Constructing k-ary Orientable Sequences with Asymptotically Optimal Length

Daniel Gabrić
University of Guelph, Canada

Joe Sawada
University of Guelph, Canada

Abstract

An orientable sequence of order n over an alphabet $\{0, 1, \ldots, k-1\}$ is a cyclic sequence such that each length-n substring appears at most once *in either direction*. When $k = 2$, efficient algorithms are known to construct binary orientable sequences, with asymptotically optimal length, by applying the classic cycle-joining technique. The key to the construction is the definition of a parent rule to construct a cycle-joining tree of asymmetric bracelets. Unfortunately, the parent rule does not generalize to larger alphabets. Furthermore, unlike the binary case, a cycle-joining tree does not immediately lead to a simple successor-rule when $k \geq 3$ unless the tree has certain properties. In this paper, we derive a parent rule to derive a cycle-joining tree of k-ary asymmetric bracelets. This leads to a successor rule that constructs asymptotically optimal k-ary orientable sequences in $O(n)$ time per symbol using $O(n)$ space. In the special case when $n = 2$, we provide a simple construction of k-ary orientable sequences of maximal length.

1   Introduction

Given a set $\mathbf{S}$ of k-ary strings of length n, a *universal cycle* is a cyclic sequence of length $|\mathbf{S}|$ that contains each string in $\mathbf{S}$ as a substring exactly once. When $\mathbf{S}$ consists of *all* k-ary strings of length n, universal cycles are known as *de Bruijn sequences*. Universal cycles have been studied for many fundamental objects including permutations, subsets, and graphs [3, 6]. Universal cycles do not exist directly for permutations; however, efficient constructions exist using a shorthand representation [21, 34]. Universal cycles for n-subsets of a k-set must satisfy the following necessary condition: k divides $\binom{k}{n}$, or equivalently n divides $\binom{k-1}{n-1}$. For, if 3-subsets of a 6-set, $\mathbf{S}$ will contain exactly one of $\{123, 132, 213, 231, 312, 321\}$. Universal cycles for subsets, are only known to exist for small values of n [23, 24, 33, 39]. This gives rise to studying subset *packings*, that is, cyclic sequences that contains each n-subset *at most* once as a substring [8, 10, 39].

In this paper we are interested in a set $\mathbf{S}$ that does not contain both a string and its reversal. Similar to the problem of subset packings, determining a maximal set $\mathbf{S}$ that admits a universal cycle is extremely challenging. A universal cycle for such a k-ary set of length-n strings is known as an *orientable sequence* of order n (an $\mathcal{OS}_k(n)$). By definition, an orientable sequence does not contain a length-n substring that is a palindrome.

Orientable sequences were introduced for binary strings by Dai, Martin, Robshaw, and Wild [9] with an application related to robotic position sensing. In particular, consider an autonomous robot with limited sensors. To determine its location on a cyclic track labeled with coloured squares, the robot scans a window of n squares directly beneath it (see the graphic on the right). For the position *and* orientation to be uniquely determined, the track is designed with the property that each length n window can appear at most once in *either direction*.
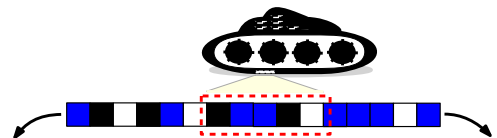
> **Example 1**   Consider the sequence $\mathcal{S} = 012013023123$ over the alphabet $\{0, 1, 2, 3\}$. In the forward direction, including the wraparound, $\mathcal{S}$ contains
>
> $$012, 120, 201, 013, 130, 302, 023, 231, 312, 123, 230, 301$$
>
> as substrings; in the reverse direction $\mathcal{S}$ contains
>
> $$321, 213, 132, 320, 203, 031, 310, 102, 021, 210, 103, 032.$$

Since each substring is unique, $\mathcal{S}$ is an $\mathcal{OS}_4(3)$ with length 12. For biological applications, $\{0, 1, 2, 3\}$ represents the four nucleotide bases $\{\texttt{A}, \texttt{C}, \texttt{G}, \texttt{T}\}$ of a DNA strand – see the discussion at the end of the section.

Recently, Mitchell and Wild developed a recursive algorithm to construct long orientable sequences for a binary alphabet [30]. Subsequently, efficient constructions of $\mathcal{OS}_2(n)$s with asymptotically optimal length have been developed based on cycle-joining [18], and a previously known existence proof [9]. However, there is no known construction of long orientable sequences for $k > 2$. In this paper we demonstrate that the binary cycle-joining approach does not naturally generalize to larger alphabets. In particular, there are special cases for $k = 3$ that arise for $n \geq 12$ (see Example 6 in Section 5), and there are additional challenges in deriving successor rules from cycle-joining trees for $k \geq 3$ (see Section 2.2). However, by deriving a new parent rule that satisfies the Chain Property (see Section 2.2), we are able to obtain the first efficient construction of long $\mathcal{OS}_k(n)$s.

Main result: For $k \geq 3$, we develop a successor rule to construct an $\mathcal{OS}_k(n)$ of asymptotically optimal length in $O(n)$ time per symbol using $O(n)$ space. For $n = 2$, we construct $\mathcal{OS}_k(2)$s of maximal length in $O(1)$ time per symbol.

Let $M_k(n)$ denote the maximum length of an $\mathcal{OS}_k(n)$. When $k = 2$, the maximum length of an orientable sequence is known only for $n \leq 7$ [9, 18]. For $n = 2$ and $k \geq 3$, we demonstrate that $M_k(2) = k\lfloor (k-1)/2 \rfloor$ by a simple construction (see Section 3). For $n \geq 3$, exhaustive search demonstrates that $M_3(3) = 9$, $M_4(3) = 20$, and $M_3(4) = 30$. Search also reveals an $\mathcal{OS}_5(3)$ of length 50 which attains the upper bound stated in [2], and thus $M_5(3) = 50$. Orientable sequences that admit these maximal lengths are given below:

- $n = 3$, $k = 3$: 001120122 (9),
- $n = 3$, $k = 4$: 00112012230130231233 (20),
- $n = 3$, $k = 5$: 00112003102210320331140142042132143043144223342344 (50),
- $n = 4$, $k = 3$: 000102001201112022101121022212 (30).

Since the number of palindromes of length n is $k^{\lfloor (n+1)/2 \rfloor}$, a trivial upper bound on $M_k(n)$ is $(k^n - k^{\lfloor (n+1)/2 \rfloor})/2$. A deeper analysis on upper bounds is given in [2].

Recall the problem of determining a robot's position and orientation on a track. Suppose now that we allow the track to be non-cyclic. The corresponding sequence that allows one to determine orientation and position is called an *acyclic orientable sequence*. One can construct an acyclic $\mathcal{OS}_k(n)$ from a cyclic $\mathcal{OS}_k(n)$ by taking the cyclic $\mathcal{OS}_k(n)$ and appending its prefix of length $n-1$ to the end. See the paper by Burns and Mitchell [5] for more on binary acyclic orientable sequences, which they call *aperiodic* 2-*orientable window sequences*. Gabric and Sawada provide some long acyclic orientable sequences in [17]. Rampersad and Shallit [32] showed that for every alphabet size $k \geq 2$ there is an infinite sequence such that for every sufficiently long substring, the reversal of the substring does not appear in the sequence. Fleischer and Shallit [15] later reproved the results of the previous paper using theorem-proving software. See [7, 29] for more work on sequences avoiding reversals of substrings.

Families of strings related to orientable sequences also appear in DNA computing. Two single strands of DNA can bind to each other if they are "reverse complements" of each other, where $\texttt{A}$ is the complement of $\texttt{T}$ and $\texttt{C}$ of $\texttt{G}$. The binding of DNA strands allows for the creation of secondary structures, which are useful in certain DNA computing techniques [31]. For example, a *stem-loop*, also known as a *hairpin*, is a DNA secondary structure that has applications in DNA computing [11, 28, 41]. Roughly speaking, a string of symbols u contains a hairpin if it has substring v and $\theta(v)$, where θ is an antimorphic involution. Hairpin-free strings, that is, strings that do not contain a hairpin of sufficient length, have been studied with the motivation of creating a large collection of DNA molecules that do not bind to themselves in undesirable ways [27]. In the case that θ is the mirror involution (i.e., reversal), hairpin-free strings are essentially

orientable sequences without the restriction that every substring of a specified length occurs at most once. Thus, every orientable sequence is a hairpin-free sequence. See [26, 40] for more on applications of long hairpin-free sequences.

Outline. In Section 2, we provide background definitions and notation, including a review of the cycle-joining technique, and k-ary successor rules. In Section 3, we present a simple construction for $\mathcal{OS}_k(2)$s and demonstrate they are of maximal length. In Section 4, we define k-ary symmetric/asymmetric necklaces and bracelets, and provide some useful properties of these objects. In Section 5, we provide a parent rule for constructing a cycle-joining tree composed of asymmetric bracelets. This leads to an $O(n)$ time per symbol successor-rule construction of $\mathcal{OS}_k(n)$s that we demonstrate has asymptotically optimal length in Section 6. An implementation of our construction is available for download at http://debruijnsequence.org/db/orientable.

2 Preliminaries

Let $\Sigma = \{0, 1, 2, \ldots, k-1\}$ be an alphabet of size $k \geq 2$. Let Σ^n denote the set of all length-n strings over Σ. Let $\alpha = \mathtt{a}_1\mathtt{a}_2\cdots\mathtt{a}_n \in \Sigma^n$ and $\beta = \mathtt{b}_1\mathtt{b}_2\cdots\mathtt{b}_m \in \Sigma^m$ for some $m, n \geq 0$, Throughout this paper, we use lexicographic order when comparing two strings. More specifically, $\alpha < \beta$ if α is a prefix of β or if $\mathtt{a}_i < \mathtt{b}_i$ for the smallest i such that $\mathtt{a}_i \neq \mathtt{b}_i$. Let α^R denote the reversal $\mathtt{a}_n\mathtt{a}_{n-1}\cdots\mathtt{a}_1$ of α; α is a *palindrome* if $\alpha = \alpha^R$. For $j \geq 1$, let α^j denote j copies of α concatenated together. If $\alpha = \gamma^j$ for some non-empty string γ and some $j > 1$, then α is said to be *periodic*; otherwise, α is said to be *aperiodic* (or *primitive*). Let $\mathrm{ap}(\alpha)$ denote the shortest string γ such that $\alpha = \gamma^t$ for some positive integer t; we say γ is the *aperiodic prefix* of α.

