NOVI SAD J. MATH.
VOL. 28, NO. 2 1998, 127-133

BINARY n-WORDS WITHOUT THE SUBWORD 1010...10

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

Olivera Marković
Faculty of Education University of Kragujevac
31000 Užice, Trg Svetog Save 36, Yugoslavia

Abstract

The paper gives a special construction of those words (binary sequences) of length n over the alphabet $\{0,1\}$ in which the subword $\underbrace{1010...10}_{2p}$ is forbidden for some natural number p, where p is fixed. This number of words is counted in two different ways, which gives some new combinatorial identities.

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

1. Definitions and notations

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

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

the only element of X^0 is the empty word, i.e. the word of the length 0. The set of all finite words over the 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 $\ell_q(p)$ we denote the number of subwords q in the word $p \in X^*$. The set N is the set of natural numbers, $N_n = \{1, 2, ..., n\}$, $N_n = \emptyset$ for $n \leq 0$, $\binom{n}{k} = 0$ iff $n < k$ and $[x]$ is the nearest integer to x. If $x_1 x_2 ... x_n \in X^n$, then $\mathbf{x_n} = x_1 x_2 ... x_n$.

2. Results and discusion

Now we shall construct and enumerate the set of words

$$A_p(n) = \{\mathbf{x_n} \mid \mathbf{x_n} \in X^n, \ (\forall i \in N_{n-2p+1})(x_i x_{i+1} ... x_{i+2p-1} \neq \underbrace{1010...10}_{2p})\}$$

for each natural number p. We shall denote $|A_p(n)|$ with $a_{p,n}$. It is obvious that

$$a_{1,n} = \mid A_1(n) \mid = n + 1$$

On the other hand it is known [4] that

$$a_{1,n} = \mid A_1(n) \mid = \sum_{i=0}^{\lfloor n/2 \rfloor} (-1)^i \binom{n-i}{i} 2^{n-2i}$$

where

$$A_1(n) = \{\mathbf{x_n} \mid \mathbf{x_n} = x_1 x_2 ... x_n \in X^n, \ (\forall i \in N_{n-1}) \ (x_i x_{i+1} \neq 10)\}.$$

Now we have the theorems

Theorem 1.

$$a_{1,n} = \sum_{i=0}^{\lfloor n/2 \rfloor} (-1)^i \binom{n-i}{i} 2^{n-2i} = n + 1.$$

Theorem 2. *[6]*

$$a_{2,n} = |A_2(n)| = n+1+\sum_{k=1}^{\lfloor\frac{n+1}{3}\rfloor}\ \sum_{i_1+i_2+\ldots+i_{k+1}=n-2k+2} i_1(i_2+1)(i_3+1)\ldots(i_k+1)i_{k+1}$$

where $i_1, i_2, \ldots, i_{k+1} \in N$ *and*

$$A_2(n) = \{\mathbf{x_n} | \mathbf{x_n} = x_1x_2\ldots x_n \in \{0,1\}^n \wedge (\forall i \in N_{n-3}) x_i x_{i+1} x_{i+2} x_{i+3} \neq 1010\}$$

Theorem 3. *[6]*

$$a_{2,n} = |A_2(n)| = \left[\frac{2\alpha^3+2\alpha-1}{2\alpha^3-2\alpha^2+6\alpha-4}\alpha^n\right] \quad \text{where}$$

$$\alpha = \frac{1+\sqrt{2}+\sqrt{2\sqrt{2}-1}}{2} \approx 1,883203506.$$

Corollary 1.

$$\lim_{n\to\infty}\frac{1}{\alpha^n}\sum_{k=1}^{\lfloor\frac{n+1}{3}\rfloor}\ \sum_{i_1+i_2+\ldots+i_{k+1}=n-2k+2} i_1(i_2+1)(i_3+1)\ldots(i_k+1)i_{k+1} =$$

$$= \frac{2\alpha^3+2\alpha-1}{2\alpha^3-2\alpha^2+6\alpha-4} \quad \text{where } i_1, i_2, \ldots, i_{k+1} \in N$$

Lemma 1.

$$a'_{2,n} = \sum_{k=0}^{\lfloor\frac{n+1}{3}\rfloor}\ \sum_{m_1+m_2+\ldots+m_{k+1}=n-2k} (m_1+1)(m_2+1)\ldots(m_{k+1}+1)$$

where $m_1, m_2, \ldots, m_{k+1} \in N$ *and* $a'_{2,n}$ *is the number of all words of the length n over the alphabet* $\{0,1\}$ *with the forbidden subword 1010 and which neither begin nor end with 10.*

Lemma 2.

$$a''_{2,n} = n+1+\sum_{k=1}^{\lfloor\frac{n+1}{3}\rfloor}\ \sum_{i_1+i_2+\ldots+i_{k+1}=n-2k+1} (i_1+1)(i_2+1)\ldots i_{k+1}$$

where $i_1, i_2, \ldots, i_{k+1} \in N$ *and* $a''_{2,n}$ *is the number of all words of the length n over the alphabet* $\{0,1\}$ *with the forbidden subword 1010 and which only do not begin with 10.*

The proofs of Lemma 1 and Lemma 2 follows from the proof of Theorem 2.

