On the Number of Closed Factors in a Word

Golnaz Badkobeh[1], Gabriele Fici[2,⋆], Zsuzsanna Lipták[3]

[1] Department of Computer Science, University of Sheffield, UK
g.badkobeh@sheffield.ac.uk
[2] Dipartimento di Matematica e Informatica, Università di Palermo, Italy
fici@math.unipa.it
[3] Dipartimento di Informatica, Università di Verona, Italy
zsuzsanna.liptak@univr.it

Abstract. A closed word (a.k.a. periodic-like word or complete first return) is a word whose longest border does not have internal occurrences, or, equivalently, whose longest repeated prefix is not right special. We investigate the structure of closed factors of words. We show that a word of length n contains at least $n+1$ distinct closed factors, and characterize those words having exactly $n+1$ closed factors. Furthermore, we show that a word of length n can contain $\Theta(n^2)$ many distinct closed factors.

Keywords: Closed word, complete return, rich word, bitonic word.

Introduction

It is known (see for example [8]) that any word w of length n contains at most $n+1$ palindromic factors. Triggered by this result, several researchers initiated a study to characterize words that can accommodate a maximal number of palindromes, called *rich* (or *full*) *words* (see, for example, [2, 10, 3, 4, 12]).

In this paper, we consider the notion of *closed word* (a.k.a. periodic-like word or complete first return). A word w is closed if and only if it is empty or has a factor $v \neq w$ occurring exactly twice in w, as a prefix and as a suffix of w. We also say in this case that w is a complete return to v. For example, *aaa*, *ababa*, *ccabcc* are all closed words (they are complete returns to *aa*, *aba* and *cc*, respectively), while *ab* and *abaabab* are not. As shown in Proposition 1, any word whose exponent is at least two is closed.

The *closed factors* of a word are its factors that are closed words. In contrast to the case of palindromic factors, we show that a word of length n contains at least $n+1$ closed factors (Lemma 3). Inspired by this property, we study the class of words that contain the smallest number of closed factors, and we call them *CR-poor words*.

As an example, *abca* is a CR-poor word, since it has length 4 and exactly 5 closed factors, namely ε, a, b, c and *abca*, whereas the word *ababa* is not CR-poor

⋆ Partially supported by Italian MIUR Project PRIN 2010LYA9RH, "Automi e Linguaggi Formali: Aspetti Matematici e Applicativi".

since it has length 5 but contains 8 closed factors: ε, a, b, aba, bab, $abab$, $baba$ and $ababa$.

However, there is some relation between rich words and CR-poor words. Bucci, de Luca and De Luca [3] showed that a palindromic word is rich if and only if all of its palindromic factors are closed. We show, in Proposition 3, that if a word w has the property that all of its closed factors are palindromes, then w is a CR-poor word, and it is also rich. CR-poor words are also connected to some problems on *privileged words* (see [11]).

While having only palindromic closed factors is a necessary and sufficient condition for a binary word to be CR-poor (Theorem 3), we prove that in a word w over an alphabet Σ of arbitrary cardinality, the set of closed factors and the set of palindromic factors of w coincide if and only if w is both rich and CR-poor (Proposition 5).

In Theorem 2, we give a combinatorial characterization of CR-poor words over an alphabet Σ of cardinality greater than one: A word over Σ is CR-poor if and only if it does not contain any closed factor that is a complete return to xy, for x, y different letters in Σ. In other words, CR-poor words are exactly those words having as their closed factors only complete returns to powers of a single letter. As a consequence, the language of CR-poor words over Σ is a regular language. In contrast, the language of closed words is not regular (Proposition 2).

We give some further characterizations of CR-poor words in the case of the binary alphabet (Theorem 3). One of them is that the binary CR-poor words are the *bitonic words*, i.e., the conjugates to words in a^*b^*. We therefore have that binary CR-poor words form a regular subset of the language of rich words.

Finally, we show that a word of length n can contain $\Theta(n^2)$ many distinct closed factors (Theorem 4).

