Construction of de Bruijn Sequences of Minimal Complexity

TUVI ETZION AND ABRAHAM LEMPEL, FELLOW, IEEE

***Abstract*—It is well known that the linear complexity of a de Bruijn sequence S of length 2^n is bounded below by $2^{n-1} + n$ for $n \geq 3$. It is shown that this lower bound is attainable for all n.**

I. Introduction

The linear complexity $C(S)$ of a sequence S is one of the measures of its predictability—S is completely determined by $2C(S)$ consecutive bits of S. Although high complexity does not necessarily mean low predictability, the converse is always true: low complexity implies high predictability.

We investigate the linear-complexity distribution of binary de Bruijn sequences of given length. Chan, Games, and Key [1] showed that if S is a de Bruijn sequence of order n, then $2^{n-1} + n \leq C(S) \leq 2^n - 1$. They also showed that the lower bound $2^{n-1} + n$ is achievable for $3 \leq n \leq 6$, and they conjectured that it can be attained for every n. Etzion [2] showed that this conjecture is true for $7 \leq n \leq 11$. In this paper we demonstrate that for each $n \geq 3$ there exists a de Bruijn sequence S of length 2^n with $C(S) = 2^{n-1} + n$, thus proving the conjecture by Chan *et al.*

In Section II we present a brief background and some definitions and notation. In Section III we derive a sufficient condition for the existence of a de Bruijn sequence S of order n with $C(S) = 2^{n-1} + n$. In Section IV we prove that this condition is satisfied when $n = 2^m$, $m \geq 3$, and in Section V we prove that this condition holds for every $n \geq 8$.

II. Definitions and Notation

In the sequel we shall need the following definitions and notation.

Let $s_1, s_2, \cdots,$ denote a string of binary digits. A cyclic, or closed, string is called a *sequence* and is denoted by $S = [s_0, s_1, \cdots, s_{k-1}]$, where $k = \text{length}\,(S)$ is the *length* of 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 $U_i = (s_i, s_{i+1}, \cdots, s_{i+n-1})$, $0 \leq i \leq k-1$, 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 U_i are successive *states* of the register (and of the sequence).

A sequence S with $\text{length}(S) = 2^n$ and order n is called a *de Bruijn sequence*. Note that each of the 2^n possible n-tuples appears exactly once as a state of every de Bruijn sequence.

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

$$s_{i+m} + \sum_{j=1}^{m} a_j s_{i+m-j} = 0, \qquad i \geq 0,$$

where m, the degree of the recursion, is less than or equal to length (S).

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

$$\boldsymbol{E}[s_0, s_1, \cdots, s_{k-1}] = [s_1, s_2, \cdots, s_{k-1}, s_0],$$

the linear recursion takes the form

$$f(\boldsymbol{E})S = \left(\boldsymbol{E}^m + \sum_{j=1}^{m} a_j \boldsymbol{E}^{m-j}\right) S = 0^k,$$

where b^k denotes a sequence of k b's.

The *period* of a sequence S is the least integer p such that $\boldsymbol{E}^p S = S$.

The (linear) *complexity* $C(S)$ of S is defined as the least integer m for which there exists a polynomial $f(\boldsymbol{E})$ of degree m such that $f(\boldsymbol{E})S = 0^{\text{length}(S)}$.

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

Fact 1: If S is a sequence whose length is a power of 2, then $C(S) = c$ if and only if $(\boldsymbol{E} + 1)^{c-1} S = 1^{\text{length}(S)}$ [1], [3].

Fact 2: Let $F(n)$ denote a maximal set of pairwise inequivalent sequences of period $2^{\lfloor \log n \rfloor + 1}$ and complexity $n + 1$. Then, every $S \in F(n)$ satisfies $(\boldsymbol{E} + 1)^n S = 1^{\text{length}(S)}$, 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 $F(n)$.

Although this fact has been known for some time, we are not aware of a published reference containing a proof. Therefore, we present a short proof below.

