Embedding and enhancing eAssessment in the leading open source VLE.

Conference or Workshop Item

Embedding and enhancing eAssessment in the leading open source VLE.

Philip Butcher
The Open University
Learning and Teaching Solutions
Milton Keynes
philip.butcher@open.ac.uk
labspace.open.ac.uk/course/view.php?id=3484

Christopher Sangwin
The University of Birmingham
School of Mathematics
University of Birmingham
c.j.sangwin@bham.ac.uk
web.mat.bham.ac.uk/C.J.Sangwin/

Tim Hunt
Moodle Quiz component lead
Learning and Teaching Solutions
Milton Keynes
t.j.hunt@open.ac.uk
tjhunt.blogspot.co.uk

Abstract

Automating the assessment of large numbers of students while providing instant personalised feedback requires a multi-faceted computer-based solution. This paper describes how the long term eAssessment developments in science and technology at the Open University (OU) and mathematics at the University of Birmingham have been brought together in Moodle, highlights key features of these developments and discusses the costs and benefits of engaging with these ever more sophisticated systems.

Keywords

eAssessment, feedback, assessment for learning, Moodle, STACK, CAS

1. Introduction

OU computer-marked assignments were transformed by the development of OpenMark (Butcher, 2006) in 2002 to use the interactive features of the internet to provide instant feedback to science students. In parallel the school of mathematics at the University of Birmingham has been using and developing mathematical automated assessments since 2000. Initially adopting the AiM system (Strickland, 2002), a variety of motivations led to the development of STACK version 1.0 in 2004 as a stand-alone system.

Observing the changing technological landscape, in late 2005 one of the OU's STEM Centres for Excellence in Teaching and Learning appointed its first fellows to research into the use of eAssessment for learning. The project was motivated by the perception that recent developments in computing and communications technologies could lead to eAssessments fulfilling a much wider role than had hitherto been possible, combining motivating interactions with rich, and instant, feedback to create engaging eAssessments across a broad

range of learning outcomes (Butcher et al, 2009). Meanwhile the decision was being taken to adopt Moodle as the university VLE with a commitment to invest university resources in the Moodle platform. Again in parallel STACK v2.0 inserted questions into Moodle's quiz using web services but retained a separate identity as a project. Substantial support was provided for these STACK developments by the University of Birmingham and the Higher Education Academy's Maths Stats and OR Network from 2000-2011.

In the ensuing years these initially separate developments combined and most recently STACK v3 was written as a collaboration between the authors as a dedicated question type for the Moodle Quiz. The results are available today to all Moodle users.

2. Adoption and embedding

Since the start of the CETL initiative and the adoption of Moodle the use of eAssessments at the OU has grown 100 fold (Figure 1 and noting that numbers for 2005-2006 are too small to register) and in 2012 over 60 OU modules in STEM subjects from first to third level use eAssessments.

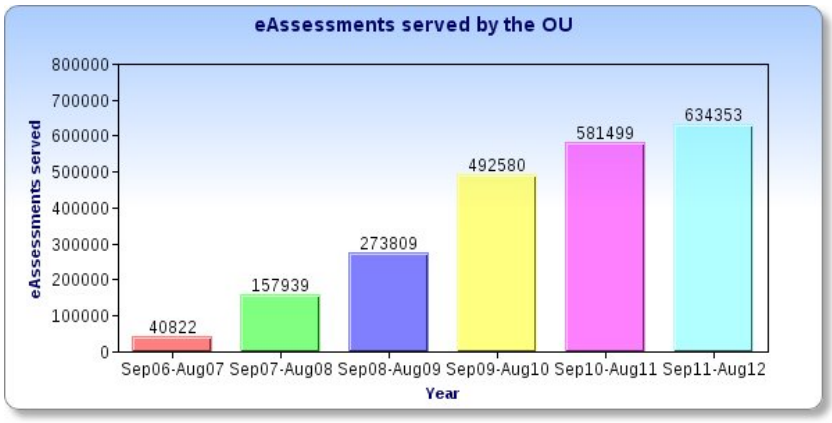

Figure 1: Overall OU usage of eAssessments by calendar year. Approximately 145,000 of the Sep11-Dec12 eAssessments were submitted as part of a formal module assessment strategy.

