Proof of Theorem 1: Let $x_k = A_k A_{-k}$. Then $x_k \in \{0,1\}$. The equations shown in (7) are a system of equations in $N-1$ variables which has the following coefficient matrix

$$B = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha^{-1} & \alpha^{-2} & \cdots & \alpha^{-(N-1)} \\ \alpha^{-2} & \alpha^{-4} & \cdots & \alpha^{-2(N-1)} \\ \vdots & & & \\ \alpha^{-(N-1)} & \alpha^{-(N-1)2} & \cdots & \alpha^{-(N-1)^2} \end{bmatrix}.$$

The submatrix consisting of the first $N-1$ rows is the Vandermonde matrix, so the rank of B is $N-1$. It follows that $x_k = 0$, $k = 1, 2, \ldots, N-1$. □

II. Linear Spans of the Sequences in C

In this section, we will derive a decomposition of a sequence in C as rth-order linear sequences where $r \mid n$ and give an achievable upper bound for the linear spans of two-level correlation sequences.

Let G be the set of all coset leaders modulo N. We denote by $f_{\alpha^v}(x)$ the minimal polynomial of α^v for $v \in G$. If $f_{\alpha^v}(x) \neq f_{\alpha^{-v}}(x)$, then we call v a *nonreciprocal-symmetry coset leader modulo N*. Otherwise, we call v a *reciprocal-symmetry coset leader modulo N*. Notice that if v is a nonreciprocal-symmetry coset leader modulo N, so is $p-v$. Let P be the set of all nonreciprocal symmetry coset leaders modulo N which satisfy $v + 2^j u \not\equiv 0 (\mathrm{mod}\, N)$ for u, v in P. Let S be the set of all reciprocal-symmetry coset leaders modulo N. Let $\bar{P} = \{k \mid -k \in P\}$. Then

$$G = P \cup \bar{P} \cup S. \tag{10}$$

The following corollary follows from Theorem 1.

Corollary 1: If $\{a_t\} \in C$, then

$$A_k A_{-k} = 0 \quad \text{if } k \in P \quad \text{and} \quad A_k = 0 \quad \text{if } k \in S. \tag{11}$$

Moreover

$$a_t = \sum_{k \in P} Tr(A_k \alpha^{-r_k t}), \qquad t = 0, 1, \ldots, N-1 \tag{12}$$

where $r_k \in P$ or $r_k \in \bar{P}$, and $A_k \in \{0,1\}$.

The linear span of $\{a_t\}$, denoted by $LS(\{a_t\})$, is the number of nonzero values in the spectrum of $\{a_t\}$. We denote by P_c the set of coset leaders in P together with their conjugates, and denote by S_m the set of all reciprocal-symmetry coset leaders modulo $2^m - 1$, where m is a proper factor of n.

Theorem 2: If $\{a_t\} \in C$, the linear span of $\{a_t\}$ is bounded by

$$LS(\{a_t\}) \leq |P_c| \tag{13}$$

where if n odd

$$|P_c| = \frac{1}{2} \sum_{d \mid n} \sum_{m \mid d} m \mu(m) 2^{d/m} \tag{14}$$

if n even

$$|P_c| = \frac{1}{2} \sum_{d \mid n} \sum_{m \mid d} m (\mu(m) 2^{d/m} - \delta_m |S_m|) \tag{15}$$

where δ_m is zero if m is odd and is one if m is even.

Proof: From Corollary 1, the formula (13) is immediate. Now we will find the number of coset leaders in P. Note that $|G|$ is the number of irreducible polynomials over $\mathrm{GF}(2)$ whose degrees divide n. Golomb [3] showed that $|S| \neq 0$ if and only if n is even. So, if n is odd, then $|S| = 0$. From [2], we get the formulae (14) and (15). □

Remark 1: If N is a prime number and $\{a_t\}$ is a quadratic sequence

$$LS(\{a_t\}) = |P_c| = \frac{1}{2}(N-1) \tag{16}$$

which achieves this upper bound.

References

[1] L. D. Baumert, *Cyclic Difference Sets*, Lecture Notes in Mathematics. New York: Springer-Verlag, 1971.

[2] S. W. Golomb, *Shift Register Sequences*, Revised ed. Laguna Hills, CA: Aegean, p. 39, 1982.

[3] S. W. Golomb, "Theory of transformation group of polynomials over $\mathrm{GF}(2)$ with applications to linear shift register sequences," *Inform. Theory*, vol. 1, no. 1, pp. 87–109, Dec. 1968.

[4] S. W. Golomb, "On the classification of balanced binary sequences of period $2^n - 1$," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 730–732, Nov. 1980.

Linear Complexity of de Bruijn Sequences—Old and New Results

T. Etzion, *Member, IEEE*

***Abstract*—The linear complexity of a de Bruijn sequence is the degree of the shortest linear recursion which generates the sequence. It is well known that the complexity of a binary de Bruijn sequence of length 2^n is bounded below by $2^{n-1} + n$ and above by $2^n - 1$ for $n \geq 3$. We briefly survey the known knowledge in this area. Some new results are also presented, in particular, it is shown that for each interval of length $2^{\lfloor \log n \rfloor + 1}$, in the above range, there exist binary de Bruijn sequences of length 2^n with linear complexity in the interval.**

***Index Terms*— Complexity distribution, de Bruijn sequences, Games and Chan algorithm, linear complexity, minimal complexity.**

I. Introduction

De Bruijn sequences have many applications in communication systems [13]. They have many desirable properties such as long period and low predictability and can be used as stream ciphers [17] in cryptographic applications. A comprehensive survey on those properties can be found in [13]. Outline of past work about construction of those sequences and other related problems can be found in [9]. The *(linear) complexity* $C(S)$ of a sequence S is one of the measures of its predictability [15]. This paper deals with the complexity distribution of de Bruijn sequences.

Let $s_1 s_2 \cdots s_{k-1}$ denote a string of k binary digits. A cyclic, or closed, string is called a *sequence* and is denoted by $S =$

Manuscript received June 29, 1997; revised April 19, 1998. This work was supported in part by the office of Naval Research under Contract N00014-84-K-0189. Part of this work was presented at the meeting in honor of S. W. Golomb for his 60th birthday, Oxnard, CA, May 1992.

