Available online at www.sciencedirect.com
SCIENCE DIRECT®
Information Processing Letters 96 (2005) 214–219
Information Processing Letters
www.elsevier.com/locate/ipl

De Bruijn sequences and De Bruijn graphs for a general language

Eduardo Moreno [1]

Centro de Modelamiento Matemático, UMR 2071, UCHILE-CNRS, Departamento de Ingeniería Matemática, Universidad de Chile, Casilla 170-3, Correo 3, Santiago, Chile

Received 12 November 2004; received in revised form 10 May 2005; accepted 10 May 2005

Available online 23 September 2005

Communicated by L. Boasson

Abstract

A de Bruijn sequence over a finite alphabet of span n is a cyclic string such that all words of length n appear exactly once as factors of this sequence. We extend this definition to a subset of words of length n, characterizing for which subsets exists a de Bruijn sequence. We also study some symbolic dynamical properties of these subsets extending the definition to a language defined by forbidden factors. For these kinds of languages we present an algorithm to produce a de Bruijn sequence. In this work we use graph-theoretic and combinatorial concepts to prove these results.

Keywords: De Bruijn sequences; De Bruijn graphs; Eulerian labeled graphs; Combinatorics on words; Graph algorithms; Combinatorial problems

1. Introduction

Given a set $\mathcal{D}$ of words of length n, a de Bruijn sequence of span n is a periodic sequence such that every word in $\mathcal{D}$ (and no other n-tuple) occurs exactly once. Its first known description appears as a Sanskrit word *yamátárájabhánasalagám* which was a memory aid for Indian drummers, where the accented/unaccented syllables represent long/shorts beats, so all possible triplets of short and long beats are included in the word. De Bruijn sequences are also known as "shift register sequences" and were originally studied by De Bruijn for $\mathcal{D} = \{0, 1\}^n$ [1]. These sequences have many different applications, such as memory wheels in computers and other technological devices, network models, DNA algorithms, pseudo-random number generation, modern public-key cryptographic schemes, to mention a few (see [2–4]). Typically, de Bruijn sequences have been studied over an arbitrary alphabet A considering the set of all the n-tuples, that is, A^n. There is an exponential number of de Bruijn sequences in this case, but only a few can be generated efficiently.

In this work we generalize the definition of de Bruijn sequence for any set $\mathcal{D}$, characterizing those sets $\mathcal{D}$ for which a de Bruijn sequence exists. In Section 3 we study some symbolic dynamical properties of these sets, extending our results to languages defined by forbidding some factors. Finally, in Section 4 we present an algorithm to produce a de Bruijn sequence for these kinds of languages.

E-mail address: emoreno@dim.uchile.cl (E. Moreno).

[1] Partially supported by ECOS C00E03 (French–Chilean Cooperation), Fundación Andes and CONICYT Ph.D. Fellowship.

0020-0190/$ – see front matter
doi:10.1016/j.ipl.2005.05.028

2. Definitions and generalizations

Let A be a finite set. A *word* w on the alphabet A is a finite sequence of elements of A. For a word w, its length is denoted by $|w|$.

A word p is said to be a *factor* of a word w if there exist words $u, v \in A^*$ such that $w = upv$. If u is the empty word (denoted by ε), then p is called a *prefix* of w, and if v is empty then is called a *suffix* of w.

Let $\mathcal{D}$ be a set of words of length $n+1$. We call this set a *dictionary*. A *de Bruijn sequence of span* $n+1$ *for* $\mathcal{D}$ is a cyclic word $B^{\mathcal{D}}$ of length $|\mathcal{D}|$ such that all the words in $\mathcal{D}$ are factors of $B^{\mathcal{D}}$. In other words,

$$\left\{(B^{\mathcal{D}})_i \dots (B^{\mathcal{D}})_{i+n \bmod |\mathcal{D}|} \mid i = 0, \dots, |\mathcal{D}| - 1\right\} = \mathcal{D}.$$

