Citation for published version:
Bradford, R, Davenport, JH & Sangwin, C 2010, 'A comparison of equality in computer algebra and correctness in mathematical pedagogy (II)', *International Journal for Technology in Mathematics Education*, vol. 17, no. 2, pp. 93-98.

Publication date:
2010

Link to publication

Reproduced here by kind permission of Research Information Limited. Journal Homepage: http://www.researchinformation.co.uk/time.php

University of Bath

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 07. Dec. 2019

A Comparison of Equality in Computer Algebra and Correctness in Mathematical Pedagogy (II)

Russell Bradford[1], James H. Davenport[1] and Chris Sangwin[2]

1 Department of Computer Science, University of Bath, Bath BA2 7AY, United Kingdom
2 School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom
R.J.Bradford@bath.ac.uk, J.H.Davenport@bath.ac.uk, C.J.Sangwin@bham.ac.uk

Received: 15 September 2009 Revised: if applicable

A perennial problem in computer-aided assessment is that "a right answer", pedagogically speaking, is not the same thing as "a mathematically correct expression", as verified by a computer algebra system, or indeed other techniques such as random evaluation. Paper I in this series considered the difference in cases where there was "the right answer", typically calculus questions. Here we look at some other cases, notably in linear algebra, where there can be many "right answers", but still there can be answers that are mathematically right but pedagogically wrong. We reformulate the problem in terms of articulating the sought-after properties, which may include both mathematical *equivalence and* algebraic *form.*

1 INTRODUCTION

Computer-aided assessment (CAA) of mathematics is a growing trend, endorsed by the American Mathematical Society (Lewis and Tucker, 2009). The first of their three recommendations is "Harness the power of technology to improve teaching and learning". There are many advantages to such schemes, as exemplified in the large Rutgers study by Weibel and Hirsch (2002): not only sheer feasibility, but also immediacy and rewarding persistence. However, once we move beyond simple multiple-choice, grading mathematical answers is not as easy as might be envisaged.

Acknowledgements. We are grateful to AJP (John Power, Bath), the teacher of CM10197, Sally Barton (University of Nottingham) for contributions to section 6, the Applications of Computer Algebra organisers and referees, the IJTME referees, and the London Mathematical Society for funding the travelling.

2 MATHEMATICAL CORRECTNESS

Once one gets beyond simple numbers, mathematical correctness is no longer textual correctness, and is not a simple task for automated assessment schemes. The student's answer is generally (with today's technology[1]) a series of characters: how does the mathematical object represented by them relate to the mathematical object that is the desired answer? Even in "arithmetic", this issue arises. Suppose the question is "*Subtract 5 from 12*". We expect the student to answer `7`, but what if the answer is `07`, or `+7`? Both alternatives should probably be regarded as correct.

The reader may well complain that this is "merely a question of parsing", and in the purely arithmetic context it would be possible to reduce issues of correctness to issues that parsing technology can solve, but what about $a+b$ versus $b+a$? In this paper we will assume that the student's typed input (or indeed handwriting) *has* been parsed and we are considering the correctness of this parse tree. For comments on the practical challenges of parsing in this context, see Sangwin & Ramsden (2007).

Lewis and Tucker (2009) survey three systems.

- **WeBWorK** (Gage *et al*., 2002) checks the student's answer against the "correct" answer by random evaluation: a technique that is, in fact, unreasonably effective for a wide range of calculus and pre-calculus problems.
- **MapleTA** uses the Maple computer algebra engine to check the student's answer against the correct one.
- **Webassign** uses built-in techniques, and no details are available to the authors.

In addition,

- **STACK** (Sangwin, 2007) uses the Maxima computer algebra engine both to establish algebraic equivalence and satisfaction of a (fixed) range of forms.

Within their scope, these techniques seem to provide good solutions to the problem of determining mathematical correctness. The importance of this should not be under-estimated: few things are more damaging to pedagogy than a wrong answer that is marked as right, and class confidence can easily be shattered by a few examples of the teacher having to say "yes, that's right even though the computer marked it wrong".

