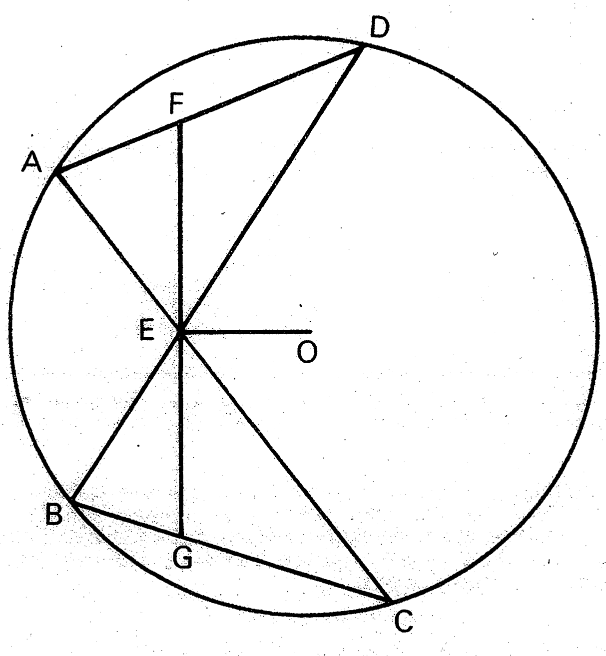
Function

Founder Editor G. B. Preston

Volume 15 Part 1 February 1991



A SCHOOL MATHEMATICS MAGAZINE

. by Aust. Post Publ. No. VBH0

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. There are articles on recent 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 back cover.

FUNCTION is published five times a year, appearing in February, April, June, August, October. Price for five issues (including postage): $16.00*; single issues $3.50. 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.

*$8.00 for *bona fide* secondary or tertiary students.

* * * * *

FUNCTION

(Founding editor: G.B. Preston)

Volume 15 *Part* 1

CONTENTS

* * * * *

Published by the department of Mathematics, Monash University

About Function

Towards the end of 1976, Professor G.B. Preston called a meeting at the Monash University department of Mathematics. He explained to those who attended that Victoria had reached a situation of having no journal of school mathematics, written for school students and giving access to good quality exposition of high-calibre mathematical material. Such journals, he said, were the norm both in the U.K. and in many countries of continental Europe; indeed, other Australian states had them. By contrast, the Victorian journals that once might have played this rôle had instead turned to other things. He saw a need for a specialist journal in this area.

Thus *Function* was born. Its name was one of many supplied to Professor Preston from a variety of people enthused by the idea of such a journal. The "supplier" in this instance was the late Dr. Len Grant, then in the Monash department of Philosophy. Professor Preston singled out this suggestion in preference to all the others because, as he said, "I can think of no idea more central to the whole endeavour of Mathematics". Thus *Function* was christened.

Over the years since then, *Function* has continued to fulfil the aims laid down by that first meeting. The founder editor was Professor Preston himself and he remained active on the editorial board throughout the 14 years that followed – jointly editing *Volume* 14, *Part* 5 with Dr. Rod Worley.

His retirement at the end of 1990 thus sees the end of an era, but *Function* itself continues and remains dedicated to the same ideals as those listed by Professor Preston back in 1976. Essentially these are:

(a) valid and interesting mathematics
(b) high-quality exposition
(c) access to all.

In pursuit of the third of these aims, *Function* has always tried to encourage Mathematics among girls.

The need for an organ like *Function* has recently been significantly increased by the introduction of the new VCE, covering precisely the years (11 and 12) which *Function* is designed to serve.

Late in 1990, the *Mathematical Association of Victoria* printed a compilation of some of its most useful articles under the title *Composite Function* (see p.14). The articles reproduced there were chosen for their especial suitability for student projects. Each year, *Function* publishes more articles, problems, news-items (even cartoons) in the same vein. Recent issues have discussed symmetries in Physics, fractals, supercircles, multi-dimensional chess, planetary paths and Mathematics for girls.

We hope teachers and students alike will continue to use *Function*, that it finds its way into more and more school-libraries and that, in particular, students and teachers write for it. *Function* can be a very good place to send that good project for wider circulation and appreciation, or to ask about that question that's always bugged you; and so on.

For details on whom to contact and how to go about it, see the inside front and back covers.

Michael A.B. Deakin
Chief Editor

THE FRONT COVER

The front cover shows a diagram used to prove and to follow a geometrical theorem known (from the appearance of the figure) as the Butterfly Theorem. It is a theorem found (probably) last century, but provable by Euclidean methods.

A, B, C, D are any four points on the circumference of a circle whose centre is 0. The line AC and the line BD intersect at E. The line $E0$ is now drawn and, perpendicular to it and passing through E, the line FG, where F lies on AD and G on BC.

The theorem states that $FE = GE$.

For a (relatively) simple proof, although the statement and notation differ a little from that given here, see *Geometry Revisited* by H.S.M. Coxeter and S.L. Greitzer (New Mathematical Library, 1967). This proceeds by dropping perpendiculars from F onto AE and DE and from G onto BE and CE. This sets up four new triangles which are then shown to be similar in pairs. By operating algebraically on the equalities so obtained, Coxeter and Greitzer show that

$$\left[\frac{EF}{EG}\right]^2 = \frac{AF \cdot DF}{CG \cdot BG}$$ (after some further work with similar triangles).

Further geometric identities are then combined with more algebra to show that this latter ratio equals 1.

The authors remark that this proof is only one of many but is simple and easy to remember. (Readers of *Function* who take the trouble to look it up, or to reconstruct it from the brief description above may not entirely agree.) They also mention a number of other proofs, one attributed to the 19th-century amateur mathematician, William George Horner, and apparently still unpublished.

Elsewhere in their book, they note that the use of coordinate geometry to prove such theorems may be much more difficult than the search for elegant Euclidean arguments.

This judgement may now need review. The Butterfly Theorem is one of a large number proved by algebraic means on a computer by Shang-Ching Chou of the University of Texas following techniques developed by Wen-Jun Wu of Beijing. (More recently, a new theorem has apparently been *discovered* by a computer using Wu's programmes. See p.8.)

In the case of the Butterfly Theorem, what is required is a simplification of seven simultaneous quadratic equations in eleven unknowns. Not to be attempted by hand!

The statement of the Butterfly Theorem given above and the description of how it can be set up for computer proof are taken from a report by Dr. Chou and kindly relayed to us by Garry Tee of the University of Auckland. It raises, as others have noted in the past, questions about the nature of proof, when theorems are proved "inside" a machine and do not lend themselves to human checking. However, in the case of the Butterfly Theorem, such checks are available.

* * * * *

A NUMERICAL STREET-CLEANING QUALITY MEASURE

Neil S. Barnett, Victoria University of Technology

You may have noticed the word 'quality' appearing more frequently on product labels these days, and particularly becoming more prevalent in advertising slogans. The dictionary definition of 'quality' has long since been dropped and in the manufacturing industries rather than referring to absolute performance has become has become tagged to that often nebulous property described as that 'thing' that keeps customers happy. Cost, of course, is very much a product feature that has the potential to upset or satisfy a customer, so disregard for product cost can only lead to 'poor quality'!

To produce 'quality', therefore, requires a careful blend of performance, reliability, appearance and cost. Much pressure is being exerted on the manufacturing industry to meet these 'quality' requirements. To achieve this has meant considerable upheaval in many industries, involving a change of emphasis not only in operations but in management practice. Many of these changes have been seen to have potential benefit in the more service-oriented industries, which of course include many government services, agencies and departments that have a predominantly public interface. In all this, one major problem is to quantify customer satisfaction. This then has to be defined and tested to see if it does in fact measure that which is intended.

A student in one of my postgraduate classes works for one of Melbourne's local suburban municipal councils and he approached me recently regarding a project that he is overseeing. The project is aimed at assessing the quality of street-cleaning performance in the municipality for which he works. So the push to improve quality has now reached street level! His reason for discussing the issue with me was that we wanted to develop a measure of performance so that he could quantitatively measure quality performance, make meaningful comparisons and gauge 'quality' improvement.

Now, as a statistician I have a certain aversion to providing the uninitiated with 'magical' measures because I have seen far too many misuses of and inappropriate deductions made from them. Be that as it may, simple numerical descriptors and indices are very much in use all around us. We hear much about the consumer price index. In manufacturing process control, much discussion revolves around process capability indices and in statistics we have the correlation coefficient, and of course the mean and variance are both numerical descriptors. All are an attempt to reduce a large amount of data to a meaningful single numerical value that measures an important facet of the data and by implication an important feature of the whole of which the data is merely a part.

There is, of course, a price to pay when radically condensing data. Whilst the creation of such measures stems from the natural human desire for simplicity, simple results are usually open to interpretation. Unscrupulous manipulators of data play this game to great advantage – which usually means to others' disadvantage! You can, therefore, appreciate my reticence when asked to formulate an appropriate measure that in effect would say of road-sweeper A: you have not performed as well (quality-wise) as road-sweeper B.

Now, when discussing numerical descriptors there are two distinct aspects at issue:

(i) the conceptual issue of providing some useful measure; and
(ii) since in most instances the actual values will need to be estimated, how do estimated values relate to the actual values themselves?

For illustration, consider the familiar concept of population mean, μ. This has a conceptual interpretation of measuring centrality and it must be calculated over the population as a whole. In any practical application, however, we have only sample values and so must estimate μ from them; for example, if the sample values are $x_1, x_2, \ldots, x_n$ then the sample mean $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ and this is an estimate of μ.

Statisticians have spent considerable energy over the years posing and answering questions of the type: "What properties does $\bar{x}$ have that make it a 'good' estimator of μ?" "What do I mean by a 'good' estimating procedure?" And so on.

Calculating $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ is a 'data collapsing' exercise since n numbers are reduced to just one. We are also familiar with the concept of the variance of a random variable. If the variable is of the same type as in the foregoing, then the variance σ^2 must be calculated over the entire population but a sample estimate is given by

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

This is another data collapsing or summarizing calculation. Conceptually, the variance provides a comparative measure of distribution spread about its mean and again the issue of merit of using s^2 to estimate σ^2 arises. These two estimates of the population indices or parameters tell us much more than one on its own, but there is still a lot about the distribution that remains untold.

