Constructing de Bruijn sequences by concatenating smaller universal cycles

Daniel Gabric* and Joe Sawada†

University of Guelph, Canada

June 23, 2018

Abstract

We present sufficient conditions for when an ordering of universal cycles $\alpha_1, \alpha_2, \ldots, \alpha_m$ for disjoint sets $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_m$ can be concatenated together to obtain a universal cycle for $\mathbf{S} = \mathbf{S}_1 \cup \mathbf{S}_2 \cup \cdots \cup \mathbf{S}_m$. When $\mathbf{S}$ is the set of all k-ary strings of length n, the result of such a successful construction is a de Bruijn sequence. Our conditions are applied to generalize two previously known de Bruijn sequence constructions and then they are applied to develop three new de Bruijn sequence constructions.

1 Introduction

Let $\Sigma_k = \{0, 1, \ldots, k-1\}$ be an alphabet of $k \geq 2$ symbol and let Σ_k^n be the set of k-ary strings of length n. Given a non-empty subset $\mathbf{S}$ of Σ_k^n, a *universal cycle* for $\mathbf{S}$ is a sequence of length $|\mathbf{S}|$ that contains every string in $\mathbf{S}$ as a substring exactly once when the sequence is viewed circularly. A universal cycle is said to be a *de Bruijn sequence* in the case that $\mathbf{S} = \Sigma_k^n$. For example,

$$00011122212110120100210220 2$$

is a de Bruijn sequence for Σ_3^3. It is well known that de Bruijn sequences are in one-to-one correspondence with directed Euler cycles in a related de Bruijn graph. However, algorithms to find Euler cycles in graphs require that the graph be stored in memory, and the de Bruijn graph is exponential in size. Amazingly, a prefer-smallest greedy approach [9, 17] generates the lexicographically smallest de Bruijn sequence [10]; however, like other preference-based methods [2], including prefer-same [5, 11] and prefer-opposite [1] in the binary case, it also requires an exponential amount of memory. As a result, there has been significant research to efficiently construct de Bruijn sequences for arbitrary n and k. The majority of this work constructs de Bruijn sequences via a successor-rule, finding one symbol at a time using the previous n symbols. Of this work, most apply only to the case when $k = 2$ [4, 7, 8, 10, 15, 16, 24], although several approaches generalize to larger alphabets [3, 6, 10, 26]. In the best case, these algorithms require $O(n)$-time per symbol and use $O(n)$-space.

The most efficient constructions of de Bruijn sequences arise from a concatenation approach, with some generating each symbol in $O(1)$-amortized time using $O(n)$ space. However, very little is known about these constructions in general. The first such construction was given by Fredericksen and Maiorana [12]. To

---

*E-mail: `dgabric@uoguelph.ca`

†E-mail: `jsawada@uoguelph.ca`

describe their approach we need the following two definitions. A *necklace* is the lexicographically smallest string in an equivalence class under rotation. The *periodic reduction* of a string $\alpha = a_1 a_2 \cdots a_n$ is $a_1 a_2 \cdots a_p$ where p is the smallest integer such that $\alpha = (a_1 a_2 \cdots a_p)^{n/p}$, where exponentiation denotes concatenation. They show that the lexicographically smallest de Bruijn sequence[1] can be constructed by concatenating together the periodic reductions of all k-ary necklaces of length n listed in lexicographic order. An analysis in [19] shows that this construction generates each symbol in $O(1)$-amortized time using $O(n)$ space.

Example 1   The set of necklaces of length $n = 4$ when $k = 2$ is $\{0000, 0001, 0011, 0101, 0111, 1111\}$. By concatenating the periodic reduction of each necklace in lexicographic order we obtain the following de Bruijn sequence for Σ_2^4

$$0 \cdot 0001 \cdot 0011 \cdot 01 \cdot 0111 \cdot 1$$

where $\cdot$ is used to denote concatenation for clarity.

There is a very subtle point in the description of this algorithm. Are the periodic reductions listed in lexicographic order, or are the necklaces first listed in lexicographic order and *then* the periodic reductions are applied? It turns out, that when using lexicographic order, it does not matter; the two listings are equivalent. This is pointed out by Ruskey [18] who also describes the algorithm as *the concatenation of all Lyndon words whose length divide n in lexicographic order*, where *Lyndon words* are necklaces equal to their periodic reductions. Interestingly, these approaches are no longer equivalent when we consider a co-lexicographic (colex) order as pointed out in [4]. If we order the periodic reductions in colex order for $n = 4$ and $k = 2$ we obtain

$$0 \cdot 1 \cdot 01 \cdot 0001 \cdot 0011 \cdot 0111 = 0101000100110111$$

which is not a de Bruijn sequence (it has no substring 1111). However, by first ordering the necklaces in colex order and then taking their periodic reductions, it is proved by Dragan et al. [3] that for any $k \geq 2$ and $n \geq 1$ the result is a de Bruijn sequence for Σ_k^n. When $k = 2$ and $n = 4$ this construction produces the following de Bruijn sequence

$$0 \cdot 0001 \cdot 01 \cdot 0011 \cdot 0111 \cdot 1.$$

When $k = 2$, by applying an algorithm in [25] each symbol can be generated in $O(1)$-amortized time using $O(n)$ space. When $k > 2$, each symbol can be produced in $O(n)$ time [3], and it remains an open problem to improve this bound. For each of these two concatenation approaches, observe that the periodic reductions correspond to universal cycles for each necklace equivalence class.

The main result of this paper is to provide more general conditions for when universal cycles $\alpha_1, \alpha_2, \ldots, \alpha_m$ for disjoint subsets $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_m$ of Σ_k^n can be concatenated together to obtain a universal cycle for $\mathbf{S} = \mathbf{S}_1 \cup \mathbf{S}_2 \cup \cdots \cup \mathbf{S}_m$. Our results can be applied to:

- generalize the lexicographic concatenation scheme by Fredricksen and Maiorana [12],

- generalize the colex concatenation scheme by Dragan et al. [3],

- obtain a new de Bruijn sequence construction that in the binary case is observed to be equivalent to the successor-rule based algorithm in [26], and

- obtain two new binary de Bruijn sequence constructions based on co-necklaces (defined in Section 4.2).

---
[1]They actually show (equivalently) their approach produces the lexicographically largest de Bruijn sequence.

These results generalize preliminary work presented at WORDS 2017 [13].

In addition to the two concatenation schemes presented earlier, two others are known to construct universal cycles for subsets of binary strings. One generalizes Fredricksen and Maiorana's approach for binary strings with a minimum specified weight (number of 1s) [23], and another is based on cool-lex order that applies to binary strings in a given weight range [20–22]. Each algorithm constructs universal cycles in $O(1)$-amortize time per bit using $O(n)$-space. The sufficient conditions presented in this paper do not apply to these algorithms.

The remainder of this paper is presented as follows. In Section 2, we present background definitions and notation. In Section 3, we present our main results which provide sufficient conditions for when smaller universal cycles can be concatenated together to create a new larger universal cycle. In Section 4, we apply our conditions to generalize previously known de Bruijn sequence constructions and to develop three new and generalized de Bruijn sequence constructions.

