CONTEMPORARY PROOFS FOR MATHEMATICS EDUCATION

FRANK QUINN

Abstract. In contemporary mathematical practice the primary importance of proof is the advantage it provides to users: proofs enable very high levels of reliability. This essay illustrates how a similar approach might have similar benefits in elementary education.

Contents

Introduction

Professional mathematical practice has changed over time to be better–adapted to the structure of the subject. In particular practices after about 1930 are significantly more effective than those of the nineteenth century and earlier, and the improvements played a large part in the explosive development in mathematics in the twentieth century.

Date: Version 0.1, December 2009.

Contemporary methodology is described in detail and depth in [7], with historical, sociological, and other background. Explicit comparisons of practices before and after the early twentieth–century transition are given in [13].

One conclusion of these studies is that mathematics education is still modeled on the primitive methodologies of the nineteenth century and before. This persistence seems to be due to historical and philosophical factors rather than intrinsic difficulty: the basic principles of contemporary methodology are actually simpler than the older versions and can be used effectively at any level. This essay illustrates elementary use of contemporary proofs. Elementary use of contemporary definitions is similarly illustrated in [12].

Contemporary proofs are first and foremost an enabling technology. Mathematical analysis can, in principle, give the right answer *every time*, but in practice people make errors. The proof process provides a way to minimize errors and locate and fix remaining ones, and thereby come closer to achieving the abstractly–possible reliability.

This view of proof is much more inclusive than traditional ones. “Show work”, for instance, is essentially the same as “give a proof”, while the annotations often associated with proofs appear here in “formal proofs”, §1.2, as aids rather than an essential part of the structure. I use the word “proof” systematically in this essay to emphasize the underlying commonalities, but synonyms such as “show work” are appropriate for use with students.

1. Proofs, Potential Proofs, and Formal Proofs

Too much emphasis on the correctness of proofs tends to obscure the features that help *achieve* correctness. Consequently I suggest that the key idea is actually “potential proof”, which does not require correctness. Variations are described in §1.1–1.2, and the role of correctness is described in §1.3. Some educational consequences are discussed in §1.4, *The Role of Diagnosis*; others occur later in the essay.

1.1. **Potential Proof.** A *potential* proof is a record of reasoning that uses reliable mathematical methods and is presented in enough detail to be checked for errors.

Potential proofs are defined in terms of what they *do* rather than what they *are*, and consequently are context–dependent. At lower levels, for instance, more detail is needed. Further, the objective is to enable individual users to get better results so even in a single class different students may need different versions. Commonalities and functionality are illustrated here, but individual needs must be borne in mind.

1.1.1. *Example, Integer Multipication I.* Multiply 24 and 47 using single–digit products.

Solution:

$$\begin{array}{cccc} & & 2 & 4 \\ & & 4 & 7 \\ \hline & & 1 & 1 \\ & 1 & 4 & \\ & 1 & 6 & \\ & 8 & & \\ \hline 1 & 1 & 1 & 1 \end{array}$$

This is a stylized format designed to efficiently support the algorithm rather than display mathematical structure (see §1.2.1 for an alternative). It is also not annotated so is not a *formal* proof in the sense of §1.2. Nonetheless it provides a clear record of the student's work that can be checked for errors, so it is a potential proof that the product is 1111.

1.1.2. *About the Example.* The example is not a proof because it contains an error. However:

- The error is localized and easily found. Ideally the student would find it during routine checking and fix it himself.
- The error is not random, and the mistake that produced the error can be diagnosed. A teacher could make a targeted suggestion of what to watch for in later work.
- in the worst case a conceptual confusion prevents the student from seeing this particular error. Again the diagnosis exposes the problem and suggests a targeted intervention to correct it.

Correcting the error produces a new outcome, and a better potential proof that this outcome is correct.

1.1.3. *About the Idea.* In the last decade I have spent hundreds of hours helping students with computer–based practice tests. In the great majority of cases they more-or-less understand how to approach the problem and have a record of the work they did on it, but something went wrong and they can't find the error. The goal is to diagnose the error, correct it, and perhaps look for changes in work habits that would avoid similar errors in the future.

Sometimes the student's work is easy to diagnose: intermediate steps are clearly and accurately recorded; the reasoning used in going from one to the next can be inferred without too much trouble; the methods used are known to be reliable; etc. In other words it is what is described here as a potential proof. In these cases the mistakes are often minor, and the student often catches them when rechecking. Sometimes I can suggest a change in procedure that will reduce the likelihood of similar mistakes in the future (see §2 on Proof Templates). The occasional conceptual confusions are well–localized and can usually be quickly set right.

In most cases my students' work does not constitute a potential proof. Problems include:

- Intermediate expressions are incomplete or unclear. For instance when simplifying a fragment of a long expression it is not necessary to copy the parts that do not change, but without some indication of what is going on it is hard to follow such steps and there are frequently errors in reassembly.
- Steps are out of order or the order is not indicated, for instance by numbering.
- Too many steps are skipped.
- The student is working "intuitively" by analogy with an example that does not apply.
- Notations used to formulate a problem (especially word problems) are not clear.

All of these increase the error rate and make finding errors difficult for either the student or a helper. If not corrected these work habits limit what the student will be able to accomplish.

The point here is that “potential proof” is to some extent an abstraction of the work habits of successful students. Exactly the same factors apply to the work of professional mathematicians, though this is obscured by technical difficulty and the fact that checking typically proceeds rapidly and almost automatically once a genuine potential proof is in hand.

1.2. **Formal Potential Proof.** A *formal* potential proof includes explicit explanation or justification of some of the steps.

The use of justifications is sometimes taken as part of the definition of proof. Here it appears as useful aid rather than a qualitatively different thing: the objective is still to make it possible to find errors, and formality helps with complicated problems and sneaky errors.

It seems to me that the best opportunities for formal proofs in elementary mathematics are in introducing and solidifying methods that in standard use will not need formality. This should improve elementary work as well as make the formal–proof method familiar and easily available when it is really needed.

1.2.1. *Example, Integer Multipication II.* Multiply 24 and 47 using single–digit products.

