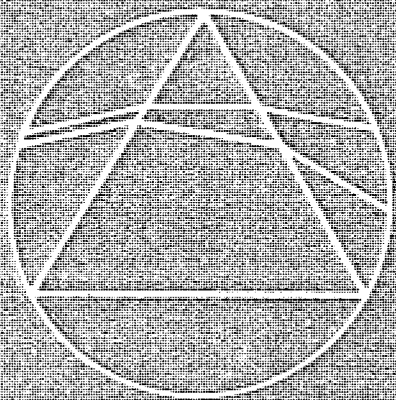
Mathematical Spectrum



Volume 2 1969/70 *Number 1*

A Magazine of *Contemporary Mathematics*

Published by *Oxford University Press*

Mathematical Spectrum is a magazine for the instruction and entertainment of student mathematicians in schools, colleges, and universities. It is published by Oxford University Press on behalf of the Applied Probability Trust, a non-profit making organization 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 2 of *Mathematical Spectrum* (1969/70) will consist of two issues, the second of which will be published in the Spring of 1970.

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; however, the Editorial Committee 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 Managing Editor, *Mathematical Spectrum*,
Hicks Building, The University, Sheffield S3 7RH.

Mathematical Competitions

MARGARET HAYMAN

Mayfield School, Putney, London

The first mathematical competitions were held in 1894 by the Mathematical and Physical Society of Hungary in honour of its founder, the physicist Loránd Eötvös. Although not all winners of the competition became mathematicians, many of the best known Hungarian mathematicians have been prize winners. Consequently the idea of national mathematical competitions has spread. Competitions have been held in Russia from 1935 onwards under the direction of the Moscow Mathematical Olympiads and particularly Professor D. O. Schklarski; and, since 1949, in Poland, organised by the Polish Mathematical Society, under the direction of the Ministry of Education. Other East and West European countries, including Sweden, have followed suit, but the idea of competitions seems less popular in Western Europe. The Mathematical Association of America began by organising a competition in 1950 for Metropolitan New York and extended this, with the cooperation of the Society of Actuaries, to the whole of America in 1957. The National Mathematical Contest was started in this country in 1964 by Mr F. R. Watson, now of Keele University. More recently Israel and Tanzania have organised competitions and there is a move to organise something in South-East Asia, and perhaps in Trinidad.

The aim of these competitions is not primarily to find a winner, although in East Europe the winner does have awards comparable to our university scholarships, but to stimulate interest in mathematics both among teachers and among pupils of varied mathematical ability. The syllabus for all these competitions is within the range of 'high school'. This means 10th to 12th grade in America, the syllabus for O-level additional mathematics in England, and the last two grades in schools in East Europe.

The form of the competitions varies from country to country. In East Europe the competition consists of three stages. The first stage lasts throughout October, November, and December. At the beginning of each month the central committee sends out to all secondary schools in the country a set of problems which the pupils are expected to solve individually, working at home, within the month in question. There is no check to see if the pupils' work is entirely their own. The school sends the solutions to a regional committee, and the pupils with the best solutions are admitted to the second stage. This takes place in spring, the pupils having to travel to a regional centre. There they have to do a written examination under supervision, and the best pupils are allowed to enter the final contest, held centrally. In each of the second and third stages candidates are given four hours to solve three or four problems which require fairly searching investigations. The examiners are looking particularly for elegance of solution.

In America the competition consists of a single 'multiple choice' paper which is set by the Mathematical Association of America and the Society of Actuaries (MAASA) and is marked in school. The competition is taken by the top 5% of the pupils in the 11th and 12th grades in American high schools, about 150,000 altogether. The questions are graded in difficulty in three sections, and marks are gained for correct answers and lost for incorrect ones. Owing to comments on the increasing difficulty of the papers, in 1968 the paper was arranged in four sections, with the first section considerably easier than the others. There are 40 questions to answer in 80 minutes, so speed rather than elegance is the major factor in this competition. Prizes (only nominal in value) are given for the best schools and for the best individual in each area.

In England we compromise between these two systems. Mr Watson asked the organisers of the MAASA competition in America to send some papers for young people in England and in 1963 about 200 pupils took one of these. This was repeated in 1964 and in that year my husband and I suggested that those who did well in the competition should be invited to sit a second paper set by us. In 1965 therefore the top 60 candidates in the National Mathematics Contest (the British branch of MAASA) took the first British Olympiad paper. There was a three hour paper consisisting of ten rather hard questions on the fifth and lower sixth form syllabus. This was a deliberate policy to avoid overlap with the Oxbridge scholarship examinations; and for the first two years there was also an upper age limit of seventeen for Olympiad competitors. However this was dropped since it was found that the best candidates were usually young. The number of candidates entering the National Mathematics Contest has risen rapidly since 1965 to 10,000 in 1968, and a wider range of schools—comprehensive and secondary modern as well as grammar and public—is now represented. As in America, the MAASA paper is marked in school and the names are forwarded to a central office which picks the high scorers for the Olympiad competition. Up to now this has been set and marked by my husband and myself, but in 1969 Dr B. Thwaites will take over from my husband. The competitions in England have been sponsored by the Guinness Science Awards who give secretarial and financial help and arrange a ceremony at which certificates are awarded to all those who qualify for the Olympiad competition and the remainder of the 100 best in the National Mathematics Contest. The best six in the Olympiad also receive a small prize. In 1968 a team of six boys from Upper New York State came to London to join in the Olympiad competition.

About ten years ago the East European countries combined for an International Olympiad which is held in a different country each year; and in 1966, as a result of conversations I had during the International Congress of Mathematicians in Moscow, Britain was invited to join the competition. In 1967, therefore, the best eight candidates in the British Olympiad went to Yugoslavia to compete with

France, Italy, Sweden, and nine 'East European' countries, an umbrella term which includes Mongolia. The International Olympiad is primarily an individual event, but the British team came fourth and one member won a gold medal and a prize for elegance for one solution. After much striving for finance and some controversy with the Department of Education and Science a similar team went to Moscow in 1968 and again came fourth with three gold medals and a prize for elegance. Considering the deliberate 'amateur' status of our competitors compared with the careful training by correspondence and Saturday morning classes of the East European candidates, we have reason to be satisfied with the standard achieved in mathematics in Britain. So far the international team has come from rather few schools, but we hope that this may change as the competitions become better known and more schools enter the National Mathematics Contest as a gateway to the international event. However, winning is less important than the interest that these competitions supply to school lessons and the new topics that they introduce.

To provide some idea of the relative standards of the various competitions a few questions from each are given below.

I. From the easy section of a MAASA test.

1. $\sqrt{\left(\frac{4}{3}\right)}-\sqrt{\left(\frac{3}{4}\right)}$ is

(A) $\frac{\sqrt{3}}{6}$, (B) $\frac{-\sqrt{3}}{6}$, (C) $\frac{\sqrt{(-3)}}{6}$, (D) $\frac{5\sqrt{3}}{6}$, (E) 1.

2. A square and an equilateral triangle have equal perimeters. The area of the triangle is $9\sqrt{3}$ square inches. The diagonal of the square in inches is

(A) $\frac{9}{2}$, (B) $2\sqrt{5}$, (C) $4\sqrt{2}$, (D) $\frac{9\sqrt{2}}{2}$, (E) none of these.

II. From the hard section of a MAASA test.

1. The limiting sum of the infinite series

$$\frac{1}{10}+\frac{2}{10^2}+\frac{3}{10^3}+\ldots+\frac{n}{10^n}+\ldots$$

is

(A) $\frac{1}{9}$, (B) $\frac{10}{81}$, (C) $\frac{1}{8}$, (D) $\frac{17}{72}$, (E) greater than any finite number.

2. If, in the triangle ABC, E is on AB so that $AE:EB=1:3$, D is on BC so that $CD:DB=1:2$ and F is the intersection of AD, CE, then $(EF/FC)+(AF/FD)$ is

(A) $\frac{4}{5}$, (B) $\frac{5}{4}$, (C) $\frac{3}{2}$, (D) 2, (E) $\frac{5}{2}$.

III. From the British Olympiad competition.

1. The faces of a tetrahedron are formed by four congruent triangles. If α is the angle between a pair of opposite edges of the tetrahedron, show that

$$\cos\alpha = \frac{\sin(B-C)}{\sin(B+C)},$$

where B, C are the angles adjacent to one of these edges in a face of the tetrahedron.

2. The sum of the reciprocals of a set of n different positive integers is equal to 1. If $n = 3$, show that there is only one such set and find it. Find also a set for $n = 4$, $n = 5$ and, in general, any $n > 3$.

IV. From the International Olympiad competition, 1968.

1. Consider the sequence (c_n) defined by

$$\begin{aligned} c_1 &= a_1+a_2+\ldots+a_8, \\ c_2 &= a_1^2+a_2^2+\ldots+a_8^2, \\ &\ldots \quad \ldots \quad \ldots \quad \ldots \\ c_n &= a_1^n+a_2^n+\ldots+a_8^n, \\ &\ldots \quad \ldots \quad \ldots \quad \ldots \end{aligned}$$

where $a_1, a_2, \ldots, a_8$ are real numbers, not all equal to zero. Being given that among the numbers of the sequence (c_n) there is an infinity equal to zero, determine all the values of n for which $c_n = 0$.

2. $ABCD$ is a parallelogram, $AB = a$, $AD = 1$, the angle DAB is α and the three angles of the triangle ABD are acute. Prove that the four circles K_A, K_B, K_C, K_D, each of radius 1 and with centres A, B, C, D, respectively, cover the parallelogram if and only if $a \leqslant \cos\alpha+\sqrt{3}\sin\alpha$.

