An algorithm for generating all CR sequences in the de Bruijn sequences of length $2^n$ where $n$ is any odd number

Hiroshi Fujisaki$^{1,a)}$

$^{1}$ Graduate School of Natural Science and Technology, Kanazawa University
Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan

$a)$ fujisaki@t.kanazawa-u.ac.jp

Abstract: For the case that $p$ is any prime number, we have already constructed all CR (complement reverse) sequences in the de Bruijn sequences of length $2^{2p+1}$. In this research, with the help of the Dyck language, we characterize CR sequences in the de Bruijn sequences of length $2^{2m+1}$ where $m$ ($\geq 4$) is a non-prime number. In virtue of this characterization, we show that for any odd number $n$, there exist CR sequences in the de Bruijn sequences of length $2^n$, which completely settles the fundamental problem posed by Fredricksen on existence of the CR sequences. Consequently, we establish an algorithm for generating all CR sequences in the de Bruijn sequences of length $2^n$ for any odd $n$.

Key Words: de Bruijn sequences, CR (complement reverse) sequences, discretized Markov transformations

1. Introduction

In view of randomness in chaotic dynamics of one-dimensional ergodic transformations, de Bruijn sequences were generalized in [1]. Although the total number of the celebrated de Bruijn sequences of length $2^n$ ($n \geq 1$) is known to be $2^{2n-1} - n$, the total number of the sequences proposed in [1] was not known. Motivated by the research in [1], we generally defined discretized Markov transformations and found an algorithm to give the total number of full-length sequences based on discretized Markov transformations in [2].

As pointed out in [3] and [4], only a few algorithms are known for generating all de Bruijn sequences. In [5], we defined the piecewise-monotone-increasing Markov transformations, which included de Bruijn sequences, and gave the bounded monotone truth-table algorithm for generating all full-length sequences which are based on the discretized piecewise-monotone-increasing Markov transformations.

From the viewpoint of discretized Markov transformations, we have studied the correlational properties of de Bruijn sequences in [6] and [7]. We provided a novel lower bound of the minimum values of the normalized auto-correlation functions for de Bruijn sequences of length $2^n$ ($n \geq 3$) in [6].

It was pointed out in [3] that for $n = 5$, there exist CR (complement reverse) sequences in the de
2. Dyck language

Following [11] and [12], we define the Dyck language \( L(D_n) \) \((n \geq 1)\) from the viewpoint of symbolic dynamics. We set \( \Sigma = \{ \alpha_m, \beta_m : 1 \leq m \leq n \} \). For each \( m \) \((1 \leq m \leq n)\), \( \alpha_m \) is called a negative symbol while \( \beta_m \) is called a positive symbol. We define an inverse monoid (with zero) \( D_n \): It has generators \( \alpha_i, \beta_j \) \((1 \leq i, j \leq n)\) and \( 1 \), whose relations are

\[
\alpha_i \cdot \beta_j = \begin{cases} 
1 & \text{if } i = j, \\
0 & \text{otherwise,}
\end{cases}
\]

and \( \gamma \cdot 1 = 1 \cdot \gamma = \gamma, \gamma \cdot 0 = 0 \cdot \gamma = 0 \) \((\gamma \in \Sigma \cup \{1\})\), \( 0 \cdot 0 = 0 \).

We call elements \( u = u_1u_2 \cdots u_k \in \Sigma^k \) words or blocks over \( \Sigma \) of length \( k \) \((k \geq 1)\). A word of length \( k \) is simply called a \( k \)-word.

The Dyck language \( L(D_n) \) is defined by

\[
L(D_n) = \{ u \in \Sigma^* : \text{red}(u) \neq 0 \}.
\]

If \( \text{red}(u) = 1 \) for \( u \in \Sigma^* \), then \( u \) is said to be balanced.\(^1\) The empty word \( \epsilon \) is balanced.

The set of balanced words in \( L(D_1) \) consists of all regular parenthesis structures. In fact, for \( n = 1 \), denoting \( \alpha_1 = ( \) and \( \beta_1 = ) \), we obtain all regular parentheses structures with up to three pairs of parentheses:

\[
( ), (()) , ()( ), ((( ))) , ( (( ))) , ( )(( ) ), ()(()) , (()()).
\]

Remark 1 It is well known that the \( k \) pairs of parentheses are enumerated by the Catalan numbers:

\[
\frac{1}{k+1} \binom{2k}{k}.
\]

For completeness, we review definitions and results stated in [7] which are closely related to this study in the next section.

\(^1\)In [13], the language with \( n \) types of balanced parentheses are said to be the Dyck language.
3. Construction of a prototype of CR graphs

Let \( G_n = (V_n,A_n) \) be the de Bruijn graph with the set \( V_n = \{0,1\}^{n-1} \) of vertices and the set \( A_n = \{0,1\}^{n} \) of arcs. De Bruijn sequences of length \( 2^n \) are exactly Eulerian circuits in the de Bruijn graph \( G_n \).

For \( a \in \{0,1\} \), we use \( \overline{a} \) to denote the binary \textit{complement} of \( a \), i.e., \( \overline{0} = 1 \) and \( \overline{1} = 0 \). We also treat a \textit{time-reversal} of sequences: For a sequence \( X = (X_i)_{i=0}^{N-1} \) over a finite alphabet \( \Sigma \), the reverse \( \overline{X} \) of \( X \) is defined by \( \overline{X} = (X_i)^0_{i=N-1} \).

A \textit{(binary) cycle} of length \( k \) is a sequence of binary \( k \)-word \( a_1a_2\cdots a_k \) taken in a circular order. In the cycle \( a_1a_2\cdots a_k \), \( a_1 \) follows \( a_k \), and \( a_2\cdots a_k a_1\cdots a_{k-1} \) are all the same cycle as \( a_1a_2\cdots a_k \). Two sequences \( X = (X_i)_{i=0}^{N-1} \) and \( Y = (Y_i)_{i=0}^{N-1} \) are said to be \textit{equivalent}, in symbols \( X \simeq Y \), if \( X \) and \( Y \) are the same cycle.

Now we can define the following.

**Definition 1** If \( X \simeq \overline{X} \) or equivalently \( X \simeq \overline{X} \overline{X} \), then \( X \) is called a \textit{CR (complement reverse)} sequence.

By the definition, if \( X \) is a CR sequence, so are \( \overline{X} \) and \( \overline{X} \overline{X} \).

In what follows, let \( X = (X_i)_{i=0}^{N-1} \) be a de Bruijn sequence of length \( 2^n \) if it is not stated otherwise. It was pointed out in [3] that for even \( n \geq 4 \), \( X \neq \overline{X} \) holds, and on the other hand that for \( n = 5 \), \( X \simeq \overline{X} \) occurs. In fact, 32 pairs of CR sequences exist for \( n = 5 \). An example of such CR sequences is given:

**Example 1** (Fredricksen [3]) \( X = 1111001000101010111010000110100 \) has \( \overline{X} \simeq \overline{X} \overline{X} \).

Naturally the following problem was posed by Fredricksen in [3]: Show that there exists a CR sequence whenever \( n \geq 3 \) is odd. In [8], the following characterization of CR sequences was presented.