A *necklace class* is an equivalence class of strings under rotation. Let $[\alpha]$ denote the set of strings in α's necklace class. We say α is a *necklace* if it is the lexicographically smallest string in $[\alpha]$. Let $\tilde{\alpha}$ denote the necklace in $[\alpha]$. For example, if $\alpha = 0201$, then $[\alpha] = \{0201, 2010, 0102, 1020\}$ and $\tilde{\alpha} = 0102$. Let $\mathbf{N}_k(n)$ denote the set of k-ary necklaces of length n. A *bracelet class* is an equivalence class of strings under rotation and reversal. We say α is a *bracelet* if it is the lexicographically smallest string in $[\alpha] \cup [\alpha^R]$. A bracelet is always a necklace, but a necklace need not be a bracelet.

Given $\mathbf{S} \subseteq \Sigma^n$, a *universal cycle* U for $\mathbf{S}$ is a cyclic sequence of length $|\mathbf{S}|$ that contains each string in $\mathbf{S}$ as a substring (exactly once). An *orientable sequence* is a universal cycle where if $\alpha \in \mathbf{S}$, then $\alpha^R \notin \mathbf{S}$. If $\mathbf{S} = \Sigma^n$ then U is known as a *de Bruijn sequence*. Given a universal cycle U for a set $\mathbf{S} \subseteq \Sigma^n$, a *successor rule* for U is a function $f : \mathbf{S} \to \Sigma$ such that $f(\alpha)$ is the symbol following α in U.

2.1 Cycle-joining trees

In this section we review how two universal cycles can be joined to obtain a larger universal cycle. Cycle joining is perhaps the most fundamental technique applied to construct universal cycles; it has graph-theoretic underpinnings related to Hierholzer's algorithm for constructing Euler cycles [20]. For some applications, see [12, 13, 14, 16, 19, 22, 25, 36, 38].

Let $\mathtt{x}, \mathtt{y}$ be distinct symbols in Σ. If $\alpha = \mathtt{x}\mathtt{a}_2\cdots\mathtt{a}_n$ and $\hat{\alpha} = \mathtt{y}\mathtt{a}_2\cdots\mathtt{a}_n$, then α and $\hat{\alpha}$ are said to be *conjugates* of each other, and $(\alpha, \hat{\alpha})$ is called a *conjugate pair*. We say γ *belongs to* a conjugate pair $(\alpha, \hat{\alpha})$ if either $\gamma = \alpha$ or $\gamma = \hat{\alpha}$. The following well-known result (see for instance Lemma 3 in [37]) based on conjugate pairs is the crux of the cycle-joining approach.

▶ **Theorem 1.** *Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be disjoint subsets of Σ^n such that $\alpha = \mathtt{x}\mathtt{a}_2\cdots\mathtt{a}_n \in \mathbf{S}_1$ and $\hat{\alpha} = \mathtt{y}\mathtt{a}_2\cdots\mathtt{a}_n \in \mathbf{S}_2$; $(\alpha, \hat{\alpha})$ is a conjugate pair. If U_1 is a universal cycle for $\mathbf{S}_1$ with suffix α and U_2 is a universal cycle for $\mathbf{S}_2$ with suffix $\hat{\alpha}$ then $U = U_1U_2$ is a universal cycle for $\mathbf{S}_1 \cup \mathbf{S}_2$.*

Let U_i denote a universal cycle for $\mathbf{S}_i \subseteq \Sigma^n$. Two universal cycles U_1 and U_2 are said to be *disjoint* if $\mathbf{S}_1 \cap \mathbf{S}_2 = \emptyset$. Theorem 1 states that two disjoint universal cycles can be joined to form a single universal cycle if they each contain one string of a conjugate pair as a substring. Note that necklaces correspond to disjoint cycles that partition the set Σ^n.

A *cycle-joining tree* $\mathbb{T}$ is an unordered tree where the nodes correspond to a disjoint set of universal cycles $U_1, U_2, \ldots, U_t$; an edge between U_i and U_j is defined by a conjugate pair $(\alpha, \hat{\alpha})$ such that $\alpha \in \mathbf{S}_i$ and $\hat{\alpha} \in \mathbf{S}_j$. For our purposes, we consider cycle-joining trees to be rooted.

> **Example 2** Let $n = 3$ and $k = 4$. Consider necklace classes $\mathbf{S}_1 = [021]$, $\mathbf{S}_2 = [011]$, and $\mathbf{S}_3 = [031]$ with corresponding universal cycles $U_1 = 210$, $U_2 = 110$ and $U_3 = 310$. The three cycles can be joined via conjugate pairs $(210, 110)$ and $(210, 310)$ to form the cycle-joining tree on the right. Joining U_1 and U_2 we obtain the larger cycle 210110; joining U_3 to this cycle we obtain the universal cycle 110210310 for $(\mathbf{S}_1 \cup \mathbf{S}_2) \cup \mathbf{S}_3$. If we join U_1 and U_3 first, we obtain a different universal cycle 310210110.
>
> ```
>        210
>       /   \
>     110   310
> ```
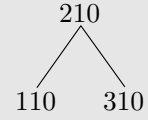
Many universal cycle constructions have a corresponding cycle-joining tree that can be defined by a rather simple *parent rule*. For example, when $\mathbf{S} = \Sigma^n$ and $\alpha = \mathtt{a}_1\mathtt{a}_2 \cdots \mathtt{a}_n \in \mathbf{N}_k(n)$, the following are four of the *simplest* parent rules that define how to construct cycle-joining trees with nodes corresponding to $\mathbf{N}_k(n)$ [19]:

- firstSymbol$(\alpha) = $ the necklace of $[(\mathtt{a}_1{-}1)\mathtt{a}_2 \cdots \mathtt{a}_n]$ with root $(k{-}1)^n$,
- lastSymbol$(\alpha) = $ the necklace of $[\mathtt{a}_1\mathtt{a}_2 \cdots \mathtt{a}_{n-1}(\mathtt{a}_n{+}1)]$ with root 0^n,
- firstNonMin$(\alpha) = 0^{i-1}(\mathtt{a}_i{-}1)\mathtt{a}_{i+1} \cdots \mathtt{a}_n$ with root 0^n, and
- lastNonMax$(\alpha) = \mathtt{a}_1 \cdots \mathtt{a}_{j-1}(\mathtt{a}_j{+}1)(k{-}1)^{n-j}$ with root $(k{-}1)^n$,

where i denotes the index of the first non-zero in α and j denotes the index of the last non-$(k{-}1)$ in α. Addition on the symbols is modulo k. From the definition of a necklace, it is straightforward to see that if α is a (non-root) necklace, then both $0^{i-1}(\mathtt{a}_i{-}1)\mathtt{a}_{i+1} \cdots \mathtt{a}_n$ and $\mathtt{a}_1 \cdots \mathtt{a}_{j-1}(\mathtt{a}_j{+}1)(k{-}1)^{n-j}$ are also necklaces. Note that if a necklace α is periodic, the corresponding universal cycle for α is its aperiodic prefix. However, for simplicity of understanding, we use the full necklace to represent a node in our cycle-joining trees. For instance, when $n = 3$, we use 000 instead of 0. As an example, Figure 1 illustrates the cycle-joining trees induced by the four parent rules above with nodes $\mathbf{N}_3(3)$. Each node α and its parent β are joined by a conjugate pair, where the highlighted bit in α is the first bit in one of the conjugates. When $k > 3$, we apply the last three of the four parent rules to construct an $\mathcal{OS}_k(n)$ with length that is asymptotically optimal. When $k = 3$, we must introduce one additional function.
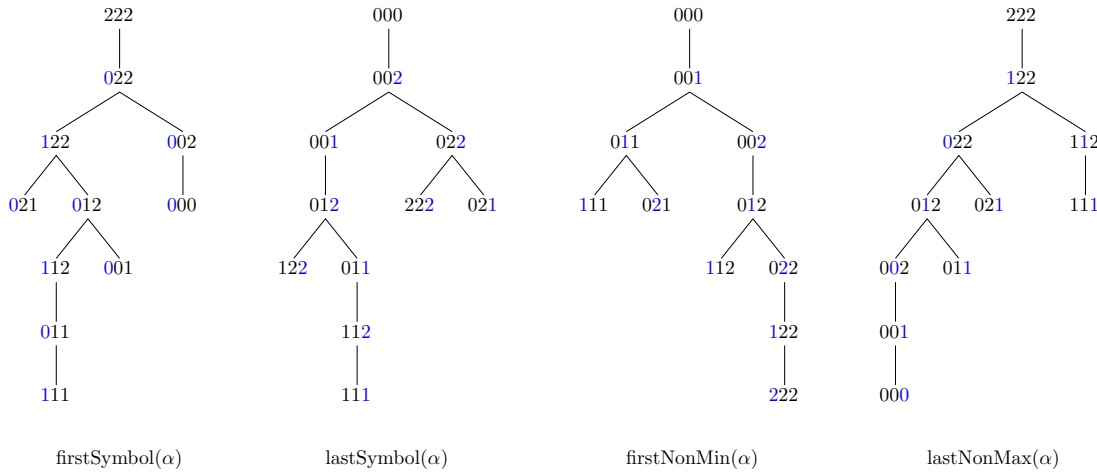
Figure 1 Cycle-joining trees for $\mathbf{N}_3(3)$ induced by four different parent rules.

2.2 Successor rules

In this section we outline how to derive a successor-rule from a cycle-joining tree $\mathbb{T}$ for an underlying set $\mathbf{S}$. In the binary case, each cycle-joining tree corresponds to a unique universal cycle; however, when $k > 2$, this is not necessarily the case.

> **Uniqueness Property**: The cycles in a cycle-joining tree $\mathbb{T}$ are joined such that no two conjugate pairs that have a string in common, i.e., there are no two conjugate pairs of the form $(\mathtt{x}\mathtt{a}_2\cdots\mathtt{a}_n, \mathtt{y}\mathtt{a}_2\cdots\mathtt{a}_n)$ and $(\mathtt{x}\mathtt{a}_2\cdots\mathtt{a}_n, \textcolor{red}{\mathtt{z}}\mathtt{a}_2\cdots\mathtt{a}_n)$.