Proof: Consider the linear recursion $(\boldsymbol{E} + 1)^{n+1} S = 0^{\text{length}(S)}$. Elspas [4] shows that there are $2^{n - \lfloor \log n \rfloor - 1}$ inequivalent sequences of period $2^{\lfloor \log n \rfloor + 1}$ that satisfy this recursion but do not satisfy the recursion $(\boldsymbol{E} + 1)^n S = 0^{\text{length}(S)}$. That is, there are $2^{n - \lfloor \log n \rfloor - 1}$ inequivalent sequences that satisfy $(\boldsymbol{E} + 1)^n S = 1^{\text{length}(S)}$. Together with Fact 1, this implies that every $S \in F(n)$ satisfies $(\boldsymbol{E} + 1)^n S = 1^{\text{length}(S)}$ and that $|F(n)| = 2^{n - \lfloor \log n \rfloor - 1}$. Since each $(n + 1)$-tuple appears exactly once in one of the sequences satisfying the recursion $(\boldsymbol{E} + 1)^{n+1} S = 0^{\text{length}(S)}$, each n-tuple appears exactly twice in these sequences. Since the

Manuscript received November 28, 1983; revised March 16, 1984.

The authors are with the Computer Science Department, Technion—Israel Institute of Technology, Haifa, Israel.

0018-9448/84/0900-0705$01.00 ©1984 IEEE

sequences satisfying $(E+1)^n S = 0^{\text{length}(S)}$ also satisfy $(E+1)^{n+1} S = 0^{\text{length}(S)}$ and since the former contain each n-tuple exactly once, it follows that each of the 2^n binary n-tuples appears exactly once in one of the members of $F(n)$. Q.E.D.

III. A Sufficient Condition for the Existence of de Bruijn Sequences of Minimal Complexity

In this section we derive a sufficient condition for the existence of de Bruijn sequences of order n with minimal complexity $2^{n-1}+n$.

The *companion* U' of a state $U = (u_1, u_2, \cdots, u_{n-1}, u_n)$ is defined by $U' = (u_1, u_2, \cdots, u_{n-1}, \bar{u}_n)$, where $\bar{x}$ denotes the complement of x.

Two sequences, S_1 and S_2, are *adjacent* if they are state-disjoint and there exists a state U on S_1 whose companion U' is on S_2.

Theorem 1: Two adjacent sequences S_1 and S_2, with U on S_1 on U' on S_2, are joined into a single sequence when the predecessors of U and U' are interchanged [5].

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

The Sufficient Condition: Consider a set $F(n)$ as defined in Fact 2. Then it is possible to choose one state (of size n) in each of the sequences of $F(n)$, designated as the first state of the sequence, and it is possible to arrange the members of $F(n)$ in pairs $P_i = (A_i, B_i)$, $1 \le i \le 2^{n-\lfloor \log n \rfloor -2}$, so that Properties 1 though 4 hold.

Property 1: For each pair P_i, the first state of A_i is the companion of the first state of B_i.

Property 2: For each i, $A_i + B_i = A_1 + B_1$, where the sum of the sequences is their bitwise sum.

Property 3: $C(A_1 + B_1) = n$.

Property 4: The graph $(V(n), E(n))$, where $V(n) = \{v_i | 1 \le i \le 2^{n-\lfloor \log n \rfloor -2}\}$ and $\{v_i, v_j\} \in E(n)$ 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.

Before presenting the proof of Theorem 2, it would be helpful to illustrate the feasibility of the sufficient condition by an example.

Example 1: Let $n = 8$. By Fact 2, $|F(n)| = 16$, so that there are 16 sequences of length 16 with complexity 9. These sequences are listed below in eight pairs that satisfy Property 1 through Property 4. It is easy to verify that this arrangement satisfies Properties 1 through 3. To check Property 4, let $POC(i,j)$ denote a position in A_i and A_j that implies $\{v_i, v_j\} \in E(n)$ according to Property 4. 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 $(V(8), E(8))$, thus validating Property 4.

$$(A_1, B_1) = ([0111111110000000], [0111111010000001])$$
$$(A_2, B_2) = ([0110111110010000], [0110111010010001])$$
$$(A_3, B_3) = ([1110111100010000], [1110111000010001])$$
$$(A_4, B_4) = ([1110011100011000], [1110011000011001])$$
$$(A_5, B_5) = ([0100111110110000], [0100111010110001])$$
$$(A_6, B_6) = ([0100101110110100], [0100101010110101])$$
$$(A_7, B_7) = ([1010111101010000], [1010111001010001])$$
$$(A_8, B_8) = ([0100110110110010], [0100110010110011]).$$

Proof of Theorem 2: Given an arrangement of a set $F(n)$ that satisfies Property 1 through Property 4, 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 Theorem 1 as follows.