**Lemma 1** (Etzion and Lempel [8]) Let \( Y = (Y_i)_{i=0}^{N-1} \) be a sequence over \( \{0,1\} \), which is not necessarily a de Bruijn sequence. The sequence \( Y \) is a CR sequence if and only if \( N \) is even and \( Y \simeq \overline{w} \) for an \( N/2 \)-word \( w \).

For words \( u \) and \( v \), we use \( wv \) to denote a concatenation of \( u \) and \( v \).

We set \( n = 2m + 1 \) \((m \geq 1) \). Since \( n - 1 = 2m \) is even, in view of Lemma 1, the set \( V_{2m+1} = \{0,1\}^{2m} \) of vertices includes all \( 2m \) CR sequences of length \( 2m \). To distinguish such CR sequences of length \( 2m \) from CR sequences in question of length \( 2^n \), we refer to such CR sequences as \textit{CR vertices} or \textit{CR \( 2m \)-words}. We use \( V_{2m}^{CR} \subset V_{2m+1} \) to denote the set of CR vertices. Since CR \( 2m \)-words are in the form of \( \overline{w} \) where \( w \in \{0,1\}^m \), a total order relation \( \leq \) on \( V_{2m}^{CR} \) is defined by the following: for any \( \overline{w} \) and \( \overline{v} \) in \( V_{2m}^{CR} \), \( \overline{w} \leq \overline{v} \) if and only if

\[
\begin{align*}
  u_1v_12^{m-1} + u_2v_22^{m-2} + \cdots + u_mv_m &\leq v_1v_12^{m-1} + v_2v_22^{m-2} + \cdots + v_nv_n,
\end{align*}
\]

where \( u = u_1u_2\cdots u_m \) and \( v = v_1v_2\cdots v_m \) are in \( \{0,1\}^m \). Thus we number all the elements in \( V_{2m}^{CR} \):

\[
\begin{align*}
  u(0) &< u(1) \ll \cdots \ll u(2^{2m-1}).
\end{align*}
\]

**Definition 2** The weight \( W(Y) \) of a sequence \( Y = (Y_i)_{i=0}^{N-1} \) over \( \{0,1\} \) is defined to be the number of nonzero digits among the \( N \) \( Y_i \)'s, i.e., \( W(Y) = \sum_{i=0}^{N-1} Y_i \).

Using this, we divide \( V_n \) into three disjoint subsets \( V_n^- = \{v \in V_n : W(v) < m\} \), \( V_n^0 = \{v \in V_n : W(v) = m\} \), and \( V_n^+ = \{v \in V_n : W(v) > m\} \). Note that \( V_n^{CR} \subset V_n^0 \) since \( W(v) = m \) for \( v \in V_n^{CR} \).

Further, we divide \( V_n^0 \) into four disjoint subsets \( V_n^{00} = \{v \in V_n^0 : v = 0w0, w \in \{0,1\}^{2(m-1)}\} \), \( V_n^{01} = \{v \in V_n^0 : v = 0w1, w \in \{0,1\}^{2(m-1)}\} \), \( V_n^{10} = \{v \in V_n^0 : v = 1w0, w \in \{0,1\}^{2(m-1)}\} \), and \( V_n^{11} = \{v \in V_n^0 : v = 1w1, w \in \{0,1\}^{2(m-1)}\} \). In the case that \( m = 1 \), we think of \( w \in \{0,1\}^0 \) as \( w = \epsilon \).

For integers \( a \) and \( b \), if \( a \) is a divisor of \( b \), we write \( a|b \). For \( m \geq 2 \), we use \( d(m) \) to denote the number of the divisors of \( m \). For a word \( w \), we use \( w^k \) to denote the concatenation of \( k \) copies of \( w \), i.e., \( w \cdots w \).
We use \([x]\) to denote the greatest integer not exceeding \(x\). We use \(S\) to denote the shift transformation on \([0, 1)^{2m}\), i.e., \(S(v_1, v_2, \ldots, v_{2m-1}, v_{2m}) = (v_2, v_3, \ldots, v_{2m}, v_1)\) for \(v = v_1 v_2 \cdots v_{2m} \in [0, 1)^{2m}\).

**Definition 3** For \(m (\geq 2), 2(d(m) - 1)\) vertices in \(\mathcal{V}^{CR}_n\) in the form of \(v^{(i(k))} = (1^k \varphi^j)\) and \(v^{(i(k))}\) with \(k \geq 2\) are called the neutral vertices, where \(|k|, m\) and \(i(k) = \frac{2^k(\frac{m}{2})+ k - 2^k}{2^k + 1}\). We use \(\mathcal{V}^{CR, \nu}_n\) to denote the set of the neutral vertices in \(\mathcal{V}^{CR}_n\). For each \(j = 1, 2, \ldots, k - 1, S^j(v^{(i(k))})\) is in \(\mathcal{V}^{CR}_n\). Such vertices in \(\mathcal{V}^{11}_n\) are also called neutral. We use \(\mathcal{V}^{11, \nu}_n\) to denote the set of the neutral vertices in \(\mathcal{V}^{11}_n\). The set \(\mathcal{V}^{00, \nu}_n\) of the neutral vertices in \(\mathcal{V}^{00}_n\) is complementarily defined.

First we construct a directed graph \(G_0^n\) associated with the de Bruijın graph \(G_n\). We set \(\mathcal{V}_n = \{\lambda\} \cup \mathcal{V}_n \setminus \mathcal{V}_n^+\). For two vertices of the forms \(u = a_1 a_2 \cdots a_{n-1}\) and \(v = a_2 a_3 \cdots a_n\) in \(\mathcal{V}_n\), the binary \(n\)-word \(a_1 a_2 \cdots a_n\) is defined as an arc from \(u\) to \(v\). The obtained subgraph of \(G_n\) is not Eulerian since two types of arcs in \(G_n\) are not presented: \(u1\) where \(u \in \mathcal{V}_n\) is in the form of \(u = 0v\); and \(1u\) where \(u \in \mathcal{V}_n^0\) is in the form of \(u = v0\). Corresponding all such arcs each in \(G_n\): \(u1\) where \(u = 0v \in \mathcal{V}_n^0\); and \(1u\) where \(u = v0 \in \mathcal{V}_n^0\), we add an arc \(u0\) from \(u\) to \(v\) for every \(u = 0v \in \mathcal{V}_n^0\); and an arc \(v0\) from \(v\) to \(u\) for every \(u = v0 \in \mathcal{V}_n^0\). The resulting directed graph is Eulerian, which we use \(G_0^n\) to denote.

