Providing opportunities to demonstrate mastery rather than memory: testing programming skills in a programming environment

Kelsey van Haaster, School of Computer Science and Software Engineering
Monash University
Kelsey.vanHaaster@csse.monash.edu.au

Abstract: First year programming units are commonly assessed by paper-based programming examinations. This component forms a large proportion of the assessment of the unit, and students report that they find both the preparation process and the environment highly stressful. Studying for a closed book, paper-based examination encourages surface learning, rather than understanding. This method of assessment tests a student's ability to perform at only the lower levels of the Cognitive Domain of Bloom's Taxonomy of Learning and does not effectively test performance at the higher levels of Synthesis and Evaluation. Often what is really being tested is a student's ability to memorise information and to perform under examination conditions. The question of whether or not a student has achieved the learning outcomes of a programming unit may be better answered by assessing a student's ability to design, code and test a solution to a real programming problem in a real programming environment.

This paper describes a situation in which students undertaking a first year programming unit are assessed using a programming examination, in a programming environment and focuses on the logistics management and security issues raised by this kind of examination. Preliminary results are analysed, and feedback from the students is documented.

Introduction

The core principles of effective assessment identified in James (2002) are:
- assessment should guide and encourage effective approaches to learning;
- assessment should validly and reliably measure the higher order learning outcomes expected at tertiary level; and
- assessment should define and protect academic standards.

A substantial proportion of the assessment of many undergraduate units takes the form of a paper-based examination contributing between 40% and 70% of the unit assessment. Examinations are often considered by both academic staff and students to be the single most significant indicator of whether the objectives of a unit have been met. This is particularly the case for many students who may have come from an educational culture that measures a student's academic success in terms of their ability to pass examinations (Kam-Cheung 2000).

An evaluation of the paper-based examination as an assessment instrument suggests that, in some circumstances, formal paper-based examinations contravene rather than support these core principles of effective assessment (Scouller 2000).

This paper describes an ongoing project that uses a reflective approach to redesigning the form and content of the final examination in a first year programming unit. This paper focuses on the practical and technical issues associated with the design and implementation of practical examinations, and sets the stage for a detailed study of the educational merits of this approach in the future.

Background and rationale

This project is a direct result of student questions and queries posted on an asynchronous anonymous discussion forum. The forum was designed to provide students with a non-threatening way to communicate with teaching staff and proved very popular amongst students as a kind of live help

desk (Sheard 2002). Just prior to the examination student questions were most frequently related to the mechanical aspects of the examination: the topics to be examined; the format of the examination; and requests for model answers. The number and tone of the posts suggested that many of the students were experiencing a great deal of stress about the examination and were concentrating on trying to recall as much material as possible. This suggests students preparing for the final examination were adopting a surface approach to learning, whilst the focus of the unit and the assignment work was aimed at a deeper approach.

Over the course of several semesters, a number of students commented that sitting a paper-based examination to assess a programming unit made little sense. The objectives of the unit required students to demonstrate evaluation and synthesis and the students themselves were keen to have an opportunity to show that they could apply what they had learned to new problems. Whilst these comments were purely anecdotal, this idea is supported by a number of studies of student attitudes towards different kinds of assessment (Gordon 2002). Students report consciously changing their approach to learning depending upon the form of assessment. Most students state a strong preference for assessment that encourages deep rather than shallow learning and allows them to demonstrate higher order cognitive skills and affective and psychomotor skills (Scouller 2000).

The extent to which a paper-based examination fulfils the third principle of effective assessment is difficult to measure. For a large unit, the examination marking task is lengthy and requires attention to detail over a sustained period. It is made more difficult by the mechanics of paper handling and the difficulties of deciphering poor handwriting or language. Assessing a student's response to a complex examination question requires subjective judgement, yet the marking task is sometimes performed by someone other than the teacher or the writer of the question. All of these difficulties contribute to uncertainty about the standard of traditionally marked paper examinations and the degree to which these define and protect academic standards in the context of learning to program.

Setting a programming examination for a programming unit
Students of a programming unit are required to change the way they study and practise to prepare for a paper-based examination. Throughout the semester students are told the only way to learn programming is to spend time at the keyboard – programming. We encourage them to make use of the compiler and other tools, but then we test their ability to do this away from the keyboard with a pencil and paper.

Software development is a process, yet most paper-based examinations test students' ability to perform isolated parts of that process and do not permit a student to demonstrate the application of the entire process to solve a problem.

An examination for a programming unit should allow students to demonstrate their abilities under the best possible circumstances. It should test the student's ability to solve a problem using the tools (compiler and editor), techniques, and ideas they have learned during their course. The examination content should be complex enough to include the key ideas introduced in the unit. The examination should also permit students to demonstrate a range of techniques and approaches to a problem.

