Discrete Mathematics 70 (1988) 295–302
North-Holland

THE NUMBER OF FIXED POINTS OF THE MAJORITY RULE

Z. AGUR, A.S. FRAENKEL and S.T. KLEIN
Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot, Israel

Received 12 September 1986
Revised 17 April 1987

Formulae are derived for the number of cyclic binary strings of length n in which no single 1 occurs between two zeros and no single 0 occurs between two ones, and for the number of cyclic binary strings without substrings of the form 000 and 111. This problem is motivated by a problem of genetic information processing.

1. The combinatorial problem

We consider the following model of information processing: at the beginning (level 0), a binary string A_0 of n bits is given. This string is transformed according to a certain rule $\mathcal{M}$, producing a new string $A_1 = \mathcal{M}(A_0)$ at level 1, and the process is repeated a sufficient number of times, i.e. $A_i = \mathcal{M}(A_{i-1})$ for $i \geq 1$. One possible choice for $\mathcal{M}$ is the so-called *majority rule*, where each bit in level i is determined by its three closest neighbors in level $i-1$ by a simple majority vote, more precisely, if $A = a_0 \ldots a_{n-1}$ and $B = b_0 \ldots b_{n-1}$ and $B = \mathcal{M}(A)$ then $b_j =$ majority(a_{j-1}, a_j, a_{j+1}) for $0 \leq j < n$, where addition and subtraction on indices are modulo n, and for every triplet of bits (x, y, z), the function majority is defined by

$$\text{majority}(x, y, z) = (x \wedge y) \vee (x \wedge z) \vee (y \wedge z).$$

We thus consider the strings as being cyclic, so that the first and last bits are neighbors. The following example, where $n = 11$, should clarify the above definitions:

level 0	A_0	01010011011
level 1	A_1	10100011111
level 2	A_2	11000011111.

We are interested in the number of possible strings of length n after a sufficiently large number of levels, in other words in the number of fixed points of the function $\mathcal{M}$. Note that if n is even, the two strings $C = 1010\ldots10$ and $D = 0101\ldots01$ are such that $C = \mathcal{M}(D)$ and $D = \mathcal{M}(C)$, so starting with one of them, the sequence oscillates with period 2. It is easy to see that for every other string A_0, there exists a string A_∞ and an integer $i_0 \leq \frac{1}{2}n$ such that

$$A_\infty = \mathcal{M}^{i_0}(A_0) \quad \text{and} \quad \mathcal{M}(A_\infty) = A_\infty.$$

The single oscillation of period 2 and fixed points otherwise, are consistent with Goles and Olivos [6, 7].

Let $\mathcal{D}_n = \{A : A \text{ has } n \text{ bits and } \mathcal{M}(A) = A\}$ be the set of fixed points of $\mathcal{M}$. We evaluate the number of elements of $\mathcal{D}_n$ using the following lemma, the proof of which is immediate.

Lemma. *A cyclic binary string is invariant under $\mathcal{M}$ if and only if it contains no "widowed" bits, i.e., no single* 0 *between two* 1's *and no single* 1 *between two* 0's.

Theorem. *The number $f(n)$ of elements of $\mathcal{D}_n$, for $n \geq 3$, is*[1]

$$f(n) = \phi^n + \hat{\phi}^n + (-1)^{\lfloor (n+1)/3 \rfloor}\left(2 - \left\lceil \frac{n \bmod 3}{2} \right\rceil\right), \tag{1}$$

where $\phi = \frac{1}{2}(1 + \sqrt{5})$ is the golden ratio and $\hat{\phi} = 1 - \varphi$.

This means that asymptotically $f(n)$ behaves like φ^n, since $|\hat{\phi}| < 1$ so that $\hat{\phi}^n$ tends rapidly to zero, and the last term is a small correction: adding or subtracting 1 or 2.

