COMPUTERISED PEER-ASSESSMENT THAT SUPPORTS THE REWARDING OF EVALUATIVE SKILLS IN ESSAY WRITING (CAP) & PROGRAMMING (COURSEMARKER)

Phil Davies and Stuart Lewis

Computerised Peer-Assessment that supports the rewarding of evaluative skills in essay writing (CAP) & programming (Coursemarker)

Phil Davies and Stuart Lewis
School of Computing
University of Glamorgan
South Wales
pdavies@glam.ac.uk and sflewis@glam.ac.uk

Abstract

This short paper details recent innovations in the development of peer-assessed environments that support the automated rewarding of the higher order skills of evaluation in two areas of higher education, namely essay writing and computer programming.

The work builds upon past uses and developments of the CAP (computerised assessment by peers) system that has been used as a means of providing a safe, anonymous and qualitative method of assessing student peer-assessment for the past four years within the School of Computing at the University of Glamorgan, South Wales.

Introduction

The use of computerised peer-assessment in the area of assessing students in their development of essay writing has been reported upon over the past few years (Davies, 2000; Davies, 2003). It has also been documented concerning the need to reward the evaluative skills shown by these students in their marking and commenting of an essay in a qualitative manner (Davies, 2003a). The basic advantages of utilising peer-assessment have been well documented in the past (Falchikov & Goldfinch, 2000). However, little work has been previously undertaken in actually quantifying the quality of the student's efforts with regard to providing a 'mark for marking'. The two studies detailed in this short paper show how peer-assessment methods have been applied to both essay writing (CAP) and computer programming (making use of Coursemarker©) in order to both enhance the learning and assessment processes.

Computerised Peer-Assessment of Essays (mark don't write)

This part of the paper reports on a study undertaken within the School of Computing at the University of Glamorgan in the Autumn term of 2003, that

evaluated the validity of assessing students via a computerized peer-marking process, rather than on their production of an essay in a particular subject area. The students were not expected to produce an essay of their own, but evaluate essays set to a previous cohort of students. The study evaluated the higher order skills shown by a student in marking and providing consistent feedback on an essay, measured by making use of a computerised peer-assessment environment. It is important to note that it is the student's ability in both commenting and providing marks that was being judged in this exercise.

The students involved were provided with the coursework pro-forma as issued to the students who were set the original essays. They were provided with basic guidelines concerning the subject area in question. Also basic web links were provided as a starting point for their research. The students were then given two weeks to research this subject area. Once this period of research time had elapsed they were then required to mark and comment the work of at least six essays. These markings were then compared against the previously produced compensated peer marks and comments. By analysing the differences between the 'expected' and 'produced' results, a 'differences' grading was produced. Part of the discussion included within the presentation will be associated with how to map these grades to actual marks.

In order to evaluate the students' abilities, their results in performing this peer-marking process were correlated against their results in a number of computerised multiple-choice exercises and also the production of an essay in a cognate area of the subject being undertaken.

The results overall showed a correlation of the expected results in all three areas of assessment being undertaken, rated by the final grades of the students undertaking the assessment. However, this correlation does not consistently match between the various types of assessment being undertaken. The results produced by quantifying the quality of the marking and commenting of the students, does map well to the overall expectations of the results produced for the cohort of students.

Also identified within this presentation are ways in which the peer-marking process can be improved with regard to marker subjectivity, by using an extension to the CAP (Computerised Assessment by Peers) system. Prior to the marking process taking place, the students provide weightings as to how important they believe each comment is in their view. Also they are able to include comments of their own making, with attached weightings.

Finally it is shown that the higher performing students achieve a greater improvement in their overall marks by performing the marking process, than those of a lower quality. This supports the claim that awarding a 'mark for marking' rewards the demonstration of higher order skills.

Peer-Assessment of Computer Programming using Coursemarker©

The problem of developing the programming skills of students is an area that has created major problems in higher education. Numerous methods have

been attempted to improve students' understanding of 'how to program' (Lewis, 1997). From our experiences at the University of Glamorgan over the past ten years, the area of teaching programming to undergraduate students is one that has produced increasing problems with regard to high failure rates. The Java programming language, although widely used as an initial language in many programming courses due to its flexibility, provides a considerable challenge to many students.

As detailed previously in this paper, the practice of using peers to support/assess each other brings into question the fair rewarding of student effort in this role as marker. This 'reward' must be of a tangible and qualitative nature. It is also necessary that this reward is both objective and automated. If this qualitative reward for providing constructive and evaluative markings of peers' work can be mirrored within the area of computer programming then the general benefits of peer-assessment that have been widely published in the past can similarly be applied to the area of computer programming. The main aim of this study is an attempt to provide an assessment environment that actively promotes the students in attaining a 'deeper understanding' of programming.

This part of the presentation introduces an "add-on" to the Coursemarker© Programming Environment that permits the use of student peer-support and peer-assessment. The students having submitted their own programming assignments for automatic checking by the system are then required to peer-comment and mark the programs of a selection of their peers. The add-on tool allows the students to select from a wide range of pre-defined comments for a particular assignment. By using this system a student is demonstrating higher order skills of evaluation. By including a weighting to each pre-defined comment, it may also be possible to measure the accuracy of a student's marking.

This session reports on the findings of the initial study, and suggests methods that can be utilised to automatically qualitatively assess the peer-marker's ability in performing the marking process. This study is at an early stage in its development process, and further work is currently being performed in providing a user interface that permits staff to 'easily' create a 'bank' of re-usable generic weighted comments.

A benefit that can be gained from the use of this system of peer-assessment is the ability to be able to recognise where plagiarism has been used to produce an initial solution, or a student has received 'excessive' help in their development process. By comparing a student's mark achieved for submitting an assignment against their ability in providing detailed markings of a peer's work, it is could be possible to automatically identify inconsistencies of ability.

By utilising these methods of peer-assessment, students are being assessed on their 'true' understanding of a programming task i.e. do they have the necessary understanding to evaluate the work of their peers, not just 'write' their own solutions.

Conclusions

Both elements described within this paper detail work in progress. However, the preliminary results indicate that the development of an automated option that provides a reward for the 'mark for marking' process is becoming closer to being a reality.

If students are using peer-assessment within their studies as both a means of learning and assessment, and it can be 'proved' that they are being rewarded for showing higher order skills of evaluation, then traditional assessment of essay writing and computer program production as the sole means of assessments may be enhanced, if not replaced?

References

Davies, P. (2000) Computerized Peer Assessment, Innovations in Education and Training International, 37, 4, 346-355.

Davies, P (2003) Closing the Communications Loop on the Computerized Peer-Assessment of Essays, ALT-J, 11, 1, 41-54.

Davies, P (2003a), Peer-Assessment: No marks required just feedback? – Evaluating the Quality of Computerized Peer-Feedback compared with Computerized Peer-Marking, in Cook, J and McConnell, D (eds). *Communities of Practice, Research Proceedings of the 10th Association for Learning Technology Conference (ALT-C 2003). Held in 8-10 Sept 2003, The University of Sheffield and Sheffield Hallam University, UK.*

Falchikov, N. and Goldfinch, J. (2000) Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks, *Review of Educational Research*, 70, 3, 287-322.

Lewis, S.F. (1997), *"Developing a Modula 2 course for Ceilidh,"* CTI Computing, Dublin.