REFLEXICONS

by Lee Sallows

Reflexiconstruction

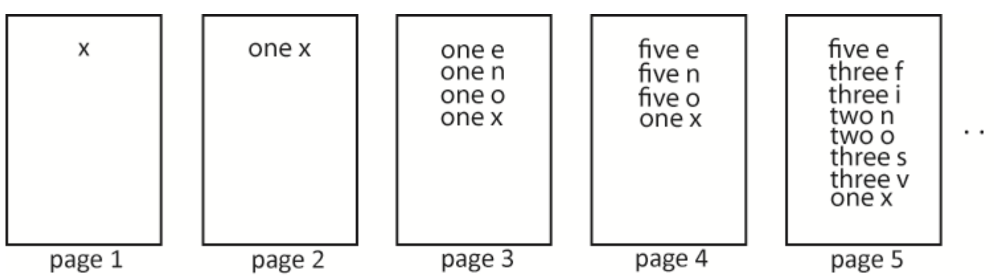
A lexicon is a dictionary or a list of words. Hence my use of "reflexive lexicon", or more crisply, *reflexicon*, for a self-enumerating word list that describes its own letter frequencies, as in Figure 1.

Fig. 1
A French
reflexicon

Immortal verity sans superfluity. Now that is what I call *belles lettres*. Below we shall look at some English examples. The focus here is thus upon self-enumerating *sets* or *lists*, rather than the more familiar self-enumerating *sentences* which may include text inessential to the basic enumeration. But first, since the answer is far from obvious, how are reflexicons produced?

Imagine a book with pages inscribed as follows. The text on page 1 might be anything - a colophon, an epigram, a dedication. For convenience we assume something short. But page 2 and all subsequent pages each comprise a descriptive list, in words, of the numbers of a's, b's, c's, etc., appearing on the previous page. Thus if page 1 were to feature a single *x* then our volume begins as in Figure 2. Here the letter counts are listed alphabetically, but any order will suffice.

Fig. 2



This approach may not be the recipe for a bestseller but the plot does have its appeal. You can hardly help wondering how it ends! Will list lengths continue to expand? Clearly 26 items, the number of letters in the alphabet, is the absolute limit. But in fact here none will exceed 16. These will be totals for E,F,G,H,I,L,N,O,R,S,T,U,V,W,X and Y, which are the sixteen letters occurring in English number words under *one hundred*, a number much higher than feasible list entries, assuming brevity in the opening text. Our example is therefore a *lipogram*, a work in which certain letters, namely A,B,C,D,J,K,M,P,Q and Z, will be absent because missing from page 1, the only page on which they could first occur. The conclusion of our volume can now be discerned.

Every new page produced shows a list of at most 16 totals, none of them large. The possible variations are thus finite. As new pages are added, sooner or later the set of entries appearing on one page will recur on another. Suppose the totals on page N are the same as those on an earlier page M. Then page N is an *anagram* of page M, meaning their letter frequencies agree. But this entails that page $N + 1$ will be *identical* to page $M + 1$, which shows that our book has wound up in a repetitive cycle: page $N+2$ will be a repeat of page $M + 2$, page $N + 3$ a repeat of page $M + 3$, and so on. And the same thing will happen whatever the starting text on page 1. Call the number of pages occurring in such a closed cycle its *period*. If the period is P, then we have a closed loop of P sequentially descriptive lists: L_1, L_2, .., L_P, in which L_P enumerates L_{P-1}, L_{P-1} enumerates L_{P-2}, L_{P-2} enumerates L_{P-3}, and so on until we reach L_1 which enumerates L_P. So if $P = 3$ they will form a triad of lists, L_1, L_2, L_3, such that L_3 enumerates L_2, L_2 enumerates L_1, and L_1 enumerates L_3. Similarly, if $P = 2$ the two lists will form a mutually enumerating pair L_1 and L_2 . While if $P = 1$, then we have a list whose description is a copy of itself: a reflexicon.

Let distinct letters stand for distinct lists. The onset of a period 1 loop, R, then looks like this: ..,*L*,*M*,*N*,*O*,*P*,*Q*,*R*,*R*,*R*,*R*,.. This shows that the reflexicon R not only describes itself, but it describes list Q, as well. So Q must be an anagram of R, most likely a re-ordering of the same set of totals. Once any of its anagrams turns up, the reflexicon itself follows at once. No reflexicon is reached *except* via one of its anagrams, unless of course we start off with the reflexicon itself on page 1.

Question: Assuming no A/B/C/D/J/K/M/P/Q/Z on page 1, how many different loops are there? Using a computer to generate further pages in a book that starts from a single *x* on page 1 shows that its pages converge on a loop of period 316. By repeating this process, but re-starting each run from different initial texts on page 1 and then extrapolating, we wind up in different loops.

The 24 English loops

A program I wrote that repeatedly extrapolated from randomly generated initial lists on page 1 discovered a total of 24 loops. However, these are in several cases only trivially distinct variants of eachother, such as the five period 1 loops (i.e. reflexicons) shown in Figure 3 in which numbers stand for their corresponding number names. Hence the top row is short for: FOURTEEN E, EIGHT F, .. (no L), .. ONE, W, ONE X, ONE Y.

E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
14	8	4	4	8	0	8	9	5	1	5	5	4	1	1	1
14	8	4	4	8	1	8	9	5	0	5	5	4	1	1	1
14	8	4	4	8	1	8	9	5	1	5	5	4	0	1	1
14	8	4	4	8	1	8	9	5	1	5	5	4	1	0	1
14	8	4	4	8	1	8	9	5	1	5	5	4	1	1	0

Fig. 3
Five reflexicons
from one

See how the 4 ONE's appear variously distributed over the letters L, S, W, X, Y to result in 5 slightly different versions of what is essentially the same reflexicon.

In all, the program found 9 period 1 loops, or reflexicons, the remaining 4 cases recorded in Figure 4 yielding no variants:

E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
5	5	0	0	5	0	0	0	0	0	0	0	5	0	0	0
12	6	0	3	7	2	2	5	5	5	6	3	6	4	4	0
16	5	3	7	6	0	3	4	8	4	8	4	3	0	3	0
20	4	1	5	3	1	10	7	7	3	9	3	4	3	1	2

Fig. 4
Four reflexicons

Figure 5 shows how the 16 all non-zero items in the bottom list can be used to fill a 4×4 square. Note that, if used to create such a square, any zeroes in a list would result in empty cells.

TWENTY E	FOUR F	ONE G	FIVE H
THREE I	ONE L	TEN N	SEVEN O
SEVEN R	THREE S	NINE T	THREE U
FOUR V	THREE W	ONE X	TWO Y

Fig. 5
A self-descriptive square

Here the order of the entries in the square reading from top left to bottom right is alphabetical, although their position has of course no effect on its self-enumerating property. This suggested to me the possibility of so placing the numbers as to create a *magic square,* but alas turns out to be impossible with these numbers.

Figure 6 shows two of the 5 loops of period 2 found by the program. The two lists in each are thus mutually descriptive, the first describing the second and the second describing the first. Here again, a trivially distinct variant is produced by switching the values of L and Y. The two loops are thus essentially the same specimen.

