

R U T C O R  
R E S E A R C H  
R E P O R T

MORE EXTREMAL PROPERTIES OF DE  
BRUIJN SEQUENCES

Endre Boros <sup>a</sup>      Vladimir Gurvich <sup>b</sup>  
Matthew Oster <sup>c</sup>

RRR 8-2009, APRIL 29, 2009

RUTCOR  
Rutgers Center for  
Operations Research  
Rutgers University  
640 Bartholomew Road  
Piscataway, New Jersey  
08854-8003  
Telephone:      732-445-3804  
Telefax:        732-445-5472  
Email:    rrr@rutcor.rutgers.edu  
<http://rutcor.rutgers.edu/~rrr>

---

<sup>a</sup>RUTCOR–Rutgers Center for Operations Research, Rutgers, The State University of New Jersey, 640 Bartholomew Road, Piscataway, NJ 08854-8003, USA; e-mail: boros@rutcor.rutgers.edu

<sup>b</sup>RUTCOR–Rutgers Center for Operations Research, Rutgers, The State University of New Jersey, 640 Bartholomew Road, Piscataway, NJ 08854-8003, USA; e-mail: gurvich@rutcor.rutgers.edu

<sup>c</sup>RUTCOR–Rutgers Center for Operations Research, Rutgers, The State University of New Jersey, 640 Bartholomew Road, Piscataway, NJ 08854-8003, USA; e-mail: osterm38@eden.rutgers.edu

# RUTCOR RESEARCH REPORT

RRR 8-2009, APRIL 29, 2009

## MORE EXTREMAL PROPERTIES OF DE BRUIJN SEQUENCES

Endre Boros

Vladimir Gurvich

Matthew Oster

**Abstract.** Given an alphabet  $\llbracket q \rrbracket := \{0, 1, \dots, q-1\}$  of cardinality  $q \geq 2$ , we consider cyclic strings or *words*  $S = (s_0, s_1, \dots, s_{n-1})$  of length  $|S| := n \geq 1$ , where  $s_i \in \llbracket q \rrbracket$  for each  $i \in \llbracket n \rrbracket$ . Furthermore, given  $S \in \llbracket q \rrbracket^n$  and  $m \in \mathbf{N}$ , let us consider in  $S$  the set of all unique  $m$ -substrings or  $m$ -subwords  $W_m(S) = \{(s_i, s_{i+1}, \dots, s_{i+m-1}) \mid i \in \llbracket n \rrbracket\}$ , where indices are taken modulo  $n$ . A word  $S$  is called  $m$ -minimal if strict containment  $W_m(S') \subset W_m(S)$  holds for no word  $S'$ , and  $W_m(S') = W_m(S)$  holds for no  $S'$  such that  $|S'| < |S|$ . We prove that the set of all  $q$ -ary  $m$ -minimal words (denoted  $M_m$ ) is in one-to-one correspondence with the chordless directed cycles (di-cycles) of the de Bruijn-Good directed graph (digraph)  $G_m^{(q)}$  and, furthermore, with the simple di-cycles of  $G_{m-1}^{(q)}$ , while the latter are in one-to-one correspondence with the closed directed walks of  $G_{m-2}^{(q)}$ . This implies each  $S \in M_m$  is such that  $|S| \leq q^{m-1}$ . In particular, the longest  $m$ -minimal words are of cardinality  $q^{m-1}$  and in one-to-one correspondence with the set of all  $q$ -ary  $(m-1)$ th order de Bruijn sequences  $B_{m-1}^{(q)}$ . Furthermore, the latter set is known to be in one-to-one correspondence with the Hamiltonian di-cycles of  $G_{m-1}^{(q)}$ , and the Eulerian directed circuits of  $G_{m-2}^{(q)}$ .