De Bruijn sequences are closely related to de Bruijn graphs. The *de Bruijn graph of span n for* $\mathcal{D}$, denoted by $G^{\mathcal{D}}$, is the directed graph with vertex set

$$V(G^{\mathcal{D}}) = \{u \in A^n \mid u \text{ is a prefix or a suffix of a word in } \mathcal{D}\}$$

and arc set

$$E(G^{\mathcal{D}}) = \left\{(\alpha v, v\beta) \mid \alpha, \beta \in A, \alpha v\beta \in \mathcal{D}\right\}.$$

This graph was first defined implicitly in 1894 by Flye [5] and it was explicitly detailed in 1946 by de Bruijn [1] and Good [6] independently. In both cases the dictionary studied was $\mathcal{D} = A^{n+1}$. The first use of this graph for a subset of A^{n+1} was given in [7].

From this definition, we can do a bijection between the arcs of $G^{\mathcal{D}}$ and the words in $\mathcal{D}$: to an arc going from αv to $v\beta$ we associate the word $\alpha v\beta$. Using this bijection we can interpret the graph $G^{\mathcal{D}}$ as the union of non-trivial components of the original de Bruijn graph for A^{n+1} after removing the arcs corresponding to words not in $\mathcal{D}$ (see Fig. 1).

We label the graph $G^{\mathcal{D}}$ using the following function l: if $e = (\alpha v, v\beta)$ then $l(e) = \beta$. This labeling has an interesting property:

Remark 1. Let $P = e_0 \dots e_m$ be a walk over $G^{\mathcal{D}}$ of length $m \geqslant n$. Then P finishes in a vertex u if and only if u is a suffix of $l(P) = l(e_0) \dots l(e_m)$.

This property is essential to understand de Bruijn graphs and will be used in all the proofs in this work. Therefore we mention a few important consequences of this property:

Corollary 2. *All the walks of length* $n+1$ *finishing at vertex* u *have label* αu *for some* $\alpha \in A$.

Corollary 3. *If* u *and* v *are vertices of a cycle* C*, then* u *and* v *are factors of the infinite word* $l(C)^\infty$.

These consequences and the bijection between arcs and words in $\mathcal{D}$ explain the relation between de Bruijn graphs and de Bruijn sequences:

Lemma 4. *There exists a de Bruijn sequence* $B^{\mathcal{D}}$ *if and only if* $G^{\mathcal{D}}$ *is an Eulerian graph. Moreover, the labels of Eulerian cycles over* $G^{\mathcal{D}}$ *are the de Bruijn sequences for* $\mathcal{D}$.

Proof. Let C be an Eulerian cycle of $G^{\mathcal{D}}$. As we explained before, any word $w \in \mathcal{D}$ has a corresponding arc e in $G^{\mathcal{D}}$. By Remark 1 any sub-walk of length $n+1$ of C finishing with the arc e has label w, therefore any word in $\mathcal{D}$ is a factor of $l(C)$. As the length of C is the number of words in $\mathcal{D}$ we conclude that $l(C)$ is a de Bruijn sequence for $\mathcal{D}$.

Conversely, let B be a de Bruijn sequence for $\mathcal{D}$. Any factor of length $n+1$ is a word of $\mathcal{D}$ so there is a corresponding arc in $G^{\mathcal{D}}$. Moreover, two consecutive factors αv and $v\beta$ have two corresponding arcs such that the head of the first is the tail of the second one. Therefore B has a corresponding closed walk over $G^{\mathcal{D}}$ with label B. Since every factor is different, every arc in the walk is different, and since every word of $\mathcal{D}$ is a factor of B, every arc of $G^{\mathcal{D}}$ is in the walk. We conclude that the closed walk over $G^{\mathcal{D}}$ is an Eulerian cycle of label B. □

By previous lemma, given a dictionary $\mathcal{D}$, the existence of a de Bruijn sequence of span $n+1$ is characterized by the existence of an Eulerian cycle over $G^{\mathcal{D}}$. A graph has an Eulerian cycle if and only if it is strongly

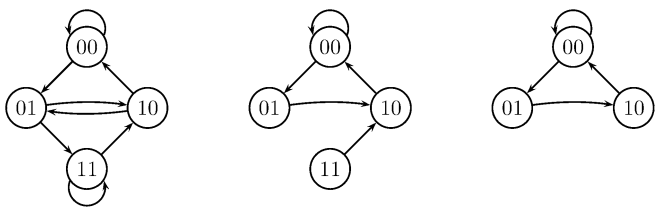

Fig. 1. Examples in a binary alphabet: De Bruijn sequence of span 2, $G^{\mathcal{D}}$ for $\mathcal{D} = \{000, 001, 010, 100, 110\}$ and its essential subgraph.

connected and at each vertex the in-degree and the out-degree are equal. Therefore we can write these conditions as restrictions over $\mathcal{D}$, characterizing the dictionaries with a de Bruijn sequence.