The Uniqueness Property is always satisfied when $k = 2$, but is not necessarily the case for $k > 2$. For example, none of the trees in Figure 1 satisfy the Uniqueness Property. If $\mathbb{T}$ has the Uniqueness Property, then a successor rule for the corresponding *unique* universal cycle is given by $f(\alpha)$, where $\alpha = \mathtt{a}_1\mathtt{a}_2\cdots\mathtt{a}_n$:

$$f(\alpha) = \begin{cases} \mathtt{y} & \text{if } \alpha \text{ belongs to some conjugate pair } (\alpha, \mathtt{y}\mathtt{a}_2\cdots\mathtt{a}_n); \\ \mathtt{a}_1 & \text{otherwise.} \end{cases}$$

When the Uniqueness Property is not satisfied, a universal cycle derived from a cycle-joining tree depends on the order that the cycles are joined together, as illustrated in Example 2. See [35] for a deeper analysis on different universal cycles that can be obtained from the same cycle-joining tree. Our challenge is to create a relatively simple successor rule that defines a universal cycle derived from a cycle-joining tree $\mathbb{T}$ that does not satisfy the Uniqueness Property. Ultimately, we require our cycle-joining tree to be defined with the following property:

> **Chain Property**: If a node in a cycle-joining tree $\mathbb{T}$ has two children joined via conjugate pairs $(\mathtt{x}\mathtt{a}_2\cdots\mathtt{a}_n, \mathtt{y}\mathtt{a}_2\cdots\mathtt{a}_n)$ and $(\mathtt{x}'\mathtt{b}_2\cdots\mathtt{b}_n, \mathtt{y}'\mathtt{b}_2\cdots\mathtt{b}_n)$, then $\mathtt{a}_2\cdots\mathtt{a}_n \neq \mathtt{b}_2\cdots\mathtt{b}_n$.

Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ denote a maximal-length path of nodes in $\mathbb{T}$ such that for each $1 \leq i < m$, the node α_i is the parent of α_{i+1} and they are joined via a conjugate pair of the form $(\mathtt{x}_i\beta, \mathtt{x}_{i+1}\beta)$, where β is the same in each conjugate pair. We call such a path a *chain* of length m. Any node with j children will belong to at least j chains. Given a chain let $\text{NEXT}(\mathtt{x}_i\beta) = \mathtt{x}_{i+1}$, where $\mathtt{x}_{m+1} = \mathtt{x}_1$.

> **Example 3**   The nodes $\alpha_1 = 112$, $\alpha_2 = 011$, and $\alpha_3 = 111$ in the first tree of Figure 1 form a chain with length $m = 3$ joined by conjugate pairs (211,011) and (011,111). $\text{NEXT}(211) = 0$, $\text{NEXT}(011) = 1$, and $\text{NEXT}(111) = 2$.

If $\mathbb{T}$ is a cycle-joining tree with the Chain Property for an underlying set $\mathbf{S}$, then the following function g is a successor rule for a universal cycle of $\mathbf{S}$ (based on theory in [19]):

$$g(\alpha) = \begin{cases} \text{NEXT}(\alpha) & \text{if } \alpha \text{ belongs to some conjugate pair;} \\ \mathtt{a}_1 & \text{otherwise.} \end{cases}$$

When $k = 2$, $f = g$. As stated, this successor rule requires exponential space to store the conjugate pairs. In our application (see Section 5.2), the corresponding successor rule will run in $O(n)$ time per symbol and use $O(n)$ space.

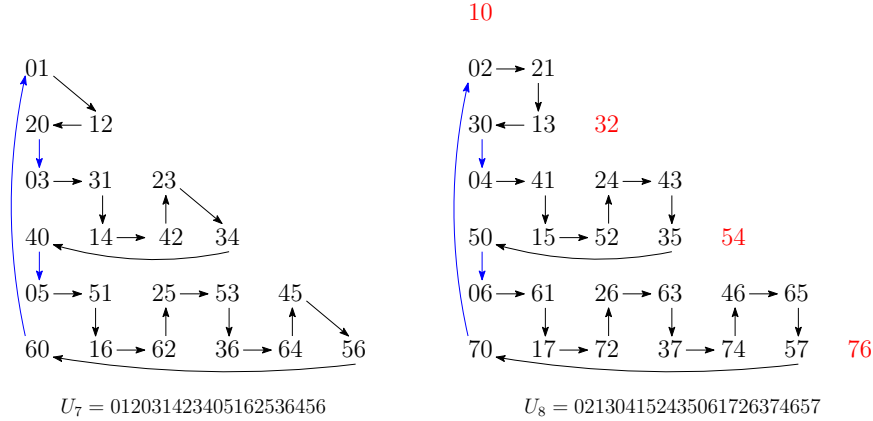## 3   A maximal-length construction for $\mathcal{OS}_k(2)$s

In this section we consider the case when $n = 2$. There does not exist an $\mathcal{OS}_1(2)$ or an $\mathcal{OS}_2(2)$, so we assume $k \geq 3$. Since there are no substrings of the form $\mathtt{xx}$ in any $\mathcal{OS}_2(2)$, each symbol in any $\mathcal{OS}_k(2)$ can appear at most $\lfloor (k-1)/2 \rfloor$ times. Thus, $M_k(2) \leq k\lfloor (k-1)/2 \rfloor$. In fact, we show that this bound is tight via a simple construction that depends on the parity of k.

For each pair of distinct symbols $\mathtt{x}$ and $\mathtt{y}$, an $\mathcal{OS}_k(2)$ can contain either $\mathtt{xy}$ or $\mathtt{yx}$, but not both. If k is odd, our $\mathcal{OS}_k(2)$ will contain exactly one such substring for each pair of symbols. If k is even, we must remove $k/2$ such pairs to meet the upper bound. In particular, we remove the pairs $\{0, 1\}, \{2, 3\}, \ldots, \{k-2, k-1\}$. We outline our choices from each pair of symbols, and illustrate $\mathcal{OS}_k(2)$s in Figure 2 for $k = 7$ and $k = 8$; it is straightforward to generalize the construction depending on the parity of k, as follows.

- k **odd**. Let $\sigma_3 = 012$. For $k \geq 5$, let $\sigma_k = \mathtt{s_1 s_2 \cdots s_{2k-3}}$ where $\mathtt{s_1 s_3 s_5 \cdots s_{2k-5}} = 012 \cdots (k{-}3)$ and $\mathtt{s_2 s_4 s_6 \cdots s_{2k-4} s_{2k-3}} = ((k{-}2)(k{-}1))^{(k-1)/2}$.

- k **even**. Let $\sigma_4 = 0213$. For $k \geq 6$, let $\tau_k = \mathtt{t_1 t_2 \cdots t_{2k-4}}$ where $\mathtt{t_1 t_3 t_5 \cdots t_{2k-5}} = 012 \cdots (k{-}3)$ and $\mathtt{t_2 t_4 t_6 \cdots t_{2k-4}} = ((k{-}2)(k{-}1))^{k/2-1}$.

Let $U_k = \sigma_3 \sigma_5 \cdots \sigma_k$ for k odd, and let $U_k = \tau_4 \tau_6 \cdots \tau_k$ for k even.

$$U_7 = 012031423405162536456 \qquad\qquad U_8 = 0213041524350617263746 57$$

Figure 2 Illustrating the construction of U_7 and U_8.

Example 4 The following illustrates U_k for $3 \leq k \leq 8$:

- $U_3 = 012$
- $U_5 = 012\,0314234$
- $U_7 = 012\,0314234\,05162536456$
- $U_4 = 0213$
- $U_6 = 0213\,04152435$
- $U_8 = 0213\,04152435\,061726374657$

▶ **Theorem 2.** *For $k > 2$, U_k is an $\mathcal{OS}_k(2)$ with length $k\lfloor (k-1)/2 \rfloor$ that can be generated in $O(1)$-time per symbol.*

Proof. We provide a high-level proof outline. Clearly U_3 is an $\mathcal{OS}_3(2)$ with length 3, and U_4 is an $\mathcal{OS}_4(2)$ with length 4. Suppose $k \geq 5$ is odd. The sequences $\sigma_3, \sigma_5, \cdots, \sigma_k$ are disjoint $\mathcal{OS}_k(2)$s that are cycle-joined to obtain U_k. Simple math shows U_k has length $k\lfloor (k-1)/2 \rfloor$. A similar analysis can be applied for even k. It is straightforward to generate U_k in constant time per symbol. ◀

An immediate consequence of this theorem is $M_k(2) = k\lfloor (k-1)/2 \rfloor$.

Observe that when n is odd, U_k is a universal cycle for the 2-subsets of a k-set. Previously, such universal cycles were claimed to exist in [6], with justification in [24]. No construction was previously provided. When n is even, U_k is a maximal 2-subset packing of a k-set.

4 Symmetric and asymmetric necklaces and bracelets

In this section, we present properties for symmetric/asymmetric necklaces and bracelets that are necessary for our main results in Section 5. A necklace α is *symmetric* if it belongs to the same necklace class as α^R, i.e., both α and α^R belong to $[\alpha]$. By this definition, a symmetric necklace is a bracelet. If a necklace or bracelet is not symmetric, it is said to be *asymmetric*. Let $\mathbf{A}_k(n)$ denote the set of all k-ary asymmetric bracelets of length n. Table 1 lists all 70 necklaces in $\mathbf{N}_4(4)$ partitioned into asymmetric necklace pairs and symmetric necklaces. The asymmetric necklace pairs belong to the same bracelet class, and the first string in each pair is an asymmetric bracelet. Thus, $|\mathbf{A}_4(4)| = 15$. In general, $|\mathbf{A}_k(n)|$ is equal

Asymmetric necklace pairs				Symmetric necklaces				
0012 , 0021	0133 , 0331	0000	0102	0222	1113	1323		
0013 , 0031	0213 , 0312	0001	0103	0232	1122	1333		
0023 , 0032	0223 , 0322	0002	0111	0303	1133	2222		
0112 , 0211	0233 , 0332	0003	0121	0313	1212	2223		
0113 , 0311	1123 , 1132	0011	0131	0323	1213	2233		
0122 , 0221	1223 , 1322	0022	0202	0333	1222	2323		
0123 , 0321	1233 , 1332	0033	0203	1111	1232	2333		
0132 , 0231		0101	0212	1112	1313	3333		

■ **Table 1** The 70 necklaces in $\mathbf{N}_4(4)$ partitioned into 15 asymmetric necklace pairs and 40 symmetric necklaces. The first (highlighted) necklace in each pair is an asymmetric bracelet in $\mathbf{A}_4(4)$.

to the number of k-ary necklaces of length n minus the number of k-ary bracelets of length n. For more background on asymmetric bracelets, see [1, 18].

Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ denote the asymmetric bracelets in $\mathbf{A}_k(n)$. Let $\mathbf{S}_k(n)$ denote the set $[\alpha_1] \cup [\alpha_2] \cup \cdots \cup [\alpha_m]$.

Important property: If $\alpha \in \mathbf{S}_k(n)$, then $\alpha^R \notin \mathbf{S}_k(n)$.

Let $L_k(n) = |\mathbf{S}_k(n)|$. Our main result defines a cycle-joining tree with nodes $\mathbf{A}_k(n)$, producing an $\mathcal{OS}_k(n)$ of length $L_k(n)$. In Section 6, we provide a formula for $L_k(n)$ and demonstrate it to be an asymptotically optimal lower bound for $M_k(n)$ (the maximal length of an $\mathcal{OS}_k(n)$). The reminder of this section is devoted to properties of symmetric/asymmetric necklaces and bracelets required to prove our main results in Section 5.

▶ **Lemma 3** ([18]). *A necklace α is symmetric if and only if there exists palindromes β_1 and β_2 such that $\alpha = \beta_1\beta_2$.*

▶ **Lemma 4.** *Let $\alpha = \mathtt{a}_1\mathtt{a}_2\cdots\mathtt{a}_n \neq 0^n$ be a bracelet, where i is the index of the first non-zero symbol. If $\mathtt{a}_i > 1$, then* firstNonMin(α) *is a bracelet; moreover, if $\alpha \in \mathbf{A}_k(n)$, then* firstNonMin$(\alpha) \in \mathbf{A}_k(n)$.

Proof. Suppose $\mathtt{a}_i > 1$. It is straightforward to apply the definition of a bracelet to verify that $\beta = $ firstNonMin$(\alpha) = 0^{i-1}(\mathtt{a}_i-1)\mathtt{a}_{i+1}\cdots\mathtt{a}_n$ is a bracelet, given α is a bracelet. Let $\alpha \in \mathbf{A}_k(n)$. Suppose β is symmetric. From Lemma 3, $\beta = \beta_1\beta_2$ for palindromes β_1 and β_2. Clearly $|\beta_1| \geq i-1$. Since α is a bracelet $\mathtt{a}_i \leq \mathtt{a}_n$, and thus $|\beta_1| \neq i-1$. If $|\beta_1| = i$, then $i = 1$ which implies that α is symmetric; otherwise, since β_1 has prefix 0^{i-1}, it must have length at least $2i-1$. If $|\beta_1| = 2i-1$, then α is symmetric; if $|\beta_1| > 2i-1$ then α is not a bracelet. Each case contradicts that $\alpha \in \mathbf{A}_k(n)$. Thus, β is an asymmetric bracelet. ◀

The following lemma follows by applying the definition of a bracelet.

▶ **Lemma 5.** *If $\alpha \neq (k-1)^n$ is a bracelet, then* lastNonMax(α) *is a bracelet.*

Note that lastSymbol$(\alpha) = $ lastNonMax(α) when $\mathtt{a}_n < k-1$.

▶ **Lemma 6.** *Let $\alpha = \mathtt{a}_1\mathtt{a}_2\cdots\mathtt{a}_n \in \mathbf{A}_k(n)$, where j is the index of the last non-$(k-1)$ and ℓ is the index of the second last non-$(k-1)$. If (i) $\mathtt{a}_n < k-1$ or (ii) $\mathtt{a}_j < k-2$ or (iii) $\mathtt{a}_1\cdots\mathtt{a}_\ell \neq (\mathtt{a}_1\cdots\mathtt{a}_\ell)^R$, then* lastNonMax$(\alpha) \in \mathbf{A}_k(n)$; *otherwise,* lastNonMax(α) *is a symmetric bracelet.*

Proof. From Lemma 5, $\beta = $ lastNonMax(α) is a bracelet. If β is symmetric, then from Lemma 3 it can be written as $\beta_1\beta_2$ for some palindromes β_1 and β_2. Since $\mathtt{a}_1 < \mathtt{a}_n$ because $\alpha \in \mathbf{A}_k(n)$, we clearly have β_1 and β_2 are non-empty.

(i) Let $\mathtt{a}_n < k-1$. Suppose β is symmetric. Let γ be β_2 with its last symbol decremented; $\alpha = \beta_1\gamma$. Then $(\gamma\beta_1)^R \leq \alpha$, contradicting that $\alpha \in \mathbf{A}_k(n)$.

(ii) Let $a_j < k-2$. Suppose β is symmetric. If $|\beta_1| > j$, then it must start and end with $(k-1)$ and include a_j. This contradicts that β is a bracelet. If $|\beta_1| = j$, then since β_1 is a palindrome, the length j prefix of α is greater than its reversal, contradicting that α is a bracelet. Thus β_2 has suffix $(a_j+1)a_{j+1}\cdots a_n$. If $|\beta_2| = 2(n-j)+1$, then the suffix of α with the same length is also a palindrome, which implies α is symmetric, contradiction. Otherwise, $|\beta_2| > 2(n-j)+1$. Let γ denote the suffix of α of length $|\beta_2|$. It must be that $\gamma > \gamma^R$, which implies that $(\gamma\beta_1)^R < \alpha$, contradicting that α is a bracelet. Thus, $\beta \in \mathbf{A}_k(n)$.

(iii) Let $a_1\cdots a_\ell \neq (a_1\cdots a_\ell)^R$. By the definition of a bracelet, $a_1\cdots a_\ell < (a_1\cdots a_\ell)^R$. If $a_j < k-2$, then β is in $\mathbf{A}_k(n)$ by the previous case. Consider $a_j = k-2$. Suppose $\beta = a_1\cdots a_\ell(k-1)^{n-\ell}$ is symmetric. We must have $|\beta_1| < \ell$. Since $a_\ell < (k-1)$, $|\beta_2| \geq 2(n-\ell)+1$. But this implies that α is not a bracelet, contradiction. Thus, $\beta \in \mathbf{A}_k(n)$.

Finally, if $a_n = k-1$, $a_j = k-2$, and $a_1\cdots a_\ell = (a_1\cdots a_\ell)^R$, then $\mathrm{lastNonMax}(\alpha) = a_1\cdots a_\ell(k-1)^{n-\ell}$, which is symmetric by Lemma 3. ◄

▶ **Lemma 7.** *If $\alpha \in \mathbf{A}_k(n)$ such that* $\mathrm{lastNonMax}(\alpha) \notin \mathbf{A}_k(n)$, $\mathrm{lastSymbol}(\alpha) \notin \mathbf{A}_k(n)$ *and* $\mathrm{firstNonMin}(\alpha) \in \mathbf{A}_k(n)$, *then* $\mathrm{lastNonMax}(\mathrm{firstNonMin}(\alpha)) \in \mathbf{A}_k(n)$ *or* $\mathrm{lastSymbol}(\mathrm{firstNonMin}(\alpha)) \in \mathbf{A}_k(n)$.

Proof. Consider $\alpha = a_1a_2\cdots a_n \in \mathbf{A}_k(n)$ where i is the index of the first non-zero, j is the index of the last non-$(k-1)$, and ℓ is the index of the second-last non-$(k-1)$. Since $\mathrm{lastNonMax}(\alpha) \notin \mathbf{A}_k(n)$, from Lemma 6, $a_j = k-2$, $a_n = k-1$, and $a_1\cdots a_\ell = (a_1\cdots a_\ell)^R$. Since $\alpha \in \mathbf{A}_k(n)$, $\alpha = a_1\cdots a_\ell(k-1)^{j-\ell-1}(k-2)(k-1)^{n-j}$ where $j-\ell-1 < n-j$. This implies that $\mathrm{lastSymbol}(\alpha) = 0a_1\cdots a_{n-1}$ has prefix 0^i; it is a necklace since α has no 0^i substring. Since $\mathrm{lastSymbol}(\alpha) \notin \mathbf{A}_k(n)$, $a_i\cdots a_{n-1} \geq (a_i\cdots a_{n-1})^R$, which means that $a_i \geq a_{n-1} \geq k-2$. Let $\beta = \mathrm{firstNonMin}(\alpha)$. If $i \geq n-1$, then $\alpha = 0^{n-2}(k-2)(k-1)$ and $\mathrm{lastNonMax}(\beta) = \alpha$. Thus, assume $i < n-1$ and consider two cases.

- Suppose $i > \ell$. Then $i = \ell+1$ and $\alpha = 0^\ell(k-1)^{j-i}(k-2)(k-1)^{n-j}$. Since $j-i < n-j$ and $a_i\cdots a_{n-1} \geq (a_i\cdots a_{n-1})^R$, $j-i = n-j-1 > 1$. Thus, $\mathrm{lastSymbol}(\beta) = 0^\ell(k-2)(k-1)^{n-i}$ which is in $\mathbf{A}_k(n)$.
- Suppose $i \leq \ell$. If the first ℓ symbols of β are not a palindrome, then by Lemma 6, $\mathrm{lastNonMax}(\beta) \in \mathbf{A}_k(n)$. Otherwise, ℓ is odd, $i = (\ell+1)/2$, and α has prefix $0^{i-1}a_i0^{i-1}$. If $a_i = k-2$, since $a_i\cdots a_{n-1} \geq (a_i\cdots a_{n-1})^R$, we have $a_{n-1} = k-2$, $\alpha = 0^{i-1}(k-2)0^{i-1}(k-2)(k-1)$, and $\mathrm{lastSymbol}(\beta) = 0^i(k-3)0^{i-1}(k-2)$ which is in $\mathbf{A}_k(n)$. Otherwise, $a_i = k-1$ and $a_{n-1} = k-1$ which means $i > 1$, $j < n-1$ and $a_{n-2} \geq k-2$. But this contradicts that $a_i\cdots a_{n-1} \geq (a_i\cdots a_{n-1})^R$.

◄

<div style="background:#F5A800; display:inline-block"> **5** </div> **A cycle-joining construction for orientable sequences**

In this section we provide a parent rule to create a cycle-joining tree with nodes $\mathbf{A}_k(n)$. We then apply the tree to derive an $O(n)$-time successor rule for a corresponding universal cycle, which is an $\mathcal{OS}_k(n)$ of length $L_k(n)$.

