MATHEMATICS OF COMPUTATION, VOLUME 25, NUMBER 113, JANUARY, 1971

Strongly Asymmetric Sequences Generated by Four Elements

By František Fiala*

Abstract. Some general properties of strongly asymmetric sequences generated by $m \geqq 1$ elements (m-SAS) are given. Computational experience with two algorithms—for listing of all 4-SASs of a given length n and for generating the smallest 4-SAS of length $n = 1, 2, \cdots$ —supports the conjecture that there exists an infinite 4-SAS. The smallest 4-SAS of length 592 is presented.

Let us put

$$\mathbf{E}_m = \{0, 1, \cdots, m - 1\}, \tag{1}$$

where $m \geqq 1$. A sequence

$$(a_1, a_2, \cdots, a_n) \tag{2}$$

is said to be a *strongly asymmetric m-sequence* (m-SAS) *of length n* if

(a) $a_i \in \mathbf{E}_m$ for $i = 1, \cdots, n$;

(b) for any $j \geqq 0, k > 0$ such that $j + 2k \leqq n$ two consecutive segments $(a_{j+1}, \cdots, a_{j+k})$, $(a_{j+k+1}, \cdots, a_{j+2k})$ do not contain the same number (frequency) of 0's, 1's, $\cdots$, $(m - 1)$'s.

For instance, if $m = 3$ then (0102010) is 3-SAS of length 7, but (01020120) is no 3-SAS (for $j = 1, k = 3$ two consecutive segments (102), (012) contain the same number of 0's, 1's and 2's).

An infinite sequence

$$(a_1, a_2, \cdots, a_n, \cdots) \tag{3}$$

is m-SAS if for an arbitrary positive integer n the sequence (2) is m-SAS.

P. Erdös [1] posed the problem of finding an infinite SAS using the minimal number of symbols. A. A. Evdokimov [2] constructed an infinite m-SAS for $m = 25$ and expressed the opinion that the number m might be reduced. It is easy to establish that there is no infinite m-SAS for $m \leqq 3$ [2]. Therefore, I investigated the case $m = 4$. It appears that the technique used in [2] is not applicable to this case. In this paper some general properties of m-SASs are given and computational experience with 4-SASs is collected. The smallest 4-SAS of length 592 is presented.

1. We shall describe some general properties of m-SAS for $m \geqq 1$ which can be easily verified.

Received June 22, 1970.

AMS 1969 *subject classifications.* Primary 1007, 0510; Secondary 6510.

Key words and phrases. Strongly asymmetric sequences, lexicographical ordering, number systems.

* This paper was written while the author was a post-doctorate fellow at the Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada.

1.1. An m-SAS

$$(a_1, \cdots, a_{n_1}) \tag{4}$$

has *a prolongation* if there is an m-SAS

$$(a_1, \cdots, a_{n_1}, a_{n_1+1}, \cdots, a_{n_2}) \tag{5}$$

of length $n_2 > n_1$. m-SAS (4) is called *an initial segment* of (5).

The set of all (finite and infinite) m-SASs can be *ordered lexicographically*, i.e. if

$$(a'_1, \cdots, a'_{n_1}), \tag{6}$$

$$(a''_1, \cdots, a''_{n_2}), \tag{7}$$

are two m-SASs then (6) is *less than* (7) if either there exists such an i, $0 \leqq i < \min(n_1, n_2)$, that

$$a''_j = a''_i \qquad \text{for} \quad j = 1, \cdots, i, \tag{8}$$

and

$$a'_{i+1} < a''_{i+1} \tag{9}$$

or (6) is an initial segment of (7). E.g. (01020) is less than (0121030), (10230) is less than (120).

Principle of duality. If (2) is an m-SAS then

$$(m - 1 - a_1, \cdots, m - 1 - a_n) \tag{10}$$

is m-SAS, too. (10) is called *dual* to (2) and vice versa.

All m-SASs of a given length n form a finite chain in the lexicographical ordering. If (2) is the smallest m-SAS of length n then its dual m-SAS is the greatest one (of the same length).