There are many other mathematical competitions at all levels, such as the monthly problem on Hungarian television and the Putnam prize problems in America, but I have dealt in detail with those which impinge to some extent on young mathematicians in England today.

It is rather early to draw any general conclusions from these competitions, but one disturbing fact does emerge. Even in countries where schools are completely coeducational and where there is no social prejudice against intelligent or scientific women, girls do markedly worse in mathematics after the age of about fourteen. In the four years of the British Olympiad there have never been more than three girls out of 60 competitors and only once has one done reasonably well. In the international competition in Yugoslavia only one of the 108 competitors was a girl (from Bulgaria)!

On the Isoperimetric Theorem: History and Strategy

G. PÓLYA

Stanford University

Of all plane figures having the same perimeter, which one has the largest area?

The circle. This answer may seem plausible, but there is a long way from plausibility to proof, and a long history, some phases of which will be sketched in this article. These sketches are intended to introduce the reader to a mathematical subject of great beauty which has several elementary facets. A few attached questions endeavour to induce him to do some work of his own and to reflect upon his ways of working.

Points of history

1. The *isoperimetric theorem* is the name we shall give, following a widespread usage, to the following statement:

The area of the circle is larger than that of any other curve with the same perimeter.

The term 'isoperimetric' means 'of equal perimeter'. A 'curve' means a closed curve which is not self-intersecting in the plane. Our statement of the theorem emphasizes its 'unicity'. The problem of finding the curve of given perimeter that has the largest area has a unique solution: the circle.

2. *Zenodorus.* The known history of the isoperimetric theorem begins with Zenodorus, who was a Greek mathematician. We know very little about his life; the experts conjecture that he lived some time (probably not very long) after Archimedes. Yet that part of his work with which we are concerned can be quite well reconstructed from the writings of later commentators. He found:

I. The area of the regular polygon is larger than the area of any other polygon having the same number of sides and the same perimeter.

II. Of two regular polygons with the same perimeter the one that has more sides has the larger area.

III. The area of the circle is larger than that of any polygon with the same perimeter.

On the intuitive level, we have little difficulty in deriving proposition III from I and II.

3. *Simon Lhuilier* was a citizen of Geneva. His work with which we are concerned, *Abrégé d'isopérimétrie élémentaire*, appeared in 1789, in the first year of the French Revolution. He proves again the propositions I, II, and III of Zenodorus, but in a very different way. Here is one of his theorems which is particularly striking:

Of all polygons having the same sides the polygon inscribed in a circle has the largest area.

'Having the same sides' means, of course, that the polygons compared have the same number of sides and the corresponding sides are of equal length.

4. *Jakob Steiner* (1796–1863) was a Swiss mathematician; he started life as a cowherd and finished it as a professor at the University of Berlin. Steiner attempted several proofs of the isoperimetric theorem (and of its analogues in space and on the surface of the sphere). All his proofs have a common scheme which we can express in the form of a problem:

Given the curve C with the perimeter L and the area A, such that C is NOT a circle, construct a curve C' with perimeter L' and area A' such that

$$L' = L, \quad A' > A.$$

Steiner devised several ingenious constructions for changing a curve which is not a circle into an isoperimetric curve with a larger area. Yet does the feasibility of such a construction prove the isoperimetric theorem?

Points of strategy

Would you like to devise a proof for the isoperimetric theorem yourself? Probably you will not be able to devise one right away. Yet, by making a serious effort you may learn something about the strategy (or tactics) of problem-solving.

One of the first precepts of this strategy is: *If you cannot solve the proposed problem, try to devise another problem.* The other problem should be, of course, more accessible than, and have a chance to contribute to the solution of, the proposed problem: it should be a help, an *auxiliary problem.*

Do you want to prove the isoperimetric theorem? Its history offers you several promising auxiliary problems. Take for granted some, or all, of the results mentioned above and try to think of what is still lacking, what should be added to obtain a full proof.

If you don't see your way, try to prove some of the results stated, for example, those of Zenodorus, of Lhuilier. Such an effort may be the best means of familiarizing yourself with the ideas involved in the problem.

If you find these proofs too hard, try to carve out some manageable piece of the difficult whole: some particular case, some consequence that looks more accessible. I mention a few possibilities; but don't look at them before trying to find something by yourself.

To Zenodorus' proposition I: Prove that the equilateral triangle has a larger area than any other triangle with the same perimeter.

To Zenodorus' proposition II: Prove that the area of a square is larger than the area of an equilateral triangle with the same perimeter.

To Lhuilier's proposition: Prove that the area of a quadrilateral inscribed in a circle is larger than the area of a quadrilateral with the same sides that is not so inscribable. (This problem is hard.)

To Steiner's problem: Given a non-convex curve C with perimeter L and area A produce a convex curve C' with perimeter L' and area A' such that

$$L' < L, \quad A' > A.$$

Here is an auxiliary to an auxiliary problem: take for C a non-convex quadrilateral.

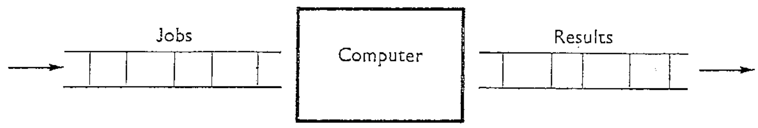
Having mastered a piece of a larger problem, return to the larger problem itself; the experience gained in working at that piece may enable you to master that larger problem too.

And even if you do not advance very far the experience gained may make you understand better your ways of working and improve your ability to solve problems. For the strategy or tactics of problem-solving, which is also termed heuristics, you may find my book *Mathematics and Plausible Reasoning*, especially Volume 1, Chapter 10 on the isoperimetric theorem, helpful.

An Introduction to Conversational Computing

DONALD MICHIE
University of Edinburgh

The diagram below depicts the 'batch-processing' style in which digital computing machines have mainly been used in the past. I want to base this article on a different style of computing, the so-called 'conversational mode', in which the user is free to conduct a rapid-fire dialogue with the machine, choosing his next response from moment to moment in the light of the latest output of the machine and of his own changing view of the situation.



In order to bring to life the essential nature of interactive computing, and the resources of a modern conversational computing language, I shall introduce the notion of a computing facility as a workshop for building internal (abstract) models of external (concrete) reality. Notice that I have framed my definition so widely that it could actually cover much of the activity which goes on inside our own skull-bound computing machines and which we call 'thought'. I have done this deliberately, since the restriction of digital computing to the processes of arithmetic is scarcely more than a historical accident, as artificial as if we were to confine the terms of reference of a metal workshop to operating upon nuts and bolts. This is not to say that nuts and bolts are not important.

I spoke above of a 'computing facility' rather than a 'computer'. This is because a naked computer, in the form in which it leaves the factory, can hardly be said to be a machine at all. It is only a potential machine. Which particular machine it will become out of an infinite range of possibilities is determined by the particular program with which it is loaded. I am going to assume that our computer has been loaded with the compiler program for a modern conversational computing language. A compiler is a program which translates from a computing language, which is oriented towards the human user's natural style of expressing himself, into the machine's own internal language. I shall further assume that this language is a truly mathematical (not just a numerical) one. In other words, in the general command schema

$$f(x) \rightarrow y$$

('apply the function f to the object x and call the result y') the user is free to define an f for x of any type whatsoever—integer, real, Boolean, word, string, list, array, record, set, group, ring, graph, etc.—and the type of y can be similarly unrestricted. Demanding this degree of mathematical generality narrows the range of programming languages from which to choose. For the illustrations which follow I shall use Burstall and Popplestone's language POP-2 with which I am most familiar.

In the real world we expect a workshop to contain

—materials

—structures

—machines (including hand-tools).

We use a machine to operate on one or more structures and/or materials to make one or more new structures. The new structure may or may not itself be a machine. Thus we apply a riveter to a pot, a handle, and some rivets to make a saucepan. A saucepan is itself a machine which, when applied to heat, water, and the contents of an appropriate packet, will make soup. The abstract world of computing correspondingly contains

—store

—data structures (e.g., lists, arrays, etc.)

—functions.

Thus we might apply a list-pairing function to a list of French words and a list of English words to make an association list or dictionary. To get the feel of the language the following POP-2 function for doing this is worth looking at in detail:

```
FUNCTION PAIROFF LIST1 LIST2;
    VARS NEWLIST;
    IF NOT (LENGTH(LIST1) = LENGTH(LIST2)) THEN PR("UNEQUAL") EXIT;
    NIL → NEWLIST;
NEXT: IF LIST1.NULL THEN NEWLIST EXIT;
    (LIST1.HD::[%LIST2.HD%])::NEWLIST → NEWLIST;
    LIST1.TL → LIST1; LIST2.TL → LIST2;
    GOTO NEXT
END;
```

The first line is synonymous with 'LET PAIROFF(LIST1,LIST2) BE ...', i.e., it names the function and introduces formal names for its arguments. In the next line we declare a local variable 'NEWLIST', which will be used to contain the result of the function.

