Automatic assessment of problem solving skills in mathematics

by

C E Beevers (email: c.e.beevers@hw.ac.uk) and J S Paterson (email: j.s.paterson@hw.ac.uk)

Department of Mathematics, Heriot-Watt University

Abstract

Since 1985 the CALM Project for Computer Aided Learning in Mathematics at Heriot-Watt University in Edinburgh [1] has delivered automatic assessment to large groups of students in engineering and science.

Initially, CALM used the computer to assess basic mathematical skills/techniques and acquired facts/knowledge. CALM has been proud of its testing capability, providing as it does a comparison of numerical and mathematical expressions as answers. This is well beyond the multiple-choice format favoured in some subjects but not appropriate in much of Mathematics. However, much progress has been made through the design of intermediate steps, the inclusion of hints and the extension of question types. In the SCHOLAR Programme [2], delivered over the Web to 600 students in 44 Scottish schools during the last academic year, more complex mathematical skills have been tested automatically using the on-line assessment engine known as CUE. This paper addresses the advances in automatic assessment with particular emphasis on Mathematics. The capability of the computer to assess more general problem solving skills in Mathematics is explored. The paper concludes with some thoughts on how the subject might expand in the near future.

1. Introduction

Assessment in Mathematics and related subjects has traditionally been carried out using examination style question papers presented in a formal setting to the students enrolled on a particular course. Increasingly, computers are being introduced to undertake some of the testing. The advantages include saving in staff costs and time, flexibility for the student, fairness of assessment and reliability in testing. Initially, the main use of computers was in formative and diagnostic testing [3 - 5] but the huge developments being made in assessment systems indicate that there is a growing confidence in allowing practice summative and even summative assessment by computer [6, 7]. Computer testing ranges from straightforward multiple choice tests which may be submitted on-line for marking and returned via the WEB to more sophisticated assessment engines which offer a variety of question types, include interaction, an immediate marking system, help and results service (see http://www.i-assess.co.uk, [1] and [8] for more details).

Some of the main general issues arising are security, data protection and access to information, matching appropriate question types to particular learning outcomes and plagiarism. More specific to Mathematics, Science and Engineering are issues like the display of mathematical equations, the input of mathematical expressions into the computer and the checking of mathematical answers. Some of the restrictions of computer delivery include availability of a computer, finance implications using Internet access and communication between human and machine in scientific notation.
This paper describes, through the CALM and Mathwise Projects [9 - 12] the role the computer can play in the automatic assessment of basic skills. Following a review of Bloom’s taxonomy [13, 14], a number of educational theories are considered as they relate to Mathematics. Section 4 seeks to match theory to practice by suggesting how the computer can assess more complex subject specific skills with reference to the SCHOLAR Programme [2]. The final section looks ahead to future developments.

2. Some early progress

Assessment is the measurement of learning and testing is often seen as a means in itself. However, assessment is an integral part of learning and viewed as such the computer can play an important role in the education of the next generation of students. Despite some natural drawbacks cited above the advantages of speed and accuracy of response, flexibility and objectivity of marking and the growing range of question types make the computer central to the process of learning. In the early years of the CALM Project students were given a weekly test throughout a twenty-five week introductory course of Calculus. It had been important in CALM from the outset that more sophisticated testing was needed well beyond that offered by multiple choice Questioning. A style of question emerged in which the student could display knowledge and basic skill through the input of mathematical expressions as answers to specific prompts. However, the best students found that there were too many parts to be answered on the way to the correct final answer. So, through projects like Mathwise [9 - 12] the notion of key parts with optional intermediate steps appeared. In this way, the weaker students, who could not always answer the key parts of a question immediately, could take more steps on the way to the correct answer. In addition, the student was given the choice of mode of delivery ranging from Examination Mode with no visible marking on the screen to Help Mode in which students could see immediate ticks or crosses and had the choice of revealing an answer if stuck.

