American Review of Mathematics and Statistics
June 2015, Vol. 3, No. 1, pp. 9-21
ISSN: 2374-2348 (Print), 2374-2356 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/arms.v3n1a2
URL: http://dx.doi.org/10.15640/arms.v3n1a2

An Algorithm for the Exhaustive Generation of Binary Words avoiding up-down Sequences

Stefano Bilotta[1], Elisabetta Grazzini[1,2] & Elisa Pergola[1]

Abstract

In this paper we propose an algorithm to generate binary words with no more 0's than 1's having a fixed number of 1's and avoiding the sequence $(10)^j1$ for any fixed $j \geq 1$. Any binary word w can be represented as a path on the Cartesian plane by associating a rise step, defined by (1,1), with each bit 1 in w and a fall step, defined by (1,-1), with each bit 0 in w. Given a path ω with *n* rise steps associated with the binary word w avoiding the sequence $(10)^j1$, we generate a given number of paths with *n*+*h* rise steps, $1 \leq h \leq j$, avoiding the same sequence, by means of constructive rules. The number and the shape of the generated paths depend on the ordinate k of the endpoint of ω and on its suffix. We will prove that this generation is exhaustive, that is, all binary words with *n* bit 1 and avoiding the sequence $(10)^j1$ are generated.

Keywords: Binary words, Exhaustive generation, Paths, Sequence avoiding

1. Introduction

Let $F \subset \{0,1\}^*$ be the set of binary words w such that $|w|_0 \leq |w|_1$, for any $w \in F$, $|w|_0$ and $|w|_1$ corresponding to the number of 0's and 1's in the word w, respectively. In this paper we study the construction of the subset F[p]$\subset F$ of binary words excluding a given sequence p= $p_0 \cdots p_{\ell-1} \in \{0,1\}^{\ell}$, that is a word $w \in F$[p] if and only if there are no consecutive indices $i, i+1, \ldots, i+\ell-1$ such that $w_i w_{i+1} \cdots w_{i+\ell-1} = p_0 p_1 \ldots p_{\ell-1}$.

If we consider the set of binary words without any restriction, the defined language is regular and we can refer to using classical results (see Guibas & Odlyzko and Sedgewick & Flajolet, [3,4,5]). When the restriction to words with no more 0's than 1's is valid, the language F[p] is not a regular one and it becomes more difficult to deal with. For example, in order to generate the language F[p] for each forbidden sequence p an "ad hoc" grammar should be defined. Our aim is to determine a constructive algorithm suggesting a more unified approach which makes it possible to generate all binary words in the set F[p].

Given $|w| = |w|_0 + |w|_1$ the *length* of $w \in F$, we denote by w^h, (*h* > 0), the word with length $h * |w|$ obtained by linking w to itself *h* times, that is $w^h = \underbrace{ww \cdots w}_{h}$ and $w^0 = \varepsilon$, ε being the empty word.

In this paper we show how to obtain all binary words belonging to F and avoiding the sequence p= $(10)^j1$, for any fixed $j \geq 1$.

---

1 Dipartimento di Matematica e Informatica "Ulisse Dini", Università di Firenze, Viale G. B. Morgagni 65, 50134 Firenze, Italy. Email:{stefano.bilotta,elisabetta.grazzini,elisa.pergola}@unifi.it
2 Corresponding author

In Bilotta, Merlini, Pergola & Pinzani [2] an algorithm for the construction of all binary words in *F* having a fixed number of 1's and excluding those containing the forbidden sequence $1^{j+1}0^{j}$, for any fixed $j \geq 1$ is introduced. That algorithm generates all the words in *F* then eliminates those containing the forbidden sequence. Basically, the construction marks in an appropriate way the forbidden sequences in the words and generates 2^{C} copies of each word having *C* forbidden sequences such that the 2^{C-1} instances containing an odd number of marked forbidden sequences are annihilated by the other 2^{C-1} instances containing an even number of marked forbidden sequences. For example, the words $00110\overline{110}$ and $00\overline{110}110$, containing two copies of the forbidden sequence p = 110, (the marked forbidden sequences are overlined) are eliminated by the words 00110110 and $00\overline{110110}$, respectively.

This is possible since no prefix of p = $1^{j+1}0^{j}$ is also a suffix of p, that is the forbidden sequences do not overlap and so they are univocally identified inside the words. Then, the algorithm in Bilotta, Merlini, Pergola & Pinzani [2] cannot be used to generate the words in *F*[p] when p = $(10)^{j}1$ since the forbidden sequences may overlap inside the words. For example, in w =110101010 there are two overlapping copies of the forbidden sequence p = $(10)^{2}1$. So, we propose a new algorithm that generates right the words in *F* avoiding the forbidden sequence p = $(10)^{j}1$, for any fixed $j \geq 1$. Recall that any word $w \in F$ can be naturally represented as a path on the Cartesian plane by associating a *rise* (or *up*) *step*, defined by (1,1) and indicated by x with each bit 1 in w and a *fall* (or *down*) *step* defined by (1,-1) and indicated by $\bar{x}$, with each bit 0 in w. For example, the word w =11011010010000101111 is represented by the path $\omega = xx\bar{x}xx\bar{x}x\bar{x}\bar{x}x\bar{x}\bar{x}\bar{x}\bar{x}x\bar{x}xxxx$ (see Figure 1). An *up-down* step is the sequence $x\bar{x}$.

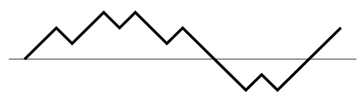

Figure 1: The Path Representing w =11011010010000101111

From now on, we refer interchangeably to words or their graphical representation on the Cartesian plane, that is paths. So by *F*[p] we denote both the set of p avoiding binary words and the set of corresponding paths.

In the rest of this paper, a path is defined as:

– *primitive* if it begins and ends at ordinate 0 and remains strictly above the *x*-axis,
– *positive* if it begins at ordinate 0 and remains above or on the *x*-axis,
– *negative* if it begins and ends at ordinate 0 and remains below or on the *x*-axis (remark that a negative path in *F* necessarily ends at ordinate 0),
– *strongly negative* if it begins and ends at ordinate –1 and remains below or on the line *y*=–1,
– *underground* if it ends with a negative suffix.

