МАТЕМАТИЧКИ ВЕСНИК
46 (1994), 93–98

UDC 519.111
оригинални научни рад
research paper

BINARY SEQUENCES WITHOUT $0\underbrace{\mathbf{11\ldots11}}_{\mathbf{k-1}}0$ FOR FIXED k

Rade Doroslovački

Abstract. The paper gives a special construction of those words (binary sequences) of length n over alphabet $\{0,1\}$ in which the subword $0\underbrace{11\ldots11}_{k-1}0$ is forbidden for some natural number k. This number of words is counted in two different ways, which gives some new combinatorial identities.

1. Definitions and notations

Let $X=\{0,1\}$ denote 2-element set of digits. X is called an alphabet. By X^n we shall denote the set of all strings of length n over alphabet X, i.e.

$$X^n=\{\,x_1x_2\ldots x_n \mid x_1\in X\wedge x_2\in X\wedge\cdots\wedge x_n\in X\,\},$$

the only element of X^0 is the empty string, i.e. the string of the length 0. The set of all finite strings over alphabet X is

$$X^*=\bigcup_{n\geq 0}X^n.$$

If S is a set, then $|S|$ is the cardinality of S. By $\lceil x\rceil$ and $\lfloor x\rfloor$ we denote the smallest integer $\geq x$ and the greatest integer $\leq x$, respectively. By $\ell_0(p)$ and $\ell_1(p)$ we denote the number of zeros and ones respectively in the string $p\in X^*$. $N_n=\{1,2,\ldots,n\}$, $N_n=\emptyset$ for $n\leq 0$, $\binom{n}{k}=0$ iff $n<k$ and $[x]$ is the nearest integer to x.

2. Results and discusion

Now we shall construct and enumerate the set of words

$$A_k(n)=\{\,\mathbf{x_n}\mid \mathbf{x_n}=x_1x_2\ldots x_n\in X^n,\ (\forall i\in N_{n-k})(x_ix_{i+1}\ldots x_{i+k}\neq 0\underbrace{1\ldots1}_{k-1}0)\,\}$$

AMS *Mathematics Subject Classification* (1991) 05A15
Key words and phrases: subword.

for each natural number k. It is known that

$$a_1(n) = |A_1(n)| = \sum_{i=0}^{\lceil n/2 \rceil} \binom{n-i+1}{i} = \left[\frac{5+3\sqrt{5}}{10} \left(\frac{1+\sqrt{5}}{2} \right)^n \right] \tag{1}$$

(Fibonacci numbers) where

$$A_1(n) = \{ \mathbf{x_n} \mid \mathbf{x_n} = x_1 x_2 \ldots x_n \in X^n,\ (\forall i \in N_{n-1})\, (x_i x_{i+1} \neq 00) \}.$$

In [5] it is shown that the following theorem is valid.

Theorem 1.

$$a_2(n) = |A_2(n)| = 1 + \sum_{i=1}^{n} \sum_{j=0}^{i-1} \binom{i-1}{j} \binom{n-i-j+1}{j+1} = \left[\frac{2\alpha^2+1}{2\alpha^2-2\alpha+3} \alpha^n \right]$$

where

$$A_2(n) = \{ \mathbf{x_n} \mid \mathbf{x_n} = x_1 x_2 \ldots x_n \in X^n,\ (\forall i \in N_{n-2})\, (x_i x_{i+1} x_{i+2} \neq 010) \}$$

and

$$\alpha = \frac{1}{6}(4 + \sqrt[3]{100+4\sqrt{621}} + \sqrt[3]{100-4\sqrt{621}}) \approx 1,754877666247.$$

$A_2(n)$ is the set of all words of length n over alphabet $\{0,1\}$ with forbidden subword 010.