3 BEYOND MATHEMATICAL CORRECTNESS

We can identify three major challenges.

1) The answer may be mathematically correct, in the sense that a "sufficiently powerful" algebra system (or mathematician) would agree that it was equivalent to the `right' answer, but no teacher would agree that it *was* the (or a) correct answer. We will call an answer that

[1] Practical recognition of hand-written mathematics will soon be upon us, but this will make the problems we are describing more, not less, salient.

teachers would agree to be correct a *pedagogically correct* answer.

2) The answer may not be in the form required, even though mathematically correct, and indeed pedagogically correct were it not for the specific stipulation in the exercise: "Express *X* in the following form …". This is the specific focus of this paper, which has not received much attention so far. For example, Lewis and Tucker (2009) write of WeBWorK "However, if a specified simplification of an expression is desired, […] WeBWorK cannot be used".
3) The answer may well be mathematically incorrect, but nevertheless be "nearly right".

We illustrate these points with possible answers to the exercise "*express* $\dfrac{x-3}{(x-2)(x^2-1)}$ *in partial fractions*".

1) $\dfrac{1}{2}\dfrac{3-1}{(3-2)(x-1)}+\dfrac{1}{2}\dfrac{-3-1}{(x+1)(1+2)}+\dfrac{-3+2}{(x-2)(2^2-1^2)}$ (probably the result of direct substitution into a canned formula) is mathematically correct but not pedagogically correct;
2) $\dfrac{x+5}{3(x^2-1)}+\dfrac{-1}{3(x-2)}$ is correct, and simplified, but not completely in partial fractions;
3) $\dfrac{2}{(x-1)}-\dfrac{4}{3(x+1)}-\dfrac{2}{3(x-2)}$ is wrong by a factor of 2, but is otherwise simplified and in complete partial fraction form.

The first challenge is fundamental to moving beyond simple "mathematical correctness" in computer assessment. A student's answer may well be *mathematically* correct, in the sense that a computer algebra system would say that it is equivalent to "the correct" answer, without being pedagogically correct, in the sense that a human teacher would award it full marks. In Bradford *et al.* (2009) we introduced the concept of classifying the rules for mathematical correctness according to their pedagogical appropriateness, and defined "the correct answer" as that which was both mathematically correct *and* simplest in the sense of Carette (2004), which *roughly speaking* means "smallest". There may in fact be several such, e.g., *a+b* or *b+a*, but they must all be equivalent under the underlying rules.

This philosophy works as long as

- there is a single (up to the application of the underlying rules) correct answer, as tends to be case in calculus, and
- there are no other constraints on the answer, such as "in partial fractions" or "reduced to upper triangular form".

Some other mathematical domains do not lend themselves to this paradigm so easily, though, and in this paper we explore the application of our philosophy to such areas, taking "undergraduate linear algebra" as our opening specific example.

4 THE CLASSIFICATION OF RULES

In Bradford *et al.* (2009) we considered the abstract model of a computer algebra system as operating via a set of rewrite rules. We stress that it is *not* necessary that the system actually operates this way: some do (e.g., Mathematica) but many, such as Maxima which underlies STACK (Sangwin, 2007), do not. Within such a model, we proposed classifying the rules into three types.

Underlying: those that the student, or the algebra system, is free to apply at will, without changing mathematical *or* pedagogic correctness. A typical example of this class would be commutativity of addition as in *a+b* or *b+a*, both of which are normally[2] "equally correct".

Venial: those that the student *ought* to have applied, but which aren't the main thrust of the subject matter in hand. Once we have got beyond basic algebra, rules such as "combining terms" and "carrying out numeric computations" fall into this class, and *might* attract a deduction of marks. It should be noted that the venial rules might be sub-divided into "degrees of incorrectness", and that this classification might evolve with the pedagogic context: see the discussion of trigonometric contraction in Bradford *et al.* (2009).