1 Closed Words

A *word* is a finite sequence of elements from a finite set Σ. We refer to the elements of Σ as *letters* and to Σ as the *alphabet*. The i-th letter of a word w is denoted by w_i. Given a word $w = w_1 w_2 \cdots w_n$, with $w_i \in \Sigma$ for $1 \leq i \leq n$, the nonnegative integer n is the *length* of w, denoted by $|w|$. The empty word has length zero and is denoted by ε. The set of all words over Σ is denoted by Σ^*. Any subset of Σ^* is called a *language*. A language is *regular* (or *rational*) if it can be recognized by a finite state automaton.

A *prefix* (resp. a *suffix*) of a word w is any word u such that $w = uz$ (resp. $w = zu$) for some word z. A *factor* of w is a prefix of a suffix (or, equivalently, a suffix of a prefix) of w. The set of prefixes, suffixes and factors of the word w are denoted by $\mathrm{Pref}(w)$, $\mathrm{Suff}(w)$ and $\mathrm{Fact}(w)$ respectively. A *border* of a word w is any word in $\mathrm{Pref}(w) \cap \mathrm{Suff}(w)$ different from w. From the definitions, we have that ε is a prefix, a suffix, a border and a factor of any word. An *occurrence* of a factor u in w is a factorization $w = vuz$. An occurrence of u is *internal* if both v and z are non-empty.

The word $\tilde{w} = w_n w_{n-1} \cdots w_1$ is called the *reversal* (or *mirror image*) of w. A *palindrome* is a word w such that $\tilde{w} = w$. In particular, the empty word is a palindrome. A *conjugate* of a word w is any word of the form vu such that $uv = w$, for some $u, v \in \Sigma^*$. A conjugate of a word w is also called a *rotation* of w.

A *period* for the word w is a positive integer p, with $0 < p \leq |w|$, such that $w_i = w_{i+p}$ for every $i = 1, \ldots, |w| - p$. Since $|w|$ is always a period for w, we have that every non-empty word has at least one period. We can unambiguously define *the* period of the word w as the smallest of its periods. The *exponent* of a word w is the ratio between its length and its smallest period. A *power* is a word whose exponent is an integer greater than 1. A word that is not a power is called *primitive*

We denote by $\mathrm{PAL}(w)$ the set of factors of w that are palindromes. A word w of length n is *rich* [10] (or *full* [2]) if $|\mathrm{PAL}(w)| = n + 1$, i.e., if it contains the largest number of palindromes a word of length n can contain.

A language L is called *factorial* if $L = \mathrm{Fact}(L)$, i.e., if L contains all the factors of its words. A language L is *extendible* if for every word $w \in L$, there exist letters $a, b \in \Sigma$ such that $awb \in L$. The language of rich words over a fixed alphabet Σ is an example of a factorial and extendible language.

We recall the definition of closed word given in [9]:

Definition 1. *A word w is* closed *if and only if it is empty or has a factor $v \neq w$ occurring exactly twice in w, as a prefix and as a suffix of w.*

The word *aba* is a closed, since its factor a appears in it only as a prefix and as a suffix. The word *abaa*, on the contrary, is not closed. Note that for any letter $a \in \Sigma$ and for any integer $n > 0$, the word a^n is closed, a^{n-1} being a factor occurring only as a prefix and as a suffix in it (this includes the special case of single letters, for which $n = 1$ and $a^{n-1} = \varepsilon$).

Remark 1. The notion of closed word is equivalent to that of *periodic-like* word [6]. A word w is periodic-like if its longest repeated prefix does not have two occurrences in w followed by different letters, i.e., if its longest repeated prefix is not right special.

The notion of closed word is also closely related to the concept of *complete return* to a factor, as considered in [10]. A complete return to the factor u in a word w is any factor of w having exactly two occurrences of u, one as a prefix and one as a suffix. Hence a non-empty word w is closed if and only if it is a complete return to one of its factors; such a factor is clearly both the longest repeated prefix and the longest repeated suffix of w (i.e., the longest border of w).