5.1 A simple parent rule

The parent rule for the binary case defined in [18] uses $0^{n-4}1011$ as the root, and the parent of each non-root node $\alpha \in \mathbf{A}_2(n)$ is the first string in the list $\langle\mathrm{firstNonMin}(\alpha), \mathrm{lastSymbol}(\alpha), \mathrm{lastNonMax}(\alpha)\rangle$ that is also in $\mathbf{A}_2(n)$. However, there are several issues when generalizing to a larger alphabet. In particular, the rule is not well-defined for $k = 3$, and the corresponding cycle-joining tree does not have the Chain Property. We will demonstrate each of these short-comings before deriving a new parent rule for alphabets of arbitrary size.

Assume $n, k \geq 3$. Let $r_{n,k} = 0^{n-2}(k-2)(k-1)$ denote the root of our upcoming cycle-joining tree. The following example is for this specific root; however, similar examples exist for any arbitrary root.

Example 5   Consider any parent rule with root $r_{n,k}$ where the the parent of $\alpha \in \mathbf{A}_k(n)$ is the first string in a list starting with $\langle \text{firstNonMin}(\alpha), \text{lastSymbol}(\alpha), \ldots \rangle$ that is also in $\mathbf{A}_k(n)$. Let $\alpha = 012202$ and $\beta = 010122$. Then the parent of α is lastSymbol$(\alpha) = 000122$ and the parent of β is firstNonMin$(\beta) = 000122$; α is joined via conjugate pair $(201220, 001220)$ and β is joined via conjugate pair $(101220, 001220)$. The two conjugate pairs share a string, and thus the corresponding cycle-joining tree does not have the Chain Property.

The next example, and following lemma, demonstrate that the four functions firstSymbol, firstNonMin, lastSymbol, and lastNonMax alone are not sufficient to define a parent rule with root $r_{n,k}$ when $k = 3$.

Example 6   Consider $\alpha = 001022010012$ which is in $\mathbf{A}_3(12)$. No matter how we change the first, last, first non-zero, or last non-$(k-1)$ symbol in α, the resulting string is not in $\mathbf{A}_3(12)$. In particular:

- firstSymbol$(\alpha) = 001220102201$ is not a bracelet,
- lastSymbol$(\alpha) = 000102201001$ is not a bracelet,
- firstNonMin$(\alpha) = 0000022010012$ is not a bracelet, and
- lastNonMax$(\alpha) = 001022010022$ is symmetric.

Such strings are uncommon. There are only 82 such strings in $\mathbf{A}_3(20)$ and they all have suffix 0012.

▶ **Lemma 8.** *Let* $\alpha = 00102^{(n-10)}010012$ *for* $n \geq 12$ *and* $k = 3$. *Then* α *is in* $\mathbf{A}_k(n)$ *and each of* firstSymbol(α), lastSymbol(α), firstNonMin(α), *and* lastNonMax(α) *is not in* $\mathbf{A}_k(n)$.

Proof. It is straightforward to observe that α is a bracelet by definition and is asymmetric by Lemma 3. Applying the definitions, firstSymbol$(\alpha) = 001220102^{(n-10)}01$, lastSymbol$(\alpha) = 000102^{(n-10)}01001$, and firstNonMin$(\alpha) = 000002^{(n-10)}010012$ are all not bracelets, and lastNonMax$(\alpha) = 00102^{(n-10)}010022$ is symmetric. ◀

Lemma 8, demonstrates that for $k = 3$, no parent rule exists for $\mathbf{A}_3(n)$ that applies only a combination of the four rules from Section 2.1. Thus, we define an additional function in order to define our parent rule for $\mathbf{A}_3(n)$. Let ℓ denotes the second last symbol in α that is not $(k-1)$, and define

- secondLastNonMax$(\alpha) = \mathtt{a}_1 \cdots \mathtt{a}_{\ell-1}(\mathtt{a}_\ell + 1)\mathtt{a}_{\ell+1} \cdots \mathtt{a}_n$.

This function is well-defined, since each $\alpha \in \mathbf{A}_k(n)$ must contain at least two symbols that are not $k-1$. Recall Example 6, where $\alpha = 001022010012$. Observe that secondLastNonMax$(\alpha) = 001022010112$, which is in $\mathbf{A}_3(12)$.

Parent rule for cycle-joining $\mathbf{A}_k(n)$ with root $r_{n,k}$.   If $\alpha = \mathtt{a}_1\mathtt{a}_2 \cdots \mathtt{a}_n \in \mathbf{A}_k(n) \setminus \{r_{n,k}\}$, then define par$(\alpha)$ to be the first string that is an asymmetric bracelet in the list

$\langle$ lastNonMax(α), lastSymbol(α), firstNonMin(α), secondLastNonMax(α) $\rangle$.

In the upcoming Lemma 9, we demonstrate that par(α) is well-defined and interestingly, that secondLastNonMax(α) is only necessary for $k = 3$. The upcoming Theorem 10 demonstrates that the above parent rule induces a cycle-joining tree, which we denote by $\mathbb{T}_k(n)$. Moreover, we demonstrate $\mathbb{T}_k(n)$ has the Chain Property in Theorem 12. In proving these results, we do not consider the case when $n = 3$ and $k = 3$, since $\mathbf{A}_3(3) = \{012\}$. Figure 3 illustrates $\mathbb{T}_3(6)$.

▶ **Lemma 9.** *Let* $\alpha = \mathtt{a}_1\mathtt{a}_2 \cdots \mathtt{a}_n \in \mathbf{A}_k(n) \setminus \{r_{n,k}\}$ *for some* $n \geq 3$, $k \geq 4$ *or* $n \geq 4$, $k = 3$ *such that* lastNonMax(α) *and* lastSymbol(α) *are not in* $\mathbf{A}_k(n)$. *If* $k \geq 4$, *then* firstNonMin$(\alpha) \in \mathbf{A}_k(n)$. *Furthermore, if* $k = 3$ *and* firstNonMin$(\alpha) \notin \mathbf{A}_3(n)$, *then* $\alpha = 0\gamma012$ *where* γ *is a palindrome, and* secondLastNonMax$(\alpha) \in \mathbf{A}_3(n)$.
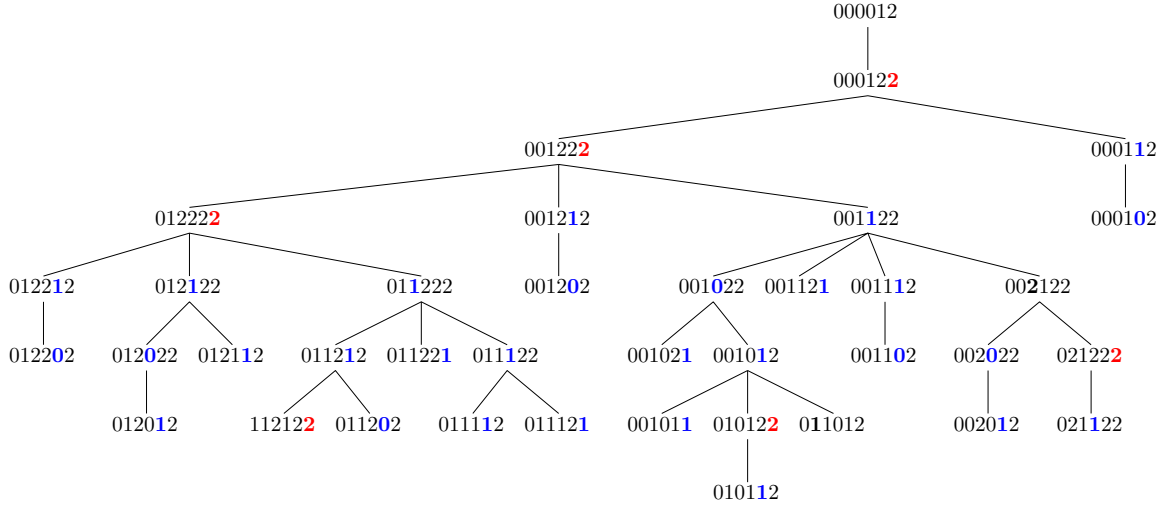
Figure 3 The cycle-joining tree $\mathbb{T}_3(6)$. Each node differs from its parent (cyclically) at the highlighted symbol. The symbols highlighted in blue indicate that $\mathrm{par}(\alpha) = \mathrm{lastNonMax}(\alpha)$; the symbols highlighted in red indicate that $\mathrm{par}(\alpha) = \mathrm{lastSymbol}(\alpha)$; the symbols highlighted in bold black indicate that $\mathrm{par}(\alpha) = \mathrm{firstNonMin}(\alpha)$. There are no nodes in this tree such that $\mathrm{par}(\alpha) = \mathrm{secondLastNonMax}(\alpha)$; the first instance of such a case arises when $n = 12$.

Proof. Consider α where i is the index of the first non-zero, i' is the index of the second non-zero, j is the index of the last non-$(k-1)$, and ℓ is the index of the second-last non-$(k-1)$. Since $\mathrm{lastNonMax}(\alpha) \notin \mathbf{A}_k(n)$, from Lemma 6, $\mathsf{a}_j = k-2$, $\mathsf{a}_n = k-1$, and $\mathsf{a}_1 \cdots \mathsf{a}_\ell = (\mathsf{a}_1 \cdots \mathsf{a}_\ell)^R$. Thus, since $\alpha \in \mathbf{A}_k(n)$, $\alpha = \mathsf{a}_1 \cdots \mathsf{a}_\ell (k-1)^{j-\ell-1}(k-2)(k-1)^{n-j}$ where $j - \ell - 1 < n - j$. Thus, if $j = n - 1$, then $\ell = n - 2$. This implies that $\mathrm{lastSymbol}(\alpha) = 0\mathsf{a}_1 \cdots \mathsf{a}_{n-1}$ with prefix 0^i, which is a necklace since α has no 0^i substring. Since $\mathrm{lastSymbol}(\alpha) \notin \mathbf{A}_k(n)$, $\mathsf{a}_i \cdots \mathsf{a}_{n-1} \geq (\mathsf{a}_i \cdots \mathsf{a}_{n-1})^R$, which means that $\mathsf{a}_i \geq \mathsf{a}_{n-1} \geq k-2$. If $k \geq 4$, then $\mathsf{a}_i > 1$, and thus by Lemma 4 $\mathrm{firstNonMin}(\alpha) \in \mathbf{A}_k(n)$. If $k = 3$, suppose $\mathrm{firstNonMin}(\alpha) \notin \mathbf{A}_k(n)$. Then by Lemma 4, $\mathsf{a}_i = 1$, and $\mathsf{a}_{i'} \cdots \mathsf{a}_n \geq (\mathsf{a}_{i'} \cdots \mathsf{a}_n)^R$. Since $1 = \mathsf{a}_i \geq \mathsf{a}_{n-1}$, $j = n-1$ and $\mathsf{a}_{n-1} = (k-2) = 1$. If $i = 1$ then α does not contain any 0s, and thus $i' = 2$. Thus $\mathsf{a}_2 = \mathsf{a}_n = 2$. The only bracelet that both starts and ends with 12 is of the form $(12)^s$ which is asymmetric, contradicting that $\alpha \in \mathbf{A}_k(n)$. Thus $i > 1$ and $\mathsf{a}_1 = \mathsf{a}_\ell = 0$. Recall $j = n - 1$ which implied $\ell = n - 2$. Thus, $\alpha = 0\gamma 012$ where γ is a palindrome. Let $\beta = \mathrm{secondLastNonMax}(\alpha)$; it has suffix 112. Observe β is a bracelet by applying the definition. Suppose β is symmetric. By Lemma 3, $\beta = \beta_1 \beta_2$ where β_1 and β_2 are both palindromes. Clearly, $|\beta_2| > 3$. Note that α does not have prefix 02, since α is a necklace containing 012 as a substring. If $\beta_2 = 2112$, then α has prefix 02, contradiction. If $\beta_2 = 21112$, then α starts and ends with 012. Since α is a necklace, this implies that $\alpha = (012)^t$ for some integer t, which contradicts the form of β_2. Thus $|\beta| > 5$. Let δ denote the suffix of α of length $|\beta_2|$. Then α is not a bracelet since $\delta > \delta^R$ and thus $\alpha^R < \alpha$. Contradiction. Thus, β is asymmetric and in $\mathbf{A}_3(n)$.   ◄

