The

Mathematical Gazette

A JOURNAL OF THE MATHEMATICAL ASSOCIATION

Vol. 70 March 1986 No. 451

Co-descriptive strings

LEE SALLOWS AND VICTOR L. EIJKHOUT

An appetiser

What sentence in this *Mathematical Gazette* contains seven a's, four c's, thirty-two e's, eight f's, five g's, ten h's, twelve i's, three l's, three m's, fifteen n's, eight o's, one q, seven r's, twenty-seven s's, twenty-five t's, four u's, eight v's, seven w's, four y's & two z's?

Introduction

A popular problem in the puzzle literature of late concerns construction of sentences which tabulate their own digits:

> In this sentence there is 1 '0', 11 '1's, 2 '2's, 1 '3', 1 '4', 1 '5', 1 '6', 1 '7', 1 '8' and 1 '9'.

A mathematical weakness of these curiosities is their dependence on the representational form employed: solutions are not invariant under translation into different number base systems or other notations.

Recently, however, in the *Gazette* McKay and Waterman [1] have investigated *self-descriptive strings* such as '6, 2, 1, 0, 0, 0, 1, 0, 0, 0', which describe themselves in the sense of being self-enumerators:

integer	0	1	2	3	4	5	6	7	8	9
number of occurrences	6	2	1	0	0	0	1	0	0	0

The tabulation of occurrences is seen to be identical to the string itself. Here string structure is independent of notation, being entirely determined by the properties of numbers rather than the digits in which they are expressed. The authors go on to give a general formula for self-descriptive strings of arbitrary length N ($N \geqslant 7$) and prove its uniqueness:

integer	0	1	2	3	4	5	...	$N-5$	$N-4$	$N-3$	$N-2$	$N-1$
occurrences	$N-4$	2	1	0	0	0	...	0	1	0	0	0

In this article we shall investigate closed cycles of strings in which, rather than describing itself, each string enumerates its predecessor in the loop until the last is enumerated by the first. The simplest such case is a loop of length 2, an illustration in sentential style being:

> The following sentence uses 1 '0', 7 '1's, 4 '2's, 1 '3', 1 '4', 1 '5', 1 '6', 1 '7', 2 '8's and 1 '9'.
>
> The previous sentence uses 1 '0', 8 '1's, 2 '2's, 1 '3', 2 '4's, 1 '5', 1 '6', 2 '7's, 1 '8' and 1 '9'.

The following couplet of *co-descriptive* strings, in which each describes the other, is analogous to the self-descriptive string of McKay and Waterman:

6, 3, 0, 0, 0, 0, 0, 1, 0, 0
7, 1, 0, 1, 0, 0, 1, 0, 0, 0

and it too can be generalised to arbitrary length N ($N \geqslant 8$) (again unique, as we shall see later):

0	1	2	3	4	5	...	$N-5$	$N-4$	$N-3$	$N-2$	$N-1$
$N-4$	3	0	0	0	0	...	0	0	1	0	0
$N-3$	1	0	1	0	0	...	0	1	0	0	0

The reader may now be wondering if there exist triplets, quadruplets, etc. of such sequentially-enumerating strings. Surprisingly, for $N \geqslant 8$ the answer is negative: the above formula embodies the only existing co-descriptive cycle. This is the main proposition we intend to prove here.

The iteration technique

We begin by developing a simple technique for deriving strings and formulae like those above. Let S_0 be any given string of numbers of a certain length N, and consider the sequence $S_0, S_1, S_2, \ldots$ where each successive term is a string of length N corresponding to an enumeration of the occurrences of $0, 1, \ldots, N-2, N-1$, respectively, in the immediately

preceding term. Hence, if $N = 7$ and S_0 is formed from the first seven digits of π, say, then we have:

