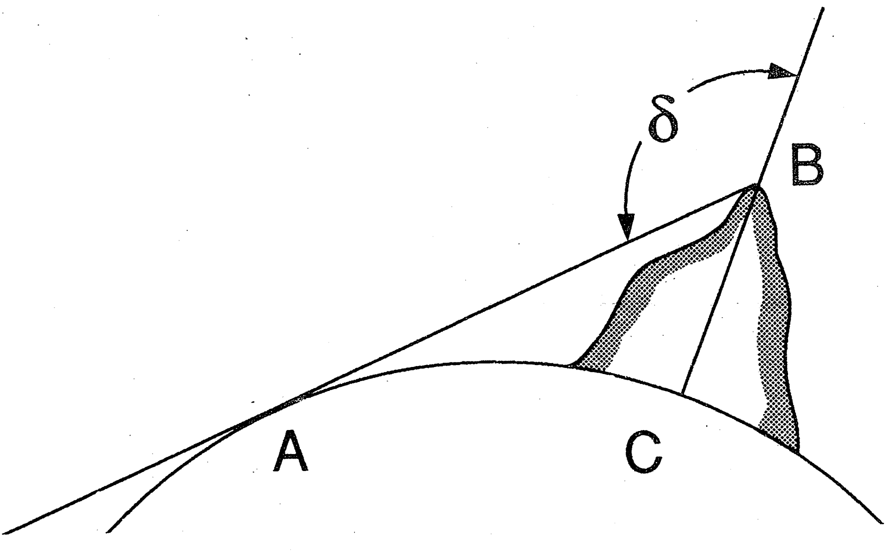
Function

Founder Editor G. B. Preston

Volume 17 Part 1 February 1993



A SCHOOL MATHEMATICS MAGAZINE

...g. by Aust. Post Publ. No. VBH0171

Registered for posting as a periodical – "Category B"
ISSN 0313 – 6825

* * * * *

Published by Monash University Mathematics Department

FUNCTION

Volume 17 *Part 1*

(Founder editor: G.B. Preston)

Welcome to new readers and welcome back to old friends. *Function* is a journal devoted to high quality Mathematics at the school level and it welcomes contributions from readers – especially student readers. This issue unravels the complexities of housing mortgage loans in its leading article, but also features Karl Spiteri's solution to an everyday problem. If we extend the ideas involved in that article, we come to the Mathematics discussed by Cristina Varsavsky in her Computer Column. Jim Mackenzie writes a guest column on the History of Mathematics and tells us how proofs began. And then there are our regular cover feature and Problems Section. Let us know if there are topics you would like to see covered or if you have ideas to share with other readers.

CONTENTS

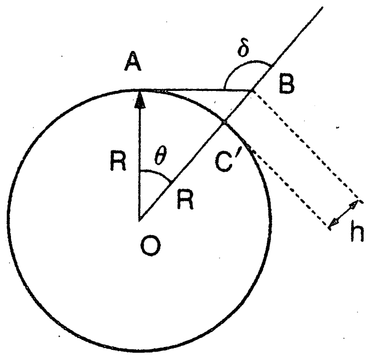
The Front Cover	Michael A.B. Deakin	2
How to work out your Home Loan	J.G. Kupka	3
Tick Tock	Karl Spiteri	11
Computers and Computing		13
History of Mathematics: How Proofs Began	Jim Mackenzie	18
Problems and Solutions		24
Coloured Caps and Right-Hand Turns		29
How about this one?		32

THE FRONT COVER

Michael A.B. Deakin, Monash University

The cover diagram for this issue is taken from an early 20th Century manual of surveying (John Whitelaw, Jun.: *Surveying as Practised by Civil Engineers and Surveyors*, 1902). It illustrates a technique for finding the heights of mountains by "Observation to the Sea Horizon".

The surveyor takes a theodolite to B, the top of the mountain, and uses the instrument to measure the angle δ between the zenith (the point directly overhead) and A a point on the horizon out to sea.



From the diagram at right we see that if R is the radius of the earth and h is the height of the mountain, then

$$R = (R + h)\cos\theta. \qquad (1)$$

It is also known from geometry that

$$\delta = 90^\circ + \theta \qquad (2)$$

as the angle OAB is a right angle.

In practice, h is much smaller than R and this will mean that θ is a very small angle. So δ will be an angle just a little greater than 90°.

As an illustration, consider the case $\delta = 92^\circ$, i.e. $\theta = 2^\circ$. We may now use Equation (1) to determine h if R is known.

The founders of the metric system so defined the metre that the circumference of the earth is approximately 40 000 km. This means that $R \approx 6\ 366$ km, or 6.366×10^6 m.

We could now use Equation (1) to determine h directly, but some extra theory makes the calculation easier. First convert θ to radians, to find that in *radian* measure

$$\theta = 2\pi/180 \approx 0.0349.$$

In radian measure, it is known that for small angles

$$\frac{1}{\cos\theta} \approx 1 + \frac{\theta^2}{2},$$

so here

$$\frac{1}{\cos\theta} \approx 1 + \left[6.09 \times 10^{-4}\right].$$

Rearrangement of Equation (1) gives

$$1 + \frac{h}{R} = \frac{1}{\cos\theta}$$

and so we have

$$h \approx (6.09 \times 10^{-4})R \approx 3880 \text{ m}.$$

The method is a quite practical one and is still in use today.

HOW TO WORK OUT YOUR HOME LOAN

J.G. Kupka, Monash University

Last year you took out a home loan for \$55,000. The interest rate was 13%, and you agreed to a 15-year repayment term. The bank manager then said that repayments would amount to \$696 per month. Can we quickly check the accuracy of this figure without making $15 \times 12 = 180$ tedious calculations?

This year the interest rate has been steadily falling. The bank asks that you keep up your monthly payments of \$696 but warmly assures you that your repayment term has become shorter. Can we quickly determine the new term without having to go down to the bank and wait in a queue every time the interest rate changes?

Now a rival bank is advertising "fabulous gains" to be made by changing your payment schedule to fortnightly repayments. There is a clear implication that they are offering a much better deal than your behind-the-times old bank, but your bank is happy to negotiate fortnightly repayments and says that its deal will "certainly" match the other bank's. But who can be believed in these days of cutthroat competition and rampant corruption in high places? Would it not be better to work out for yourself the effects of rescheduling?

To make a proper scientific investigation of these questions, we must first identify and give names to the various quantities which are involved in any home loan situation.

Let P be the principal, or the amount which you originally borrowed. In your case, $P = 55{,}000$.

Let M be the (fixed) amount you are required to pay at the end of each payment period. In your case, $M = 696$.

Let N be the number of repayments made per year. While N could in theory be any positive integer, the only common values of N are $N = 12$ in the case of monthly repayments (your case) and $N = 26$ in the case of fortnightly repayments.

Let I be the annual interest rate, expressed as a proportion. Thus, if your interest rate is 13%, then $I = 0.13$.

Let C be the number of times the interest is compounded each year. Savings accounts may compound the interest annually, which means that $C = 1$, or quarterly ($C = 4$), or monthly ($C = 12$). Thanks to computers, home loan accounts nowadays are typically compounded *daily*, so that $C = 365$.

Let t be the total number of payments which have to be made. In your case, $t = 12 \times 15 = 180$.

We shall treat the variables P, M, N, I, C, t as primary. Other variables of interest are

T = the term of the loan in years = $t/N = 15$

A = the amount paid annually to the bank = $MN = 8352$

R = the grand total repaid to the bank = $tM = TA = 125{,}280$.

Since R is always very much larger than P, the banks are unlikely to volunteer its exact value. It can, however, change dramatically with changes in the primary variables and is therefore something of a hypothetical figure. What with variable interest rates, unscheduled payments (or non-payments) on your part, bank fees, government charges, and unpredictable policy changes (for example, in the C-value), the final cost of your home loan has to remain at best an approximation, a "ball park" figure.

We are going to derive a single equation (equation (*) below) which relates the primary variables. It will give quick and reasonably accurate answers to most questions about home loans and can be used repeatedly as circumstances change. And you won't have to pester your bank manager too much or take his word for everything.

First we need to review how the interest on your loan is actually calculated. The year is divided into C equal *compounding periods* of duration $1/C$ each. Suppose you owe an amount D at the beginning of a compounding period. The interest accumulating during this period will be one C-th part of the annual interest rate I, or I/C. At the end of the period the bank will increase your debt by the amount $D(I/C)$, so you then owe a total of $D + DI/C = D(1 + I/C)$. After the next compounding period you owe $[D(1 + I/C)](1 + I/C) = D(1 + I/C)^2$, and after the k-th compounding period you would owe $D(1 + I/C)^k$. After one year, or $k = C$ compounding periods, assuming no repayments during this time, your debt would be multiplied by the factor

$$\kappa = \left(1 + \frac{I}{C}\right)^C,$$

which we shall henceforth call the *compounding factor* determined by I and C. For example, when $I = 0.13$ and $C = 365$, we get $\kappa = 1.138802$. The effect on your debt is the same as if the interest had been calculated once only at an annual rate of about 13.88%. Banks sometimes refer to this sort of figure as the "effective" annual interest rate, but only in connection with savings accounts (i.e. interest they are paying to *you*), never in connection with loan accounts (i.e. interest you are paying to *them*).

An increase in C always produces an increase in κ. So you get a better deal with a *high* C-value in a savings account and a *low* C-value in a loan account. Fortunately, there are limits to the extent that banks can squeeze you by excessive compounding. Even if they ran amok and started compounding by the minute, the compounding factor could never get beyond the limiting value of $\kappa^* = e^I = 1.138828$, where $e = 2.71828 \ldots$ is the base of the natural logarithm. (To show this, you would need to use the mathematical theory of limits.) As there is negligible difference here between κ and κ^*, we see that daily compounding takes us close to the maximum squeeze. Using κ^* itself as the compounding factor has been referred to as "instantaneous compounding".

