Accepted,
IJMEST 2003

New opportunities for encouraging higher level mathematical learning by creative use of emerging computer aided assessment

C J Sangwin
LTSN Maths, Stats & OR Network,
School of Mathematics and Statistics, University of Birmingham,
Birmingham, B15 2TT, UK
Telephone 0121 414 6197, Fax 0121 414 3389

Email: C.J.Sangwin@bham.ac.uk http://www.mat.bham.ac.uk/C.J.Sangwin/

Abstract: This article defines 'higher level mathematical skills' and details an important class: that of *constructing instances* of mathematical objects satisfying certain properties. Comment is made on the frequency of higher level tasks in undergraduate work. We explain how such questions may be assessed *in practice* without the imposition on staff of an onerous marking load. Included are examples which have been implemented on a *free* computer aided assessment system. Lastly we report an investigation of students' reactions to these questions and discuss their design and impact.

Keywords: high level mathematical skills, computer aided assessment, deep learning

1 Introduction

Firstly in this article we examine and define what is meant by 'higher level mathematical skills'. We argue that one particular type of mathematical task, that of *constructing instances* of mathematical objects satisfying certain properties, requires the use of these higher level skills. This class includes generating examples of mathematical objects. Secondly, by drawing on a recent analysis of current undergraduate work we examine and remark on how often, in practice, students are asked to perform tasks that require the use of higher level mathematical skills. Thirdly we explain how such questions may be assessed *in practice* without the imposition on staff of an onerous marking load by the judicious use of a *free* computer aided assessment system known as `AIM`. How to author questions is briefly addressed. Lastly we report an investigation of students' initial reactions to these questions during a group interview and discuss their design and impact. This discussion concludes with suggestions on how this approach may be incorporated into the overall learning cycle.

2 Higher level mathematical skills

A skill is defined literally as a *'practised ability, expertness, technique, craft, art, etc.'* Any attempt to elaborate on what is meant by mathematical skills must be based on an analysis

Group A	**Group B**	**Group C**
1. Recall factual knowledge	4. Information transfer	6. Justifying and interpreting
2. Comprehension	5. Application in new situations	7. Implications, conjectures and comparisons
3. Routine use of procedures		8. Evaluation

Table 1: Smith's [4] mathematical question taxonomy

of what *in reality* we ask the students to do. That is to say, the nature and purpose of the course are inherently influenced by the assessment criteria, which are in turn fulfilled by the successful demonstration of various skills. In contraposition, if we fail to ask the students to practice some technique, or to develop expertise in some craft – such as writing a mathematical proof – we can hardly complain afterwards that they have failed to develop this skill. If we genuinely covet other skills – use of computer technology, presentation, research and investigation, for example – our schemes of work, both formative and summative should reflect this. It is widely known that assessment drives what and how mathematics is learned and that students tend to adopt surface approaches to learning [1]. Furthermore, mathematics education research indicates that students who adopt a deeper approach to learning are more successful [2]. Possessing a vocabulary with which to describe tasks is a necessary precursor to developing an understanding of their potential effect in various situations. This allows more conscious choices to be made during course design [3, 1].

There have been various attempts to classify mathematical tasks, including [4] wherein mathematical examination questions are classified into the eight categories shown in Table 1. The three groups broadly show an increase in 'level' of the skills. Of course, all such classification schemes are fraught: no attempt is made to classify difficulty (Smith, rather tongue-in-cheek proposed *'easy' = questions I can do and 'difficult' = questions I can't do*). Similarly such classifications are entirely context dependent; demonstration of the skill *5. Application in new situations* by definition may only occur once in each mathematical context. Furthermore we might not be assessing what we intend to assess: a proof requiring only verbatim transcription or minor modification may only test factual recall, not true understanding.

Central to a recent evaluation of current first year undergraduate core mathematics material, in the light of prototype hand held computer algebra technology, was a classification of course work questions [5]. Independent of [4] the question classes shown in Table 2 were identified. Elaboration of these classes is given, with examples in Section 3. Section 4 gives details of the question type on which we concentrate effort in this paper: that of *constructing instances*. These two classification schemes show remarkable similarities, furthermore, when identical question sets were evaluated independently with the scheme shown in Table 2 the results showed a high degree of correlation, providing evidence of robustness in the procedure, despite its apparent subjectivity.

Given that we may classify the majority of course work tasks using one of the two schemes in Tables 1 or 2, one comes to the issue of defining 'higher level skills'. The classification scheme of [4] does this in the three groups A, B and C shown in Table 1. The TELRI project [6] differentiated between two forms of learning as follows.

1.	Factual recall
2.	Carry out a routine calculation or algorithm
3.	Classify some mathematical object
4.	Interpret situation or answer
5.	Prove, show, justify – (general argument)
6.	Extend a concept
7.	Construct example/instance
8.	Criticize a fallacy

Table 2: Alternative mathematical question taxonomy

They defined *adoptive* learning as an essentially reproductive process requiring the application of well-understood knowledge in bounded situations. Mathematically this may be typified by successful completion of tasks as laid out as 1–4 in Table 2. Proofs requiring verbatim transcription or minor alterations from a template are similarly adoptive in nature. This behaviour, they claim, would be typical of the *competent practitioner*.

