

A Construction of Short Sequences Containing All Permutations of a Set as Subsequences

Saša Radomirović

Institute of Information Security
ETH Zürich
Zürich, Switzerland

sasa.radomirovic@inf.ethz.ch

Submitted: Feb 9, 2012; Accepted: Oct 29, 2012; Published: Nov 22, 2012

Mathematics Subject Classifications: 68R15, 05A05

Abstract

A sequence over a fixed finite set is said to be *complete* if it contains all permutations of the set as subsequences. Determining the length of shortest complete sequences is an open problem. We improve the existing upper bound and introduce tools to manually prove the completeness of sequences.

1 Introduction

A sequence over a fixed finite set $A = \{a_1, \dots, a_n\}$ is said to be *complete*, if it contains all permutations of $(a_1, \dots, a_n)$ as a (not necessarily consecutive) subsequence. The problem of finding the shortest complete sequences seems to have first been posed by R.M. Karp and has been stated as a research problem in [2, Problem 36].

Newey [7], Adleman [1], Galbiati and Preparata [3], and Mohanty [6] have constructed algorithms which produce complete sequences of length $n^2 - 2n + 4$. Newey has shown that for $3 \leq n \leq 7$, the shortest complete sequences have indeed length $n^2 - 2n + 4$. For instance, two shortest complete sequences over $\{1, 2, 3\}$ are $(1, 2, 1, 3, 1, 2, 1)$ and $(1, 2, 3, 1, 2, 3, 1)$. Newey and Koutas-Hu have conjectured [7, 5] that $n^2 - 2n + 4$ is the shortest possible length of complete sequences over sets with at least three elements. Kleitman and Kwiatkowski [4] have shown a lower bound of $n^2 - C_\epsilon n^{7/4+\epsilon}$ for $\epsilon > 0$, where C_ϵ is a constant depending on ϵ .

Zălinescu [8] has shown that $n^2 - 2n + 3$ is an upper bound for $n \geq 10$. In the present note we provide a tool which can be used to prove the completeness of sequences produced by a variety of algorithms and we show a construction which improves the upper bound to $\lceil n^2 - \frac{7}{3}n + \frac{19}{3} \rceil$ for $n \geq 7$. This bound is lower than Zălinescu's bound for $n \geq 13$.

2 Preliminaries

All sequences considered in this paper are finite sequences $\sigma = (a_1, \dots, a_m)$ over a finite set and are denoted by Greek letters, while their elements are denoted by lowercase Latin letters. We write $(\sigma_1, \dots, \sigma_k)$ for the concatenation of the sequences $\sigma_1, \dots, \sigma_k$. An element c is said to *appear* in the sequence $\sigma = (a_1, \dots, a_m)$ if there is $1 \leq i \leq m$ such that $c = a_i$. We denote the number of appearances of c in σ by $\#_c \sigma$ and we write $\text{del}_c(\sigma)$ for the sequence obtained from σ by deleting all appearances of c .

Let A be a finite set. A sequence is said to be *complete over A* if it contains all permutations of elements of A as a (not necessarily consecutive) subsequence.

We write $[A]_k$ for the set consisting of all sequences over A of length k which contain k distinct elements, that is $[A]_k = \{(a_1, \dots, a_k) \mid a_1, \dots, a_k \in A, a_i \neq a_j \text{ for } 1 \leq i \neq j \leq k\}$. We say that a sequence is *k -complete over A* if it contains all elements of $[A]_k$ as a (not necessarily consecutive) subsequence.

We write $[A]$ instead of $[A]_k$ when $k = |A|$. We write $\text{set}(\sigma)$ for the set of elements appearing in the sequence σ and $[\text{set}(\sigma)]$ or simply $[\sigma]$ for the set of all permutations of $\text{set}(\sigma)$. We abuse notation by writing $\{(\rho, [\sigma], \tau)\}$ for the set of sequences $\{(\rho, \sigma', \tau) \mid \sigma' \in [\sigma]\}$.

