Maths-CAA Series
March 2004

CAA in Context: A Case Study

Contents

Introduction 2

Features and Issues of CAA (AiM) 3

Some key features of AiM 3
Randomisation of questions 5
Feedback 6
And also... 7

CAA: Early Experiences 8

Support Module 8

Philosophy 9
Teaching activities 10
Learning activities 11
Evaluation 13

Future Developments 13

Conclusions 14

References 14

CAA in Context: A Case Study

Dirk F. M. Hermans
School of Mathematics and Statistics, The University of Birmingham
Edgbaston, Birmingham. B15 2TT.
Email: d.f.m.hermans@bham.ac.uk

Abstract: An introduction into AiM will focus on the benefits the coupling of a Computer Aided Assessment (CAA) package with a Computer Algebra System (CAS) generates. The type of questions possible in AiM will be discussed to illustrate how systems like AiM will empower the educator to go beyond what can be achieved with most standard CAA packages. Case studies are presented of the use of AiM in first year core mathematics courses at the University of Birmingham, in particular its use in an unusual type of module aimed at supporting weaker students. The paper will list in detail how CAA was used to complement traditional approaches and how different CAA elements were constructed to serve different purposes. This should illustrate that the use of CAA should not serve its own purpose but should reflect the desire of the teacher to use all tools creatively to provide the desired learning environment.

1. Introduction

The on-line assessment of mathematics has suffered for a long time from the restrictions imposed by standard computer aided assessment packages. Besides the more general issue of displaying mathematical expressions on-line, the inability to input mathematical expressions and check the equivalence of two mathematical expressions has been a major hurdle. In recent years, several software packages have become available that combine a computer algebra package with computer aided assessment. These new packages not only lift many of the restrictions mentioned above, but also provide many new opportunities for on-line assessment of mathematics.

AiM is an example of a nearly-free package where computer aided assessment is combined with the computer algebra package Maple. This not only allows for the entry of mathematical expressions using Maple syntax, but also addresses the problem of evaluating equivalent answers. In addition, many more tactics for evaluating answers to questions have been generated, as well as opportunities to provide tailored feedback. Randomisation of parameters allows for generic questions, but care needs to be taken that such questions continue to assess the intended learning outcomes of the task.

Software like AiM also opens up a whole new area of educational research in how higher level skills can be assessed on-line (see, e.g., [1], [2]). For example, open-ended "example" questions take students away from the routine execution of sets of rules and can provide valuable data to feed into discussions. New opportunities now arise to integrate on-line assessment of mathematics into a virtual learning environment, increasing the quality of e-learning opportunities in mathematics.

This paper contains elements already published in [12].

2. Features and Issues of CAA (AiM)

2.1 Some key features of AiM

AiM was developed at Ghent University, Belgium (AiM-plain: [3], [4]) and has been modified at the University of Sheffield, U.K. (AiM-tth: [5]). It is currently being taken forward by an international team of developers as a freely available system [6]. It links the Computer Algebra System (CAS) Maple [7] with a web based Computer Aided Assessment (CAA) system. This link allows a number of new features, i.e.:

- free text input, using Maple syntax;
- testing equivalence of different mathematical expressions;
- evaluation of answers;
- marking questions with multiple correct answers;
- randomisation of parameters in questions;
- answer-related feedback;
- random selection of questions from a question bank;
- partial marks for correct features in the answer;

The link with Maple allows the input of mathematical expressions as text, using the Maple syntax, e.g. "`cos(3*x+8)`", "`exp(-x)*(x^2-5)`". This eliminates the need for selecting answers of a mathematical nature using multiple choice questions. So, assessment strategies like elimination of possible answers or guessing answers can be avoided completely. However, students using AiM will need some introduction in the Maple syntax. For undergraduate students in a mathematical discipline this is not a problem, but it may be a more significant drawback for students at schools or in a discipline where access to a CAS like Maple is not normally provided.

However, AiM allows students to "Validate" their answers, so that they are given the opportunity to resolve any syntax errors before the answer is marked. This is an essential feature, as students are bound to make entry mistakes. In

addition, AiM has a parser that will make helpful suggestions when the syntax cannot be fully understood by Maple, e.g., it will indicate possible unclosed brackets. Typically, students will not be penalised for detectable entry mistakes.