These educational features had emerged through a series of evaluations in the 1980s and 1990s [3, 6, 7] and were built into the package Interactive PastPapers [15]. IPP had many features including the question types known as ordered and unordered lists. The former enabled vector or co-ordinate answers in which order was important and the latter allowed questions that, for example, asked for the factors of a polynomial or the roots of an equation in which order was unimportant. Moreover, IPP (as Mathwise had done before) carried an Input Tool which graphically showed students how the computer was interpreting their one-line input of a mathematical expression. Then, just when delivery on a PC in authoring systems like Authorware had solved rendering issues for the display of mathematics, the Input Tool had resolved issues of mathematical entry and key parts with steps had gone some way to solving issues of partial credit along came the Internet!

Over the last few years CALM has collaborated closely with UCLES (University of Cambridge Local Examination Syndicate) and the commercial software company EQL of Livingston in West Lothian. The result of that collaboration has been an assessment system known as CUE to acknowledge the tri-partite work. In CUE there are the following question types:

- Multiple choice including multiple response;
- Judged mathematical expression as in the original CALM Project and an integral part of Mathwise and IPP;
- Word match which is most useful in testing basic knowledge in subjects like biology, Chemistry and Computing;
- Multiple hotspot questions which can be employed to test knowledge of maps, diagrams, tables or graphs;
• Essay that provides an expert answer for comparison when used as self-assessment; and,
• Hint to provide helpful feedback for wrong answers.

CUE also supports multimedia within the questions; a feature of particular value for automatic and web based delivery.

3. Theories of assessment

Bloom et al [13, 14] provides a good starting point for the categorisation of questions. For completeness the six levels are:

• Level 1: Knowledge
• Level 2: Comprehension
• Level 3: Application
• Level 4: Analysis
• Level 5: Synthesis and,
• Level 6: Evaluation.

However, his six levels of educational objectives give no more than a guideline and have frequently been the subject of discussion particularly when related to problem solving and skills assessment. Freeman and Lewis [16] who devote a chapter to assessing problem solving, suggest that Bloom’s taxonomy is not helpful in identifying which levels of learning are involved. They, however, give an alternative which divides into headings not too far removed from Bloom’s:

• Routines
• Diagnosis
• Strategy
• Interpretation
• Generation

See Freeman and Lewis [16, page 238]

It is not clear that they are totally in favour of this approach either as they immediately identify the problems with it. They conclude that Polya’s [17] approach is the most useful but it is interesting to note that subsequently under the heading of Methods for Assessing Problem Solving, Bloom’s objectives reappear when Freeman and Lewis admit that:

‘most problems of the calculate/solve/prove variety are tests of application or of analysis.’

Freeman and Lewis [16, page 243]

McBeath [18] suggest that all six levels can be tested using objective test questions and so automatic assessment would then be possible for problem solving. This may not be the case in every subject at all levels of attainment. This paper considers the validity of such an assertion at the school/university interface for students of Mathematics.

Objective tests can assess a wide range of learning and are very good at examining knowledge (recall of facts), comprehension (making sense of the facts) and application (using knowledge and understanding appropriately). Such questions can also be used where numerical responses, mathematical expressions or short text answers are required. As a simple example of all three answer types consider the question:
Question 1

Find the equation of the normal to the curve with equation \( f(x) = x^3 - 2x \) at the point where \( x = 1 \).

Here is one way in which the question could appear:

**KEY PART 1.1: What is the equation of the normal to the curve \( y = x^3 - 2x \) where \( x = 1 \)? \( y = ? \)**

*Answer: \(-x\)*

Note that in this question and any further examples, all answers which appear within stars (**)* would be hidden and available on pressing the reveal button.

