# EXTENDING FLEXIBILITY IN AN EXISTING ON-LINE ASSESSMENT SYSTEM

H S Ashton and C E Beevers

# Extending Flexibility in an Existing On-line Assessment System

H S Ashton C E Beevers

#### Scottish Centre for Research into On-Line Learning and Assessment School of Mathematical and Computing Sciences Heriot-Watt University Edinburgh EH14 4AS

h.s.ashton@hw.ac.uk

# Abstract

This paper seeks to extend the work of seventeen years of research and development in the field of computer aided assessment. Work began on the Computer Aided Learning in Mathematics (CALM) Project in 1985 at Heriot-Watt University (http://www.calm.hw.ac.uk/). But, more general issues of automatic assessment are now being considered in collaboration with UK Examination Boards, commercial companies (http://www.i-assess.co.uk/) and Scottish academics through the forum of the Scottish Centre for Research into On-Line Learning and Assessment (SCROLLA) (http://www.scrolla.ac.uk/). To set the debate in context some of the main results of the CALM Project will be briefly reviewed.

Two of the areas of research in SCROLLA are investigation into automatic assessment of higher order skills (Beevers *et. al.* 2003) and how traditional paper based questions translate into an equivalent on-line version using existing question types (Fiddes *et. al.* 2002). For example, the rewording of a question to allow it to be delivered in an assessment engine can mean that the student is provided with additional information that may not be available in the traditional paper based version, or the rewording may change the skills that are being tested. Some of these issues can be addressed by creating new question types to increase the flexibility given to the author in creating questions, and to increase familiarity of the students in the responses they can provide.

The assessment system CUE has been used in UK programs such as SCHOLAR (http://scholar.hw.ac.uk/), which provides a wide range of questions in a variety of subjects (Higher, Advanced Higher and A Level in Biology, Chemistry, Computer Science, Mathematics and Physics). This provides a stable base from which CUE can expand. SCROLLA intends to provide CUE as a research and development resource for education institutions in Scotland from schools through to higher education over a range of subject areas.

CUE provides an opportunity for exploring the research issues further and investigating possible solutions that more flexible question types may provide. Using examples chosen from Computer Science and Mathematics this paper will illustrate where some of the restrictions with currently available question types occur, and offer potential solutions for discussion and comment.

### What is CUE?

The assessment system known as CUE is the result of a number of projects carried out at Heriot-Watt University over the last 17 years. Through projects like CALM (Computer Aided Learning in Mathematics) (http://www.calm.hw.ac.uk/; Beevers *et. al.* 1988; Beevers *et. al.* 1991), Mathwise (Harding *et. al.* 1996; Beevers *et. al.* 1998) and Interactive Past Papers (Beevers *et. al.* 1997) CUE has developed into a sophisticated assessment engine. CUE, and its predecessors, have also been used in a wide variety of subjects and at Secondary School (through the SCHOLAR project (http://scholar.hw.ac.uk/)) and University level (Beevers *et. al.* 1995).

In 1999 a collaboration between the CALM group, the University of Cambridge Local Examination Syndicate (UCLES) and the commercial firm EQL of Livingston moved the existing assessment technology from the PC to web delivery and the CUE assessment engine was born (Fiddes *et. al.* 2000).

CUE has many of the question types expected from an assessment engine, including multiple choice, multiple response, hotspot and word or phrase match. In addition CUE has a judged mathematical expression (JME) question type that enables mathematical expressions to be marked. This feature allows equivalent expressions to be marked correctly, or unsimplified expressions to be marked incorrectly if desired. This is one of the features, which makes CUE more appropriate for scientific testing than other systems. More details on the CUE assessment system can be found in Paterson (Paterson 2002a).

Many of the features of CUE have been developed with educational issues in mind. For example, CUE enables authors to set questions which have optional built in steps which the student can choose if they cannot answer the question in one step. This is useful in formative and summative assessment. In the former case it helps the student to move forward in their learning and construct strategies for solving problems, whilst in the latter case it enables partial marks to be awarded if they are unable to complete a whole question. The reader who is sceptical of the crucial part partial credit plays in the life of a student should consult the recent article by Strickland (Strickland, 2002). For a fuller discussion on the issues of partial credit within a CAA system the interested reader is directed to the article by Beevers (Beevers *et. al.* 1999).

