A true calculator

Harold Thimbleby

Existing calculators are unreliable and difficult to use, and they are marketed misleadingly. A new calculator is described that is simple, reliable and very powerful.

Originally published in *Engineering Science and Education Journal*, **6**(3):128–136, 1997.

CALCULATORS are used in all aspects of work, from laboratories to banks, as well as in everyday life, from paying loans to passing exams. Handheld electronic calculators seem impressive, and they have been around long enough for us to take them for granted. They are ubiquitous: they can be found in wrist watches or in rulers; they may be credit-card sized, pocket sized, or 'desk top' with their own printer. Calculators are required in school curriculums, and no office would be complete without one. They are probably the most commonly used and experienced of all modern 'information technology' electronic gadgets.

Calculators are also at the centre of controversy. Our children are not good at mathematics, and calculators get some of the blame. Is this because calculators stop children thinking about numbers, so they never learn any insight into mathematics?

This article suggests that things are worse: the maths that calculators do is damaged, and learning to use a calculator is really more about learning to cope with unnecessarily bad technology than about learning about numbers.

The first article on the front page article of *The Times* (December 7, 1995) entitled, "Calculators banned in exams after maths standards fall," claimed that School Curriculum and Assessment Authority officials were worried about childrens' falling standards in decimal places, fractions and "lack of basic understanding of percentage." Ironically these problems children have with maths are identical to problems calculators themselves have, as we shall see. Maybe the problems are not calculators *per se*, but with the way they are conventionally designed? This article will not only criticise current calculators, but suggest a workable alternative.

Background

After the abacus, more-or-less unchanged through antiquity, the first main development in devices for making calculation easier was John Napier's invention of logarithms around 1594. (He was also one of the first people to make systematic use of the decimal point.) Napier later published a book, *Study of Divining Rods*, concerning his 'bones,' which were sticks for helping do multiplication and division. Napier's bones were successful and there have been various (what would now be called) human factors improvements to them. In 1622 William Oughtred produced the slide rule, using the principles of logarithms: less accurate but far more convenient than tables.

Around 1642, as a teenager, Blaise Pascal made an ingenious mechanical adding machine that relied on rotating gears and gravity to perform carries. After seeing one of Pascal's calculators, Gottfried von Leibnitz tried for twenty years to make a machine that could multiply by repeated addition. When he finally succeeded his invention formed the heart of many calculators for the next 300 years.

The rapid innovation in earlier centuries has been eclipsed by the developments of the present century. Electronics can of course process information far faster and more conveniently than any mechanical arrangement of bones or gears. By the 1950s electronics was being widely used in calculators. Since then things have got smaller, faster and more reliable — but not easier to use.

In this paper, I will present an alternative approach to calculator design that overcomes the problems of conventional calculators (which we will discuss). Though much easier to use, the new design is technically comparable to existing calculators, so there are no electronic or mechanical hurdles to manufacture it. It proves that the problems we shall identify in other calculators are avoidable.

Modern technology usually looks very different to early technology. Thus today's cars and aeroplanes are very different from the first working examples. Today's calculators, albeit smaller and more feature-laden, are still used in the same old ways despite the huge advances in microelectronics and in computer programming that have occurred over the last three decades. We do not expect to drive a modern car with a tiller or any other vintage feature; so why should we expect to 'drive' modern calculators the same way as museum pieces? Likewise, our new design is different.

Today's calculators are bad

Clearly calculators are ubiquitous, widely used and apparently extremely successful. And so they must surely be good! We are bombarded with adverts and claims that suggest they are easy and powerful. The following, is part of an advertisment by a market leader, Casio:

> "Thanks to V.P.A.M. (Visually Perfect Algebraic Method) calculations have become a lot easier. It may sound confusing, but all it means is that the scientific calculators in the Casio VPAM range perform calculations exactly as you would write them or read them in a text book.
>
> For example, using the FX115s illustrated here, you would enter the calculation 5×2+sin30 = 10.5 exactly as you would write it."*

It is a minor quibble to mention that you cannot enter the 10.5 *exactly* as suggested (because it is the answer); however, more seriously, the fx-115s

---

* Casio Electronics Company Ltd., "For Complicated Calculations. Uncomplicated Calculators." *The Times* , Sept. 7, p13, 1994.

cannot do either 5+2sin30 or 5×2+sin–30. Both of these are surely "calculations exactly as you would write them or read them in a text book."

How else would you read or write 5×2+sin–30 in a text book? The fx-115s gets –20 using the – key, and either 9.5 or 10.5 when using its ± key; and *in no case is an error reported*. The only case where it does get the right answer is nothing like anything I have seen in a textbook. Certainly, in my view, the calculator does not merit being called "visually perfect algebraic method"! Implicit multiplication isn't understood, and negative numbers aren't understood.

We might quibble that you needn't expect a calculator to do 2sin30, but to imply in advertising it can when it can't is deceitful. There are calculators on the market in the same price range that can do both of these problems, in exactly the form written above.

The advertisement is not specific about the calculator's complex number abilities, but it is clearly a "complex" number calculator that "contains all the features needed for study up to and including key stage 4." It can do some complex arithmetic, like squaring *i*, but it cannot do what I would have thought essential complex arithmetic like taking the square root of minus one.

