Engineering Sciences International Research Journal
ISSN (Print) : 2330 – 4338 ; Volume 1 Issue 1 ǁ February. 2013 ǁ

Combinatorial Properties of Sturmian Arrays

N. Jansirani[1], V. Rajkumar Dare[2]

Abstract: In this paper we consider the concept of a Sturmian array that extends the well-known notion of a Sturmian word. A general construction to obtain a Sturmian array from a Sturmian word is given. In this paper we introduce balanced arrays, bordered arrays and study some of their combinatorial properties.

Keywords: Balanced array, Bordered Array, Fibonacci array, Fibonacci word, Sturmian array.

1. INTRODUCTION

The field of "Combinatorics on words" has grown rapidly after a unified treatment of this theory appeared in the book by Lothaire [6]. In recent years, the combinatorial properties of finite and infinite words have become significantly important in the fields of physics, biology, mathematics and computer science. One of the first impulse for extensive research in this field was the discovery of quasi-crystals. Aldo de Luca introduced the combinatorial method for the analysis of finite words, for the study of biological molecules [2]. Combinatorial properties of finite and infinite words such as the Thue-morse word, Fibonacci word, Lyndon words, Sturmian words and several others have been explored in the literature. Sturmian words appear in many areas of study like combinatorics on words, algorithms on words, symbolic dynamics, crystallography, pattern recognition and so on. Sturmian words are used in computer graphics as digital approximation of straight lines. A Sturmian word over a binary alphabet is an infinite non-ultimately periodic word which has exactly n+1 factors of length n for each natural number n. They represent the simplest family of quasi-crystals[1], [2], [3].

Infinite arrays have been considered in as extensions of infinite words in the context of the study of infinite computations [4]. The extension of a word to two dimensions is an array which gives a set of pictures, where a picture is a 2-dimensional array of characters over an alphabet. The study of picture language was initially motivated by the problems of patterns recognition and image processing. As a consequence in this paper we extend one-dimensional Sturmian word into two-dimensional Sturmian array. Here we consider the notion of a Sturmian array by extending the Sturmian feature of infinite words to infinite arrays. A general method of constructing a Sturmian array from a Sturmian word is given based on a known Sturmian word. In particular a specific Sturmian array called Fibonacci array is obtained from the famous Fibonacci word which is also a Sturmian word [7]. Certain properties of such Sturmian arrays are obtained.The well-known Sturmian words are characterized by balance property. The balance property is a fine tool on words appearing in combinatorics on words [3], in billiard theory and dynamical systems. Based on the notion of a balanced word a balanced array is defined and is related to Sturmian arrays [1]. In this paper we introduce balanced arrays, bordered arrays study some of their combinatorial properties.

2. BASIC DEFINITIONS AND NOTATIONS

Let Σ be a finite alphabet. The set of all words over Σ is denoted by Σ*. The empty word is denoted by λ. We write $\Sigma^+ = \Sigma^* - \{\lambda\}$.

An infinite word w over a finite alphabet Σ is a mapping from positive integers into Σ. We write $w = a_1\, a_2 \ldots a_i \ldots$ where a_i ε Σ. The set of all infinite words over Σ is denoted by Σ^ω. An infinite word w is ultimately periodic if $w = uv^\omega$, where $u \in \Sigma^*$ and $v \in \Sigma^+$.

An infinite word w over a binary alphabet $\{a, b\}$ which is not ultimately periodic such that for any positive integer n, the number $g_x(n)$ of its factors of length n is minimal *i.e* $g_x(n) = n + 1$, is called a Sturmian word [1, 2, 5].The Fibonacci word f which is an important example of a Sturmian word is the fixed point of a morphism

$\varphi : B^* \rightarrow B^*$ where $B = \{a,b\}$ and $\varphi(a) = ab$, $\varphi(b) = a$ *i.e.* $f = \varphi^\omega(a)$. In fact the first few symbols of the Fibonacci word is *abaababaabaababaababa* …[3],[4],[8].

An (m, n) array $A = (a_{ij})_{mxn}$ over an alphabet Σ is a rectangular arrangement of symbols of Σ in m rows and in n columns. The size of the array A is the ordered pair (m, n). The set of all arrays over Σ is denoted by Σ** . The empty array is also denoted by λ. We adopt the convention that for an (m, n) array $A = (a_{ij})_{mxn}$ the bottom most row is the first row and the left most column is the first column. Also we write $\Sigma^{++} = \Sigma^{**} - \{\lambda\}$. A factor or sub array of an array A is also an array which is a part of A (formed by the intersection of certain consecutive rows and certain consecutive columns of A).

An infinite array u has an infinite number of rows and infinite number of columns. The collection of all infinite arrays over Σ is denoted by $\Sigma^{\omega\omega}$. Row and column concatenation of arrays in Σ^{**} are partial operations. For row catenation of two arrays A and B , denoted by A® B, the number of columns in A and B should be equal and for column concatenation A© B of A and B, the number of rows in A and B should be equal. The collection of all arrays with finite number of rows and an infinite number of columns is denoted by $\Sigma^{*\omega}$ and the collection of all arrays with an infinite number of rows and a finite number of columns is denoted by $\Sigma^{\omega*}$.