They further defined *adaptive* learning as requiring higher cognitive processes such as creativity, reflection, criticism and so on. These attributes are typically tested by questions from classes 4–8 in Table 2. Such behaviour would be typical of the *expert*.

They introduce these terms to draw a distinction between deep adoptive and deep adaptive learning. The former learner may be adept at complex tasks requiring a number of separate skills, see Example question 1 below for such a mathematical task. The latter learner will be able to generate examples and proofs more independently showing insight, creativity and a higher degree of conceptualization.

We define the skills which typify such an adaptive approach as the *higher level skills*. This clearly does not divide tasks into disjoint classes – there is room for subjectivity. The creativity needed to genuinely construct an example could manifestly be replaced by the appropriate factual knowledge. That is inevitable and presents one particular challenge for assessment design which discriminates between surface and deep learners.

3 The question classification scheme

This section provides brief explanation of the classification scheme outlined in Table 2. Some illustrative examples of questions are also given, more information may be found in [5].

1. Factual recall. All questions require factual recall to some extent. However, on occasion questions may only require the recall, usually verbatim, of some knowledge.

2. Carry out a routine calculation or algorithm. Such questions may be complex and involved. They may require the recall of some facts, such as the details of the algorithm itself at university level. They certainly pre-suppose a certain level of algebraic fluency. Often such tasks may be performed by a computer algebra system.

Example question 1 *Solve the wave equation on the square two dimensional domain D subject to the boundary conditions...* [Separation of variables followed by Fourier series methods]

3. Classify some mathematical object. There are essentially two parts to these questions. Firstly to recall the definitions, which may be posed as a separate task in its own right. Secondly to perform any algebraic tasks or provide justification to show the specific object satisfies this. Often this is a question of the form 'Show that ... is a ...'. Furthermore, correctly applying a theorem, for example, tacitely assumes the case to which the theorem is being applied satisfies the hypotheses. This latter example is a case of object classification.

Example question 2 *Show that* $(\{3, 6, 9, 12\}, \times \bmod 15)$ *is a group, and identify the identity element.*

Clearly there will be some factual recall, and possibly some routine calculation. The essential distinction between factual recall, use of routine procedures, and classification of objects is the presence of a specific object. The distinction between the latter is typified by the two questions

> 'Find a solution to the differential equation ...'
> and
> 'Show that ... is a solution to the differential equation ...'

4. Interpret situation or answer. In this case a problem is posed with reference to a physical situation and requires modelling prior to the application of routine procedures. Typically the solution will require interpretation from the mathematical model to the terms in which the question was posed. The student must communicate their solution in the terms used by the question.

Example question 3 *[Kreyszig] Find the current in the RLC-circuit shown in Fig 60. when R=50 ohms, ..., I(0)=0, I'(0)=0.*

5. Prove, show, justify – (general argument). Such a question requires a general argument involving abstract or general objects rather than specific examples. In posing such questions phrases such as 'prove', 'show', 'with justification' are often used. Although used interchangeably, the solution to a 'prove' question typically requires Boolean logic, meta-logical structures (such as a contradiction, induction, exhaustive cases) and formal definitions. The solution to a 'show' question would, in general require an algebraic demonstration of a general but routine result.

Example question 4 *Show for all* $\mu \in \mathbb{R}$ *the matrix* $\begin{bmatrix} 1 & \mu \\ 0 & 1 \end{bmatrix}$ *has repeated eigenvalue* 1.

This is a general argument in the sense that the μ is an arbitrary parameter, but there is little in the way of logical structure in the simplest correct solution. Care needs to be exercised when 'with justification' is encountered as this may be code for 'we want to see your working' in an otherwise routine calculation.

Either direct factual recall and verbatim transcription of a proof, or adaptation of a known proof as template for the task in hand is not uncommon.

6. Extend a concept. Students are asked to evaluate previously acquired knowledge in a new situation. Such a classification is highly context dependent, temporary and thus fraught. There may be an overlap with other classifications, such as classification of some object. Nevertheless it is a useful distinction, particularly for the course designer.

Example question 5 *Show that the set of polynomials in x is a vector space over the real numbers. Show that the process of formal differentiation defines a linear transformation on this vector space.* [In the context of a level 1 vector spaces course in which a limited number of examples of finite dimensional vector spaces have been introduced.]

7. Construct example/instance. Questions of this type will occupy the remainder of this paper and are addressed separately in Section 4 below.

8. Criticize a fallacy. Finding mistakes in supposed proofs or criticizing reasoning does not fit within the scheme above. Such questions are posed, and classified separately. Such questions require higher level skills and a good understanding of the subject matter.