Theorem 2.

$$|A_3(n)| = 1 + \sum_{i=1}^{n} \sum_{j=0}^{n-i} \sum_{k=0}^{\lfloor \frac{n-i-j}{3} \rfloor} \binom{i-1}{j} \binom{i-1-j}{k} \binom{n-i-j-2k+1}{k+1}$$

where

$$A_3(n) = \{ \mathbf{x_n} \mid \mathbf{x_n} = x_1 x_2 \ldots x_n \in X^n,\ (\forall i \in N_{n-3})(x_i x_{i+1} x_{i+2} x_{i+3} \neq 0110) \}.$$

Proof. Now we shall construct this set of words $A_3(n)$ in some special way, which gives the result for $|A_3(n)|$. We make a partition of the set $A_3(n)$ into subsets $A_3^i(n)$ which contain exatly i zeros.

$$A_3^i(n) = \{ \mathbf{x_n} \mid \mathbf{x_n} \in A_3(n),\ (\forall s \in N_{n-3})\, (x_s x_{s+1} x_{s+2} x_{s+3} \neq 0110),\ \ell_0(\mathbf{x_n}) = i \}.$$

First we write i zeros and then we write one of the letters from the set $\{p, q, \lambda\}$ on the $i-1$ $(1 \leq i \leq n)$ places between i zeros where $p = 1$, $q = 111$ and λ is the empty letter. Letter λ is the letter with property that if λ is writen between two zeros then actually nothing is written. Let j be the number of appearances of the letter p and k is the number of appearances of the letter q. We choose j places from

$i-1$ places for letters p and after that we choose k places from $i-1-j$ places for leters q. This we can do in

$$\binom{i-1}{j}\binom{i-1-j}{k} \tag{2}$$

different ways. Now we have only $n-i-j-3k$ ones, which must be put on k places where we have subwords 111 as well as into the regions in front of and behind the word, that is into $k+2$ regions in all. It can be done in

$$\binom{n-i-j-2k+1}{k+1} \quad \text{ways .} \tag{3}$$

Thus from (2), (3) and

$$|A_3(n)| = \sum_{i=0}^{n} |A_3^i(n)| = 1 + \sum_{i=1}^{n} |A_3^i(n)|$$

Theorem 2 follows. ■

Theorem 3.

$$a_3(n) = |A_3(n)| = \left[\frac{2\alpha^3+1}{2\alpha^3-3\alpha+4}\alpha^n\right]$$

where

$$\alpha = \frac{1}{2}\left(1+\sqrt{3+2\sqrt{5}}\right) \approx 1,866760399.$$

Proof. Words $\mathbf{x_n} \in A_3(n)$ are obtained from other words $\mathbf{x_{n-1}} \in A_3(n-1)$ by appending 0 or 1 in front of them. Let $\mathbf{x_{n-1}} \in A_3(n-1)$, $\mathbf{x_{n-3}} \in A_3(n-3)$ and $\mathbf{x_{n-4}} \in A_3(n-4)$. Then $1\mathbf{x_{n-1}} \in A_3(n)$, $0110\mathbf{x_{n-4}} \notin A_3(n)$ and $0111\mathbf{x_{n-4}} \in A_3(n)$, which means that $011\mathbf{x_{n-3}} \in A_3(n)$ if and only if $\mathbf{x_{n-3}}$ begins with letter 1. This implies the reccurence relation

$$a_3(n) = 2a_3(n-1) - a_3(n-3) + a_3(n-4)$$

whose characteristic equation is $x^4 - 2x^3 + x - 1 = 0$ and whose roots are

$$\alpha = \frac{1}{2}\left(1+\sqrt{3+2\sqrt{5}}\right), \quad \beta = \frac{1}{2}\left(1-\sqrt{3+2\sqrt{5}}\right)$$

$$\gamma = \frac{1}{2}\left(1+i\sqrt{2\sqrt{5}-3}\right) \quad \text{and} \quad \delta = \frac{1}{2}\left(1-i\sqrt{2\sqrt{5}-3}\right).$$

The explicit formula for $a_3(n)$ is

$$a_3(n) = C_1\alpha^n + C_2\beta^n + C_3\gamma^n + C_4\delta^n$$

where

$$C_1 = \frac{2\alpha^3+1}{2\alpha^3-3\alpha+4}, \quad C_2 = \frac{2\beta^3+1}{2\beta^3-3\beta+4}$$

$$C_3 = \frac{2\gamma^3+1}{2\gamma^3-3\gamma+4}, \quad \text{and} \quad C_4 = \frac{2\delta^3+1}{2\delta^3-3\delta+4}.$$

Since $|\beta| < 1$, $|\gamma| < 1$ and $|\delta| < 1$ we obtain Theorem 3. ■

Thus from Theorem 2 and Theorem 3 follows

Corollary 1.

