Note

Palindrome positions in ternary square-free words

James D. Currie

Department of Mathematics and Statistics, University of Winnipeg, Canada

Received 4 May 2006; received in revised form 14 December 2006; accepted 16 September 2007

Communicated by M. Crochemore

Abstract

We answer a question of Brešar et al. about the structure of non-repetitive words: For any sequence $A$ of positive integers with large enough gaps, there is a ternary non-repetitive word having a length 3 palindrome starting at each position $a \in A$. In fact, we can find ternary non-repetitive words such that for each $a \in A$, the length 3 subword starting at position $a$ is a palindrome or not, as one chooses. This arbitrariness in the positioning of subwords contrasts markedly with the situation for binary overlap-free words.

Keywords: Non-repetitive sequences; Overlap-free sequences; Palindromes

1. Introduction

Counter-examples in many areas, including formal language theory, logic, partial orders, group theory, algebra, and dynamical systems [4,5,9,1,6,8] have been built from infinite non-repetitive ternary sequences, sometimes with additional structure. Brešar et al. [2], studying repetitiveness in trees, pose the following question:

Is there a set $A$ of positive integers with gaps of size at least $k$, so that whenever $S$ is a non-repetitive sequence over \{1, 2, 3\}, then a length 3 palindrome is guaranteed to occur in $S$ at some position $a \in A$?

A word $w$ is repetitive if it can be factored as $w = uvvz$, where $v$ is a non-empty word. Otherwise, it is non-repetitive. In this note we show that for any set $A$ with large gaps, there are ternary non-repetitive sequences such that, for each $a \in A$, the length 3 subword starting at position $a$ is a palindrome or non-palindrome, just as one chooses. Thus length 3 palindromes may appear or not at any desired locations arbitrarily, as long as the specified locations are not too close together.

The analogous behaviour does not occur for binary overlap-free words: In binary words, the length 2 subwords divide into squares (viz. 00, 11) and non-squares (viz. 01, 10). Let $w$ be a binary overlap-free sequence. It is well known (see [3] for example) that if a length 2 square starts at position $n \geq 3$ in $w$, then the length 2 subword starting at position $m > n$ is a non-square whenever $m \not\equiv n \pmod{2}$. In the analog of Brešar’s question, one may thus choose $A = \{3, 4 + 2[k/2]\}$. Any overlap-free binary sequence contains a non-square starting at some position $a \in A$. 

E-mail address: currie@uwinnipeg.ca.

0304-3975/8 - see front matter © 2008 Published by Elsevier B.V.
doi:10.1016/j.tcs.2007.09.015
We freely use notions of combinatorics on words, as found in [7], for example. For a positive integer \( n \) we say that word \( v \) occurs in position \( n \) in word \( w \) if we can write \( w = uvz \), where \( |u| = n - 1 \). If \( A \) is a set of integers, the gaps of \( A \) are the differences \( \{|a - b| : a, b \in A, a \neq b\} \). A palindrome of length 3 will be called a 3-palindrome; a non-palindrome of length 3 will be called a 3-non-palindrome.

2. Constructing some non-repetitive sequences

Let \( f : \{1, 2, 3\}^* \to \{1, 2, 3\}^* \) be the morphism generated by
\[
\begin{align*}
f(1) &= 123 \\
f(2) &= 13 \\
f(3) &= 2.
\end{align*}
\]
Let \( w = \lim_{n \to \infty} f^n(1) \) be the fixed point of \( f \). It is well known that \( w \) is non-repetitive, and that neither 121 nor 323 is a subword of \( w \). Let \( u = 23213231232 \), \( \bar{u} = 232131232 \). Thus \( u \) is obtained from \( \bar{u} \) by replacing the second occurrence of 3 by 323. Since \( u \) contains 323 as a subword, \( u \) is not a subword of \( w \). On the other hand, \( \bar{u} \) is a subword of \( f^5(1) \), and hence appears in \( w \) infinitely often, with bounded gaps.

**Lemma 2.1.** Let \( v \) be obtained from \( w \) by replacing some occurrences of \( \bar{u} \) with \( u \). Then \( v \) is non-repetitive.