Fatal: those rules which are fundamental to the subject matter in hand, and, if the computer algebra system needs to apply them to get the "correct answer", the student hasn't truly answered the question. In the detailed example of Bradford *et al.* (2009), which was differentiation, the actual differentiation rules fell into this category.

We should note that this classification, and the marking penalties to be applied in the case of the venial rules, is very dependent on what we have called "the subject matter in hand".

An answer is then: *correct* if it is equivalent to the correct answer under just the underlying rules; *partially correct* if it is equivalent to the correct answer under just the underlying and venial rules; and (pedagogically) *incorrect* otherwise. If it is equivalent to the correct answer under the underlying, venial and fatal rules, then the feedback ought to be along the lines of "OK. But that isn't really a finished answer is it?", as with the student who, on being asked to differentiate xe^x, replies with $x'\exp x + x(\exp x)'$. Conversely, if it is not equivalent to the correct answer under even the fatal rules, then it is *mathematically* incorrect. "Buggy rules" (Brown & Burton, 1978) may then be used to identify possible causes for the error, from which feedback may be generated. An example of a buggy rule would be to rewrite $(x+y)^n = x^n + y^n$ when $n \neq 1$.

[2] But commutativity of addition would probably *not* be "underlying" for the terms of a power series: $1-x+x^2-x^3+\cdots$ is acceptable, but the alternative $1-x^3-x+x^2+\cdots$ is not. This emphasises the dependence of pedagogic correctness on context.

© 2009 Research Information Ltd. All rights reserved.

5 UNDERGRADUATE LINEAR ALGEBRA

5.1 The subject matter

Two subjects clearly separate school mathematics from work at university: real analysis and linear algebra. Since the latter is covered world-wide, and contains a significant methods-based component we shall consider this here. Our precise subject matter is the linear algebra component of the course CM10197 "Analytical mathematics for applications" at Bath[3].

5.2 Applicability of the paradigm

Many of the questions one might want to set in teaching such material *do* have precise answers: in our terminology answers which are unique up to the application of the underlying rules. Hence, when teaching matrix multiplication, one would generally classify the rules for scalar expressions as underlying or venial, but the rule for matrix multiplication itself would be classified as fatal. Hence a student who answered

$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \times \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$ is marked as correct,

whereas a student who answers

$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \times \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1\times5+2\times7 & 1\times6+2\times8 \\ 3\times5+4\times7 & 3\times6+4\times8 \end{pmatrix}$ is

regarded as having committed (several) venial omissions, and a student who reduced the problem of multiplying 4×4 matrices to 2×2 matrices but no further, as having fatal omissions in the answer.

Much the same applies to questions like "compute the inverse of …", or "solve this set of (fully determined) linear equations". In this case, we note that classifying matrix multiplication and inverse as fatal would require the complete solution, whereas classifying them as venial would allow an answer of the form $M^{-1}.\mathbf{v}$ to be "partially correct". This stresses the point that the classification of the rules is dependent on the pedagogic aims[4].

Equally, though, we may choose to relegate mathematical correctness to the algebra system altogether, and just apply the underlying/venial/fatal classification to the individual scalar entries.

5.3 Lack of "a correct answer"

However, many questions in this domain do not have an unambiguous "right answer". A typical example, taken from the CM10197 problem sheets, is "reduce M to (upper) triangular form". AJP informs us that this is the only "open-ended" question in his sheets since he wishes to provide very rapid response to student homework. This is typical: Pointon and Sangwin (2003), for example, found very few such questions in their analysis of first year university coursework. While a human teacher "looks at the answer", a CAS "establishes mathematical properties". The human makes many judgements rapidly, the author of CAA has to articulate these and encode them. All question formats distort the assessment process: paper-based assessments do not encourage teachers to set open-ended questions, cited as one of the great advantages of automatic marking schemes by Lewis and Tucker (2009), and so we are not surprised that, in principle, AJP wishes he could set more such questions. In this question the pedagogic context is implicit, but one would assume that "elementary row operations", are the methods to be used.