Next comes a test for an error condition (list lengths unequal) which, if positive, causes exit from the function with a printed message 'UNEQUAL'. Now the real business starts, with the creation of an empty association list, which is assigned to the variable NEWLIST. The infixed operator '::' is a joining function, and HD and TL are functions which select the head and the tail of a list respectively. Thus [FEE FI FO].HD (a shorthand way of writing HD([FEE FI FO])) is FEE. [FEE FI FO].TL is [FI FO]. [FEE FI FO].HD::[FI FO FUM] is [FEE FI FO FUM]. In the 'pairoff' example above, the brackets of one of the lists have been decorated with per cent signs. This is done when the elements of the list are expressions, which need to be evaluated, as opposed to constants such as FEE or 2. [%"CHARGE", "IS", SQRT (X), "POUNDS"%] is therefore equivalent, if X happens to have the value 4 at the time, to [CHARGE IS 2 POUNDS]. We now repeatedly go round the loop, taking the heads off LIST1 and LIST2, joining these heads in pairs and chaining each pair in turn to the top of NEWLIST. When LIST1 and LIST2 are exhausted exit occurs with NEWLIST as the value of the function.

Suppose I type this function definition on a conversational terminal and follow it with

```
VARS FRENCH ENGLISH VOCAB;
[VACHE CHIEN CHAT HOMME MANGER DORMIR] → FRENCH;
[COW DOG CAT MAN EAT SLEEP] → ENGLISH;
```

(square brackets are used by convention to enclose lists).

I can now construct a French–English vocabulary by

```
PAIROFF(FRENCH,ENGLISH) → VOCAB;
```

and take a look at it by typing

```
VOCAB ⇒
```

(the symbol ⇒ is a print command).

The teletype will reply

```
**[[DORMIR SLEEP] [MANGER EAT] [HOMME MAN]
  [CHAT CAT] [CHIEN DOG] [VACHE COW]],
```

Note that in POP-2 the machine always prefaces its response to the ⇒ print command with the symbols **. If I don't like the reversal of the original order of items I can modify my 'pairoff' definition, for example by giving REV(NEWLIST), rather than NEWLIST, as the result. This illustrative exercise may appear at first sight to be a little vacuous, since all that PAIROFF does is to replace 'row pairing' by 'column pairing'. The usefulness of simple manipulations of this kind is seen if we now imagine writing a look-up function for French–English translation, such as the following:

```
    FUNCTION LOOKUP ITEM ASSOCLIST;
LOOP: IF ASSOCLIST.NULL THEN "UNKNOWN" EXIT;
        IF ASSOCLIST.HD = ITEM THEN
            ASSOCLIST.TL.HD EXIT;
        ASSOCLIST.TL → ASSOCLIST;
        GOTO LOOP
    END;
```

We would test this in the above example by typing, say,

LOOKUP(HOMME,VOCAB) ⇒, receiving the reply
** MAN, or
LOOKUP(FEMME,VOCAB) ⇒, with the reply
** UNKNOWN,

If we now update our list by

[FEMME WOMAN]::VOCAB → VOCAB;

we can again type

LOOKUP(FEMME,VOCAB) ⇒, this time with a different result:
** WOMAN,

Now let us press the workshop analogy a bit harder: let me imagine myself in a workshop for mixing and spraying paints. My task is to provide myself with a few dozen pots, to fill the first few with some starting colours, say red, blue, and yellow, and to use a paint mixer and a paintspray to apply paints of desired colours to specified objects—say pegs. First I set up a couple of dozen empty pots, by creating a one-dimensional array:

VARS POT;
NEWARRAY([%1,24%], INITIAL) → POT;

'NEWARRAY' is a constructor function of two arguments, the first—'[%1,24%]'—a list of lower and upper bounds and the second—'INITIAL'—a function which assigns starting values to the cells of an array. If we wanted the cells of the paintpot array to start empty we might previously define 'INITIAL' as follows:

FUNCTION INITIAL N; "BLANK" END;

to produce the result 'BLANK' for all values of N.

To check that all is well we now type

POT(2) ⇒

and on receiving the answer

**BLANK,

we proceed to fill the first three pots.

"RED" → POT (1);
"BLUE" → POT (2);
"YELLOW" → POT (3);

Typing the statement

POT (2) ⇒ now produces the result
**BLUE.

The next need is for a colour-mixing machine, which samples paint from two pots and puts a blend of the two into a third pot. For this purpose I have invented some simple laws of colour blending and have designed the mixing machine accordingly:

```
FUNCTION PAINTMIX I J K;
  IF NOT (POT(K) = "BLANK") THEN "FULL" => EXIT;
  VARS COLOUR1 COLOUR2;
  POT (I) -> COLOUR1; POT(J) -> COLOUR2;
  IF COLOUR1 = "MESS" OR COLOUR2 = "MESS" THEN "MESS" -> POT(K) EXIT;
  IF COLOUR1 = COLOUR2 OR COLOUR1 = "BLANK" THEN COLOUR2 -> POT(K) EXIT;
  IF COLOUR2 = "BLANK" THEN COLOUR1 -> POT(K) EXIT;
  IF COLOUR1 = "BROWN" OR COLOUR2 = "BROWN" THEN "BROWN" -> POT(K) EXIT;
  IF COLOUR1 = "BLUE" AND COLOUR2 = "YELLOW" THEN "GREEN" -> POT(K) EXIT;
  IF COLOUR1 = "RED" AND COLOUR2 = "BLUE" THEN "PURPLE" -> POT(K) EXIT;
        ...
        ...
        ...
    ELSE "MESS" -> POT(K) CLOSE
END
```

The details need hardly concern us, but the intended effect is to make plausible colour-products from any mixture of which one component is either brown or a primary colour, and otherwise to classify the result as a 'mess'.

Now to check that the mixer is working:

```
PAINTMIX(1,2,4);
POT(4) =>
**PURPLE,
PAINTMIX(3,4,4);
**FULL,
PAINTMIX(2,3,5); PAINTMIX(1,3,6);
POT(5) =>
**GREEN,
POT(6) =>
**ORANGE,
```

Remember that the machine's responses are determined by the information written into the definition of the PAINTMIX function. The user is in a position similar to someone carrying out a physical experiment on a tool of his own design, except that he has complete knowledge, in principle, of the laws governing the results, and the freedom to change these laws at will.

Now to construct a paintspray, or, to revert to the language of abstractions, to define a 'paintspray' function. This is an abstract model of a machine which when applied to a paint and a peg transfers the former to the latter's head

```
FUNCTION PAINTSPRAY PEG PAINT;
   PAINT -> PEG.HD
END
```

Create some objects for painting: VARS PEG1 PEG2 PEG3 PEG4; and give them 'blank', i.e., uncoloured, heads to start with.

```
[BLANK PEG] → PEG1;
[BLANK PEG] → PEG2;
[BLANK PEG] → PEG3;
[BLANK PEG] → PEG4;
```

Now we try painting them:

```
PAINTSPRAY(PEG1,POT(4));
PEG1 =>
**[PURPLE PEG],
PAINTMIX(1,4,8);
PAINTSPRAY(PEG2,POT(8));
PEG2 =>
**[MAGENTA PEG]
PAINTMIX(5,6,7);
POT(7) =>
**MESS,
```

so we empty pot 7: "BLANK" → POT(7); and re-fill it:

```
PAINTMIX(3,4,7);
POT(7) =>
**BROWN
```

We get tired of spelling out 'PAINTMIX' each time and give it a shorter name:

VARS MIX; PAINTMIX → MIX; now apply the new name: MIX(3,7,9);

We now need an instrument for displaying the contents of the first N pots:

```
FUNCTION PRINTPOTS N;
    VARS COUNT; 0 → COUNT
NEXT: IF COUNT = N THEN NL(3) EXIT;
    COUNT+1 → COUNT;
    NL(2); PR(COUNT); SP(1); PR(POT(COUNT));
    GOTO NEXT
END;
```

'SP(N)' means 'print N blank spaces' and 'NL(N)' means 'do carriage return and line feed N times'.

We try this function out. The command provokes the following response:

```
PRINTPOTS(10);
 1 RED
 2 BLUE
 3 YELLOW
 4 PURPLE
 5 GREEN
 6 ORANGE
 7 BROWN
 8 MAGENTA
```

```
  9 BROWN
 10 BLANK
```

and again:

```
MIX(1,8,10); MIX(2,8,11); MIX(3,8,12);
MIX(2,5,13); MIX(3,13,14); MIX(1,6,15);
PRINTPOTS(16);
  1 RED
  2 BLUE
  3 YELLOW
  4 PURPLE
  5 GREEN
  6 ORANGE
  7 BROWN
  8 MAGENTA
  9 BROWN
 10 RED
 11 PURPLE
 12 BROWN
 13 TURQUOISE
 14 GREEN
 15 VERMILION
 16 BLANK
```

Clearly any amount of elaboration and further fun can be extracted from this nursery exercise, without adding much didactic value. But before leaving it there is just one more example worth exhibiting of the parallelism between abstract model-building and building real models in real workshops. Suppose I have a tool or machine which normally requires two inputs to work on, just as a paintspray needs to have its back end applied to a paint source and its front end to a target. What happens if I apply it to just one of its inputs, and in some way freeze this input into its structure (as I might apply a paintspray to a paintpot and then weld the join solid)? The commonsense answer is that the result will be a machine of correspondingly restricted application, no longer a paintspray but, for example, a redspray, which now requires only one input (its target) to be supplied for it to do its work. There is not only commonsense but mathematical sense in this, for there is a real basis for regarding, say *add* (3,1) → 4 as proceeding via the creation of an intermediate function '*add*1' (i.e., the successor function in this example) which is then applied to 3 to give the final answer. The first of these two successive steps is called in POP-2 'partial application' of a function.