The author is with the Computer Science Department, Technion, Israel Institute of Technology, Haifa 32000, Israel. Part of this work was done while the author was with the Department of Electrical Engineering-Systems, University of Southern California.

Communicated by D. Stinson, Associate Editor for Complexity and Cryptography.

Publisher Item Identifier S 0018-9448(99)01423-6.

$[s_0 s_1 \cdots s_{k-1}]$, where $k = l(S)$ is the *length* of the sequence (or the string). For a string S let S^r denote concatenation of r occurrences of S. The *weight* of S, $W(S)$, is the number of "1's" in S. The *order* of a sequence $S = [s_0 s_1 \cdots s_{k-1}]$ is the least integer n such that the n-tuples $V_i = (s_i, s_{i+1}, \cdots, s_{i+n-1})$, $0 \le i \le k-1$, with subscripts taken modulo k, are all distinct. Such sequences can be viewed as k-cycles from a feedback shift-register of n-stages, where the n-tuples V_i are successive *states* of the register (or of the sequence). Two sequences S_1 and S_2 are said to be *equivalent*, $S_1 \simeq S_2$, if one is a cyclic shift of the other. The *complement* $\mathbf{c}S$ and the *reverse* $\mathbf{r}S$ of a string $S = s_0 s_1 \cdots s_{k-1}$ are defined by $\mathbf{c}S = \overline{s}_0 \overline{s}_1 \cdots \overline{s}_{k-1}$, where $\overline{s}_i$ is the binary complement of s_i, and $\mathbf{r}S = s_{k-1} \cdots s_1 s_0$. Note that the operators $\mathbf{c}$ and $\mathbf{r}$ commute. S is called a *CR-sequence* if $\mathbf{c}S \simeq \mathbf{r}S$, or equivalently $\mathbf{cr}S \simeq S$. A sequence S of length 2^n and order n is called a *de Bruijn sequence*. Note, that each of the possible 2^n n-tuples appears exactly once as a state of S. The set of all de Bruijn sequences of order n will be denoted by $DS(n)$. It is well known [3] that the cardinality of $DS(n)$ is $2^{2^{n-1}-n}$.

Every sequence $S = [s_0 s_1 \cdots s_{k-1}]$ satisfies a linear recursion of degree $m \le k$

$$s_{i+m} + \sum_{j=1}^{m} a_j s_{i+m-j} = 0, \qquad \text{for all} \quad i \ge 0. \tag{1}$$

In terms of a *shift operator* $\boldsymbol{E}$, defined by

$$\boldsymbol{E}[s_0, s_1, \cdots, s_{k-1}] = \boldsymbol{E}[s_1, \cdots, s_{k-1}, s_0] \tag{2}$$

i.e., $\boldsymbol{E}s_i = s_{i+1}$, for all $i \ge 0$ the linear recursion takes the form

$$\left(\boldsymbol{E}^m + \sum_{j=1}^{m} a_j \boldsymbol{E}^{m-j}\right) s_i = 0, \qquad \text{for all} \quad i \ge 0. \tag{3}$$

Let $f(\boldsymbol{E})s_i = 0$, $i \ge 0$, be the linear recursion of least degree satisfied by S. Then the *complexity* $C(S)$ of S is defined as the degree of $f(\boldsymbol{E})$ viewed as a polynomial in $\boldsymbol{E}$. An efficient algorithm to compute the complexity of sequences with length 2^n was given by Games and Chan [12].

The input to the Games and Chan algorithm is a sequence S of length $l(S) = 2^n$. If $S \ne 0^{2^n}$, the complexity of S is computed recursively as follows. Initially, set $c_n = 0$ and $A_n = S$. At a typical step of the algorithm the left half of A_m, $L(A_m) = [a_0 \cdots a_{2^{m-1}-1}]$, is added to the right half, $R(A_m) = [a_{2^{m-1}} \cdots a_{2^m-1}]$, the result being a sequence B_m, of length 2^{m-1}. If $B_m = 0^{2^{m-1}}$, A_m is replaced by $A_{m-1} = L(A_m)$ and the complexity is left unchanged, i.e., $c_{m-1} = c_m$. If $B_m \ne 0^{2^{m-1}}$, A_m is replaced by $A_{m-1} = B_m$ and c_m is replaced by $c_{m-1} = c_m + 2^{m-1}$. The complexity of S is given by $C(S) = c_0 + 1$.

Chan, Games and Key [2] proved that the complexity of a de Bruijn sequence of length 2^n, $n \ge 3$, is bounded below by $2^{n-1} + n$ and above by $2^n - 1$. They gave complete complexity distribution tables for de Bruijn sequences of order n, where $4 \le n \le 6$. They had four conjectures on $\gamma(c, n)$, the number of de Bruijn sequences of order n and complexity c.

Description:

- C.1) For $n \ge 3$, $\gamma(2^{n-1} + n + 1, n) = 0$.
- C.2) The lower bound for the complexities of de Bruijn sequences of order n is attained for all n, that is, $\gamma(2^{n-1} + n, n) > 0$.
- C.3) At least half of the de Bruijn sequences of order n have maximum complexity, that is, $\gamma(2^n - 1, n) \ge 2^{2^{n-1}-n-1}$.
- C.4) For $n \ge 4$, $\gamma(c, n) \equiv 0 (\mathrm{mod}\, 4)$.

Games [10] proved C.1), i.e., he showed that there are no de Bruijn sequences of order n and complexity $2^{n-1} + n + 1$. Etzion and Lempel [8] showed how to construct de Bruijn sequences of order n and minimal complexity $2^{n-1} + n$, thus proving C.2). For most of the other values of c between $2^{n-1} + n$ and $2^n - 1$, it is not known whether $\gamma(c, n) > 0$, although we conjecture.

Conjecture 1: For $n \ge 3$ and for all c in the range between $2^{n-1} + n$ and $2^n - 1$, except for $2^{n-1} + n + 1$, there exist de Bruijn sequences of order n and complexity c, that is, $\gamma(c, n) > 0$.

C.3) is an open problem and it is not known if it is true for $n = 7$, since it is not possible yet to make an exhaustive computer search on all the de Bruijn sequences of order seven and maximal complexity 127. Etzion [5] found that C.4) is not true in general, but there are values of c and n for which $\gamma(c, n) \equiv 0 (\mathrm{mod}\, 4)$.