In 2012 a typical module will have eAssessments run by the Moodle *Quiz* scattered through the online course *Calendar*. *Reviews* for students give feedback and marks and are extended to *Reports* with classical test statistics for tutors. Marks are summarised in the *Gradebook* for onward transfer to the student record system. In other words eAssessment facilities are now embedded into the fabric of Moodle. All of these tools have been redeveloped in recent years and are included in the latest version of Moodle 2 (v2.4 was released in December 2012) (Hunt, 2012).

3. STEM influences

The major similarity between the projects was their desire to engage students in an interactive exchange around one or more learning outcomes, with the computer providing instant feedback and multiple attempts for students who answer incorrectly. At the OU the overall STEM CETL project was titled *eAssessment for Learning* and all projects included teaching feedback, often with course references, to persuade students to revisit topics where their answers are incorrect, before attempting the question again. The theoretical

and evaluative studies arising from these projects contributed to the blue-print for the OU's Moodle redevelopments which highlighted:

- *The emphasis placed on feedback.* All OU students are distance learners and the university emphasises the importance of giving feedback on written assessments. The redevelopments enable feedback, perhaps at multiple levels, to be included.
- *Allowing multiple attempts.* Computers are interactive systems, and consequently can ask students to act on feedback that is given 'there and then', while the problem is still in their mind. If their first answer is incorrect, they can have an immediate second, or third, attempt. A key development has been that of 'question behaviours' which support both instant and delayed feedback.
- *The breadth of interactions supported.* The full capabilities of modern multimedia computers should be used to create engaging assessments. The addition of STACK described below exemplifies this approach.

The framework that has been developed for the Moodle Quiz by the 3rd author of this paper is now available to all question types as well as those provided by the OU and the University of Birmingham (Moodle, 2012).

3.1 A typical user interaction with a STACK question

STACK (System for Teaching and Assessment using a Computer algebra Kernel), as developed by the authors, is now a dedicated question type for the Moodle quiz. As such it is able to take advantage of the options in the quiz, such as the behaviours and reporting. A typical question is shown in figure 2, in adaptive mode. A student has made an attempt at this question and the figure contains the feedback provided to the student.

The central goals of the question type are to (i) randomly generate versions of questions in a structured mathematical way; (ii) accept answers which are mathematical expressions; (iii) establish the mathematical properties of these answers automatically and (iv) generate outcomes which can be used for formative, summative and evaluative purposes. Although this question has only one *input element,* i.e. answer box, the type supports an arbitrary number of input elements, enabling the teacher to break a complex task into pre-determined steps and where appropriate implement follow-through marking.

In many areas question validity is seriously undermined by adopting multiple-choice, or similar, question formats. This is particularly problematic in mathematics. Our motivation was to significantly broaden the range of valid assessments available for mathematics by accepting *a mathematical expression as the answer* from the student, while acknowledging that automatic assessment of arbitrary mathematical arguments (i.e. proof) in general remains unfeasible for the foreseeable future.

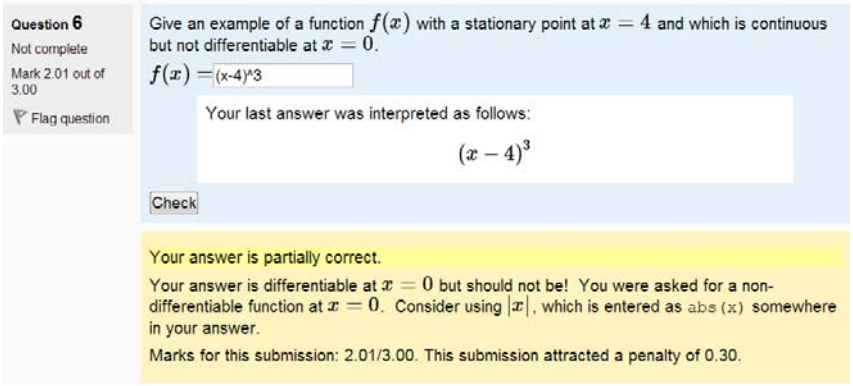

Figure 2: a typical STACK question.