2   Background Definitions and Notation

Let $\alpha = a_1 a_2 \cdots a_m$ and $\beta = b_1 b_2 \cdots b_n$ be two distinct k-ary strings. Then α comes before β in *lexicographic* (lex) order if α is a proper prefix of β or if $a_i < b_i$ for the smallest i where $a_i \neq b_i$. We say that α comes before β in *colexicographic* (colex) order if α is a proper suffix of β or if $a_i < b_i$ for the largest i where $a_i \neq b_i$. Given a set $\mathbf{S}$ of strings of arbitrary length, let

- $\text{lex}(\mathbf{S})$ denote the strings of $\mathbf{S}$ listed in lex order,

- $\text{revlex}(\mathbf{S})$ denote the strings of $\mathbf{S}$ listed in reverse lex order,

- $\text{colex}(\mathbf{S})$ denote the strings of $\mathbf{S}$ listed in colex order, and

- $\text{revcolex}(\mathbf{S})$ denote the strings of $\mathbf{S}$ listed in reverse colex order.

Example 2   Let $\mathbf{S} = \{0101, 21201, 12020, 000, 220, 02102\}$. Then

$$\text{lex}(\mathbf{S}) = 000, 0101, 02102, 12020, 21201, 220, \quad \text{colex}(\mathbf{S}) = 000, 12020, 220, 0101, 21201, 02102,$$

$$\text{revlex}(\mathbf{S}) = 220, 21201, 12020, 02102, 0101, 000, \quad \text{revcolex}(\mathbf{S}) = 02102, 21201, 0101, 220, 12020, 000.$$

Let $\alpha = a_1 a_2 \cdots a_s$ and $\beta = b_1 b_2 \cdots b_t$ be two strings with $s, t \geq n > 0$. Let $\text{suff}_n(\alpha)$ be the length n suffix of α and $\text{pre}_n(\alpha)$ be the length n prefix of α. For example, $\text{suff}_3(0032233) = 233$ and $\text{pre}_3(0032233) = 003$. The set of necklaces when $n = 5$ and $k = 2$ is

$$\mathbf{S} = \{00000, 00001, 00011, 00101, 00111, 01011, 01111, 11111\}.$$

When you order $\mathbf{S}$ in colex order, an interesting property between adjacent necklaces becomes apparent,

$$\text{colex}(\mathbf{S}) = 00000, 00001, 00101, 00011, 01011, 00111, 01111, 11111.$$

Observe that for any two adjacent necklaces σ and τ in the above listing, if j is the smallest index where τ is not 0 at index j, then σ and τ share a length $n - j$ suffix. For example, consider the adjacent necklaces 00111

and 01111. At index 2, 01111 is not 0, and in fact this is the smallest index for which this is true. With this index, we predict that the longest matching suffix is of length $5 - 2 = 3$, which we can easily verify to be true. The following two properties generalize this idea, and extend it to prefixes.

Suffix-related: Let $x \in \Sigma_k$, and let j be the smallest index of β such that $x \neq b_j$, or ∞ if no such j exists. Then the ordered pair of strings (α, β) is said to be *suffix-related* with respect to (x, n) if $j \leq n$ and $\text{suff}_{n-j}(\alpha) = \text{suff}_{n-j}(\beta)$.

Prefix-related: Let $x \in \Sigma_k$, and let j be the smallest index of α such that $x \neq a_{s-j}$, or ∞ if no such j exists. Then the ordered pair of strings (α, β) is said to be *prefix-related* with respect to (x, n) if $j \leq n$ and $\text{pre}_{n-j-1}(\alpha) = \text{pre}_{n-j-1}(\beta)$.

Example 3 Let $\alpha = 00001200$, $\beta = 02000200$, $x = 0$ and $n = 5$,

The smallest index j of β such that $b_j \neq x$ is $j = 2$. Note, $\text{suff}_{5-2}(00001200) = \text{suff}_{5-2}(02000200) = 200$. Thus, (α, β) are suffix-related with respect to $(0, 5)$.

The smallest $j > 0$ such that $a_{s-j} \neq x$ is $j = 2$. Note, $00 = \text{pre}_{5-2-1}(00001200) \neq \text{pre}_{5-2-1}(02000200) = 02$. Thus, (α, β) are not prefix-related with respect to $(0, 5)$.

Let $\text{ext}_n(\alpha) = \alpha^t$, where t is the smallest integer so $t|\alpha| \geq n$. Let $\mathbf{sub}_n(\alpha)$ be the set of all length n substrings in the cyclic string α. For example, $\mathbf{sub}_3(01201) = \{012, 120, 201, 010, 101\}$ and $\mathbf{sub}_5(02) = \{02020, 20202\}$.

3 Concatenating Universal Cycles

Let $\mathbf{S}$ be a non-empty subset of Σ_k^n. A partition of $\mathbf{S}$ into subsets $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_m$ is called a *UC-partition* if there is a universal cycle α_i for each $\mathbf{S}_i$, $1 \leq i \leq m$.

Example 4 The following sets

$$\mathbf{S}_1 = \{00000, 00001, 00011, 00111, 01111, 11111, 11110, 11100, 11000, 10000\},$$
$$\mathbf{S}_2 = \{00100, 01001, 10011, 00110, 01101, 11011, 10110, 01100, 11001, 10010\},$$
$$\mathbf{S}_3 = \{00010, 00101, 01011, 10111, 01110, 11101, 11010, 10100, 01000, 10001\},$$
$$\mathbf{S}_4 = \{01010, 10101\},$$

together form a UC-partition of Σ_2^5 with universal cycles

$$\alpha_1 = 0000011111, \ \alpha_2 = 0010011011, \ \alpha_3 = 0001011101, \ \alpha_4 = 01.$$

The sets are pairwise disjoint and their union is Σ_2^5.

Given a UC-partition for a set $\mathbf{S}$ along with their corresponding universal cycles, we present conditions for when the smaller universal cycles can be concatenated together to obtain a universal cycle for $\mathbf{S}$.

Theorem 3.1 *Let* $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_m$ *be a UC-partition of* $\mathbf{S} \subseteq \Sigma_k^n$ *with universal cycles* $\alpha_1, \alpha_2, \ldots, \alpha_m$ *where* x *is the first symbol in* α_1 *and* $\mathcal{U}_{m,n} = \alpha_1 \alpha_2 \cdots \alpha_m$. *If the following three conditions hold,*

1. $|\alpha_1| \geq n$,

2. α_1 *has a largest prefix of consecutive* x's *out of all* $\alpha_i, 1 \leq i \leq m$,

3. *For each* $1 \leq i < m$, $(ext_n(\alpha_i), ext_n(\alpha_{i+1}))$ *are suffix-related with respect to* (x, n),

then $\mathcal{U}_{m,n}$ *is universal cycle for* $\mathbf{S}$ *and* $suff_n(\mathcal{U}_{m,n}) = suff_n(ext_n(\alpha_m))$.