Example question 6 *The following* `FORTRAN` *code [omitted] will not compile. Explain why not and suggest corrections.*

4 Constructing examples/instances

The task classification *constructing examples/instances* requires the student to provide an object satisfying certain mathematical properties. Typically there will be many correct solutions but no general method by which such a solution may be constructed. However an example may be only a matter of factual recall, as in the case of many famous examples, such as the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ as an example of a divergent series generated by a non-negative sequence which converges to zero. In general the desired properties will generate a set of criteria and hence a mechanical procedure for testing whether a solution is correct. That no unique solution exists means that, in practice, such questions are difficult to mark, requiring close attention from the marker.

This category is mentioned by [4] who, differing from the usual pattern in that paper, provides no examples of such questions. In fact, the recent analysis of tasks set to level one undergraduates [5] revealed questions requiring the construction of examples are rare in undergraduate work.

In general exemplification is the process where something specific is taken to represent the general. Previous work [7] identified four classes of examples; *startup examples*, *reference examples*, *model examples*, and *counter examples*. Recent work [8] encourages the construction of *boundary examples* as a method of promoting higher level skills. We will argue that [8]'s *boundary examples*, and other similar techniques, are but one part of a more general higher level skill of *constructing instances*. Much of interest has already been conjectured and written about the rôle examples play in the teaching and learning of mathematics, for example [7, 9, 10, 11], and in the development of the subject as a whole [12, 13]. Research [2, pg 293] concluded that *'The generation of and reflection upon examples provided powerful stimuli for eliciting learning events'*, but that students show a reluctance to generate their own examples

[14]. Absent from these, and other works, are practical methods for marking complex examples generated as solutions to course work problems at undergraduate level, some of which might involve significant calculations. This is the issue we address.

We concentrate on this question classification category for the remainder of this paper, and provide a number of sub-classifications illustrating the sorts of tasks that may be classified under this heading. When carefully posed so that appropriate factual knowledge is neither pivotal or decisive, success with such questions demonstrates genuine creativity on the part of the student. It is precisely the lack of routine methods, or prototype worked examples, which require students to develop for themselves techniques to solve these kinds of questions.

It is also conjectured that the skill of instance creation will aid students in checking conditions in other situations, such as when applying a theorem, when the hypothesis needs to be satisfied by the case in consideration.

4.1 Examples of...

One may ask students to directly provide examples of objects satisfying certain properties, such as:

Example question 7 *Find a cubic polynomial with a critical point at* $x = 1$ *passing through the co-ordinate* $(-1, 2)$.

Example question 8 *Find a singular* 5×5 *matrix with no repeated entries.*

Of course, one may illustrate the multiplicity of correct solutions by asking for more than one example. One could ask a student to provide 'as many different examples as possible'. Given one example this might well provide an opportunity to consider *dimensions-of-possible-variation*, ('what can be changed?') and within each dimension, the *range-of-permissible-change*, ('how much can be changed?') without destroying these properties [11]. The objects at the extremities of these ranges in some sense provide the *boundary examples*. As a specific example consider the harmonic series as a boundary example when we consider the convergence or non-convergence of series of the form $\sum_{n=1}^{\infty} \frac{1}{n^p}$, where $p \in \mathbb{R}$.

The boundary examples of [8] are a particular form of this type of question in which a list of criteria are given which are progressively restrictive. This list is then reversed and the final condition negated to demonstrate the boundaries between concepts. For example given criteria A, B, and C:

1. Find an x satisfying A.
2. Find an x satisfying A and B.
3. Find an x satisfying A and B and C.
4. Find an x satisfying A and B but not C.
5. Find an x satisfying A but not B and not C.

It has been suggested that an effective way to pose such questions would be to reveal the criteria one at a time, but in such a way that a typical solution to part 1 would satisfy A–C inclusive. When part 4 is posed the real learning begins as the original answer becomes invalid.

4.2 Counter examples to...

In creating a counter example to a proposition or theorem such as

$$\text{If } \underline{A} \text{ then } \underline{B} \tag{1}$$

there are two distinct stages. Firstly deriving the criteria, that is stating and simplifying the criteria

$$\underline{A} \text{ and } \underline{\text{not } B} \tag{2}$$

which the example must satisfy. Secondly providing the actual example itself and showing this indeed satisfies (2).

Example question 9 *Prove that if $n \in \mathbb{N}$ then $n^2 - n + 41$ is prime, or find a counter example.*

Example question 10 *Provide a counter example to the following statement. 'If (a_n) is a sequence of non-negative real numbers that converges to zero then $\sum_{n=1}^{\infty}(-1)^n a_n$ converges.'*

4.3 Facilitator objects

Some theorems or techniques in mathematics require the use of *facilitator objects*. Given a question, the technique requires a complementary object to be constructed. These two must interact either, to satisfy a set of criteria, or to simplify the original question to some canonical form, to allow the application of a standard technique. Typically this may be a substitution of some kind, or a change of co-ordinates. Where this is not provided or a hint is absent this question presents quite a challenge. In fact, questions that begin

Example question 11 *Using the substitution $u = \ldots$ show that ...,*

are ubiquitous. It is contended here that without providing an explicit substitution they would be classified as requiring the creation of an instance.

