Corrigendum to "On the theorem of Fredricksen and Maiorana about de Bruijn sequences" [Adv. in Appl. Math. 33 (2) (2004) 413-415]

Eduardo Moreno[a], Dominique Perrin[b]

[a] Faculty of Engineering and Sciences, Universidad Adolfo Ibañez, Santiago, Chile
[b] Université Paris-Est, LIGM, France

Abstract

In (Discrete Math. 23 (1978) 207–210), Fredricksen and Maiorana prove that the concatenation of Lyndon words of length dividing n in lexicographic order produces a de Bruijn sequence of span n, and they state that this word is lexicographically minimal among all de Bruijn sequences of span n. An alternative proof was presented in (Adv. App. Math. 33 (2004) 413–415). The purpose of this corrigendum is twofold. We give a complete proof, clarifying some ambiguities of the previous proof. Additionally, we include a proof of the minimality of the de Bruijn sequence obtained in this way.

Keywords: De Bruijn sequence, Lyndon words, Necklaces
2000 MSC: 05A05

Let A be a finite ordered alphabet and let a, z be respectively the least and the largest letters of A. Denote by σ the shift operator defined on a word $w = a_1 a_2 \cdots a_n$ by $\sigma(w) = a_2 \cdots a_n a_1$. The words $\sigma^i(w)$ for $0 \leq i \leq n-1$ are the *conjugates* of w. A word w of length n is *primitive* if it has n distinct conjugates. A word w of length n is *minimal* if it is lexicographically minimal among its conjugates. It is a *Lyndon word* if it is minimal and primitive. For the rest of the notation, we refer the reader to [4]. The following technical lemma is useful to clarify the proof of the main result in [4].

Lemma 1. *Let w be a prefix of length $n-i$ with $i > 0$ of a minimal word of length n, with $w \neq a^{n-i}$. Let v the smallest minimal word of length n having w as a prefix. Let u be the largest minimal word of length n smaller than v. Then z^i is a suffix of u.*

Proof. Note that the condition $w \neq z^{n-i}$ guarantees the existence of v. Note also that $u_1 \ldots u_{n-i} < w$. Suppose that z^i is not a suffix of u. Hence, $u_1 \ldots u_{n-i} z^i$

Email addresses: `eduardo.moreno@uai.cl` (Eduardo Moreno), `perrin@univ-mlv.fr` (Dominique Perrin)

is a minimal word of length n, larger than u and smaller than v, which is a contradiction. $\square$

The following result is due to Fredericksen and Maiorana [2] (see also the expositions in Knuth's book [3] where de Bruijn cycles are presented in the context of the generation of all n-tuples and in [5]). Let $\mathcal{L}_n$ be the set of Lyndon words of length dividing n.

Theorem 1 ([2]). *For any $n \geq 1$, the lexicographic concatenation of the words of $\mathcal{L}_n$ generates a de Bruijn sequence of span n.*

Proof. Let $\mathcal{B}$ be the sequence obtained by concatenation of the words in $\mathcal{L}_n$ in lexicographic order[1].

We will prove that for any minimal word w of length n, all its conjugates $\sigma^i(w)$, $i = 0 \dots n-1$, are substrings of $\mathcal{B}$. Let $w = w_1 \dots w_j z^{n-j}$ be a minimal word, with $w_j < z$ for some $j \in \{0, \dots n\}$.

First, we prove this for the first j conjugates $\sigma^i(w)$, $i = 0, \dots, j-1$. Note that these words have the form $w_{i+1} \dots w_j z^{n-j} w_1 \dots w_i$.

If w is not a Lyndon word (that is, it is not primitive), let $\hat{w}$ be the primitive root of w and let l be its length. Note that $\hat{w}$ has the form $\hat{w} = w_1 \dots w_{j'} z^{l-j'}$ with $l - j' = n - j$. In this case, we only need to prove that the first $j' - 1$ conjugates appear in $\mathcal{B}$ (the other ones appear among the last $n-j$ conjugates). Since the word x of $\mathcal{L}_n$ next to $\hat{w}$ in lexicographic order[1] has the form $x = \hat{w}^{\frac{n}{l}-1} w_1 \dots w_{j'-1}(w_{j'}+1) b_{j'+1} \dots b_l$, it follows that $\sigma^i(w)$ is a substring of $\hat{w}x$ for $i = 0, \dots, j' - 1$.

If w is primitive, let x be the minimal word of length n (not necessarily primitive), next to w in lexicographic order. Therefore x has the form $w_1 \dots w_{j-1}(w_j + 1) b_{j+1} \dots b_n$ and in this case $\sigma^i(w)$ is a substring of wx for $i = 0, \dots, j-1$. If x is primitive, then wx is a substring of $\mathcal{B}$ and we are done. Otherwise, by the previous argument, x is a prefix of $\hat{x}y$ where $\hat{x}$ if the primitive root of x and y is the word of $\mathcal{L}_n$ next to x in lexicographic order. Hence, wx is a substring of $w\hat{x}y$ and therefore it is a substring of $\mathcal{B}$.