First, we form S_1 by joining all the A_i sequences via the companion pairs that define the edges of $(V(n), T)$. Then we form S_2 by joining all the B_i sequences via the corresponding companion pairs whose existence is guaranteed by Property 2. We designate the first states of A_1 and B_1 to be the first states of S_1 and S_2, respectively. It is easy to verify that under this convention the following holds.

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

As a result, it follows that

$$S_1 + S_2 = (A_1 + B_1)^k,$$

where $k = 2^{n-\lfloor \log n \rfloor -2}$ and S^k is a concatenation of k occurrences of S. Also, by Property 3, $C(S_1 + S_2) = n$. Finally, we join S_1 and S_2 via their respective first states to form a de Bruijn sequence

$$S = [S_1, S_2].$$

Because of this form of S and the fact that $C(S_1 + S_2) = n$, it follows directly from the Games and Chan algorithm [6] that $C(S) = 2^{n-1}+n$. Q.E.D.

The authors were unable to resolve whether the sufficient condition is also a necessary one.

IV. Validity of the Sufficient Condition for $n = 2^m$

In this section we show that for every $n = 2^m$, $m \ge 3$, there exists a valid set $F(n)$. That is, there exists an arrangement for the set $F(n)$ that satisfies Property 1 through Property 4. This is done by demonstrating a recursive construction, namely, by showing how a specific valid set for $n = 2^m$, $m \ge 3$, leads to one for $n = 2^{m+1}$.

First, we need some more definitions and facts.

The *weight* $W(S)$ of a string S is the number of ONES in S.

The *complement* $\bar{S}$ of a string $S = s_0 s_1, \cdots, s_{k-1}$ is defined by $\bar{S} = \bar{s}_0 \bar{s}_1, \cdots, \bar{s}_{k-1}$.

A sequence S is *self-dual* if $S = [X\bar{X}]$ for some X.

Fact 3: 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 [1].

Fact 4: Let S be a sequence of length 2^m. Then $C(S) = 2^m$ if and only if $W(S)$ is odd [1].

For the recursive construction we also need the ***D*-morphism** operator $\boldsymbol{D}$ for de Bruijn graphs and its inverse $\boldsymbol{D}^{-1}$ as defined by Lempel [7]. When applied to a sequence, $\boldsymbol{D}$ can be viewed as being equivalent to the operator $\boldsymbol{E} + 1$ (see [1]). That is, for $S = [s_0, s_1, s_2, \cdots, s_{k-1}]$,

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

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 $\boldsymbol{D}_1^{-1}$. When applied to a sequence $S = [s_0, s_1, s_2, \cdots, s_{k-1}]$ of even weight, they yield a pair of complementary sequences:

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

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

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:

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

and

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

It also follows from the definition of $\boldsymbol{D}$ (see [1]) 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 Fact 5.

Fact 5: Let length(S) be a power of 2. Then $C(\boldsymbol{D}^{-1}S) = C(S) + 1$, where $\boldsymbol{D}^{-1}$ stands for either $\boldsymbol{D}_0^{-1}$ or $\boldsymbol{D}_1^{-1}$.

The length of all the sequences considered from now on is assumed to be a power of 2.

Construction 1: The initial valid set for the recursive construction is the set $F(2^3)$ of Example 1. (Note that each of the 16 sequences of Example 1 is self-dual.) Given a valid set $F(2^m)$, all of whose members are self-dual, construct a valid set $F(2^{m+1})$ as follows. Let $Y(m)$ denote the set of 2^{2^m-1} elements consisting of the $2^{2^m-1} - 1$ elements of the form $(0, y_1, y_2, \cdots, y_{2^m-1})$, where at least one of the y_i is not zero, and the element $(1, 1, \cdots, 1)$ of 2^m ones. For each $S = [X\overline{X}] \in F(2^m)$ and for every $Y \in Y(m)$, let S_Y =• $[YX + Y\overline{Y}X + \overline{Y}]$. Then

$$F(2^{m+1}) = \dot{\bigcup_{S \in F(2^m)}} S(m),$$

where

$$S(m) = \dot{\bigcup_{Y \in Y(m)}} S_Y.$$

Lemma 1: The set $F(2^{m+1})$ obtained via Construction 1 satisfies the defining properties given in Fact 2.