# 1 Introduction

The de Bruijn-Good directed graph (di-graph) of order  $n$ , or simply de Bruijn Graph [1, 6], is a di-graph  $G_n^{(q)}$  with a vertex set  $V_n := V(G_n^{(q)})$  consisting of all  $q$ -ary  $n$ -tuples, and an arc  $(u, v) \in E_n := E(G_n^{(q)})$  exists if and only if  $u = (u_1, \dots, u_n), v = (v_1, \dots, v_n) \in V_n$  are such that  $u_{i+1} = v_i$  for  $i = 1, 2, \dots, n-1$  (see Figures 1 and 2 for small binary de Bruijn graphs). One can easily see that, by definition, each vertex has in-degree and out-degree equal to  $q$ , and thus  $|E_n| = q|V_n| = q^{n+1}$ . It can also be seen that  $G_{n+1}^{(q)}$  is the directed line graph of  $G_n^{(q)}$  [11]. A  $q$ -ary  $n$ th order de Bruijn sequence  $S$  is one where every  $q$ -ary  $n$ -tuple appears uniquely as a contiguous subsequence. The set of all de Bruijn sequences for fixed  $n$  and  $q$  is denoted  $B_n^{(q)}$ . Clearly each  $S \in B_n^{(q)}$  has size  $q^n$ , and such a sequence can be found as a Hamiltonian di-cycle in  $G_n^{(q)}$ , or an Eulerian directed circuit in  $G_{n-1}^{(q)}$  [4, 5]. The number of such binary maximum length di-cycles, i.e. de Bruijn sequences in  $G_n^{(2)}$ , is known to be  $2^{2^{n-1}-n}$  [1, 3], and in general  $(q!)^{q^{n-1}} q^{-n}$  [4], which are found by counting the number of non-linear shift registers which may generate such sequences. In the next section, we show that there is a one-to-one correspondence between  $B_{m-1}^{(q)}$  and all  $m$ -minimal words of maximum size.

Of particular interest is the question of existence of (simple) di-cycles, not necessarily those of maximum length, but ones of length  $k$  such that  $1 \leq k \leq q^n$  in  $G_n^{(q)}$ . By induction on the number of edges and the well known fact that  $G_n^{(q)}$  is Eulerian, i.e. the digraph can be decomposed into disjoint di-cycles, it is shown in [9] that all di-cycles of arbitrary length  $1 \leq k \leq q^n$  exist for each  $q \in \mathbf{N}$ . In the following section, we show the one-to-one correspondence between simple di-cycles  $S$  of  $G_{m-1}^{(q)}$  and minimal  $q$ -ary words of  $M_m$ , and thus the existence of simple di-cycles of all lengths in  $G_{m-1}^{(q)}$  implies the existence of minimal  $q$ -ary  $S \in M_m$ , for each  $1 \leq |S| \leq q^{m-1}$ .

There is extensive numerical data and asymptotically tight bounds in [12] for the number of di-cycles of length  $k \leq q^n$  for all values of  $q$  and  $n$  (denoted  $\beta^{(q)}(n, k)$ ). Of course some exact values are known, such as  $\beta^{(q)}(n, q^n) = (q!)^{q^{n-1}} q^{-n}$  and  $\beta^{(q)}(n, q^n - 1) = \frac{q}{q-1} \beta^{(q)}(n, q^n)$  [12]. It is also known that for each  $n \geq k - 1$ ,  $k$  fixed,  $\beta^{(q)}(n, k) = \frac{1}{k} \sum_{k|d} \mu(d/k) q^d$ , which is the total number of  $q$ -ary di-cycles of length  $k$  with  $\mu(\cdot)$  denoting the möbius function. For example, the binary case with  $n \geq k - 1$ ,  $\{\beta^{(2)}(n, k)\}_{k \geq 1}$  presents us with the sequence  $(2, 1, 2, 3, 6, 9, 18, 30, 56, 99, \dots)$ . Note that  $\beta^{(2)}(n, k)$  is a nondecreasing function with respect to  $n$  when  $k$  is fixed.

For applications of de Bruijn sequences in topics such as cryptography, interconnection networks, and pseudo-random number generation, see [5, 8].

# 2 Main Results

We call  $S \in \llbracket q \rrbracket^n$ , where  $\llbracket q \rrbracket = \{0, 1, \dots, q-1\}$ , a *word*. For fixed  $m \in \mathbf{N}$ , each contiguous subsequence of  $S$  of length  $m$  is called an  *$m$ -subword*. The set of all unique  $m$ -subwords of a given  $S$  is denoted  $W_m(S) = \{(s_i, s_{i+1}, \dots, s_{i+m-1}) \mid i \in \llbracket n \rrbracket\}$ , and by definition  $|W_m(S)| \leq$

$q^m$ . We let  $S_i^{(m)} = (s_i, s_{i+1}, \dots, s_{i+m-1})$  be the  $i$ th  $m$ -subword of  $S$ . We say that a word  $S$  is  $m$ -minimal (or simply *minimal* when  $m$  is clear from context) if there is no other  $S'$  such that  $W_m(S') \subset W_m(S)$  and no  $S'$  such that  $W_m(S') = W_m(S)$  with  $|S'| < |S|$ . If such an  $S'$  exists, then we say  $S'$  *dominates*  $S$ . Let  $M_m = \{S \mid S \text{ } m\text{-minimal}\}$  be the set of  $m$ -minimal words.