Example question 12 In dynamical systems theory proving the stability of a continuous dynamical system can be achieved by finding a Lyapunov function. This requires the generation of a facilitator object: namely the Lyapunov function itself.

Once the object has been found, the working is typically straight forward and relatively easy to mark. Factual knowledge about the general form of the required facilitator object is often indispensable but such knowledge does not immediately provide the solution. For example, while experience shows that 'almost all Lyapunov functions are quadratic-like' such questions can still be very difficult to answer. Furthermore, application of such knowledge demonstrates creativity, insight and the ability to transfer reasoning – all typical of higher level mathematical skills.

4.4 General *vs* specific instances

Some questions in mathematics require instances which are general. This apparent contradiction is perhaps best explained by the following example. When trying to prove, from first principles, that the sequence (a_n) defined by $a_n := \frac{1}{n^2+1}$ converges to zero, the solution requires, in part, the solution to the problem

Example question 13 *For each $\epsilon > 0$, find an example of an $N = N(\epsilon)$ such that if $n \geq N$ then $|\frac{1}{n^2+1}| < \epsilon$.*

Notice that, in the context of the specific sequence, the solution requires for each ϵ the generation of a specific $N = N(\epsilon)$ that 'works'. That is, the N must be an instance with certain properties. Compare this with the specific instance

Example question 14 *Find an example of an N such that if $n \geq N$ then $|\frac{1}{n^2+1}| < \frac{1}{100}$.*

Thus a distinction may be drawn between creating specific and general instances, as typified by the Example questions 13 and 14.

5 Analysis of current undergraduate work

Current undergraduate course work has recently been evaluated, in order to better understand the effects of prototype hand held calculators equipped with `Maple V`* [5]. This analysis generated the question classification scheme detailed in the previous sections. The results of this study, which involved an analysis of some 489 individual questions, show that 84.5% of course work questions, and 71.2% of examination questions are classified as either routine procedures or proofs. Many of these are proofs chasing definitions via an algebraic demonstration. Some require verbatim transcription or only minor modifications. This is not intended as a criticism – these proofs provide 'proof templates' from which more complex proofs in subsequent years may be built. It is entirely appropriate for first year courses that students' confidence be build by requiring simple adaptations of prototype proofs. In one of the courses examined by the study the proofs are more substantial with additional logical structure, such as contradiction or induction. Questions identified as requiring students to use the higher-level skills (6–8) account for only 3.4% of the questions asked. The conclusion drawn from this preliminary analysis is that (i) *the vast majority of current work may be successfully completed by routine procedures or minor adaption of results learned verbatim* and (ii) *the vast majority questions asked may be successfully completed* without *the use of higher skills.*

6 Marking using computer algebra

It is evident that the type of questions of Section 4 generally do not have unique answers and so such questions are often difficult or time consuming to mark, particularly so at university

*`Maple V` and `Maple` are registered trademarks of Waterloo Maple Software.

level. Furthermore, questions which require the construction of an instance might appear to be necessarily advanced, such as Example question 10. This need not be the case. Consider the following example questions which test a simple understanding of complex numbers.

Example question 15 *Given the complex number $z := 7 + 3i$ find:*

1. *a complex number $w = a + ib$ such that $z + w$ is real,*
2. *a complex number $w = a + ib$ such that $z.w$ is real,*
3. *a complex number $w = a + ib$ such that both $z + w$ and $z.w$ are real.*

When is your answer to (3) unique?

When a routine algebraic procedure exists to mark such a question this may be performed by a computer algebra system. When computer algebra and computer aided assessment are linked, this may be done automatically. An example of such a system is `AIM`, which uses the `Maple` computer algebra system to grade students' answers. Such automatic marking provides new opportunities to pose questions which educational theory has demonstrated to be greatly beneficial, but which would be impractical to set if marked by hand.

The `AIM` system is web based, both students and staff access the system using a normal browser and enter their responses using `Maple` syntax. All administrative tasks are available via the web, including authoring questions, creating quizzes (access, due date), generating grade sheets & statistics. Questions are (relatively) easy to author. There are two versions[†] of `AIM`, the original system [15] and the more recent version [16]. The system is *free*. Specific details of the system are available elsewhere [17, 18, 19].

The use of computer algebra allows the expression entered by the student to be manipulated and tested symbolically, using the impressive collection of genuine mathematical operations available within `Maple`. The key to this system is the marking procedure which must cope robustly with a variety of correct answers. This challenge can be met in some cases by computer algebra which can certainly test equivalence of expressions, including trigonometric forms; perform calculus operations; perform matrix operations; substitute values and simplify expressions, (e.g. substitute a function into a differential equation and simplify to check its' status as a solution) etc. As a specific example consider the following,

Example question 16 *Give a differentiable function with a critical point at $x = 1$.*

If $f(x)$ represents the student's answer the criteria we derive to check a solution is that $f'(1) = 0$. Thus to check the answer is correct we perform the following procedure

