2014 International Symposium on Nonlinear Theory and its Applications
NOLTA2014, Luzern, Switzerland, September 14-18, 2014

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

Hiroshi Fujisaki

Graduate School of Natural Science and Technology
Kanazawa University
Kakuma-machi, Kanazawa, Ishikawa, 920-1192 Japan
Email: 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}. 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. Then, 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.

1. Dyck Language

Following [4], we define the Dyck language $\mathcal{L}(D_n)$ ($n \geq 1$) from the viewpoint of symbolic dynamics. We set $\Sigma = \{\alpha_m, \beta_m : 1 \leq m \leq n\}$. For each m ($1 \leq m \leq n$), α_m is called a *negative symbol* while β_m is called a *positive symbol*. We define an inverse monoid (with zero) $\mathcal{D}_n$: It has generators α_i, β_j ($1 \leq i, j \leq n$) and $\mathbf{1}$, whose relations are $\alpha_i \cdot \beta_j = \begin{cases} \mathbf{1} & \text{if } \ i = j, \\ 0 & \text{otherwise,} \end{cases}$ and $\gamma \cdot \mathbf{1} = \mathbf{1} \cdot \gamma = \gamma$, $\gamma \cdot 0 = 0 \cdot \gamma = 0\,(\,\gamma \in \Sigma \cup \{\mathbf{1}\}\,)$, $0 \cdot 0 = 0$.

We call elements $u = u_1 u_2 \cdots u_k \in \Sigma^k$ *words* or *blocks* over Σ of length k ($k \geq 1$). A word of length k is simply called a *k-word*. We use Σ^* to denote the collection of all words over Σ and the empty word ϵ. We use red() to denote a mapping from Σ^* to the inverse monoid $\mathcal{D}_n$ by letting for $\gamma = \gamma_1\gamma_2\cdots\gamma_k \in \Sigma^*$ ($k \geq 1$), $\mathrm{red}(\gamma) = \gamma_1 \cdot \gamma_2 \cdot \ \cdots \ \cdot \gamma_k$ and $\mathrm{red}(\epsilon) = \mathbf{1}$.

The Dyck language $\mathcal{L}(D_n)$ is defined by $\mathcal{L}(D_n) = \{u \in \Sigma^* : \mathrm{red}(u) \neq 0\}$. If $\mathrm{red}(u) = \mathbf{1}$ for $u \in \Sigma^*$, then u is said to be *balanced*.[1] The empty word ϵ is balanced.

The set of balanced words in $\mathcal{L}(D_1)$ 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:

$$(),\ \ (()),\ ()(),\ \ ((())),\ (()()),\ (())(),\ ()(()),\ ()()(). \tag{1}$$

[1] In [5], the language with n types of balanced parentheses are said to be the Dyck language.

Remark 1 *It is well known that the k pairs of parentheses are enumerated by the Catalan numbers:* $\frac{1}{k+1}\binom{2k}{k}$.

2. Construction of a Prototype of CR Graphs

Let $G_n = (\mathcal{V}_n, \mathcal{A}_n)$ be the de Bruijn graph with the set $\mathcal{V}_n = \{0,1\}^{n-1}$ of vertices and the set $\mathcal{A}_n = \{0,1\}^n$ 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 *complement* of a, i.e. $\overline{0} = 1$ and $\overline{1} = 0$. We also treat a *time-reversal* of sequences: For a sequence $\boldsymbol{X} = (X_i)_{i=0}^{N-1}$ over a finite alphabet Σ, the *reverse* ${}^r\boldsymbol{X}$ of $\boldsymbol{X}$ is defined by ${}^r\boldsymbol{X} = (X_i)_{i=N-1}^{0}$.

A (*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 $a_2\cdots a_ka_1, \cdots, a_ka_1\cdots a_{k-1}$ are all the same cycle as $a_1a_2\cdots a_k$. Two sequences $\boldsymbol{X} = (X_i)_{i=0}^{N-1}$ and $\boldsymbol{Y} = (Y_i)_{i=0}^{N-1}$ are said to be *equivalent*, in symbols $\boldsymbol{X} \simeq \boldsymbol{Y}$, if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are the same cycle.