Proof: From the form of S_Y, $S \in F(2^m)$, $Y \in Y(m)$, it follows immediately that S_Y is self-dual, its period equals length$(S_Y) = 2^{m+2}$, and $(\boldsymbol{E} + 1)^{2^{m+1}}S_Y = \boldsymbol{E}^{2^{m+1}}S_Y + S_Y = 1^{2^{m+2}}$. Hence, by either Fact 1 or Fact 3, it also follows that $C(S_Y) = 2^{m+1} + 1$ for all $S \in F(2^m)$ and all $Y \in Y(m)$. We proceed to show now that no two members of $F(2^{m+1})$ are equivalent. Assume, to the contrary, that for $S, S^* \in F(2^m)$ and $Y, \hat{Y} \in Y(m)$, $S_Y \simeq S^*_{\hat{Y}}$. That is, $S^*_{\hat{Y}} = \boldsymbol{E}^k S_Y$ for some $0 < k \leq 2^{m+1}$. (Note that k need not exceed half the length and that $k = 0$ implies $Y = \hat{Y}$ and $X = X^*$.) Since equivalence is preserved under shift-and-add, it follows that

$$\begin{aligned}S^2 = (\boldsymbol{E} + 1)^{2^m} S_Y \simeq (\boldsymbol{E} + 1)^{2^m} S^*_{\hat{Y}} &= (\boldsymbol{E} + 1)^{2^m} \boldsymbol{E}^k S_Y\\ &= \boldsymbol{E}^k (\boldsymbol{E} + 1)^{2^m} S_Y = \boldsymbol{E}^k S^2.\end{aligned}$$

Since the period of S equals length$(S) = 2^{m+1}$, k must be a multiple of *length*(S) and thus we must have $k = 2^{m+1}$. This implies $\overline{Y} = \hat{Y}$, which by the definition of $Y(m)$, is impossible. Hence, no two members of $F(2^{m+1})$ can be equivalent. A simple count now shows that $|F(2^{m+1})| = 2^{2^{m+1} - m - 2}$, which, by Fact 2, proves maximality. Q.E.D.

We now proceed to show that the elements of $F(2^{m+1})$ can be arranged so that Property 1 through Property 4 are satisfied. To this end, we add Property 5 (below) to be satisfied only when $n = 2^m$.

For a valid set $F(2^m)$ and for $A_i, A_j \in F(2^m)$, let $d(A_i, A_j)$ denote the first position in which A_i differs from A_j, and let $d_m = \{d(A_i, A_j) | \{v_i, v_j\} \in T(2^m)\}$, where $T(2^m)$ is a tree of $(V(2^m), E(2^m))$ for $F(2^m)$ (see Property 4).

The new property required of a valid set $F(2^m)$ is Property 5.

Property 5: $d_m = \{0, 1, 2, \cdots, 2^m - 2\}$.

Example 2: For the set of Example 1 we have

$$d(A_1, A_2) = 3, \quad d(A_2, A_3) = 0, \quad d(A_3, A_4) = 4,$$
$$d(A_2, A_5) = 2, \quad d(A_5, A_6) = 5, \quad d(A_3, A_7) = 1,$$
$$d(A_5, A_8) = 6.$$

Hence, $d_3 = \{0, 1, 2, 3, 4, 5, 6 = 2^3 - 2\}$.

Lemma 2: If $F(2^{m+1})$ is obtained via Construction 1 from a valid set $F(2^m)$ satisfying Property 1 through Property 5, then $F(2^{m+1})$ can be arranged to satisfy Property 1 through Property 5.