Proof. For a given bit-string $x = x_0 \ldots x_{n-1}$, let $S(x)$ denote the bit-string of length n obtained by cyclically shifting the bits of x one to the right, i.e. $S(x) = x_{n-1}x_0 \ldots x_{n-2}$, and let $\mathcal{T}(x)$ be defined by

$$\mathcal{T}(x) = x \quad \text{XOR} \quad S(x).$$

For example, $\mathcal{T}(00011) = 10010$ and $\mathcal{T}(10111) = 01100$. In other words, when passing cyclically over the string x from left to right, the changes in its bits are recorded by the 1-bits of $\mathcal{T}(x)$. Hence x and its binary complement $\bar{x}$ will have the same image, $\mathcal{T}(x) = \mathcal{T}(\bar{x})$, and so $\mathcal{T}$ is not one-to-one. Not every binary string of length n is in the range of $\mathcal{T}$ since when passing cyclically over a string x, the number of changes is even. The function $G(x_0x_1 \ldots x_{n-1}) = \bar{x}_0x_1 \ldots x_{n-1}$ is a bijection between the set of n-bit strings with an even number of 1's and the one with an odd number of 1's, thus each of these sets has 2^{n-1} elements. We therefore restrict the domain of $\mathcal{T}$ to the 2^{n-1} n-bit strings with leading zero, and the range to the 2^{n-1} n-bit strings with an even number of 1's. Then $\mathcal{T}$ is one-to-one and onto.

It is easy to see that x has no widowed 0's or 1's if and only if $\mathcal{T}(x)$, considered as a cyclic string, has no adjacent 1's. We thus look for the number $g(n)$ of cyclic

[1] We use the conventional symbols $\lfloor x \rfloor$ and $\lceil x \rceil$ to denote resp. the largest integer $\leq x$ and the smallest integer $\geq x$, and $n \bmod k$ to denote the remainder of the division of n by k, i.e. $n - k\lfloor n/k \rfloor$.

n-bit strings with an even number of 1's and without adjacent 1's, from which we obtain, using the lemma, the solution to our problem, namely

$$f(n) = 2g(n). \tag{2}$$

Let us first ignore the cyclicity constraint and evaluate the number $h(n)$ of (linear) n-bit strings with an even number of 1's and no adjacent 1's. From the set of these strings, we shall then have to discard all those starting and ending with a 1-bit. That is, for $n \geq 4$, the strings to be excluded must be of the form $10x_2 \ldots x_{n-3}01$, where the substring $x_2 \ldots x_{n-3}$ contains an even number of 1's and no adjacent 1's. Thus there are exactly $h(n-4)$ elements to be discarded, which yields

$$g(n) = h(n) - h(n-4) \qquad \text{for } n \geq 4. \tag{3}$$

Let $\mathcal{H}(n)$ denote the set of n-bit strings satisfying the required constraints. The null-string belongs to $\mathcal{H}(0)$, so $h(0) = 1$. $\mathcal{H}(1) = \{0\}$, $\mathcal{H}(2) = \{00\}$ *and* $\mathcal{H}(3) = \{000, 101\}$, hence $h(1) = h(2) = 1$ and $h(3) = 2$. Generally, for $n \geq 2$, let us partition the $h(n)$ elements of $\mathcal{H}(n)$ into $\mathcal{H}_0(n)$, the set of those ending with zero, and $\mathcal{H}_1(n)$, the set of those ending with 1. If a string ends with zero, the constraints on its $n-1$ leftmost bits are: no adjacent 1's and even number of 1's, so there are $h(n-1)$ such strings. If a string ends with 1, the next to last bit must be zero and the constraints on the $n-2$ leftmost bits are: no adjacent 1's and *odd* number of 1's.

We therefore consider the set $\mathcal{F}(n)$ of n-bit strings with no adjacent 1's, but without condition on the parity of the number of 1's. Then $\mathcal{H}_1(n)$ is the complement of $\mathcal{H}(n-2)$ in $\mathcal{F}(n-2)$, where to each string the suffix 01 has been appended, thus

$$h(n) = h(n-1) + |\mathcal{F}(n-2)| - h(n-2) \quad \text{for } n \geq 2.$$

The set $\mathcal{F}(n)$ is related to the *binary Fibonacci numeration system* (see Fraenkel [4] or Knuth [10, Exercise 1.2.8–34]). Let $F(n)$ be the Fibonacci sequence defined by $F(0) = 0$, $F(1) = 1$, $F(n) = F(n-1) + F(n-2)$ for $n \geq 2$. Every integer K in the range $0 \leq K \leq F(n+2)$ has a unique binary representation of precisely n bits, $K = k_n k_{n-1} \ldots k_1$ such that $K = \sum_{i=1}^{n} k_i F(i+1)$ and such that there are no adjacent 1's in this representation of K ($n \geq 1$). Thus $|\mathcal{F}(n)| = F(n+2)$ and we get the following recurrence for $h(n)$:

