BINARY SEQUENCES WITHOUT $011\ldots110$

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 $011\ldots110$ 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.

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

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 $\lfloor x \rfloor$ and $\lceil x \rceil$ 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 $\lfloor x \rfloor$ is the nearest integer to $x$.

2. Results and discussion

Now we shall construct and enumerate the set of words

$$A_k(n) = \{ x_n \mid x_n = x_1x_2\ldots x_n \in X^n, \ (\forall i \in N_{n-k})(x_i x_{i+1} \ldots x_{i+k} \neq 01\ldots10) \}$$

AMS Mathematics Subject Classification (1991) 05A15

Key words and phrases: subword.
for each natural number $k$. It is known that

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

(Fibonacci numbers) where

$$A_1(n) = \{ x_n \mid x_n = x_1 x_2 \ldots x_n \in X^n, (\forall i \in N_{n-1}) (x_i x_{i+1} \neq 0) \}.$$

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

**Theorem 1.**

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

where

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

and

$$\alpha = \frac{1}{6}(4 + \sqrt{100 + 4 \sqrt{621} + \sqrt{100 - 4 \sqrt{621}}}) \approx 1,754877666247.$$
Binary sequences without $011\ldots110$ for fixed $k$

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

\[
\binom{i-1}{j} \binom{i-1-j}{k}
\]

(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

\[
\binom{n - i - j - 2k + 1}{k + 1}
\]

(3)
ways.

Thus from (2), (3) and

\[
|A_3(n)| = \sum_{i=0}^{n} |A^j_3(n)| = 1 + \sum_{i=1}^{n} |A^j_3(n)|
\]

Theorem 2 follows.

THEOREM 3.

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

where

\[
\alpha = \frac{1}{2} \left( 1 + \sqrt{3 + 2\sqrt{5}} \right) \approx 1.866760399.
\]

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

\[
|A_2(n)| = 1 + \sum_{i=1}^{n} \sum_{j=0}^{n-i} \sum_{k=0}^{\lfloor \frac{n-i-j}{2} \rfloor} \binom{i}{j} \binom{i-j}{k} \binom{n-i-j-2k+1}{k+1} \\
= \left[ \frac{2\alpha^2 + 1}{2\alpha^3 - 3\alpha + 4}\alpha^n \right], \\
\text{where } \alpha = \frac{1}{2} \left( 1 + \sqrt{3 + 2\sqrt{5}} \right).
\]

**Theorem 4.**

\[
|A_k(n)| = 1 + \sum_{i=1}^{n} \sum_{j_1=0}^{n-i} \sum_{j_2=0}^{n-i-j_1} \cdots \sum_{j_{k-1}=0}^{n-i-j_1-j_2-\ldots-j_{k-2}} \sum_{\ell=0}^{\lfloor \frac{n-i-j_1-j_2-\ldots-j_{k-1}}{\ell} \rfloor} \prod_{m=0}^{m-k-3} \binom{i-1-s_m}{j_{m+1}} \binom{i-1-s_{k-2}}{\ell} \binom{n-i-S_{k-2}-(k-1)\ell+1}{\ell+1} \\
\text{where } s_k = j_1 + j_2 + \cdots + j_k, \ s_0 = 0, \ S_k = j_1 + 2j_2 + \cdots + kj_k \text{ and } \\
A_k(n) = \{ x_n | x_n = x_1x_2 \ldots x_n \in X^n, (\forall s \in N_{n-k})(x_{s]\cdots x_{s+k} \neq 01\cdots 10) \}.
\]

**Proof.** We partition the set \(A_k(n)\) into subsets \(A^n_k(n)\) which contain exactly \(i\) zeros i.e.

\[
A^n_k(n) = \{ x_n | x_n = x_1x_2 \ldots x_n \in X^n, \\
(\forall s \in N_{n-k})(x_{s}\cdots x_{s+k} \neq 01\cdots 10), \ell_0(x_n) = i \}.
\]

Now we shall construct words from \(A^n_k(n)\) 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 \(q_m = \underbrace{11\ldots1}_{m\text{ copies}}, \text{ for } m \in \{1, 2, \ldots, k-2\}\), \(r = 11\ldots1\) and \(\lambda\) is the empty letter. Let \(j_m\) be the number of letters \(q_m\), and \(\ell\) 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 \(q_2, \ldots, j_{k-2}\) places from \(i-1-s_{k-3}\) places for letters \(q_{k-2}\) and \(\ell\) places from \(i-1-s_{k-2}\) places for letters \(r\). It can be done in

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

(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 \(\ell\) 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
\[
\binom{n - i - S_{k-2} - (k - 1)\ell + 1}{\ell + 1}
\]
different ways, where \( S_k = j_1 + 2j_2 + \cdots + kj_k \). Thus from (4), (5) and \( |A_k(n)| = \sum_{i=0}^{n} |A_i^k(n)| \) Theorem 4 follows. ■

**Theorem 5.**
\[
|A_k(n)| = \left[ C(k, \alpha)\alpha^n \right]
\]
for large enough values of \( n \), where \( \alpha \) is the unique real root of equation
\[
x^{k+1} - 2x^k + x - 1 = 0
\]
which lies between 1 and 2 and \( C(k, \alpha) \) is the rational function of \( \alpha \) and \( k \).

**Proof.** Words \( x_n \in A_k(n) \) are obtained from other words \( x_{n-1} \in A_k(n-1) \) by appending 0 or 1 in front of them. Let
\[
x_{n-1} \in A_k(n-1), \quad x_{n-k} \in A_k(n-k) \quad \text{and} \quad x_{n-k-1} \in A_k(n-k-1).
\]
Then
\[
1x_{n-1} \in A_k(n), \quad 011\ldots1x_{n-k-1} \in A_k(n), \quad 011\ldots1x_{n-k-1} \notin A_k(n)
\]
which means that \( 011\ldots1x_{n-k} \in A_k(n) \) if and only if \( x_{n-k} \) begins with the letter 1. This implies the recurrence 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 \( \alpha \) 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 \( \alpha \) and \( \beta \) 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 \( \alpha \) and \( k \). ■

**Corollary 2.**
\[
|A_k(n)| = 1 + \sum_{i=1}^{n} \sum_{j=0}^{n-i} \sum_{k=0}^{n-i-j} \sum_{\ell=0}^{n-i-j-k} \frac{(i - 1)!}{j!} \frac{(i - 1 - j)!}{k!} \frac{(n - i - j - 2k)!}{\ell!} \frac{1}{\ell + 1}
\]
\[
= \left[ \frac{4\alpha^4 - \alpha + 2}{4\alpha^4 - 4\alpha^2 + 3\alpha + 2\alpha^n} \right],
\]
i.e. \( C(4, \alpha) = \frac{4n^5 - 2n^3 + 2}{4n^5 - 4n^3 + 8n^2 + 2} \) and \( \alpha \) is unique real root of equation \( x^5 - 2x^4 + x - 1 = 0 \). whose complex roots are with modules less than 1.

REFERENCES


[3] Doroslovački, R., The set of all words over alphabet \( \{0, 1\} \) of length \( n \) with the forbidden subword \( 1\ldots1 \), Rev. of Res., Fac. of Sci. math. ser., Novi Sad, 14, 2 (1984) 167–173


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


{received 14.09.1994.}

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