E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
16	5	2	3	6	0	7	7	4	6	6	3	6	3	3	1
17	3	1	5	7	0	5	4	6	8	7	2	4	2	6	1
16	5	2	3	6	1	7	7	4	6	6	3	6	3	3	0
17	3	1	5	7	1	5	4	6	8	7	2	4	2	6	0

Fig. 6
A co-descriptive pair and variant

Figure 7 records the remaining 3 loops of period 2, none of which give rise to variants.

E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
13	9	2	3	7	0	4	7	6	3	5	5	6	2	2	0
14	4	1	4	7	0	6	5	5	5	8	2	5	4	3	0
16	6	2	3	5	2	9	9	6	4	4	5	5	2	2	1
11	6	1	2	9	1	7	8	4	4	7	3	4	5	4	1
19	5	2	6	6	1	9	7	8	3	8	4	3	2	2	1
16	3	3	5	8	1	9	7	4	4	9	2	3	4	3	1

Fig. 7
Three
co-descriptive pairs

As with the self-descriptive square of Figure 5, two of the loops in Figure 7 feature 2×16 non-zero entries. Either can thus be used to fill the cells of a pair of 4×4 co-descriptive squares, as exampled in Figure 8.

SIXTEEN E	THREE F	THREE G	FIVE H
EIGHT I	ONE L	NINE N	SEVEN O
FOUR R	FOUR S	NINE T	TWO U
THREE V	FOUR W	THREE X	ONE Y

NINETEEN E	FIVE F	TWO G	SIX H
SIX I	ONE L	NINE N	SEVEN O
EIGHT R	THREE S	EIGHT T	FOUR U
THREE V	TWO W	TWO X	ONE Y

Fig. 8 A pair of co-descriptive squares

No loops of period 3 were discovered by the program, but two of period 4 were. The latter are instructive in showing how even longer loops can display trivial variants of themselves. Figure 9 shows one such case, while Figure 10 shows its variant in which L and Y have again switched values.

	E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
1	17	4	2	5	9	1	5	4	4	6	7	1	5	2	5	0
2	14	8	1	1	7	1	8	8	4	4	4	4	7	3	2	0
3	15	6	4	5	4	1	7	10	7	3	7	6	3	2	1	0
4	17	6	1	3	5	1	8	6	5	6	6	3	5	2	3	0
1	17	4	2	5	9	1	5	4	4	6	7	1	5	2	5	0

Fig. 9
A loop of period 4

	E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
1	17	4	2	5	9	0	5	4	4	6	7	1	5	2	5	1
2	14	8	1	1	7	0	8	8	4	4	4	4	7	3	2	1
3	15	6	4	5	4	0	7	10	7	3	7	6	3	2	1	1
4	17	6	1	3	5	0	8	6	5	6	6	3	5	2	3	1
1	17	4	2	5	9	0	5	4	4	6	7	1	5	2	5	1

Fig. 10
A variant of Fig. 9

Note that the bottom list shown is a repeat of the first, as should indeed be the case since both are a description of the fourth. I find it convenient to include such a repeat of the first list as an aid while mentally checking a loop for accuracy.

Thus far we have looked at a total of 16 loops with periods of 1, 2 and 4. Besides these however, the 24 loops identified by the program include examples showing periods of 5, 7, 11, 16, 26, 45 and 316. The loop of period 5 is shown in Figure 11.

	E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
1	15	9	3	6	7	0	2	6	8	1	8	5	5	2	1	0
2	14	5	3	4	9	0	7	5	2	4	7	1	4	3	3	0
3	17	7	1	4	4	0	7	7	8	3	6	5	5	2	1	0
4	18	5	2	3	5	0	8	6	4	6	5	3	7	2	2	0
5	14	5	3	5	8	0	3	5	4	4	8	2	5	4	3	0
1	15	9	3	6	7	0	2	6	8	1	8	5	5	2	1	0

Fig. 11
A period 5 loop

Figure 12 shows one of the two loops of period 7 found, the other a near identical copy but again with totals for L and Y exchanged.

	E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
1	16	10	2	4	9	1	4	5	5	3	6	3	7	2	2	0
2	14	5	1	3	6	1	7	7	5	4	8	3	4	4	3	0
3	18	7	2	5	5	1	6	7	8	4	6	5	5	1	2	0
4	15	6	3	3	9	1	6	6	2	5	5	2	7	3	3	0
5	17	5	1	5	8	1	6	4	5	5	8	1	4	3	4	0
6	16	8	3	4	8	1	6	7	5	3	5	4	6	1	2	0
7	15	5	3	5	8	1	5	6	5	5	7	3	4	2	4	0
1	16	10	2	4	9	1	4	5	5	3	6	3	7	2	2	0

Fig. 12
One of the two loops of period 7

Loops with lengths greater than period 7 would be too space-consuming to record here in full. But since any loop is easily re-created simply by extrapolating from any of its constituent lists, it suffices to show such a sample list for each of the remaining cases; see Figure 13.

Period	E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
11	12	7	2	2	8	1	7	9	4	4	5	4	5	4	3	1
16	19	6	2	5	5	1	6	7	7	5	8	4	6	3	2	1
26	19	3	3	8	8	1	6	5	7	4	10	2	2	2	4	1
45	21	4	3	7	7	1	9	5	6	5	8	2	3	1	3	1
316	19	6	1	3	5	2	8	7	5	4	6	3	7	3	2	2

Fig. 13

Randomly chosen lists from loops of period > 7

This completes a survey of the 24 loops identified by a program that mimics the book-creating analogy described, starting repeatedly from a random set of 16 small integers on page 1 and then extrapolating until processing falls into a repetitive cycle. If we count trivially distinct variants as essentially the same loop then the 24 cases reduce to 17.

Recalling that each trial made by the program starts from a randomly generated page 1, then an obvious question that arises is: How many trials are necessary before we can be certain that every existent loop has been visited? To which must be answered, we can never be certain. In practice the program detects several loops within the first few seconds of starting (the discovery of each announced by a bleep). Thereafter, the emergence of new loops becomes increasingly infrequent, eventually slowing to a seeming halt in that no new finds are made even following days of running time. Even then an elusive specimen may suddenly turn up. With this in mind, it is tempting for me to suppose that the 24 specimens here reported account for all existing loops. But the fact remains that, alas, certainty is impossible.

Singular ONE and Plural S

An alternative type of reflexicon not yet mentioned is one in which a plural *s* is appended to list items where appropriate, i.e., when the letter count is greater than one. Clearly even such a small change in content will give rise to an associated set of loops different to those presented above. Figure 14 shows three such specimens, while Figure 15 shows a sample list from each of the 12 loops discovered. Half of these loops are of period 1, the remaining six showing periods of 2, 7, 13, 16, 33 and 39. The complete loops can of course be reconstructed by extrapolating from the sample lists shown. The addition of plural s's is reflected in the increased totals for letter S.

thirteen e's, five f's, two g's, five h's, eight i's, two l's, three n's, six o's, six r's, twenty s's, twelve t's, three u's, four v's, six w's, four x's, two y's.

fifteen e's, seven f's, four g's, six h's, eight i's, four n's, five o's, six r's, eighteen s's, eight t's, four u's, three v's, two w's, three x's.

sixteen e's, five f's, three g's, six h's, nine i's, five n's, four o's, six r's, eighteen s's, eight t's, three u's, three v's, two w's, four x's.