Remark 2. Let w be a non-empty word over Σ. The following characterizations of closed words follow easily from the definition:

1. w has a factor $v \neq w$ occurring exactly twice in w, as a prefix and as a suffix of w;

2. the longest repeated prefix (resp. suffix) of w does not have internal occurrences in w, i.e., occurs in w only as a prefix and as a suffix;
3. the longest repeated prefix (resp. suffix) of w does not have two occurrences in w followed (resp. preceded) by different letters;
4. w has a border that does not have internal occurrences in w;
5. the longest border of w does not have internal occurrences in w;
6. w is a complete return to its longest repeated prefix;
7. w is a complete return to its longest border.

For more details on closed words and related results see [6, 3, 9, 5, 7, 1, 13]. We end this section by exhibiting some properties of closed words.

Proposition 1. *Any word whose exponent is at least* 2 *is closed.*

Proof. Let $w = v^n v'$ for $n \geq 2$, v a primitive word, and v' a prefix of v such that the exponent of w is equal to $n + |v'|/n$. Then $v^{n-1}v'$ is a border of w. If $v^{n-1}v'$ has an internal occurrence in w, then there exists a proper prefix u of v such that $uv = vu$, and it is a basic result in Combinatorics on Words that two words commute if and only if they are powers of the same word, in contradiction with our hypotheses on u and v. □

Moreover, it is easy to see that for any rational number x between 1 and 2, there exists a closed word having exponent x (it is sufficient to take a word over $\{a, b\}$ ending with b and with only one other occurrence of b, placed in the first half of the word).

Proposition 2. *Let Σ be an alphabet of cardinality $|\Sigma| \geq 2$. The language of closed words over Σ is not regular.*

Proof. Let L be the language of closed words over Σ and let $a, b \in \Sigma$ be different letters. Let us assume that L is regular. This implies that also $L \cap a^*b^*a^*$ is regular, since $a^*b^*a^*$ is a regular language and the intersection of two regular languages is regular. We claim that $L \cap a^*b^*a^* = \{a^n b^m a^n \mid n, m \geq 0\}$, which is not a regular language, and so we have a contradiction.

Clearly, every word in $\{a^n b^m a^n \mid n, m \geq 0\}$ is closed. Suppose now that w belongs to $a^*b^*a^*$. Hence, $w = a^n b^m a^k$, for some $n, m, k \geq 0$. If $n \neq k$, say $n < k$, then the longest repeated prefix of w is a^n and it has at least one internal occurrence in w. By Remark 2, w is not closed. The case $n > k$ is symmetric. □

Finally, we recall two results from [7].

Lemma 1. *[7, Lemma 4] Let w be a non-empty word over Σ. Then there exists at most one letter $x \in \Sigma$ such that wx is closed.*

Lemma 2. *[7, Lemma 5] Let w be a closed word. Then wx, $x \in \Sigma$, is closed if and only if wx has the same period of w.*

2 Closed Factors

Let w be a word. A factor of w that is a closed word is called a *closed factor* of w. The set of closed factors of the word w is denoted by $C(w)$.

Lemma 3. *For any word w of length n, one has $|C(w)| \geq n + 1$.*

Proof. We show that every position of w is the ending position of an occurrence of a distinct closed factor of w. Thus w contains at least n non-empty closed factors, and the claim follows. Indeed, let v be the longest non-empty closed factor ending in position i, so that $w_{i-|v|+1} \cdots w_i = v$. Since a is closed for every $a \in \Sigma$, such a factor always exists. If v did not occur before in w, then we are done. Otherwise, let j be the largest position smaller than i such that $w_{j-|v|+1} \cdots w_j = v$. Set $v' = w_{j-|v|+1} \cdots w_i$ and observe that v' is a closed factor ending in i, with longest border v. But $|v'| > |v|$, in contradiction to the choice of v. □

