a set of requirements that we believe an assessment system should fulfill. This study motivated our work in this field, and we considered all those requirements in the development of the assessment system. The system presents many interesting features, it can be extended easily with other required functionalities (e.g., new types of questions), and provides immediate feedback to students and teachers through an easy-to-use graphical user interface.

A possible line of future work is to extend the system to support new types of questions (e.g., fill-in or ordering exercises) and new representation formats for the interchange of information among e-learning platforms, such as *QTI -Question and Test Interoperability-* (IMS Global Learning Consortium, Inc., 2006).

REFERENCES

Conallen, J. (2000). *Building Web applications with UML*. Addison-Wesley Longman.

Dalziel, J. (2001). Enhancing web-based learning with computer assisted assessment: Pedagogical and technical considerations. In *International Computer Assisted Assessment Conference, Loughborough (England)*.

Fasli, M. and Michalakopoulos, M. (2005). Supporting active learning through game-like exercises. In *Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05), Kaohsiung (Taiwan)*, pages 730–734. IEEE Computer Society.

Iahad, N., Dafoulas, G. A., Kalaitzakis, E., and Macaulay, L. A. (2004). Evaluation of online assessment: The role of feedback in learner-centered e-learning. In *37th Annual Hawaii International Conference on System Sciences (HICSS'04), Big Island (Hawaii)*. IEEE Computer Society.

IMS Global Learning Consortium, Inc. (2006). IMS Question & Test Interoperability Specification Version 2.1. `http://www.imsglobal.org/question/qtiv2p1pd2/imsqti_oviewv2p1pd2.html`, [Accessed October 19, 2006].

Juwah, C., Macfarlane-Dick, D., Matthew, B., Nicol, D., Ross, D., and Smith, B. (2004). *Enhancing Student Learning Through Effective Formative Feedback*. The Higher Education Academy Generic Centre.

Marais, E., Argles, D., and Solms, B. (2006). Security issues specific to e-assessments. *The International Journal for Infonomics Special issue: e-Learning Security*.

Rockwell, W. (2001). *XML, XSLT, Java, and JSP: A Case Study in Developing a Web Application*. New Riders Press.

Shubin, H. and Perkins, R. (1998). Web navigation: resolving conflicts between the desktop and the Web. In *Human factors in computing systems (CHI'98), Los Angeles, California (USA)*. ACM Press.