Now we need to examine the effects of your regular payments on the debt. Let P_n be the size of the debt after your n-th payment (of M dollars). As with C, we assume (with negligible inaccuracy) that the year is divided into N equal *payment periods* of duration $1/N$ each. With C equal compounding periods per year and N equal payment periods per year (say $C = 365$, $N = 73$), there are exactly C/N $(= 5)$ compounding periods in each payment period. Hence, during the n-th payment period, your debt at the start of that period, which was P_{n-1}, is multiplied by the factor

$$r = \left(1 + \frac{I}{C}\right)^{C/N} = \kappa^{1/N} \;(= 1.001782).$$

which we shall henceforth call the *rate factor* determined by I, C, and N. To get from P_{n-1} to P_n the banks use the rule: *First* add in the full amount of the interest on P_{n-1} accruing during the entire interval, *then* subtract the payment M. Hence

$$P_n = rP_{n-1} - M.$$

This equation is even valid when $n = 1$, provided we define $P_0 = P =$ the original sum borrowed. It is also the original amount owing, prior to any repayments.

In order to derive workable formulas, we shall have to treat r as constant from one payment period to another. And so it is! – but only in the somewhat artificial situation when C/N, the number of compounding periods in each payment period, is a whole number. In the most commonly occurring real-life situation, where $C = 365$ and $N = 12$, we have $C/N = 30.4167$ but will still use $r = \left[1 + \frac{.13}{365}\right]^{365/12} = 1.010890$ as our rate factor. The banks would use $\left[1 + \frac{.13}{365}\right]^m$ for a month consisting of m days, hence $r_1 = 1.01\underline{1100}$ for months of 31 days, $r_2 = 1.01\underline{0740}$ for months of 30 days, and $r_3 = 1.01\underline{0021}$ for most Februarys. Our use of the constant intermediate value of $r = 1.01\underline{0890}$ will cause the relatively small errors in P_n virtually to cancel out over the course of a year (since $r^{12} = \kappa =$ the product of the r_1's, r_2's and r_3). In your case ($P_0 = 55{,}000$, $M = 696$), 12 successive calculations using r give $P_{12} = 53{,}763.24$, whereas the bank's calculations using r_1, r_2, r_3 give $P_{12} = 53{,}760.71$. By comparison with the large sums involved in a home loan, the difference between these two figures is insignificant.

Anyone who actually calculates P_n in this way, even on a calculator, will appreciate how tedious it would be to get all the way out to P_{180} to see if it really is zero. An explicit formula for P_n would save a lot of time. At the start we only know the value of $P_0 = P$. We then have to calculate, in succession:

$$\begin{aligned}
P_1 &= Pr - M \\
P_2 &= P_1r - M = [Pr - M]r - M \\
P_3 &= [[Pr - M]r - M]r - M,
\end{aligned}$$

and so on. At this rate, it looks as if an explicit formula for P_n in terms of P, M, r (and hence ultimately in terms of P, M, I, C, N) would be hopelessly complicated. However, a little bit of algebra and educated guesswork will produce a simpler-than-expected formula for P_n. Multiply out the formula for P_3 to get

$$\begin{aligned}
P_3 &= Pr^3 - Mr^2 - Mr - M \\
&= Pr^3 - M(1 + r + r^2).
\end{aligned}$$

The advantage of this version of P_3 is that it contains a clearly discernible *pattern.* We *hope* that this pattern will be preserved in P_4, P_5, etc., and, if so, then we *ought* to have

$$\begin{aligned} P_n &= Pr^n - M(1 + r + r^2 + \dots + r^{n-1}) \\ &= Pr^n - M\left(\frac{1 - r^n}{1 - r}\right) \\ &= \boxed{r^n\left(P - \frac{M}{r-1}\right) + \frac{M}{r-1}}\,. \end{aligned}$$

Our original expression for P_n may now be used to show that the above formula for P_n is correct for any value of n. Of course, we have to show it first for P_1, then for P_2, then for P_3, and, in general, if we have shown it for any particular P_{n-1}, then

$$\begin{aligned} P_n &= P_{n-1}r - M \\ &= \left[\, Pr^{n-1} - M(1 + r + \dots + r^{n-2}) \,\right] r - M \\ &= Pr^n - Mr(1 + r + \dots + r^{n-2}) - M \\ &= Pr^n - M(1 + r + r^2 + \dots + r^{n-1}), \end{aligned}$$

which shows it for P_n as well. Please observe how much easier it is to get $P_{12} = 53{,}763.24$ when you use the formula in the box rather than the 12 successive calculations performed earlier.

Before proceeding, let us take a moment to see what this formula is telling us about P_n. We can think of $r - 1$ as the "effective" interest rate operating during one payment period. Hence an M-value of $P(r-1) = 598.96$ would indicate an "interest only" loan in which $P_n = P$ for all n. In other words, the amount you owe remains fixed, and you are making zero progress in paying off the principal. If $M < P(r-1) = P\left[\left(1 + \frac{I}{C}\right)^{C/N} - 1\right]$, then $P - \frac{M}{r-1} > 0$, and the formula for P_n tells us that

$$P_0 < P_1 < P_2 < \dots < P_n < \dots,$$

i.e. your debt is *increasing* with every payment period. If the interest rate ever blew out to the point where $M < P(r-1)$, you should expect an urgent call from your bank manager! On the other hand, if $M > P(r-1)$, then the formula for P_n tells us that the debt is *decreasing* with every payment period, and the bigger the difference, the faster it decreases, i.e. the shorter the term.

Remember now that P_n is the amount you still owe after the n-th payment. The only value of n which really interests us is $n = t =$ the total number of payments. After you have made your t-th payment, you should owe absolutely *nothing*, or, in other words, $P_t = 0$. Set $n = t$ in the formula for P_n, simplify slightly, and we get our key equation:

$$(*) \qquad \boxed{M = (M - P(r-1))r^t}$$

(Remember that $r = \left[1 + \frac{I}{C}\right]^{C/N}$.)

Let us look again at your original loan, where $P = 55{,}000$, $M = 696$, $N = 12$, $I = 0.13$, $C = 365$, and $T = 15$, so that $t = 12 \times 15 = 180$. We shall take $\boxed{C = 365}$ in all the examples to follow. However, to illustrate the slightly adverse effects on the borrower of the frequent compounding of interest, we shall recalculate some figures changing $C = 365$ to $C = N$. The case $C = N$ is "fair" in the sense that, in this case, the payment and compounding periods have exactly the same duration. Moreover, all the formulas are simpler, since $r = 1 + I/N$.

The easiest quantities to compare are M and P, with the other variables fixed. Performing a little algebra on equation (*) gives

$$(1) \qquad P = M\,\frac{1 - r^{-t}}{r - 1}\,.$$

This formula will let you check how much you could have borrowed, under the original terms, for various M-values. If, for example, you had only been able to afford $M = 500$ a month, then formula (1) gives $P = 39{,}378$ (rounded to the nearest dollar). It is also clear from (1) that doubling your monthly repayments would double the amount you could borrow. Hence a \$1000 monthly payment would permit you to borrow anything up to \$78,756. And setting $M = 696$ gives $P = 54{,}814$ (or 55,009 when $C = N$). Has the bank been a bit generous here, allowing you to borrow more than the calculated P-value?

To find out, we solve for M in (1) to get

$$(2) \qquad M = P\,\frac{r - 1}{1 - r^{-t}}.$$

This formula will let you check how your periodic repayments are influenced by the other factors. When $P = 55{,}000$, formula (2) gives an M-value rounding to \$698.36 (or \$695.89 when $C = N$). Interesting. Is the bank being generous, or is it just using the simpler formulas of the $C = N$ case to make its reckoning?

Suppose now that one year (or 12 payments) after the original loan agreement was made, the interest rate has dropped to 10.5%. Putting $I = 0.105$ into (2) (with $P = 55{,}000$ and $t = 180$) gives $M = 609.49$, so if you were strapped for cash, you might theoretically argue the bank down to a monthly payment of \$610. However, your bank manager is not amused by this and announces with a tone of irritation that you will have to "re-establish" the loan for a fee of \$500. Furthermore, the monthly payments will be \$614 and not \$610. After you angrily threaten to take your business to that other bank, he agrees to waive the \$500 fee but holds firm on the \$614 payment. Does this come from a legitimate calculation, or is it a peevish \$4-a-month rort? Let us see that it is legitimate. After your 12 payments at the 13% rate, the amount you currently owe is \$53,760.71 (as we worked out earlier using the bank's method of reckoning). In re-establishing your loan, the manager has simply used the figure $P = 53{,}761$ together with $t = 180 - 12 = 168$, $I = 0.105$, and $N = 12$. Inserting these into formula (2) gives $M = 613.49$, which, of course, the manager rounds up to \$614. (If he had used the $C = N$ case, he would have offered \$612, but he doesn't want you bothering him like this.)

If you were really strapped for cash and were able to persuade the manager to extend the term of your re-established loan from 14 to 25 years (so that $t = 25 \times 12 = 300$), then your monthly payment would reduce further to \$510 (or \$508 if $C = N$).

If you are able to maintain or even increase your monthly payments, and if the interest rate continues to decline, then you have the satisfaction of knowing that it will all be over sooner. To find out how much sooner, we rearrange equation (*) to get

$$r^t = \frac{M}{M-P(r-1)} = \frac{1}{1-\frac{P}{M}(r-1)}.$$

Taking logs of both sides gives

$$t \log r = \log M - \log(M - P(r-1)) = -\log\left[1 - \frac{P}{M}(r-1)\right].$$

(Any type of log will do here, either the natural log or the log to the base 10, so long as all logs are of the *same* type.) Hence

$$(3) \qquad t = \frac{\log M - \log(M-P(r-1))}{\log r} = -\frac{\log\left[1 - \frac{P}{M}(r-1)\right]}{\log r}.$$

Using the figures of the original home loan gives $t = 181.9044$. The fact that t is not equal to 180 comes from your payment of \$696 instead of, say, \$699, which would give $t = 179.4904$. (To get $t = 180$ exactly, you would need to set $M = 698.357963$, which was the exact (calculator) value obtained earlier in our first application of formula (2).)

