Pattern 1^j0^i avoiding binary words

Stefano Bilotta   Elisa Pergola   Renzo Pinzani

Dipartimento di Sistemi e Informatica
Università degli Studi di Firenze
Viale G.B. Morgagni 65, 50134 Firenze, Italy

bilotta@dsi.unifi.it   elisa@dsi.unifi.it   pinzani@dsi.unifi.it

In this paper we study the enumeration and the construction, according to the number of ones, of particular binary words avoiding the fixed pattern $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$. The growth of such words can be described by particular jumping and marked succession rules. This approach enables us to obtain an algorithm which constructs all binary words having a fixed number of ones and then kills those containing the forbidden pattern.

1 Introduction

The problem of determining the appearance of a fixed *pattern* in long sequences of observation is interesting in many scientific problems.

For example in the area of computer network security, intrusions are becoming increasingly frequent and their detection is very important. Intrusion detection is primarily concerned with the detection of illegal activities and acquisitions of privileges that cannot be detected with information flow and access control models. There are several approaches to intrusion detection, but recently this subject has been studied in relation to pattern matching (see [1, 9, 12]).

In the area of computational biology, for example, it could be interesting to detect the occurrences of a particular pattern in a genomic sequence over the alphabet $\{A, G, C, T\}$, see for instance [16, 18].

These kinds of applications are interested in the study concerning both the enumeration and the construction of particular words avoiding a given pattern over an alphabet Σ.

In particular, binary words avoiding a fixed pattern $\mathfrak{p} = p_0...p_{h-1} \in \{0,1\}^h$ constitute a regular language and can be enumerated in terms of the number of bits 1 and 0 by using classical results (see, e.g., [10, 11, 17]). Recently, in [2, 13], this subject has been studied in relation to the theory of Riordan arrays.

In [5], the authors study the enumeration and the construction, according to the number of ones, of the class $F^{[\mathfrak{p}(j)]}$, that is, the class $F \subset \{0,1\}^*$ of binary words w excluding the fixed pattern $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$, such that $|w|_0 \leq |w|_1$ for any $w \in F$, $|w|_0$ and $|w|_1$ being the number of zeroes and ones in the word w, respectively. The enumeration problem, according to the number of ones, is solved algebraically by means of Riordan arrays theory. This approach gives a *jumping and marked succession rule* describing the growth of such words. Moreover, in [5] was introduced an algorithm for constructing all binary words having a fixed number of ones and excluding those containing the forbidden pattern $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$.

In this paper, we focus on the generalization of the fixed forbidden pattern $\mathfrak{p}$, passing from $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$ to $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$.

In this case the theory of Riordan arrays is not applicable, while it is possible to adapt the succession rule for the class $F^{[\mathfrak{p}(j)]}$ with $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$, to the class $F^{[\mathfrak{p}(j,i)]}$ for any $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$.

The paper is organized as follows. In Section 2 we give some basic definitions and notation related to the notions of succession rule and generating tree. In particular, we introduce the concept of *jumping and marked succession rules* (see [7, 8]) which are succession rules acting on the combinatorial objects of a class and producing sons at different levels where appear marked or non-marked labels.

In Section 3, we give a construction, according to the number of ones, for the set $F^{[\mathfrak{p}(j,i)]}$ for any fixed forbidden pattern $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$, by means of particular jumping and marked succession rules related to the form of the words in F.

2   Basic definitions and notations

A *succession rule* Ω is a system constituted by an *axiom* (a), with $a \in \mathbb{N}$, and a set of *productions* of the form:

$$(k) \rightsquigarrow (e_1(k))(e_2(k))\ldots(e_k(k)), \quad k \in \mathbb{N}, \, e_i : \mathbb{N} \to \mathbb{N}.$$

A production constructs, for any given label (k), its *successors* $(e_1(k)),(e_2(k)),\ldots,(e_k(k))$. In most of the cases, for a succession rule Ω, we use the compact notation:

$$\begin{cases} (a) \\ (k) \quad \rightsquigarrow (e_1(k))(e_2(k))\ldots(e_k(k)) \end{cases} \tag{1}$$

The rule Ω can be represented by means of a *generating tree*, that is a rooted tree whose vertices are the labels of Ω; where (a) is the label of the root and each node labelled (k) has k sons labelled $(e_1(k)),(e_2(k)),\ldots,(e_k(k))$, respectively. As usual, the root lies at level 0, and a node lies at level n if its parent lies at level $n-1$. If a succession rule describes the growth of a class of combinatorial objects, then a given object can be coded by the sequence of labels met from the root of the generating tree to the object itself. We refer to [3] for further details and examples.

The concept of a succession rule was introduced in [6] by Chung et al. to study reduced Baxter permutations, and was later applied to the enumeration of permutations with forbidden subsequences (for details see [4, 19]).

We remark that, from the above definition, a node labelled (k) has precisely k sons. A succession rule having this property is said to be *consistent*. However, we can also consider succession rules, introduced in [7], in which the value of a label does not necessarily represent the number of its sons, and this will be frequently done in the sequel.

Regular succession rules are not sufficient to handle all the enumeration problems and so we consider a slight generalization called *jumping succession rule* [8]. Roughly speaking, the idea is to consider a set of productions acting on the objects of a class and producing sons at different levels.

The usual notation to indicate a jumping succession rule is the following:

$$\begin{cases} (a) \\ (k) \quad \overset{1}{\rightsquigarrow} (e_1(k))(e_2(k))\ldots(e_k(k)) \\ (k) \quad \overset{j}{\rightsquigarrow} (d_1(k))(d_2(k))\ldots(d_k(k)) \end{cases} \tag{2}$$

The generating tree associated with (2) has the property that each node labelled (k) lying at level n has two sets of sons, the first set at level $n+1$ and having labels $(e_1(k)),(e_2(k)),\ldots,(e_k(k))$ and the second one at level $n+j$, with $j > 1$, and having labels $(d_1(k)),(d_2(k)),\ldots,(d_k(k))$, respectively.