Second, we show that the last $n - j$ conjugate words are substrings of $\mathcal{B}$. Note that these words have the form $z^i w_1 \dots w_j z^{n-j-i}$, for $i = 1, \dots, n-j$. Let v be the first minimal word of length n with prefix $w_1 \dots w_j z^{n-j-i}$. By previous steps, we know that v appears in $\mathcal{B}$. By Lemma 1, the previous minimal word of length n has the form $u = u_1 \dots u_{n-i} z^i$ with $u_1 \dots u_{n-i} < w_1 \dots w_j z^{n-j-i}$. Hence, the largest word in $\mathcal{L}_n$ which is smaller than the root of v has suffix z^i, and thus $z^i w_1 \dots w_j z^{n-j-i}$ is a substring of $\mathcal{B}$. Note that Lemma 1 cannot be applied if $w_1 \dots w_j = a^j$ and $i = n - j$ (that is, $\sigma^i(w) = z^{n-j}a^j$), but it is easy to see that these words and the remaining case $w = z^n$ appear in the concatenation of the end and the beginning of $\mathcal{B}$, which is $z^n a^n$.

Finally, since the length of $\mathcal{B}$ is exactly $\mathrm{Card}(A)^n$, all words of length n appear only once. $\square$

[1] A correction to the proof in [4] is made here

The *cyclic factors* of a word are the factors of its conjugates. A *partial* de Bruijn word of span n is a word of length N which has N distinct cyclic factors and such that its set of cyclic factors is closed under conjugacy. Thus an ordinary de Bruijn word corresponds to the case $N = \mathrm{Card}(A)^n$.

The following result is proved in [4].

Corollary 1. *Let $\ell_1, \ldots, \ell_m$ be the words of $\mathcal{L}_n$ in lexicographic order, then for any $s < m$, $\ell_s \ell_{s+1} \ldots \ell_m$ is a partial de Bruijn sequence of span n.*

For example, for $A = \{a, b\}$ and $n = 4$, we have $m = 6$. Taking $s = 3$, we obtain the partial de Bruijn sequence *aabb ab abbb b* with set of cyclic factors of length 4, the 11 conjugates of *aabb*, *abab*, *abbb* and *bbbb*.

Let us mention another, recently obtained, variant of Theorem 1: the concatenation of the Lyndon words of length n in lexicographic order produces a sequence in which all primitive words of length n appear exactly once as a cyclic factor [1]. For example, the cyclic factors of the word *aaab aabb abbb* are the 12 primitive words of length 4.

Finally, in [2] the lexicographical minimality of the de Bruijn sequence described by Theorem 1 is stated, but not formally proved. We use Corollary 1 to provide a simple proof of this minimality.

Theorem 2. *The de Bruijn sequence obtained by Theorem 1 is lexicographically minimal among all de Bruijn sequences of span n.*

Proof. By contradiction, suppose that there exists another de Bruijn sequence $\hat{\mathcal{B}}$ that is minimal lexicographically. Let wu and wv be the substrings of length n where these two sequences first differ, with $w \in A^{n-1}$ and $u, v \in A$, $u < v$.

Let us denote $M(z)$ be the minimal word of length n conjugate of a given $z \in A^n$. Note first that $M(wu) < M(wv)$. In fact, let k be such that $M(wv) = \sigma^k(wv)$. Then $M(wu) \leq \sigma^k(wu) < \sigma^k(wv) = M(wv)$.

Second, note that wu appears after wv in $\mathcal{B}$. Let us assume that $M(wu)$ and $M(wv)$ are Lyndon words (that is, they are primitive). By Corollary 1, this means that wu appears in the partial de Bruijn sequence constructed from the concatenation of the Lyndon word associated with wv and all the subsequent words of $\mathcal{L}_n$ in lexicographical order, which contradicts the fact that $M(wu) < M(wv)$. Note that the same arguments apply if $M(wu)$ or $M(wv)$ are not primitive, because both words appear in the concatenation of their corresponding Lyndon word and the next word of $\mathcal{L}_n$ in the lexicographical order. $\square$

Acknowledgments. We thank Jeffrey Shallit for suggesting improvements in the proof of the main result of [4].

[1] Yu Hin Au. Shortest sequences containing primitive words and powers. 2013. arXiv:0904.3997.

[2] Harold Fredricksen and James Maiorana. Necklaces of beads in k colors and k-ary de Bruijn sequences. *Discrete Math.*, 23(3):207–210, 1978.

[3] Donald E. Knuth. *The Art of Computer Programming*, volume 4A Combinatorial Algorithms, Part 1. Addison Wesley, 2011.

[4] Eduardo Moreno. On the theorem of Fredricksen and Maiorana about de Bruijn sequences. *Adv. in Appl. Math.*, 33(2):413–415, 2004.

[5] Dominique Perrin and Antonio Restivo. Enumerative combinatorics on words. In Miklos Bona, editor, *Handbook of Enumerative Combinatorics*. CRC Press, 2014.