Corollary 5. *A dictionary $\mathcal{D} \subseteq A^{n+1}$ has a de Bruijn sequence of span $n+1$ if and only if*

(1) *For any $u, v \in \mathcal{D}$ there exists a word $w \in A^*$ such that u is a prefix of w, v is a suffix of w and any factor of length $n+1$ of w is in $\mathcal{D}$.*
(2) *For any word $x \in A^n$ there exists a bijection between words in $\mathcal{D}$ having x as a suffix, and words in $\mathcal{D}$ having x as a prefix.*

Proof. By the bijection between arcs and words in $\mathcal{D}$, the first condition assures the existence of a walk (of label w) between any two arcs. Hence $G^{\mathcal{D}}$ is strongly connected. For any word x, a word in $\mathcal{D}$ having x as suffix (prefix) has a corresponding arc terminating (starting) at x. Therefore, the second condition assures that the in-degree and the out-degree at any vertex are equal. □

3. Symbolic dynamics

Symbolic dynamics gives a natural framework to study the sets $\mathcal{D}$ with a de Bruijn sequence.

A first class of dictionaries with a de Bruijn sequence is given by the set of factors of length $n+1$ in a bi-infinite sequence $\mathbf{u}$ over an alphabet A. We denote this set by $\mathcal{L}_{n+1}(\mathbf{u})$.

A factor v of length n in $\mathbf{u}$ is *right extensible* (respectively left extensible) if $v\alpha$ (respectively αv) is in $\mathcal{L}_{n+1}(\mathbf{u})$ for some $\alpha \in A$. These concepts have an important relation with the complexity of the sequence (see [8]).

For any sequence $\mathbf{u}$, is easy to see that the dictionary $\mathcal{D} = \mathcal{L}_{n+1}(\mathbf{u})$ satisfies the first condition of Corollary 5. Also, the second condition is satisfied if and only if the numbers of left and right extensions of any factor of length n are equal. Therefore, we obtain the next theorem.

Theorem 6. *Let $\mathbf{u}$ be a bi-infinite sequence. For any n, the dictionary $\mathcal{D} = \mathcal{L}_{n+1}(\mathbf{u})$ has a de Bruijn sequence if and only if any factor of length n has equal number of left and right extensions.*

Another class of dictionaries with a de Bruijn sequence is given by the language of subshifts. Given an alphabet A, a full shift $A^{\mathbb{Z}}$ is the set of all bi-infinite sequences of symbols from A. Let $\mathcal{F}$ be a collection of (finite) words, we call these words "forbidden words". A shift $X = X_{\mathcal{F}}$ is the subset of sequences of $A^{\mathbb{Z}}$ which do not contain any factor from $\mathcal{F}$. If $\mathcal{F}$ is finite, we say that X is a subshift of finite type.

Let $\mathcal{L}_n(X)$ be the set of factors of sequences in X of length n. The language of a shift X is the set $\mathcal{L}$ of the factors of any finite length of sequences in X.

$$\mathcal{L}(X) = \bigcup_{n=0}^{\infty} \mathcal{L}_n(X).$$

A shift X is irreducible if for every pair of words $u, v \in \mathcal{L}(X)$, there is a $w \in \mathcal{L}(X)$ such that $uwv \in \mathcal{L}(X)$.

Given a labeled graph G, let X_G be the set of labels of all bi-infinite walks over G. It is known that X_G is a (sofic) shift [9], however in the case of de Bruijn graphs, we show that $X_{G^{\mathcal{D}}}$ is a subshift of finite type.