- V.1) For even $n \ge 4$, $\gamma(c, n) \equiv 0 (\mathrm{mod}\, 4)$ [2], [9].
- V.2) For even c and $n \ge 3$, $\gamma(c, n) \equiv 0 (\mathrm{mod}\, 4)$ [7].
- V.3) For all $n \ge 4$ and c in the range $2^{n-1} + n \le c \le 2^{n-1} + 2^{n-2}$, $\gamma(c, n) \equiv 0 (\mathrm{mod}\, 4)$ [4].
- V.4) For all $n \ge 4$, $\gamma(2^n - 1, n) \equiv 0 (\mathrm{mod}\, 8)$ [7].
- V.5) For all $k \ge 3$, $\gamma(2^{2k} - 1, 2k) \equiv 0 (\mathrm{mod}\, 16)$ [7].

For later reference, we also state the following known facts.

Fact 1 [2], [9]: If S is a sequence whose length is a power of two then $C(S) = c$ if and only if $(\boldsymbol{E} + \mathbf{1})^{c-1} S = 1^{l(S)}$.

Fact 2 [8]: Let $F(n)$ denote a maximal set of pairwise inequivalent sequences of period $2^{\lfloor \log n \rfloor + 1}$ and complexity $n + 1$. Every $S \in F(n)$ satisfies $(\boldsymbol{E} + \mathbf{1})^n S = 1^{l(S)}$, the cardinality of $F(n)$ is $|F(n)| = 2^{n - \lfloor \log n \rfloor - 1}$ and each of the 2^n binary n-tuples appears exactly once in one of the members of $F(n)$.

An interesting application of de Bruijn sequences with minimal complexity and the sequences of $F(n)$, in the design of perfect maps, can be found in [6].

Fact 3: Let S_1 and S_2 be two sequences of length 2^m. If $C(S_1) < C(S_2)$ then $C(S_1 + S_2) = C(S_2)$.

The main objective of this paper is to survey the known knowledge in this area and to present some new related results. In Section II we will give some new results on the distribution of complexities of de Bruijn sequences. In Section III we modify the construction in [8] to obtain a sufficient condition for the existence of de Bruijn sequences of order n with complexity $2^{n-1} + n + \sum_{i=\lfloor \log n \rfloor + 1}^{n-3} \alpha_i 2^i$ for $n \ge 9$ and any choice of $\alpha_i \in \{0, 1\}$. In Section IV we will prove that the sufficient condition of Section III may be realized, and show that the attained values of complexities for de Bruijn sequences are dense.

II. Complexity Distribution

For $n = 7$ it is difficult to compute $\gamma(c, 7)$ for every c, since there are 2^{57} de Bruijn sequences of order seven. By using computer search we found that $\gamma(71, 7) = 477240$, $\gamma(73, 7) = 688$, $\gamma(74, 7) = 696$, $\gamma(75, 7) = 5760$, and $\gamma(76, 7) = 1232$.

Etzion and Lempel [7] defined two operators on every $S \in DS(n)$. $\mathbf{z}S$ (respectively, $\mathbf{u}S$) denotes the sequence obtained from S, by interchanging the positions of the unique runs of n and $n-2$ "0's" (respectively, "1's"). It is readily verified that $\mathbf{z}S$, $\mathbf{u}S \in DS(n)$, and it was proved [7] that $C(S) = 2^n - 1$ if and only if $C(\mathbf{z}S) = 2^n - 1$ and $C(\mathbf{u}S) = 2^n - 1$. They defined the group $G = \{\mathbf{e}, \mathbf{r}, \mathbf{z}, \mathbf{u}, \mathbf{rz}, \mathbf{ru}, \mathbf{zu}, \mathbf{rzu}\}$, where $\mathbf{e}$ is the identity operator, and proved that if $k \ge 3$ and $S \in DS(2k)$ then $(G \cup G\mathbf{c})S$ consists of 16 pairwise inequivalent de Bruijn sequences of order n. For odd $n \ge 5$ and for $S \in DS(n)$, we can conclude from [7] that one of the following cases holds.

Case 1: $(G \cup G\mathbf{c})S$ consists of sixteen pairwise inequivalent de Bruijn sequences, each of them is not a CR-sequence.

Case 2: $GS = G\mathbf{c}S$ consists of eight pairwise de Bruijn sequences, four of them are CR-sequences.

Games [11] mentioned the following lemma on $\delta(c, n)$, the number of de Bruijn CR-sequences of order n and complexity c.

Lemma 1: For $n \geq 5$, $\delta(2^n - 1, n) \equiv 0 (\mathrm{mod}\, 2^{(n+1)/2})$.

Now, we can extend V.4) and V.5) to the following theorem.

Theorem 1: For $n \geq 5$, $\gamma(2^n - 1, n) \equiv 0 (\mathrm{mod}\, 16)$.

Proof: From Lemma 1 we have that the number of sequences which belong to Case 2 is congruent to 0 modulo $2 \cdot 2^{(n+1)/2}$. Together with Case 1 and V.5) we have that for $n \geq 5$, $\gamma(2^n - 1, n) \equiv 0 (\mathrm{mod}\, 16)$.

Q.E.D.

III. Sufficient Conditions for the Existence of De Bruijn Sequences with low Complexity

Although we did not prove Conjecture 1, we were able to prove that the complexity distribution of de Bruijn sequences cannot be too sparse. In this section we derive a sufficient condition for the existence of de Bruijn sequences of order n with complexity $2^{n-1} + n + \sum_{i=\lfloor \log\, n \rfloor + 1}^{n-3} \alpha_i 2^i$ for any selection of $\alpha_i \in \{0, 1\}$.

The *companion* U' of a state $U = (u_1, u_2, \cdots, u_{n-1}, u_n)$ is defined by $U' = (u_1, u_2, \cdots, u_{n-1}, \overline{u}_n)$. Two sequences, S_1 and S_2 are said to be *adjacent* if they are state-disjoint and there exists a state U on S_1 whose companion U' is on S_2. The following lemma is a well-known observation [13].

Lemma 2: Two adjacent sequences S_1 and S_2, with U on S_1 and U' on S_2, are joined into a single sequence when the predecessors of U and U' are interchanged.

The following theorem is a straightforward generalization of the results from [8].

