Mathematical Spectrum

Volume 13 1980/81 *Number 3*

A Magazine of *Contemporary Mathematics*

Published by the *Applied Probability Trust*

Mathematical Spectrum is a magazine for the instruction and entertainment of student mathematicians in schools, colleges and universities, as well as the general reader interested in mathematics. It is published by the Applied Probability Trust, a non-profit making organisation established in 1963 with the support of the London Mathematical Society. The object of the Trust is the encouragement of study and research in the mathematical sciences.

Volume 13 of *Mathematical Spectrum* consists of three issues, of which this is the third. The first issue was published in September 1980 and the second in January 1981.

Articles published in *Mathematical Spectrum* deal with the entire range of mathematical disciplines (pure mathematics, applied mathematics, statistics, operational research, computing science, numerical analysis, biomathematics). Both expository and historical material may be included, as well as elementary research and information on educational opportunities and careers in mathematics. There is also a section devoted to problems.

Articles are normally commissioned by the Editors; the Editorial Committee also welcomes the submission of suitable material, including correspondence, queries and solutions to problems, for publication in *Mathematical Spectrum*. All correspondence about the contents should be sent to:

The Editor, *Mathematical Spectrum*,
Hicks Building, The University, Sheffield S3 7RH.

The Dawn of the Computer Age—The Life and Work of Alan Turing

KEITH AUSTIN
University of Sheffield

Keith Austin is a Lecturer in Pure Mathematics at the University of Sheffield. He was an undergraduate and postgraduate student at Manchester and has taught in a college of education. His interests include logic, combinatorics, computer science, mathematical recreations and the understanding and teaching of mathematics.

Outside mathematics his main occupation is service with his local Methodist church. His recreation is watching old (1930–1955) movies on television.

Alan Turing is probably one of the best-known British mathematicians of the twentieth century. His concept of the Turing machine is fundamental to computer science and mathematical logic. The Association for Computing Machinery, with over 40000 members, makes the Turing Award, its most prestigious technical award, to an individual who has made contributions of lasting and major technical importance to the computing community. (Recently, when the Academy of Sciences of the U.S.S.R. asked for permission to reprint the Turing lectures given since 1973, six of the seven Turing lecturers involved withheld their permission as a protest against the treatment of Soviet dissidents.)

In order to see the significance of Turing's work we must consider the history of mathematics. There are two basic questions which mathematicians have had to face over the years. The first is 'what is a proof?'. This question falls into logic and we can see the answer developing as we move from Aristotle and Euclid through to Lobachevsky, Boole and Frege in the nineteenth century.

The second question is 'what is an algorithm?'. Traditionally, an algorithm is an effective procedure for calculating the value of some quantity or for finding the solution of some mathematical problem. It usually consists of a set of instructions about how to behave. If we follow the instructions properly then we are sure to end up with the answer. The algorithm reduces the problem of finding the solution from a task of intellectual discovery to a mere matter of effort.

The best-known examples of algorithms are the four for adding, subtracting, multiplying and dividing numbers written in the decimal notation. Another example is the procedure known as Euclid's algorithm which can be used to find the highest common factor of two integers. In the same way that the division algorithm consists of repeated applications of the subtraction algorithm, so Euclid's algorithm consists of repeated applications of the division algorithm.

On the whole, proofs follow a common pattern which was set by the Greeks and remains to this day. Algorithms, however, appear to come in all shapes and sizes, and no pattern had emerged by the early part of this century. The best that could be said at this time was that, if anyone produced an algorithm, then it was clear that it was one. In 1936 Alan Turing answered the question 'what is an algorithm?'. His answer is deceptively simple. Using an algorithm is 'mechanical' work, so an algorithm is a machine.

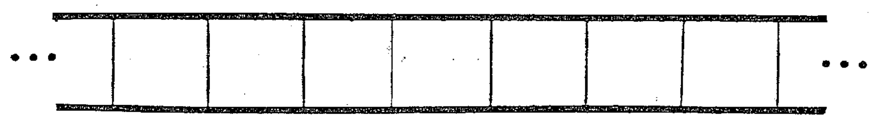
Alan Mathison Turing was born on 23 June 1912 in London. He was educated at Sherborne and then, from 1931, at Cambridge. In 1936 he published his most famous work, in which he introduced the concept of the Turing machine. The problem that Turing had been working on was to find an algorithm to decide whether any statement of that branch of logic called the predicate calculus was valid or not. Logicians had worked on the problem for many years without success, and so Turing decided to try and prove that there was no such algorithm. However, if his proof was to be precise, then it was necessary first to give a precise definition of an algorithm. This he did in two stages:

(a) He gave a precise definition of a Turing machine.
(b) He stated that an algorithm is simply a Turing machine.

This statement is known as 'Turing's Thesis'. He then produced evidence and arguments to support his thesis.

From this basis Turing proceeded to prove that no algorithm exists which can solve the predicate calculus problem and so the problem is unsolvable.

The easiest way to think of a Turing machine is as a computer which has had a program put in and is now waiting for the data to be put in, when it will proceed with the calculation. A Turing machine consists of an infinite tape of blank squares,



a finite list of symbols which may be written on the tape, a finite set of cards with symbols on them, and a table of operating instructions. The cards are called *state cards*, and one is called the *starting state card.*

To begin an operation, a finite number of squares of the tape have symbols written in them, and the starting state card is placed below the filled-in square which is farthest to the left. This corresponds to putting the data into the computer.

The operation proceeds by a sequence of steps. At each step, look at the symbol on the state card below the tape and the symbol—possibly blank—in the tape square above the state card. In the operating table you will find, corresponding to these two symbols, an instruction to be carried out. The instruction will consist of three parts.

(i) Replace the state card by another, specified, state card.
(ii) Rub out the symbol in the tape square above the state card and write in another, specified, symbol.

(iii) One of the following:
    (a) move the state card one square to the left,
    (b) move the state card one square to the right,
    (c) leave the state card where it is.

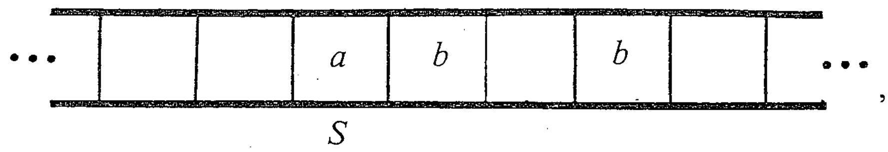
Carry out the instruction. This completes the step. Proceed to the next step and work in the same way.

As an example, consider the Turing machine with tape symbols a, b, states S, T, where S is the starting state, and operating table

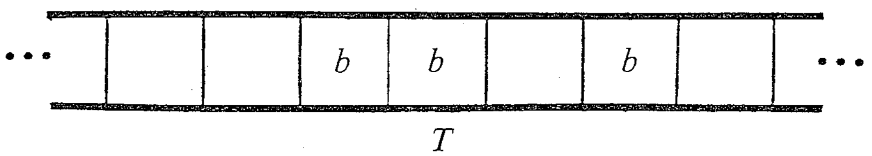
$S\ a$	$T\ b \rightarrow$
$S\ b$	$S\ a \leftarrow$
$S\ B$	$S\ B \rightarrow$
$T\ a$	$S\ a\ 0$
$T\ b$	$T\ B \rightarrow$
$T\ B$	$S\ b\ 0$

Note that B denotes blank. The table is used in the following way. Suppose the symbol on the state card below the tape is S and the symbol in the tape square above the state card is a. By looking at the left-hand two columns of the table we find the row corresponding to Sa; it is the first row. The instruction in that row is $Tb \rightarrow$ which means replace the S card by the T card, rub out the a and write b, and move the state card one square to the right. The instructions in the other rows are of a similar kind, except that $\leftarrow$ means move the state card to the left, and 0 means do not move the state card.

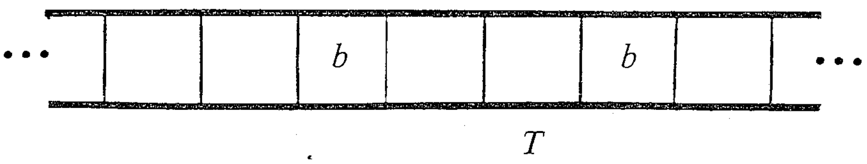
Thus suppose we start with the following situation:

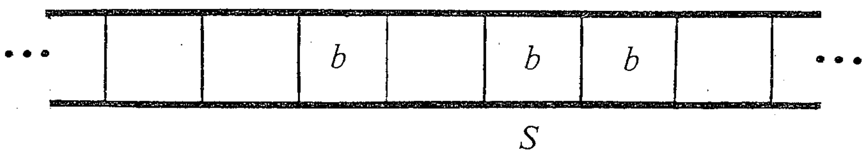


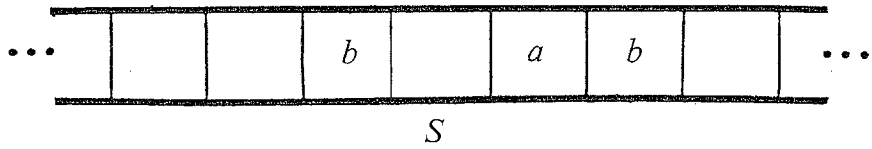
then after one step of the operation we reach the situation

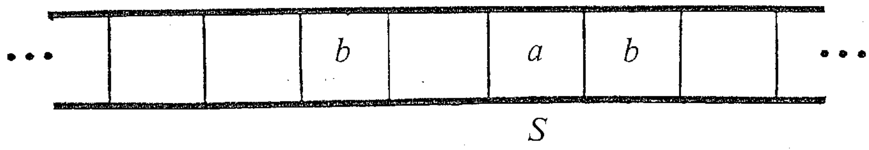


Continuing with the operation we obtain in turn:









etc.