Given a node α in $\mathbb{T}_k(n)$, we say that γ is an *ancestor* of α if $\gamma = \alpha$ or $\gamma = \mathrm{par}^t(\alpha)$ for some $t \geq 1$.

▶ **Theorem 10.** *For $n, k \geq 3$, the parent rule $\mathrm{par}(\alpha)$ for $\mathbf{A}_k(n)$ induces a cycle-joining tree $\mathbb{T}_k(n)$ with nodes $\mathbf{A}_k(n)$ rooted at $r_{n,k}$.*

Proof. Let $\alpha = \mathsf{a}_1 \mathsf{a}_2 \cdots \mathsf{a}_n \in \mathbf{A}(n) \setminus \{r_{n,k}\}$, where i is the index of the first non-zero, j is the index of the last non-$(k-1)$, and ℓ is the index of the second last non-$(k-1)$ symbol in α. We demonstrate that $r_{n,k}$ is an ancestor of α. If $i = n - 1$ then $\mathrm{par}^{(k-1-\mathsf{a}_n)+(k-2-\mathsf{a}_{n-1})}(\alpha) = r_{n,k} = 0^{n-2}(k-2)(k-1)$, where each application of par uses lastNonMax. For $i \leq n - 2$, we demonstrate that α has an ancestor β such that $\beta < \alpha$, and thus must have an ancestor where $i = n - 1$. If $i = j$, then $\alpha = 0^{i-1}\mathsf{a}_i(k-1)^{n-i}$ and $\mathrm{par}^{k-1-\mathsf{a}_i}(\alpha) = 0^i(k-2)(k-2)^{n-i-1} < \alpha$. Otherwise $i < j$. If $k = 3$, then from Lemma 9, $\mathrm{par}(\alpha) = \mathrm{secondLastNonMax}(\alpha)$ occurs only when α has suffix 012. Thus, repeated applications of only lastNonMax and secondLastNonMax to α will never change any of the first i symbols. Since both operations only increment symbols to at most $(k-1)$, one of lastSymbol or firstNonMin must eventually be applied by repeated

application of par starting with α. If firstNonMin is applied then it will decrement $\mathtt{a}_i$ leading to an ancestor that is less than α. Similarly, lastSymbol will change the last symbol from $(k{-}1)$ to 0 leading to a node with prefix 0^i which is also less than α. ◄

By repeatedly joining cycles from $\mathbb{T}_k(n)$ via conjugate pairs, we can construct an $\mathcal{OS}_k(n)$ (universal cycle) of length $L_k(n)$ using exponential time per symbol (to search for the conjugates) and exponential space. Table 2 illustrates the lengths of the $\mathcal{OS}_k(n)$s constructed for small n, k. In Section 6, we present an exact formula for $L_k(n)$.

▶ **Theorem 11.** *There exists a universal cycle for* $\mathbf{S}_k(n)$*, which is an* $\mathcal{OS}_k(n)$*, of length* $L_k(n)$*.*

n ╲ k	3	4	5	6	7	8
3	3	12	30	60	105	168
4	12	60	180	420	840	1512
5	60	360	1260	3360	7560	15120
6	225	1608	6750	21150	54831	124320
7	819	7308	36890	135450	403389	1034264
8	2676	30300	187980	821940	2844408	8315496
9	8778	126516	962580	5003970	20101326	66961608
10	27180	511680	4836300	30097620	140902440	536135040
11	84579	2074644	24328150	181141950	988016337	4293525544
12	257205	8327808	121790490	1087414170	6917824935	34352668560
13	782964	33447960	609843780	6528527460	48439152216	274864275504
14	2361177	133931952	3050119450	39175228260	339088485771	2198957209792
15	7125423	536379792	15255860130	235079896440	2373737520945	17592060218208
16	21419076	2146175580	76284577980	1410507942900	16616280850008	140736884449896
17	64402800	8587706400	381453125040	8463244062000	116314913988000	1125898765992000
18	193357350	34353845664	1907295914700	50779660926240	814205346309138	9007193819084160
19	580569795	137428992036	9536650390670	304679295576630	5699444909171768	72057583837380680
20	1742213832	549729612720	47683422899280	1828077103852860	39896121850134479	576460703985782112

■ **Table 2** Lower bounds $L_k(n)$ on the maximal length of an $\mathcal{OS}_k(n)$ for $n \le 20$ and $k \le 8$.

▶ **Theorem 12.** $\mathbb{T}_k(n)$ *has the chain property.*

Proof. By contradiction. Suppose $\mathbb{T}_k(n)$ has a node γ with two children α and β joined via conjugate pairs $(\mathtt{x}\sigma, \mathtt{y}\sigma)$ and $(\mathtt{x}\sigma, \mathtt{y}'\sigma)$, respectively, for some string σ. Note that $\alpha, \beta, \gamma \in \mathbf{A}_k(n)$. Since $\alpha \ne \beta$, the functions they apply to obtain their parent γ cannot both increment a symbol. Without loss of generality, assume $\mathrm{par}(\beta) = \mathrm{firstNonMin}(\beta)$ and $\mathrm{par}(\alpha)$ applies one of lastNonMax, lastSymbol, or secondLastNonMax. Let $\alpha = \mathtt{a}_1\mathtt{a}_2 \cdots \mathtt{a}_n$ and consider the three possible cases for $\mathrm{par}(\alpha)$.

- $\mathrm{par}(\alpha) = \mathrm{lastNonMax}(\alpha)$. Let j denote the index of the last non-$(k{-}1)$ in α. Since $\gamma = \mathtt{a}_1 \cdots \mathtt{a}_{j-1}(\mathtt{a}_j{+}1)\mathtt{a}_{j+1} \cdots \mathtt{a}_n$, $\beta = \mathtt{a}_1 \cdots \mathtt{a}_{j-1}(\mathtt{a}_j{+}2)\mathtt{a}_{j+1} \cdots \mathtt{a}_n$ where $\mathtt{a}_j{+}2 \le (k{-}1)$. However, from Lemma 4, either $\mathrm{par}(\beta) = \mathrm{lastNonMax}(\beta)$, or β is not in $\mathbf{A}_k(n)$. Contradiction.
- $\mathrm{par}(\alpha) = \mathrm{lastSymbol}(\alpha)$. As noted in the proof of Lemma 9, $\gamma = 0\mathtt{a}_1 \cdots \mathtt{a}_{n-1}$. Therefore $\beta = 1\mathtt{a}_1 \cdots \mathtt{a}_{n-1}$. Since $\beta \in \mathbf{A}_k(n)$, $1\mathtt{a}_1 \cdots \mathtt{a}_{n-2} < (1\mathtt{a}_1 \cdots \mathtt{a}_{n-2})^R$; otherwise, β is not a bracelet or it is symmetric (see Lemma 3). Since $\mathrm{par}(\beta) \ne \mathrm{lastNonMax}(\beta)$, from Lemma 6, $\mathtt{a}_{n-1} = k{-}1$. However, this means that $\mathrm{lastSymbol}(\beta) = 01\mathtt{a}_1 \cdots \mathtt{a}_{n-2} \in \mathbf{A}_k(n)$, which contradicts that $\mathrm{par}(\beta) = \mathrm{firstNonMin}(\beta)$.
- $\mathrm{par}(\alpha) = \mathrm{secondLastNonMax}(\alpha)$. From Lemma 9, α has suffix 012 and thus $\gamma = \mathtt{a}_1 \cdots \mathtt{a}_{n-3}112$. Therefore $\beta = \mathtt{a}_1 \cdots \mathtt{a}_{n-3}212$ with its first non-0 at index $n - 2$, which means β is symmetric. Contradiction.

◄

5.2 An $O(n)$-time successor rule

In this section, we apply the generic successor rule $g(\alpha)$ defined in Section 2.2 to the cycle-joining tree $\mathbb{T}_k(n)$ that admits the Chain Property. In particular, we determine whether or not α belongs to a conjugate pair, and if so, how to efficiently compute the function $\text{NEXT}(\alpha)$.