$$|A_3(n)| = 1 + \sum_{i=1}^{n} \sum_{j=0}^{n-i} \sum_{k=0}^{\lfloor \frac{n-i-j}{3} \rfloor} \binom{i-1}{j} \binom{i-1-j}{k} \binom{n-i-j-2k+1}{k+1}$$
$$= \left[\frac{2\alpha^2+1}{2\alpha^3-3\alpha+4} \alpha^n \right], \qquad \textit{where} \quad \alpha = \frac{1}{2}\left(1+\sqrt{3+2\sqrt{5}}\right).$$

Theorem 4.

$$|A_k(n)| = 1 + \sum_{i=1}^{n} \sum_{j_1=0}^{n-i} \sum_{j_2=0}^{\lfloor \frac{n-i-S_1}{2} \rfloor} \sum_{j_3=0}^{\lfloor \frac{n-i-S_2}{3} \rfloor} \cdots \sum_{j_{k-2}=0}^{\lfloor \frac{n-i-S_{k-3}}{k-2} \rfloor} \sum_{\ell=0}^{\lfloor \frac{n-i-S_{k-2}}{k} \rfloor}$$
$$\prod_{m=0}^{m=k-3} \binom{i-1-s_m}{j_{m+1}} \binom{i-1-s_{k-2}}{\ell} \binom{n-i-S_{k-2}-(k-1)\ell+1}{\ell+1}$$

where $s_k = j_1 + j_2 + \cdots + j_k$, $s_0 = 0$, $S_k = j_1 + 2j_2 + \cdots + kj_k$ *and*

$$A_k(n) = \{\mathbf{x_n} \mid \mathbf{x_n} = x_1x_2\ldots x_n \in X^n, (\forall s \in N_{n-k})(x_sx_{s+1}\ldots x_{s+k} \neq 0\underbrace{1\ldots1}_{k-1}0)\}.$$

Proof. We partition the set $A_k(n)$ into subsets $A_k^i(n)$ which contain exactly i zeros i.e.

$$A_k^i(n) = \{\, \mathbf{x_n} \mid \mathbf{x_n} = x_1x_2\ldots x_n \in X^n,$$
$$(\forall s \in N_{n-k})(x_sx_{s+1}\ldots x_{s+k} \neq 0\underbrace{1\ldots1}_{k-1}0), \ell_0(\mathbf{x_n}) = i\,\}.$$

Now we shall construct words from $A_k^i(n)$ in the following way. First we write i zeros and then we write one of the letters from the alphabet $\{q_1, q_2, \ldots, q_{k-2}, r, \lambda\}$ on $i-1$ places between i zeros where $q_m = \underbrace{11\ldots1}_{m}$, for $m \in \{1, 2, \ldots, k-2\}$, $r = \underbrace{11\ldots1}_{k}$ and λ is the empty letter. Let j_m be the number of letters q_m, and ℓ the number of letters r. We choose j_1 places from $i-1$ places for letters q_1, j_2 places from $i-1-j_1$ places for letters q_2, ... , j_{k-2} places from $i-1-s_{k-3}$ places for leters q_{k-2} and ℓ places from $i-1-s_{k-2}$ places for letters r. It can be done in

$$\prod_{m=0}^{m=k-3} \binom{i-1-s_m}{j_{m+1}} \binom{i-1-s_{k-2}}{\ell} \tag{4}$$

different ways, where $s_k = j_1 + j_2 + \cdots + j_k$ and $s_0 = 0$. There remains to write $n-i-S_{k-2}-k\ell$ letters 1 on ℓ regions which already contain r, as well as into the regions in front of and behind the word, that is into $\ell+2$ regions in all. It can be done in

$$\binom{n-i-S_{k-2}-(k-1)\ell+1}{\ell+1} \tag{5}$$

different ways, where $S_k = j_1 + 2j_2 + \cdots + kj_k$. Thus from (4), (5) and $|A_k(n)| = \sum_{i=0}^{n} |A_k^i(n)|$ Theorem 4 follows. ∎

Theorem 5.

$$|A_k(n)| = [C(k,\alpha)\alpha^n]$$

for large enough values of n, where α is the unique real root of equation

$$x^{k+1} - 2x^k + x - 1 = 0$$

which lies betwen 1 and 2 and $C(k,\alpha)$ is the rational function of α and k.