Theorem 2: If the sufficient condition, stated below, holds for a given n then there exists a de Bruijn sequence of order n with complexity $2^{n-1} + n + g$.

A. The Sufficient Condition

Consider a set F of 2^r sequences of length 2^{n-r}, which contains all the binary n-tuples as states. Then it is possible to choose one state in each of the sequences of F, designated as the first state of the sequence, and it is possible to arrange the members of F in pairs $P_i = (A_i, B_i)$, $1 \leq i \leq 2^{r-1}$, so that properties p.1)–p.4) hold.

- p.1) For each pair P_i, the first state of A_i is the companion of the first state of B_i.
- p.2) For each i, $A_i + B_i = A_1 + B_1$, where the sum of the sequences is their bitwise sum.
- p.3) $C(A_1 + B_1) = n + g$.
- p.4) The graph (V, E), where $V = \{v_i | 1 \leq i \leq 2^{r-1}\}$ and $\{v_i, v_j\} \in E$ if and only if A_i and A_j have a pair of companion states in the same position (relative to their respective first states), is a connected graph.

The proof of Theorem 2 is the same as the proof of Theorem 2 in [8].

Example 1: Let $n = 7$, we take sequences of length 16. These sequences are listed below in four pairs that satisfy p.1)–p.4). It is easy to verify that this arrangement satisfies p.1)–p.3). To check p.4), let $POC(i, j)$ denote a position in A_i and A_j which implies $\{v_i, v_j\} \in E$ according to p.4). It is easy to see now that the three edges implied by $POC(1, 2) = 4$, $POC(1, 3) = 7$, $POC(1, 4) = 10$ form a tree of (V, E), thus validating p.4)

$$(A_1, B_2) = ([1000000010110100], [1000001100011101])$$
$$(A_2, B_2) = ([1111000010010111], [1111001100111110])$$
$$(A_3, B_3) = ([0111001010110010], [0111000100011011])$$
$$(A_4, B_4) = ([0101000011110100], [0101001101011101]).$$

The join is performed by applying Lemma 2 to the seven states

$$(0000101), (0101101), (1101001), (0011001)$$
$$(1000111), (0111011), (1000000)$$

and their companions. The generated de Bruijn sequence is one of the combinations to form a de Bruijn sequence with complexity 74 from those eight sequences.

If $g = 0$ and $F = F(n)$ then this sufficient condition becomes the same as the one for constructing de Bruijn sequences with minimal complexity [8]. By making some modifications in the sufficient condition of Theorem 2, we can obtain a stronger result.

Theorem 3: If the sufficient condition, stated below, holds for a given n then there exists a de Bruijn sequence S of order n with $C(S) = 2^{n-1} + n + \sum_{i=\lfloor \log\, n \rfloor + 1}^{n-3} \alpha_i 2^i$ for any selection of $\alpha_i \in \{0, 1\}$.

B. The Sufficient Condition

Consider a set $F(n)$ as defined in Fact 2 and let $k = 2^{n - \lfloor \log\, n \rfloor - 3}$. Then it is possible to choose one state (of size n) in each of the sequences $F(n)$, designated as the first state of the sequence, and it is possible to arrange the members of $F(n)$ in pairs $P_i = (A_i, B_i)$, $1 \leq i \leq 2k$, so that properties q.1)–q.8) hold.

- q.1) For each pair P_i, $1 \leq i \leq 2k$, the first state of A_i is the companion of the first state of B_i.
- q.2) For each i, $1 \leq i \leq 2k$, $A_i + B_i = A_1 + B_1$.
- q.3) $C(A_1 + B_1) = n$.
- q.4) For each i, $1 \leq i \leq k$, $A_i + A_{i+k} = A_1 + A_{k+1}$.
- q.5) $C(A_1 + A_{k+1}) = n$.
- q.6) For each i, $1 \leq i \leq k$, $C(A_i + B_{k+i}) < n$.
- q.7) For each i, $1 \leq i \leq k$, A_i and A_{k+i} have a pair of companion states in position $2^{\lfloor \log (n-1) \rfloor} + 1$.
- q.8) The graph $(V(n), E(n))$, where $V(n) = \{v_i | 1 \leq i \leq k\}$ and $\{v_i, v_j\} \in E(n)$ iff A_i and A_j have a pair of companion states in the same position is a connected graph.

Example 2: Let $n = 9$. By Fact 2, $|F(n)| = 32$, so that there are 32 sequences of length 16 with complexity ten. These sequences are listed below in eight pairs that satisfy q.1)–q.8). It is easy to verify that this arrangement satisfies q.1)–q.7). To check q.8), it is easy to see now that the seven edges implied by $POC(1, 2) = 4$, $POC(2, 3) = 1$, $POC(3, 4) = 5$, $POC(2, 5) = 3$, $POC(5, 6) = 6$, $POC(3, 7) = 2$, $POC(5, 8) = 7$ form a tree of $(V(9), E(9))$, thus validating q.8), as shown in (4) at the bottom of the next page.

Before we prove the theorem we remark that properties q.1)–q.8) imply many other properties, e.g., we have the following result.

Lemma 3: For each i, $1 \leq i \leq k$, $A_i + B_{k+i} = A_1 + B_{k+1}$.

Proof: $A_i + B_{k+i} = A_1 + B_{k+1} + A_1 + B_{k+1} + A_i + B_{k+i}$ and by q.4) we have that $A_i + B_{k+i} = A_1 + B_{k+1} + A_1 + B_{k+1} + A_i + B_{k+i} + A_i + A_{k+i} + A_1 + A_{k+1} = A_1 + B_{k+1} + (A_1 + A_1) + (A_i + A_i) + (A_{k+1} + B_{k+1} + A_{k+i} + B_{k+i}) = A_1 + B_{k+1}$.

Q.E.D.

Proof of Theorem 3: Given an arrangement of a set $F(n)$ that satisfies q.1)–q.8), let $(V(n), T)$ denote a tree of $(V(n), E(n))$. We join the members of $F(n)$ to form a single sequence S by applying Lemma 2 as follows.