Now we can define the following.

Definition 1 *If* $\boldsymbol{X} \simeq {}^r\overline{\boldsymbol{X}}$ *or equivalently* $\overline{\boldsymbol{X}} \simeq {}^r\boldsymbol{X}$*, then* $\boldsymbol{X}$ *is called CR (complement reverse) sequence.*

By the definition, if $\boldsymbol{X}$ is a CR sequence, so are $\overline{\boldsymbol{X}}$ and ${}^r\boldsymbol{X}$.

In what follows, let $\boldsymbol{X} = (X_i)_{i=0}^{N-1}$ be a de Bruijn sequence of length 2^n if it is not stated otherwise.

It was pointed out in [1] that for even $n \geq 4$, $\boldsymbol{X} \not\simeq {}^r\overline{\boldsymbol{X}}$ holds, and on the other hand that for $n = 5$, $\boldsymbol{X} \simeq {}^r\overline{\boldsymbol{X}}$ occurs. In fact, 32 pairs of CR sequences exist for $n = 5$. Naturally the following problem was posed by Fredricksen in [1]: Show that there exists a CR sequence whenever $n\,(\geq 3)$ is odd. In [2], the following characterization of CR sequences were presented.

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

For words u and v, we use uv 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 $\mathcal{V}_{2m+1} = \{0,1\}^{2m}$ of vertices includes all 2^m 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 *CR vertices* or *CR $2m$-words*. We use $\mathcal{V}_n^{CR}\,(\subset \mathcal{V}_n)$ to denote the set of CR vertices. Since CR $2m$-words are in the form of ${}^{r}\overline{w}w$ where $w \in \{0,1\}^m$, a total order relation $\le$ on $\mathcal{V}_n^{CR}$ is defined by the following: for any ${}^{r}\overline{u}u$ and ${}^{r}\overline{v}v$ in $\mathcal{V}_n^{CR}$, ${}^{r}\overline{u}u \le {}^{r}\overline{v}v$ if and only if

$$u_1 2^{m-1} + u_2 2^{m-2} + \cdots + u_m \le v_1 2^{m-1} + v_2 2^{m-2} + \cdots + v_m,$$

where $u = u_1 u_2 \cdots u_m$ and $v = v_1 v_2 \cdots v_m$ are in $\{0,1\}^m$. Thus we number all the elements in $\mathcal{V}_n^{CR}$: $v^{(0)} < v^{(1)} < \cdots < v^{(2^m-1)}$.

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

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

Further, we divide $\mathcal{V}_n^0$ into four disjoint subsets $\mathcal{V}_n^{00} = \{v \in \mathcal{V}_n^0 : v = 0w0,\ w \in \{0,1\}^{2(m-1)}\}$, $\mathcal{V}_n^{01} = \{v \in \mathcal{V}_n^0 : v = 0w1,\ w \in \{0,1\}^{2(m-1)}\}$, $\mathcal{V}_n^{10} = \{v \in \mathcal{V}_n^0 : v = 1w0,\ w \in \{0,1\}^{2(m-1)}\}$, and $\mathcal{V}_n^{11} = \{v \in \mathcal{V}_n^0 : v = 1w1,\ w \in \{0,1\}^{2(m-1)}\}$. 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 $m \ge 2$, 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., $\underbrace{w\cdots w}_{k}$. 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., $S(v_1, v_2, \cdots, v_{2m-1}, v_{2m}) = (v_2, v_3, \cdots, v_{2m}, v_1)$ for $v = v_1 v_2 \cdots v_{2m} \in \{0,1\}^{2m}$.