Another generalization is introduced in [14], where the authors deal with *marked succession rules*. In this case the labels appearing in a succession rule can be marked or not, therefore *marked* labels are

considered together with usual ones. In this way a generating tree can support negative values if we consider a node labelled $(\overline{k})$ as opposed to a node labelled (k) lying on the same level.

A *marked generating tree* is a rooted labelled tree where appear marked or non-marked labels according to the corresponding succession rule. The main property is that, on the same level, marked labels kill or annihilate the non-marked ones with the same label value, in particular the enumeration of the combinatorial objects in a class is given by the difference between the number of non-marked and marked labels lying on a given level.

For any label (k), we introduce the following notation for generating tree specifications:

$$(\overline{\overline{k}}) = (k);$$
$$(k)^n = \underbrace{(k)\dots(k)}_{n}, \ \ n > 0.$$

Each succession rule (1) can be trivially rewritten as (3)

$$
\begin{cases}
(a) \\
(k) & \rightsquigarrow (e_1(k))(e_2(k))\dots(e_k(k))(k) \\
(k) & \rightsquigarrow (\overline{k})
\end{cases}
\tag{3}
$$

For example, the classical succession rule for Catalan numbers can be rewritten in the form (4) and Figure 1 shows some levels of the associated generating tree.

$$
\begin{cases}
(2) \\
(k) & \rightsquigarrow (2)(3)\dots(k)(k+1)(k) \\
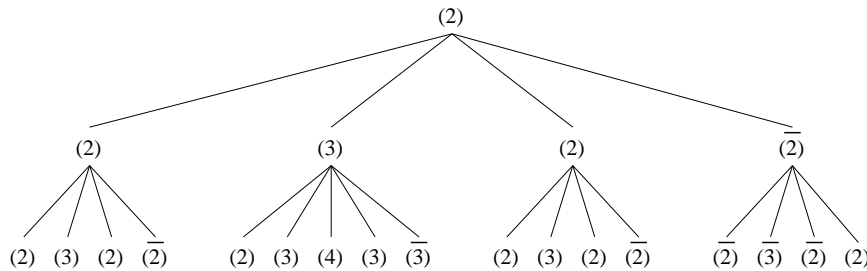(k) & \rightsquigarrow (\overline{k})
\end{cases}
\tag{4}
$$

Figure 1: Three levels of the generating tree associated with the succession rule (4)

The concept of marked labels has been implicity used for the first time in [15], then in [7] in relation with the introduction of the signed ECO-systems. In Section 3, we show how marked succession rules appear in the enumeration of a class of particular binary words according to the number of ones. Let $F \subset \{0,1\}^*$ be the class of binary words w such that $|w|_0 \le |w|_1$ for any $w \in F$, $|w|_0$ and $|w|_1$ being the number of zeroes and ones in w, respectively.

In this paper we are interested in studying the subclass $F^{[\mathfrak{p}]} \subset F$ of binary words excluding a given pattern $\mathfrak{p} = p_0 \dots p_{h-1} \in \{0,1\}^h$, i.e. the word $w \in F^{[\mathfrak{p}]}$ that does not admit a sequence of consecutive indices $i, i+1, \dots, i+h-1$ such that $w_i w_{i+1} \dots w_{i+h-1} = p_0 p_1 \dots p_{h-1}$. Each word $w \in F$ can be naturally represented as a lattice path on the Cartesian plane by associating a *rise step*, defined by $(1,1)$ and denoted by x, to each 1's in F, and a *fall step*, defined by $(1,-1)$ and denoted by $\overline{x}$, to each 0's in F. From now on, we refer interchangeably to words or their graphical representations on the Cartesian plane, that is paths.

3   A construction for the class $F^{[\mathfrak{p}(j,i)]}$

In this section, we study the enumeration and the construction for the set $F^{[\mathfrak{p}(j,i)]}$, where $\mathfrak{p}(j,i) = x^j\overline{x}^i = 1^j0^i$, $0 < i < j$, by setting jumping and marked succession rules describing the growth of the set. The succession rules, according to the number of rise steps or equivalently the number of ones, are related to the form of the lattice paths in F.

First of all, we define a *marked forbidden pattern* $\mathfrak{p}(j,i)$ as a pattern $\mathfrak{p}(j,i) = x^j\overline{x}^i$, $0 < i < j$, whose steps cannot be divided, they must lie always in that defined sequence. Therefore, a cut operation is not possible within a marked forbidden pattern.

We denote a marked forbidden pattern by marking its peak. We say that a point is strictly contained in a marked forbidden pattern if it is between two steps of the pattern itself.

In order to study the enumeration and the construction for the class $F^{[\mathfrak{p}(j,i)]}$, we have to distinguish two cases depending on the form of the paths in F.

