Unbordered Pictures: Properties and Construction

Marcella Anselmo[1], Dora Giammarresi[2], and Maria Madonia[3](✉)

[1] Dipartimento di Informatica, Università di Salerno, Via Giovanni Paolo II, 132-84084 Fisciano, SA, Italy
anselmo@dia.unisa.it

[2] Dipartimento di Matematica, Università Roma "Tor Vergata", via Della Ricerca Scientifica, 00133 Roma, Italy
giammarr@mat.uniroma2.it

[3] Dipartimento di Matematica e Informatica, Università di Catania, Viale Andrea Doria 6/a, 95125 Catania, Italy
madonia@dmi.unict.it

Abstract. The notion of unbordered picture generalizes to two dimensions the notion of unbordered (or bifix-free) string. We extend to two dimensions Nielsen's construction of unbordered strings ([23]) and describe an algorithm to construct the set $U(m, n)$ of unbordered pictures of fixed size (m, n). The algorithm recursively computes the set of quasi-unbordered pictures $Q(m, n)$, i.e. pictures that can possibly have some "large" borders.

Keywords: Bifix-free strings · Unbordered pictures

1 Introduction

The study of the structure and special patterns of the strings plays an important role in combinatorics of strings, both from theoretical and applicative side. Given a string s, a *bifix* or a *border* of s is a substring x that is both prefix and suffix of s. A string s is *bifix-free* or *unbordered* if it has no other bifixes besides the empty string and s itself.

Bifix-free strings are connected with the theory of codes [9] and are involved in the data structures for pattern matching algorithms [15,19]. From a more applicative point of view, bifix-free strings are suitable as synchronization patterns in digital communications and similar communications protocols [23]. The combinatorial structure of bifix-free strings over a given alphabet was studied by P.T. Nielsen in [23]: he provided an algorithm to enumerate recursively all bifix-free strings of the same length n over a given alphabet. A set of strings X in which no prefix of any string is the suffix of any other string in X is

Partially supported by MIUR Projects *"Formal Languages and Automata: Mathematical Structures and Applicative Directions"* and *"PRISMA PON04a2 A/F"*, and by FARB Projects of University of Catania, Roma "Tor Vergata", Salerno.

A. Maletti (Ed.): CAI 2015, LNCS 9270, pp. 45–57, 2015.
DOI: 10.1007/978-3-319-23021-4_5

called a *cross-bifix-free code*. Constructive methods for cross-bifix-free codes are investigated in [7,10,13].

The increasing interest for pattern recognition and image processing has motivated the research on two-dimensional languages of pictures. A two dimensional string is called *picture* and it is given by a rectangular array of symbols taken from a finite alphabet Σ. The set of all pictures over Σ is usually denoted by Σ^{**}. Extending results from the formal (string) languages theory to two dimensions is a very challenging task. The two-dimensional structure in fact imposes some intrinsic difficulties even in the basic concepts. For example, between two pictures we can define two concatenation operations (horizontal and vertical concatenations) but they are only partial operations and do not induce a monoid structure to the set Σ^{**}. The definition of "prefix" can be extended to a picture by considering its rectangular portion in the top-left corner: nevertheless, if one deletes a prefix from a picture, the remaining part is not a picture anymore.

Several results from string language theory can be worthy extended to pictures. Many researchers have investigated how the notion of recognizability by finite state automata can be transferred to two dimensions to accept picture languages ([2,4–6,11,17,18,20,24,25]). Two dimensional codes were studied in different contexts ([1,8,12,21] and recently two-dimensional prefix codes were introduced as the two-dimensional counterpart of prefix string codes ([3,6]). Matrix periodicity plays a fundamental role in two-dimensional pattern matching (see e.g. [15,22]), while two-dimensional quasi-periodicity was very recently studied in [16].

In this paper we investigate the notion of *unbordered* picture that is somehow connected both to picture codes and to two-dimensional pattern matching. Observe that the notion of border extends very naturally from strings to pictures since it is not related to any scanning direction. Informally we can say that a picture p is *bordered* if a copy p' of p can be overlapped on p by putting a corner of p' somewhere on some position in p. The border of p will be the subpicture corresponding to the portion where p and p' match. The two dimensions of the structure allow several possibilities to specialize this notion. The simplest one is when the matching is checked only by sliding the two picture copies with a horizontal or a vertical move: in this case we allow only borders with the same number of columns or rows of the picture p itself. Notice that this case is not really interesting, since pictures can be handled as they were thick strings on the alphabet either of the columns or of the rows: then the string algorithm by Nielsen can be directly applied to calculate all unbordered pictures. A more intriguing case is taking square pictures and allow only overlaps that put a corner of p' on positions of the diagonals of p. This corresponds to consider only square borders as defined in ([14]). Also in this special case some properties of string borders still hold for pictures.