While it would be feasible to handle "elementary row operations" by rewrite rules, the rules would look relatively cumbersome, e.g.,
matrix(rows1,rowA,rows2,rowB,rows3)→
matrix(rows1,rowA,rows2,rowB-(b/a)*rowA,rows3) where
rowA=(k zeros,a,elements) and $a\neq0$ and
rowB=(k zeros,b,elements) and $b\neq0$.

Hence it may be more appropriate simply to rely on the underlying algebra system for *mathematical* correctness. We would still suggest the approach of the previous section for pedagogical correctness, asking whether each component of the student's answer is reducible by the venial rules.

We propose that, if the parse tree representing the student's answer is N, there are three distinct stages of verification that N should undergo.

Syntactic: does N actually represent an upper triangular matrix of the right size? Note that the correct test for "upper triangular", i.e., all elements below the main diagonal being zero, is a purely syntactic one, and does not involve the rule classification outlined in the previous section. This means that we would regard the matrix $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ as being upper triangular, but the mathematically equivalent $\begin{pmatrix} 1 & 2 \\ 3-3 & 1 \end{pmatrix}$ would not be, since it is not *manifestly* upper triangular.

Mathematical: is N actually equivalent to M? If M is non-singular, then the test is "is $\det(M^{-1}.N)=\pm1$?" If M is singular (and the teacher may well not intend to teach this case!), then the correct test will depend on how the teacher has taught this case, but will certainly include "rank(M)=rank(N)?"

Pedagogic: this is where the classification of rules comes in, but yet again the pedagogic context is vital. If pivoting is not required, then the elements of N are

---

[3] http://people.bath.ac.uk/masdr/23nov08.pdf and http://www.bath.ac.uk/catalogues/2009-2010/cm/CM10197.htm

[4] We also note that expressing matrix operations such as multiplication or inverse as rewrite rules is perfectly possible, but may well, in practice, be *implemented* via the internal operations of the underlying algebra system, e.g.,

matmul(A,B) → internal_matmul(A,B).

given in closed form, e.g., from $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ we have

$$\begin{pmatrix} a & b & c \\ 0 & e-\dfrac{bd}{a} & f-\dfrac{cd}{a} \\ 0 & 0 & i-\dfrac{gc}{a}-\dfrac{(ah-gb)(af-dc)}{a(ae-bd)} \end{pmatrix}$$

but would the teacher be content simply to see numeric values substituted in unevaluated? Hence "carrying out numeric computations", described above as venial, might, in this context, be regarded as fatal, or carry a penalty of, say, 50%.

5.4 More challenging "correct answers"

A question which is not often set, in the authors' experience, largely because of the difficulty of verifying the solution, is "Give all the solutions to $M.\mathbf{x}=\mathbf{v}$", where M is not (necessarily) of full rank. Here a major problem is syntactic verification, and we probably need to force the format of the students' answers somewhat, e.g., "express in vector notation, $\mathbf{v}_0+a\mathbf{v}_1+b\mathbf{v}_2+\ldots$, where a, b etc. are parameters, *all* solutions to $M.\mathbf{x}=\mathbf{v}$". Once this is done, the three stages of verification would look as follows.

Syntactic: given the restrictions imposed, this is not too hard, but actually becomes a parsing problem, as we need to identify the various components.

Mathematical: here it is tempting to observe that, by analogy with the "evaluate at five points" strategy of WeBWorK (Gage *et al.*, 2002), we need only check that $\mathbf{v}_0$, $\mathbf{v}_0+\mathbf{v}_1$, $\mathbf{v}_0+\mathbf{v}_2$ etc. satisfy the system of the question. Indeed, since we are doing linear algebra, this is indeed sufficient, not merely a heuristic, to show that the student's answer does indeed describe solutions. However, we only know it describes *all* solutions if we also verify that dim($\mathbf{v}_1$, $\mathbf{v}_2$,…)=codim(M). In practice one would stray into "wrong answer but" territory if it was only this test that failed.