To illustrate using POP-2, we could say

```
ADD(%1%) -> SUCCESSOR;
```

(decoration of the argument brackets with 'per cent' signs is used as a signal that partial application is being used)

```
SUCCESSOR(3) =>
** 4,
```

or in terms of the paintshop

```
VARS SPRAY1 SPRAY2;
PAINTSPRAY(%POT(1)%) -> SPRAY1;
PAINTSPRAY(%POT(6)%) -> SPRAY2;
```

Now I can do the following

```
SPRAY1(PEG4); SPRAY2(PEG3);
PEG4 =>
**[RED PEG],
PEG3=>
**[ORANGE PEG];
```

Notice that if I now change the contents of pot 1 to green, e.g.,

```
"GREEN" -> POT(1);
```

the SPRAY1 function will continue to produce 'RED' and will not switch to 'GREEN'. The device of partial application will allow us to freeze into the function the value which the relevant part of its environment (in this case pot 1) had *at the time*—a most useful property when manipulating abstract models of a changing world.

The use of computers for model-building in industrially important applications is being extended today to the simulation of entire chemical plants, traffic systems, telephone networks, aeroplane structures and so forth, requiring computer programs costing dozens of man-years to write, test and maintain in use. At a simpler level a very wide range of non-numerical computer uses can be regarded as paralleling familiar categories in the outside world, in a sense which I hope my toy example has helped to make clear.

Mathematical Aspects of Smoking

P. CHADWICK

University of East Anglia, Norwich

1. Introduction

When a cigarette is lit the burning of the tobacco produces smoke which consists partly of gases and partly of minute liquid droplets. The latter part (called the particulate phase) is highly complex in composition; it contains several hundred distinct chemical compounds including those which have been shown to cause cancer in animals and which are now widely believed to contribute to the incidence of lung cancer in cigarette smokers. These so-called carcinogenic substances are formed in chemical reactions which take place at the high temperatures (about 850°C) prevailing at the burning end of the cigarette. They are not present in unburnt tobacco.

As the cigarette is smoked some of the smoke escapes into the air but most of it is drawn through the unburnt tobacco by the smoker. In their passage through the cigarette the particles present in the smoke undergo absorption in the sense that some of them are deposited on the tobacco. As the cigarette burns down re-distillation of compounds in the particulate phase therefore occurs and the concentration of these substances in the newly-formed smoke steadily increases.

It is evident from this simple description of the smoking process that the total amount of a specified compound which has been taken in by the smoker at a given time after lighting up depends upon the length of the unburnt portion of the cigarette and its effectiveness as a filter, and the possibility of a mathematical analysis of the situation presents itself.

2. Mathematical description. The conservation law

With reference to Figure 1, let x be distance measured from the unlit end O of the cigarette and let $X(t)$ be the length of the cigarette at time t after it is lit.

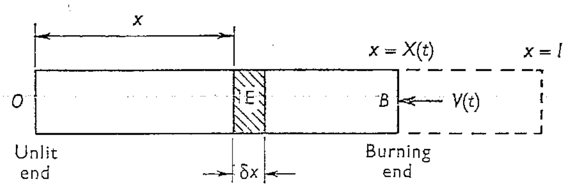


Figure 1

We denote by l the original length of the cigarette, by $V(t)$ the rate at which the length is reduced, and by T the total time of burning, that is, the time at which the butt-end is discarded. It follows from these definitions that

$$l = X(0) \quad \text{and} \quad V = -\frac{dX}{dt}. \tag{1}$$

Suppose now that we fix attention on the element E of the cigarette which lies between the planes distant x and $x + \delta x$ from O and consider the mass of a particular compound $\mathscr{C}$ which would be produced if the tobacco in E at time t were to be burned. This mass depends upon both x and t; for, as described in Section 1, smoke particles are being deposited on the unburnt tobacco and it is hardly to be expected that deposition will take place uniformly along the length of the cigarette. We therefore denote by $c(x, t)\,\delta x$ the mass of $\mathscr{C}$ produced by burning the element E at time t and we call c the *concentration* of $\mathscr{C}$ in the cigarette.[1]

The concentration is a function of the two *independent variables* x and t and we now introduce a second function of these variables. This time we station ourselves, in imagination, on the plane at distance x from O and measure the mass of compound $\mathscr{C}$ which is carried by the smoke during the interval of time I

[1] Note that if we were to detach E from the adjoining portions of the cigarette at time t and carry out a chemical analysis, the mass of $\mathscr{C}$ would be found to be much less than $c(x, t)\,\delta x$. This is because the full amount of $\mathscr{C}$ is released only when the tobacco in E is burned.

starting at t and ending at $t+\delta t$. Again the details of the smoking process given in Section 1 imply that this mass depends upon x and t and we denote it by $f(x, t)\,\delta t$; f is called the *flux* of $\mathscr{C}$ along the cigarette.

The fact that the compound $\mathscr{C}$ is absorbed as it is carried through the cigarette by the smoke means that, at a given time, the flux decreases as x decreases and that, at a given place, the concentration increases as t increases. We now show that these rates of change are the same by calculating in two different ways the total mass δm of $\mathscr{C}$ deposited in E during the time interval I. Firstly, δm is the difference between the mass of $\mathscr{C}$ entering E across the plane $x+\delta x$ and the mass leaving E across the plane x:

$$\delta m = f(x+\delta x, t)\,\delta t - f(x, t)\,\delta t. \tag{2}$$

Secondly, δm is the difference between the total masses of $\mathscr{C}$ which would be obtained by burning the tobacco in E at times $t+\delta t$ and t:

$$\delta m = c(x, t+\delta t)\,\delta x - c(x, t)\,\delta x. \tag{3}$$

On combining equations (2) and (3) we obtain

$$\frac{c(x, t+\delta t) - c(x, t)}{\delta t} = \frac{f(x+\delta x, t) - f(x, t)}{\delta x},$$

and letting δt and δx tend to zero yields the required result

$$\frac{\partial c}{\partial t} = \frac{\partial f}{\partial x}. \tag{4}$$

Equation (4) is called a *conservation law*. It expresses the balance of the compound $\mathscr{C}$ during the smoking of the cigarette and is quite independent of the mechanisms which govern the production and absorption of $\mathscr{C}$. Conservation laws similar in form to (4) are of fundamental importance in many branches of mathematical physics.

3. Construction of a mathematical model

The aim of the analysis which follows is to determine the distribution of the concentration c in the unburnt part of the cigarette (that is, for $0 \leqslant x \leqslant X(t)$) at all times in the interval $0 \leqslant t \leqslant T$. The only information so far at our disposal, equation (4), merely connects c to the flux f which is also unknown. Additional equations are therefore needed and they are obtained by setting up a mathematical representation of the smoking process. The aspects which particularly concern us are the production of the compound $\mathscr{C}$ at the burning end B of the cigarette and the absorption of $\mathscr{C}$ as the smoke filters through the unburnt tobacco. The model which we adopt is based upon the following assumptions.

ASSUMPTION 1. A constant fraction α (where $0 < \alpha < 1$) of the compound $\mathscr{C}$ produced at B is carried by the smoke into the unburnt portion OB of the cigarette. The remaining fraction $1-\alpha$ escapes.

ASSUMPTION 2. The flux f decreases as x decreases at a rate proportional to f.

Assumption 1 will be strictly accurate only when the cigarette is smoked at a constant rate. This of course is never the case in practice and we do not assume in the later analysis that V is constant.

Expressed mathematically Assumption 2 takes the form

$$\frac{\partial f}{\partial x} = qf, \tag{5}$$

where the constant of proportionality, q, is the *absorption coefficient* characterizing the deposition of $\mathscr{C}$ within OB. Integration of equation (5) yields[2]

$$f(x, t) = \phi(t)\exp(qx), \tag{6}$$

and we note that since the left-hand side of (5) is a *partial* derivative with respect to the independent variable x the usual *constant* of integration is replaced by a *function* of the second independent variable t; this function, ϕ, is as yet unknown. Let $F(t)$ be the flux of $\mathscr{C}$ at B, that is the mass of $\mathscr{C}$ which is drawn into the unburnt part of the cigarette per unit time. Then

$$F(t) = f(X(t), t) = \phi(t)\exp\{qX(t)\},$$

using (6), and we deduce that

$$\phi(t) = F(t)\exp\{-qX(t)\}.$$

Substitution of this expression for ϕ into (6) gives

$$f(x, t) = F(t)\exp(-q\{X(t)-x\}), \tag{7}$$

a formula for the distribution of the flux in OB in terms of the flux at B.

4. An integral equation for the concentration

Since the burning end B approaches the unlit end O with speed V, the length of the cigarette is reduced by an amount $V(t)\,\delta t$ in the time interval I. The weight of $\mathscr{C}$ produced during this time is found by multiplying this length by the concentration at B, namely $c(X(t), t)$. Recalling Assumption 1 we conclude that the mass of $\mathscr{C}$ entering the unburnt part OB of the cigarette during the time interval I is $\alpha c(X(t), t)\,V(t)\,\delta t$. But this mass is also given by $F(t)\,\delta t$. Hence

$$F(t) = \alpha c(X(t), t)\,V(t),$$

and the formula (7) for the flux can be rewritten as

$$f(x, t) = \alpha c(X(t), t)\,V(t)\exp(-q\{X(t)-x\}). \tag{8}$$

We now make use of the conservation law (4). Substituting the expression (8) for f into the right-hand side we obtain

$$\frac{\partial c}{\partial t} = \alpha q c(X(t), t)\,V(t)\exp(-q\{X(t)-x\}),$$

[2] $\exp x$ is an alternative way of writing e^x; it is preferred by printers for fairly obvious reasons.

and integration with respect to t between 0 and t gives[3]

$$c(x,t) = c(x,0) + \alpha q \exp(qx) \int_0^t c(X(\tau),\tau) \exp\{-qX(\tau)\} V(\tau)\, d\tau. \tag{9}$$

Apart from $c(x,0)$, which is the concentration of $\mathscr{C}$ in the unsmoked cigarette, c is the only unknown function in equation (9). We might therefore hope to determine c from this equation when the initial concentration is specified.