Solution:

Explanation	Result
write as polynomials in powers of 10	$(2\times 10^1+4\times 10^0)(4\times 10^1+7\times 10^0)$
set up blank form for output	$10^2(\qquad)+10^1(\qquad\qquad)+10^0(\qquad)$
enter products in the form, without processing	$10^2(2\times 4\;)+10^1(2\times 7+4\times 4)+10^0(4\times 7\;)$
compute coefficients	$10^2(8\quad)+10^1(30\qquad)+10^0(28\qquad)$
recombine as a single integer	$800+300+28=1128$

1.2.2. *Comments.* This uses a “structured” format for proof, see [4], [5]. I have not had enough experience to judge the benefits of a standardized structure.

The procedure follows the “template” for multiplication of polynomials described in §2.1, and see §3.1 for a version used to multiply large numbers.

Writing in expanded form with explanations clarifies the procedure. Once the procedure is mastered a short–form version can be used:

$$10^2(\underbrace{2\times 4}_{8}\qquad)+10^1(\underbrace{\underbrace{2\times 7}_{14}+\underbrace{4\times 4}_{16}}_{30}\;\;)+10^0(\underbrace{4\times 7}_{28}\qquad)$$
$$800+300+28=1128$$

In this form:

- The numbers are not rewritten explicitly as polynomials because the coefficients can be read directly from the decimal form. Some students may have to number the digits to do this reliably.
- The extra space in the outer parentheses after the powers of ten indicates that the blank template was set up first.

- The products for the coefficients were entered without on-the-fly arithmetic (the reason for this is explained in §2.1).
- Individual steps in the arithmetic are indicated, as is the final assembly.

The conclusion is that when the method is familiar this compressed notation provides an effective potential proof that the outcome is correct.

1.2.3. *Example, Solutions of Linear Systems.* For which values of a is the solution of the system *not* unique?

$$\begin{gathered} x + ay + 2z = -1 \\ 3y + az = 2 - a \\ 4x + y = 13 \end{gathered}$$

Solution:

The solution to a square linear system is not unique exactly when the determinant of the coefficient matrix is zero. The coefficient matrix here is

$$\begin{pmatrix} 1 & a & 2 \\ 0 & 3 & a \\ 4 & 1 & 0 \end{pmatrix}$$

Row operations $R_3 = R_3 - 4R_1$ and $R_3 = R_3 - \frac{1-4a}{3}R_2$ do not change the determinant and reduce this to a triangular matrix with $R_3 = (0, 0, -8 - a\frac{1-4a}{3})$. The determinant of a triangular matrix is the product of the diagonal entries, so the determinant is

$$(1)(3)(-8 - a\frac{1-4a}{3}) = -24 - a(1-4a) = 4a^2 - a - 24$$

This is zero for $a = (-1 \pm \sqrt{385})/8$.

1.2.4. *Comments.* This is a bit less detailed than the previous example because some calculations (effects of the row operations and application of the quadratic formula) are not recorded. Presumably they are on a separate paper, but because the operations themselves are recorded the calculations can be completely reconstructed. At this level the student should be able to reliably handle the hidden steps and this should be acceptable as a potential proof.

An alternative evaluation of the determinant might be: "Cramer's rule applied to the second row gives $(+1)(3)(-4 \times 2) + (-1)(a)(1 - 4a)\ldots$".

Cramer's rule involves adding up: a sign times the entry times the determinant of the matrix obtained by omitting the row and column containing the entry. The expression reflects this structure, with the 2×2 determinants evaluated. Giving relatively unprocessed expressions like this both reduces errors (by separating organization from calculation) and allows quick localization of them when they occur. For example it would be possible to distinguish between a sign error in the second term due to a misunderstanding of Cramer's rule from a sign error in the evaluation of the sub-determinant.

Students will not give this sort of explanation without examples to copy and quite a bit of guidance. This guidance might include:

- If you are using a theorem (e.g. nonzero det if and only if unique solutions) say enough about it to inspire confidence that you know a precise statement and are using it correctly. Confused statements indicate that conceptual errors are likely in the future even if this wasn't the problem in this case.

- In particular, mention of the theorem is an essential part of the work and must be included even in short–form versions. There is additional discussion of style in short–form proofs in [11].
- In lengthy calculations, rather than showing all details, describe the steps and carry out details on a separate sheet. The descriptions should be explicit enough to enable reconstruction of the details. Organizing work this way both reduces errors and makes it easier to check.

It can be helpful to have students check each others' work and give explicit feedback on how well the layout supports checking. The eventual goal is for them to diagnose their own work, and trying to make sense of others' can give insight into the process.

1.2.5. *Further Examples.* For further discussion and examples of elementary formal proofs concerning fractions and area see *Proof Projects for Teachers*, [11].

1.3. **Proof and Correctness.**

A *proof* is a potential proof that has been checked for errors and found to be error–free.

The intent is that work that does not qualify as a potential proof cannot be a proof even if the conclusion is known to be correct. In education the reason is simple: the goal is not a correct answer but to develop the ability to routinely get correct answers, and facility with potential proofs is the most effective way to do this. Too much focus on correctness may undercut development of this facility.

This is not an issue with weak students because potential proofs are an enabling technology and they cannot succeed without it. They tend to have the opposite problem: the routines are so comforting and the success so rewarding that it can be hard to get them to compress notation (e.g. avoid recopying) or omit minor details even when they have reached the point where it is safe to do so. Similarly some persist in writing out formal justifications even after they have thoroughly internalized the ideas.

Strong students are more problematic because the connection between good work habits and correct answers is less direct. I have had many students who were very successful in high–school AP courses but a focus on correctness rather than methodology let them get by with sloppy work. Most of these have trouble in engineering calculus.

- The better ones figure it out, especially with diagnostic support and good templates (§2). Most probably never fully catch up to where they might have been, but they are successful.
- Unfortunately there are a significant number who were good enough to wing it in high school, and good enough to have succeeded in college with good methodological preparation, but not good enough to recover from poor preparation.

The conclusion is that all students stand to benefit from a potential–proof oriented rather than a correctness–oriented curriculum, but for different reasons. Gains by weak and mid–range students are likely to be clearest.