First, we form S_1 by joining all the A_i's, $1 \le i \le k$, sequences via the companion pairs that define the edges of $(V(n), T)$. Then, we form S_2 by joining all the B_i's, $1 \le i \le k$, sequences via the corresponding companion pairs whose existence is guaranteed by q.2). We form S_3 by joining all the A_i's, $k+1 \le i \le 2k$, sequences via the companion pairs whose existence is guaranteed by q.4) and the edges of $(V(n), T)$. Finally, we form S_4 by joining all the B_i's, $k+1 \le i \le 2k$, sequences via the corresponding companion pairs whose existence is guaranteed by q.2) and q.4). We designate the first states of S_1, S_2, S_3, and S_4, to be the first states of A_1, B_1, A_{k+i}, and B_{k+i}, respectively. It is easy to verify that, under this convention, the following holds:

- r.1) Two states occupying the same position in an (A_i, B_i) pair are also located opposite each other in either S_1 and S_2 (respectively, S_3 and S_4).
- r.2) Two states occupying the same position in (A_i, A_{k+i}) (respectively, (B_i, B_{k+i})), are also located opposite each other in either S_1 and S_3 (respectively, S_2 and S_4).
- r.3) The position of each state in S_1 and S_3 (respectively, S_2 and S_4) is congruent to its original A_i-position (respectively, B_i-position) modulo $2^{\lfloor \log n \rfloor + 1}$.

We can write S_1 as $x_1 \hat{S}_1 = x_1 y_1 x_2 y_2 \cdots x_k y_k$, where $l(x_i) = 2^{\lfloor \log (n-1) \rfloor}$, $1 \le i \le k$. The length of the $y_i's$ depend whether n is a power of two or not. If $n = 2^m$ for some m, then $l(y_i) = 3 \cdot 2^{\lfloor \log (n-1) \rfloor}$, $1 \le i \le k$, and otherwise $l(y_i) = 2^{\lfloor \log (n-1) \rfloor}$, $1 \le i \le k$. Similarly, we can write S_3 as $a_1 \hat{S}_3 = a_1 b_1 a_2 b_2 \cdots a_k b_k$, where $l(a_i) = l(x_i)$ and $l(y_i) = l(b_i)$, $1 \le i \le k$. The first states of $\hat{S}_1$ and $\hat{S}_3$ are companions as guaranteed by q.7). As a result, it follows from Lemma 2 that $S = \hat{S}_1 S_2 x_1 \hat{S}_3 S_4 a_1$ is a de Bruijn sequence, and by q.2), q.4), r.2), and r.3)

$$\hat{S}_1 S_2 x_1 + \hat{S}_3 S_4 a_1 \simeq S_1 S_2 + S_3 S_4 = (A_1 + A_{k+1})^{2k}.$$

Also, by q.5) we have that $C(A_1 + A_{k+1}) = n$. Due to this form of S, it follows directly from the Games and Chan algorithm that $C(S) = 2^{n-1} + n$. This is an alternative method for constructing de Bruijn sequences with minimal complexity and the solution if all the α_i's are "0's."

$S_1 S_2$ may be written in the form $x_1 y_1 x_2 y_2 \cdots x_k y_k x_1 z_1 x_2 z_2 \cdots x_k z_k$ as a consequence of q.1), r.1), and r.3). Similarly, $S_3 S_4$ may be written in the form $a_1 b_1 a_2 b_2 \cdots a_k b_k a_1 c_1 a_2 c_2 \cdots a_k c_k$. Now, note that by q.2), q.4), r.2), and r.3), we have that $x_i y_i + a_i b_i = x_i z_i + a_i c_i = A_1 + A_{k+1}$, $1 \le i \le k$ (let $\hat{d} \triangleq A_1 + A_{k+1}$). Similarly, by q.2), r.2), r.3), and Lemma 3, we have that $x_i y_i + a_i c_i = x_i z_i + a_i b_i = A_1 + B_{k+1}$, $1 \le i \le k$ (let $\hat{e} \triangleq A_1 + B_{k+1}$). Finally, define $\hat{f} \triangleq \hat{d} + \hat{e}$. By q.5) we have that $C(\hat{d}) = n$, and by q.7) we have that $C(\hat{e}) < n$, and hence by Fact 3 $C(\hat{f}) = n$. Recall that the sequence

$$S = y_1 x_2 y_2 \cdots x_k y_k x_1 z_1 x_2 z_2 \cdots x_k z_k x_1 b_1 a_2 b_2 \cdots a_k b_k a_1 c_1 a_2 c_2 \cdots a_k c_k a_1$$

is a de Bruijn sequence (note that the first $n-1$ bits of $y_1 x_2$ and $b_1 a_2$ are identical). Now, note that since by q.1) and r.1) the first $n-1$ bits of $a_i b_i$ and $a_i c_i$ are identical, it follows that the sequence

$$S_3' S_4' = a_1 b_1 a_2 u_2 \cdots a_k u_k a_1 c_1 a_2 v_2 \cdots a_k v_k$$

where u_i and v_i are either b_i or c_i, $u_i \ne v_i$, includes the same states as $S_3 S_4$. Therefore,

$$S' = y_1 x_2 y_2 \cdots x_k y_k x_1 z_1 x_2 z_2 \cdots x_k z_k x_1 b_1 a_2 u_2 \cdots a_k u_k a_1 c_1 a_2 v_2 \cdots a_k v_k a_1$$

is a de Bruijn sequence. u_i and v_i should be chosen in such a way that $C(S) = 2^{n-1} + n + \sum_{i=\lfloor \log n \rfloor + 1}^{n-3} \alpha_i 2^i$, $\alpha_i \in \{0, 1\}$ (the procedure for choosing them is given below). We now show how the computation of the complexity of S' with the Games and Chan algorithm should look like. Let $S' = L(S')R(S')$, where $L(S')$ and $R(S')$ are the left and right halves of S', as defined in the Introduction. Define $T_1 = L(S') + R(S')$. To obtain the desired complexity of S' the minimal period of T_1 must be 2^{r+1}. Then, we can write $T_1 = (T_2)^{2^{n-r-2}}$, where $l(T_2) = 2^{r+1}$ and T_2 also have minimal period 2^{r+1}. T_2 is a concatenation of strings of the form $\hat{d}$ and $\hat{e}$. Let $T_2 = L(T_2)R(T_2)$ and define $T_3 = L(T_2) + R(T_2)$. Clearly, T_3 is a concatenation of sequences of the form $\hat{f}$ and $0^{2^{\lfloor \log n \rfloor + 1}}$. Based on these facts we have the following procedure to generate the de Bruijn sequence S' with the desired complexity $2^{n-1} + n + \sum_{i=\lfloor \log n \rfloor + 1}^{n-3} \alpha_i 2^i$, $\alpha_i \in \{0, 1\}$, i.e., we just take the steps of the Games and Chan algorithm backwards, starting with the

