Combinatorics of Finite Words and Suffix Automata*

Gabriele Fici

Dipartimento di Informatica e Applicazioni
Università di Salerno, Italy
fici@dia.unisa.it

Abstract. The suffix automaton of a finite word is the minimal deterministic automaton accepting the language of its suffixes. The states of the suffix automaton are the classes of an equivalence relation defined on the set of factors. We explore the relationship between the combinatorial properties of a finite word and the structural properties of its suffix automaton. We give formulas for expressing the total number of states and the total number of edges of the suffix automaton in terms of special factors of the word.

1 Introduction

The suffix automaton, also called DAWG (Directed Acyclic Word Graph), is a data structure largely used in many text processing problems, like pattern matching and data compression. It is an indexing structure on a text which solves the following problem: given a text (string) X, preprocess it in order to search all the occurrences of a pattern (substring) P in linear time with respect to the size of the pattern. From an algorithmic point of view, the efficiency of the suffix automaton is a consequence of the fact that it can be constructed (on-line) in linear time and space with respect to the size of the input text X [1,2].

We introduce the suffix automaton from an algebraic point of view. The states of the suffix automaton of w are in fact the classes of an equivalence relation defined on the set of factors of w by means of their occurrences in w. The terminal states of the suffix automaton are then the classes containing a suffix of w. The classes of this equivalence have several well known remarkable properties. Indeed, any two classes are either in an inclusion relation or disjoint. Moreover, the total number of classes, i.e. the number of states of the suffix automaton, is smaller than twice the length of w.

Our investigation deals with this question: what are the relations between the structure of the suffix automaton and the combinatorics of the word w?

We show that each class contains as longest element either a prefix of w or a left special factor which is not a prefix. A left special factor is a factor u such that au and bu are both factors of w, for a and b distinct letters.

* Partially supported by the **MIUR** Project *"Aspetti matematici e applicazioni emergenti degli automi e dei linguaggi formali"* (2007).

S. Bozapalidis and G. Rahonis (Eds.): CAI 2009, LNCS 5725, pp. 250–259, 2009.

As a consequence, we can characterize the words having suffix automaton of minimal size as the words such that every left special factor is a prefix.

Then we focus on binary words. In this case we can express the total number of classes in terms of two parameters, H_w and P_w. The first is the length of the shortest prefix of w that has no repetitions in w, while the second is the length of the longest prefix of w which is left special. We then derive a new combinatorial characterization of standard sturmian words in terms of these two parameters.

Finally, we give a formula that expresses the total number of edges of the suffix automaton of a binary word w in terms of its special factors.

The paper is organized as follows. In Section 2 we fix the notation and recall some definitions and basic facts about words. In Section 3 we introduce the suffix automaton and we recall some of its properties. In Section 4 we deal with the size of the suffix automaton and the number of its terminal states. In Section 5 we focus on binary words and we derive the size of the suffix automaton in terms of two combinatorial parameters, H_w and P_w. Finally, we give a formula for the computation of the number of edges of the suffix automaton of a binary word.

2 Notation and Background

An *alphabet*, denoted by A, is a finite set of symbols. The size of A is denoted by $|A|$. A *word* over A is a sequence of symbols from A. The *length* of a word w is denoted by $|w|$. The set of all words over A is denoted by A^*. The empty word has length zero and is denoted by ε. The set of all words over A having length $n \geq 0$ is denoted by A^n. A *language* over A is a subset of A^*. For a finite language L we denote by $|L|$ the number of its elements.

Let $w = a_1 a_2 \dots a_n$, $n > 0$, be a nonempty word over the alphabet A. Any i such that $1 \leq i \leq n$ is called a *position* of w, and the letter $a_i \in A$ is called *the letter in position* i.

The *reversal* of the word $w = a_1 a_2 \dots a_n$ is the word $\widetilde{w} = a_n a_{n-1} \dots a_1$.

A *prefix* of w is any word v such that $v = \varepsilon$ or v is of the form $v = a_1 a_2 \dots a_i$, with $1 \leq i \leq n$. A *suffix* of w is any word v such that $v = \varepsilon$ or v is of the form $v = a_i a_{i+1} \dots a_n$, with $1 \leq i \leq n$. A *factor* (or *substring*) of w is a prefix of a suffix of w (or, equivalently, a suffix of a prefix of w). Therefore, a factor of w is any word v such that $v = \varepsilon$ or v is of the form $v = a_i a_{i+1} \dots a_j$, with $1 \leq i \leq j \leq n$.