Lemma 7. *Let $\mathcal{D} \subseteq A^{n+1}$ be a dictionary. Then $X_{G^{\mathcal{D}}}$ is a subshift of finite type. Moreover,*

$$X_{G^{\mathcal{D}}} = X_{\mathcal{F}} \quad \text{with } \mathcal{F} = A^{n+1} \setminus \mathcal{D}.$$

Proof. Since $\mathcal{L}_{n+1}(X_{G^{\mathcal{D}}}) \subseteq \mathcal{D}$ we have that $X_{G^{\mathcal{D}}} \subseteq X_{\mathcal{F}}$. Let $x \in X_{\mathcal{F}}$, any factor of length $n+1$ of x is in $\mathcal{D}$ so each factor has a corresponding arc in $G^{\mathcal{D}}$. Moreover, two consecutive factors αv and $v\beta$ of length $n+1$ have two corresponding arcs in $G^{\mathcal{D}}$ such that v is the head of the first and the tail of the second one. Therefore there exists a walk over $G^{\mathcal{D}}$ with label x, so $X_{\mathcal{F}} \subseteq X_{G^{\mathcal{D}}}$. □

Corollary 8. *Let $\mathcal{F}$ be a set of forbidden words of length at most $n+1$. Then for $\mathcal{D} = \mathcal{L}_{n+1}(X_{\mathcal{F}})$ we have that $X_{\mathcal{F}} = X_{G^{\mathcal{D}}}$.*

Proof. We can extend $\mathcal{F}$ to a subset $\mathcal{F}' \subseteq A^{n+1}$ such that $X_{\mathcal{F}} = X_{\mathcal{F}'}$. Since $\mathcal{D} = A^{n+1} \setminus \mathcal{F}'$ we conclude. □

A vertex v is *stranded* if either no arc starts at v or no arc terminates at v. A subgraph is *essential* if no vertex of the graph is stranded (see Fig. 1). Obviously a bi-infinite walk does not use stranded vertices, so for any graph G there exists an essential subgraph G' such that $X_G = X_{G'}$. Therefore in the rest of this work we only consider sets $\mathcal{D}$ such that $G^{\mathcal{D}}$ is essential.

Note that if $G^{\mathcal{D}}$ is essential then for any word $w \in \mathcal{D}$ there exists a walk over $G^{\mathcal{D}}$ with label w.

In order to obtain sets $\mathcal{D}$ with a de Bruijn sequence, $G^{\mathcal{D}}$ needs to be an Eulerian graph, in particular it needs to be strongly connected. This property has an interpretation in symbolic dynamics:

Lemma 9. *Let $\mathcal{D}$ be a dictionary. $X_{G^{\mathcal{D}}}$ is irreducible if and only if $G^{\mathcal{D}}$ is strongly connected.*

Proof. Since $G^{\mathcal{D}}$ is essential, the strongly connected components have size at least 2. Let $X_{G^{\mathcal{D}}}$ be an irreducible subshift of finite type. For any two arcs e, f of $G^{\mathcal{D}}$ there are two corresponding words w_e, w_f in $\mathcal{D}$. Since $X_{G^{\mathcal{D}}}$ is irreducible, there exists a word $\hat{w}$ such that $w_e\hat{w}w_f$ is a factor of $X_{G^{\mathcal{D}}}$. In other words, there exists a walk over $G^{\mathcal{D}}$ with label $w_e\hat{w}w_f$. Therefore there exists a walk with label $\hat{w}w_f$ connecting e to f, so $G^{\mathcal{D}}$ is strongly connected.

Conversely, if $X_{G^{\mathcal{D}}}$ is not irreducible, there exist factors w_1, w_2 such that $\forall z \in A^*$, w_1zw_2 is not a factor of $X_{G^{\mathcal{D}}}$. But w_1 is the label of a walk over $G^{\mathcal{D}}$ finishing at a vertex v_1 and w_2 is the label of a walk starting at a vertex v_2, therefore there is no walk over $G^{\mathcal{D}}$ connecting v_1 to v_2, hence $G^{\mathcal{D}}$ is not strongly connected. □

Let $X_{\mathcal{F}}$ be an irreducible subshift of finite type. If $\mathcal{D} = \mathcal{L}_{n+1}(X_{\mathcal{F}})$ then the corresponding graph $G^{\mathcal{D}}$ is not necessarily an Eulerian graph.