$$\begin{aligned}
(A_1, B_1) &= ([00101010\ 10000000], [00101010\ 01111111]) \\
(A_2, B_2) &= ([11011010\ 10001111], [11011010\ 01110000]) \\
(A_3, B_3) &= ([01011010\ 11110000], [01011010\ 00001111]) \\
(A_4, B_4) &= ([10100010\ 11110111], [10100010\ 00001000]) \\
(A_5, B_5) &= ([00111010\ 10010000], [00111010\ 01101111]) \\
(A_6, B_6) &= ([11000110\ 10010011], [11000110\ 01101100]) \\
(A_7, B_7) &= ([10011010\ 11001111], [10011010\ 00110000]) \\
(A_8, B_8) &= ([11000100\ 10010001], [11000100\ 01101110]) \\
(A_9, B_9) &= ([11010101\ 10000000], [11010101\ 01111111]) \\
(A_{10}, B_{10}) &= ([00100101\ 10001111], [00100101\ 01110000]) \\
(A_{11}, B_{11}) &= ([10100101\ 11110000], [10100101\ 00001111]) \\
(A_{12}, B_{12}) &= ([01011101\ 11110111], [01011101\ 00001000]) \\
(A_{13}, B_{13}) &= ([11000101\ 10010000], [11000101\ 01101111]) \\
(A_{14}, B_{14}) &= ([00111001\ 10010011], [00111001\ 01101100]) \\
(A_{15}, B_{15}) &= ([01100101\ 11001111], [01100101\ 00110000]) \\
(A_{16}, B_{16}) &= ([00111011\ 10010001], [00111011\ 01101110])
\end{aligned} \tag{4}$$

sequence $\hat{f}$ whose complexity is n.

$len := 2^{\lfloor \log\ n \rfloor + 1}$;
$j := \lfloor \log\ n \rfloor + 1$;
let r be the largest integer such that $\alpha_r = 1$;
$S := \hat{f}$; $\{S$ is of length $len; C(S) = n\}$
while $len <> 2^r$ do
begin
if $\alpha_j = 1$ then $S := S0^i$
else $S := SS$;
$len := 2 \cdot len$;

$$\left\{ S \text{ is of length } len;\ C(S) = n + \sum_{i=\lfloor \log\ n \rfloor + 1}^{j} \alpha_i 2^i \right\}$$

$j := j + 1$;
end;

$$\left\{ S \text{ is of length } 2^r; C(S) = n + \sum_{i=\lfloor \log\ n \rfloor + 1}^{r-1} \alpha_i 2^i \right\}.$$

Now, let $t = 2^{r - \lfloor \log\ n \rfloor - 1}$ and $S = w_1 w_2 \cdots w_t$, where $l(S) = 2^r$ and $l(w_i) = 2^{\lfloor \log\ n \rfloor + 1}$.

for $i := 1$ to $2^{r - \lfloor \log\ n \rfloor - 1}$ do
if $w_i = 0^{2^{\lfloor \log\ n \rfloor + 1}}$ then $u_i := \hat{d}$
else $\{w_i = \hat{f}\}\ u_i := \hat{e}$;
$S := \hat{d}^t u_1 u_2 \cdots u_t$;

$$\left\{ S \text{ is of length } 2^{r+1};\ C(S) = n + \sum_{i=\lfloor \log\ n \rfloor + 1}^{r} \alpha_i 2^i \right\}$$

for $len := r + 2$ to $n - 2$ do $S := SS$;

$$\left\{ S \text{ is of length } len;\ C(S) = n + \sum_{i=\lfloor \log\ n \rfloor + 1}^{r} \alpha_i 2^i \right\}$$

Now, let $S = m_1 m_2 \cdots m_k$. $\{l(S) = 2^{n-2}; l(m_i) = 2^{\lfloor \log\ n \rfloor + 1}\}$.
for $i := 1$ to k do
if $m_i = \hat{d}$ then begin $u_i := b_i$; $v_i := c_i$ end
else $\{p_i = \hat{e}\}$ begin $u_i := c_i$; $v_i := b_i$ end

Note, that $m_1 = \hat{d}$ and hence $u_1 = b_1$ and $v_1 = c_1$. Now, let

$$S_3' S_4' = a_1 b_1 a_2 u_2 \cdots a_k u_k a_1 c_1 a_2 v_2 \cdots a_k v_k$$

and by the Games and Chan algorithm

$$S' = y_1 x_2 y_2 \cdots x_k y_k x_1 z_1 x_2 z_2 \cdots x_k z_k x_1 b_1 a_2 u_2 \cdots a_k u_k a_1 c_1 a_2 v_2 \cdots a_k v_k a_1$$

is a de Bruijn sequence with the desired linear complexity.

Q.E.D.

IV. Validity of the Sufficient Condition

In this section we show that for every $n \geq 9$, there exists a valid set $F(n)$. That is, there exists an arrangement for the set $F(n)$ that satisfies q.1)–q.8). This is done by demonstrating a recursive construction which is similar to the construction in [8]. For the recursive construction we need the ***D*-morphism** operator $\boldsymbol{D = E+1}$ defined in [16] for de Bruijn graphs and its inverse $\boldsymbol{D}^{-1}$ as defined by Lempel [16]. When applied to a sequence, $\boldsymbol{D}$ can be viewed as being equivalent to the operator $\boldsymbol{E} + 1$. That is, for $S = [s_0, s_1, s_2, \cdots, s_{t-1}]$

$$\begin{aligned}\boldsymbol{D}S &= (\boldsymbol{E} + 1)S \\ &= [s_0 + s_1, s_1 + s_2, \cdots, s_{t-2} + s_{t-1}, s_{t-1} + s_0].\end{aligned}$$

