Advancing Electronic Assessment

Nikolaos Doukas, Antonios Andreatos

Abstract: A computer-aided assessment system is presented that has been designed to produce and deliver tests to the Hellenic Air Force Academy students and assess their performance. The system is called e-Xaminer and is intended for use in both undergraduate courses and distance learning post-graduate programs of the Academy. The e-Xaminer uses meta-language concepts to automatically generate tests, based on parametrically designed questions. Tests intended for different students may entail differences in the arithmetic parameters. Additionally, different tests may be composed from different but equivalent and randomly chosen sub-questions. The system may also present each student with a scrambled sequence of the same questions, as a counter-measure against cheating. Examinations are delivered via a web-based interface; an automatically generated program marks the answers submitted by each student. e-Xaminer allows the implementation of question parameterisation and counter cheating measures, so that electronic tests become significantly different and more powerful than traditional ones. Sample problems are presented which show the additional features of the e-Xaminer, intended to facilitate the work of the course organiser in issuing and marking the tests, as well as in combating cheating. This paper focuses on some new, advanced types of questions enabled by electronic assessment; it then compares paper-and-pencil exams to electronic exams; results from a small student poll on the electronic exams are also presented. Finally, the directions for planned future work are outlined.

Keywords: e-Xaminer, computer-assisted assessment, domain specific languages, HAFA.

1 Introduction

1.1 The revolution of e-learning

E-learning may be defined as the technology and the academic services related to the teaching and learning processes [1]. It is a wide term covering all the range of previous educational applications such as Computer Based Training (CBT) and Web Based Training (WBT), as well as more recent technologies such as LMS, virtual classrooms or labs and digital collaboration [2]. Historically, the rush of e-learning may be located in the decade 1990-2000, a period characterised by a boom in the Information & Communication Technologies (ICT) and the invent and evolution of the Web. For better exploitation of the above technologies, standards which will allow the interoperability of various platforms and the re-use of educational material are being developed. The first such standards appeared in 2001. The above evolution, along with other reasons, not only led to the foundation of many Open Universities around the world [3], but also pushed many traditional universities to offer distance learning courses [1]. It is estimated that, as far as continuing education in higher education institutions is concerned, distance learning will grow ten times faster than on-campus learning over the next ten years [4]. According to Burns, "Up to 45 percent of colleges and university enrolment is from adult learners, many of which sign-up for distance learning classes rather than on-campus classes. Revenues for continuing education rose 67% at responding institutions since the previous survey in 2004. The trend is expected to grow distance learning 10 times faster than campus classes over the next decade. The growth in distance learning is driven by the growth of interactive marketing" [4]. The Web has been considered a means of education and knowledge since its early days. Today it is widely used for educational purposes due to its worldwide spread and penetration and because it supports a lot ways for the representation of information [5].

The recent proliferation of distance learning (delivered by Open Universities worldwide, as well as traditional institutions) has encouraged many higher education organisations to develop software for supporting this type of learning. A well known such package is Claroline [6], which is freeware and supports 32 languages. With the expansion of e-learning, the need for electronic examinations has become more imperative. A significant number of Computer Aided Assessment (CAA) systems has appeared as a result of the work of both academic institutions and commercial companies. Examples of such systems are given in [6], [7], [8], [9], [10], [11] and [12]. These systems have already been extensively tested and are being widely used.

1.2 The merits of electronic testing

In ref. [13], the public-private coalition known as the *'Partnership for 21st Century Skills'* gives a vision of how students should be prepared to face the challenges of the 21st century. Within this report, the benefits of using technology in order to give immediate feedback on student assessments is underlined. Electronic testing has been accused of bringing non-technology students to a disadvantage when they are forced to use a keyboard to type their answers, rather than writing them on paper [14]. In response, public and private sector experts state in [13] that 21st century literacy is much more than the basic computer skills required for typing an answer. It is pointed out that the new tools for learning including computers, telecommunications and audio or video based media are critical enablers of learning in a number of realms, even for subjects that have nothing to do with technoology. They hence conclude that there is a need for assessment tools that measure those essential skills that will never be captured by traditional tests. While the value of traditional testing (like portfolio assessment) as a means for classroom level assessment of individual students is not questioned, electronic tests provide rapid feedback on students performance that can be compared across classrooms or schools [13].