We denote by $\sigma_{c^i \dots d^j}$ the consecutive sequence of elements in σ between the i -th appearance of c and the j -th appearance of d in σ and *not* including c^i and d^j . If $\sigma_{c^i \dots d^j}$ is not defined (because c or d does not appear sufficiently many times in σ), we set $\sigma_{c^i \dots d^j} = \infty$, with the convention that ∞ is not complete over any set. By leaving out c^i or d^j in $\sigma_{c^i \dots d^j}$, we indicate the entire prefix of σ up to, but not including d^j , or the entire suffix of σ from, but not including, c^i , respectively. Thus, $\sigma_{\dots d^j}$ denotes the consecutive subsequence starting with the first element of σ up to, but not including, the j -th appearance of d in σ and it denotes ∞ if d appears fewer than j times in σ . We also write $\sigma_{\dots c}$ for $\sigma_{\dots c^\ell}$ when $\ell = \#_c \sigma$ and we write $\sigma_{c \dots}$ for $\sigma_{c^1 \dots}$.

Finally, let $\text{el}_i(\sigma)$ be the i -th element of σ , with $\text{el}_1(\sigma)$ being the first element. We write $\text{el}_{-1}(\sigma)$ for the last element of σ .

3 Properties of complete sequences

We first state the simple observation that a sequence is a complete sequence if and only if for each first appearance of an element, the remaining subsequence is a complete sequence over a smaller set.

Lemma 1. *A sequence σ over A is complete if and only if for all $c \in A$, the sequence $\sigma_{c \dots}$ is complete over $A \setminus \{c\}$.*

Proof. Let $n = |A|$. Let σ be a complete sequence over A and fix $c \in A$. Then σ contains all permutations $(c, c_2, \dots, c_n)$ over A . Write $\sigma = (\sigma_1, c, \sigma_2)$, where σ_1 is a sequence over $A \setminus \{c\}$. Since σ_1 does not contain c , it follows that $(c_2, \dots, c_n)$ is a subsequence of σ_2 for all permutations $(c_2, \dots, c_n)$ over $A \setminus \{c\}$. Therefore $\sigma_{c \dots} = \sigma_2$ is a complete sequence over $A \setminus \{c\}$.

Conversely, let σ be a sequence over A with the property that for all $c \in A$ the sequence $\sigma_{c\dots}$ is complete over $A \setminus \{c\}$. Then for every permutation $(c_1, \dots, c_n)$ over A , we can find σ_1, σ_2 such that $\sigma = (\sigma_1, c_1, \sigma_2)$ and $\sigma_{c_1\dots} = \sigma_2$ is a complete sequence over $A \setminus \{c_1\}$. Thus σ_2 contains $(c_2, \dots, c_n)$, hence σ contains $(c_1, \dots, c_n)$. $\square$

While the preceding lemma characterizes complete sequences over A in terms of first appearance of each element of A , the following lemma and theorem characterize them in terms of all appearances of a *single* element.

In the following we work with the set of sequences $\{(\sigma_{\dots c^i}, [\sigma_{c^i \dots c^{i+1}}], \sigma_{c^{i+1} \dots})\}$. This set denotes all sequences obtained from σ by permuting the set of elements between the i -th and $(i+1)$ -st appearance of c .

Example 2. Let $\sigma = (x, 1, 2, 3, x, 1, 2, x, 3, 1, 2, x)$, then

$$\begin{aligned} & \{(\sigma_{\dots x^2}, [\sigma_{x^2 \dots x^3}], \sigma_{x^3 \dots})\} \\ &= \{(x, 1, 2, 3, [1, 2], 3, 1, 2, x)\} \\ &= \{(x, 1, 2, 3, 1, 2, 3, 1, 2, x), (x, 1, 2, 3, 2, 1, 3, 1, 2, x)\}. \end{aligned}$$

Lemma 3. Let $c \in A$. If σ is a complete sequence over A , then for every $1 \leq i < \#_c \sigma$, every sequence in $\{(\sigma_{\dots c^i}, [\sigma_{c^i \dots c^{i+1}}], \sigma_{c^{i+1} \dots})\}$ is a complete sequence over $A \setminus \{c\}$.