Second we modify the directed graph \(G_0^n\) to obtain a prototype of CR graphs. Except the neutral vertices in \(\mathcal{V}^{CR}_n \cup \mathcal{V}^{11}_n\), we split every vertex \(v \in \mathcal{V}_n^0\) into two vertices: \(v\) with arcs \(0v\) and \(v0\); and \(v^+\) with arcs \(1v+\) and \(v+1\), as in the diagram:

\[
\begin{array}{c cc c c c}
0v & / & v0 & 0v & v+1
\end{array}
\]

Then, other than the neutral vertices, for every \(v \in \mathcal{V}_n^0\), the copied vertex \(v^+\) occurs in a single loop \(\lambda v^+\lambda v\) where \(0 \leq i + j \leq m\). We delete all such single loops. On the other hand, for each pair of neutral vertices \(\overline{v^{(i(k))}}\) and \(\overline{v^{(i(k))}}\) in \(\mathcal{V}^{CR}_n\), we have an arc \(v^{(i(k))}\varphi^j\lambda v^{(i(k))}\) from \(\overline{v^{(i(k))}}\) to \(v^{(i(k))}\), where \(|m|\) with \(k \geq 2\), and \(i(k)\) is as in Definition 3. For each \(k\), we delete such an arc from \(\overline{v^{(i(k))}}\) to \(v^{(i(k))}\). Then we add an arc from \(\lambda\) to \(v^{(i(k))}\) and label it as \(\lambda v^{(i(k))}\) while we add an arc from \(\overline{v^{(i(k))}}\) to \(\lambda\) labeled as \(\overline{v^{(i(k))}}\). Thus we obtain an Eulerian graph with the vertex set \(\{\lambda\} \cup (\mathcal{V}_n \setminus \mathcal{V}^{00, \nu}_n) \cup \mathcal{V}^\sim_n\), which we use \(G^-_n\) to denote. We call it the prototype of CR graphs.

**4. Construction of CR graphs**

Now we are in a position to construct CR graphs by modifying the directed graph \(G^-_n\). For the case \(m = p\) where \(p\) is a prime number, we have already constructed the set of CR graphs, which yields all CR sequences in the de Bruijn sequences of length \(2^{2p+1}\) in [7]. Hence, in what follows, we suppose \(m (\geq 2)\) is a non-prime number, which implies \(m \geq 4\).

First, we replace the vertex \(\lambda\) and its all \(4(d(m) - 1)\) arcs labeled \(\lambda v^{(i(k))}\) or \(\overline{v^{(i(k))}}\lambda\), where \(v^{(i(k))} \in \mathcal{V}^{CR, \nu}_n\), by \(2(d(m) - 1)\) arcs from \(v^{(i(k))}\) to \(v^{(i(k))}\), where \(|m|\) with \(k \geq 2\), and \(i(k)\) is as in Definition 3. For each \(k\), the resulting two arcs from \(v^{(i(k))}\) to \(v^{(i(k))}\) are labeled the same as \(\overline{v^{(i(k))}}\lambda v^{(i(k))}\).

Choose \(v^{(i)} \in \mathcal{V}^{CR}_n\) in \(G^-_n\) and fix it. If \(v^{(i)}\) is not the neutral vertex, i.e., \(v^{(i)} \in \mathcal{V}^{CR}_n \setminus \mathcal{V}^{CR, \nu}_n\), then we add a loop, an arc from \(v^{(i)}\) to \(v^{(i)}\), labeled \(v^{(i)}\lambda v^{(i)}\). If \(v^{(i)}\) is the neutral vertex, i.e., \(v^{(i)} = v^{(i)}\) or \(\overline{v^{(i)}}\) \(v^{(i)}\), then do nothing since two arcs labeled \(v^{(i)}\lambda v^{(i)}\) from \(v^{(i)}\) to \(v^{(i)}\) when \(v^{(i)} = v^{(i)}\), or from \(v^{(i)}\) to \(\overline{v^{(i)}}\) when \(v^{(i)} = \overline{v^{(i)}}\), are already provided in the graph in construction.

Next, if \(v^{(i)} \in \mathcal{V}^{CR}_n \setminus \mathcal{V}^{CR, \nu}_n\), then we split every pair of neutral vertices \(v^{(i(k))}\) and \(\overline{v^{(i(k))}}\) in \(\mathcal{V}^{CR, \nu}_n\) each into two vertices similarly as in the above diagram, which leads to

\[
\begin{aligned}
0v^{(i(k))} & \rightarrow \overline{v^{(i(k))}} \\
\overline{v^{(i(k))}} & \rightarrow v^{(i(k))}\lambda v^{(i(k))} \\
v^{(i(k))}\lambda v^{(i(k))} & \rightarrow v^{(i(k))} \\
v^{(i(k))} & \rightarrow 0v^{(i(k))}
\end{aligned}
\]

On the other hand, if \(v^{(i)}\) is the neutral vertex, i.e., \(\exists k_0 (k_0 \geq 2, k_0 | m)\), \(v^{(i)} = v^{(i(k_0))}\) or \(v^{(i)} = \overline{v^{(i(k_0))}}\), then we split both neutral vertices \(v^{(i(k_0))}\) and \(\overline{v^{(i(k_0))}}\) in \(\mathcal{V}^{CR, \nu}_n\) each into two vertices as in the following diagram:

\[
\begin{aligned}
0v^{(i(k_0))} & \rightarrow \overline{v^{(i(k_0))}} \\
\overline{v^{(i(k_0))}} & \rightarrow v^{(i(k_0))}\lambda v^{(i(k_0))} \\
v^{(i(k_0))}\lambda v^{(i(k_0))} & \rightarrow v^{(i(k_0))} \\
v^{(i(k_0))} & \rightarrow 0v^{(i(k_0))}
\end{aligned}
\]
while we split the other pairs of neutral vertices $v(i(k))$ and $\overline{v(i(k))}$ ($k \neq k_0$) in $\mathcal{V}_{n}^{CR, \nu}$ each into two vertices in the same way as in the above diagram (2).

Eventually, for each $v(i) \in \mathcal{V}_{n}^{CR}$, we obtain an Eulerian graph with the vertex set $(\mathcal{V}_{n}^{CR}) \setminus \mathcal{V}_{n}^{\bullet, \nu} \cup \mathcal{V}_{n}^{\nu+}$, where $\mathcal{V}_{n}^{CR, \nu+} = \{v(i(k)) + , v(i(k)) - ; v(i(k)), \overline{v(i(k))} \in \mathcal{V}_{n}^{CR, \nu}\}$, which we use $H_{v(i)}$ to denote. We call the CR graph associated with $v(i)$ since Eulerian circuits in $H_{v(i)}$ yield CR sequences. Noting that the vertex sets are the same for all $v(i) \in \mathcal{V}_{n}^{CR}$, we write $W_{n}^{CR} = (\mathcal{V}_{n}^{CR}) \setminus \mathcal{V}_{n}^{\bullet, \nu} \cup \mathcal{V}_{n}^{\nu+}$. Using $E_{v(i)}$ to denote the set of arcs in $H_{v(i)}$, we write $H_{v(i)} = (W_{n}^{CR}, E_{v(i)})$. At this stage we have $2^{m}$ CR graphs. It is worth noting that $H_{v(i)}$ and $H_{\overline{v(i)}}$ are graph isomorphic. In symbols, we write $H_{v(i)} \simeq H_{\overline{v(i)}}$.

5. An algorithm for generating all CR sequences

Using the notion of CR vertex, in [7], we obtain a refinement of Lemma 1 as follows, which plays crucially important roles in constructions of CR sequences.