We consider the more general situation when the overlaps can be made on any position in p and therefore the borders can be of any size. This leads to a different scenario with respect to the string case. It can be proved that if a string

s of lenght n has a border, then it can be written in the form $s = xvx$, i.e. s admits also a "small" border of length less than or equals to $\frac{n}{2}$. Unfortunately this property does not hold in general in two dimensions. Borders of a picture p of size (m, n), can be of three types: borders with dimensions both greater or both smaller than the half of the corresponding dimensions of p (say "large" or "small" borders) and borders with only one dimension greater than the half of the corresponding dimension of p (say a "medium" border). We can only prove that the presence of a "large" border implies also a medium or a small border. For this reason it is not possible to directly generalize Nielsen's construction for unbordered strings to pictures. In this paper we use *quasi-unbordered* pictures as intermediate concept: they can have only certain types of borders that become unlikely when the size of the pictures grows. We describe a recursive procedure to calculate all quasi-unbordered pictures of a given size: the (pure) unbordered pictures can then be easily extracted from this set.

The paper is organized as follows: Section 2 reports the recursive construction of bifix-free strings given by Nielsen in [23], together with all the needed notations and definitions on pictures. In Section 3 the notion of unbordered picture is introduced as two-dimensional extension from the string case. Some related properties are stated together with some examples. Section 4 contains the recursive construction for the set of all unbordered pictures of a given size. Some conclusions together with a table of experimental results are given in Section 5.

2 Preliminaries

In this section we first report the formal definition of unbordered strings together with their recursive construction given by Nielsen. Then, we recall all definitions on pictures needed for the main results of the paper.

2.1 Unbordered Strings and Nielsen's Construction

A string is a sequence of zero or more symbols from an alphabet Σ. A string w of length h is a substring of s if $s = uwv$ for $u, v \in \Sigma^*$. Moreover we say that a string w occurs at position j of s if and only if $w = s_j \dots s_{j+h}$. A string x of length $m < n$ is a *prefix* of s if x is a substring that occurs in s at position 1; a string y is a *suffix* of s if it is a substrings that occurs in s at position $n - m + 1$. A string x that is both prefix and suffix of s is called a *border* or a *bifix* of s. The empty string and s itself are *trivial* borders of s. A string s is *unbordered* or *bifix-free* if it has no borders unless the trivial ones.

Unbordered strings have received very much attention since they occur in many applications as message synchronization or string matching. In [23] P. T. Nielsen proposed a recursive procedure to generate all bifix-free strings of a given length that is based on a property of string borders. We report briefly the main steps that will be used as base for the results of this paper.

The *bifix indicator* h_i of a string s of length n, $1 \le i < n$, is equal to 1 if s has a border of size i, and $h_i = 0$ otherwise. Then the following results hold.

Lemma 1. *A string $s \in \Sigma^*$ is unbordered if and only if $h_i = 0$ for $1 \leq i \leq \lfloor n/2 \rfloor$.*

Saying differently, the previous lemma states that if a string is *not* unbordered, then it must have a "short" border, i.e. of length less than the half of the length of the string. Let $s = s_1 s_2 \dots s_n \in \Sigma^*$ be a unbordered string of even length n, $s_L = s_1 s_2 \dots s_{n/2}$ and $s_R = s_{n/2+1} \dots s_n$. Consider now the strings $s' = s_L a s_R$ and $s'' = s_L ab s_R$, with $a, b \in \Sigma$. Then Lemma 1 is used to prove the following one.

Lemma 2. *The string s is unbordered if and only if s' is unbordered. If s'' is unbordered then s is unbordered. If s is unbordered, then s'' has a border if and only if the following conditions are satisfied: $a = s_n$, $b = s_1$ and $s_2 \dots s_{n/2} = s_{n/2+1} \dots s_{n-1}$ for $n \geq 4$.*

Lemma 2 is then exploited to construct all bifix-free strings of length n from bifix-free strings of shorter length, by inserting extra symbols in the central positions. The starting set of bifix-free strings of length 2 is simply the set of all strings ab with $a, b \in \Sigma$ and $a \neq b$. Remark that Lemmas 1 and 2 and the deriving construction hold for alphabets of any cardinality.

2.2 Basic Notations on Pictures

We recall some definitions about pictures (see [18]). A *picture* over a finite alphabet Σ is a two-dimensional rectangular array of elements of Σ. Given a picture p, $|p|_{row}$ and $|p|_{col}$ denote the number of rows and columns, respectively while $size(p) = (|p|_{row}, |p|_{col})$ denotes the picture *size*. The pictures of size $(m, 0)$ or $(0, n)$ for all $m, n \geq 0$, called *empty* pictures, will be never considered in this paper. The set of all pictures over Σ of fixed size (m, n) is denoted by $\Sigma^{m,n}$, while the set of all pictures over Σ is denoted by Σ^{**}.

Let p be a picture of size (m, n). The set of coordinates $dom(p) = \{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$ is referred to as the *domain* of a picture p. We let $p(i, j)$ denote the symbol in p at coordinates (i, j). We assume the top-left corner of the picture to be at position $(1, 1)$. Moreover, to easily detect border positions of pictures, we use initials of words "top", "bottom", "left" and "right": then, for example, the *tl-corner* of p refers to position $(1, 1)$ while the *br-corner* refers to position (m, n).