Lemma 4. *For any words u, v one has $|C(u)| + |C(v)| \leq |C(uv)| + 1$.*

Proof. Clearly, $C(u) \subseteq C(uv)$. In order to prove the statement, it is sufficient to prove that for any non-empty z in $C(v)$, there exists an $f(z)$ in $C(uv) \setminus C(u)$ and f is injective. So let $z \in C(v)$, $uv = w = w_1 \cdots w_n$, and let j be the smallest integer greater than $|u|$ such that $z = w_j \cdots w_{j+|z|-1}$. If j is the smallest integer such that $z = w_j \cdots w_{j+|z|-1}$, then set $f(z) = z$. Otherwise, there is in w a closed z' to z ending in position $w_{j+|z|-1}$. If this is the first occurrence of z' in w, then set $f(z) = z'$, otherwise repeat the construction for z'. Eventually, we will find a closed factor $f(z) = z^{(k)}$ whose first occurrence in w ends in position $w_{j+|z|-1}$.

By construction, f has the desired properties. □

Proposition 3. *Let w be a word of length n. If $C(w) \subseteq \mathrm{PAL}(w)$, then $C(w) = \mathrm{PAL}(w)$ and $|C(w)| = |\mathrm{PAL}(w)| = n + 1$. In particular, w is a rich word.*

Proof. On the one hand, from Lemma 3, one has $|C(w)| \geq n + 1$. On the other hand, one has $|\mathrm{PAL}(w)| \leq n + 1$. Hence, if $C(w) \subseteq \mathrm{PAL}(w)$, then it must be $C(w) = \mathrm{PAL}(w)$ and $|C(w)| = |\mathrm{PAL}(w)| = n + 1$, and so w is a rich word. □

Bucci et al. [3, Proposition 4.3] showed that a word w is rich if and only if every closed factor v of w has the property that the longest palindromic prefix (or suffix) of v is unrepeated in v. Moreover, they proved the following remarkable result:

Theorem 1 (Bucci et al. [3, Corollary 5.2]). *A palindromic word w is rich if and only if $\mathrm{PAL}(w) \subseteq C(w)$.*

In Section 4, we will prove that the condition $\mathrm{PAL}(w) = C(w)$ characterizes the CR-poor words over a binary alphabet.

3 CR-poor Words

By Lemma 3, we have that $n+1$ is a lower bound on the number of closed factors of a word of length n. We introduce the following definition:

Definition 2. *A word $w \in \Sigma^*$ is* CR-poor *if $|C(w)| = |w| + 1$. We also set*

$$\mathcal{L}_\Sigma = \{w \in \Sigma^* : |C(w)| = |w| + 1\}$$

the language of CR-poor words over the alphabet Σ.

Remark 3. If $|\Sigma| = 1$, then $\mathcal{L}_\Sigma = \Sigma^*$. So in what follows we will suppose $|\Sigma| \geq 2$.

Note that, for any alphabet Σ, the language $\mathcal{L}_\Sigma$ of CR-poor words over Σ is closed under reversal. Indeed, it follows from the definition that a word $w \in \Sigma^*$ is closed if and only if its reversal $\tilde{w}$ is closed.

Proposition 4. *The language $\mathcal{L}_\Sigma$ of CR-poor words over Σ is a factorial language.*

Proof. We have to prove that for any word CR-poor w and any factor v of w, v is a CR-poor word. Suppose by contradiction that there exists a CR-poor word w containing a factor v that is not a CR-poor word, i.e., $w \in \mathcal{L}_\Sigma$, $w = uvz$ and $|C(v)| > |v| + 1$. By Lemma 4, $|C(w)| \geq |C(u)| + |C(v)| + |C(z)| - 2 > |u| + |z| + |v| + 1 = |w| + 1$ and therefore w cannot be a CR-poor word. □

The following technical lemma will be used in the proof of the next theorem.