Proof. Let $n = |A|$ and fix an $i \in \{1, \dots, \#_c \sigma - 1\}$. Let $\sigma_1 = \sigma_{\dots c^i}$, $\sigma_2 \in [\sigma_{c^i \dots c^{i+1}}]$, $\sigma_3 = \sigma_{c^{i+1} \dots}$. Let $\rho = (\sigma_1, \sigma_2, \sigma_3)$ and consider a permutation $\phi = (c_1, \dots, c_{n-1})$ over $A \setminus \{c\}$. Let j be the largest non-negative integer such that $(c_1, \dots, c_j)$ is a subsequence of σ_1 . We distinguish two cases:

- $c_{j+1} \in \sigma_2$

Consider the permutation $(c_1, \dots, c_{j+1}, c, c_{j+2}, \dots, c_{n-1})$. By completeness of σ and maximality of j , this permutation is a subsequence of $(\sigma_1, \sigma_2, c, \sigma_3)$.

More specifically, let $\tau = \sigma_{c^i \dots c^{i+1}}$ and note that $\sigma = (\sigma_1, c, \tau, c, \sigma_3)$. By construction, $(c_1, \dots, c_j)$ is a subsequence of σ_1 , but $(c_1, \dots, c_{j+1})$ is not. The element c_{j+1} appears in σ_2 and therefore also in τ . Since c does not appear in τ , $(c_{j+2}, \dots, c_{n-1})$ must be a subsequence of σ_3 by completeness of σ .

It follows that $(c_1, \dots, c_{n-1})$ is a subsequence of ρ .

- $c_{j+1} \notin \sigma_2$

Consider the permutation $(c_1, \dots, c_j, c, c_{j+1}, \dots, c_{n-1})$. By completeness of σ and maximality of j , this permutation is a subsequence of $(\sigma_1, c, \sigma_2, \sigma_3)$.

It follows analogously to the preceding case that $(c_1, \dots, c_{n-1})$ is a subsequence of ρ . $\square$

The preceding lemma states that a complete sequence σ over a set A , with $c \in A$, gives rise to a number of complete sequences over $A \setminus \{c\}$ with the property that the elements between consecutive appearances of c can be arbitrarily permuted without losing the completeness property.

Example 4. The sequence $(x, 1, 2, 3, x, 1, 2, x, 3, 1, 2, x)$ is complete over the set $\{x, 1, 2, 3\}$. Thus, all of the 14 sequences in the following set are complete over $\{1, 2, 3\}$:

$$\{([1, 2, 3], 1, 2, x, 3, 1, 2, x), (x, 1, 2, 3, [1, 2], 3, 1, 2, x), (x, 1, 2, 3, x, 1, 2, [3, 1, 2])\}$$

The following theorem states under which conditions the preceding lemma is not only necessary, but also sufficient for a sequence σ to be complete over A .

Theorem 5. *Let $c \in A$. A sequence σ over A is complete if and only if both of the following conditions are satisfied.*

- (1) $\sigma_{c\dots}$ and $\sigma_{\dots c}$ are complete over $A \setminus \{c\}$.
- (2) For every $1 \leq i < \#_c \sigma$, every sequence in $\{(\sigma_{\dots c^i}, [\sigma_{c^i \dots c^{i+1}}, \sigma_{c^{i+1} \dots}])\}$ is complete over $A \setminus \{c\}$.

Proof. Let $n = |A|$. If σ is a complete sequence, then (1) follows from Lemma 1 (and the fact that the reverse of a complete sequence is a complete sequence) and (2) follows from Lemma 3.

Conversely, let σ be a sequence over A satisfying conditions (1) and (2). Let $\phi = (c_1, \dots, c_n)$ be a permutation over A . We distinguish the following cases.

- c is the first element of ϕ , i.e., $c = c_1$.

Then by condition (1), the sequence $(c_2, \dots, c_n)$ is a subsequence of $\sigma_{c\dots}$, thus σ contains ϕ .

- c is the last element of ϕ , i.e., $c = c_n$.

Then by condition (1), the sequence $(c_1, \dots, c_{n-1})$ is a subsequence of $\sigma_{\dots c}$, thus σ contains ϕ .