Pedagogic: the same points as above largely apply here.

Similar considerations arise with such questions as "What is the equation for the intersection of the planes …". Again, we probably need to force[5] the format.

5.5 Eigenvectors

Similar questions arise when it comes to eigen*vectors* (eigen*values* are a simple issue and the methodology of Bradford *et al.* (2009) applies). The teacher may think that the eigenvector is (1,2,3), but it is hard to mark as "wrong" the student's (17,34,51), whereas we might want to mark as (at least venially) wrong a partially-worked out answer.

When the eigenspace has dimension greater than one, the challenge is greater. We need to check that the eigenvectors are *correct*, *complete* (both mathematically and pedagogically) and *irredundant*. At this point, we have essentially the same problem as in the previous section, since we are actually asking for the solutions of $(A-\lambda I).\mathbf{v}=\mathbf{0}$.

6 OTHER SYNTACTIC ISSUES

We can see the same three-fold classification of Syntactic/ Mathematical/ Pedagogical correctness arising elsewhere, when what is required is an answer *in a certain form.*

6.1 Factorization

If it were not for the syntactic constraint "factored", of course a student would be able to return the question as the answer. This shows the importance of the "syntactic" check. However, the syntactic check is more subtle than might appear. It is insufficient to ask for a top-level product, else a student could "factor" x^2-1 as $1\times(x^2-1)$, which is of course *a* factored form, but not what the teacher intended.

The question "what *did* the teacher intend" is very important here. We can imagine at least five possible meanings, which we illustrate with respect to the polynomial $f=x^8+16x^4+48$.

1. Any non-trivial factorization, e.g., $(x^4+4)(x^4+12)$.
2. A factorization into irreducible (over the integers) factors, i.e., $(x^2+2x+2)(x^2-2x+2)(x^4+12)$.
3. A factorization over the reals such as (x^2+2x+2) $(x^2-2x+2)(x^2+2\sqrt[4]{3}x+2\sqrt{3})(x^2-2\sqrt[4]{3}x+2\sqrt{3})$
4. A factorization into irreducible (over the Gaussian integers, with i allowed) polynomials, i.e., $(x+1+i)(x+1-i)(x-1+i)(x-1-i)(x^4+12)$.
5. An absolute factorization over the complex numbers, where the (x^4+12) factor would also be split into the four terms $x\pm\sqrt[4]{3}(1\pm i)$.

While it is unlikely that a teacher would use our terminology with a class, the teacher, and the assessment scheme, must be clear what is intended. There is also the issue of what to do with the over-eager pupil, who returns, say, answer 3 or answer 4 when the teacher was only expecting answer 2.

Syntactic: For option 1 above, the only syntactic requirement is for the factorization to be non-trivial,

[5] Manual marking tends to prefer fixed formats, but without forcing them.

© 2009 Research Information Ltd. All rights reserved.

with the semantic requirement being mathematical correctness.

Mathematical: For this family of questions, mathematical correctness is probably the easiest aspect — (i) are the factors irreducible and (ii) do the factors multiply back to the original input? However, the issue of "buggy rules" and diagnosis of (mathematically) incorrect answers is also important here. Typical errors would include getting the sign, or another constant factor, wrong, which could be relatively easily tested for.

Pedagogic: Suppose the "true" answers from our computer algebra system are the factors f_i and the student's answers are the g_j: then each g_j must correspond to *some* f_i, or to some product of the f_is. Here "correspond" would normally mean "equal up to a constant factor", unless the teacher is trying to impose specific rules about the distribution of constants. Having established that g_j corresponded to f_i, i.e., that $g_j = cf_i$, then the methodology of Bradford *et al.* (2009) can be applied to see whether it is equivalent under the underlying rules, the venial rules, or the fatal rules (which would include division). We should note that the "mathematical correctness" branch of the test will ensure that the *c* multiply out correctly, or not as the case may be. If not, we have the necessary material to look for a "buggy rule" (Brown & Burton, 1978).