Lemma 5. *Let w be a CR-poor word over the alphabet Σ and $x \in \Sigma$. The word wx (resp. xw) is CR-poor if and only if it has a unique suffix (resp. prefix) that is closed and is not a factor of w.*

Proof. We prove the statement for wx, the one for xw will follow by symmetry. The "if" part is straightforward. For the "only if" part, recall from the proof of Lemma 3 that there is at least one new closed factor ending in every position, so in particular wx has at least one suffix that is closed and is not a factor of w. □

Remark 4. Suppose that a word w contains as a factor a complete return to some word u. Then for every factor u' of u, the word w contains as a factor a complete return to u'.

We now give a characterization of CR-poor words.

Theorem 2. *A word w over Σ is CR-poor if and only if for any two different letters $a, b \in \Sigma$, w does not contain any complete return to ab. In other words,*

$$\mathcal{L}_\Sigma = \Sigma^* \setminus \bigcup_{a \neq b} \Sigma^* ab \Sigma^* ab \Sigma^*.$$

Proof. Let u be a complete return to ab for $a, b \in \Sigma$ different letters. We claim that u is not CR-poor. Since by Proposition 4, a CR-poor word cannot contain a factor that is not CR-poor, once the claim is proved the "only if" part of the theorem follows. So let u' be the longest suffix of u that is closed and starts with the letter b. Such a suffix exists since u contains at least two occurrences of b. Then u' is unioccurrent in u, and since u is a closed suffix of itself we have, by Lemma 5, that u is not CR-poor.

Conversely, suppose that the word w is not CR-poor. Then, analogously as in the proof of Lemma 3, it follows that there is a position i of w such that there are at least two different closed factors u and u' of w that end in position i and do not occur in $w_1 \cdots w_{i-1}$. If both u and u' are complete returns to a power of the letter w_i, then one of them must occur in $w_1 \cdots w_{i-1}$, so this situation is not possible, and we can therefore suppose that there is a factor ending in position i that is a complete return to a word containing at least two different letters. The statement then follows from Remark 4. □

Corollary 1. *A word w over Σ is CR-poor if and only if every closed factor of w is a complete return to a power of a single letter.*

Corollary 2. *The language $\mathcal{L}_\Sigma$ of CR-poor words over Σ is a regular language.*

We can now state the following result:

Proposition 5. *Let w be a word over Σ. Then $C(w) = \mathrm{PAL}(w)$ if and only if w is rich and CR-poor.*

Proof. If $C(w) = \mathrm{PAL}(w)$, then $|C(w)| = |\mathrm{PAL}(w)|$, and since $|C(w)| \geq |w| + 1$ (by Lemma 3) and $|\mathrm{PAL}(w)| \leq |w| + 1$, then it must be $|C(w)| = |\mathrm{PAL}(w)| = |w| + 1$, and hence by definition w is rich and CR-poor.

Conversely, suppose that w is rich and CR-poor. Let $v \in C(w)$. By Corollary 1, v is a complete return to a power of a single letter, so v is a complete return to a palindrome. It is known (see [10, Theorem 2.14]) that a word is rich if and only if all of its factors that are complete returns to a palindrome are palindromes themselves. Therefore, v is a palindrome, and hence we proved that $C(w) \subseteq \mathrm{PAL}(w)$. By Proposition 3, $C(w) = \mathrm{PAL}(w)$ and we are done. □

4 The Case of Binary Words

In this section we fix the alphabet $\Sigma = \{a, b\}$. For simplicity of exposition, we will denote the language of CR-poor words over $\{a, b\}$ by $\mathcal{L}$ rather than by $\mathcal{L}_{\{a,b\}}$. We first recall the definition of bitonic word.

Definition 3. *A word $w \in \{a, b\}^*$ is* bitonic *if it is a conjugate of a word in a^*b^*, i.e., if it is of the form $a^i b^j a^k$ or $b^i a^j b^k$ for some integers $i, j, k \geq 0$.*

By Theorem 2, it is easy to see that a binary word is in $\mathcal{L}$ if and only if it is bitonic.