Proof. The proof is by induction on m. In the base case when $m = 1$, $\mathcal{U}_{1,n} = \alpha_1$ is a universal cycle for $\mathbf{S}_1$, and by assumption $|\alpha_1| \geq n$ so $\text{suff}_n(\mathcal{U}_{1,n}) = \text{suff}_n(\text{ext}_n(\alpha_1))$. Inductively, assume $\mathcal{U}_{m-1,n}$ is a universal cycle for $\mathbf{S} - \mathbf{S}_m$, and $\text{suff}_n(\mathcal{U}_{m-1,n}) = \text{suff}_n(\text{ext}_n(\alpha_{m-1}))$, for $m > 1$. Consider $\mathcal{U}_{m,n} = \mathcal{U}_{m-1,n}\alpha_m$. Let:

- $\alpha_1 = g_1 g_2 \cdots g_q$,

- $\text{suff}_n(\text{ext}_n(\alpha_{m-1})) = a_1 a_2 \cdots a_n$,

- $\text{ext}_n(\alpha_m) = x^{j-1} b_j b_{j+1} \cdots b_s$,

where j is the smallest index where $b_j \neq x$. First we show that $\text{suff}_n(\mathcal{U}_{m,n}) = \text{suff}_n(\text{ext}_n(\alpha_m))$. If $|\alpha_m| \geq n$, clearly $\text{suff}_n(\alpha_m)$ is a suffix of $\mathcal{U}_{m-1,n}\alpha_m$. If $|\alpha_m| < n$, $\text{suff}_n(\text{ext}_n(\alpha_{m-1}))$ appears as a suffix of $\mathcal{U}_{m-1,n}$ by the inductive hypothesis. By assumption $\text{suff}_{n-j}(\text{ext}_n(\alpha_m)) = \text{suff}_{n-j}(\text{ext}_n(\alpha_{m-1}))$, so a suffix of $\mathcal{U}_{m,n}$ will be $\beta = \text{suff}_{n-j}(\text{ext}_n(\alpha_m))\alpha_m$, which shares a suffix of length $\min(|\beta|, |\text{ext}_n(\alpha_m)|)$ with $\text{ext}_n(\alpha_m)$. Since $j \leq |\alpha_m|$, it must be the case that $|\beta| = n - j + |\alpha_m| \geq n$, and thus $\min(|\beta|, |\text{ext}_n(\alpha_m)|) \geq n$. Therefore, $\text{suff}_n(\mathcal{U}_{m,n}) = \text{suff}_n(\text{ext}_n(\alpha_m))$. Now we prove that $\mathcal{U}_{m,n}$ is a universal cycle for $\mathbf{S}$. By the inductive hypothesis, $\mathcal{U}_{m,n}$ will contain all of the strings in $\mathbf{S} - \mathbf{S}_m$ except for possibly the strings $\{a_2 a_3 \cdots a_n g_1, a_3 a_4 \cdots a_n g_1 g_2, \ldots, a_n g_1 \cdots g_{n-1}\}$ which were involved in the wraparound. However, we know that $\text{suff}_{n-j}(\text{ext}_n(\alpha_m)) = \text{suff}_{n-j}(\text{ext}_n(\alpha_{m-1}))$ and $\text{suff}_n(\text{ext}_n(\alpha_m)) = \text{suff}_n(\mathcal{U}_{m,n})$. This implies that each string in $\{a_{j+1} a_{j+2} \cdots a_n g_1 \cdots g_j, a_{j+2} a_{j+3} \cdots a_n g_1 \cdots g_{j+1}, \ldots, a_n g_1 \cdots g_{n-1}\}$ occurs as a substring in the wraparound of the cyclic $\mathcal{U}_{m,n}$. Furthermore, the strings $\{a_2 a_3 \cdots a_n x, a_3 a_4 \cdots a_n x x, \ldots, a_j a_{j+1} \cdots a_n x^{j-1}\}$ exist within $\mathcal{U}_{m,n}$ because α_m has prefix $x^{j-1} = g_1 \cdots g_{j-1}$. Thus, the cyclic $\mathcal{U}_{m,n}$ contains each string in $\mathbf{S} - \mathbf{S}_m$ as a substring. Finally, we show that all strings in $\mathbf{S}_m$ occur as a substring in $\mathcal{U}_{m,n}$ (when considered cyclically). Those that are not trivially substrings of α_m occur either in the wrap-around or have their prefix as a suffix in $\mathcal{U}_{m-1,n}$ and suffix in a prefix of α_m. Let $t = s - n$, $i = t + j$, and

- $\mathbf{T}_1 = \{b_{t+2} b_{t+3} \cdots b_s x, b_{t+3} b_{t+4} \cdots b_s x x, \ldots, b_i b_{i+1} \cdots b_s x^{j-1}\}$,

- $\mathbf{T}_2 = \{b_{i+1} b_{i+2} \cdots b_s x^{j-1} b_j, b_{i+2} b_{i+3} \cdots b_s x^{j-1} b_j b_{j+1}, \ldots, b_{z+1} \cdots b_n x^{j-1} b_j b_{i+1} \cdots b_z\}$,

where $z = n$ if $|\alpha_m| > n$ and $z = |\alpha_m|$ otherwise. Notice that $\mathbf{T}_1$ and $\mathbf{T}_2$ together cover all length n substrings in the wraparound of α_m. Each string in $\mathbf{T}_1$ occurs in the wraparound of $\mathcal{U}_{m,n}$ since $\text{suff}_n(\text{ext}_n(\alpha_m)) = \text{suff}_n(\mathcal{U}_{m,n})$, and the prefix of α_1 has a run of x at least as big as the run in the prefix of α_m. Since $\text{suff}_{n-j}(\text{ext}_n(\alpha_m)) = \text{suff}_{n-j}(\text{ext}_n(\alpha_{m-1}))$, each string in $\mathbf{T}_2$ has a prefix in $\mathcal{U}_{m-1,n}$ and a suffix in α_m, so each string in $\mathbf{T}_2$ occurs as a substring of $\mathcal{U}_{m,n}$. We have shown that $\mathbf{S} \subseteq \text{sub}_n(\mathcal{U}_{m,n})$ and by construction $|\mathcal{U}_{m,n}| = |\mathbf{S}|$. Thus, $\mathcal{U}_{m,n}$ is a universal cycle for $\mathbf{S}$. $\qquad\square$

Example 5   From Example 4, the UC-partition $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \mathbf{S}_4$ of Σ_2^5 with universal cycles $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ satisfies the conditions of Theorem 3.1. Thus, the concatenation

$$\alpha_1\alpha_2\alpha_3\alpha_4 = 0000011111 \cdot 0010011011 \cdot 0001011101 \cdot 01$$

is a universal cycle for Σ_2^5.

Let $\text{rev}(\alpha)$ denote the reverse of the string α. Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be a list of strings of length at least n where (α_i, α_{i+1}) are prefix-related with respect to some (x, n) for $1 \le i < m$. Then $(\text{rev}(\alpha_{i+1}), \text{rev}(\alpha_i))$ are suffix-related with respect to (x, n). Thus Corollary 3.2 follows from Theorem 3.1.

Corollary 3.2 *Let $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_m$ be a UC-partition of $\mathbf{S} \subseteq \Sigma_k^n$ with universal cycles $\alpha_1, \alpha_2, \ldots, \alpha_m$ where x is the last symbol in α_m and $\mathcal{U}_{m,n} = \alpha_1\alpha_2\cdots\alpha_m$. If the following three conditions hold,*

1. *$|\alpha_m| \ge n$,*

2. *α_m has a largest suffix of consecutive x's out of all $\alpha_i, 1 \le i \le m$,*

3. *For each $1 \le i < m$, $(ext_n(\alpha_i), ext_n(\alpha_{i+1}))$ are prefix-related with respect to (x, n),*

then $\mathcal{U}_{m,n}$ is universal cycle for $\mathbf{S}$ and $pre_n(\mathcal{U}_{m,n}) = pre_n(ext_n(\alpha_1))$.