A *subdomain* of $dom(p)$ is a set d of the form $\{i, i+1, \dots, i'\} \times \{j, j+1, \dots, j'\}$, where $1 \leq i \leq i' \leq m$, $1 \leq j \leq j' \leq n$, also specified by the pair $[(i, j), (i', j')]$. The portion of p corresponding to positions in subdomain $[(i, j), (i', j')]$ is denoted by $p[(i, j), (i', j')]$. Then a non-empty picture x is *subpicture of* p if $x = p[(i, j), (i', j')]$, for some $1 \leq i \leq i' \leq m$, $1 \leq j \leq j' \leq n$; we say that x *occurs* at position (i, j) (its tl-corner).

Several operations can be defined on pictures (cf. [18]). Let $p, q \in \Sigma^{**}$ be pictures of size (m, n) and (m', n'), respectively, the *column concatenation* of p and q ($p \oplus q$) and the *row concatenation* of p and q ($p \ominus q$) are partial operations, defined only if $m = m'$ and if $n = n'$, respectively, as:

$$p \oplus q = \boxed{\begin{array}{c|c} p & q \end{array}} \qquad\qquad p \ominus q = \boxed{\begin{array}{c} p \\ \hline q \end{array}}$$

The reverse operation on strings can be generalized to pictures and give rise to two different mirror operations (called *row-* and *col-mirror*) obtained by reflecting with respect to a vertical and a horizontal axis, respectively. Another operation that has no counterpart in one dimension is the *rotation*. The rotation of a picture p of size (m, n), is the clockwise rotation of p by $90°$, denoted by $p^{90°}$. Note that $p^{90°}$ has size (n, m). All the operations defined on pictures can be extended in the usual way to sets of pictures.

We conclude by remarking that any string $s = y_1 y_2 \cdots y_n$ can be identified either with a single-row or with a single-column picture, i.e. a picture of size $(1, n)$ or $(n, 1)$. In the sequel it will be used the notation $[y_1 y_2 \cdots y_n]$ to indicate a single-row picture, while a single-column picture will be denoted by $[y_1 y_2 \cdots y_n]^{90°}$.

3 Bordered and Unbordered Pictures

We first generalize the notion of border from strings to pictures. Note that the notions of prefix and suffix of a string implicitly assume the left-to-right reading direction. On the other hand the notion of border is completely independent from any preferred direction. A string has a border when we can find the same substring at the two ends of the string. We extend these concepts to two dimensions.

Informally we say that a picture p is bordered when we can find the same rectangular portion at two opposite corners. Remark that there are two different kinds of borders depending on the pair of opposite corners that hold the border.

More formally we state the following definition.

Definition 3. *Given pictures $p \in \Sigma^{m,n}$ and $x \in \Sigma^{m',n'}$, with $1 \leq m' \leq m$ and $1 \leq n' \leq n$, the picture x is a* tl-border *of p, if x is a subpicture of p occurring at position $(1, 1)$ and at position $(m - m' + 1, n - n' + 1)$; picture x is a* bl-border *of p, if x is a subpicture of p occurring at position $(m - m' + 1, 1)$ and at position $(1, n - n' + 1)$ Moreover x is a* border *of p if it is either a tl- or a bl-border.*

As special cases, p is a *trivial border* of itself, and x is a *proper border* of p if it is not trivial. A tl-border is called a diagonal border in [14]. Notice that a tl-border x of a picture p of size (m, n) can be univocally detected either by giving the position where it occurs in p (besides position $(1, 1)$) or by giving its size. The same holds for bl-borders. Examples of pictures together with their borders are given below.

$$p = \begin{array}{|cccccc|} \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ \hline \end{array} \quad q = \begin{array}{|ccccc|} \hline 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline \end{array} \quad r = \begin{array}{|ccc|} \hline 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ \hline \end{array} \quad s = \begin{array}{|cccc|} \hline 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ \hline \end{array}$$

Note that if a picture p has a tl-border x, then the rotation $p^{90°}$ has a bl-border (that coincides with $x^{90°}$). In the figure above $q = p^{90°}$.

Definition 4. *A picture* $p \in \Sigma^{m,n}$ *is* bordered *if there exists a picture* x *that is a proper border of* p. *Picture* p *is* unbordered *(or* border-free*) if it is not bordered.*

The set of all unbordered pictures of size (m, n) over an alphabet Σ is denoted by $U_\Sigma(m, n)$, or simply $U(m, n)$, when the alphabet can be omitted.

Few simple results can be immediately listed.

Proposition 5. *Let* Σ *be an alphabet. For any* $m, n \geq 1$, *the set* $U_\Sigma(m, n)$ *is closed with respect to the rotation, col- and row-mirror operations, and with respect to permutation or renaming of symbols in* Σ. *Moreover,* $U_\Sigma(m, n)^{90°} = U_\Sigma(n, m)$.

Remark 6. The opposite corners of an unbordered picture p of size (m, n) must contain different symbols otherwise p would have a border of size $(1, 1)$. Moreover the first row (column, resp.) must be different from the last one: otherwise p would have a border of size $(1, n)$ ($(m, 1)$, respectively).