Lemma 6. *Let w be a bitonic word. Then $C(w) \subseteq \mathrm{PAL}(w)$.*

Proof. Since w is bitonic, a closed factor of w can only be the complete return to a power of a single letter. So a closed factor u of w is of the form $u = a^n$, $u = b^n$, $u = a^n b^m a^n$ or $u = b^n a^m b^n$, for some $n, m > 0$, and these words are all palindromes. □

Thus, by Proposition 3, any bitonic word w of length $n > 0$ contains exactly $n + 1$ closed factors and so is a CR-poor word. We therefore have the following characterizations of CR-poor binary words.

Theorem 3. *Let $w \in \{a, b\}^*$. The following are equivalent:*

1. *$w \in \mathcal{L}$;*
2. *w does not contain any complete return to ab or ba;*
3. *$C(w) \subseteq \mathrm{PAL}(w)$;*
4. *$C(w) = \mathrm{PAL}(w)$;*
5. *w is a bitonic word.*

Notice that the condition $C(w) \subseteq \mathrm{PAL}(w)$ does not hold in general for CR-poor words over alphabets larger than two. As an example, the word *abca* is CR-poor but contains a closed factor (*abca*) that is not a palindrome. In view of Theorem 1, a natural question would be that of establishing whether a palindrome w is CR-poor if and only if $C(w) = \mathrm{PAL}(w)$, i.e., whether the characterization in Theorem 3 can be generalized to larger alphabets at least for palindromes. However, the answer to this question is negative since, for example, the word $w = abcacba$ is a CR-poor palindrome and contains the non-palindromic closed factor *abca*. Note that, coherently with Theorem 1 (and with Proposition 5), w is not rich. However, in the case of a binary alphabet, we have, by Theorem 3 and Proposition 3, that every CR-poor word is rich. Since by Theorem 2 it follows that the language $\mathcal{L}_\Sigma$ is extendible for any alphabet Σ, the language $\mathcal{L}$ is therefore a factorial and extendible subset of the language of (binary) rich words.

In the following proposition we exhibit a closed enumerative formula for the language $\mathcal{L}$.

Proposition 6. *For every $n > 0$, there are exactly $n^2 - n + 2$ distinct words in $\mathcal{L}$.*

Proof. Each of the $n - 1$ words of length $n > 0$ in a^+b^+ has n distinct rotations, while for the words a^n and b^n all the rotations coincide. Thus, there are $n(n - 1) + 2$ bitonic words of length n, and the statement follows from Theorem 3. □

5 How Many Closed Factors Can a Word Contain?

We showed in Lemma 3 that any word of length n contains at least $n+1$ distinct closed factors. But how many closed factors, at most, can a word contain? We provide an answer in the following theorem.

Theorem 4. *For every $n > 4$, there exists a word $w \in \{a, b\}^n$ with quadratically many closed factors.*

Proof. Let $n > 4$ be fixed. We construct a word w of length n such that $|C(w)| \geq (k+1)(k+2)/2$, where $k = \lfloor n/4 \rfloor$.

Let $w = a^k b^k a^k b^k a^{n-4k}$. Clearly $|w| = n$. Let $v_{i,j} = w_i \cdots w_j$, $1 \leq i \leq j \leq n$, be a factor of w. We claim that for every $i = 1, 2, \ldots, k-1$, every factor $v_{i,j}$, with $3k - 1 + i \leq j \leq 4k$, is closed. Indeed, fixed i between 1 and $k-1$, the factor $v_{i,3k-1+i}$, of length $3k$, is equal to $a^{k-i+1}b^k a^k b^{i-1}$, and therefore it is closed since it is a complete return to $a^{k-i+1}b^{i-1}$. Then, for every j such that $3k-1+i \leq j \leq 4k$, the factor $v_{i,j}$ has the same period of $v_{i,3k-1+i}$, and therefore is closed by Lemma 2.

Finally, notice that whenever (i', j') is different from (i, j), for i' and j' in the same range of i and j, respectively (that is, $1 \leq i \leq k-1$ and $3k-1+i \leq j \leq 4k$), the factor $v_{i',j'}$ is different from the factor $v_{i,j}$.