1. Differentiate the solution with respect to x,
2. Substitute the value $x = 1$,

[†]The working system may be viewed at `http://mat111.bham.ac.uk/` or may be downloaded from `http://aim.shef.ac.uk/aimsource/`. There is also a JISC mailing list, `aim@jiscmail.ac.uk`, for staff and developers. See `http://www.jiscmail.ac.uk/lists/aim.html`

3. Test the resulting expression equals zero.

This may be done with the following (simple) Maple code:

```
Mark := ans -> evalb( simplify( subs(x=1,diff(ans,x)) ) ) = 0);
```

Briefly, `evalb(·)` evaluates the enclosed expression as a Boolean, `simplify` is self-explanatory, `subs` is a procedure that substitutes $x = 1$ into the second expression, and `diff(·,·)` differentiates the first expression with respect to the second. In this case `Mark(x^2-2*x+3);` and `Mark((x-1)^4);` both evaluate to true. Other conditions may be checked using Boolean connectives, so that to mark Example question 7 we could use a Maple procedure

```
MarkAgain := ans -> evalb( simplify( subs(x=1,diff(ans,x)) ) ) = 0)
and degree(ans)=3 and subs(x=-1,ans) = 2);
```

This uses the Maple procedure `degree(·)` to check the answer is a cubic and another substitution to ensure the function passes through the point $(-1, 2)$. Maple correctly marks the answer $(x + 1)(x - 1)^2 + 2$ for example.

Note the system only tests differentiability at the point $x = 1$. A function, such as $|x|(x-2)$, which is differentiable at $x = 1$ and has a local minimum there, is marked as correct. This function is not differentiable (everywhere). The marking scheme actually only considers the following question.

Give an example of a function $f(x)$ that satisfies $f'(1) = 0$.

Furthermore, there are many answers which are difficult to express using Maple syntax, such as functions defined piecewise. Thus the questions need to be designed carefully and tested to ensure the system is robust.

In the case of Example question 8 we could use a marking procedure such as

```
Mark := ans -> evalb (det(ans)=0 and nops(convert(ans,set))=25);
```

This uses the function `det` from the `linalg` package to check the matrix is singular and that the *set* of elements is of the correct size – the idea here is that sets do not contain duplicate elements.

In the case of multiple conditions, such as in Example question 7, we would like to provide high quality tailored feedback and may want to give partial credit. If only one condition is violated, for example, the automatic procedure could indicate which. In the case of Example question 8 this feedback might extend to giving a list of repeated entries using code such as

```
Feedback := ans -> FindRepetitions(Flatten(convert(ans,listlist)));
```

(which uses procedures from the Maple 7 package `ListTools`). This is possible in the AIM system, and has been done. In Section 7 students' reactions to questions of this type as implemented in AIM are given. Further discussion of how the system may be used to initiate a debate and to begin the process of example generation is given in Section 8.

It is not claimed that AIM is the only computer aided assessment system on which it is possible to assess in this way, nor that AIM represents the current best such system. However, competitors are not evident, nor does this question strategy seem to be encouraged. It is firmly believed that in the near future *all* computer aided assessment systems will link computer algebra and assessment to perform automatic marking [18].

7 Students' reactions to these questions

The AIM system is currently being integrated into the core of a first year, UK university, mathematics degree [20], with the help of an institutional Learning Development Unit grant.

Analysis of the current work showed how rarely students were asked to think in creative ways, and how success could be achieved largely using mechanistic or rote learning. Since it was the intention to introduce questions into the course which departed from this, by demanding the creation of instances, there was concern that this was 'radical' and might alienate a large section of the student body. Such an outcome would be entirely counter productive. In an attempt to better understand students' reactions to questions of this type a group of six second year students were recruited to answer the trial questions detailed in Appendix A. This is a small study, never the less the results are revealing and interesting. Further work in ascertaining how students react and consequential refinement in posing such questions would no doubt be of great benefit.

The members of the group were *selected* and invited personally to join this study. These were students with a variety of experience but all were known to have been previously successful in the examinations for the core course. Furthermore, they were known to be comfortable talking in a group about their work. Two of the group had been previously employed to help with secretarial level coding work on the initial installation of the AIM system itself. The group was mixed gender.

The group was perhaps a-typical of the student body. This was deliberate. Firstly they could be better relied upon to report their thoughts honestly in a group – it was impractical to conduct individual interviews. Secondly, these students had succeeded on previous summative assessments – the material itself should present less difficulty than for other students. For two individuals at least, prior experience of using AIM removed any barriers the system itself might present the students.

Students were recruited to 'test some questions I have written for the A core course. ... They are set using the computer aided assessment package AIM.' Further, the students were reassured that 'I am testing the questions, not you. You may use any notes, books, calculators or other aids you like. However, please try to do the questions on your own, not as a committee!' This was to set students at ease, not patronize successful second year students by 'testing' them with level one material, and thirdly notes or calculators should in any case not help by providing an 'answer'.