The aim of the rest of the paper will be to construct all the unbordered pictures of a fixed size (m, n). The unbordered pictures of size $(1, n)$ or $(m, 1)$ coincide with the unbordered strings and therefore can be calculate using techniques described in Section 2.1. Before studying the general case let us consider the case of the binary alphabet $\Sigma = \{0, 1\}$ and of pictures of "small" size. It is immediate to see that there are no unbordered pictures of size $(2, 2)$: there is no way to have different opposite corners and different first and last row (see Remark 6). For similar reasons there are no unbordered pictures of sizes $(2, 3), (3, 2)$ and $(3, 3)$. The "smallest" unbordered pictures are of size $(4, 2)$ and are all listed below:

$$\boxed{\begin{matrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{matrix}}, \quad \boxed{\begin{matrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{matrix}}, \quad \boxed{\begin{matrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{matrix}}, \quad \boxed{\begin{matrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{matrix}}.$$

Notice that they can be obtained from the first one by applying mirror operations.

Then the 40 unbordered pictures of size $(4, 3)$ can be obtained by somehow generalizing Nielsen's construction of unbordered strings: it is possible to construct them by inserting a suitable middle column in the unbordered pictures of size $(4, 2)$ listed above. Unfortunately this procedure does not work anymore when the size of pictures grows, as shown by the following example.

Example 7. The picture of size $(5, 4)$ below is unbordered. Nevertheless all the pictures obtained by deleting some columns in the "middle" of the picture (the second column or the third one or both) are all bordered ones. Note that also by deleting the middle (the third) row, one obtains a bordered picture.

$$\boxed{\begin{matrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{matrix}}$$

The main reason why Nielsen's construction of unbordered strings can not be directly generalized to pictures (as in Example 7), is that it is based on Lemma 1, that does not hold in two dimensions. For pictures we have the following weaker result.

Lemma 8. *Let $p \in \Sigma^{m,n}$. If p has a border of size (i,j) with $i \geq \lfloor m/2 \rfloor + 1$ and $j \geq \lfloor n/2 \rfloor + 1$ then p has a border of size (h,k) with $h \leq \lfloor m/2 \rfloor$ or $k \leq \lfloor n/2 \rfloor$.*

Proof. Let b be a border of size (i,j) with $i \geq \lfloor m/2 \rfloor + 1$ and $j \geq \lfloor n/2 \rfloor + 1$. Then p has a border x of size $(h,k) = (2i-m, 2j-n)$. The border x is given by the "intersection" of the two occurrences of the border b in p. More formally, if b is a tl-border then $x = p[(r,s),(r',s')]$ where $[(r,s),(r',s')] = [(1,1),(i,j)] \cap [(m-i+1, n-j+1),(m,n)]$. The case of bl-border is analogous.

Note that $h < i$ and $k < j$. Now, if x is still "large" (i.e. $h \geq \lfloor m/2 \rfloor + 1$ and $k \geq \lfloor n/2 \rfloor + 1$) one can iterate the reasoning until a border, with at least one of the dimension that satisfies the desired inequality, is obtained. $\square$

Informally Lemma 8 claims that if a picture has a "large" border then it necessarily has a "small" or a "middle" border. Indeed, according to its size, a border of a picture p can be of three types: a border with both dimensions greater (smaller, resp.) than the half of the corresponding dimensions of p, say a "large" ("small", resp.) border; or a border with only one dimension greater than the half of the corresponding dimension of p, say a "medium" border. It is the presence of these medium borders that does not allow a simple generalization.

4 Construction of Unbordered Pictures

In this section we present a construction of the class $U(m,n)$ of all unbordered pictures of given size (m,n), that takes inspiration from Nielsen's construction of unbordered strings given in [23] (see Section 2.1). With this aim we introduce the class of *quasi-unbordered* pictures and present its recursive construction. The set $U(m,n)$ will be extracted from the set of quasi-unbordered pictures.

Informally a picture is quasi-unbordered if it has no border occurring in its right side.

Definition 9. *A picture $p \in \Sigma^{m,n}$ is* quasi-unbordered *if p has no border at position (i,j) with $1 \leq i \leq m$ and $\lceil n/2 \rceil + 1 \leq j \leq n$.*

The set of all quasi-unbordered pictures of size (m, n) over an alphabet Σ is denoted by $Q_\Sigma(m, n)$, or simply $Q(m, n)$, when the alphabet can be omitted. Examples of quasi-unbordered pictures can be found in Example 13. Observe that $U(m, n) \subseteq Q(m, n)$.

In the following the set $Q(m, n)$ is constructed in a recursive way by the insertion of one column in the middle of pictures in $Q(m, n-1)$. We introduce first some formal notations. For any picture $p \in \Sigma^{m,n}$, the *left side* of p is the subpicture $p_L = p[(1,1),(m, \lceil n/2 \rceil)]$, containing the first $\lceil n/2 \rceil$ columns of p, and the *right side* of p is the subpicture $p_R = p[(1, \lceil n/2 \rceil + 1),(m, n)]$ containing the remaining columns. Hence $p = p_L \varobar p_R$.