Given $\alpha = \mathtt{a}_1 \mathtt{a}_2 \cdots \mathtt{a}_n \in \mathbf{S}_k(n)$, let i be the largest index such that $\mathtt{a}_i > 0$, let j be the smallest index greater than 1 such that $\mathtt{a}_j < k-1$, and let ℓ be the second smallest index greater than 1 such that $\mathtt{a}_\ell < k-1$. Recall that $\tilde{\alpha}$ denotes the necklace of $[\alpha]$. If α belongs to some conjugate pair (possibly more than one) used to create $\mathbb{T}_k(n)$, then we consider the possibilities for α depending on whether or not it belongs to a parent or child node joined by a given conjugate pair. If α belongs to the child, let $\beta = \tilde{\alpha}$. If $\text{par}(\beta) = \text{lastNonMax}(\beta)$, then $\beta = \mathtt{a}_j \mathtt{a}_{j+1} \cdots \mathtt{a}_n \mathtt{a}_1 (k-1)^{j-2}$. In other words, $\mathtt{a}_1$ corresponds to the last non-$(k-1)$ symbol in β. If α belongs to the parent, let γ denote the child node. If $\text{par}(\gamma) = \text{lastNonMax}(\gamma) = \tilde{\alpha}$, then it must be that $\gamma = \mathtt{a}_j \mathtt{a}_{j+1} \cdots \mathtt{a}_n (\mathtt{a}_1 - 1)(k-1)^{j-2}$. A similar analysis holds for the other three cases of the parent rule $\text{par}(\alpha)$ giving rise to the definitions of the following eight strings:

$$\begin{array}{lll}
\beta_1 = \mathtt{a}_j \mathtt{a}_{j+1} \cdots \mathtt{a}_n \mathtt{a}_1 (k-1)^{j-2} & \gamma_1 = \mathtt{a}_j \mathtt{a}_{j+1} \cdots \mathtt{a}_n (\mathtt{a}_1-1)(k-1)^{j-2} & \text{(lastNonMax)} \\
\beta_2 = \mathtt{a}_2 \mathtt{a}_3 \cdots \mathtt{a}_n \mathtt{a}_1 & \gamma_2 = \mathtt{a}_2 \mathtt{a}_3 \cdots \mathtt{a}_n (\mathtt{a}_1-1) & \text{(lastSymbol)} \\
\beta_3 = 0^{n-i} \mathtt{a}_1 \mathtt{a}_2 \cdots \mathtt{a}_i & \gamma_3 = 0^{n-i}(\mathtt{a}_1+1)\mathtt{a}_2 \cdots \mathtt{a}_i & \text{(firstNonMin)} \\
\beta_4 = \mathtt{a}_\ell \mathtt{a}_{\ell+1} \cdots \mathtt{a}_n \mathtt{a}_1 \mathtt{a}_2 \cdots \mathtt{a}_{\ell-1} & \gamma_4 = \mathtt{a}_\ell \mathtt{a}_{\ell+1} \cdots \mathtt{a}_n (\mathtt{a}_1-1)\mathtt{a}_2 \cdots \mathtt{a}_{\ell-1} & \text{(secondLastNonMax).}
\end{array}$$

Assume addition on the symbols is modulo k; i.e., $(k-1) + 1 = 0$ and $0 - 1 = (k-1)$. The above strings can be tested to determine whether or not α belongs to a conjugate pair. For instance if $\gamma_1 \in \mathbf{A}_k(n)$ and $\text{par}(\gamma_1) = \text{lastNonMax}(\gamma_1)$, then α belongs to the conjugate pair $(\alpha, (\mathtt{a}_1-1)\mathtt{a}_2 \cdots \mathtt{a}_n)$. If α is found to belong to some conjugate pair, then the second issue is to efficiently compute the function $\text{NEXT}(\alpha)$. If α belongs to some parent node of a conjugate pair, then $\text{NEXT}(\alpha)$ is simply the incremented or decremented value of $\mathtt{a}_1$ defined by the parent rule for the corresponding γ_i. If α does not belong to a parent in any conjugate pair, then it belongs to the last node in its corresponding chain which contains some β_i; to compute $\text{NEXT}(\alpha)$ we must determine the first node in the chain. Naïvely, we can repeatedly check the ancestors α until we reach the top of the chain. In the worst case this will take $O(kn)$ time. With a deeper analysis of the four cases, we can remove the factor k. Let $\sigma = \mathtt{a}_j \mathtt{a}_{j+1} \cdots \mathtt{a}_n (k-2)(k-1)^{j-2}$.

- Suppose $\beta_1 \in \mathbf{A}_k(n)$ and $\text{par}(\beta_1) = \text{lastNonMax}(\beta_1)$. Then Lemma 6 implies that $\sigma \in \mathbf{A}_k(n)$ and $\text{NEXT}(\alpha)$ is either $(k-2)$ or $(k-1)$, depending on whether or not $\text{par}(\sigma) = \text{lastNonMax}(\sigma)$.

- Suppose $\beta_2 \in \mathbf{A}_k(n)$ and $\text{par}(\beta_2) = \text{lastSymbol}(\beta_2) = \delta$. Since $\text{lastNonMax}(\beta_2) \notin \mathbf{A}_k(n)$, from Lemma 6, $a_1 = k-1$. Furthermore, $\delta = 0\mathtt{a}_2 \cdots \mathtt{a}_n$ since from Lemma 6, $\mathtt{a}_n$ is either $(k-2)$ or $(k-1)$. It is easy to see that δ (if not the root) is not joined to $\text{par}(\delta)$ via a conjugate pair containing $0\mathtt{a}_2 \cdots \mathtt{a}_n$. Thus, $\text{NEXT}(\alpha) = 0$.

- Suppose $\beta_3 \in \mathbf{A}_k(n)$ and $\text{par}(\beta_3) = \text{firstNonMin}(\beta_3) = \delta$. Then Lemma 7 implies that $\text{par}(\delta) = \text{lastNonMax}(\delta)$ or $\text{par}(\delta) = \text{lastSymbol}(\delta)$. Thus, δ is not joined to $\text{par}(\delta)$ via a conjugate pair containing the string $(\mathtt{a}_1-1)\mathtt{a}_2 \cdots \mathtt{a}_n$ because it implies $\text{par}(\delta) = \tilde{\alpha}$. Thus $\text{NEXT}(\alpha) = \mathtt{a}_1 - 1$.

- Suppose $\beta_4 \in \mathbf{A}_k(n)$ and $\text{par}(\beta_4) = \text{secondLastNonMax}(\beta_4) = \delta$. Then $k = 3$ and by Lemma 9, β_4 starts with 0 and has suffix 012. This implies δ starts with 0 and has suffix 112. By Lemma 6, $\text{lastNonMax}(\delta) \in \mathbf{A}_3(n)$ and hence $\text{par}(\delta) = \text{lastNonMax}(\delta)$. Thus, $\text{NEXT}(\alpha) = \mathtt{a}_1 + 1$.

The above analysis gives rise to the following successor rule $h(\alpha) = g(\alpha)$ based on $\mathbb{T}_k(n)$.

Successor-rule based on $\mathbb{T}_k(n)$ to construct an $\mathcal{OS}_k(n)$ of length $L_k(n)$ Apply the conditions top down:

$$h(\alpha) = \begin{cases} \texttt{a}_1 - 1 & \text{if } \gamma_1 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\gamma_1) = \mathrm{lastNonMax}(\gamma_1); \\ \texttt{a}_1 - 1 & \text{if } \gamma_2 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\gamma_2) = \mathrm{lastSymbol}(\gamma_2); \\ \texttt{a}_1 + 1 & \text{if } \gamma_3 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\gamma_3) = \mathrm{firstNonMin}(\gamma_3); \\ \texttt{a}_1 - 1 & \text{if } \gamma_4 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\gamma_4) = \mathrm{secondLastNonMax}(\gamma_4); \\[6pt] k-1 & \text{if } \beta_1 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\beta_1) = \mathrm{lastNonMax}(\beta_1) \text{ and } \mathrm{par}(\sigma) = \mathrm{lastNonMax}(\sigma); \\ k-2 & \text{if } \beta_1 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\beta_1) = \mathrm{lastNonMax}(\beta_1); \\ 0 & \text{if } \beta_2 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\beta_2) = \mathrm{lastSymbol}(\beta_2); \\ \texttt{a}_1 - 1 & \text{if } \beta_3 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\beta_3) = \mathrm{firstNonMin}(\beta_3); \\ \texttt{a}_1 + 1 & \text{if } \beta_4 \in \mathbf{A}_k(n) \text{ and } \mathrm{par}(\beta_4) = \mathrm{secondLastNonMax}(\beta_4); \\[6pt] \texttt{a}_1 & \text{otherwise.} \end{cases}$$

▶ **Theorem 13.** *For $n, k \geq 3$, $h(\alpha)$ is a successor rule for an $\mathcal{OS}_k(n)$ of length $L_k(n)$ that runs in $O(n)$ time and uses $O(n)$ space.*

Proof. Let $\alpha = \texttt{a}_1\texttt{a}_2 \cdots \texttt{a}_n \in \mathbf{S}_k(n)$. Our previous analysis demonstrates that $h(\alpha) = g(\alpha)$ for the cycle joining tree $\mathbb{T}_k(n)$. Determining whether or not a string is a symmetric/asymmetric necklace or bracelet can be computed in $O(n)$ time and $O(n)$ space [4, 18]. Thus, all of the membership tester and functions required in the definition of $h(\alpha)$ can be computed in $O(n)$ time using $O(n)$ space. ◀

In the next section, we demonstrate that length of the $\mathcal{OS}_k(n)$ generated by our successor rule $h(\alpha)$ is asymptotically optimal.

6 Bounds on the maximal length of an $\mathcal{OS}_k(n)$

Recall that $L_k(n) = |\mathbf{S}_k(n)|$ and $M_k(n)$ is the maximal length of an $\mathcal{OS}_k(n)$. Our construction in Section 5 demonstrates that $L_k(n) \leq M_k(n)$ or $k \geq 3$. Dai et al. [9] provide the following lower bound $L_2(n)$ for $M_2(n)$, where μ is the Möbius function:

$$L_2(n) = \frac{1}{2}\left(2^n - \sum_{d|n}\mu(n/d)\frac{n}{d}H_2(d)\right), \quad \text{where} \quad H_2(d) = \frac{1}{2}\sum_{i|d}i\left(2^{\lfloor\frac{i+1}{2}\rfloor} + 2^{\lfloor\frac{i}{2}\rfloor+1}\right).$$

Applying the same techniques, this formula can be generalized to $L_k(n)$,

$$L_k(n) = \frac{1}{2}\left(k^n - \sum_{d|n}\mu(n/d)\frac{n}{d}H_k(d)\right), \quad \text{where} \quad H_k(d) = \frac{1}{2}\sum_{i|d}i\left(k^{\lfloor\frac{i+1}{2}\rfloor} + k^{\lfloor\frac{i}{2}\rfloor+1}\right).$$

Exact values of $L_k(n)$ for some small n, k are given in Table 2.