Furthermore, computer based training can enable the evolvment of novel concepts that instructors would have never contemplated of delivering a few years ago [15]. The authors present a training platform which automatically designs courses that not only adapt to the students capabilities and previous knowledge, but also dynamically adjusts the contents, according to the students performance during the progress of the course. The statistical results presented show that computer delivered courses are extremely efficient in promoting learning.

An automated computer-assisted assessment (CAA) system has already been presented [12], [16], which has been designed for the needs of the Hellenic Air Force Academy (HAFA). This system is called **e-Xaminer**, a title which emphasises the fact that this system extends e-Learning in the field of examinations. The principle reasons which dictated the development of a new system from scratch, in order to cover the needs of the HAFA, where the following: (1) The limited number of types of questions supported by the existing systems. (2) The intention of HAFA to experiment with cheating countermeasures so that the final system could eventually be used in its distance learning program [17]. (3) Most existing systems do not support the Greek language, while teaching in HAFA is done exclusively in Greek. This fact creates some special requirements that needed to be met.

In ref. [16] we have presented the nine different problem types and partial credit possibilities supported by the e-Xaminer so far and we have briefly discussed the main advantages of CAA. In ref. [12] we have presented the DSL approach adopted by the e-Xaminer, cheating countermeasures and some experimental results obtained by the pilot application of the e-Xaminer.

1.3 Making the difference

Multiple choice tests suffer from the drawback that examinees may aquire free-marks by taking lucky guesses [18]. Furthermore, a way to exploit electronic exams in order to overcome this problem is proposed. Given the evolution of web technologies since then, the need to make electronic exams

significantly different is currently more pronounced.

Another challenge for the design of electronic examination systems is to make exams more pleasant for students. Students are naturally reluctant to sit exams, since exams are a source of stess for them. On the contrary, they are keen to spend large periods of time in front of the computer screen, surfing the Internet. The question is hence posed: *If exams are delivered using a friendly web interface, can they become a less stressful experience?*

This paper focuses on the implementation of test parameterisation and counter-cheating measures on the e-Xaminer, that distinguish e-Xams from paper and pencil exams and make them a better tool from promoting learning; it also lists some empirical results and presents a comparison between conventional paper-and-pencil exams and electronic exams given by means of the e- Xaminer (and therefore, referred to as 'e-Xams'). The way in which questions are parameterised is firstly presented. Examples of parametric questions are shown. Counter-measures against exam-time cheating are listed and the way in which these have been implemented in the new system is described. The new system has been used for mid-term tests of students of all four years of the Academy, for all specialisations. Statistical results from these tests are given. Finally conclusions are drawn and the directions for the planned future work are presented.