**Proof.** We note first that \( v \) cannot contain the specific repetition 123123. Suppose nevertheless that \( v \) is repetitive, and write \( v = pxxs, x \neq \epsilon \). Let \( m \) occurrences of \( u \) overlap this occurrence of \( xx \) in \( v \). Without loss of generality, we may assume that these are the only occurrences of \( u \) in \( v \). Assume further that \( v \) is chosen to make \( m \) as small as possible. Since \( w \) is non-repetitive, we must have \( m \geq 1 \).

Let \( \hat{x} \) be obtained by replacing each occurrence of 323 in \( x \) by 3. If \( \hat{x} \neq x \), then \( \hat{v} = px\hat{x}x \) contains the non-empty square \( \hat{x}\hat{x} \), but is obtained from \( w \) by replacing at most \( m - 2 \) occurrences of \( \bar{u} \) with \( u \). This contradicts our choice of \( v \). We may assume then that \( \hat{x} = x \), and thus that 323 is not a subword of \( x \).

There is a 323 in the center of \( u \), but 323 is not a subword of \( x \); therefore, a copy of \( u \) overlapping \( xx \) must either (i) overlap the beginning of \( xx \), (ii) span the two copies of \( x \), or (iii) overlap the end of \( xx \):

```
       u
      /   \
     x   x
```
or

```
       u
      /   \
     x   x
```
or

```
  x   x
       u
```

**Claim 2.2.** Either \( x \) starts with 231232 or \( x \) ends in 232132.

**Proof of Claim 2.2.** Suppose that \( u \) overlaps the beginning of \( xx \) in a suffix \( y \). Suppose that \( |y| \leq 5 \). Write \( v = qryzxs \) where \( u = ry, x = yz \).

```
     g  r  y  z
    /   /   \
   p   x   x
```

Every suffix of \( u \) of length 5 or less is also a suffix of \( \bar{u} \). Thus \( y \) is also a suffix of \( \bar{u} \), so that the word \( q\bar{u}zxs \) also contains \( xx \). However, \( q\bar{u}zxs \) only contains \( m - 1 \) occurrences of \( u \). This contradicts our choice of \( v \). We conclude that \( |y| \geq 6 \).
Any suffix of \( u \) of length 7 or more contains 323, which is not in \( x \). It follows that \( |y| = 6 \); thus \( y = 231232 \).

We have shown that the present claim is true in the case where \( u \) overlaps the beginning of \( xx \). Symmetrically, the claim is true if \( u \) overlaps the end of \( xx \). Suppose then, that \( u \) overlaps neither the beginning nor the end of \( xx \), but spans the two copies of \( x \). Write \( x = qr = yz \) where \( ry = u \).

Since neither \( r \) nor \( y \) can contain 323, we must have either

\[
  r = 23213 \quad \text{and} \quad y = 231232
\]

or

\[
  r = 232132 \quad \text{and} \quad y = 31232.
\]

This establishes the claim. \( \square \)

Suppose \( x \) begins with 231232. Word 1x then begins with 123123, which is not a subword of \( v \). We conclude that every occurrence of \( x \) in \( w \) must be preceded by a 3. Since \( x \) precedes \( x \) in \( xx \), word \( x \) must end in a 3. We may therefore write \( x = 231232y3 \), for some word \( y \), and \( v = q3x|x = q3231232y32312323y3s \) where \( p = q3 \). Relabelling with \( P = q \), \( X = 321232y \), \( S = 3s \), we find \( v = PXXS \), but with 323 a subword of \( X \). As we have seen, this contradicts the minimality of \( m \) in our choice of \( v \). A similar contradiction arises if \( x \) ends in 232132. \( \square \)

3. Putting 3-palindromes in specified positions

Fix \( N_0 \) such that every subword of \( w = \lim_{n \to \infty} f^n(1) \) of length at least \( N_0 \) contains seven or more occurrences of \( \bar{u} = 232131232 \). These occurrences will necessarily be mutually non-overlapping.