Mathematically equivalent but different looking answers can be evaluated. AiM does allow the manipulation of student answers, hence, there is no need to compare them with a model solution. An answer can be subjected to a series of tests. A solution to a differential equation with given boundary conditions can be substituted in the differential equation to verify it is a solution and then the boundary conditions can be substituted in the answer to verify that they are satisfied as well. If one requires a vector normal to another vector, the inner product between the answer and the given vector can be checked. This allows for questions with multiple correct answers, or questions where the answers need only to satisfy certain properties, see an example in Figure 1. Maple functions can be used to check properties, so one can use the Maple command "`isprime()`" to verify that the student has indeed given you a prime number when requested.

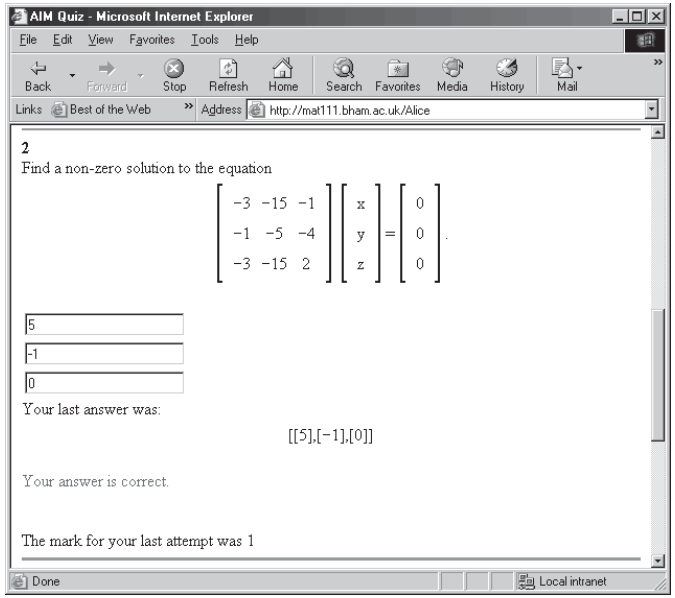

Figure 1: AiM Question involving matrix input and multiple possible solutions

The analysis of the answer can also be used to enforce good practice. An indefinite integral is only known up to an indeterminate constant. In AiM, one can check the answer for the presence of this constant and deduct marks if necessary, see e.g. Figure 2.

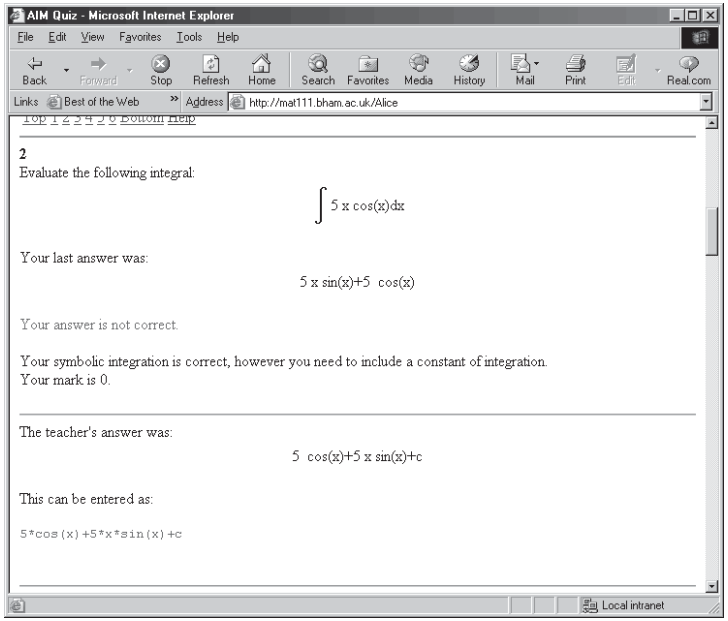

Figure 2: AiM question on integration with system response and mark

2.2 Randomisation of questions

AiM allows for the use of random variables in questions, which in turn allows for setting generic questions. A significant advantage of this is that the same test may feature slightly different question versions for individual students. This not only eliminates cheating (by looking at another screen) in exam situations, it also addresses the problem of students copying work in quizzes that are running over a period of time. Instead of asking a friend, "What is the answer to question 3?", they now have to ask, "How did you get your answer to question 3?", which steers the interaction among peers in an educationally far more positive format.