Obviously, *there is an m-*SAS *of length n if and only if there is the smallest m-*SAS *of length n.*

1.2. If (2) is an m-SAS of length n and p is an arbitrary permutation of $\mathbf{E}_m$ then

$$(p(a_1), p(a_2), \cdots, p(a_n)) \tag{11}$$

is m-SAS, too. If (2) contains each element of $\mathbf{E}_m$ and p is nonidentical then

$$(p(a_1), \cdots, p(a_n)) \neq (a_1, \cdots, a_n). \tag{12}$$

1.3. Let $m \leqq 3$. Direct check shows that there is no m-SAS of length $n > 7$. Therefore, any 4-SAS of length $n \geqq 8$ contains all numbers 0, 1, 2, 3. Consequently, the number S_n^4 of all 4-SASs of length $n \geqq 8$ is divisible by $4! = 24$. For arbitrary $m > 1$, the number S_n^m of all m-SASs of length $n \geqq 2$ is divisible by $m(m - 1)$.

1.4. *Denote by*

$$(a_1, \cdots, a_{2^m-1}) \tag{13}$$

*the smallest m-*SAS *of length* $2^m - 1$. *Then* (13) *has no m-*SAS *prolongation and*

$$(a_1, \cdots, a_{2^m-1}, m, a_1, \cdots, a_{2^m-1}) \tag{14}$$

is the smallest $(m + 1)$-SAS (*of length* $2^{m+1} - 1$).

Thus, 0, 010, 0102010, 010201030102010, 0102010301020104010201030102010 are the smallest 1-, 2-, 3-, 4- and 5-SASs, respectively.

Proof by induction. For $m = 1$ the proposition obviously holds. Assume that it is true for $m \geqq 1$, i.e. (13) has no m-SAS prolongation and (14) is the smallest $(m+1)$-SAS. We prove that the proposition holds for $m + 1$.

Let $(a_1, \cdots, a_{2^m-1}, m, a_1, \cdots, a_{2^m-1}, a)$ be an $(m+1)$-SAS prolongation of (14). If $a < m$ then $(a_1, \cdots, a_{2^m-1}, a)$ is an m-SAS prolongation of (13), contrary to our assumption. $a = m$ is obviously impossible. Thus (14) has no $(m+1)$-SAS prolongation.

It is easy to check that

(15) $$(a_1, \cdots, a_{2^m-1}, m, a_1, \cdots, a_{2^m-1}, m+1, a_1, \cdots, a_{2^m-1}, m, a_1, \cdots, a_{2^m-1})$$

is an $(m+2)$-SAS of length $2^{m+2} - 1$. Let

(16) $$(b_1, \cdots, b_{2^{m+2}-1})$$

be the smallest $(m+2)$-SAS (of length $2^{m+2} - 1$). Then $(b_1, \cdots, b_{2^{m+1}-1})$ is an $(m+2)$-SAS and therefore

(17) $$(b_1, \cdots, b_{2^{m+1}-1}) \leqq (a_1, \cdots, a_{2^m-1}, m, a_1, \cdots, a_{2^m-1}).$$

(14) is the smallest $(m+1)$-SAS and also the smallest $(m+2)$-SAS. Hence, in the relation (17), the equal sign holds. (14) has no $(m+1)$-SAS prolongation, thus $b_{2^{m+1}} = m + 1$. Again

(18) $$(b_{2^{m+1}+1}, \cdots, b_{2^{m+2}-1}) \geqq (a_1, \cdots, a_{2^m-1}, m, a_1, \cdots, a_{2^m-1}).$$

The minimal property of (16) causes again that in (18) the equal sign holds. Thus (15) is the smallest $(m+2)$-SAS.

2. We consider the case $m = 4$. Two algorithms have been prepared. The first one lists all 4-SASs of a given length n in their lexicographical order. The second algorithm generates the smallest 4-SAS of length $n = 1, 2, \cdots$. Both of them were written in FORTRAN IV for IBM 360/67.

2.1. We recall that the set of all 4-SASs of a given length n is ordered lexicographically (1.1). A 4-SAS of length n

(2) $$(a_1, a_2, \cdots, a_n)$$

may be seen as an (n-positional) integer written in the number system with the radix 4. Therefore, we can also handle (2) as a number and add to it another number (in the number system with the radix 4). This feature is used in the following algorithms both for finding all 4-SASs of length n in their ascendent order and for generating the smallest 4-SAS (2.2).

The first algorithm runs as follows:

2.1.1. Start with the smallest 4-SAS of the form (2).

2.1.2. Write the current 4-SAS.

2.1.3. Add 1 to the current sequence (in the number system with the radix 4). If a transfer from the most left position occurs, go to 2.1.5.