Since c appears under the integral sign (as well as on the left-hand side), (9) is called an *integral equation* for c. While differential equations are frequently encountered in elementary applied mathematics, particularly in mechanics, integral equations usually make their appearance only in more advanced work. Fortunately for us the integral equation (9) can be turned into a differential equation which is easy to solve.

5. Solution of the integral equation

For simplicity we shall suppose that the concentration of $\mathscr{C}$ in the original cigarette is uniform, so that $c(x,0) = c_0$, a constant. Equation (9) can then be written in the form

$$c(x,t)\exp(-qx) = c_0 \exp(-qx) - \alpha q \int_0^t c(X(\tau),\tau) \exp\{-qX(\tau)\} \frac{dX}{d\tau}\, d\tau, \tag{10}$$

where use has been made of the second of equations (1). Our next step is to regard t as a function of X (reversing the state of affairs which has held up to now[4]) and to change the variable of integration in (10) from τ to X. Defining the function $C(X)$ by

$$C(X) = c(X, t(X)) \exp(-qX) \tag{11}$$

and remembering that $X = l$ at $t = 0$, equation (10) becomes

$$c(x,t)\exp(-qx) = c_0 \exp(-qx) + \alpha q \int_{X(t)}^{l} C(\xi)\, d\xi. \tag{12}$$

On evaluating the first two terms of (12) at $x = X$ we obtain the much simplified integral equation

$$C(X) = c_0 \exp(-qX) + \alpha q \int_X^l C(\xi)\, d\xi. \tag{13}$$

Differentiating each term of equation (13) with respect to X now gives

$$\frac{dC}{dX} + \alpha q C = -c_0 q \exp(-qX),$$

[3] See Appendix.

[4] Since X is a steadily decreasing function of t, the functional relation $X = X(t)$ can be inverted to give $t = t(X)$. In effect we are now using the length of the cigarette to measure the passage of time.

a differential equation for C which is solved by multiplying through by an integrating factor, in this case $\exp(\alpha q X)$. We then have

$$\frac{d}{dX}\{C\exp(\alpha q X)\} = -c_0 q \exp\{-(1-\alpha)qX\},$$

and integration with respect to X yields

$$C\exp(\alpha q X) = \frac{c_0}{1-\alpha}\exp\{-(1-\alpha)qX\}+\text{constant.} \tag{14}$$

To determine the constant of integration we return to the definition (11) and put $X = l$. Since $t = 0$ when $X = l$ we find that $C(l) = c(l, 0)\exp(-ql) = c_0\exp(-ql)$ and the constant in (14) is hence $-c_0\alpha(1-\alpha)^{-1}\exp\{-(1-\alpha)ql\}$. It follows that

$$C(X) = \frac{c_0}{1-\alpha}[\exp(-qX)-\alpha\exp\{-(1-\alpha)ql-\alpha qX\}], \tag{15}$$

and the solution of the integral equation (10) is completed by substituting this expression for C into equation (12), carrying out the integration and then reverting to the original independent variables, x and t. The result is

$$c(x,t) = c_0\left[1+\frac{\alpha}{1-\alpha}\exp(-q\{X(t)-x\})\left\{1-\exp(-(1-\alpha)q\{l-X(t)\})\right\}\right]. \tag{16}$$

Note that $X(t)-x$ is distance from the burning end B and that $l-X(t)$ is the length of cigarette which has been burned at time t.

6. Application of the solution

With the derivation of equation (16) the main objective of our analysis has been reached. We now use this solution to investigate a feature of the smoking process which is of particular interest, namely the smoker's total intake of the compound $\mathscr{C}$. This is the mass M of $\mathscr{C}$ which crosses the plane $x = 0$ during the time interval $0 \leqslant t \leqslant T$ and it is obtained by integrating the flux at O with respect to t over this interval[5]:

$$M = \int_0^T f(0,\tau)\,d\tau.$$

Using the formula (8) for f and the second of equations (1) there follows

$$M = -\alpha\int_0^T c(X(\tau),\tau)\exp\{-qX(\tau)\}\frac{dX}{d\tau}\,d\tau.$$

We now change the variable of integration from τ to X exactly as in Section 5.

[5] The reader will find it instructive to derive the alternative expression

$$M = \int_{-\infty}^0 \{c(x,T)-c_0\}\,dx$$

for M and to verify, using equation (16), that it also leads to the formula (18).

Recalling the definition (11) we then obtain

$$M = \alpha \int_{X^*}^{l} C(\xi)\, d\xi, \tag{17}$$

where $X^* = X(T)$ is the length of the butt-end which the smoker throws away. The total intake of $\mathscr{C}$ is now found by entering into equation (17) the expression (15) for C and carrying out the integration:

$$M = \frac{c_0 \alpha}{(1-\alpha)q} \exp(-ql)\,[\exp\{q(l-X^*)\} - \exp\{\alpha q(l-X^*)\}]. \tag{18}$$

The total mass of $\mathscr{C}$ which would be produced by burning the entire cigarette is $c_0 l$. Expressing M as a fraction of this quantity we can put equation (18) into the form

$$\frac{M}{c_0 l} = \frac{\alpha}{1-\alpha} \frac{1}{ql} [\exp(-ql\beta) - \exp\{-ql(1-\alpha+\alpha\beta)\}], \tag{19}$$

where $\beta = X^*/l$. We see that $M/c_0 l$ depends upon the three numbers α, β, and ql which are, in turn, the fraction by mass of the smoke which is drawn into the cigarette, the fraction of the cigarette which is discarded, and (see equation (7)) the reciprocal of the length of cigarette, expressed as a multiple of l, which would be needed to reduce the flux of $\mathscr{C}$ by the factor $1/e = 0{\cdot}37$.

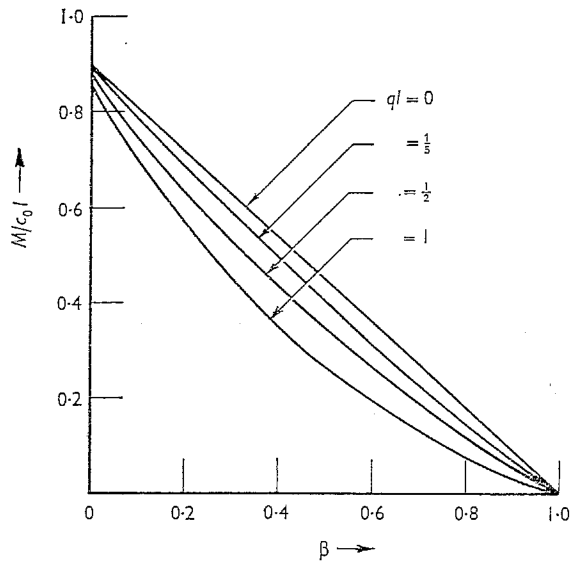


Figure 2

Taking α to be 0·9 the variation of $M/c_0 l$ with β for different values of ql can easily be computed from (19). Some results are shown in Figure 2 and inspection of these graphs shows that, in reducing the smoker's intake of $\mathscr{C}$, the advantage to be gained by discarding a greater fraction of the cigarette (say 0·4 or 0·5 rather than the more usual 0·2 or 0·3) becomes progressively greater as the absorption coefficient increases. Unfortunately a high value of q would impede the flow of smoke and spoil the smoker's enjoyment and the extent to which the desirable aim of increasing q can be achieved in practice is severely limited by this consideration.

7. Concluding remarks

A mathematical analysis of the filtration process accompanying the smoking of a cigarette was first developed in 1950 by M. S. Klamkin and published some years later (reference 1). An alternative approach to Klamkin's theory has been given in a recent book by B. Noble (reference 2) which contains a wealth of novel applications of elementary mathematics. The present article is based upon Professor Noble's treatment but differs from it in making explicit the role of the conservation law (4). Also, because the construction of models is one of the essential skills involved in the application of mathematics, this aspect of the problem has been discussed in rather greater detail than in references 1 and 2. An extended account of mathematical model building with interesting applications to industrial problems can be found in reference 3.

It will be evident to the reader that the theory of cigarette smoking which has been presented here is, in certain respects, over-simplified. In particular, we have left aside the important question of the effects of a filter tip on the smoker's intake of $\mathscr{C}$. The theory can be extended to deal with this and other complications and some details are given in references 1 and 2.

The author's thanks are due to Harold and Lionel Clarke for their helpful comments on an earlier version of this article.

References

1. M. S. Klamkin, A moving boundary filtration problem or 'The cigarette problem'. *American Mathematical Monthly*, **64** (1957), 710–715.
2. B. Noble, *Applications of Undergraduate Mathematics in Engineering* (Macmillan, New York, 1967), pp. 153–157.
3. J. Crank, *Mathematics and Industry* (Oxford University Press, London, 1961).

Appendix

In equation (9) τ has made its appearance as the variable of integration. It can easily be seen that τ has in fact taken the place of t in the previous equation for $\partial c/\partial t$. The technique which is being used here is a standard one and it can be explained in the following way.

Suppose that g is a function of a single variable, say x, and that we know that

$$\frac{dg}{dx} = G(x),$$

where G is a further function of x. Then, if we wish to integrate that equation and to display g itself in terms of x we shall write

$$g(x) - g(0) = \int_0^x G(\xi)\, d\xi,$$

where, because we wish to use x as a limit of integration, we must use a different symbol for the variable under the integral sign.

