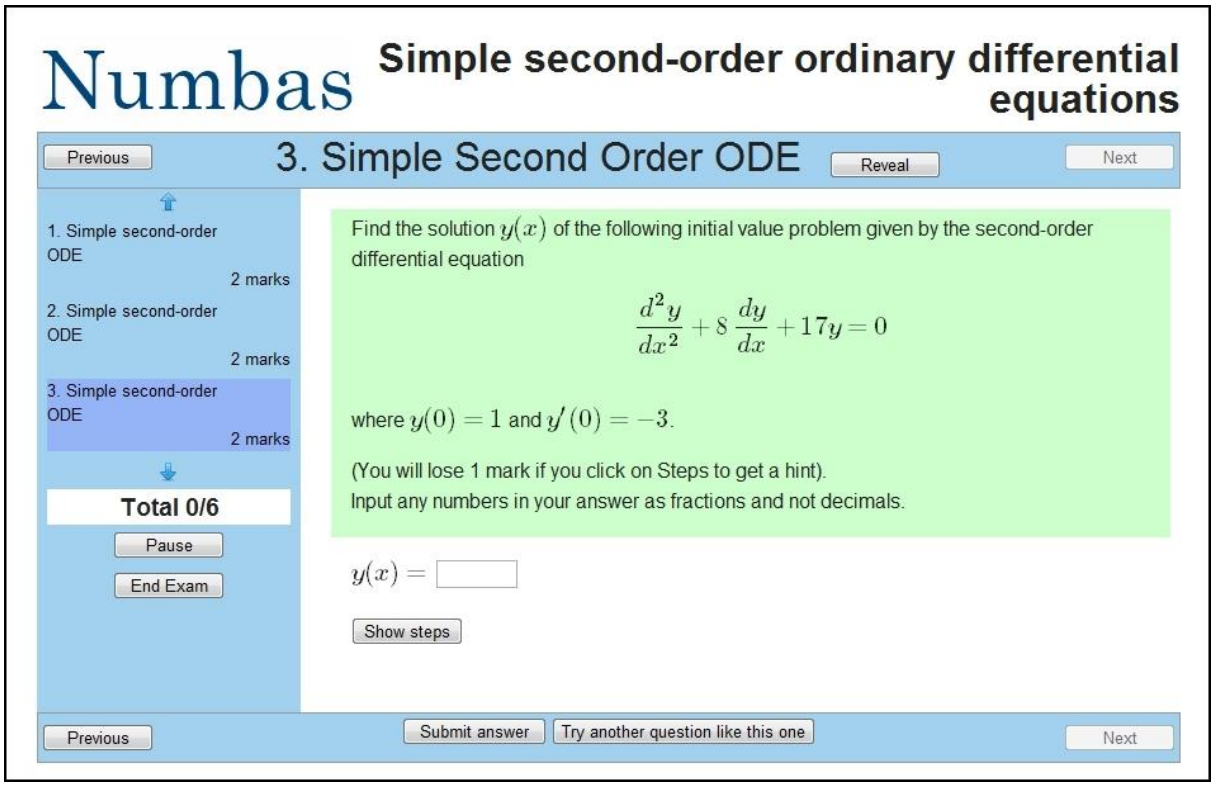
A completely client-side approach to e-assessment and e-learning of Mathematics and Statistics

Bill Foster, Christian Perfect and Anthony Youd,
E-learning Unit of the School of Mathematics and Statistics, Newcastle University.
http://www.ncl.ac.uk/maths/outreach/elearning/
e-mails: Bill.Foster@ncl.ac.uk , Christian.Perfect@ncl.ac.uk , a.j.youd@ncl.ac.uk
Numbas website: http://www.ncl.ac.uk/maths/numbas/

Abstract.

This article introduces Numbas: a new SCORM-2004 compliant open-source and multi-platform e-assessment and e-learning system developed and used at Newcastle University. The main focus of implementation is on rich formative e-assessment and learning; blending powerful mathematical and statistical functionality with its unique browser and client-based design, bringing into play the full capability and resources of the internet. It can be used for all numerate disciplines in education and training and builds upon and extends successful designs and implementations used for many years in HE, FE and secondary education. As well as describing its unique features and their future development, the reasons behind its design are explained and contrasted with other powerful open-source assessment systems aimed at numerate disciplines.