4   New Universal Cycle Concatenation Constructions

In this section we apply the results from the previous section to produce five new universal cycles based on concatenating together smaller universal cycles. A direct consequence of each result is a unique de Bruijn sequence construction; the first two were previously known and the last three are new.

Each of our constructions follows the approach outlined in Section 1 of concatenating the periodic reductions of a listing of strings. The function UC is defined on a listing of strings $\mathcal{L} = \alpha_1, \alpha_2, \ldots, \alpha_j$ as follows, where $pr(\alpha)$ is the periodic reduction of α:

$$\text{UC}(\mathcal{L}) = pr(\alpha_1)pr(\alpha_2)\cdots pr(\alpha_j).$$

Example 6   If $\mathcal{L} = 1111, 1212, 13213, 23131, 32312, 331331$, then

$$\text{UC}(\mathcal{L}) = 1 \cdot 12 \cdot 13213 \cdot 23131 \cdot 32312 \cdot 331.$$

In the next subsection we outline three concatenation constructions based on necklaces. Then, we define co-necklaces and use them to outline two more constructions.

4.1 Necklaces

Recall that a necklace is the lexicographically smallest string in an equivalence class of strings under rotation. Let $\mathbf{Neck}_k(n)$ denote the set of k-ary necklaces of length n. For example, $\mathbf{Neck}_3(3) = \{000, 001, 002, 011, 012, 021, 022, 111, 112, 122, 222\}$. It is well known that $\{\mathbf{sub}_n(\alpha) : \alpha \in \mathbf{Neck}_k(n)\}$ is a partition of Σ_k^n. By applying the results from Section 3, we obtain three de Bruijn sequence constructions that are generalized to some subsets of Σ_k^n.

4.1.1 Necklaces in Lex Order

Theorem 4.1 *For $n \geq 2$ and $m \geq 2$, let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the last m strings in $lex(\mathbf{Neck}_k(n))$. Then $\mathcal{U} = UC(\alpha_1, \alpha_2, \ldots, \alpha_m)$ is a universal cycle for $\mathbf{S} = \bigcup_{i=1}^{m} \mathbf{sub}_n(\alpha_i)$.*

Proof. Observe that $\alpha_m = (k-1)^n$ and $\alpha_{m-1} = (k-2)(k-1)^{n-1}$, so the last symbol in the sequence $\mathcal{U}$ is $x = k-1$. Note that $\alpha'_{m-1} = pr(\alpha_{m-1})pr(\alpha_m) = (k-2)(k-1)^n$ is a universal cycle for $\mathbf{S}_{m-1} = \mathbf{sub}_n(\alpha_{m-1}) \cup \mathbf{sub}_n(\alpha_m)$. Let $\mathbf{S}_i = \mathbf{sub}_n(\alpha_i)$ and $\alpha'_i = pr(\alpha_i)$ for $1 \leq i < m-1$. Then $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_{m-1}$ is a UC-partition of $\mathbf{S}$ with universal cycles $\alpha'_1, \alpha'_2, \ldots, \alpha'_{m-1}$ where $x = k-1$. To prove that $\mathcal{U}$ is a universal cycle for $\mathbf{S}$ we show that the three conditions of Corollary 3.2 hold.

1. Clearly $|\alpha'_{m-1}| \geq n$.

2. α'_{m-1} has suffix x^n, which must be maximal since all of the universal cycles are disjoint.

3. We must show that consecutive strings in $\text{ext}_n(\alpha'_1), \text{ext}_n(\alpha'_2), \ldots, \text{ext}_n(\alpha'_{m-1})$ are prefix-related with respect to (x, n). Notice that α_{m-1} is a length n prefix of $\text{ext}_n(\alpha'_{m-1})$ and $\text{ext}_n(\alpha'_i) = \alpha_i$ for $1 \leq i < m-1$. So we only need to show that (α_i, α_{i+1}) are prefix-related with respect to (x, n) for $1 \leq i < m-1$. The proof is by contradiction. Let $\alpha_i = a_1 a_2 \cdots a_n$ and $\alpha_{i+1} = b_1 b_2 \cdots b_n$ for some $1 \leq i < m-1$. Let j be the smallest index of α_i such that $a_{n-j} \neq x$. Suppose $a_1 a_2 \cdots a_{n-j-1} \neq b_1 b_2 \cdots b_{n-j-1}$. Then there exists some smallest $s < n-j$ such that $a_s \neq b_s$. Since α_i comes before α_{i+1} in lex order, then $a_s < b_s$. However, since α_i is a necklace, then $\gamma = a_1 a_2 \cdots a_s (k-1)^{n-s}$ will also be a necklace. But this means that γ comes between α_i and α_{i+1} in lex order, which is a contradiction. So $a_1 a_2 \cdots a_{n-j-1} = b_1 b_2 \cdots b_{n-j-1}$, which implies $\text{pre}_{n-j-1}(\alpha_i) = \text{pre}_{n-j-1}(\alpha_{i+1})$. Thus the pair of strings (α_i, α_{i+1}) are prefix-related with respect to (x, n). □

When $m = |\mathbf{Neck}_k(n)|$, the above theorem yields the following corollary which describes an equivalent construction to the (lexicographically smallest) de Bruijn sequence construction from [12].

Corollary 4.2 *For $n \geq 2$, $UC(lex(\mathbf{Neck}_k(n)))$ is a de Bruijn sequence for Σ_k^n.*

Example 7 Consider the set $\mathbf{Neck}_2(6)$ listed in lex order:

$000000, 000001, 000011, 000101, 000111, 001001, 001011, 001101, 001111, 010101, 010111, 011011, 011111, 111111$.

By Corollary 4.2, the following is a de Bruijn sequence,

$$0 \cdot 000001 \cdot 000011 \cdot 000101 \cdot 000111, 001 \cdot 001011 \cdot 001101 \cdot 001111 \cdot 01 \cdot 010111 \cdot 011 \cdot 011111 \cdot 1.$$

By Theorem 4.1, $011 \cdot 011111 \cdot 1$, is a universal cycle for $\mathbf{sub}_6(011) \cup \mathbf{sub}_6(011111) \cup \mathbf{sub}_6(1)$.