6.2 Partial Fractions

In elementary algebra there are relatively few *forms*, the normal pedagogic order being given by Barnard (1999). At the macro level, interesting comments are made about form in Miller (1995). After factored form, perhaps the most important is that of *partial fractions.*

Here is a typical example of an expression in both rational and partial fraction form.

$$\frac{1}{1-4n^2} = \frac{1}{4n+2} - \frac{1}{4n-2} = \frac{n+1}{2n+1} + \frac{n}{1-2n}$$

What would a teacher say to the third form on the right? It is certainly a sum of terms, each of which has a linear denominator, and for which the order of the numerator is no greater than[6] that of the denominator, and where the numerator is coprime to the denominator.

Syntactic: This example is chosen to illustrate that recognizing form is non-trivial.

Mathematical: Again, here mathematical correctness is algebraic equivalence with the original expression.

Pedagogic: In this example the unary minus can cause significant technical problems. Teachers are likely to condone differences such as the following, which involve significantly more than commutativity of addition.

$$\frac{1}{4n+2} - \frac{1}{4n-2} = \frac{1}{2+4n} + \frac{1}{2-4n}$$

In the terminology of Bradford *et al.* (2009), we have to regard $a - \frac{b}{c} = a + \frac{-b}{c}$ and $a - \frac{b}{c} = a + \frac{b}{-c}$ as further underlying rules. Further comments on the technical aspects of this are given by Heeren and Jeuring (2009).

7 OTHER ISSUES

Rewrite rules cannot capture *all* the aesthetic judgements that teachers would like to make: notational conventions which require consistency aid human recognition, as in $ax^2 + bx + c$ versus $Bx^2 + ax + \gamma$. Similarly in section 5.5, choosing a unit eigenvector, or an eigenvector with integer coefficients, is essentially an aesthetic judgement. A practical teacher may well choose problems which obviate the need for such decisions to be made, and this can be made relatively systematic (Steele, 2005).

Our goal in this paper, as in the previous, was to examine judgements made by teachers in terms of abstract rewrite rules. While it is not our goal here to design a complete system, we note that run-time complexity is normally not a major issue for the sort of problems that teachers *actually* pose to students.

A more relevant complexity question is how to describe the rules, and to provide a user interface which enables the teacher to assign each rule, or group of rules, to one of our three categories in a practical on-line assessment system.

8 CONCLUSIONS

We have seen that the methodology of Bradford *et al.* (2009) is useful in more general contexts than the calculus context in which it was introduced, where "the right answer" was a meaningful concept (even if not always as well-articulated as one would like). There are several topics in linear algebra where "the right answer" is not as well-defined, and in section 6 we have touched on some other questions where the same is true.

We regard the methodology of Bradford *et al.* (2009) as answering the third leg of our three-legged test:

- Is it **syntactically correct**, i.e., does it meet the requirements of the question posed;
- Is it **mathematically correct**, i.e., is it a correct answer to the question;
- Is it **pedagogically correct**, i.e., has the student completed the work (no venial rules need to be

[6] Of course, the teacher *may* have insisted on having the numerator degree strictly less than the denominator degree, in which case the third solution would fail the "syntactic" check.

applied) and not tried to answer the question with a re-statement (no fatal rules need to be applied).

We note that different questions have very different balances between the syntactic, mathematical and pedagogic aspects of their verification: in some the syntactic part is trivial, whereas for under-determined linear systems (and multiple eigenvectors) it is the most challenging component.

There is a real distinction, in our view, between the **syntactic** and **pedagogic** checks. Consider asking the student to triangularize $\begin{pmatrix} 1 & 5 & 3 \\ 2 & 8 & 6 \\ 1 & 3 & 2 \end{pmatrix}$. The "correct" answer is $A := \begin{pmatrix} 1 & 5 & 3 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, but the zero in position $A_{2,3}$ is fundamentally different from the others. They *have* to be zero, else the matrix is not syntactically upper triangular, whereas writing $A_{2,3}$ as, say, $6-6$, is a venial error of failing to carry through the arithmetic.