An array $w \varepsilon \Sigma^{\omega\omega}$ is said to be row periodic if there exist two arrays u and v in $\Sigma^{*\omega}$ such that $w = u$® $v^{®\omega}$, $v^{®\omega}$ denotes an infinite number of row catenation of v with v. An array $w \in \Sigma^{\omega\omega}$ is said to be column ultimately periodic if there exist two arrays u and v of $\Sigma^{\omega*}$ such that $w = u$ © $v^{©\omega}$, $v^{©\omega}$ denotes an infinite number of column catenation of v with v **.** An array $w \in \Sigma^{\omega\omega}$ is said to be ultimately periodic if it is both row periodic and column periodic.

3. STURMIAN ARRAY

In this section we first define a Sturmian array and give a general method of constructing a Sturmian array from a given Sturmian word.

Definition 1

Let $u = (u_{ij}) \epsilon \Sigma^{**}$, $1 \leq i \leq m$, $1 \leq j \leq n$. The string or the word $u_{11}u_{12}u_{13} \ldots u_{1n}u_{2n}u_{3n} \ldots u_{mn}$ is called the right boundary of the array u and the string or the word $u_{11}u_{21}u_{31}u_{41} \ldots u_{m1}u_{m2}u_{m3} \ldots u_{mn}$ is called the left boundary of u.

Note that the boundary of an array is formed by taking the right boundary and "joining" to it the reverse of the left boundary. By "joining" we mean that a word aub joined to bva gives $aubv$. For an array $u \varepsilon \Sigma^{**}$ of size (m, n) the length of the right boundary or the left boundary is $m+n-1$.

Definition 2

An infinite array u which is not ultimately periodic is said to be a Sturmian array if $g_u(m, n) = m + n$ where $g_u(m, n)$ denotes the number of distinct sub arrays of size (m, n).

4. CONSTRUCTION SA

Let $x = a_1a_2a_3\ldots$ be a Sturmian word over a binary alphabet. We construct an array U from x as follows:
$U = (u_{jk})$, where for , $u_{jk} = a_{j+k-1}$.

Theorem 1

The infinite array U formed as in the construction SA is a Sturmian array.

Proof : In the array U constructed by SA we notice that any subarray of size (m, n) is determined by the right boundary which is a word of length $m+n-1$. In fact it is a factor of the Sturmian word x and is of the form $a_{i+1}a_{i+2} \ldots a_{i+m}a_{i+m+1} \ldots a_{i+m+n-1}$ for some $i \geq 0$. Such factors of length $m+ n -1$ in the Sturmian word x can only be $m+ n$ in number. This means that there are only $m + n$ number of distinct sub arrays of size (m, n) in U and hence $g_U(m, n) = m + n$. Note that each row (or column) being a suffix of the Sturmian infinite word x given in the construction, is not ultimately row (column) periodic.

Remark 1

If U is a Sturmian array as in the construction SA, then every row (column) of U is a Sturmian word.

Definition 3

An array W is a finite Sturmian array if there exists an infinite Sturmian array U such that W is a sub array of U. The set of all finite sub arrays of U is denoted by SFSA(U)

Balance property characterization of Sturmian words is important because it leads to many generalizations used in computer science. Based on this we studied Balance property on arrays. Two finite arrays of the same size on the alphabet $\{a, b\}$ have the balance property if the number of a in the two arrays are almost same. More precisely the difference between the number of a in the two arrays are bounded by 1.

Definition 4

If $U \in \Sigma^{**}$ is an array of size (m ,n) then the size of U is denoted by $|U|$ and size value of U is mn and it is denoted by $||U||$.

Remark 2

We note that two arrays with equal size value are need not be of same size.

Definition 5

An infinite array W on $\Sigma = \{a ,b\}$ is balanced if the following condition is satisfied $|U| = |V|$ implies $|\,|U|_a - |V|_a| \leq 1$ where U and V are sub arrays of W and $|U|_a$ denotes number os a's in U.

Theorem 2

If $W \in \Sigma^{\omega\omega}$ is a Sturmian array over $\Sigma = \{a ,b\}$ then the language set SFSA(W) satisfies the property $||U|_a - |V|_a| \leq 1$ for any pair (U,V) of sub arrays of W having the same size.

Theorem 3

An infinite array W on $\Sigma = \{a, b\}$ is balanced if and only if for any two nonempty sub arrays U and V of W, we have $\big|\, |U|_a \, \|V\| - |V|_a \, \|U\| \,\big| \leq \|U\| + \|V\| - \gcd(\|U\|, \|V\|)$

Proof:

The proof is given in [5].

Theorem 4

Let $U \in \Sigma^{\omega\omega}$ over $\Sigma = \{a, b\}$.Then the following conditions are equivalent.

1) U is Sturmian.
2) U is balanced and non-ultimately periodic.

