Scan 7039
7040

Moser

7039
7040

Report from the / Rapport du

Department of Mathematics and Statistics

ISSN 0824-4944

91-10

On cyclic binary n-bit strings

W.O.J. Moser

McGill University

Burnside Hall, 805 ouest, rue Sherbrooke

Montréal, Québec, Canada H3A 2K6

91-10

On cyclic binary n-bit strings

W.O.J. Moser

Department of Mathematics and Statistics
McGill University
Burnside Hall
805 Sherbrooke Street West
Montreal, Canada H3A 2K6

Received: 14 May, 1991
Legal deposit: 2nd quarter, 1991, National Library of Canada
Dépot Légal: 2me trimestre, 1991, Bibliothèque nationale du Québec

91-10

On cyclic binary n-bit strings

W. O. J. Moser

Abstract

Recurrences are obtained for the number of cyclic binary n-bit strings restricted by the non occurrence of certain substrings. These simplify and generalize the counts obtained by Agur et al (Discrete Math. 70 (1988) 295–302).

Résumé

Nous obtenons des récurrences pour le nombre de suite "n-bit" cycliques binaires qui sont contraintes par la non occurrence de certaines sous-suites. Ceci simplifie et généralise les décomptes obtenus par Agur et al (Discrete Math. 70 (1988) 295–302).

On cyclic binary n-bit strings

W. Moser

McGill University, Montreal, Canada H3A2K6

A binary $n-$bit *cyclic* string (briefly $n-$CS) is a sequence of n 0's and 1's (the bits), with the first and last bits considered to be adjacent (i.e., the first bit follows the last bit). This condition is visible when the string is displayed in a circle with one bit "capped": the capped bit is the first bit and reading clockwise we see the second bit, the third bit, and so on to the n^{th} bit (the last bit). A sequence of consecutive bits is a *substring*. Motivated by a problem of genetic information processing, Agur, Fraenkel and Klein [1] derived formulae for the number of n-CSs with no substrings 010 nor 101 (i.e., no alternating substring has length > 2 or, equivalently, all alternating substrings have length ≤ 2) and for the number with no substrings 0 0 0 nor 1 1 1 (i.e., no subtrings of like bits has length > 2, or, equivalently, all substrings of like bits have length ≤ 2).

In this note we shall generalize the counts to "no alternating substring has length $> w$ (equivalently "all alternating substrings have length $\leq w$") and to "no substring of like bits has length $> w$ "(equivalently, all substrings of like bits have length $\leq w$).

For this purpose we first investigate (for $n \geq 1$, $k \geq 0$, $w \geq 1$) the number $(n : k)_w$ of $n-$CSs which have exactly k 1's ($n - k$ 0's) and every 1 followed by $\leq w$ 0's (equivalently, every substring of 0's has length $\leq w$). The parenthetical remark suggests that we take

$$(n : 0)_w = \begin{cases} 1, & 1 \leq n \leq w, \\ 0, & w + 1 \leq n. \end{cases} \tag{1}$$

Of course

$$(n : k)_w = \begin{cases} \binom{n}{k}, & 1 \leq n \leq w + k, \\ 0, & 1 \leq k,\ k(w + 1) < n, \end{cases} \tag{2}$$

where

$$\binom{n}{k} = \begin{cases} n!/k!(n-k)!, & 0 \leq k \leq n, \\ 0, & 0 \leq n < k. \end{cases}$$

Consider a $n-$CS counted in $(n : k)_w$, $n \geq w + 2$, $k \geq 2$. If the first bit is 1 (i.e., capped bit is $\hat{1}$), and the last 1 is followed by exactly i 0's ($0 \leq i \leq w$), delete this last 1 and the i 0's which follow it and then we have a $(n - 1 - i)-$CS with $k - 1$ 1's, first bit 1, and every 1 followed by $\leq w$ 0's. If the first bit is 0, and the first 1 is followed by i 0's ($0 \leq i \leq w$), delete this first 1 and the i 0's which follow it and we then have a $(n - 1 - i)-$CS with $k - 1$ 1's, first bit 0, and every 1 followed by $\leq w$ 0's. Hence

$$(n : k)_w = (n - 1 : k - 1)_w + (n - 2 : k - 1)_w + \cdots + (n - 1 - w : k - 1)_w, \qquad k \geq 2, \quad n \geq w + 2. \tag{3}$$

The number of $n-$CSs with every 1 followed by $\leq w$ 0's is $F_w(n) = \sum_{k=0}(n:k)_w$, while the number of $n-$CSs with an even (resp. odd) number of 1's, each followed by $\leq w$ 0's, is

$$F_w^e(n) = \sum_{k=0}(n:2k)_w \qquad \text{resp.} \qquad F_w^o(n) = \sum_{k=1}(n:2k-1)_w. \tag{4}$$