We denote by $\mathit{Pref}(w)$, $\mathit{Suff}(w)$, $\mathit{Fact}(w)$ respectively the set of prefixes, suffixes, factors of the word w.

The *factor complexity* of a word w is defined as $p_n(w) = |\mathit{Fact}(w) \cap A^n|$, for every $n \geq 0$. The *maximal factor complexity* of w is defined as $p(w) = \max_{n \geq 0}\{p_n(w)\}$, which represents the maximum number of distinct factors of w having the same length. Note that $p_1(w)$ is the number of distinct letters occurring in w. A *binary word* is a word w such that $p_1(w) = 2$.

A factor u of w is said *left special* if there exist $a, b \in A$, $a \neq b$, such that $au, bu \in \mathit{Fact}(w)$. A factor u of w is said *right special* if there exist $a, b \in A$,

$a \neq b$, such that $ua, ub \in Fact(w)$. A factor u of w is said *bispecial* if it is both left special and right special. We denote by $LS(w)$ (resp. $RS(w)$, $BS(w)$) the set of left special (resp. right special, bispecial) factors of the word w. We note $S_n^l(w)$ (resp. $S_n^r(w)$) the number of left (resp. right) special factors of w which have length n. We note $S^l(w)$ (resp. $S^r(w)$) the total number of left (resp. right) special factors of w, i.e. $S^l(w) = \sum_{n\geq 0} S_n^l(w) = |LS(w)|$ (resp. $S^r(w) = \sum_{n\geq 0} S_n^r(w) = |RS(w)|$). Since a word w has exactly one factor of length zero (the empty word ε), one has $S_0^l(w) = S_0^r(w) = 1$ if and only if $p_1(w) > 1$.

More about combinatorics on words can be found in [6].

3 The Suffix Automaton

Let $w = a_1 a_2 \dots a_n$, $n > 0$, be a nonempty word over the alphabet A. For any $v \in Fact(w)$ we can define the *set of ending positions* of v in w. It is the set $Endset_w(v)$ of the positions of w in which an occurrence of v ends. We assume that $Endset_w(\varepsilon) = \{0, 1, \dots, n\}$.

Example 1. Let $w = aabaab$. Then one has $Endset_w(ba) = \{4\}$, $Endset_w(aab) = Endset_w(ab) = \{3, 6\}$.

In the next proposition we recall some properties of the sets of ending positions (see [3]):

Proposition 1. *[3] Let $u, v \in Fact(w)$. Then one of the three following conditions holds:*

1. $Endset_w(v) \subseteq Endset_w(u)$
2. $Endset_w(u) \subseteq Endset_w(v)$
3. $Endset_w(v) \cap Endset_w(u) = \emptyset$

Moreover, if $u \in Suff(v)$ then $Endset_w(v) \subseteq Endset_w(u)$. If $Endset_w(v) = Endset_w(u)$ then $v \in Suff(u)$ or $u \in Suff(v)$.

On the set $Fact(w)$ we can thus define the following equivalence:

$$u \equiv_w v \quad \Leftrightarrow \quad Endset_w(u) = Endset_w(v).$$

The set $Fact(w)$ is then partitioned into a finite number of classes with respect to this equivalence. These classes are called *right-equivalence classes*.

We note $[u]_w$ (or simply $[u]$, if the context does not make it ambiguous) the right-equivalence class of u in w. So $[u]_w = \{v \in Fact(w) : Endset_w(v) = Endset_w(u)\}$.

In the following proposition we gather some useful facts about the right-equivalence classes, that we will use in next sections.

Proposition 2. *Let $[u]$ be a right-equivalence class of factors of the word w. Then:*

1. *Two distinct elements in* $[u]$ *cannot have the same length. If* v *is the longest element in* $[u]$*, then any other element in* $[u]$ *is a proper suffix of* v*.*
2. *The class* $[u]$ *contains at most one prefix of* w*; this prefix is the longest element in* $[u]$ *and we call* $[u]$ *a* prefix class*.*
3. *If* $v \in [u]$ *is a suffix of* w*, then all the elements in* $[u]$ *are suffixes of* w*. In this case we call* $[u]$ *a* suffix class*.*

We now recall the definition and the basic properties of the suffix automaton (for more details see, for instance, [3]).