Notice that we allow for a subword to wrap around itself, e.g.  $m = 4$  and  $S = (0, 1) \in \llbracket 2 \rrbracket^2$  implies  $W_4(S) = \{(0, 1, 0, 1), (1, 0, 1, 0)\}$ . This is for convenience. One could easily define minimality similarly, but restricting  $S \in M_m$  such that  $|S| \geq m$ . First we denote  $S^t := (S, S, \dots, S)$  where  $S$  is appended to itself  $t$  times. Then we find  $M_m$  as in our definition above, and finally let  $t = \lceil m/|S| \rceil$ . Clearly  $W_m(S^t) = W_m(S)$  and there is no  $t'$  such that  $S^{t'}$  dominates  $S^t$  with  $m \leq t' < t$ . We let our new set of minimal words be  $M'_m$ , and thus we have  $S^t \in M'_m$  if and only if  $S \in M_m$  where  $t$  is as we defined above. For example take  $(\ell)$ , where  $\ell \in \llbracket q \rrbracket$ . Clearly  $(\ell) \in M_m$  for any  $m \geq 1$ , but with our second definition we have  $(\ell)^m \in M'_m$ . One could further develop an analysis with restricted attention to lexicographically minimal di-cycles and words. Our focus in this section is to demonstrate the one-to-one correspondence between each minimal word  $S \in M_m$  and an  $|S|$ -di-cycle in  $G_m^{(q)}$ .

For  $v, v' \in V_m$  we denote the sequence obtained by traversing an arc  $(v, v') \in E_m$  as  $\langle v, v' \rangle = (v_0, \dots, v_{m-1}, v'_0, \dots, v'_{m-1}) = (v_0, v'_0, \dots, v'_{m-1})$ . Thus  $(v, v') \in E_m$  if and only if  $\langle v, v' \rangle \in V_{m+1}$ . Furthermore, for  $v, v' \in V_m$ , we say  $v'$  is a *successor* of  $v$  if and only if  $(v, v') \in E_m$ , and we write  $v \rightarrow v'$ . A di-cycle  $C$  of size  $n \geq 1$  ( $n$  can take on the value 1 since there exist loops, but not parallel arcs, in our digraphs) in  $G_m^{(q)}$ , can be written as  $C = (v^0, v^1, \dots, v^{n-1})$ , where  $v^i \rightarrow v^{i+1}, i \in \llbracket n \rrbracket$  and  $v^{n-1} \rightarrow v^0$ , or as the (cyclic) sequence  $\langle v^0, v^1, \dots, v^{n-1} \rangle$  which corresponds to the word  $(v_0^0, v_0^1, \dots, v_0^{n-1})$  where  $v^i = (v_0^i, \dots, v_{m-1}^i)$  for each  $i \in \llbracket n \rrbracket$ . Thus the one-to-one correspondence between feasible words  $S = (s_0, \dots, s_{n-1}) \in \llbracket q \rrbracket^n$  and di-cycles  $(v^0, \dots, v^{n-1})$  of  $G_m^{(q)}$  is clear, namely  $(s_0, \dots, s_{n-1}) \equiv \langle v^0, \dots, v^{n-1} \rangle$  where  $v^i := S_i^{(m)} \in V_m$  for all  $i \in \llbracket n \rrbracket$ , and now we can interchangeably talk about (minimal) di-cycles or (minimal) words. We also refer to a di-cycle in  $G_{m-1}^{(q)}$  as the *projection* of the corresponding di-cycle in  $G_m^{(q)}$ .

**Lemma 1.** *Fix  $m \in \mathbf{N}$ , and let  $S \in \llbracket q \rrbracket^n$ . If  $S \in M_m$ , then  $S_i^{(m)} = S_j^{(m)} \in W_m(S)$  implies  $i = j$ .*

*Proof.* Suppose  $S = (s_0, \dots, s_n) \in M_m$ , but also suppose, to arrive at a contradiction,  $\exists i \neq j \in \llbracket n \rrbracket$  such that  $S_i^{(m)} = S_j^{(m)}$ . Without loss of generality, suppose  $i < j$ . Clearly  $S$  is a feasible di-cycle of  $G_m^{(q)}$  by our previous discussion where  $S_k^{(m)} \in V_m$  for all  $k \in \llbracket n \rrbracket$ . We now construct a smaller feasible di-cycle, by taking a subsequence of vertices  $S' = (S_i^{(m)}, \dots, S_{j-1}^{(m)})$  if  $j - i \leq n/2$ , otherwise  $S' = (S_j^{(m)}, \dots, S_{i+n-1}^{(m)})$ . Since  $S_i^{(m)} = S_j^{(m)}$  we have  $S_{j-1}^{(m)} \rightarrow S_i^{(m)}$  and  $S_{i-1}^{(m)} \rightarrow S_j^{(m)}$ , and so  $S'$  is a feasible di-cycle of  $G_m^{(q)}$ . Furthermore,  $W_m(S') \subseteq W_m(S)$  and  $|S'| \leq |S|/2$ . Thus  $S'$  dominates  $S$ , so  $S \notin M_m$ , a contradiction. Hence the claim follows.  $\square$