	0	1	2	3	4	5	6
S_0	3	1	4	1	5	9	2
S_1	0	2	1	1	1	1	0
S_2	2	4	1	0	0	0	0
S_3	4	1	1	0	1	0	0
⋮				⋮			

Notice that for all N, whatever S_0 is selected, S_1 must always comprise N numbers $\leqslant N$, so that the sum of the elements in S_2 (i.e. the total number of numbers in S_1) as well as in all subsequent terms equals N. In short, the integers appearing in S_i, $i > 1$, form partitions of N. Furthermore, strings built up from different orderings of the same N numbers will all give rise to an identical successor. However, since the number of distinct partitions of N is finite, it is clear that the series cannot extend indefinitely without repetition of some term. In other words, every such series must loop. Continuing the above, for instance, gives a loop of length 3 as shown; i.e. a co-descriptive triplet:

$S_3 = 4$	1	1	0	1	0	0
$S_4 = 3$	3	0	0	1	0	0
$S_5 = 4$	1	0	2	0	0	0
$S_6 = 4$	1	1	0	1	0	0
$S_7 = 3$	3	0	0	1	0	0
$S_8 = 4$	1	0	2	0	0	0
$S_9 = 4$	1	1	0	1	0	0
⋮			⋮			

Now for small N we can investigate all possible repeating loops by setting S_0 to each distinct partition of N in turn and extending the sequence until repetition occurs. Hence for $N = 4$ we would try $S_0 = 4, 0, 0, 0$; 3, 1, 0, 0; 2, 2, 0, 0; 2, 1, 1, 0 and 1, 1, 1, 1, in turn, the order of the numbers in each not being critical. This pencil and paper exercise soon reveals that besides loops of length 1, the only co-descriptive cycles for $N < 8$ are a couplet ($N = 6$) and the triplet ($N = 7$):

length	string	type
$N = 4$	1 2 1 0	self-descriptive
$N = 4$	2 0 2 0	self-descriptive
$N = 5$	2 1 2 0 0	self-descriptive
$N = 6$	3 1 1 1 0 0	co-descriptive
	2 3 0 1 0 0	
$N = 7$	3 2 1 1 0 0 0	self-descriptive
	3 3 0 0 1 0 0	
$N = 7$	4 1 0 2 0 0 0	co-descriptive
	4 1 1 0 1 0 0	

Similar exhaustive searches for N larger than 7 are tedious and thus better computer-generated, but soon reveal an overall characteristic: after a few iterations strings become composed mainly of zeros, with non-zero elements being confined to the first four or so positions except for a single 1 occurring near the end. This recurrent pattern soon suggests the idea of applying iteration to a generalised string in which the middle terms are all assumed to be zeros. Choosing S_0 to represent some simple partition of the string of unspecified length N, such as $N-1$, 1, 0, 0, . . ., 0, we proceed as usual:

	0	1	2	3	4	5	...	$N-5$	$N-4$	$N-3$	$N-2$	$N-1$
S_0	$N-1$	1	0	0	0	0	...	0	0	0	0	0
S_1	$N-2$	1	0	0	0	0	...	0	0	0	0	1
S_2	$N-3$	2	0	0	0	0	...	0	0	0	1	0
S_3	$N-3$	1	1	0	0	0	...	0	0	1	0	0
S_4	$N-4$	3	0	0	0	0	...	0	0	1	0	0
S_5	$N-3$	1	0	1	0	0	...	0	1	0	0	0
S_6	$N-4$	3	0	0	0	0	...	0	0	1	0	0
S_7	$N-3$	1	0	1	0	0	...	0	1	0	0	0
⋮							⋮					

The result of this tactic is that the pair of terms emerging in the loop of length 2 encountered corresponds to that appearing in the co-descriptive formula quoted earlier. We see then that the iterative method can be used to generate both individual strings and general formulae.

Readers may care to repeat the above using different initial partitions; a few trials will suffice to suggest the truth of our unexpected finding that for $N \geqslant 8$, all series terminate in the self-descriptive string (which we shall call S_{I}) or the co-descriptive pair above (which we shall call S_{II}).