It is obvious that $a''_{2,n}$ is the number of all words of the length n over the alphabet $\{0,1\}$ with the forbidden subword 1010 and which only do not end with 10.

Theorem 4.

$$a_{3,n} = a_{2,n} + \sum_{k=1}^{\lfloor \frac{n+1}{5} \rfloor} \sum_{i_1+i_2+\ldots+i_{k+1}=n-4k} a''_{2,i_1} a'_{2,i_2} \ldots a'_{2,i_k} a''_{2,i_{k+1}}$$

where i_1, $i_{k+1} \in N \cup \{0\}$ *and* $i_2, i_3, \ldots, i_k \in N$

Proof. Let us count the number of all words of the length n over the alphabet $\{0,1\}$ with the forbidden subword 101010, i.e. the number of words in set $A_3(n)$. We make a partition of the set $A_3(n)$ into subsets $A_3^k(n)$, where $A_3^k(n)$ is the set of all those words of the length n over the alphabet $\{0,1\}$ which contain exactly k subwords 1010 ($\mathbf{x_n} \in A_3^k(n) \Rightarrow l_{1010}(\mathbf{x_n}) = k$) and do not contain the subword 101010. In the same way as in Theorem 2 we complete the proof by using Lemma 1 and Lemma 2. □

Theorem 5.

$$a_{3,n} = |A_3(n)| = \left[\frac{\alpha^6 + \alpha^4 + \alpha^2}{-\alpha^4 + 4\alpha^3 - 3\alpha^2 + 8\alpha - 5} \alpha^n \right]$$

where $\alpha = 1,974818708297706\ldots$

Proof. In the same way as in Theorem 3 we have the recurrence relation

$$a_{3,n} = 2a_{3,n-1} - a_{3,n-2} + 2a_{3,n-3} - a_{3,n-4} + 2a_{3,n-5} - a_{3,n-6},$$

whose characteristic equation is

$$x^6 - 2x^5 + x^4 - 2x^3 + x^2 - 2x + 1 = 0,$$

and whose roots we denote with $\alpha, \beta, \gamma, \delta, \epsilon, \zeta$. This equation has two real roots α and β,where $\alpha \in (1,2)$ and $\beta \in (0,1)$. The complex roots $\gamma, \delta, \epsilon, \zeta$ have modules equal 1.

The explicit formula for $a_{3,n}$ is

$$a_{3,n} = C_1\alpha^n + C_2\beta^n + C_3\gamma^n + C_4\delta^n + C_5\epsilon^n + C_6\zeta^n$$

where

$$C_1 = \frac{\alpha^6+\alpha^4+\alpha^2}{-\alpha^4+4\alpha^3-3\alpha^2+8\alpha-5},\quad C_2 = \frac{\beta^6+\beta^4+\beta^2}{-\beta^4+4\beta^3-3\beta^2+8\beta-5},$$

$$C_3 = \frac{\gamma^6+\gamma^4+\gamma^2}{-\gamma^4+4\gamma^3-3\gamma^2+8\gamma-5},\quad C_4 = \frac{\delta^6+\delta^4+\delta^2}{-\delta^4+4\delta^3-3\delta^2+8\delta-5},$$

$$C_5 = \frac{\epsilon^6+\epsilon^4+\epsilon^2}{-\epsilon^4+4\epsilon^3-3\epsilon^2+8\epsilon-5},\quad C_6 = \frac{\zeta^6+\zeta^4+\zeta^2}{-\zeta^4+4\zeta^3-3\zeta^2+8\zeta-5},$$

Since $|\beta| < 1$ and $|\gamma| = |\delta| = |\epsilon| = |\zeta| = 1$ the theorem is proved. □

By using Theorem 4 and Theorem 5 we have

Theorem 6.

$$a_{3,n} = a_{2,n} + \sum_{k=1}^{\lfloor\frac{n+1}{5}\rfloor} \sum_{i_1+i_2+\ldots+i_{k+1}=n-4k} a''_{2,i_1}a'_{2,i_2}\ldots a'_{2,i_k}a''_{2,i_{k+1}} =$$

$$= \left[\frac{\alpha^6+\alpha^4+\alpha^2}{-\alpha^4+4\alpha^3-3\alpha^2+8\alpha-5}\alpha^n\right]$$

where i_1 *and* $i_{k+1} \in N\cup\{0\}$, $i_2, i_3, \ldots, i_k \in N$ *and* $\alpha = 1,974818708297706\ldots$