Proof: Consider the pairs $P_i = (A_i, B_i)$ of $F(2^m)$ (see Property 1). Let $A_i = [X_i, \overline{X}_i]$. Since X_i is the first state of A_i, it follows from Property 1 that $B_i = [X_i', \overline{X}_i']$, and thus $A_i + B_i = (0^{2^m-1}10^{2^m-1}1)$. For each P_i of $F(2^m)$ and for every $Y \in Y(m)$, we form the pair $P_{iY} = (A_{iY}, B_{iY})$ as described in Construction 1. Therefore, $A_{iY} + B_{iY} =$

$(0^{2^{m+1}-1}10^{2^{m+1}-1}1)$, which immediately implies Properties 1 through 3 for $F(2^{m+1})$.

To complete the proof we have to show that given a tree $T(2^m)$ for $F(2^m)$ with $d_m = \{0,1,2,\cdots,2^m-2\}$, the graph $(V(2^{m+1}), E(2^{m+1}))$ for $F(2^{m+1})$ is connected and it has a tree $T(2^{m+1})$ with $d_{m+1} = \{0,1,2,\cdots,2^{m+1}-2\}$.

Consider $A_i, A_j \in F(2^m)$ such that $\{v_i, v_j\} \in T(2^m)$. Then since both A_i and A_j are self-dual and $A_j \neq B_i$, it follows from Property 4 that $A_i + A_j = (0^k10^{2^m-1}10^{2^{m-1-k}})$, where $k = d(A_i, A_j)$. Note that $POC(i, j) = k + 1$ (see Example 1).

For every $Y \in Y(m)$ let $G(Y)$ be the subgraph of $(V(2^{m+1}), E(2^{m+1}))$ that is spanned by the vertices v_{iY} corresponding to the sequences A_{iY}, $1 \leq i \leq 2^{2^m-m-2}$. It can be easily verified that $\{v_i, v_j\} \in T(2^m)$ implies $\{v_{iY}, v_{jY}\} \in E(2^{m+1})$ with $d(A_{iY}, A_{jY}) = 2^m + d(A_i, A_j)$ and $POC(iY, jY) = 2^m + POC(i, j)$. Therefore, the set $T_{m+1}(Y) \triangleq \{\{v_{iY}, v_{jY}\} | \{v_i, v_j\} \in T(2^m)\}$ forms a tree of $G(Y)$, isomorphic to $T(2^m)$, and

$$\begin{aligned} d_{m+1}(Y) &\triangleq \left\{d\left(A_{iY}, A_{jY}\right) | \{v_{iY}, v_{jY}\} \in T_{m+1}(Y)\right\} \\ &= \{2^m, 2^m+1, \cdots, 2^{m+1}-2\}. \end{aligned}$$

We now show that the 2^{2^m-1} trees, $T_{m+1}(Y)$, $Y \in Y(m)$ can be embedded in a tree $T(2^{m+1})$ of $(V(2^{m+1}), E(2^{m+1}))$ so that the corresponding set d_{m+1} will include the set $\{0,1,\cdots,2^m-1\}$ along with $d_{m+1}(Y)$. To this end, consider $2^m - 1$ pairs $(A_i, A_j)_k$, one for each $k \in d_m$ such that $A_i, A_j \in F(2^m)$, $\{v_i, v_j\} \in T(2^m)$ and $d(A_i, A_j) = k$. For each such pair $(A_i, A_j)_k$ we form the pair $(A_{iY_k}, A_{jY_{k+1}})$, where $Y_r = 0^r1^{2^m-r}$, $r \in d_m$. As before it is easy to see that $\{v_i, v_j\} \in T(2^m)$ implies $\{v_{iY_k}, v_{jY_{k+1}}\} \in E(2^{m+1})$ with $d(A_{iY_k}, A_{jY_{k+1}}) = k$. Moreover, the union G^*_{m+1} of these $2^m - 1$ members of $E(2^{m+1})$ and of the 2^{2^m-1} trees $T_{m+1}(Y)$, $Y \in Y(m)$ contains no cycle.