Now when people request a suitable descriptive measure, it has been my experience that they don't want the complication of more than one. They want one and only one and they want it to answer all questions unequivocably. If only life was that simple! So throwing caution to the wind I set about creating, as if by magic, that numerical combination that would condemn road-sweeper *B* as a lesser mortal than sweeper *A*. In all fairness, considerable effort had been expended in defining quality and in collecting and processing data. Fundamental to the whole exercise was the five-level grading of road cleanliness 1, 2, 3, 4, 5, ranging from a unit value of 'yuk' to a 'squeaky-clean' value of 5. (The descriptors are my own!) This rating, of course, is very subjective but just one person was involved in grading at the initial stage so it was hoped that this subjectivity was minimal and would reduce with experience. Several steps had been taken to eliminate obvious factors that would unfairly affect the rating.

Data collection involved sampling the streets cleaned by a particular crew and assessing their cleanliness (on the five point scale) both before cleaning and immediately afterwards. Tables of frequencies of gradings before and after cleaning were then assembled and from these it was hoped to define a suitable quality' measure. The table below is one example of the frequencies obtained:

Date	Number of Samples	Rating Frequencies					
			1	2	3	4	5
13/3/90	60	B	11	30	14	5	0
		A	4	22	21	13	0

The issue was compounded by the fact that in a number of instances the number of streets examined after cleaning was less than those examined before. Also, the frequency tables don't trace an individual street for performance before and after. Despite these deficiencies, do the frequency tables tell us anything useful?

If we denote the frequencies before cleaning by f_1, f_2, f_3, f_4 and f_5 and after it by F_1, F_2, F_3, F_4, F_5 then a 'quality' cleaning performance would be evidenced by, for example, having $F_4 > f_4$, $F_5 > f_5$, and $f_1 > F_1$, $f_2 > F_2$, i.e. high frequencies are shifted along from the lower ratings to the higher, following cleaning.

As a measure,

$$\left(f_1 - F_1\right) + \left(f_2 - F_2\right) + \left(F_4 - f_4\right) + \left(F_5 - f_5\right)$$

has some merit in that the larger the result, the 'better' the performance. However, such raw calculations can favour larger samples rather than improved performance and what about the significance of the relationship between f_3 and F_3?

Besides, what if

$$\sum_{i=1}^{5} f_i \neq \sum_{i=1}^{5} F_i?$$

Perhaps rather than consider raw frequencies, we ought to consider relative frequencies,

$$\frac{f_i}{\sum_{i=1}^{5} f_i} \text{ and } \frac{F_i}{\sum_{i=1}^{5} F_i},\ i = 1,2,\ldots,5.$$

There are still many flaws in this as a basis for describing quality, not least of all the lack of significance of moving frequencies into the immediately adjoining rating category. Without further pursuit, it is painfully obvious that this approach is doomed to failure. Now there are some standard statistical procedures for handling frequencies, but here maximum achievable performance in shifting frequencies to the right is very much dependent on the f_i's. Also, we are not just concerned with establishing the existence of a shift to the right but with quantifying this shift meaningfully.

Working with frequencies alone is fraught with all manner of risks of mis-interpretation, so I recommended looking at individual street performance, examining these and scoring them on the following basis:

	Rating Before	Rating After	Score
Street	i	j	$j-i$

Individual street performances range from $-4 \rightarrow 4$. On no account will the full range of possible scores be available.

For any given i there are only 5 possible scoring options. Implicit in this method of scoring is that the effort to shift from rating 1 to rating 2 is compatible with the effort to shift from rating 4 to rating 5 which, if I were a road sweeper (and I'm not), I would probably object to. Similarly, the scoring scheme equates moving from 1 to 3 with moving from 3 to 5, etc.

Consider now, ignoring these niceties, that a sample of streets before cleaning has produced the frequencies f_1, f_2, f_3, f_4 and f_5. The maximum possible score on the above basis for cleanliness improvement is $4f_1 + 3f_2 + 2f_3 + f_4$.

If the actual score achieved after cleaning is S then an index of performance that has some merit is:

$$\frac{S}{4f_1+3f_2+2f_3+f_4}$$ where S is scaled on the basis of the maximum performance possible.

The index has a maximum possible value of 1 and a minimum possible value of

$$\frac{-(4f_5+3f_4+2f_3+f_2)}{4f_1+3f_2+2f_3+f_4}$$

although realistically we would not expect it to fall below 0 (unless rubbish was being deliberately emptied into the streets)!

For comparison purposes it would seem most equitable to ensure that the number of streets sampled remains near enough constant, since the values of f_i are a function of both sample size and frequency distribution.

An additional question also springs to mind and that pertains to issue (ii) considered earlier. What does $\frac{S}{4f_1+3f_2+2f_3+f_4}$ say about $\frac{\Sigma}{4F_1+3F_2+2F_3+F_4}$, the index considered over all the streets cleaned by this same crew in, say, a week? Does it in fact have to say anything about this? (Note F_i being used here differently than previously.)

Doesn't $\frac{S}{4f_1+3f_2+2f_3+f_4}$ stand on its own as a measure of performance?

The scoring basis can of course be adjusted to allow for the extra effort that is, in all likelihood, expended in going from grade $1 \to 2$ as opposed to going from grade 4 to 5. Of course, what the index doesn't broach is the issue of time and efficiency and failure to consider this will undoubtedly lead to injustice and the making of unreasonable comparisons.

One measure cannot possibly answer all questions, so it is dangerous and usually unjust to evaluate performance on just a single measure alone. That does seem obvious, but in a complex world – a world of data 'overkill', the craving for simplicity often overshadows justice and equity. Can you think of a better measure than mine?

* * * * *

A COMPUTER-GENERATED THEOREM IN ELEMENTARY GEOMETRY?

Michael A.B. Deakin, Monash University

Computers are now indispensable in Mathematics and through their use we now know things that we could not know without them. The most celebrated example is perhaps the proof of the Four Colour Theorem (see *Function*, *Vol.* 1, *Part* 1, p.9), but this is really only one case among a very great many.

Nevertheless, it is often claimed that computers have the power not only to compute and check with great accuracy, but also to "think" heuristically and so to take over not merely the hackwork but some of the creative rôle of the mathematician as well.

To some extent this also happens already – computer chess programs, while not infallible, can be very hard to beat. But this is not quite a comparable case. From time to time, the following question is raised: *Is there a clear case of a computer's generating a new theorem in elementary geometry? Or perhaps of producing a genuinely new proof of a previously known such theorem?*

Every so often, claims are made that such an event has occurred. Each time I have looked at these claims, the result has turned out to be previously known, not in fact by the computer at all, or trivial in the extreme. One of my regular correspondents, Garry Tee of the University of Auckland, writes that he has checked many such claims and these criticisms apply, severally or together, to every one of them.

Recently, however, I've learned of a case that may be an exception. I'll tell you what I know of it, which is not everything, but probably enough. It then becomes a matter for your judgement. The story goes as follows.

Recently I attended the First Asian Mathematical Conference. This was held in Hong Kong in August of 1990, and one of the principal speakers was Professor Wu Wenjun of the Academia Sinica, Beijing. Professor Wu's special expertise lies in the field of computer proof and here is one of his examples.

Look at Figure 1 opposite. Z, X, Y, U, V are the vertices of a square pyramid whose vertical axis OZ is perpendicular to the base $XYUV$. We now imagine a plane cutting through the pyramid. As Figure 1 illustrates, this plane may so intersect the various edges and faces of the pyramid as to form a pentagon $PQSTR$. The question is: *Can* (*and if so in what circumstances*) $PQSTR$ *be a regular pentagon*? That is to say, can it happen that the sides PQ, QS, ST, TR, RP are all equal and that the angles PQS, QST, STR, TRP, RPQ are also all equal; if this can come about, how and when is it achieved?

This question, incidentally, was not asked by a computer, but by a mathematician. This may be an important point, depending on quite what view you take of mathematical discovery.

To answer the question, we need to use 3–D coordinate geometry if we are to have that answer supplied by a computer. In my own analysis of the matter, I chose the coordinate system of Figure 1 and put $OX = OY = OU = OV = 1$ and $OZ = a$. Clearly this is not the only choice available, but it seemed to me to be the simplest that still retained all the generality of the original problem.

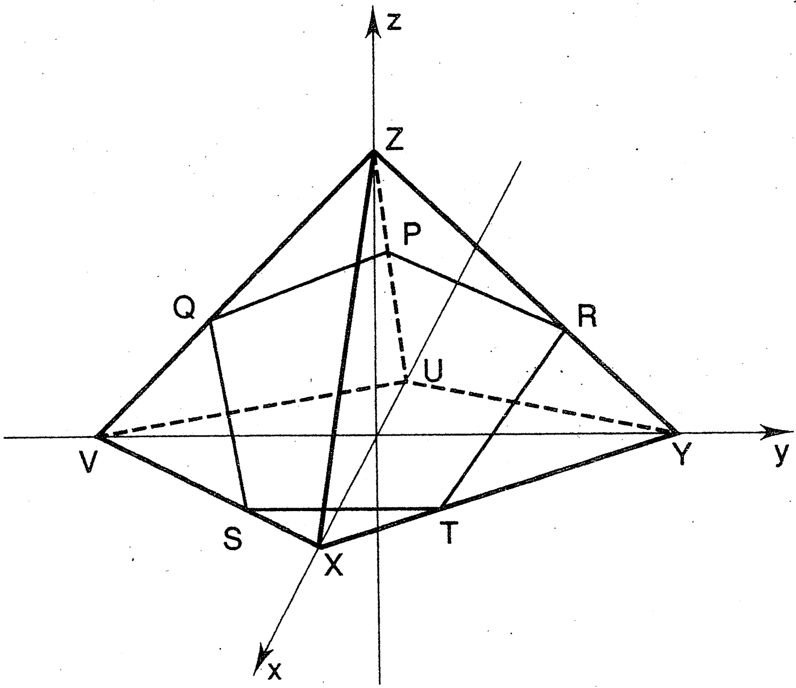


Figure 1

Something like this, perhaps precisely this, must have been done with the computer proof. Professor Wu told me that this was carried out by a graduate student at Beijing University and published in his M.Sc. thesis – in Chinese, and in any case not available to me. (Professor Wu did say that a discussion of the problem appears in a text either recently or soon to be published in the US. He did not, however, have the details. Nor could he say whether the student who programmed the computer and wrote the thesis was the same person as the mathematician who had first posed the problem, but this is a very minor detail.)