Fig. 14

Three reflexicons using plural s

Incidentally, although the inclusion of plural *s* in reflexicons was a practice I once applauded, with the passage of time my enthusiasm has cooled. They are, I now see, really a relic left over from the days of self-enumerating *sentences*, in which grammar and other literary conventions are rightly observed, but which become irrelevant in the minimalistic world of self-enumerating sets, lists and matrices with which we are here concerned.

Loop no.	period	E	F	G	H	I	L	N	O	R	S	T	U	V	W	X	Y
1	1	16	6	1	3	9	0	9	5	5	16	5	3	4	1	4	0
2	1	16	5	1	3	6	1	9	7	6	20	7	4	4	3	4	2
3	1	18	2	2	2	6	2	12	9	1	17	9	1	5	7	2	1
4	1	13	5	2	5	8	2	3	6	6	20	12	3	4	6	4	2
5	1	15	7	4	6	8	0	4	5	6	18	8	4	3	2	3	0
6	1	16	5	3	6	9	0	5	4	6	18	8	3	3	2	4	0
7	2	18	4	2	5	6	1	9	7	6	19	11	3	3	4	3	2
8	7	16	5	2	5	7	1	7	6	7	20	9	4	3	3	4	2
9	13	16	5	2	5	8	1	6	6	6	20	10	3	3	3	5	2
10	16	25	4	4	8	7	2	7	5	5	16	12	1	6	4	1	2
11	33	19	7	2	2	5	2	10	9	5	19	6	5	7	4	1	1
12	39	15	5	1	2	8	1	14	8	4	15	7	3	3	4	2	2

Fig. 15

Sample lists from the 12 plural-s-using loops.

Note that besides the three examples of Figure 14, three further reflexicons are listed in Figure 15, namely loops 1, 2 and 3. However, each of these include items in which the total involved is 1, which is to say, ONE, a total unlike any other in that the letter (or character) it counts occurs uniquely at that very point in the list and nowhere else. In consequence, it is replaceable by any other random letter (or character) not already present elsewhere in the reflexicon. For example, in the case of loop 1 seen in Figure 15, ONE G could be changed to ONE L, say, with impunity because L is absent from the list. And similarly for many alternatives such as ONE Y, ONE A, ONE $, ONE ?, etc. Since the number of alternative characters we might use is effectively limitless, the reflexicon of loop 1 is revealed as merely one case among a potentially infinite number of trivially distinct specimens that could be substituted instead. People will perhaps differ in their views on this, but mine is that the presence of ONE in a reflexicon is to be seen as a slightly unfortunate impurity made necessary because without it a perfect self-enumeration could not be achieved.

NINE O	SEVEN I	TWO G	TWO H
TWO L	ONE FOX	FIVE F	TWELVE E
TWO R	SIX N	EIGHT T	FOUR X
SEVEN W	SIX V	FIVE S	TWO U

Fig. 16

A self-enumerating magic square

Nevertheless, the divergent character of ONE can have its uses. On page 3, we looked at a reflexicon composed of 16 non-zero totals and thus able to fill the cells of a 4×4 grid; see Figure 5. I commented that the 16 numbers used were unable to form a *magic* square. In Figure 16, however, a cunning fox has contributed three extra letters to result in a new set of totals that *are* able to form a magic square. Each row, column and long diagonal sums to 20, and every letter used is correctly counted. Moreover, the fox itself is correctly counted; there is just one of him, as stated. But how could the author have known that the addition of F, O and X would do the trick? And what if the letters required had been different? Trade secrets I'm afraid, you'll have to ask the fox.

Self-enumerating pangrams

A further special type of reflexicon is one that includes every letter of the alphabet at least once and is thus a *pangram*. The following example was found in 1998 by Gilles Esposito-Farese, in collaboration with Éric Angelini and Nicolas Graner.

> ONE A, ONE B, ONE C, ONE D, TWENTYEIGHT E, SEVEN F, FIVE G, FIVE H, EIGHT I, ONE J, ONE K, ONE L, ONE M, EIGHTEEN N, EIGHTEEN O, ONE P, ONE Q, FOUR R, TWO S, TEN T, FOUR U, FIVE V, FOUR W, ONE X, TWO Y, ONE Z

The following two co-descriptive pairs are among other loops found that Greg Ross kindly included on his ever engaging weblog [Futility Closet](). The second uses plural S's.

ONE A, ONE B, ONE C, ONE D, THIRTYONE E, FOUR F, ONE G, FIVE H, FIVE I, ONE J, ONE K, ONE L, ONE M, TWENTYTWO N, SEVENTEEN O, ONE P, ONE Q, SEVEN R, FOUR S, ELEVEN T, THREE U, FIVE V, FOUR W, ONE X, THREE Y, ONE Z.

ONE A, ONE B, ONE C, ONE D, THIRTYTWO E, SEVEN F, ONE G, FOUR H, FIVE I, ONE J, ONE K, TWO L, ONE M, TWENTY N, NINETEEN O, ONE P, ONE Q, SEVEN R, THREE S, NINE T, FOUR U, SEVEN V, THREE W, ONE X, THREE Y, ONE Z .

ONE A, ONE B, ONE C, ONE D, TWENTYSEVEN ES, SIX FS, ONE G, THREE HS, SIX IS, ONE L, TWENTY NS, SIXTEEN OS, ONE P, ONE Q, SIX RS, NINETEEN SS, TWELVE TS, FOUR US, FOUR VS, FIVE WS, THREE XS, FOUR YS, ONE Z.

ONE A, ONE B, ONE C, ONE D, TWENTYNINE ES, FIVE FS, ONE G, THREE HS, SEVEN IS, ONE J, ONE K, TWO LS, ONE M, TWENTY NS, SIXTEEN OS, ONE P, ONE Q, SIX RS, TWENTY SS, TEN TS, FOUR US, FOUR VS, FOUR WS, FIVE XS, THREE YS, ONE Z.

A further period 2 loop found is shown here with numerals standing for number names.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	1	1	25	4	3	5	8	1	1	1	1	19	19	1	1	5	3	13	3	2	5	3	3	1
1	1	1	1	36	6	2	8	9	1	1	1	1	20	14	1	1	8	1	14	2	5	3	1	2	1

Two loops of length 3 were discovered. The first list is enumerated by the second, the second by the third and the third by the first. Again, the first loop is composed of minimal pangrams, the second, pangrams with plural S's: .

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	1	1	31	5	1	5	9	1	1	1	1	20	16	1	1	5	5	11	1	4	3	4	2	1
1	1	1	1	28	7	1	3	8	1	1	2	1	20	18	1	1	5	2	8	3	6	3	2	3	1
1	1	1	1	31	2	5	9	7	1	1	1	1	16	15	1	1	5	3	16	1	3	6	2	3	1

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	1	1	32	5	2	3	7	1	1	1	1	22	18	1	1	3	19	14	2	6	7	2	3	1
1	1	1	1	32	3	2	6	6	1	1	1	1	20	18	1	1	6	19	16	2	4	7	2	3	1
1	1	1	1	27	2	2	5	8	1	1	1	1	19	17	1	1	5	21	14	2	2	6	5	3	1

A pangrammatic length 4 loop completes this selection (no equivalent using plural s found).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	1	1	25	4	2	4	7	1	1	2	1	16	18	1	1	5	5	11	3	4	5	4	2	1
1	1	1	1	28	9	2	3	7	1	1	2	1	16	18	1	1	6	3	9	5	7	5	2	2	1
1	1	1	1	30	3	3	5	9	1	1	1	1	20	15	1	1	3	5	12	1	5	6	3	2	1
1	1	1	1	30	6	1	6	8	1	1	2	1	17	14	1	1	6	2	12	1	5	4	2	3	1

Besides the above, minimal pangrammatic loops with periods of 10, 33, and 55 (no plural s's) and periods of 15, 22, 23, 61, 207 and 220 (with plural s's) were found. This completes what I hope is an exhaustive survey of all self-enumerating minimal pangrammatic loops.