Keywords: formative e-assessment, e-learning, open-source, client-side, SCORM, mathematics and statistics.



Introduction and Overview

Numbas is a new, unique, powerful web-based e-assessment and e-learning system which is freely available to all, subject to an Apache 2.0 permissive licence. [1] It has been developed by the E-Learning Unit of the School of Mathematics and Statistics, Newcastle University, based upon the extensive successful use and experience of e-assessment and e-learning within the School since 2006 using the i-assess [2] system. See the Numbas web-site for more details and information on how to get hold of Numbas.

There have been over a quarter of a million student sessions using i-assess and this use has been spread across all Faculties within the University. In particular, the use of formative e-assessment within the School of Mathematics and Statistics is fully integrated into the curriculum for the first two years of the mathematics and statistics degree courses. Detailed descriptions of the use at Newcastle can be found in [3,4,5].

The success of this implementation has been recognised by the award from the E-Assessment Scotland Conference (2010) for Formative e-Assessment

Numbas has many more features than i-assess and is replacing it at Newcastle. Special features include:

- can be delivered across many platforms;
- enabling higher levels of mathematics and statistics applications if needed;
- adapted and personalised feedback for students;
- a rich multi-media environment;
- easy and almost instantaneous installation.

Newcastle is distributing Numbas on Blackboard for use in our courses as well as including it in e-learning and revision DVD packages (with accompanying video and text) for international students, engineering students, business studies students, mature students and also for new entrants as transition aids.

We have an open-source GitHub repository focused on the further development of Numbas. For example, the present method of authoring using a mark-up language is to be replaced by a web based authoring system. Also we plan to co-operate with other departments within Newcastle and other HE institutions on the creation of a database of questions, tests and e-learning materials.

The issues that Numbas addresses include:

- Adaptive, high quality and effective feedback to large classes. This is a major issue in the National Student Survey[17]. The present system, i-assess, has feedback appreciated by students but we need more personalized and immediate feedback which Numbas provides.
- Engagement. The present i-assess system has good engagement levels for in-course assessment where it is used; the introduction and spread of Numbas will consolidate and improve this.
- Student expectation. Come 2012 there will be an increasing pressure for a range of high quality teaching and assessment materials, which are under the control of HE and can be quickly adapted following student feedback.
- Meeting the requirements of differing student markets especially international and mature students by producing good quality e-learning packages which fill gaps in required skills.

Numbas can be seen as belonging to the well-established family of computer-based assessments descended from the original CALM system developed at Heriot-Watt[6]. These include i-assess and

the computer based assessment systems within SCHOLAR[7] both used widely and successfully in various sectors of education and training. As a result, Numbas has benefited greatly in both the technical and pedagogic design from over 25 years of experience and application.

The Drivers for the Development at Newcastle University

The implementation of i-assess since 2006 had been a success, well received by the students, recognised as of high quality by the HE and e-Assessment community and firmly integrated into the curriculum in the School of Mathematics & Statistics as formative assessment for both first- and second-year courses as well as large service courses.[3,4,5]

Why Change?

We increasingly found we needed extra functionality and more control over the content, delivery and feedback than i-assess provided. We also wanted the ability to deliver e-assessments to local schools and colleges using a variety of platforms as part of an outreach initiative.

Display and input of mathematics was an issue, in particular we wanted to be able to monitor and advise on student input of mathematical expressions. Also we wanted to easily include rich content such as videos, graphs and other interactive elements in questions.

Client-side assessment functionality

An important decision motivated by the above drivers for change was to have all assessment functionality handled by the client machine As computers have become more powerful, they can easily and quickly handle the computations necessary in mathematical and statistical assessments. There is no need for the server to handle such calls and it is a natural step to separate out the assessment engine from the record keeping and question database components. Furthermore, such a splitting increases interoperability: you only need, for example, a SCORM-compliant system, not some standard which looks at both maths and session-tracking.

As a result, all Numbas assessment and computational functionality is carried out on the client machine and data storage of user records etc. can be either on a server for assessed work, or on the client for local use.

Some other advantages:

- Numbas can be used straight away without installing any new software on the client or on servers.
- The marking of mathematics can involve a lot of (potentially unstable) computation and it is better that, if it goes wrong, only one client crashes rather than the server administering every student's session.
- Much less network communication takes place, so many more simultaneous student sessions are possible in examination conditions.