Hewlett Packard describe their HP 20S as "meeting the high quality standards that [their] customers have come to expect" and as having an

> "**Easy-to-use algebraic entry system**. Enter data in algebraic notation, just as you would on paper."

Surprisingly, then, it cannot handle negative numbers "just as you would on paper." Instead of writing –100 in a sum, you often have to use 100±. In fact, this unnecessary convention has become so commonplace with calculators many people accept it as inevitable! There are plenty of 'simple' calculators that do not have a ± key on them, and dealing with negative numbers on them is quite a challenge for anyone!

We've only given a few examples so far. This paper will show, more generally, all conventional handheld calculators are surprisingly bad. We shall also discuss at greater length whether the technical limitations of calculators we identify amount to no more than quibbles in their design, or whether it is a more serious problem.

Errors and idiosyncrasies in conventional calculators

Try the sum "4 times minus 5 equals" on almost any calculator[1,2,4,5,6,7,8,9,11] and you get the result –1.* The calculators ignore the multiply and do a subtraction instead, though some can report an error.[2] Some, but not all, calculators provide two minus signs[3] or a change sign key[4] to work around this problem, introducing unusual conventions. That leaves a number of calculators[1] that, despite apparently having all the necessary keys, are *incapable* of performing this elementary sum.

---

* The IEE's own promotional calculator shows 1– or 1M if there is a number in the memory.

The change sign key changes the sign of the last number, so 1–5= gives the same result as 1+5±= (that is, 1 plus *minus* 5). On one calculator[11] the % key changes the change sign rule: 1–5% gets –80 but 1+5±% gets 0.95. In other words, the 5± followed by % does not mean –5%. On another calculator, [4] we find that, though 3.14± is minus 3.14, using the calculator's constant π produces a different result. It seems that π± is treated like an implicit multiplication: so π±1 is treated as π×–1.

The sum 1+5% gives different answers on different calculators made by the same manufacturer.[1,3,5,9] One of these[9] is a financial calculator that can display results shown to two digits (useful for pounds and pence, dollars and cents); in this mode, 1+5% is displayed as 1.05, apparently 1 plus 5% of 1, but in fact it has worked out the number 1.0526315, which is the number 5% less of which is 1 — except that it is rounded incorrectly (the correct value is 1.0526316 to this number of digits). The table below summarises:

Model	4×–5	1–5%	1+5%	
1	–1	–80	1.0526315	
3	–20	–80	120	
5	–1	0.95	1.05	
9	–1	–80	1.05	*in 2 digit mode*

One calculator[12] gets –1, 5 (or 6), 5 (or –4) respectively for these sums in this table — when it is in 'complex' mode. In complex mode on this calculator, % just stops working, and *there is no error indication*. (There is no technical reason why % should not work with complex numbers. If the manufacturers decide that % should not work, there is no reason why the calculator should not indicate an error. As it is, you simply get the wrong answer *with no warning*.) Moreover, the manual gives a specific list of functions that only work with real numbers (i.e., a list of functions that do not work with complex numbers): and % is *not* mentioned in this list.

It seems that calculator designers, not just school children, don't understand percentage.

The display 1,234 means either 1.234 or 1234, a problem that would have been avoided if the calculator[6] adhered to the long-established international standard, SI. The comma/dot mode seems to me like a feature designed to make the calculator attractive rather than more reliable, and in this case the advantages of having the choice seem to be outweighed by the potential problems. Many calculators provide degree, radian and grade modes. There are 400 grades in 360 degrees, so the user — typically a school child — risks mixing up degrees and grades, since many answers will be too close to easily distinguish. Again it looks more like a feature for feature's sake rather than a feature to make the calculator more practical.

A number well within the range of a calculator (as *defined* by the manufacturers in its manual) but entered with more digits than the display (for example, 10^{11}, entered as 1 and eleven zeroes) goes wrong

without warning.[7] Or calculate 254!, blink, and 10^{500}, a wrong answer appears.[6]

Many calculators have buttons to perform calculations that cannot be used in any obvious way for certain calculations. For example $100 \div \binom{10}{5}$ (i.e., 100 divided by 10 choose 5, the number of ways of taking things 5 at a time from 10) fails because division and combinations *both* use the 'swap' key and therefore *cannot* work together.[7] (When writing this article I decided to double-check how to do combinations on this calculator. I was not able to work out what to do without the manual.) Yet there are calculators[4] where the combinations button works as a simple operator, like + or ×. These calculators can be used more reliably without continual reference to the manual.

The key sequence[4] that correctly calculates $2^{-3.14}$ fails when π (using the special π key) is substituted for the exponent to calculate $2^{-\pi}$. Ironically, these calculators[4,7] specifically claim to be 'algebraic.' Some calculators[4] provide implicit multiplication, so 2π can be entered directly, rather than 2×π. But 2RANDOM won't work, nor π2.

All calculators have obscure limitations, such as[7] providing only three memory registers using the first three digits (0, 1, 2) despite — *obviously*! — having ten digit keys. Or there are calculators where memories are corrupted by certain sorts of calculation.[6] To put the current number in the memory requires a precise four key incantation[1] — rather defeating the object of having a memory if a sheet of paper is needed to work out how to use it.