This lemma shows us that a minimal word  $S \in M_m$ , cannot have repeated  $m$ -subwords, and thus contains any  $m$ -subword at most once. Since each  $m$ -subword is a vertex in  $G_m^{(q)}$ , we see that the corresponding di-cycle contains no repeated vertices, and thus is a simple, finite di-cycle of length  $|S| \leq q^m$ . Similarly, a minimal word  $S \in M_m$  cannot repeat edges in  $G_{m-1}^{(q)}$ . By the previous lemma, we have a simple way of reducing a non-simple di-cycle in  $G_m^{(q)}$  of arbitrary length, to a di-cycle which is at most half the length and dominates the original. The upper bound to the size of a simple di-cycle is reached only when our word  $S \in B_m^{(q)}$ , but this is not to say that such an  $S$  is a minimal word. This can be seen by considering the word  $S' := (\ell) \in \llbracket q \rrbracket$ . This is clearly a feasible word,  $|W_m(S')| = 1$ , and thus dominates all other words  $S''$  with  $(\ell)^m \in W_m(S'')$ . Of course each de Bruijn sequence  $S$  must consist of all  $x \in \llbracket q \rrbracket^m$ , but yet  $W_m(S') \subset W_m(S)$ . So for each  $S \in M_m$  we must have  $|S| < q^m$ . The next result will provide us with another useful property of minimal words.

**Lemma 2.** *Fix  $m \in \mathbf{N}$  and let  $S \in \llbracket q \rrbracket^n$ .  $S \in M_m$  if and only if for each  $i \in \llbracket n \rrbracket$ ,  $S_i^{(m)} \rightarrow S_j^{(m)}$  implies  $j = i + 1$ .*

*Proof.* For necessity, suppose  $S = (s_0, \dots, s_{n-1}) \in M_m$  and  $\exists i : S_i^{(m)} \rightarrow S_j^{(m)}$ , but  $j \neq i + 1$ . Clearly, if  $S_k^{(m)} = S_\ell^{(m)}$  for some  $k \neq \ell \in \llbracket q \rrbracket$  then by Lemma 1, we have a repeated vertex in the corresponding di-cycle in  $G_m^{(q)}$ , and thus a contradiction. So assume there are no repeated vertices in  $S$  and, without loss of generality, suppose  $i + 1 < j$ . Clearly if  $S_i^{(m)} \rightarrow S_j^{(m)}$ , then we have an arc between the two  $m$ -subwords. We now form a strictly smaller di-cycle by taking  $S' = (S_j^{(m)}, S_{j+1}^{(m)}, \dots, S_i^{(m)})$ , where indices are taken modulo  $n$ . Since we have  $i + 1 < j$  and no repeated vertices,  $|S'| < n$  and  $W_m(S') \subset W_m(S)$ , which clearly implies that  $S \notin M_m$ , a contradiction.

For sufficiency, suppose for each  $i \in \llbracket n \rrbracket$ ,  $S_i^{(m)} \rightarrow S_j^{(m)}$  implies  $j = i + 1$ , but  $S \notin M_m$ . So  $\exists S'$  which dominates  $S$ , i.e.  $W_m(S') \subset W_m(S)$  or  $W_m(S') = W_m(S)$  and  $|S'| < |S|$ . If the latter is true, then  $S'$  must be a strict sub-di-cycle along the same vertices. This clearly implies that we have a repeated vertex in  $S$ , and thus  $\exists k : S_k^{(m)} \rightarrow S_\ell^{(m)}$  with  $k \neq \ell + 1$ , a contradiction. Now suppose  $W_m(S') \subset W_m(S)$ , which implies  $\exists k : S_k^{(m)} \in W_m(S) \setminus W_m(S')$ . Now choose the largest index  $i < k$  in  $S$  such that  $S_i^{(m)} = S_\ell^{(m)}$  for some  $\ell \in \llbracket |S'| \rrbracket$ . Since  $S_k^{(m)} \neq S_r^{(m)}$  for each  $r \in \llbracket |S'| \rrbracket$ , we must have  $S_i^{(m)} \rightarrow S_j^{(m)}$  where  $i < k < j$  or  $j < i$  and  $S_j^{(m)} = S_{\ell+1}^{(m)}$ . In either case  $j \neq i + 1$ , a contradiction, and thus the claim follows.  $\square$