There is an immediate issue of security if assessment takes place on the client. It is possible to find the correct answer for a question on the client machine by various methods if the browser is not locked down. For high-stakes invigilated assessments this lock-down is perfectly possible and is routine at Newcastle, for example, using the OLAF lock-down browser. There are many such examples of such browsers.

For non-invigilated assessments there has always been the possibility of collusion or cheating and the issue is one of policy. At Newcastle we will be using Numbas in its present state for the purpose it

excels at -- formative e-assessment with a very strong e-learning emphasis and high-quality feedback -- and this is the optimum way to use such a rich system. To use such systems only for one-off high-stakes summative assessments is clearly inappropriate and a waste of valuable learning resource. If security is required in any mode then a wrapped browser can be easily developed and accessed through a Remote Access Server.

Improved Feedback.

Following on from the observation above about the emphasis on formative e-assessment, the evolution of the formative assessment system at Newcastle since 2006 has followed and can be mapped onto most of the seven principles laid out in Nicol and Macfarlane-Dick(2006)[8] for self-regulation. However, weaknesses have existed in providing timely contextual information to individual students (for example, why they are doing the assessment and what is it needed for in terms of the learning objectives of the course) as well as promoting teacher-student and peer dialogue.

Given the issues raised by the National Student Survey [17] on feedback, these have to be addressed using an extensible and reactive assessment and feedback environment under our control.

Basic Requirements

Our requirements of any new system boiled down to the following criteria:

- It must be possible to carry over our large bank of questions (about 600 question types) to the new system more-or-less as they were originally designed and without too much investment of time or effort.
- It must provide the opportunities for enhanced advice and feedback features including more adaptability to individual student input.
- Given the success of the CALM-style formative e-assessment methodology, to continue this basic CALM design.
- There must be robust mathematics capability, including the ability to recognise, parse and evaluate symbolic mathematical input, to randomise mathematical expressions and to use LaTeX for their display.
- It should be possible to use rich content such as videos and graphs and any available web resource.
- The interface presented to the student should be clear and intuitive and easily configurable.
- Assessments should be accessible on all the platforms students use; these days that means not just PCs and Macs but also mobile devices such as the iPad.

Other Systems and Interoperability

There exist several excellent, well-established and well used sophisticated mathematical assessment systems which accept the input of and process symbolic mathematics, such as STACK [9], Maple TA[10]and DEWIS [11] but since they work so differently to the CALM design it would not have been feasible to convert our question banks to these systems.

MathAssess[12] has been developed as a system which attempts to extend the QTI toolkit to MathQTI and the aim is to achieve interoperability between the various systems, and in particular those which use algebraic input. However, this is an ambitious project and one as yet not realised in terms of interoperability as translations still have to be built between the various data structures representing the parsed algebraic expressions in the various systems. This will be difficult.

Apart from technical issues of interoperability there arises an important point about the design of e-assessment of mathematics: since a wide range of skills is being assessed, not just recall of facts or

ability to perform computations accurately, there is no one obvious way of designing questions and marking schemes which would provide meaningful feedback to both students and teachers.

As a result, the major maths e-assessment systems in use are largely incompatible. For example, Mathletics[13] uses rich, randomly-generated question statements on top of multiple-choice answers, with distractors representing common errors, to provide useful and detailed feedback to the student while compromising on the kinds of questions which can be asked. STACK uses decision trees together with trapping student input in order to generate feedback error messages and advice. On the other hand, in DEWIS every question is a Perl script, which allows very complicated question design and marking schemes but sets the barrier to entry for question authors very high.