Random parameters, however, need to be used in a considerate way. Most problems with AiM result from a randomisation that renders a problem impossible to solve. Also, it is necessary to reflect whether different parameter values may lead to drastically different types of solutions. In such cases, controlled randomisation may keep the problem in the chosen class. For example, randomisation of a matrix of coefficients in a two by two system of linear equations will mix up the cases where there is a unique solution, an infinite set of solutions or no solution at all. It may be necessary to restrict the randomisation in such a way that the learning objectives for the question are adhered to. At the lowest level of testing, each individual possibility may need to be tested separately. Higher level practice may well find the more general randomisation acceptable. In other words, randomisation may lead to different classifications under a particular taxonomy.

When using randomisation, it is important to realise the context in which the tests take place. First year students in mathematics at Birmingham are banned from using pocket calculators at examinations. Hence, random parameters used should be such that questions can continue to be answered by mental or brief manual calculations.

2.3 Feedback

Feedback is a crucial ingredient in any closed learning loop. In AiM, feedback can be controlled in various ways. At the question level, the system will automatically provide a correct/incorrect response. However, the question author is able to provide more information following a detailed analysis of the answer. For example, where a normalised eigenvector of a matrix is required, and the answer is an eigenvector, but not a normalised one, the author can make sure the student is told this. In addition, the question author can detail a model solution to the problem, shown only when correct answers are revealed. This can be automated to an extent. For example, Strickland [5] has provided some sample questions involving integration by parts, where students can be shown how a successful integration by parts can be performed on a given problem (see also Figure 3). These question source files only need to be edited at one line for all these features to apply to the new question.

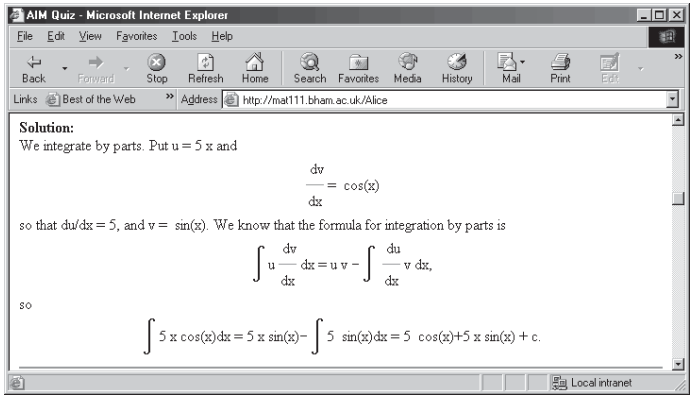

Figure 3: Detailed solution provided by AiM

AiM also controls the feedback given in a test at the managerial level. Various modes can be used to run a test. Examinations can be conducted in "Strict" mode, where the student is only told that the answer is syntactically correct. The student will only know which questions are correctly done or be able to see other feedback on the questions after a set date. Assessments performed over a period of time use the "Assessment" mode. Students will be told if their answer is right or wrong, and may be given another attempt at a prescribed penalty. They may be given other feedback, but would not see the solution or how it has been derived until after a set date. Self-assessment during revision can be set-up using the "Practice" mode, where students receive full feedback when they want and can try new versions of the questions/quizzes at their own pace. The same question can be used in all three modes.

2.4 And also...

AiM also allows the random selection of questions in a quiz, such that different students are given different questions. This, of course, demands some classification of the question bank in terms of keywords, duration, level of difficulty or learning outcome. Various classification schemes have been studied for use in mathematics, starting with Bloom's taxonomy [8,9] and further developed by many others. A good introduction to these theories can be found in [10]. The issue of classifying mathematical questions is crucial for the development of shared question banks and well balanced tests. The opening up of hitherto impossible question types will require further research in this area.

The analysis of the answers using Maple also allows for the award of partial marks. CAA is often disliked by students because there are no "method" marks. This will always be true to some extent. But in some cases, the achievement of a correct final solution may indeed be your learning objective, testing persistence, accuracy in the development and knowledge. However, AiM does allow the award of marks for solutions that are partially correct, and can be recognised as such. Also, the ability to allow several attempts at a given question, with pre-imposed penalties, allows the students to recover from mistakes. Indeed, it may be a much better educational experience to develop strategies for determining the error in their own work, rather then rely on a comparison with a teacher provided model solution.

A type of question made possible by software like AiM is the "Create Example" type. It is a type of question often avoided on written tests or examinations because of the heavy toll on marking it may produce. A possibly vast array of answers needs to be tested against the set criteria. Such marking is easily automated in AiM, and moreover, AiM will provide a list of all answers given for further discussion/analysis. It has been argued [1] that such questions encourage learning at a higher level. It is clear that a different strategy is needed to create an example, e.g. of a quadratic function with a maximum or minimum at x=1, than to verify whether or not a given quadratic function satisfies this property. This is just one example of how software like AiM may engender new research into assessment and learning.