1.4. **The Role of Diagnosis.**

The thesis of this article is that the reliability possible with mathematics can be realized by making mathematical arguments that can be checked for errors, *checking them*, and correcting any errors found. Other sections describe how checkable arguments could become a routine part of mathematics education. However they won't produce benefits unless *checking* also becomes a routine part. To be explicit: diagnosis and error correction should be key focuses in mathematics education.

- Answers are important mainly as proxies for the work done. Incorrect answers indicate a need for diagnosis and correction and ideally *every* problem with a wrong answer should be diagnosed and corrected.
- Mathematics uniquely enables quality so the emphasis should be on quality, not quantity. In other words, doing fewer problems to enable spending more time on getting them right is a good tradeoff.
- An important objective is to teach students to routinely diagnose their own work. The fact that diagnosis is possible and effective is the essence of mathematics, so teaching self-diagnosis is mathematics education in the purest sense.

Ideally teachers would regularly go through students' work with them so they can see the checking process in action. Students should be required to redo problems when the work is hard to check, not just when the answer is wrong. As explained in the previous section the goal is to establish work habits that will benefit students, but students respond to feedback from teachers, not long–term goals.

1.5. **Other Views of Proof.** There are many other views of the role of proof, c.f. [1], [14], [15]. These are quite different from the viewpoint here, and generally emphasize proofs as sources of understanding and insight or as repositories of knowledge.

The basic difference is that I have emphasized proofs as an enabling technology for users while the other views focus on "spectator proofs" that readers should benefit from but are not intended as templates for emulation. Both views are valid and important, but the reasons are very different and this should be kept in mind when considering specific situations.

What counts as user–oriented or spectator–oriented, and the mix in practice, varies enormously with level. At elementary levels—as illustrated here—almost everything is designed for emulation and spectator proofs play little or no role. Issues that might be addressed with spectator proofs (e.g. how do we know the multiplication algorithm really works?) are simply not addressed. At intermediate levels spectator proofs play a large role, and at the research frontier the primary focus is again on user–oriented work.

For example in [15] Thurston justified his failure to provide a proof of a major claim by observing that the technology needed for a good spectator proof was not yet available. This point found resonance among educators. However he was responding to criticism in [2] that he had failed to provide a user–oriented proof. These proofs may be technical, ugly, and nearly insight–free, but if they are complete they provide templates for emulation and a starting point for development of insight. Thurston's inability to provide a good spectator proof was not accepted in the research community as justifying the failure to provide a technical proof. The problem was declared unsolved, and eventually complete proofs were provided by others.

The point in [2], incidently, was that Thurston had made many spectacular and technically complete advances, but his refusal to withdraw a claim of full credit in this particular instance reduced his long–term influence in the field and promoted an unnecessarily negative view of his other work. We were trying to analyze what had gone wrong and suggest ways to avoid similar problems in the future.

2. Proof Templates

Students learn mainly by abstraction from examples and by imitating procedures. It is important, therefore, to carefully design examples and procedures to guide effective learning.

A "proof template" is a procedure for working a class of problems. Design considerations are:

- Procedures should clearly reflect the mathematical structures they exploit. This makes them more reliable and flexible, and often provides subliminal preparation for more complex work.
- Procedures should minimize problems with limitations of human cognitive abilities.
- Efficient short–form versions should also be provided.

Examples in this section explain and illustrate these points.

2.1. **Polynomial Multiplication.** [1]

2.1.1. *Problem.* Write $(3z^2 - z + 5a)(z^3 + (2-a)z^2 - a)$ as a polynomial in z. Show steps.

2.1.2. *Step 1: Organization.* There are three terms in each factor so there will be nine terms in the product. Some organizational care is needed to be sure to get them all. Further we would like to have them sorted according to exponent on z rather than producing them at random and then sorting as a separate step. To accomplish this we set up a blank form in which to enter the terms. A quick check of exponents shows that all exponents from 0 to 5 will occur, so the appropriate blank form is:

$$\begin{aligned} &z^5[\qquad\qquad\qquad\qquad] + z^4[\qquad\qquad\qquad\qquad] + z^3[\qquad\qquad\qquad\qquad]+ \\ &z^2[\qquad\qquad\qquad\qquad] + z^1[\qquad\qquad\qquad\qquad] + z^0[\qquad\qquad\qquad\qquad] \end{aligned}$$

Next scan through all possible combinations of terms, one from each factor. (Use a finger to mark your place in one term while scanning the other.) For each combination, write the product of coefficients in the row with the right total exponent. The result is:

$$\begin{aligned} &z^5[(3)(1)\qquad\qquad\qquad] + z^4[(3)(2-a) + (-1)(1)] + z^3[(-1)(2-a) + (5a)(1)]+ \\ &z^2[(3)(-a) + (5a)(2-a)] + z^1[(-1)(-a)\qquad\qquad] + z^0[(5a)(-a)\qquad\qquad] \end{aligned}$$

Note the products were recorded with *absolutely no* arithmetic, not even writing (3)(1) as 3. Reasons are:

- Organization and arithmetic are cognitively different activities. Switching back and fourth increases the error rate in both, with sign errors being particularly common.

---

[1] This material is adapted from the polynomial problem list in [9].

- This form can be diagnosed. We can count the terms to see that there are nine of them and the source of each term can be identified. The order of scanning can even be inferred, though it makes no difference.

Note also that every term is enclosed in parentheses. This is partly to avoid confusion because juxtaposition is being used to indicate multiplication. The main point, however, is to avoid thinking about whether or not it is necessary in each case. Again, this is cognitively different from the organizational task and may interfere with it.

2.1.3. *Step 2: Calculation.* Simplify the coefficient expressions to get the answer:

$$3z^5 + (5 - 3a)z^4 + (6a - 2)z^3 + (7a - 5a^2)z^2 + az + (-5a^2)$$

In this presentation the only written work is the organizational step and the answer. More–complicated coefficient expressions, or less–experienced students, would require recording some detail about the simplification process. A notation for this is shown in the arithmetic example in §1.2.1.