Definition 3.1 *A path ω in F is a Δ-path if:*

- *it ends on the x-axis (see Figure 2.a));*

- *the ordinate of its endpoint is greater than 0 and its rightmost suffix ρ begins from the x-axis by a rise step and strictly remains above the x-axis itself. The suffix ρ can contain marked forbidden patterns $\mathfrak{p}(j,i)$ (see Figure 2.b)) or not (see Figure 2.c)). If ρ contains marked forbidden patterns $\mathfrak{p}(j,i)$, then their marked points have ordinate $b \geq j$.*

Definition 3.2 *A path ω in F is a Γ-path if the ordinate of its endpoint is greater that 0 and its rightmost suffix ρ begins from the x-axis by a fall step and contains a marked forbidden pattern $\mathfrak{p}(j,i)$ with ordinate b, $i < b < j$ (see Figure 2.d)).*

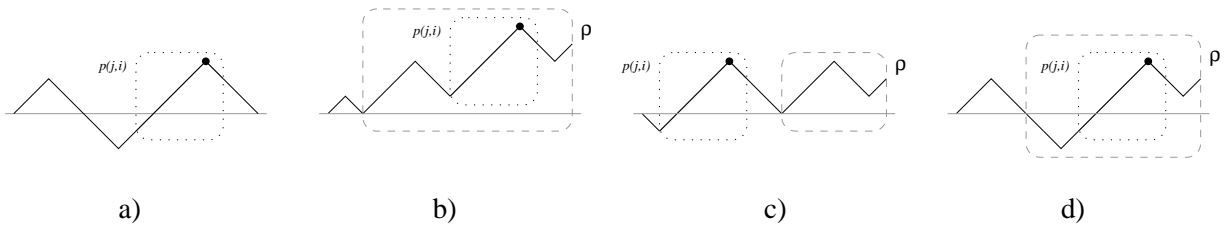a)   b)   c)   d)

Figure 2: Some examples of paths in F

3.1 Δ-paths in F

For each Δ-path ω in F having k as the ordinate of its endpoint, we apply the succession rule (5), for each $k \geq 0$:

$$\begin{cases} (0) & \\ (k) & \overset{1}{\rightsquigarrow} (0)^2(1)(2)\cdots(k)(k+1) \\ (k) & \overset{j}{\rightsquigarrow} (\overline{0})^{j-i+1-a}(\overline{1})^{j-i-a}(\overline{2})^{j-i-1-a}\ldots(\overline{j-i-1-a})^2(\overline{j-i-a})\ldots(\overline{k+j-i}) \end{cases} \quad (5)$$

In the second production of (5), the parameter a, with $0 \leq a \leq j-i-1$, is related to the form of the Δ-path ω and the way to set a will be described later in this section.

At this point, we define an algorithm which associates a Δ-path in F to a sequence of labels obtained by means of the succession rule (5).

The axiom (0) is associated to the empty path ε.

A Δ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $k+3$ lattice paths, with $n+1$ rise steps, according to the first production of (5) having $0,0,1,\dots,k+1$ as endpoint ordinate, respectively. The last $k+2$ labels are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-y$, $0 \le y \le k+1$, fall steps (see Figure 3).

Each lattice path so obtained has the property that its rightmost suffix beginning from the x-axis, either remains strictly above the x-axis itself or ends on the x-axis by a fall step. Note that in this way, the paths ending on the x-axis and having a rise step as last step are never obtained. These paths are bound to the first label (0) of the first production in (5) and the way to obtain them will be described later in this section.
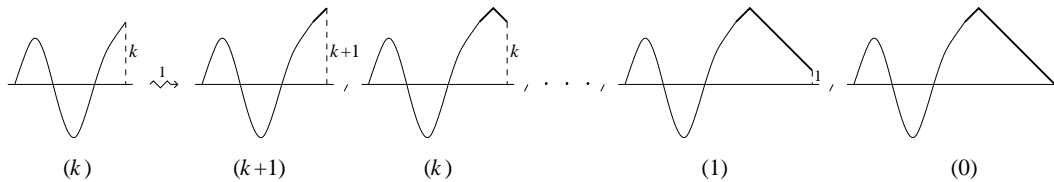

$$(k) \qquad\qquad (k+1) \qquad\qquad (k) \qquad\qquad\qquad (1) \qquad\qquad\qquad (0)$$

Figure 3: The mapping associated to $(k) \overset{1}{\rightsquigarrow} (0)(1)(2)\dots(k+1)$ of (5)

Let the parameter a be fixed, a Δ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $1+k+j-i+\sum_{m=0}^{j-i-a-1} j-i-a-m$ lattice paths, with $n+j$ rise steps, according to the second production of (5) such that $1+k+j-i$ lattice paths having $0,1,2,\dots,j-i-a,\dots,k+j-i$ as endpoint ordinate, respectively, and $j-i-a-m$ lattice paths having m as endpoint ordinate, $0 \le m \le j-i-a-1$. The first $1+k+j-i$ lattice paths are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j \overline{x}^i$ followed by $k+j-i-y$ fall steps, $0 \le y \le k+j-i$ (see Figure 4).

Each lattice path so obtained has the property that its rightmost suffix beginning from the x-axis, either remains strictly above the x-axis itself or ends on the x-axis by a fall step. At this point the first label (0) according to the first production of (5) and the $j-i-a-m$ labels $(\overline{m})$, $0 \le m \le j-i-a-1$, according to the second production of (5), must give lattice paths which either do not contain marked forbidden pattern in its rightmost suffix and end on the x-axis by a rise step or having the rightmost marked point with ordinate less than j.

In order to obtain the first label (0) according to the first production of (5), we consider the lattice path ω' obtained from ω by adding a sequence of steps consisting of one rise step followed by k fall steps. By applying the previous actions, a path ω' can be written as $\omega' = v\varphi$, where φ is the rightmost suffix in ω' beginning from the x-axis and strictly remaining above the x-axis.

We distinguish two cases: in the first one φ does not contain any marked point and in the second one φ contains at least one marked point.

If the suffix φ does not contain any marked point, then the desired label (0) is associated to the path $v\varphi^c x$, where φ^c is the path obtained from φ by switching rise and fall steps (see Figure 5).