**Lemma 2 ([7])** Let $X \simeq \{w\}$ be a CR sequence in the de Bruijn sequence of length $2^{2m+1}$, where $w = w_{1}w_{2} \cdots w_{2m} \in \{0,1\}^{2m}$. Then there exists a unique CR vertex $v \in \mathcal{V}_{2m+1}^{CR}$ such that

$$v = \gamma_{w_{1}w_{2} \cdots w_{m}w_{1}w_{2} \cdots w_{m} = w_{2m-m+1} \cdots w_{2m-1}w_{2m} \cdots w_{2m-1}w_{2m}}.$$ (3)

Moreover, the unique $v$ occurs in $X$ twice in the form of $0v1$ and $1v0$ while the other CR vertices $u \in \mathcal{V}_{2m+1}^{CR}$ occurs only once in $w$ in the form of $1uv1$ or $0uv0$.

As in the previous section, we suppose $m \geq 4$ is a non-prime number. For a fixed $v(i) \in \mathcal{V}_{n}^{CR}$, since $H_{v(i)}$ is Eulerian, we obtain an Eulerian circuit in $H_{v(i)}$. The circuit exhibits one of $(2d(m) - 1)!$ circular permutations of elements in $\mathcal{V}_{n}^{CR, \nu}$. Apart from the case $m = p$ where $p$ is a prime number, all the circuits do not yield CR sequences if $m$ is a non-prime number. To construct all CR sequences from the Eulerian circuits in CR graphs, we introduce

**Definition 4** For each neutral vertex $v(i(k)) \in \mathcal{V}_{n}^{CR, \nu}$, where $k|m$ with $k \geq 2$, and $i(k)$ is as in Definition 3, the pair $0v(i(k))\lambda v(i(k))0$ and $1v(i(k))\lambda v(i(k))1$ are said to be balanced. Similarly, the pair $0\overline{v(i(k))}\lambda \overline{v(i(k))}1$ and $1\overline{v(i(k))}\lambda \overline{v(i(k))}0$ are said to be balanced.

We observe there exist $d(m) - 1$ balanced pairs in every Eulerian circuit in $H_{v(i)}$. We think of the set of such balanced pairs as the alphabet $\Sigma$ for the Dyck language $L(D_{d(m)-1})$. If $v(i)$ is not the neutral vertex in $\mathcal{V}_{n}^{CR}$, for each $k$ where $k|m$ with $k \geq 2$, there is a one-to-one correspondence between such $k$’s and $j(k)$’s with $1 \leq j(k) \leq d(m) - 1$ such that

$$\{0v(i(k))\lambda v(i(k))0, 1v(i(k))\lambda v(i(k))1\} = \{\alpha_{j(k)}, \beta_{j(k)}\}. \quad (4)$$

If $v(i)$ is the neutral vertex, i.e., $\exists k_{0}$ ($k_{0} \geq 2$, $k_{0}|m$), $v(i) = v(i(k_{0}))$ or $v(i) = \overline{v(i(k_{0}))}$, where $i(k_{0})$ is as in Definition 3, we have

$$\{0\overline{v(i(k_{0}))}\lambda \overline{v(i(k_{0}))}1, 1\overline{v(i(k_{0}))}\lambda \overline{v(i(k_{0}))}0\} = \{\alpha_{j(k_{0})}, \beta_{j(k_{0})}\}.$$ For other $k \neq k_{0}$, the correspondence is the same as in the case that $v(i)$ is not the neutral vertex, which is given by (4). In either case, we obtain $2^{d(m)-1}$ one-to-one correspondences between the set of the balanced pairs and $\Sigma$.

Let us consider all regular parentheses structures with $d(m) - 1$ pairs of parentheses as in (1). Its total number is given by $\frac{1}{d(m)} \left(\frac{2d(m) - 1}{d(m) - 1}\right)$ from Remark 1. In such a regular parentheses structure
of length $2(d(m) - 1)$, we have $d(m) - 1$ open brackets $\)$. We freely arrange $d(m) - 1$ negative symbols $\alpha_1, \cdots, \alpha_{d(m) - 1}$ in the position of $d(m) - 1$ open brackets. Its total number is given by $(d(m) - 1)!$. To obtain a balanced Dyck word from the regular parentheses structure of length $2(d(m) - 1)$, if we choose such an arrangement of $d(m) - 1$ negative symbols in the regular parentheses structure, the position of positive symbols $\beta_1, \cdots, \beta_{d(m) - 1}$ is uniquely determined. Taking account of the equivalence relation in the cycle, we eventually obtain

$$\frac{1}{d(m)} \binom{2(d(m) - 1)}{d(m) - 1} (d(m) - 1)! 2^{d(m) - 1}$$

circular permutations of elements in the set of the balanced pairs in Definition 4 which correspond to balanced Dyck word of length $2(d(m) - 1)$ in $\mathcal{L}(D_{d(m) - 1})$. Such a circular permutation of elements in the set of the balanced pairs in Definition 4 is said to have a balanced parenthesis structure of length $2(d(m) - 1)$ with $d(m) - 1$ types of pairs of parentheses. We will see the Eulerian circuits which exhibit such circular permutations in CR graphs only admit CR sequences. The existence of such an Eulerian circuit in each CR graph is guaranteed by

**Lemma 3** For each $v^{(i)} \in \mathcal{V}^{CR}_n$, there exists an Eulerian circuit in $H_{v^{(i)}}$ which exhibits a balanced parenthesis structure of length $2(d(m) - 1)$ with $d(m) - 1$ types of pairs of parentheses.

**Proof:** See Appendix A.

Henceforth we may suppose that, once given a CR graph $H_{v^{(i)}}$, we obtain all Eulerian circuits in $H_{v^{(i)}}$, each of which exhibits the balanced parenthesis structure stated above. In fact, we preliminarily select all such Eulerian circuits by checking the balanced parenthesis structure in all Eulerian circuits in $H_{v^{(i)}}$. Let $Y$ be such an Eulerian circuit in $H_{v^{(i)}}$. We identify the circuit $Y$ as a sequence over $\{\lambda, 0, 1\}$, where we define $X = \lambda$.

Let us consider a periodic sequence generated by the sequence $Y$, which we use $Y^\infty$ to denote. We use $\Phi : \Sigma \rightarrow \hat{\Phi}(\Sigma)$ to denote one of the above-mentioned $2^{d(m) - 1}$ one-to-one correspondences for $Y$.

The following observation plays an important role in constructions of CR sequences.