The next point that occurred to me to ask was whether the plane $PQSTR$ had been supposed to cut the pyramid symmetrically or asymmetrically. That is to say, was it assumed that $XS = XT$ or not? If it is not assumed that $XS = XT$, then the problem leads to a much more complicated and difficult set of equations involving 6 unknowns instead of 4. By taking this full set, I was able (after much work) to show algebraically that if the pentagon $PQSTR$ was to be regular then we did indeed require that $XS = XT$.

When I next spoke with Professor Wu, I asked him if the computer program had assumed symmetry or not. His reply was that it *had* assumed symmetry, because this was an obvious requirement if the pentagon was to be regular. At first I wondered if it really was so "obvious", but it was in fact I and not he who was overlooking a point. In my algebraic proof that $XS = XT$, at one point I produced a combination of equations that had the geometric interpretation $PS = PT$. Clearly we must have this if the pentagon is to be regular. Almost equally clearly, the requirement $PS = PT$ implies that $XS = XT$. I leave the details to the reader.

We are thus ready to put our problem into algebraic form. We need to give algebraic expression to the points P, Q, S, T, R and this is readily done. Let:

$-s$ be the x-coordinate of P
$-t$ be the y-coordinate of Q
v be the x-coordinate of S.

Then the requirement that $PQSTR$ be a regular pentagon may be expressed by means of four equations:

$$s^2 + t^2 + a^2(s-t)^2 = 4(v-1)^2 \tag{1}$$
$$v^2 + (1-v-t)^2 + a^2(t-1)^2 = 4(v-1)^2 \tag{2}$$
$$sv + t(1-v-t) + a^2(t-1)(s-t) = 2(1-v-t)(v-1) \tag{3}$$
$$-s^2 + t^2 - a^2(s-t)^2 = 2(1-v-t)(v-1). \tag{4}$$

Equation (1) expresses the condition that $|PQ| = |ST|$; Equation (2) expresses the condition that $|QS| = |ST|$; Equation (3) expresses the condition that $\angle PQS = \angle TSQ$; Equation (4) expresses the condition that $\angle TSQ = \angle QPR$. These four equations are necessary and sufficient to ensure that the pentagon $PQSTR$ is regular. For complete information on the problem, all we have to do is to solve them (or else to show that they have no solution). There are four equations and we may say there are four unknowns s, t, v, a; alternatively, we may say that we have four equations in three unknowns s, t, v but involving a parameter a, whose value may need to be determined.

I played around with Equations (1)-(4) for a good many hours and, though I made some progress, a solution eluded me. Eventually I realised that I could include another geometric condition. The points P, Q, S necessarily lie in the plane of the pentagon, and because we assumed symmetry ($XS = XT$), that plane is parallel to the y-axis. This gives a fifth equation, mercifully simpler than the others:

$$s(1-t+v) = tv. \tag{5}$$

Adjoining Equation (5) to Equations (1)-(4) allows a solution to be found.[†1] Nor is it particularly difficult. Once you find it and throw out all the false scents and blind alleys, it takes a little less than a page to write out. I'll give the result later.

Clearly something like this was done by the Beijing computer. Some set of equations, equivalent in strict logic to the set (1)-(4), must have been solved, but some such equivalent sets are more readily solved than others and we're not sure quite which one was solved, nor whether the computer had human help such as the provision of Equation (5). Professor Wu did tell me that the initial setting up of the equations depended on a clever choice of coordinates, so possibly something much more subtle was involved.

After leaving Hong Kong, I travelled to Kyoto (Japan) for another large Mathematics conference – the International Congress of Mathematicians. Here a display of the software package *Mathematica* was in progress and as part of that display they undertook at my request to see if *Mathematica* could crack the problem. It couldn't. It spent a day to no purpose on Equations (1)-(4) and another day, again to no purpose, on Equations (1)-(5). By contrast, I had, in $2\frac{1}{2}$ *working* days, derived the equations, wasted time on the asymmetrical case, decided on the subset (1)-(4) of all those that I might have used instead, adjoined Equation (5) when the going got rough, and finally solved the equations.

[1†]My colleague, Dr C. McIntosh, has found algebraic solutions of Equations (1)-(4) that do not satisfy Equation (5). They are not, however, geometrically realistic as they involve $a = 0$.

Quite how long the Beijing computer took I don't know, nor whether it got stuck and needed help along the way. I would like, in view of the fact that I succeeded where *Mathematica* failed, to say that this shows the superiority of human over computer. This would, however, be less than honest, for what did help me enormously was that I knew, in broad outline at least, the answer in advance. Professor Wu had given it in his lecture. If I had not known this, I'd probably have given up the search. So the answer was supplied by the computer and before it gave that answer neither Professor Wu nor any others in the Beijing team knew it. Nor would I nor anybody else at the conference have known it, had not Professor Wu told it to us.

It is now time to look at the answer in some detail. The solution to Equations (1)-(4), equivalently Equations (1)-(5), is:

$$a = 1,\ s = 1 + \phi,\ t = v = -\phi, \tag{6}$$

where

$$\phi = \tfrac{1}{2}(1 \pm \sqrt{5}). \tag{7}$$

First consider the implications of the value $a = 1$. The regular pentagon can exist only in that case where the height of the pyramid is the same as the distance from O to each of the base vertices. In other words, the original square pyramid has to be half of a solid known as a regular octahedron. This result was produced by the Beijing computer.

Secondly, note Equation (7). There are two values of ϕ (essentially the golden ratio). It is actually the value $\frac{1}{2}(1-\sqrt{5})$ that produces the solution we expect and which is diagrammed in Figure 1. But what of the value $\frac{1}{2}(1+\sqrt{5})$?

This was the real surprise for the Beijing researchers. Look at Figure 2.

The points P, Q, S, T, R lie on the lines ZU, VZ, YZ, XV, XY respectively, just as in Figure 1, but now each of these lines is extended beyond the confines of the original pyramid. We still have $PQ = QS = ST = TR = RP$ and we still have $\angle PQS = \angle QST = \angle STR = \angle TRP = \angle RPQ$, exactly as in that case, but now our "regular pentagon" is a star, or, more precisely, a regular penta*gram*. Again the computer told the researchers something that previously they had not known.

These results then are the latest candidate for a computer-generated theorem in elementary geometry. True, the computer had a lot of human help, but the humans who helped it were surprised by the results. There is also one other question: Are these results new? Neither I nor anyone else at the conference knew of them, nor did the Beijing team, nor has anyone I've spoken to since; but it could be that (say) somewhere in a 19th-century text on Solid Geometry they do appear – perhaps even among the exercises. [See *Function*, *Vol.* 14, *Part* 5, p.151 for a not dissimilar case.] Then we'd have to say the result was re-discovered by a computer.

That is, of course, if you admit this story as a valid case of computer discovery. What do you think? Keep the letters flowing in.

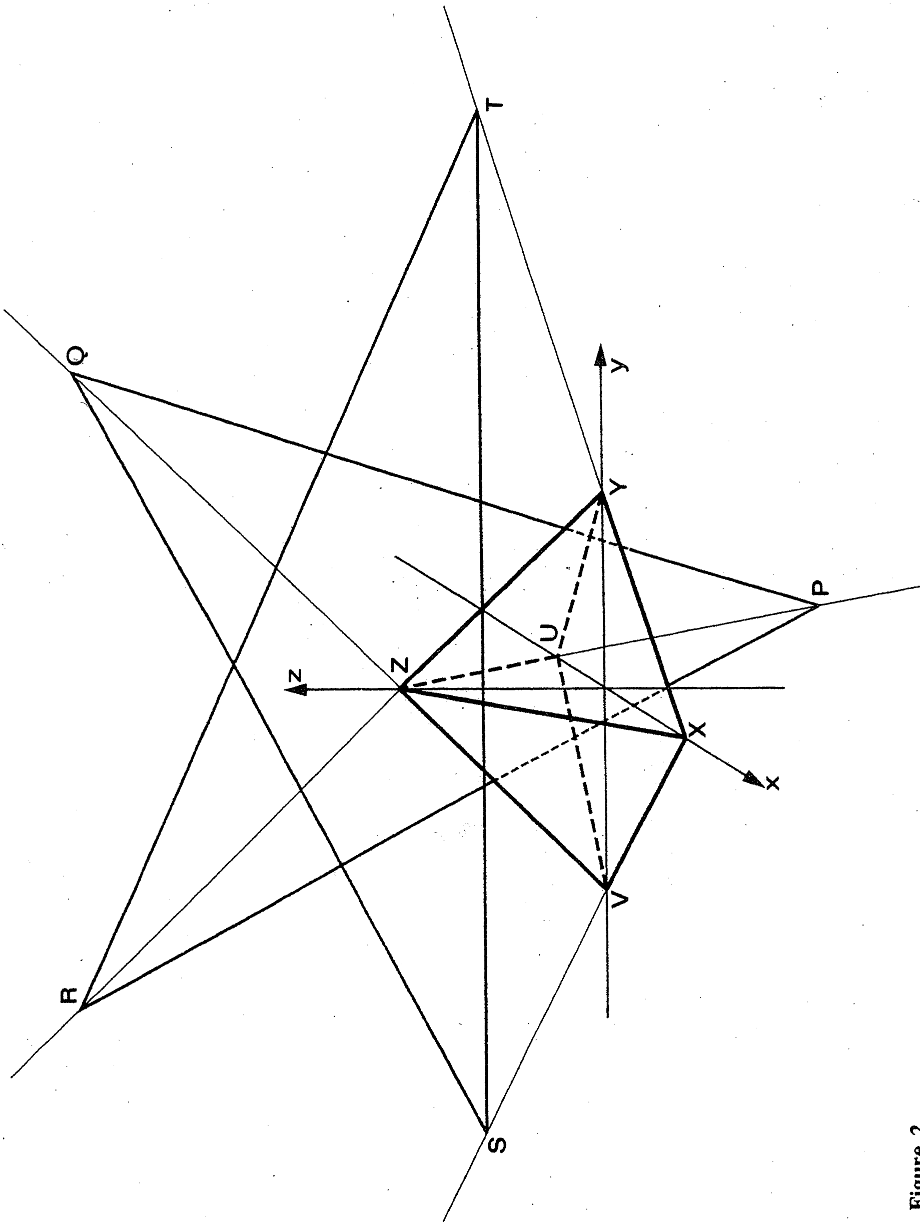


Figure 2

A TRIGONOMETRIC IDENTITY

K.R.S. Sastry, Addis Ababa, Ethiopia

Normally the conditional trigonometric identity

$$\sin 2A + \sin 2B + \sin 2C = 4\sin A \sin B \sin C, \text{ where } A + B + C = \pi \qquad (*)$$

is established by changing the sums on the left-hand side into products or *vice versa* if one chooses the right-hand side to begin with. Here is a geometrical look at (*) which incidentally shows that each side of it represents twice the area of triangle ABC inscribed in a *unit circle*. We use the following well-known or easily established results:

The size of an angle at the centre of a circle is twice the size of an angle at the circumference provided both angles are on the same arc. (1)

If a chord of length l subtends an angle θ at the centre of a unit circle then $l = 2\sin(\theta/2)$. (2)

Area $(\Delta ABC) = \frac{1}{2}(BC)(CA)\sin\angle ACB$. (3)

We shall consider separately the cases of triangle ABC being acute, right or obtuse angled.