Proof. Words $\mathbf{x_n} \in A_k(n)$ are obtained from other words $\mathbf{x_{n-1}} \in A_k(n-1)$ by appending 0 or 1 in front of them. Let

$$\mathbf{x_{n-1}} \in A_k(n-1),\ \mathbf{x_{n-k}} \in A_k(n-k) \quad \text{and} \quad \mathbf{x_{n-k-1}} \in A_k(n-k-1).$$

Then

$$1\mathbf{x_{n-1}} \in A_k(n),\ 0\underbrace{11\ldots1}_{k}\mathbf{x_{n-k-1}} \in A_k(n),\ 0\underbrace{11\ldots1}_{k-1}0\mathbf{x_{n-k-1}} \notin A_k(n)$$

which means that $0\underbrace{11\ldots1}_{k-1}\mathbf{x_{n-k}} \in A_k(n)$ if and only if $\mathbf{x_{n-k}}$ begins with the letter 1. This implies the reccurence relation

$$a_k(n) = 2a_k(n-1) - a_k(n-k) + a_k(n-k-1)$$

whose characteristic equation is $x^{k+1} - 2x^k + x - 1 = 0$ which has only one real root α for $k = 2m$, $m \in N$. This real root lies between 1 and 2. If $k = 2m+1$, $m \in N \cup \{0\}$, then the characteristic equation has only two real roots α and β where $\alpha \in (1,2)$ and $\beta \in (-1,0)$. The complex roots have modules less than 1. Because of that it follows that $a_k(n) = [C(k,\alpha)\alpha^n]$ for large enough values of n, where $C(k,\alpha)$ is rational function of α and k. ∎

Corollary 2.

$$|A_4(n)| = 1 + \sum_{i=1}^{n}\sum_{j=0}^{n-i}\sum_{k=0}^{\lfloor\frac{n-i-j}{2}\rfloor}\sum_{\ell=0}^{\lfloor\frac{n-i-j-2k}{4}\rfloor} \binom{i-1}{j}\binom{i-1-j}{k}\binom{i-1-j-2k}{\ell}\binom{n-i-j-2k-3\ell+1}{\ell+1}$$
$$= \left[\frac{4\alpha^4-\alpha+2}{4\alpha^4-4\alpha^2+3\alpha+2}\alpha^n\right],$$

i.e. $C(4,\alpha)=\frac{4\alpha^4-\alpha+2}{4\alpha^4-4\alpha^2+3\alpha+2}$ *and* α *is unique real root of equation* $x^5-2x^4+x-1=0$ *whose complex roots are with modules less than 1.*

REFERENCES

[**1**] Austin Richard and Guy Richard, *Binary sequences without isolated ones*, The Fibonacci Quarterly, **16**, 1 (1978) 84–86

[**2**] Cvetković, D., *The generating function for variations with restrictions and paths of the graph and self complementary graphs*, Univ. Beograd, Publ. Elektrotehnički fakultet, serija Mat. Fiz. **320–328** (1970) 27–34

[**3**] Doroslovački, R., *The set of all words over alphabet* $\{0,1\}$ *of lenght* n *with the forbidden subword* $11\dots1$, Rev. of Res., Fac. of Sci. math. ser., Novi Sad, **14**, 2 (1984) 167–173

[**4**] Doroslovački, R., *The set of all words of length* n *over any alphabet with a forbidden good subword*, Rev. of Res., Fac. of Sci. math. ser., Novi Sad, **23**, 2 (1993) 239–244

[**5**] Doroslovački, R., *The set of all the words of length* n *over alphabet* $\{0,1\}$ *with any forbidden subword of lenght three*, Rev. of Res., Fac. of Sci. math. ser., Novi Sad. (in print)

[**6**] Einb, J. M., *The enumeration of bit sequences that satisfy local criteria*, Publ. de L'Instit. Math. Beograd, **27 (41)** (1980) 51–56.

(received 14.09.1994.)

Department of Mathematics, Faculty of Engineering University of Novi Sad, 21000 Novi Sad, Trg Dositeja Obradovića 6, Yugoslavia