For example, for $A = \{0, 1\}$ and $\mathcal{F} = \{11\}$ a vertex $0w1$ has two in-going arcs (corresponding to words $00w1$ and $10w1$) but only one out-going arc (corresponding to the word $0w10$). Therefore we will study the subset of *periodic* words in $\mathcal{L}_{n+1}(X_{\mathcal{F}})$ because for this set we obtain an Eulerian de Bruijn graph.

Let $w \in A^*$ be a word, we say that w is a *periodic* word of $X_{\mathcal{F}}$ if and only if the bi-infinite sequence w^{∞}, obtained by infinite concatenations of w, is in $X_{\mathcal{F}}$. The set of periodic words of length n is denoted by $\mathcal{P}_n(X_{\mathcal{F}})$.

Theorem 10. *Let $\mathcal{F}$ be a set of forbidden words of length at most $n+1$ and $\mathcal{D} = \mathcal{P}_{n+1}(X_{\mathcal{F}})$. If $X_{G^{\mathcal{D}}}$ is irreducible then there exists a de Bruijn sequence for the dictionary $\mathcal{D}$.*

Proof. By Lemma 9, $G^{\mathcal{D}}$ is strongly connected. Let $u \in A^n$ be a vertex of $G^{\mathcal{D}}$. Any arc leaving u with label α corresponds to a word $u\alpha \in \mathcal{D}$. Since $u\alpha$ is a periodic word, αu is also in $\mathcal{D}$. Therefore there exists an arc going into u corresponding to the word αu, which implies that the in-degree of u is greater or equal to out-degree of u. The same argument proves that the out-degree of u is greater than or equal to the in-degree of u, concluding that $G^{\mathcal{D}}$ is an Eulerian graph. □

Note that not all irreducible subshifts of finite type have a de Bruijn sequence for $\mathcal{D} = \mathcal{P}_{n+1}(X_{\mathcal{F}})$. For example, for $A = \{0, 1\}$ and $\mathcal{F} = \{010\}$ the subshift of finite type $X_{\mathcal{F}}$ is irreducible but $X_{G^{\mathcal{D}}}$ is not irreducible, because $G^{\mathcal{D}}$ has two strongly connected components.

4. Constructing a de Bruijn sequence for subshifts

Let $X_{\mathcal{F}}$ be a subshift of finite type and $\mathcal{D} = \mathcal{P}_{n+1}(X_{\mathcal{F}})$ such that $X_{G^{\mathcal{D}}}$ is irreducible. In this section we study an efficient generation of a de Bruijn sequence for $\mathcal{D}$.

Even in the unrestricted case (where $\mathcal{F} = \emptyset$) this is an interesting problem (see [10] for a survey on this subject). One of the most elegant and efficient solutions in the unrestricted case is given in [11] and uses *Lyndon words*.

Let $<$ be a linear order over alphabet A. The set A^* of all words on the alphabet A is linearly ordered by the lexicographical order induced by the order $<$ on A. A word w is a Lyndon word if and only if $\forall u, v$ such that $w = uv$, then $w < vu$.

The algorithm of Fredricksen and Maiorana consists of to concatenate in increasing lexicographical order the Lyndon words of length dividing n. This is a linear time algorithm because the Lyndon words can be generated efficiently (see [12]).

We always can construct the graph $G^{\mathcal{D}}$ and apply one of the known results about constructing an Eulerian cycle to obtain a de Bruijn sequence, however the construction of $G^{\mathcal{D}}$ is not efficient. Therefore in this section we study the structure of $G^{\mathcal{D}}$ in order to obtain an algorithm to construct a de Bruijn sequence only using the words in $\mathcal{D}$.

The set of arcs of an Eulerian graph can be partitioned in cycles. In the particular case of $G^{\mathcal{D}}$ these cycles have a given length.