When applied to individual states, $\boldsymbol{D}$ effects a two-to-one map from B^n (the set of all binary n-tuples) onto B^{n-1}. Thus, $\boldsymbol{D}^{-1}$ actually consists of two maps $\boldsymbol{D}_0^{-1}$ and $\boldsymbol{D}_1^{-1}$ which, when applied to a sequence $S = [s_0, s_1, s_2, \cdots, s_{t-1}]$ of even weight yield a pair of complementary sequences

$$\boldsymbol{D}_0^{-1} S = \left[0, s_0, s_0 + s_1, \cdots, \sum_{i=0}^{t-2} s_i\right].$$

$$\boldsymbol{D}_1^{-1} S = \left[1, 1 + s_0, 1 + s_0 + s_1, \cdots, 1 + \sum_{i=0}^{t-2} s_i\right].$$

while when $W(S)$ is odd, the images under $\boldsymbol{D}_0^{-1}$ and $\boldsymbol{D}_1^{-1}$ are self-dual and are cyclic shifts of one another

$$\boldsymbol{D}_0^{-1} S = \left[0, s_0, s_0 + s_1, \cdots, \sum_{i=0}^{t-2} s_i, 1, 1 + s_0, 1 + s_0 + s_1, \cdots, 1 + \sum_{i=0}^{t-2} s_i\right]$$

and

$$\boldsymbol{D}_1^{-1} S = \left[1, 1 + s_0, 1 + s_0 + s_1, \cdots 1 + \sum_{i=0}^{t-2} s_i, 0, s_0, s_0 + s_1, \cdots, \sum_{i=0}^{t-2} s_i\right].$$

It also follows from the definition of $\boldsymbol{D}$ (see [2]) that if $f(\boldsymbol{E})S = 0$ and $\boldsymbol{E} + 1$ is a factor of $f(\boldsymbol{E})$ then $C(\boldsymbol{D}S) = C(S) - 1$. Hence, by Fact 1, we obtain the following.

Fact 4: Let $l(S)$ be a power of two. Then $C(\boldsymbol{D}^{-1}S) = C(S)+1$, where $\boldsymbol{D}^{-1}$ stands for either $\boldsymbol{D}_0^{-1}$ or $\boldsymbol{D}_1^{-1}$.

We start with a set $F(n)$, $n = 2^m$, $m \geq 3$, arranged to satisfy the sufficient condition of Theorem 2 to obtain de Bruijn sequences with minimal complexity $2^{n-1} + n$. The existence of such valid set was proven in [8]. This set does not have to satisfy the sufficient condition of Theorem 3. The next step is to find a valid set $F(n+1)$ which satisfies the sufficient condition of Theorem 3. This is done with the following construction.

Construction 1: Given a positive integer $n = 2^m$, $m \geq 3$ with the set $F(n)$ which satisfies the sufficient condition of Theorem 2 with $g = 0$, construct a valid set $F(n+1)$ by applying $F(n+1) = \boldsymbol{D}_0^{-1}F(n) \cup \boldsymbol{D}_1^{-1}F(n)$, where $\boldsymbol{D}_i^{-1}F(n) = \cup_{S \in F(n)} \boldsymbol{D}_i^{-1}S$, $i = 0, 1$.

Lemma 4 [8]: The set $F(n+1)$ obtained via Construction 1 satisfies the defining properties given in Fact 2.

To continue with the arrangement of the set $F(n)$, we need the following fact and lemmas.

Fact 5 [2]: Let S be a sequence of length 2^{m+1}. Then $C(S) = 2^m + 1$ if and only if $S = [X\overline{X}]$ for some X.

Lemma 5 [8]: If the first states of S_1 and S_2 are companions, then the first states of $\boldsymbol{D}_i^{-1}S_1$ and $\boldsymbol{D}_i^{-1}S_2$ are companions, $i \in \{0, 1\}$.

Lemma 6 [8]: The mth state of $\boldsymbol{D}_i^{-1}S_1$ is the companion of either the mth state of $\boldsymbol{D}_i^{-1}S_2$ or the mth state of $\boldsymbol{D}_i^{i-1}S_2$, $i \in \{0, 1\}$ if and only if the mth states of S_1 and S_2 are companions.

Lemma 7 [8]: If $l(S_1) = l(S_2)$ and $W(S_1) \equiv W(S_2) (\mathrm{mod}\, 2)$, then $\boldsymbol{D}_0^{-1}(S_1 + S_2) = \boldsymbol{D}_i^{-1}S_1 + \boldsymbol{D}_i^{-1}S_2$, $i \in \{0, 1\}$.

Now, given the pair $P_i = (A_i, B_i)$ of $F(n)$ it follows by q.1) and Lemma 5, that the pairs $P_{i0} = (\boldsymbol{D}_0^{-1}A_i, \boldsymbol{D}_0^{-1}B_i)$ and $P_{i1} = (\boldsymbol{D}_1^{-1}A_i, \boldsymbol{D}_1^{-1}B_i)$ of $F(n+1)$ satisfy q.1). By Lemma 7, Construction 1 preserves q.2) and, by Fact 4 and Lemma 7, Construction 1 preserves q.3). Now, for either $j = 0$ or $j = 1$ let $P_i' = (\boldsymbol{D}_j^{-1}A_i, \boldsymbol{D}_j^{-1}B_i)$ and $P_{k+i}' = (\boldsymbol{D}_{1-j}^{-1}B_i, \boldsymbol{D}_{1-j}^{-1}A_i)$. By Fact 5, $A_i = [X\overline{X}]$ for some X and by q.1), $B_i = [X'\overline{X}']$. Therefore, $\boldsymbol{D}_j^{-1}A_i + \boldsymbol{D}_{1-j}^{-1}B_i = [1^{2^{\lfloor \log n \rfloor}}0^{2^{\lfloor \log n \rfloor}}]$ and since $C([1^{2^{\lfloor \log n \rfloor}}0^{2^{\lfloor \log n \rfloor}}]) = 2^{\lfloor \log n \rfloor} + 1$, it follows that q.4), q.5), and q.7) are satisfied. $\boldsymbol{D}_0^{-1}A_i + \boldsymbol{D}_1^{-1}A_i = \boldsymbol{D}_0^{-1}B_i + D_1^{-1}B_i = [1^{2^{\lfloor \log n \rfloor + 1}}]$ and hence q.6) is satisfied. By Lemma 6, q.8), and Construction 1, the existence of a tree $T(n)$ for $F(n)$, implies the existence of a corresponding pair of trees $T_1(n+1)$ and $T_2(n+1)$ isomorphic to $T(n)$. In the construction of these two trees we will decide for each i, which j is taken for $P_i' = (\boldsymbol{D}_j^{-1}A_i, \boldsymbol{D}_j^{-1}B_i)$. Thus, q.8) is satisfied.