Refining the algorithm

In a previous section we looked at the 24 loops found by a program that mimics the book composition process with which we started. Thus, in English, all the multitude of different possible starting positions (alternative page 1 entries) lead inexorably into one of these 24 whirlpools or *attractors* as they are known to mathematicians (see Doug Hofstadter's admirably lucid exposition in [1]). This convergence is easy to understand. A 16 element list has 16! = 20,922,789,890,000 distinct permutations (= anagrams), all of them giving rise to a common description which is itself one among 16! new lists having a common successor, and so on. The resultant funnelling effect carries interesting implications.

Consider a computer program able to generate pages for such a book, starting from any text. A basic routine scans TEXTIN = page *N*, initially page 1, counts its letters and writes their totals in the form of number-words to TEXTOUT = page *N*+1. TEXTOUT is now substituted for TEXTIN, the routine reiterated, and so on. I like to picture this process as a machine that takes text as input and yields text as output, the latter coupled back to the former via a feedback path. This makes it easier to see that a reflexicon is effectively a *virus*: a code sequence able to subvert the machine so as to get itself perpetually reproduced. One way to hunt for reflexicons is therefore to set such a machine going and just wait for contagion to set in. However, there are still other viruses that may easily usurp it first. These are the loops of longer period, all of them similarly infectious. How can we immunize the device against these unwanted invaders? How do we write a book that ends specifically in a loop of period 1?

One answer is to alter the mechanism so as to neutralize longer cycles. Instead of updating the totals for every letter on every page, suppose the next page results from correcting the total for a single letter chosen at random each time. The resulting haphazard behaviour is loop-free by definition except in one case: when updating a total entails no change in the subsequent list because it is already correct — because the list *is* already a self-descriptor. In this way the program is forever free to keep juggling numbers until it eventually succumbs to a self-reproducer. The only snag is that anagrams of a solution then pass unheeded, which means 16! chances lost every time.

But not if we *alternate* methods: all totals updated on one pass, one random correction the next, and so on repeated. Now the former will catch any anagram, while the latter prevents latch-up in loops. A few million iterations (mutations) normally suffice to evolve (naturally select) a viable solution (virus). Assuming one exists, of course, failing which the process grinds on unchecked.

I should like to add that the key idea of loop-busting through inclusion of a *random* factor in the iteration process was the invention of John R. Letaw, a consultant in the areas of high-energy physics and astrophysics. Letaw had been the first to respond to a foolhardy challenge of mine that appeared in *Scientific American*, by coming up with an algorithm that yielded:

This computer-generated pangram contains six a's, one b, three c's, three d's, thirty-seven e's, six f's, three g's, nine h's, twelve i's, one j, one k, two l's, three m's, twenty-two n's, thirteen o's, three p's, one q, fourteen r's, twenty-nine s's, twenty-four t's, five u's, six v's, seven w's, four x's, five y's, and one z.

In fact, Letaw's algorithm worked quite differently to the one described above, successive approximations being determined by a weighted averaging process. Later, I experimented extensively with his algorithm, ending up with the method here outlined, which retains a random component although quite differently applied. Details of Letaw's algorithm can be found in [2], which appeared in response to an article of mine in [4].

Foreign language examples

In English, the word we use for the number 40 is FORTY and not FOURTY, as a non-native speaker might wrongly suppose. Similar opportunities for producing howlers lie in wait for English speakers attempting to create self-enumerators in foreign languages. Hence a certain coyness on my part in presenting the finds below, concious as I am that they could contain mistakes. I restrict myself to specimens that can fill squares or rectangles without leaving empty cells.

We begin with a couple of French reflexicons in Figures 17 and 18, each of them fitting snugly into 4×4 squares. Note the absence of any UN's. The fact that in each case the sum of the totals is not divisible by 4 is enough to show that they cannot be rearranged to form a magic square.

Fig. 17

TROIS A	TROIS C	TROIS D	NEUF E
QUATRE F	DEUX H	NEUF I	SIX N
QUATRE O	DEUX P	CINQ Q	SIX R
SEPT S	HUIT T	NEUF U	CINQ X

Fig. 18

TROIS A	DEUX C	QUATRE D	HUIT E
TROIS H	DIX I	TROIS N	SIX O
TROIS P	QUATRE Q	SEPT R	HUIT S
ONZE T	SEPT U	CINQ X	DEUX Z

A Dutch reflexicon follows in Figure 19, but in which D, F and L could be replaced by Þ, ð and ŋ, say, without injury to the square's standing as a self-enumerator:

TWEE A	TWEE C	EEN D	VIEREN TWINTIG E
EEN F	TWEE G	TWEE H	ZES I
EEN L	ZES N	VIER R	ZES S
ACHT T	VIER V	ZES W	ZES Z

Fig. 19

Next a German reflexicon in Figure 20, followed by a German co-descriptive pair in Figure 21.

EIN A	ZWEI B	ZWEI C	ZWEI D	NEUNZEHN E	SECHS F
VIER H	DREIZEHN I	ZWEI L	ZWÖLF N	ZWEI Ö	FÜNF R
VIER S	EIN T	FÜNF U	VIER V	SIEBEN W	NEUN Z

Fig. 20

ZWEI A	EIN B	ZWEI C
DREI D	SIEBZEHN E	SIEBEN F
ZWEI G	VIER H	FÜNFZEHN I
NEUN N	EIN O	NEUN R
DREI S	EIN T	VIER Ü
SIEBEN V	FÜNF W	ACHT Z

ZWEI A	VIER B	ZWEI C
DREI D	ZWANZIG E	FÜNF F
EIN G	VIER H	VIERZEHN I
VIERZEHN N	EIN O	FÜNF R
VIER S	ZWEI T	FÜNF Ü
DREI V	VIER W	SECHS Z

Fig. 21

According to the righthand list, there are five Ü's in its lefthand neighbour. Which is true provided we count U's as Ü's. But should we? Je ne sais pas.You'll have to ask the fox. One way around this would be to count both the Us and the Üs separately.

A pleasingly simple Italian reflexicon of just 12 entries in Figure 22 contains no UNOs:

TRE A	TRE C	QUATTRO D	OTTO E
SEI I	DUE N	OTTO O	QUATTRO Q
SEI R	TRE S	DODICI T	CINQUE U

Fig. 22

While in Figures 23 and 24 are seen Italian squares each including as many as four UNOs:

QUATTRO A	SEI C	SETTE D	SEI E
UNO F	DODICI I	OTTO N	DODICI O
CINQUE Q	QUATTRO R	UNO X	QUATTRO S
UNDICI T	UNO Y	UNDICI U	UNO Z

Fig. 23

DUE A	QUATTRO C	SETTE D	UNDICI E
UNO G	DICI I	SEI N	SEI O
DUE Q	DUE R	UNO W	SEI S
SETTE T	UNO X	DICI U	UNO Y

Fig. 24

Lastly, I conclude this little collection of (European) foreign language self-enumerators with a speciman in Latin. Alas a rectangle, rather than the "Latin square" one would have preferred.