Corollary 2.

$$\lim_{n\to\infty}\frac{1}{\alpha^n}\left(a_{2,n} + \sum_{k=1}^{\lfloor\frac{n+1}{5}\rfloor} \sum_{i_1+i_2+\ldots+i_{k+1}=n-4k} a''_{2,i_1}a'_{2,i_2}\ldots a'_{2,i_k}a''_{2,i_{k+1}}\right) =$$

$$= \frac{\alpha^6+\alpha^4+\alpha^2}{-\alpha^4+4\alpha^3-3\alpha^2+8\alpha-5}$$

where i_1 *and* $i_{k+1} \in N\cup\{0\}$, $i_2, i_3, \ldots, i_k \in N$ *and* $\alpha = 1,974818708297706\ldots$

Lemma 3.

$$a'_{m,n} = a'_{m-1,n} + \sum_{k=1}^{\lfloor \frac{n+1}{2m-1} \rfloor} \sum_{i_1+i_2+\ldots+i_{k+1}=n-2(m-1)k} a'_{m-1,i_1} a'_{m-1,i_2} \cdots a'_{m-1,i_k} a'_{m-1,i_{k+1}}$$

where $i_1, i_2, \ldots, i_{k+1} \in N$ *and* $a'_{m,n}$ *is the number of all words of the length* n *over the alphabet* $\{0,1\}$ *with the forbidden subword* $\underbrace{1010\ldots10}_{m\ times\ 10}$ *and which ineither begin nor end with 10.*

Lemma 4.

$$a''_{m,n} = a''_{m-1,n} + \sum_{k=1}^{\lfloor \frac{n+1}{2m-1} \rfloor} \sum_{i_1+i_2+\ldots+i_{k+1}=n-2(m-1)k} a''_{m-1,i_1} a'_{m-1,i_2} \cdots a'_{m-1,i_k} a'_{m-1 i_{k+1}}$$

where $i, 1 \in N \cup \{0\}$, $i_2, \ldots, i_{k+1} \in N$ *and* $a''_{m,n}$ *is the number of all words of the length* n *over the alphabet* $\{0,1\}$ *with the forbidden subword* $\underbrace{1010\ldots10}_{m\ times\ 10}$ *and which only do not begin with 10.*

It is obvious that $a''_{m,n}$ is the number of all words of the length n over the alphabet $\{0,1\}$ with the forbidden subword $\underbrace{1010\ldots10}_{m\ times\ 10}$ and which only do not end with 10.

Theorem 7.

$$a_{m+1,n} = a_{m,n} + \sum_{k=1}^{\lfloor \frac{n+1}{2m+1} \rfloor} \sum_{i_1+i_2+\ldots+i_{k+1}=n-2mk} a''_{m-1,i_1} a'_{m-1,i_2} \cdots a'_{m-1,i_k} a''_{m-1,i_{k+1}}$$

where $i_1, i_{k+1} \in N \cup \{0\}$ *and* $i_2, i_3, \ldots, i_k \in N$

Proof. Let us count the number of all words of the length n over the alphabet $\{0,1\}$ with the forbidden subword $\underbrace{1010\ldots10}_{m\ times\ 10}$ i.e. the number of words in the set $A_m(n)$. We make a partition of the set $A_m(n)$ into subsets $A_m^k(n)$, where $A_m^k(n)$ is the set of all those words of the length n over the alphabet $\{0,1\}$ which contain exactly k subwords $\underbrace{1010...10}_{m-1 \text{ times } 10}$ $(\mathbf{x_n} \in A_m^k(n) \Rightarrow l_{\underbrace{1010...10}_{m-1}}(\mathbf{x_n}) = k)$ and wich do not contain the subword $\underbrace{1010...10}_{m}$. In the same way as in Theorem 2, by using Lemma 3 and Lemma 4, we complete the proof. □

References

[1] Austin Richard and Guy Richard, Binary sequences without isolated ones, The Fibonacci Quarterly, Volume 16, Number 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. No 320–No 328 1970, 27–34.

[3] Doroslovački, R., The set of all words over the alphabet $\{0,1\}$ of length n with the forbidden subword $11\ldots1$, Rev. of Res., Fac. of Sci. Math. Ser., Novi Sad, Vol. 14, Num. 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.23, 2 (1993), 239-244 Novi Sad.

[5] Doroslovački, R., Binary sequences without $0\overbrace{11...11}^{k-1}0$ for fixed k, Matematički Vesnik 46 (1994), 93-98, Beograd.

[6] Doroslovački, R., The binary n-words with forbidden 4-subwords, Novi Sad Journal of Mathematics (to appear).

[7] Einb, J. M., The enumeration of bit sequences that satisfy local criteria, Publications de l'Institut Mathématique Beograd, tome 27(41) (1980), 51–56.

Received by the editors December 10, 1998.