The development process

Redesigning the unit examination required consideration of both academic and practical issues. The academic issues included the design of the examination question and the marking scheme. The examinations in both the cases described took the form of a textual description of a small business problem and required students to demonstrate key skills covered in the unit. A version of the examination question was made available to the students one week before the examination and students were advised that the final question would be identical in form and similar in content to the supplied version. This approach was designed to encourage students to direct their study towards developing an understanding of the principles of programming as introduced in the unit and to

practice applying these principles in order to solve a programming problem. Students were asked to prepare a generic solution to the problem presented and then implement their solution to a specific problem in the computer laboratories under examination conditions. Marks were awarded for code that fulfilled the requirements of the specification as well as style and adherence to the principles covered in the course. Elegance and efficiency were rewarded but were not required to obtain a passing grade. Practical issues to be considered were the effective distribution and collection of examination submissions, and the prevention of cheating.

Stage 1 CSE1203 summer

The summer semester version of the unit had an enrolment of less than twenty students and was less rigidly timetabled than units run during the main teaching periods. This provided an ideal environment in which to introduce the new examination format. The examination question was provided to students one week prior to the examination. The question took the form of a written specification for a module of code and required that input data be read from a text file, analysed and then the results of the analysis written to the terminal. Two computer laboratories were set aside to conduct the examination, the machines in these laboratories were checked for faults, and the operating environment was re-imaged. The laboratories were then locked prior to the examination. Due to the small number of students the software and data required for the examination was installed on each machine individually and the completed examination submissions were collected using floppy disks. The laboratories were disconnected from the university network to prevent students from accessing external resources or communicating with one another during the examination. After completion of the examination the students were asked to complete a web-based anonymous survey on their experiences.

Stage 2 CSE1203 semester 1 2003

Ninety students were enrolled in the unit in semester 1 2003. The larger group meant the practical issues of distributing the examination question, supervision of the examination and collecting the resulting examination submissions needed to be addressed. The unit was using *WebCT* as a vehicle to distribute learning resources, conduct tests, and manage assignment work and it was felt this would provide a convenient and secure way to distribute and collect the examinations. To use *WebCT* for this purpose it was necessary for the machines to remain connected to the university network. This meant the students could potentially communicate with one another, with outside parties and have access to their personal storage space and courseware materials. This clearly compromised the examinations ability to meet the third principle of effective assessment.

A number of approaches were considered and discussed with the faculty's Technical Operations Group. Any proposed solution had to prevent students from accessing disallowed materials or forms of communication with each other, but permit communication with a specific unit hosted on the University's *WebCT* server. The solution also needed to be fast and simple to implement for the technical staff. Because a large number of computer laboratories (6) would be needed, it was also important to minimise disruption to other users of the facilities by limiting the amount of time the laboratories would be unavailable.

The solution, involved the creation of a special examination image, which contained only the software required for the examination and redirected the machines to access the network via a special proxy server. This proxy server permitted access only to the *WebCT* server and students were permitted access only to a special *WebCT* course containing a single submission tool that allowed them to submit their examination solution. The use of the proxy server also permitted the collection of data about which machines attempted to access other sites on the university network and the wider Internet.

The examination was of three hours duration and was held within the official University examination period. As student computer laboratories are routinely re-imaged at the conclusion of each semester much of the infrastructure and many of the procedures required for this project were already in place. The image was able to be rolled out across multiple computer laboratories in less than one business day. This process also permitted the checking of all machines to ensure that they were fully functional. In addition, the maximum allocation of students to each lab was fixed at 12 which allowed for a 25% redundancy rate amongst machines. In order to manage any technical problems the units' teaching staff invigilated the examination. These staff members were able to deal with any software or hardware problems that arose and were able to assist any student who had difficulties submitting their completed examination to *WebCT*.

As an added security measure, students were allocated seats in particular laboratories, this allocation was done by student ID number rather than by name or tutorial group to prevent students from arranging to sit near a friend or other supporter.

As in the summer semester, students were provided with the examination question a week prior to the examination. To discourage memorisation of the solution, students were advised that the question they would be asked to answer on the day of the examination would differ slightly from the preparation question in terms of the kind of analysis they were asked to perform.

Once the examination was complete the submissions were downloaded from the *WebCT* server and checked for completeness. Once all submissions were verified the laboratories were returned to their original configuration and made available for general use, this process being completed in less than one business day. Again, students were asked to complete a short anonymous survey about their experiences.

Evaluation and results

The post examination survey asked the students to indicate whether they had found the process of preparing for this type of examination more or less stressful than preparing for a paper-based examination, the survey asked whether they had found the process of preparing for the examination helped them to consolidate the material covered in the unit and if they had spent more or less time on preparation than they would have for a paper-based examination. Students were also invited to make comments about the examination and the process.