The picture obtained by inserting in the "middle" of p a column $c \in \Sigma^{m,1}$ is denoted $p^{\|c} = p_L \varobar c \varobar p_R$. We also define the inverse operation of removing the central column in a picture. More exactly, if n is odd, then $p^{\nparallel}$ denotes the picture obtained by removing the $\lceil n/2 \rceil$-th column; if n is even, then $p^{\nparallel}$ denotes the picture obtained by removing the $(n/2+1)$-th column.

Let us now focus on quasi-unbordered pictures and show the properties used for their recursive construction.

Proposition 10. *Let $p \in \Sigma^{m,n}$. If p is quasi-unbordered then $p^{\nparallel}$ is quasi-unbordered.*

Proof. Suppose by contradiction that $p^{\nparallel}$ has a tl-border x that occurs at position (i, j) in its right side. It is easy to see that the same tl-border x occurs at position $(i, j+1)$ of p, contradicting the hypothesis that p is quasi-unbordered (note that $\lceil n/2 \rceil + 1 \leq j+1 \leq n$). The case of bl-borders is analogous. $\square$

Proposition 11. *Let p be a quasi-unbordered picture, $p \in Q(m, n)$, and c be a column, $c \in \Sigma^{m,1}$.*

1. If n is even then $p^{\|c} \in Q(m, n+1)$

2. If n is odd then $p^{\|c}$ has a border in its right side if and only if the border occurs at a position in c.

Proof. 1. Arguing by contradiction, suppose that there exist i and j, with $1 \leq i \leq m$ and $\lceil (n+1)/2 \rceil + 1 \leq j \leq n+1$, such that $p^{\|c}$ has a tl-border x that occurs at the position (i, j). It is easy to see that the same tl-border border x occurs at the position $(i, j-1)$ of p contradicting the hypothesis that p is quasi-unbordered (note that $\lceil n/2 \rceil + 1 \leq j-1 \leq n$). The case of bl-borders is analogous.

2. Suppose first that $p^{\|c}$ has a border x in its right side, and suppose w.l.o.g. that x is a tl-border. If x occurs at a position (i, j) not in c, then we can find the same tl-border x at position $(i, j-1)$ of p, that is a position in the right side of p, and this contradicts the assumption p quasi-unbordered. Suppose now that $p^{\|c}$ has a border that occurs at a position in c. Since n is odd, then all the positions of c belong to the right side of $p^{\|c}$ and this concludes the proof. $\square$

Consider now the basis case of the recursion, that is quasi-unbordered pictures with one or two columns. Quasi-unbordered pictures with one column are indeed unbordered strings. Quasi-unbordered pictures with two columns can be characterized in terms of special unbordered strings, that we call *heart-free*. An unbordered string of even length $w \in \Sigma^{2m}$ is *heart-free* if $w = w_1 w_2$, with $|w_1| = |w_2| = m$ and there exists no $x \in \Sigma^*$ that is a suffix of w_1 and a prefix of w_2. In other words both $w_1 w_2$ and $w_2 w_1$ are unbordered.

Proposition 12. *Let $p \in \Sigma^{m,2}$ for some $m \geq 2$, $p = c_1 \oplus c_2$ with $c_1, c_2 \in \Sigma^{m,1}$. Then p is quasi-unbordered if and only if $(c_1 \ominus c_2)^{90^\circ}$ is a heart-free unbordered string.*

Proof. By definition p is quasi-unbordered if p has no border in its right side, i.e. c_2. Then p has a tl-border of size $(i, 1)$ iff $(c_1 \ominus c_2)^{90^\circ}$ is a string with a border of length i; and p has a bl-border of size $(i, 1)$ iff $(c_2 \ominus c_1)^{90^\circ}$ has a border of length i. □

Thanks to Proposition 12, all quasi-unbordered pictures in $\Sigma^{m,2}$, for any $m \geq 2$, can be constructed as follows. Use Nielsen's construction to obtain all unbordered strings over Σ of length $2m-2$. For any unbordered string $w = w_1 w_2$, with $w_1, w_2 \in \Sigma^{m-1}$, insert in the middle only pairs of symbols (a, b) with $a \neq b$, that satisfy the heart-free and unbordered requirements. Then $w_1 a$ and $b w_2$ are the columns of the pictures.

We are now ready to sketch the algorithm that provides the set $Q(m, n)$ of quasi-unbordered pictures of a given size (m, n). It consists in the following two steps.

1. Construct $Q(m, 2)$ (following Proposition 12).
2. Recursively construct $Q(m, n)$ from $Q(m, n-1)$ as follows.

If n is odd then define $Q(m, n)$ as the set of all pictures $p^{\|c}$ for all $p \in Q(m, n-1)$, $c \in \Sigma^{m,1}$.

If n is even then define $Q(m, n)$ as the set of all pictures $p^{\|c}$ for all $p \in Q(m, n-1)$, $c \in \Sigma^{m,1}$, such that $p^{\|c}$ has no border occurring at a position in c.

