Work in Progress – Development and Use of a Software Tool for Improving the Average Student Capacity in the Greek Higher Education System

Catherine C. Marinagi, Vassilis Th. Tsoukalas, and Vassilis G. Kaburlasos
Department of Industrial Informatics, Technological Educational Institution of Kavala,
GR-65404 Kavala, Greece {kmarin, vtsouk, vgkabs}@teikav.edu.gr

Abstract **- An inherent problem of higher education in Greece regards the absence of regular homework assignments due, at large, to limited teaching resources. Work is currently in progress for overcoming the aforementioned problem by the development, followed by implementation, of a customized software tool namely *Platform for Adaptive and Reliable Evaluation of Students* or PARES for short. The basic idea behind PARES is to regularly give students short exams, including multiple-choice questions, in response to homework assignments. This work presents technical features as well as our first experience of using PARES at a large-scale in practice. A statistical evaluation of PARES is also described.**

Index Terms – Greek higher education, student evaluation software, multi-user application, statistical hypothesis testing.

INTRODUCTION

An inherent problem of the state-funded, undergraduate higher education in Greece regards the absence of regular homework assignments due, at large, to limited teaching resources. The absence of homework assignments on a regular basis during the semester reflects adversely on the average student capacity especially in rapidly evolving fields of technology. Note that the aforementioned problem has been further exacerbated, lately, by the on-going expansion of higher education in Greece as the number of registered students increases rapidly every semester while new academic departments proliferate throughout the country. Under the circumstances there is an urgent need to keep providing high quality technological education to undergraduate students.

The basic *work hypothesis* (*WP*) here is that "keeping the body of students actively involved during a semester will improve the average student capacity". Nevertheless, it is hard to test the "basic WP" using traditional ways of delivering education due to the limited teaching resources.

Work is currently in progress for testing the aforementioned basic WH by the development and use of a customized software tool, namely *Platform for Adaptive and Reliable Evaluation of Students* or PARES for short. A version of PARES has been completed and it is already in use as described below.

TECHNICAL FEATURES

PARES is a client/server, multi-user software application which includes three modules, namely the Secretary module, the Exam Composer module, and the Exam Conduction module. Hypertext markup language (HTML) pages are included containing text as well as figures.

For a course 'C' an instructor is required to supply a *question Bank* (qB_C) including N_C multiple-choice questions. The questions are organized hierarchically in a tree structure of sections/subsections, moreover a vector of attributes is attached to each question. The latter attributes currently include the question type, e.g. design /mathematical /comprehension, as well as the level of question difficulty.

A test, including n_C questions, may be generated randomly in response to a homework assignment. Note that an instructor has the liberty to define the domain of an exam by marking "on" or "off" a branch in a qB_C's tree structure. It turns out that a maximum number of $\binom{N_C}{n_C} = \frac{N_C!}{n_C!(N_C - n_C)!}$ different tests with n_C questions can be generated, furthermore it is up to the instructor to define both the number n_C of questions in a test as well as the estimated average time t_Q a student requires to answer a question.

As soon as a student finishes an exam he/she is expected to submit his/her answers by activating a button. Nevertheless, if the exam time elapses then the student answers are submitted automatically. In either case a local printer produces a hard copy of the student answers.

The grading is carried out automatically; furthermore, a record of the student performance is stored in the server. After submitting his/her answers a student can only view the answers, i.e. the student is not allowed to change his/her answers. Note that PARES can support simultaneously any number of exams in the department.

A first implementation of PARES has been completed in the Spring semester of 2003 in the Delphi programming language and used in a pilot study [1]. After an encouraging response from the students, in the Fall semester of 2003 we have switched successfully to the Java programming language to attain a more convenient, Internet implementation.

USING PARES IN PRACTICE

PARES has been installed in two laboratories each one of which contains around 25 personal computers (PCs). The same server is used by the PCs in both labs. The students enter a supervised lab "asynchronously", at their convenience, within a predefined time span to take an exam. A student uses a password to activate a specific exam from a list of exams available. Only registered students are eligible for taking an exam only once.

As soon as the exam commences, a list of n_C questions is displayed on the computer monitor. The display with the questions can be scrolled up or down. After submitting his/her answers, a student is required as well to turn in a signed hard copy of his/her answers.