In general, a non-integer value of t is to be expected if you determine the other variables at will. It reflects the fact that no home loan can ever be engineered to work out so perfectly that your very last payment will be exactly equal to M. Take the t-value of 179.4904 resulting from $M = 699$. The next-to-last payment will be the 179-th, and the formula for P_n gives $P_{179} = 340.02$. Suppose you decided to close off your loan exactly .4904 of one month past your next-to-last payment. In a month of 31 days, this would represent 15.2024 days, and so you make your payment on day 16. The bank will not officially credit your payment until the end of the current compounding period, which means midnight of that day. Hence, your payout figure is $340.02 \left[1 + \frac{.13}{365}\right]^{16} = 341.96$.

(Remember that this calculation assumes the 13% interest rate throughout the entire life of the loan.) Now compare this with the figure $.4904M = .4904 \times 699 = 342.79$. The two figures are practically identical. So we may interpret the fractional part of $t = 179.4904$ as that fraction of a payment period when you would close down the loan, assuming that you keep paying *at the same rate*. In any event, the non-integer t-value will certainly give you a very accurate idea of how long, under current conditions, your payments will last (using $T = t/N$) and how much you will ultimately have to pay (using $R = tM$).

In your original situation, the t-value of 181.9044 gives $T = t/12 = 15.16$ years, which exceeds the original 15-year term by about 2 months. This was owing to the "generosity" of the bank in allowing you to pay \$696 instead of \$699. Naturally the grand total payment of $R = 696t = 126,606$ slightly exceeds the original estimate of 125,280. (Compare this with the $M = 699$ multiplied by the $t = 179.4904$, which amounts to 125,464.) We shall use $T = 15.16$ and $R = 126,606$ as benchmark figures in assessing the impact which the interest rate and the periodic payment figure have on your home loan.

First, the interest rate. Suppose the 13% is reduced to 11% throughout the life of the loan. Formula (3) (with no other figure of the original home loan changed) gives $t = 141.8913$, whence $T = 11.82$ years and $R = 98{,}757$, a saving of about $4\frac{1}{4}$ years and \$28,000. Clearly the I-value has a big impact on the total repayment.

As for the M-value, suppose we increase it to \$1,000 a month, leaving everything else unchanged. Formula (3) gives $t = 84.3572$, whence a 7-year term and a total repayment of \$84,357. This is a marvellous saving, if you can cop the M-value!

What now of the "fabulous gains" which Mr. and Mrs. Jones, that ecstatic couple in the ad for the rival bank, were supposed to have made when they changed from monthly to fortnightly repayments? This amounts to changing the N-value from 12 to 26. To gauge the impact of this on its own, we must adjust M to keep the total annual payment A the same, and we must adjust t to keep the term T the same. (Obviously, paying \$696 per fortnight rather than monthly would have a dramatic effect on both your home loan and your lifestyle.) One way to do this is to eliminate M and t from equation (*). So we replace M by A/N, t by NT, r by $\kappa^{1/N}$ (where $\kappa = \left(1 + \frac{I}{C}\right)^C$ is the compounding factor), and multiply both sides by N to get the alternative version

(**) $$A = (A - PN(\kappa^{1/N} - 1))\kappa^T.$$

Solve for A and we get

(4) $$A = P \frac{\kappa^T}{\kappa^T - 1} N\left(\kappa^{1/N} - 1\right).$$

For purposes of illustration, let us just take $P = 55{,}000$, $T = 15$, and $\kappa = \left(1 + \frac{.13}{365}\right)^{365} = 1.138802$. Then $N = 12$ gives $A = 8380.30$. (Notice that when this figure is divided by $N = 12$ we get the M-figure of \$698.36, which is required to ensure a term of exactly 15 years.) Now $N = 26$ gives $A = 8355.86$, a saving of about \$25 a year or \$375 over the life of the loan. Not quite in the "fabulous" category. But it *is* a saving. What if we keep going, say to weekly or even daily repayments? Perhaps you have a compromising photo of your bank manager in a motel room with a young girl and can persuade him to allow minuscule payments once every nanosecond. No matter. If he knows the score, he's smiling. True, an increase in N always produces a decrease in A. But no matter how much you run amok with the N-value, the annual repayment figure could never get beyond the limiting value of

$$A^* = P \frac{\kappa^T}{\kappa^T - 1} \ln \kappa = 8334.99.$$

(This is the *natural* logarithm here.) Your additional gain is not even worth another \$25 per year. (And under the more lenient $C = N$ regime, the savings would only be about one-third of this.)

Thus, shortening your payment period gives you about the same amount of benefit as shortening the compounding period gives to your bank. In either case the benefits are relatively slight. This is why we suggest that equal payment and compounding periods, the $C = N$ case, would probably make a fair compromise between you and your bank.

At this point we can say that the ad for the rival bank is at best misleading. Is it an outright lie? Probably not. If you went, all starry-eyed, to the rival bank, the rival manager would almost certainly tell you that a fortnight is about half a month, so your fortnightly payment will be half of \$696, or \$348. And since \$348 sounds less painful than \$696, he might even try to squeeze a bit more out of you. He hopes you won't notice that 26 fortnightly payments of \$348 will increase your total annual outlay by \$696. And this is where the advertised gains are really coming from. Put $P = 55{,}000$, $I = 0.13$, $M = 348$, $N = 26$ into formula (3) and we get $t = 314.1641$, whence $T = t/26 = 12.08$ years and $R = tM = 109{,}329$, a gain to you (over the benchmark figures) of just around 3 years and \$17,000. Not bad. To see that this is due mainly to the extra \$696 you are forking out each year, let us see what happens if, in your original loan arrangement, you change $M = 696$ to $M = 348 \times 26 \div 12 = 754$. Then formula (3) gives $t = 146.0330$, whence $T = t/12 = 12.17$ years and $R = 110{,}109$. So you would pay an extra \$780 or so (or \$320 when $C = N$) over the entire life of the loan for the privilege of repaying monthly rather than fortnightly.

It seems safe to conclude that fortnightly repayment schemes serve mainly to trick people into higher annual loan repayments.

There is one final type of question which might be asked about home loans. I, too, want a home loan for \$55,000 and am pleased by the current interest rate, which has just dropped to 10%. However, they hardly pay me anything at all to write articles like this, so I can only afford repayments of \$500 a month tops. I use formula (3) to get $t = 303.6855$, whence $T = 25.31$ years. A bit longish, perhaps. But I am an important person. I work in a university. So I approach my bank manager with confidence. He raises his eyebrows, peers at me over his spectacles, strains out a Mona Lisa smile reserved for people whose IQ does not exceed the room temperature, and says: "No way, José. You seem to forget that you are also frightfully old. We cannot possibly authorize a term longer than 15 years. At that rate, even your precious formula (1) only gives $P = 46{,}422$. So we can only offer \$47,000 tops. Take it or leave it." I respond in kind and depart his office with loud shoes.

We have seen that fiddling with N has a negligible effect on the outcome, so my only hope for a deal would be a lower interest rate. To determine how low it would have to go to make \$500 a month possible, it would be ideal to solve for r, and hence I, in (*). But I'm not going to try it. I'm in a bad mood, and they didn't pay me enough. If you can do it, and your price is right, the editors of *Function* would be delighted to hear from you. Meanwhile, I shall content myself with plugging various I-values into formula (2) just to get a rough idea. With $P = 55{,}000$, $N = 12$, $t = 180$, an I-value of 0.07 gives $M = 494.96$. Just about good enough.

* * * * *

[The situations described in the above article are, of course, fictitious, as the style makes clear. Nonetheless, we have been at some pains to reflect current banking practice. There are, however, changes that take place, not just from time to time but quite frequently, and even while the article was being prepared for publication, changes were made to it in an attempt to be as up-to-date as possible. Nor do all banks and lending institutions adopt the same policies. Some, for example, have extra charges that only come out in the "fine print"; others would not so readily dismiss our aging author's proposal. To keep up, one needs to do the calculation not once but over and over again, hence the utility of having nice formulas; the subject is inexhaustible!

For more on the connection between compound interest and natural logarithms, see, e.g., W.W. Sawyer's Mathematician's Delight, *pp. 74-77. Eds.]*

TICK TOCK

Karl Spiteri, Student, University of Melbourne

Tick tock, hands of a clock,
At what times will you interlock?

By "interlock", I mean "align". What times of day have both the minute and the hour hand together?

Let α be the angle *in degrees* between the hour hand and 12 o'clock; similarly let β be the angle *in degrees* between the minute hand and 12 o'clock. (See Figure 1.) E.g. at 1 o'clock $\alpha = 30$, $\beta = 0$.

Then

$$\beta = 12\alpha. \tag{1}$$

Clearly the hands are together if $\alpha = \beta = 0$, i.e. at 12 o'clock itself.

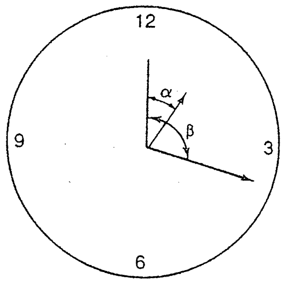


Figure 1

More generally, if the hands are to be together, we require

$$\beta = \alpha + 360n \tag{2}$$

where n is an integer.†

Take now the case $n = 1$. We then have the simultaneous equations

$$\beta = 12\alpha$$
$$\beta = \alpha + 360.$$

† See Cristina Varsavsky's computer column in this issue (Eds.).