2.1.4. *Comments.*

(1) The separation of organization and computation makes this procedure reliable and relatively easy to use.
(2) The close connection to mathematical structure makes it flexible. It is easy to modify to handle questions like "find the coefficient on z^3" or "write a product involving both x and y as a two–variable polynomial".
(3) Variations provide methods for by–hand multiplication of integers, §1.2.1, and multiplication of large integers using a calculator, §3.1.
(4) If the baby version in §1.2.1 is used to multiply integers then students will find the polynomial version familiar and easy to master.
(5) Similarly students who work with polynomials this way will find some later procedures, e.g. products of sums that may not be polynomials, or iterated products like the binomial theorem, essentially familiar and easier to master.

This should be contrasted with the common practice of restricting to multiplication of binomials, with the "FOIL" mnemonic. This is poorly organized even for binomials, inflexible, and doesn't connect well even with larger products. In particular, students trained with FOIL are often unsuccessful with products like the one in the example.

2.2. **Solving Equations.** This is illustrated with a very simple problem so the structuring strategies will be clear.

2.2.1. *Problem.* Solve $5x - 2a = 3x - 7$ for x.

Annotated Solution:

Explanation	Result
Collect terms: move to other side by adding negatives	$5x - 3x = -7 + 2a$
calculate	$\underbrace{(5-3)}_{2}x = -7 + 2a$
move coefficient to other side by multiplying by inverse	$x = \frac{1}{2}(-7 + 2a)$

2.2.2. *Comments.* The primary goals in this format are efficiency and separation of different cognitive activities (organization and calculation).

The first step is organizational: we decide that we want all x terms on one side and all others on the other. Collecting x terms can be accomplished by adding $-3x$ to each side. However it is inefficient to do this as a separate calculation step because we know ahead of time what will happen on the right side: we have chosen the operation exactly to cancel the $3x$ term. Instead we think of it as a purely organizational step: "move $3x$ to the other side…". To keep it organizational we refrain from doing arithmetic (combining coefficients) and include "by adding negatives" to the mental description.

The second step is pure calculation.

The final step is again organizational, and the description is designed to emphasize the similarity to the first step.

Finally note that the steps are guided by pattern–matching: the given expression is manipulated to become more like the pattern $x =?$. See the next section for another example.

2.3. **Standardizing Quadratics.** This is essentially "completing the square" with a clear goal.

2.3.1. *Problem.* Find a linear change of variables $y = ax + b$ that transforms the quadratic $5x^2 - 6x + 21$ into a standard form $r(y^2 + s)$ with s one of $1, 0 - 1$, and give the standard form.

This is done in two steps each of which brings the expression closer to the desired form. A short–form version is given after the explanation.

2.3.2. *First Step.* Eliminate the first–order term with a change of the form $y_0 = x + t$.

Begin with $5y_0^2 = 5x^2 + 10tx + 5t^2$ to get an expression with the same second-order term as that of the given quadratic. To match the first-order term as well we need $10t = -6$, so $t = -3/5$ and $y_0 = x - 3/5$. Moving the constant term to the other side gives $5y_0^2 - 5t^2 = 5x^2 - 6x$. Use this to replace the first and second-order terms in the original to transform it to

$$5(y_0)^2 \underbrace{-5(-3/5)^2 + 21}_{-\frac{9}{5}+\frac{105}{5}=\frac{96}{5}} \tag{1}$$

2.3.3. *Second Step.* Factor out a *positive* number to make the constant term standard.

$$5y_0^2 + \frac{96}{5} = \frac{96}{5}(\underbrace{\frac{5^2}{96}y_0^2}_{(\frac{5}{\sqrt{96}}y_0)^2} +1) \tag{2}$$

The number factored out must be positive because we had to take the square root of it.

Comparing with the goal shows the standard form is $\frac{96}{5}(y^2+1)$ with $y = \frac{5}{\sqrt{96}}y_0 = \frac{5}{\sqrt{96}}(x - \frac{3}{5})$.

2.3.4. *Short Form.*

$$5(\underbrace{x+t}_{y_0})^2 = 5x^2 + \underbrace{10t}_{-6}\, x + 5t^2$$

So $t = -3/5$.

$$\begin{array}{l} \qquad\underbrace{5x^2-6x}\underbrace{+21} \\ 5y_0^2 \underbrace{-5(3/5)^2+21}_{\frac{96}{5}} \quad = \frac{96}{5}(\underbrace{\frac{5}{96}5y_0^2}_{(\frac{5}{\sqrt{96}}y_0)^2} +1) \end{array}$$

So $y = \frac{5}{\sqrt{96}}y_0 = \frac{5}{\sqrt{96}}(x - \frac{3}{5})$ and the form is $\frac{5}{96}(y^2+1)$.

Methods must be introduced with explanations but compression is necessary for routine use. It is important for teachers to provide a carefully–designed short format because the compressions student come up with on their own are rarely effective.

For example it is often necessary to simplify a fragment of an expression. The notation here indicates precisely which fragment is involved and connects it to the outcome. I have never seen a student do this. The fragments either appear without reference or the student rewrites the whole expression.

Experience often reveals errors that need to be headed off by the notation. In the work above the notation

$$5y_0^2 \underbrace{-5(3/5)^2+21}_{\frac{96}{5}}$$

clearly indicates that the sign on $-5(3/5)^2$ is part of the fragment being simplified. Many students seem to think of this sign as the connector between the sub–expression and the whole thing, and do not include it in the sub–expression. This is a very common source of errors, and would probably have resulted in an error in this case. Providing a clear notation and being consistent in examples will avoid such errors.

2.3.5. *Pattern Matching.* Routine success requires that at any point the student can figure out "what should I do next?" This procedure illustrates the use of pattern matching to guide the work.

There is a more routine way to solve this problem: plug $y = ax+b$ into the given quadratic, set it equal to $r(y^2+s)$, and solve for a, b, r, s. This can be simplified by doing it in two steps, as above, but even so it requires roughly twice as much calculation as the method given. This is a heavy price to pay for not having to think.

The approach above can be summarized as "what do we have to do to the given quadratic to get it to match the standard pattern?" In the first step we note that the given one has a first–order term and the pattern does not. We get closer to the pattern by eliminating this term, getting something of the form $Ay_0^2 + B$. If B is not 1, 0, or -1 we can get closer to the pattern by factoring something out to get $C(Dy_0^2 + s)$ with standard s. The only thing remaining to exactly match the pattern is to rewrite Dy_0^2 as a square, and whatever we get is the y we are seeking.