- c is neither the first nor the last element of ϕ .

If c appears exactly once in σ , then we are done by the preceding two items. Thus we may assume that c appears at least twice in σ .

Let $c_j = c$, for $1 < j < n$, and write $\sigma = (\sigma_1, \sigma_2)$, where σ_1 is the *shortest consecutive* subsequence of σ which contains $(c_1, \dots, c_{j-1})$ as a subsequence. Such a sequence exists by (1). If c does not appear in σ_1 , then σ contains ϕ by (1). Else, write $\sigma_1 = (\sigma'_1, c, \sigma''_1)$ with σ''_1 being a sequence in which c does not appear. Write $\sigma_2 = (\sigma'_2, c, \sigma''_2)$ with σ'_2 being a sequence in which c does not appear. (Note that c appears in σ_2 by (1).)

Since $\sigma = (\sigma'_1, c, \sigma''_1, \sigma'_2, c, \sigma''_2)$, by (2), each of the sequences in $S = \{(\sigma'_1, [\sigma''_1, \sigma'_2], \sigma''_2)\}$ is a complete sequence over $A \setminus \{c\}$.

Consider $\rho \in [\sigma'_1, \sigma'_2]$ with $\text{el}_{-1}(\rho) = c_{j-1}$. Then $(\sigma'_1, \rho, \sigma''_2) \in S$ and therefore contains the permutation $(c_1, \dots, c_{j-1}, c_{j+1}, \dots, c_n)$. However, $(c_1, \dots, c_{j-1})$ is not a subsequence of σ'_1 , because $\sigma_1 = (\sigma'_1, \sigma''_1)$ is by construction the shortest sequence

containing $(c_1, \dots, c_{j-1})$ and σ''_1 is not the empty sequence. Since ρ is a permutation, $c_{j-1} = \text{el}_{-1}(\rho)$ appears exactly once in ρ . Thus the sequence $(c_1, \dots, c_{j-1}, c_{j+1})$ cannot be a subsequence of (σ'_1, ρ) and hence $(c_{j+1}, \dots, c_n)$ must be a subsequence of σ''_2 . $\square$

Theorem 5 is our main tool to prove that certain algorithms indeed produce complete sequences.

4 Constructing complete sequences

Let σ be a sequence over A . We call a sequence obtained from σ by eliminating consecutive repetitions of elements a *simplification* of σ . Formally, a simplification of σ is the maximal subsequence $(a_1, \dots, a_m)$ such that $a_i \neq a_{i+1}$ for all $1 \leq i < m$. It is obvious that the simplification of a complete sequence is still a complete sequence.

The sequence σ is said to be *quasi-palindromic* if there is a relabeling of its elements which is equal to the reverse of σ . Formally, let $\sigma = (s_1, \dots, s_m)$ and let $\pi : A \rightarrow A$ be a bijective function. Then σ is quasi-palindromic if $\pi(s_i) = s_{m-i+1}$ for all $1 \leq i \leq m$.

We start with a trivial lemma, stating that the concatenation of k permutations of A is k -complete over A .

Lemma 6. *Let $1 \leq k \leq |A|$ and let $\phi_1, \dots, \phi_k \in [A]$ be any permutations of the elements in A . Then the sequence $\sigma = (\phi_1, \dots, \phi_k)$ is a k -complete sequence over A .*

Note that we may choose the permutations $\phi_1, \dots, \phi_k$ such that the last element of one permutation equals the first element of the following permutation, i.e. $\text{el}_{-1}(p_i) = \text{el}_1(p_{i+1})$. Thus after simplification we obtain k -complete sequences over A of length $k|A| - k + 1$. The following corollary states a slightly more general fact for future reference.

Corollary 7. *Let $1 \leq k \leq |A|$ and let $\sigma = (\sigma_1, \dots, \sigma_k)$ be such that $\text{set}(\sigma_1) = A$ and for $1 < i \leq k$*

$$\text{set}(\text{el}_{-1}(\sigma_{i-1}), \sigma_i) = A. \quad (1)$$

Then σ is a k -complete sequence over A .

