Reprinted from *Mathl Gaz.*, 1982, **66**, 1–4

SELF-DESCRIPTIVE STRINGS

Michael D. McKay and Michael S. Waterman

The

Mathematical Gazette

A JOURNAL OF THE MATHEMATICAL ASSOCIATION

Vol. 66 March 1982 No. 435

Self-descriptive strings

MICHAEL D. McKAY AND MICHAEL S. WATERMAN

Take a look at the 'string' (or finite sequence)

6, 2, 1, 0, 0, 0, 1, 0, 0, 0.

Count the occurrences of the numbers:

Number	0	1	2	3	4	5	6	7	8	9
Occurrence	6	2	1	0	0	0	1	0	0	0

Notice that the bottom row is the string we started with! In fact the above string is the unique answer to the popular question (posed, for example, in a recent edition of the Trans World Airlines magazine, *The Ambassador*, August 1978) to find a string

$$a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9$$

of integers such that, for $0 \leqslant i \leqslant 9$, a_i equals the number of occurrences of i in the string. There is, of course, nothing special about the number 9. Other examples of such strings are

2, 0, 2, 0

1, 2, 1, 0

and 2, 1, 2, 0, 0.

In general we'll call

$$a_0, a_1, a_2, \ldots, a_{N-1}$$

a *self-descriptive string* if, for $0 \leqslant i \leqslant N-1$, i occurs precisely a_i times in the string. We leave it to the reader to show that there are no self-descriptive strings for $N \leqslant 3$. It is also straightforward (but tedious) to show that the above strings are the only self-descriptive ones for $N = 4$ or 5. To make the work easier one can answer the following question first:

Question 1 If $a_0, a_1, \ldots, a_{N-1}$ is a self-descriptive string, then what are the following sums:

(a) $a_0 + a_1 + \cdots + a_{N-1}$

(b) $a_1 + 2a_2 + \cdots + (N-1)a_{N-1}$?

Answer Both sums are N. For the total number of occurrences is $a_0 + a_1 + \cdots a_{N-1}$ and is N. Similarly there are a_0 occurrences of 0, a_1 occurrences of 1, a_2 occurrences of 2, etc, and so their total contribution to this sum of N is $a_1 + 2a_2 + \cdots + (N-1)a_{N-1}$.

So, for example, in the case $N = 3$ we are looking for a_0, a_1, a_2 with

$$a_0 + a_1 + a_2 = 3 \quad \text{and} \quad a_1 + 2a_2 = 3.$$

The only non-negative integral solutions are 1, 1, 1 and 0, 3, 0; neither of which is self-descriptive.

Now, by a sequence of questions (suitable, we hope, as class-room exercises) we are going to investigate the self-descriptive strings for $N \geqslant 6$. So for the remainder of this article

$$a_0, a_1, a_2, \ldots, a_{N-1}$$

is self-descriptive and $N \geqslant 6$.

Question 2 Let $n = [N/2]$ (so that, for example, $[7/2] = 3$ and $[8/2] = 4$). What is the sum

$$a_n + a_{n+1} + \cdots + a_{N-1}?$$

Answer Just 1. For if this sum is at least 2, then at least 2 of the terms in the string must be chosen from $n, n+1, \ldots, N-1$. Also, to keep the total $a_1 + 2a_2 + \cdots + (N-1)a_{N-1}$ equal to N these two terms of at least n must be a_0, a_1 or a_2. But then

$$N = a_0 + a_1 + \cdots + a_{N-1} \geqslant n + n + a_n + a_{n+1} + \cdots + a_{N-1} \geqslant 2n + 2 > N$$

which is a contradiction.

On the other hand, if $a_n + a_{n+1} + \cdots + a_{N-1} = 0$, then each of $a_n, a_{n+1}, \ldots, a_{N-1}$ is 0 and so a_0 (= number of zeros in the string) is at least n and $a_{a_0} \neq 0$, which gives another contradiction

$$0 = a_n + a_{n+1} + \cdots + a_{N-1} \geqslant a_{a_0} > 0.$$

It is clear from that answer that all but one of $a_n, a_{n+1}, \ldots, a_{N-1}$ is zero and that the remaining one is 1. Bearing in mind that one of $a_n, a_{n+1}, \ldots, a_{N-1}$ is 1, answer the following question:

Question 3 What is the smallest possible value of a_1?

Answer The smallest possibility is 2. For there is at least one occurrence of 1 (namely amongst $a_n, a_{n+1}, \ldots, a_{N-1}$ as above) so $a_1 \geqslant 1$. But then if $a_1 = 1$ it follows that a_1 is *itself* a second occurrence of 1! So $a_1 \geqslant 2$.