Proof of uniqueness

To prove our assertion that for all string lengths $N \geqslant 8$ the only existing loops are S_{I} and S_{II}, we focus on the changes in 'nz_i', the number of non-zero elements in the i-th string of the iteration process. We show that if, for some $i \geqslant 5$, $nz_{i+1} > nz_i$ holds, then the string S_{i+4} is in S_{II}; if however $nz_{i+1} = nz_i$, then either S_{i+2} or S_{i+4} is S_{I}, or S_{i+3} is in S_{II}. The proof is then complete since for a loop other than S_{I} or S_{II} to occur, all strings composing it would have to satisfy $nz_{i+1} < nz_i$, which is clearly impossible.

We begin by establishing the general structure of S_i ($N \geqslant 8$) as already described informally in the previous section. Denoting the i-th string in the iteration by

$$S_i = a_0^{(i)}, a_1^{(i)}, a_2^{(i)}, \ldots, a_{N-1}^{(i)},$$

recall that for $i \geqslant 2$ the sum of the coefficients equals N, i.e.

$$\sum_{j=0}^{N-1} a_j^{(i)} = N;$$

and that since $a_j^{(i)}$ is by definition equal to the number of occurrences of j in S_{i+1}, the above equation implies that for $i \geqslant 3$,

$$\sum_{j=0}^{N-1} j a_j^{(i)} = N.$$

Three important properties of all S_i, $i \geqslant 5$, can now be identified. The techniques needed are not unlike those used by McKay and Waterman, and are left as exercises for the interested reader:

(1) At least half of the coefficients are zero.
(2) The first coefficient is at least $[N/2]$.
(3) The only non-zero element in the latter half of S_i is a single 1.

We now follow the proof outlined above by investigating how a repeated non-zero number in some S_i influences nz_{i+1}, the number of non-zeros in S_{i+1}. Notice first that a non-zero number occurring just once in S_i contributes a 1 in S_{i+1} but all occurrences of a repeated number in S_i are replaced by a single non-zero integer in S_{i+1}. For example

$$S_i = ⑥\ 2\ 3\ \boxed{1}\ 0\ 0\ 0\ \boxed{1}\ 0\ 0$$

$$S_{i+1} = 5\ \ 2\ \ 1\ \ 1\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0.$$

It follows that, if there are no repeated non-zeros in S_i, then $nz_{i+1} = nz_i + 1$ and, apart from $a_0^{(i+1)}$, all the coefficients in S_{i+1} are 0's or 1's. An instance of this would be:

$$S_i = 6\ 0\ 3\ 2\ 0\ 0\ 0\ 1\ 0\ 0 \qquad nz_i = 4$$

$$S_{i+1} = 6\ 1\ 1\ 1\ 0\ 0\ 1\ 0\ 0\ 0 \qquad nz_i = 5$$

It is also clear that the absence of repeated non-zero coefficients is the *only* condition under which $nz_{i+1} > nz_i$.

So as soon as $nz_{i+1} > nz_i$ the following iterations yield

$S_i = \ldots$ (nz_i distinct non-zeros) . . . (length $N \geqslant 8$)

$S_{i+1} = \alpha \ldots$ (nz_i 1's, rest 0's) . . .
↑
($\neq 0$ or 1)

$S_{i+2} = \beta\,\gamma\,0 \ldots 0\ \ 1\ \ 0 \ldots 0$
($\neq 0$ or 1) (Because of the α in S_{i+1})

We can deduce $\beta \neq \gamma$ from the three general properties of strings listed above. Then

$S_{i+3} = N-3$	1	0...0	1	0...0	1	0...0				
$S_{i+4} = N-4$	3	0..............0			1	0	0			
$S_{i+5} = N-3$	1	0	1	0......0	1	0	0	0		
$S_{i+6} = N-4$	3	0..............0			1	0	0			

and we have reached S_{II}.