We cannot emphasise too strongly our view that the pedagogic context has to drive the classifications we have outlined above, and any marking penalties to be applied in the case of the venial rules. Indeed, the case of partial fractions shows that the precise definition adopted by the teacher has to control even the syntactic phase of answer checking.

REFERENCES

Barnard, T. (1999) A Pocket Map of Algebraic Manipulation, The Mathematical Association.

Bradford, R. J., Davenport, J. H. and Sangwin, C. J. (2009) A Comparison of Equality in Computer Algebra and Correctness in Mathematical Pedagogy, in Carette, J. *et al.* (eds.), *Proceedings of Intelligent Computer Mathematics 2009,* Springer Lecture Notes in Artificial Intelligence **5625**, Berlin–Heidelberg: Springer-Verlag, 75–89.

Brown, J. S. & Burton, R. R. (1978) Diagnostic models for procedural bugs in basic mathematical skills, *Cognitive Science* **2**, 155–192.

Carette, J. (2004) Understanding Expression Simplification, in Gutierrez, J. (ed.), *Proceedings ISSAC 2004*, New York, NY: ACM Press, 72–79.

Gage, M., Pizer, A. and Roth, V. (2002) WeBWorK: Generating, delivering, and checking math homework via the Internet, in *Proceedings ICTM2 international congress for teaching of mathematics at the undergraduate level.* http://www.math.uoc.gr/~ictm2/Proceedings/pap189.pdf.

Heeren, B. and Jeuring, J. (2009) Canonical Forms in Interactive Exercise Assistants, in Carette, J. *et al.* (eds.), *Proceedings of Intelligent Computer Mathematics 2009,* Springer Lecture Notes in Artificial Intelligence **5625**, Berlin–Heidelberg: Springer-Verlag, 325–340. http://people.cs.uu.nl/bastiaan/CanonicalForms.html

Lewis, J. and Tucker, A. (2009) Report of the AMS First-Year Task Force. *Notices AMS,* **56,** 754–760.

Miller, B.R. (1995) An Expression Formatter for Macsyma, http://math.nist.gov/~BMiller/computer-algebra/

Pointon, A. and Sangwin, C.J. (2003) An analysis of undergraduate core material in the light of hand held computer algebra systems, *International Journal of Mathematical Education in Science and Technology*, **34**(5), 671–686.

Sangwin, C. J. (2007) STACK: making many fine judgements rapidly, in *Proceedings CAME 2007.* http://web.mat.bham.ac.uk/C.J.Sangwin/Publications/2007CAME_Sangwin.pdf.

Sangwin, C.J. & Ramsden, P. (2007) Linear syntax for communicating elementary mathematics, *J. Symbolic Comp.,* **42,** 902–934.

Steele, J. D. (2005) Setting Linear Algebra Problems, at http://web.maths.unsw.edu.au/~jds/Papers/linalg.pdf.

Weibel, C. and Hirsch, L. (2002) WeBWorK Effectiveness in Rutgers Calculus, at http://www.math.rutgers.edu/~weibel/webwork.html.

BIOGRAPHICAL NOTES

Russell Bradford is a Senior Lecturer in the Department of Computer Science at the University of Bath, U.K. His research interests include the underpinnings of computer algebra systems, in particular ensuring they correctly simplify multivalued algebraic expressions.

James Davenport is Hebron & Medlock Professor in the Departments of Mathematical Sciences and Computer Science at the University of Bath, U.K. His research focuses on applying mathematics to computer science, or *vice versa*, notably in computer algebra, where he was the lead author of one of the first textbooks. He was a principal designer of the Axiom computer algebra system, the first to have a formal underpinning, in terms of Universal Algebra, for the mathematical operations performed.

Chris Sangwin is a Senior Lecturer in the School of Mathematics at the University of Birmingham, UK. His research interests include mathematics education, in particular using computer algebra systems to support online automatic assessment of mathematics.

© 2009 Research Information Ltd. All rights reserved.