Pattern matching is a powerful technique, a highly–touted feature of computer algebra systems, and humans can be very good at it. Much of the work in a calculus course can be seen as pattern–matching. I believe students could use it more effectively if the idea was presented more explicitly.

2.4. Summary.

Carefully–designed procedures and templates for students to emulate can greatly improve success and extend the range of problems that can be attempted. Important factors are:

- Procedures should follow the underlying mathematical structure as closely as possible. This reveals connections, provides flexibility, and expands application. It also ensures upward–compatibility with later work, and frequently provides subliminal preparation for this work.
- Ideas that guide the work, pattern matching for example, should be abstracted and made as explicit as possible for the level.
- Procedures should separate different cognitive tasks. In particular organizational work should be kept separate from computation.
- Short–form formats should be provided that show the logical structure (i.e. is checkable) and encourage good work habits.

Good test design can also encourage good work habits. For example:

- Ask for a single coefficient from a good–sized product like the example in §2.1. This rewards understanding the organizational step well enough to pick out only the terms that are needed.
- A computer–based test with built–in functionality might ask for an algebraic expression that *evaluates* to give the coefficient[2]. The student could then enter the unevaluated output from the organizational step. This rewards a careful separation of organization and calculation by reducing the time required and reducing exposure to errors in computation.

3. LONG PROBLEMS

Current mathematics education is almost entirely concerned with short, routine problems. At higher levels the focus may shift to short tricky problems. However much of the power of mathematics comes from its success with long *routine* problems. Because the conclusions of each step can be made extremely reliable, many steps can be put together and the combination will still be reliable. Further, when the methods for dealing with short problems are carefully designed then they will apply to long problems as well.

I believe that long problems have an important place in elementary education. They give a glimpse into the larger world and illustrate the power of the methods being learned. They also reveal the need for care and accuracy with short problems. It is not clear how long problems might be incorporated into a curriculum, but group projects are a possibility and the examples here are presented this way.

The examples concern multiplication and addition of large integers (with calculators), and logic puzzles.

3.1. Big Multiplications.

The goal is to exactly multiply two large (say 14 or 15–digit) integers using ordinary calculators. This cannot be done directly so the plan is to break it into smaller pieces, e.g. 4–digit multiplications, that can be done on a calculator and then assemble the answer from these pieces. The method is the same

---

[2] Tests with this kind of functionality is one of the goals of the EduTEX project [10].

as the by–hand method for getting multi–digit products from single–digit ones, and we use a notation like that of §1.2.1 modeled on polynomial multiplication.

The number of digits in each piece depends on capability of the calculators being used. The product of two 4–digit numbers will generally have 8 digits. We will be adding a list of these, but no more than nine so the outcome will have 9 or fewer digits. Four–digit blocks will therefore work on calculators that can handle nine digits. Eight–digit calculators would require the use of three–digit blocks.

3.1.1. *Problem.* Multiply 638521988502216 and 483725147602252 using (9–digit) calculators by breaking them into 4–digit blocks.

3.1.2. *Step 1: Organize the Data.* Write the numbers as polynomials:

$$\begin{aligned} 638521988502216 &= 2216 + 8850x + 5219x^2 + 638x^3 \\ 483725147602252 &= 2252 + 4760x + 7251x^2 + 483x^3 \end{aligned}$$

where $x = 10^4$.

This notation should be used even with pre–algebra students because it is a powerful organizational aid. The exponent records the number of blocks of four zeros that follow these digits.

3.1.3. *Step 2: Organize the Product.* The product of two sums is gotten from all possible products using one piece from each term. Individual terms follow the rule $(ax^n)(bx^k) = (ab)x^{n+k}$ and we use this to organize the work. The product will have terms x^r for $r = 0, \ldots, 6$ and seven individuals or teams could work separately on these.

For instance the x^2 team would collect the pairs of terms whose exponents add to 2: x^0 (x not written) from the first number and x^2 from the second, then x^1 from the first and x^1 from the second, etc. They would record:

$$x^2(2216 \times 7251 + 8850 \times 4760 + 5219 \times 2252)$$

This is an organizational step and no arithmetic should be done. The students can infer how the pieces were obtained, and can double–check each other to see that nothing is out of place and no pieces were left out.

3.1.4. *Step 3: Compute the Coefficient.* Carry out the arithmetic indicated in the second step, using calculators. If the students can use a memory register to accumulate the sum of the successive products then the output is the answer, $x^2(69947404)$. If the multiplications and addition have to be done separately then the notation of §1.2.1 can be used:

$$x^2(\underbrace{(\underbrace{2216 \times 7251}_{16068216} + \underbrace{8850 \times 4760}_{42126000} + \underbrace{5219 \times 2252}_{11753188}}_{69947404})$$

Again different students or teams should double–check the outcomes.

3.1.5. *Step 4: Assemble the Answer.* At this point the group has found the product of polynomials,

$$4990432 + 30478360x + 69947404x^2 + 91520894x^3 + 45154399x^4 + 7146915x^5 + 308154x^6$$

and the next step is to evaluate at $x = 10^4$, or in elementary terms translate the powers of x back to blocks of zeros, and add the results. The next section gives a way to carry out the addition.

3.2. **Big Additions.** The goal is to add a list of large integers using ordinary calculators. This cannot be done directly so the plan is to break it into smaller pieces, e.g. 6–digit blocks, that can be done on a calculator and then assemble the answer from these pieces. The procedure is illustrated with the output from the previous section.

3.2.1. *Problem.* Use calculators to add $4990432+30478360\times 10^4+69947404\times 10^8+91520894\times 10^{12}+45154399\times 10^{16}+7146915\times 10^{20}+308154\times 10^{24}$ using 6–digit blocks.

3.2.2. *Step 1: Setup.*

			4	990432
			304783	600000
		6994	740400	
	91	520894		
	451543	990000		
714	691500			
308154				

Here we have written the seven numbers to be added in a column with aligned digits. Vertical lines are drawn to separate the six-digit blocks, and we omit blocks that consist entirely of zeros.