2   Design of the system

2.1   System Architecture

The architecture of the e-Xaminer is depicted in Figure 1. The system operates as follows. The instructor sets examination questions and model answers using a Domain Specific Language (DSL) [19]. These files are given as input to the core of the e-Xaminer which produces two programs, the **examination agent** and the **marking agent**. The examination agent is installed in an appropriate HTTP server and is the only part of the system that is ever exposed to a public terminal. The examinees open the 'e-Xam' pages and answer the test questions by filling in forms, using their favorite web browser. Student answers are stored and finally passed as input to the examination agent that runs on the instructors terminal (e.g. a laptop computer only temporarily connected to the student network). A marking report and statistics are produced.
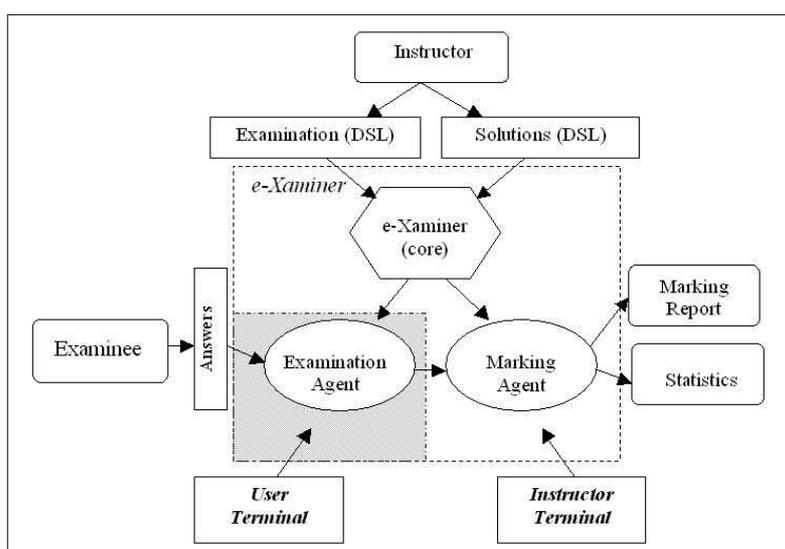
Figure 1: Architecture of the e-Xaminer

2.2   The examination procedure

Figure 2 depicts the examination procedure. The correspondance to the system architecture described earlier on is straightforward. It is interesting to note however that it is possible for the instructor, to re-iterate the process with a revised marking scheme. A marking (grading) policy is implicitely determined by the instructor and is included in his/her marking scheme. The marking scheme may include alternative answers that carry different marks. For example, the correct answer to the question "What is the maximum value that can be represented in an N-bit unsigned binary number", is $2^N - 1$. The instructor may however decide to grant students a partial credit for the (wrong) answer 2^N. This iterative procedure, which would have been more tedious for a paper-and-pencil test, represents a significant advantage of e-assessment.

As already pointed out, the main drawback of existing systems, as far as their use by the HAFA is concerned, was the fact that they supported only a limited number of question types; furthermore, these questions were often fixed (same for all students). It was required that the new system supported parameterisation. Additionally, the whole process needed to be automated. By using the e-Xaminer, the entire class is examined simultaneously, but each student sees a different worksheet [12].
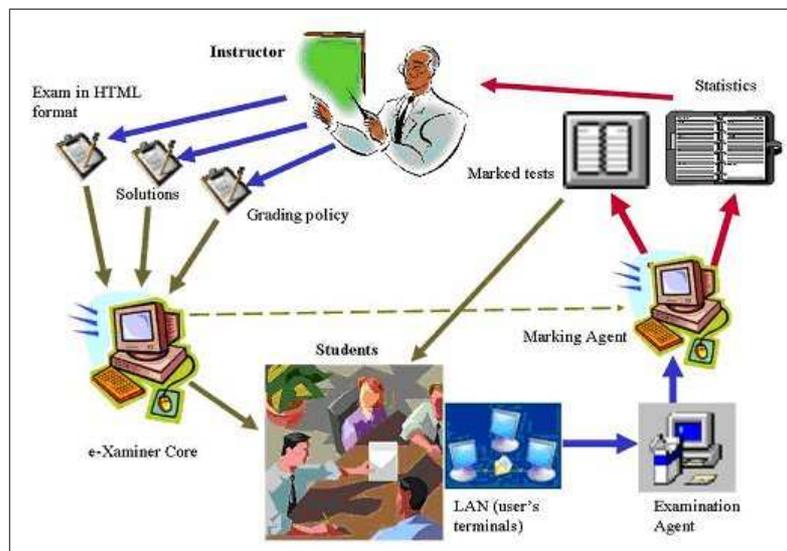
Figure 2: Flowchart of the examination procedure