If the suffix φ contains marked points, let r be the rightmost and highest marked point in φ and let t be the nearest and highest point on the right of the marked forbidden pattern containing r not being strictly within a marked forbidden pattern. We then consider the straight line s through the point t and the leftmost and highest point z in φ lying above or on the line s and which is not strictly within a marked
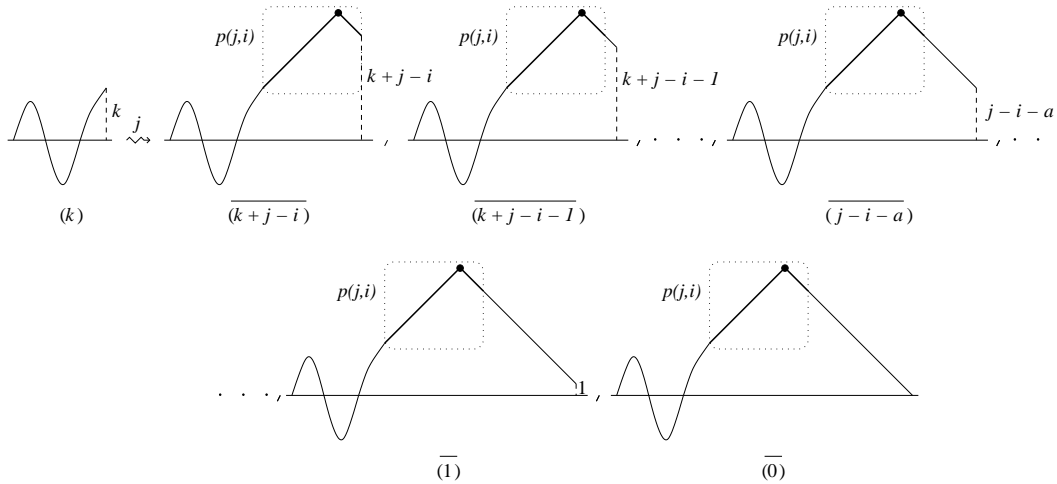
Figure 4: The mapping associated to $(k) \overset{j}{\rightsquigarrow} (\overline{0})(\overline{1})(\overline{2})\dots(\overline{j-i-a})\dots(\overline{k+j-i})$ of (5)
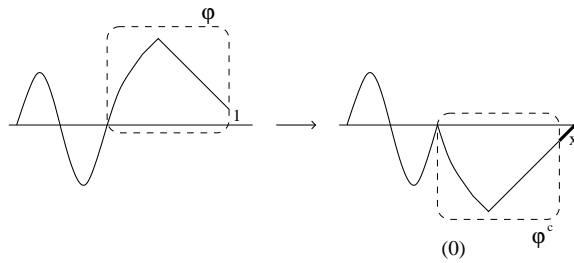

Figure 5: A graphical representation of the actions giving the first label (0) in case of no marked points in φ

forbidden pattern (see Figure 6). Obviously if the straight line s does not intersect any point on the left of t or intersects only points lying strictly within a marked forbidden pattern, then $z \equiv t$.
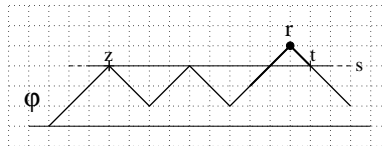

Figure 6: Marked point in the suffix φ: an example with the pattern $\mathfrak{p}(j,i) = x^2\overline{x}$

The desired label (0) is associated to the path obtained by applying the cut and paste actions which consist on the concatenation of a fall step $\overline{x}$ with the path in φ running from z to the endpoint of the path, say α, and the path running from the initial point in φ to z, say β (see Figure 7).

In order to obtain the $j-i-a-m$ labels $(\overline{m})$, $0 \leq m \leq j-i-a-1$, according to the second production of (5), we consider the paths ω'' obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j\overline{x}^i$ followed by y fall steps, $k+a \leq y \leq k+j-i-1$. Therefore, we consider the just obtained paths labelled with $(\overline{k+j-i-y})$, $k+a \leq y \leq k+j-i-1$, which are represented in Figure 4.

By performing on each ω'' the cut and paste actions, we obtain $j-i-a$ paths labelled with $(\overline{k+j-i-y-1})$, $k+a \leq y \leq k+j-i-1$. By adding g fall steps, $0 < g \leq k+j-i-y-1$, to each of such paths (see Figure 8), we obtain the complete mapping associated with the second production of (5).
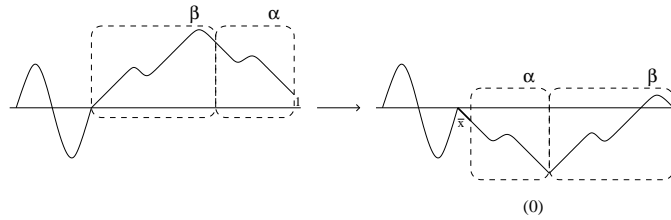
Figure 7: A graphical representation of the cut and paste actions giving the first label (0) in case of marked points in φ

Note that, we apply the cut and paste actions to the paths ω'' exclusively. Indeed, by performing the cut and paste actions to the paths obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j \overline{x}^i$ followed by y fall steps, $0 \le y < k+a$, we have repeated paths.

In the following is explained the way to set the parameter a related to the form of the Δ-path ω in F having k as ordinate of its endpoint.

- If the Δ-path ω in F has the ordinate k of its endpoint equal to 0 (or equivalently ends on the x-axis), then $a = 0$ and we apply to the path ω the production (6) for the second production of (5).

$$(0) \stackrel{j}{\leadsto} (\overline{0})^{j-i+1}(\overline{1})^{j-i}(\overline{2})^{j-i-1}\ldots(\overline{j-i-1})^2(\overline{j-i}) \qquad (6)$$

  We can observe that, for $k = 0$, the paths ω'' related to the previous construction are the paths obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j \overline{x}^i$ followed by y fall steps, for any y with $0 \le y \le j-i-1$. In this case the point z for the cut and paste actions is always the endpoint of the path $\omega x^j \overline{x}^i$ having ordinate $j-i$. Figure 9 shows the complete mapping associated to (6) on an example with the pattern $\mathfrak{p}(j,i) = x^5 \overline{x}^2$.