3. CAA: Early Experiences

At The University of Birmingham, AiM has been used in various courses over the past few years. These courses span levels 0 (or F) up to level 3 (H) and were typically mixed with other types of assessment. Most of these courses involved small class sizes. More recently, one experience helped to shape our current provision.

All students taking the core module in mathematical techniques were given access to a large set of AiM practice questions, mirroring questions set on paper exercise sheets. They were then given eight summary tests, introduced at a supervised computer laboratory, with a deadline about a fortnight after the introduction. Students were given three attempts at each question, with a penalty applied for each incorrect attempt. The marks for these tests would count only if they improved the students overall grade on the module.

57% of the 188 students attempted all eight tests and of this group, only three failed the core course. Nearly 80% of all students tried at least 6 tests. 14% of

students attempted less then half of the tests and of these, more then 50% failed the module. The high participation rate is indicative of the view among students that this type of practice does help. The problem, as usual, is that those who need it most often do not use it. It was also possible to glimpse login habits, with several students accessing the system late at night from home. This provides a justification for choosing a web-based system that allows the students access whenever and wherever it suits them.

4. Support Module

Partly encouraged by the widening participation agenda, possible support mechanisms for students with a weaker qualification in Mathematics were considered. Out of this emerged the support module in which a blended approach was used to address the many problems students may have with the first year core material. CAA was included by design in such a way as to complement the more traditional elements in student support. This module referred to the content delivered in a traditional way in the core modules and as such added no new content.

4.1 Philosophy

In designing the support of core mathematics the emphasis is on the acquisition of basic manipulative and calculus skills. Conceptual problems are tackled through exposure to many different examples so that an understanding, based on experience, is developed.

Since the mathematical topics concerned are taught in the core course, attended by all students in the support module, the focus is on the assimilation of that knowledge. Only where feedback from student work indicated a need were topics explained again in lectures. Most of the contact time is given to addressing solution strategies explicitly, possible pitfalls and useful associations between mathematical entities. Pervading all these face to face activities is the instigation, encouragement and appreciation of individual practice.

Students are expected to tackle a significant amount of problems, and will on the whole only engage with such amounts of work if it is assessed and if mechanisms are in place to provide quick and useful feedback. One established mechanism is the traditional paper based exercise feeding into small group supervisions, but this has a limited capacity, particularly in terms of staff time. Computer aided assessment provides an additional set-up, which can be used to complement the more traditional one. This combination allows exercises to be set that are not ideal for computer aided assessment. Together, set exercises

can cover a much wider parameter range, so students can build up their experience and understanding of different scenarios. In addition, mechanisms can be devised that ensure that students engage successfully with all of the mathematical topics in the curriculum.

Because of the emphasis on student activity, the module mainly assesses student effort rather then deep understanding. There is no final examination, as the mathematical content will be fully examined in the core module anyway. Since module marks in the first year do not contribute to any degree classification, there is no real issue in the awarding of "effort marks" in this way.

4.2 Teaching activities

The support module combined the following activities:

- Summary lectures, where common problems could be addressed and common queries raised. Issues with concepts or types of exercises, which emerged from the other activities, could be addressed here. Intended to be interactive, a typical lecture slot would move through three phases:
  - Phase 1: the lecturer addresses common problems or important issues that may have surfaced from an analysis of the other activities in the module.
  - Phase 2: students can field any questions that might be of interest to the other students, regarding mathematical issues encountered either in this module or in the core module.
  - Phase 3: students can leave, while others have a chance to ask questions individually.

  These lectures proved a useful focal point for the module as a whole and also allowed for an efficient processing of administrative matters.
- Supervisions, where Postgraduate Teaching Assistants (PGTAs) or Academic Staff would spend one hour each week with a small group of less then 6 students each. Students would have been set a number of exercises to do on paper, which would have been submitted in time for the supervisor to mark and annotate and list any issues that needed discussion at the supervision. In addition, students could introduce problems with other set tasks or with the content covered so far. The small group supervisions did allow for a more interactive setting, particularly for students who found it hard to contribute in a bigger class. Partnerships forged during such supervisions often persisted after the module had finished. Marking work by hand also allows assessment to focus on methodology rather then result, complementing computer aided assessment elsewhere in the module.