Good students could answer this question in one key part but for those needing more support then the optional STEPS button is available. On pressing it the following steps to key part 1 would appear:

**KEY PART 1.1: What is the equation of the normal to the curve \( y = x^3 - 2x \) where \( x = 1 \)? \( y = ? \)**

**Step 1.1.1: What is the derivative of \( f \)? \( f'(x) = ? \)**

*Answer: \(3x^2 - 2\)*

**Step 1.1.2: In the equation of a straight line given as \( y - b = m(x - a) \), what is \( m \) called?**

*(Answer: gradient|slope)*

**Step 1.1.3: Complete this line:
\( m = -1/f'(x) = \)

*(Answer: -1)*

**Step 1.1.4: In the equation for this curve, \( y - b = m(x - a) \), give the value of \( b \) when \( a = 1 \)**

*(Answer: -1)*

*Answer: \(-x\)*

The above type of question has been employed successfully at Heriot-Watt in both formative and summative assessment. In the former the steps give the student who is unable to proceed, some helpful guidance towards the correct answer whereas in the latter case it provides some partial credit to the student who can make some progress towards the correct answer.

The above example also illustrates how numerical and mathematical expressions can be submitted as answers in CUE and that word match can have more than one correct response as in step 1.1.2. It is clear that knowledge of the process of differentiation, comprehension of the equation of a normal to a curve and the application of this to a particular example test the three basic levels of learning objectives articulated by Bloom et al [13, 14].

It is far from certain, though, that objective tests can assess learning beyond basic understanding. However, questions that are constructed imaginatively can challenge students and may test higher learning levels. For example, students can be presented with electronic tools to manipulate or construct objects on a screen (see the graph drawing package produced by McGuire in the Mathwise Calculus cluster of modules at http://www.bham.ac.uk/mathwise). It may also be possible in the preparation of complicated mathematical proofs to put statements onto the screen randomly and ask a student to re-arrange them in the correct order. Problem
solving skills may also be assessed with the right type of questions. Assessing higher learning levels with objective test questions is considered more closely in Section 4. For now let us consider some of the learning theories that might be relevant to the acquisition of problem solving skills.

In Mathematics and in particular in the automatic assessment of Mathematics, five theories are particularly relevant. The theories and their founders are shown.

1. Problem Solving Strategies: Polya;
2. Mathematical Problem Solving: Schoenfeld;
3. Information Processing: Miller;
4. Conditions of Learning: Gagne; and,
5. Operant Conditioning: Skinner

One of the simplest theories came from the mathematician Polya [17] who, in 1957, put forward a four-point approach to learning:

1. Understand the problem;
2. Devise a plan to solve it;
3. Carry out the plan; and,
4. Check the result

This theory relates well to the process of constructing computer-assisted assessments from an author’s point of view. Although authors of assessments are faced with problem posing tasks, they too have to deal with problem solving in constructing a meaningful test. Polya’s approach provides a good framework. Understanding is needed to identify what the questions should ask. The author then must devise a logical question set, write them and test them. The theory in effect demonstrates the required skills of the author in problem solving. It also offers a very sound approach to problem solving in general.

At this point it is worth considering whether problem solving, which nowadays is assumed to be a skill in its own right, is actually a collection of skills as defined by Bloom [13, 14]. For example the following may be said about Polya’s four-point strategy:

1. To identify the problem in the first place requires knowledge of the subject matter.
2. This leads to use of the higher skill of comprehension of the problem in context (and may require analysis skills to break down the problem in the first place).
3. A student can then apply these skills (knowledge, comprehension and perhaps analysis) to carry out the plan for solving the problem (application and synthesis).
4. A sensible student will then try to assess whether the solution is appropriate. (Unlike one of us (JSP) who, at a very young age, was ridiculed for an answer of 533/4 women in a national exam!) (evaluation).

Put in this way McBeath [18] may well be correct in believing that all skills can be tested by objective assessment. The strategy allows for the problem to be sectioned into considerably more manageable questions and introduces a stepped approach to solving the problem - a feature of some computer based assessment systems. Moving on, there is more than a passing resemblance of Polya’s theory in that of Schoenfeld’s Mathematical Problem Solving Theory [19]. He advocates that four categories of knowledge/skills are needed to be successful in mathematics. These are

- resources (procedural knowledge);
- heuristics (strategies and techniques);
Wilson et al [20] in their research on problem-solving recognise the contributions of both Polya [17] and Schoenfeld [19] in this field. They firmly believe that the most important part of these theories is self-reflection:

'It is what you learn after you have solved the problem that really counts.' Wilson et al [20] Schoenfeld [19] calls this 'metacognition' in his theory.