$$h(n) - h(n-1) + h(n-2) = F(n) \quad \text{for } n \geq 2.$$

The solution of this recurrence relation is

$$h(n) = A\alpha_+^n + B\alpha_-^n + \tfrac{1}{2}F(n+2) \quad \text{for } n \geq 0,$$

where $\alpha_+ = (1 + i\sqrt{3})/2$ and $\alpha_- = \frac{1}{2}(1 - i\sqrt{3})$ are the roots of the characteristic equation $\alpha^2 - \alpha + 1 = 0$, and where the coefficients A and B are constants fixed by the boundary conditions $h(0) = h(1) = 1$:

$$A = \frac{3 + i\sqrt{3}}{12} \quad \text{and} \quad B = \frac{3 - i\sqrt{3}}{12}.$$

Returning now to Equation (3), we get for $n \geq 4$:

$$g(n) = A(\alpha_+^4 - 1)\alpha_+^{n-4} + B(\alpha_-^4 - 1)\alpha_-^{n-4} + \tfrac{1}{2}(F(n+2) - F(n-2)), \tag{4}$$

but

$$\alpha_+^4 - 1 = -6A \quad \text{and} \quad 6A^2 = \tfrac{1}{2}\alpha_+ \tag{5}$$

and similarly

$$\alpha_-^4 - 1 = -6B \quad \text{and} \quad 6B^2 = \tfrac{1}{2}\alpha_-. \tag{6}$$

It is well-known that $F(n) = (1/\sqrt{5})(\phi^n - \hat{\phi}^n)$ so the last term on the right hand side of (4) becomes

$$\frac{1}{2\sqrt{5}}\phi^{n-2}(\phi^4 - 1) - \frac{1}{2\sqrt{5}}\hat{\phi}^{n-2}(\hat{\phi}^4 - 1) \tag{7}$$

but

$$\frac{\phi^4 - 1}{\sqrt{5}} = \phi^2 \quad \text{and} \quad \frac{\hat{\phi}^4 - 1}{\sqrt{5}} = -\hat{\phi}^2, \tag{8}$$

so we get from (2) and (4)–(8)

$$f(n) = \phi^n + \hat{\phi}^n - \alpha_+^{n-3} - \alpha_-^{n-3} \quad \text{for } n \geq 4. \tag{9}$$

Incidentally, this formula holds also for $n = 3$. One can obtain a more compact form of the solution by noting that α_+ and α_- are primitive 6th roots of unity: $\alpha_+ = e^{i\pi/3}$ and $\alpha_- = e^{-i\pi/3}$. Therefore

$$(\alpha_+, \alpha_+^2, \alpha_+^3, \alpha_+^4, \ldots) = (\alpha_+, -\alpha_-, -1, -\alpha_+, \alpha_-, 1, \alpha_+, \ldots)$$
$$(\alpha_-, \alpha_-^2, \alpha_-^3, \alpha_-^4, \ldots) = (\alpha_-, -\alpha_+, -1, -\alpha_-, \alpha_+, 1, \alpha_-, \ldots),$$

thus $\{\alpha_+^j + \alpha_-^j\}_{j=0}^{\infty}$ is the periodic sequence $(2, 1, -1, -2, -1, 1, 2, 1, -1, \ldots)$. The period is of length 6, the first two and last elements of each period being positive, the other three negative; $|\alpha_+^j + \alpha_-^j| = 2$ if j is a multiple of 3, and for the other values of i, $|\alpha_+^i + \alpha_+^j| = 1$. This suggests the representation

$$\alpha_+^j + \alpha_-^j = (-1)^{\lfloor (i+1)/3 \rfloor}\left(2 - \left[\frac{j \bmod 3}{2}\right]\right) \quad \text{for } i \geq 0, \tag{10}$$

from which (1) follows by $\alpha_+^3 = \alpha_-^3 = -1$. □

Using the same technique, we can also evaluate the number of cyclic strings of length n having no consecutive sequence of three 0's or three 1's.