We describe the above as the operation of the Turing machine on the string *abBb*. We call the string *abBb* the starting string of the operation.

For our second example consider the Turing machine with tape symbol |, states S, T, where S is the starting state and operating table

S |	T | $\rightarrow$
$S\ B$	$S\ B\ 0$
T |	S | $\rightarrow$
$T\ B$	$T\ B\ 0$

The symbol | (stroke) is used as this is an example from stroke arithmetic, where each positive integer n is denoted by a string of n strokes, i.e. we are using the unary system.

We consider the operation of the Turing machine on a string of n strokes. e.g. $n = 5$. We obtain the following sequence of situations:

||||| ,	||||| ,	||||| ,	||||| ,	||||| ,	||||| ,	||||| ,	...
S	T	S	T	S	T	T	

The situation remains the same from this point on.

For general n it is not difficult to see that the Turing machine eventually settles permanently at state T if n is odd and at state S if n is even. Thus the Turing machine is able to distinguish between odd and even numbers.

Our third Turing machine also deals with stroke arithmetic. The tape symbols are | and $-$, the states are S, T, U, V, where S is the starting state and the operating table is

S |	S | $\rightarrow$
$S\ -$	$S - \rightarrow$
$S\ B$	$T\ B \leftarrow$
T |	$U\ B \leftarrow$
$T\ -$	$T\ B\ 0$
$T\ B$	$T\ B\ 0$
U |	U | $\leftarrow$
$U\ -$	$U - \leftarrow$
$U\ B$	$V\ B \rightarrow$
V |	$S\ B \rightarrow$

We consider the Turing machine operating on strings of the form

$$\underbrace{|\ |\ |\dots|\ |}_{\leftarrow n\text{ strokes}\rightarrow} - \underbrace{|\ |\ |\dots|\ |}_{\leftarrow m\text{ strokes}\rightarrow}$$

where $n > m$. In such operations, situations requiring rows of the table beginning $V-$ and VB do not occur and so we have omitted such rows from the table.

If the Turing machine operates on the string $|\ |\ -\ |$ then we obtain the following sequence of situations:

$$\begin{array}{ccccccc} |\,|-|\,, & |\,|-|\,, & |\,|-|\,, & |\,|-|\,, & |\,|-|\,, & |\,|-|\,, & |\,|\,|-B\,, \\ S & S & S & S & S & T & U \end{array}$$

$$\begin{array}{cccccccccc} |\,|-\,, & |\,|-\,, & |\,|-\,, & |\,|-\,, & B\,|-\,, & |-\,, & |-\,, & |-\,, & |\,B\,, & |\,B\,,\ \dots \\ U & U & U & V & S & S & S & T & T & T \end{array}$$

The situation remains the same from this point on.

It is not difficult to see that, if the starting string is

$$\underbrace{|\ |\ |\dots|\ |}_{\leftarrow n\text{ strokes}\rightarrow} - \underbrace{|\ |\ |\dots|\ |}_{\leftarrow m\text{ strokes}\rightarrow},$$

then the Turing machine eventually settles permanently into a situation where the string on the tape is

$$\underbrace{|\ |\ |\dots|\ |}_{\leftarrow (n-m)\text{ strokes}\rightarrow}$$

Thus the Turing machine is able to perform subtraction.

A question about Turing machines and multiplication occurs in the problems section.

After a little practice with Turing machines, it becomes clear that they are as powerful as any computer. Any calculation a computer can do can be done by a Turing machine, although the latter may take a very long time.

A computer is described as a general-purpose machine because it can be programmed to do various different tasks. We said above that a Turing machine corresponds to a computer which has already been programmed. Thus we say that a Turing machine is a special-purpose machine. Turing was, however, able to construct a special Turing machine, now called a universal Turing machine (UTM), which operates like a general-purpose computer. The UTM is very large, with about 100 states and 100 tape symbols and hence about 10000 rows in its operating table.

To see how the UTM works, select any other Turing machine M. As an example we will take M to be the first example given above. We first standardize M by replacing all symbols by numbers, so:

$$a - 1, b - 2, B - 3, S - 4, T - 5, \rightarrow - 6, \leftarrow - 7, 0 - 8.$$

Write M's operating table in a single string, so:

$$4152642417\dots53428.$$

Put this string on to the UTM's tape. This corresponds to putting the program into the computer; it guarantees that the UTM will operate like M.

Next select a starting string for M, e.g. *abBb*, which we considered above. Now the complete starting situation is

$$\begin{array}{c} abBb \\ S \end{array}$$

and we shall write this as *SabBb*. Finally replace the symbols by numbers, so *SabBb* becomes 41232. Write this onto the UTM's tape immediately to the right of M's table, which has already been written on. This corresponds to putting the data into the computer.

Put the UTM's starting state card under the left-hand number on the tape and proceed with the operation of the UTM on this string:

table section ⟵⟶	string section ⟵⟶
4 1 5 2 . . . 5 3 4 2 8	4 1 2 3 2

□

(As the UTM's operating table is not given in this article you will not be able actually to carry out the operation, but it is not difficult to imagine as it follows the same procedure as any other Turing machine.) The UTM operates so as to change the string section of the tape, using the table section as instructions on what changes to make. The UTM is so devised that it changes the string section in precisely the same way that M changes

$$\begin{array}{c} abBb \\ S \end{array}$$

in our example. Thus it changes 41232 to 25232, the numerical form of *bTbBb*, which is the way we write

$$\begin{array}{c} bbBb, \\ T \end{array}$$

the second situation in our example. Note that we write the state symbol immediately to the left of the tape symbol below which it occurs. Continuing, the UTM will change 25232 to 23532 which represents

$$\begin{array}{c} bBBb \\ \;\;T \end{array}$$

and so on.

Of course the UTM makes the changes rather slowly as, at each stage, it has to search along the table section of the tape to find the relevant instruction, and this takes many steps. It reads and remembers the changes it has to make and then goes back to the string section and makes the changes.

Thus the operation of the UTM on 415 . . . 2841232 simulates the operation of M on *abBb*. The situation is the same whatever Turing machine M and starting string for M we choose.

Returning to general Turing machines, we shall now add a new type of instruction. We shall allow the right-hand side of a row of the operating table to consist of the instruction 'halt'. This allows us to formulate the following halting problem.

Choose any Turing machine M. Write out M's operating table as a string of numbers, as above. This will include replacing 'halt' by a number. Select any tape symbol of M other than blank, say a. Transform the string of numbers into a string of a's and blanks by replacing each number n by the string

$$\underbrace{Baa\ldots aB}_{\leftarrow\ n\,a\text{'s}\ \rightarrow}$$

The resulting string is described as M's table in coded form.

Consider the operation of M on the starting string consisting of its own table in coded form. The question is, does M halt in this operation? Clearly for some M the answer is yes and for others it is no. Can we construct a Turing machine which tells us which, for any M? Consider the two possibilities. If M is going to halt then we can obtain the answer to our question by watching M operate and waiting for it to halt. However, if M is not going to halt, then it is useless just to watch M as we shall never get an answer. So what we want is a Turing machine A which does the following:

> Given any Turing machine M, if we set A to operate on M's table i.e. on the starting string consisting of M's table in coded form, then A will halt if and only if M does not halt when set to operate on its own table.

The Turing machine A is precisely what we want to answer the halting problem, for we simply set both M and A operating on M's table. If M halts the answer is yes and if A halts the answer is no.

However, Turing proved that no such machine A exists. The reason is easy to see. Suppose there was such an A. Set A operating on its own table. Then A halts if and only if A does not halt when set to operate on its own table.

This result, the unsolvability of the halting problem, is the starting point for all work on proving the unsolvability of particular problems. That a problem is unsolvable means there is no Turing machine which will solve it. In 1931 the Austrian logician Kurt Gödel considered all the statements about the arithmetic of the numbers 0, 1, 2, 3, . . . with addition and multiplication. With the advent of the Turing machine in 1936 it was clear that Gödel's work readily showed that there is no Turing machine which decides whether any given statement of this arithmetic is true or false.

At the International Congress of Mathematicians in Paris in 1900, the German mathematician David Hilbert asked if there was an algorithm which decided whether or not any given polynomial equation with integer coefficients, such as

$$1 + 2x^2 = 7y^4 + 4xyz^6$$

has an integer solution. The Russian mathematician Yuri Matiyasevic showed in 1970 that there is no Turing machine which answers this question.

In general, if we wish to show that a particular problem P is unsolvable, then the method of working is to try to connect P to the halting problem in some way, so that, if we had a Turing machine which solved P, then we could use it to construct a Turing machine to solve the halting problem. It then follows that no Turing machine for P exists. Having proved P is unsolvable we may be able to prove that another problem P' is unsolvable by showing that a Turing machine for P' could be used to solve P; and so on.

We now consider the argument in support of Turing's thesis. Why can a person carrying out an algorithm be compared to a Turing machine operating? The restriction to a one-dimensional tape as against two-dimensional sheets of paper is clearly of no significance. The restriction to a finite number of tape symbols is reasonable, as we can use finite sequences of symbols in place of extra symbols. The fact that we only look at one tape square at a time corresponds to the fact that a person has a bounded range of observation. The individual operations are obtained by simply breaking down the person's operations as far as possible. The states correspond to the states of mind of the person as he carries out the algorithm. For example, at one point he may have in his mind that he is to add 3 to the last number he wrote down, while at another point that he is to subtract 7 from the last but one number he wrote down. Although an infinite number of states of mind appears possible, as the person is carrying out an algorithm one would not expect a large number of different states of mind to occur. The essence of an algorithm is that the person is saved too much thinking and the task is performed, to a large extent, by means of routine paper work.