For $k = 2^m - 1$ take any $P_i = (A_i, B_i)$ and form the pair $A_{iY}, B_{iY} \in F(2^{m+1})$, where $Y = 0^{2^m-2}11$ and $Y' = 0^{2^m-2}10$. Since $A_i + B_i = (0^{2^m-1}10^{2^m-1}1)$, we have $A_{iY} + B_{iY'} = (0^{2^m-1}10^{2^{m+1}-1}10^{2^m})$, which implies $\{v_{iY}, v_{iY'}\} \in E(2^{m+1})$ with $d(A_{iY}, B_{iY'}) = 2^m - 1$. Adding $\{v_{iY}, v_{iY'}\}$ to G^*_{m+1} creates the cycle-free graph G^* and completes the construction of $d_{m+1} = \{0,1,\cdots,2^{m+1}-2\}$. Since we have used a B-sequence (rather than an A-sequence) with Y', we have to interchange $A_{jY'}$ with $B_{jY'}$ for every $j = 1,\cdots,2^{2^m-m-2}$. That is, the original pairs $(A_{1Y'}, B_{1Y'})$, $(A_{2Y'}, B_{2Y'})$, etc. associated with Y' now become the pairs $(B_{1Y'}, A_{1Y'})$, $(B_{2Y'}, A_{2Y'})$, etc. (It is easy to verify that the graph $G(Y')$ obtained from the $B_{iY'}$ is isomorphic to the one obtained from the $A_{iY'}$.)

So far we have shown that $G(Y')$ and the $G(Y_k)$, $k \in d_m$, form a connected graph G^*. To see that the rest of the $G(Y)$ are all connected to G^*, consider $G(Y_1)$ of G^* and any maximum weight Y such that $G(Y)$ is not in G^*. Then $W(Y + Y_1) = 1$, i.e., $Y + Y_1 = 0^r10^{2^m-r-1}$ for some $1 \leq r < 2^m - 1$. Let $A_i, A_j \in F(2^m)$ be such that $\{v_i, v_j\} \in T(2^m)$ and $d(A_i, A_j) = r$. Then, as before, it follows that $\{v_{iY}, v_{jY_1}\} \in E(2^{m+1})$. This procedure can be repeated until all of the $G(Y)$ are shown to be connected by considering, at each step, the current proven connected piece G^* and a maximum weight Y such that $G(Y)$ is not in G^*. For any such Y there exists a $\hat{Y}$ such that $G(\hat{Y})$ is in G^* and $W(Y + \hat{Y}) = 1$. As before, this proves that $G(Y)$ is also connected to G^*. Q.E.D.

V. Validity of the Sufficient Condition for Every n

In this section we show that for every $n \geq 8$, there exists a valid set $F(n)$. This is done by showing that the validity of the sufficient condition for $2^{\lfloor \log n \rfloor}$ and $2^{\lceil \log n \rceil}$ implies the validity for n.

Construction 2: Given a positive integer $n \geq 8$ that is not a power of 2, construct a valid set $F(n)$ by repeatedly applying the recursion $F(k+1) = \mathbf{D}_0^{-1}F(k) \cup \mathbf{D}_1^{-1}F(k)$, where $\mathbf{D}_i^{-1}F(k) = \cup_{S \in F(k)} \mathbf{D}_i^{-1}S$, $i = 0,1$, beginning with the valid set $F(2^{\lfloor \log n \rfloor})$ obtained by Construction 1.

Lemma 3: The set $F(n)$ obtained via Construction 2 satisfies the defining properties given in Fact 2.