3.2.3. *Step 2: Add 6–digit Columns.*

			4	990432
			304783	600000
		6994	740400	
	91	520894		
	451543	990000		
714	691500			
308154				
308868	1	517888	1	590432
1	143134	1	045187	

Each column is added separately, for instance by five different students, and again the outcomes should be double–checked.

Most of the sums overflow into the next column and we have written the sums of the even–numbered columns one level lower to avoid overlaps. Note that since there are fewer than nine entries in each column the sum can overflow only into the first digit of the next column to the left.

3.2.4. *Step 3: Final Assembly.* Add the sums of the individual columns:

308868	1	517888	1	590432
1	143134	1	045187	
308869	143135	517889	045188	590432

The final addition is easy in this case because the overflow from one column only changes one digit in the next. This happens most of the time, and if examples are chosen at random it is very unlikely that students will see more than two digits change due to overflow.

Students should realize, however, that digits in sums are unstable in the sense that very rarely an overflow will change *everything* to the left. Teachers should ensure that students encounter such an example, or perhaps challenge them to contrive an example that makes the simple–minded approach crash. This illustrates the difference between extremely unlikely and mathematically impossible, and the "low–probability catastrophic failures[3]" that can occur when the difference is ignored.

3.3. **Digits in Big Products.** The goal here is to find a specific digit in a product of big numbers, and be sure it is correct. An attractive feature of the formulation is that careful reasoning and understanding of structure are rewarded by a reduction in computational work.

The (least thought)/(most work) approach is to compute the entire number and then throw away all but one of the digits. I give three variations with increasing sophistication and decreasing rote computation. In practice students (or groups) could be allowed to choose the approach that suits their comfort level. More–capable students will enjoy exploiting structure to achieve efficiency. Less–capable ones will be aware of the benefits of elaborate reasoning, but may see additional rote computation as a safer or more accessible alternative.

3.3.1. *Problem.* Find the eighteenth digit (from the right, i.e. in the 10^{17} place) in the product $52498019913177259058 \times 33208731911634712456$.

3.3.2. *Plan A.* We approach this as before, by breaking the numbers into 4–digit blocks and writing them as coefficients in a polynomial in powers of $x = 10^4$. These are 20–digit numbers so there are five 4–digit blocks and this gives polynomials of degree 4 (powers of x up to x^4). The product has terms up to degree 8.

The eighteenth digit is the second digit in the fifth 4–digit block ($18 = 4 \times 4 + 2$). When working with polynomials in $x = 10^4$ this means it will be determined by the terms of degree x^4 and lower (the coefficient on x^5 gets 20 zeros put after it, so cannot effect the 18th digit).

Plan A is to compute the polynomial coefficients up to x^4, combine as before to get a big number, and see what the 18th digit is. This gives a significant savings over computing the whole number because we don't find the $x^5 \dots x^8$ coefficients.

3.3.3. *Plan B.* This is a refinement of Plan A that reduces the work done on the x^4 coefficient.

We only need the 18th digit, so only need the second (from the right) digit in the coefficient on x^4. To get this we only need the product of the lowest two digits in each term. To make this explicit, the terms in the coefficient on x^4 are:

$$x^4(9058 \times 3320 + 7725 \times 8731 + 9131 \times 9116 + 8019 \times 3471 + 5249 \times 2456)$$

But we only need the next-to-last digit of this. If we write the first term as $(9000 + 58) \times (3300 + 20)$ then the big pieces don't effect the digit we want so it is sufficient just to compute 58×20.

This modification replaces the x^4 coefficient by

$$x^4(58 \times 20 + 25 \times 31 + 31 \times 16 + 19 \times 71 + 49 \times 56)$$

---

[3] A term from the computational software community where this is a serious problem.

Lower coefficients are computed and the results are combined to give a single number as before. This number will have the same lower 18 digits as the full product, and in particular will have the correct 18th digit.

3.3.4. *Plan C, Idea.* Plans A and B reduce work by not computing unneeded higher digits. Here we want to reduce work by not computing unneeded *lower* digits. This is trickier because it is not guaranteed to work and some careful estimation is needed to check this.

1. The coefficients in the product polynomial have at most nine digits (products of 4–digit numbers have at most 8 digits, and we are adding fewer than nine of these in each coefficient). The x^2 term therefore has at most $9 + 2 \times 4 = 17$ digits. This can effect the 18th digit only through addition overflow.
2. The plan, therefore, is to compute the coefficients on x^4 and x^3, combine these to get a number, and see how big a 17–digit number can be added before overflow changes the 18th. We will then have to estimate the x^2 and lower terms and compare this to the overflow threshold.
    - If the lower order terms cannot cause overflow into the 18th digit then the 18th digit is correct.
    - If lower terms might cause overflow then we will have to compute the x^2 coefficient exactly, combine with the part already calculated, and see what happens. In this case we will also have to check that the degree 0 and 1 terms cannot cause overflow that reaches all the way up to the 18th digit. This is extremely unlikely: these terms have at most $9 + 1 \times 4 = 13$ digits, so overflow to the 18th can only happen if the 14th through 17th digits are all 9.
    - In this unlikely worst–case scenario we will have to compute the lower–order terms too.

3.3.5. *Plan C, Setup and Compute.* The x^3 coefficient and Plan–B version of the x^4 coefficient are:

$$\begin{array}{ll} x^4 & (58 \times 20 + 25 \times 31 + 31 \times 16 + 19 \times 71 + 49 \times 56) \\ x^3 & (9058 \times 8731 + 7725 \times 9116 + 9131 \times 3471 + 8019 \times 2456) \end{array}$$

Computing gives $200894863x^3 + 6524x^4$. Substituting $x = 10^4$ gives

$$(200894863 + 65240000) \times 10^{12} = 266134863 \times 10^{12}.$$

The 18th digit (from the right) is 1. It is not yet certain, however, that this is the same as the digit in the full product.

3.3.6. *Plan C, Check for Overflow.* The 17th digit in 266134863×10^{12} is 3. If the top (i.e. 9th) digit in the x^2 coefficient is 5 or less then adding will not overflow to the 18th digit. ($3 + 5 = 8$, and overflow from the x^1 and x^0 terms can increase this by at most one).

The next step is to estimate the top digit in this coefficient.