Accurate and helpful communication with the computer is also vital. In Mathematics this poses special challenges since many mathematical expressions are normally displayed over several lines on paper but the computer usually requires a one-line input. In the Mathwise Project (Beevers *et. al.* 1998) and for the Interactive PastPapers series (Beevers *et. al.* 1997) an Input Tool was created to aid this communication. It has not yet been

possible to construct an Input Tool for the web version but plans are underway to resolve this omission. Indeed, an intermediate solution already exists in CUE in the form of a rendered version of the one-line input, although this is not as good as a dynamic display. However, since CUE has used the MathML mark up it has also been possible to go some way towards making CUE accessible for blind and partially sighted users. Any mathematical expression which is rendered on screen to display as it would appear on paper can be "read" by screen reading packages like JAWS.

From an early stage of development it became clear from educational evaluations that it would be helpful to be able to deliver the same questions in a number of modes. CUE has three basic modes, help, practice and exam, with the test author being able to configure this to their own requirements. In help mode the questions are delivered with maximum feedback including hints, visible marking on screen and the chance to reveal a correct answer. This has proved popular with students when they start out on their learning. Practice Mode restricts the help to just visible ticks and crosses on screen for right and wrong responses. Finally, exam Mode presents questions with no option for revealing answers and no ticks/crosses appearing. It was also necessary in one research project with UCLES to restrict the browse facility and force students to take the questions in a certain order. These options for test delivery are now available as choices for the author of a CUE test.

In Science, Engineering and Mathematics another powerful driver in CAA is the use of randomisation. This feature enables the inclusion of numeric variables in questions that CUE will randomise on delivery (within an authors specified range). By simply changing two or three parameters within a problem numerous questions can be generated from one original question. This means that every time CUE delivers that question it is likely that the student will see a different set of numbers. This has two effects: in formative assessment the students are given plenty of practice and in grading tests it prevents copying from the screens of adjacently seated students.

The insight and experience from years of development and use, the development of technology, and the development of authoring tools for questions and tests has resulted in a sophisticated, easy to use, computer based assessment system. However, there is still a lot of room for further development. Some of the current ideas for development are discussed in the remainder of this paper.

#### An Equation Type

Recent work by colleagues in the area of automatic assessment of mathematics (Paterson 2001, Paterson 2002b, McGuire *et. al.* 2002) has highlighted some potential areas of development for CUE. For example, the introduction of a question type to deal with equations.

A typical question may be



A student would be expected to supply an answer such as y=2x+1. Currently this is not possible in CUE as equations cannot be entered (however mathematical expressions can). One possible solution has been to express the question in a slightly different style, i.e.

A STRAIGHT LINE HAS A GRADIENT OF 2 AND PASSES THROUGH THE POINT X=0, Y=1. FIND THE EQUATION OF THIS LINE. EXPRESS YOUR ANSWER IN THE FORM Y=MX+C. **KEYPART 1: EQUATION OF THE LINE** Y=?

In this case the student could provide an answer of 2x+1, or any mathematically equivalent expression, and this can be marked by CUE. However, there are differences in these two questions.

Firstly, the information provided by the question is different. In the second question we have provided the student with the form of the equation of a

straight line. This seemingly innocent change could prompt the student into being able to answer the question, or change the learning objectives being tested (Paterson 2001).

Secondly is the issue of familiarity. Students would be used to expressing the answer in the form y=mx+c, not just mx+c. Changes of this kind could confuse some students.

Thirdly, this may not be the only, or the most appropriate way to express the answer. For example, consider the same question, but with a gradient of 1/2. A student may wish to express the answer in the form y=x/2 + 1. However, equally correct is the response 2y=x+2. Expressing the question in the second format could rule out this form of the answer.

The checking of equations, rather than merely expressions, does pose automatic testing systems some difficulties. These have been overcome in the Treefrog system designed by Strickland (Strickland, 2001) and could be overcome in CUE. What is needed is the ability to rearrange the equation so that all the terms are on one side and then test the correctness using existing CUE procedures. In addition the equation type would need to have the ability to remove common factors in the equation. The student would also need to be informed of the required variables (x and y in this case). The CUE system could then set the question as:



This would allow several answers to be marked correctly, y=2x+1, y-1=2x, (y-1)/2=x, y/2 = x+1/2, etc.

The removal of common factors in mathematical expressions is already a feature of CUE so the extension to handle equations is possible.

# Hidden Multiple Choice

Multiple Choice questions can be very powerful and have been used extensively on paper for a number of years. They are also frequently used in computer assessment systems as they are easy to mark automatically. Currently the CUE author has a number of layout options for multiple choice questions including using a drop-down list or radio buttons, altering the number of rows and columns for the display of the possible answers, and the inclusion or absence of a submit button. However, multiple choice questions are often criticised as the student can adopt a "guessing" strategy to reach the most likely correct answer. Estimates in Bull and McKenna (Bull *et. al.* 2001) on the odds of students successfully guessing their way through a test are not always convincing especially if the student chooses to adopt alternative strategy to that envisaged by the setter of the questions (Lawson, 2001).