Let us roughly estimate the complexity of the algorithm. Observe that Step 2 when n is odd requires no comparisons. On the other hand, for $k = 2, \cdots \lfloor m/2 \rfloor$, the pictures in $Q(m, 2k)$ are obtained by inserting in any $p \in Q(m, 2k-1)$ a column $c = [c_m c_{m-1} \ldots c_1]^{90^\circ}$; symbols in c must be taken so that no border occurs at c. First consider tl-borders. To avoid a tl-border of size (i, k), for $i = 1, \cdots, m$, the algorithm does ik comparisons at most. The same number of comparisons is then necessary to avoid also bl- borders at positions in c. Hence the algorithm does $2 \sum_{i=1,\cdots,m} ik \leq 2km^2$ comparisons for any picture in $Q(m, 2k-2)$. The construction of $Q(m, n)$ from $Q(m, 2)$, needs in total a number of comparisons $C(m, n) \leq \sum_{k=1,\cdots,n/2} |Q(m, 2k-2)| 2km^2$.

A simple bound on $|Q(m, n)|$ is $|Q(m, n)| \leq 1/4 |\Sigma^{m,n}|$, for any $m, n \geq 2$, since opposite corners in quasi-unbordered pictures must be different (in an analogous way as for unbordered ones, Remark 6). Applying this bound and some mathematical formulas on summations, one can obtain $C(m, n) \leq \frac{1}{2|\Sigma^{2m}|} m^2 \sum_{k=1,\cdots,n/2} |\Sigma^{2m}|^k k$ and finally $C(m, n) = O(m^2 n |\Sigma|^{mn})$.

Example 13. As an example of the algorithm sketched above, let us show how to obtain some pictures in $Q(3,4)$ for $\Sigma = \{0,1\}$. Note that $|Q(3,4)| = 196$ (see Section 5).

The basis case is the construction of $Q(3,2)$. Applying Proposition 12, take the 20 unbordered binary strings, extract the 6 heart-free unbordered strings and obtain:

$$Q(3,2) = \left\{ \boxed{\begin{array}{cc} 0 & 1 \\ 0 & 0 \\ 0 & 1 \end{array}}, \boxed{\begin{array}{cc} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{array}}, \boxed{\begin{array}{cc} 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{array}}, \boxed{\begin{array}{cc} 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{array}}, \boxed{\begin{array}{cc} 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{array}}, \boxed{\begin{array}{cc} 1 & 0 \\ 0 & 0 \\ 1 & 0 \end{array}} \right\}.$$

Then, using Proposition 11 (case 1) we have:

$$Q(3,3) = \left\{ \begin{array}{|c|c|c|} \hline & & \\ p_L & c & p_R \\ & & \\ \hline \end{array} \text{ with } p = p_L \oplus p_R \in Q(3,2) \text{ and } c \in \Sigma^{3,1} \right\}.$$

Let us give now the construction of some pictures in $Q(3,4)$ from pictures in $Q(3,3)$. Consider for example pictures $p = \boxed{\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{array}}$ and $q = \boxed{\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{array}}$ in $Q(3,3)$ that show a different behavior. From Proposition 11 (case 2), we know that, for any $c \in \Sigma^{3,1}$, $p^{\|c}$ has a border in its right side if and only if the border occurs at a position in c. Observing the picture $p^{\|c}$, one notes that no border can occur at a position in c. Hence, $p^{\|c} \in Q(3,4)$, for any $c \in \Sigma^{3,1}$, i.e. $\boxed{\begin{array}{cccc} 0 & 0 & x & 1 \\ 0 & 0 & y & 0 \\ 0 & 0 & z & 1 \end{array}} \in Q(3,4)$, for any $x, y, z \in \Sigma$.

Consider now q and $q^{\|c} = \boxed{\begin{array}{cccc} 0 & 0 & x & 1 \\ 0 & 0 & y & 0 \\ 0 & 1 & z & 1 \end{array}}$ with $c = \boxed{\begin{array}{c} x \\ y \\ z \end{array}} \in \Sigma^{3,1}$. No tl-border of size $(1,2)$, $(2,2)$, and $(3,2)$ can occur in $q^{\|c}$, for any choice of z, y, x. On the other hand, in order to have no bl-border of size $(1,2)$, necessarily $x = 1$, while y, z can be chosen arbitrarily.

Let us now come back to unbordered pictures. The unbordered pictures of a given size can be obtained from the quasi-unbordered ones of the same size. All pictures in $Q(m,n)$ have no border in their right side. Then the bordered pictures in $Q(m,n)$ to be removed are the ones with a border in their left side. From Lemma 8, it can be argued that it is sufficient to remove pictures with borders of size (i,j), with $i \leq \lceil m/2 \rceil$ and $j > \lfloor n/2 \rfloor$. So only a limited number of comparisons are needed on pictures in the set $Q(m,n)$.

