Infinite words without palindrome

Jean Berstel, Luc Boasson, Olivier Carton, Isabelle Fagnot

March 13, 2009

Abstract

We show that there exists an uniformly recurrent infinite word whose set of factors is closed under reversal and which has only finitely many palindromic factors.

1 Notations

For a finite word \( w = w_1 \cdots w_n \), the reversal of \( w \) is the word \( \tilde{w} = w_n \cdots w_1 \). This notation is extended to sets by setting \( F^\sim = \{ \tilde{w} \mid w \in F \} \) for any set \( F \) of finite words. A word \( w \) is a palindrome if \( \tilde{w} = w \). A set \( F \) of finite words is closed under reversal if \( F^\sim = F \).

For an infinite word \( x \), we denote respectively by Fac(\( x \)) and Pal(\( x \)), the set of factors of \( x \) and the set of factors of \( x \) which are palindrome.

An infinite word is uniformly recurrent if each of its factors occurs infinitely many times with bounded gap. Equivalently, \( x \) is uniformly recurrent if for any integer \( m \), there is an integer \( n \) such that any factor of \( x \) of length \( n \) contains all factors of \( x \) of length \( m \).

If \( x \) uniformly recurrent and if Pal(\( x \)) is infinite, the set of factors of \( x \) is closed under reversal, that is Fac(\( x \)) = Fac(\( x \)^\sim). In the following examples, we show that the converse does not hold.

2 Over a 4-letter alphabet

Let \( A \) be the alphabet \( A = \{0, 1, 2, 3\} \). Define by induction the sequence \( (x_n)_{n \geq 0} \) of words over \( A \) by \( x_0 = 01 \) and \( x_{n+1} = x_n23\tilde{x}_n \). The first values are \( x_1 = 012310, x_2 = 01231023013210 \) and \( x_3 = 012310230132102301231032013210 \). We denote by \( x \) the limit of the sequence \( (x_n)_{n \geq 0} \).

We claim that the word \( x \) has the following properties

- \( x \) is uniformly recurrent,
- Fac(\( x \)) is closed under reversal : Fac(\( x \)^\sim) = Fac(\( x \)),
- Pal(\( x \)) is finite : Pal(\( x \)) = A.
It can be easily shown by induction on \( n \) that there is a sequence \((x'_i)_{i \geq 1}\) of words from \(\{23,32\}\) such that

\[
x_{p+n} = x_p x'_1 x_p x'_2 x_p x'_3 x_p \cdots x_p x'_{2n-1} x_p.
\]

Since each factor of \(x\) is factor of \(x_n\) for \(n\) large enough, the word \(x\) is uniformly recurrent. If \(w\) is a factor of \(x_n\), then \(\tilde{w}\) is a factor of \(x_{n+1}\). This shows that \(\text{Fac}(x) = \text{Fac}(x)\). The word \(x\) belongs to \(((01 + 10)(23 + 32))\). Therefore, it has no factor of the form \(aa\) of \(aba\) for \(a, b \in A\). This shows that \(\text{Pal}(x) = A\).

### 3 Over a 2-letter alphabet

Define the morphism \(h\) from \(A^*\) to \(\{0,1\}^*\) as follows

\[
h: \begin{cases} 
0 &\mapsto 101 \\
1 &\mapsto 1001 \\
2 &\mapsto 10001 \\
3 &\mapsto 100001
\end{cases}
\]

Note the image of each letter is a palindrome and \(h(\tilde{w}) = h(w)\). Let \(y\) be the infinite word \(h(x)\). The beginning of \(y\) is the following.

\[y = 101100110001100110110001100001101\cdots\]

We claim that \(y\) has the following properties

- \(y\) is uniformly recurrent,
- \(\text{Fac}(y)\) is closed under reversal : \(\text{Fac}(y)^\sim = \text{Fac}(y)\),
- \(\text{Pal}(y)\) is finite.

Since \(y\) is the image by a morphism of uniformly recurrent word, it is also uniformly recurrent. Since \(h(\tilde{w})\) is the mirror image of \(h(w)\) for each word \(w\), equality \(\text{Fac}(y)^\sim = \text{Fac}(y)\) holds. Each word \(w\) from \(\text{Pal}(y)\) is a factor of a word of the form \(h(aub)\) where \(u\) belongs to \(\text{Pal}(x)\) and \(a, b \in A\). Since \(\text{Pal}(x)\) is finite, \(\text{Pal}(y)\) is also finite.

Define by induction the sequence \((z_n)_{n \geq 0}\) of words over \(\{0,1\}\) by \(z_0 = 01\) and \(z_{n+1} = z_n 01 \tilde{z}_n\). The first values are \(z_1 = 010110\), \(z_2 = 01011001011010\) and \(z_3 = 01011001011010010101101001101010\). We denote by \(z\) the limit of the sequence \((z_n)_{n \geq 0}\). Note that \(z\) is also equal to \(g(x)\) where the morphism \(g\) is given by \(g(0) = g(2) = 0\) and \(g(1) = g(3) = 1\). We claim that the word \(z\) has the following properties

- \(z\) is uniformly recurrent,
- \(\text{Fac}(z)\) is closed under reversal : \(\text{Fac}(z)^\sim = \text{Fac}(z)\),
- \(\text{Pal}(z)\) is finite.
The first two properties are proved as for $x$. We claim that each word $w$ in $\text{Pal}(z)$ satisfies $|w| \leq 12$. Note that it suffices to prove that $\text{Pal}(z)$ contains no word of length 13 or 14. We prove by induction on $n$ that no palindrome of length 13 or 14 occurs in $z_n$. An inspection proves that no palindrome of length 13 or 14 occurs in either $z_3 = z_20\hat{1}\tilde{z}_2$ or in $\tilde{z}_201z_2$. For $n \geq 3$, the word $z_n$ can be factorized $z_n = z_2t_n\tilde{z}_2$ and the word $z_{n+1}$ is equal to $z_2t_n\tilde{z}_201z_2\tilde{t}_n\tilde{z}_2$. Since $z_2$ is of length 14 a palindrome of length 13 or 14 which occurs in $z_{n+1}$ occurs either in $z_n$ or in $\tilde{z}_201z_2$. The result follows from the induction hypothesis.

4 Links with paperfolding

We point out a few links between the words we have introduced and the so-called folding word. For a finite word $w = w_1 \cdots w_n$ over $\{0, 1\}$, denote by $\tilde{w}$ the word $\tilde{w}_n \cdots \tilde{w}_1$ where $\tilde{0} = 1$, $\tilde{1} = 0$.

Define by induction the sequence $(t_n)_{n \geq 0}$ of words over $\{0, 1\}$ by $t_0 = 0$ and $t_{n+1} = t_n0\tilde{t}_n$. We denote by $t$ the limit of the sequence $(t_n)_{n \geq 0}$. The set of factors of $t$ is not closed under reversal. Indeed, the word $01000$ is a factor of $t$ whereas $00010$ is not. The word $y$ is equal to $f(t)$ where the morphism $f$ is given by $f(0) = 01$ and $f(1) = 10$. 

3