Thus domain knowledge -the basic facts and skills which constitute Bloom's lower order cognitive skills - is required to undertake any form of problem solving which may be seen to encompass all six of Bloom's levels. This is an important point that seems to be overlooked by many enthusiasts of problem-based learning but recognised by Schoenfeld.

Another learning theory that Schoenfeld draws upon is that developed by Miller in 1956. Miller called it 'Information Processing Theory' [22]. He believed that learning takes place in small chunks and that humans process information in a similar manner to computers. He argued that short-term memory could 'process' at most seven 'chunks' of information. This is an interesting theory and fits in well with computer-based assessment where lengthy questions are not common. Extended questions tend to be divided into parts which not only help with automatic marking but which, if Miller is to be believed, assists the student with the fundamental learning of a topic.

The psychologist Gagne in 1965 [23] thought of problem solving as a type of learning which involved thinking. It seems rather extreme to suggest this definition which infers that other skills do not involve 'thinking'. Gagne's ideas however also form a sound base for learning. He believes in assessment, feedback and positive reinforcement - a structure that most computer assessment applications would adopt. He put forward the following structure for teaching [24] consisting of nine separate events:

1. Gain attention;
2. Inform learners of the objective;
3. Recall prior learning;
4. Present the stimulus;
5. Provide guidance;
6. Elicit performance;
7. Provide feedback;
8. Assess performance; and,
9. Enhance retention.

These criteria form a good strategy on which to base any assessment. In keeping with Bloom's cognitive levels, Gagne also believed that skills need to be learned one at a time and that lower level skills must be mastered before higher level skills can be considered. Computer-aided assessment can achieve this through the availability of a set of carefully constructed tests which
allow help steps, reveal buttons and random parameters at a level appropriate to the learning
level of the student.

As mentioned earlier, the CALM project [1] at Heriot-Watt University used an early assessment
system of its own until the Mathwise assessment mechanism came along. Then, CALM
embarked on the development of Interactive Past Papers (IPP)(Beevers et al [15]) following the
award of the Bank of Scotland Prize to Higher Education in 1995. This system allows, in
addition to earlier features, partial credit to be given for answers that are correct but in the
wrong format. Thus positive reinforcement of the learning is implemented; the student feels
rewarded and learns quickly. This was one of the principles put forward by Skinner in his
Operant Conditioning Theory [25]. In keeping with another of Skinner’s principles the
introduction of steps in IPP promotes quicker positive learning by allowing the student, who is
unsure of the final answer, to choose the steps that give the question in a more structured
manner. Fundamental to Skinner’s theory is the concept of stimulus/response and that is one of
the strengths of automatic assessment. The students can receive help through the steps and
immediate relevant feedback at the moment when they need it most (on submitting the answer).
This not only promotes active learning but also a motivation within the student to continue.

4. Linking theory and practice

From the theories described in the last section it is appropriate to ask if it is possible to match
the assessment to learning outcomes: can the computer accurately assess what is needed? Are
there areas that cannot be assessed adequately?

At the school/university interface traditional examinations like ‘A’ Level in England and
Advanced Higher in Scotland have undergone changes recently to reflect a modular structure.
In Scotland, for example, minimum competence is being measured after forty hours of
classroom teaching. Passes in three such units of study then permit the candidates to take a
traditional end of course examination that sets out to assess extended competence. The
SCHOLAR programme [2] has been delivering via the web, learning resources aimed at the
Advanced Higher in Scotland. For the purposes of this paper the questions that address
extended competence will be considered as appropriate tests of problem solving skills i.e. the
ability to analyse, synthesise and evaluate with competence in the three lower order cognitive
skills of knowledge, comprehension and application. Three questions are chosen from this
year’s SQA Advanced Higher Mathematics paper [26] and set out below in a style capable of
computer delivery and automatic marking. It is this style of question that has been delivered by
the Mathematics team in the SCHOLAR programme. The three questions have been modified
slightly to accommodate automatic assessment but an attempt has been made to keep as close
as possible to the style in which the official paper was set.