On calculators[6,12] that 'support' complex numbers, you can add $i+i$ but not calculate i^i which should be about 0.2*. Even more bizarre[12] is that you cannot work out $\sqrt{-1}$, that is i itself, yet you *can* do i^2, the converse operation. Although you *can* use the 'x^2' button to square i (i.e., press i then x^2), the calculator *cannot* do i^2 should we use the 'x^y' button (i.e., press 'i', 'x^y', then 2). Such rules, what is available, what is not, are mysterious.

It is perhaps no surprise that these rules for what is and what is not possible with complex numbers are not correctly reported in the manuals, which claim that *neither* square root nor squaring work, though squaring

---

* In fact, $i^i = \frac{1}{\sqrt{e^\pi}} \approx 0.20788$, which is a real number. Since this cannot be worked out on a calculator, even ones that claim to handle complex numbers, some of the interest of experimenting with complex numbers is denied — here, we have a complex power of a complex number being a real number. In fact, just working out the value of $\frac{1}{\sqrt{e^\pi}}$ (supposing you know this is the value required) is challenge enough on most calculators.

does work. (We already mentioned the confusion with complex numbers and % above.)

I think it must be difficult to write a good, clear manual for any such calculator, and indeed we do find that calculator manuals are very badly written, obscure and under-specified. (What manual specifies exactly what % does? What manual specifies the range and distribution of the random numbers?) Some manuals have important sections written in alternate Spanish and English sentences.[8,12]

One calculator[10] has a manual and other promotional material that is illustrated with problems it cannot solve: the display shown in one picture is a digit longer than the calculator can actually handle. Many operations cannot safely be performed without reference to a manual (e.g., combinations,[7] complex numbers[6,12]).

The problems of calculators do not arise because of some sort of limitations in electronics or what can be squeezed into a small hand held package. One calculator can solve differential equations and plot graphs:[13] but it still needs two buttons for minus signs, using one for subtract and one for negate. Unnecessarily, you have to push buttons 4×(-)5 ENTER if you want to know 4×–5. (See Thimbleby, 1996, for a review of this calculator.)

Are there *really* problems?

The previous section briefly showed that calculators suffer from many specific problems. I believe it is unimaginative use of language merely just to say they are appalling and egregious. There is almost a conspiracy of approval, perhaps exploiting most people's innate fear of mathematics, and resulting in most people being made to feel incompetent. They are a case study in the 'woes' of technology that Weiner (1993) covers.

Beyond the specific problems (only some of which were listed above), which obviously differ from manufacturer to manufacturer, and even from model to model, calculators of all sorts uniformly suffer from several common problems. These common problems of conventional calculators include:

- Many operations on calculators are undefined or inconsistent. (Try the simple sum 4×–5: you should get –20, but you are more likely to get –1; or try 1+5% on several calculators — even from the same manufacturer — and you'll get various answers, not all of them can be right!)

- You have to rearrange any problem so the answer comes last, to get it with a final hit of the = button. (To see, try finding the power of two that is 100.) Sometimes the rearrangement is so complex as to pose its own problems for entry into the calculator, even for seemingly innocuous calculations like $2^{-\pi}$.

- You can't reuse a calculation (except in some special cases). Experimenting with numbers is very difficult (yet essential to help check answers).

- Key-pressing mistakes are very difficult, often impossible, to correct without starting over. Some calculators provide keys that sometimes work (e.g., the delete digits, but not in an exponent, and anyway don't delete operators like +) — as if proving that the manufacturer has noticed the problem but is not prepared to really fix it. On no calculator is there an 'undo' button.

- You have to remember what you have done and how this relates to what the calculator is doing. If a calculator shows 457 on its display: what would it show when you press a key? You can't say unless you can remember *why* it shows 457.

- You generally can't find out what is in memory registers without spoiling the current calculation. Some calculators don't let you assign numbers to memories without going to contortions or first clearling the entire calculator (which loses the very numbers you want to store).

- The problems are compounded by manufacturer's hyperbole, which misleads users. If taken in, users will have to blame themselves when they cannot use calculators claiming to work "just as you would on paper."

We will shortly show how to solve these problems and do new and interesting things very naturally.

But it is important to be clear that there ***really*** *are problems rather than just technical inconsistencies or limitations.*

People usually deny there are problems with calculators.

Ask anyone what the % key does on their calculator, and they are likely to say they don't *need* it. I suspect that they decided they didn't need it sometime after they discovered that it didn't make much sense to them. (The previous section gave some examples that seem to indicate % doesn't make sense to manufacturers either.)

Many people say, "Even if your calculator cannot do $2^{-\pi}$ correctly, mine can, so what's the problem? Get a real calculator!" Surely, if some calculator can do $2^{-\pi}$ correctly, then the manufacturers selling ones that cannot — but which pretend to — are negligent, even exploitative? Or, taking a cynical view: if there is a calculator that cannot do $2^{-\pi}$ (there is), but yours can (you say), what else may be wrong with yours?