The reader can show similarly that $nz_{i+1} = nz_i$ if and only if exactly two identical non-zero numbers appear in S_i. Two or more 1's then occurring in S_{i+1} will eventually lead to S_I, a single 1 in S_{i+1} implies S_{i+3} and S_{i+4} constitute S_{II}; all this is left as an exercise.

Since it is impossible that all strings in a loop satisfy $nz_i > nz_{i+1}$, it follows that for some i, $nz_i \leqslant nz_{i+1}$ must hold, and as this has been covered by the above cases we have shown that S_I and S_{II} are the only self- or co-descriptive loops reached for $N \geqslant 8$.

Concluding remarks

Besides co-descriptive strings, the iterative method here introduced can be used to produce a variety of related loop phenomena. The sentential counterparts of co-descriptive strings, for instance, differ significantly from the latter only in the extra appearance of each number listed. As such they can be easily generated by adding 1 to the true enumeration at every step in the iteration. Starting again with the first seven digits of π, for example:

	0	1	2	3	4	5	6	
S_0	3	1	4	1	5	9	2	
S_1	1	3	2	2	2	2	1	(add 1 to each coefficient)
S_2	1	3	5	2	1	1	1	(add 1 to each coefficient)
⋮				⋮				⋮

soon generates a sentential analogue of the co-descriptive triplet ($N = 7$):

The second sentence employs 1 '0', 4 '1's, 4 '2's, 1 '3', 1 '4', 2 '5's and 1 '6'.
The third sentence employs 1 '0', 5 '1's, 2 '2's, 2 '3's, 1 '4', 2 '5's and 1 '6'.
The first sentence employs 1 '0', 5 '1's, 2 '2's, 1 '3', 3 '4's, 1 '5' and 1 '6'.

Whimsical extension of this idea soon leads us to unsuspected literary canons:

In general, in a self-descriptive sentence there are $(N-3)$ '1's, 3 '2's, 2 '3's, 1 '4', 1 '5', ..., 1 '$(N-5)$', 1 '$(N-4)$', 2 '$(N-3)$'s, 1 '$(N-2)$', 1 '$(N-1)$' and 1 'N'.

A similar parallel to S_{II} can be derived of course; N is assumed to be more than 8.

Making a further departure from the rules governing co-descriptive strings, the number(s) added to the enumeration at each step in the iteration may vary. With a little ingenuity, loops of any desired length can then be produced for all N. Here is an example with loop length 5:

(1) Sentence 2 uses 1 '0', 6 '1's, 4 '2's, 2 '3's, 2 '4's, 1 '5', 2 '6's, 1 '7', 1 '8' and 1 '9'.
(2) Sentence 3 uses 1 '0', 6 '1's, 4 '2's, 2 '3's, 2 '4's, 1 '5', 1 '6', 2 '7's, 1 '8' and 1 '9'.
(3) Sentence 4 uses 1 '0', 7 '1's, 3 '2's, 2 '3's, 1 '4', 2 '5's, 1 '6', 1 '7', 2 '8's and 1 '9'.
(4) Sentence 5 uses 1 '0', 8 '1's, 3 '2's, 1 '3', 1 '4', 2 '5's, 2 '6's, 1 '7', 1 '8' and 1 '9'.
(5) Sentence 1 uses 1 '0', 6 '1's, 5 '2's, 1 '3', 2 '4's, 1 '5', 2 '6's, 1 '7', 1 '8' and 1 '9'.

In this illustration the first sentence refers to the second, the second to the third, etc., and the last to the first. Different orderings could be used however, and each will result in its distinct set of sentences. We leave the exact mechanics of their generation for interested readers to unravel.