(We'll now let n stand for $[N/2]$ throughout.) The number $a_n + a_{n+1} + \cdots + a_{N-1}$ represents the total number of occurrences of $n, n+1, \ldots, N-1$. So, by Question 2, there is precisely one occurrence of a number n or more.

Question 4 Which of $a_0, a_1, \ldots, a_{N-1}$ is at least n?

Answer $a_0 \geqslant n$ and $a_1, a_2, \ldots, a_{N-1} < n$. For if $a_j \geqslant n$ and $a_0, a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_{N-1} < n$, where $j > 0$, then there are a_j occurrences of j in the string. So (as only one of $a_n, a_{n+1}, \ldots, a_{N-1}$ is non-zero) there are at least $a_j - 1$ ($\geqslant n-1$) occurrences of j amongst $a_0, a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_{n-1}$. But there are only $n-1$ terms here altogether! So they are all j and there is at least one j amongst $a_n, \ldots, a_{N-1}$. Since $a_n + \cdots + a_{N-1} = 1$ we know that $j = 1$ and that there is exactly one j amongst $a_n, \ldots, a_{N-1}$. This means that $a_n = 1$ and that the string is

a_0	a_1	a_2	$\cdots$	a_{j-1}	a_j	a_{j+1}	$\cdots$	a_{n-1}	a_n	a_{n+1}	$\cdots$	a_{N-1}
1	1	1		1	n	1		1	1	0		0

But then $1 = a_0 =$ occurrences of $0 = N - n - 1$: but $N - n = 2$ is impossible for $N \geqslant 6$.

Question 5 What's the least number of non-zero terms in the string $a_0, a_1, \ldots, a_{N-1}$?

Answer There are at least 4 non-zero terms. For $a_0 \geqslant n$, one of $a_n, a_{n+1}, \ldots, a_{N-1}$ is non-zero, $a_1 \geqslant 2$ and $a_{a_1} \geqslant 1$ (where $1 < a_1 < n$ from the previous exercises).

Question 6 What's the largest number of non-zero terms in the string $a_0, a_1, \ldots, a_{N-1}$?

Answer Again the answer is 4! For assume that there are r non-zero terms. Then $a_0 = N - r$ and $a_{N-r} = 1$ (the one non-zero term amongst

$a_n, a_{n+1}, \ldots, a_{N-1}$). Hence, as the r non-zero terms include a_0 and a_{N-r},

$$\begin{aligned} N &= a_1 + 2a_2 + \cdots + (N-1)a_{N-1} \\ &\geqslant (1 + 2 + \cdots + (r-2)) + (N-r) \\ &= \tfrac{1}{2}(r^2 - 5r + 2) + N. \end{aligned}$$

Therefore $r^2 - 5r + 2 \leqslant 0$, which means that $r \leqslant 4$ as claimed.

So there are *precisely* 4 non-zero terms in our self-descriptive string. It follows that $a_0 = N - 4$. We've also seen that $a_0 \geqslant n$. Hence $n \leqslant N - 4$ and so $N \geqslant 7$. So, for a start, there are no self-descriptive strings for $N = 6$. But if $N \geqslant 7$ the 4 non-zero terms are

$$a_0 = N - 4, \quad a_1\ (\geqslant 2), \quad a_{a_1} \quad \text{and} \quad a_{N-4} = 1.$$

Hence (their sum being N) we have that $a_1 + a_{a_1} = 3$ and the only solution is $a_1 = 2$ and $a_{a_1} = 1$. So the unique self-descriptive string for $N \geqslant 7$ is as in the second row of our final table:

Number (i)	0	1	2	3	4	...	$N-5$	$N-4$	$N-3$	$N-2$	$N-1$
Occurrence (a_i)	$N-4$	2	1	0	0	...	0	1	0	0	0

We wish to thank K. Campbell for her useful comments.

MICHAEL D. McKAY
and MICHAEL S. WATERMAN

Los Alamos National Laboratory, New Mexico 87545, U.S.A.

Some connected problems on probability

FRANK BUDDEN

The purpose of this article is threefold. First to suggest other approaches to the 'Glass rod problem' (*Gazette* note **65.6**, March 1981). Second, to show the unexpected connection between that and the article 'Triangle of triangles' by H. B. Griffiths on page 10 of the same issue and with the article on 'Random triangles' by T. Easingwood in the December 1981 edition. And third, to explore further problems which arise from a consideration of the above.

First to remind the reader of the problems under investigation:

Problem I A rod is cut at random into three pieces. What is the probability that the three pieces may be reassembled to form a triangle? (G. Haigh's 'Glass rod' problem.)