Question 2

a) Obtain partial fractions for $\frac{x}{x^2 - 1}, x > 1$

b) Use the result of (a) to find $\int \frac{x^2}{x^2 - 1} \, dx, x > 1$

A possible framework in which this question can be set for delivery in automated formative
assessment follows:

Part a)
**KEY PART 2.1:** Express \( \frac{x}{x^2 - 1} \) in partial fractions in the form \( \frac{a}{x + c} + \frac{b}{x + d} \) where \( c < d \)

**STEP 2.1.1:** Factorise the denominator of \( \frac{x}{x^2 - 1} \)
*Answer: \((x - 1)(x + 1)\)*

**STEP 2.1.2:** If \( \frac{x}{x^2 - 1} = \frac{a}{x + c} + \frac{b}{x + d} \) where \( c < d \), then by finding the common denominator, state the right hand side of the equation which equates the numerators.
*Answer: \(a(x + 1) + b(x - 1)\)*

**STEP 2.1.3:** By equating coefficients of \( x \), state the expression equal to 1
*Answer: \(a + b\)*

**STEP 2.1.4:** By equating the constants, state the expression equal to 0
*Answer: \(a - b\)*

**STEP 2.1.5:** What is the value of \( a \)?
*Answer: \(\frac{1}{2}\)*

**STEP 2.1.6:** What is the value of \( b \)?
*Answer: \(\frac{1}{2}\)*

*Answer: \(\frac{1}{2(x - 1)} + \frac{1}{2(x + 1)}\)*

**Part b)**

**KEYPART 2.2:** Use the first result to find \( \int \frac{x^3}{x^2 - 1} \) dx, \( x > 1 \) quoting + C as the constant of integration.

**STEP 2.2.1:**
State the quotient when \( x^3 \) is divided by \( (x^2 - 1) \)
*Answer: \(x\)*

**STEP 2.2.2:** State \( \int x \) dx using +C as the constant of integration
*Answer: \(\frac{x^2}{2} + C\)*

*Answer: \(\frac{x^2}{2} + \log (x^2 - 1) + C\)*

This question 2 has made use of numerical and algebraic answers. Further steps could be set to assist the students in identifying that the remainder upon division is the answer to the first part of the question. Similarly it is possible to include a stepped approach to break down the integration of the first answer.
Question 3

A function \( f(x) \) is defined by

\[
\frac{x^2 + 6x + 12}{x + 2}, \quad x \neq -2
\]

a) Express \( f(x) \) in the form \( ax + b + \frac{b}{x + 2} \) stating the values of \( a \) and \( b \);

b) Write down an equation for each of the two asymptotes;

c) Show that \( f(x) \) has two stationary points;

d) Sketch the graph of \( f(x) \); and,

e) State the range of values of \( k \) such that the equation \( f(x) = k \) has no solution.

In this case the question already offers some steps in the first three parts towards the completion of the task to sketch the graph at d). An automated approach for formative assessment could take the following lines. Note that there is a slight change to the key parts in order to accommodate the two answers to part b) and part c).

### MAIN QUESTION TEXT

A function \( f(x) \) is defined by

\[
\frac{x^2 + 6x + 12}{x + 2}, \quad x \neq -2
\]

<table>
<thead>
<tr>
<th>KEY PART 3.1: Express ( f(x) ) in the form ( ax + b + \frac{b}{x + 2} ) stating the values of ( a ) and ( b )</th>
<th>STEP 3.1.1 Use long division to find ( f(x) ) in this form</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Answer: (1, 4)</em></td>
<td>REVEAL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>KEY PART 3.2: Write down an equation for the slant asymptote, ( y = ? )</th>
<th>STEP 3.2.1: The slant asymptote is given by the quotient upon division</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Answer: ( x + 4 )</em></td>
<td>REVEAL</td>
</tr>
</tbody>
</table>