Example 14. (continued) Unbordered pictures in $U(3,4)$ are obtained from pictures in $Q(3,4)$. Consider again pictures p and q in Example 13. We noted that

$p^{\|c} \in Q(3,4)$, for any $c \in \Sigma^{3,1}$, i.e. $\boxed{\begin{array}{cccc} 0 & 0 & x & 1 \\ 0 & 0 & y & 0 \\ 0 & 0 & z & 1 \end{array}} \in Q(3,4)$, for any $x, y, z \in \Sigma$. For $x \neq z$, these pictures in $Q(3,4)$ are unbordered pictures. Moreover, the pictures $\boxed{\begin{array}{cccc} 0 & 0 & 1 & 1 \\ 0 & 0 & y & 0 \\ 0 & 1 & z & 1 \end{array}}$ obtained from q belong to $U(3,4)$ if and only if $z = 0$.

We conclude the section with a simple result that in fact can be obtained as corollary of Proposition 10, but it sheds light on the original motivation to introduce the class of quasi-unbordered pictures, when interested in the construction of unbordered pictures.

Proposition 15. *Let $p \in \Sigma^{m,n}$. If p is unbordered then $p^{\nparallel}$ is quasi-unbordered.*

Moreover it is worthwhile to remark that in the special case of one-row pictures, that are strings, $Q(1,n) = U(1,n)$ for any $n \geq 1$, thanks to Lemma 1. Hence the construction presented here coincides with Nielsen's construction when applied to strings.

5 Final Remarks

We presented general definitions for unbordered pictures by imposing that all possible overlaps between two copies of such pictures are forbidden. This exploits the "bi-dimensionality" of the structures. As a result, the definition imposes many constrains to the pictures. Below we present a table reporting the cardinality of the sets of unbordered and quasi-unbordered pictures over a 2-letters alphabet. The rate with respect to the whole set of pictures of corresponding size is also shown.

Few considerations can be done. First of all, notice that unbordered pictures of size (m,n) are very few with respect to the whole set $\Sigma^{m,n}$ (remember that we had only a rough estimation of $1/4$ given by Remark 6). Regarding the basic step of our recursive construction, Proposition 12 states an interesting bijection between quasi-unbordered pictures with two columns and heart-free unbordered strings. In particular it allows to estimate $|Q(m,2)|$ as the cardinality of heart-free unbordered strings of length $2m$. By some clever considerations on Nielsen's construction it can be observed that the heart-free unbordered strings of given length are at most $1/2$ than the unbordered strings of same size. Moreover denote $v_n = \frac{|U(1,n)|}{|\Sigma^{1,n}|}$ and recall that v_n is a not increasing sequence with $v_4 = \frac{3}{8}$ ([23]). Hence $\frac{|Q(m,2)|}{|\Sigma^{m,2}|} \leq \frac{1}{2} \cdot v_{2m} \leq \frac{1}{2} \cdot \frac{3}{8} = \frac{3}{16}$. This bound is completely reflected in the table.

Finally, observe that the table reports also the rate $\frac{|U(m,n)|}{|Q(m,n)|}$: this is important to estimate the overhead complexity of calculating set $Q(m,n)$ as intermediate step for $U(m,n)$. Notice that, already for those small values of n, the two sets are not so different in size. This can be easily understood if we think that the probability that a picture has a border with more than $n/2$ columns sensibly decreases when n grows.

m	n	$\|U(m,n)\|$	$\|Q(m,n)\|$	$\|\Sigma^{m,n}\|$	$\frac{\|U(m,n)\|}{\|\Sigma^{m,n}\|}$	$\frac{\|Q(m,n)\|}{\|\Sigma^{m,n}\|}$	$\frac{\|U(m,n)\|}{\|Q(m,n)\|}$
2	2	0	2	16	0,00%	12,50%	0,00%
2	3	0	8	64	0,00%	12,50%	0,00%
2	4	4	18	256	1,56%	7,03%	22,2%
2	5	24	72	1024	2,34%	7,03%	33,3%
2	6	120	200	4096	2,93%	4,88%	60,0%
2	7	528	800	16384	3,22%	4,88%	66,0%
2	8	2220	2734	65536	3,39%	4,17%	81,2%
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
3	2	0	6	64	0,00%	9,38%	0,00%
3	3	0	48	512	0,00%	9,38%	0,00%
3	4	40	196	4096	0,98%	4,79%	20,4%
3	5	512	1568	32768	1,56%	4,79%	32,7%
3	6	5048	8542	262144	1,93%	3,26%	59,1%
3	7	44880	68336	2097152	2,14%	3,26%	65,7%
3	8	376768	465266	16777216	2,25%	2,77%	81,0%
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
4	2	4	22	256	1,56%	8,59%	18,2%
4	3	40	352	4096	0,98%	8,59%	11,4%
4	4	864	2720	65536	1,32%	4,15%	31,8%
4	5	16712	42920	1048576	1,59%	4,09%	38,9%
4	6	303976	472990	16777216	1,81%	2,82%	64,3%
4	7	5164176	7567840	268435456	1,92%	2,82%	68,2%
4	8	85346944	103001874	4294967296	1,99%	2,40%	82,9%
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
5	2	24	80	1024	2,34%	7,81%	30,0%
5	3	512	2560	32768	1,56%	7,81%	20,0%
5	4	16712	39646	1048576	1,59%	3,78%	42,2%
5	5	563584	1268672	33554432	1,68%	3,78%	44,4%
5	6	19057664	27609768	1073741824	1,77%	2,57%	69,0%