Definition 1 ([1,2]). *The* suffix automaton *(or* Direct Acyclic Word Graph*) of a word* $w \in A^*$ *is the minimal deterministic automaton accepting the language Suff*(w)*. It is denoted by* $\mathcal{A}(w)$*.*

The states of $\mathcal{A}(w)$ are in fact the right-equivalence classes of factors of the word w. For each state q of the suffix automaton, the elements of the class $[u]_q$ associated to q are the labeled paths starting at the initial state and ending in q. Hence one can associate to each state q of $\mathcal{A}(w)$ the set of ending positions of factors in $[u]_q$.

There is an edge from the class q to the class q' labeled by the letter $a \in A$ if q' is the class of ua for any u in the class q.

An example of suffix automaton is displayed in Figure 1.

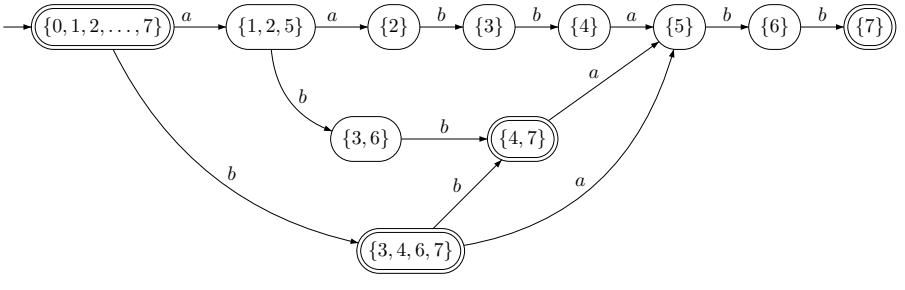

Fig. 1. The suffix automaton of the word $w = aabbabb$. Terminal states are double circled.

The *size* of $\mathcal{A}(w)$, denoted by $|Q_w|$, is the number of its states. Therefore, $|Q_w|$ is the number of right-equivalence classes of factors of w.

The bounds on $|Q_w|$ are well known. The following proposition can be found in [3].

Proposition 3. *[3] Let* w *be a word over* A*. If* $|w| = 0$ *then* $|Q_w| = 1$*; if* $|w| = 1$ *then* $|Q_w| = 2$*. If* $|w| \geq 2$ *then* $1 + |w| \leq |Q_w| \leq 2|w| - 1$*, and the upper bound is reached when* w *has the form* $ab^{|w|-1}$*, for* a *and* b *distinct letters.*

The set of edges of the suffix automaton of the word w is denoted by $\mathcal{E}_w$. In [3] we can find the following bounds.

Proposition 4. *[3] Let w be a word over A. If $|w| = 0$ then $|\mathcal{E}_w| = 0$; if $|w| = 1$ then $|\mathcal{E}_w| = 1$; if $|w| = 2$ then $2 \leq |\mathcal{E}_w| \leq 3$. If $|w| \geq 3$ then $|w| \leq |\mathcal{E}_w| \leq 3|w| - 4$, and the upper bound is reached when w has the form $ab^{|w|-2}c$, for a, b and c pairwise distinct letters.*

4 The Size of the Suffix Automaton

In this section we give some formulas for the computation of the number of states of the suffix automaton of a word w.

Definition 2. *Let w be a word. We denote by $D(w)$ the set of factors u of w such that u is not a prefix of w and u is left special.*

Proposition 5. *Let w be a word. Any $u \in D(w)$ is the longest element in its class $[u]$. In particular, then, the elements of $D(w)$ are each in a distinct class.*

Proof. By contradiction, suppose that there exists a factor v of w such that $v \in [u]$ and $|v| > |u|$. By Proposition 2, u is a proper suffix of v. Let us write $v = zau$, with $z \in A^*$, $a \in A$. Since u and v are in the same class, this implies that every occurrence of u in w is an occurrence of zau, and so u appears in w always preceded by the letter a, against the hypothesis that u is left special.

Since the longest element of a class is unique (by Proposition 2), each $u \in D(w)$ is in a distinct class. □

Proposition 6. *Let w be a word over the alphabet A such that $|w| > 2$. Then the suffix automaton of w has size:*

$$|Q_w| = 1 + |w| + |D(w)|.$$

Proof. From Proposition 2 we know that each right-equivalence class contains at most one prefix of w, so the prefixes of w are each in a distinct class. Since a word w has exactly $|w| + 1$ prefixes the suffix automaton of w has $|w| + 1$ prefix classes. It rests to prove that the number of classes which are not prefix classes is $|D(w)|$.