Consider an acute angled triangle ABC inscribed in a unit circle, centre O, shown in Figure 1.

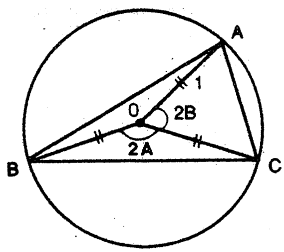
Then

from (1), $\angle BOC = 2A$, $\angle COA = 2B$, $\angle AOB = 2C$,

from (2), $BC = 2\sin A$, $CA = 2\sin B$.

Now

area(ΔBOC) + area$(\Delta COAB)$ + area(ΔAOB) = area(ΔABC).

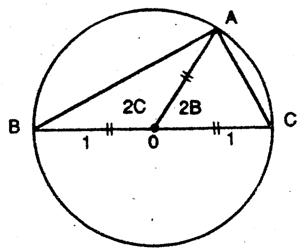


Using (3) in the preceding equation we obtain

$$\frac{1}{2}\sin 2A + \frac{1}{2}\sin 2B + \frac{1}{2}\sin 2C = \frac{1}{2}(2\sin A)(2\sin B)\sin C$$

which yields (*), showing at the same time that each side of (*) is simply twice the area of triangle ABC.

If angle A of triangle ABC happens to be a right angle, as Figure 2 shows, we have

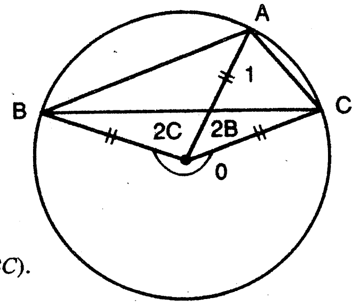


$$0 + \frac{1}{2}\sin 2B + \frac{1}{2}\sin 2C = \frac{1}{2}(2)(2\sin B)\sin C \qquad (4)$$

or

$$\sin 2B + 2\sin C = 4\sin B\sin C.$$

Since $\sin 2A = \sin\pi = 0$ and $\sin A = \sin(\frac{\pi}{2}) = 1$ the above equation is equivalent to (*).



If angle BAC of triangle ABC happens to be obtuse then Figure 3 shows that

$$BC = 2\sin\left[\frac{2\pi-2A}{2}\right] = 2\sin A, \text{ on using (2)}$$

and

$$-\text{area}(\Delta BOC) + \text{area}(\Delta COA) + \text{area}(\Delta AOB) = \text{area}(\Delta ABC).$$

Using (3), the above equation yields

$$-\frac{1}{2}\sin(2\pi-2A) + \frac{1}{2}\sin 2B + \frac{1}{2}\sin 2C = \frac{1}{2}(2\sin A)(2\sin B)\sin C$$

which again leads to (*) because $\sin(2\pi-2A) = -\sin 2A$.

We now observe that, using (*) as a starting point, a number of interesting identities can be derived. The transformation

(i) $A \to \pi-2A, B \to \pi-2B, C \to \pi-2C$ (which preserves $A + B + C \to 3\pi - 2(A+B+C) = \pi$)
   yields $\sin 4A + \sin 4B + \sin 4C = -4\sin 2A \sin 2B \sin 2C$.

(ii) $A \to \frac{\pi}{2} - \frac{A}{2}, B \to \frac{\pi}{2} - \frac{B}{2}, C \to \frac{\pi}{2} - \frac{C}{2}$ yields
   $\sin A + \sin B + \sin C = 4\cos\frac{A}{2} \cdot \cos\frac{B}{2} \cdot \cos\frac{C}{2}$.

(iii) $A \to \frac{\pi}{4} - \frac{A}{4}, B \to \frac{\pi}{4} - \frac{B}{4}, C \to \frac{\pi}{4} - \frac{C}{4}$ yields

$$\cos\frac{A}{2} + \cos\frac{B}{2} + \cos\frac{C}{2} = 4\sin\left[\frac{\pi}{4} + \frac{A}{4}\right]\sin\left[\frac{\pi}{4} + \frac{B}{4}\right]\sin\left[\frac{\pi}{4} + \frac{C}{4}\right].$$

* * * * *

COMPOSITE FUNCTION

* * * * *

LETTERS TO THE EDITOR

Generating Functions

I offer some comments on Marta Sved's article, "Generating Functions", which appeared on page 49 of the April 1990 issue of *Function*.

The problem on which the article is based is restated in general terms as follows:

There are n couples, that is, $2n$ members in an association. A committee of r members is to be formed; membership being awarded to couples. A single membership entitles one representative of the couple to attend a meeting, while a double membership means that both husband and wife are to attend. How many such committees are possible?

I believe there is a much simpler solution:

Consider any one of the n couples. There are three possibilities:

(a) both parents will attend (double membership);
(b) only 1 of the 2 parents will attend (single membership);
(c) neither will attend (the couple is not elected).

Therefore the numbers of different committees must be the coefficient of x^r in the expansion of $(x^0 + x^1 + x^2)^n$ because this coefficient arises out of all the different ways in which n of the indices 0, 1, 2 may be taken in order to form r by addition. We therefore require the coefficient of x^r in the expansion of $(1 + x + x^2)^n$ or, in the particular example quoted, the coefficient of x^6 in the expansion of $(1 + x + x^2)^{200}$.

$$(1 + x + x^2)^{200} = \left[\frac{1-x^3}{1-x}\right]^{200}$$

$$= (1-x^3)^{200}(1-x)^{-200}$$

$$= (1 - 200x^3 + \frac{200\cdot 199}{2!}x^6 - \ldots)(1 + 200x + \frac{200\cdot 201}{2!}x^2 + \ldots)$$

$$= \left\{1 - \binom{200}{1}x^3 + \binom{200}{2}x^6 - \ldots\right\}\left\{1 + \ldots + \binom{202}{3}x^3 + \ldots + \binom{205}{6}x^6 + \ldots\right\}.$$

The coefficient of $x^6 = \binom{205}{6} + \binom{200}{2} - \binom{200}{1}\binom{202}{3}$

$$= 95\ 476\ 299\ 600,$$

which is not the answer recorded by Marta Sved because I believe it has been miscalculated.

David Shaw,
217 Noble St.,
Newtown, Vic.

[*Mr. Shaw is right. His number is correct and that published by* Function *was wrong. His method is also very elegant. In fairness to Marta Sved, however, it should be noted that the method she used was not the one we printed, nor Mr. Shaw's, but a different one again. She did get the right answer. Because her technique was rather advanced for* Function*'s readers, this section of her paper was rewritten by another member of the editorial board, whom (to save embarrassment) we won't identify. In the course of this rewrite, the answer was "corrected". We thank Mr. Shaw for reinstating the right figure and for supplying a most elegant method for finding it.*

Eds.]

Areas of Supercircles

I was very surprised that Michael Deakin ("Running round in Squares and getting Cross", *Function, Vol.* 14, *Part* 4, pp. 101-106) nowhere considered the areas enclosed by the various curves he discussed. The value of the area is, in fact, a much easier determination than the length of the perimeter, which he *did* raise.

Let $A(n)$ be the area enclosed by the curve

$$|x|^n + |y|^n = 1.$$

It is possible, though difficult, to give a formula for $A(n)$. The following cases, however, are well known, and readers will be familiar with all except perhaps the second.

1. $n = 0$ (the cross): $A(0) = 0$
2. $n = 2/3$ (the astroid): $A(2/3) = 3\pi/8$
3. $n = 1$ (the tilted square): $A(1) = 2$
4. $n = 2$ (the circle): $A(2) = \pi$
5. $n = \infty$ (the upright square): $A(\infty) = 4$.

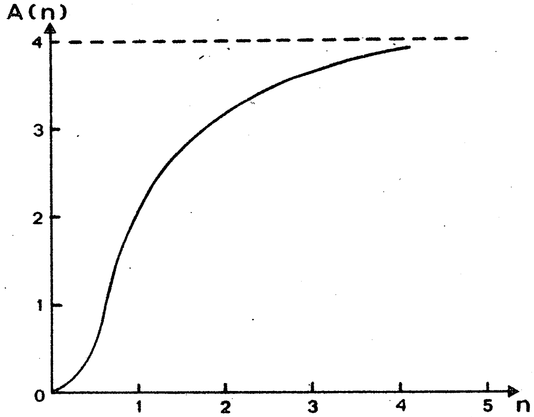
If these are graphed, the diagram opposite is the result.

Kim Dean,
Union College, Windsor

ICM-90

In August 1990 the world's mathematicians met for 10 days at the Conference Centre in Kyoto, Japan. Such gatherings are held every four years under the auspices of the International Mathematical Union – the body through which all national bodies (such as the Australian Mathematical Society) keep in touch with one another.

Out of a total attendance of almost 4000, some dozen or so Australian mathematicians attended, including your intrepid editor, M.A.B. Deakin. The total attendance was somewhat down on that for recent years, as Japan is perceived as an expensive country to visit. 55% of the delegates came, in fact, from Japan itself, 10% from the US and 4% from the USSR. The only other delegations of any size came from France, the UK and (East and West) Germany at about 2.5% each.



* * * * *

Sadly, there was no Pakistani delegation, as the IMU has finally expelled Pakistan for persistent non-payment of annual dues. Many of the third-world delegates who did manage to attend, however, deplored this move on the part of the international body.