Proof: By definition, $C(S) = 2^{\lfloor \log n \rfloor} + 1$ for each $S \in F(2^{\lfloor \log n \rfloor})$ and, by Fact 5, $C(\mathbf{D}_0^{-1}S) = C(\mathbf{D}_1^{-1}S) = C(S) + 1$. Hence, the complexity of every sequence of $F(k)$ is $k + 1$, $2^{\lfloor \log n \rfloor} \leq k \leq n$. Since no k in the range $2^{\lfloor \log n \rfloor} < k \leq n$ is a power of 2, it follows that the complexity of the sequences of $F(k-1)$ is not a power of 2. Hence, by Fact 4, $W(S)$ is even for each $S \in F(k-1)$, which implies preservation of length and period under $\mathbf{D}_i^{-1}$, $i = 0,1$. That is, the period of S equals *length*$(S) = 2^{\lfloor \log n \rfloor + 1}$ for each $S \in F(k)$, $2^{\lfloor \log n \rfloor} < k \leq n$. We now show that no two members of the obtained set are equivalent. Assume, to the contrary, that for some $S, S^* \in F(k-1)$ and for some $i, j \in \{0,1\}$, $\mathbf{D}_i^{-1}S \simeq \mathbf{D}_j^{-1}S^*$. Then $\mathbf{D}_j^{-1}S^* = \mathbf{E}^m\mathbf{D}_i^{-1}S$ for some $0 < m \leq 2^{\lfloor \log n \rfloor}$ (Note that $m = 0$ implies $i = j$ and $S^* = S$). Since equivalence is preserved under shift-and-add, it follows that

$$\begin{aligned} S = (\mathbf{E}+1)\mathbf{D}_i^{-1}S &\simeq (\mathbf{E}+1)\mathbf{D}_j^{-1}S^* = (\mathbf{E}+1)\mathbf{E}^m\mathbf{D}_i^{-1}S \\ &= \mathbf{E}^m(\mathbf{E}+1)\mathbf{D}_i^{-1}S = \mathbf{E}^mS, \end{aligned}$$

which is impossible because m is smaller then the period of S. As in the proof of Lemma 1, a simple count and Fact 2 imply maximality. Q.E.D.

To continue with the arrangement of the set $F(n)$, we need Lemmas 4 through 6, given below, whose validity can be easily verified.

Lemma 4: If the first states of S_1 and S_2 are companions, then the first states of $\mathbf{D}_i^{-1}S_1$ and $\mathbf{D}_i^{-1}S_2$ are companions, $i \in \{0,1\}$.

Lemma 5: The mth state of $D_i^{-1}S_1$ is the companion of either the mth state of $\mathbf{D}_i^{-1}S_2$ or the mth state of $\mathbf{D}_{1-i}^{-1}S_2$, $i \in \{0,1\}$ if and only if the mth states of S_1 and S_2 are companions.

Lemma 6: If $\text{length}(S_1) = \text{length}(S_2)$ and $W(S_1) \equiv W(S_2) \pmod 2$, then $\mathbf{D}_0^{-1}(S_1 + S_2) = \mathbf{D}_i^{-1}S_1 + \mathbf{D}_i^{-1}S_2$, $i \in \{0,1\}$.

Now, given the pair $P_i = (A_i, B_i)$ of $F(k-1)$, it follows, by Lemma 4, that the pairs $P_{i0} = (\mathbf{D}_0^{-1}A_i, \mathbf{D}_0^{-1}B_i)$

and $P_{i1} = (\boldsymbol{D}_1^{-1}A_i, \boldsymbol{D}_1^{-1}B_i)$ of $F(k)$, $2^{\lfloor \log n \rfloor} < k \leq n$, satisfy Property 1. By Lemma 6, Construction 2 preserves Property 2 and, by Fact 5 and Lemma 6, Construction 2 preserves Property 3.