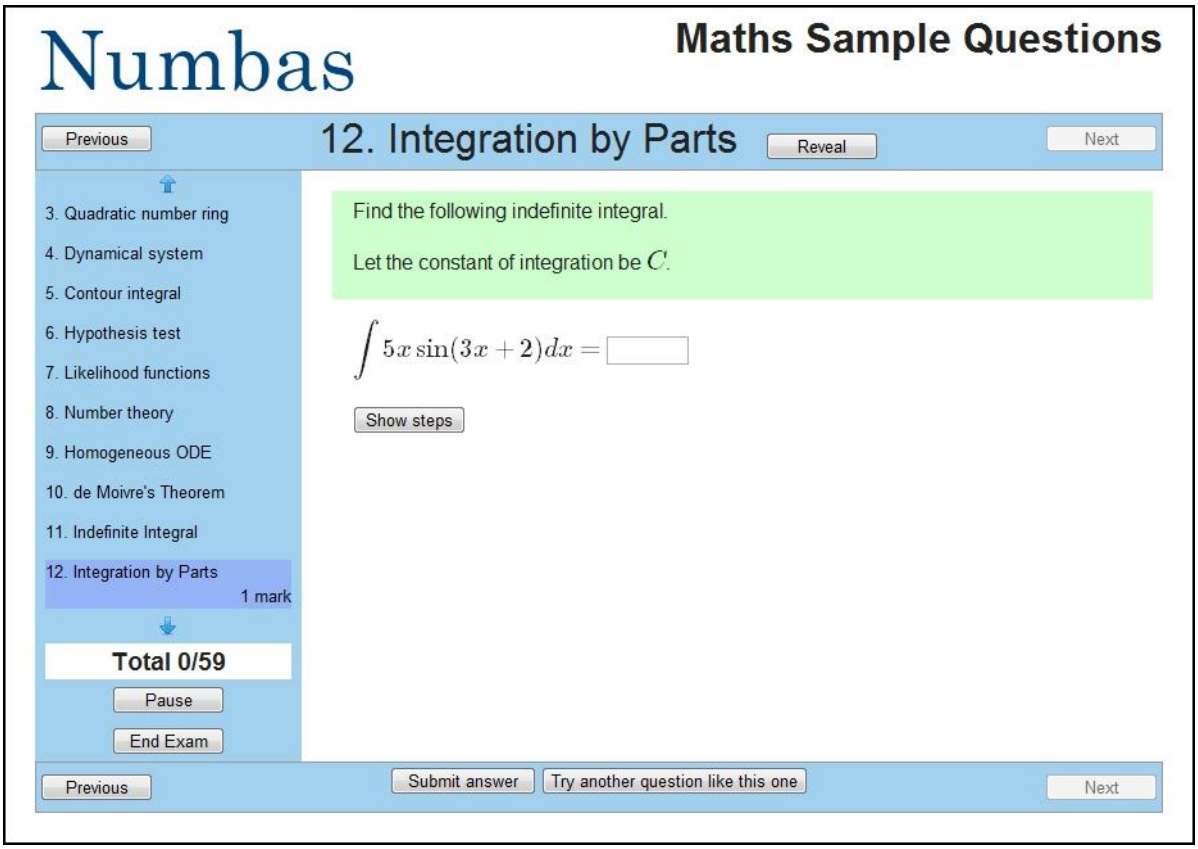
Considering this, we felt that the well-used and proven CALM-style design made the optimal compromise between difficulty of authoring and robustness of design. Since no CALM-style systems matching the above criteria were available to us, we decided to start from scratch on our own system.