Theorem 11. *Let $\mathcal{F}$ be a set of forbidden words of length at most $n+1$ and $\mathcal{D} = \mathcal{P}_{n+1}(X_{\mathcal{F}})$ such that $G^{\mathcal{D}}$ is the de Bruijn graph of span n for $\mathcal{D}$. Then the cycles of length dividing $n+1$ partition the set of arcs of $G^{\mathcal{D}}$.*

Proof. We prove that any arc of the graph is in one and only one cycle of length dividing n.

Let e be an arc from the vertex au to the vertex ub with $a, b \in A$ (then, $l(e) = b$). By construction of the graph, there is a walk of length n from vertex ub to vertex au with label au. Therefore, the union of this walk with the arc e produces a closed walk of length $n+1$ with label aub corresponding to one or more repetitions of a cycle of length dividing $n+1$, proving the existence of one cycle.

Let us suppose now that there are two cycles C and C' of lengths dividing $n+1$ using the arc e. Let f be an arc of C and g an arc of C' with tail at the same vertex u and different heads. Since e is in both cycles, by Corollary 2 the walks of length n from the head of e to the tail of e using only the arcs of C and C' must

INPUT: $L = \{L^1, \ldots, L^k\}$ Lyndon words in $\mathcal{L}(X_{\mathcal{F}})$ of length dividing $n+1$.
(1) $Size \leftarrow \sum |L^i|$
(2) $u \leftarrow L^i$ for any $L^i \in L$ such that $|L^i| = n+1$
(3) $L \leftarrow L \setminus u$
(4) $B \leftarrow uu$
(5) **while** $L \neq \emptyset$
(6) **for** $\alpha = 1$ **to** $|A| - 1$
(7) $w \leftarrow B_{j-n-1} \ldots B_j \overline{B_{j+1}}^{\alpha}$
(8) $w' \leftarrow \text{LYNDON}(w)$
(9) **if** $w' \in L$ **then**
(10) $B \leftarrow B_1 \ldots B_j w_n w_1 \ldots w_{|w'|-1} B_{j+1} \ldots$
(11) $L \leftarrow L \setminus w'$
(12) **end if**
(13) **end for**
(14) **end for**
(15) $B \leftarrow B_1 \ldots B_{Size}$
where $\overline{a}^{\alpha} = a + \alpha \bmod |A|$ and $\text{LYNDON}(w)$ return the Lyndon word z such that $z^{\infty} = w^{\infty}$.

Algorithm 1. Produce a de Bruijn sequence using the Lyndon words of the language.

have the same label. Therefore the label of $l(f) = l(g)$ but in this case the head of f and the head of g are the same vertex, producing a contradiction. This proves the uniqueness of the cycles. □

Corollary 12. *The set of Lyndon words of length dividing $n+1$ in $\mathcal{L}(X_{\mathcal{F}})$ corresponds to a partition of the set of arcs of $G^{\mathcal{D}}$.*

Proof. Let C be a cycle of length d with d dividing $n+1$ and let us label it w in such a way that $\forall u, v$ such that $w = uv$, we have that either $w = vu$ or $w < vu$. We only have to prove that w is not a repetition of a smaller word u.

Let us assume that $w = u^i$ for an integer $i \geqslant 2$ and let x and y be two vertices of C at distance $|u|$ over C such that the walk of C from x to y has label u. Since both vertices are in C, x and y are factors of length n of the word $w^{(n+1)/d}$. Since the walk from x to y has label u, u is a suffix of y. Moreover, since $w^{(n+1)/d} = u^{i(n+1)/d}$, uu is a suffix of y then u is also a suffix of x, concluding that $x = y$.

Therefore, every cycle in the partition has a different label which is a Lyndon word of length dividing $n+1$.

It remains to prove that to each Lyndon word, one can associate a cycle. But this can be proved using cardinality considerations. Indeed, a periodic word of length $n+1$ has either least period $n+1$ or least period d with d dividing $n+1$. Therefore,

$$\left|\mathcal{P}_{n+1}(X_{\mathcal{F}})\right| = \sum_{d|n+1} \left|\{\text{words with least period } d\}\right|.$$

Now, a word with least period d is a Lyndon word or one of the $d-1$ rotations of a Lyndon word of length d. Hence,

$$\sum_{d|n+1} \left|\{\text{words with least period } d\}\right| = \sum_{d|n+1} d \cdot \left|\{\text{Lyndon words of length } d\}\right|.$$

Since $|E(G^{\mathcal{D}})| = |\mathcal{P}_{n+1}(X_{\mathcal{F}})|$ we conclude. □

Now we are prepared to construct an algorithm producing a de Bruijn sequence for $\mathcal{D} = \mathcal{P}_{n+1}(X_{\mathcal{F}})$.