Underpinning this is the computer algebra system (CAS) Maxima. These tools enable teachers to work backwards from a randomly generated answer, through the steps in a model answer, to create a feasible and fair question version. The original motivation was to reduce plagiarism and impersonation and to provide opportunities for repeated practice. Our experience suggests that many students also appreciate the opportunity to collaborate solving different versions without cheating themselves, something which is not possible in traditional fixed formats (Sangwin, 2013).

Students require some knowledge of syntax. This is arguably a valuable independent skill in the 21st century, but sensitive design supports this process and seeks to mitigate the consequences of invalid input rather than "incorrectness", which is particularly important in high-stakes settings. Specific comments about this important topic are given in Sangwin et al (2007) and also in Sangwin (2013).

The core role of the CAS is to establish the mathematical properties relevant to the question. Some of these may be central, and others underlying conventions such as avoiding the use of floating point approximations, or using particular algebraic forms. Testing for algebraic equivalence with the "correct answer" (where this is unique) is the prototype test. STACK also enables the teacher to check whether answers are in a correct form, such as factored. In general correctness might require a number of separate properties, and feedback can be provided for each. In figure 2 there is not a unique correct answer. Instead STACK seeks to establish each property separately. The formative feedback shown is designed to help students improve on the task, something which research such as Kluger et al, 1966, suggest is key. In this behaviour students are encouraged to try again, although the teacher in this situation has decided to impose a small numerical "penalty" for each incorrect attempt.

4. Costs and benefits

Online eAssessment was first introduced, tentatively, on one OU science module in 2002. Since then multiple studies have been undertaken to establish student attitudes to, and use of, eAssessments. Jordan (2011) in particular has undertaken an extended study of their use on a range of Science modules and reports that "...e-assessment can be extremely effective

in supporting learning...". The results of these studies have been sufficiently persuasive that the use of eAssessments has spread to 60 STEM modules at the OU. But being 'extremely effective' is only one part of the equation. A rudimentary glance at the example question in section 3 will reveal that the author has thought about how students might go wrong and included feedback to help. In fact the author observed students using an early version of this question, collected their responses and developed his feedback scheme on the basis of his analysis of their answers and comments; a costly undertaking. But now the question exists it can be used by any number of students, 24/7, year round and year on year from wherever they might be. And it is this balance of costs and benefits that drives our colleagues in choosing when to develop eAssessments and how sophisticated to make them.

Realising the learning gains afforded by the sophistication now available in Moodle and exemplified by STACK questions requires considerable academic investment. Consequently readers will be pleased to hear that 450 questions in calculus, algebra and other procedural topics initially released for STACK 2 as outcomes of various previous projects are included with STACK 3 under a Creative Commons licence.

5. Conclusions

Moodle is the leading open source VLE and for both sets of authors has proven to be a suitable vehicle for developing innovations over a number of years. In recent years the increasing adoption and demands by a growing number of modules at both universities have led the authors to combine their expertise and resources to fully embed and thereby enhance the eAssessment facilities available in Moodle. The results are available today to all Moodle users who are invited to join us in this ongoing collaboration.

6. References

Butcher, P. G., (2006), *OpenMark Examples*, http://www.open.ac.uk/openmarkexamples

Butcher, P. G., Swithenby, S. J., & Jordan, S. E. (2009), eAssessment and the independent learner. *23rd ICDE World Conference on Open Learning and Distance Education, Maastricht*.

Hunt, T., (2012), Computer-Marked Assessment in Moodle: Past, Present and Future. *CAA 2012 International Conference, University of Southampton* (ed. Whitelock, D., et al)

Jordan, S. (2011), Using interactive computer-based assessment to support beginning distance learners of science. *Open Learning*, *26*(2), 151-68.

Kluger, A. N. and DeNisi, A., (1996), Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory., *Psychological Bulletin* 119(2), 254–284.

Moodle, (2012), *Question types*, https://moodle.org/plugins/browse.php?list=category&id=28

Sangwin, C. J., (2013), *Computer Aided Assessment of Mathematics*, OUP.

Sangwin, C. J. and Ramsden, P., (2007), Linear syntax for communicating elementary mathematics, *Journal of Symbolic Computation* 42(9), 902–934.

Strickland, N., (2002), Alice interactive mathematics, *MSOR Connections* 2(1), 27–30.