This technique is made use of in obtaining equation (9), where τ has replaced t, and the same technique is used again in obtaining equation (12), where ξ has replaced X.

Forecasting Trends

G. A. WATTERSON[1]
University of Sheffield

1. Introduction

A typical problem in meteorology is the following. We know the yearly rainfalls for the past n years, which we write as $X_1, X_2, \ldots, X_n$, where X_1 refers to the first recorded year, X_n to the last completed year. Farmers, reservoir managers, umbrella manufacturers, and others are interested in the question: How much rain will fall this year? So the meteorologist studies ways to predict X_{n+1} from the available data. An obvious and simple predictor would be obtained by averaging the observed values,

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} \tag{1}$$

and use $\bar{X}$ as predicted value for X_{n+1}. Of course, as rainfall is such a variable quantity, we would be surprised indeed if X_{n+1} turned out to be exactly as predicted! However, we would like predictors which were as good as possible in some sense, perhaps those whose squared-deviation from the actual value is as small as possible on average.

It may be that X_{n+1} would tend to be closer to results of recent years than to hose of more remote times, in which case we would give the former more weight than the latter, using a weighted average of the form

$$c_1 X_1 + c_2 X_2 + \ldots + c_n X_n \quad \left(\sum_{i=1}^{n} c_i = 1\right) \tag{2}$$

to predict X_{n+1}. Here, the coefficients $c_1, c_2, \ldots, c_n$ are decided upon by methods studied in the subject called 'Time Series Analysis'. We need not go into details, for actual data suggest that rainfall from one year is very little dependent on that of previous years, so that a predictor like (2) is little better than (1) itself.

For some purposes, it may be sufficient to predict which of $X_{n+1} > X_n$, $X_{n+1} = X_n$, or $X_{n+1} < X_n$ will hold, that is, whether the 'trend' is for wetter, equally wet, or drier weather respectively, without being more precise as to the actual value of X_{n+1}. We study this problem in Section 3, but first we look at a similar problem in gambling.

2. Choosing the right game to play

Peter and Paul are playing a game with two dice, and Peter notices that he is very rapidly losing his capital. And yet the game seems fair enough: first Peter tosses a die and both can see the result, then Paul predicts whether his toss will achieve a higher, equal, or lesser result. And Paul is predicting much better than

[1] Present address: Monash University, Clayton, Victoria, Australia.

50% correctly, to Peter's concern, in spite of the fact that the coins seem unbiased and fairly tossed.

There is no paradox here, as a little analysis of the game shows. If Peter's toss is a 1, then the probability is 5/6 that Paul will toss more than a 1 (namely, a 2, 3, 4, 5, or 6). Similarly, if Peter's toss is a 2, there is a 4/6 chance that Paul will toss more, and if Peter's toss is a 3, there is a 3/6 chance that Paul will toss more. Therefore Paul decides that in any of these cases he will be correct in about half the time, *or better*, if he predicts that he will toss more than Peter did. On the other hand, if Peter tosses a 4, 5, or 6, then there is at least 50% chance, *or better*, that Paul will toss less, and this is the way he will predict the outcome then.

The probability that Paul predicts correctly is the average of the probabilities given the various possible scores for Peter's die. We take an unweighted average assuming that Peter's die is as likely to turn up one face as another.

Pr (Paul predicts correctly)

$$= \tfrac{1}{6}\cdot\tfrac{5}{6}+\tfrac{1}{6}\cdot\tfrac{4}{6}+\tfrac{1}{6}\cdot\tfrac{3}{6}+\tfrac{1}{6}\cdot\tfrac{3}{6}+\tfrac{1}{6}\cdot\tfrac{4}{6}+\tfrac{1}{6}\cdot\tfrac{5}{6}$$
$$= \tfrac{2}{3}.$$

Thus for this game, and this prediction rule, Paul should win about two-thirds of the games played, without cheating. Adding to this the possibilities of cheating, and Paul is on to a good thing! In the next section, we show (in a sense) that even without cheating but using dice with many sides, Paul's chances can go as high as 75% correct. (A similar problem about an n-sided die will be found in the Problems section.)

3. Trends in rainfall data

Our farmer, who was mentioned in Section 1, is concerned with whether this year he should buy more cattle, keep the same number, or sell some cattle, compared with last year, and his decision might be influenced by whether $X_{n+1} > X_n$, $X_{n+1} = X_n$, or $X_{n+1} < X_n$ will hold. X_{n+1} is the rainfall to be obtained this year, and X_n is that for last year. If the farmer knows the median rainfall, m, fairly accurately from past records, where m is that quantity such that

$$\Pr(X_{n+1} > m) = \Pr(X_{n+1} < m) = \tfrac{1}{2},$$

then he is in a good position to predict the trend. (Notice that in the above we have assumed $\Pr(X_{n+1} = m) = 0$, with the idea that rainfall is continuously variable; for X_{n+1} to be exactly equal to m, or any other real number, is an event with zero probability, while not being impossible.)

From the definition of m, we see that if $X_n = x < m$, then the event $X_{n+1} > x$ has probability at least $\frac{1}{2}$, while if $X_n = x > m$, then the event $X_{n+1} < x$ has probability at least $\frac{1}{2}$. Hence we will be correct with at least a probability of $\frac{1}{2}$ if we use the prediction rule:

$$\text{when } X_n < m, \quad \text{predict that } X_{n+1} > X_n \text{ will hold,}$$

and,

$$\text{when } X_n > m, \quad \text{predict that } X_{n+1} < X_n \text{ will hold.}$$

Notice that this is similar to Paul's prediction rule in the previous section, especially if we take $m = 3\frac{1}{2}$, say. In the unlikely event that $X_n = m$ exactly, we might decide to predict that X_{n+1} will be larger, although the opposite would have the same probability.

To investigate the probability that our prediction rule gives us a correct conclusion, we assume that the probability distribution of X_{n+1} is unaffected by the value taken by X_n. We denote the distribution function by

$$F(x) = \Pr(X_{n+1} < x) = \Pr(X_{n+1} \leqslant x) = 1 - \Pr(X_{n+1} > x)$$

which we assume possesses a density function $f(x)$ where

$$f(x) = \frac{dF(x)}{dx} \quad \text{and} \quad F(x) = \int_0^x f(y)\, dy.$$

When $X_n = x \leqslant m$, we predict that $X_{n+1} > x$ according to our rule; this has probability $1 - F(x)$. But when $X_n = x > m$ we predict that $X_{n+1} < x$, which has probability $F(x)$. Averaging these probabilities over possible values for X_n itself (according to the probability density $f(x)$ which we assume applies to X_n just as much as to X_{n+1}), the long-term prediction rate is
Pr(correct prediction)

$$\begin{aligned}
&= \int_0^m (1 - F(x)) f(x)\, dx + \int_m^\infty F(x) f(x)\, dx, \\
&= \int_0^m f(x)\, dx - \int_0^{\frac{1}{2}} F\, dF + \int_{\frac{1}{2}}^1 F\, dF, \\
&= \tfrac{1}{2} - \tfrac{1}{2}[(\tfrac{1}{2})^2 - 0] + \tfrac{1}{2}[1^2 - (\tfrac{1}{2})^2], \\
&= 3/4.
\end{aligned}$$

In this derivation, we have used the facts that

$$\Pr(X_n < m) = F(m) = \int_0^m f(y)\, dy = \tfrac{1}{2},$$

and that

$$\Pr(X_n < \infty) = F(\infty) = \int_0^\infty f(y)\, dy = 1.$$

Our conclusion, then, is that for continuously variable quantities X_n and X_{n+1}, which have the same distribution and are independent of each other, it is possible to predict the 'trend' with a 75% success rate. I know of one eminent scientist who was very pleased to have achieved nearly this success rate after hours of computer calculations, in which prediction was based on an expression like (2). He thought (unfortunately, incorrectly) that he had made a great advance in the science of forecasting trends.

Solving Equations in the Algebra of Classes

R. L. GOODSTEIN

University of Leicester

The need to solve equations arises in many parts of mathematics. Simple equations like $3x = 6$ or $3x+4 = 10$ are as old as algebra itself and the problems presented by equations like $3x+4 = 1$ and $3x = 7$ led to the extension of the number system from positive integers to negative integers and fractions. Some equations, like those we have already mentioned, have unique solutions, others like $2x+3y = 11$ have more than one solution, finitely many in positive integers, but infinitely many if we accept both positive and negative solutions.

In this note I shall consider equations, not between numbers but between sets, the solutions to which are themselves sets. I shall denote sets by capital letters, and for any two sets A and B, I shall write AB for their intersection (common part) and $A \cup B$ for their union[1]; the complement of a set A (the set of elements not in A) will be denoted by A'. I shall write 0 for the empty set (the set with no members) and 1 for the universal set (the set from which all other sets are selected). In the sequel we shall speak of classes rather than sets, treating the two terms as equivalent.

We shall find necessary and sufficient conditions for a class equation to have a solution, and when the conditions are satisfied we shall show how to find all the solutions. The solution of class equations has applications in electrical engineering but I shall not consider these applications.

I shall assume that it is known that union and intersection are commutative and associative, that each distributes over the other, and that the De Morgan laws for complements

$$(AB)' = A' \cup B', \quad (A \cup B)' = A'B'$$

hold. I note that since A and A' have no elements in common, $AA' = 0$, and since every element belongs either to A or to A', $A \cup A' = 1$. Clearly $A \cup 0 = A$ and $A1 = A$ since the empty class contributes nothing to a union and the universal class contains the whole of A.