**Remark 2** For each correspondence $\Phi(\gamma) = a\pi b\alpha v$, where $\gamma \in \Sigma$, $a, b \in \{0, 1\}$, and $v \in \mathcal{V}^{CR}_m$, we define $\hat{\Phi}(\gamma) = bv\alpha\pi a$. Then we obtain $^\tau\Phi(\alpha_j) w \Phi(\beta_j) = \hat{\Phi}(\alpha_j) \hat{\Phi}(\beta_j)$ for $1 \leq j \leq d(m) - 1$, where $w \in \{0, 1, \lambda\}^*$.\[ \]

i) If $v^{(i)}$ is not the neutral vertex in $\mathcal{V}^{CR}_n$, then, taking account of the equivalence relation which arises from the cycle, $Y^\infty$ may be written uniquely in the form of

$$v^{(i)}0f^\Phi(\alpha_{j_1})g^\Phi(\beta_{j_1})h0v^{(i)}\lambda v^{(i)}0f \cdots \tag{5}$$

where $f \in \{0, 1\}^*$, $g$ and $h$ are in $\{0, 1, \lambda\}^*$, $\alpha_{j_1}$ is the leftmost negative symbol in the corresponding balanced Dyck word, and $v^{(i)}$ appears exactly twice in

$$v^{(i)}0f^\Phi(\alpha_{j_1})g^\Phi(\beta_{j_1})h0v^{(i)}.$$  

We have to consider two cases, namely $\Phi(\alpha_{j_1}) = 0v^{(i)(k_1)}\lambda v^{(i)(k_1)}0$ and $\Phi(\beta_{j_1}) = 1v^{(i)(k_1)}\lambda v^{(i)(k_1)}1$, or $\Phi(\alpha_{j_1}) = 1v^{(i)(k_1)}\lambda v^{(i)(k_1)}1$ and $\Phi(\beta_{j_1}) = 0v^{(i)(k_1)}\lambda v^{(i)(k_1)}0$. However, we consider only the former case since the processes of constructing a CR sequence from $Y$ are exactly the same in both cases. We transform

$$v^{(i)}0f0v^{(i)(k_1)}\lambda v^{(i)(k_1)}0g1v^{(i)(k_1)}\lambda v^{(i)(k_1)}10v^{(i)}\lambda$$

in $Y^\infty$ into

$$v^{(i)}0f0v^{(i)(k_1)}\lambda v^{(i)(k_1)}0g1v^{(i)(k_1)}\lambda v^{(i)(k_1)}1h0v^{(i)}\lambda.$$  

Noting that $v^{(i)(k_1)}$ and $\epsilon v^{(i)(k_1)}$ are CR words, we obtain

$$v^{(i)}0f0v^{(i)(k_1)}\lambda v^{(i)(k_1)}0g1v^{(i)(k_1)}\lambda v^{(i)(k_1)}1h0v^{(i)}\lambda.$$
After deleting two \( \lambda \)'s, replace repetitions \( v^i(k_1) \) and \( v^i(k_1) \) each by single words \( \bar{v}^i(k_1) \) and \( \bar{v}^i(k_1) \) respectively, then we obtain
\[
v^i(0)f_0\bar{v}^i(k_1)0\gamma_1\bar{v}^i(k_1)1h0v^i(1),
\]
which we use \( Z^{(1)} \) to denote.

Next, \( Z^{(1)} \) may be written in the form of
\[
v^i(0)f^{(2)}\Phi(\alpha_{\gamma_2})g^{(2)}\Phi(\beta_{\gamma_2})h^{(2)}0v^i(1)
\]
depending on \( \Phi(\alpha_{\gamma_2}) \) and \( \Phi(\beta_{\gamma_2}) \) appear in \( g \) or \( h \), respectively, in (5), respectively, where \( f^{(2)} \), \( g^{(2)} \) and \( h^{(2)} \) are in \( \{0, 1, \lambda\}^* \), and where \( \alpha_{\gamma_2} \) is the second leftmost negative symbol in the corresponding balanced Dyck word. Since \( Y \) has the balanced parenthesis structure, both \( \Phi(\alpha_{\gamma_2}) \) and \( \Phi(\beta_{\gamma_2}) \) appear together in either \( g \) or \( h \) but not separately in \( g \) and \( h \) in (5). If \( \Phi(\alpha_{\gamma_2}) \) and \( \Phi(\beta_{\gamma_2}) \) appeared separately in \( g \) and \( h \), then they would break the balanced parenthesis structure. Even if the both appear in \( g \), the above transformation from \( g \) into \( \gamma_1 \) does not affect the balanced parenthesis structure in \( Y \) in view of Remark 2.

On repeating the above transformations without changing the balanced parenthesis structure, we inductively obtain \( Z^{(d(m)−1)} \). Noting again that \( v^i(k_1) \) and \( \bar{v}^i(k_1) \) are CR words, we obtain
\[
Z^{(d(m)−1)} ▶ \gamma Z^{(d(m)−1)} = v^i(0)f_0\bar{v}^i(k_1)0\ldots 0\bar{v}^i(1)v^i(1)\ldots 1\bar{v}^i(k_1)1\gamma_1v^i(1).
\]
Replacing the repetition \( v^i(v^i) \) that occurs twice in a circular order each by single word \( v^i(1) \) respectively, we obtain a CR sequence
\[
X = v^i(0)f_0\bar{v}^i(k_1)0\ldots 0\bar{v}^i(1)\ldots 1\bar{v}^i(k_1)1\gamma_1v^i(1)
\]
of length \( 2^{2m+1} \). It is easy to check that the obtained CR sequence \( X \) is in the de Bruijn sequences of length \( 2^{2m+1} \).

ii) We consider the case that \( v^i(0) \) is the neutral vertex, i.e., \( v^i = v^i(k_0) \) or \( v^i = \bar{v}^i(k_0) \). We have to consider both cases. However, we only consider the case that \( v^i = v^i(k_0) \) since we have \( H_{v^i} = H_{\bar{v}^i} \). Then, \( Y^{\infty} \) may be written in the form of
\[
\Phi(\alpha_{\gamma_2})f\Phi(\beta_{\gamma_2})g\Phi(\alpha_{\gamma_2})f \cdots
\]
where \( f, g, \in \{0, 1, \lambda\}^* \), \( \alpha_{\gamma_2} \) is the leftmost negative symbol in the corresponding balanced Dyck word, and \( v^i(k_0) \) and \( \bar{v}^i(k_0) \) appear exactly twice in \( \Phi(\alpha_{\gamma_2})f\Phi(\beta_{\gamma_2})g \). We have to examine two cases, namely \( \Phi(\alpha_{\gamma_2}) = 1v^i(k_0)\lambda v^i(k_0) \) \( \text{and} \) \( \Phi(\beta_{\gamma_2}) = 0v^i(k_0)\lambda v^i(k_0) \), or \( \Phi(\alpha_{\gamma_2}) = 0v^i(k_0)\lambda v^i(k_0) \) \( \text{and} \) \( \Phi(\beta_{\gamma_2}) = 1v^i(k_0)\lambda v^i(k_0) \). However, we only consider the former case since the processes of constructing a CR sequence from \( Y \) are exactly the same in both cases. Then, taking account of the equivalence relation which arises from the cycle, \( Y^{\infty} \) may be written uniquely in the form of
\[
v^i(k_0)0f_0\bar{v}^i(k_0)1g_1v^i(k_0)\lambda v^i(k_0)0f \cdots
\]
We transform
\[
v^i(k_0)0f_0\bar{v}^i(k_0)1g_1v^i(k_0)\lambda
\]
in \( Y^{\infty} \) into
\[
v^i(k_0)0f_0\bar{v}^i(k_0)1g_1v^i(k_0)\lambda = v^i(k_0)0f_0\bar{v}^i(k_0)1g_1v^i(k_0)\lambda r_0v^i(k_0)
\]
After deleting two \( \lambda \)'s, replace the repetition \( v^i(k_0) \) and \( \bar{v}^i(k_0) \) by single words \( \bar{v}^i(k_0) \), then we obtain
\[
v^i(k_0)0f_0\bar{v}^i(k_0)0r_0v^i(k_0),
\]
which we use \( Z^{(1)} \) to denote. By using exactly the same procedure as in the case i) above, we inductively obtain \( Z^{(d(m)−1)} \). Modifying \( Z^{(d(m)−1)} \) \( \text{and} \) \( Z^{(d(m)−1)} \) similarly as in the case i) above, we obtain a CR sequence \( X \) in the de Bruijn sequences of length \( 2^{2m+1} \).
Conversely, when we are given a CR sequence $X$ in the de Bruijn sequences of length $2^{2m+1}$, in view of Lemma 3, we find in $X$ or $\overline{X}$ a unique $v^{(i)} \in \mathcal{V}^C_{2m+1}$ that satisfies the condition (3). Depending on whether $v^{(i)}$ is neutral or not, if we reverse the above procedure for the case i) or ii), we obtain an Eulerian circuit in $H_{v^{(i)}}$ from $X$ or $\overline{X}$. This correspondence is two-to-one and onto. Since $X \not\equiv \overline{X}$ for $n \geq 3$ [14], corresponding to a CR sequence $X$, $\overline{X}$ gives a distinct CR sequence. Hence the above procedures for $v^{(i)} \in \mathcal{V}^C_{2m+1}$ as a whole exhaust all pairs $(X, \overline{X})$ of CR sequences in the de Bruijn sequences of length $2^{2m+1}$.