2.3   Question parameterisation

Apart from the innovative types of questions, a key feature of the e-Xaminer is question parameterisation. It was required that the system be able to receive as input from the course instructor a skeleton question and produce a series of different questions from it. The process of defining and designing such skeleton questions is the question parameterisation incorporated in the e-Xaminer [12]. In order to implement this parameterisation, our system employs the concept of Domain Specific Languages or DSL [19]. A skeleton question can hence be defined by its textual parts, a series of parameters that it requires and a rule which is used for assembling the final question. An additional parameterisation for the entire test is question order scrambling, i.e., the randomisation of the order in which the questions in the test of a particular student is arranged. This randomisation may be expanded to the arrangement of sub-questions of a particular question. Two examples of parameterised questions which were used during the past academic year (2005-6) are given below.

The first question was designed to examine a chapter on binary arithmetic. Each student was given

six different random numbers. Each student was hence required to perform a number of operations on the numbers assigned to them. These operations were selected at random for each student from a set of predefined ones (binary to decimal conversion, decimal to binary conversion, addition, subtraction and negation). The setup of this question eliminated any chance that a student might have copied from their collegue sitting next to them.

The second question was designed as part of a course on Computer Networks [20]. The skeleton of question was the following: "A company owns a set of class B IP addresses. The company has X LANs with an increase rate of Y% during the next 10 years. Each LAN connects W computers, with an increase rate of Z% during the next 10 years. Which is the best way of sharing the class B host bits between subnets and computers?". All students sitting this examination had to answer this question, however each student had a different set of values for the X, Y, Z and W parameters. These values where automatically assigned by the e-Xaminer.

Tests delivered during the pilot application phase, were just automated versions of what a paper and pencil test would have looked like. In order to better exploit the potentials of the CAA web interface, the decision was made to include various innovative features in the tests. The first such feature was the electronic processing of an (electronic) document, linked to the e-Xam by a simple hyperlink. The question had as follows [20]:

"Given the specification of HTTP 1.1 [RFC 2616] in electronic format, answer the following questions:

1. What is the signalling mechanism used to terminate a persistent connection?

2. Who may terminate such a connection? (a) the client, (b) the server, (c) both of them

3. What does an error code 415 signify?

4. What encryption services are provided by HTTP?"

RFC 2616 is a 176-page document describing the HTTP standard. Students are obviously not asked to memorise this text and it is impractical to copy the entire document for each student for the purpose of the exam. They are however expected to be familiar with such documents and able to extract information from them using various text/word processing tools (such as 'Search' or 'Find'). Testing of this skill is only possible via electronic exams. These questions may further combined with parameterisation, in order to produce a different error code for each student. The next planned test requires the students to use some special services available on the Internet, get the answer on their browser and copy it on their 'e-Xam'.

Telecommunication engineering students in HAFA are required to be able to write, assemble and run machine language programs. In another type of question therefore, examinees were required to assemble their programme and submit the assembler listing file. This is another example showing how e-assessment can test skills that would have been impossible to examine with paper-and-pencil exams.

2.4   Facilitating the process of in-classroom examinations

An additional requirement for the system was that it should shield both instructors and students from having to comprehend any complicated software engineering concepts. As far as students are concerned, the target was met by virtue of the fact that a plain web interface was used for delivering the tests. Students were pleasantly surprised to notice that each one of them had a different set of numbers. This made them adopt an attitude such as "I will mind my own business and finish my questions in time", rather than the more traditional "I will try to see if one of my colleagues sitting next to me can help me with the exam questions". Hence learning was promoted and student and instructor psychology was boosted. In order to maximise the benefits for course instructors as well, the concept of Domain

Specific Languages was employed. Instructors compose the skeleton examination questions and their model (correct) solutions in HTML format (Figures 1, 2). They are not required to be familiar with HTML (low-level) coding [12]. They can use their favourite HTML editor to complete the task, as current editors have a simple user interface (similar to a word processor such as Microsoft Word).