Corollary. *The number $f'(n)$, for $n \geqslant 3$, of binary strings $x_0 \ldots x_{n-1}$ such that $x_i x_{i\oplus 1} x_{i\oplus 2} \neq 000$ and $x_i x_{i\oplus 1} x_{i\oplus 2} \neq 111$ for $0 \leqslant i < n$, where $\oplus$ denotes addition modulo n is*

$$f'(n) = \phi^n + \hat{\phi}^n + (-1)^{\lceil (n \bmod 3)/2 \rceil}\left(2 - \left\lceil \frac{n \bmod 3}{2} \right\rceil\right). \tag{11}$$

Proof. We define $\bar{\mathcal{T}}(x)$ as the complement of $\mathcal{T}(x) = x$ XOR $S(x)$, i.e. when passing cyclically over the string x from left to right, the changes are now recorded as zeros. Again the number of changes is even for every x, so that $\bar{\mathcal{T}}(x)$ has an even number of zeros. A cyclic string x contains no 000 or 111 substring if and only if there are no adjacent 1's in the cyclic string $\bar{\mathcal{T}}(x)$. We thus define $g'(n)$ as the number of cyclic binary strings of length n with no adjacent 1's and having an even number of zeros. Then we have as before.

$$f'(n) = 2g'(n). \tag{12}$$

For n even, a string has an even number of zeros if and only if it has an even number of 1's. Thus

$$g'(n) = g(n) \quad \text{for } n \text{ even.} \tag{13}$$

For n odd, we again consider the set $\mathcal{F}(n)$ of $F(n+2)$ strings of length n with no adjacent 1's (ignoring the cyclicity constraint). We first discard all the strings starting and ending with 1, their number being $F(n-2)$. From the remaining strings, $g(n)$ have an even number of 1's, so the number of strings with an even number of zeros is

$$g'(n) = F(n+2) - F(n-2) - g(n) \quad \text{for } n \text{ odd.} \tag{14}$$

From (7) and (8) we already know that $F(n+2) - F(n-2) = \phi^n + \hat{\phi}^n$, thus we get from (9) and (12)–(14) (using again $\alpha_+^3 = \alpha_-^3 = -1$)

$$f'(n) = \phi^n + \hat{\phi}^n + (-1)^n(\alpha_+^n + \alpha_-^n) \quad \text{for } n \geqslant 3.$$

Now $\{(-1)^i(\alpha_+^i + \alpha_-^i)\}_{i=0}^{\infty}$ is the periodic sequence $(2, -1, -1, 2, -1, -1, 2, \ldots)$, the period being of length 3. This suggests the representation

$$(-1)^j(\alpha_+^j + \alpha_-^j) = (-1)^{\lceil (j \bmod 3)/2 \rceil}\left(2 - \left\lceil \frac{j \bmod 3}{2} \right\rceil\right) \quad \text{for } j \geqslant 0,$$

from which (11) follows. □

For even n, there is a simple direct way to see that $f'(n) = f(n)$, considering the function $\mathcal{R}$ on the set of n-bit strings defined by

$$\mathcal{R}(x_0 x_1 x_2 x_3 \ldots x_{n-2} x_{n-1}) = x_0 \bar{x}_1 x_2 \bar{x}_3 \ldots x_{n-2} \bar{x}_{n-1},$$

i.e. $\mathcal{R}$ complements the bit in odd-indexed positions. Clearly $\mathcal{R}$ is a bijection and a cyclic string x contains no widowed 0's or 1's if and only if the cyclic string $\mathcal{R}(x)$ contains neither 000 nor 111 as substring.

Table 1

n	3	4	5	6	7	8	9	10	11	12	13	14	15
$f(n)$	2	6	12	20	30	46	74	122	200	324	522	842	1362
$f'(n)$	6	6	10	20	28	46	78	122	198	324	520	842	1366

Table 1 gives the first few values of $f(n)$ and $f'(n)$.

The techniques of this note can be generalized to evaluate the number of cyclic or non-cyclic strings with forbidden substrings of the form $1010\ldots$ and $0101\ldots$, or $0\ldots0$ and $1\ldots1$. After applying $\mathcal{T}$ or $\bar{\mathcal{T}}$, one of the constraints becomes that no m consecutive 1's appear in the string, for some $m \geqslant 2$. These strings are related to the mth order Fibonacci numeration system which is based on generalized Fibonacci numbers satisfying the recurrence

$$F^{(m)}(n) = F^{(m)}(n-1) + F^{(m)}(n-2) + \cdots + F^{(m)}(n-m)$$

(see [4] or Knuth [11, Exercise 5.4.2–10]).

2. Background and biological motivation

The majority rule considered above is a special case of discrete iteration models studied in Robert [13], which includes an extensive literature on the subject. Robert considers various transition functions $\mathcal{M}$ applied to general graphs and parallel as well as serial modes of operation. The question of the number of fixed points in these more general models is left open. Our model corresponds to a simple closed path of n vertices, where at each step, the rule is applied simultaneously on all the vertices. A serial mode of operation would mean applying the transition rule consecutively on the vertices, following some predetermined ordering. Our result on the number of fixed points holds also for serial application of the majority rule, if a linear order of scanning is chosen. What changes is the rate of convergence for a given configuration to its fixed point, and the fact that now also the above mentioned strings C and D converge. More properties of the majority rule can be found in Poljak and Sura [12] and Goles [5].