**Theorem 3.1.** Let \( A = \{a_n\} \) be a sequence of positive integers such that \( a_{n+1} - a_n \geq N_0 \) for each \( n \). Let \( \pi : \mathbb{N} \to \{P, N\} \) be a function. There exists a ternary non-repetitive sequence \( v = \{v_i\}, v_i \in \{1, 2, 3\}, \) such that for each \( m, v_{a_m}v_{a_m+1}v_{a_m+2} \) is a 3-palindrome if and only if \( \pi(m) = P \).

**Proof.** Since \( w \) contains both 3-palindromes and 3-non-palindromes, by deleting a prefix of \( w \) we can obtain a ternary sequence \( u_1 \) which has a 3-palindrome in position \( a_1 \) if and only if \( \pi(1) = P \). We create a sequence of ternary non-repetitive sequences \( u_1, u_2, u_3, \ldots \) such that for each \( n \geq 1 \)

1. each \( u_n \) is obtained by replacing finitely many occurrences of \( \bar{u} \) with \( u \) in \( u_1 \),
2. the only occurrences of subwords 121 or 323 in \( u_n \) are in its prefix of length \( a_n \),
3. for each \( m \leq n, u_n \) contains a 3-palindrome at position \( a_m \) if and only if \( \pi(m) = P \),
4. \( u_n \) and \( u_{n+1} \) are identical on a prefix of length \( a_n \).

Suppose that word \( u_n \) has been obtained. It may already be the case that

\[ u_n \text{ contains a 3-palindrome at position } a_{n+1} \text{ if and only if } \pi(n+1) = P. \]

In this case, we specify \( u_{n+1} = u_n \). Otherwise, write \( u_n = pqr \), where \( |p| = a_n, |q| = a_{n+1} - a_n + 2 \). Note that since \( a_{n+1} - a_n \geq N_0 \), word \( q \) contains at least seven occurrences of \( \bar{u} \).

The following claim is easily verified by hand:

**Claim 3.2.** Let \( q \) be a non-repetitive word over \{1, 2, 3\} of length at least 8, and not containing either of 121 or 323 as a subword.

- There is no \( s \) such that \( q \) has 3-palindromes in positions \( s, s+2 \) and \( s+4 \).
- There is no \( s \) such that \( q \) has 3-non-palindromes in each position \( s+2i, 0 \leq i \leq 7 \).

**Remark 3.3.** The word \( q = 21312313213231 \) has 3-non-palindromes in each position \( 1+2i, 0 \leq i \leq 6 \).
Suppose that $\pi(n + 1) = N$, but that a 3-palindrome occurs at position $a_{n+1}$ in $u_n$. This means that $q$ contains a 3-palindrome in position $a_{n+1} - a_n$. Alter $u_n$ by replacing the first occurrence of $\bar{u}$ in $q$ with $u$. Call the resulting sequence $u'_n$. If $u'_n$ has a 3-non-palindrome in position $a_{n+1}$, let $u_{n+1} = u'_n$. Otherwise, since $|u| = |\bar{u}| + 2$, $q$ contains a 3-palindrome in position $a_{n+1} - a_n - 2$ also. If this is the case, alter $u'_n$ by now replacing the next occurrence of $\bar{u}$ from $q$ with $u$. Call the resulting sequence $u''_n$. This new sequence $u''_n$ will contain a 3-palindrome in position $a_{n+1}$ if and only if $q$ contains a 3-non-palindrome in position $a_{n+1} - a_n - 4$. This would violate our Claim, with $s = a_{n+1} - a_n - 4$. We conclude that $u''_n$ contains a 3-non-palindrome in position $a_{n+1}$, and we let $u_{n+1} = u''_n$.

The analogous construction is undertaken in the case where $\pi(n + 1) = P$, but a 3-non-palindrome occurs at position $a_{n+1}$ in $u_n$. In this case, by the Claim, we find the desired $u_{n+1}$ after at most seven replacements of $\bar{u}$ by $u$.

This completes the construction of the sequences $u_n$, $n \in \mathbb{N}$. Let the prefix of $u_n$ of length $a_n$ be $z_n$. The sequence $v = \lim_{n \to \infty} z_n$ is our desired sequence. □

References