- Computer laboratories, where students could do computer-based work under the supervision of PGTAs or a member of the Academic Staff. During these sessions, students were introduced to both Maple and Matlab. Some introduction to Maple helped with the use of AiM, but also allowed the use of Maple-based worksheets to complement other learning resources. The introduction to Matlab was required as a pre-requisite for later modules. There was also time to allow students to engage with CAA (AiM) quizzes set while supervisors were present.

4.3 Learning activities

All these activities support and complement private practice, a key requirement aimed at developing mathematical techniques, gaining confidence in manipulating expressions and solving typical problems like integration and derivation. However much one might try to argue the case for private practice, assessments seem to be the most efficient driving force behind student's work. In this module, several forms of assessment were used to achieve the required effect:

- Class tests are useful in preparing students for tackling tasks in a timed environment. They require revision practice and also provide secure individual marks to complement other assessments where collaborations may have made an impact. Class tests are delivered in a traditional examination format as well as using AiM in a "strict mode" during a computer laboratory session, where students authentication can be done in a traditional way. Computer laboratory based examinations of this sort may raise questions with regard to the use of other software on the system (like Maple) or the use of electronic communications. The use of a web browser in full screen mode, combined with the presence of sufficient invigilators, made any attempt to engage with other software easily detectable even from a distance. However, for examinations that contribute more to the final mark, it might be possible to prepare computer clusters to make any such "cheating" technically impossible.
- Paper based exercises allow for the critical analysis of student notation and methodology. Method marks can easily be given and they link up quite naturally with the dominant assessment methods in other courses. By focussing on exercises of a more intermediate level of difficulty, they neatly complement CAA based assessment of basic techniques. These assessments, however, are most prone to plagiarism, since typically, the same exercises are set for all students and mere copying can be difficult to detect, unless where uncommon errors are made. One way to make copying less likely is to use AiM to set randomised questions, but require

complete worked out solutions to be handed in. This was not implemented, but may be considered in future.

- CAA using AiM allowed for individual homework to be set over a longer period. In one series of quizzes, students were given randomised questions of various levels of difficulty, with a given deadline, typically 2 weeks after the quiz was issued. Wrong attempts would be penalised by 1/3 of the total mark. This encouraged students to think carefully when issuing their first answer, but also to continue after the first failed attempt to investigate where they may have gone wrong. The penalty is set so high to discourage guessing. Students were able to print out the quiz and enter their answers at a later stage. Even though AiM will pick up and report on many syntax errors, students were advised to try similar questions in a bank of free practice questions first. Most, however, did not take up this option. If there was any student concern it was with the fact that the questions used marked only the result. However, it was explained to them that this was partly by design, to encourage careful and complete work and good practice in checking their answers before entering them into the system.
- A new element consisted of a series of AiM delivered quizzes of basic exercises where students were given an infinite number of attempts (a zero penalty was applied). However, the context was that they needed to supply the correct answer to every question in every quiz in this series in order to obtain any marks at all. The "win or loose" contribution to the overall module mark was set at 25%. The argument for such a type of assessment was that there should be no "opting out" of basic exercises across the curriculum. Past performances indicated that students often did not even attempt to engage with parts of the curriculum, assuming (correctly) that a good performance on the other parts would see them pass. Where a course is a pre-requisite for another, this approach to learning is quite clearly counter-productive. Initially, the deadlines for these quizzes was set at the final day of term, but this tended to lead to an almost incredible last minute rush. One student even managed to start his very first of ten quizzes in this series on the Thursday before the Friday 6:00pm deadline. In the current version of this module, shorter deadlines are set, with a "resit" option over the Easter revision period. Maybe, to my own surprise, students did accept this type of assessment as sensible. Because of the zero penalty, it is evident that students tend to answer by trial and error sometimes. AiM allows the teacher to look at all attempts made by students and it was clear that certain students needed far too many attempts. A more thorough analysis of how the number of attempts needed correlates with other scores for students would be interesting, as would be a more in depth analysis of the wrong answers supplied.

4.4 Evaluation

This support module was taken, in its pilot form, by 22 students (reduced to 14 in the second term) who were expected to face some difficulties. The main negative feedback on this course related to the selection of these students and the fact they felt "labelled" by being directed to this course. But even students who started this module with some resentment agreed that it had helped them cope with the first year course.