Proof. Let $\sigma'_1 = \sigma_1$ and for $1 < i \leq k$ let $\sigma'_i = (\text{el}_{-1}(\sigma_{i-1}), \sigma_i)$. Note that the simplification of the sequence $(\sigma'_1, \dots, \sigma'_k)$ is equal to σ . By equation (1), for each i , $1 \leq i \leq k$, there is a subsequence ϕ_i of σ'_i which is a permutation of the elements of A . Thus the simplification of $(\phi_1, \dots, \phi_k)$ is a subsequence of σ which by Lemma 6 and the observation following it is a k -complete sequence. $\square$

The conditions in the following theorem are related to the constructions of Newey [7] and Adleman [1].

Theorem 8. Let $n = |A|$. Let $\sigma = (\sigma_0, \sigma_1, \dots, \sigma_n, \sigma_{n+1})$ be such that σ_0, σ_{n+1} are arbitrary, possibly empty, sequences over A and for $1 \leq i \leq n$

$$\text{set}(\text{el}_{-1}(\sigma_{i-1}), \sigma_i) = \text{set}(\sigma_i, \text{el}_1(\sigma_{i+1})) = A. \quad (2)$$

Then $(\sigma_0, x, \sigma_1, x, \sigma_2, x, \dots, x, \sigma_n, x, \sigma_{n+1})$ is a complete sequence over $A \cup \{x\}$.

Proof. We apply Theorem 5 with respect to x . We observe first that $(\sigma_0, \dots, \sigma_n)$ and $(\sigma_1, \dots, \sigma_{n+1})$ are complete sequences over A by Corollary 7: The former immediately due to the hypothesis $\text{set}(\text{el}_{-1}(\sigma_{i-1}), \sigma_i) = A$ for $1 \leq i \leq n$. The latter by using the hypothesis $\text{set}(\sigma_i, \text{el}_1(\sigma_{i+1})) = A$ for $1 \leq i \leq n$ and considering the reverse sequence. Thus condition (1) of Theorem 5 is satisfied.

For $1 \leq j \leq n$, let $\bar{\sigma}_j$ denote the concatenation of the sequences $\sigma_1, \dots, \sigma_n$, in which σ_j is replaced by any permutation $\phi(\sigma_j) \in [\sigma_j]$ of the elements in σ_j . That is $\bar{\sigma}_j = (\sigma_1, \dots, \sigma_{j-1}, \phi(\sigma_j), \sigma_{j+1}, \dots, \sigma_n)$. Then $(\sigma_1, \dots, \sigma_{j-1}, \phi(\sigma_j))$ is a j -complete sequence over A by Corollary 7 and $(\sigma_{j+1}, \dots, \sigma_n)$ is $(n-j)$ -complete by applying Corollary 7 to the reverse of this sequence. Thus $\bar{\sigma}_j$ is a complete sequence since it is the concatenation of a j -complete and an $(n-j)$ -complete sequence. This satisfies condition (2) of Theorem 5. $\square$

Shortest known sequences which satisfy the conditions of the preceding theorem are given by the constructions of Newey [7] and Adleman [1] and have length $m^2 - 2m + 4$, where $m > 2$ is the number of elements over which the sequences are complete. For $m \geq 10$ we can produce shorter sequences with the construction given in the Main Theorem below.

Before stating the Main Theorem and its proof, we illustrate the idea in the following example.

Example 9. We claim that the sequence σ over the set $\{0, \dots, 9, a, \mathbf{x}, \mathbf{y}\}$ of 13 elements shown below is complete. The sequence has length 145, which is one element shorter than the length obtained with the construction in [8].