| KEY PART 3.3: Write down an equation for the second asymptote beginning with the variable | STEP 3.3.1: Choose the left hand side of the equation of the second asymptote: 
- \( x = \)
- \( y = \) |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Answer: ( x = -2 )</em></td>
<td>REVEAL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>KEY PART 3.4: Show that ( f(x) ) has two stationary points and state the coordinates of the turning point which has the greatest ( x ) coordinate value in the form ( (x, y) ).</th>
<th>STEP 3.4.1: At turning points the derivative of ( f(x) ) is zero. What is the derivative of ( f(x) )?</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Answer: ( (0, 6) )</em></td>
<td>REVEAL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>KEY PART 3.5: Is this turning point a</th>
<th>STEP 3.5.1: The second derivative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
maximum or minimum?
determines the nature of the turning point. What is the second derivative?
*Answer: $8(x + 2)^3$ *

STEP 3.5.2: Use the values of $x$ at the turning points to find the value of the second derivative. When does a minimum occur:
- when $f''(x) = 0$
- when $f''(x) > 0$
- when $f''(x) < 0$
*Answer: when $f''(x) > 0$*

*Answer: minimum*

KEY PART 3.6: State the coordinates of the turning point with the smallest $x$-coordinate value in the form $(x, y)$.
STEP 3.6.1: This is one of the two values of $x$ for which $f'(x) = 0$
*Answer: $(-4, -2)$*

KEY PART 3.7: Is this turning point a maximum or a minimum?
STEP 3.7.1: Use the value of the second derivative at this value of $x$ to determine its nature.
*Answer: maximum*

KEY PART 3.8: Sketch the graph and hand it to your tutor.
*Answer:

KEY PART 3.9: Find the range of values of $k$ such that the equation $f(x) = k$ has no solution and state the minimum value.
STEP 3.9.1: Examine the graph.
*Answer: $-2$*

KEY PART 3.10: Find the range of values of $k$ such that the equation $f(x) = k$ has no solution and state the maximum value.
*Answer: $6$*

This question 3 offers a wider range of question types by including the following:
It would also have been possible to use hotspot questions. There are alternative methods for finding the range in keypart 3.9 and 3.10 such as using the discriminant of a particular quadratic. The keyparts and thus the steps could be written to guide the student to use this technique if it was required instead of using the graph. The graph shown in the answer would also be annotated accordingly in a computer test and is shown here only for reference in outline.

### Question 4

A chemical plant food loses effectiveness at a rate proportional to the amount present in the soil. The amount M grams of plant food effective after t days satisfies the differential equation

\[
\frac{dM}{dt} = kM, \quad \text{where } k \text{ is a constant.}
\]

a) Find the general solution for M in terms of t where the initial amount of plant food is M_0 grams.

b) Find the value of k if, after 30 days, only half the initial amount of plant food is effective.

c) What percentage of the original amount of plant food is effective after 35 days?

d) The plant food has to be renewed when its effectiveness falls below 25%. Is the manufacturer of the plant food justified in calling its product “sixty day super food”?  

#### KEY PART 4.1: In the general solution, to what does \( M/M_0 \) equate?

**Step 4.1.1:** \( \frac{dM}{M} = k \). With C a constant, to what does \( kt + C \) equate?

*Answer: \( \ln(M) \)*

**Step 4.1.2:** At \( t = 0 \), \( M = M_0 \). What is the value of \( C \)?

*Answer: \( \ln(M_0) \)*

**Step 4.1.3:** Hence, give the value of \( \ln(M/M_0) \)

*Answer: \( kt \)*

*Answer: \( e^{(kt)} \) REVEAL

#### KEY PART 4.2: Find the value of \( k \) if, after 30 days, only half the initial amount of plant food is effective.

**Step 4.2.1:** When \( M = M_0/2 \) then \( M_0/2 = M_0 e^{(30k)} \)

Rearrange and state the value of \( 30k \)

*Answer: \( -\ln(2) \)*

*Answer: \( -\ln(2) \)* REVEAL

#### KEY PART 4.3: What percentage of the original amount of plant food is effective after 35 days?

**Step 4.3.1:** When \( t = 35 \), \( M/M_0 = e^p \)

State the value of \( p \) as a fraction.