Let $[u]$ be a class which is not a prefix class, and let u be its (unique) longest element. Thus, by Proposition 2, u is not a prefix of w (and in particular this implies that $|u| > 0$). So there exists a letter $a \in A$ such that au is factor of w. From Proposition 1 we have $Endset_w(au) \subseteq Endset(u)$. Since we supposed that u is the longest element in its class, au cannot belong to $[u]$, and so $Endset(au) \subset Endset(u)$. Hence there exists a position i such that $i \in Endset_w(u)$ but $i \notin Endset_w(au)$. Since u is not a prefix of w this implies that there exists a letter $b \in A$, $b \neq a$, such that $bu \in Fact(w)$, and so u is left special. Thus $u \in D(w)$.

The statement then follows from Proposition 5. □

An interesting consequence of Proposition 6 is the following.

Corollary 1. *Let w be a word over the alphabet A such that $|w| > 2$. Then the suffix automaton of w has minimal number of states $|Q_w| = |w| + 1$ if and only if every left special factor of w is a prefix of w.*

Let us say that a (finite or infinite) word w over A has property LSP if every left special factor of w is a prefix of w.

Sciortino and Zamboni showed in [11] that finite words over a binary alphabet having property LSP are exactly the prefixes of standard sturmian words. Recall that a right infinite binary word w is a *sturmian word* if, for every $n \geq 0$, w has exactly $n+1$ distinct factors of length n. A *standard sturmian word* is a sturmian word having property LSP.

To the best of our knowledge property LSP does not characterize known sets of finite words in the case of larger alphabets.

A right infinite word is an *episturmian word* if it has at most one left special factor (or equivalently right special factor) for each length and the set of its factors is closed under reversal. A *standard episturmian word* is an episturmian word having property LSP.

More generally, if in the definition of episturmian word we substitute the reversal operator with any involutory antimorphism ϑ (i.e. a map $\vartheta : A^* \mapsto A^*$ such that $\vartheta(uv) = \vartheta(v)\vartheta(u)$ and $\vartheta \circ \vartheta = id$), we obtain a *$\vartheta$-episturmian word*. Once again, a *standard ϑ-episturmian word* is a ϑ-episturmian word having property LSP (see [8]).

The word $w = abcaaba$ has property LSP but has two right special factors of the same length (a and b). This implies that w cannot be a factor of a ϑ-episturmian word for any involutory antimorphism ϑ [7]; in particular, it cannot be a factor of an episturmian word.

5 Binary Words

In this section we show that for binary words the number of states of the suffix automaton can be expressed in terms of two combinatorial parameters related to the structure of the word.

We introduce the two parameters H_w and P_w.

Definition 3. *Let w be a word over A.*
We note H_w the minimal length of a prefix of w which occurs only once in w.
We note P_w the maximal length of a prefix of w which is left special.

We now show a property of binary words that underlines the relationship between H_w and the total number $S^l(w)$ of left special factors of w.

The next proposition shows that there is a close relation between the number of left special factors and the factor complexity of w.

Lemma 1. *Let w be a binary word such that $|w| > 2$. Then $S^l_n(w) = p_{n+1}(w) - p_n(w)$ if $0 \leq n < H_w$ and $S^l_n(w) = p_{n+1}(w) - p_n(w) + 1$ if $H_w \leq n \leq |w| - 1$.*

Proof. Let $0 \leq n < H_w$. Among the $p_n(w)$ factors of w of length n there are $S^l_n(w)$ factors that can be extended to the left with two letters, and $p_n(w) - S^l_n(w)$ factors that can be extended to the left with only one letter. If $H_w \leq n \leq |w| - 1$, there is one factor (the prefix of w of length n) that cannot be extended to the left by any letter, since it appears in w only as a prefix.

Thus, the number of factors of w having length $n+1$, that is $p_{n+1}(w)$, is $2S_n^l(w)+p_n(w)-S_n^l(w)$ when $1 \leq n < H_w$, and it is $2S_n^l(w)+p_n(w)-S_n^l(w)-1$ when $H_w \leq n \leq |w|-1$. □

The following lemma gives us a simple formula for the computation of the total number of left special factors of a binary word.