Construction 2: Given a positive integer $n \geq 9$ which is not a power of 2, construct a valid set $F(n+1)$ by repeatedly applying the recursion $F(t+1) = \boldsymbol{D}_0^{-1}F(t) \cup \boldsymbol{D}_1^{-1}F(t)$, beginning with the valid set $F(2^{\lfloor \log n \rfloor} + 1)$ obtained by Construction 1.

The proof that the set $F(n+1)$ of Construction 2 satisfies q.1)–q.8) is similar to the one in [8], with some additional proofs for the properties not needed in [8] as given after Construction 1.

Theorem 4: For every $n \geq 4$ there exists a de Bruijn sequence of order n and complexity $2^{n-1} + n + \sum_{i=\lfloor \log n \rfloor + 1}^{n-3} \alpha_i 2^i$ for any selected $\alpha_i \in \{0, 1\}$.

The following two theorems, from [2] and [4], respectively, are useful for proving our main result.

Theorem 5: $\gamma(2^{n-1} + c, n) \geq 2\gamma(c, n-1)$.

Theorem 6: $\gamma(2^{n-1} + 2^{n-2}, n) > 0$.

Combining Theorems 4–6 and computer search which shows that conjecture 1 is true for $n = 7$ and $n = 8$, we have the following theorem.

Theorem 7: For any given t, $2^{n-1} + n \leq t \leq 2^n - 2^{\lfloor \log n \rfloor + 1}$ there exists a de Bruijn sequence of order n and linear complexity c, where $t \leq c < t + 2^{\lfloor \log n \rfloor + 1}$.

V. Conclusions

We have proven that in each interval of length $2^{\lfloor \log n \rfloor + 1}$, between $2^{n-1} + n$ and $2^n - 1$, there is some integer c and de Bruijn sequences of length 2^n and linear complexity c. It is still desired to prove that for every c between $2^{n-1} + n$ and $2^n - 1$, $c \neq 2^{n-1} + n + 1$ there is a de Bruijn sequences of length 2^n and linear complexity c.

Another direction in this area is to find the complexity distribution of nonbinary de Bruijn sequences. A pioneer work in this direction can be found in [1]. More results can be found in [14].

Acknowledgment

The author thanks one of the referees for his valuable suggestions.

References

[1] S. R. Blackburn, T. Etzion, and K. G. Paterson, "Permutation polynomials, de Bruijn sequences, and linear complexity," *J. Combin. Theory*, Ser. A, vol. 76, pp. 55–82, 1996.

[2] A. H. Chan, R. A. Games, and E. L. Key, "On the complexities of de Bruijn sequences," *J. Combin. Theory*, Ser A, vol. 33, pp. 233–246, 1982.

[3] N. G. de Bruijn, "A combinatorial problem," in *Nederl. Akad. Wetench. Proc.*, 1946, vol. 49, pp. 758–764.

[4] T. Etzion, "On the distribution of de Bruijn sequences of low complexity," *J. Combin. Theory*, Ser A, vol. 38, pp. 241–253, 1985.

[5] ——, "On the distribution of de Bruijn CR-sequences," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 422–423, 1986.

[6] ——, "Constructions for perfect maps and pseudorandom arrays," *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 1308–1316, 1988.

[7] T. Etzion and A. Lempel, "On the distribution of de Bruijn sequences of given complexity," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 611–614, 1984.

[8] ——, "Construction of de Bruijn sequences of minimal complexity," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 705–709, 1984.

[9] H. M. Fredricksen, "A survey of full length nonlinear shift register cycle algorithms," *SIAM Rev.*, vol. 24, pp. 195–221, 1982.

[10] R. A. Games, "There are no de Bruijn sequences of span n with complexity $2^{n-1} + n + 1$," *J. Combin. Theory*, Ser A, vol. 34, pp. 248–251, 1983.

[11] ——, "A generalized recursive construction for de Bruijn sequences," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 843–850, 1983.

[12] R. A. Games and A. H. Chan, "A fast algorithm for determining the complexity of a binary sequence with a period 2^n," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 144–146, 1983.

[13] S. W. Golomb, *Shift Register Sequences.* Laguna Hills, CA: Aegean, 1982.

[14] P. A. Hines, "Characterizing the linear complexity of span 1 de Bruijn sequences over finite fields," *J. Combin. Theory*, Ser A, vol. 81, pp. 140–148, 1998.

[15] E. L. Key, "An analysis of structure and complexity of nonlinear binary sequence generators," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 732–736, 1976.

[16] A. Lempel, "On a homomorphism of the de Bruijn graph and its applications to the design of feedback shift registers," *IEEE Trans. Computers*, vol. C-19, pp. 1204–1209, 1970.

[17] ——, "Cryptology in transition," *Computing Surveys*, vol. 11, pp. 285–303, 1979.

Every Binary $(2^m - 2, 2^{2^m - 2 - m}, 3)$ Code Can Be Lengthened to Form a Perfect Code of Length $2^m - 1$

Tim Blackmore

***Abstract*—We answer a problem posed by Etzion and Vardy by showing that a binary code of length $N = 2^m - 2$ with 2^{N-m} codewords and minimum distance three can always be lengthened to form a perfect code of length $2^m - 1$.**

***Index Terms*—Lengthened code, perfect code.**

I. Introduction

We write $\mathbb{F}_2$ for the binary field and $\mathbb{F}_2^n$ for the set of binary vectors of length n. Throughout, by distance we mean Hamming distance. For $v, w \in \mathbb{F}_2^n$ and a code $C \subseteq \mathbb{F}_2^n$ we write $d(v, w)$ for the distance

Manuscript received July 13, 1998. This work was supported by the United Kingdom Engineering and Physical Sciences Research Council under Grant L88764.

The author is with the Algebraic Coding Research Group, Centre for Communications Research, University of Bristol, Bristol, U.K.

Communicated by A. Barg, Associate Editor for Coding Theory.

Publisher Item Identifier S 0018-9448(99)01412-1.