The above argument comes from Turing's paper, but since then further evidence has arisen in that alternative definitions of algorithm have been proposed, and these have all been proved to be equivalent to Turing's definition. In the forty years since Turing proposed his thesis no one has produced an algorithm which cannot be performed by a Turing machine. Turing's thesis is now generally accepted by workers in computer science and mathematical logic.

In 1936 Turing went to Princeton to work with the American logician Alonzo Church. He remained in America until 1938 when he returned to Cambridge.

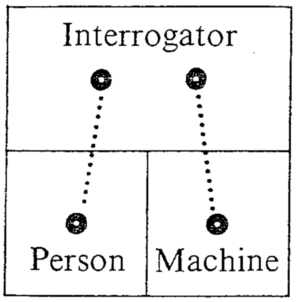
With the outbreak of war in 1939, Turing joined the Foreign Office's Department of Communications at Bletchley Park in Buckinghamshire. Although much of this work is still officially secret, some details have been released during the last year or so and these suggest that important work on codes and ciphers was involved.

One of the most historically important developments at Bletchley Park was the building of the Colossus series of computers. The first of the series went into operation in 1943 and was the world's first programmable electronic computer. The design of the Colossus computers was apparently influenced by Turing's pre-war work. Also working at Bletchley Park was the mathematician M. H. A. Newman who was to be involved with Turing after the war in the development of computers at

Manchester University. The actual construction of the Colossus computers was under the control of a team from the Post Office's Research Establishment at Dollis Hill and the government's Telecommunications Research Establishment, led by T. H. Flowers. When the war ended in 1945, Turing was awarded the O.B.E. for his work.

Turing next joined the National Physical Laboratory and produced, in 1945, a report containing one of the earliest complete designs for a stored-program computer. Neither the Colossi nor the American ENIAC (1945) was a stored-program computer; they were programmed by switches and plugs. The first stored-program computer was a small experimental machine built at Manchester University in 1948 by F. C. Williams and T. Kilburn from the Telecommunications Research Establishment. Turing moved to Manchester University in 1948 and was involved in the initial work on the design of programs for the computer. The first practical stored-program computer was built at Cambridge between 1947 and 1949 and was based on the development of the American EDVAC, including John von Neumann's 1945 'Draft Report' which contained a description of the planned machine and the reasoning behind the various design decisions.

In 1950 Turing considered the question 'can machines think?'. To decide whether or not a particular machine can think, he proposed that it should be subjected to the following test, which has since become known as Turing's test.



The test involves the machine, two people and three rooms. The machine is put in one room and the two people are put separately in the other two rooms. One person is called the interrogator and he can communicate with the occupants of the other two rooms by teleprinter. The interrogator does not know which teleprinter connects with the machine and which with the person. The interrogator asks questions by means of the teleprinters, and the person and the machine must answer. The questions can be of any kind and are not restricted to mathematics. If the interrogator is unable to find out which teleprinter connects with the machine then the machine has passed the test and we say that it can think.

Alan Turing died on 7 June 1954.

Odd Binomial Coefficients

D. G. NORTHCOTT
University of Sheffield

Professor Northcott is Town Trust Professor of Pure Mathematics at the University of Sheffield. He has made important contributions to the abstract theory of commutative algebra. He is also interested in presenting mathematics to a wider, non-specialist, audience; this article is an attractive example.

If n is a positive whole number and

$$(1+x)^n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

then $a_0, a_1, \ldots, a_n$ are the well-known *binomial coefficients*. In fact

$$a_r = \frac{n!}{r!\,(n-r)!}.$$

Binomial coefficients are rich in properties. What concerns us here is that they are whole numbers and therefore each is either an odd number or an even number. Since $a_0 = 1$ and $a_n = 1$, there are at least two odd numbers among $a_0, a_1, \ldots, a_n$. Our aim will be to determine the number and distribution of the odd coefficients. One interesting fact which will emerge is that the number of odd coefficients is always a power of 2.

The key to our problem lies in what may be termed the *binary structure* of n. In the binary scale n is represented by a sequence $\varepsilon_v \varepsilon_{v-1} \cdots \varepsilon_1 \varepsilon_0$ of 0's and 1's. If $\varepsilon_{m_1}, \varepsilon_{m_2}, \ldots, \varepsilon_{m_q}$ are the 1's in the sequence, then

$$n = 2^{m_1} + 2^{m_2} + \cdots + 2^{m_q} \tag{A}$$

and, of course, $m_1, m_2, \ldots, m_q$ are all different. Since a number can be represented in the binary scale in exactly one way, each positive whole number has a single representation of type (A). Note that the number q of terms on the right-hand side is the number of 1's among $\varepsilon_v, \varepsilon_{v-1}, \ldots, \varepsilon_0$ and therefore *q is the sum of the digits in the binary representation of n.*

Consider the numbers that can be obtained by striking out a collection of terms in

$$2^{m_1} + 2^{m_2} + \cdots + 2^{m_q} \tag{B}$$

and summing the rest. (We allow the possibility that all the terms may be deleted in which case the sum of the remainder is taken to be 0. We also permit the case where no term is removed.) The numbers so obtained will be called the *binary partial sums* of n. When striking out terms in (B) we have the choice, for each term, as to whether we leave it in or take it out. This gives $2 \times 2 \times \cdots \times 2 = 2^q$ selections. Since the representation of a positive whole number in the form (A) is unique, we obtain 2^q

different binary partial sums of n. Accordingly if $k_1, k_2, \ldots, k_s$ are the binary partial sums, then $s = 2^q$, the k_i are all different, and

$$(1 + x^{2^{m_1}})(1 + x^{2^{m_2}}) \cdots (1 + x^{2^{m_q}}) = x^{k_1} + x^{k_2} + \cdots + x^{k_s}. \tag{C}$$

For example suppose that $n = 21$. In the binary scale this is 10101 so $21 = 2^4 + 2^2 + 2^0$. The binary partial sums in descending order are

$$\begin{array}{rr} 2^4 + 2^2 + 2^0 = 21 & 2^2 + 2^0 = 5 \\ 2^4 + 2^2 = 20 & 2^2 = 4 \\ 2^4 + 2^0 = 17 & 2^0 = 1 \\ 2^4 = 16 & 0 \end{array} \tag{D}$$

the last one coming from the empty sum. The digits in the binary representation of 21 add up to 3, and $2^3 = 8$ is the number of different binary partial sums.

It will next be shown that

$$(1 + x)^{2^m} = 1 + x^{2^m} + 2f_m(x), \tag{E}$$

where $f_m(x)$ is a polynomial whose coefficients are whole numbers. This is certainly true when $m = 0, 1$. If it is true when $m = k$, then on squaring

$$(1 + x)^{2^k} = 1 + x^{2^k} + 2f_k(x)$$

we find that

$$(1 + x)^{2^{k+1}} = 1 + x^{2^{k+1}} + 2f_{k+1}(x),$$

where

$$f_{k+1}(x) = 2f_k^2(x) + 2f_k(x)(1 + x^{2^k}) + x^{2^k}$$

and this too has whole numbers as coefficients. The general assertion follows by induction.

We are now in a position to solve our problem. If n is expressed in the form (A), then

$$\begin{aligned}(1 + x)^n &= (1 + x)^{2^{m_1}}(1 + x)^{2^{m_2}} \cdots (1 + x)^{2^{m_q}} \\ &= \{1 + x^{2^{m_1}} + 2f_{m_1}(x)\}\{1 + x^{2^{m_2}} + 2f_{m_2}(x)\} \cdots \{1 + x^{2^{m_q}} + 2f_{m_q}(x)\}\end{aligned}$$

by virtue of (E), and therefore

$$(1 + x)^n = (1 + x^{2^{m_1}})(1 + x^{2^{m_2}}) \cdots (1 + x^{2^{m_q}}) + 2g(x),$$

where $g(x)$ is a polynomial whose coefficients are whole numbers. Moreover if we use (C) this reduces to

$$(1 + x)^n = x^{k_1} + x^{k_2} + \cdots + x^{k_s} + 2g(x),$$

where $k_1, k_2, \ldots, k_s$ are the binary partial sums of n. Since $k_1, k_2, \ldots, k_s$ are all different, it follows that $x^{k_1}, x^{k_2}, \ldots, x^{k_s}$ are the powers of x which in the expansion of $(1 + x)^n$ have an odd coefficient. Thus our problem is solved and we may sum up as follows.

If $k_1, k_2, \ldots, k_s$ *are the binary partial sums of* n, then $x^{k_1}, x^{k_2}, \ldots, x^{k_s}$ *are precisely the powers of* x *which in the expansion of* $(1+x)^n$ *have an odd coefficient. If* q *is the sum of the digits in the binary representation of* n, *then* 2^q *is the total number of odd coefficients. In particular, the number of odd coefficients is always a power of* 2.

By way of example consider $(1+x)^{21}$. The binary partial sums of 21 are set out in (D). Hence if

$$(1+x)^{21} = a_0 + a_1 x + a_2 x^2 + \cdots + a_{21} x^{21},$$

then $a_0, a_1, a_4, a_5, a_{16}, a_{17}, a_{20}, a_{21}$ are odd and the remaining coefficients are even.

A New Paradox?

DAVID SINGMASTER
Polytechnic of the South Bank

Let x be the measurement of an angle in radians. Then the corresponding measurement of the angle in degrees, $y = 360x/2\pi = \alpha x$, where we have set $\alpha = 360/2\pi$ for convenience. Now $d/dx \sin x = \cos x$ and so $d/dy \sin y = d/dx \sin \alpha x (dx/dy) = \alpha \cos \alpha x (1/\alpha) = \cos y$, and these hold for any x in radians and any y in degrees. Now imagine that x is a function of t, such that $dx/dt = V$ for all t and $V \neq 0$. Thus $dy/dt = \alpha V$ for all t. Now

$$\frac{d}{dt}\sin x = \cos x \frac{dx}{dt} = V\cos x, \qquad \frac{d}{dt}\sin y = \cos y \frac{dy}{dt} = \alpha V \cos y.$$

When y in degrees $= \alpha x$ with x in radians, then $\sin y = \sin x$, so the derivatives just computed are equal. In particular, when $x = y = 0$, we get $V = \alpha V$???