Definition 3 *For $m\,(\ge 2)$, $2(d(m)-1)$ vertices in $\mathcal{V}_n^{CR}$ in the form of $v^{(i(k))} = (1^k 0^k)^{\frac{m}{k}}$ and $\overline{v^{(i(k))}}$ with $k \ge 2$ are called the neutral vertices, where $k|m$ and $i(k) = \frac{2^{2k\left[\frac{m}{k}\right]+k} - 2^k}{2^k+1}$. We use $\mathcal{V}_n^{CR,\nu}$ to denote the set of the neutral vertices in $\mathcal{V}_n^{CR}$. For each $j = 1, 2, \cdots, k-1$, $S^j(v^{(i(k))})$ is in $\mathcal{V}_n^{11}$. Such vertices in $\mathcal{V}_n^{11}$ are also called neutral. We use $\mathcal{V}_n^{11,\nu}$ to denote the set of the neutral vertices in $\mathcal{V}_n^{11}$. The set $\mathcal{V}_n^{00,\nu}$ of the neutral vertices in $\mathcal{V}_n^{00}$ is complementarily defined.*

First we construct a directed graph G_n^0 associated with the de Bruijn graph G_n. We set $\mathcal{W}_n = \{\lambda\} \cup \mathcal{V}_n \setminus \mathcal{V}_n^+$. 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 $\mathcal{W}_n$, 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 $u \in \mathcal{V}_n^0$ 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 $u\lambda$ from u to λ for every $u = 0v \in \mathcal{V}_n^0$; and an arc λu from λ to u for every $u = v0 \in \mathcal{V}_n^0$. The resulting directed graph is Eulerian, which we use G_n^0 to denote.

Second we modify the directed graph G_n^0 to obtain a prototype of CR graphs. Except the neutral vertices in $\mathcal{V}_n^{CR} \cup \mathcal{V}_n^{11}$, 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:

$$\begin{array}{ccc} 0v \searrow & & \nearrow v0 \\ & \circ v & \\ 1v \nearrow & & \searrow v1 \end{array} \quad \text{into} \quad \begin{array}{c} \xrightarrow{0v} \overset{v}{\circ} \xrightarrow{v0} \\ \\ \xrightarrow{1v^+} \overset{v^+}{\circ} \xrightarrow{v^+1} \end{array}.$$

Then, other than the neutral vertices, for every $v \in \mathcal{V}_n^0$, the copied vertex v^+ occurs in a single loop $\lambda 1^i v^+ 1^j \lambda$ where $0 \le i + j \le m$. We delete all such single loops. On the other hand, for each pair of neutral vertices $v^{(i(k))}$ and $\overline{v^{(i(k))}}$ in $\mathcal{V}_n^{CR}$, we have an arc $\overline{v^{(i(k))}} 0^k v^{(i(k))}$ from $\overline{v^{(i(k))}}$ to $v^{(i(k))}$, where $k|m$ with $k \ge 2$, 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 λ to $v^{(i(k))}$ and label it as $\lambda v^{(i(k))}$ while we add an arc from $\overline{v^{(i(k))}}$ to λ labeled as $\overline{v^{(i(k))}}\lambda$. Thus we obtain an Eulerian graph with the vertex set $\{\lambda\} \cup (\mathcal{V}_n^0 \setminus \mathcal{V}_n^{00,\nu}) \cup \mathcal{V}_n^-$, which we use G_n^- to denote. We call it the prototype of CR graphs.

3. 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 [3]. Hence, in what follows, we suppose $m\,(\ge 2)$ is a non-prime number, which implies $m \ge 4$.

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

Choose $v^{(i)} \in \mathcal{V}_n^{CR}$ in G_n^- and fix it. If $v^{(i)}$ is not the neutral vertex, i.e., $v^{(i)} \in \mathcal{V}_n^{CR} \setminus \mathcal{V}_n^{CR,\nu}$, 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., $v^{(i)} = v^{(i(k))}$ or $v^{(i)} = \overline{v^{(i(k))}}$, then do nothing.