4.1.2 Necklaces in Colex Order

> **Theorem 4.3** *For $n \geq 2$ and $m \geq 2$, let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the first m strings in $colex(\mathbf{Neck}_k(n))$. Then $\mathcal{U} = UC(\alpha_1, \alpha_2, \ldots, \alpha_m)$ is a universal cycle for $\mathbf{S} = \bigcup_{i=1}^{m} \mathbf{sub}_n(\alpha_i)$.*

Proof. Observe that $\alpha_1 = 0^n$ and $\alpha_2 = 0^{n-1}1$, so the first symbol in the sequence $\mathcal{U}$ is $x = 0$. Note that $\alpha_1' = pr(\alpha_1)pr(\alpha_2) = 01^n$ is a universal cycle for $\mathbf{S}_1 = \mathbf{sub}_n(\alpha_1) \cup \mathbf{sub}_n(\alpha_2)$. Let $\mathbf{S}_i = \mathbf{sub}_n(\alpha_{i+1})$ and $\alpha_i' = pr(\alpha_{i+1})$ for $2 \leq i < m$. Then $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_{m-1}$ is a UC-partition of $\mathbf{S}$ with universal cycles $\alpha_1', \alpha_2', \ldots, \alpha_{m-1}'$ where $x = 0$. To prove that $\mathcal{U}$ is a universal cycle for $\mathbf{S}$ we show that the three conditions of Theorem 3.1 hold.

1. Clearly $|\alpha_1'| \geq n$.

2. α_1' has prefix x^n, which must be maximal since all of the universal cycles are disjoint.

3. We must show that consecutive strings in $\text{ext}_n(\alpha_1'), \text{ext}_n(\alpha_2'), \ldots, \text{ext}_n(\alpha_{m-1}')$ are suffix-related with respect to (x, n). Notice that α_2 is a length n suffix of $\text{ext}_n(\alpha_1')$ and $\text{ext}_n(\alpha_i') = \alpha_{i+1}$ for $2 \leq i < m$. So we only need to show that (α_i, α_{i+1}) are suffix-related with respect to (x, n) for $2 \leq i < m$. The proof is by contradiction. Let $\alpha_i = a_1 a_2 \cdots a_n$ and $\alpha_{i+1} = b_1 b_2 \cdots b_n$ for some $1 < i < m$. Let j be the smallest index of α_{i+1} such that $b_j \neq x$. Suppose $a_{j+1} a_{j+2} \cdots a_n \neq b_{j+1} b_{j+2} \cdots b_n$. Then there exists some largest $s > j$ such that $a_s \neq b_s$. Since α_i comes before α_{i+1} in colex order, then $a_s < b_s$. However, since α_{i+1} is a necklace, then $\gamma = 0^{s-1} b_s b_{s+1} \cdots b_n$ will also be a necklace. But this means that γ comes between α_i and α_{i+1} in colex order, which is a contradiction. So $a_{j+1} a_{j+2} \cdots a_n = b_{j+1} b_{j+2} \cdots b_n$, which implies $\text{suff}_{n-j}(\alpha_i) = \text{suff}_{n-j}(\alpha_{i+1})$. Thus the pair of strings (α_i, α_{i+1}) are suffix-related with respect to (x, n). $\qquad\square$

When $m = |\mathbf{Neck}_k(n)|$, the above theorem yields the following corollary which describes a construction equivalent to the de Bruijn sequence construction from [3].

Corollary 4.4 *For $n \geq 2$, $UC(colex(\mathbf{Neck}_k(n)))$ is a de Bruijn sequence for Σ_k^n.*

> **Example 8** Consider the set $\mathbf{Neck}_2(6)$ listed in colex order:
>
> $000000, 000001, 001001, 000101, 010101, 001101, 000011, 001011, 011011, 000111, 010111, 001111, 011111, 111111.$
>
> By Corollary 4.4, the following is a de Bruijn sequence:
>
> $\quad 0 \cdot 000001 \cdot 001 \cdot 000101 \cdot 01 \cdot 001101 \cdot 000011 \cdot 001011 \cdot 011 \cdot 000111 \cdot 010111 \cdot 001111 \cdot 011111 \cdot 1.$
>
> By Theorem 4.3, $0 \cdot 000001 \cdot 001$, is a universal cycle for $\mathbf{sub}_6(0) \cup \mathbf{sub}_6(000001) \cup \mathbf{sub}_6(001)$.

4.1.3 Rotations of Necklaces in Reverse Lex Order

Consider $\mathbf{Neck}_2(5)$ listed in reverse lex order: $L = 11111, 01111, 01011, 00111, 00101, 00011, 00001, 00000$. Observe that $UC(L) = 1011110101100111001010100011000010$ is not a de Bruijn sequence since, when considered in a cyclic way, it does not contain the substring 00000. However by looking at specific rotations

of these strings in reverse lex order we obtain positive results. Let $\mathbf{R}_k(n)$ be the set of all $\alpha = a_1a_2\cdots a_n$ such that $a_{i+1}a_{i+2}\cdots a_n a_1 \cdots a_i$ is in $\mathbf{Neck}_k(n)$, where i is the largest index of α such that $a_i \neq 0$. For example, $\mathbf{R}_2(5) = \{11111, 11110, 10110, 11100, 10100, 11000, 10000, 00000\}$.

Theorem 4.5 *For $n \geq 2$ and $m \geq 2$, let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the last m strings in revlex$(\mathbf{R}_k(n))$. Then $\mathcal{U} = UC(\alpha_1, \alpha_2, \ldots, \alpha_m)$ is a universal cycle for $\mathbf{S} = \bigcup_{i=1}^m \mathbf{sub}_n(\alpha_i)$.*

Proof. Observe that $\alpha_m = 0^n$ and $\alpha_{m-1} = 10^{n-1}$, so the last symbol in the sequence $\mathcal{U}$ is $x = 0$. Note that $\alpha'_{m-1} = pr(\alpha_{m-1})pr(\alpha_m) = 10^n$ is a universal cycle for $\mathbf{S}_{m-1} = \mathbf{sub}_n(\alpha_{m-1}) \cup \mathbf{sub}_n(\alpha_m)$. Let $\mathbf{S}_i = \mathbf{sub}_n(\alpha_i)$ and $\alpha'_i = pr(\alpha_i)$ for $1 \leq i < m-2$. Then $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_{m-1}$ is a UC-partition of $\mathbf{S}$ with universal cycles $\alpha'_1, \alpha'_2, \ldots, \alpha'_{m-1}$ where $x = 0$. To prove that $\mathcal{U}$ is a universal cycle for $\mathbf{S}$ we show that the three conditions of Corollary 3.2 hold.

1. Clearly $|\alpha'_{m-1}| \geq n$.

2. α'_{m-1} has suffix x^n, which must be maximal since all of the universal cycles are disjoint.

3. We must show that consecutive strings in $\text{ext}_n(\alpha'_1), \text{ext}_n(\alpha'_2), \ldots, \text{ext}_n(\alpha'_{m-1})$ are prefix-related with respect to (x, n). Notice that α_{m-1} is a length n prefix of $\text{ext}_n(\alpha'_{m-1})$ and $\text{ext}_n(\alpha'_i) = \alpha_i$ for $1 \leq i < m-1$. So we only need to show that (α_i, α_{i+1}) are prefix-related with respect to (x, n) for $1 \leq i < m-1$. The proof is by contradiction. Let $\alpha_i = a_1a_2\cdots a_n$ and $\alpha_{i+1} = b_1b_2\cdots b_n$ for some $1 \leq i < m-1$. Let j be the smallest index of α_i such that $a_{n-j} \neq x$. Suppose $a_1a_2\cdots a_{n-j-1} \neq b_1b_2\cdots b_{n-j-1}$. Then there exists some smallest $s < n-j$ such that $a_s \neq b_s$. Since α_i comes before α_{i+1} in reverse lex order, then $a_s > b_s$. Let $\gamma = a_1a_2\cdots a_s 0^{n-s}$. Clearly γ is between α_i and α_{i+1} in reverse lex order, and $0^{n-s}a_1a_2\cdots a_s$ is a necklace since 0^{n-s} is the largest run of zeroes within the string, a contradiction. So $a_1a_2\cdots a_{n-j-1} = b_1b_2\cdots b_{n-j-1}$, which implies $\text{pre}_{n-j-1}(\alpha_i) = \text{pre}_{n-j-1}(\alpha_{i+1})$. Thus the pair of strings (α_i, α_{i+1}) are prefix-related with respect to (x, n). $\qquad\square$

Corollary 4.6 *For $n \geq 2$, $UC(revlex(\mathbf{R}_k(n)))$ is a de Bruijn sequence for Σ_k^n.*

The concatenation scheme described in the above corollary was originally motivated by considering the successor-rule based construction in [26]. Although we do not prove it here, the two constructions produce the same de Bruijn sequences when $k = 2$, but produce different sequences for $k > 2$.