The solution of these equations gives $\alpha = 32\frac{8}{11}$, and it is not difficult to see that this corresponds to a time of $1 : 05 : 27\frac{3}{11}$.

Substituting successively $n = 2, 3, ..., 11$ produces the table below. When $n = 11$, we are back to 12 o'clock. There are thus 11 different times over a twelve-hour period, at which the hands align. These times may also be found by adding $1 : 05 : 27\frac{3}{11}$ to each previous time. Thus $n = 2$ corresponds to a time of

$$1 : 05 : 27\tfrac{3}{11} + 1 : 05 : 27\tfrac{3}{11} = 2 : 10 : 54\tfrac{6}{11}$$

and if we add $1 : 04 : 27\frac{3}{11}$ to this, we find $3 : 16 : 21\frac{9}{11}$, etc. We may this construct Table 1.

Value of n	Value of α	Alignment time
0	0	12:00:00
1	$32\frac{8}{11}$	$1{:}05{:}27\frac{3}{11}$
2	$65\frac{5}{11}$	$2{:}10{:}54\frac{6}{11}$
3	$98\frac{2}{11}$	$3{:}16{:}21\frac{9}{11}$
4	$130\frac{10}{11}$	$4{:}21{:}49\frac{1}{11}$
5	$163\frac{7}{11}$	$5{:}27{:}16\frac{4}{11}$
6	$196\frac{4}{11}$	$6{:}32{:}43\frac{7}{11}$
7	$229\frac{1}{11}$	$7{:}38{:}10\frac{10}{11}$
8	$261\frac{9}{11}$	$8{:}43{:}38\frac{2}{11}$
9	$294\frac{6}{11}$	$9{:}49{:}5\frac{5}{11}$
10	$327\frac{3}{11}$	$10{:}54{:}32\frac{8}{11}$
11	360	12:00:00

Table 1

We may also ask at what times the hands form a given angle φ, or what happens if the gearing is incorrect and $\beta/\alpha \neq 12$, or what goes on with a 24-hour clock, and so on.†

* * * * *

† Or for that matter, what happens if the clock has a second hand as well as a minute hand and an hour hand. See Problem 4.4.3 (solved *Vol.* 4, *Part* 2).

COMPUTERS AND COMPUTING

EDITOR: CRISTINA VARSAVSKY

Modular Arithmetic Keeps The Numbers Small

Although the name *modular arithmetic* may seem new to you, you have been applying its principles for some time without having any mathematical language to describe what you were doing. A typical example is the clock. As long as its source of power remains intact, a clock keeps on counting hours. But there is something different in this counting: even though it counts for ever (ideally) it never gets up to any large number.

Clock arithmetic is very odd looking. Let us suppose it is 9 o'clock and we want to add 7 hours. We quickly find the answer, namely 4 o'clock. Then, from the clock point of view,

$$9 + 7 = 4.$$

In the same way we find that

$$9 - 10 = 11 \quad \text{and} \quad 8 + 11 = 7.$$

What we are doing is counting in sets of 12 and considering only the remainder. This is what mathematicians call *modular arithmetic*.

Going back to the first addition, $9 + 7$, in standard arithmetic we get $9 + 7 = 16$. To relate this to the clock arithmetic we usually write

$$16 \equiv 4 \pmod{12}$$

and this relationship is read "*16 is congruent to 4 modulo 12*". This notation is due to Gauss†, and 12 is called the *modulus*.

But there is nothing special about the number 12. By fixing any integer $m (> 1)$, we define

$$b \equiv a \pmod{m}$$

if the remainders of b and a when divided by m are the same, which is equivalent to saying that m divides $b - a$.

Thus, $49 \equiv 1 \pmod 4$ says that 49 has a remainder of 1 when counting in sets of four. Also, $49 \equiv -3 \pmod 4$ because 4 divides $49 - (-3) = 52$.

The *modulo* m *system* uses a finite set of integers: the possible remainders when dividing by m, namely $0, 1, 2, ..., m-1$. Three of the fundamental operations, addition, subtraction and multiplication, are the same as those of ordinary arithmetic, except that if the result of the operation is greater than $m-1$, it is divided by m, and the remainder (r, where $0 \leq r < m$) is used in place of the ordinary result. For example,

† Carl Friedrich Gauss (1777-1855), one of the greatest mathematicians of all time.

$$7 + 3 \equiv 2 \pmod 8$$
$$4 - 6 \equiv 7 \pmod 9$$
$$4 \times 10 \equiv 7 \pmod{11}.$$

The operation of division has to be treated more carefully. Let us try to find $1-3$ in the modulo 6 system. This means we have to find a number s such that $3 \times s \equiv 1$ (mod 6). This number s is either 0, or 1, or 2, or 3, or 4, or 5. Since there are only 6 possibilities, we can try each of them:

$$\begin{aligned} 0 \times 3 &\equiv 0 \pmod 6 \\ 1 \times 3 &\equiv 3 \pmod 6 \\ 2 \times 3 &\equiv 0 \pmod 6 \\ 3 \times 3 &\equiv 3 \pmod 6 \\ 4 \times 3 &\equiv 0 \pmod 6 \\ 5 \times 3 &\equiv 3 \pmod 6. \end{aligned}$$

Therefore it is impossible to divide 1 by 3 in the modulo 6 system. But it is possible to divide 1 by 5, because $5 \times 5 \equiv 1$ (mod 6), and consequently $1 + 5 \equiv 5$ (mod 6). You may wonder why, but if you have a close look at the numbers involved, there is something important that distinguishes them: while 3 and 6 have a non-trivial common factor†, 5 and 6 do not. This is the general restriction when inverting a number in modular arithmetic: the number and the modulus cannot have common non-trivial factors. Once the problem of inversion is sorted out, the division follows immediately. For example (from the previous table):

$$3 \div 5 \equiv 3 \times (1 \div 5) \equiv 3 \times 5 \equiv 3 \pmod 6.$$

An important fact is that both sides of a congruence equation can be added to, subtracted from, multiplied (and divided by, in some cases), or raised to a power, and still remain true. For example, from $57 \equiv 1$ (mod 8), we can conclude that

$$57^{123456789} \equiv 1 \pmod 8.$$

Hence the remainder if $57^{123456789}$ is divided by 8 is 1.

One of the many applications of modular arithmetic is an ancient error-detection technique known as *casting out nines*. This technique will be illustrated by means of an example:

$$\begin{array}{r} 74 \\ +\ 112 \\ +\ \ 89 \\ \hline 275 \end{array} \qquad \text{Now} \qquad \begin{array}{r} 7 + 4 \equiv 2 \pmod 9 \\ 1 + 1 + 2 \equiv 4 \pmod 9 \\ 8 + 9 \equiv 8 \pmod 9 \\ \\ 2 + 7 + 5 \equiv 5 \pmod 9. \end{array}$$

Notice that the numbers on the right satisfy

$$2 + 4 + 8 \equiv 5 \pmod 9.$$

Had a wrong answer (say 265) been given, we could detect it by means of such a check. The method applies also to multiplication, but in neither case is it 100% reliable. (Can you see why?)

† I.e. a common factor other than 1.

The same idea may be applied to check if a large number is divisible by 9. We can safely say that 7432891 is not divisible by 9 because

$$7 + 4 + 3 + 2 + 8 + 9 + 1 \equiv 7 \pmod 9.$$

There are many other applications of congruences. An important one is a clever handling of large integers. Before showing an example, it is necessary to introduce the following powerful result.

Chinese Remainder Theorem: Let $m_1, m_2, m_3, \ldots, m_k$ be a pairwise relatively prime set of integers all exceeding 1 (that is to say, no two of the numbers have a common factor other than 1) and let $M = m_1 \times m_2 \times m_3 \times \ldots \times m_k$. Then there is a unique x, $0 \le x < M$, which is a solution of the simultaneous congruences

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ x &\equiv a_3 \pmod{m_3} \\ &\vdots \\ x &\equiv a_k \pmod{m_k}, \end{aligned}$$

where $a_1, a_2, a_3, \ldots, a_k$ are given integers.

This theorem is used in an old mind-reading trick in which you ask someone to think of a number between 1 and 30 and give you only the remainders of the division of that number by 2, 3 and 5. If those remainders were 0, 2 and 1 respectively, to get the number x we have to solve the system

$$x \equiv 0 \pmod 2 \tag{1}$$
$$x \equiv 2 \pmod 3 \tag{2}$$
$$x \equiv 1 \pmod 5. \tag{3}$$

The Chinese Remainder Theorem guarantees the existence and uniqueness of x (observe that $M = 2 \times 3 \times 5 = 30$). Let us find it using the constructive proof of the theorem. The idea is to take the equations in pairs. From (1) we can write

$$x = 2 \times q. \tag{4}$$

To determine q, we use (2):

$$2 \times q \equiv 2 \pmod 3.$$

Since 3 is a prime number it is possible to divide by 2 both sides of this relation:

$$q \equiv 1 \pmod 3.$$

Then $q = 1 + 3r$ and by substituting this into (4) we get

$$x = 2 \times (1+3r) = 2 + 6r.$$

Using this fact and (3), we proceed as before,

$$2 + 6r \equiv 1 \pmod 5$$

which implies

$$6r \equiv -1 \equiv 4 \pmod 5.$$

Dividing both sides by 6 we have

$$r \equiv 4 \times (1+6) \equiv 4 \times 1 \equiv 4 \pmod 5.$$

Then $r = 4 + 5s$ and

$$x = 2 + 6r = 2 + 6(4 + 5s) = 26 + 30s.$$

Therefore

$$x \equiv 26 \pmod{30},$$

since we know that x does not exceed 30, we conclude that $x = 26$.

You may well think that it takes longer to work out x this way than dividing each odd number between 1 and 30 by 3 and by 5. Well, you might be right, but this is no longer true if the number were between 1 and 210 (in which case also the remainder of the division by 7 should be given) or between 1 and 2310 (210×11).