One possible extension to multiple choice, and multiple response, questions is that of developing hidden multiple choice (HMC). With this setting the student would be presented with the options one at a time and would have to make a decision as to the correctness of the answer one at a time. Once the student has made a decision it cannot be changed without resitting the question. The student will not see all the options at once and so cannot reject some answers and reduce the options before deciding on their answer. With HMC the student would not be permitted to view all the choices unless the correct option is the last alternative presented to the student. This format forces the student to recognise the correct answer when it appears on screen from any working made on rough paper. An issue with such a question type is how it is dealt with in help and practice modes. Should the student see information (tick or cross) at every stage in a multiple choice/response, or only once all decisions have been made? Would the student then see all options and have a chance to change their mind, or be presented with each option in turn again? HMC was a feature of the work from Sunderland University in the late 1980s (Middleton et. al. 1990).

#### Extended Randomised Multiple Choice

Currently CUE has the ability to randomise multiple choice and multiple response questions. That is, the order of the options presented to the students can be randomised. In addition randoms can be used in these options allowing differing numeric values to be displayed automatically by the assessment system. This is extremely useful in mathematical based subjects. However, there is a further extension that may prove useful in multiple choice/response questions – the ability to have the assessment system choose from a bank of options, allowing more flexibility in none numeric based subjects.

For example, consider the following question:



At present each student would see the same four options presented to them, but in a different order. It would be useful if the options in the question could be randomised in a similar manner to randomising numeric options. In this situation there are a number of other correct and incorrect answers that could be presented to the student. What is needed is a new feature to enable this.

This would then allow the author to specify a list of correct answers, and a list of incorrect answers. The author would then specify how many options should be displayed from each list, i.e. display two correct answers and two incorrect answers. The order these options are presented to the student would then be randomised. This would provide benefits in the same manner that allowing random numeric variables in questions enables the generation of a number of questions from the same template, allowing multiple questions for practice, or eliminating copying.

# Follow Through

Throughout the educational life of students they are often told that they will not be penalised for the same mistake twice in one question. For example, consider the situation where the student answers the first part of a question incorrectly but the remainder of the question requires them to use the answer to the first part. In this situation the student should still be able to obtain full marks for the rest of the question if the technique used is correct. The notion of "follow through" may provide the ability to take account of such situations. For example: A GREENGROCER BUYS NINE ORANGES AT THE MARKET FOR £1.50 THEN SELLS THEM AT 29P EACH.

KEYPART 1: HOW MUCH (IN PENCE) DOES THE GREENGROCER HAVE WHEN ALL NINE ORANGES ARE SOLD?

KEYPART 2: HOW MUCH PROFIT (IN PENCE) IS THERE WHEN ALL NINE ORANGES HAVE BEEN SOLD?

KEYPART 3: WHAT IS THE PERCENTAGE PROFIT?

In this example the correct answers are 261, 111 and 74. However, a typical miscalculation in the first keypart could mean the student continues with an incorrect answer, for example 252 might be a common mistake. For the remaining two keyparts of the question the student uses the answer of 252 and gives answers of 102 and 68. Notice here that although the student has incorrectly calculated the amount of pence at keypart 1 the student has then gone on to correctly work out the profit (252 - 150) = 102 and determined correctly the percentage profit. Currently this student would obtain no marks for this question.

If "follow through" has been applied keypart 1 would be marked as wrong, however keyparts 2 and 3 would be correct even though the values do not compare with our model solution. In this situation a human marker might give the student 2 out of 3 marks.

When writing a question that could be used in follow through mode the author would need to specify the answers to questions in terms of the answer to the other parts of the question. It is likely that this would be done by assigning a variable name to the student answer to a specific question. For example, the author would specify the correct answer to keypart 1 as 261. The author would also be able to give the student answer to this a name, i.e. sale. The answer to the second keypart would then be given as (sale -150), and the third part as 100\*(sale - 150)/150. When a test is set the test author would choose whether follow through is active. If follow through is not active then the answers are calculated by the system using a value of sale as 261. However, if follow through is activated then the answers would be calculated using the students answer to keypart 1 as the value of sale. In this way the same question can be used with and without follow through. In addition the author could choose to have follow through on keypart 3 from keypart 2, for example the answer to keypart 2 would still be (sale-150), but this would be given the name profit. Keypart 3 could then have the answer 100\*profit/150. This would enable complete follow through to be used.