This lemma implies that a minimal di-cycle  $S$  in  $G_m^{(q)}$  cannot have any two nonadjacent vertices sharing an edge. Otherwise we can find a smaller di-cycle  $S'$  which dominates  $S$ . If  $S$  has this property, we say it is *chordless*, or *1-pathless*. In general, if there exists no directed di-path of length  $k$  or less between any two vertices of a cycle  $S$  (other than the di-paths on the di-cycle), we say it is *k-pathless*. If a di-cycle is not *k-pathless*, we say that each nonadjacent pair sharing a di-path of length  $\ell$  (or  $\ell$ -path),  $1 \leq \ell \leq k$ , are *successive  $\ell$ -pairs*. For the simple case of  $\ell = 1$ , we call a 1-path a *chord* and the vertices sharing it *successive pairs*. The next result is our main result, and proves the one-to-one correspondence between minimal words and simple di-cycles in  $G_{m-1}^{(q)}$ .

**Theorem 1.** Fix  $m \in \mathbf{N}$  and let  $S \in \llbracket q \rrbracket^n$ . Then  $S \in M_m$  if and only if  $S$  is a simple di-cycle in  $G_{m-1}^{(q)}$ .

*Proof.* We need only to establish a one-to-one correspondence between the simple di-cycles of  $G_{m-1}^{(q)}$  and the chordless di-cycles of  $G_m^{(q)}$ . The statement will then follow from Lemma 2 and the previously established one-to-one correspondence between all feasible words of  $\llbracket q \rrbracket^n$  and all di-cycles of length  $n$  in  $G_m^{(q)}$ .

A simple di-cycle  $C$  in  $G_{m-1}^{(q)}$  is one which does not repeat any vertices, i.e.  $C = (v^0, \dots, v^{n-1})$  where  $v^i \neq v^j \in V_{m-1}$  for all  $i \neq j \in \llbracket n \rrbracket$ .  $C$  can also be written as a di-cycle in  $G_m^{(q)}$ :  $C = (\langle v^0, v^1 \rangle, \dots, \langle v^{n-1}, v^0 \rangle)$ . Suppose there exists a chord  $e = (\langle v^i, v^{i+1} \rangle, \langle v^j, v^{j+1} \rangle)$  with  $j \neq i$  and  $j \neq i + 1$ . This clearly implies that  $v^i \rightarrow v^j$ , i.e.  $(v^i, v^j) \in E_{m-1}$ . But since  $j \neq i$ ,  $j \neq i + 1$ , and  $(v^i, v^{i+1}) \in E_{m-1}$ , we have two separate edges traversed in  $C$  with  $v^i$  as the tail. This implies  $v^i$  is a repeated vertex in  $C$ , and so we have contradicted  $C$  being di-simple in  $G_{m-1}^{(q)}$ .

Suppose now that  $C$  is a chordless cycle in  $G_m^{(q)}$ . Let  $C = (\langle v^0, v^1 \rangle, \dots, \langle v^{n-1}, v^0 \rangle)$ , where each  $v^i \in V_{m-1}$  for all  $i \in \llbracket n \rrbracket$ . Suppose  $C$  is not simple in  $G_{m-1}^{(q)}$ . This implies some  $v^i = v^j$  for  $i < j$ . But then  $v^{i-1} \rightarrow v^i = v^j \rightarrow v^{j+1}$ . So we have  $e^{i-1} := \langle v^{i-1}, v^i \rangle$  and  $e^j := \langle v^j, v^{j+1} \rangle \in V_m$ . Furthermore,  $e^{i-1}$  and  $e^j$  are vertices of  $C$  but  $(e^{i-1}, e^j) \in E_m$  is not an edge of  $C$ . Thus  $C$  has a chord, a contradiction.  $\square$

**Corollary 1.** Fix  $m \in \mathbf{N}$  and let  $S \in \llbracket q \rrbracket^n$ . Then  $S \in M_m$  if and only if  $S$  is a closed directed walk (di-walk) in  $G_{m-2}^{(q)}$ .

*Proof.* We need only to prove the one-to-one correspondence between the closed di-walks of  $G_{m-2}^{(q)}$  and the simple di-cycles of  $G_{m-1}^{(q)}$ . Then using Theorem 1, the statement follows. But this is trivially true by definitions of closed di-walk and simple di-cycle, and also since  $G_{m-1}^{(q)}$  is the directed line graph of  $G_{m-2}^{(q)}$ .  $\square$

Notice that  $q$  is irrelevant in our proofs. This is partly because for any  $q$  our graph  $G_m^{(q)}$  is Eulerian and also since de Bruijn graphs of order  $m$  and  $m + 1$  are strongly related. Another corollary follows our theorem, generalizing the structure of a di-cycle in higher orders.

**Corollary 2.** Fix  $m \in \mathbf{N}$  and let  $S \in \llbracket q \rrbracket^n$ . Then  $S \in M_m$  if and only if  $S$  is a  $k$ -pathless di-cycle in  $G_{m+k-1}^{(q)}$  for  $k \geq 1$ .