Design Decisions and SCORM Standards

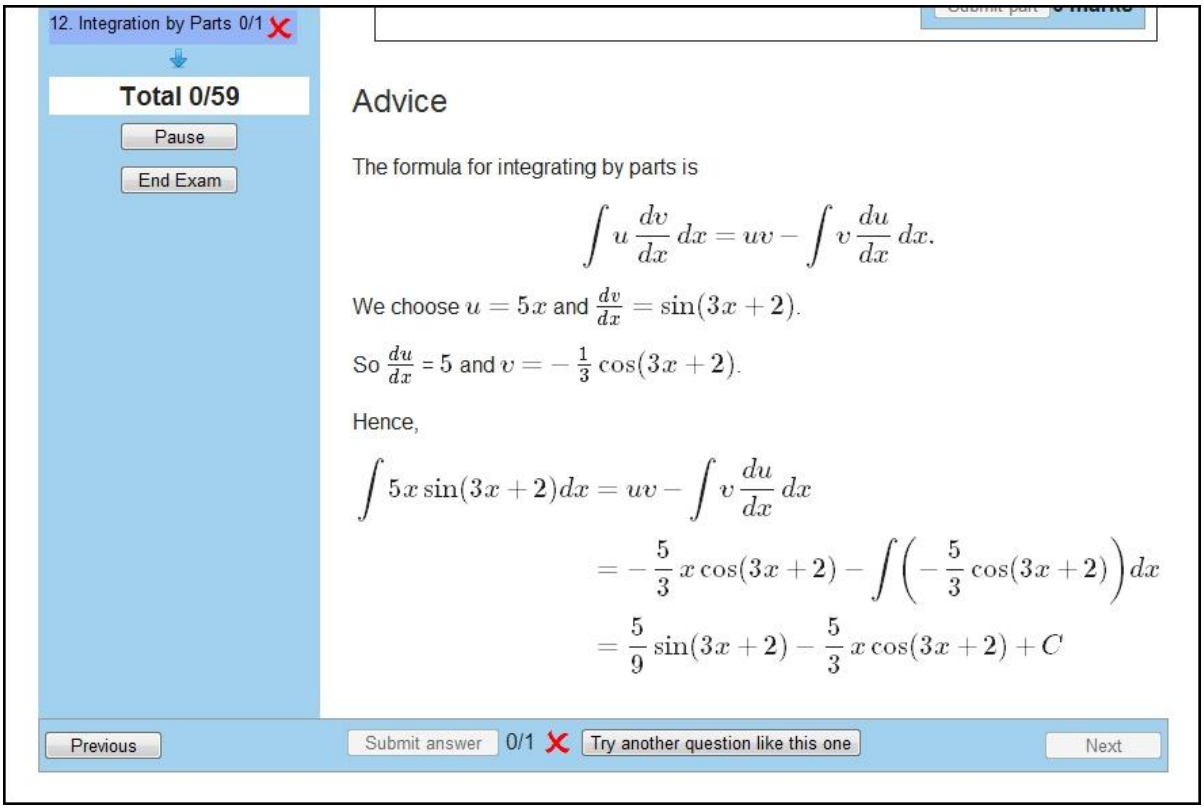
The CALM-Like Design

We have briefly rehearsed above the reasons why we chose the CALM -like design and functionality. For example, in CALM's Judged Mathematical Expression (JME) question type all mathematical input by the user is evaluated over a range and compared to the author's solution. This we have retained as it has been shown to be very robust and works well. However, we have extended JME into a unique pattern-matching expression simplification system which is entirely controlled by the examination or question author, and can be extended with new rules very easily. This is particularly useful in authoring partial solutions, hints or full solutions. We have also added more mathematical functionality, in particular support for complex numbers, matrices and lists of data.

Numbas question-design is as in CALM and i-assess, comprising Parts, Steps and Advice. The question types include all the standard question types and have been enhanced.



This question is in formative mode, with one Part. The Reveal button at the top gives a full and detailed answer. There is a “Show steps” button which may provide a hint or may break the question down into smaller sub-questions. Note that there is an extra button at the bottom; this is to give the user another question of this format. All questions are randomised, so this gives the opportunity for as much practice as wanted.



On pressing Reveal the user is shown a full worked solution.

Note that this solution is displayed in LaTeX, with JME used to evaluate randomised expressions within the LaTeX and simplification rules used to present the randomised expressions as an examination or question author would.

Added Mathematical Functionality.

As well as being able to write their own extensions to the system (a major improvement on i-assess; for example the Statistics extension as described in the Numbas manual, section 10, gives authors access to standard statistical distributions and functionality not available in i-assess), we included the ability for authors to write their own functions, using a simple programming language (full details in the Numbas manual, section 5.4).

The following is a screen-shot of a solution to a question which has a variable number of steps in the algorithm used to calculate the greatest common divisor of two integers, and does not restrict the author to preparing a question with a fixed number of steps. This was generated by a function defined by the author and gives the appropriate detailed feedback. This was not possible with i-assess as there was no facility to write such functions. Such functionality greatly increases the scope and sophistication of the questions that can be authored.

Advice

a)

On applying the standard method for finding the $\gcd$ of two numbers we have the following sequence:

Divide 62491 by 8151 with remainder 5434: $62491 = 8151 \times 7 + 5434,$

Divide 8151 by 5434 with remainder 2717: $8151 = 5434 \times 1 + 2717,$

Finally, 2717 divides into 5434 exactly 2 times: $5434 = 2717 \times 2.$

The last non-zero remainder is 2717, so this is the $\gcd$ of 62491 and 8151.

Now work backwards through those steps, rearranging them to find the remainders as linear combinations of the other numbers.