Modular arithmetic is very useful when dealing with large integers. Let us show this by means of an example. In the calculation of the determinant

$$D = \begin{vmatrix} 1234 & 965 \\ 1212 & 948 \end{vmatrix}$$

that is, in the computation of

$$D = 1234 \times 948 - 1212 \times 965,$$

we have to perform two multiplications and a subtraction:

$$D = 1234 \times 948 - 1212 \times 965 = 1169832 - 1169580 = 252.$$

We observe that although the result is in the order of hundreds, the intermediate calculations involve numbers in the order of millions. But what if we calculate the determinant in a modulo system? Take for example the modulus to be 9. In this case

$$1234 \times 948 - 1212 \times 965 \equiv 1 \times 3 - 6 \times 2 \equiv 3 - 12 \equiv 3 - 3 \equiv 0 \pmod 9.$$

This tells us that D is a multiple of 9, and this alone is not enough to determine it.

But let us assume that before calculating the determinant we have the additional information that the determinant is less than 300 and greater than zero. In that case we could operate with the modulus 300:

$$D = 1234 \times 948 - 1212 \times 965 \equiv 34 \times 48 - 12 \times 65 \equiv 132 - 180 \equiv 252 \pmod{300}.$$

Now we know that the remainder of the determinant when divided by 300 is 252. Using the additional information, we can safely conclude that $D = 252$ and in obtaining that conclusion, our intermediate calculations are bounded by the modulus 300.

A much more interesting way to obtain the same result with even smaller numbers involved is by using the Chinese Remainder Theorem. In this case we do not operate with only one modulus but with a few of them, namely 5, 7, and 11. We choose these three because $5 \times 7 \times 11 = 385 > 300$. The determinant in terms of the three systems gives

$$D \equiv 4 \times 3 - 2 \times 0 \equiv 2 \pmod 5$$
$$D \equiv 2 \times 3 - 1 \times 6 \equiv 0 \pmod 7$$
$$D \equiv 2 \times 2 - 2 \times 8 \equiv 10 \pmod{11}.$$

Now, according to the Chinese Remainder Theorem there is a unique x, $0 \le x \le 385$, which solves the system of congruences

$$x \equiv 2 \pmod 5 \tag{5}$$
$$x \equiv 0 \pmod 7 \tag{6}$$
$$x \equiv 10 \pmod{11}. \tag{7}$$

That unique solution must be D. Let us obtain it. From (5), $x = 2 + 5q$. Then, using (6) we have

$$2 + 5q \equiv 0 \pmod 7$$

and we derive

$$5q \equiv -2 \equiv 5 \pmod 7.$$

Consequently $q \equiv 1 \pmod 7$ and therefore

$$x = 2 + 5(1 + 7r) = 7 + 35r.$$

We now make use of (7)

$$7 + 35r \equiv 10 \pmod{11}$$
$$2r \equiv 3 \pmod{11}$$
$$r \equiv 3 \times (1+2) \equiv 3 \times 6 \equiv 18 \equiv 7 \pmod{11}.$$

And finally,

$$x = 7 + 35(7 + 11s) = 252 + 385s.$$

The unique value of x, that determines D, is achieved for $s = 0$. So

$$D = 252.$$

Observe that we have chosen only prime moduli so division is possible within each system.

This example, although very simple, shows the power of using modular arithmetic to avoid intermediate swell in calculations, and could be applied to any calculation involving integers where an upper bound for the final result is known. The same technique is also applied in the implementation of algorithms for exact computation involving polynomials (there is a polynomial version for the Chinese Remainder Theorem). As was shown in a previous edition of *Function* (February 1992, p. 9) the computation of a greatest common divisor of two polynomials may involve huge integers in the intermediate calculations, even when the starting polynomials and the final result are quite simple. This results in greater computer time and sometimes in the impossibility of handling those large integers.

* * * * *

HISTORY OF MATHEMATICS

EDITOR: M.A.B. DEAKIN

The article below is a version, especially prepared for *Function,* of a talk given back in 1980 to the First Australian Conference on the History of Mathematics. The proceedings of that conference were published and are still available from Professor J.N. Crossley (Department of Mathematics, Monash University) for a small sum. The original article appears on pp. 159-167 of that volume, under the title "Dialogue and Proof".

How Proofs Began

Jim Mackenzie, University of Sydney

Most cultures have some mathematical knowledge. Even so, there is still a sense in which Mathematics was invented by the Greeks of classical times. The Greeks not only knew mathematical facts, they had begun to produce mathematical *proofs*; for, so it is said, without proofs the ability to do sums and solve problems does not constitute mathematical *knowledge* and properly belongs to the prehistory of the subject. In this connection the name of Euclid is usually invoked. Euclid certainly produced a book containing *proofs*; and very good and rigorous ones they were, for it was not until the late nineteenth century that mathematicians were able to find real fault with them [1]. Not all the proofs in Euclid's book need have been original with him; he probably stitched together bits of proof coming from other Greek thinkers. But this possibility is of course consistent with the claim that the Greeks were the first to produce proofs. It is not clear that what Euclid thought he was doing was inventing Mathematics; indeed, there is a historical tradition that what Euclid set out to do was not to invent Mathematics but to provide a systematic solution to the problems of Plato's cosmology, as developed in the *Timaeus* [2]. But that possibility is also consistent with the claim that Euclid and other Greek thinkers did invent Mathematics.

But if the notion of a proof is the important contribution of the Greeks to Mathematics, if it is such an essential notion that without it mathematical discoveries are to be dismissed as premathematical, what *is* a proof and how did the Greeks come to discover (or invent) proof? Our modern notion of a proof, as a sequence of statements such that each member of the sequence is either an axiom of the system under investigation or follows from earlier members of the sequence by means of rules of inference of the system, is not of much help here. It is not clear either why such sequences of statements are important, nor how their importance could have been discovered in little communities around the Mediterranean two and a half millenia ago.

Let us have a look at how the notion of a proof could have arisen, and at what its importance would have been in that society, by examining the conversation Socrates had with Meno, as reported (whether accurately or not is of course irrelevant, for our concern is not with the particular event but with whether the people of that society were familiar with that *kind* of event) by Plato [3]. The dialogue begins with Meno asking Socrates whether virtue (or excellence – the Greek word means both) can be taught – a typical philosophical question of the time.

Socrates in this dialogue is, as usual, a very slippery debater. After giving Meno some lessons in the matter of definition, which so confuse the poor chap that he compares Socrates' effect on interlocutors with the numbing effect of torpedo (or as we would call it, electric eel), Socrates confesses that *he* does not even know what virtue is. Meno tries a familiar rhetorical puzzle of the time, intended to show that one can never learn anything: for either one knows it already, in which case one has no need to learn it, or one does not, in which case how could one recognise it when it is found? Socrates says that this is a trick argument and no good, and Meno asks, very reasonably, how it fails. Socrates, in a manner quite out of character with his subsequent reputation, appeals to the authority of theologians – of the "men and women who understand the truths of religion" – for the doctrine that the soul is immortal and has therefore seen the eternal truths and already knows them. So, Socrates argues, what is commonly called *learning*, or acquiring new knowledge, is in fact only *recollection* of what one (or one's soul) already knows but has forgotten. He then adds that to accept the conclusion that learning is impossible would make us lazy. The first of these points may answer the allegedly trick argument in the special case of eternal truths, but not in general; and the second, which adverts only to the evil effects of believing it to be valid, is no answer at all. Nonetheless, Meno does not challenge Socrates in either of these ways, but asks Socrates to teach him that learning is impossible and really only recollection. Socrates replies that *if* learning is recollection, then teaching – inducing learning – would also be impossible, and complains that Meno is trying to catch him out in a contradiction. Meno pleads ingrained habits of speech and asks Socrates to make clear in any way he can that learning is only recollection. Socrates then asks to borrow one of Meno's slaves to use in demonstration or exhibition of what he means.

We should remember here (as Klein [4] reminds us) that there are three demonstrations going on simultaneously. (a) There is the conversation between Socrates and the slave-boy, in which the latter is supposed to "learn" something and to show, by the way in which he does it, that his learning is simply recollection. (b) There is the conversation between Meno and Socrates, in which this first demonstration occurs, and in which the demonstration is supposed to enable Meno to learn (that is, to recollect) that learning is recollection. And finally, (c) there is the dialogue between Plato and ourselves, the readers, in which the whole conversation between Socrates and Meno is presented to enable *us* to learn (again, to recollect) the lesson which Plato wishes to convey about human excellence.

The problem for the slave is this: given a square of side two, what is the length of side of a square with double the area of the given square? There is a catch here; for the answer is a magnitude *incommensurable* with the given side, a matter which had then recently caused concern among mathematicians and familiar enough to Meno, who had already displayed in the discussion of definitions a knowledge of technical geometry and who makes it clear in an aside to Socrates' discussion with the slave that he knows what the answer is.

Socrates' dialogue with the slave consists of three parts. In the first he establishes that the slave speaks the same language that he does and draws the diagram (Figure 1) with two feet to each side and an area of four, and that the double square will therefore have an area of eight.

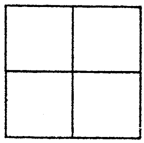


Figure 1

Then he sets the slave-boy up: "Now then, try to tell me how long each of its sides will be. The present figure has a side of two feet. What will be the side of the double-sized one?" and the boy answers, as expected, "Double". Here there is a digression between Socrates and Meno, which shows that Meno knows that this answer is wrong.

Socrates then adds to the diagram to produce a larger diagram (Figure 2) and asks whether this diagram does not contain four squares each equal to the original four-foot one; and is therefore four times as big, and sixteen, not eight, in areas; and then whether the answer must lie between two feet and four feet.