The *complement* of a path ω is the path ω^{c} obtained from ω by switching rise and fall steps.

In Section 2 we give an algorithm for the construction of the set *F*[p], i.e. the set of binary words excluding the sequence p = $(10)^{j}1$, for any fixed $j \geq 1$, and such that the number of 0's in each word is at most equal to the number of 1's. In Section 3 we prove that the construction given in Section 2 allows us to obtain an exhaustive and univocal generation of binary words in *F*[p] containing *n* 1's.

2. A construction for the set *F*[p]

In this section we show the constructive algorithm to generate the set *F*[p], p= $(x\bar{x})^{j}x = (10)^{j}1$ for any fixed $j \geq 1$, according to the number of rise steps, or equivalently to the number of 1's. Given a path $\omega \in$ *F*[p] with *n* rise steps, we generate a given number of paths in *F*[p] with *n*+*h* rise steps, $1 \leq h \leq j$, by means of constructive rules. The number and the shape of the generated paths depend on the ordinate *k* of the endpoint of ω and on its suffix. With regard to *k*, we can point out three cases: *k* =0, *k* =1, and $k \geq 2$ whereas for the suffix we consider whether it is equal to $(x\bar{x})^{j}$ or not. When *k* =0, we must pay attention also to the case in which ω is an underground path ending with $(x\bar{x})^{j-1}x$.

As we will show further on, for each $\omega \in F[p]$ such that $k=0$ or $k \geq 2$, the generating algorithm produces for each h, $1 \leq h \leq j$, two or more positive paths and one underground path with $n+h$ rise steps, $1 \leq h \leq j$, whereas, when $k=1$, it produces only one positive path with $n+h$ rise steps. The generating algorithm of the set $F[p]$ with p= $(x\bar{x})^j x = (10)^j 1$, for any fixed $j \geq 1$, is described in the following sections. The constructive rules related to the special cases in which the suffix of ω is $(x\bar{x})^j$ or $(x\bar{x})^{j-1}x$ are described in Sections 2.2 and 2.3, respectively, whereas in Section 2.1 we examine all the other simple cases. The starting point of the algorithm is the empty word ε

2.1 Simple Cases

In this section we describe the constructive rules to be applied when the suffix of ω is neither $(x\bar{x})^j$ nor $(x\bar{x})^{j-1}x$. We point out three cases for the ordinate k of the endpoint of ω: $k=0$, $k=1$, and $k \geq 2$. Let us denote by $\omega_{|k}$ a path with endpoint at ordinate k.

$k=0$. A path $\omega \in F[p]$, with n rise steps and such that its endpoint has ordinate 0, generates, for each h, $1 \leq h \leq j$, three paths with $n+h$ rise steps: (i) a path ending at ordinate 1 by adding to ω a rise step and a sequence of h–1 up-down steps; (ii) a path ending at ordinate 0 by adding to ω a rise step, a sequence of h–1 up-down steps, and a fall step; (iii) an underground path obtained by the one generated at (ii) mirroring on x-axis its rightmost primitive suffix.

Figure 2 shows the above described operations; the number above the right arrow corresponds to the value of h. Both in this figure and in the following ones we consider j=4, that is p = $(x\bar{x})^4 x = (10)^4 1$.

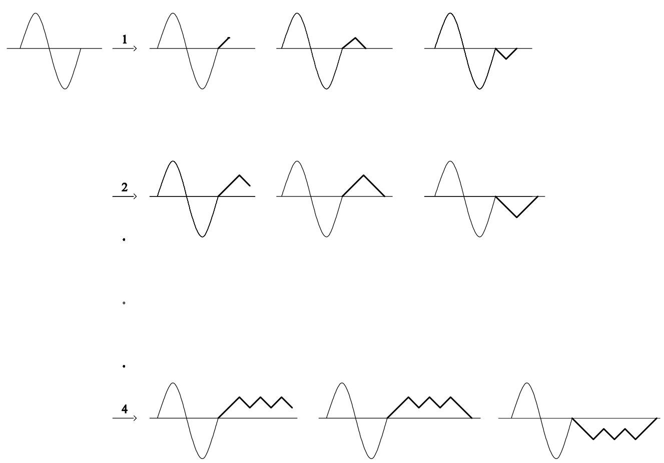

Figure 2: The Paths Generated by $\omega_{|0}$

Therefore

$$\omega_{|0} \rightarrow \begin{cases} \omega_{|0} x (x\bar{x})^{h-1} \\ \omega_{|0} x (x\bar{x})^{h-1} \bar{x} \\ \omega_{|0} \bar{x} (\bar{x}x)^{h-1} x \end{cases} \qquad (1)$$

$k=1$. A path $\omega \in F[p]$, with n rise steps and such that its endpoint has ordinate 1, generates, for each h, a path with $n+h$ rise steps with endpoint at ordinate 2 obtained by adding to ω a rise step and a sequence of h–1 up-down steps (see Figure 3).

Therefore

$$\omega_{|1} \to \omega_{|1} x (x\overline{x})^{h-1} \qquad (2)$$

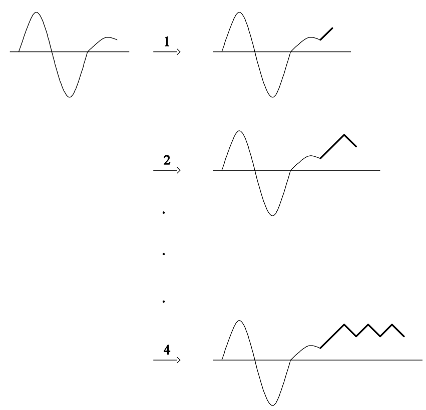

Figure 3: The Paths Generated by $\omega_{|1}$

$k \geq 2$. A path $\omega \in F$[p], with n rise steps and such that its endpoint has ordinate k, $k \geq 2$, generates, for each h, k+2 paths with n+h rise steps: (i) a path ending at ordinate (k+1) by adding to ω a rise step and a sequence of h–1 up-down steps; (ii) k–1 paths ending at ordinate (k–1), (k–2), …, 1, respectively, by adding to ω a rise step, a sequence of m, $2 \leq m \leq k$, fall steps and a sequence of h–1 up-down steps; (iii) a path ending at ordinate 0 by adding to ω a rise step, a sequence of k fall steps, a sequence of h–1 up-down steps, and a fall step; (iv) an underground path which will be described in Section 2.4. Figure 4 shows the above described operations.