$\mathbf{x}$ 0123456789a $\mathbf{yx}$ 0123456789a $\mathbf{xy}$ 0123456789 $\mathbf{x}$
 a01234567y8 $\mathbf{x}$ 9a01234567 $\mathbf{x}$ 8y9a0123456 $\mathbf{x}$
 789a01234y5 $\mathbf{x}$ 6789a01234 $\mathbf{x}$ 5y6789a0123 $\mathbf{x}$
 456789a012 $\mathbf{yx}$ 3456789a012 $\mathbf{xy}$ 3456789a012 $\mathbf{x}$

The sequence is displayed in a manner that highlights the construction pattern. The element $\mathbf{x}$ separates blocks of the repeated enumeration of the elements $(0, \dots, 9, a)$ in accordance with the structure given in Theorem 8. The savings in terms of the overall length of the sequence, compared to previously published constructions, are achieved with the large gaps between subsequent appearances of $\mathbf{y}$.

To see that the sequence is complete over $\{0, \dots, 9, a, \mathbf{x}, \mathbf{y}\}$, we apply Theorem 5 with respect to element $\mathbf{y}$. We verify the two conditions stated in the theorem separately and take advantage of the fact that σ is quasi-palindromic in order to reduce the number of case distinctions. The sequence's quasi-palindromicity can be seen by applying the relabeling $(0, \dots, 9, a) \mapsto (2, 1, 0, a, 9, \dots, 3)$ to its elements.

- Condition (1) of Theorem 5 is satisfied because $\sigma_{\mathbf{y}\dots}$ and $\sigma_{\dots\mathbf{y}}$ are complete over $A \setminus \{\mathbf{y}\}$.

The sequence $\sigma_{\mathbf{y}\dots}$ is complete over $A \setminus \{\mathbf{y}\}$ because $\text{del}_{\mathbf{y}}(\sigma_{\mathbf{y}\dots})$ is complete over $A \setminus \{\mathbf{y}\}$ by Theorem 8. The same reasoning applies to the sequence $\sigma_{\dots\mathbf{y}}$.

- To verify condition (2) of Theorem 5, we show that all sequences in the sets $\{(\sigma_{\dots\mathbf{y}^i}, [\sigma_{\mathbf{y}^i\dots\mathbf{y}^{i+1}}], \sigma_{\mathbf{y}^{i+1}\dots})\}$, for $i \in \{1, 2, 3, 4\}$, are complete over $A \setminus \{\mathbf{y}\}$.

Completeness of sequences in the sets $\{(\sigma_{\dots\mathbf{y}^i}, [\sigma_{\mathbf{y}^i\dots\mathbf{y}^{i+1}}], \sigma_{\mathbf{y}^{i+1}\dots})\}$, for $i \in \{5, 6, 7\}$, follows from the fact that σ is quasi-palindromic.

In the following sequences $\mathbf{y}$ is annotated with the ordinal of its appearance in σ for ease of reading. Furthermore, we frequently repeat the last element of a line on the next line. This does not change the completeness property, since such sequences are identical with the original sequence after simplification.

- The sequences in $(\sigma_{\dots\mathbf{y}^1}, [\sigma_{\mathbf{y}^1\dots\mathbf{y}^2}], \sigma_{\mathbf{y}^2\dots})$ and $(\sigma_{\dots\mathbf{y}^3}, [\sigma_{\mathbf{y}^3\dots\mathbf{y}^4}], \sigma_{\mathbf{y}^4\dots})$ are, respectively,

$\mathbf{x}0123456789a [\mathbf{x}0123456789a\mathbf{x}] 0123456789\mathbf{x}a$
 $01234567\mathbf{y}^38\mathbf{x}9a 01234567\mathbf{x}8\mathbf{y}^49 a0123456\mathbf{x}78$
 $89a01234\mathbf{y}^55\mathbf{x}67 89a01234\mathbf{x}5\mathbf{y}^66 789a0123\mathbf{x}45$
 $56789a012\mathbf{y}^7\mathbf{x}34 56789a012\mathbf{x}\mathbf{y}^83 456789a012\mathbf{x}$

and

$\mathbf{x}0123456789a \mathbf{y}^1\mathbf{x}0123456789 a\mathbf{x}\mathbf{y}^2012345678$
 $89\mathbf{x}a01234567 [8\mathbf{x}9a01234567\mathbf{x}8] 9a0123456\mathbf{x}78$
 $89a01234\mathbf{y}^55\mathbf{x}67 89a01234\mathbf{x}5\mathbf{y}^66 789a0123\mathbf{x}45$
 $56789a012\mathbf{y}^7\mathbf{x}34 56789a012\mathbf{x}\mathbf{y}^83 456789a012\mathbf{x}$.