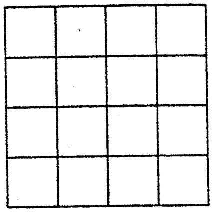


Figure 2

The boy's next guess is, naturally, three. And Socrates draws that (Figure 3) and the slave obediently counts its area as nine, and as therefore different from eight.

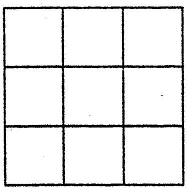


Figure 3

Then Socrates pauses to point out to Meno that the slave was confident at the beginning and is now confused as if stung by a torpedo, but nevertheless better off, for he has now come to admit his ignorance. Then Socrates goes back to the diagram of Figure 2 and draws the diagonals of each subsquare, producing a figure each quarter of which has half the area of a two-foot square (Figure 4), and asks questions to establish that therefore the new figure has just twice the area of the original and that it is, moreover, square. Socrates' point is that the slave has not been taught anything, but answered each question with his own opinion; and that if now knows the correct answer, which he did not know at the beginning, then he must have recovered it from within himself, that is, recollected it.

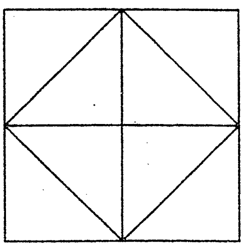


Figure 4

From the point of view of dialogue, Socrates asks two formally different kinds of questions: those which are straight bipolar ("Yes/No") questions and those which require a numerical (or, in one case, an ostensive†) answer. The latter, however, play two different roles: the original question, how long is the side of the double square, which is repeated several times, and the other numerical (and the ostensive) questions, all of which require the result of some *very* simple calculation, e.g. "What is twice two?", or the result of counting figures marked on the diagram. These are questions which, we may suppose, are very difficult to answer incorrectly; their answers are, we may say, evident from the question itself and the diagram and the use of language, they are *immediate*.

To frame a formal model for this dialogue is not difficult. We need, first of all, a language which contains statements, bipolar questions, numerical questions of the form "How many ...?", ostensive questions of the form "Which ...?" numerals and gestures of ostension ("This one"). There are three replies by the boy which do not fit any of these patterns – his emphatic assertion of ignorance (the third time Socrates has asked the original question, after the three-foot hypothesis has been refuted); his helplessness ("I do not understand" when first faced with the square formed by the four diagonals of Figure 4); and his pointing to the diagonal ("This one") as the base of the double square in answer to Socrates' ostensive question. Of these, "I do not know" is a necessary point of order in any realistic system of dialogue which permits questions, "I do not understand" is a necessary point of order still more generally in any system of dialogue in which perfect communication cannot be assumed; and Socrates' ostensive question "On which base?" which elicits the third response is about an aspect of the diagram and immediately answerable therefrom. Every question which Socrates asks, except the general question, can be answered by some fairly immediate inference from either the diagram or what the slave has already said. The interesting thing is that in the first two attempts (four-foot side and three-foot side) the answers deducible from what the boy has said and those deducible from the diagram are evidently inconsistent. We need, then, some relations of immediate inference and immediate inconsistency, which can be syntactically specified without much difficulty.

The rules for this sort of encounter provide that a dialogue begins with a question, the *problem*. Each subsequent question must be answerable either from inspection of the (evolving) diagram or from the answerer's previous remarks (since the last time the problem was put) by an immediate inference step. When the answerer is committed to an immediate inconsistency, the problem may be repeated, thereby wiping out the answerer's commitments and allowing him to begin again with a clear slate. The dialogue continues until the questioner is satisfied that the problem has been answered.

These rules are much simpler than required for ordinary Platonic dialogues – for the *Meno* itself as a whole, for instance. But there are significant parallels between the slave's progress and Meno's own progress in the larger dialogue, some of which Socrates points out in his asides, and others of which are indicated by parallelisms of expression presumably intentionally inserted by Plato. The rules are also much simpler than is required for mathematical proofs in general. There is no provision for what is commonly called "natural deduction", in which a theorem is stated and then proved; for separation of cases, for conditional proof, for constructive dilemma, for mathematical induction, and so on. All we can do are very simple *reductio* arguments; and these will be successful only if either the questioner knows exactly what he is doing, or if the participants are very lucky.

† I.e. by reference to a direct demonstration.

Nonetheless, within the very simple framework of dialogues of the slave-boy kind, Socrates' technique is devastating. There is some hope that an improved version of the same approach will give equally devastating results even in less restricted kinds of dialogue. This would be a clue to dialectical success – a crib or script which would enable a participant to succeed in any dialogue about a given subject matter.

With a couple of reservations, this is what Plato's successor Euclid was later to provide. Euclid uses most of what have since become the standard techniques of deduction. These can be partly modelled in dialogue by permitted *challenges*, i.e. locutions of the form "What is the evidence that ...?" or "Why should I accept that ...?". An obvious restriction on replies to challengers is that the answerer must not use as his reply (= ask to be granted as a premiss) anything which has already been challenged in that dialogue – he must not beg the question, in other words. This is, of course, a familiar restriction on dialogues [5]. Consider the dialogue:

n	Student	Why P_0?
$n+1$	Euclid:	P_1
$n+2$	Student:	Why P_1?
$n+3$	Euclid:	P_2
$n+4$	Student:	Why P_2?
. . .		
$n+2k-2$	Student:	Why P_k-1?
$n+2k-1$	Euclid:	P_k.

In this growing chain, Euclid cannot use any of the statements previously challenged by the student as his reply. Otherwise he would beg the question. If we write out the statements discussed in reverse order, we have:

$$P_k, P_k-1, \ldots, P_2, P_1, P_0.$$

And this, of course, is recognisable as a proof in deductive style of P_0 from Euclid's last defence, P_k. When so written, the prohibition against begging the question appears as a prohibition against using a *later* theorem in the proof of an *earlier* one – a standard rule in deductive systems, but a rule whose *raison d'être* is rarely adequately explained.

What Euclid had done, in fact, was to provide a comprehensive crib† for all dialogues about geometry for more than two thousand years. I may mention a dialogue in which Euclid successfully engaged the seventeenth century philosopher Hobbes:

> "According to his friend John Aubrey, Hobbes chanced one day to find a copy of Euclid lying open in a gentleman's library at the page containing Pythagoras' theorem. He reads the proposition. "By God", sayd he, "this is impossible". So he reads the demonstration of it, which referred him back to such a proposition; which proposition he read. That he referred him back to another, which he also read. *Et sic deinceps*†† that at last he was demonstratively convinced of that trueth. This made him in love with geometry." [6]

† A reference manual.   †† And so on.

(I should mention that the love for the subject stimulated by Euclid was on occasion blind. Hobbes later believed himself to have squared the circle.)

Euclid has engaged in dialogues with many people, sometimes causing them to love geometry, sometimes I fear causing them to hate it, but usually demonstratively convincing them of the "trueth" of its theorems. In one case, you will remember, he failed to convince an interlocutor named Lobachevsky; this unpleasant fellow did not object to any of Euclid's steps, but in effect challenged P_k, one of the last statements in the dialogue, one for which Euclid had provided no defence – the fifth postulate. (An axiom, we might say, is a statement we have been given no reason to accept.) No logician will be surprised that every crib must contain some statements for which the crib provides no defence.

So construed, Euclid's *Elements* is an elaborate defence in depth of its "last" theorem, the first which would be challenged in dialogue, the cosmologically important theorem that there are but five regular solids. This theorem might have been found even more impressive by the Greeks if they connected it with the doctrine of the five elements (earth, water, air, fire and the non-terrestrial element aether, which is what, according to them, the stars are made of). A fully mathematical chemistry (or perhaps theory of the states of matter) would have seemed almost within their grasp. And Euclid's book should, of course, be read backwards. For read in that way, it does provide a crib for conversations about its subject matter. In conversations, even competitive conversations, the Greeks were very interested. This interest was encouraged by the fact that skill in persuasion was important for the ambitious in their political system, the system which gave us the word democracy. Aristotle, who came between Plato and Euclid in time and whose considerable contribution to the development of deductive argument I have not had time to include in this sketch, mentions those who dispute "to the death" [7]. Even allowing that stodgy old Aristotle must have been exaggerating there, the Greeks clearly took dialectical disputation very seriously. Thus the possibility of cribs which would guarantee success in conversations about a certain subject matter would excite them; and these cribs are what we take to be the distinctive Greek contribution to mathematics, *proofs*.

References

1. HILBERT, D., *Grundlagen der Geometrie*, Leipzig, 1899.
2. POPPER, K. R., *Conjectures and Refutations* (London, 1969), p.88, citing Proclus (*Procli Diadochi in primum Euclidis Elementorum librum commentarii*, ed. G. Friedlein (Leipzig, 1873), Prologus II, p.71, 2-5.
3. PLATO, *Protagoras and Meno*, trans. W.K.C. Guthrie (Harmondsworth, Middlesex, 1956) is the most freely available edition. References will be given by Stephanus numbers, which are indicated in most editions.
4. KLEIN, J., *A Commentary on Plato's Meno* (Chapel Hill, North Carolina, 1965).
5. MACKENZIE, J.D., "Why do we number Theorems?", *Australasian J. Philosophy* LVIII (1980), 135-149.
6. KNEALE, W. and M. KNEALE, *The Development of Logic* (Oxford, 1962), p.311.
7. ARISTOTLE, *Sophistical Refutations* iii, 165b13.

* * * * *

PROBLEMS AND SOLUTIONS

EDITOR: H. LAUSCH

We cordially welcome our subscribers in 1993 and hope they will enjoy this section of Function *by trying their hands and minds on the problems or by puzzling fellow subscribers with their own problems. Let us start the new year lightheartedly, with a problem that Professor Phill Schultz (University of Western Australia) has communicated to* Function.

Problem 17.1.0 What is a computer scientist?