Example 9 Consider the set $\mathbf{R}_2(6)$ listed in reverse lex order:

$111111, 111110, 111100, 111000, 110110, 110100, 110000, 101110, 101100, 101010, 101000, 100100, 100000, 000000.$

By Corollary 4.6, the following is a de Bruijn sequence:

$$1 \cdot 111110 \cdot 111100 \cdot 111000 \cdot 110 \cdot 110100 \cdot 110000 \cdot 101110 \cdot 101100 \cdot 10 \cdot 101000 \cdot 100 \cdot 100000 \cdot 0.$$

By Theorem 4.5, $100 \cdot 100000 \cdot 0$, is a universal cycle for $\mathbf{sub}_6(100) \cup \mathbf{sub}_6(100000) \cup \mathbf{sub}_6(0)$.

4.2 Co-necklaces

For this subsection we will be working over the binary alphabet $\Sigma_2 = \{0, 1\}$. Let α be a binary string and let $\overline{\alpha}$ denote its bitwise complement. We say that α is a *co-necklace* if $\alpha\overline{\alpha}$ is a necklace. The set of all co-necklaces of length 5 is $\{00000, 00010, 00100, 010101\}$. If α is a co-necklace, then we call $\alpha\overline{\alpha}$ an *extended co-necklace*. Let $\mathbf{coN}(n)$ denote the set of all extended co-necklaces for co-necklaces of length n. For example, $\mathbf{coN}(5) = \{0000011111, 0001011101, 0010011011, 0101010101\}$. It is well known that $\{\mathbf{sub}_n(\alpha) : \alpha \in \mathbf{coN}(n)\}$ is a partition of Σ_2^n; they correspond to the partition obtained from the complemented cycling register [14]. Using extended co-necklaces, we apply Theorem 3.1 and Corollary 3.2 to construct new universal cycles for subsets of Σ_2^n and ultimately to produce two new binary de Bruijn sequence constructions. Unlike necklaces, we cannot simply use lex or colex orderings. For example, neither

$$UC(\mathrm{lex}(\mathbf{coN}(5))) = 0000011111 \cdot 0001011101 \cdot 0010011011 \cdot 01, \text{ or}$$

$$UC(\mathrm{colex}(\mathbf{coN}(5))) = 01 \cdot 0001011101 \cdot 0010011011 \cdot 0000011111$$

are de Bruijn sequences, since neither contains the substring 10101.

4.2.1 Extended Co-necklaces in Reverse Colex Order

The following lemma will be useful in the proof of our universal cycle construction using the reverse colex order of extended co-necklaces.

Lemma 4.7 *Let $\alpha = a_1 a_2 \cdots a_{2n}$ and $\alpha' = b_1 b_2 \cdots b_{2n}$ be consecutive strings in the listing $\mathrm{revcolex}(\mathbf{coN}(n))$. Then $pre_n(\alpha)$ comes before $pre_n(\alpha')$ in colex order.*

Proof. Since $\alpha, \alpha' \in \mathbf{coN}(n)$, $\alpha = \beta\overline{\beta}$ and $\alpha' = \beta'\overline{\beta'}$ for some $\beta, \beta' \in \Sigma_2^n$. Because α comes before α' in reverse colex order and $\overline{\beta} \neq \overline{\beta'}$, there exists a largest index $n + 1 \leq i \leq 2n$ where $a_i \neq b_i$ and $a_i > b_i$. This implies that the largest index $1 \leq j \leq n$ such that $a_j \neq b_j$ is $j = i - n$. This means that $a_j = \overline{a_i}$ and $b_j = \overline{b_i}$ and thus $a_j < b_j$. Therefore $a_1 a_2 \cdots a_n = \beta = pre_n(\alpha)$ comes before $b_1 b_2 \cdots b_n = \beta' = pre_n(\alpha')$ in colex order. $\quad\square$

Theorem 4.8 *For $n \geq 2$ and $m \geq 1$, let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the first m strings in $\mathrm{revcolex}(\mathbf{coN}(n))$. Then $\mathcal{U} = UC(\alpha_1, \alpha_2, \ldots, \alpha_m)$ is a universal cycle for $\mathbf{S} = \bigcup_{i=1}^m \mathbf{sub}_n(\alpha_i)$.*

Proof. Observe that $\alpha_1 = 0^n 1^n$, so the first symbol in the sequence $\mathcal{U}$ is $x = 0$. Let $\mathbf{S}_i = \mathbf{sub}_n(\alpha_i)$ and $\alpha_i' = pr(\alpha_i)$ for $1 \leq i \leq m$. Then $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_m$ is a UC-partition of $\mathbf{S}$ with universal cycles $\alpha_1', \alpha_2', \ldots, \alpha_m'$ where $x = 0$. To prove that $\mathcal{U}$ is a universal cycle for $\mathbf{S}$ we show that the three conditions of Theorem 3.1 hold.

1. Clearly $|\alpha_1'| \geq n$.

2. α_1' has prefix x^n, which must be maximal since all of the universal cycles are disjoint.

3. We must show that consecutive strings in $\mathrm{ext}_n(\alpha_1'), \mathrm{ext}_n(\alpha_2'), \ldots, \mathrm{ext}_n(\alpha_m')$ are suffix-related with respect to (x, n). Notice that $pre_n(\mathrm{ext}_n(\alpha_i')) = pre_n(\alpha_i)$ for $1 \leq i \leq m$. So we only need to show that $(pre_n(\alpha_i), pre_n(\alpha_{i+1}))$ are suffix-related with respect to (x, n) for $1 \leq i < m$. The proof is by

contradiction. Let $\text{pre}_n(\alpha_i) = a_1 a_2 \cdots a_n$ and $\text{pre}_n(\alpha_{i+1}) = b_1 b_2 \cdots b_n$ for some $1 \le i < m$. Let j be the smallest index of α_{i+1} such that $b_j \ne x$. Suppose $a_{j+1} a_{j+2} \cdots a_n \ne b_{j+1} b_{j+2} \cdots b_n$. Then there exists some largest $i > j$ such that $a_i \ne b_i$. Since $\text{pre}_n(\alpha_i)$ comes before $\text{pre}_n(\alpha_{i+1})$ in colex order by Lemma 4.7, $a_i < b_i$. However, since $\text{pre}_n(\alpha_{i+1})$ is a co-necklace, then $\gamma = 0^{i-1} b_i b_{i+1} \cdots b_n$ will also be a co-necklace. But this means γ comes between $\text{pre}_n(\alpha_i)$ and $\text{pre}_n(\alpha_{i+1})$ in colex order, which is a contradiction. So $a_{j+1} a_{j+2} \cdots a_n = b_{j+1} b_{j+2} \cdots b_n$, which implies $\text{suff}_{n-j}(\text{pre}_n(\alpha_i)) = \text{suff}_{n-j}(\text{pre}_n(\alpha_{i+1}))$. Thus the pair of strings $(\text{pre}_n(\alpha_i), \text{pre}_n(\alpha_{i+1}))$ are suffix-related with respect to (x, n), which implies (α_i, α_{i+1}) are suffix-related with respect to (x, n). $\qquad\square$