Eventually, we obtain the following.

**Theorem 1** For the case that $m \geq 4$ is a non-prime number, there exists at least $2^{m+1}$ CR sequences in the de Bruijn sequences of length $2^{2m+1}$.

Together with the previous result in [7], we have completely solved the fundamental problem posed by Fredricksen in [3] on existence of CR sequences in the de Bruijn sequences of length $2^{2m+1}$ ($m \geq 1$).

### 6. On generating all CR sequences

For each CR graph $H_{v^{(i)}}$ associated with $v^{(i)}$ ($v^{(i)} \in \mathcal{V}^C_{n}$, $n = 2m + 1$, $m \geq 1$, $i = 0, 1, \ldots, 2m - 1$), we use $\Delta(v^{(i)})$ to denote the cofactor of the $(1, 1)$-th entry in the matrix of admittance of $H_{v^{(i)}}$. Since the CR graphs are Eulerian, by virtue of the theorem by van Aardenne-Ehrenfest and de Bruijn in [15] together with so-called matrix tree theorem [16], the number of Eulerian circuits in $H_{v^{(i)}}$ is given by $\Delta(v^{(i)})$. Denoting the total number of CR sequences generated by the CR graph $H_{v^{(i)}}$ associated with $v^{(i)}$ by $\Upsilon(v^{(i)})$, we obtain $\Upsilon(v^{(i)}) = \Delta(v^{(i)})$ if $m$ is a prime number, and

$$1 \leq \Upsilon(v^{(i)}) \leq \Delta(v^{(i)})$$

if $m$ is a non-prime number.

For any odd number $n = 2m + 1$, based on the algorithm proposed in this research, we first generate $\Delta(v^{(i)})$ Eulerian circuits in $H_{v^{(i)}}$. If $m$ is a prime number, using the algorithm given in [7], we only transform all $\Delta(v^{(i)})$ Eulerian circuits to CR sequences. On the other hand, if $m$ is a non-prime number, as noted just after Lemma 3, we have to select all $\Upsilon(v^{(i)})$ Eulerian circuits with the balanced parenthesis structure in all $\Delta(v^{(i)})$ Eulerian circuits in $H_{v^{(i)}}$. Then we transform all $\Upsilon(v^{(i)})$ Eulerian circuits with the balanced parenthesis structure to CR sequences. Eventually, for the case that $m$ is a non-prime number, the algorithm requires further process of selection compared to the case where $m$ is a prime number.

If the total number $\Upsilon(v^{(i)})$ was known in advance for the case that $m$ is a non-prime number, we might construct CR sequences by adaptively checking the balanced parenthesis structure in Eulerian circuits in $H_{v^{(i)}}$. Regrettably, it seems to be a hard task to obtain the number $\Upsilon(v^{(i)})$ explicitly. We leave the enumeration of $\Upsilon(v^{(i)})$ to be an open problem.

Besides, for any odd number $n$, we might not need the total number $\Delta(v^{(i)})$ of Eulerian circuits in $H_{v^{(i)}}$ for generating all CR sequences of length $2^n$. If fact, we gave the bounded monotone truth-table algorithm for generating all full-length sequences without computing their total number, which are based on the discretized piecewise-monotone-increasing Markov transformations in [5]. Unfortunately, however, we could not apply directly apply this algorithm to generation of all CR sequences based on the CR graphs since we could not find in the CR graphs the expanded cycles defined in [5], which plays an essential role in the algorithm.

We successfully generate all CR sequences of length $2^n$ for $n \leq 7$. We give an example of CR sequences of length $2^7$ below. For $n \geq 9$, further research is needed to develop a computationally efficient algorithm for generating all CR sequences of length $2^n$.

**Example 2** In the case that $n = 7$, for the CR graphs associated with 111000, 011001, 101010, and 110100, we obtain $\Delta(111000) = 696320$ and $\Delta(011001) = \Delta(101010) = \Delta(110100) = 417792$, respectively. In view of the graph isomorphism $H_{v^{(i)}} \cong H_{v^{(i)}}$ and the pairs $(X, \overline{X})$ of CR sequences, the number of CR sequences in the de Bruijn sequences of length $2^7$ is 7798784. For the CR graph associated with 101010, transforming $Z(1) \rightarrow Z(1)$ as stated in the previous section, where
\[ Z^{(1)} = 1010100101000101100011001000100100001000000011101001111000101010, \]
we obtain a CR sequence \( X \).

7. Summary

With the help of the Dyck language, we characterized CR sequences in the de Bruijn sequences of length \( 2^{2m+1} \) where \( m \geq 4 \) is a non-prime number. In virtue of this characterization, we have shown that for any odd number \( n \), there exist CR sequences in the de Bruijn sequences of length \( 2^n \), which completely has settled the fundamental problem posed by Fredricksen on existence of the CR sequences. Consequently, we established an algorithm for generating all CR sequences in the de Bruijn sequences of length \( 2^n \) for any odd \( n \).

Acknowledgments

This study was supported by the Grant-in-Aid for Scientific Research (C) under Grant No. 24560445 from the Japan Society for the Promotion of Science.