- If the Δ-path $\omega = \upsilon\rho$ in F has the ordinate of its endpoint greater than 0, then we must have to distinguish two cases: in the first one the rightmost suffix ρ in ω does not contain any marked points and in the second one ρ contains at least one marked point.

  The suffix ρ in ω does not contain any marked point. We denote by h the ordinate of the peak in ρ having highest height. We consider the endpoint of the path $\omega x^j \overline{x}^i$ having ordinate $k+j-i$ obtained from ω by applying the mapping $(k) \stackrel{j}{\leadsto} \overline{(k+j-i)}$ (see Figure 10) and we distinguish three subcases: $k+j-i \le h$, $h < k+j-i < h+j-i$ and $k+j-i \ge h+j-i$.

  ○ If $k+j-i \ge h+j-i$ (or equivalently $h-k \le 0$), then $a = 0$ and we apply to the path ω the production (7) for the second production of (5) similarly to the case $k = 0$.

$$(k) \stackrel{j}{\leadsto} (\overline{0})^{j-i+1}(\overline{1})^{j-i}(\overline{2})^{j-i-1}\ldots(\overline{j-i-1})^2(\overline{j-i})\ldots(\overline{k+j-i}) \qquad (7)$$

  Note that, if the suffix ρ in ω does not contain any peak (or equivalently $h = 0$) then we apply the production (7) for the second production of (5). We can observe that the paths ω'' related to the previous construction are the paths obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j \overline{x}^i$ followed by y fall steps, for any y with $k \le y \le k+j-i-1$. In this case, the point z for the cut and paste actions is always the endpoint of the path $\omega x^j \overline{x}^i$ having ordinate $k+j-i$.
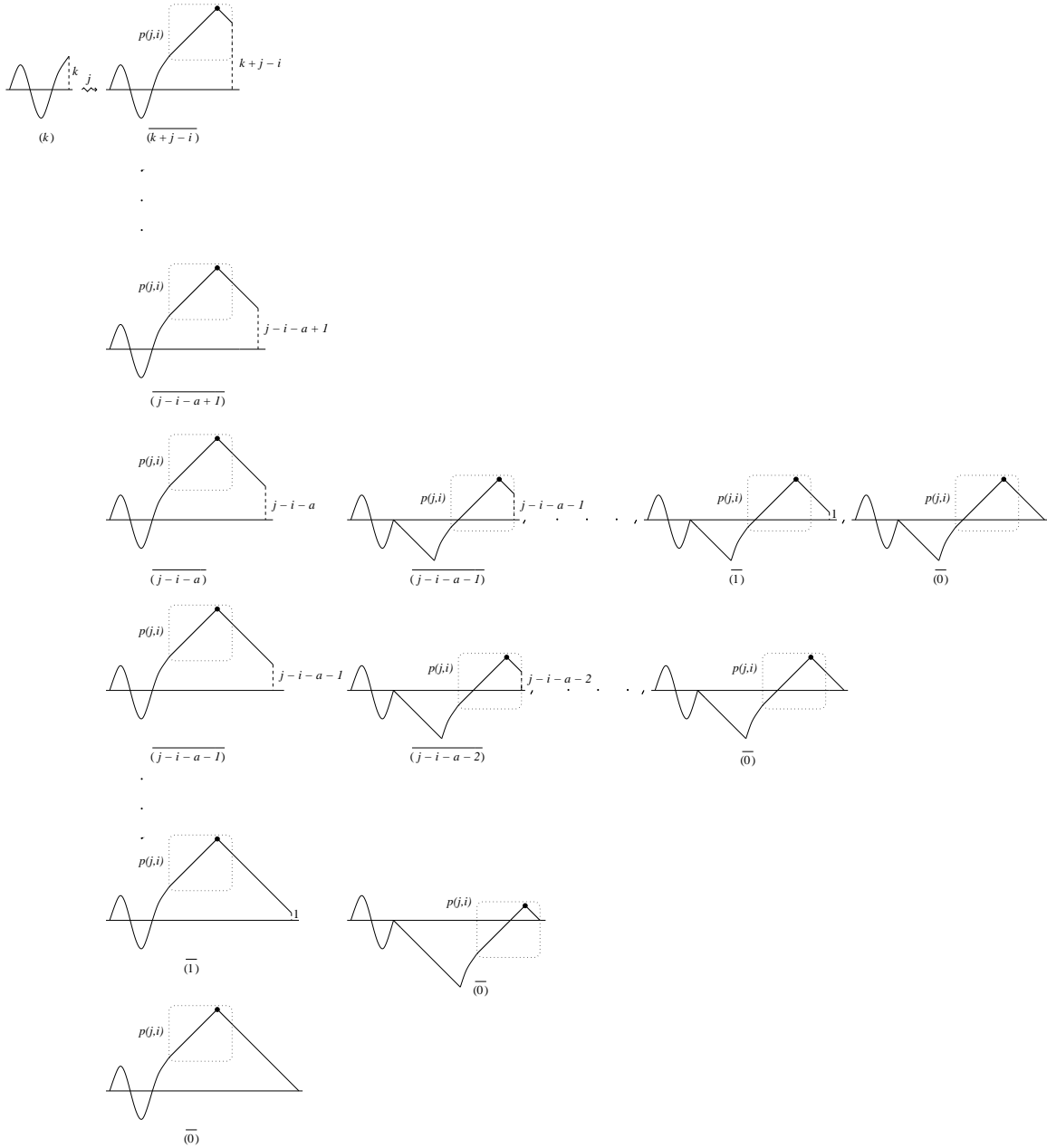
Figure 8: The mapping associated to
$(k) \overset{j}{\leadsto} (\overline{0})^{j-i+1-a}(\overline{1})^{j-i-a}(\overline{2})^{j-i-1-a}\ldots(\overline{j-i-1-a})^2(\overline{j-i-a})\ldots(\overline{k+j-i})$ of (5)