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

$$\begin{array}{c} \xrightarrow{0\overline{v^{(i(k))}}} \overset{\overline{v^{(i(k))}}}{\circ} \xrightarrow{\overline{v^{(i(k))}}\lambda v^{(i(k))}} \overset{v^{(i(k))}}{\circ} \xrightarrow{v^{(i(k))}0} \\ \\ \xrightarrow{1\overline{v^{(i(k))+}}} \overset{\overline{v^{(i(k))+}}}{\circ} \xrightarrow{\overline{v^{(i(k))+}}\lambda v^{(i(k))+}} \overset{v^{(i(k))+}}{\circ} \xrightarrow{v^{(i(k))+}1} \end{array}. \tag{2}$$

On the other hand, if $v^{(i)}$ is the neutral vertex, i.e., $\exists\, k_0\ (k_0 \ge 2,\ k_0|m)$, $v^{(i)} = v^{(i(k_0))}$ 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 $\mathcal{V}_n^{CR,\nu}$ each into two vertices as in the following diagram:

$$\begin{array}{c} \xrightarrow{1\overline{v^{(i(k_0))+}}} \underset{\circ}{\overline{v^{(i(k_0))+}}} \xrightarrow{\overline{v^{(i(k_0))+}}\lambda v^{(i(k_0))}} \underset{\circ}{v^{(i(k_0))}} \xrightarrow{v^{(i(k_0))}0} \\ \xrightarrow{0\overline{v^{(i(k_0))}}} \underset{\circ}{\overline{v^{(i(k_0))}}} \xrightarrow{\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))+}} \underset{\circ}{v^{(i(k_0))+}} \xrightarrow{v^{(i(k_0))+}1} \end{array},$$

while we split the other pairs of neutral vertices $v^{(i(k))}$ and $\overline{v^{(i(k))}}$ $(k \neq k_0)$ in $\mathcal{V}_n^{CR,v}$ 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 $(\mathcal{V}_n^0 \backslash \mathcal{V}_n^{00,v}) \cup \mathcal{V}_n^- \cup \mathcal{V}_n^{CR,v+}$, where $\mathcal{V}_n^{CR,v+} = \{v^{(i(k))+}, \overline{v^{(i(k))+}} : v^{(i(k))}, \overline{v^{(i(k))}} \in \mathcal{V}_n^{CR,v}\}$, which we use $H_{v^{(i)}}$ to denote. We call it 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 $\mathcal{W}_n^{CR} = (\mathcal{V}_n^0 \backslash \mathcal{V}_n^{00,v}) \cup \mathcal{V}_n^- \cup \mathcal{V}_n^{CR,v+}$. Using $\mathcal{B}_{v^{(i)}}$ to denote the set of arcs in $H_{v^{(i)}}$, we write $H_{v^{(i)}} = (\mathcal{W}_n^{CR}, \mathcal{B}_{v^{(i)}})$. 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 $H_{v^{(i)}} \simeq H_{\overline{v^{(i)}}}$.

4. An Algorithm for Generating All CR sequences

Using the notion of CR vertex, in [3], we obtain a refinement of Lemma 1 as follows, which plays crucially important roles in constructions of CR sequences.

Lemma 2 ([3]) *Let $\boldsymbol{X} \simeq {}^r\overline{w}w$ be a CR sequence in the de Bruijn sequence of length 2^{2m+1}, where $w = w_1 w_2 \cdots w_{2^{2m}} \in \{0,1\}^{2^{2m}}$. Then there exists a unique CR vertex $v \in \mathcal{V}_{2m+1}^{CR}$ such that*

$$\begin{aligned} v &= {}^r\overline{w_1 w_2 \cdots w_m} w_1 w_2 \cdots w_m \\ &= w_{2^{2m}-m+1} \cdots w_{2^{2m}-1} w_{2^{2m}} \, {}^r\overline{w_{2^{2m}-m+1} \cdots w_{2^{2m}-1} w_{2^{2m}}}. \end{aligned} \tag{3}$$

Moreover, the unique v occurs in $\boldsymbol{X}$ 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 $1u1$ or $0u0$.