*Answer: \( \frac{-7 \ln(2)}{6} \)*

**Step 4.3.2:** Evaluate \( e^p \) to three d.p.
**KEY PART 4.4:** The plant food has to be renewed when its effectiveness falls below 25%. After how many days will this happen?

<table>
<thead>
<tr>
<th>Answer: 60*</th>
</tr>
</thead>
</table>

**Step 4.4.1:** What is the value of $M/M_0$?

| Answer: 0.25* |

**Step 4.4.2:** Use the equation found in the first key part to solve for $t$. In terms of $t$, to what does $\ln(0.25)$ equate?

| Answer: $\frac{-\ln(2)}{30}$ |

This question also naturally sub-divides into four key parts each of which can be stepped for the less confident student. The answers are either numerical or require a mathematical expression as an answer. The question as stated would not be appropriate for computer delivery as the last sentence expects a yes/no answer and the marks would be given to the method leading to the correct conclusion. However, a slight re-wording of the question would render it capable of computer exploitation without losing any of its impact. $M_0$ is used as one of the parameters in the key parts and steps to keep as close to the original question as possible. In reality, if this question was randomised for computer delivery, a different parameter would be used to make the subsequent automatic checking of the answer easier.

### 5. Looking Ahead

In the examples shown it is apparent that certain improvements would make automated assessment much easier. For example, ordered and unordered lists extend the capabilities of questioning on graph work and more complex word match may be a help in areas such as finding the equations of the asymptotes as in question 3.2 above. Gap fill is possible in the product I-assess and this allows for ordered and unordered lists so important in Mathematics. Drag and drop question types can also help but in some ways these offer the solution albeit in a semi-hidden manner. Essay marking is more complex although a marking system using norm referencing may be possible with a huge essay bank on which to base results. There are many problems in this area and to date, no progress on criteria referenced essay marking is apparent.

However, one obvious issue that perhaps is demonstrated with these questions is the structured way in which the student is directed through a complex series of tasks. This is commendable for formative assessment particularly at the lower cognitive skills levels where there is normally a standard method by which the question should be approached. Once the higher order cognitive skills are to be tested and incorporated in a problem solving context then the questions arise: Is this approach too structured? Does the structure allow for true testing of problem solving abilities? Does the emphasis on one approach suggest that an alternative approach is less well founded?… and so on. It is in this area that developments in the field of computer adaptive testing (C.A.T.) may be of most use.

Computer adaptive testing is a means by which the next question or part of a question in the test is determined by the answer to the current question or part in a question. Work is being carried out mainly in America and in particular in languages to try to extend the capabilities and the use of C.A.T. [27]
In Britain there have been some experiments that suggest that such a system is possible [5] but it has some way to go before it will be in general use. Automatic problem solving in Mathematics will certainly benefit from C.A.T. by allowing the students freedom to work through a complex question in a variety of ways, all equally correct.

This gives the students more control over their own learning and encourages them to learn actively and with imagination. Returning to question 3 above, it is clear to see that the order in which some of the questions set does not matter and that the student may well prefer a different order. It is also evident that key parts 9 and 10 have been answered from the information given on the graph which was sketched in answer to key part 8. However changing the order of the key parts would in fact strengthen this question if the main aim of key parts 9 and 10 was to test students’ knowledge of the discriminant. These are the types of issue that C.A.T. may be able to address.

The Scottish Centre for Research into On-Line Learning and Assessment (SCROLLA) started in April 2001 with a Research and Development grant of £560,000 from the Scottish Higher Education Funding Council. Three universities, Heriot-Watt, Glasgow and Edinburgh will carry out the research over three years in three main areas: on line assessment, networked learning and ICT policy. The purpose of the centre is to build an interdisciplinary research infrastructure in Scottish Higher Education in using ICT at all levels of education. At Heriot-Watt, the research will encompass some of the issues and developments highlighted in this paper.

References