2.1.4. Ask if the new sequence is 4-SAS. If the answer is "Yes", go to 2.1.2. If the answer if "No", go to 2.1.3.

2.1.5. Stop.

A complete listing of 4-SASs was made for $n = 1, \cdots, 9$. The numbers S_n^4 of all 4-SASs for $n = 10, \cdots, 18$ were found using 1.3.

The following tables show some of the results. Table 1 contains the first 25 smallest 4-SASs of lengths 14, 15, 16. Table 2 shows the numbers S_n^4 and the ratio of two consecutive S_n^4, S_{n+1}^4. For comparison, similar data for $m = 3$ are added.

For $n < 14$, every 4-SAS has at least one prolongation, but for $n \geqq 14$, there are some 4-SAS without any prolongation (denoted by * in Table 1).

2.2. *The algorithm for generating the smallest* 4-SAS *for* $n = 1, 2, \cdots$ *can be described in the following way* (cf. 2.1):

2.2.1. Start with $n = 1$, $(a_1) = (0)$.

2.2.2. Write n, $(a_1, \cdots, a_n)$.

2.2.3. Multiply the number $a_1 \cdots a_n$ by 4 (in the number system with the radix 4). Add 1 to n.

2.2.4. Ask if the new sequence is 4-SAS. Yes: go to 2.2.2. No: go to 2.2.5.

TABLE 1

Order No.	$n = 14$	$n = 15$	$n = 16$
1	01020103010201	010201030102010*	0102010301021013
2	01020103010210	010201030102101	0102010301021230
3	01020103010212	010201030102123	0102010301021231
4	01020103010213	010201030102131	0102010301021232
5	01020103012010*	010201030102132	0102010301021310
6	01020103012013	010201030120131	0102010301021312
7	01020103012023	010201030120230	0102010301021320
8	01020103012101	010201030120232	0102010301021321
9	01020103012103	010201030121012	0102010301021323
10	01020103012130	010201030121013	0102010301201310
11	01020103012131	010201030121031	0102010301201312
12	01020103012132	010201030121301	0102010301202302
13	01020103012310	010201030121303	0102010301202303
14	01020103012312	010201030121310	0102010301202320
15	01020103012313	010201030121312	0102010301202321
16	01020103012320	010201030121321	0102010301210123
17	01020103012321	010201030121323	0102010301210130
18	01020103020102	010201030123101	0102010301210131
19	01020103020120	010201030123103	0102010301210132
20	01020103020121	010201030123121	0102010301210310
21	01020103020123	010201030123130	0102010301210312
22	01020103021013	010201030123202	0102010301210313
23	01020103021020*	010201030123203	0102010301213010*
24	01020103021023	010201030123212	0102010301213012
25	01020103021202	010201030201020*	0102010301213013

2.2.5. Add 1 to the current sequence (in the number system with the radix 4). If a transfer from the most left position occurs, go to 2.2.6. Elsewhere, go to 2.2.4.

2.2.6. Stop.

In other words, if $(a_1, \cdots, a_n)$ is the smallest 4-SAS of length n, we will try to prolong it successively by adjoining to the end 0, 1, 2, 3. The first prolongation

TABLE 2

Length	S_n^4	S_{n+1}^4/S_n^4	S_n^3	S_{n+1}^3/S_n^3
1	4		3	
		3		2
2	12		6	
		3		2
3	36		12	
		$2.\overline{6}$		1.5
4	96		18	
		2.75		$1.\overline{6}$
5	264		30	
		$2.\overline{45}$		1
6	648		30	
		$2.\overline{4}$		0.6
7	1584		18	
		$2.\overline{257}$		
8	3576		None	
		2.201342282		
9	7872			
		1.951219512		
10	15360			
		1.9		
11	29184			
		1.751644737		
12	51120			
		1.768075117		
13	90384			
		1.753053638		
14	158448			
		1.806876704		
15	286296			
		1.78070249		
16	509808			
		1.773797194		
17	904296			
		1.721011704		
18	1556304			

TABLE 3

010201030102101310121320210130102030201202310120 23

(if any) is the smallest 4-SAS of the length $n + 1$. If there is no prolongation of $(a_1, \cdots, a_n)$ then we take instead of it the following greater 4-SAS of length n (with regard to the lexicographical order—cf. 2.1) and proceed in a similar way. The procedure stops if no 4-SAS of length n has any prolongation.

Some of the results obtained by using this algorithm follow.