*Proof.* We prove the statement by induction on  $k$ . For  $k = 1$ , the statement holds, due to Theorem 1. Now suppose the set of minimal di-cycles  $M_m$  is one-to-one with all  $j$ -pathless di-cycles of  $G_{m+j-1}^{(q)}$  for all  $1 \leq j < k$ . The  $(k - 1)$ -pathless property of any minimal  $S$  of  $G_{m+k-2}^{(q)}$  implies that each di-path between two nonadjacent vertices (arcs), not using arcs of  $S$ , has at least  $k - 1$  arcs between them. Since  $G_{m+k-1}^{(q)}$  is the line graph of  $G_{m+k-2}^{(q)}$ , the same  $(k - 1)$ -path, now in  $G_{m+k-1}^{(q)}$ , is one between two nonadjacent vertices on the di-cycle, and has at least  $k - 1$  vertices between them. Thus for any pair of vertices of  $S$  in  $V_{m+k-1}$ , each di-path between them not on  $C$  has at least  $k$  arcs between them. But this satisfies the  $k$ -pathless condition, and the statement follows.  $\square$

**Corollary 3.** *The longest  $q$ -ary  $m$ -minimal words are of cardinality  $q^{m-1}$  and in one-to-one correspondence with the set of  $q$ -ary  $(m - 1)$ th order de Bruijn sequences  $B_{m-1}^{(q)}$ .*

*Proof.* All maximum length minimal  $q$ -ary words with respect to  $m$ -subwords have length  $q^{m-1}$ . Since the set of minimal words  $M_m$  is in one-to-one correspondence with the simple di-cycles of  $G_{m-1}^{(q)}$  and all simple di-cycles of maximum size in  $G_{m-1}^{(q)}$  are Hamiltonian di-cycles of length  $q^{m-1}$ , our statement holds trivially.  $\square$

### 3 Examples

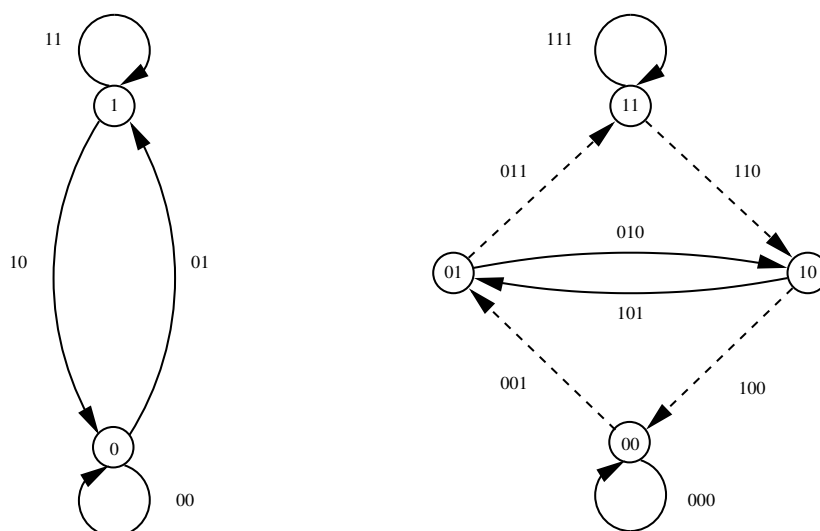


Figure 1:  $G_1^{(2)}$  (left);  $G_2^{(2)}$  (right)

In this section we give examples to demonstrate how  $q$ -ary  $m$ -minimal words correspond to di-cycles in the de Bruijn graphs  $G_m^{(q)}, G_{m-1}^{(q)}, G_{m-2}^{(q)}$  for  $m = 3$  and  $m = 4$ . Let us consider four graphs  $G_m^{(2)}$  for  $m \leq 4$ , given in Figures 1-3. We label each graph's vertices and arcs as strings of 0 and 1 valued entries (e.g. we abbreviate  $(i_1, i_2, \dots, i_n) \in \llbracket 2 \rrbracket^n$  with  $i_1 i_2 \dots i_n$ ). Also note that, for the sake of brevity, the decimal integers from 0 to  $q^m - 1$  denote the binary  $m$ -bit words.