The support module as a whole was a significant success. Those students who engaged with the various aspects of the course also successfully completed the core course (which was the main purpose of the support). Students acknowledge the greater workload but admitted it worked to their advantage by giving them more practice at mathematical problems. The highly structured way of setting tests encouraged them into a consistent learning effort throughout the term, which benefited their overall attitude towards learning at this level. Students felt that the instant feedback provided by AiM was very helpful, but did not like the lack of marks for their workings when their final answer was wrong. Overall, this support module scored the best mark of all first year modules in the annual feedback provided by the students at the end of the year.

5. Future Developments

The successful use of AiM in the support module is an encouragement to exploit further possible applications in the learning environment. At Birmingham such developments are supported by the Learning Development Unit, which has just sponsored its second project involving AiM. The purpose of this project is to develop systems and procedures that allow the efficient use of learning technologies in mathematics. In practice, this means looking at integrating mathematical tools into WebCT, the Virtual Learning Environment currently supported.

One aspect of this project is to integrate the learning aids into a single environment. Students accessing pieces of notes should, from there, be able to start interactive demonstrations and relevant questions. This would complement the archives of practice questions and demonstrations that might still be useful if students like to go to a particular demonstration or quiz quickly. WebCT also allows the use of mathematical symbols in its discussion board.

AiM is one example of a CAA package with links to a computer algebra package. Another early example of such a package is Calmaeth [11], which uses Mathematica. Commercial developers have recently started to release prod-

ucts of this sort, notably Maple TA (see [7] also). AiM is currently linked to Maple software and hence restricted in its use by the Maple licences. At Birmingham, AiM quizzes have now been integrated successfully with WebCT, allowing for a single authentication.

6. Conclusions

The introduction of AiM has enabled the creation of a learning environment at Birmingham, which more efficiently supported weaker students. Computer Aided Assessment of this type has enormous potential, illustrated in various ways in this paper. Educators, wary of the pitfalls of multiple choice questions, now have an alternative tool that not only allows them to move a large amount of questions onto a web-based CAA system, but also allows for the efficient delivery of question types that resources would have otherwise ruled out.

However, a re-surfacing element is surely that systems like AiM are only a tool in the hands of the educator. Good judgement needs to be exercised when constructing questions or tests. One should never loose sight of the learning objectives intended and should give careful thought to how these tools integrate with other learning provision. CAA can be used in different ways in different contexts; the support module shows only a few examples. It is important to gain a better understanding of how and why CAA is used in certain contexts.

The educational implications of software like AiM are only just being recognised. Neither have the limits of the system, in terms of answer analysis and feedback or generation of more complex questions, been fully explored. Just like with Multiple Choice Questions, there is a whole new theory of CAA in Mathematics waiting to be developed.

References

[1] Sangwin, C. J. 2002. Encouraging higher level mathematical learning using computer aided assessment. *Proceedings of the International Congress of Mathematicians*, Satellite Conference on Mathematics education, Lhasa, Tibet.

[2] Hermans, Dirk F. M. 2002. Embedding Intelligent Web Based Assessment in a Mathematical Learning Environment. *International Symposium on Symbolic and Algebraic Computation (ISSAC2002)*, Internet Accessible Mathematical Computation Workshop, Lille.

[3] AiM-plain: http://allserv.rug.ac.be/~nvdbergh/aim_rug/docs/

[4] Klai, S., Kolokolnikov, T. and Van den Bergh, N. 2000. Using Maple and the Web to Grade Mathematics Tests. *International Workshop on Advanced Learning technologies*, Palmerston, New Zealand.

[5] Strickland, N. 2002. Alice Interactive Mathematics. *MSOR Connections*, **2**(1), 27-30.

[6] AiM-tth: http://aimmath.sourceforge.net/

[7] Maple, available from Waterloo Maple Software, http://www.maplesoft.com/

[8] Bloom, B. S. *et al.* 1956. *Taxonomy of Educational Objectives: the classification of educational goals. Volume 1*, London: Longman.

[9] Bloom, B. S. *et al.* 1964. *Taxonomy of Educational Objectives: the classification of educational goals. Volume 2*, London: Longman.

[10] Beevers, C. E. and Paterson, J. S. 2001. Automatic Assessment of problem Solving Skills in Mathematics. Available online at http://mathstore.ac.uk/articles/maths-caa-series/july2001/index.shtml.

[11] Calmaeth: http://calmaeth.maths.uwa.edu.au/

[12] Hermans, D. F. M. 2003. Intelligent on-line Assessment of Mathematics using AIM. *Proceedings of the 6th International Conference on Technology in Mathematics Teaching*, Volos, Greece.