When $m = |\mathbf{C}(n)|$, the above theorem yields the following corollary which describes a construction equivalent to the de Bruijn sequence construction from [13]. In that paper an algorithm is provided that generates the de Bruijn sequence in $O(1)$-amortized time per bit.

Corollary 4.9 *For $n \ge 2$, $UC(revcolex(\mathbf{coN}(n)))$ is a de Bruijn sequence for Σ_2^n.*

Example 10   Consider the set $\mathbf{coN}(6)$ listed in reverse colex order:

$$000000111111, 000100111011, 001100110011, 000010111101, 001010110101, 000110111001.$$

By Corollary 4.9, the following is a de Bruijn sequence:

$$000000111111 \cdot 000100111011 \cdot 0011 \cdot 000010111101 \cdot 001010110101 \cdot 000110111001.$$

By Theorem 4.8, $000000111111 \cdot 000100111011 \cdot 0011$, is a universal cycle for $\mathbf{sub}_6(000000111111) \cup \mathbf{sub}_6(000100111011) \cup \mathbf{sub}_6(0011)$.

4.2.2   Rotations of Extended Co-necklaces in Lex Order

Let $\mathbf{C}(n)$ be the set of all $\alpha = a_1 a_2 \cdots a_{2n}$ such that $a_{i+1} a_{i+2} \cdots a_{2n} a_1 \cdots a_i \in \mathbf{coN}(n)$, where i is the largest index of α such that $a_i \ne 0$. For example, $\mathbf{C}(5) = \{1111100000, 1011101000, 1001101100, 1010101010\}$.

Lemma 4.10 *Let $\alpha = a_1 a_2 \cdots a_{2n}$ and $\beta = b_1 b_2 \cdots b_{2n}$ be in $\mathbf{C}(n)$. If $(\text{pre}_n(\alpha), \text{pre}_n(\beta))$ are prefix-related with respect to (x, n) where $x \in \{0, 1\}$, then (α, β) are prefix-related with respect to $(1 - x, n)$.*

Proof. Assume that the pair of strings $(\text{pre}_n(\alpha), \text{pre}_n(\beta))$ are prefix-related with respect to (x, n). There exists a smallest index $1 \le i \le n$ such that $x \ne a_{n-i}$ and $\text{pre}_{n-i-1}(\text{pre}_n(\alpha)) = \text{pre}_{n-i-1}(\text{pre}_n(\beta))$. Since $\alpha = \text{pre}_n(\alpha)\overline{\text{pre}_n(\alpha)}$ and $\beta = \text{pre}_n(\beta)\overline{\text{pre}_n(\beta)}$, we have that the smallest index j such that $1 - x \ne a_{2n-j}$ is $j = i$ and we already know from before than $\text{pre}_{n-i-1}(\alpha) = \text{pre}_{n-i-1}(\beta)$. So (α, β) are prefix-related with respect to $(1 - x, n)$. $\qquad\square$

Theorem 4.11 *For $n \ge 2$ and $m \ge 1$, let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the last m strings in $lex(\mathbf{C}(n))$. Then $\mathcal{U} = UC(\alpha_1, \alpha_2, \ldots, \alpha_m)$ is a universal cycle for $\mathbf{S} = \bigcup_{i=1}^m \mathbf{sub}_n(\alpha_i)$.*

Proof. Observe that $\alpha_m = 1^n 0^n$, so the last symbol in the sequence $\mathcal{U}$ is $x = 0$. Let $\mathbf{S}_i = \mathbf{sub}_n(\alpha_i)$ and $\alpha_i' = pr(\alpha_i)$ for $1 \le i \le m$. Then $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_m$ is a UC-partition of $\mathbf{S}$ with universal cycles $\alpha_1', \alpha_2', \ldots, \alpha_m'$ where $x = 0$. To prove that $\mathcal{U}$ is a universal cycle for $\mathbf{S}$ we show that the three conditions of Corollary 3.2 hold.

1. Clearly $|\alpha'_m| \geq n$.

2. α'_m has suffix x^n, which must be maximal since all of the universal cycles are disjoint.

3. We must show that consecutive strings in $\text{ext}_n(\alpha'_1), \text{ext}_n(\alpha'_2), \ldots, \text{ext}_n(\alpha'_m)$ are prefix-related with respect to (x, n). Notice that $\text{pre}_n(\text{ext}_n(\alpha'_i)) = \text{pre}_n(\alpha_i)$ for $1 \leq i \leq m$. By Lemma 4.10 we only need show that $(\text{pre}_n(\alpha_i), \text{pre}_n(\alpha_{i+1}))$ are prefix-related with respect to $(1-x, n)$ for $1 \leq i < m$. The proof is by contradiction. Let $\text{pre}_n(\alpha_i) = a_1 a_2 \cdots a_n$ and $\text{pre}_n(\alpha_{i+1}) = b_1 b_2 \cdots b_n$ for some $1 \leq i < m$. Let j be the smallest index of $\text{pre}_n(\alpha_i)$ such that $a_{n-j} \neq 1 - x$. Suppose $a_1 a_2 \cdots a_{n-j-1} \neq b_1 b_2 \cdots b_{n-j-1}$. Then there exists some smallest $s < n - j$ such that $a_s \neq b_s$. Since $\text{pre}_n(\alpha_i)$ comes before $\text{pre}_n(\alpha_{i+1})$ in lex order, then $a_s < b_s$. Let $\gamma = a_1 a_2 \cdots a_s 1^{n-s}$. Clearly γ is between $\text{pre}_n(\alpha_i)$ and $\text{pre}_n(\alpha_{i+1})$ in lex order, and $0^{n-s} a_1 a_2 \cdots a_s$ is a co-necklace since 0^{n-s} is the largest run of ones within the string, a contradiction. So $a_1 a_2 \cdots a_{n-j-1} = b_1 b_2 \cdots b_{n-j-1}$, which implies $\text{pre}_{n-j-1}(\text{pre}_n(\alpha_i)) = \text{pre}_{n-j-1}(\text{pre}_n(\alpha_{i+1}))$. Thus the pair of strings $(\text{pre}_n(\alpha_i), \text{pre}_n(\alpha_{i+1}))$ are prefix-related with respect to $(1-x, n)$, which implies (α_i, α_{i+1}) are prefix-related with respect to (x, n).

