A Characterization of Periodicity of Bi-infinite Words

Tero Harju
Arto Lepistö
Dirk Nowotka

Turku Centre of Computer Science (TUCS),
Department of Mathematics,
University of Turku

Turku Centre for Computer Science
TUCS Technical Report No 545
August 2003

ISBN 952-12-1204-7
ISSN 1239-1891

Abstract

A finite word is called bordered if it has a proper prefix which is also a suffix of that word. Costa proves in (Theoret. Comput. Sci., 290(3):2053–2061, 2003) that a bi-infinite word w is of the form ${}^{\omega}fgf^{\omega}$, for some finite words f and g, if, and only if, there is a factorization $w = suv$, with $u \in A^*$ such that every factor $s'uv'$, with $s' \preccurlyeq s$ and $v' \leq v$, is bordered. We present a shorter proof of that result in this paper.

Keywords: combinatorics on words, bi-infinite words, unbordered factors, periodicity

TUCS Laboratory
Discrete Mathematics for Information Technology

We give a shorter proof for a strong result due to J. Costa (see Theorem 1 below). Let us fix our notations first. We refer to [3] for more basic and general definitions on combinatorics of words. We consider a finite alphabet A of letters. Let A^* denote the monoid of all finite words over A including the empty word, denoted by ε. Let ${}^{\omega}A$ and A^{ω} and ${}^{\omega}A^{\omega}$ denote the set of all left- and right- and bi-infinite words over A, respectively. We denote the length of a word w by $|w|$. Suppose $w = uv$, if u is finite, it is called a *prefix* of w, denoted by $u \leq w$, and, if v is finite, it is called a *suffix* of w, denoted by $v \preccurlyeq w$. A nonempty word $u \in A^*$ is a *border* of a word $w \in A^*$, if $w = uv = v'u$ for some suitable nonempty words v and v' in A^*. Note, that every bordered word w has a *shortest border* u such that $w = uvu$, where u is unbordered. Let $\delta(w)$ denote the length of the shortest border of w where $\delta(w) = 0$ if w is unbordered.

Costa proved the following result in [1].

Theorem 1. *Let w be a bi-infinite word. Then there exist $f, g \in A^*$ such that $w = {}^{\omega}fgf^{\omega}$, if, and only if, there is a factorization $w = suv$, with $u \in A^*$ such that every factor $s'uv'$, with $s' \preccurlyeq s$ and $v' \leq v$, is bordered.*

The rest of the paper is devoted to the proof of Theorem 1. The next lemma follows from the critical factorization theorem, cf. [2].

Lemma 2. *Let w be unbordered. There exists a factorization $w = u_0u_1$, where u_0 and u_1 are not empty words, such that for any word x, we have u_ixu_{i+1}, where the indices are modulo 2, is either unbordered or has a minimum border z with $|z| \geq |w|$.*

Such a factorization is called *critical* and $|u_0|$ is called a critical point of w. Let $w = suv$ be a bi-infinite word, where $u \in A^*$, such that $s'uv'$ is bordered for all $s' \preccurlyeq s$ and $v' \leq v$. Clearly, for every finite suffix t of s there exists a minimal m'_t such that $\delta(tuv') \leq m'_t$, for all $v' \leq v$, since $\delta(tuv') \leq |tu|$, for all $v' \leq v$, otherwise there is an unbordered prefix w' of w such that $|w'| > |tu|$ contradicting our assumption on the shape of w. Moreover, there is a maximum integer m_t such that $m_t = \delta(tuv')$ for infinitely many $v' \leq v$. Let $\chi(t)$ denote the prefix of length m_t of tu. Note, that $\chi(t)$ is unbordered.

Lemma 3. *Let $w = suv$ be a bi-infinite word, where $u \in A^*$, such that $s'uv'$ is bordered for all $s' \preccurlyeq s$ and $v' \leq v$. There exists an integer k such that for every suffix t of s longer than k there is a critical point p in $\chi(t)$ with $p \leq |t|$.*

Proof. The case is clear if there are only finitely many suffixes t of s such that $p > |t|$ for all critical points p in $\chi(t)$.

Assume there are infinitely many suffixes t of s such that $p > |t|$ for all critical points p in $\chi(t)$. Surely, there is a prefix u' of u such that there are infinitely many suffixes t of s with $\chi(t) = tu'$ and $p > |t|$ for all critical points p in $\chi(t)$. Then tu' occurs in v inifinitely often. Let $v'tu' \leq v$ denote any such occurence. Now, let t' be a suffix of s such that $\chi(t') = t'u'$ and $|t'| \geq |uv'tu'|$ and $p > |t'|$ for all critical points p in $\chi(t')$. Surely, the shortest border z of $t'uv'tu'$ is shorter than $uv'tu'$, and hence, shorter than t'. However, z is longer than tu' since $t'u'$ is unbordered and $tu' \preccurlyeq t'u'$. So, u' occurs in t', and hence, $t'u'$ has no critical point larger than $|t'|$; a contradiction. $\square$

Lemma 4. *Let $w = suv$ be a bi-infinite word, where $u \in A^*$, such that $s'uv'$ is bordered for all $s' \preccurlyeq s$ and $v' \leq v$. Then $w = {}^{\omega}xyz^{\omega}$.*

Proof. Firstly, we show that there exists a suffix t of s such that $\chi(t)$ has a critical factorization t_0t_1, with $|t_0| \leq |t|$ and $t = t_0\hat{t}$, and $t_0t_1t_0 \leq tu$ and $\chi(\hat{t}) = \chi(t)$. Consider the integer k from Lemma 3.