Setting the questions themselves presented a challenge. The characteristics sought were that questions should test simple core material and have non-unique correct answers which the student would have to create. Questions from a variety of topic areas, such as algebra, calculus, complex numbers, linear algebra, differential equations etc were set. Isolated from lectures this was found to be significantly more challenging than embedding questions in a coherent course. Discussion of the questions in Appendix A will be given below.

The web based nature of the AIM system allowed the students to take the tests when and where was most convenient. Two students worked from home over a weekend, the others used an open lab during the week. All students completed all the material, the vast majority (91%) correctly. Two students split their time into two sessions. By using the data logged by the AIM system itself, the time students took varied between an hour and an hour and a half. This information was subsequently confirmed during the interviews. All students met and discussed their reactions to these questions. With their permission this discussion was tape recoded, transcribed verbatim and analysed. Question 5 in Appendix A contained a bug in the computer code and correct answers were incorrectly marked by the system. Care also needs to be taken that a guess will not suffice. Unfortunately the automatic feedback given for question 9 promoted this behaviour.

> CY: You actually get marks for guessing as well because it tells you its lower. 'This needs to be lower.' If you get it wrong it tells you you need to be lower.

Students reported difficulties in recalling basic definitions. All took advantage of their notes, one student showed particular resourcefulness in this respect.

> CY: When you said you could use anything, when I did that [Q8] I didn't even use my notes, I looked up symmetric matrix on the Internet.

Many of these questions exposed a lack of student confidence with concepts which resulted in over complex solutions to problems:

> D: To begin with I thought you wanted a particular formula [for Q2], then I thought about it and thought, e just on its own - why not?

This student went on to say

> D: At first I thought this [i.e. Q2] is a good question let's have a look at it, then I thought that if n isn't part of it, it won't matter so it won't depend on n at all.
> J: It depends on your confidence of how you can put it in.
> Interviewer: Why is that a cheat?
> D: I just felt a bit tricked by it personally.

When asked why she had entered x^3+x^2+x+4 in answer to Question 1 one student responded

> J: I don't view x^3 as a proper cubic as a polynomial has more than one term as far as I'm concerned. I know technically it doesn't. But I assume a polynomial has a couple of terms.

That such questions have non-unique correct solutions should be exposed. This could be achieved by asking students to give more than one *different* solution to a particular problem. 'Different' here is taken to mean more than trivial modifications. For example, the answers to Question 2 supplied by the students were

$$a_n := e, \text{ [1 student]} \quad a_n := e + \frac{1}{n}, \text{ [2 students]} \quad a_n := e + \frac{1}{n} + \frac{2}{n} + \frac{3}{n} + \frac{7}{n^{10}}, \text{ [1 student]}$$

$$a_n := e^{\frac{1+n}{n}}, \text{ [1 student]} \quad a_n := \left(1 + \frac{1}{n}\right)^n, \text{ [1 student]}.$$

These students have responded with four identifiably different solutions: (i) the constant sequence, (ii) $e + p(n)$ where $p(n) \to 0$ as $n \to \infty$, (iii) $e^{p(n)}$ where $p(n) \to 1$ as $n \to \infty$, and (iv) Euler's particular sequence $a_n := \left(1 + \frac{1}{n}\right)^n$.

Asking each student to provide more of these, perhaps by asking 'Give an example, give a *different* example. Give a third example.' would require them to think more deeply about the concepts. That is to say to expand the dimensions-of-possible-variation in which the student operates [11]. Example question 5 expands the domain of variation of the concept of a vector space, for instance. At least one student responded

> B: I don't know why I didn't put e, it didn't occur to me.

But, with hindsight, neither did the question encourage him to seek for examples other than his correct $a_n := e + \frac{1}{n}$. Assessing what constitutes 'difference' in a particular circumstance could result in valuable discussion in tutorials and might generate the sort of discussions the students engaged in over Question 3.

> JL: Ok, just take the parabola and shift it one.
> ...
> B: I said, $x - 1 = 0$, then integrated it.

These two approaches caused some friction in the group when the discussion moved to the general characteristics of the questions themselves. Student P was also 'an integrator' while D was a 'graph shifter'. Only reluctantly does P acknowledge an alternative method, which his voice intonation and manner revealed he had not considered.

> P: Because if it says, give an example of a differentiable function which has a turning point at $x = 1$ you have got to know what it means for a function to have a turning point at 1 and then infer back that means you can integrate this function so that it has ...
> D: You don't have to do it that way! You could do it graphically like I did and consider x^2 and move it along.
> P: Yes, it is not a simple case of saying - this is a polynomial, where are its turning points? We know how to do that. You can differentiate it and set it equal to zero. This is saying 'we are giving you a turning point what is the function?' You have then got to understand the process behind how you work out and apply it backwards ... or in this case you could sort of draw it and realize it is a shift.

The technique needed to succeed may pre-suppose fluency with the basic techniques. However, if a technique has not been mastered hitherto it is precisely the *need* to use the technique as part of a wider problem than will first expose any weakness and secondly encourage remedy. Question 9 exposed a weakness in understanding solutions to second order differential equations and generated much discussion.