Another response is, "Who wants to do $2^{-\pi}$ anyway?" This question is misguided. If I found that a car could not operate on the B1000 road, nobody would ask, "Who wants to drive on the B1000?" Clearly, the B1000 is a "typical" road, and if a car can't run on it, it isn't going to be reliable on many other roads. The calculation $2^{-\pi}$ is a typical calculation in the way the B1000 is a typical road.

Some people have responded to the problems by saying that you could of course use a mathematics program (of which many are available) running on a PC. This is true, but it invokes the fallacy that bad technology

is acceptable when there is *another* technology that solves the problems. The 'solution' is more expensive, and perhaps neither available nor appropriate to the people who suffer with the bad technology. Moreover, such a view does not help anyone believing they have a handheld calculator they can rely on.

One argument I have had deserves comment because it came from a distinguished mathematician. He said the problems were very well known, and that it was a truism that a calculator should be used with care else you'd get bad results. If indeed the problems are well known, then why not do something about them? Certainly calculators do need to be used with care, but, more so, they ought to be made with integrity.

Finally, a sophisticated denial of the problems is to request empirical study of the seriousness, or otherwise, of them. This line of argument assumes that if most people can cope most of the time, then what is the problem? Well, suppose a study was done, and we discovered that 99.99% of the population have no problems. The argument would be like saying that most aeroplanes don't crash, so we don't need to worry about the ones that do. Furthermore, this sort of argument confuses people's problems (which are real, but only *sometimes* due to their difficulty with mathematics) with the technology's problems (also real, but avoidable).

So we do need to worry, and if the problems can be fixed (we argue they can) and can be fixed at more-or-less no cost, why wait until an accident? There are probably many cases of people making mistakes with calculators, but if they knew they had made a mistake, they may not have been using the calculator in the first place! Or they'd assume they did not understand how to use the calculator properly, when in fact most calculators are difficult to understand however clever you are. So we would naturally expect under-reporting of errors with calculators.

Buying status?

It seems that *saying* a calculator does complex numbers helps sell it, and that manufacturers have discovered that so few people understand or actually want complex numbers and all those functions, nobody will notice if they hardly work. In support of this hypothesis, I note that the Casio fx-115s has a lid that serves — or rather might have served — as a quick reference card, but apparently is more concerned with reminding the owner that his or her calculator has got 252 functions (wow!) and has an auto power-off feature. It does not provide any serious help on how to actually *use* it. Some calculators have several thousand functions (e.g., the HP 48GX has 2300).

Obviously calculators have found an ecological niche they can exploit by evolving into ever more complex and obscuritanist devices. They are not calculators but status symbols. It would appear that market forces, plus consumer acquiesence, leads to ever more sophisticated looking products, with less and less genuine sophistication. Worse, calculators are a symptom of a larger problem, as Weiner's *Digital Woes* book (1993) makes clear.

Background to a new calculator

The main innovation of the new calculator is that it has a "declarative user interface," in contrast to the conventional "imperative" approach. We do not expand on these ideas directly here, as they are very well known in the computing literature. Arguably the inconsistencies with conventional calculators arise because their designers did not have a simple consistent model of what they were supposed to do. They look like they grow functions year by year, and that the functions are grafted on partially and without regard for the overall coherence of the design. In contrast, a declarative calculator (along the lines we propose) is easier to design coherently, because its conceptual model is much simpler — and familiar.

Computer scientists should be interested in this paper because it shows what is probably the first non-trivial declarative user interface, and one that is arguably superior to the imperative approach. See Thimbleby (1996) for a technical review.

The new calculator

It is often easiest to explain new ideas by a metaphor. The metaphor of the new calculator is *a blackboard*. The board is shared between the user and the calculator. The calculator, like an eager and helpful teacher, ensures that *anything* written is true. The corrections or completions *are* the answers.

Calculations are entered using a conventional text editor: one can move to any position in a calculation, and insert or erase any symbol. (Because the calculator allows you to edit calculations, it is very easy to do a whole range of related calculations without starting over.) All the time the calculator corrects the calculation to ensure that it is correct in every way, but using a different 'chalk' (colour or style of text) so its answers can be seen easily. In this article, outline has been used to show corrections and completions.

If you tried 4× = 5, which isn't complete, it is corrected to 4×1.25 = 5. Notice how the calculator automatically filled in the blank between the × and = with the answer that was wanted. The new calculator's style of work is literally a 'fill in the blanks' approach familiar from primary school maths books.

What power of 2 is 100? Just enter 2↑ = 100, and the correction is 2↑6.6438562 = 100. Note that formulae do not need to be rearranged to get the answer after an = sign. In this example, an ordinary calculator would have required some considerable expertise, beyond understanding the problem itself, to get the correct button pressing: AC 100 log / 2 log =.

If we tried 4×–5=, which, as we complained above, most calculators get wrong, the new calculator displays 4×–5=–20 — correctly. The same sum could have been entered as =4×–5, with the equals sign first, or even just as 4×–5, which gets the answer =–20, including the missing = sign.