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 $(2(d(m)-1)-1)!$ circular permutations of elements in $\mathcal{V}_n^{CR,v}$. 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 $v^{(i(k))} \in \mathcal{V}_n^{CR,v}$, where $k|m$ with $k \geq 2$, and $i(k)$ is as in Definition 3, the pair $0\overline{v^{(i(k))}}\lambda v^{(i(k))}0$ and $1\overline{v^{(i(k))}}\lambda v^{(i(k))}1$ are said to be balanced. Similarly, the pair $0\overline{v^{(i(k))}}\lambda v^{(i(k))}1$ and $1\overline{v^{(i(k))}}\lambda v^{(i(k))}0$ 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 Σ for the Dyck language $\mathcal{L}(D_{d(m)-1})$. 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

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

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 $\{0\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}1,\ 1\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}0\} = \{\alpha_{j(k_0)}, \beta_{j(k_0)}\}$. 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 Σ.

Let us consider all regular parentheses structures with $d(m)-1$ pairs of parentheses as in (1). Its total number is given by $\frac{1}{d(m)}\binom{2(d(m)-1)}{d(m)-1}$ 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 $\frac{1}{d(m)}\binom{2(d(m)-1)}{d(m)-1}\frac{(d(m)-1)!}{2(d(m)-1)} 2^{d(m)-1}$ 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 $v^{(i)} \in \mathcal{V}_n^{CR}$, 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.*

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 $\boldsymbol{Y}$ be such an Eulerian circuit in $H_{v^{(i)}}$. We identify the circuit $\boldsymbol{Y}$ as a sequence over $\{\lambda, 0, 1\}$, where we define $\overline{\lambda} = \lambda$.

Let us consider a periodic sequence generated by the sequence $\boldsymbol{Y}$, which we use $\boldsymbol{Y}^\infty$ to denote. We use $\Phi : \Sigma \to \Phi(\Sigma)$ to denote one of the above-mentioned $2^{d(m)-1}$ one-to-one correspondences for $\boldsymbol{Y}$. The following observation plays an important role in constructions of CR sequences.

Remark 2 *For each correspondence* $\Phi(\gamma) = a\overline{v}\lambda vb$ *where* $\gamma \in \Sigma$, $a, b \in \{0, 1\}$, *and* $v \in \mathcal{V}_n^{CR,v}$, *we define* $\widehat{\Phi}(\gamma) = av\lambda\overline{v}b$. *Then we obtain* ${}^{r}\overline{\Phi(\alpha_j)w\Phi(\beta_j)} = \widehat{\Phi}(\alpha_j)\,{}^{r}\overline{w}\,\widehat{\Phi}(\beta_j)$ *for* $1 \le j \le d(m) - 1$, *where* $w \in \{0, 1, \lambda\}^*$.

i) If $v^{(i)}$ is not the neutral vertex in $\mathcal{V}_n^{CR}$, then $\boldsymbol{Y}^\infty$ may be written 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 α_{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 $\Phi(\alpha_{j_1}) = 0\overline{v^{(i(k_1))}}\lambda v^{(i(k_1))}0$ and $\Phi(\beta_{j_1}) = 1\overline{v^{(i(k_1))}}\lambda v^{(i(k_1))}1$, or $\Phi(\alpha_{j_1}) = 1\overline{v^{(i(k_1))}}\lambda v^{(i(k_1))}1$ and $\Phi(\beta_{j_1}) = 0\overline{v^{(i(k_1))}}\lambda v^{(i(k_1))}0$. However, we consider only the former case since the processes of constructing a CR sequence from $\boldsymbol{Y}$ are exactly the same in both cases. We transform $v^{(i)}0f0\overline{v^{(i(k_1))}}\lambda v^{(i(k_1))}0g1\overline{v^{(i(k_1))}}\lambda v^{(i(k_1))}10v^{(i)}\lambda$ in $\boldsymbol{Y}^\infty$ into $v^{(i)}0f0\overline{v^{(i(k_1))}}\lambda\,{}^{r}\overline{v^{(i(k_1))}0g1\overline{v^{(i(k_1))}}}\,\lambda v^{(i(k_1))}1h0v^{(i)}\lambda$. Noting that $v^{(i(k_1))}$ and $\overline{v^{(i(k_1))}}$ are CR words, we obtain $v^{(i)}0f0\overline{v^{(i(k_1))}}\lambda\overline{v^{(i(k_1))}}0\,{}^{r}\overline{g}1v^{(i(k_1))}\lambda v^{(i(k_1))}1h0v^{(i)}\lambda$. After deleting two λ's, replace repetitions $\overline{v^{(i(k_1))}}\,\overline{v^{(i(k_1))}}$ and $v^{(i(k_1))}v^{(i(k_1))}$ each by single words $\overline{v^{(i(k_1))}}$ and $v^{(i(k_1))}$ respectively, then we obtain $v^{(i)}0f0\overline{v^{(i(k_1))}}0\,{}^{r}\overline{g}1v^{(i(k_1))}1h0v^{(i)}$, which we use $\boldsymbol{Z}^{(1)}$ to denote.