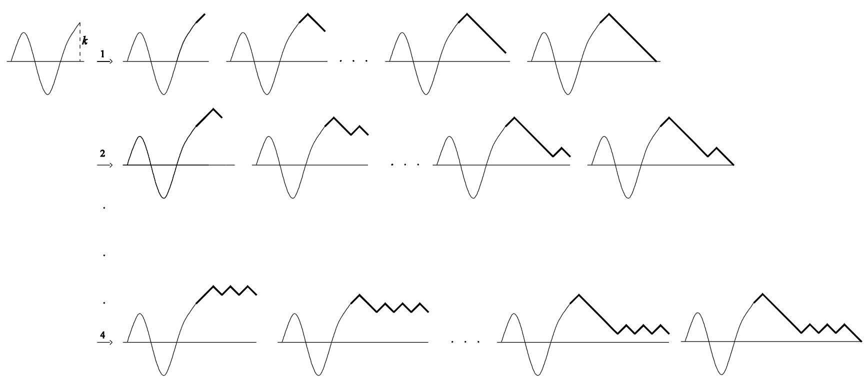

Figure 4: The Paths Generated by $\omega_{|k}$, $k \geq 2$

Therefore

$$\omega_{|k} \to \begin{cases} \omega_{|k} x (x\overline{x})^{h-1} & \\ \omega_{|k} x (\overline{x})^{m} (x\overline{x})^{h-1} & 2 \leq m \leq k \\ \omega_{|k} x (\overline{x})^{k} (x\overline{x})^{h-1} \overline{x} & \end{cases} \qquad (3)$$

At this point it is clear that:

1. when the path ω ends with the suffix $(x\overline{x})^{j}$ the paths obtained by means of the constructions (1), (2), and (3) contain the forbidden sequence p= $(x\overline{x})^{j} x$. So, we will act as described in Section 2.2;
2. when ω is an underground path ending with $(x\overline{x})^{j-1} x$, some paths generated by means of the above constructions might contain the forbidden sequence p $= (x\overline{x})^{j} x$. So, we will follow a different procedure described in Section 2.3.

2.2. Paths Ending with $(x\bar{x})^j$

Even when the path ω ends with the suffix $(x\bar{x})^j$, the number and the shape of the generated paths depend on the ordinate k of the endpoint of ω. Let $\rho = (x\bar{x})^j$ be the suffix of ω. k =0. A path $\omega \in F[\mathrm{p}]$, with n rise steps and such that its endpoint has ordinate 0, generates, for each h, $1 \le h \le j$, three paths with $n+h$ rise steps (see Figure 5): (i) a path ending at ordinate 1, by inserting a sequence of h–1 up-down steps and a rise step on the left of ρ; (ii) a path ending at ordinate 0, by inserting a sequence of h–1 up-down steps and a rise step on the left of ρ, and adding a fall step at the end of ω; (iii) an underground path, obtained by mirroring on x-axis the rightmost primitive suffix of the path generated at (ii). Therefore

$$\omega_{|0}\rho \rightarrow \begin{cases} \omega_{|0}(x\bar{x})^{h-1}x\rho \\ \omega_{|0}(x\bar{x})^{h-1}x\rho\bar{x} \\ \omega_{|0}(\bar{x}x)^{h-1}\overline{xx}(x\bar{x})^{j-1}xx \end{cases} \tag{4}$$

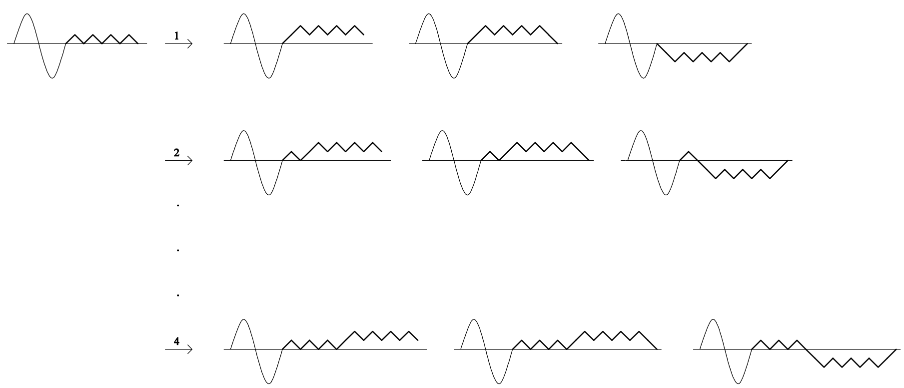

Figure 5: The Paths Generated by $\omega_{|0}(x\bar{x})^j$

k =1. A path $\omega \in F[\mathrm{p}]$, with n rise steps and such that its endpoint has ordinate 1, generates, for each h, a path with $n+h$ rise steps with endpoint at ordinate 2 obtained by inserting a sequence of h–1 up-down steps and a rise step on the left of the suffix ρ (see Figure 6). Therefore

$$\omega_{|1}\rho \rightarrow \omega_{|1}(x\bar{x})^{h-1}x\rho \tag{5}$$

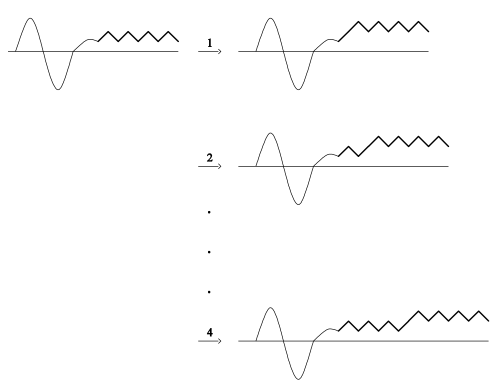

Figure 6: The Paths Generated by $\omega_{|1}(x\bar{x})^j$

$k \geq 2$. A path $\omega \in F[\mathrm{p}]$, with n rise steps and such that its endpoint has ordinate k, $k \geq 2$, generates, for each h, k+2 paths with n+h rise steps (see Figure 7): (i) a path ending at ordinate (k+1) by inserting a sequence of h–1 up-down steps and a rise step on the left of the suffix ρ; (ii) k–1 paths ending at ordinate (k–1), (k–2),…, 1, respectively, by inserting a sequence of h–1 up-down steps, a rise step, and a sequence of m, $2 \leq m \leq k$, fall steps on the left of ρ; (iii) a path ending at ordinate 0, by inserting a sequence of h–1 up-down steps, a rise step, and a sequence of k fall steps on the left of ρ, and then adding a fall step at the end of ω; (iv) an underground path which will be described in Section 2.4. Therefore

$$\omega_{|k}\rho \to \begin{cases} \omega_{|k}(x\bar{x})^{h-1}x\rho & \\ \omega_{|k}(x\bar{x})^{h-1}x(\bar{x})^{m}\rho & 2 \leq m \leq k \\ \omega_{|k}(x\bar{x})^{h-1}x(\bar{x})^{k}\rho\bar{x} & \end{cases} \tag{6}$$