Let's consider a more substantial example. What is the radius of a 1kg lead sphere? Assume we know the density of lead is 11343 $kg \cdot m^{-3}$ and the

formula for the volume of a sphere is $\frac{4\pi}{3}r^3$. On an ordinary calculator the problem would now have to be manipulated to get the answer on one side of an equation. On our true calculator, however, we might easily proceed as follows:

Density is mass divided by volume. So we enter 11343 = 1/ simply omitting the volume we want the calculator to find out. This is displayed immediately as 11343 = 1/8.8160099e–5. (So the volume is 8.8×10^{-5} m^3.) The volume of a sphere is $\frac{4\pi}{3}r^3$, so we just carry on and enter (4π/3× ↑3) where the calculator shows the volume. Notice how we miss out the answer we want, and how the calculator completes the sum for us, providing the answer — just like a helpful teacher. Thus the calculator now finds the radius: 11343 = 1/(4π/3×0.027609666↑3). So a 1kg lead sphere is about 2.8cm radius.

In effect, we *directly* entered the problem $\rho=\frac{m}{\frac{4\pi}{3}r^3}$ by just substituting the values we knew, ρ = 11343, m = 1, and using the calculator's buttons × and ↑ for multiply and raising to a power. What we didn't know we left for the calculator to fix for us, and we didn't need to struggle to rearrange the sum or move the answers so that they came after an = sign.

We didn't know r so we just missed it out, and the calculator fixed the omission, finding the answer we wanted.

Naturally this is not a difficult problem, but it clearly shows how solutions can be found without rearranging equations. This flexibility becomes more important when the calculations become more sophisticated, and especially if the same formula needs to be reused. If we want the radius of a 2.5kg sphere, simply editing the original numerator from 1 to 2.5 gets the new radius immediately: 11343 = 2.5/(4π/3×0.037472082↑3). It is as easy as using an ordinary word processor to edit a sentence — and gives as much freedom to work out what you want to say.

Compare this with what would have to be done on a conventional calculator. To start with, you would have to rearrange the sum to use cube roots. Then you'd have to know the 'algebraic' laws the calculator used. Then I recommend you do the sum several times with similar numbers, just in case — as we saw above was likely — the calculator you were using couldn't cope with π or something else in the wrong place.

The new calculator gives an excellent feel for how functions work. The sum of two logs is the log of their product, which can be tested by typing log 4+log 3 = log. The calculator shows log(4)+log(3) = log(12). It can also correct log 4+log = log 12, effectively answering the question "log of 4 plus log of what is log of 12?" It shows: log(4)+log(3) = log(12).

The relation between logs and powers can be explored directly in many ways:

log(345) = 2.53782	2×log(4) = log(16)
345 = 10↑2.53782	10↑(4×3) = 10↑12
345 = 10↑log(345)	3×log(4) = log(4↑3)
log(10↑345) = 345	log(4)–log(3) = log(4/3)

Seeing equations 'work' *while* numbers are entered is very convincing. Take the last example of the formulae above: if you continued typing 3, 4, 5, you'd get successively the logs of 3, 34 and 345, with results as follows:

log(4)–log(3) = log(4/3), then
log(4)–log(34) = log(4/34), then
log(4)–log(345) = log(4/345)

Think of it as a word processor, with a "maths checker" that is running all the time. At *every* stage the display is correct, even if you have not finished. This is a huge advantage. The continual correction is a significant aid to getting an understanding of the mathematics. (The calculator is very much easier to use than to appreciate from a written description!)

The calculator handles independent calculations one per line. Also, it allows you to select any text and to make a table replacing it successively with 0, 1, 2 … 9 on separate lines. The simple sum ×7 is corrected to 1×7 = 7; so far, not much, but you can use this basis to create a multiplication table. Select the 1 and press TABLE. The calculator creates ten lines: 0×7 = 0, 1×7 = 1, 2×7 = 14 … to 9×7 = 63.

For a table of sines at 10 degree intervals, type sin 0 and place the cursor in front of the 0 before pressing TABLE. The calculator makes ten lines, sin 00, sin 10, sin 20, to sin 90. These are corrected to sin(00) = 0, sin(10) = 0.173648, to sin(90) = 1.

The 0 to 9 of the table need not make 'proper' numbers. A table for arcsines can be made typing 'sin = .' and pressing TABLE just *after* the decimal point. This gets a table running from 0.0 to 0.9.

A formula like 146 = × has many correct answers. The calculator normally makes corrections change things as little as possible. Since the correction to 146 alone is 146 = 146, typing × makes it 146 = 146×1, which has only appended 1. Other choices are available, however, and they can be used to help solve interesting problems.

Choosing 'find factors' as an option for multiplication makes the answer 146 = 2×73 (typing × again gets 146 = 1×2×73). This is an easy way of working with factors. Or instead of finding factors, choosing 'find roots' would obtain 146 = 5.26564×5.26564×5.26564, which finds the cube roots easily. Typing another multiply sign would find fourth roots, and so on.

Other functions have their own options. Thus division can find rational approximations: 22/7–π = 803/635039.

Addition can find powers of two. So 146 = 2+16+128. Clearly, 146 in binary is 10010010. The conversion to binary is easier to see using powers,

which can be set to prefer powers of 2. Type 146 = ↑, corrected to 146 = 2↑7.1898246. Type +↑ again, obtaining 146 = 2↑4.169925+2↑7. Continuing until all the exponents are integers results in 146 = 2↑1+2↑4+2↑7.