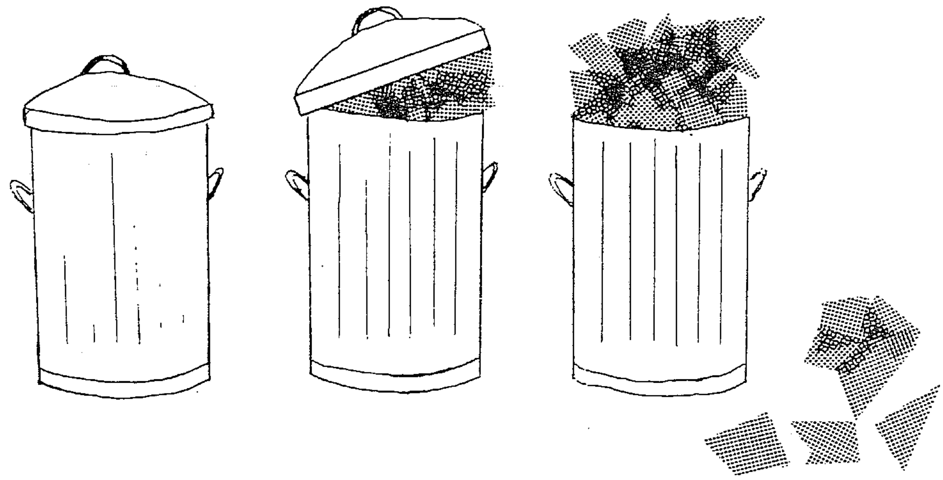
*   *   *

The explanation is given on p. 83 to give you time to think.

The Bin-Filling Problem

MICHAEL HOLCOMBE
The Queen's University of Belfast

Mike Holcombe is a Lecturer in Pure Mathematics at the Queen's University of Belfast. He studied at the University of Sheffield from 1962–1966 and did research in algebra at the University of Leeds from 1966–1968. He moved to Belfast in 1968. Apart from an interest in automata theory and its biological applications, he is also active in the field of near-rings and related systems.



Imagine that we have a number of objects of various weights and we have to put them into some identical containers. The containers are all the same size and there is a limit to the total weight of objects that each container can hold. We want to find a method of packing all the weights using as few containers (or bins) as possible. For example, we have objects with weights 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 5, 5 and some bins with capacity 10. We can pack these objects using 4 bins (Figure 1), and this is

2	2	3	2
3	4	3	3
5	4	4	5

Figure 1

the smallest number of bins of this size that will do. We call it the *optimal solution of the problem*. It is fairly easy to find the optimal solution when there are not many objects, by trial and error, but what do we do if this method is impracticable?

Before we discuss some possible strategies, it might be worth asking ourselves how useful this type of problem is. In many industrial situations, especially those involving stockholding and warehousing, a solution to this type of problem can reduce the amount of stock that is needed, and thus cut the overheads. Another application is in allocating various tasks to a team of workers; each bin represents a worker, and the weights are given by the times taken for each task to be completed. Or perhaps you have to make a wooden fence and need various lengths of wood. The woodyard sells wood in standard lengths. How do you measure up your wood so that you buy as little as possible? I am sure that you can think up other applications yourselves.

So let us try to develop some strategies for solving this problem. We have weights $w_1, w_2, \ldots, w_n$, and bins of size k (where $w_i \leq k$ for every i). Let us choose one bin, put w_1 in it, now look along the list to find the next w_i that fits into the bin, and carry on until the bin is as full as possible, then move to the next bin and repeat the process. In the above example this gives Figure 2, and we need 5 bins. (The shaded areas represent wasted space.) However, let us write the weights in *decreasing order of magnitude*, 5, 5, 4, 4, 4, 3, 3, 3, 3, 2, 2, 2 and try a similar process. This gives Figure 3

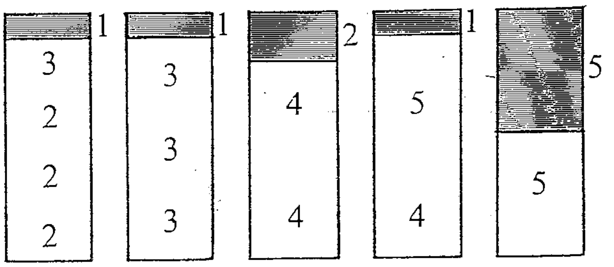


Figure 2

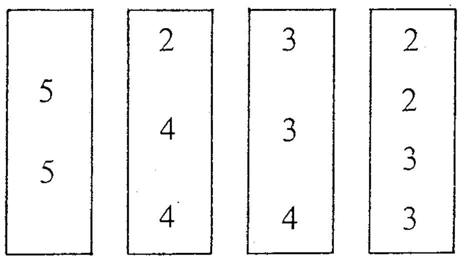


Figure 3

and uses only 4 bins. We now have an optimal solution to this example (there is no wasted space!). This technique is called *First Fit Decreasing* (FFD). However, it does not always give an optimal solution. Let us take a similar problem—5, 4, 4, 4, 3, 3, 3, 2, 2 are the weights and the bin size is again 10. (We have just removed three weights from the previous example.) The FFD strategy gives Figure 4, i.e. 4 bins, but the optimal solution is shown in Figure 5.

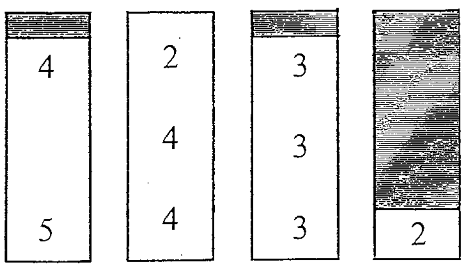


Figure 4

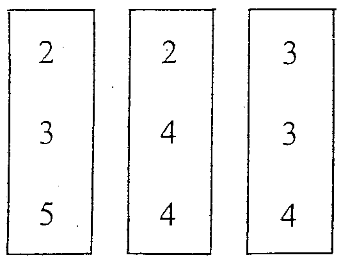


Figure 5

So the FFD strategy does *not* give us the optimal solution in every case, but it is the best general method we have at present. Recently a research mathematician in America proved that, if m is the optimal number of bins and n is the number of bins needed for the FFD strategy, then

$$n \leq \frac{11}{9} \cdot m + 4.$$

(We shall not give the proof here; it takes 75 pages!) This gives us a guide to the maximum 'wastage' involved in using FFD.

One might ask if a computer could be used to sort through all the arrangements systematically to find an optimum one. Unfortunately, this is an example of an 'NP-problem', that is one for which the computer time involved is probably 'not polynomial' but an exponential function of the size of the problem (i.e. the number of weights). Much research, these days, is concerned with these NP-problems, and even great advances in computer technology are unlikely to help here. It is quite simple to find a bin-filling problem involving only 100 weights which cannot be solved by existing computers.

Finally we end with a 'mind-blowing' example taken from the article by R. S. Graham in reference 1. The 33 weights are: one of size 442, seven of size 252, five of size 127, four of size 106, one of size 85, one of size 84, one of size 46, two of size 37, three of size 12, six of size 10 and two of size 9. The bin capacity is 524. Using FFD we get Figure 6.

This is thus an *optimal solution* using 7 bins. If we now *remove* the weight of size 46 and apply FFD to the 32 weights remaining we need 8 bins!

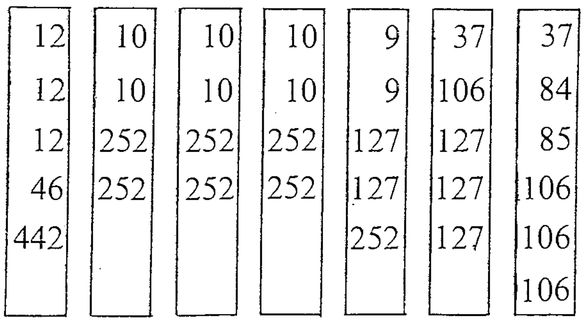


Figure 6

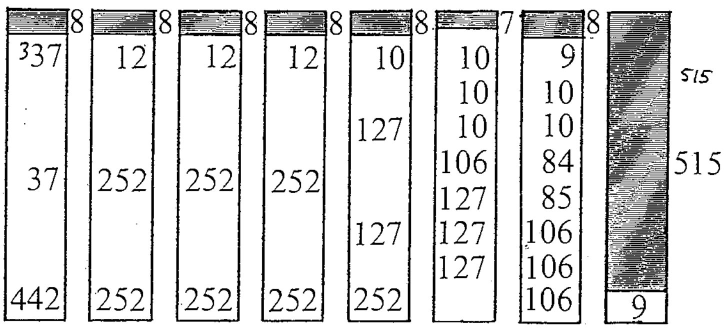


Figure 7

References and further reading

1. An excellent survey book covering many aspects of modern mathematics is *Mathematics Today*, by L. A. Steen and published by Springer (1978).
2. For a discussion of the computational aspects of this type of problem see *Computers and Intractability—a guide to the theory of* NP-*Completeness* by M. R. Garey and D. S. Johnson (W. H. Freeman, 1979).

When Squares are Triangles

DOUG AVERIS
Manor Farm School, Walsall
DAVID SHARPE
University of Sheffield

As a student, Doug Averis attended lectures by David Sharpe on elementary algebra at the University of Sheffield. Despite that ordeal, he is now Head of the Mathematics Department at Manor Farm School. He can be seen many evenings running round the environs of Lichfield Cathedral followed by (or following) his three young sons.