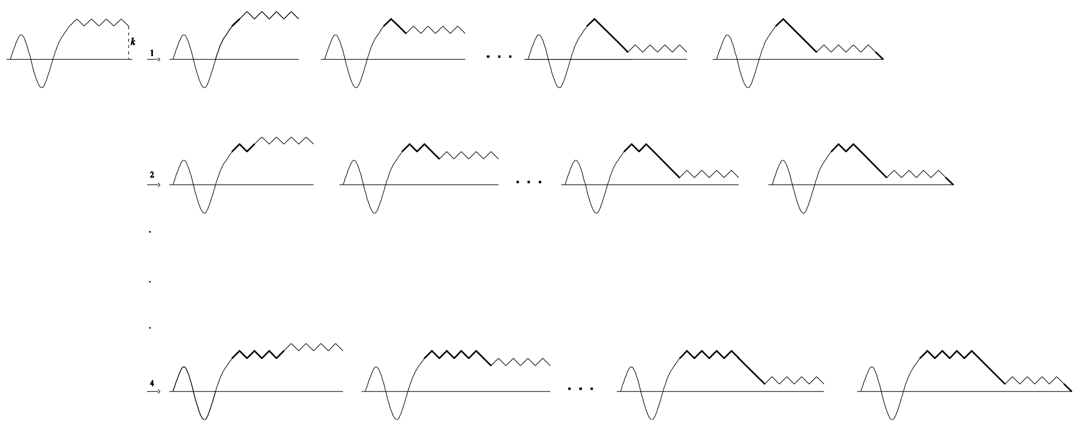

Figure 7: The Paths Generated by $\omega_{|k}(x\bar{x})^{j}$, $k \geq 2$

2.3 Paths Ending with $(x\bar{x})^{j-1}x$

The paths $\omega \in F[\mathrm{p}]$, ending on the x-axis with the sequence $(x\bar{x})^{j-1}x$ have the following shape $\omega_{|0} = \mu\bar{x}\eta x(\bar{x}x)^{j-1}$ where μ is a path ending on the x-axis and η is either the empty path ε or is a strongly negative path. The constructions applied to paths ending at ordinate 0 described in (1) (see Figure 2) can be used even for the paths ending with the sequence $(x\bar{x})^{j-1}x$ when $h \geq 2$, or to generate the paths ending at ordinate 1 or on the x-axis with a positive suffix, when h=1. However, when h=1, we would get an underground path which contains the forbidden sequence p $= (x\bar{x})^{j}x$. Therefore if the path ends with the sequence $(x\bar{x})^{j-1}x$ and h=1, in order to generate the underground path we proceed as follows. Two cases must be taken into consideration.

1) μ does not end with $x\bar{x}$. The underground path generated from $\omega_{|0} = \mu\bar{x}\eta x(\bar{x}x)^{j-1}$ is obtained by adding the path $\bar{x}x$ to $\omega_{|0}$, mirroring on x-axis the rightmost suffix $(\bar{x}x)^{j}$ of $\omega_{|0}\bar{x}x$, and shifting the sequence $(\bar{x}x)^{j}$ between μ and the sub-path $\bar{x}\eta x$ So the path $\omega_{|0} = \mu\bar{x}\eta x(\bar{x}x)^{j-1}$ generates the underground path $\mu(x\bar{x})^{j}\bar{x}\eta x$ (see Figure 8). It should be noticed that this construction applies to ω even if $\mu = \varepsilon$.

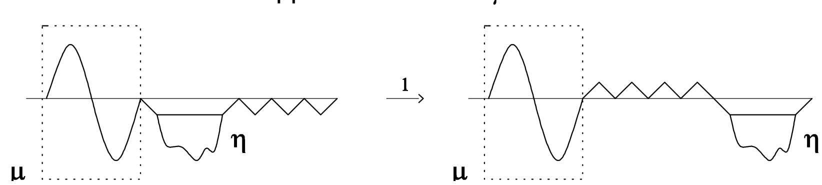

Figure 8: The underground path generated by $\omega_{|0}$ in the case 1)

2) μ ends with $x\bar{x}$. When the path μ ends with $x\bar{x}$, that is $\mu = \mu' x\bar{x}$, the insertion of the sequence $(x\bar{x})^j$ between μ and the sequence $\bar{x}\eta x$ produces the forbidden sequence p $=(x\bar{x})^j x$. Let us consider the following subcases: $\eta \neq \varepsilon$ and $\eta = \varepsilon$.

2.1) $\eta \neq \varepsilon$. The underground path is obtained by performing on $\omega_{|0} = \mu' x\bar{x}\bar{x}\eta x(\bar{x}x)^{j-1}$ the following operations: shift the rightmost up-down step $x\bar{x}$ of μ to the right of the subpath $\bar{x}\eta x$, mirror on x-axis the sequence $(\bar{x}x)^{j-1}$, and add to such path the steps $\bar{x}x$.

So, when h=1, the underground path with negative suffix generated by $\omega_{|0} = \mu' x\bar{x}\bar{x}\eta x(\bar{x}x)^{j-1}$ is $\mu'\bar{x}\eta x(x\bar{x})^j \bar{x}x$ (see Figure 9).