References

1. Aigrain, P., Beauquier, D.: Polyomino tilings, cellular automata and codicity. Theoretical Computer Science **147**, 165–180 (1995)
2. Anselmo, M., Giammarresi, D., Madonia, M.: Deterministic and unambiguous families within recognizable two-dimensional languages. Fund. Inform. **98**(2–3), 143–166 (2010)
3. Anselmo, M., Giammarresi, D., Madonia, M.: Strong prefix codes of pictures. In: Muntean, T., Poulakis, D., Rolland, R. (eds.) CAI 2013. LNCS, vol. 8080, pp. 47–59. Springer, Heidelberg (2013)

4. Anselmo, M., Giammarresi, D., Madonia, M., Restivo, A.: Unambiguous recognizable two-dimensional languages. RAIRO -ITA **40**(2), 227–294 (2006)
5. Anselmo, M., Giammarresi, D., Madonia, M.: A computational model for tiling recognizable two-dimensional languages. Theor. Comput. Sci. **410**(37), 3520–3529 (2009)
6. Anselmo, M., Giammarresi, D., Madonia, M.: Prefix picture codes: a decidable class of two-dimensional codes. Int. J. Found. Comput. Sci. **25**(8), 1017–1032 (2014)
7. Bajic, D., Loncar-Turukalo, T.: A simple suboptimal construction of cross-bifix-free codes. Cryptography and Communications **6**(1), 27–37 (2014)
8. Beauquier, D., Nivat, M.: A codicity undecidable problem in the plane. Theoret. Comp. Sci **303**, 417–430 (2003)
9. Berstel, J., Perrin, D., Reutenauer, C.: Codes and Automata. Cambridge University Press (2009)
10. Bilotta, S., Pergola, E., Pinzani, R.: A new approach to cross-bifix-free sets. IEEE Transactions on Information Theory **58**(6), 4058–4063 (2012)
11. Blum, M., Hewitt, C.: Automata on a 2-dimensional tape. In: SWAT (FOCS), pp. 155–160 (1967)
12. Bozapalidis, S., Grammatikopoulou, A.: Picture codes. RAIRO - ITA **40**(4), 537–550 (2006)
13. Chee, Y.M., Kiah, H.M., Purkayastha, P., Wang, C.: Cross-bifix-free codes within a constant factor of optimality. IEEE Transactions on Information Theory **59**(7), 4668–4674 (2013)
14. Crochemore, M., Iliopoulos, C.S., Korda, M.: Two-dimensional prefix string matching and covering on square matrices. Algorithmica **20**(4), 353–373 (1998)
15. Crochemore, M., Rytter, W.: Jewels of stringology. World Scientific (2002). http://www-igm.univ-mlv.fr/mac/JOS/JOS.html
16. Gamard, G., Richomme, G.: Coverability in two dimensions. In: Dediu, A.-H., Formenti, E., Martín-Vide, C., Truthe, B. (eds.) LATA 2015. LNCS, vol. 8977, pp. 402–413. Springer, Heidelberg (2015)
17. Giammarresi, D., Restivo, A.: Recognizable picture languages. Int. Journal. Pattern Recognition and Artificial Intelligence **6**(2–3), 241–256 (1992)
18. Giammarresi, D., Restivo, A.: Two-dimensional languages. In: Rozenberg, G.(ed.) Handbook of Formal Languages, vol. III, pp. 215–268. Springer Verlag (1997)
19. Gusfield, D.: Algorithms on Strings, Trees, and Sequences - Computer Science and Computational Biology. Cambridge University Press (1997)
20. Kari, J., Salo, V.: A survey on picture-walking automata. In: Kuich, W., Rahonis, G. (eds.) Algebraic Foundations in Computer Science. LNCS, vol. 7020, pp. 183–213. Springer, Heidelberg (2011)
21. Kolarz, M., Moczurad, W.: Multiset, set and numerically decipherable codes over directed figures. In: Smyth, B. (ed.) IWOCA 2012. LNCS, vol. 7643, pp. 224–235. Springer, Heidelberg (2012)
22. Na, J.C., Ferragina, P., Giancarlo, R., Park, K.: Indexed two-dimensional string matching. In: Encyclopedia of Algorithms (2015)
23. Nielsen, P.T.: A note on bifix-free sequences (corresp.). IEEE Transactions on Information Theory **19**(5), 704–706 (1973)
24. Otto, F., Mráz, F.: Extended two-way ordered restarting automata for picture languages. In: Dediu, A.-H., Martín-Vide, C., Sierra-Rodríguez, J.-L., Truthe, B. (eds.) LATA 2014. LNCS, vol. 8370, pp. 541–552. Springer, Heidelberg (2014)
25. Pradella, M., Cherubini, A., Crespi-Reghizzi, S.: A unifying approach to picture grammars. Inf. Comput. **209**(9), 1246–1267 (2011)