**Example 1.** *The binary 2nd-order de Bruijn sequence  $S = (0, 0, 1, 1)$  is a 3-minimal word of maximum size.*

We see that  $W_3(S) = \{(0, 0, 1), (0, 1, 1), (1, 1, 0), (1, 0, 0)\}$ . Looking at the graph in Figure 2 (dashed arcs), we see that the elements of  $W_3(S)$  are precisely the vertices of the chordless di-cycle  $(1, 3, 6, 4)$  in  $G_3^{(2)}$ . So our corresponding cyclic sequence  $\langle 1, 3, 6, 4 \rangle \equiv (0, 0, 1, 1)$  is 3-minimal, by Theorem 1. Now looking at  $W_2(S) = \{(0, 0), (0, 1), (1, 1), (1, 0)\}$ , we notice

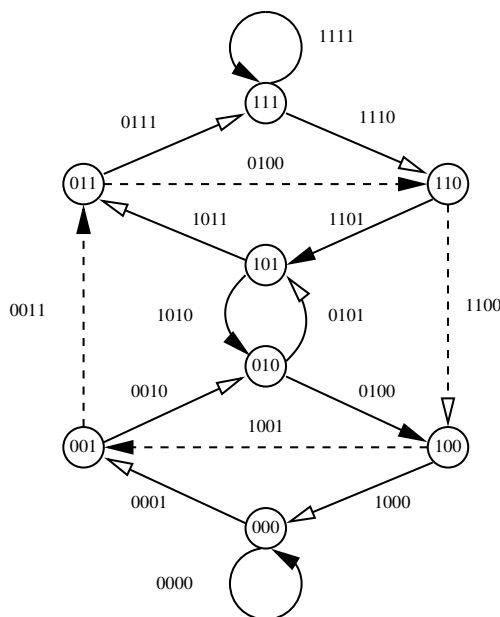


Figure 2:  $G_3^{(2)}$

that  $S \in B_2^{(2)}$ , and so we have a Hamiltonian di-cycle in  $G_2^{(2)}$ , namely  $(0, 1, 3, 2)$  (see dashed arcs of Figure 1). Furthermore, we see that the cycle’s projection is an Eulerian di-cycle in  $G_1^{(2)}$ .

**Example 2.** *The binary 3rd-order de Bruijn sequence  $S = (0, 0, 0, 1, 0, 1, 1, 1)$  is a 4-minimal word of maximum size.*

Our set of 4-subwords is  $W_4(S) = \{1, 2, 5, 11, 7, 14, 12, 8\}$ . This set is precisely the set of vertices traversed in the order shown (see white-arrowed arcs of Figure 3) of the di-cycle in  $G_4^{(2)}$ . Now in Figure 2, we have the corresponding di-cycle in its 4-subword (edge) representation  $(1, 2, 5, 11, 7, 14, 12, 8)$  or its 3-subword (vertex) representation  $(0, 1, 2, 5, 3, 7, 6, 4)$ . The cyclic sequence  $\langle 0, 1, 2, 5, 3, 7, 6, 4 \rangle \equiv S$  and is in  $B_3^{(2)}$ , and so we have a Hamiltonian di-cycle in  $G_3^{(2)}$  (see white-arrowed arcs in Figure 2). Finally, it is clear that the vertices traversed in the previously mentioned Hamiltonian di-cycle, are now the arcs traversed of the Eulerian di-cycle of  $G_2^{(2)}$ .

**Example 3.** *The binary word  $S = (0, 0, 1)$  is 3-minimal.*

We have  $W_3(S) = \{(0, 0, 1), (0, 1, 0), (1, 0, 0)\}$ . The corresponding cycle in  $G_3^{(2)}$  shares the elements of  $W_3(S)$  as its vertices (see Figure 2) and is a chordless di-cycle, since no two nonadjacent pairs of vertices share an edge (i.e. in this case no loops are present). Furthermore, our cycle in  $G_2^{(2)}$  is simple (but not Hamiltonian) and the projection into  $G_1^{(2)}$  is a closed di-walk (but not an Eulerian di-cycle).