### Conclusion

CUE currently provides a powerful assessment engine for the testing of a wide range of skills particularly in Science, Engineering and Mathematics. In this paper we have discussed a number of extensions to question types and marking schemes that have the potential to expand the usefulness and flexibility of the CUE online assessment system, and computer based assessment in general.

#### References

Beevers, C.E., Cherry, B.G., Clark, D.E.R., McGuire, G.R. and Renshaw, J.H. (1988) *The CALM Before the Storm!*, Computers and Education, 12, 43-47.

Beevers, C.E., Cherry, B.G., Foster, M.G. and McGuire, G.R. (1991) *Software Tools for Computer Aided Learning in Mathematics*, Ashgate Publishing Company.

Beevers, C.E., McGuire, G.R., Stirling, G. and Wild, D.G. (1995) *Mathematical Ability Assessed by Computer*, J. Comp. Educ. 25, 123-132.

Beevers, C.E., Fiddes, D.J., McGuire, G.R., Wild, D.G. and Youngson, M.A. (1997) *Interactive Past Papers for A Level/Higher Mathematics*, Lander Educational Software, Glasgow, ISBN 1-85572-431-6.

Beevers, C.E., Bishop, P. and Quinney, D.A. (1998) *Mathwise Diagnostic Testing and Assessment*, Information Services and Use 18, 191-205.

Beevers, C.E., Youngson, M.A., McGuire, G.R., Wild, D.G. and Fiddes, D.J. (1999) *Issues of Partial Credit in Mathematical Assessment by Computer*, ALT-J 7, 26-32.

Beevers, C.E. and Paterson, J.S. (2003) *Automatic Assessment of Problem Solving Skills in Mathematics*, accepted for publication by Active Learning in Higher Education.

Bull, J. and McKenna, C. (2001) *Blueprint for Computer-assisted Assessment* ISBN 1-904020-00-3 http://www.caacentre.ac.uk/bp/ (16 May 2002).

Fiddes, D.J., Jackson, D.H. and Beevers, C.E. (August 2000) *Assessments* over the Web using CUE, MSOR Newsletter, **0**(3) 17-20.

Fiddes, D.J., Korabinski, A.A., McGuire, G.R., Youngson, M.A. and McMillan, D. (2002) *Does the mode of delivery affect Mathematics Examination results?* ALT-J 10, 62-69.

Harding, R.D. and Quinney, D. (1996) *Mathwise and the UK Mathematics Courseware Consortium* Active Learning 4, 53 http://www.ilt.ac.uk/public/cti/ActiveLearning/issue4/harding/index.html (16 May 2002).

Lawson, D. (October 2001) *Computer Assisted Assessment in Relation to Learning Outcomes*, MSOR series: Computer-Aided Assessment in Maths http://ltsn.mathstore.ac.uk/articles/maths-caa-series/oct2001/index.shtml (16 May 2002).

McGuire, G. R., Youngson, M. A. and Korabinski, A. A. (2002) *Partial credit in mathematics exams – a comparison of traditional and CAA exams* to be presented at CAA2002, Loughborough.

Middleton, W., Curran, D. and Moscardini, A.O. (1990) *Remedial Mathematics in Higher Education - A Computer Based Approach*, Presented at the 7th International Conference on Technology and Education Brussels.

Paterson, J.S. (2001) *Which of Bloom's cognitive skills can be tested by automatic assessment in mathematics* Napier University MSc CEME (Computer Enhanced Mathematics Education).

Paterson, J.S. (April 2002a) *The CUE Assessment System* MSOR Connections, **2**(2) Birmingham http://ltsn.mathstore.ac.uk/articles/maths-caaseries/apr2002/index.shtml (16 May 2002).

Paterson, J.S. (2002b) *Linking on-line Assessment in Mathematics to Cognitive Skills* to be presented at CAA2002, Loughborough.

Strickland, P. (August 2001) *How should a perfect computer aided assessment package in Mathematics behave?* MSOR Connections **1**(3)

http://ltsn.mathstore.ac.uk/newsletter/aug2001/pdf/treefrog\_editor.pdf (16 May 2002).

Strickland, N.P. (February 2002) *ALICE Interactive Mathematics*, MSOR Connections **2**(1), 27-30 http://ltsn.mathstore.ac.uk/newsletter/feb2002/pdf/aim.pdf (16 May 2002).