Summer semester 2003
A total of 18 students sat the online examination in summer 2003, of these only six students (33%) chose to return to the unit web site and complete the examination and unit survey. Amongst this group the results of those who completed all the requirements of the unit were very good, with only 3 students failing the examination itself and only 4 failing the unit overall. This represents a pass rate of 83%, which is higher than generally seen during the standard teaching period. This may be because students who take a summer course are often those who are fast tracking their degree and who display a higher than average level of commitment to their studies. The results from this very small sample suggest that the response was very positive. Of particular interest in this context is that 100% of the survey respondents agreed or strongly agreed that preparing for this examination assisted them in consolidating their understanding of the unit materials, and none of the respondents indicated a preference for a paper-based examination. A number of the respondents chose to make comments about their experience, all of which were positive and similar in tone to these examples.

I felt that the online examination was an excellent way to test programming and problem solving skills. It meshes the theory studied in the lectures and tutorials with the application of the newly learnt knowledge. It's the doing of a thing that cements the 'skills' that the course is providing. …

It was good, much better than a written examination. But the only problem I could forsee would be that some students cannot type as quickly as required to finish a question in time. But other than that I think it was a much better way to examine our knowledge.

Semester 1 2003

A total of 84 students sat the online examination in semester 1 2003 and 21 (25%) have chosen to complete the examination survey to date. The survey for semester 1 students is still open and a further invitation to complete the survey will be issued shortly. The final results for those who completed all the requirements of the unit showed some deviation from the usual pattern. Of these students 26 (31%) failed the actual examination and 31 failed the unit overall. This is a slightly higher failure rate (37%) than has been seen in previous semesters. Of those students who passed the examination yet failed the unit, all had failed to submit one or more other assessable component.

The technical arrangements for the examination proved to be very successful and no problems were experienced with any part of the process. A number of students (10) were logged by the proxy server as attempting to defeat the security measures and access external web sites or hotmail accounts, all of these attempts failed. The examination image itself was well suited to its purpose and could quickly and easily be modified to include additional software in order to support online examinations in other units.

It is not possible to provide a reliable analysis of the survey data at this point in time. However responses to date are very similar to those observed in the earlier survey: 76% indicated that they found the process of preparing for the examination helped them consolidate their learning in the unit. Student comments are generally positive but show more variation than in the earlier survey.

The idea of having a week to prepare for the examination was I believe a very fair and great idea for assessment, and to be able to go over any weak points in that time made for great revision and consolidation of the semesters work, being able to do things online meant that if you were nervous and a method name or a semicolon you could compile and if you knew the basics get the code working. …

at first I was a bit uneasy about doing it online but I found out that it was ok. I think we get a second chance when we do it online as we get to see the results of our work before submit it. I think online examinations for programming subjects is a very good idea

Although we're allowed to use earplug, it's still inconvenient sitting on the examination, because the sound of the keyboard being typed was loud. So I felt like in a typing competition. Practical examination is a really good in assessing the students' skills, and it is even better if the lab consists of only few people, max 8-10 people.

Conclusion and further work

The initial motivation for this project was to increase the effectiveness of the final examination as an assessment instrument. However a number of additional benefits have been realised: marking the examination submissions was partially automated by using a test driver and was therefore less time consuming and easier than marking paper-based examinations; accuracy was improved by using a spreadsheet and paper handling overhead was eliminated; the university examinations department estimate that the cost to the school of running a paper-based examination is generally $25.00 per student, whereas the direct cost of the online examination was $12.00 per student; and there is no requirement to store and manage large numbers of examination scripts since online submissions can be stored in electronic format on CD-ROM (This supports easy retrieval should a student wish to view their paper and reduces the workload of the school's administrative staff).

The results obtained to date will be used to inform the continued development of this kind of assessment. A number of areas for improvement were identified: these include refinement of the

examination questions to discourage the practice of memorising code, the refinement of the examination image to improve security, and the development of a submission vehicle to produce a generic solution and reduce dependence on *WebCT*. Improvements need to be made to the speed and reliability of the hardware used to provide the proxy server, and the image roll-out needs to be automated so that the process can be completed quickly and minimise the amount of time computer laboratories are unavailable for general use.

A number of staff members have shown interest in incorporating this kind of examination into their unit assessment. One is planning to do this in semester 2 2003, which will provide an opportunity to evaluate the practicality of this mode of assessment for more than one unit, and to collect sufficient data to evaluate the educational implications of this kind of examination.

References

Gordon, M. (2002) *Changes in approaches to learning: A qualitative investigation of international students at an Australian University*. Brisbane Australia: Australian Association for Research in Education.

James, R., McInnis C. and Devlin, M. (2002) *Assessing learning in Australian Universities*. Melbourne, Australia: The University of Melbourne.

Kam-Cheung, W. (2000) Problems faced by Hong Kong in developing world class education. In *World Class Curriculum 2000*. Victoria, Australia: Curriculum Corporation seventh national conference, 40-46.

Scouller, K. (2000) The influence of assessment on student learning. Melbourne, Australia: Conference of the Australian Association for Research in Education, 4-7.

Sheard, J., Miller, J. and Ramakrishnan, S. (2002) Student satisfaction with a web-based anonymous feedback system. Sunshine Coast, Queensland: The Eighth Australian World Wide Web Conference.