Lemma 2. *Let w be a binary word such that $|w| > 2$. Then the total number of left special factors of w is $S^l(w) = |w| - H_w$.*

Proof. In view of Lemma 1 we have:

$$\begin{aligned} S^l(w) &= \sum_{i=0}^{|w|-1} S_i^l(w) \\ &= \sum_{i=0}^{H_w-1} S_i^l(w) + \sum_{i=H_w}^{|w|-1} S_i^l(w) \\ &= \sum_{i=0}^{H_w-1} (p_{i+1}(w) - p_i(w)) + \sum_{i=H_w}^{|w|-1} (p_{i+1}(w) - p_i(w) + 1) \\ &= \sum_{i=0}^{|w|-1} (p_{i+1}(w) - p_i(w)) + (|w| - 1 - H_w + 1) \\ &= p_{|w|}(w) - p_0(w) + |w| - H_w \\ &= |w| - H_w \end{aligned}$$

□

Analogous results hold for right special factors. If we denote by K_w the minimal length of a suffix of w which occurs only once in w, we get the following result:

Lemma 3. *Let w be a binary word such that $|w| > 2$. Then $S_n^r(w) = p_{n+1}(w) - p_n(w)$ if $0 \leq n < K_w$ and $S_n^r(w) = p_{n+1}(w) - p_n(w) + 1$ if $K_w \leq n \leq |w| - 1$.*

The total number of right special factors of w is $S^r(w) = |w| - K_w$.

Proof. The proof is very similar to that of Proposition 1 and Lemma 2. In fact, one reaches the result by using a symmetric argument in which "right" is replaced by "left", and K_w by H_w. □

The previous technical results (Lemmas 1,2 and 3) are rather easy observations on the factor complexity of finite binary words. For a deep study on the combinatorics of finite words over alphabets of arbitrary size see [4].

For binary words we can then express the number of states of the suffix automaton, $|Q_w|$, in terms of H_w and P_w.

Theorem 1. *Let w be a binary word. Then the number of states of the suffix automaton of w is*

$$|Q_w| = 2|w| - H_w - P_w$$

Proof. Let first $|w| = 2$. Since w is a binary word then either $w = ab$ or $w = ba$ for a and b distinct letters. In both cases we have $H_w = 1$ and $P_w = 0$. Moreover, in both cases $|Q_w| = 3$, so the claim holds.

Let now $|w| > 2$. From Proposition 6 we have $|Q_w| = 1 + |w| + |D(w)|$, where $D(w)$ is the set of left special factors of w that are not prefixes of w.

The total number of left special factors of w is given by the formula $S^l(w) = |w| - H_w$ by Lemma 2. In order to obtain the number of left special factors which are not prefixes, i.e. $|D(w)|$, we have to subtract the number of left special prefixes of w from $S^l(w)$.

If u is the longest prefix of w which is left special, then all the other left special prefixes of w are the prefixes of u. Since, by definition, $|u| = P_w$, we have that the number of left special factors of w which are not prefixes is $|D(w)| = S^l(w) - (|u| + 1) = |w| - H_w - (P_w + 1)$. Hence, the total number of distinct right-equivalence classes of the suffix automaton of w is given by:

$$\begin{aligned} |Q_w| &= 1 + |w| + |D(w)| \\ &= 1 + |w| + |w| - H_w - (P_w + 1) \\ &= 2|w| - H_w - P_w \end{aligned}$$

□

The previous result does not hold for words over an arbitrary alphabet. As an example, consider the word $w = abbccb$. The set of left special factors of w which are not prefixes of w is $D(w) = \{b, c\}$. Hence, by Proposition 6, one has $|Q_w| = 9$. Nevertheless, $H_w = 2$ and $P_w = 0$.

In fact, for words over alphabets larger than two, one has $S^l(w) \leq |w| - H_w$ (see [4]), and so Lemma 2 does not hold in general.

Another formula, involving left special factors, can be derived from the previous theorem and Lemma 2.

Corollary 2. *Let w be a binary word. Then the number of states of the suffix automaton of w is*

$$|Q_w| = |w| + S^l(w) - P_w$$

As a consequence of Theorem 1, we get another characterization of the prefixes of standard sturmian words.

Proposition 7. *Let w be a binary word. Then w is a prefix of a standard sturmian word if and only if $|w| = H_w + P_w + 1$.*

Remark. The previous proposition does not apply to words of the form $w = a^n$, $n \geq 0$. Indeed, such a word belongs to the set of prefixes of standard sturmian words, but $H_w + P_w + 1 = |w| + 1$.

We now deal with the number of edges of the suffix automaton of binary words. The following proposition gives a formula for the computation of the number of edges $|\mathcal{E}_w|$ of the suffix automaton of a binary word.