Other combinations of preferences can be very effective. Powers of prime factors (×) with integer powers (↑) achieves, for example, 144 = 2↑4×3↑2: that is, $144 = 2^2 3^4$.

This section has done two things. It has shown an alternative way to organise a simple calculator, and shown some interesting things can be done this way. Further worked examples are given below.

We have not directly shown that this new calculator can be implemented any better than a conventional calculator could be. We can argue, however, that the declarative approach is conceptually simpler for the designers, and that the vast, long-standing literature on parsing arithmetic expressions (e.g., Aho, Sethi & Ullman, 1985) is more easily applied correctly when the design is seen in this light. That means any calculator implemented this way can be implemented correctly, and would not have the bizarre properties (such as handling π inconsistently) we saw was common in the conventional *ad hoc* designs. We also have a prototype of the calculator which is available from the author, and you can try it for yourself.

The only argument I've had against the calculator that is almost unanswerable is that it doesn't help children break down sums, because it can work with complex sums directly. I don't think conventional calculators 'help' children break down sums into simple steps, because those steps are often bizarre and more to do with bad calculator design than with how to do mathematics. On the other hand, if you want a 'calculator' (rather, a teaching aid) to be used to help teach children, a different sort of approach is required. Flynn, Dekker and Richards (1996) discuss some ideas; alternatively, if you believe people should break down complex sums into their individual steps, then why not use a reverse Polish calculator, which at least does everything consistently? The true calculator was not, in fact, designed for teaching arithmetic; it was designed for doing it properly.

Summary

We should have come along way from the abacus. The abacus is now a child's toy. Indeed, we should have developed from the conceptual traditions of the earliest electronic calculators. Instead it looks like calculators are getting worse, and are substituting features for quality. Users of calculators seem caught up in buying attractive-looking but technically-incompetent designs, and they blame themselves for their difficulties. In turn, our culture assumes that problems with calculators are not the manufacturers' responsibility.

Although there are really no limitations on what modern calculators can do or how they might work, they work in a way dictated by conformity to obsolete and awkward calculators. They have lots of problems. Even if we accepted the imperative style of use, no calculator works consistently. This article has shown that the problems can be solved: by getting the calculator to show calculations in full; by allowing the user to make any changes they wish; and by maintaining the numbers so that the calculations are always true, moment-to-moment. Sums do not need to be rearranged, hence avoiding a major source of calculator problems. That is why it is called the *true calculator*. This is a declarative approach, and it better matches the way the calculator works with how arithmetic is actually done. It results in a calculator that truly can do calculations exactly as they would be written without any non-standard conventions — and of course it gets the calculations right.

All this can be done very simply and intuitively: the true calculator is direct and surprisingly powerful. People who try the calculator become really excited with its possibilities. It is impossible not to play with numbers. The final problems? People blame *themselves* when they can't use obfuscated and incorrect calculators; and manufacturers will continue to sell incompetent technology to an uncritical public.