The Japanese government minted a special set of stamps for the occasion, featuring origami models of the regular polyhedra – indeed, classes were conducted (in conjunction with the conference) on origami and how to fold polyhedra and much else besides.

Four Fields medals (the mathematical equivalent of Nobel prizes – see *Function*, *Vol.* 11, *Part* 1) were awarded. They went to:

(a) Vladimir Drinfeld of the USSR for work on generalised number theory and other work on quantum groups (these latter relate to the currently much studied "Yang-Baxter" operators, named in part after Prof. Baxter of the Australian National University, an invited speaker at the conference);

(b) Vaughan Jones, a New Zealander now resident in the US, for work on knots (see *Function*, *Vol.* 12, *Part* 5) including the Jones index theorem and the Jones polynomial, with applications in many areas including quantum groups;

(c) Shigefumi Mori of Japan for work on algebraic geometry in higher (i.e. three or four) dimensions; work in this field has been dominated by the so-called "Mori programme" for over a decade; some "feel" for what is involved may be seen from the special case of the 2-D problem on the cover of *Function*, *Vol.* 9, *Part* 2;

(d) Edward Witten of the US, a mathematical physicist who has given a sound mathematical basis to earlier work by the Nobel laureate Richard Feynman and who seems to have solved the long-standing problem of representing energy in Einstein's general theory of relativity.

Readers may note the heavy concentration on mathematical physics [(a), (b) and (d) above]. Witten has been widely tipped as a future Nobel prize-winner and thus could become the first person ever to receive both the Nobel prize and the Fields medal.

As well as the Fields medals, the *Nevanlinna Prize* (named after the Finnish mathematician Rolf Nevanlinna and awarded for work in Computer Science) was awarded to Alexander Razborov of the USSR for work in the theory of algorithms (mechanical solution procedures). In particular, Razborov has shown that some of the outstanding problems of graph theory (see, e.g., *Function, Vol.* 13, *Part* 1, p.20) are "inherently difficult" (see *Function, Vol.* 4, *Part* 3).

The opening plenary address was delivered by Karen Uhlenbeck of the USA. Her work explores the interaction of analysis (higher calculus) and topology. She is only the second woman ever to have given a plenary address at such a conference (the other was the algebraist Emmy Noether in 1932).

Vaughan Jones was another plenary speaker and Edward Witten exercised his right (as a Fields medallist) to deliver a plenary address. Another plenary speaker was Richard Melrose, an expatriate Tasmanian now at M.I.T. (USA).

Much else went on besides. One speaker spent quite some time on the $(3n+1)$-conjecture (see *Function, Vol.* 7, *Part* 3, p.26). In Japan, it is known as the *Kakutami Conjecture*. It is still unsolved.

W.M. Kahane, a Canadian computer scientist now working at Berkeley, California, gave a very interesting and lively talk on deficiencies in computer design and software. Here is one of his examples.

Define

$$f_n(x) = \left[x^{2^{-n}} \right]^{2^n}$$

as worked out by your computer or calculator. That is, for some number x, take the square root n times and then square the answer n times. The result should be x. It often isn't. On my CASIO FX-570, $f_{29}(2) = 1.805$, $f_{30}(2) = 1$. Try this test on your favourite machine. Plot the results of $f_n(x)$ *versus* n for different x. The results can be very interesting – they deserve your study and consideration.

Professor Kahane claims that the CRAY supercomputers do particularly badly on this test.

There then are some of the highlights of the Kyoto conference: interesting and successful and held in an interesting city with many splendid and beautiful shrines, palaces and temples to visit and experience.

* * * * *

Utility Theory

"A ton of ore contains an almost infinitesimal amount of gold, yet its extraction proves worthwhile. So if only a microscopic part of pure mathematics proves useful, its production would be justified."

L.J. Mordell

HISTORY OF MATHEMATICS SECTION

EDITOR: M.A.B. DEAKIN

The Pre-Pre-History of Mathematics

In earlier columns, we have looked at the mathematics of the past: sometimes the recent past, sometimes the rather more remote past. The earliest figure mentioned was Zeno of Elea, who lived about 2500 years ago. Zeno's own words have not come down to us but we do have accounts by Aristotle and Plato of what Zeno is supposed to have said. This then is history, but only just; history is what is vouched for in writing. The history of mathematics goes back a little beyond the Greeks of Zeno's day. We have some records of the mathematics of the Babylonians (about 2000 B.C.) and the Egyptians (at least as old and probably older).

To find out what happened before this we rely on the techniques of pre-history. This issue's column will be about the very dawn of mathematics and what we can know about it. The earliest mathematics we can imagine is the mathematics of counting, and all counting systems use a base, a special number that is used to express other numbers. Thus we say (e.g.) "forty-three" and mean "four times ten plus three". Similarly we write 43 to the same end. Our base is ten.

In fact all cultures of advanced numeracy today use a base of ten; only two such cultures have ever used bases other than ten. The Babylonians are usually said to have had a base of sixty, a loose and not entirely accurate account of the true situation. The Mayans, pre-Columbian inhabitants of Central America, used a base of twenty (see *Function, Vol.* 12, *Part* 4).

As the example of forty-three shows, the structure of our own language, English, reflects our adoption of a base ten system. English is one of a large family of languages, called the Indo-European family. All such languages descend ultimately from a single language called *Proto-Indo-European*, or PIE for short.

Linguists have been able to reconstruct PIE to a very impressive extent. Written records in Greek, Latin and Sanscrit take us back a long way and scholars then deduce what happened before that.

Take the example of the word "eight". The Greek (Ancient Greek) for "eight" is *oktō*; the Latin is *octō*, the Sanscrit *astau*. From these and other pieces of evidence, linguists have reconstructed the PIE **oktou* (the asterisk is a device to show that the word is reconstructed and not directly attested). The *k* sound has turned into an *s* in the Sanscrit. This is in line with a general principle that consonants tend to move forward in the mouth as time goes by. Thus *k*, made at the back of the mouth, down in the throat, became an *s*, made at the front of the mouth, with the tip of the tongue.

Another line of evidence comes from languages still spoken today but which preserve many archaic features. The best example is Lithuanian. The Lithuanian word for eight is *aštuoni*, and again we see the $k \to s$ shift that was observed in Sanscrit. The *k* is indeed preserved, but in an odd way, in English. True, we don't pronounce it, but we write eight with a *-gh-*. This corresponds to a *-ch-* in the German *acht*. We have dropped the guttural sound, as have the Italians, who have *otto*, from the Latin *octō*. The difference is that our spelling continues to reflect the older pronunciation.

Other Indo-European languages show variations on these themes. Thus we have in Old Irish *ocht*, Gothic *ahtau*, Old English *eahta*, Old Slav *osmi*, Armenian *ut'*, *Hindi āth*, Persian *hasht*, Tocharian *okät* and so on. All these words derive from **oktou*.

This is just one example of how linguists can enter into a pre-historical world and deduce how things must have been before written records existed. By such means, very much of PIE has been reconstructed. Apart from a few details we know the numerals in PIE, and how its speakers counted. They very clearly used base ten.

There is some doubt as to who spoke PIE and when and where it was spoken. Probably about five to six thousand years ago and probably in Western Asia or the Middle East. This is about as much as we can say.

But now consider that PIE itself must have come from somewhere. And that the base ten system must have evolved from something less developed.

If we look at cultures today in which number notions are less important than in our own, we find a vast array of ways in which numbers are described. All languages have a word for "two", but by no means all have made what one authority (Menninger) calls "the step to three". There are languages (some Australian Aboriginal languages have this character) where the numeral concepts are *one*, *two* and *many*.

Where this happens, there are three forms of the word representing the objects being counted: singular, dual and plural. In English, by contrast, we have singular and plural (but no dual). We say, for instance, "one horse", "many horses", but we also say "two horses". It's horse*s* as long as there are more than one of them.

Interestingly enough PIE had a singular-, dual-, plural system and this was preserved in (e.g.) Sanscrit. The Sanscrit word for "horse" is *áśvas*, but for "horses" the word is *áśvau* if there are two and *áśvās* if there are three or more. Similarly in Ancient Greek where the words are (respectively) *híppos*, *híppō*, *híppoi*. The dual survives up to a point in modern Lithuanian, but in large measure we have lost it from today's Indo-European languages.

But not entirely; about a year ago I set out to find vestiges of the dual in modern English usage. I ended up finding five such.

The first and most obvious one comprises the words *both* and *either*. These can only be used in the dual. For more than two items, they must be replaced by the words *all* and *each* which are not normally used if the situation is a dual one. Similarly with "the other" and "another".

A like situation exists with the prefixes *ambi-* and *amphi-*, as in "ambidextrous" and "amphibious". These prefixes have a clear dual implication, although in one case, it has partially extended to cover plurality as well: if a sentence were so unclear as to be capable of three or more interpretations, we would still speak of it as *ambiguous*. (This is to say there can be several *alternative* meanings. *This* word also was, until very recently, exclusively dual.) Thus that pressure which for thousands of years has been acting to exclude the dual is still operating today.

A somewhat more subtle case of a dual-plural distinction occurs with the difference between the comparative and the superlative forms of the adjective. The comparative form ("more", "larger", etc.) is used where there are two objects under discussion; the superlative ("most", "largest", etc.) when there are more than two.

Fourthly, there are special words signifying pairs. These differ from specialist words relating to other numbers ("quartet", for example) in that the word *pair*, the word *yoke* and the word *brace*, etc. do not derive from the same root as the word *two*. "Quartet", by contrast, does derive from the PIE word **kwetwores*, which is also the remote ancestor of our word *four*. (See *Function*, *Vol.* 8, *Part* 2, p.25.)

Finally, even when a word does derive from the same root as the word *two*, e.g. *twin*, our language allows it a wider use than it would (say) *triplet*. So we have *twin-tubs* and *twin-sets*, where we would say "3-stage", "3-piece", etc. almost universally, were the number *three* instead of *two*.

These then are the vestiges of the singular-dual-plural system that still stay with us today. Linguistic habits take a long time to die. Back in about 4000 B.C. when PIE was spoken, those linguistic habits were stronger and the dual was a more powerful influence. Where then did it come from?

Well, of course, we can't know, but we can put up the most plausible available hypotheses and see if these make sense. We are, if you like, asking questions about the pre-pre-history of our number-words.