1. The x^2 coefficient has three terms (from x^0x^2, x^1x^1, and x^1x^0).
2. Each term is a product of two 4-digit numbers so has at most 8 digits. In other words the contribution of each term is smaller than 10^9. Adding three such gives a total coefficient smaller than 3×10^9.

(3) When we substitute $x = 10^4$ we get a number less than 3×10^{17}. The top digit is therefore at most 2.
(4) Since the top digit of the lower–order term is smaller than the threshold for overflow ($2 \leq 5$) we conclude that the 18th digit found above is correct.

We were fortunate: if the 17th digit coming from the higher–order terms had been 7, 8, or 9 then we could not rule out overflow with this estimate.

The actual coefficient on x^2 is 131811939. Knowing this we see that a 17th digit 7 would not have caused an overflow while 9 would have increased the 18th by 1, and 8 is uncertain. This conclusion can be sharpened by using more digits: if digits 15–17th were 867 or less then there is no overflow, 869 or more then there is an overflow of 1, and a small interval remains uncertain. As noted above there are rare cases where the lower–order terms have to be computed completely to determine whether or not overflow occurs.

3.3.7. *Grand Challenge.* Use this method to find the 25th digit of the product of two fifty–digit numbers!

3.4. **Puzzles.** We will not explore it here but logic puzzles deserve mention as opportunities for mathematical thinking, see Wanko [16] and Lin [3]. These should incorporate an analog of proof: a record of moves that enables reconstruction of the reasoning and location of errors. The notation for recording chess moves (see the Wikipedia entry) may be a useful model.

A minor problem is that the rules of many puzzles are contrived to avoid the need for proof–like activity and should be de–contrived.

For example, the usual goal in Sudoku is to fill entries to satisfy certain conditions. The final state can be checked for correctness and—unless there is an error—would seem to render the record of moves irrelevant. A better goal is to find *all* solutions. If the record shows that every move is forced then the solution is unique. However if at some point no forced moves can be found and a guess is made, all branches must be followed. If a branch that leads to an error, that branch can be discarded (proof by contradiction). If a branch leads to a solution then other branches still have to be explored to determine whether they also lead to solutions. This would be made more interesting by a source of Sudoku puzzles with multiple solutions.

Notations and proof also enable collaborative activity. All members of a group would be given a copy of the puzzle, and one appointed “editor”. Whenever a member finds a move he would send the notation to the editor as a text message. The editor would check for correctness and then forward the move to the rest of the group. Maintaining group engagement might require a rule like: someone who submits a move must wait for someone else to send one before he can send another.

4. Word Problems and Applications

This essay concerns the use of contemporary mathematical methodology in education. Up to this point the ideas have been unconventional and possibly uncomfortable but more-or-less compatible with current educational philosophy.

There are, however, genuine conflicts where both contemporary methodology and direct experience suggest that educational practices are counterproductive, not just inefficient. Some of the methodological conflicts are discussed in this section. A more systematic comparison is given in [13], and conflicts in concept formation are discussed in [12]. Historical analysis in [7] indicates that many educational practices are modeled on old professional practices that were subsequently found to be ineffective and discarded from professional use.

4.1. **Word Problems and Physical–World Applications.** The old view was that mathematics is an abstraction of patterns seen in the physical world, and there is no sharp division between the two. The contemporary view is that there is a profound difference, and the articulation between the two worlds is a key issue. The general situation is described in [7]; here I focus on education.

4.1.1. *Mathematical Models.* In the contemporary approach, physical–world phenomena are approached indirectly: a *mathematical model* of the phenomenon is developed, and then analyzed mathematically. The relationship between the phenomenon and the model is not mathematical, and is not accessible to mathematical analysis.

4.1.2. *Example.* A beaker holds 100 cc. of water. If 1 cc of X is added, what is the volume of the result?

Expected solution: $100 + 1 = 101$ cc.

4.1.3. *Discussion.* The standard expected solution suppresses the modeling step. Including it gives:

Model: volumes add.

Analysis: $100 + 1 = 101$ so the model predicts volume 101 cc.

The analysis of the model is certainly correct so correctly predicts the outcome when the model applies, eg. if X is water. If X is sand, salt, or alcohol then the volume will be more than 100 cc. but significantly less than the predicted value of 101 cc. If X is metallic sodium a violent reaction takes place. If the beaker is still intact when the smoke clears it will probably contain considerably less than 100 cc.

In the latter instances the prediction fails because the model is not appropriate. This is not a mathematical difficulty, and in particular no amount of checking the written work can reveal an error that accounts for the failure. One might try to avoid the problem in this case by specifying that X should be water, but discrepancies could result from differences in temperature. Even elaborately legalistic descriptions of the physical circumstances cannot completely rule out reality/model disconnects.

The conclusion is that the reality/model part of real–world applications is essentially non–mathematical. I believe that applications have an important place in mathematics courses, but this aspect should not be represented as mathematics.

Another important point is that modeling and analysis of the model are different cognitive activities. They should be separated, and the model formulated as an intermediate step, because mixing the two increases error rates just as with organization and calculation in §2.1. This may not be necessary when the mathematical component is trivial, as is often the case, but it grows rapidly in importance as complexity grows.

4.2. **Applications.** Mathematics is brought to life through applications. In this context the word "application" is usually understood to mean "physical–world application", but in fact these do a poor job of bringing elementary mathematics to life, quite apart from the modeling problems described above. After explaining why, I suggest that there are better opportunities using applications within mathematics.

4.2.1. *Difficulties with the Real World.* The main difficulty with physical–world applications is a complexity mismatch:

1. There are impressive applications of elementary mathematics but these usually require significant preparation in other subjects. For example:
    - One can do interesting chemistry with a little linear algebra, but the model–building requires a solid grasp of atomic numbers, bonding patterns, etc. The preparation required is probably beyond most high–school chemistry courses and certainly beyond what one could do in a mathematics course.
    - There are nice applications of trig functions to oscillation and resonance in mechanical systems, electric circuits, and acoustics, but again subject knowledge requirements makes these a stretch even in college differential equations courses.
    - Multiplication of big integers, as in §3.1, plays an important role in cryptography, but it is not feasible to develop the subject enough to support cryptographic "word problems".
2. Coming from the other side, there are easily–modeled real–world problems but the mathematical analysis tends to be either completely trivial, or highly nontrivial.
    - Some are best seen as calculus problems though they are often too hard for educational use there as well.
    - Special cases may have non–calculus solutions but these tend to be tricky and can hardly be said to give insight into the problem.
    - In another direction, our world is at least three–dimensional and many real problems require vectors in all but the most artificially contrived and physically–boring cases.