Further reading

A. V. Aho, R. Sethi and J. D. Ullman, *Compilers: Principles, Techniques and Tools*, Addison-Wesley, 1985.

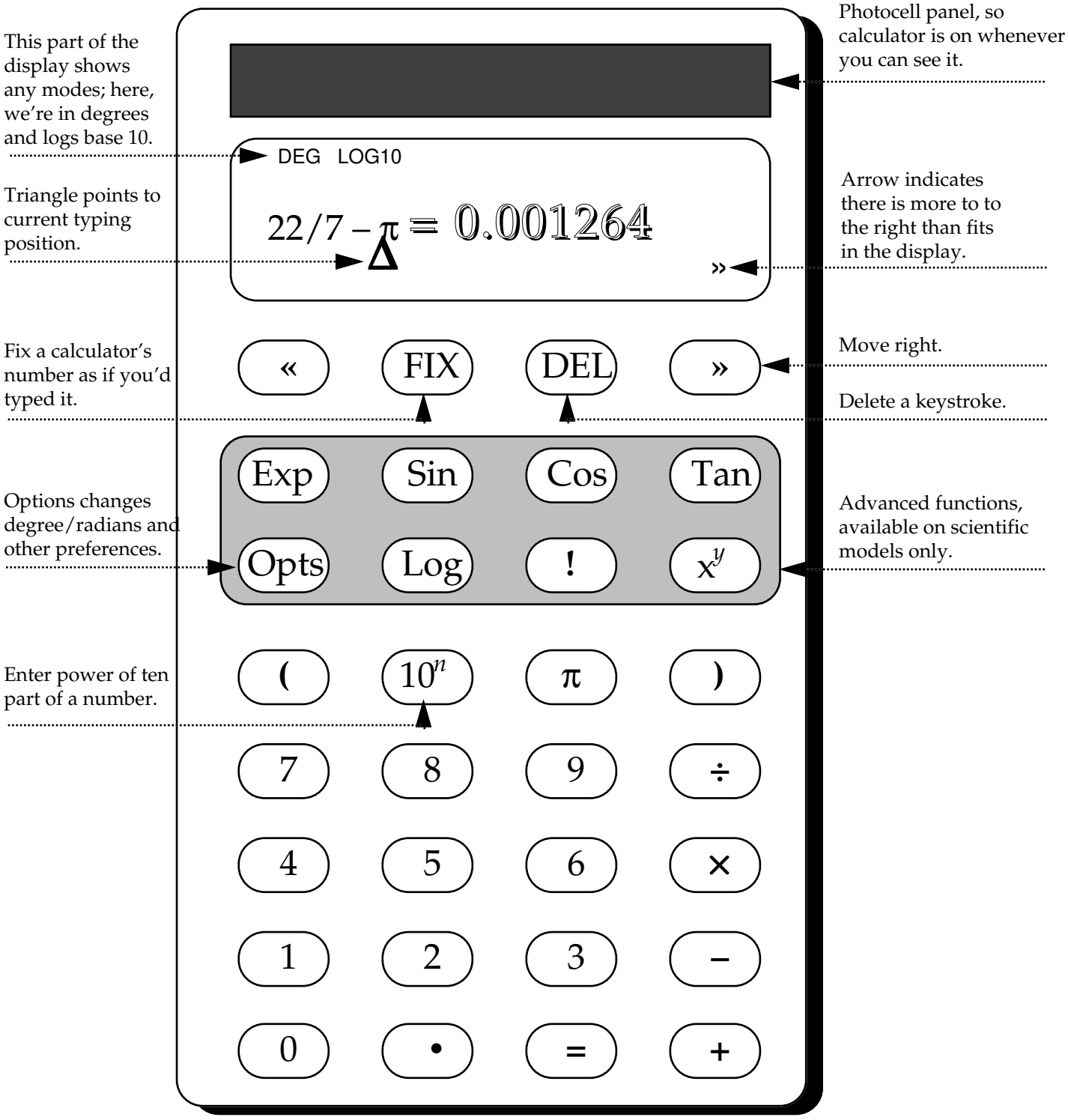
H. Thimbleby, "A new calculator and why it is necessary," *Computer Journal*, **38**(6), pp418–433, 1996.

H. Thimbleby, "Still a few problems to solve," *Times Higher Education Supplement Multimedia*, vi, June 14, 1996.

S. J. Flynn, K. P. Dekker & C. J. Richards, "The design and evaluation of an interactive calculator for children," *Proceedings Eurographics 14th Annual Conference*, edited by H. Jones, R. Raby & D. Vicars, **2**, pp173–181, 1996.

L. R. Wiener, *Digital Woes*, Addison-Wesley, 1993.

Calculators cited: 1: Casio MS-70L. 2: Sharp EL-9200. 3: Casio fx-P401. 4: Sharp EL-531GH. 5: Casio JW-8L. 6: Hewlett Packard HP 20S. 7: Texas Instruments TI-30X. 8: Casio fx-82LB. 9: Casio MS-270L. 10: Innovations Mathspad. 11: Canon WS-121H. 12: Casio fx-115s. 13: Texas Instruments TI-92.



Handheld version of the true calculator. Different models would have different selections of features and arrangements of buttons. (Elementary calculators would have fewer buttons.) A larger display area makes tables practical. An Apple Macintosh prototype provides a menu to control radians/degree modes, logarithm base, and preferences such as requesting division to find rational approximations, powers to find integer solutions, and so forth. The prototype also has tables of constants and standard formulas.

The design shown here is cramped, to fit the features into a small area. More visually appealing designs are of course possible, and indeed there is considerable evidence that well designed layout can have a dramatic effect on users' performance (see Flynn, Dekker & Richards, 1996).

— Worked Examples —

A financial calculation

A credit card agreement states that, "Interest is calculated on a daily basis at the rate of 1.585% per month." If I have £276 credit for a week, how much do I pay?

We will need to work out the daily rate and compound it over seven days. The formula for compound interest is $(1+r)^n$. If r is the daily rate, then the annual rate is $(1+r)^{365}-1$. The daily rate can be worked out from the monthly rate: $(1+r)^{365} = (1+0.01585)^{12}$, since both sides of this must be the same annual rate.

We can enter this formula directly: (1+0.000517143)↑365 = (1+0.01585)↑12, using the button ↑ for raising to a power. As can be seen, the formula does not need to be rearranged to get the daily rate, r, to one side. The daily rate is evidently 0.000517143 or 0.0517143%.

Move the cursor to the daily rate correction, and press FIX. Now the daily percentage rate is fixed, so it will not alter when anything else is changed. We now want to change the term to 7 days. The cursor is moved to the 365 which we delete and replace with 7.

At this point, the calculator shows (1+0.000517143)↑7+0.20406312 = (1+0.01585)↑12. The correction of 0.20406312 arises because we are asking for the impossible combination of seven days at 0.05% daily and twelve months at 1.585% monthly! It is clearest to delete the old monthly calculations, to get (1+0.000517143)↑7 = 1.0036256.