▶ **Theorem 14.** *For $n \geq 3$ and $k \geq 3$, $\displaystyle\lim_{n \to \infty} \frac{M_k(n) - L_k(n)}{L_k(n)} = 0$.*

Proof. Since $\mu(n/d) \leq 1$ and $H_k(d) = \frac{1}{2}\sum_{i|d}i\left(k^{\lfloor\frac{i+1}{2}\rfloor} + k^{\lfloor\frac{i}{2}\rfloor+1}\right) \leq \sum_{i=1}^{d}ik^{\lfloor\frac{i}{2}\rfloor+1} \leq d^2k^{\lfloor\frac{d}{2}\rfloor+1}$, we have

$$\sum_{d|n}\mu(n/d)\frac{n}{d}H_k(d) \leq \sum_{d=1}^{n}\frac{n}{d}d^2k^{\lfloor\frac{d}{2}\rfloor+1} \leq n^3k^{\lfloor\frac{n}{2}\rfloor+1}.$$

Thus, $L_k(n) = \frac{1}{2}\left(k^n - \sum_{d|n} \mu(n/d)\frac{n}{d}H_k(d)\right) \geq \frac{1}{2}\left(k^n - n^3k^{\lfloor\frac{n}{2}\rfloor+1}\right)$. Recalling from Section 1 that $M_k(n) \leq \frac{1}{2}(k^n - k^{\lfloor(n+1)/2\rfloor})$, we have

$$\frac{M_k(n) - L_k(n)}{L_k(n)} \leq \frac{\frac{1}{2}(k^n - k^{\lfloor(n+1)/2\rfloor}) - \frac{1}{2}\left(k^n - n^3k^{\lfloor\frac{n}{2}\rfloor+1}\right)}{\frac{1}{2}\left(k^n - n^3k^{\lfloor\frac{n}{2}\rfloor+1}\right)}$$
$$= \frac{n^3k^{\lfloor\frac{n}{2}\rfloor+1} - k^{\lfloor(n+1)/2\rfloor}}{k^n - n^3k^{\lfloor\frac{n}{2}\rfloor+1}}.$$

The result follows.   ◄

▶ **Corollary 15.** *The $\mathcal{OS}_k(n)$ generated by the successor rule $h(\alpha)$ has asymptotically optimal length.*

──── **References** ────

1   ADAMSON, D., GUSEV, V. V., POTAPOV, I., AND DELIGKAS, A. Ranking bracelets in polynomial time. In *32nd Annual Symposium on Combinatorial Pattern Matching (CPM 2021)* (Dagstuhl, Germany, 2021), P. Gawrychowski and T. Starikovskaya, Eds., vol. 191 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, pp. 4:1–4:17.

2   ALHAKIM, A., MITCHELL, C. J., SZMIDT, J., AND WILD, P. R. Orientable sequences over non-binary alphabets. *manuscript, 2023*.

3   BLANCA, A., AND GODBOLE, A. P. On universal cycles for new classes of combinatorial structures. *SIAM Journal on Discrete Mathematics 25*, 4 (2011), 1832–1842.

4   BOOTH, K. S. Lexicographically least circular substrings. *Inform. Process. Lett. 10*, 4/5 (1980), 240–242.

5   BURNS, J., AND MITCHELL, C. J. Position sensing coding schemes. In *Cryptography and Coding III (M.J.Ganley, ed.)* (1993), Oxford University Press, pp. 31–66.

6   CHUNG, F., DIACONIS, P., AND GRAHAM, R. Universal cycles for combinatorial structures. *Discrete Mathematics 110*, 1 (1992), 43 – 59.

7   CURRIE, J., AND LAFRANCE, P. Avoidability index for binary patterns with reversal. *Electronic J. Combinatorics 23*, 1 (2016), 1–14. #P1.36.

8   CURTIS, D., HINES, T., HURLBERT, G., AND MOYER, T. Near-universal cycles for subsets exist. *SIAM Journal on Discrete Mathematics 23*, 3 (2009), 1441–1449.

9   DAI, Z. D., MARTIN, K. M., ROBSHAW, M. J. B., AND WILD, P. R. Orientable sequences. In *Cryptography and Coding III (M.J.Ganley, ed.)* (1993), Oxford University Press, pp. 97–115.

10 DĘBSKI, M., AND LONC, Z. Universal cycle packings and coverings for k-subsets of an n-set. *Graphs and Combinatorics 32*, 6 (Nov 2016), 2323–2337.

11 DOMARATZKI, M. Hairpin structures defined by DNA trajectories. In *DNA Computing* (2006), C. Mao and T. Yokomori, Eds., vol. 4287 of *Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg, pp. 182–194.

12 ETZION, T. An algorithm for generating shift-register cycles. *Theoret. Comput. Sci. 44*, 2 (1986), 209–224.

13 ETZION, T. Self-dual sequences. *J. Combin. Theory Ser. A 44*, 2 (1987), 288–298.

14 ETZION, T., AND LEMPEL, A. Algorithms for the generation of full-length shift-register sequences. *IEEE Trans. Inform. Theory 30*, 3 (1984), 480–484.

15 FLEISCHER, L., AND SHALLIT, J. O. Words that avoid reversed factors, revisited. Arxiv preprint arXiv:1911.11704 [cs.FL], available at http://arxiv.org/abs/1911.11704, 2019.

16 FREDRICKSEN, H. A survey of full length nonlinear shift register cycle algorithms. *SIAM Review 24*, 2 (1982), 195–221.

17 GABRIĆ, D., AND SAWADA, J. Construction of orientable sequences in $O(1)$-amortized time per bit. Arxiv preprint arXiv:2401.14341v3, available at https://arxiv.org/abs/2401.14341, 2024.

18 GABRIĆ, D., AND SAWADA, J. Efficient construction of long orientable sequences. In *35th Annual Symposium on Combinatorial Pattern Matching (CPM 2024)* (2024), pp. 1–12. to appear.

19 GABRIĆ, D., SAWADA, J., WILLIAMS, A., AND WONG, D. A successor rule framework for constructing k-ary de Bruijn sequences and universal cycles. *IEEE Trans. Inform. Theory 66*, 1 (2020), 679–687.

20 HIERHOLZER, C. Ueber die Möglichkeit, einen Linienzug ohne Wiederholung und ohne Unterbrechung zu umfahren. *Math. Annalen 6* (1873), 30–32.

21 HOLROYD, A. E., RUSKEY, F., AND WILLIAMS, A. Shorthand universal cycles for permutations. *Algorithmica 64*, 2 (2012), 215–245.

22 HUANG, Y. A new algorithm for the generation of binary de Bruijn sequences. *J. Algorithms 11*, 1 (1990), 44–51.

23 HURLBERT, G. On universal cycles for k-subsets of an n-set. *SIAM Journal on Discrete Mathematics 7*, 4 (1994), 598–604.

24 JACKSON, B. W. Universal cycles of k-subsets and k-permutations. *Discrete mathematics 117*, 1 (1993), 141–150.

25 JANSEN, C. J. A., FRANX, W. G., AND BOEKEE, D. E. An efficient algorithm for the generation of DeBruijn cycles. *IEEE Trans. Inform. Theory 37*, 5 (1991), 1475–1478.

26 KARI, L., KONSTANTINIDIS, S., LOSSEVA, E., SOSÍK, P., AND THIERRIN, G. Hairpin structures in DNA words. In *DNA Computing* (2006), A. Carbone and N. A. Pierce, Eds., vol. 3892 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 158–170.

27 KARI, L., KONSTANTINIDIS, S., AND SOSÍK, P. Bond-free languages: Formalizations, maximality and construction methods. *Internat. J. Found. Comp. Sci. 16*, 05 (2005), 1039–1070.

28 KARI, L., KONSTANTINIDIS, S., SOSÍK, P., AND THIERRIN, G. On hairpin-free words and languages. In *Developments in Language Theory* (2005), C. De Felice and A. Restivo, Eds., vol. 3572 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 296–307.

29 MERCAŞ, R. On the aperiodic avoidability of binary patterns with variables and reversals. *Theoret. Comput. Sci. 682* (2017), 180–189.

30 MITCHELL, C. J., AND WILD, P. R. Constructing orientable sequences. *IEEE Trans. Inform. Theory 68*, 7 (2022), 4782–4789.

31 PĂUN, G., ROZENBERG, G., AND SALOMAA, A. *DNA Computing: New Computing Paradigms*. Springer Berlin, Heidelberg, 1998.

32 RAMPERSAD, N., AND SHALLIT, J. O. Words that avoid reversed subwords. *J. Combin. Math. Combin. Comput. 54* (2005), 157–164.

33 RUDOY, Y. An inductive approach to constructing universal cycles on the k-subsets of $[n]$. *The Electronic Journal of Combinatorics 20* (2013), P18.

34 RUSKEY, F., AND WILLIAMS, A. An explicit universal cycle for the (n-1)-permutations of an n-set. *ACM Trans. Algorithms 6*, 3 (July 2010), 1–12.

35 SAWADA, J., SEARS, J., TRAUTRIM, A., AND WILLIAMS, A. Concatenation trees: A framework for efficient universal cycle and de Bruijn sequence constructions. Arxiv preprint arXiv:2308.12405 [math.CO], available at https://arxiv.org/abs/2308.12405, 2023.

36 SAWADA, J., AND WILLIAMS, A. Constructing the first (and coolest) fixed-content universal cycle. *Algorithmica 85*, 6 (2023), 1754–1785.

37 SAWADA, J., WILLIAMS, A., AND WONG, D. Universal cycles for weight-range binary strings. In *Combinatorial Algorithms - 24th International Workshop, IWOCA 2013, Rouen, France, July 10-12, 2013, LNCS 8288* (2013), pp. 388–401.

38 SAWADA, J., AND WONG, D. Efficient universal cycle constructions for weak orders. *Discrete Math. 343*, 10 (2020), 112022.

39 STEVENS, B., BUSKELL, P., ECIMOVIC, P., IVANESCU, C., MALIK, A. M., SAVU, A., VASSILEV, T. S., VERRALL, H., YANG, B., AND ZHAO, Z. Solution of an outstanding conjecture: the non-existence of universal cycles with $k = n - 2$. *Discrete Mathematics 258*, 1 (2002), 193–204.

40 TAKAHASHI, N., KAMEDA, A., YAMAMOTO, M., AND OHUCHI, A. Aqueous computing with DNA hairpin-based RAM. In *DNA Computing* (2005), C. Ferretti, G. Mauri, and C. Zandron, Eds., Springer Berlin Heidelberg, pp. 355–364.

41 TAKINOUE, M., AND SUYAMA, A. Hairpin-DNA memory using molecular addressing. *Small 2*, 11 (2006), 1244–1247.