When you reach the last line you will have found s and t such that

$$62491s + 8151t = 2717$$

1. $2717 = 8151 \times 1 - 5434 \times 1$
2. $2717 = -62491 \times 1 + 8151 \times 8$

So $s = -1$ and $t = 8$.

SCORM

The aim to provide assessments to schools and colleges through outreach was another important motivating factor for replacing i-assess, so we decided very early-on that we should make use of the SCORM [14] standard to create self-contained assessments which could be integrated with whatever learning management system our partners used.

We made a very clear decision early-on to conform to relevant standards as much as possible. By sticking to widely-accepted web standards, the resulting examinations would be accessible to as many potential users as possible for as long as possible. Using proprietary plug-ins, libraries and technologies would restrict the ways in which examinations could be accessed. In the long term, systems using non-standard technology become unusable as the required software inevitably becomes less commonly available or compatible.

Since SCORM objects are required to be self-contained and run entirely in the browser, the other criteria were immediately satisfied: rich content, especially videos, can very easily be included in HTML; a lot of attention has been paid to producing clean and intuitive interfaces in web pages; most importantly, every device on which one might wish to take a mathematics test has a browser, meaning assessments gain cross-platform compatibility as a matter of course.

Conversely, the restrictions of SCORM meant that we could not rely on external servers to perform mathematical calculations, so it was necessary to create, from scratch, a computer algebra system entirely in Javascript. This system is, as far as we know, unique to Numbas.

Implementation

Numbas is entirely written in Javascript, along with some small Python scripts which compile examination packages for distribution. The system is modular and easily extensible; for example we have quickly written our own extensions to perform statistical tests and to plot interactive graphs.

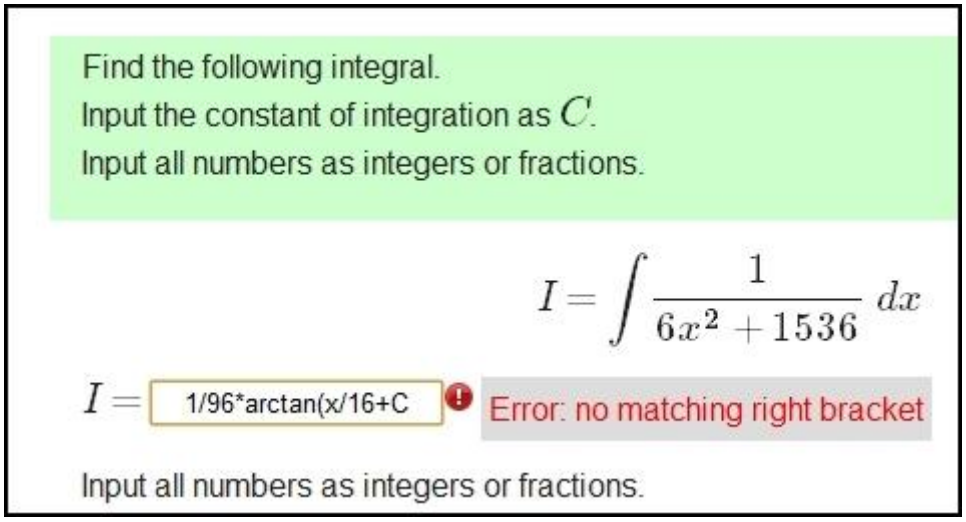
Mathematics Display and Input.

We have made use of several excellent new browser-based technologies which have recently appeared. In particular, MathJax has solved the problem of displaying mathematical notation on the web. It doesn't require the user to install anything and runs on any browser.

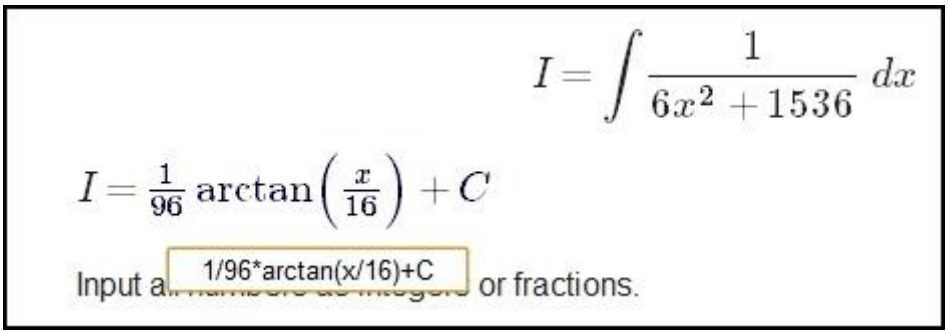
A huge benefit over other browser-based mathematics renderers is that LaTeX input is accepted, meaning mathematics can be quickly written by hand instead of relying on a graphical equation editor or MathML, which is effectively not human-readable.