I start by considering the equation

$$AX \cup BX' = 0 \tag{1}$$

in the single unknown X.

If the equation has a solution, say $X = X_0$, then

$$AX_0 \cup BX_0' = 0$$

and so $AX_0 = 0$ and $BX_0' = 0$; but $X_0 \cup X_0' = 1$ and therefore

$$AB = AB(X_0 \cup X_0') = ABX_0 \cup ABX_0' = 0,$$

[1] The union of A and B is the set whose numbers are the members of A or of B and no others.

which shows that $AB = 0$ is a necessary condition for equation (1) to have a solution. The condition is also sufficient: when $AB = 0$, then $X = B$ is a solution; for if $X = B$, then

$$AX \cup BX' = AB \cup BB' = 0 \cup 0 = 0.$$

Thus $AB = 0$ is a necessary and sufficient condition for equation (1) to have a solution. It remains to find all the solutions of the equation when this condition is satisfied. I shall show that, when $AB = 0$ (so that B is outside A and therefore contained in A'), *then any class contained in the 'interval'* (B, A'), *i.e., contained in* A' *and containing* B, *and only such classes, are solutions of* (1).

I observe first that if X_0 is a class in the interval (B, A'), then because B is contained in X_0, no part of B is in X_0' and so $BX_0' = 0$. Similarly, since X_0 is contained in A', $AX_0 = 0$, whence $AX_0 \cup BX_0' = 0$, i.e., X_0 is a solution of (1).

Conversely, let X_0 be a solution of (1), so that $AX_0 = BX_0' = 0$. Then no part of X_0 is in A so that X_0 is contained in A', and no part of B is outside X_0 so that B is contained in X_0. Thus X_0 lies in the interval (B, A'), as was to be proved.

The aggregate of classes which lie in a given interval may be given a simple representation. I shall show that all classes which lie in an interval (A, B) are given by the formula

$$AP \cup BP',$$

where P is arbitrary.

We note first that, as A is contained in B (which is entailed in calling (A, B) an interval), we have $AB = A$ and $AB' = 0$. Hence

$$A(BP \cup AP') = ABP \cup AP' = AP \cup AP' = A(P \cup P') = A,$$

which shows that $BP \cup AP'$ contains A; and

$$B'(BP \cup AP') = BB'P \cup AB'P' = 0,$$

which shows that $BP \cup AP'$ is contained in B. Accordingly for any P, $AP \cup BP'$ lies in the interval (A, B).

Suppose now that X_0 is any class contained in the interval (A, B). Let $P = BX_0'$. Then $AP = 0$, since A lies in X_0 and so outside X_0', so that $AX_0' = 0$. Also, $BP' = B(B' \cup X_0) = BX_0 = X_0$. Hence

$$AP \cup BP' = X_0.$$

Thus a class lies in the interval (A, B) if and only if it may be represented in the form $AP \cup BP'$. In particular therefore *any solution of the equation*

$$AX \cup BX' = 0$$

may be expressed in the form

$$A'P \cup BP'.$$

We have so far considered only one very special equation $AX \cup BX' = 0$, but we shall now show that in class algebra every equation $F = G$ in an unknown X (where F and G are built up from X and other letters by union, intersection, and complementation) may be reduced to the form $AX \cup BX' = 0$ and so solved.

Since a union $A \cup B$ is empty if and only if each of A and B is empty, therefore a pair of equations

$$A = 0, \quad B = 0$$

is equivalent to a single equation $A \cup B = 0$. Thus in the algebra of classes it is unnecessary to consider simultaneous systems of equations, since such a system is equivalent to a single equation. Moreover an equation $A = B$ holds if and only if A is contained in B and B is contained in A; but A is contained in B if and only if A has nothing in common with B', i.e., $AB' = 0$. Hence $A = B$ if and only if both $AB' = 0$ and $A'B = 0$ hold, and these hold simultaneously if and only if

$$AB' \cup A'B = 0.$$

It follows that any system of equations $A_i = B_i$ $(i = 1, 2, \ldots, n)$ is equivalent to a single equation

$$C = 0.$$

Using the De Morgan laws to work out complements we can transform any expression built up from class letters by union, intersection, and complementation to a string of unions and intersections of class letters with or without dashes. For instance

$$(A' \cup BC)'(A \cup D)$$

becomes in turn

$$A(BC)'(A \cup D) = A(B' \cup C')(A \cup D).$$

Next we use the distributive law to remove brackets round unions and we arrive at an expression of the form

$$I_1 \cup I_2 \cup \ldots \cup I_k,$$

where each I is just an intersection. To return to our previous example,

$$A(B' \cup C')(A \cup D)$$

becomes

$$A(AB' \cup AC' \cup B'D \cup C'D) = AB' \cup AC' \cup AB'D \cup AC'D.$$

Suppose now that we start with an equation $F = 0$ which contains an unknown X; F may contain both X and X'; arrange F in the form

$$I_1 \cup I_2 \cup \ldots \cup I_k.$$

If any of the factors I does not contain X or X' we may replace it by $IX \cup IX'$ since $X \cup X' = 1$ and so $IX \cup IX' = I(X \cup X') = I$. Thus F may be arranged in the form

$$I_1 \cup I_2 \cup \ldots \cup I_p \cup J_1 \cup J_2 \cup \ldots \cup J_q,$$

where the I's all contain X and the J's contain X'. (We do not need to include terms containing XX' since this intersection is empty.)

Using the distributive law again we can take the X outside a bracket to give $I_1 \cup I_2 \cup \ldots \cup I_p = AX$ say, and similarly $J_1 \cup J_2 \cup \ldots \cup J_q = BX'$, say. Therefore, finally, we have reduced the equation $F = 0$ to the form

$$AX \cup BX' = 0,$$

which we know how to solve.

To sum up, we have seen that in class algebra the condition for any equation or system of equations to have a solution is known and when this condition is satisfied all the solutions can be found. The method may be extended to equations in any number of unknowns, but I shall not consider the extension here.

Problems and Solutions

Readers who have not yet reached the age of 20 on 1 October 1969 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 the postal address of your school, college or university.

Problems

12. Show that, if z is a complex number such that $-1 \leqslant \mathscr{R}z \leqslant 1$, then

$$|1+z^2| \geqslant 2(\mathscr{R}z)^2.$$

13. Let $0 \leqslant a \leqslant 1$, $0 \leqslant b \leqslant \frac{1}{2}$. Show that there exists an interval I of length 1 such that

$$|x|+|ax+b| \leqslant 1$$

whenever x belongs to I.

14. Let c be a positive number. The sequence $a_1, a_2, a_3, \ldots$ is such that $a_1 > 0$ and $a_{n+1}-a_n \geqslant c$ for all n. Prove that

$$\frac{a_{n+1}}{(a_{n+1}-a_n)\,a_n^2} \to 0 \quad \text{as} \quad n \to \infty. \qquad (*)$$

Find an increasing sequence $a_1, a_2, a_3, \ldots$ of positive numbers such that $a_n \to \infty$ as $n \to 0$, but (*) does not hold.

15. An n-sided die has faces numbered $1, 2, 3, \ldots, n$. Peter tosses the die first, and then Paul predicts whether his own toss will be more than, equal to, or less than Peter's in outcome. Assuming that the die is fair, and first taking n to be even, find the rule that Paul should use to maximise his correct-prediction probability, and find that probability. Solve the problem also when n is odd.

Solutions to Problems in Volume 1, Number 2

7. Evaluate the integral

$$\int_0^{\pi} \log \sin \theta \, d\theta.$$

Solution. Denote the integral by I. Putting $\phi = \theta - \frac{1}{2}\pi$ we have

$$\int_{\frac{1}{2}\pi}^{\pi} \log \sin \theta \, d\theta = \int_0^{\frac{1}{2}\pi} \log \cos \phi \, d\phi.$$

Hence

$$I = \int_0^{\frac{1}{2}\pi} \log \sin \phi \, d\phi + \int_0^{\frac{1}{2}\pi} \log \cos \phi \, d\phi$$

$$= \int_0^{\frac{1}{2}\pi} \log(\sin \phi \cos \phi) \, d\phi$$

$$= \int_0^{\frac{1}{2}\pi} \log(\tfrac{1}{2} \sin 2\phi) \, d\phi.$$

The change of variable $\psi = 2\phi$ now gives

$$I = \tfrac{1}{2}\int_0^{\pi} \log(\tfrac{1}{2} \sin \psi) \, d\psi = \tfrac{1}{2}I - \tfrac{1}{2}\int_0^{\pi} \log 2 \, d\psi$$

and so

$$I = -\pi \log 2.$$

Solved, in essentially the above way, by G. B. Chaplin (Beauchamp Grammar School, Oadby), H. E. Clarke (Attleborough, Norfolk), J. Cummings (University of Glasgow), P. D. Dickenson (March Grammar School), R. Graham (Grammar School for Boys, Cambridge), R. J. Hall (The Grammar School, Ebbw Vale), G. Robertson (Leeds Grammar School), A. Russell (King George V School, Southport).

8. Let ABC be a triangle such that $AC = BC$ and $A\hat{C}B = 90°$. Let P be a point in the plane of ABC such that P and C lie on opposite sides of AB. Show that, if PC bisects $A\hat{P}B$ and if $AP \neq BP$, then $A\hat{P}B = 90°$.