The HTML files are processed by the e-Xaminer core and two sets of programs are produced: the examination agent and the marking agent (Figure 1). The examination agent is an active page combining HTML and Perlscript. The testing agent is a Perl script which processes the answers submitted by students. The outputs are the marked tests and statistics of the class performance.

3   Contribution to the virtual classroom

The fact that questions are parameterised has several positive implications for the virtual classroom. Firstly, tests can become personalised (a different set of questions for each student) without compromising test fairness. Additionally, tests can be given more often than usual (the customary midterm exam may be replaced by weekly exams), since the effort required by the instructor is minimal. Both the above achievements are recommended practices against examination-time cheating and in favour of promoting learning [21], [22]. It has additionally been observed that the web interface makes the idea of frequent exams, more agreeable than usual for students. Marking papers is a fairer process, since, even at early stages, the system has caught grading errors made by humans.

As it was mentioned earlier, the electronic tests in particular have been accused of bringing some students to a disadvantage [14]. The e-Xaminer experience has shown that the e-test marks have large correlation to both the oral mark awarded by the instructor, as well as the mark attained during the paper-and-pencil semester final exam. Furthermore the HAFA, being a technically oriented institution, demands from its students (which, as Officers of the Hellenic Air Force, are going to handle electronic equipment worth millions of euros) to be familiar and at ease with technology.

Accusations have been extended to state that "test-takers are unable to underline text, scratch out eliminated choices and work out math problems - all commonly-used strategies" [14]. Current web technology allows not only underlining on-screen texts, but also much more elaborate formatting to take place by incorporating a simple HTML editor (a practice used by many web-based email services).

As far as scratching is concerned, we have provided special scratch-pad areas next to each question (implemented as HTML textareas). We have also encouraged our students to get better help by using electronic dictionaries and calculators, as well as their old paper notebook which is sitting next to their keyboard.

4   Pilot application and statistics

The e-Xaminer is being used in testing digital electronics, computer science, microprocessors and computer network courses taught in HAFA since the beginning of the past academic year (2005-6). The CAA tests were addressed to students of all disciplines taught in HAFA (pilots, aircraft, telecommunications and civil engineers, as well as air-traffic controllers). During the first six months all student answers were being marked by both the e-Xaminer and the instructor. During this first period, automatically assigned marks deviated by up to 10% from the instructor marks. This deviation declined steadily as the system evolved and staff acquired experience in assigning better model answers. Currently, only a limited number of tests, selected at random, are being manually graded and the deviation between the two sets of marks is is less than 1%. The grading resolution (that is, the minimum grade associated to questions), was set to 0.5% during recent e-Xams. After the initial 6-month deployment period, the e-Xaminer has been systematically catching human marking errors in up to 10% of student papers. Instructors have very often exploited the fact that the system is automatic and have revised their model

answers (along with their marking scheme) so as to better assess class performance.

4.1 Comparing paper exams to electronic exams

In this subsection electronic exams delivered by the e-Xaminer will be compared to traditional paper-and-pencil exams. Rather than answering a question of the form 'which one is better', our experience from the use of the e-Xaminer during the current (2006-7) and the past academic year (2005-6) will be presented. Each approach has its own merits that need to be considered by course instructors before making a choice. The aim of this section is to support such a choice.

The **advantages** of paper-and-pencil exams include:

- They can be easily adopted to any subject.

- They may be used by digitally illiterate students.

- No technological infrastructure is necessary.

- No computer-related background skills are necessary.

Their major **disadvantages** are:

- The grading process is usually tedious and time-consuming.

- It is difficult to embed in them multimedia (other than figures).

On the other hand, the **advantages** of electronic assessment are:

- Tests may be generated automatically from a pool of questions and problems.

- Grading is easy, automatic and hence tests may be delivered more frequently.