We have seen that all languages of today allow a "one-two-many" distinction and some have no numerical concepts beyond this. This structure is mirrored in the singular-dual-plural system which was well-preserved in PIE. It is thus plausible to suppose that the grammatical structure of PIE was a relict of an earlier counting system: a "one-two-many" system pre-dating PIE's fully developed use of base ten.

Very recently the work of some influential Soviet linguists has started to become known in the west. Let us see what these linguists suggest.

The Indo-European languages fall into several sub-groupings, such as the Romance languages, the Slavic languages, the Germanic, the Celtic, and so on. These different sub-groupings together make up the Indo-European family. Other families exist: the Uralic (of which Finnish and Hungarian are examples), the Afro-Asian (Hebrew, Arabic, etc.), the Dravidian (Tamil and its relatives) and so on.

The Soviet linguists propose that some *families* are related to form "super-families", just as the sub-groupings together form families. In particular, they proposed that all the families mentioned above, together with some others (but also excluding yet others like Chinese) all derive from a very remote common ancestor called *Nostratic*.

This is regarded by many as rather speculative – but what little has been published in the West does lend some support to the historic base two notion in that proponents of Nostratic theory regard the expression of the numeral "two" as one of the strongest elements in their theory. Other numerals (with the possible exceptions of three and four, of which I will say more in a later column) do not seem to have been part of Nostratic (and the evidence for three and four is much more tenuous).

For further reading in this difficult but fascinating area, see *Scientific American*, October 1989 and March 1990.

To close, I quote an English translation of a poem written in Nostratic by the Soviet linguist Illič-Svitič. (In the original it rhymes and has a regular rhythm!)

"Language is a ford through the river of time,
It leads us to the dwelling of ancestors.
But he doesn't arrive there
Who fears deep water."

COMPUTER SECTION

R.T. WORLEY

Late last year there was a number of messages on the electronic news about computer generation of mazes. There was even a maze-generating program for which the program itself when printed looked like a maze. It turns out that there are a variety of ways to generate a maze using a computer. One way, though not the simplest, relates to what mathematicians call *graph theory*, and may be of special interest to VCE students in Victoria. We will look at this method, and relate it to graph theory. However it requires some sophisticated programming methods to run efficiently enough to run in interpreted BASIC on a PC, so we will look briefly at another method for which we give a program and some example mazes.

For simplicity, we consider a square maze, as illustrated in Figure 1b, which we regard as having been formed from the 4×4 grid of cells illustrated in Figure 1a by removing some interior walls. For convenience the cells of the grid will be labelled

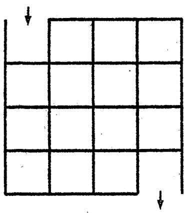


Figure 1a

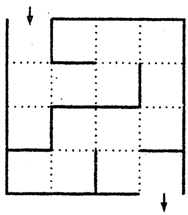


Figure 1b

Cell(1,1)...Cell(1,4) left to right across the top row, and so on, with the bottom right cell being Cell(4,4). The vertical interior walls are regarded as belonging to the cell on their left, and the horizontal interior walls belong to the cell above them. To change the grid of Figure 1a into the maze of Figure 1b we must remove (in arbitrary order) the bottom wall of Cell(1,1), the bottom wall of Cell(2,1), the right wall of Cell(2,1), the right wall of Cell(2,2), and so on. If we are to write a computer program to generate a maze we must have some means of deciding which walls are to be removed so that the resulting diagram is a maze. Indeed, we should actually specify what features characterise a maze. For example, will we call any of the diagrams in Figure 2 a maze?

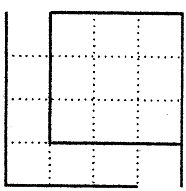


Figure 2a

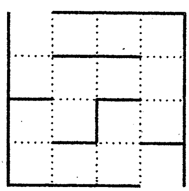


Figure 2b

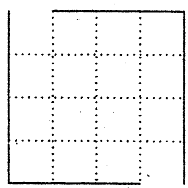


Figure 2c

Observe that in Figure 2a, there are some cells (those in the upper right region) which cannot be reached from the starting point, while in Figure 2b there are some "islands" of interior wall, and consequently there is more that one way through the maze. Figure 2c is just the most extreme case in which all interior walls have been removed. For convenience we will not allow islands or unreachable cells – it seems to me that these in fact make it easier to find a path from the start to the end.

Now that we have decided what sort of diagram we call a maze, we can try to see how to construct one. We look at Figure 1 again. Inititally, in Figure 1a, there is only one cell, Cell(1,1), that is reachable from the starting point. If we remove the bottom wall of Cell(1,1) then we make Cell(2,1) reachable. Likewise, if we now remove the bottom wall of Cell(2,1) then we make Cell(3,1) reachable. Removing the right wall of Cell(2,1) now makes Cell(2,2) reachable. It is clear that if at each step we remove a wall that separates a reachable cell from an unreachable cell then at each step we increase the number of reachable cells. Clearly, if we remove 15 such walls of the 4×4 grid, then we increase the number of reachable cells from 1 to 16. Since there are only 16 cells in the grid, all cells are reachable, so we have a maze. Clearly a computer program can generate a n×n maze as follows

° Initially Cell(1,1) is reachable and all other cells are unreachable.

° Repeat n^2-1 times: remove a wall separating a reachable cell from an unreachable cell (the unreachable cell then becomes reachable).

° Draw the maze.

The difficulty with this algorithm lies in the selection of the wall to be removed. This is because we want to produce random looking mazes, and so need some sort of randomness in the wall selection. One way of doing this is to allocate initially a random number to each wall. Then the wall selection procedure just has to select the wall with the lowest number (subject of course to the constraint that the wall separates a reachable cell from an unreachable cell). The only problem with this approach is that the walls and their random number must be stored in the computer in a way that makes this selection procedure efficient. If the wall selection procedure has to examine every wall to find the one with the lowest number, then the wall selection will require examining each of the $2n^2$-2n walls. Since we have to do this n^2-1 times, the total number of wall examinations will be $(2n^2-2n)(n^2-1)$, which is approximately $2n^4$. For a 8×8 maze this amounts to only 2×8^4 = 8192 wall examinations. Suppose these 8192 wall examinations take 1 minute (when programmed in interpreted BASIC on a PC). This is not an unreasonable amount of time. However, for an 16×16 maze there will be 2×16^4 = 131072 wall examinations - 16 times as many as for a 8×8 maze – which will take say 16 minutes. On the same basis a 32×32 maze will take over 4 hours. While it is possible to make the wall selection procedure more efficient by storing the walls and their random numbers in a suitable way, the resulting program will look quite daunting for a student who has not studied computer data structures.

To have a reasonably simple program, I chose a different method of wall selection. When a cell becomes reachable, it is put in a list. To select an wall, I take the most recent cell put in the list. If it has no walls separating it from an unreachable cell, it is thrown away and the next cell taken from the list. Otherwise one of the walls separating the cell from an unreachable cell is chosen at random. The orignal cell is put back on the list (if it has more suitable walls), and the newly reachable cell is also put on the list. The program is slightly complicated by the fact that the probabilities used in selecting the wall are chosen to give a preference to choosing to remove a wall opposite an already removed wall – this is to try to produce a maze that has recognisable passages instead of being quite jagged.

The following are mazes produced by both methods described above. Method A is the method which ascribes random numbers to the walls (modified to try to produces mazes with recognisable passages), and Method B is the list method (using the accompanying program).