By Lemma 5 and Construction 2, the existence of a tree $T(k-1)$ for $F(k-1)$, $2^{\lfloor \log n \rfloor} < k \leq n$, implies the existence of a corresponding pair of trees $T_1(k)$ and $T_2(k)$, isomorphic to $T(k-1)$, which form subtrees of the graph $(V(k), E(k))$ for $F(k)$. $T_1(k)$ and $T_2(k)$ are disjoint and together include every element of $V(k)$. Thus, all we need to complete the proof that Construction 2 yields a valid set $F(n)$ is to demonstrate the existence of an edge in $E(k)$ that connects $T_1(k)$ with $T_2(k)$, $2^{\lfloor \log n \rfloor} < k \leq n$. That is, we have to show that it is possible to find two sequences, $\boldsymbol{D}_i^{-1}A_r$ and $\boldsymbol{D}_j^{-1}A_s$, $i, j \in \{0,1\}$, such that $\boldsymbol{D}_i^{-1}A_r$ corresponds to a vertex of $T_1(k)$, $\boldsymbol{D}_j^{-1}A_s$ corresponds to a vertex of $T_2(k)$, and the mth state of $\boldsymbol{D}_i^{-1}A_r$ is the companion of the mth state of $\boldsymbol{D}_j^{-1}A_s$ for some positive integer m. Assume to the contrary, that $k_0 \geq 2^{\lfloor \log n \rfloor} + 1$ is the least integer for which there is no edge in $E(k_0)$ that connects $T_1(k_0)$ with $T_2(k_0)$. Then, by Lemma 5, none of the sets $F(k_0)$ through $F(2^{\lceil \log n \rceil} - 1)$, obtained from $F(k_0 - 1)$ via Construction 2, correspond to a connected graph. Consider the set $F \triangleq F(2^{\lceil \log n \rceil} - 1)$. This set can be partitioned into F_0 and F^*, where $F_0 = \boldsymbol{D}_0^{-(2^{\lfloor \log n \rfloor} - 1)} F(2^{\lfloor \log n \rfloor})$ and $F^* = F - F_0$. Given that the graph (V, E) for F is not connected, it follows that the graph $(\hat{V}, \hat{E})$ for $\hat{F} \triangleq \boldsymbol{D}_1^{-1}F_0 \cup \boldsymbol{D}_0^{-1}F^*$ is not connected. However, it is easy to verify that the set $\hat{F}$ is identical to the set $F(2^{m+1})$ of Construction 1 for $m = \lfloor \log n \rfloor$. Since by Lemma 2 the graph $(\hat{V}, \hat{E})$ of $F(2^{m+1}) = \hat{F}$ is connected (the commutation of some of the (A_i, B_i) pairs in the proof of Lemma 2 does not affect connectivity), we have a contradiction. This invalidated our assumption regarding k_0 and completes the proof of the following result.

Theorem 3: For every $n \geq 3$ there exists a de Bruijn sequence of order n and complexity $2^{n-1} + n$.

Note that although Construction 2 can replace Construction 1, a combination may speed up the process.

References

[1] 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, Nov. 1982.

[2] T. Etzion, "On the distribution of de Bruijn sequences of low complexity," *J. Combin. Theory.* ser. A, to be published.

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

[4] B. Elspas, "Theory of autonomous linear sequential networks," *IRE Trans. Circuit Theory*, vol. CT-6, pp. 45–60, March 1959.

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

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

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

Duadic Codes

JEFFREY S. LEON, JOHN M. MASLEY, AND VERA PLESS

Abstract—**A new family of binary cyclic $(n,(n+1)/2)$ and $(n,(n-1)/2)$ codes are introduced, which include quadratic residue (QR) codes when n is prime. These codes are defined in terms of their idempotent generators, and they exist for all odd $n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ where each p_i is a prime $\equiv \pm 1 \pmod 8$. Dual codes are identified. The minimum odd weight of a duadic $(n,(n+1)/2)$ code satisfies a square root bound. When equality holds in the sharper form of this bound, vectors of minimum weight hold a projective plane. The unique projective plane of order 8 is held by the minimum weight vectors in two inequivalent $(73,37,9)$ duadic codes. All duadic codes of length less than 127 are identified, and the minimum weights of their extensions are given. One of the duadic codes of length 113 has greater minimum weight than the QR code of that length.**

Manuscript received July 30, 1983; revised February 21, 1984. This work was supported in part by the National Science Foundation under Grants MCS-8201311 and MCS-8002261. This paper was presented in part at the American Mathematical Society Meeting, Evanston, IL, November 11, 1983, and in part at the Canadian Mathematical Society Winter Meeting, Montréal, PQ, Canada, December 11, 1983.

The authors are with the Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Box 4348, Chicago, IL 60681.

I. Introduction

WE DEFINE a new infinite family of binary cyclic codes called "duadic" codes. This family includes the binary quadratic residue (QR) codes of prime length. At a given prime length where QR codes exist other duadic codes are often present (see Section IV for examples). However, duadic codes exist for composite lengths also. Further, self-dual extended duadic codes exist at certain lengths where extended QR codes exist but are not self-dual,

0018-9448/84/0900-0709$01.00 ©1984 IEEE