- It is easy to embed multimedia (such as designs, circuits, video and sound clips for courses such as signal processing, electronics and telecoms).

- Students with bad, illedgible handwriting can produce a well formed document.

- Results and statistics are immediately generated automatically and students obtain rapid feedback on their performance.

- Counter-cheating methods may be employed.

- Supplementary programs may be used (such as dictionaries, assemblers, compilers, MATLAB etc.).

- Exams may be easily stored and retrieved, results may be further processed with other computer programs such as Excel and SPSS.

- Finally it should be noted that e-Xams are environment-friendly in that little paper need ever be used.

The **disadvantages** of electronic assessment are [16]:

- It might be difficult to adopt it to some types of questions (such as math formulae, where special equation editors are needed for typing the formulae and complex software for interpreting them).

- It might be difficult to adopt it to some subjects.

- It requires additional equipment such as PCs, LAN, software etc. and is hence more complicated to give, as well as, vulnerable to power or system failures.

An instructor must take all the above factors into account before making a decision.

4.2 Student opinions

It is important to consider the opinions of users in evaluating any e-Learning system [22], [23]. The authors of ref. [22] define such an evaluation procedure based on hypothesis testing. The hypotheses that needed to be tested in the case of the e-Xaminer where (1) "e-Xams are perceived as equally difficult to the corresponding paper ones", (2) "e-Xams promote learning by making tests more agreable experience for students" and (3) "e-Xams promote learning by helping students accept their test marks as a fair assessment of their performance". In order for us to investigate these hypotheses, three extra questions were added at the end of the most recent tests. The questions where:

1. Do you consider the electronic test to be more difficult than what you would expect from a paper test?

2. Do you believe that automatic assessment you will get for this test will be fairer than the one you would have got if the test was marked by your instructor?

3. Do you prefer this test to a traditional paper one?

HAFA students answered by an overwhelming majority (over 90%) that the electronic test was equally difficult and more preferable to the traditional test, while they expected their automatically assigned marks to better reflect their performance. Answers to these questions are still being collected, as more classes sit electronic exams and will be presented in future publications once the required statistical significance is attained.

5   Summary and conclusions

The implementation of parametric questions and counter cheating measures of the 'e-Xaminer', our CAA system, was outlined. This implementation was shown to offer advantages to university course organisers, in that it facilitated their work and eliminated much of the tediousness involved in the grading process. The new system was also shown to promote learning by the students, by making exam taking a frequent, fair and agreeable procedure. Some contributions were also made to theoretical issues concerning the electronic classroom and its acceptance in general. Additionally, e-Xaminer was shown to offer a number of advantages in the fight against examination-time cheating. Statistics from the pilot application were given, were it was shown that e-Xaminer was capable of effectively marking student answers. Innovative types of exam questions were used, that could not have been set without the existence of an e-assessment platform. Initial statistical results were presented on the acceptance of the system by students. These results are part of a study on the usefulness of this system in promoting learning. During the next academic year we plan to further develop the system in order to:

- Include some additional types of questions/problems.

- Support multimedia in both questions and answers. This is a clear advantage of CAA over paper-and-pencil exams.

- Include a timer and enforce automatic submission upon time-out.

- Strengthen security and fight cheating attempts by adding an authentication module that will monitor users login and logout [12].

References

[1] The Univ. of Iowa in cooperation with HBO Systems Inc., *E-Learning Assessment Summary Report*, Available online at: www.hbosystems.com, 2004, retrieved on March 31, 2006.

[2] F. Pantano-Rokou, Educational design for e-learning: models, meaning and impact on learning (in Greek), *Open Education*, vol. 1, pp. 45-68, 2005.

[3] D. Vergidis, A. Lionarakis, A. Lykourgiotis, B. Makrakis and Ch. Matralis, *Open and distance learning, vol. 1, Institution and Operation*, Hellenic Open University, Patra (in Greek; title of book translated by authors), 1998.