From (1), (2) and (3) we deduce that

$$F_w^e(n) = \begin{cases} 2^{n-1}, & 1 \leq n \leq w, \\ 2^{n-1}-1, & n = w+1, \\ F_w^o(n-1) + F_w^o(n-2) + \cdots + F_w^o(n-1-w), & n \geq w+2, \end{cases} \tag{5}$$

$$F_w^o(n) = \begin{cases} 2^{n-1}, & 1 \leq n \leq w+1, \\ F_w^e(n-1) + \cdots + F_w^e(n-1-w) + n - 2(w+1), & w+2 \leq n \leq 2w+1, \\ F_w^e(n-1) + F_w^e(n-2) + \cdots + F_w^e(n-1-w), & n \geq 2w+2, \end{cases} \tag{6}$$

and

$$F_w(n) = \begin{cases} 2^n, & n = 1,2,\ldots,w, \\ 2^n - 1, & n = w+1, \\ F_w(n-1) + \cdots + F_w(n-1-w) + n - 2(w+1), & w+2 \leq n \leq 2w+1, \\ F_w(n-1) + F_w(n-2) + \cdots + F_w(n-1-w), & n \geq 2w+2. \end{cases} \tag{7}$$

We will need the numbers $F_w^e(n)$ and $F_w^o(n)$. These can be computed from (5) and (6). A simpler way of getting them is to observe that the numbers $D_w(n) = F_w^e(n) - F_w^o(n)$ satisfy

$$D_w(n) = \begin{cases} 0, & 1 \leq n \leq w, \\ -1, & n = w+1, \\ w+1, & n = w+2, \\ -1, & w+3 \leq n \leq 2w+2, \\ D_{n-2-w}, & n \geq 2w+3. \end{cases} \tag{8}$$

Thus, we may compute $F_w(n)$ from (7), $D_w(n)$ from (8) and then

$$2F_w^e(n) = F_w(n) + D_w(n), \qquad 2F_w^o(n) = F_w(n) - D_w(n). \tag{9}$$

The following Tables show $F_w(n)$, $D_w(n)$, $2F_w^e(n)$ and $F_w^o(n)$ for $w = 1,2,3$ and $n = 1,2,3,\ldots,15$. The boldface entries are the initial values.

(A1610)−1
a rational fraction

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$F_1(n)$	**2**	**3**	**4**	7	11	18	29	47	76	123	199	322	521	843	1364
$D_1(n)$	**0**	**-1**	**2**	**-1**	-1	2	-1	-1	2	-1	-1	2	-1	-1	2
$2F_1^e(n)$	2	2	6	6	10	20	28	46	78	122	198	324	520	842	1366
$2F_1^o(n)$	2	4	2	8	12	16	30	48	74	124	200	320	522	844	1362

$$F_1(n) = F_1(n-1) + F_1(n-2),\ n \geq 4; \quad D_1(n) = D_1(n-3),\ n \geq 5;$$
$$2F_1^e(n) = F_1(n) + D_1(n); \quad 2F_1^o(n) = F_1(n) - D_1(n), \quad n \geq 1.$$

Table 1

another rat fraction

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$F_2(n)$	**2**	**4**	**7**	**11**	**21**	39	71	131	241	443	815	1499	2757	5071	9327
$D_2(n)$	**0**	**0**	**-1**	**3**	**-1**	**-1**	-1	3	-1	-1	-1	3	-1	-1	-1
$2F_2^e(n)$	2	4	6	14	20	38	70	134	240	442	814	1502	2756	5070	9326
$2F_2^o(n)$	2	4	8	8	22	40	72	128	242	444	816	1496	2758	5072	9328

$$F_2(n) = F_2(n-1) + F_2(n-2) + F_2(n-3),\ n \geq 6; \quad D_2(n) = D_2(n-4),\ n \geq 7;$$
$$2F_2^e(n) = F_2(n) + D_2(n); \quad 2F_2^o(n) = F_2(n) - D_2(n), \quad n \geq 1.$$

Table 2

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$F_3(n)$	**2**	**4**	**8**	**15**	**26**	**51**	**99**	191	367	708	1365	2631	5071	9775	18842
$D_3(n)$	**0**	**0**	**0**	**-1**	**4**	**-1**	**-1**	**-1**	-1	4	-1	-1	-1	-1	4
$2F_3^e(n)$	2	4	8	14	30	50	98	190	366	712	1364	2630	5070	9774	18846
$2F_3^o(n)$	2	4	8	16	22	52	100	192	368	704	1266	2632	5072	9776	18838

$$F_3(n) = F_3(n-1) + F_3(n-2) + F_3(n-3) + F_3(n-4), n \geq 8; \quad D_3(n) = D_3(n-5), n \geq 9;$$
$$2F_3^e(n) = F_3(n) + D_3(n); \quad 2F_3^o(n) = F_3(n) - D_3(n), \quad n \geq 1.$$