$\square$

When $m = |\mathbf{C}(n)|$, the above theorem yields the following corollary which describes a new de Bruijn sequence construction.

Corollary 4.12 *For $n \geq 2$, $UC(lex(\mathbf{C}(n)))$ is a de Bruijn sequence for Σ_2^n.*

Example 11   Consider the set $\mathbf{C}(6)$ listed in lex order:

$$100111011000, 101011010100, 101111010000, 110011001100, 110111001000, 111111000000.$$

By Corollary 4.12, the following is a de Bruijn sequence:

$$100111011000 \cdot 101011010100 \cdot 101111010000 \cdot 1100 \cdot 110111001000 \cdot 111111000000.$$

By Theorem 4.11, $1100 \cdot 110111001000 \cdot 111111000000$, is a universal cycle for $\mathbf{sub}_6(1100) \cup \mathbf{sub}_6(110111001000) \cup \mathbf{sub}_6(111111000000)$.

---

5   Conclusions and open problems

In this paper we presented conditions for when a listing of universal cycles can be concatenated together to produce longer universal cycles and ultimately, de Bruijn sequences. By applying the conditions, we generalized two previously known de Bruijn sequence concatenation-based constructions and discovered three new ones. De Bruijn sequences from each of these five constructions for $n = 6$ and $k = 2$ are given below.

Construction	de Bruijn Sequence for $n = 6, k = 2$
$UC(lex(\mathbf{Neck}_2(6)))$	0000001000011000101000111001001011001101001111010101110110111111
$UC(colex(\mathbf{Neck}_2(6)))$	0000001001000101010011010000110010110110001110101110011110111111
$UC(revlex(\mathbf{R}_2(6)))$	1111110111100111000110110100110000101110101100101010001001000000
$UC(revcolex(\mathbf{coN}(6)))$	0000001111110001001110110011000010111101001010110101000110111001
$UC(lex(\mathbf{C}(6)))$	1001110110001010110101001011110100001100110111001000111111000000

We conclude by posing some open problems.

1. Can k-ary necklaces be generated in $O(n)$-amortized time per string in colex order? If so, then the de Bruijn sequence UC(colex($\mathbf{Neck}_k(n)$)) can be generated in $O(1)$-amortized time per symbol. There is an efficient algorithm in the binary case [25].

2. Can the de Bruijn sequences UC(revlex($\mathbf{R}_k(n)$)) and UC(lex($\mathbf{C}(n)$)) be generated in $O(1)$-amortized time per symbol?

References

[1] A. Alhakim. A simple combinatorial algorithm for de Bruijn sequences. *The American Mathematical Monthly*, 117(8):728–732, 2010.

[2] A. Alhakim. Spans of preference functions for de Bruijn sequences. *Discrete Applied Mathematics*, 160(7-8):992 – 998, 2012.

[3] P. B. Dragon, O. I. Hernandez, J. Sawada, A. Williams, and D. Wong. Constructing de Bruijn sequences with co-lexicographic order: The k-ary Grandmama sequence. *European Journal of Combinatorics, to appear*, 2017.

[4] P. B. Dragon, O. I. Hernandez, and A. Williams. The grandmama de Bruijn sequence for binary strings. In *Proceedings of LATIN 2016: Theoretical Informatics: 12th Latin American Symposium, Ensenada, Mexico*, pages 347–361. Springer Berlin Heidelberg, 2016.

[5] C. Eldert, H. Gray, H. Gurk, and M. Rubinoff. Shifting counters. *AIEE Trans.*, 77:70–74, 1958.

[6] T. Etzion. An algorithm for constructing m-ary de Bruijn sequences. *Journal of Algorithms*, 7(3):331 – 340, 1986.

[7] T. Etzion. Self-dual sequences. *Journal of Combinatorial Theory, Series A*, 44(2):288 – 298, 1987.

[8] T. Etzion and A. Lempel. Construction of de Bruijn sequences of minimal complexity. *IEEE Transactions on Information Theory*, 30(5):705–709, September 1984.

[9] L. Ford. A cyclic arrangement of M-tuples. *Report No. P-1071, Rand Corporation, Santa Monica, California, April 23*, 1957.

[10] H. Fredricksen. The lexicographically least de Bruijn cycle. *Journal of Combinatorial Theory*, 9:1–5, 1970.

[11] H. Fredricksen. A survey of full length nonlinear shift register cycle algorithms. *Siam Review*, 24(2):195–221, 1982.

[12] H. Fredricksen and J. Maiorana. Necklaces of beads in k colors and k-ary de Bruijn sequences. *Discrete Math.*, 23:207–210, 1978.

[13] D. Gabric and J. Sawada. A de Bruijn sequence construction by concatenating cycles of the complemented cycling register. In *Combinatorics on Words - 11th International Conference, WORDS 2017, Montréal, QC, Canada, September 11-15, 2017, Proceedings*, pages 49–58, 2017.

[14] S. W. Golomb. *Shift Register Sequences*. Aegean Park Press, Laguna Hills, CA, USA, 1981.

[15] Y. Huang. A new algorithm for the generation of binary de Bruijn sequences. *J. Algorithms*, 11(1):44–51, 1990.

[16] C. J. A. Jansen, W. G. Franx, and D. E. Boekee. An efficient algorithm for the generation of de Bruijn cycles. *IEEE Transactions on Information Theory*, 37(5):1475–1478, Sep 1991.

[17] M. H. Martin. A problem in arrangements. *Bull. Amer. Math. Soc.*, 40(12):859–864, 1934.

[18] F. Ruskey. *Combinatorial Generation*. Working version (1i) edition, 1996.

[19] F. Ruskey, C. Savage, and T. M. Y. Wang. Generating necklaces. *J. Algorithms*, 13:414–430, 1992.

[20] F. Ruskey, J. Sawada, and A. Williams. De Bruijn sequences for fixed-weight binary strings. *SIAM Journal on Discrete Mathematics*, 26(2):605–617, 2012.

[21] J. Sawada, B. Stevens, and A. Williams. De Bruijn sequences for the binary strings with maximum specified density. In *Proceeding of 5th International Workshop on Algorithms and Computation (WALCOM 2011), New Dehli, India, LNCS*, 2011.

[22] J. Sawada, A. Williams, and D. Wong. Universal cycles for weight-range binary strings. In *Combinatorial Algorithms - 24th International Workshop, IWOCA 2013, Rouen, France, July 10-12, 2013, Revised Selected Papers*, pages 388–401, 2013.

[23] J. Sawada, A. Williams, and D. Wong. The lexicographically smallest universal cycle for binary strings with minimum specified weight. *Journal of Discrete Algorithms*, 28:31 – 40, 2014. StringMasters 2012 and 2013 Special Issue (Volume 1).

[24] J. Sawada, A. Williams, and D. Wong. A surprisingly simple de Bruijn sequence construction. *Discrete Math.*, 339:127–131, 2016.

[25] J. Sawada, A. Williams, and D. Wong. Necklaces and Lyndon words in colexicographic and binary reflected Gray code order. *Journal of Discrete Algorithms*, 46-47(Supplement C):25 – 35, 2017.

[26] J. Sawada, A. Williams, and D. Wong. A simple shift rule for k-ary de bruijn sequences. *Discrete Mathematics*, 340(3):524 – 531, 2017.