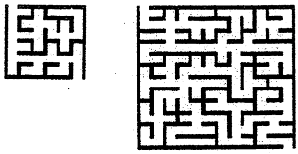


Figure 3a. Mazes produced by Method A

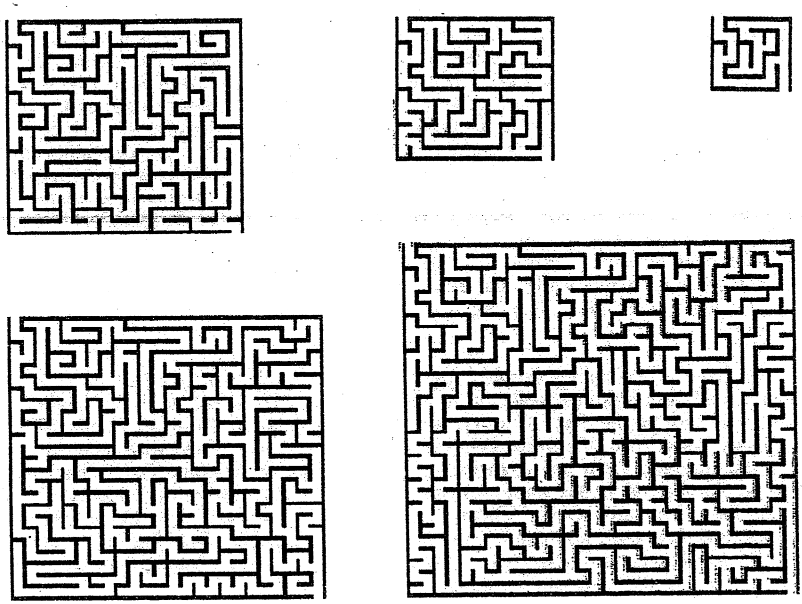


Figure 3b. Mazes produced by Method B

```
100 REM Maze Generating Program
110 REM R.T. Worley 01 Feb 1991
999 DEFINT A-Z
1000 INPUT "Enter size of maze (2 to 30)"; MAZDIM
1010 IF (MAZDIM<2) OR (MAZDIM>30) THEN PRINT "Invalid size entered" : GOTO 1000
1020 DIM CELL(MAZDIM,MAZDIM), RWALL(MAZDIM,MAZDIM-1), BWALL(MAZDIM-1,
     MAZDIM)
1021 DIM STACK(MAZDIM*MAZDIM)
1022 DIM LOWER(14),UPPER(14),FIRST(14),SECOND(14),THIRD(14)
1023 DATA 0,0,1,2,0,    0,0,0,0,0,    2,8,1,4,0,    1,8,2,4,0
1024 DATA 0,0,4,1,2,    0,0,0,0,0,    2,4,1,8,0,    1,4,2,8,0
1025 DATA 0,0,8,1,2,    0,0,4,8,0,    0,0,1,4,8,    0,0,2,4,8
1030 GOSUB 2000 : REM To initialise variables
1040 GOSUB 3000 : REM Wall selection
1070 GOSUB 4000 : REM Draw maze
1075 ZZ$=INKEY$: IF ZZ$="" THEN GOTO 1075
1080 STOP
1990 REM ----- initialisation subroutine ------
2000 UNREACHABLE=0: REACHABLE=1 : REM values the cell type flag cell can take
2005 SELECTED=0: UNSELECTED=1: REM flag values for selected/unselected walls
2010 FOR I=1 TO MAZDIM: FOR J=1 TO MAZDIM: CELL(I,J) = UNREACHABLE
2020 NEXT J: NEXT I
2029 REM set cell at (1,1) as starting cell
2030 CELL(1,1)=REACHABLE: TOP=1: STACK(TOP)=0
2040 FOR I=3 TO 14: READ LOWER(I),UPPER(I),FIRST(I),SECOND(I),THIRD(I):NEXT I
2050 FOR I=1 TO MAZDIM: FOR J=1 TO MAZDIM-1: RWALL(I,J)=UNSELECTED:NEXT J,I
2060 FOR I=1 TO MAZDIM-1: FOR J=1 TO MAZDIM: BWALL(I,J)=UNSELECTED:NEXT J,I
2070 RANDOMIZE
2150 RETURN
2990 REM ------ wall selection subroutine -------
3000 IF TOP=0 THEN GOTO 3990
3010 CELLNO=STACK(TOP)
3020 ROW = 1 + (CELLNO\MAZDIM): REM 1 + int(cellno/mazdim)
3030 COL = 1+CELLNO MOD MAZDIM: REM cellno-(row-1)*mazdim
3040 TOERASE=0
3050 REM determine which walls of cellno can be erased
3060 IF ROW>1 THEN IF CELL(ROW-1,COL)=UNREACHABLE THEN TOERASE=TOERASE+
3070 IF ROW<MAZDIM THEN IF CELL(ROW+1,COL)=UNREACHABLE THEN
     TOERASE=TOERASE+2
3080 IF COL>1 THEN IF CELL(ROW,COL-1)=UNREACHABLE THEN TOERASE=TOERASE+4
3090 IF COL<MAZDIM THEN IF CELL(ROW,COL+1)=UNREACHABLE THEN
     TOERASE=TOERASE+8
3100 IF TOERASE=0 THEN TOP=TOP-1: GOTO 3000 : REM no removable wall
3110 ON TOERASE GOTO 3200, 3200, 3300, 3200, 3400, 3400, 3500, 3200, 3400, 3400,
     3500, 3300, 3500, 3500
3190 REM routines to select which wall to erase
3197 REM This routine handles case where only one wall can be erased
3199 REM The cell is removed from the list as it will have no more walls
3199 REM that can be removed
3200 TOP=TOP-1: GOTO 3600
3218 REM This routine handles case where there are two walls that can be
3219 REM erased. The walls are opposite each other. Make a random choice.
3300 IF RND<.5 THEN TOERASE=FIRST(TOERASE) ELSE TOERASE=SECOND(TOERASE)
3310 GOTO 3600
3397 REM This routine handles case where there are two walls that can be
3398 REM erased. The walls are adjacent so see if one is opposite an
3399 REM already removed wall.
3400 REM we first find which walls of cellno have been removed
```

```
3405 GONE=0
3410 IF ROW>1 THEN IF BWALL(ROW-1,COL)=SELECTED THEN GONE=GONE+1
3420 IF ROW<MAZDIM THEN IF BWALL(ROW,COL)=SELECTED THEN GONE=GONE+2
3430 IF COL>1 THEN IF RWALL(ROW,COL-1)=SELECTED THEN GONE=GONE+4
3440 IF COL<MAZDIM THEN IF RWALL(ROW,COL)=SELECTED THEN GONE=GONE+8
3448 REM Choose probabilities depending on which walls have been removed
3349 REM Then select randomly which of the two walls to remove
3450 IF GONE=LOWER(TOERASE) THEN BOUND!=.6 ELSE IF GONE=UPPER(TOERASE) TH
     BOUND!=.4 ELSE BOUND!=.5
3455 RAND!=RND
3460 IF RAND!<BOUND! THEN TOERASE = FIRST(TOERASE) ELSE TOERASE =
     SECOND(TOERASE)
3470 GOTO 3600
3497 REM In this case there are three walls that could be remove. Choose
3498 REM probabilities to give higher probability to one opposite an
3499 REM already removed wall, and select wall
3500 RAND!=RND
3510 IF RAND!<.4 THEN TOERASE=FIRST(TOERASE) ELSE IF RAND!<.7 THEN
     TOERASE=SECOND(TOERASE) ELSE TOERASE=THIRD(TOERASE)
3520 GOTO 3600
3598 REM We now remove the selected wall
3600 REM PRINT "toerase after modification";TOERASE
3601 IF TOERASE=1 THEN ROW=ROW-1: BWALL(ROW,COL)=SELECTED: GOTO 3700
3610 IF TOERASE=2 THEN BWALL(ROW,COL)=SELECTED:ROW=ROW+1: GOTO 3700
3620 IF TOERASE=4 THEN COL=COL-1: RWALL(ROW,COL)=SELECTED: GOTO 3700
3630 IF TOERASE=8 THEN RWALL(ROW,COL)=SELECTED:COL=COL+1: GOTO 3700
3700 CELL(ROW,COL)=REACHABLE
3709 REM add newly reachable cell to list
3710 TOP=TOP+1: STACK(TOP)=MAZDIM*(ROW-1)+COL-1
3800 GOTO 3000
3990 RETURN
3999 REM ------- subroutine to draw maze --------
4000 CLS
4010 SCREEN 1
4030 REM draw top line (bottom of imagined cells in row 0)
4035 REM the 3 and 4 in the following try to cope with the aspect ratio
4036 REM of the IBM PC screen
4040 Y=0: FOR I=2 TO MAZDIM: X=4*I: GOSUB 4220 : NEXT I
4050 REM draw bottom line
4060 Y=3*MAZDIM: FOR I=1 TO MAZDIM-1: X=4*I: GOSUB 4220 : NEXT I
4070 REM draw left side
4080 X=0: FOR I=1 TO MAZDIM: Y=3*I: GOSUB 4210 : NEXT I
4090 REM draw right side
4100 X=4*MAZDIM: FOR I=1 TO MAZDIM: Y=3*I: GOSUB 4210 : NEXT I
4110 REM draw all right walls
4120 FOR I=1 TO MAZDIM: FOR J=1 TO MAZDIM-1: Y=3*I: X=4*J
4125 IF RWALL(I,J)<>SELECTED THEN GOSUB 4210
4130 NEXT J: NEXT I
4140 REM draw all bottom walls
4150 FOR I=1 TO MAZDIM-1: FOR J=1 TO MAZDIM: Y=3*I: X=4*J
4155 IF BWALL(I,J)<>SELECTED THEN GOSUB 4220
4160 NEXT J: NEXT I
4209 REM subroutine to draw wall at right of cell
4210 LINE (X+6,Y+2)-(X+6,Y+5)
4213 RETURN
4219 REM subroutine to draw wall at bottom of cell
4220 LINE (X+2,Y+5)-(X+6,Y+5)
4223 RETURN
```

How does the description of maze generation relate to graph theory? A graph consists of elements called vertices, and some pairs of vertices are called edges. A graph is often represented on paper by drawing dots to represent the vertices, and by drawing lines joining vertices to represent the edges. For example, the following are pictures of graphs.

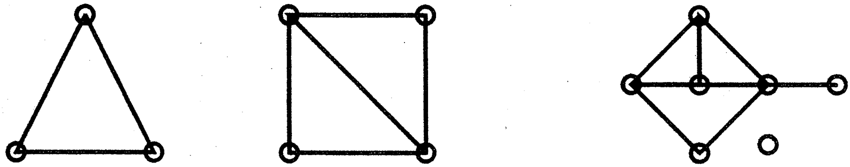


Figure 4. Some graphs

Now we regarded a maze as a consisting of cells, and some pairs of cells (adjacent ones) have walls connecting them. This is just a graph - the cells are the vertices of the graph and the edges are the walls between adjacent cells. We can represent this diagrammatically be putting a dot in the middle of each cell (to represent the vertices) and by joining the dots in adjacent cells with lines (the lines represent the edges). This is illustrated in Figure 5a, which shows the graph corresponding to the grid of Figure 1a. In Figure 5b we show the graph of the maze of Figure 1b (in this case we have only joined dots where the wall between the corresponding cells has been removed).

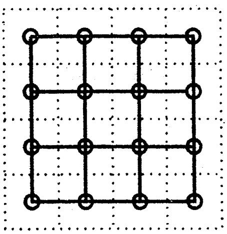


Figure 5a

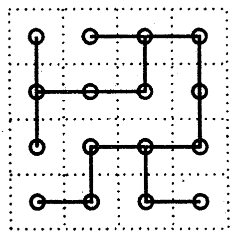


Figure 5b

If we take a graph, and choose only some of the edges, we get what is called a *subgraph* of the original graph. If we look carefully we see that the graph in Figure ?b is a subgraph of the graph of the grid. It is a special type of subgraph, called a *spanning tree* of the graph. With some graphs, a number (called a *cost*) may be associated with each edge. In this case there are special spanning trees (called *minimal spanning trees*) for which the total cost of all the edges of the tree is least. Our first method for generating a maze, involving selection of the wall with the smallest random number associated with it, is really just one of two best-known algorithms for constructing a minimal spanning tree of a graph.

PROBLEMS AND SOLUTIONS

EDITOR: H. LAUSCH

We cordially welcome our subscribers in 1991 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.

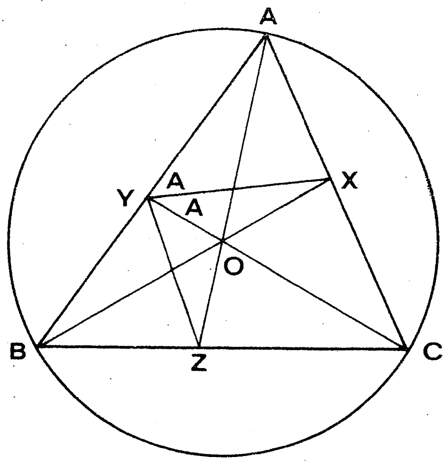
Solutions

K.R. Sastry of Addis Ababa, Ethiopia, writes: "The October 1990 issue of FUNCTION was rushed to me by air mail as I authored the note AN IDEA. The fortunate by-product is that I saw your 'PROBLEM OF THE YEAR' sooner. Here is my solution to it."