By combining MathJax with the Javascript symbolic algebra system we were able to create a mathematical input method which takes the student's linear keyboard input and instantly displays their expression neatly typeset and checked for syntax errors. Using this entry method, students can be sure that the computer has interpreted their answer the way they intended before submitting it for marking.

The following example shows the display of the expression as the user inputs it together with a warning that a bracket is needed.



On completing the expression the following is displayed.



The formula becomes syntactically correct once the bracket is included by the user, and is typeset using mathematical notation so that the student can check that it has been interpreted as they intended. This method of input using "algebraic calculator syntax" with immediate mathematical display of the input, without the additional syntax analysis given by Numbas, has been used at Newcastle University since 2006 using the i-assess system and is completely accepted by all users. An initial questionnaire survey of all mathematics and statistics students in 2007 on the use of i-assess confirmed this acceptance. Subsequently, although never directly surveyed in questionnaires, there have been no complaints on this method of input in student feedback (from all disciplines) or from student-staff committees. The dynamic display of mathematical input was one of the requests made by students at an early stage of the implementation of e-assessment at Newcastle in 2006.

User Interface

Other Javascript libraries allowed the rapid production of an attractive, clean, intuitive user-interface. Our present design has emphasised this uncluttered interface. The stylesheets, resources and Javascript code used to produce the display are separated from the rest of the system, so it is possible to completely change the look and feel of examinations by creating a new 'theme'.

Authoring

Authors use a very simple structured data format, similar to JSON, to create exams. All content displayed to the user is written as simple HTML or Textile, with LaTeX used for mathematical notation. See the manual on the Numbas web-site for more information on this mark-up script. However, we are also developing a "higher-level" browser-based editor which can be used instead. This is to be used and tested in workshops in early 2012.

SCORM Implementation

The SCORM standard was not too hard to implement, using the excellent documentation available from www.scorm.com. Under SCORM, an LMS (Learning Management System) controls access to the Numbas exam and takes care of user management, session tracking and data storage and reporting, so we saved a lot of time not implementing these things. Unlike other interoperability standards, SCORM makes very few assumptions about the design and marking of questions, leaving us free to implement question types and marking algorithms to suit our needs.

Many of the early design decisions had unexpected benefits. The decision to use SCORM kept us honest about not relying on external services. Calling a backend server for any calculations would mean that anybody else using Numbas would also need to set up their own server. Many users are either unwilling or unable to do this. Additionally, running everything on the client means that sessions with large numbers of simultaneous users don't cause any problems through network over-use.

Implementing in a browser

An immense advantage of implementing Numbas as a browser-based system is the large and growing range of high-quality material and libraries which can be included in browser-based apps. We have already made use of several such resources, including using Vimeo to embed videos of lecturers explaining mathematical methods, and open-source graphing libraries such as JSXGraph and flot for generating interactive graphics to accompany question statements. Of course it is important to choose those applications which are well supported and sustainable, but this supply will only be increasing in quantity and we expect quality as browser-based applications become the norm.

Numbas in action.

For the academic year 2010-2011 a DVD was produced containing several Numbas examinations preparing first-year maths students for a diagnostic test administered in the first week. We received good feedback from the students and no reports of problems with use.

In October 2011, Numbas was used to create a diagnostic delivered to first-year biomedics over Blackboard. The test involved 200+ simultaneous users, scaled fine, no problems with use, test scores up on last year.

Numbas was used to create revision resources [15] for Maths-Aid, the Mathematics Support Centre at Newcastle University. This was created and up on the web site and available to students the same day we decided to create it.

A set of slides created to accompany a talk at e-Assessment Scotland 2011, with an accompanying Numbas exam is available at http://www.e-assessment-scotland.org/numbas/index.html [16].