Solution. A computer scientist is a person who does not know the difference between Halloween and Christmas: OCT 31 = DEC 25.

Solutions to other problems

Here are solutions to some problems that were published in the two most recent volumes of Function.

Problem 15.1.9. Prove that the product of the first fifty odd numbers is less than a tenth of the product of the first fifty even numbers.

Solution (John Barton, North Carlton). We put

$$\rho = \frac{2.4.6.8. \ \dots \ .2n \ . \ \dots \ .100}{1.3.5.7. \ \dots \ .(2n-1). \ \dots \ .99} \cdot$$

Then

$$\rho = \frac{2}{\sqrt{1}.\sqrt{3}} \cdot \frac{4}{\sqrt{3}.\sqrt{5}} \cdot \frac{6}{\sqrt{5}.\sqrt{7}} \cdot \dots \cdot \frac{2\nu}{\sqrt{2\nu-1}\cdot\sqrt{2\nu+1}} \cdot \dots \cdot \frac{98}{\sqrt{97}.\sqrt{99}} \cdot \frac{100}{\sqrt{99}} \cdot$$

The general factor $\dfrac{2\nu}{\sqrt{2\nu-1}\cdot\sqrt{2\nu+1}} = \dfrac{2\nu}{\sqrt{(2\nu)^2-1}}$ for each ν. Hence $\rho > \dfrac{100}{\sqrt{99}} > 10$.

John Barton adds: "The result is fairly 'fine' in the sense that 10 cannot be increased by very much. We can see this by using Stirling's formula:

$$n! \sim e^{-n} \cdot n^{n+\frac{1}{2}} \cdot \sqrt{2\pi};$$

$$2.4.6. \ \dots \ . \ 100 = 2^{50} \cdot 50!$$

$$1.3.5. \ \dots \ . \ 99 = \frac{100!}{2^{50} \cdot 50!} \cdot$$

$$\textit{Hence} \quad \rho = \frac{2^{100}(50!)^2}{100!}$$

$$\sim \frac{2^{100} \cdot e^{-100} \cdot 50^{101} \cdot 2\pi}{e^{-100} \cdot 10 \cdot 100^{100} \cdot \sqrt{2\pi}}$$

$\sim 5\sqrt{2}\pi$

$\sim 12.533.$

Problem 15.3.4 (University entrance examination, Akademisches Gymnasium Salzburg, Austria). Given are two circles K_1 and K_2, with centres C_1 and C_2 respectively, same radius R and a common tangent t such that K_1 and K_2 lie on the same side of t. Let t touch K_1 at T_1 and K_2 at t_2. Let K_3 be another circle, with centre C_3, radius R and tangent to K_1 and K_2 such that K_1 and K_2 lie in the exterior of K_3. How should the circles be placed in the plane in order to make the area of the pentagon $C_1T_1T_2C_2C_3$ as large as possible? Express area and perimeter of this pentagon as functions of R.

Solution (modified version of a solution submitted by Seung-Jin Bang, Republic of Korea). Let $\theta = \angle C_1C_2C_3$. Then the area of $C_1T_1T_2C_2C_3$ is

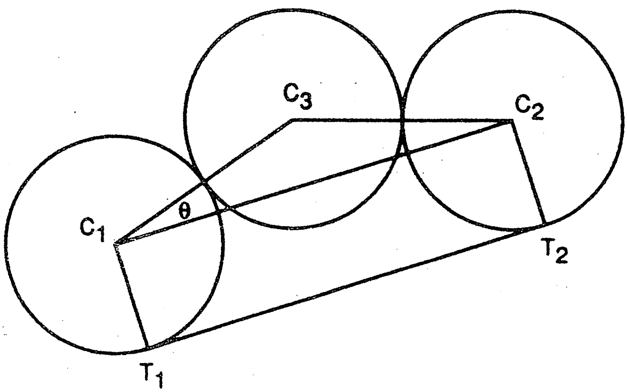
$$f(\theta) = 4R^2\cos\theta(\sin\theta + 1),\ 0 < \theta \le \frac{\pi}{3}.$$

Since $f'(\theta) = -4R(2\sin^2\theta + \sin\theta - 1)$, $f(\theta)$ has a maximum at $\sin\theta = \frac{1}{2}$, i.e. at $\theta = \frac{\pi}{6}$.

Answer: $\angle C_1C_2C_3 = \frac{\pi}{6}$.

Area: $3R^2\sqrt{3}$,

Perimeter: $2R(3 + \sqrt{3})$.



Problem 16.4.2 (adapted to Australian conditions from Alexander Yakovlevič Halameiser's entertaining reader *Mathematics ? – Entertaining !*, Moscow 1990). "Some years ago", remembered Hansel, "I encountered a Tasmanian tiger in the Dandenongs when I had a barbecue with Red Riding Hood. It was on April 1." "Really", laughed Gretel mockingly, "and you would, of course, remember the day of the week on which you had your very strange encounter with the big bad wolf, ... pardon me, I mean ... with your Tasmanian tiger and that girl whatshername." "I am afraid I don't remember the day of the week", came Hansel's reply, who felt quite embarrassed, "but wait, ... I do remember that there were three Sundays in this month falling on even-numbered days. Is that of any help to you?" "It is indeed", said Gretel, making herself sound very important. "I don't believe a word of your Tasmanian tiger story, but at least I now know the day of the week on which you and that girl whatshername had a barbecue in the Dandenongs! It was on a ... ." She could not finish her sentence because a ferocious looking animal was sitting down beside her and ... .

What day of the week was Gretel just about to mention?

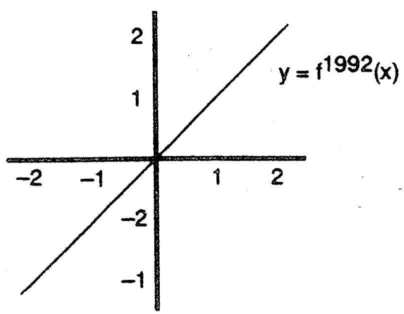
Solution (Zeev Ragadol, Doncaster). Since no two successive Sundays can fall on even-numbered days, a month with three Sundays falling on even-numbered daus must have five Sundays, the first Sunday of the month falling on an even-numbered day. Hence, in any such month this Sunday must fall on the second day of this month. – The fourth day would be "too late", as the third Sunday of the month would then fall on the thirty-second day of this month. – Therefore Gretel was just about to say: "Saturday!"

Problem 16.4.5 (from the Slovenian magazine *Presek*, Volume 18 (1990/91), issue #2). Let f be the function which is defined by $f(x) = \frac{x-1}{x+1}$. For each positive integer n, let

$$f^n(x) = \underbrace{f(f(\ldots f(x)\ldots))}_{\text{n times}} .$$

Draw the graph of $f^{1992}(x)$.

Solution. $f^2(x) = \dfrac{\frac{x-1}{x+1} - 1}{\frac{x-1}{x+1} + 1} = -\frac{1}{x}$, hence $f^4(x) = -\dfrac{1}{-\frac{1}{x}} = x$. It follows that $f^n(x) = x$ if 4 divides n. Since 4 divides 1992, we have $f^{1992}(x) = x$. The graph of $f^{1992}(x)$ is thus



PROBLEMS

Towards the VCE: multiple-choice year-twelve problems

"What is the VCE ?", Function *readers, indeed, may wonder – especially those subscribers who live outside Victoria. The abbreviation stands for Victorian Certificate of Education. Normally Victorian secondary students obtain this certificate at the end of their year 12. Admission to Victorian universities is usually based on the score shown on the certificate. Here are a few multiple-choice questions addressing the subject "cluster".*

Change and Approximation: Integration Techniques

(a) An anti-derivative of $\cos(x^2 + 1)$ is:

A $\dfrac{\sin(x^2+1)}{2x}$

B $-\dfrac{\sin(x^2+1)}{2x}$

C $2x \cdot \sin(x^2+1)$

D $-2x \cdot \sin(x^2+1)$

E $\sin(x^2+1)$

F none of the above.

(b) The derivative of $\dfrac{\sin(x^2+1)}{2x}$ is:

A $\cos(x^2+1)$

B $\cos(x^2+1) - \dfrac{1}{2x^2}\sin(x^2+1)$

C none of the above.

(a) $\displaystyle\int_4^9 e^{\sqrt{x}}dx$ is:

A $e^3 - e^2$

B $3e^3 - 2e^2$

C $\dfrac{2e^3-3e^2}{12}$

D $6e^3 - 4e^2$

E $4e^3 - 2e^2$

F none of the above.

(b) The derivative of $2\sqrt{x}e^{\sqrt{x}}$ is:

A $e^{\sqrt{x}}$

B $e^{\sqrt{x}}(1 + \dfrac{1}{\sqrt{x}})$

C none of the above.

(a) $\displaystyle\int \frac{x^4}{x^5+1}dx$ can be rewritten as:

A $\displaystyle\int\frac{1}{u}du$

B $\displaystyle\int\frac{u^4}{u^5+1}dx$

C $\displaystyle 5\int\frac{1}{u}du$

D $\displaystyle \frac{1}{5}\int\frac{1}{u}du$

E $\displaystyle 1993\int\frac{1}{u}du.$

(b) If $u = \sqrt[5]{x^5+1}$, then $\displaystyle\int\frac{1}{u}du$ equals

A $\displaystyle\int\frac{x^4}{x^5+1}dx$

B some other integral of the form $\displaystyle\int f(x)dx$

If $f'(x) = \frac{1}{1-2x}$ and $f(\frac{1-e}{2}) = 1$, then $f(x)$ is equal to:

A $-\frac{\log_e(1-2x)}{2} + 1$

B $-\frac{\log_e(1-2x)+3}{2}$

C $\frac{2}{(1-2x)^2} - \frac{1-e}{2}$

D $-\frac{1}{(1-2x)^2} + \frac{e^2+1}{e^2}$

E $-\frac{\log_e(1-2x)-3}{2}$

F none of the above.