Therefore we conclude that w contains at least $(k+1)(k+2)/2 = \Theta(n^2)$ many different closed factors, and we are done. □

Since a word of length n contains $O(n^2)$ distinct factors, the previous theorem tells us that there exist words in which almost all factors (asymptotically) are closed.

One could also give a formula for the precise value of the maximal number of closed factors in a word of length n, but we think this adds nothing to the general picture provided by Theorem 4. Moreover, the words realizing the upper bound do not have a nice characterization, contrarily to the case of words realizing the lower bound, discussed in the previous sections. However, for completeness, we report in Table 1 the first values of the sequence of the maximum number of closed factors for binary words.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
max	2	3	4	6	8	10	12	15	18	21	25	29	33	37	42	47	52	58	64	70

Table 1. The sequence of the maximum number of closed factors in a binary word.

6 Conclusion and Open Problems

This paper is a first attempt to study the set of closed factors of a finite word. In particular, we investigated the words with the smallest number of closed factors, which we referred to as CR-poor words. We provided a combinatorial characterization of these words and exhibited some relations with rich words.

An enumerative formula for rich words is not known, not even in the binary case. A possible approach to this problem is to separate rich words in subclasses

to be enumerated separately. Our enumerative formula for binary CR-poor words given in Proposition 6 could constitute a step towards this direction.

The set of closed factors could be investigated for specific (finite or infinite) words or classes of words, and could be a tool to derive new combinatorial properties of words.

Finally, the notion of closed factor has recently found applications in string algorithms [1], hence a better understanding of the structure of closed factors of a word could lead to some applications.

References

1. G. Badkobeh, H. Bannai, K. Goto, T. I, C. S. Iliopoulos, S. Inenaga, S. J. Puglisi, and S. Sugimoto. Closed factorization. In *Proceedings of the Prague Stringology Conference 2014*, pages 162–168, 2014.
2. S. Brlek, S. Hamel, M. Nivat, and C. Reutenauer. On the palindromic complexity of infinite words. *Internat. J. Found. Comput. Sci.*, 15:293–306, 2004.
3. M. Bucci, A. de Luca, and A. De Luca. Rich and Periodic-Like Words. In *DLT 2009, 13th International Conference on Developments in Language Theory*, volume 5583 of *Lecture Notes in Comput. Sci.*, pages 145–155. Springer, 2009.
4. M. Bucci, A. De Luca, A. Glen, and L.Q. Zamboni. A new characteristic property of rich words. *Theoretical Computer Science*, 410(30):2860–2863, 2009.
5. M. Bucci, A. De Luca, and G. Fici. Enumeration and Structure of Trapezoidal Words. *Theoretical Computer Science*, 468:12–22, 2013.
6. A. Carpi and A. de Luca. Periodic-like words, periodicity and boxes. *Acta Inform.*, 37:597–618, 2001.
7. A. De Luca and G. Fici. Open and Closed Prefixes of Sturmian Words. In *Proceedings of the 9th International Conference on Words*, volume 8079 of *Lecture Notes in Computer Science*, pages 132–142. Springer Berlin Heidelberg, 2013.
8. X. Droubay, J. Justin, and G. Pirillo. Episturmian words and some constructions of de Luca and Rauzy. *Theoret. Comput. Sci.*, 255(1-2):539–553, 2001.
9. G. Fici. A Classification of Trapezoidal Words. In *WORDS 2011, 8th International Conference on Words*, number 63 in Electronic Proceedings in Theoretical Computer Science, pages 129–137, 2011.
10. A. Glen, J. Justin, S. Widmer, and L. Q. Zamboni. Palindromic richness. *European J. Combin.*, 30:510–531, 2009.
11. J. Peltomäki. Introducing privileged words: Privileged complexity of Sturmian words. *Theoret. Comput. Sci.*, 500:57–67, 2013.
12. A. Restivo and G. Rosone. Burrows–Wheeler transform and palindromic richness. *Theoretical Computer Science*, 410(30):3018–3026, 2009.
13. N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. Available electronically at http://oeis.org. Sequence A226452: Number of closed binary words of length n.