The longest 4-SAS I found is the following smallest 4-SAS of length 592:

The preceding 4-SAS was not found at once and Table 4 shows the necessary computing time for generating the smallest 4-SASs for different lengths n:

We can see that for increasing n, the indicated computing time increases very rapidly. This increase is partially due to the increase in length, but mostly it is caused by the fact that we do not have available for a given n all lexicographically ordered 4-SASs of length n. If a 4-SAS has no prolongation then we have to determine the next greater 4-SAS of the same length and try to prolong that one. This procedure is very costly with regard to time. Table 5 lists the lengths $n < 592$ having the prop-

TABLE 4

Length n	50	100	200	261	299	376	434	466
Time (min)	0.17	0.60	2.57	5.57	10.57	25.57	40.57	55.57

Length n	509	530	588	592
Time (min)	70.57	85.57	100.57	115.57

TABLE 5

15(1), 23(1), 36(1), 62(2), 63(3), 85(1), 94(1),
106(1), 119(2), 122(2), 124(1), 130(1), 135(1), 136(1), 140(1), 146(2), 149(1), 153(1), 154(1), 167(1), 172(1), 176(1), 177(1), 187(4), 188(1), 189(1),
205(1), 216(3), 222(1), 224(1), 239(1), 240(1), 241(1), 243(1), 248(1), 249(1), 282(1), 283(4), 284(2), 289(2), 290(1), 299(3),
305(1), 306(1), 308(1), 310(2), 314(1), 321(1), 323(3), 326(1), 331(6), 332(1), 335(1), 339(1), 343(1), 358(1), 363(1), 369(2), 370(4), 373(1), 375(1), 376(1), 390(1), 392(2), 394(2), 395(3),
404(1), 412(1), 415(1), 420(1), 424(1), 428(2), 434(1), 435(1), 445(2), 446(1), 447(2), 449(1), 456(1), 462(2), 463(1), 464(1), 470(3), 473(2), 478(2), 484(1), 485(1), 490(1), 491(2),
508(1), 509(1), 510(2), 515(1), 525(1), 529(1), 530(1), 532(3), 542(1), 543(1), 544(1), 549(1), 554(2), 559(1), 574(1), 583(1), 588(1), 589(1), 591(1), 592(?)

erty that the smallest 4-SAS of length n has no prolongation. The number i in parentheses shows that i smallest 4-SASs of length n (with regard to the lexicographical order) have no prolongation.

The most unfavorable case occurs for $n = 331$ when 6 smallest 4-SASs have no prolongation. But the most frequent case is that only one smallest 4-SAS has no prolongation.

Frequency of symbols in the smallest 4-SAS of length 592 and several consecutive segments is given in Table 6. We can see from the second column that the frequency of an arbitrary symbol from $\mathbf{E}_4$ in a 4-SAS (or its segment) of length 100 varies between 17 and 35.

TABLE 6

Symbol	Length						
	1–100	101–200	201–300	301–400	401–500	501–592	1–592
0	35	29	26	24	20	26	160
1	26	25	27	27	27	26	158
2	22	23	28	24	24	21	142
3	17	23	19	25	29	19	132

Conclusions. Unfortunately, the problem of determining an infinite 4-SAS remains open. This paper is only a small contribution toward its solution. From Table 2 we can see that the number of 4-SASs of length n grows quickly for small n and will probably continue. In spite of the fact that the time for generating the smallest 4-SAS increases for greater n (Table 4) the prolongation is not more difficult in essence (Tables 5, 6). That leads to the following conjecture:

The algorithm 2.2 *will never stop, i.e. for any* $n = 1, 2, \cdots$ *there exists a* 4-SAS.

Acknowledgements. I am very indebted to Dr. I. Rosenberg for introducing me to this problem and for many suggestions and discussions about it.

Department of Computing Science
University of Alberta
Edmonton, Alberta, Canada

1. ERDÖS, "Some unsolved problems," *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, v. 6, 1961, pp. 221–254; Russian transl., *Matematika*, v. 7, 1963, no. 4, p. 128. MR **31** #2106.

2. A. A. EVDOKIMOV, "Strongly asymmetric sequences generated by a finite number of symbols," *Dokl. Akad. Nauk SSSR*, v. 179, 1968, pp. 1268–1271 = *Soviet Math. Dokl.*, v. 9, 1968, pp. 536–539. MR **38** #3156.