A major problem in biology is the potential number of forms, or *phenotypes*, of a given system. For example, our immune response depends to a great extent on the number of different antibody forms that our body can produce. Each antibody is the result of a multi-level processing of a DNA sequence, that is, a concatenation of several basic gene segments. Since this processing is highly nonlinear, no simple relation exists normally between the number of basic genes and the resulting set of phenotypes, say the antibodies.

"Genetic nets" of the above form have been studied by Kauffman (e.g. [8, 9]).

Each element in Kauffman's network receives k inputs and is assigned a Boolean rule at random out of the 2^{2^k} possible rules. These networks are difficult to analyse and their study proceeds by simulations. In another model, the regulation of biological systems is described as asynchronous Boolean networks whose elements affect the rate of operation of each other; see Thomas [14]. Conditions for the existence of fixed points and stable cycles are derived. However, this model appears to be intractable for formal analysis for any reasonably large number of elements; see Weisbuch [15].

The majority rule considered in Section 1 constitutes a simplified model of information processing, in which the initial binary string A_0 corresponds to the initial DNA sequence and $A_\infty = \mathcal{M}^{t_0}(A_0)$ is a phenotype. This model of biological processing of genetic information has been initiated in Agur and Kerszberg [2] and is further applied in Agur [1]. Although the majority rule is a relatively simple rule, it incorporates the most fundamental properties of biological information handling, namely an error-damping property and a many/one and one/many type of hierarchy (for further details see [1, 2]).

From the above theorem it follows that for the simplified model the ratio between the size of the set of basic DNA sequences of length n to that of the set of resulting phenotypes is asymptotically $(2/\phi)^n$. In fact this ratio is $(2/\phi)^n + O(c^{-n})$ for a constant $c > 1$.

For more complicated transformation rules, even the form of A_∞ may be very hard to predict. A case in point is Conway's "game" of *life*, with its two-dimensional transformation rule, for which this prediction is as hard as some of the hardest problems in mathematics [3, Ch. 25].

References

[1] Z. Agur, Resilience and variability in pathogens and hosts, *IMA J. Math. Applied Med. and Biol.* 4 (1987) 295–307

[2] Z. Agur and M. Kerszberg, The emergence of phenotypic novelties through progressive genetic change, *Amer. Natur.* 129 (1987) 862–875.

[3] E.R. Berlekamp, J.H. Conway and R.K. Guy, Winning ways for your mathematical plays (Academic Press, London, 1982).

[4] A.S. Fraenkel, Systems of numeration, Amer. Math. Monthly 92 (1985) 105–114.

[5] E. Goles, Dynamics of positive automata networks, Theor. Comp. Sc. 41 (1985) 19–32.

[6] E. Goles and J. Olivos, Comportement périodique des fonctions à multiseuil, Inf. and Control 45 (1980) 300–313.

[7] E. Goles and J. Olivos, Periodic behaviour of generalized threshold functions, Discrete Math. 30 (1980) 187–189.

[8] S.A. Kauffman, Gene regulation networks: a theory for their global structure and behaviour, in Current Topics in Development Biology 6 (Academic Press, New York, 1971) 145–182.

[9] S.A. Kauffman, Emergent properties in random complex automata, Physica 10D (1984) 145–156.

[10] D.E. Knuth, The Art of Computer Programming, Vol I, Fundamental Algorithms (Addison-Wesley, Reading, MS, 1973).

[11] D.E. Knuth, The Art of Computer Programming, Vol III, Sorting and Searching (Addison-Wesley, Reading, MA, 1973).

[12] S. Poljak and M. Sura, On periodical behaviour in societies with symmetric influences, Combinatorica 3 (1983) 119–121.

[13] F. Robert, Discrete Iterations (Springer, Berlin, 1986).

[14] R. Thomas, Fully asynchronous logical description of networks comprising feedback loops, in Lecture Notes in Biomathematics 49 (Springer, Berlin, 1983) 189–201.

[15] G. Weisbuch, Networks of automata and biological organization, J. Theor. Biol. 121 (1986) 255–267.