QUATTUOR A	SEX C	QUATTUOR D	QUINDECIM E	QUINQUE I
SEX M	TRES N	OCTO O	TRES P	SEPTEM Q
SEPTEM R	OCTO S	QUATTUOR- DECIM T	DECEM U	TRES X

Fig. 25

Reflexiconography

The method of reflexicon construction having been explained, what else can we do with it? For a start, note that a self-descriptive *sentence* is really a sugar-coated reflexicon, the essential kernel, the list of totals, made more digestible through the addition of what I call “dummy text” such as, "This sentence contains ..". Thus, on appending these constant ballast letters to successive counts, our standard process again issues in an associated self-enumerating list, provided it exists. If not, change "contains" to "employs", say, and try again. Passing over the simplest instances, a few special finds made after adapting the mechanism to suit the purpose, deserve notice here. These are seen in Figures 26 - 29 below. Details of the program changes entailed by these special types would occupy us unduly, the basic mechanism remains the same.

Fig. 26

A (British) letter-totalling sentence

This sentence contains one
hundred and ninety-seven letters:
four a’s, one b, three c’s, five d’s,
thirty-four e’s, seven f’s, one g, six h’s,
twelve i’s, three l’s, twenty-six n’s, ten o’s,
ten r’s, twenty-nine s’s, nineteen t’s,
six u’s, seven v’s, four w’s,
four x’s, five y’s, and one z.

This pangram contains
two hundred nineteen letters:
five a’s, one b, two c’s, four d’s,
thirty-one e’s, eight f’s, three g’s, six h’s,
fourteen i’s, one j, one k, two l’s, two m’s,
twenty-six n’s, seventeen o’s, two p’s,
one q, ten r’s, twenty-nine s’s, twenty-
four t’s, six u’s, five y’s, nine w’s,
four x’s, five y’s, and one z.

Fig. 27

An (American) letter-totalling pangram

The right-hand
sentence contains
four a’s, one b, three c’s,
three d’s, thirty-nine e’s,
ten f’s, one g, eight h’s,
eight i’s, one j, one k,
four l’s, one m, twenty-three n’s,
fifteen o’s, one p, one q,
nine r’s, twenty-three s’s,
twenty-one t’s, four u’s,
seven v’s, six w’s,
two x’s, five y’s, and one z.

The left-hand
sentence contains
four a’s, one b, three c’s,
three d’s, thirty-five e’s,
seven f’s, four g’s, eleven h’s,
eleven i’s, one j, one k,
one l, one m, twenty-six n’s,
fifteen o’s, one p, one q,
ten r’s, twenty-three s’s,
twenty-two t’s, four u’s,
three v’s, five w’s,
two x’s, five y’s, and one z.

Fig. 28

A (trans-Atlantic) mutually-descriptive (pangrammatic) pair

The adjacent text utilizes
four a’s, one b, two c’s,
three d’s, thirty-six e’s, five f’s,
three g’s, nine h’s, eleven i’s,
two j’s, one k, four l’s, one m,
eighteen n’s, thirteen o’s,
one p, one q, eight r’s,
twenty-seven s’s, twenty-four t’s,
four u’s, four v’s, seven w’s,
three x’s, four y’s, and two z’s.

The adjacent text utilizes
four a’s, one b, two c’s,
three d’s, thirty-two e’s, nine f’s,
three g’s, eight h’s, eleven i’s,
two j’s, one k, three l’s, one m,
seventeen n’s, fifteen o’s,
one p, one q, eleven r’s,
twenty-six s’s, twenty-one t’s,
eight u’s, six v’s, six w’s,
three x’s, four y’s, and two z’s.

Fig. 29

A mutually-descriptive pair with identical dummy text

Returning to reflexicons proper, in line with French practice above (Figure 1), plural S is dispensible. Two instances are then found, one trivial :

FIVE F, FIVE I, FIVE V, FIVE E.

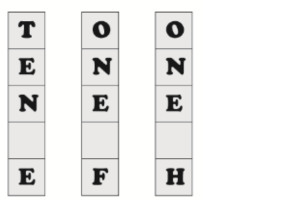
The other less dull:

TWELVE E, FIVE R,
SIX F, FIVE S,
THREE H, SIX T,
SEVEN I, THREE U,
TWO L, SIX V,
TWO N, FOUR W,
FIVE O, FOUR X.

This is condensed, but logologists like their alphabet soup really thick. Plural S has been dropped. Is there any other way to increase the semantic density through discarding still further redundant symbols? There is.

Consider a list in which the stated letter counts are in each case exactly one short of the true total: `TEN E, ONE F, ONE H, TWO I, SIX N, SEVEN O, ONE R, TWO S, FIVE T, TWO V, THREE W, ONE X.` That is, in fact this list contains *eleven* e's, not ten, *two* f's and not one, etc. Each of the twelve items on the list can now be written on a strip of card, on one side running from left to right:

on the other, from top to bottom:



Using trial and error, an arrangement must now be sought such that the strips overlap eachother in a self-descriptive crossword pattern that eliminates excess letters. Figure 30 shows my own very first attempt at making a *self-intersecting* reflexicon:

Fig. 30
Eliminating an excess letter on each intersection yields a self-enumerating crossword.

This was a good start, but OTT, NWW, EOO, OO, NNFIVE, and EE, are pseudowords and thus serious blemishes. A different layout wont help either since ONE H and ONE R must always remain bonded together with HR in THREE. No, to escape this problem called for a new set of items involving less intersections per strip so as to win elbow-room. This brings us to a key insight.

Twelve strips bearing 12 excess letters implies 12 intersections. Yet N strips can cross at most *N*-1 times unless linked to include a *closed chain*. Look at ONE X, SIX N, SEVEN O and TWO S in Figure 30. Contriving such a loop is the major constraint in devizing solution layouts. Thus, a new list requiring fewer intersections than strips makes for a big gain in layout flexibility (and vice versa), although two or more fewer will imply a non-connected pattern. To avoid this, the obvious course then is to seek an *N* item list involving *N*-1 excess letters = intersections. An example is seen in Figure 31.

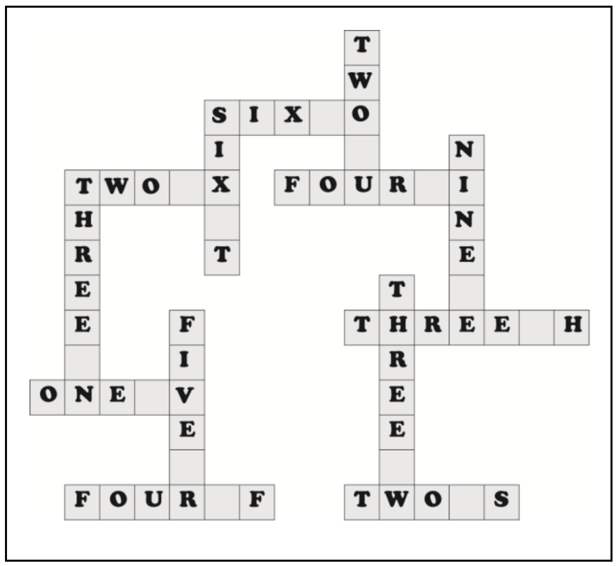


Fig. 31
A crossword with 13 strips and 12 intersections

This is more like it: no pseudowords and 3.846 letters per word or 0.26 words per letter, which, with the words now spatially interlocking, is virtually alphabet jelly! The trouble is that now one letter (F) is alone in not occupying an intersection, a niggling asymmetry. At some loss in semantic density, however, restoring plural S is another way to win room for maneuver, as in Figure 32.