Solution by S. R. Blake (Rugby School)

In the triangles APC and BPC, CP is common, $CA = CB$ and $A\hat{P}C = B\hat{P}C$. Hence the triangles are congruent or else the angles $C\hat{A}P$, $C\hat{B}P$ are supplementary. Since $AP \neq PB$, the triangles are not congruent. Thus, in the quadrilateral $CAPB$, $C\hat{A}P + P\hat{B}C = 180°$, $B\hat{C}A = 90°$. It follows that $A\hat{P}B = 90°$.

Alternative solution. Assume that $BP < AP$. Produce PB to A' so that $PA' = PA$. The triangles $PA'C$, PAC are clearly congruent and so $A'C = AC = BC$. Thus A, B, A' lie on a circle with centre C and so $A\hat{A}'B = \frac{1}{2}A\hat{C}B = 45°$. Since $PA = PA'$, $P\hat{A}A' = P\hat{A}'A$ $(= B\hat{A}'A) = 45°$. Hence $A\hat{P}B$ $(= A\hat{P}A') = 90°$.

Also solved by G. B. Chaplin (Beauchamp Grammar School, Oadby), H. E. Clarke (Attleborough, Norfolk), A. Mingay (Grammar School for Boys, Cambridge), G. Robertson (Leeds Grammar School), P. Turner (The Grammar School, Ebbw Vale).

9. Let C denote the class of positive integers which, when written in the scale of 3, do not require the digit 2. Show that no three integers in C are in arithmetic progression.

Solution by S. R. Blake (Rugby School)

Let $d\,(>0)$ be the common difference of a given arithmetic progression. Suppose that, when d is written in the scale of 3, the first non-zero digit, counted from the right, is in the nth position. Then any three consecutive members of the arithmetic progression, written in the scale of 3, have different digits in the nth place. One of these digits is 2 and the corresponding number does not belong to C.

Alternative solution. Let x, y, $z \in C$ and write

$$x = \sum_{k=0}^{\infty} x_k 3^k, \quad y = \sum_{k=0}^{\infty} y_k 3^k, \quad z = \sum_{k=0}^{\infty} z_k 3^k,$$

so that each x_k, y_k, z_k is 0 or 1 (and all except a finite number are 0). Suppose that $x+y = 2z$. Then

$$\sum_{k=0}^{\infty} (x_k + y_k)\, 3^k = \sum_{k=0}^{\infty} (2z_k)\, 3^k.$$

Since $0 < x_k + y_k < 2$ and $0 < 2z_k < 2$,

$$x_k + y_k = 2z_k$$

for all k. Also $2z_k = 0$ or 2 and therefore either $x_k = y_k = 0$, in which case $z_k = 0$, or else $x_k = y_k = 1$, in which case $z_k = 1$. Thus $x_k = y_k = z_k$ for all k, i.e., $x = y = z$.

Also solved by R. P. Allen (Grammar School for Boys, Cambridge), G. B. Chaplin (Beauchamp Grammar School, Oadby), H. E. Clarke (Attleborough, Norfolk), R. J. Hall (The Grammar School, Ebbw Vale).

10. Let a, b, c, d be real numbers and write

$$f(x, y) = axy + bx + cy + d.$$

Let R be a rectangle with its sides parallel to the coordinate axes. Show that, if $f(x, y) \geqslant 0$ at the vertices of R, then $f(x, y) \geqslant 0$ throughout R.

Does this inference remain valid if the requirement that the sides of R should be parallel to the coordinate axes is dropped?

Solution. For fixed y, $f(x, y)$ is linear in x. Hence $f(x, y) \geqslant 0$ on the sides of R parallel to the x-axis. Now let (x_0, y_0) be any point in R. The line through (x_0, y_0) parallel to the y-axis meets the sides of R parallel to the x-axis at points where $f \geqslant 0$. Since $f(x_0, y)$ is linear in y, $f(x_0, y_0) \geqslant 0$.

The inference is false if the sides of R are not required to be parallel to the axes. For take R to be the square with vertices at $(1, 0)$, $(-1, 0)$, $(0, 1)$, $(0, -1)$ and let $f(x, y) = xy$. Then $f \geqslant 0$ at the vertices of R, but $f(\frac{1}{2}, -\frac{1}{2}) < 0$.

Also solved by G. B. Chaplin (Beauchamp Grammar School, Oadby), H. E. Clarke (Attleborough, Norfolk), R. J. Hall (The Grammar School, Ebbw Vale).

11. A system has three types of components. The cost c_i of the ith component and the probability q_i that this component will be satisfactory are shown in the table below.

Component i	c_i	q_i
1	4	0·9
2	1	0·8
3	2	0·7

Find the number of components of each type that the system should have if it is to achieve a reliability of at least 0·95 at minimum cost.

Solution. The required system is (2, 3, 3), which gives a reliability of 0·955 at a cost of 17. The development of the solution is as follows.

System	Reliability	Cost
(1, 1, 1)	0·504	7
(1, 2, 1)	0·605	8
(1, 2, 2)	0·786	10
(1, 2, 3)	0·841	12
(1, 3, 3)	0·896	13
(2, 3, 3)	0·955	17

Solved by S. R. Blake (Rugby School), H. E. Clarke (Attleborough, Norfolk), D. J. Custerson (Grammar School for Boys, Cambridge), R. Dobbs (The Grammar School, Ebbw Vale), A. Mingay (Grammar School for Boys, Cambridge), C. Robertson (Leeds Grammar School), J. M. Williams (The Grammar School, Ebbw Vale).

Mathematics for the Liberal Arts Student. By F. W. RICHMAN, C. N. WALTER and R. J. WISNER. Brooks-Cole Publishing Co., Belmont, California, 1967. Pp. 190. 75*s*.

The title of the book is self-explanatory, but it does not reveal that it is an American book written for the American college freshman. Translating into English terms, it would seem that the book might be appropriate for a mathematics course in a sixth form general studies scheme in an English school.

Much of the book contains an American version of what is in standard English modern syllabuses for 'O' level. One or two topics in number theory are taken a little further; there is a chapter on Euclid's algorithm and unique factorization and another on congruences involving zero divisors and Fermat's theorem. Each chapter has a set of questions for which there are a few answers, and the book ends with some general appendices.

This is not a book that could be used where the standard school course is a modern one and the treatment is not attractive enough to commend the attention of the average sixth former who has done 'trad' mathematics.

And finally, at 75*s*. for 190 pages this book prices itself right out of the market.

St Paul's School, London, S.W.13 HUGH NEILL

Mathematics on Vacation. By JOSEPH S. MADACHY. Thomas Nelson, London, 1968. 35*s*.

This book has been compiled by Joseph S. Madachy, sometime editor of *Recreational Mathematics Magazine*. The book begins with a chapter on geometrical dissections, and continues with chapters on chessboard problems, magic squares, number 'recreations' and alphametics, concluding with a chapter on sundry odds and ends.

The first chapter is promising, with a discussion of dissections of various geometric figures and their reassembly into other patterns, but the early promise is not fulfilled. There follows, at excessive length, a chapter on various chessboard problems—'How many queens are needed to attack every square of a chessboard?', which is followed up by the same questions about rooks, knights, bishops, etc. Such questions should have been left to the reader to investigate for himself. This chapter, too, sees the first glimmerings of the author's apparent obsession with large numbers, which reaches its culmination in a rather uninspiring chapter on 'number recreations'. The same comment applies to the discussion of magic squares. It could, for example, have been instructive to show that there *is* only one magic square of order 3 (page 87). This might have helped to give some of the insight into mathematics which the author claims for his book. Not only is there little opportunity for insight into mathematics, there is too little mathematics (a reason for the title of the book?).

I found my interest declining as I became overwhelmed by a stream of unsupported facts. There are interesting sections in this book, but, on the whole, it cannot be recommended.

University of Nottingham J. A. ANDERSON

Notes on Contributors

Margaret Hayman teaches at Mayfield School, Putney. She is involved in research into methods of examining, and is also a member of the teaching committee of the Mathematical Association. Her husband is W. K. Hayman FRS, Professor of Mathematics at Imperial College.

George Pólya is Professor Emeritus of Mathematics at Stanford University, California. He is a leading mathematical analyst and has written over 200 papers in research journals, as well as a number of books on the art of mathematical discovery.

Donald Michie is Director of the Experimental Programming Unit and Professor of Machine Intelligence in the University of Edinburgh, Department of Machine Intelligence and Perception. He is Editor of the annual *Machine Intelligence* series of volumes and his personal research is concerned with application of concepts of trial-and-error learning to computer control and problem-solving.

P. Chadwick is Professor of Mathematics in the School of Mathematics and Physics at the University of East Anglia. He is one of the two Executive Editors of the *Quarterly Journal of Mechanics and Applied Mathematics* and has made important contributions to the study of the elastic and plastic deformation of solids.

G. A. Watterson is Reader in the Department of Mathematics at Monash University, Melbourne. He wrote the article in this issue while visiting Sheffield University. His main interest is mathematical genetics, but he has also consulted widely on statistical problems with other scientists, including meteorologists.

R. L. Goodstein, who is Professor of Mathematics in the University of Leicester, is a mathematical logician. He is the author of a long series of books mainly on logic, the foundations of mathematics, and algebra. For a good many years he edited the *Mathematical Gazette.*

Contents

PRICES (*postage included*)

	Individual annual subscription	Bulk annual subscriptions (5 or more copies)	Single issues
U.K. only	6/–	5/– each	3/6 each
Australia	$A1.10	$A0.90	$A0.60
U.S.A.	$1.20	$1.00	$0.70

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

Oxford University Press, Education Department,
Walton Street, Oxford.

or to Oxford University Press Branches Overseas.

Printed in England by John Wright & Sons Ltd. Bristol