Appendix

A. Proof of Lemma 3

For a word \( w = w_1 w_2 \cdots w_{|w|} \in \{0,1\}^* \) with \( |w| \geq 1 \), we use \( w_{[i,j]} \) to denote \( w_i w_{i+1} \cdots w_j \) where \( 1 \leq i \leq j \leq |w| \). We write \( w_{[i]} \) rather than \( w_{[i,j]} \) if \( i = j \). By the definition, \( w_{[i]} = w_i \).

As in the beginning of Sect. 3, we use \( G_n = (V_n, A_n) \) to denote the de Bruijn graph where \( V_n = \{0, 1\}^{n-1} \) and \( A_n = \{0, 1\}^n \). We set \( n = 2m + 1 \) where \( m \geq 4 \) is a non-prime number.

We are concerned with the cycles in \( G_n \). Given a cycle \( c = c_1 c_2 \cdots c_{|c|} \) in \( G_n \), the length \( |c| \) satisfies \( 2m + 1 \leq |c| \leq 2^n \). Since \( c = c_1 c_2 \cdots c_{|c|} \) starts and terminates at vertex \( c_1 c_2 \cdots c_{2m} \), we may think of \( c = c_1 c_2 \cdots c_{|c|} \) as a sequence of vertices: \( c_1 c_2 \cdots c_{2m}, c_2 c_3 \cdots c_{2m+1}, \ldots, c_{|c|} | c_{|c|} - 2m + 1 | c_{|c|} - 2m + 2 \cdots c_{|c|} \).

Two cycles in an Eulerian graph are said to be edge-disjoint if they have no common arcs. Similarly, two cycles in an Eulerian graph are said to be vertex-disjoint if they have no common vertices. A set of edge-disjoint cycles in \( G_n \) that cover every arc in \( A_n \) is called a factor of \( G_n \). As pointed out in [3], such a factor always exists for \( G_n \). Furthermore, by using the truth table of \( G_n \) defined in [5], it is easy to check that, even if we specify a number of edge-disjoint cycles in \( G_n \) beforehand, such a factor with the listed cycles always exists for \( G_n \). For a CR graph \( H_v(\cdot) \) where \( v(\cdot) \in V_n^{CR} \), a factor of \( H_v(\cdot) \), and the truth table of \( H_v(\cdot) \) are similarly defined. With the help of the truth table of \( H_v(\cdot) \), existence of a factor for \( H_v(\cdot) \) is easily verified even if several edge-disjoint cycles in \( H_v(\cdot) \) are listed for a factor in advance.

Let \( v(\cdot) \in V_n^{CR} \). There are two cases. Namely, \( v(\cdot) \) is not the neutral vertex in \( V_n^{CR} \), or \( v(\cdot) \) is the neutral vertex in \( V_n^{CR} \). We have to consider both cases. However, we only consider the former case since the processes of reasoning are exactly the same in both cases. In what follows, we assume \( v(\cdot) \in V_n^{CR} \setminus V_n^{CR, v} \).

First, we designate \( d(m) - 1 \) cycles in \( H_v(\cdot) \). For each neutral vertex \( v(i(k)) \in V_n^{CR, v} \), where \( k \mid m \) with \( k \geq 2 \), and \( i(k) \) is as in Definition 3, we pick out a cycle in \( H_v(\cdot) \) in the form of a sequence of vertices

\[
\begin{align*}
\varepsilon(i(k))_{[2,m]} = 0, \varepsilon(i(k))_{[3,m]} = 01, \varepsilon(i(k))_{[4,m]} = 01^2, \ldots, \varepsilon(i(k))_{[k+1,m]} = 01^{k-1}, \varepsilon(i(k))_{[k+2,m]} = 0 \varepsilon(i(k))_{[1,k]}, \ldots, \varepsilon(i(k))_{[m-1,1]} = 0 \varepsilon(i(k))_{[1,m-1]}, \varepsilon(i(k))_{[2,m]} = 0 \varepsilon(i(k))_{[1,2]}, \varepsilon(i(k))_{[3,m]} = 0 \varepsilon(i(k))_{[1,3]}, \ldots, \varepsilon(i(k))_{[k,m]} = 0 \varepsilon(i(k))_{[1,k]}, \varepsilon(i(k))_{[k+1,m]} = 0 \varepsilon(i(k))_{[1,k+1]}, \ldots. \end{align*}
\]  

(A-1)

Recall that we have an arc \( \varepsilon(i(k))_{[j,m]}V(i(k))_{[j,m]} \) from \( \varepsilon(i(k))_{[j,m]} \) to \( \varepsilon(i(k))_{[j,m]} \) in \( H_v(\cdot) \), which occurs twice in the cycle (A-1). For later use, we remark here that the vertices traversed by the second path from \( \varepsilon(i(k)) \) to \( \varepsilon(i(k)) \) in (A-1) have alternative expressions:

\[
\begin{align*}
\varepsilon(i(k))_{[j,m]} = 0 \varepsilon(i(k))_{[j-1,m-j+1]}, 1 \leq j \leq k-1. \end{align*}
\]  

(A-2)

Recalling \( \varepsilon(i(k)) = (1^k 0^k) \), we see that the obtained \( d(m) - 1 \) cycles are mutually edge-disjoint. Moreover, we note that \( v(j) \) never appears in the \( d(m) - 1 \) cycles.
If we use \(e^{(1)}, e^{(2)}, \ldots, e^{(d(m)-1)}\) to denote such \(d(m) - 1\) edge-disjoint cycles, then we are given a factor for \(H_{\psi(i)}\) including the \(d(m) - 1\) cycles:

\[
\mathcal{F}_0 = \{e^{(0)}, e^{(1)}, e^{(2)}, \ldots, e^{(d(m)-1)}\},
\]

where \(M \geq d(m) - 1\) and \(v^{(i)}_\lambda w^{(i)}\) appears in \(e^{(0)}\).

We note here that, for every vertex \(w \in V_n\), \(w\) appears exactly twice in the cycles in \(\mathcal{F}_0\). More precisely, for every \(w \in V_n\), there exists a cycle \(e^{(j)}(w)\) in \(\mathcal{F}_0\) such that \(w\) appears exactly twice in \(e^{(j)}(w)\), or there exist two cycles \(e^{(j)}(w)\) and \(e^{(j')}\) in \(\mathcal{F}_0\) such that \(w\) appears exactly once in each cycle of \(e^{(j)}(w)\) and \(e^{(j')}\).

The vertex in the latter case plays an important role in the following verification. In the latter case, if \(awb\) appears in \(e^{(j)}(w)\), then \(\pi awb\) appears in \(e^{(j')}\), where \(a, b \in \{0, 1\}\). Taking account of the equivalence relation in the cycle, we may write \(e^{(j)}(w) = awbf\) and \(e^{(j')} = \pi awbg\) where \(f \in \{0, 1\}\) and \(g \in \{0, 1\}\) such that we can join them into a single cycle \(awbg\).

Without loss of generality, we may suppose each of \(M - d(m) + 1\) cycles \(e^{(d(m))}, \ldots, e^{(M)}\) in \(\mathcal{F}_0\) has the maximal length, which implies that they are mutually vertex-disjoint. In fact, if two of them have a common vertex, we can join them as stated above.