> JL: I approached the problem, ok I did the auxiliary quadratic equation then I thought ok its got to do something with the discriminant. And then is it less than zero, more than zero or equal to zero?

B: It requires you to have knowledge of what erm....

CY: ... what changes the general solution.

B: The form of the general solution you get from the auxiliary equation. So like, $A\cos(\alpha t) + B\sin(\alpha t)$... I thought that was a prime candidate for the first graph. It looks like it could be described by that equation.

At the risk of exposing a failure on our part, students were adept at solving second order differential equations of a variety of types but uniformly had not considered the shape of the graph of such a solution.

B: Last year we didn't really look at graphs of differential equations did we?

CY: We were just told that if we manipulate the auxiliary equation we can get these different types of general solutions to these differential equations.

Interviewer: No attempt to relate the solution....

CY: Not to a graph.

D: The first graphs we did were Bessel's functions weren't it?

[General agreement - laughter]

Such behaviour is corroborative evidence that assessment drives learning: students had not been asked to consider the graphs of their solutions. With hindsight it would also have been interesting to ask them to ascertain the initial conditions *from the graph*.

Having discussed each question individually, the students were asked 'I designed these [questions] differently, that is why I've got you together. If I'd been doing normal things I wouldn't have done it right? But this is slightly different'. To which B responded.

B: Question 9 was certainly asking us stuff that we had to think about.

Interviewer: In what way?

B: You didn't give us an equation and then say 'solve it'. You have got to really think about what it means. You have to get a solution and then you think, ok that's the answer. Doing a question like that you think, argh, right, that is the shape of the graph.

As a group the students demonstrated a mature understanding of the purposes of these questions. The student generally agreed with P that such questions

P: ... test your understanding of the subject, rather than your ability to turn a handle.

8 Conclusion

This article has defined the higher level skill of *constructing instances* of mathematical objects satisfying certain properties. It has been indicated that students are asked to perform these tasks rarely. It has been explained how such questions may be assessed *in practice* without the imposition on staff of an onerous marking load. There is great potential in the simple ideas contained in this paper: having the students generate their examples is now feasible at undergraduate level in some cases where a computer algebra system may check their solutions automatically. However, the use of computer algebra marking is only envisaged as one part in the learning cycle. It occurs that as a tool for generating *tutorial debate* the system has great potential.

A number of practical questions present themselves. What topics could be automatically marked by the system when question requiring student generated examples's are posed? How 'difficult' should the questions be? How many different examples should each student generate? How should these results be used subsequently in lectures or to generate debate amongst students? What automatically generated feedback can be provided to support the process of eliciting examples? The pilot trial has shown the feasibility of using the new technology. Further research is required to better understand effective methods of use.

8.1 Acknowledgements

The author would like to thank Joe Kyle, Grant Keady, John Mason and Anne Watson for valuable discussions and encouragement. Also thanks to the students who tested the questions for their diligence and permission to quote their comments. Introduction of the AIM system described in this project has been funded in part by a grant from the Learning Development Unit at the University of Birmingham, UK. The author gratefully acknowledges Dr Mike Finney, School of Mathematics and Statistics Computer Support Officer, for technical support.

A Student trial group questions

1. In the following questions we ask you to consider a polynomial $p(x)$ and the following properties:

   (a) $p(x)$ is a cubic

   (b) $p(a) = 0$, [Random: $a \in \{1 \cdots 5\}$]

   (c) $p(b) = 0$, [Random: $b \in \{1 \cdots 4\}$+a]

   (d) $p(0) = a \times b$, [Generated by $a \times b$]

   Give an example of a polynomial satisfying:
   1. (a),
   2. (a) and (b),
   3. (a), (b) and (c),
   4. (a) – (d) inclusive,
   5. only (b) – (d).

2. By giving the formula for the general term a_n in a sequence, give an example of a sequence with limit
$$\lim_{n\to\infty} a_n = e.$$

3. Give an example of a differentiable function $f(x)$ which has a critical point at $x = 1$.

4. Given the complex number $z := 1 + 10i$, [Random: real and imaginary parts in $\{1 \cdots 10\}$] In each case find a nonzero complex number $w := a + bi$ such that

   (a) $z + w$ is real.

   (b) $z.w$ is real.

   (c) *both* $z + w$ and $z.w$ are real.

   Given a nonzero complex number z, is there a unique nonzero complex number w that satisfies both $z + w$ and $z.w$ are real?
   Multiple choice from:
   ☐ No, there are more than one,
   ☐ Not unique if z is real,
   ☐ Yes, always unique,
   ☐ Don't know.

5. Give a complex number z with nonzero imaginary part such that

   (a) z^3 is real,

   (b) z^3 is real and negative,

   (c) z is a solution to $z^3 = -1$.