Proof:

The proof follows from the construction SA.

5. BALANCE PROPERTY WITH A WORD

We study the balance property of Sturmian array not only with a single letter but with sub arrays of W. We show that for a Sturmian array W the difference between the number of occurrences of a array U sub array of Win two factors of same size of W is balanced by the size of U.

Definition 6

A cell is an array of size (1,1) containing a symbol from an alphabet. The cell at the i^{th} row and j^{th} column of an array C is denoted by $C(i, j)$. For $i \geq 2$, $j \geq 2$, the left, right, down, up, left down, left up, right down and right up neighboring cells of a cell $C(i, j)$ are respectively $C(i, j\text{-}1)$, $C(i, j+1)$, $C(i\text{-}1, j)$, $C(i+1, j)$, $C(i\text{-}1, j\text{-}1)$, $C(i+1, j\text{-}1)$, $C(i\text{-}1, j+1)$ and $C(i+1, j+1)$.

Theorem 5

Let $F \in \Sigma^{\omega\omega}$ be the Fibonacci array. Then every $F(i, j) \in F$ satisfies one of the first three conditions:

(1) $F(i\text{-}1, j) = F(i, j) \neq F(i+1, j)$
(2) $F(i\text{-}1, j) \neq F(i, j) = F(i+1, j)$
(3) $F(i\text{-}1, j) = F(i+1, j) \neq F(i, j)$
(4) There exists no cell $F(i, j)\ \varepsilon\ F$ such that
$F(i\text{-}1, j) = F(i, j) = F(i+1, j)$

Proof:

The result is a consequence of the definition of *F*. Reader can refer [5] for the Fibonacci array.

Proposition 1

The sequence of minimal and maximal values of $g_u(m,n)$ forms an Arithmetic Progression and Geometric Progression respectively, for a fixed row and for a for fixed column where the maximal number for $g_x(m,n)$ is 2^{m^n}

6. BORDERED ARRAYS

In this section we study border property. It is interesting to note that in a Sturmian array the border property can be easily checked with the equality of the right boundary of the prefix and suffix of that array.

Definition 7

Let $u \in \Sigma^{**}$. A prefix of u is a rectangular sub array that contains one corner element, of u. A suffix is a rectangular sub array that contains the diagonally opposite corner. Notice that the choice of the corner in which to put the prefix is arbitrary. Because of the symmetry the prefix may be assigned to either the upper left or lower left corner of u.

Definition 8

If a prefix and a suffix of an array as equal then the array is said to be bordered.

Example 1

The following example shows an array with border of size (3x2)

a	b	a	a	**b**	**a**
b	b	a	a	**b**	**a**
b	**a**	b	b	**b**	**a**
b	**a**	b	b	a	b
b	**a**	b	b	a	b

Definition 9

An array is bordered by a prefix of size 2x2 then it is said to be tile bordered

Example 2

The following example shows a tile border array of size (5x5)

Theorem 5

Let W be a Sturmian array .Then for every X ε SFSA W contains infinite number of sub arrays such that for each array with X as border.

Proof:

The proof form a very remarkable property related to Conway's theorem stating that given an arbitrary portion of a

Fibonacci word we will find a replica of it at a distance smaller than twice its length. Since the right boundary of any sub array is a Fibonacci word the proof is immediate.

7. CONCLUSION

In this paper we have considered the notion of a Sturmian array and obtained bordered arrays properties of a Sturmian array constructed from a Sturmian word. It remains to explore further other properties of Sturmian arrays.

8. REFERENCES

[1] Allouche and J. Shallit, Automatic Sequences**",** Theory, Applications, Generalizations, Cambridge University Press, ISBN 978052183326, (2003).

[2] A. De Luca and F. Mignosi, "Some Combinatorial Properties of sturmian words", Theoretical Computer Science, 136 (1994) 361 – 385.

[3] J.Berstel and P.Séébold, "Sturmian words", In M. Lothaire (Eds), Algebraic combinatorics on Words. Cambridge University Press, Cambridge U.K., 2002.

[4] V.R. Dare, K.G. Subramanian, D.G. Thomas and R. Siromoney "Infinite arrays attern Recognition and Artificial Intelligence, Vol. 14, 4(2000) 525 – 536.

[5] N. Jansirani, V. Rajkumar Dare and K.G. Subramanian ,**"** Sturmian Arrays" ", Advances in Image Analysis and Applications, Spain (2011) ISBN-13:978-08-7923-5, ISBN-10:381-08-7923-7 (2011)

[6] M. Lothiare, "Combinatorics on words", Addison-Wesley, 1983.

[7] F. Mignosi, "On the number of factors of Sturmian words", Theoretical Computer Science, 82 (1991) 71 – 84.

* * *

[1]*N. Jansirani, Department of Mathematics, Queen Mary's College, Chennai 600 004, India. njansirani@gmail.com*

[2]*V. Rajkumar Dare, Department of Mathematics, Madras Christian College, Chennai, 600 059, India*