If $|\chi(t')|$ is bounded for all $t' \preccurlyeq s$, then let t be a suffix of s such that $|t| \geq k$ and $\chi(t)$ is maximal. If $|\chi(t')|$ is not bounded for all $t' \preccurlyeq s$, then let t be the shortest suffix of s such that $\chi(t) \geq k$ and $\chi(t)$ has a critical point p with $p \leq |t|$.

We have that $tu = \chi(t)u'$ and there is a critical factorization $\chi(t) = t_0t_1$ with $|t_0| < |t|$. We have that $v't_0t_1 \leq v$ for infinitely many prefixes v' of v. Let $t = t_0\hat{t}$. We have $\delta(t_1u'v't_0) \geq |t_0t_1|$ for every occurence of t_0t_1 in v by Lemma 2, and hence, $\chi(\hat{t}) \geq \chi(t)$. In fact, $\chi(\hat{t}) = \chi(t)$, by the choice of t, and $tu = t_0t_1t_0\hat{u}$.

Now, since the number of different shortest borders of tuv' for all prefixes v' of v is bounded by $|tu|$, there exists a prefix v_0 of v such that $\delta(tuv_0') \leq |\chi(t)|$ for every $v_0' \leq v$ with $v_0 \leq v_0'$. Note, that $\chi(t) = t_0t_1$ occurs infinitely often in v. Let $v_0 \leq v_0'$ such that $v_0't_0t_1 \leq v$. Consider the shortest border z' of $t_1t_0\hat{u}v_0't_0$, where we have that $|z'| \geq |t_1t_0|$ by Lemma 2. Since $\chi(\hat{t}) = \chi(t)$, there are only finitely many prefixes of the kind of v_0' such that $|z'| > |t_1t_0|$. So, let $v_1 \leq v$ such that $v_0 \leq v_1$ and $z' = t_1t_0$ for every $v_1't_0t_1 \leq v$ for every $v_1 \leq v_1'$. Now, $tuv_1't_0t_1 = t_0t_1t_0\hat{u}v_1''t_1t_0t_1$ for every occurence of t_0t_1 in v right of v_1. By Lemma 2, the shortest border of $t_0t_1t_0\hat{u}v_1''t_1$ is t_0t_1. Now, we have that every of the inifinitely many occurences of t_0t_1 in v right of v_1 is immediately preceded by t_0t_1, and hence, $v = v_1''t_1(t_0t_1)^{\omega}$.

The claim $w = {}^{\omega}xyz^{\omega}$ follows by symmetry. $\square$

Lemma 5. *If $w = {}^{\omega}xyz^{\omega}$, where $xyz \in A^*$, such that $x'yz'$ is bordered for all $x' \preccurlyeq x$ and $z' \leq z$, then $w = {}^{\omega}fgf^{\omega}$, where $fg \in A^*$.*

Proof. Certainly, the word w can be factored into ${}^{\omega}fgf'^{\omega}$ such that every factor containing g is bordered and f and f' are Lyndon words w.r.t. some order $\lhd$ where $a \in A$ is minimal in $\lhd$. Surely, we can assume that a occurs both in f and f' and that $a \leq f$ and $a \leq f'$. Assume that $|f| \leq |f'|$ by symmetry. It is easy to see that every factor of w containing fgf' has to have a shortest border that is not longer than $|f|$. Assume that $f \neq f'$.

If $f \lhd f'$. Then the shortest border of fgf' implies that a prefix f_0 of f is a suffix of f', and $f_0 \lhd f'$ implies that f' is not minimal in $\lhd$; a contradiction.

If $f' \lhd f$. Then $f = f_0cf_1$ and $f' = f_0bf_1'$ for some $b, c \in A$ and $b \neq c$ and $b \lhd c$. It is clear that f_0b does not occur in ff otherwise f is not minimal w.r.t. $\lhd$. Let $f_0'b$ be the longest unbordered suffix of f_0b. Consider the factor $f_0'cf_1fgf'f_0b$ of w with the shortest border s'. If $|s'| \leq |f_0'|$ then $f_0'b$ is bordered; a contradiction. If $s' = f_0'b$ then $b = c$; a contradiction. If $|f_0'b| < |s'| \leq |f_0b|$ then $f_0'b$ is not maximal; a contradiction. If $|f_0b| < |s'|$ then f_0b occurs in ff; a contradiction.

Therefore, $f = f'$ and $w = {}^{\omega}fgf^{\omega}$. □

Proof of Theorem 1. ($\Rightarrow$) Clearly, if $w = {}^{\omega}fgf^{\omega}$ then there is a factorization $w = suv$, with $u \in A^*$ such that every factor $s'uv'$, with $s' \preccurlyeq s$ and $v' \leq v$, is bordered. Take for example $u = fgf$.

($\Leftarrow$) The claim follows from Lemma 4 and 5. □

References

[1] J. C. Costa. Biinfinite words with maximal recurrent unbordered factors. *Theoret. Comput. Sci.*, 290(3):2053–2061, 2003.

[2] T. Harju and D. Nowotka. Periodicity and unbordered words. technical report 523, Turku Centre of Computer Science (TUCS), Turku, Finland, April 2003. submitted.

[3] M. Lothaire. *Algebraic Combinatorics on Words*, volume 90 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, United Kingdom, 2002.

Turku Centre for Computer Science
Lemminkäisenkatu 14
FIN-20520 Turku
Finland

http://www.tucs.fi

University of Turku
- **Department of Information Technology**
- **Department of Mathematics**

Åbo Akademi University
- **Department of Computer Science**
- **Institute for Advanced Management Systems Research**

Turku School of Economics and Business Administration
- **Institute of Information Systems Science**