[4] E. Burns, Continuing Education drives Distance-Learning enrollment, *Available online at: www.clickz.com/stats/sectors/education/article.php/3605321*, retrieved on May 25, 2006.

[5] Ch. Fidas, Ch. Tranoris, V. Kapsalis and N. Avouris, System design for synchronous support and monitoring in web-based educational systems (in Greek), *Proceedings of the 3rd International Conference on Open and Distance Learning*, Propombos, Athens, vol. A, pp. 577-585, 2005.

[6] Claroline documentation, *http://www.claroline.net/documentation.htm*, retrieved on June 21, 2006.

[7] Blackboard documentation, *http://library.blackboard.com/docs/as/bb_academic_suite_brochure_single.pdf*, retrieved on June 21, 2006.

[8] Univ. of Loughborough CAA site, *http://www.lboro.ac.uk/service/pd/caa/index.htm*, retrieved on June 21, 2006.

[9] Quia site, *http://www.quia.com/company/quia-presentation.pdf*, retrieved on June 21, 2006.

[10] Test Assessments project, *http://www.scribestudio.com/home/ inAction/Flash/ss_in_action.jsp?cm=TestsAssessments_Project*, retrieved on June 21, 2006.

[11] Web CT site, *http://www.webct.com*, retrieved on July 10, 2006.

[12] N. T. Doukas and A. S. Andreatos, Implementation of a Computer Aided Assessment System Based on the Domain Specific Language Approach., *WSEAS Transactions on Advances in Engineering Education*, vol. 3(5), pp. 382-388, 2006.

[13] J. Salpeter, 21st Century Skills: Will Our Students Be Prepared?, *http://www.techlearning.com*, Oct. 2003, retrieved on March 22, 2006.

[14] Fairtest site, *http://www.fairtest.org/facts/computer.htm*, retrieved on June 23, 2006.

[15] A. D. Styliadis, I. D. Karamitsos and D. I. Zachariou, Personalized e-Learning Implementation - The GIS Case, *International Journal of Computers, Communications and Control*, vol. I, no. 1, pp. 59-67 (2006).

[16] A. Andreatos and N. Doukas, e-Xaminer: Electronic Examination System, *Proceedings of the 3rd WSEAS / IASME International Conference on Engineering Education, Vouliagmeni, Athens, Greece*, July 2006.

[17] A. Andreatos, Distance e-learning for the Hellenic Air Force, *Proceedings of EDEN'03*, Rhodes, Greece, pp. 428-433, 2003.

[18] M. Bush, Alternative marking schemes for on-line multiple choice tests, *7th Annual Conference on the Teaching of Computing, Belfast,* CTI Computing 1999.

[19] D. Spinellis, Notable design patterns for domain specific languages, *Journal of Systems and Software*, 56 (1), pp. 91-99, Feb. 2001.

[20] J. F. Kurose and K. W. Ross, *Computer Networking - a top-down approach featuring the Internet*, 3rd ed., Addison-Wesley, 2005.

[21] A. Angeletou et al., Assessment techniques for e-learning process, *Proceedings of the 3rd International Conference on Open & Distance learning*, vol. B, pp. 47-54, Patra, Greece, 2005.

[22] D. Spinellis, P. Zaharias and A. Vrechopoulos, Coping with plagiarism and grading load: Randomized programming assignments and reflective grading, *Computer Applications in Engineering Education*, to appear in 2007.

[23] R. Guidorzi and M. L. Giovannini, E-learning tools in higher education: users' opinions, *Proceedings of EDEN'03*, Rhodes, Greece, pp. 201-206, 2003.

Nikolaos Doukas, Antonios Andreatos
Dept. of Aeronautical Sciences
Div. of Computer Engineering and Informatics
Hellenic Air Force Academy
Dekeleia, Attica, TGA-1010 GREECE
E-mail: nikos@doukas.net.gr, andreatos@hafa.gr