There are two obvious choices for proceeding: We could multiply the compound interest formula by our credit, £276, which would get 276×(1+0.000517143)↑7 = 277.00067. Alternatively and for clarity, we might prefer to FIX the calculated value 1.0036256, delete the rest of the formula (the prototype in fact provides a single action that combines these operations), and then multiply it by our credit. This obtains 1.0036256×276 = 277.00067.

Either way, it is clear we will be paying £277 (perhaps £278!), and the credit is costing us just over a pound: 277.00067–276 = 1.00067.

Ramanujan's 1729 problem

Ramanujan's 1729 problem is to find the two ways in which 1729 can be expressed as the sum of two integer cubes.

You could enter 1729 = ××+××, corrected to 1729 = 1728×1×1+1×1×1, and it looks like there is a long search ahead. You can do better by asking the calculator to find the roots automatically: this gets 1729 = 12×12×12+1×1×1, because × has found the cube root of 1728 (it has also shown you that 1 is the cube root of 1). One solution found right away! Next, replace the 12s with 11s, which gets 1729 = 11×11×11+7.35576×7.35576×7.35576 — certainly a sum of two cubes, but

7.35576 is not an integer. Decrease the 11s to 10s, and you get 1729 = 10×10×10+9×9×9, the second solution.

The TABLE facility makes things easier and more reliable. Enter 1729 = ↑3+↑3, a formula with two slots for two numbers whose cubes add to 1729. This is corrected to 1729 = 9↑3+10↑3 giving you the first solution immediately — coincidentally, ↑ prefers powers of 10. Let's ignore that, because we can't always be so lucky! Select the 9 and press TABLE:

1729 = 0↑3+12.0023↑3
1729 = 1↑3+12↑3
1729 = 2↑3+11.9838↑3
1729 = 3↑3+11.9395↑3
1729 = 4↑3+11.8524↑3
1729 = 5↑3+11.7058↑3
1729 = 6↑3+11.4801↑3
1729 = 7↑3+11.1495↑3
1729 = 8↑3+10.6765↑3
1729 = 9↑3+10↑3

To present the results neatly, delete all the lines that don't have integer solutions, and use FIX on the answers:

1729 = 1↑3+12↑3
1729 = 9↑3+10↑3

Using built in tables

The calculator has built-in tables of interesting number facts. There are three basic tables:

- simple number facts, for example that 102↑7=12↑7+35↑7+53↑7+58↑7+64↑7+83↑7+85↑7+90↑7 or that 4!+0!+5!+8!+5!=40585;
- simple data facts, for example that the earth's oceans have a volume of 1.37×10↑18 cubic metres;
- simple formulas, for example that 32+9/5×1=33.8 converts Celsius to Fahrenheit.

All of these facts, constants and formulas, make full use of the calculator: they can be edited and played with by the user.

A brief specification of the true calculator

The true calculator is a combination of an editor and a calculator.

The editor part is conventional, like any WYSIWYG word processor. It requires a left and right arrow for moving the cursor (the position at which all typing is inserted). When any normal key is pressed, text to the right of the cursor shuffles right to make room, the symbol appears in the newly created position and the cursor moves right, ending up just to the right of the symbol just entered. A delete key erases the symbol to the left of the cursor.

Corrections are done in two stages. First syntactical corrections are made, providing missing operands and operators. (This process generates a complete parse tree.) Next arithmetical corrections are made. Some arithmetical corrections require the insertion of additional numbers, and these are inserted into the parse tree. The parse tree is then display, taking care to display the user's original expression and any corrections in different styles.

A common case of arithmetic correction occurs when the final equation does not balance; a positive numerical correction is added. All undefined expressions and overflow conditions are also corrected. The main part of the article gave "sensible examples" but extreme cases has to be handled too. Here are some examples:

4 = (5	*is corrected to*	4+1 = (5).
1/0	*to*	1/(1+0)
π!	*to*	(0.95493π)!
π = !	*to*	π=3!−2.8584073
sin = 5	*to*	5×sin(90) = 5
log(–2)	*to*	log(—2)
0↑0	*to*	(0+1)↑0

Arguably the correction to π! should have been something like (π/π)!, since (0.95493π) is still not an integer (which ! requires). The calculator is still not perfect. Numerical problems are tricky to avoid, unless the calculator is restricted to simpler functions (e.g., for a primary level calculator).

Corrections the calculator make *never* change anything the user has keyed. However it is entered, corrections for the same final calculation are *always* the same. The cursor cannot be positioned inside or immediately to the right of corrections, it skips over them when an attempt is made to move into them.

A key FIX changes the correction the cursor is just to the right of into text exactly as if typed by the user.

A key TABLE copies the current line ten times, and inserts successively 0 to 9, as if the user had keyed them at the cursor's position on each line. (The table facility requires a means for scrolling the display vertically in addition to the left/right arrows.)

Two keys, STORE and RECALL, permit copy-and-paste editing. STORE copies the number to the left of the cursor. RECALL enters the last number stored, exactly as if the user had keyed it directly.

Harold Thimbleby is Professor of Computing Research and Faculty of Technology Director of Research, Middlesex University, Bounds Green Road, LONDON, N11 2NQ. He can be contacted by email: Harold@mdx.ac.uk. This work was supported by EPSRC Grant GR/J43110.