      ○ If $h < k+j-i < h+j-i$ (or equivalently $0 < h-k < j-i$), then $a = h-k$ and we apply to
         the path ω the production (8) for the second production of (5).

$$(k) \overset{j}{\leadsto} (\overline{0})^{j-i+1-(h-k)}(\overline{1})^{j-i-(h-k)}(\overline{2})^{j-i-1-(h-k)}\ldots$$
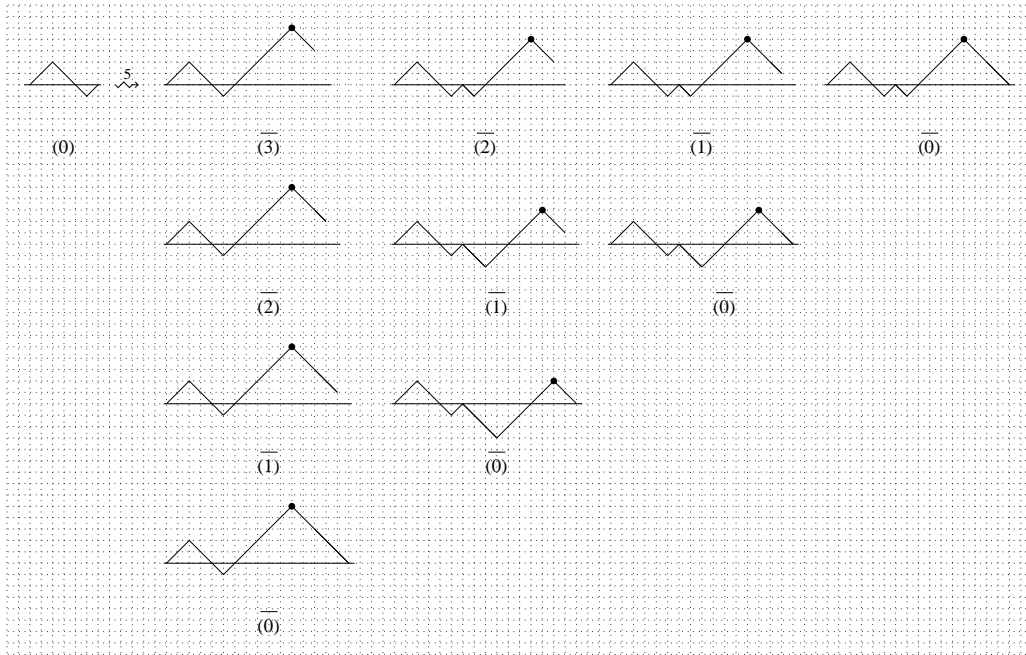$$\ldots(\overline{j-i-1-(h-k)})^2(\overline{j-i-(h-k)})\ldots(\overline{k+j-i}) \tag{8}$$

Figure 9: The mapping associated to $(0) \overset{5}{\leadsto} (\overline{0})^4 (\overline{1})^3 (\overline{2})^2 (\overline{3})$ on an example
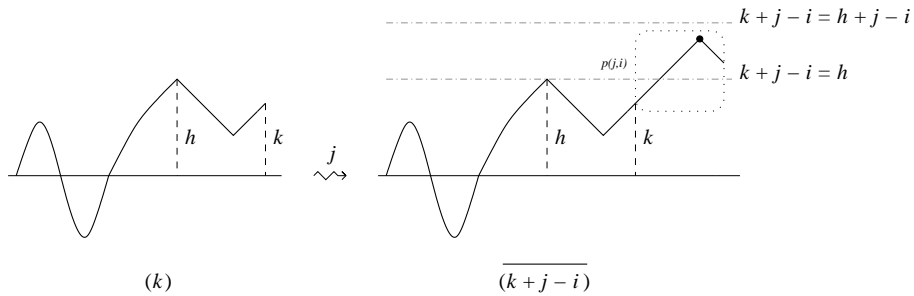

Figure 10: A graphical representation of the path $\omega x^j \overline{x}^i$ when the suffix ρ in ω does not contain any marked point

Also in this case, the point z for the cut and paste actions is always the endpoint of the path $\omega x^j \overline{x}^i$ having ordinate $k + j - i$. The paths ω'' related to the previous construction are the paths obtained from ω by adding a sequence consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j \overline{x}^i$ followed by y falls steps, $h \leq y \leq k + j - i - 1$.

○ If $k + j - i \leq h$ (or equivalently $h - k \geq j - i$), then $a = j - i - 1$ and we apply to the path ω the production (9) for the second production of (5).

$$(k) \overset{j}{\leadsto} (\overline{0})^2 (\overline{1})(\overline{2}) \ldots (\overline{k+j-i}) \tag{9}$$

In this case the point z for the cut and paste actions is always the point of peak in ρ having ordinate h, so we have only one path ω'' related to the previous construction, that is the path obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j \overline{x}^i$ followed by $k + j - i - 1$ falls steps.