Note that every line in the sequences shown above is 3-complete over $A \setminus \{\mathbf{y}\}$. This follows from Corollary 7 after deleting all appearances of $\mathbf{y}$ in the sequences above. Therefore the sequences are concatenations of four 3-complete sequences over $A \setminus \{\mathbf{y}\}$, thus 12-complete sequences over a set of 12 elements.

- The sequences in $(\sigma_{\dots\mathbf{y}^2}, [\sigma_{\mathbf{y}^2\dots\mathbf{y}^3}], \sigma_{\mathbf{y}^3\dots})$ and $(\sigma_{\dots\mathbf{y}^4}, [\sigma_{\mathbf{y}^4\dots\mathbf{y}^5}], \sigma_{\mathbf{y}^5\dots})$ are, respectively,

$\mathbf{x} 0123456789a\mathbf{y}^1 \mathbf{x} 0123456789a \mathbf{x} [0123456789\mathbf{x}a]8 \mathbf{x}$
 $9a01234567 \mathbf{x} 8\mathbf{y}^49a0123456 \mathbf{x} 789a01234\mathbf{y}^55 \mathbf{x}$
 $6789a01234 \mathbf{x} 5\mathbf{y}^66789a0123 \mathbf{x} 456789a012\mathbf{y}^7 \mathbf{x}$
 $3456789a012 \mathbf{x} \mathbf{y}^83456789a012 \mathbf{x}$

and

$$\begin{aligned} & \mathbf{x} \ 0123456789a \ \mathbf{y}^1 \mathbf{x} \ 0123456789a \ \mathbf{x} \mathbf{y}^2 \ 0123456789 \ \mathbf{x} \\ & a01234567\mathbf{y}^3 8 \ \mathbf{x} \ 9a01234567 \ \mathbf{x} \ 8[0123456\mathbf{x}789a]5 \ \mathbf{x} \\ & 6789a01234 \ \mathbf{x} \ 5\mathbf{y}^6 6789a0123 \ \mathbf{x} \ 456789a012 \ \mathbf{y}^7 \mathbf{x} \\ & 3456789a012 \ \mathbf{x} \mathbf{y}^8 \ 3456789a012 \ \mathbf{x}. \end{aligned}$$

The completeness of these sequences over $A \setminus \{\mathbf{y}\}$ follows from Theorem 8 after deleting all appearances of $\mathbf{y}$ and the appearance of $\mathbf{x}$ inside the brackets.

For ease of presentation of the Main Theorem and its proof, let σ^x denote the sequence obtained from σ by introducing x as the second-to-last element. Similarly, let $^x\sigma$ denote the sequence obtained from σ by introducing x as the second element. That is, for $\sigma = (a_1, \dots, a_m)$, $\sigma^x = (a_1, \dots, a_{m-1}, x, a_m)$ and $^x\sigma = (a_1, x, a_2, \dots, a_m)$.

Main Theorem. *Let $A = \{0, 1, 2, \dots, n-1\}$ be such that $n > 4$, and $n+1$ is divisible by 3. Let $\sigma = (\sigma_1, \dots, \sigma_{n+1})$ be such that*

$$\begin{aligned} \sigma_1 &= \sigma_2 = (0, 1, \dots, n-1) \\ \sigma_i &= (3-i, 4-i, \dots, n+2-i) \pmod{n}, \text{ for } 3 \leq i \leq n-1 \\ \sigma_n &= \sigma_{n+1} = (3, \dots, n-1, 0, 1, 2) \end{aligned}$$

Then the sequence

$$\begin{aligned} \tau &= (x, \sigma_1, y, x, \sigma_2, x, y, \sigma_3, x, \\ & \sigma_{3j-2}^y, x, \sigma_{3j-1}, x, \sigma_{3j}^y, x, \quad (\text{for } 1 < j < n/3) \\ & \sigma_{n-1}, y, x, \sigma_n, x, y, \sigma_{n+1}, x) \end{aligned}$$

is a complete sequence over $A \cup \{x, y\}$. Its length is $m^2 - \frac{7}{3}m + \frac{19}{3}$, where $m = n+2$ is the number of elements in $A \cup \{x, y\}$.