6. Give an example of a nonzero singular 3 by 3 matrix.

7. This question tests that you can apply the definition of 'permutation matrix'. (Reminder. A permutation matrix is a square matrix of 0s and 1s with precisely one 1 in each row and precisely one 1 in each column.) Enter a 5 by 5 permutation matrix P which has a 1 in its $(1, 2)$-place [Random] and for which the smallest power of P which gives the identity matrix is P^6. (Hint: Consider 'combining' a 2 by 2 permutation matrix with a 3 by 3 one.)

   There are several correct different solutions. The question has been designed so that there should be a block diagonal permutation matrix solution.

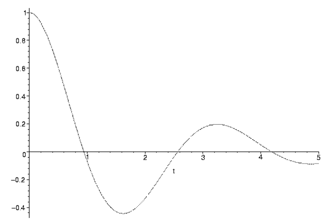
8. Enter a 2 by 2 matrix $A = [a_{ij}]$ which is symmetric and orthogonal, but which is not a diagonal matrix.

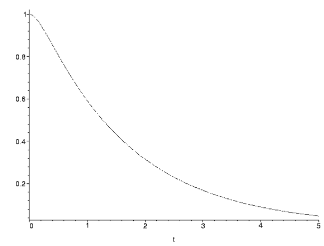
9. Consider the solutions to the differential equation

$$y'' + ay' + 4y = 0$$

   subject to the initial conditions $y(0) = 1, y'(0) = 0$. Here the prime denotes differentiation with respect to t.

   Examine the graphs below. For each give a value of a so that the solution satisfies the initial conditions and has qualitatively the same behaviour as the solution in the graph.





References

[1] G. Gibbs and C. Simpson. How assessment influences student learning – a literature review. *Preprint*, 2002.

[2] R. P. Dahlberg and R. P. Housman. Facilitating learning events through example generation. *Educational Studies in Mathematics*, 33:283–299, 1997.

[3] G. Gibbs. Using assessment strategically to change the way students learn. In *Assessment Matters in Higher Education: choosing and using diverse approaches*, pages 41–53. Society for Research into Higher Education & Open University Press, 1999.

[4] G. Smith, L. Wood, M. Coupland, and B. Stephenson. Constructing mathematical examinations to assess a range of knowledge and skills. *International Journal of Mathematics Education in Science and Technology*, 27(1):65–77, 1996.

[5] A. Pointon and C. J. Sangwin. An analysis of undergraduate core material in the light of hand held computer algebra systems. *International Journal for Mathematical Education in Science and Technology*, accepted.

[6] M. Roach, P. Blackmore, and J. Dempster. Supporting high level learning through research-based methods. *Innovations in Education and Training International*, 38(4), 2001. http://www.telri.ac.uk/ (viewed August 2002).

[7] E. R. Michener. Understanding understanding mathematics. *Cognitive Science*, 2:361–381, 1978.

[8] J. Mason and A. Watson. Getting students to create boundary examples. *MSOR Connections*, 1(1):9–11, 2001. http://ltsn.mathstore.ac.uk/ (viewed August 2002).

[9] G. Polya. *How to solve it*. Princeton University Press, 1973.

[10] D. Tall. *Advanced mathematical thinking*. Dordrecht: Kluwe, 1992.

[11] A. Watson and J. Mason. Student-generated examples in the learning of mathematics. *Canadian Journal for Science, Mathematics and Technology Education*, 2(2):237–249, 2002.

[12] B. R. Gelbaum and J. M. H. Olmsted. *Counterexamples in analysis*. Holden-Day, 1964.

[13] I. Lakatos. *Proofs and refutations*. Cambridge University Press, 1976.

[14] R. C. Moore. Making the transition to formal proof. *Educational Studies in Mathematics*, 27:249–266, 1994.

[15] S. Klai, T. Kolokolnikov, and N. Van den Bergh. Using Maple and the web to grade mathematics tests. In *Proceedings of the International Workshop on Advanced Learning Technologies, Palmerston North, New Zealand, 4–6 December*, 2000. http://allserv.rug.ac.be/~nvdbergh/aim/docs/ (viewed August 2002).

[16] N. Strickland. Alice interactive mathematics. *MSOR Connections*, 2(1):27–30, 2002. http://ltsn.mathstore.ac.uk/newsletter/feb2002/pdf/aim.pdf (viewed December 2002).

[17] D. F. M. Hermans. Embedding intelligent web based assessment in a mathematical learning environment. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation (ISSAC2002), Internet Accessible Mathematical Computation 2002 workshop, Lille*, July 2002.

[18] C. J. Sangwin. Assessing higher skills with computer algebra marking. *JISC Technology and Standards Watch*, TSW 02–04, September 2002. http://www.jisc.ac.uk/techwatch/ (viewed September 2002).

[19] C. J. Sangwin. Encouraging higher level mathematical learning using computer aided assessment. In *Proceedings of the International Congress of Mathematicians Satellite Conference on Mathematics Education, Lhasa, Tibet, August 12–17*, 2002.

[20] J. Kyle and C. J. Sangwin. AIM – a parable of dissemination. In *Proceedings of the 2nd International Conference on the Teaching of Mathematics, Hersonissos, Crete, Greece, July 1–6*, 2002.