1. Squares and triangular numbers

Everyone is familiar with the sequence of squares

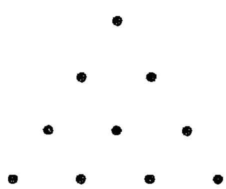
$$1, 4, 9, 16, 25, 36, 49, \ldots .$$

The ancient Greek mathematicians also considered the sequence

$$1, 1+2, 1+2+3, 1+2+3+4, \ldots,$$

or

$$1, 3, 6, 10, 15, 21, 28, 36, 45, \ldots$$



1
2
3
4

In view of the diagram, these were termed *triangular numbers;* the nth term of this sequence is $\frac{1}{2}n(n+1)$.

The question that was posed to pupils (and staff!) of Manor Farm School was:

Which numbers are both squares and triangular numbers?

We can see straight away that 1, 36 are the first two, and calculators produced 1225, 41616 as the next two; in fact,

$$1225 = 35^2 = \tfrac{1}{2} \times 49 \times 50,$$

$$41616 = 204^2 = \tfrac{1}{2} \times 288 \times 289.$$

But can we determine all numbers which are both square and triangular? What we need are natural numbers k, m such that

$$k^2 = \tfrac{1}{2}m(m+1) \qquad \text{or} \qquad m^2 + m - 2k^2 = 0.$$

This is a quadratic equation in m, with solution

$$m = \tfrac{1}{2}(-1 \pm \sqrt{(1 + 8k^2)}), \tag{1}$$

but of course not every value of k will make m an integer in formula (1). If $1 + 8k^2$ is a square, then it will be odd and the right-hand side of (1) will be an integer. Conversely, if (1) gives m as an integer, then $1 + 8k^2$ must be a square; in fact

$$1 + 8k^2 = (2m + 1)^2.$$

Thus we are looking for those natural numbers (i.e. positive integers) k such that $1 + 8k^2$ is a square, say l^2. This amounts to having to solve the equation

$$l^2 - 8k^2 = 1 \tag{2}$$

in natural numbers. We note that $l = 3$, $k = 1$ is a solution, and this gives the first square triangular number 1; also that $l = 17$, $k = 6$ is another solution, which gives the next square triangular number $6^2 = 36$. The third square triangular number comes from the solution $l = 99$, $k = 35$, the fourth from $l = 577$, $k = 204$. The question is: what are *all* the solutions of equation (2) in natural numbers?

2. Pell's equation

The equation

$$x^2 - dy^2 = 1,$$

where d is a positive integer, is very famous in mathematics, and is commonly called *Pell's equation*. The name is actually one of the great mistakes in mathematics, because it seems certain that the English mathematician John Pell (1611–1685) had little or nothing to do with the equation; it should more correctly be called *Fermat's equation*. The great Swiss mathematician Euler wrongly attributed work on the solution of the equation to Pell.

Pell's equation has facinated mathematicians throughout the ages. Archimedes posed a problem called 'the cattle problem' to Alexandrian scholars, the solution of which depended on solving a particular Pell's equation (see reference 1, Problem 1). The problem was concerned with the composition of a herd of cattle on the island of Sicily. The solution was of the order of 79 billion cattle, which it would be impossible to fit on to the island! In the 7th Century AD, an Indian mathematician, Brahmagupta, said that a person who can within a year solve the equation $x^2 - 92y^2 = 1$ is a mathematician!

Fermat was probably the first mathematician to show that Pell's equation has infinitely many solutions when d is not a perfect square. (It is clear that it has no solutions when d is a perfect square, because 1 is not the difference of the squares of two natural numbers.) In 1657 Fermat, in an attempt to interest mathematical colleagues in his theory of numbers, challenged the English mathematicians of his day to find a solution of the equation, if not for a general value of d then for certain specified values (he specified $d = 149, 109, 433$ as a test). The difficulty can be judged from the fact that the smallest values of y which solve the equation in these three cases have, respectively, 10, 14 and 29 digits in their decimal expansions! The English mathematician John Wallis published a method of solution which he attributed to William, Viscount Brouncker. Wallis was President of the Royal Society at the time, and it is thought that the method was really due to Wallis himself, who was trying to curry favour with his patron in crediting him with the solution. The method had the disadvantage that it was not guaranteed to work! Interestingly, the Hindu mathematician Bhascara Acharya, born way back in the year 1114, invented the 'cyclic method' of solution, which had the same disadvantage as the English method. The problem was picked up by Euler and Lagrange, and it was Lagrange who in 1768 published the first rigorous proof which obtained all the solutions of the equation, using continued fractions. Pell's equation has a fascinating history, fascinatingly told, for example, in reference 2.

3. A solution to our problem

The particular form of Pell's equation that we are attempting to solve is

$$l^2 - 8k^2 = 1$$

(see (2)), and we need *all* the solutions in natural numbers. Suppose that the natural numbers l, k provide a solution. Then

$$(l + 2\sqrt{2}k)(l - 2\sqrt{2}k) = 1,$$

so

$$(l + 2\sqrt{2}k)^n(l - 2\sqrt{2}k)^n = 1 \tag{3}$$

for all natural numbers n. Put

$$(l + 2\sqrt{2}k)^n = l_n + 2\sqrt{2}k_n$$

for $n = 1, 2, 3, \ldots$. Then l_n, k_n are natural numbers. Also,

$$(l - 2\sqrt{2}k)^n = l_n - 2\sqrt{2}k_n,$$

and we deduce from (3) that

$$l_n - 8k_n^2 = 1.$$

Thus l_n, k_n provide a solution for all values of n. This provides a way of generating infinitely many solutions from a single solution.

But suppose we begin with the solution with smallest values of l, k, in our case $l = 3$, $k = 1$. It is not a difficult matter to show that the values l_n, k_n for $n = 1, 2, 3, \ldots$ give *all* solutions of the equation (see reference 2, p. 339). This means that the squares which are also triangular numbers are precisely $k_1^2, k_2^2, k_3^2, \ldots$, where

$$l_n + 2\sqrt{2}k_n = (3 + 2\sqrt{2})^n.$$

To see how this works, we have $k_1 = 1$. Next,

$$l_2 + 2\sqrt{2}k_2 = (3 + 2\sqrt{2})^2 = 17 + 12\sqrt{2},$$

so $k_2 = 6$;

$$l_3 + 2\sqrt{2}k_3 = (3 + 2\sqrt{2})^3 = (17 + 12\sqrt{2})(3 + 2\sqrt{2}) = 99 + 70\sqrt{2}$$

and $k_3 = 35$;

$$l_4 + 2\sqrt{2}k_4 = (3 + 2\sqrt{2})^4 = (99 + 70\sqrt{2})(3 + 2\sqrt{2})$$

and $k_4 = 204$; and so on.

References

1. H. Dörrie, 100 *Great Problems of Elementary Mathematics* (Dover, New York, 1965).
2. D. M. Burton, *Elementary Number Theory* (Allyn and Bacon, London, 1980).

Explanation of the Paradox on p. 76

The fallacy is that sin y with y in degrees is a much different function from sin x with x in radians. Let $\sin x$ be the ordinary sine function for x in radians. Then the sine function for y in degrees is $\sin y/\alpha$ which we denote by $S(y)$. Then $(d/dy)S(y) = d/dy \sin y/\alpha = \cos y/\alpha(1/\alpha) = C(y)/\alpha$, where $C(y) = \cos y/\alpha$ is the cosine function for y in degrees. Carrying these clarifications through the problem eliminates the contradiction.

This problem arose from a calculus problem on related rates which gave the rate of change of angle as 5 degrees per minute. I think it clearly shows the dangers of writing $\sin x$ and $\sin y$ indiscriminately when x is in radians and y is in degrees.

I do not find this fallacy in any of the references below.

References

1. V. M. Bradis, V. L. Minkovskii and A. K. Kharcheva, *Lapses in Mathematical Reasoning* (Pergamon, Oxford, 1963).
2. W. Leitzmann, *Wo steckt der Fehler?*, 3rd edn. (Teubner, Stuttgart, 1953).
3. E. A. Maxwell, *Fallacies in Mathematics* (Cambridge University Press, London, 1963).

Polyspermy: An Example of Mathematical Modelling in Biology

BYRON J. T. MORGAN
University of Kent

Byron Morgan was an undergraduate at Imperial College, London, and a postgraduate at Churchill College, Cambridge. He is now a Senior Lecturer in statistics at the University of Kent, where he has worked for eight years; prior to this he was statistician to the Applied Psychology Unit at Cambridge. His research interests lie in applications of probability and statistics, mainly in biology and psychology.

1. Introduction

Polyspermy, which occurs when an egg may be fertilised by more than one sperm, has been examined in sea-urchins by Rothschild and Swann (reference 6) and more recently by Presley and Baker (reference 5), using modern biological techniques. Morgan (reference 2) presented a set of data resulting from experiments by Presley and Baker (see Table 1).

TABLE 1

	No. of eggs in experiment	Length of experiment (in seconds)	Number of sperm in the egg					
			0	1	2	3	4	5
Experiment 1	100	5	89	11	0	0	0	0
Experiment 2	84	15	42	36	6	0	0	0
Experiment 3	80	40	28	44	7	1	0	0
Experiment 4	100	180	2	81	15	1	1	0

Thus, for example, in the second of the four experiments, 84 eggs were exposed to sperm at time 0, and after 15 seconds fertilisation was stopped by the addition of a suitable chemical and it was found that 42 eggs had not been fertilised, 36 eggs had been fertilised once, and 6 eggs had been fertilised twice. The eggs in each experiment were different, and all four experiments took place under uniform conditions of temperature and sperm density.