Table 3

Consider for any $n-$CS

$$x = x_1 x_2 x_3 \dots x_n$$

the $n-$CS

$$T(x) = y_1 y_2 \dots y_n, \qquad y_i = \begin{cases} 0, & \text{if } x_i = x_{i-1}, \quad i = 1, 2, \dots, n, \\ 1, & \text{if } x_i \neq x_{i-1}, \quad x_0 = x_n. \end{cases}$$

For example

$$T(0011100101100001111) = 1010010111010010000$$
$$T(1011000111010101111) = 0110100100111111100$$
$$T(1100110011001000000) = 1010101010101100000$$

Thus, when passing over the bits of x, $T(x)$ records the changes (from 0 to 1 or from 1 to 0) by a 1, and records no change (from 0 to 0 or from 1 to 1) by a 0.

Of course

$$T(\tilde{x}) = T(x),$$

where $\tilde{x}$ is the complementary $n-$CS

$$\tilde{x} = z_1 z_2 z_3 \dots z_n, \qquad z_i = \begin{cases} 1 & \text{if } x_i = 0, \\ 0 & \text{if } x_i = 1. \end{cases}$$

However, for any two <u>different</u> $n-$CSs u and v, <u>both with first bit 1</u>, $T(u) \neq T(v)$. Indeed T is bijective between the set of 2^{n-1} $n-$CSs with first bit 1 and the set of 2^{n-1} $n-$CSs with an even number of 1's.

Thus, a $n-$CS x with first bit 1 corresponds to a n-CS $T(x)$ with an even number of 1's, and then a substring of w like bits in x corresponds to a substring of $w-1$ 0's in $T(x)$, while a substring of w alternating bits in x corresponds to a substring of $w-1$ 1's in $T(x)$. Hence

x is a $n-$CS with first bit 1 and no alternating substring
has length $> w$ (all alternating substrings have length $\leq w$)

if and only if

$T(x)$ is a $n-$CS with an even number of 1's and
all substrings of 1's have length $\leq w-1$

if and only if

$\widetilde{T(x)}$ is a $n-$CS with an even number of 0's
and all substrings of 0's have length $\leq w-1$

if and only if

n is even and $\widetilde{T(x)}$ has an even number of 1's
and all substrings of 0's have length $\leq w-1$

or

n is odd and $\widetilde{T(x)}$ has an odd number of 1's
and all substrings of 0's have length $\leq w-1$.

We have therefore that the number $a_{\leq w}(n)$ of $n-$CSs with no alternating substring of length $> w$ (equivalently, every alternating substring has length $\leq w$) is

$$a_{\leq w}(n) = \begin{cases} 2F^{e}_{w-1}(n) & \text{if } n \text{ is even,} \\ 2F^{o}_{w-1}(n) & \text{if } n \text{ is odd.} \end{cases}$$

In the case $w = 2$ we have that the number of $n-$CSs with no occurrence of 101 nor 010 is

$$a_{\leq 2}(n) = \begin{cases} 2F^{e}_{1}(n) & \text{if } n \text{ is even,} \\ 2F^{o}_{1}(n) & \text{if } n \text{ is odd.} \end{cases}$$

7039

We read these from the Table 1:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$a_{\leq 2}(n)$	2	2	2	6	12	20	30	46	74	122	200	324	522	842	1362

in agreement with Agur et al [1].

Also

x is a $n-$CS with first bit 1 and no substring of like bits has length $> w$ (all substrings of like bits have length $\leq w$)

if and only if

$T(x)$ is a $n-$CS with an even number of 1's and all substrings of 0's have length $\leq w - 1$.

Hence the number $\ell_{\leq w}(n)$ of $n-$CSs with no substring of like bits having length $> w$ (all substrings of like bits have length $\leq w$) is

$$\ell_{\leq w}(n) = 2F^{e}_{w-1}(n).$$

In the case $w = 2$ we have that the number of $n-$CSs with no occurrence of 000 nor 111 is

$$\ell_{\leq 2}(n) = 2F^{e}_{1}(n),$$

which we read from Table 1:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\ell_{\leq 2}(n)$	2	2	6	6	10	20	28	46	78	122	198	324	520	842	1366,

in agreement with Agur et al [1].

7040

References

Agur, Z., A. S. Fraenkl & S. T. Stein, The number of fixed points of the majority rule, Discrete Math. 70 (1988) 295–302.