Given a partition in cycles of an Eulerian graph, the following strategy produces an Eulerian cycle: we can start from an arc and follow the corresponding cycle in the partition, until we reach an intersection with another cycle in the partition. At this point we follow the other cycle and when we return to the intersection we continue with the original cycle. Using this procedure recursively we construct an Eulerian cycle.

By Corollary 12, we can reproduce this strategy in terms of the Lyndon words of length diving $n+1$ in $\mathcal{L}(X_{\mathcal{F}})$ obtaining Algorithm 1 producing a de Bruijn sequence for $\mathcal{D} = \mathcal{P}_{n+1}(X_{\mathcal{F}})$ without constructing the graph $G^{\mathcal{D}}$.

The function LYNDON() in the algorithm can be implemented with an on-line automata accepting when a suffix of B is a factor of length n of rotations of the words in L, allowing to do this step in a constant time (see [13]). Hence, steps (7)–(12) in the algorithm have complexity $\mathcal{O}(n)$ and these steps are repeated at most $|A| \cdot |L|$ times. Therefore, the complexity of the algorithm is $\mathcal{O}(|A| \cdot |L| \cdot n)$. Since $Size = \sum_L |L^i|$ is the

size of the input (and also the size of the output) and *Size* is at most $n \cdot |L|$, we conclude that our procedure is a linear time algorithm. Note that the input of the algorithm can also be constructed in an efficient way (see [14]).

Acknowledgement

The author wishes to thank Dominique Perrin and the members of the Institute Gaspard Monge for helpful comments and an anonymous referee for the constructive suggestions.

References

[1] N.G. de Bruijn, A combinatorial problem, Nederl. Akad. Wetensch., Proc. 49 (1946) 758–764.
[2] S.K. Stein, The mathematician as an explorer, Sci. Amer. 204 (5) (1961) 148–158.
[3] J.-C. Bermond, R.W. Dawes, F.Ö. Ergincan, De Bruijn and Kautz bus networks, Networks 30 (3) (1997) 205–218.
[4] F. Chung, P. Diaconis, R. Graham, Universal cycles for combinatorial structures, Discrete Math. 110 (1–3) (1992) 43–59.
[5] C. Flye Sainte-Marie, Question 48, L'Intermédiaire Math. 1 (1894) 107–110.
[6] I.J. Good, Normal recurring decimals, J. London Math. Soc. 21 (1946) 167–169.
[7] G. Rauzy, Suites à termes dans un alphabet fini, in: Seminar on Number Theory, 1982–1983 (Talence, 1982/1983), Univ. Bordeaux I, Talence, 1983, Exp. No. 25, 16 pp.
[8] J. Cassaigne, Complexité et facteurs spéciaux, Bull. Belg. Math. Soc. 4 (1997) 67–88.
[9] D. Lind, B. Marcus, Symbolic Dynamics and Codings, Cambridge University Press, Cambridge, 1995.
[10] H. Fredricksen, A survey of full length nonlinear shift register cycle algorithms, SIAM Rev. 24 (2) (1982) 195–221.
[11] H. Fredricksen, J. Maiorana, Necklaces of beads in k colors and k-ary de Bruijn sequences, Discrete Math. 23 (1978) 207–210.
[12] F. Ruskey, C. Savage, T.M. Wang, Generating necklaces, J. Algorithms 13 (3) (1992) 414–430.
[13] M. Crochemore, C. Hancart, T. Lecroq, Algorithmique du texte, Vuibert, 2001.
[14] F. Ruskey, J. Sawada, Generating necklaces and strings with forbidden substrings, in: Lecture Notes in Comput. Sci., vol. 1858, Springer, Berlin, 2000, pp. 330–339.