Next, depending on $\Phi(\alpha_{j_2})$ and $\Phi(\beta_{j_2})$ appear in g or h in (5), where α_{j_2} is the second leftmost negative symbol in the corresponding balanced Dyck word, $\boldsymbol{Z}^{(1)}$ may be written in the form of $v^{(i)}0f^{(2)}\widehat{\Phi}(\alpha_{j_2})g^{(2)}\widehat{\Phi}(\beta_{j_2})h^{(2)}0v^{(i)}$ or $v^{(i)}0f^{(2)}\Phi(\alpha_{j_2})g^{(2)}\Phi(\beta_{j_2})h^{(2)}0v^{(i)}$ respectively.

On repeating the above transformations without changing the balanced parenthesis structure, we inductively obtain $\boldsymbol{Z}^{(d(m)-1)}$. Noting again that $v^{(i(k))}$ and $\overline{v^{(i(k))}}$ are CR words, we obtain $\boldsymbol{Z}^{(d(m)-1)}\,{}^{r}\overline{\boldsymbol{Z}^{(d(m)-1)}} = v^{(i)}0f0\overline{v^{(i(k_1))}}0\cdots0v^{(i)}\,v^{(i)}1\cdots1\overline{v^{(i(k_1))}}1\,{}^{r}\overline{f}1v^{(i)}$. Replacing the repetition $v^{(i)}v^{(i)}$ that occurs twice in a circular order each by single word $v^{(i)}$ respectively, we obtain a CR sequence $\boldsymbol{X} = v^{(i)}0f0\overline{v^{(i(k_1))}}0\cdots0v^{(i)}1\cdots1\overline{v^{(i(k_1))}}1\,{}^{r}\overline{f}1$ of length 2^{2m+1}. It is easy to check that the obtained CR sequence $\boldsymbol{X}$ is in the de Bruijn sequences of length 2^{2m+1}.