The conclusion is that real–world problems may have to be part of a serious development of a scientific subject to be genuinely useful. The difficulties encountered when this is not possible are considered next.

4.2.2. *Bad Problems.* The practical outcome of the complexity mismatch described above is that most word problems—in the US anyway—have trivial or very constrained mathematical component and the main task is formulation of the model. The example in §4.1.1 above illustrates this.

I have heard of elementary–education programs that exploit this triviality with a "keyword" approach: "When a problem has two numbers then the possibilities are multiplication, division, addition or subtraction. Addition is indicated by words 'added', 'increased by' . . . ". The calculator version is even more mindless because the operations have become keystrokes rather than internalized structures that might connect to the problem: "Press the "+" key if you see 'added', 'increased by', ..."

The higher–level version of this can be thought of as “reverse engineering”. There are only a few techniques in play: look for keywords or other commonalities to figure out which method to use and where to put the numbers.

Other problem types amount to translating jargon: replace “velocity” with “derivative”, “acceleration” with “second derivative”, ...

- In other words, there is so little serious contact with any real–world subject that translation and reverse–engineering approaches that *avoid* engagement are routinely successful as well as being fast and reliable. Students who master this skill may like word problems because the trivial math core makes success easy.
- The errors I see make more sense as translation problems than conceptual problems. A common example is that when modeling the liquid in a container, liquid flowing *out* acquires a negative sign because it is being lost from the system. Translators miss the sign, students who actually envision the situation should not.
- Some of my students despise word problems, regarding them as easily–solved math problems made hard by a smokescreen of terminology and irrelevant material. These students may be weak at this cognitive skill, or they may be thinking too much and trying to engage the subject. In any case, like it or not, the only effective help I can offer is to show them how to think of it as an intelligence–free translation problem.
- Finally, many problems are so obviously contrived that they cannot be taken seriously. The one that begins “if a train leaves Chicago at 2:00. . . ” has been the butt of jokes in comic strips.

It is frequently said that word problems engage students and provide an important connection to real–world experience. This is abstractly attractive, but for the reasons described above I don’t believe it is effective in practice. Further, a curriculum justified by, or oriented toward, word problems is likely to be weak because weak development is good enough for immediately–accessible problems.

4.3. **Mathematical Applications.** A common justification for word problems is that mathematics is important primarily for its applications, and math without applications is a meaningless formal game. I might agree, with the following reservations:

- Goals should include preparation for applications that will not be accessible for years, not just those that are immediately accessible.
- “Application” should be interpreted to include applications *in mathematics* as well as real–world topics.

The examples to have in mind for this last point are the application of polynomial multiplication to multiplication of big integers in §3.1, and the refinements developed in §3.3 to minimize the computation required to find individual digits.

- These topics clearly have genuine substance and support extended development.
- Unlike real–world topics they are directly accessible because they concern mathematical structure that has already been extensively developed.
- The multiplication algorithm does have real–world applications, even if they are not accessible to students, and in any case is a good illustration of the kind of mathematical development that has applications.

- The single–digit refinement is a very good illustration of a major activity in computational science: carefully exploiting structure to minimize the computation required to get a result.
- Version C provides an introduction to numerical instability and "low–probability catastrophic failure" of algorithms. This is a major issue in approximate (decimal) computation but is completely ignored in education.
- These projects would also significantly deepen understanding of the underlying mathematical structure and develop mathematical intuition.

The main objection to mathematical applications is that, because they lack contact with real–world experience, they do not engage students. I believe this underestimates the willingness of children to engage with almost anything if they can be successful with it. Further, the more obviously nontrivial the material the more pride and excitement they get from successful engagement. Success is the key, and the key to success is methods and templates carefully designed to minimize errors.

References

[1] Hanna, G.; Barbeau, E. *Proofs as bearers of mathematical knowledge*, Mathematics Education **40** (2008) pp. 345–353.
[2] Jaffe, A.; Quinn, F. *Theoretical Mathematics: Towards a synthesis of mathematics and theoretical physics*, Bulletin of the American Mathematical Society, (1993)
[3] Lin, Chun-Chi *How can the game of hex be used to inspire students in learning mathematical reasoning?*, Proceedings of ICMI Study 19 (2009).
[4] Mannila, Linda; Wallin, Solveig, *Promoting students' justification skills using structured derivations*, Proceedings of ICMI Study 19 (2009).
[5] Peltomaki, Mia; Back, Ralph-Johan, *An Empirical Evaluation of Structured derivations in high school mathematics*, Proceedings of ICMI Study 19 (2009).
[6] Quinn, Frank *Education web page* http://www.math.vt.edu/people/quinn/education/
[7] ________ *The Nature of Contemporary Core Mathematics*, Draft December 2009, retrieve from [6].
[8] ________ *Beneficial High–Stakes Tests: An Example*, preprint; retrieve from [6].
[9] ________ *AMS Working Group on Preparation for College Mathematics* web site, http://amstechnicalcareers.wikidot.com
[10] ________ *The* EduTEX *project*, Wiki at http://www.edutex.tug.org.
[11] ________ *Proof Projects for Teachers*, preprint June 2009; retrieve from [6].
[12] ________ *Contemporary Definitions for Mathematics Education*, preprint January 2010; retrieve from [6].
[13] ________ *Mathematics Education is Poorly Adapted to Mathematics*, January 2010; retrieve from [6].
[14] Rav, Y. *Why do we prove theorems?*, Philosophia Mathematica **7**(3), (1999) pp. 5–41.
[15] Thurston, W. *On proof and progress in mathematics* Bull. Amer. Math. Soc. **30**, (1994) pp. 161–177.
[16] Wanko, Jeffrey J. *talking points: experiencing deductive reasoning through puzzle discussions*, Proceedings of ICMI Study 19 (2009).