The suffix ρ in ω contains at least one marked point. We denote by h the ordinate of the no marked peak in ρ having highest height and by h^* the ordinate of the marked peak in ρ having

highest height. We consider the endpoint of the path $\omega x^j \overline{x}^i$ having ordinate $k + j - i$ obtained from ω by applying the mapping $(k) \overset{j}{\leadsto} \overline{(k + j - i)}$ and we distinguish three subcases: $h^* - h < i$ (see the left side of Figure 11), $h^* - h > i$ (see the right side of Figure 11) and $h^* - h = i$.
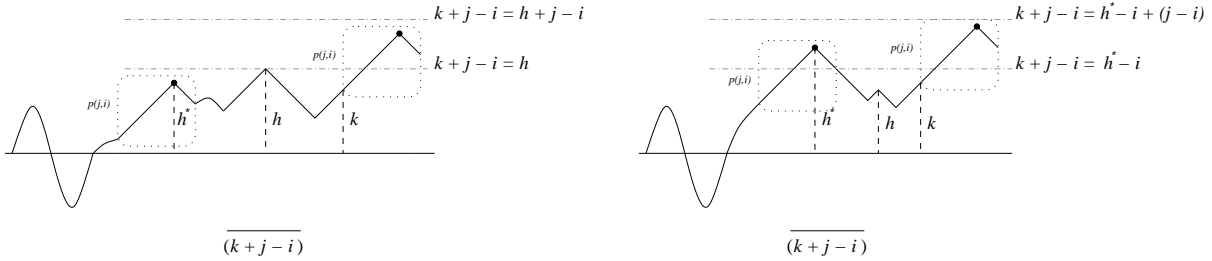

Figure 11: A graphical representation of the path $\omega x^j \overline{x}^i$ when the suffix ρ in ω contains at least one marked point

○ If $h^* - h < i$, just consider the no marked peak having ordinate h to set the parameter a, then we apply to the path ω the productions related to the case in which the suffix ρ in ω does not contain any marked point.

○ If $h^* - h > i$, then we consider three subcases: $k + j - i \geq h^* - i + (j - i)$, $h^* - i < k + j - i < h^* - i + (j - i)$ and $k + j - i \leq h^* - i$.

⋄ If $k + j - i \geq h^* - i + (j - i)$ (or equivalently $h^* - h \leq i$), then $a = 0$ and we apply to the path ω the production (7) for the second production of (5). We can observe that the paths ω'' related to the previous construction are the paths obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j, i) = x^j \overline{x}^i$ followed by y fall steps, for any y with $k \leq y \leq k + j - i - 1$. In this case the point z for the cut and paste actions is always the endpoint of the path $\omega x^j \overline{x}^i$ having ordinate $k + j - i$.

⋄ If $h^* - i < k + j - i < h^* - i + (j - i)$ (or equivalently $i < h^* - h < j$), then $a = h^* - k - i$ and we apply to the path ω the production (10) for the second production of (5).

$$(k) \overset{j}{\leadsto} (\overline{0})^{j - i + 1 - (h^* - k - i)} (\overline{1})^{j - i - (h^* - k - i)} (\overline{2})^{j - i - 1 - (h^* - k - i)} \ldots$$

$$\ldots (\overline{j - i - 1 - (h^* - k - i)})^2 (\overline{j - i - (h^* - k - i)}) \ldots (\overline{k + j - i}) \tag{10}$$

Also in this case, the point z for the cut and paste actions is always the endpoint of the path $\omega x^j \overline{x}^i$ having ordinate $k + j - i$. The paths ω'' related to the previous construction are the paths obtained from ω by adding a sequence consisting of the marked forbidden pattern $\mathfrak{p}(j, i) = x^j \overline{x}^i$ followed by y falls steps, $h^* - i \leq y \leq k + j - i - 1$.

⋄ If $k + j - i \leq h^* - i$ (or equivalently $h^* - h \geq j$), then $a = j - i - 1$ and we apply to the path ω the production (9) for the second production of (5). In this case the point z for the cut and paste actions is always the first point, having ordinate $h^* - i$, on the right of the marked forbidden pattern $\mathfrak{p}(j, i) = x^j \overline{x}^i$ having ordinate h^*, so we have only one path ω'' related to the previous construction, that is the path obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j, i) = x^j \overline{x}^i$ followed by $k + j - i - 1$ falls steps.

○ If $h^* - h = i$ and the no marked peak having ordinate h is on the left of the marked peak having ordinate h^* then we apply to the path ω the productions related to the case in which the suffix ρ in ω does not contain any marked point, otherwise if $h^* - h = i$ and the no marked

peak having ordinate h is on the right of the marked peak having ordinate h^* then we apply to the path ω the productions related to the case $h^* - h > i$.

3.2 Γ-paths in F

For each Γ-path ω in F having k as ordinate of its endpoint, we apply the following succession rule (11), for each $k \geq 1$:

$$\begin{cases} (k) & \overset{1}{\rightsquigarrow} (0)(1)(2)\cdots(k)(k+1) \\ (k) & \overset{j}{\rightsquigarrow} (\overline{0})(\overline{1})(\overline{2})\cdots(\overline{k+j-i-1})(\overline{k+j-i}) \end{cases} \tag{11}$$

A Γ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $k+2$ lattice paths, with $n+1$ rise steps, according to the first production of (11) having $0, 1, \ldots, k+1$ as endpoint ordinate, respectively. Those labels are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-y$, $0 \leq y \leq k+1$, fall steps.

Moreover, a Γ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $1+k+j-i$ lattice paths, with $n+j$ rise steps, according to the second production of (11) having $0, 1, 2, \ldots, k+j-i$ as endpoint ordinate, respectively. Those labels are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p} = x^j \overline{x}^i$ followed by $k+j-i-y$ fall steps, $0 \leq y \leq k+j-i$.