The simplest probability model for data of this kind is the *Poisson process* (see, e.g. reference 1, p. 274). Whatever the number of sperm in any egg, further sperm enter as a Poisson process at a constant rate, λ, say. In this article we consider three probability models for polyspermy and fit each of these models to the above set of data. The first model is the simple Poisson process; the second and third models are successive modifications of the Poisson process, the second model having also been presented in reference 2.

2. The Poisson process

The prediction of the Poisson process is that at any time t, the probability $p_k(t)$ that an egg has been fertilised k times in the interval $(0, t)$, is given by the *Poisson distribution*,

$$p_k(t) = \frac{e^{-\lambda t}(\lambda t)^k}{k!} \qquad (k = 0, 1, 2, \ldots). \tag{1}$$

Two questions now arise.

(i) If we assume that fertilisation can be described by a Poisson process, what value of λ, the fertilisation rate, should we take to describe this particular set of data?

(ii) Once that λ is chosen, is the description of the data provided by the model a reasonable one?

We answer (i) by using the method of *maximum likelihood:* under the model, the likelihood of the data, or the probability that the events represented by the data occur, can be written as

$$L = \prod_{i=1}^{4} \prod_{k=0}^{\infty} p_k(t_i)^{n_k(t_i)} \tag{2}$$

in which $t_1 = 5, t_2 = 15, t_3 = 40$ and $t_4 = 180, n_0(5) = 89, n_1(5) = 11, n_0(15) = 42, n_1(15) = 36, n_2(15) = 6$, and so on. In writing down this likelihood we have assumed that the fertilisations of different eggs, between and within experiments, are all independent, which does not seem unreasonable. L is a function of λ, and the method of maximum likelihood, invented by Sir Ronald Fisher, estimates λ by the value $\hat{\lambda}$ of this parameter that maximises the likelihood L. It is simpler (and equivalent, as $\log(x)$ is a monotonic function of x) to maximise $\log L$:

$$\begin{aligned}\log L &= \sum_{i=1}^{4} \sum_{k=0}^{\infty} n_k(t_i) \log p_k(t_i) \\ &= C - \lambda \sum_{i=1}^{4} t_i \sum_{k=0}^{\infty} n_k(t_i) + (\log \lambda) \sum_{i=1}^{4} \sum_{k=0}^{\infty} k n_k(t_i),\end{aligned} \tag{3}$$

in which the constant term C does not involve λ. In search of stationary values of $\log L$ we set $d(\log L)/d\lambda = 0$, to obtain the maximum likelihood estimator

$$\hat{\lambda} = \sum_{i=1}^{4} \sum_{k=0}^{\infty} k n_k(t_i) \Big/ \sum_{i=1}^{4} t_i \sum_{k=0}^{\infty} n_k(t_i), \tag{4}$$

of the fertilisation rate λ. It is simple to verify that $\hat{\lambda}$ does in fact maximise $\log L$, and that for our data $\hat{\lambda} = 0{\cdot}0104$.

For an experiment at time t_i, involving $N(t_i) = \sum_{j=0}^{\infty} n_j(t_i)$ eggs altogether, the expected number of eggs with k fertilisations is given by

$$N(t_i) \frac{e^{-\hat{\lambda} t_i}(\hat{\lambda} t_i)^k}{k!} \qquad (k = 0, 1, 2, \ldots). \tag{5}$$

This gives, for the Poisson process model, the expected numbers shown in Table 2.

TABLE 2

	No. of eggs in experiment	Length of experiment (in seconds)	Number of sperm in the egg						
			0	1	2	3	4	5	6
Experiment 1	100	5	95	5	0	0	0	0	0
Experiment 2	84	15	72	11	0	0	0	0	0
Experiment 3	80	40	53	22	5	1	0	0	0
Experiment 4	100	180	15	29	27	17	8	3	1

(The expected numbers are rounded to the nearest integer, which explains the small round-off error present.)

In answer to Question (ii) above about goodness of fit, the model very clearly does not provide an adequate description of the data. On the one hand λ should be larger than 0·0104 in order to reduce the expected number of eggs that have still to be fertilised; on the other hand that would, particularly for the last experiment, result in expected values of $n_k(t_i)$, for $k > 1$, which are far too large.

A possible explanation is that once an egg has been fertilised for the first time then subsequent fertilisations take place at some *reduced* rate, μ, say. This is, in fact, what the biologists believe to take place; the form for the $p_k(t)$ under this new model was given in reference 2, and an alternative derivation is presented in the following section.

3. Poisson process with two rates

The Poisson process has two important properties. The first, that the number of events (fertilisations in our case) in a fixed time interval has a particular Poisson distribution, was given in (1), and used above. The second equivalent property, which we shall use shortly, is that the time intervals between events are continuous random variables with an exponential probability density function:

$$\beta e^{-\beta t}, \qquad t > 0, \tag{6}$$

for a process of rate $\beta > 0$.

Returning now to the model, clearly $p_0(t) = e^{-\lambda t}$, precisely as before. However if the first fertilisation takes place at some time θ, $0 \leq \theta \leq t$, then, conditional upon this event,

$$\begin{aligned} p_n(t) &= \Pr(n-1 \text{ events occur in the remaining time interval, } (t-\theta)) \\ &= \frac{e^{-\mu(t-\theta)}(\mu(t-\theta))^{n-1}}{(n-1)!} \qquad (n \geq 1), \end{aligned}$$

where the fertilisation rate is now μ. But of course θ is itself a random variable, with density function given by (6), when $\beta = \lambda$. To obtain the unconditional form for $p_n(t)$ we therefore average the above expression with respect to this density, to obtain the solution

$$\begin{aligned} p_n(t) &= \int_0^t \frac{e^{-\mu(t-\theta)}(\mu(t-\theta))^{n-1}}{(n-1)!} \lambda e^{-\lambda\theta}\, d\theta \\ &= \frac{\lambda\mu^{n-1}e^{-\mu t}}{(n-1)!}\int_0^t e^{(\mu-\lambda)\theta}(t-\theta)^{n-1}\, d\theta \qquad (n \geq 1). \end{aligned} \tag{7}$$

The problem now is to choose λ and μ to fit this model to the data. Once again the method of maximum likelihood may be used, but now explicit expressions for $\hat{\lambda}$ and $\hat{\mu}$ are difficult to obtain. It is a simple matter, however, to draw contours of the log-likelihood surface, (the integral can be written as a finite sum, to aid computation) revealing a maximum at $\hat{\lambda} = 0{\cdot}0287$, and $\hat{\mu} = 0{\cdot}0022$. (Here, and later, values of the parameters were also checked by using the numerical maximisation procedure of Nelder and Mead (reference 4), a convenient algorithm for which is provided in reference 3.)

As anticipated, $\hat{\lambda}$ has now increased, in fact to three times its previous value, and $\hat{\mu}$ is much smaller (less than 1/10th of $\hat{\lambda}$). The (rounded) expected values for the Poisson process with two rates are shown in Table 3.

TABLE 3

	No. of eggs in experiment	Length of experiment (in seconds)	Expected number of sperm in the egg				
			0	1	2	3	4
Experiment 1	100	5	87	13	0	0	0
Experiment 2	84	15	56	28	0	0	0
Experiment 3	80	40	26	51	3	0	0
Experiment 4	100	180	1	72	23	4	0

This model provides a much better description of the data, though there remain obvious discrepancies. In reference 2 these discrepancies are reduced by allowing λ and μ to be functions of time. Further investigations of the underlying biology reveals, however, that while the biologists anticipate a reduction in fertilisation rate, from λ to μ, they also expect the secondary fertilisations (those at rate μ) to last only for some limited time, τ, say, after the time of the first fertilisation. A 'fertilisation membrane', detectable under the microscope, is then formed which bars all further fertilisations. In the next section we see how such a membrane may be built into the model considered above, and how the modification affects the parameter estimates and the fit of the model to the data.

4. A model with a fertilisation membrane

Although τ is itself likely to be a random variable, we shall simply assume here that it takes some fixed value, which is constant for all eggs. (There are biological difficulties in determining the time of the first fertilisation for any egg, and consequently, a precise value for τ.) Under these conditions, we have

$$p_0(t) = e^{-\lambda t},$$

as before.

If $t \le \tau$ then the values for $\{p_n(t)\}$ remain as in (7), since the membrane has not had time to operate. But if $t \ge \tau$ then one of two events may occur:

(i) Either the first fertilisation takes place before time $(t - \tau)$, which from (6) has the probability

$$\int_0^{t-\tau} \lambda e^{-\lambda x}\, dx = 1 - e^{-\lambda(t-\tau)},$$

and then $p_n(t)$ is just the probability of $(n-1)$ fertilisations in the time τ, or $e^{-\mu\tau}(\mu\tau)^{n-1}/(n-1)!$ for $n = 1, 2, \ldots$, by (1).

(ii) Alternatively, the first fertilisation takes place after time $(t - \tau)$, at some time θ, say, leaving just a time $(t - \theta)$ for further fertilisations, exactly as in the previous model.

Taking account of the time of the first fertilisation, we find:

$$p_n(t) = (1 - e^{-\lambda(t-\tau)})\frac{e^{-\mu\tau}(\mu\tau)^{n-1}}{(n-1)!} + \frac{\lambda\mu^{n-1}}{(n-1)!}e^{-\mu t}\int_{t-\tau}^{t} e^{(\mu-\lambda)\theta}(t-\theta)^{n-1}\, d\theta \qquad (n = 1, 2, \ldots). \tag{8}$$

A useful check is that if we set $\tau = t$, we return to the model of the previous section.

This model was fitted to the data for selected values of τ. Interestingly, τ had no effect on $\hat{\lambda}$, and little effect on the expected values; its effect on $\hat{\mu}$ was appreciable, however, as Table 4 indicates.