PARES discourages plagiarism since each student receives, in essence, a different set of questions. Two students, sitting next to each other, may occasionally receive a few identical questions; nevertheless the latter questions will, almost certainly, appear in a different location in their corresponding exams. In addition, plagiarism is discouraged since the time to answer a question has been kept short.

In the Spring semester of 2004, PARES is already in use in each one of the junior/senior level courses *Intelligent Systems*, *Artificial Intelligence*, and *Robotics* both in the theoretical- and in the laboratory- part of the aforementioned courses. Since the Spring semester in Greece ends at the end of June we will keep using PARES until mid-June. The effectiveness of PARES will be assessed statistically based on the final exam grades as explained in the following.

STATISTICAL EVALUATION

In a specific course 'C', in a specific semester 'S', the final exam for a student will be treated as an independent Bernoulli (pass/fail) trial with probability of success (pass) p. It follows that the probability $P_n(k)$ of 'k' successes in 'n' trials is given by the *binomial distribution* $P_n(k)=\binom{n}{k}p^k(1-p)^{n-k}$. Let n_p students pass the final exam of a course C in the S semester. It is clear that the probability p of success could be any number in the interval [0,1]. An optimal estimate $\hat{p}$ of p can be computed from the unimodal *likelihood function* $L(p;n,n_p)=\binom{n}{n_p}p^k(1-p)^{n-n_p}$ with parameters n and n_p. The aforementioned function attains its maximum value at the *maximum likelihood estimate* $\hat{p}$ of p, given by $\hat{p}=n_p/n$.

A statistical test will be considered involving two different groups of students, namely group-1 and group-2 with population sizes n_1 and n_2, respectively. Both groups of students have been taught a course C the same way – the only difference is that the n_1 students of group-1 have been taught course C without using PARES, whereas the n_2 students of group-2 are currently taught course C using PARES. Let n_{p1} and n_{p2} denote, respectively, the numbers of students who pass course C; furthermore let $\hat{p}_1=n_{p1}/n_1$ and $\hat{p}_2=n_{p2}/n_2$ be the corresponding maximum likelihood estimates of the Bernoulli probabilities of success p_1 and p_2, respectively, for group-1 and group-2. Consider the following *null hypothesis* H0.

H0: The (Bernoulli) probability p_1 of success for group-1 of students equals the (Bernoulli) probability p_2 of success for group-2 of students; i.e. the null hypothesis is $p_1=p_2$.

Using standard *hypothesis testing* it is expected to conclude that we cannot either reject or accept the null hypothesis at a certain confidence level. In other words, based on statistical evidence it is expected to conclude, at a certain confidence level, whether the use of PARES has improved the chances of a student to pass a course or not.

CONCLUSIONS AND FUTURE WORK

At this stage of its development PARES can be used as a tool for monitoring reliably (and "inexpensively") student progress during a semester. An improvement in the average student performance is reasonably expected by motivating the average student to study "during" the semester instead of studying only at the final exams as it currently occurs, at large.

The student response to PARES has been encouraging since 80% of the registered students in a class, or more, have participated in an exam with PARES. The only recorded student complaint has been their request for more time during an exam. We are planning on increasing the time (t_Q) allocated per question by as much as 50% to 1.5 minutes. Furthermore, in the long run, we plan on considerably increasing the number N_C of the questions in the question Base qB_C.

It is of practical interest an increase of the Greek higher education system's *throughput*, we define the latter as the number of students passing a course, while further improving the quality of provided education. On the one hand, we believe that retaining a question Base qB_C with high quality questions can contribute to maintaining a high quality education. On the other hand, a further improvement in the quality of provided education may be feasible using a "distant-learning" software module in PARES driven by a user model; the latter can be developed based on machine-learning techniques as it will be detailed elsewhere.

ACKNOWLEDGMENT

This work has been supported in part by the third European framework programme.

REFERENCES

[1] Kaburlasos, V.G., Moussiadis, L., Tsoukalas, V., Iliopoulou, A., and Alevizos, T., "Adaptive technological education delivery and student examination based on machine-learning tools", *Suppl. Proc. Intl. Conf. on Artificial Neural Networks & Intl. Conf. on Neural Information Processing (ICANN/ICONIP 2003) - Special Session on Machine Learning Advances for Engineering Education*, Istanbul, Turkey, 26–29 June 2003, pp. 478-481.