Proposition 8. *Let w be a binary word. Let $G(w) = (Pref(w) \cap RS(w)) \cup BS(w)$. Then:*

$$|\mathcal{E}_w| = |Q_w| + |G(w)| - 1$$

Proof. Let q be a state of the suffix automaton. If q is the state corresponding to the class of w itself then the outgoing degree of q is 0. Else the outgoing degree of q is either 1 or 2. If it is 2 we call the class corresponding to q a right special class. It is worth noting that all the factors in a right special class are right special factors.

Hence the total number of edges of the suffix automaton of w is $|\mathcal{E}_w| = |Q_w| - 1 + |G'(w)|$, where $G'(w)$ is the set of right special classes. So we have proved the claim once we prove that the sets $G(w)$ and $G'(w)$ are in bijection.

The set of right special classes $G'(w)$ is the union of the set of right special classes which are prefix classes and the set of ones which are not prefix classes.

We know that the longest element of a prefix class is unique and it is a prefix of w.

By Proposition 5 a class which is not a prefix class contains as longest element a left special factor which is not a prefix. So a right special class which is not a prefix class contains as longest element a bispecial factor of w.

Moreover, two different bispecial factors cannot share the same class, since two different left special factors cannot do it (by Proposition 5).

Thus, each right special class contains as longest element an element of $G(w)$ and the elements of $G(w)$ are each in a different class. □

6 Conclusions and Open Problems

This work is an attempt to investigate the combinatorics of a finite word by looking at the structure of its suffix automaton and vice versa.

The characterization of the set of prefixes of standard sturmian words in terms of their suffix automaton given by Sciortino and Zamboni ([11]) does not seem to easy generalize to larger alphabets. A more general question is to characterize the set of words having property LSP, both in the finite and in the infinite case.

Our formulas on the number of states and edges of the suffix automaton can also be used in the study of the average size of the suffix automaton, at least for binary words (this subject is treated for example in [10]). Indeed, both the parameters H_w and P_w are smaller than or equal to the repetition index of w (recall that, for a finite word w, the *repetition index* $r(w)$ is the length of the longest factor of w that has at least two occurrences in w). And it is known that for a word w randomly generated by a memoryless source with identical symbol probabilities, $r(w)$ is logarithmic in the length of w [5,9].

Another direction of research may consist in considering other data structures in place of the suffix automaton (e.g. factor oracles, suffix tries, suffix arrays, etc.). For example, a characterization of the class of words having factor automaton with minimal number of states is still lacking.

References

1. Blumer, A., Blumer, J., Haussler, D., Ehrenfeucht, A., Chen, M.T., Seiferas, J.I.: The smallest automaton recognizing the subwords of a text. Theor. Comput. Sci. 40, 31–55 (1985)
2. Crochemore, M.: Transducers and repetitions. Theor. Comput. Sci. 45(1), 63–86 (1986)
3. Crochemore, M., Hancart, C.: Automata for matching patterns. In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages. Linear Modeling: Background and Application, vol. 2, ch. 9, pp. 399–462. Springer, Berlin (1997)
4. de Luca, A.: On the combinatorics of finite words. Theor. Comput. Sci. 218, 13–39 (1999)
5. Fici, G., Mignosi, F., Restivo, A., Sciortino, M.: Word assembly through minimal forbidden words. Theor. Comput. Sci. 359(1), 214–230 (2006)
6. Lothaire, M.: Algebraic Combinatorics on Words. Encyclopedia of Mathematics and its Applications. Cambridge Univ. Press, New York (2002)
7. De Luca, A.: Private communication (2009)
8. De Luca, A., Bucci, M., de Luca, A., Zamboni, L.Q.: On different generalizations of episturmian words. Theor. Comput. Sci. 393, 23–36 (2008)
9. Mignosi, F., Restivo, A., Sciortino, M.: Forbidden Factors and Fragment Assembly. RAIRO Theoret. Inform. Appl. 35(6), 565–577 (2001)
10. Raffinot, M.: Asymptotic estimation of the average number of terminal states in dawgs. Discrete Appl. Math. 92(2-3), 193–203 (1999)
11. Sciortino, M., Zamboni, L.Q.: Suffix automata and standard sturmian words. In: Harju, T., Karhumäki, J., Lepistö, A. (eds.) DLT 2007. LNCS, vol. 4588, pp. 382–398. Springer, Heidelberg (2007)