Now we turn to the verification. In view of (5), we write \(e^{(0)} = v^{(i)}_\lambda f^{(0)}v^{(i)}_\lambda\). If \(e^{(0)}\) and \(e^{(1)}\) have a common vertex, we can join them into a single cycle. Otherwise, there exists a cycle \(e^{(j)}(d(m) \leq j_1 \leq M)\) such that we can join them into a single cycle by joining \(e^{(0)}\) and \(e^{(j_1)}\) while joining \(e^{(1)}\) and \(e^{(j_1)}\). In either case, the resulting single cycle is in written in the form of

\[
v^{(i)}_\lambda f^{(1)}(0v^{(i)}_\lambda)^{\ell(0)}g^{(1)}(0v^{(i)}_\lambda)\lambda e^{(i)}(k) 1h^{(1)}v^{(i)}_\lambda \quad \text{(A-3)}
\]

or

\[
v^{(i)}_\lambda f^{(1)}(0v^{(i)}_\lambda)^{\ell(1)}g^{(1)}(0v^{(i)}_\lambda)\lambda e^{(i)}(k) 0h^{(1)}v^{(i)}_\lambda \quad \text{(A-4)}
\]

in the light of (A-1) and (A-2). Thus, we see that this joining process does not affect the balanced parenthesis structure. Using \(c^{(0,1)}\) to denote the resulting single cycle, we obtain a factor for \(H_{\psi(i)}\):

\[
\mathcal{F}_1 = \{c^{(0,1)}, c^{(2)}, \ldots, c^{(d(m)-1)}, \ldots, c^{(M)}\} \setminus \{c^{(j_1)}\},
\]

where \(\{c^{(j_1)}\} = \emptyset\) if \(e^{(0)}\) and \(e^{(1)}\) have a common vertex.

Next, with a cycle \(e^{(j)}(d(m) \leq j_2 \leq M)\) if necessary, we join \(c^{(0,1)}\) and \(c^{(2)}\) into a single cycle \(c^{(0,1,2)}\) and obtain factor for \(H_{\psi(i)}\):

\[
\mathcal{F}_2 = \{c^{(0,1,2)}, c^{(3)}, \ldots, c^{(d(m)-1)}, \ldots, c^{(M)}\} \setminus (\{c^{(j_1)}\} \cup \{c^{(j_2)}\}),
\]

where \(\{c^{(j_2)}\} = \emptyset\) if \(c^{(0,1)}\) and \(c^{(2)}\) have a common vertex.

Repeating this joining process, we obtain a factor for \(H_{\psi(i)}\):

\[
\mathcal{F}_{d(m)-1} = \{c^{(0,1,\ldots,d(m)-1)}, c^{(d(m))}, \ldots, c^{(M)}\} \setminus \bigcup_{\ell=1}^{d(m)-1} \{c^{(j_{\ell})}\},
\]

where \(\{c^{(j_{\ell})}\} = \emptyset\) if \(c^{(0,1,\ldots,\ell-1)}\) and \(c^{(\ell)}\) have a common vertex. If the factor \(\mathcal{F}_{d(m)-1}\) only has a single cycle, the resulting cycle \(c^{(0,1,\ldots,d(m)-1)}\) is an Eulerian circuit in \(H_{\psi(i)}\). Otherwise, we join \(c^{(0,1,\ldots,d(m)-1)}\) and all cycles in \(c^{(d(m))}, \ldots, c^{(M)}\) \(\bigcup_{\ell=1}^{d(m)-1} \{c^{(j_{\ell})}\}\) into a single cycle, which is an Eulerian circuit in \(H_{\psi(i)}\).

In each joining process, a common vertex appears in a subword \(f\) of a word of the form \(\Phi(\alpha_j)f\Phi(\alpha_j), \Phi(\beta_j)f\Phi(\beta_j)\), or \(\Phi(\alpha_m)f\Phi(\beta_m)\), where neither \(\Phi(\alpha_m)\) nor \(\Phi(\beta_m)\) appears in \(f\). Since \(f\) changes similarly as \(f^{(0)}\) changes into (A-3) or (A-4) on joining \(c^{(0)}\) and \(c^{(1)}\), each joining process does not affect the balanced parenthesis structure. This completes the proof. \(\square\)

References

[1] A. Tsuneda, Y. Kuga, and T. Inoue, “New maximal-period sequences using extended nonlinear feedback shift registers based on chaotic maps,” IEICE Trans. on Fundamentals, vol. E85-A, pp. 1327–1332, 2002.
[2] H. Fujisaki, “Discretized Markov transformations – An example of ultradiscrete dynamical systems –,” IEICE Trans. Fundamentals, vol. E88-A, pp. 2684–2691, 2005.
[3] H. Fredricksen, “A survey of full length nonlinear shift register cycle algorithm,” SIAM Review, vol. 24, pp. 195–221, 1982.
[4] S.W. Golomb, Shift Register Sequences, Revised Ed., Aegean Park Press, 1982.
[5] H. Fujisaki, “An algorithm for generating all full-length sequences which are based on discretized Markov transformations,” NOLTA, vol. 1, pp. 166–175, 2010.
[6] H. Fujisaki and Y. Nabeshima, “On auto-correlation values of de Bruijn sequences,” NOLTA, vol. 3, pp. 400–408, 2011.
[7] H. Fujisaki, “A construction of all CR sequences in the de Bruijn sequences of length $2^{2p+1}$ where $p$ is a prime number,” NOLTA, vol. 5, pp. 235–249, 2014.
[8] T. Etzion and A. Lempel, “On the distribution of de Bruijn sequences of given complexity,” IEEE Trans. on Information Theory, vol. 30, pp. 611–614, 1984.
[9] P.A. Pevzner, H. Tang, and M.S. Waterman, “An eulerian path approach to DNA fragment assembly,” Proc. Natl. Acad. Sci. USA, vol. 98, pp. 9748–9753, 2001.
[10] D.R. Zerbino and E. Birney, “Velvet: algorithms for de novo short read assembly using de Bruijn graphs,” Genome Research, vol. 18, pp. 821–829, 2008.
[11] T. Hamachi and K. Inoue, “Embedding of shifts of finite type into the dyck shift,” Monatshefte für Mathematik, vol. 145, pp. 107–129, 2005.
[12] T. Meyerovitch, “Tail invariant measures of the Dyck shift,” Israel J. of Math., vol. 163, pp. 61–83, 2008.
[13] J.E. Hopcroft and J.D. Ulman, Introduction to Automata Theory, Languages, and Computation, Addison-Wesley, 1979.
[14] A.H. Chan, R.A. Games, and E.L. Key, “On the complexities of de Bruijn sequences,” J. Comb. Theory, Ser. A, vol. 33, pp. 233–246, 1982.
[15] T. van Aardenne-Ehrenfest and N.G. de Bruijn, “Circuits and trees in oriented linear graphs,” Simon Stevin, vol. 28, pp. 203–217, 1951.
[16] W.T. Tutte, “The dissection of equilateral triangles into equilateral triangles,” Proc. Cambridge Phil. Soc., vol. 44, pp. 463–482, 1948.