(a) $\int_{3\pi/4}^{\pi} \tan^2(x)dx$ is equal to:

A $\frac{1}{3}$

B $-\frac{1}{3}$

C $-\frac{\pi+4}{4}$

D $-\frac{\pi-4}{4}$

E 2

F none of the above.

(b) If f is a function such that $f(x) \geq 0$ for all values of x between $\frac{3x}{4}$ and π, and $\int_{3\pi/4}^{\pi} f(x)dx$ can be evaluated, then

A $\int_{3\pi/4}^{\pi} f(x)dx \geq 0$

B it is possible that $\int_{3\pi/4}^{\pi} f(x)dx < 0$.

(a) $G'(x) = g(x)$ and $g(x) = (1-x^4)^3$.
Hence $G(x)$ is:

A $-\dfrac{(1-x^4)^4}{16x^3}$

B $-\dfrac{(1-x^4)^4}{x^3}$

C $-4x^3(1-x^4)^4$

D $-\dfrac{3(1-x^4)^2}{4x^3}$

E $-12x^3(1-x^4)^2$

F $x - \frac{3}{5}x^4 + \frac{1}{3}x^9 - x^{13}$

G none of the above.

(b) Suppose that f is a function which has a derivative, f'. Then the statement "$f(x)/[4f'(x)]$ is an anti- derivative of $f(x)^3$" is

A always true

B sometimes true

C never true.

* * * * *

COLOURED CAPS AND RIGHT-HAND TURNS

This account of an old chestnut follows the version first printed many years ago in the education pages of the now defunct Melbourne morning paper, *The Argus*.

Three boys, Veale, Merrick and Wizzer, are seated in such a way that Veale can see both Merrick and Wizzer, Merrick can see Wizzer but not Veale, and Wizzer can see neither of the others. Each boy has on his head a cap which is either red or blue and it is known that the three caps have been drawn from a pool of five: 3 blue, 2 red.

A fourth boy (Herring, in the *Argus* account) asks Veale if he knows what colour his own cap is – but Veale is unable to say. Herring next asks Merrick, who likewise is unable to say what colour *his* cap is. Wizzer, when in turn *he* is asked, answers correctly. How? What is his answer?

Well, the reasoning goes as follows. If Veale had seen two red caps, he would have known that his own cap was blue. But he did not know this – hence he saw at least one blue cap. If now Wizzer's cap were red, then Merrick would have been able to deduce that the cap on his own (Merrick's) head was blue. However, Merrick did not know the colour of his own cap. This allowed Wizzer to conclude that *his* own cap could not be red and must therefore be blue.

Problems like this – in which a deduction depends upon somebody else's deductive abilities – abound. Problem 1.2.6 of *Function* is a case in point. We give it again here.

A person A is told the product xy and a person B is told the sum $x + y$ of two integers x, y, where $2 \leq x \leq 200$, $2 \leq y \leq 200$. A knows that B knows the sum and B knows that A knows the product. The following dialogue develops:

A: I do not know $\{x, y\}$.
B: I could have told you so!
A: Now I know $\{x, y\}$.
B: So do I.

What is $\{x, y\}$?

This truly beautiful problem was solved by several readers and the solution appeared in *Volume* 2, *Part* 1. We will not give it here, but you may care to try to find it out for yourself.

However, it seems to me that such problems have some limited, but not total, application in real life. Wizzer, in the first problem, relies on deductions by both Veale and Merrick. Now, admittedly, Veale, if he saw two red caps, would have to be pretty thick not to realise that his own cap was blue. So Merrick would be on firm ground in his own deduction that either his or Wizzer's cap was blue. It isn't a difficult deduction to make – but Merrick might not have made it. It takes rather more sophistication than Veale required, and Merrick may not have possessed this.

Suppose, in other words, that Wizzer had worn a red cap and Merrick a blue one. Veale would have had no way of knowing the colour of his own cap and this should have told Merrick that he (Merrick) had on a blue cap. However, if Merrick failed to make this deduction, then Wizzer would have concluded wrongly.

This is ironical in a way. *The Argus* had Herring using this demonstration to show how bright Wizzer was (brighter than Merrick or Veale). In fact, Wizzer could only demonstrate his own intellectual ability by relying implicitly on the intelligence of his classmates.

When it comes to the dialogue between A and B, the point has even more force. By B's first comment ("I could have told you so!") we are already attributing considerable mathematical ability to B (and some mathematical skill to A). And B knows that this still exists and may be relied upon; the subsequent dialogue depends on this point. The later comments demand more, and indeed all comments would have had to be preceded by lengthy thought and calculation. The "dialogue" could not have taken place in real time.

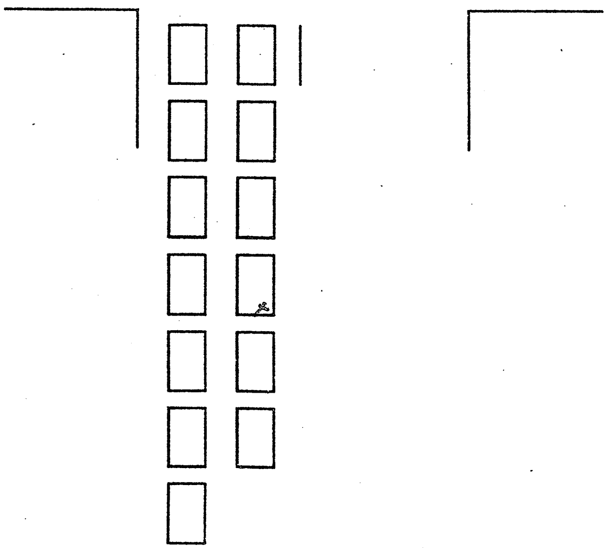
Thus there are limitations to the sort of argument being discussed here. In real life, people do not always make deductions that in theory are there to be made. We cannot always rely on the Merricks of this world to get things right.

Nonetheless, limited though they may be, such arguments can be useful in practice. I employ such reasoning at times when I'm driving. It not uncommonly happens that two lanes of traffic pull up at a red light. It is usually the case that the right-hand lane is the faster of the two, and so, if I'm anxious to cut down my travel time, I would choose to drive in that lane.

However, there is a complication. The right-hand lane is also the lane for those cars wishing to turn right and *these* cars must give way to oncoming traffic and so they delay the cars behind them. If I'm in a hurry, I don't want to have to wait while the cars on my left merrily go on their way, but with me stuck behind a stationary car.

Now suppose I draw up behind a line of cars. If any one of them wants to turn right, then I may be delayed. But often the cars between me and such a car will block my vision so that I'm unable to see the indicator and won't know if the driver intends a right-hand turn or not.

See the figure which illustrates the case in which there are 6 cars already in line and Car No. 4 wants to turn right. This is not evident, however, as Cars No. 5 and No. 6 block the view.



In practice such a situation is relatively rare. If, to start at the front, Car No. 1 were indicating a right-hand turn, then Car No. 2 would not draw up behind it if it wished to proceed straight ahead. Only if Car No. 2 wished to turn right would it draw up behind a right-turning Car No. 1; and in *this* case, Car No. 2 would indicate a right turn also.

Then Car No. 3 could apply the same logic. If Car No. 3 wishes to proceed straight, then its driver will look at Car No. 2 – a right-turn indicator would be a signal not to stay in the right-hand lane. But suppose there were no such signal. Then the driver could argue "Car No. 2 sees no impediment to forward travel, so Car No. 1 can't be turning right".

And so the argument continues on down the line. So in the diagrammed position, Car No. 5 has made a poor decision. This could be due to stupidity or to inattention. It may not have been possible to get into the left lane at the appropriate time. Car No. 4 may have one of those bad drivers who are late in indicating turns. And so on.

Nonetheless, a useful rule is to look at the car in front and argue forward. "Car No. 6 thinks Car No. 5 thinks, etc.".

It doesn't always work, but I usually get away with it.

* * * * *

How about this one?

It is said that a student, trying to solve the quadratic equation

$$(x + 3)(2 - x) = 4$$

argued as follows.

$$\begin{aligned} (x + 3)(2 - x) &= 4 \\ \Rightarrow \text{Either } \; x + 3 = 4 \;&\text{ or }\; 2 - x = 4 \\ \Rightarrow \text{Either } \; x = 1 \;&\text{ or }\; x = -2. \end{aligned}$$

The reasoning is of course incorrect. The conclusion is not, for 1 and -2 are indeed the roots of the original equation.

You may care to investigate the circumstances under which the student's "method" works.

In any case, it makes a nice illustration of the mathematical principle that a right answer is not everything: it also matters how you reach it.

* * * * *

FUNCTION is a mathematics magazine addressed principally to students in the upper forms of secondary schools.

It is a 'special interest' journal for those who are interested in mathematics. Windsurfers, chess-players and gardeners all have magazines that cater to their interests. FUNCTION is a counterpart of these.

Coverage is wide — pure mathematics, statistics, computer science and applications of mathematics are all included. Recent issues have carried articles on advances in mathematics, news items on mathematics and its applications, special interest matters, such as computer chess, problems and solutions, discussions, cover diagrams, even cartoons.

* * * * *

Articles, correspondence, problems (with or without solutions) and other material for publication are invited. Address them to:

The Editors,
FUNCTION,
Department of Mathematics,
Monash University,
Clayton, Victoria, 3168.

Alternatively correspondence may be addressed individually to any of the editors at the mathematics departments of the institutions listed on the inside front cover.

FUNCTION is published five times a year, appearing in February, April, June, August, October. Price for five issues (including postage): $17.00*; single issues $4.00. Payments should be sent to the Business Manager at the above address: cheques and money orders should be made payable to Monash University. Enquiries about advertising should be directed to the business manager.

*$8.50 for *bona fide* secondary or tertiary students.

* * * * *