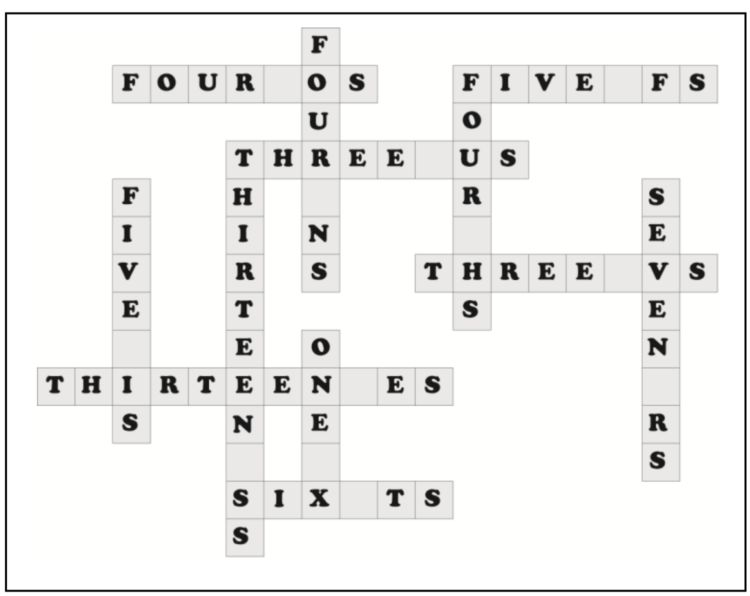
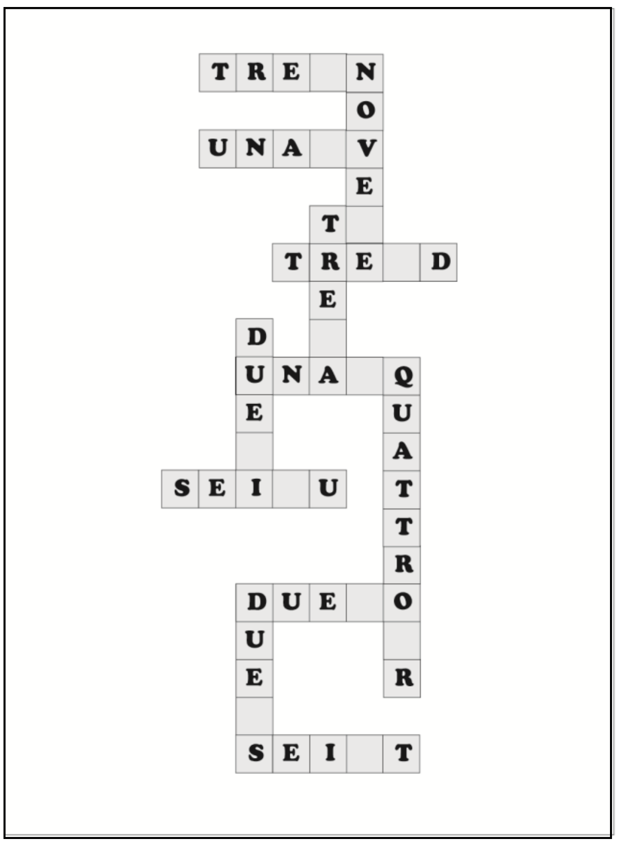


Fig. 32
Adding plural S wins more space to manoeuvre

Here we are back to 12 strips and 12 intersections (necessitating a loop), each occupied by one of the 12 letters occurring. On consideration, this is a remarkable property, more so than first sight suggests, since it depends on finding a list in which the letters outnumber their totals by one exactly, the excess then vanishing on intersects. The list used in Figure 32 is thus exceptional. For example, no French or Italian equivalent exists. Unusually, however, English enjoys two such lists (that include plural s), the second comprising 13 words, although its internal peculiarities hinder construction of elegant crosswords. Some readers may like to try their hand; the totals are as follows: E:15, F:8, G:1, H:3, I:5, L:1, N:4, O:5, R:5, S:11, T:4, U:3, V:4.

Figures 33 and 34 show Italian and French reflexicons, neither of which languages call for plural S. Of course, there is nothing against letters appearing on intersects more or less than once, as with U (twice) in the French example.

Fig. 33



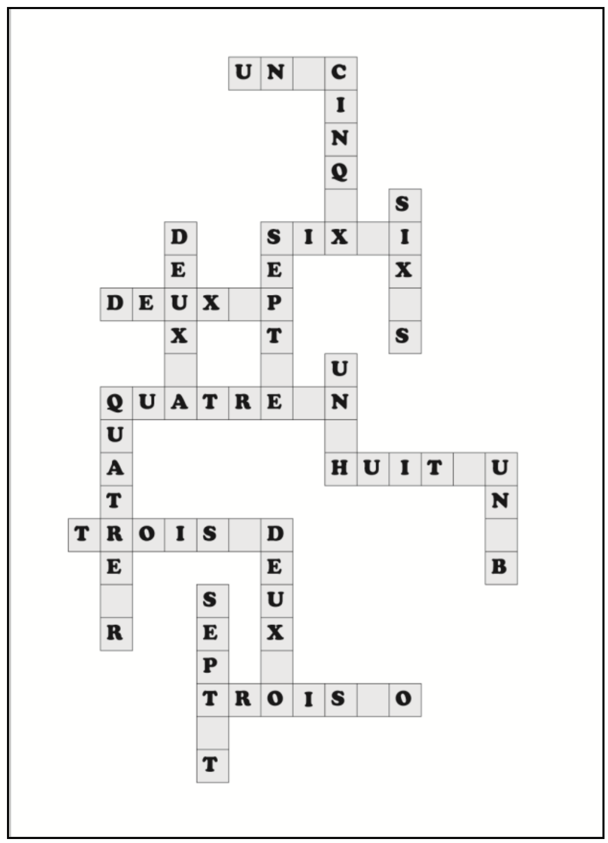


Fig. 34

Both of these illustrate a further trick in the reflexiconographer's repertoire: the use of "ONE #" (here for example, "UN B") as unobtrusively appended dummy text. This is a useful stratagem when "pure" solutions cannot otherwise be found, although the arbitrariness of letter used (UN B could equally be changed to UN Z) detracts from their logological elegance, a point to bear in mind when assessing the merits of different specimens. Dummy text may take more conventional forms of course, as in Figure 35, where intersects outnumber strips, a fact reflected in multiple loops.

However, the construction of multi-loop specimens is demanding.