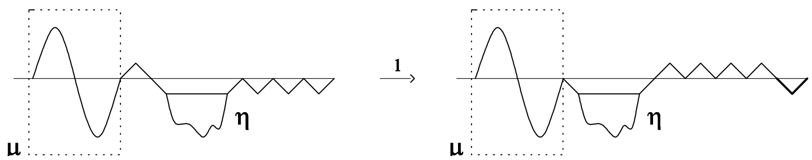

Figure 9: The Underground path Generated by $\omega_{|0}$ in the case 2.1)

2.2) $\eta = \varepsilon$. In this case, the underground path obtained by means of the construction described in 2.1) is $\omega' = \mu'\bar{x}x(x\bar{x})^j \bar{x}x$ and it contains the forbidden sequence p= $(x\bar{x})^j x$ if μ' ends with the sequence $(\bar{x}x)^j$ or with the sequence $\bar{x}\eta' x(\bar{x}x)^{j-1}$, where η' is a not empty and strongly negative path. Let us take the longest suffix of $\omega_{|0} = \mu' x\bar{x}(\bar{x}x)^j$ into account so that $\omega_{|0} = \varphi \nu_1 \nu_2 \ldots \nu_k$, where

$$\nu_1 = \bar{x}\lambda x(\bar{x}x)^{j-1} x\bar{x}$$

$$\nu_i = (\bar{x}x)^j x\bar{x} \qquad 1 < i < k$$

$$\nu_k = (\bar{x}x)^j$$

and λ is the empty path or is a strongly negative path. Every sequence ν_i, $1 \leq i \leq k$, will be changed into $\bar{\nu}_i$ in the following way:

2.2.1) if φ is a path that does not end with $x\bar{x}$, then

$$\bar{\nu}_1 = (x\bar{x})^j \bar{x}\lambda x$$

$$\bar{\nu}_i = (x\bar{x})^j \bar{x}x \qquad 1 < i < k$$

$$\bar{\nu}_k = (x\bar{x})^j \bar{x}x$$

and the underground path generated by $\omega_{|0}$ is $\varphi\bar{\nu}_1 \cdots \bar{\nu}_k$ (see Figure 10);

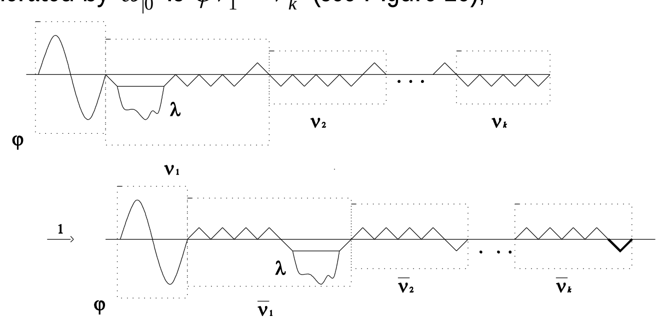

Figure 10: The Underground path Generated by $\omega_{|0}$ in the case 2.2.1)

2.2.2) if φ ends with $x\overline{x}$, that is $\varphi = \varphi' x\overline{x}$ then

$\overline{v}_1 = \overline{x}\lambda\, x(x\overline{x})^j\, \overline{x}x$

$\overline{v}_i = (x\overline{x})^j\, \overline{x}x \qquad 1 < i < k$

$\overline{v}_k = (x\overline{x})^j\, \overline{x}x$

and the underground path generated by $\omega_{|0}$ is $\varphi'\overline{v}_1 \cdots \overline{v}_k$ (see Figure 11).

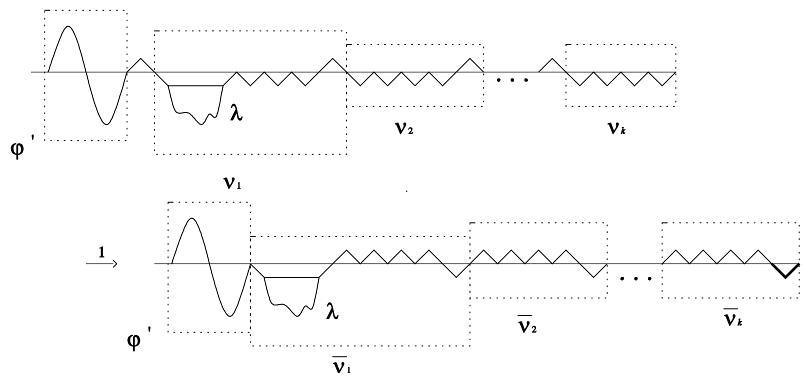

Figure 11: The Underground path Generated by $\omega_{|0}$ in the case 2.2.2)

2.4 The Underground path Generated by $\omega_{|k}$

Now let us describe how to obtain the underground path generated by $\omega_{|k}$ when $k \geq 2$. For each h, $1 \leq h \leq j$, let $\omega' = \nu\varphi$ be the path obtained from $\omega_{|k}$ and ending on the x-axis with a positive suffix, φ is the rightmost suffix in ω' which is primitive. If the path φ^c does not contain the forbidden sequence p, the underground path generated by $\omega_{|k}$ is $\nu\varphi^c$ If the path φ^c contains the forbidden sequence p, we must apply a *swap* operation Φ in order to obtain a path $\varphi_1 = \Phi(\varphi^c)$ avoiding the forbidden sequence. The underground path generated by $\omega_{|k}$ is $\nu\varphi_1$. Before describing the Φ operation on φ^c, let us consider the following proposition.

Proposition 1: Let $\mu \in F$[p] a primitive path; μ^c contains the forbidden sequence p = $(x\overline{x})^j x$ if and only if μ contains the sequence p' = $(\overline{x})^2 (x\overline{x})^j \overline{x}$. From Proposition 1 it follows that, if φ^c contains the forbidden sequence p, then it is preceded and followed by at least a rise step. Operation Φ must generate a path φ_1 avoiding the forbidden sequence p = $(x\overline{x})^j x$ and such that $\varphi_1{}^c \in F \setminus F$[p]; in this way φ_1 is not the complement of any path in F[p]. The path $\varphi_1 = \Phi(\varphi^c)$ is obtained in the following way:

i. consider the straight line r from the beginning of the sequence p = $(x\overline{x})^j x$ and let t_1 be the rightmost point in which r intersects φ^c on the left of p such that t_1 is preceded by at least two fall steps;
ii. let $\delta_2 = (x\overline{x})^m$, $0 \leq m < j$, the subsequence on the right of t_1, followed by at least a fall step;
iii. *swap* the initial subsequence $\delta_1 = (x\overline{x})^j$ of p and δ_2. Let us remark that δ_2 can not be equal to $(x\overline{x})^j$ as φ does not contain the forbidden sequence p = $(x\overline{x})^j x$ (see Figure 12.a)). When m=0, that is δ_2 is the empty word, we simply insert δ_1 into t_1 (see Figure 12.b)).