Another type of cycle deserving of mention here is what we call a *Fibonacci loop*. In these we begin with a pair of strings S_0 and S_1 and develop the series by enumerating both S_{n-1} and S_{n-2} in S_n:

	0	1	2	3	4	5	6
S_0	6	3	2	2	0	1	0
S_1	4	3	3	1	1	0	2
S_2	3	3	3	3	1	0	1
S_3	2	4	1	6	1	0	0
S_4	3	4	1	4	1	0	1
⋮				⋮			

A loop occurs when any *consecutive pair* of terms is repeated:

S_5	3	5	1	1	3	0	1
S_6	2	6	0	3	2	1	0
S_7	3	4	2	3	0	1	1
S_8	3	3	3	3	1	0	1
S_9	2	4	1	6	1	0	0
⋮				⋮			

Gleanings far and near

Monkeying about with fractions

"The chimpanzee appeared to be able to count up to four objects with accuracy. When the number was between five and six the number of mistakes increased." From *The Times* of 6 May 1985, sent in by John Backhouse.

In this example for $N = 7$, $(S_8, S_9) = (S_2, S_3)$; a Fibonacci loop of length 6. Again, a special case of this kind of cycle is when loop length shrinks to its minimum; in a Fibonacci loop this is necessarily a length of 3:

	0	1	2	3	4	5	6	7	8	9
F_0	6	7	3	0	0	1	1	1	0	1
F_1	7	6	2	1	0	1	1	1	0	1
F_2	5	9	1	1	0	0	2	2	0	0

Here $(F_3, F_4) = (F_0, F_1)$ and thus each string is automatically an enumeration of its two companion strings. Still further special instances of a length 3 loop occur when two of the strings are identical:

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

or even all three:

0	1	2	3	4	5	6
6	0	4	0	2	0	2
6	0	4	0	2	0	2
6	0	4	0	2	0	2

As before, sentential counterparts of such loops may also be formed:

> The following pair of sentences employ 2 '0's, 2 '1's, 9 '2's, 5 '3's, 5 '4's, 4 '5's, 5 '6's, 2 '7's, 3 '8's and 3 '9's.
> The sentences above and below employ 2 '0's, 2 '1's, 8 '2's, 6 '3's, 5 '4's, 6 '5's, 3 '6's, 2 '7's, 2 '8's and 4 '9's.
> The previous pair of sentences employ 2 '0's, 2 '1's, 9 '2's, 5 '3's, 4 '4's, 6 '5's, 4 '6's, 2 '7's, 3 '8's and 3 '9's.

It is a matter of regret that the number of distinct Fibonacci loops is strictly finite since as string length grows the number of zeros to be enumerated rapidly exceeds $N - 1$, thus violating the natural closure condition $0 \leqslant a_j^{(i)} \leqslant N - 1$. A general formula for these loops therefore does not exist.

Come clean!

"Last year Persil celebrated 75 years, and Hotpoint over 60 years, of caring for your family wash. That's over 135 years of expertise." From a Persil packet, sent in by Ken McKelvie.

Imperial jewel

"The metric equivalent of the calorie is the kilojoule." From *Cooking for slimmers* by Carol Bowen, and spotted by Brian Head.

Starting from all possible pairs of partitions of $2N$, computer-generated iterations have identified the only 23 existing Fibonacci loops.