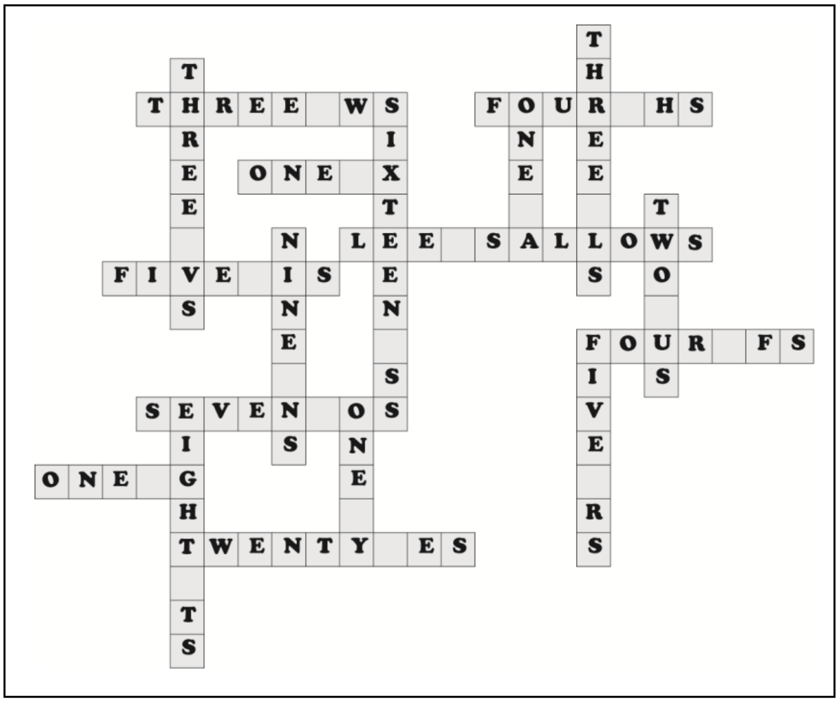


Fig. 35
A further example
of self-reference

Some loops are not what they seem. Figure 36 exhibits pseudoloops and the two ways they arise: via intersection on a blank; viz. THIRTEEN SS and FIVE FS, and via abutment onto a blank, viz. THIRTEEN ES and FOUR HS. The single real loop here is formed by FOUR OS, FOUR NS, FIVE FS, and FIVE IS .

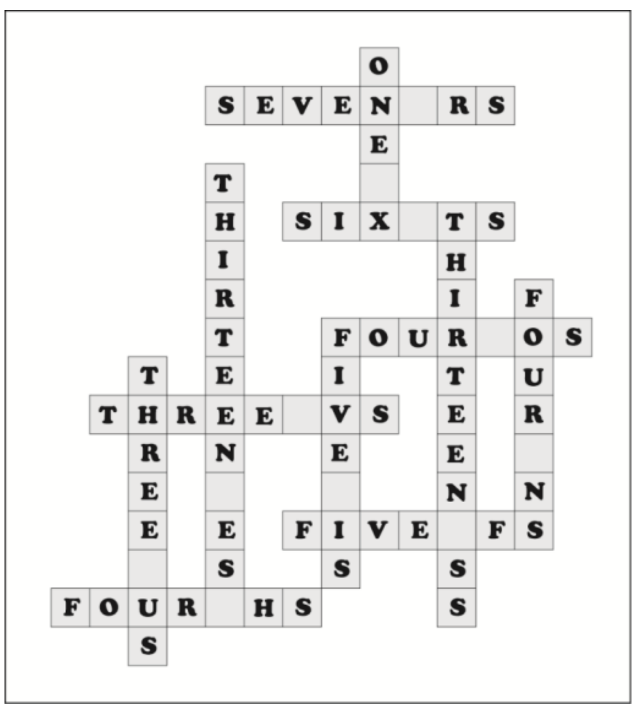


Fig. 36
A compact
layout of 14×16.

Pseudoloops can make for compacter layouts, a fact seen in comparing Figures 35 and 31, both of which employ the same entries (the special set of 12 strips), whereas the two patterns occupy rectangles of 14 × 16 versus 14 × 18, respectively. Two natural questions then arising are: How many distinct (fully connected) self-intersecting reflexicons can be formed from this set of strips? and, Which of them is the most compact?

To seek answers, Victor Eijkhout, a mathematical friend, wrote a recursive strip-shuffling computer program able to scan for solutions. However, although several days running on a mainframe computer produced thousands of alternative solution layouts, it became clear there was no chance of the job terminating within any feasible time-scale. The two questions thus remain unanswered. Figure 35, which was hand-produced, is the most compact specimen known.

Nevertheless, at my suggestion Victor set his (slightly modified) program to work on a new but related search that was to bear fantastic fruit. Figures 37 and 38 embody two jewels of logology (we seem to have reached alphabet ice). Here are the classic 12 strips again, the loop now realized as the entire set holding hands in a single twelve-linked bracelet! The pair shown are among 18 such specimens found by the program, not counting rotations and reflections, but including trivial variations such as when FIVE IS is switched with FIVE FS inFigure 38.

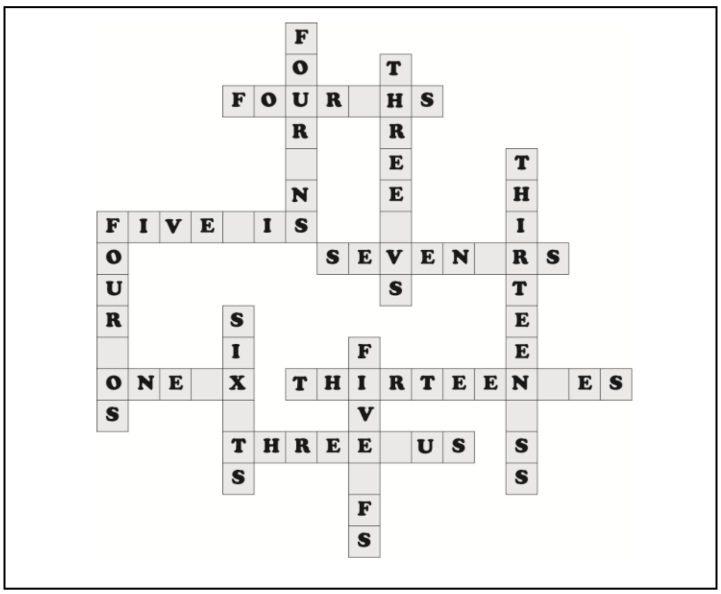


Fig. 37
The 12 strips holding hands to complete a circuit

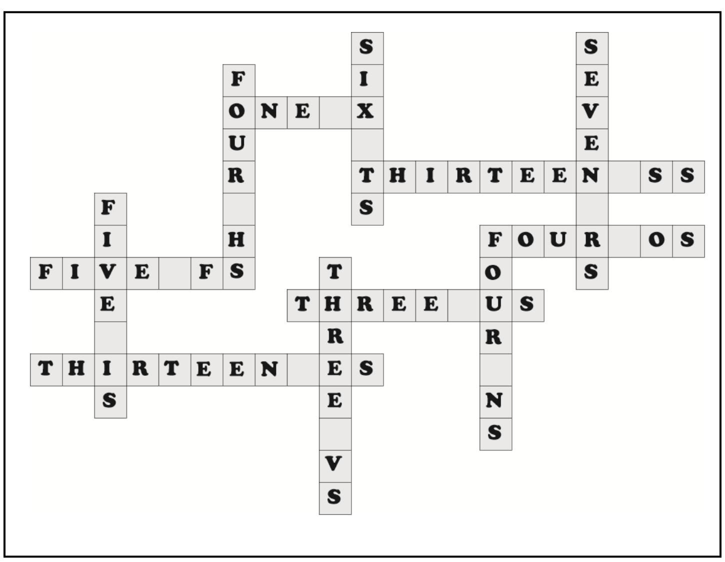


Fig. 38
Another unbroken chain discovered by Victor Eijkhout's program

Marvellous as Eijkhout's finds are, further collector's pieces probably await discovery. For example, might there exist a reflexicon with a truly symmetrical layout? A congruent pair showing distinct solution entries? A 3-dimensional bracelet (that forms a knot)? A (possibly interlacing) co-descriptive pair? A pangrammatic reflexicon (without dummy text)? The list is easily extended. In the meantime one special specimen has passed unmentioned. Figure 38 again features 12 intersections each occupied by one of the 11 letters occurring, although now there is no plural S. It is a relative of Figure 30, the first self-intersector examined, where the number of excess letters also matches the number of items, but which cannot be solved without creating “pseudowords". As with the list used in Figure 30, a second list with the same property (but minus plural S) has been found. (A third trivial case is FOUR [F], FOUR [O], FOUR [U], FOUR [R]). The analogous question then arises: How many distinct solutions can be formed from the entries in Figure 39?