Operation Φ is applied to each forbidden sequence in φ^c.

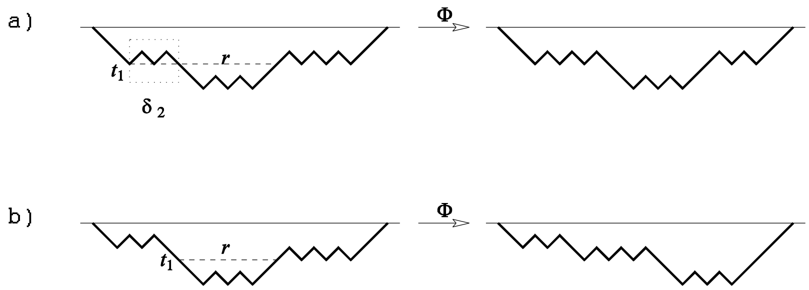

Figure 12: Some Examples of the Φ Operation, p= $(x\bar{x})^3 x$

Proposition 2: Let $\varphi_1 = \Phi(\varphi^c)$, then $\varphi_1^c \in F\backslash\, F$[p].

Proof.: The Φ operation transforms the subsequence $\rho_1 = (\bar{x})^m \delta_2 \bar{x}$, ($m \geq 2$), of φ^c into the subsequence $\rho_2 = (\bar{x})^m \delta_1 \bar{x} = (\bar{x})^m (x\bar{x})^j \bar{x}$ of φ_1.The complement of ρ_2 is

$$\rho_2^c = (x)^m (\bar{x}x)^j x = (x)^{m-1} (x\bar{x})^j xx.$$

So φ_1^c contains the forbidden sequence p = $(x\bar{x})^j x$.

Proposition 3: Let $\mu \in F\backslash\, F$[p] a primitive path such that $\mu^c \in F$[p]. Then there exists a path $\eta \in F$[p] such that $\mu^c = \Phi(\eta^c)$.

Proof: If $\mu \in F\backslash\, F$[p] and $\mu^c \in F$[p] then μ^c contains the sequence $\bar{x}\bar{x}(x\bar{x})^j \bar{x}$; we apply to μ^c the following operation Φ^{-1}:

i. consider the straight line *r* from the end of the sequence $(x\bar{x})^j$ and let t_2 be the leftmost point where *r* intersects μ^c on the right of $(x\bar{x})^j$ such that t_2 is followed by at least two rise steps;
ii. let $\delta_2 = (x\bar{x})^m$, $0 \leq m < j$, the subsequence on the left of t_2, preceded by at least a rise step;
iii. swap the subsequence $(x\bar{x})^j$ and δ_2. When m =0, that is δ_2 is the empty word, we simply insert $(x\bar{x})^j$ into t_2.

Figure 13 shows the initial steps of the generating algorithm of the paths corresponding to words in *F*[p], p = $(x\bar{x})^2 x = (10)^2 1$. Dotted lines are related to *h*=2.