0	1	2	3	4	5	6	7	8	9	10	11	12	N	loop length
4	0	2	0	4									5	3
4	0	2	0	4										
3	0	3	2	2									5	3
2	0	4	4	0										
4	0	3	2	3	0								6	3
4	0	2	4	2	0									
6	0	4	0	2	0	2							7	3
6	0	4	0	2	0	2								
2	4	4	2	0	2	0							7	3
3	2	5	1	2	1	0								
3	3	3	3	1	0	1							7	6
2	4	1	6	1	0	0								
4	2	5	1	3	1	0	0						8	3
4	4	2	2	2	2	0	0							
3	6	2	1	2	1	1	0						8	16
3	5	3	2	1	1	1	0							
5	3	4	0	1	1	2	0						8	35
4	5	2	1	1	2	1	0							
4	6	1	1	1	1	2	0						8	37
3	7	1	1	2	0	2	0							
5	4	2	3	2	1	1	0	0					9	46
4	5	2	3	2	2	0	0	0						
6	6	2	2	0	2	0	2	0	0				10	3
7	3	5	1	0	1	2	1	0	0					
6	7	3	0	0	1	1	1	0	1				10	3
7	6	2	1	0	1	1	1	0	1					
6	4	3	4	1	0	0	2	0	0				10	3
7	3	3	2	3	0	1	1	0	0					
6	4	4	4	0	0	0	0	2	0				10	3
8	2	3	2	3	0	1	0	1	0					
5	7	2	1	2	1	1	1	0	0				10	31
5	7	3	0	1	2	1	1	0	0					
8	4	3	4	1	0	0	0	0	2	0			11	3
9	3	3	2	3	0	0	0	1	1	0				
6	10	0	2	0	2	0	0	0	2	0			11	3
9	5	3	1	0	1	1	0	0	1	1				
8	4	4	4	0	0	0	0	0	0	2			11	3
10	2	3	2	3	0	0	0	1	0	1				
9	5	3	1	0	0	0	2	2	0	0			11	6
10	3	3	2	0	1	0	1	1	1	0				
11	5	3	1	0	1	0	0	1	0	1	1		12	3
8	10	0	2	0	2	0	0	0	0	0	2			
10	4	3	4	1	0	0	0	0	0	0	2		12	3
11	3	3	2	3	0	0	0	0	0	1	1			
8	6	4	4	0	0	0	0	0	0	0	2		12	3
11	3	3	2	2	0	1	0	1	0	0	1			

The complete loops are readily reconstructed from the consecutive pairs given in the table.

Finally, it is worth noting that although numerals replace number-names in all the above sentential forms, this is merely for abbreviation. Written out in full, a self-enumerating sentence, for instance, can be seen as partially enumerating its own *words*. From this vantage we can see better how the creation of a sentence enumerating *all* its own words entails only the familiar iteration process in combination with the addition of appropriate constants, contingent upon the initial text selected:

> This self-descriptive sentence employs two 'this's, two 'self-descriptive's, two 'sentence's, two 'employs's, one 'zero', seven 'one's, eight 'two's, one 'three', one 'four', one 'five', one 'six', two 'seven's, two 'eight's, one 'nine' and two 'and's.

Co-descriptive word-enumerators are of course similarly constructable. Indeed, embodied in a computer program, the iterative method has proved itself a powerful technique in producing self-descriptive patterns of a far more convoluted kind [2], [3], [4]:

> This sentence consists of two a's, three c's, two d's, twenty-five e's, six f's, two g's, six h's, thirteen i's sixteen n's, nine o's, five r's, twenty-eight s's, nineteen t's, two u's, five v's, seven w's, four x's, three y's and one z.

The present article is intended as a tentative step in the direction of a complete mathematical theory of all such self-enumerators.

References

1. Michael D. McKay and Michael S. Waterman, Self-descriptive strings, *Math. Gaz.* **66**, 1–4 (1982).
2. A. K. Dewdney, Computer recreations, *Scient. Amer.* **251**, 18–22 (October 1984).
3. A. K. Dewdney, Computer recreations, *Scient. Amer.* **252** (January 1985), 13.
4. Lee Sallows, In quest of a pangram, *Abacus* **2** (Spring 1985), 22–40.
5. John R. Letaw, Pangrams: a non-deterministic approach, *Abacus* **2** (Spring 1985) 42–47.

LEE SALLOWS
and VICTOR L. EIJKHOUT

Psychology Lab., University of Nijmegen, Montessorilaan 3, Postbus 9104, 6500 HE Nijmegen, The Netherlands

---

A wee dram

"By all the rules the whisky should have gone 'woody'—and indeed it should have evaporated, because a loss of 2 per cent per annum over 88 years should yield less than nothing." From *The century companion to whiskies* by Derek Cooper, sent in by Professor R. A. Rankin.