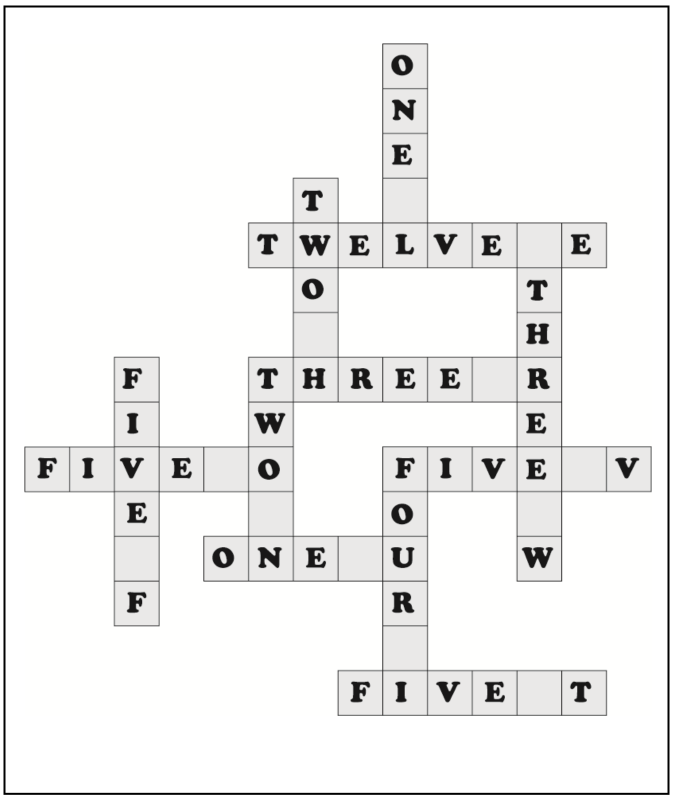


Fig. 39
11 intersections,
each occupied by
one of the letters
employed

Lastly, to conclude this brief review, in Figure 40 I offer a final example of the state of the art, a reflexicon that incorporates its own letter-total. Here the 17 intersections are occupied by 16 different letters, one of which, O, occurs twice. From this can be inferred a single closed chain of strips, in this case formed by the five strips: 106 LETTERS - 3 HS – 19 SS – 1 Y – 21 ES – 5 WS – 4 VS - 3 US - . Can a similar specimen be found using still fewer than 106 letters? Here is another tough challenge for the computational logologist.

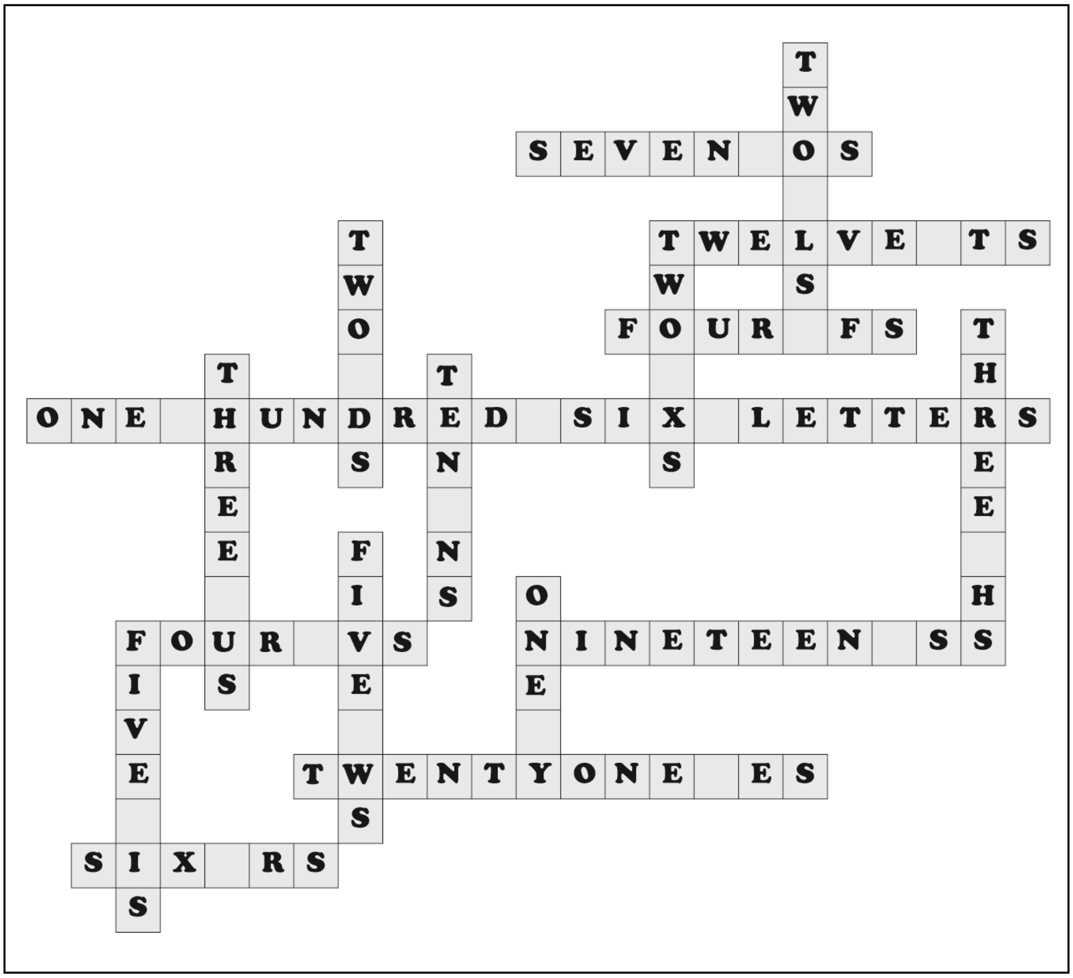


Fig. 40

* * * * * *

The original version of this now expanded article was first published in *Word Ways*, August 1992

<u>References</u>

[1] Hofstadter D. Metamagical Themas, Basic Books, 1985. p.364
[2] Letaw John R., Abacus, Vol. 2, No.3, pp 42-7
[3] Brooke Maxey, Pangrams, Word Ways Vol.20 No.2, p.90
[4] Sallows L., In Quest of a Pangram, Abacus, Vol.2, No.3, pp 22-40