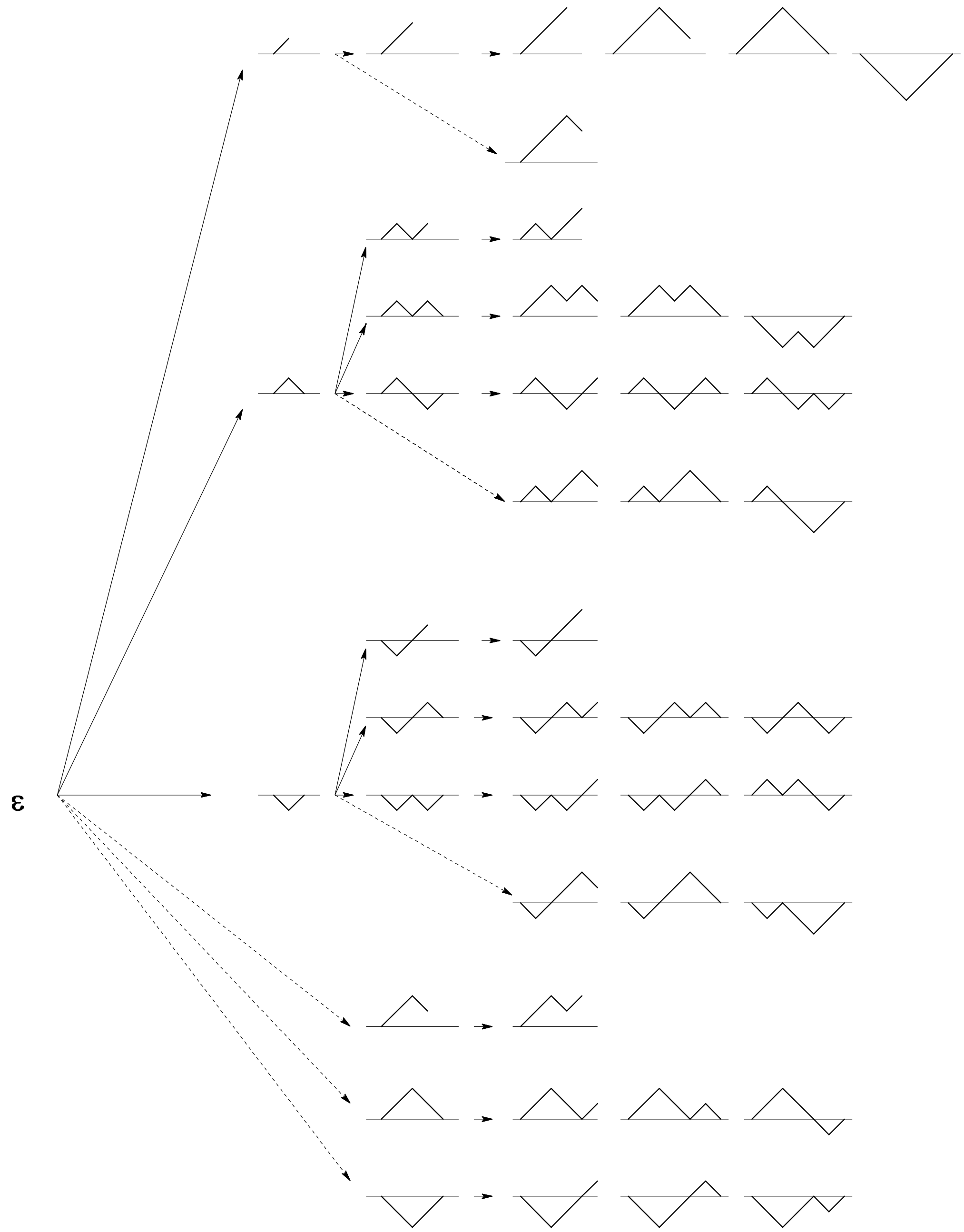

Figure 13: The Initial steps of the Generating Algorithm of the Paths corresponding to words in *F*[p], p= $(x\bar{x})^2 x = (10)^2 1.$

3. Exhaustive Generation

In this section we prove that the construction described in Section 2 allows to generate the set *F*[p] exhaustively for any fixed forbidden sequence p= $(10)^j 1$, $j \geq 1$, in the sense that all the words in *F*[p] with *n* 1's, $n \geq 0$, can be generated.

Theorem 1: Given a fixed forbidden sequence p $= (10)^j 1$, $j \geq 1$, the construction described in Section 2 generates all the paths with *n*, $n \geq 0$, rise steps representing the binary words in *F*[p] with *n* 1's.

Let $\omega \in$ *F*[p], then

$$\omega = \varphi_0 \varphi_1 \varphi_2 \cdots \varphi_s$$

that is, ω is made of s+1 subpaths such that:

- φ_0 is the empty path ε,
- φ_i , $1 \leq i < s$ is a path in F[p] beginning from and ending on the x-axis,
- φ_s is a path in F[p] beginning from the x-axis with endpoint at ordinate $k \geq 0$.

The proof of Theorem 1 is obtained by induction on the number of subpaths.

Proof: The empty path ε is generated as the starting point of the algorithm. Let us assume that all the possible subpaths $\varphi_1\varphi_2\cdots\varphi_i$ of ω are generated. We prove that the algorithm generates the subpath $\varphi_1\varphi_2\cdots\varphi_i\varphi_{i+1}$ for any path $\varphi_{i+1} \in F$[p].

Let φ_i be a path that does not end with $(x\overline{x})^j$. In this case the path φ_{i+1} may be either positive or negative.

If φ_{i+1} is a positive path we have to prove that all the positive paths are generated and this will be demonstrated in Section 3.1. In the case of a negative path, denoting d the largest value of its absolute ordinate, let us remark that:

- the negative paths with d =1 are $(\overline{x}x)^\ell$, $1 \leq \ell \leq j$, and they are generated by iterating the construction (1);
- negative paths with d =2 are $\overline{x}(\overline{x}x)^\ell x$, $1 \leq \ell \leq j$, and they are generated by means of (1) when $1 \leq \ell < j$, or by means of (4) when $j = \ell$;
- let γ be a negative path with d >2; if $\gamma^c \in F$[p] then γ is the underground path generated by a positive path with endpoint at ordinate $k \geq 2$, (see Section 2.4), otherwise $\gamma = \Phi(\eta^c)$ for a positive path η in F[p] with endpoint at ordinate $k \geq 2$ (see Proposition 3).

Note that when j =1 and φ_i ends with $\overline{x}x$, the only possible negative path φ_{i+1} with d =2 is $\overline{x}\overline{x}xx$ and it is generated by applying the construction (4) to the path $\varphi_i x\overline{x}$. When the suffix of φ_i is $(x\overline{x})^j$, φ_{i+1} must be a negative path and the path $\varphi_1\varphi_2\cdots\varphi_i\varphi_{i+1}$ is the underground path obtained by means of the construction described in case 1) in Section 2.3 (see Figure 8).

In the same way, when the subpath $\varphi_i\varphi_{i+1}$ is of type

$$(x\overline{x})^j\,\overline{x}\lambda x((x\overline{x})^j\,\overline{x}x)^r, \qquad r \geq 1$$

or

$$\overline{x}\lambda x((x\overline{x})^j\,\overline{x}x)^r, \qquad r \geq 1$$

where λ is the empty path ε or is a strongly negative path, then it is generated by the constructions described in case 2) in Section 2.3 (see Figure 9, 10, and 11).

Then, if we show that all the possible positive paths are generated, then we can claim that Theorem 1 is proved. Moreover, we observe that for each path ω in F[p] with n rise steps there exists one and only one path ω' in F[p] with n–h rise steps, $1 \leq h \leq j$, such that ω is obtained from ω' by means of the construction described in Section 2. This assertion is a direct consequence of the construction, since the actions described are univocally determined.

3.1 Positive Paths

In this section we prove that all the positive paths with n rise steps are generated by means of the construction described in Section 2. In the sequel of this section we analyze only *positive* paths. The proof is obtained by induction on n. There are only two paths with n =1 rise step, that is x and $x\overline{x}$, and they are generated by means of construction (1) applied to the empty path ε. Let us assume that all the paths with $n' < n$ rise steps are generated; we will prove that all the paths with n rise steps are generated.

Note that, following the construction given in Barcucci, Del Lungo, Pergola & Pinzani, [1], a path with n rise steps can be obtained from a Dyck path ω with n –1 rise step by inserting one rise step in each point at ordinate i of its last descent followed by q fall steps, $0 \le q \le i+1$. We will prove that all the paths obtained with this procedure are also generated following the constructions given in the above section. Let us denote by ω_q a paths ending with q fall steps. Let m be the number of fall steps in the last descent of ω. First of all, we note that for each value of m the paths obtained by inserting a rise step in the point at ordinate 0 are generated by means of the constructions (1) or (4) applied to the path ω. Here we give the proof for the case with $m>2$, distinguishing three cases: $i=1$, $(1<i<m-1) \vee (i=m)$, and $i=m-1$. The analogous and simple cases $m=1$ and $m=2$ are left to the reader.