The 1990 PROBLEM OF THE YEAR, its title being derived from its popularity as readers of Volume 14 witnessed, is:

Question 4, Function, Vol. 13, Part 3 (June 1989), p.96. Let O be the circumcentre of the triangle ABC, and let X and Y be the points on AC and AB respectively such that BX intersects CY in O. Suppose $\angle BAC = \angle AYX = \angle XYC$; determine the size of this angle.

Solution. Extend AO to meet BC at Z. Now in ΔAYC, YX bisects $\angle AYC$. So, by the Angle-Bisector Theorem, $\frac{CX}{XA} = \frac{CY}{YA}$. But, by Ceva's Theorem, $\frac{BZ}{ZC} \cdot \frac{CX}{XA} \cdot \frac{AY}{YB} = 1$. So $\frac{BZ}{ZC} = \frac{BY}{YC}$, and thus (by the converse of the Angle-Bisector Theorem) YZ bisects $\angle BYC = 180^\circ - 2A$. This gives $\angle OYZ = 90^\circ - A = \angle OBZ$ (as $\angle BOC = 2A$ and ΔBOC is isosceles). The angles $\angle OBZ$ and $\angle OYZ$ are thus subtended by the chord OZ; hence the quadrilateral $BZOY$ is cyclic. It follows that $B + 2B = 180^\circ$, so $B = 60^\circ$.



From ΔAYC, $90^\circ - B = 180^\circ - 3A$. Therefore $A = 50^\circ$.

Problem 14.5.4. ABC is a triangle, O its incentre. Show that AO passes through the circumcentre of BOC.

First solution (John Barton, North Carlton). Draw the circumcircle of the triangle ABC and let AO intersect it again in Y. We note that the incentre O is necessarily inside the triangle ABC and that Y is outside it, so that $\angle OBY$ is the sum of $\angle OBC$ and $\angle CBY$. Join YB, YB. Since $\angle BAY = \angle CAY$, (data), $BY = YC$, these being chords of the congruent arcs BY, YC. Also

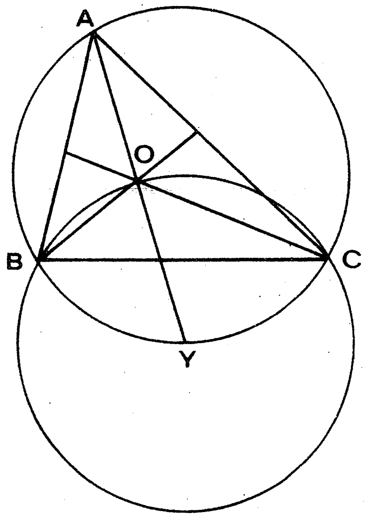
$$\angle BAY = \angle BCY \quad \text{(angles in same segment)}$$

$$\angle CAY = \angle CBY \quad \text{(angles in same segment).}$$

Hence $\angle BCY = \angle CBY = \frac{1}{2}A$.

$$\angle BOY = \angle OAB + \angle OBA \quad \text{(exterior angle theorem)}$$
$$= \tfrac{1}{2}A + \tfrac{1}{2}B\ ;$$

$$\angle OBY = \angle OBC + \angle CBY, \text{ (refer to the above note),}$$
$$= \tfrac{1}{2}B + \tfrac{1}{2}A.$$



That is, $\angle BOY = \angle OBY$, whence

$$BY = OY \quad \text{(sides opposite equal angles).}$$

Thus we have $BY = YC = OY$, so that the circle, centre Y, passing through B also passes through O and C, that is, it is the circumcircle of the triangle BOC, q.e.d.

Second solution (Garnet J. Greenbury, Brisbane). The internal bisector of angle A is extended to meet the circumcircle at Y, the midpoint of the arc BC not containing A. Then YM is the diameter perpendicular to BC.

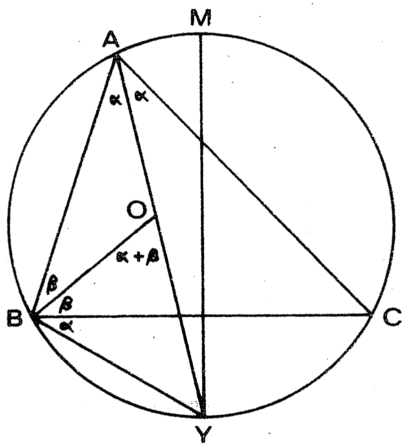
Let $\alpha = \frac{1}{2}A$ and $\beta = \frac{1}{2}B$.

Now, $\angle YBC = \angle YAC = \alpha$ (same arc),

$\angle BOY = \alpha + \beta$ (external angle),

$YB = YO$ (isosceles),

$YB = YC$ (perpendicular bisector).



Thus Y is the circumcentre of BOC.

Problem 14.5.6. Let x, y, z be three non-zero real numbers which are distinct from each other. If $x + y + z = 0$, what is the product

$$\left\{ \frac{y-z}{x} + \frac{z-x}{y} + \frac{x-y}{z} \right\} \cdot \left\{ \frac{x}{y-z} + \frac{y}{z-x} + \frac{z}{x-y} \right\} ?$$

Solution (John Barton). We use here the classical identity

$$A^3 + B^3 + C^3 - 3ABC = (A + B + C)(A^2 + B^2 + C^2 - AB - BC - CA).$$

(i) Bringing the terms of the first factor to a common denominator, the first factor is

$$\frac{1}{xyz} \Sigma\, xy(x-y) = \frac{1}{xyz} \Sigma\, (x^2y - xy^2) \qquad \text{[the sums being cyclic sums]}$$

$$= -\frac{1}{xyz}(x-y)(y-z)(z-x) \qquad \text{[as can be checked by multiplying out].} \qquad (1)$$

(ii) Dealing similarly with the second factor, it is

$$\frac{1}{(x-y)(y-z)(z-x)} = \Sigma z(y-z)(z-x). \qquad (2)$$

Now $\Sigma z(y-z)(z-x) = \Sigma(-xyz + z^2x + yz^2 - z^3)$

$$= \Sigma(-xyz - 2z^3), \text{ since } z^2x + yz^2 + z^3 = z^2(x+y+z) = 0,$$

$$= -3xyz - 2z^3. \qquad (3)$$

Since $x + y + z = 0$, we have, from the above identity,

$$\Sigma z^3 = 3xyz.$$

Hence the right side of (3) is

$$-3xyz - 2(3xyz) = -9xyz. \qquad (4)$$

Now from (1), (2), (3), (4) the product is 9.

Problems

David Shaw (Newtown, Victoria), whom we have known as a contributor to FUNCTION for many years, writes: "This is a problem from the Mathematics Tripos examination. I don't know the year!" Is there a reader who could help us with the date?

Problem 15.1.1. Given a set of n straight lines whose lengths are $1, 2, 3, ..., n$ inches [*ed.: 1 inch = 2.54 cm; the inch used to be an official unit of length measurement in Australia – ask your teacher what it was like to cope with a non-metric system!*] respectively. Prove: the number of ways in which four may be chosen from this set so that they form a quadrilateral in which a circle may be inscribed is

$$\frac{1}{48}\{2n(n-2)(2n-5) - 3 + 3(-1)^n\}.$$

David Shaw notes: "Compare this problem with Question 4 (Paper 1) of the Australian Mathematical Olympiad 1989".

We re-open this problem for the benefit of new subscribers to FUNCTION:

Problem 15.1.2. Let n be even. Four different numbers a, b, c, d are chosen from the integers $1, 2, ..., n$ in such a way that $a + c = b + d$.

Show that the number of such selections is $\dfrac{n(n-1)(2n-5)}{24}$.

Apropos of Mathematical Tripos problems, David Shaw let us have another:

Problem 15.1.3. A person sits for an examination in which there are four papers with a maximum of m marks for each paper; show that the number of ways in which a total of $2m$ marks may be obtained is $\frac{1}{3}(m+1)(2m^2+4m+3)$.

In connection with his solution to the 1990 PROBLEM OF THE YEAR, K.R.S. Sastry provided the following:

Problem 15.1.4. Let O be the circumcentre of an acute-angled triangle ABC. Extended, BO, CO meet the sides AC, AB at the points X and Y respectively.

(i) Characterize triangle ABC if YX bisects $\angle AYC$.
(ii) Determine the angles of ΔABC if $\angle AYC = 2A$.

Problem 15.1.5 (Garnet J. Greenbury). If the bisectors of two angles of a triangle are equal, the triangle is isosceles. We want a *Euclidean proof.*

Trigonometric proofs are acceptable.

Problem 15.1.6 (Garnet J. Greenbury). Let O and I be the circumcentre and incentre, respectively, of a triangle with circumradius R and inradius r ; let d be the distance OI. Show that $d^2 = R^2 - 2rR$.

Year Twelve International

The following problems are from a sixth form course in Great Britain.

Problem 15.1.5. A young man has two girlfriends who live in diametrically opposite directions from his house and whom he normally sees about an equal number of times each month. The bus route which goes to both their houses passes the front of his house. He decides to leave his house at random times and to take the first bus which comes along. Since the buses run with the utmost regularity he thinks he will see his girlfriends an approximately equal number of times. He follows this regime for a few months and finds that he has been to see one girlfriend five times as often as he has been to see the other! Why is this?

Problem 15.1.6. Each day a wife leaves home by car to collect her husband at the station when it is 18.00 hours. Today her husband arrives at 17.00 hours and sets out walking at 4 kilometres per hour. The wife sets out at the usual time, meets him on the road, and they get home 20 minutes earlier than usual. Find the average speed of the car.

Problem 15.1.7. Show that if $\frac{m}{n}$ is an approximation of $\sqrt{2}$, then $\frac{m+2n}{m+n}$ is a better one.

Problem 15.1.8. An infinite set of points in the plane is such that the distance between any two is an integer. Prove that the points are all collinear.

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.

Problem 15.1.10. Show that no integer of the form $8k + 7$ is the sum of three perfect squares.

* * * * *

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

* * * * *

Published by Monash University Mathematics Department