ii) We consider the case that $v^{(i)}$ is the neutral vertex, i.e., $v^{(i)} = v^{(i(k_0))}$ or $v^{(i)} = \overline{v^{(i(k_0))}}$. We have to consider both cases. However, we only consider the case that $v^{(i)} = v^{(i(k_0))}$ since we have $H_{v^{(i)}} \simeq H_{\overline{v^{(i)}}}$. Then, $\boldsymbol{Y}^\infty$ may be written in the form of $\Phi(\alpha_{j_1})f\Phi(\beta_{j_1})g\Phi(\alpha_{j_1})f\cdots$, where α_{j_1} is the leftmost negative symbol in the corresponding balanced Dyck word, and $\overline{v^{(k_0)}}\lambda v^{(k_0)}$ appear exactly twice in $\Phi(\alpha_{j_1})f\Phi(\beta_{j_1})g$. We have to examine two cases, namely $\Phi(\alpha_{j_1}) = 1\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}0$ and $\Phi(\beta_{j_1}) = 0\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}1$, or $\Phi(\alpha_{j_1}) = 0\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}1$ and $\Phi(\beta_{j_1}) = 1\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}0$. However, we only consider the former case since the processes of constructing a CR sequence from $\boldsymbol{Y}$ are exactly the same in both cases. Then, $\boldsymbol{Y}^\infty$ may be written uniquely in the form of $v^{(i(k_0))}0f0\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}1g1\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}0f\cdots$. We transform $v^{(i(k_0))}0f0\overline{v^{(i(k_0))}}\lambda v^{(i(k_0))}1g1\overline{v^{(i(k_0))}}\lambda$ in $\boldsymbol{Y}^\infty$ into $v^{(i(k_0))}0f0\overline{v^{(i(k_0))}}\lambda\,{}^{r}\overline{v^{(i(k_0))}1g1\overline{v^{(i(k_0))}}}\,\lambda = v^{(i(k_0))}0f0\overline{v^{(i(k_0))}}\lambda\overline{v^{(i(k_0))}}0\,{}^{r}\overline{g}0v^{(i(k_0))}\lambda$. After deleting two λ's, replace the repetition $\overline{v^{(i(k_0))}}\,\overline{v^{(i(k_0))}}$ by single words $\overline{v^{(i(k_0))}}$, then we obtain $v^{(i(k_0))}0f0\overline{v^{(i(k_0))}}0\,{}^{r}\overline{g}0v^{(i(k_0))}$, which we use $\boldsymbol{Z}^{(1)}$ to denote. By using exactly the same procedure as in the case i) above, we inductively obtain $\boldsymbol{Z}^{(d(m)-1)}$. Modifying $\boldsymbol{Z}^{(d(m)-1)}\,{}^{r}\overline{\boldsymbol{Z}^{(d(m)-1)}}$ similarly as in the case i) above, we obtain a CR sequence $\boldsymbol{X}$ in the de Bruijn sequences of length 2^{2m+1}.

Conversely, when we are given a CR sequence $\boldsymbol{X}$ in the de Bruijn sequences of length 2^{2m+1}, in view of Lemma 3, we find in $\boldsymbol{X}$ or $\overline{\boldsymbol{X}}$ a unique $v^{(i)} \in \mathcal{V}_{2m+1}^{CR}$ 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 $\boldsymbol{X}$ or $\overline{\boldsymbol{X}}$. This correspondence is two-to-one and onto. Since $\boldsymbol{X} \not\simeq \overline{\boldsymbol{X}}$ for $n \ge 3$ [6], corresponding to a CR sequence $\boldsymbol{X}$, $\overline{\boldsymbol{X}}$ gives a distinct CR sequence. Hence the above procedures for $v^{(i)} \in \mathcal{V}_{2m+1}^{CR}$ as a whole exhaust all pairs $(\boldsymbol{X}, \overline{\boldsymbol{X}})$ of CR sequences in the de Bruijn sequences of length 2^{2m+1}.

Eventually, we obtain the following.

Theorem 1 *For the case that* $m\,(\ge 4)$ *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 [3], we have completely solved the fundamental problem posed by Fredricksen in [1] on existence of CR sequences in the de Bruijn sequences of length 2^{2m+1} $(m \ge 1)$.

References

[1] H. Fredricksen, "A Survey of Full Length Nonlinear Shift Register Cycle Algorithm," *SIAM Review*, vol. 24, pp. 195–221, 1982.

[2] 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.

[3] 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, IEICE*, vol. 5, pp. 235–249, 2014.

[4] T. Hamachi and K. Inoue, "Embedding of Shifts of Finite Type into the Dyck Shift," *Monatshefte für Mathematik*, vol. 145, pp. 107–129, 2005.

[5] J. E. Hopcroft and J. D. Ulman, *Introduction to Automata Theory, Languages, and Comuputation*, Addison-Wesley, 1979.

[6] 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.