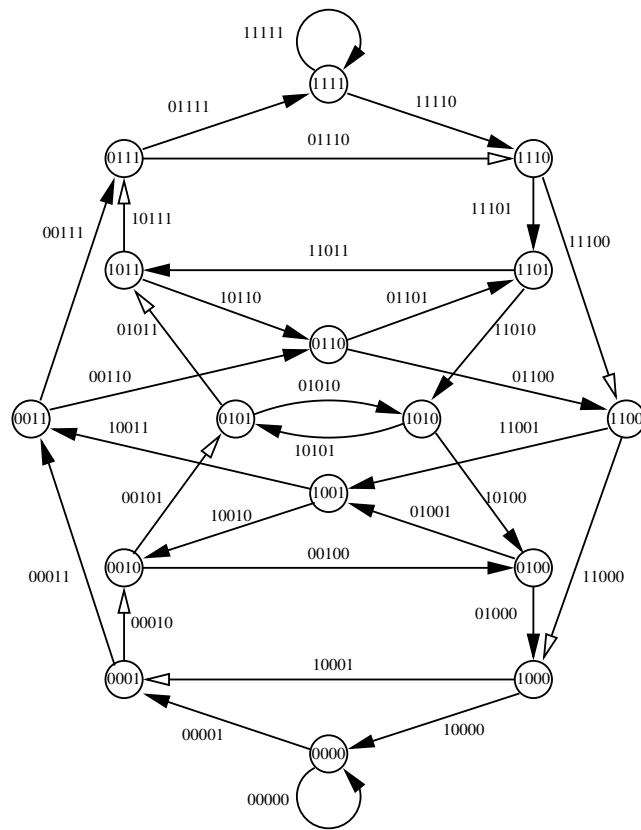


Figure 3:  $G_4^{(2)}$

**Example 4.** *The binary word  $S = (0, 0, 1, 0, 1, 1)$  is 4-minimal.*

We see that  $W_4(S) = \{2, 5, 11, 6, 12, 9\}$ , and the elements correspond to vertices of the chordless di-cycle  $(2, 5, 11, 6, 12, 9)$  in  $G_4^{(2)}$  (see Figure 3). The vertices of this di-cycle become the edges of di-cycle  $(2, 5, 3, 6, 4, 1)$ , which is a simple (but not Hamiltonian) di-cycle in  $G_3^{(2)}$ . Finally, the projection of this di-cycle into  $G_2^{(2)}$  is a closed di-walk (but not an Eulerian di-cycle).

## 4 Conclusion

We have presented a short survey regarding the properties of de Bruijn-Good digraphs. We also introduced a new concept of  $m$ -minimal words and constructed a one-to-one correspondence between them and simple di-cycles of simple di-cycles of  $(m - 1)$ th order de Bruijn graphs, chordless di-cycles of  $m$ th order de Bruijn graphs, and more generally,  $k$ -pathless di-cycles of  $(m + k - 1)$ th order de Bruijn digraphs. In particular, the longest  $q$ -ary  $m$ -minimal words are of cardinality  $q^{m-1}$  and in one-to-one correspondence with the Hamiltonian di-cycles of the  $q$ -ary  $(m - 1)$ th order de Bruijn graph. Furthermore, these di-cycles are known to be in one-to-one correspondence with the set of de Bruijn sequences  $B_{m-1}^{(q)}$ .

We further propose investigating  $q$ -ary  $m$ -minimal words over  $S \in \llbracket q \rrbracket^n$ ,  $n \in K \subseteq \mathbf{N}$ , where  $K$  is chosen with respect to some desired restriction on the size of our words. We also propose considering  $m$ -minimal words with respect to some other homomorphic properties such as flips, rotations, and translations. Some equivalence relations within cyclic words are discussed in [7] but without our notion of minimality in mind. Although the existence of di-cycles of arbitrary length in  $q$ -ary  $m$ th order de Bruijn graphs, it is still an open question as to what the exact number of di-cycles of each length is (see [12] for partial results).

## References

- [1] N.G. de Bruijn, *A combinatorial problem*, Proc. Nederl. Akad. Wetensch. 49 (1946) 758–764.
- [2] F. Chung, P. Diaconis, R. Graham, *Universal cycles for combinatorial structures*, Discrete Math. 110 (1992) 43–59.
- [3] C. Flye Sainte-Marie, *Question 48*, L’Intermédiaire Math., 1 (1894) 107–110.
- [4] H. Fredricksen, *A survey of full length nonlinear shift register cycle algorithms*, SIAM Review, 24 (2) (1982) 195–221.
- [5] S.W. Golomb, *Shift Register Sequences*, revised ed., Aegean Park Press, Laguna Hills, CA, 1982.
- [6] I.J. Good, *Normal recurring decimals*, J. London Math. Soc., 21 (1946) 169–172.
- [7] S.G. Hartke, *Binary de Bruijn cycles under different equivalence relations*, Discrete Math. 215 (2000) 93–102.
- [8] A. Lempel, *On a homomorphism of the de Bruijn graph and its application to the design of feedback shift registers*, IEEE Trans. Comput. C-19 (12) (1970) 1204–1209.
- [9] A. Lempel, *m-ary closed sequences*, J. Combin. Theory A-10 (1971) 253–258.
- [10] A. Lempel, *On extremal factors of the de Bruijn Graph*, J. Comb. Theory B-11 (1971) 17–27.
- [11] J.H. van Lint, R.M. Wilson, *A Course in Combinatorics*, Cambridge University Press, Cambridge, 1992, pp. 56–61.
- [12] U.M. Maurer, *Asymptotically-tight bound on the number of cycles in generalized de Bruijn-Good graphs*, Discrete Applied Mathematics, 37 (1992) 421-436.