TABLE 4

τ (seconds)	$\hat{\lambda}$	$\hat{\mu}$
15	0·029	0·0140
40	0·029	0·0064

For $\tau = 15$ seconds, the (rounded) expected values for this final model with a fertilisation membrane are given in Table 5.

The fit to the data is clearly improved, though some discrepancies, mainly in the data of the second experiment, remain.

TABLE 5

	No. of eggs in experiment	Length of experiment (in seconds)	Expected number of sperm in the egg				
			0	1	2	3	4
Experiment 1	100	5	87	13	0	0	0
Experiment 2	84	15	54	27	3	0	0
Experiment 3	80	40	25	46	8	1	0
Experiment 4	100	180	1	81	17	2	0

5. Discussion

Further investigation of these and other models is currently taking place, both for additional experiments of the type considered here, and for other experiments in which nicotine, which is thought to increase μ, is added. For the standard experiments, biologists believe that the ratio of $\lambda:\mu$ is in the region of 10:1, as we found in Section 3. However, the results of Section 4 suggest that this ratio might well only be 4:1, or even 2:1. Further work will clarify the situation, but these examples show the power of a mathematical model in predicting and quantifying the consequences of a simple biological mechanism. The effect of a fertilisation membrane is to confine the secondary fertilisations to a limited time, with the obvious result that μ must be larger than if no such membrane were present.

References

1. G. M. Clarke and D. Cooke, *A Basic Course in Statistics* (Edward Arnold, London, 1978).
2. R. W. Morgan, Some stochastic models to describe the fertilization of an egg, *Applied Statistics* **24** (1975) 137–138.
3. N.A.G. Fortran Library Manual Mark 6, Volume 4. (N.A.G., Oxford, 1977).
4. J. A. Nelder and R. Mead, A simplex method for function minimization, *Computer Journal* 7 (1965) 308–313.
5. R. Presley and P. F. Baker, Kinetics of fertilization in the sea-urchins: a comparison of methods. *Journal of Experimental Biology* **52** (1970) 455–468.
6. Lord Rothschild and M. M. Swann, The fertilization reaction in the sea-urchin. The block to polyspermy. *Journal of Experimental Biology* **29** (1952) 469–483.

1 + 1 = 0?

It used to be fun
To add one and one
But now I'm unsure
What sum to secure
I'm told it may even be none.

Letters to the Editor

Dear Editor,

The problem of Buffon's needle

With reference to Gani's article on the Buffon needle problem, the case $l \leq s$ needs no calculus. Suppose the needle (length $2l$) to be bent round to form a circle of radius $r = l/\pi$. Because of the assumptions about uniformity of distributions of both position and orientation, this makes no difference to the chance of any elementary section of the needle intersecting a join. However, whenever one element intersects, another must now intersect, whereas with a straight needle one only can intersect.

The probability for the circular needle being obviously $2r/2s = l/\pi s$, the probability for the straight needle is $2l/\pi s$, since with a straight needle some one element intersects twice as often as some two intersect with the circular needle.

Yours sincerely,
M. T. L. BIZLEY
(c/o Bacon & Woodrow, 55 High Street, Epsom, Surrey KT21 2NW)

Dear Editor,

Desert island theorems†

I don't know whether Stirling's method for calculating $\pi^2/6$ is fair; it works well, and can be used without a calculator, especially on the sand.

We have

$$\frac{1}{x^2} = \frac{1}{x(x+1)}\left(\frac{x+1}{x}\right) = \frac{1}{x(x+1)}\left(1+\frac{1}{x}\right).$$

Repeating the process,

$$\frac{1}{x^2} = \frac{0!}{x(x+1)} + \frac{1!}{x(x+1)(x+2)} + \frac{2!}{x(x+1)(x+2)(x+3)} \text{etc.}$$

Hence

$$\sum_{y}^{\infty} \frac{1}{x^2} = \frac{0!}{1y} + \frac{1!}{2y(y+1)} + \frac{2!}{3y(y+1)(y+2)} \text{etc.},$$

using difference methods. This series converges rapidly, and may be *proved* to do so by Gauss's test.

Choose a suitable y, say 20, and calculate the first $(y-1)$ terms of $\pi^2/6$ ordinarily. I get 1·5936632 for the first 19 terms and 0·0512708 for the remainder, using 6 terms of the tail.

(Reference. E. T. Whittaker and G. Robinson, *The Calculus of Observations*, p. 368.)

Yours sincerely,
J. L. G. PINHEY
(The Perse School, Cambridge)

† See David W. Sharpe's article in Volume 12 No. 3.

Dear Editor,

The series swindle

Having received the latest *Spectrum* (Volume 13 Number 2) a few days ago and scanned through it, I want to tell you right away that I was appalled by the inclusion of the contribution by Professor Pym.

I am well aware of the need for an imaginative presentation of mathematical ideas and that their relevance and usefulness in real life needs to be brought out, and the use of a story to embed the mathematics in is a good idea. However I strongly object to the nature of the material used in this article, particularly in view of the fact that you are aiming to produce something for sixth formers to read. I would never be willing to give a pupil something to read that made light of drinking and looseness in regard to the opposite sex, and most of all I object to the frequent swearing and taking of the name of God in vain.

I was really appalled to see this in what is supposed to be suitable reading matter for young people at school. Only one thing deters me from an immediate decision to stop my subscription and have nothing more to do with *Spectrum* in any way: the fact that so far as I know nothing like this has appeared before. I hope it won't again, because if it does I will regretfully feel I can no longer support or receive *Spectrum* despite the usefulness and interesting nature of most of its contributions.

Yours sincerely,

C. W. PURITZ
(Royal Grammar School, High Wycombe, Buckinghamshire HP13 6QT)

Problems and Solutions

Sixth formers and students are invited to submit solutions to some or all of the problems below: the most attractive solutions will be published in subsequent issues. When writing to the Editorial Office, please state your full name and home address and also the postal address of your school, college or university.

Problems

13.7. (See the article by A. K. Austin in this issue.) Construct a Turing machine which includes | and $\times$ among its tape symbols and which performs multiplication, i.e. if it is started with the string

$$\underbrace{|\,|\,|\cdots|\,|}_{\leftarrow n \rightarrow} \times \underbrace{|\,|\,|\cdots|\,|}_{\leftarrow m \rightarrow}$$

then, after operating for some time, it will eventually settle permanently into a situation where the string on the tape is

$$\underbrace{|\,|\,|\,|\cdots|\,|\,|}_{\leftarrow n \times m \rightarrow}$$

13.8. A confectioner's shop sells three sorts of fruit pie, apple, blackberry and cherry, and their quality is measured on a scale 1–6 (1 = low, 6 = high). The probabilities of the quality of each pie are given by

Pie	Quality	Probability
Apple	3	1
Blackberry	1	0·51
	5	0·49
Cherry	2	0·56
	4	0·22
	6	0·22

If you do not like cherries, which pie should you choose to maximize your probability of obtaining the best quality? If you like apple, blackberry and cherry pie equally, which should you now choose?

13.9. At a party consisting of at least four people, among any four guests there is one who has previously met the other three. Prove that, among any four guests, there is one who has previously met every other person at the party.

Solutions to Problems in Volume 13, Number 1

13.1. Let n be a natural number. Show that every set of n natural numbers contains a non-empty subset the sum of whose elements is divisible by n.

Solution

This is a nice instance of the use of the Pigeonhole Principle, which is surprisingly effective in often unlikely situations. Stated in its popular form, if n letters have to be put into fewer than n pigeonholes, then there must be two letters which end up in the same pigeonhole. Consider n natural numbers $a_1, \ldots, a_n$, and put $s_i = a_1 + \cdots + a_i$ for $1 \leq i \leq n$. If one of the s_i is divisible by n, fine. If not, then their remainders on division by n, of which there are n, are all among the integers $1, \ldots, n-1$, of which there are fewer than n. Lo and behold, the Pigeonhole Principle says that two of these remainders must be the same! Thus two of the s_i, say s_i and s_j $(j > i)$, have the same remainders on division by n. But now

$$s_j - s_i = a_{i+1} + \cdots + a_j$$

is divisible by n, and the problem is solved.

13.2. How may 17 straight lines be drawn in the plane so that they meet in exactly 101 points?

Solution

This is a famous problem which was first posed by the distinguished mathematician Joseph Fourier (see R. J. Webster's article in Volume 13 Number 1). The problem is something of a cheat, because it seems to require a geometrical solution, yet it turns out not to be a geometrical problem at all, at least in essence. Suppose we draw n_1 distinct lines in the plane, all parallel to one another, then n_2 parallel lines not parallel to the first n_1, then n_3 parallel lines not parallel to the first n_1+n_2 etc. until we finish with n_k parallel lines not parallel to the previous $n_1+\cdots+n_{k-1}$. It is certainly possible to do this so that no three lines meet in a point. After all, a plane is a big place! This is all the geometry that comes in. We require

$$n_1+\cdots+n_k=17.$$

We shall now count the number of points of intersection of these lines. Each one of the first n_1 lines meets $17-n_1$ other lines, so there are $n_1(17-n_1)$ points of intersection on these n_1 lines. The same is true of the other lines. But obviously each point of intersection lies on two lines, so we require

$$n_1(17-n_1)+\cdots+n_k(17-n_k)=202$$

to obtain 101 points of intersection. Thus we have to choose $n_1,\ldots,n_k$ to satisfy the two equations

$$n_1+\cdots+n_k=17,$$
$$n_1^2+\cdots+n_k^2=87.$$

This is easily done, e.g.

$$k=4,\quad n_1=2,\quad n_2=3,\quad n_3=5,\quad n_4=7$$
$$k=4,\quad n_1=1,\quad n_2=5,\quad n_3=5,\quad n_4=6$$
$$k=6,\quad n_1=n_2=n_3=1,\quad n_4=2,\quad n_5=4,\quad n_6=8.$$

This gives three solutions to the problem.