- $i=1$. Let $\omega = \gamma_{|m-1} x(\bar{x})^m$. The insertion of a rise step in the point at ordinate 1 gives three paths:
  - $\omega_0 = \gamma_{|m-1} x(\bar{x})^{m-1} x$, which is generated by means of (2) applied to the prefix $\gamma_{|m-1} x(\bar{x})^{m-1}$ of ω,
  - $\omega_1 = \gamma_{|m-1} x(\bar{x})^{m-1} x\bar{x}$ and $\omega_2 = \gamma_{|m-1} x(\bar{x})^{m-1} x\bar{x}\bar{x}$, which are the paths with endpoints at ordinate 1 and 0, respectively. If $j>1$, then ω_1 and ω_2 are generated by the construction (3), where $k=m-1$, applied to the path $\gamma_{|m-1}$ with $h=2$, otherwise, if $j=1$, they are generated by means of (6) with $k=m-1$ and $h=1$ applied to the path $\gamma_{|m-1} x\bar{x}$.
- $(1<i<m-1) \vee (i=m)$. The insertion of a rise step in the point at ordinate i gives $i+2$ paths ω_q, $0 \le q \le i+1$. The paths ω_q with $q \ne 1$ are all the positive paths generated by means of (3) with $k=i$ and $h=1$ applied to the prefix of ω of length $|\omega|-i$. The path $\omega_1 = \gamma_{|m-1} x(\bar{x})^{m-1} x\bar{x}$ is the path with endpoint at ordinate i. When $j>1$, ω_1 is generated by means of construction (3) with $k=m-1$ and $h=2$ applied to the path $\gamma_{|m-1}$, whereas, when $j=1$, it is generated by means of (6), with $k=m-1$ and $h=1$, applied to the path $\gamma_{|m-1} x\bar{x}$.
- $i=m-1$. The insertion of a rise step in the point at ordinate $m-1$ generates $m+1$ paths ω_q, $0 \le q \le m$. The paths ω_q with $q \ne 1$ are all the positive paths generated by means of (3) with $k=m-1$ and $h=1$ applied to the prefix of ω of length $|\omega|-m+1$. As far as the generation of ω_1 is concerned, we have to distinguish three cases:

1. If $\omega = \gamma_{|m-2} x(x\bar{x})^{\ell-1}(\bar{x})^{m-1}, 1 \le \ell-1 < j$, then $\omega_1 = \gamma_{|m-2} x(x\bar{x})^{\ell}$. If $\ell < j$, then ω_1 is the path with endpoint at ordinate $(k+1)$ generated by means of (3) with $k=m-2$ and $h=\ell+1$ applied to the path $\gamma_{|m-2}$. If $\ell = j$, then ω_1 is the path with endpoint at ordinate $(k+1)$ generated by means of (6) with $k=m-2$ and $h=1$ applied to the path $\gamma_{|m-2}(x\bar{x})^{\ell}$. Note that, when $m=3$, the endpoint of the prefix $\gamma_{|m-2}$ has ordinate 1, and the path ω_1 is obtained applying the construction (2) (or (5)) instead of (3) (or (6)).
2. If $\omega = \gamma_{|m+m'-2} x(\bar{x})^{m'}(x\bar{x})^{\ell-1}(\bar{x})^{m-1}, 1 \le \ell-1 < j$ and $m'>2$, then $\omega_1 = \gamma_{|m+m'-2} x(\bar{x})^{m'}(x\bar{x})^{\ell}$. If $\ell < j$, then ω_1 is the path with endpoint at ordinate $(k-m'+1)$ generated by means of (3) with $k=m+m'-2$ and $h=\ell+1$ applied to the path $\gamma_{|m+m'-2}$. If $\ell = j$, then ω_1 is generated by means of (6) with $k=m+m'-2$ and $h=1$ applied to the path $\gamma_{|m+m'-2}(x\bar{x})^{\ell}$.
3. If $\omega = \gamma_{|m-2}(x\bar{x})^r x(x\bar{x})^{\ell-1}(\bar{x})^{m-1}, 0 \le r < j$, and $1 \le \ell-1 < j$, then $\omega_1 = \gamma_{|m-2}(x\bar{x})^r x(x\bar{x})^{\ell}$. If $\ell < j$, then ω_1 is the path with endpoint at ordinate $(k+1)$ generated by means of (3), with $k=m-2$ and $h=\ell+1$ applied to the path $\gamma_{|m-2}(x\bar{x})^r$. If $\ell = j$, then ω_1 is generated by means of (6), with $k=m-2$ and $h=r+1$ applied to the path $\gamma_{|m-2}(x\bar{x})^{\ell}$.

4. Conclusions and Further Developments

In this paper we propose an algorithm for the construction of particular binary words, according to the number of 1's, excluding a fixed sequence p = $(10)^j 1$, $j \geq 1$.

Successive studies should take into consideration binary words avoiding different forbidden sequences both from an enumerative and a constructive point of view. Moreover, it would be interesting to study words avoiding sequences which have a different shape, that is not only sequences of rise and fall steps. This could be the first step in the study of a possible universal generating algorithm for sequence avoiding words. Another interesting field of study is to determine a sort of invariant class of avoiding sequences that is the words $p_1, p_2, \ldots, p_l$, such that $|F[p_1]| = |F[p_2]| = \ldots = |F[p_l]|$ with consequent bijective problems. One could also consider a forbidden sequence on an arbitrary alphabet and investigate words avoiding that sequence, or study words avoiding more than one sequence and the related combinatorial objects, considering various parameters.

References

Barcucci, E., Del Lungo, A., Pergola, E., & Pinzani, R. (1999). ECO: a methodology for the enumeration of combinatorial objects. J. Difference Equ. Appl, 5, 435–490.

Bilotta, S., Merlini, D., Pergola, E., & Pinzani, R. (2012). Pattern $1^{j+1}0^j$ avoiding binary words. Fund. Inform., 117, 35–55.

Guibas, L. J., & Odlyzko, M. (1980). Long repetitive patterns in random sequences. Zeitschrift für Wahrscheinlichke its theorie, 53, 241–262.

Guibas, L. J., & Odlyzko, M. (1981). String overlaps, pattern matching, and nontransitive games. Journal of Combinatorial Theory, Series A, 30, 183–208.

Sedgewick, R., & Flajolet, P.(1995). An Introduction to the Analysis of Algorithms. London: Chapman-Hall.