The just described construction, both for Δ-paths and Γ-paths in F, generates 2^C copies, $C \geq 0$, of each path having C forbidden patterns such that 2^{C-1} are coded by a sequence of labels ending by a marked label, say $(\overline{k})$, and contain an odd number of marked forbidden pattern, and 2^{C-1} are coded by a sequence of labels ending by a non-marked label, say (k), and contain an even number of marked forbidden pattern.

For brevity sake, we omit the proof of the fact that the described algorithm is a construction for $F^{[\mathfrak{p}(j,i)]}$, where $\mathfrak{p}(j,i) = x^j \overline{x}^i = 1^j 0^i$, $0 < i < j$. In order to prove the theorem we should have to show that the described actions are uniquely invertible.

4   Conclusions and further developments

In this paper we propose an algorithm for the construction, according to the number of ones, of particular binary words excluding a fixed pattern $\mathfrak{p}(j,i) = 1^j 0^i$, $0 < i < j$.

Successive studies should take into consideration binary words avoiding different forbidden patterns both from an enumerative and a constructive point of view.

Afterwards, it should be interesting to study words avoiding patterns having a different shape, that is not only patterns consisting of a sequence of rise steps followed by a sequence of fall steps. This could be the first step to investigate a possible uniform generating algorithm for pattern avoiding words.

One could also consider a forbidden pattern on an arbitrary alphabet and investigating words avoiding that pattern.

Finally, we could think to study words avoiding more than one pattern and the related combinatorial objects, considering various parameters.

References

[1] A. Apostolico & M. Atallah (2002): *Compact recognizers of episode sequences.* Information and Computation 174(2), pp. 180-192, doi:`10.1006/inco.2002.3143`.

[2] D. Baccherini, D. Merlini & R. Sprugnoli (2007): *Binary words excluding a pattern and proper Riordan arrays.* Discrete Mathematics 307, pp. 1021-1037, doi:`10.1016/j.disc.2006.07.023`.

[3] E. Barcucci, A. Del Lungo, E. Pergola & R. Pinzani (1999): *ECO: a methodology for the Enumeration of Combinatorial Objects.* Journal of Difference Equations and Applications 5, pp. 435-490, doi:`10.1080/10236199908808200`.

[4] E. Barcucci, A. Del Lungo, E. Pergola & R. Pinzani (2000): *From Motzkin to Catalan permutations.* Discrete Mathematics 217, pp. 33-49, doi:`10.1016/S0012-365X(99)00254-X`.

[5] S. Bilotta, D. Merlini, E. Pergola & R. Pinzani (2010): *Binary words avoiding a pattern and marked succession rule.* In: Lattice Path Combinatorics and Applications, Siena. Available at `http://arxiv.org/abs/1103.5689`.

[6] F. R. K. Chung, R. L. Graham, V. E. Hoggatt & M. Kleimann (1978): *The number of Baxter permutations.* Journal of Combinatorial Theory, Series A 24, pp. 382-394, doi:`10.1016/0097-3165(78)90068-7`.

[7] S. Corteel (2000): *Séries génératrices exponentielles pour les ECO-systèmes signés.* In: Formal Power Series and Algebraic Combinatorics, Moscow.

[8] L. Ferrari, E. Pergola, R. Pinzani & S. Rinaldi (2003): *Jumping succession rules and their generating functions.* Discrete Mathematics 271, pp. 29-50, doi:`10.1016/S0012-365X(02)00868-3`.

[9] P. Flajolet, W. Szpankowski & B. Valle (2006): *Hidden word statistics.* Journal of the ACM 53(1), pp. 147-183, doi:`10.1145/1120582.1120586`.

[10] L. J. Guibas & M. Odlyzko (1980): *Long repetitive patterns in random sequences.* Zeitschrift fur Wahrscheinlichkeitstheorie 53, pp. 241-262, doi:`10.1007/BF00531434`.

[11] L. J. Guibas & M. Odlyzko (1981): *String overlaps, pattern matching, and nontransitive games.* Journal of Combinatorial Theory, Series A 30, pp. 183-208, doi:`10.1016/0097-3165(81)90005-4`.

[12] S. Kumar & E.H. Spafford (1994): *A pattern matching model for misuse intrusion detection.* In: Computer Security Conference, pp. 11-21.

[13] D. Merlini & R. Sprugnoli (2011): *Algebraic aspects of some Riordan arrays related to binary words avoiding a pattern.* Theoretical Computer Science 412(27), pp. 2988-3001, doi:`10.1016/j.tcs.2010.07.019`.

[14] D. Merlini, R. Sprugnoli & M. C. Verri (2000): *An Algebra for proper generating tree.* In: Algorithms, trees, combinatorics and probabilities. Trends in Mathematics, Mathematics and Computer Science, pp. 127-139.

[15] D. Merlini & M. C. Verri (2003): *Generating trees and proper Riordan Arrays.* Discrete Mathematics 218, pp. 167-183, doi:`10.1016/S0012-365X(99)00343-X`.

[16] I. Rigoutsos, A. Floratos, L. Parida, Y. Gao & D. Platt (2000): *The emergence of pattern discovery techniques in computational biology.* Metabolic Engineering 2, pp. 159-177, doi:`10.1006/mben.2000.0151`.

[17] R. Sedgewick & P. Flajolet (1996): *An Introduction to the Analysis of Algorithms.* Addison- Wesley, Reading, MA.

[18] M. Waterman (1995): *Introduction to Computational Biology.* Chapman-Hall, London.

[19] J. West (1995): *Generating trees and the Catalan and Schröder numbers.* Discrete Mathematics 146, pp. 247-262, doi:`10.1016/0012-365X(94)00067-1`.