13.3. Solve the matrix equation $X=Y+AXB$ for X when $A^k=0$ for some positive integer k.

Solution

Sorry about the misprint in this question, where the second X in the equation came out as a multiplication sign. That made the problem rather easy to solve! The main things to beware of when solving matrix equations are that you cannot change the order of terms in a matrix product, you cannot cancel and you cannot divide. $A^k=0$ certainly looks strange, but it will tell us that

$$A^{k-1}X=A^{k-1}Y+A^kXB=A^{k-1}Y$$

when we multiply both sides of the given equation (on the left!) by A^{k-1}. If $k>1$, we then multiply both sides this time by A^{k-2} to give

$$\begin{aligned}A^{k-2}X&=A^{k-2}Y+A^{k-1}XB\\&=A^{k-2}Y+A^{k-1}YB.\end{aligned}$$

If $k>2$, we multiply both sides by A^{k-3} to give

$$\begin{aligned}A^{k-3}X&=A^{k-3}Y+A^{k-2}XB\\&=A^{k-3}Y+A^{k-2}YB+A^{k-1}YB^2.\end{aligned}$$

It should be clear how to continue, to give the final solution

$$X=Y+AYB+A^2YB^2+\cdots+A^{k-1}YB^{k-1}.$$

Finite Mathematics for Business, Social Sciences and Liberal Arts. By LOUIS M. ROTANDO. Van Nostrand Reinhold Company Ltd., Wokingham, 1980. Pp. viii + 519. £12·70.

The distinctive feature of this book is the choice of its mathematical subject matter. The title *Finite Mathematics* is presumably chosen to imply the exclusion of all calculus topics, that is anything involving limit concepts, and this principle brings together such diverse topics as logic, probability, linear equations and linear programming, and mappings. Not surprisingly, this wide spectrum of material has found at least one of the present reviewer's blind-spots: I can say, as a guinea pig, that I found the chapters dealing with formal logic an interesting introduction to a subject I have wanted to read about, and to which I should hope to return for a second reading. The remaining topics are well covered and pleasantly readable. An interesting feature is a detailed study of several gambling games as applications of probability. Many textbooks cite such games as examples, but frequently the treatment is very shallow, with neither a proper explanation of the game, nor a full exploration of its possibilities. Not so here, where full details are given. The book follows modern forms of presentation with lavish provision of spacing and diagrams, marginal subheadings, and the use of background shading for definitions. This makes for clarity, although at times appears a little gimmicky. Certainly the author seems very conscious of the need to hold the interest of his readers; he expects these to be students of 'business, social science, economics, or the liberal arts'. The level of the material, however, is well within the grasp of a sixth-former who is prepared to work through it with pen and paper. It would serve to broaden interest in many directions, since no syllabus will cover all these topics. For this purpose it could make a good addition to a school library. For the scientific or mathematical specialist in schools I doubt that it has any lasting value, as the depth of treatment is not sufficiently great.

Royal Grammar School, Newcastle upon Tyne P. MITCHELL

The Analysis of Time Series: An Introduction (2nd edn). By C. CHATFIELD. Chapman and Hall, London, 1980. Pp. 268. £6·50.

In its simplest form a time series is a sequence of observations on some quantity of interest, recorded at equally spaced time intervals. Such series are often seen in the financial columns of 'quality' newspapers, referring to such quantities as daily share prices, monthly unemployment figures, annual company profits. Dr Chatfield's book provides a very readable introduction to the statistical analysis of time series data. Although the book has been written primarily as a textbook for university students, much of it should be accessible to anybody who has studied A-level mathematics together with some elementary probability and statistics.

The book first discusses simple descriptive techniques, which include plotting the data (always a good idea!) and 'smoothing' the plot in such a way as to highlight important features which may otherwise be obscured or distorted by random variation. For example, when the Chancellor of the Exchequer quotes a figure for the current rate of inflation, he is (or should be) calculating an average increase in prices over the preceding twelve-month period, rather than multiplying by twelve the increase in the latest one-month period.

Dr Chatfield then describes various types of probability models for time series data and shows how these can be used to forecast future values of the series in question. The author is rightly suspicious of automatic forecasting schemes, and includes several examples to show how a purely statistical approach to forecasting should be tempered with common sense and expert knowledge of the particular context in which forecasts are required.

Other topics covered include spectral analysis, which seeks to identify important periodic effects in a time series, and bivariate methods, which involve the inter-relationships between two parallel series.

For a first reading, I should recommend Chapters 1 to 5 and 10, which require a working knowledge of series summation, difference equations and the concept of covariance. For the remaining chapters, some knowledge of complex numbers and of Fourier analysis is also needed.

University of Newcastle upon Tyne PETER J. DIGGLE

Calculus. By JERROLD MARSDEN and ALAN WEINSTEIN. The Benjamin/Cummings Publishing Company, Inc., Menlo Park, California, 1980. Pp. xxvii + 837 + 168 (appendix). £12·95.

Calculus with Analytic Geometry. By JOHN B. FRALEIGH. Addison–Wesley, London, 1980. Pp. xiii + 866. £9·95.

Both these attractively produced large American books give a treatment of calculus from the very beginning up to partial derivatives and multiple integration. Both the books are written in a helpful style for self study, both assume calculators are available, both have answers to odd-numbered questions only, both have instructors' guides available and both have a students' supplement. Either book would make a splendid addition to a school library, and give ideal material for able students to browse through when they finish the immediate task on hand.

University of Durham H. NEILL

The Sunday Times Book of Brain-Teasers (Book 1). By VICTOR BRYANT and RONALD POSTILL. Unwin Paperbacks, London, 1980. Pp. 160. £1·95.

This is the first of two books which include brain-teasers largely chosen from those published in the *Sunday Times* from January 1974 to December 1979. Fifty problems are provided in the first half of the book and solutions and answers make up the second half. Problems are classified in sections including clock problems, prime numbers, codes, football league tables and many others, some totally devoid of numbers but requiring logical processing of given information. Most questions provide a mass of interwoven information so that a pencil, paper and a good deal of undivided attention are needed to make any progress.

To those readers who look for the quick, elegant proof, this is not the book for you. Each situation could be likened to the knotted string which must be unravelled. There is no end to pull which makes all the knots miraculously drop out, it is more a case of finding the best knot to pick at first and then moving along the line from there. Some degree of trial and error is required in a number of the solutions.

One advantage of such problems is that they can be tackled by people of all ages and backgrounds blessed with sufficient determination, but if you are looking for problems with some sort of algebraic solution try the questions involving clocks or circles or distance, speed and time. These require 'O'-level type algebra but are generally more taxing than the usual 'O'-level problems, so while the book would be a useful acquisition for any school library it should be recommended only to the more able problem solvers.

Mathematics Department, St. Aidan's Comprehensive School, Sunderland T. D. ROSS

Adventures with your Pocket Calculator. By LENNART RÅDE and BURT A. KAUFMANN. Pelican Books, London, 1980. Pp. 139. £0·95 (paperback).

This book was first reviewed in *Mathematical Spectrum* in Volume 10, page 99. It has proved difficult to obtain in Britain, but is now readily available from Pelican Books.

We reproduce our original review below:

This is a superb book which is aimed at a readership of a slightly lower age than the average reader of *Mathematical Spectrum*, but it would have interest for those readers who own a programmable calculator.

The book consists of twenty 'adventures' which are activities designed to interest and stimulate at the eleven to sixteen age range. It does this very well, having just the right amount of open-ended work to be stimulating rather than daunting, and would be excellent to give to the pupil who finishes first and who has an odd ten minutes or quarter of an hour to fill at the end of a lesson. Different adventures will appeal to different readers; I enjoyed those based on a simple random number generator and my son and I now play backgammon using a programmable calculator instead of dice.

The interest for older readers will probably come in the challenge of fitting some of the exercises onto a programmable calculator. For instance, which integers between 0 and 100 have a perfect square all of whose digits are odd? And having got the answer, prove it!

University of Durham H. NEILL

The Penguin Book of Mathematical and Statistical Tables. By R. D. NELSON. Penguin Books Ltd., London, 1980. Pp. 63. £0·95 (paperback).

In the calculator age, it is difficult to decide what to include in a new set of tables. If the functions which appear on a scientific calculator are included, they seem superfluous; on the other hand their omission can also be criticised. This set of tables includes all the standard functions, but the main use is in the statistical tables which do seem to live up to the author's claim that they are sufficiently comprehensive for the teaching of mathematics and statistics in schools, polytechnics and universities, as well as being clearly and helpfully laid out.

University of Durham H. NEILL

International Bibliography of Journals in Mathematical Education (1980), obtainable from Universität Bielefeld, Institut für Didaktik der Mathematik, Universitätsstrasse, 4800 Bielefeld 1, West Germany.

This is a comprehensive international list of 253 magazines such as *Mathematical Spectrum* which come under the broad title of journals of mathematical education. Enough for most of our readers to know that we are not alone!

Contents

ISSN 0025–5653

PRICES (*postage included*)

Prices for Volume 13 (*Issues Nos.* 1, 2, *and* 3):

Subscribers in Britain and Europe: £2·00
Subscribers overseas: £4·00 (US\$9.00; \$A8.00)
(These prices apply even if the order is placed by an agent in Britain.)
A discount of 10% is allowed on all orders for five or more copies.

Back issues:

All back issues are available; information concerning prices may be obtained from the Editor.

Enquiries about rates, subscriptions and advertisements should be directed to:

Editor—*Mathematical Spectrum*,
Hicks Building,
The University,
Sheffield S3 7RH, England

Printed in England by Galliard (Printers) Ltd, Great Yarmouth