Future Plans and Developments

Numbas has built upon the basic functionality and proven CALM- like design of i-assess. This has been thoroughly used and tested since 2006. The extra functionality and capability that Numbas provides is now to be fully exploited and tested in the implementation at Newcastle University when it replaces i-assess in the academic year 2012-2013. Use of Numbas is also planned to be extended to the engineering schools at Newcastle: it has been trialled in Electronic Engineering.

Two other universities, Bradford and Kingston, are adopting Numbas for use in 2012-2013 as part of the HE STEM Practice Transfer projects.

We will be collecting usage data and student feedback in the year 2012-2013 from Newcastle and all other implementations elsewhere to inform our developments of Numbas.

We are running internal workshops on using Numbas for staff at Newcastle and an external workshop *Using Numbas for Open Educational Resources* funded by HEA-JISC in April 2012.

Of particular interest in these workshops is the feedback obtained from the demonstration and use of a recently developed question and examination editor which can be used instead of the mark-up script for writing questions and examinations. This feedback will be used to further inform the development of the editor so that it can best serve the HE community.

Summary

We have shown that it is possible to create a feature-packed, robust, easy-to-use, reliable mathematics and statistics e-assessment entirely in the browser, based upon proven design principles.

Given flexibility of use, together with the ever increasing availability of on-line resource, the system can be extended easily and efficiently to other disciplines, especially in Science, Technology and Engineering as well as numerate disciplines such as Economics, Business Studies and Psychology.

The system is ideally suited for challenging longitudinal formative e-assessment as part of a course, module or as an e-learning package which can be deployed anywhere.

Numbas is extremely easy to implement, deploy and use.

Numbas is designed to be as flexible and as transparent as possible and is ideal for collaborative development.

Collaborators can be involved in many ways, including:

- The technical development of Numbas, possibly to suit their own environments or to look at issues such as interoperability.
- Informing the development and extensions of Numbas through pedagogic advice and feedback.
- By implementing the systems and reporting on their experiences and requirements.
- The sharing of resources, including question databanks

References

1. http://www.apache.org/licenses/LICENSE-2.0.html (Apache 2.0 licence)
2. http://www.goalonline.co.uk/Web/EDI/assessment_solutions/i-assess.aspx (i-assess)
3. http://www.mathstore.ac.uk/headocs/Foster_B.pdf (*Using computer based assessment in first year mathematics and statistics degree courses at Newcastle University*, Foster 2007)
4. http://www.mas.ncl.ac.uk/~nlf8/Publications/CBApaper.pdf (*Using computer based assessments in a large statistics service course*, Fawcett, Foster, Youd 2008)
5. https://lectopia.ncl.ac.uk/lectopia/casterframe.lasso?fid=49966&cnt=true&usr=not-indicated&name=not-indicated (Presentation to Newcastle QuILT: *Formative eAssessment using i-assess*. 9 minutes, 15 seconds in).
6. http://www.calm.hw.ac.uk/ (CALM)
7. http://scholar.hw.ac.uk/ (SCHOLAR)
8. http://tltt.strath.ac.uk/REAP/public/Resources/DN_SHE_Final.pdf (*Formative assessment and self-regulated learning: A model and seven principles of good feedback practice*, Nicol and McFarlane-Dick 2006)
9. http://mathstore.ac.uk/headocs/WhoUsesSTACK.pdf (*Who uses STACK?*, Sangwin 2010)
10. http://www.maplesoft.com/products/mapleta/ (Maple TA)
11. http://www.cems.uwe.ac.uk/caa/welcome/index.html (DEWIS)
12. http://www.jisc.ac.uk/whatwedo/projects/mathssasses.aspx (MathAssess)
13. http://www.mathstore.ac.uk/headocs/8307_greenhow_m_mathletics.pdf (*Mathletics – a suite of computer-assisted assessments*, Greenhow 2008)
14. http://www.adlnet.gov/capabilities/scorm (SCORM)
15. http://www.ncl.ac.uk/students/mathsaid/resources/academic/ode_homogeneous.htm (Maths-Aid revision resources with accompanying Numbas examination)
16. http://www.e-assessment-scotland.org/numbas/index.html (Presentation on Numbas by Christian Perfect, e-Assessment Scotland 2011)
17. http://www.hefce.ac.uk/learning/nss/ (National Student Survey)