Proof. We first compute the length of the generated sequence, then prove its completeness.

To obtain the count of $m^2 - \frac{7}{3}m + \frac{19}{3}$ elements in τ , note that there are $m = n+2$ appearances of x , $2(n+1)/3$ appearances of y , $n-1$ elements in $\sigma_3, \dots, \sigma_{n-1}$, and n elements in $\sigma_1, \sigma_2, \sigma_n, \sigma_{n+1}$.

To prove completeness, we apply Theorem 5 with respect to element y in the sequence. We note that the sequence is quasi-palindromic, as can be seen by applying the relabeling $(0, 1, 2, \dots, n-1) \mapsto (2, 1, 0, n-1, \dots, 3)$ to its elements.

For condition (1) of Theorem 5, we note that the sequences $\tau_{y\dots}$ and $\tau_{\dots y}$ are complete by Theorem 8. Written out explicitly, these are the following two sequences.

$$\begin{aligned} \tau_{y\dots} &= (x, \sigma_2, x, \sigma_3, x, \\ & \sigma_{3i-2}, x, \sigma_{3i-1}, x, \sigma_{3i}, x, \quad (\text{for } 1 < i < n/3) \\ & x, \sigma_{n-1}, x, \sigma_n, x, \sigma_{n+1}, x) \end{aligned}$$

and

$$\begin{aligned}\tau_{\dots y} = & (x, \sigma_1, x, \sigma_2, x, \sigma_3, x, \\ & \sigma_{3i-2}, x, \sigma_{3i-1}, x, \sigma_{3i}, x, \quad (\text{for } 1 < i < n/3) \\ & x, \sigma_{n-1}, x, \sigma_n, x).\end{aligned}$$

For condition (2) of Theorem 5, we distinguish three cases, according to the positions of y^i . For convenience and ease of reference, we explicitly enumerate y in the following expression.

$$\begin{aligned}\tau = & (x, \sigma_1, y^1, x, \sigma_2, x, y^2, \sigma_3, x, \\ & \sigma_{3j-2}^{y^{2j-1}}, x, \sigma_{3j-1}, x, y^{2j} \sigma_{3j}, x, \quad (\text{for } 1 < j < n/3) \\ & \sigma_{n-1}, y^{2/3(n+1)-1}, x, \sigma_n, x, y^{2/3(n+1)}, \sigma_{n+1}, x)\end{aligned}$$

The three cases are as follows. Condition (2) of Theorem 5 is satisfied for

1. the first and last pair of y 's.

Let $L_1 = \{(\tau_{\dots y^1}, [\tau_{y^1 \dots y^2}], \tau_{y^2 \dots})\}$ be the set of sequences obtained by eliminating the first pair of y 's. Note that the sequence

$$\begin{aligned}\text{del}_y(\tau_{\dots y^1}, x, \tau_{y^2 \dots}) = & (x, \sigma_1, x, \sigma_3, x, \\ & \sigma_{3i-2}, x, \sigma_{3i-1}, x, \sigma_{3i}, x, \quad (\text{for } 1 < i < n/3) \\ & x, \sigma_{n-1}, x, \sigma_n, x, \sigma_{n+1}, x)\end{aligned}$$

is complete over $A \cup \{x\}$ by Theorem 8 and is a subsequence of all sequences in L_1 .

Similarly, let the set of sequences obtained by eliminating the last pair of y 's be $L_{2/3(n+1)-1} = \{(\tau_{\dots y^{2/3(n+1)-1}}, [\tau_{y^{2/3(n+1)-1} \dots y^{2/3(n+1)}}], \tau_{y^{2/3(n+1)} \dots})\}$. Since τ is quasi-palindromic, the sequences in $L_{2/3(n+1)-1}$ are complete over $A \cup \{x\}$, too.

2. even to odd appearance of y .

Let $L_{2i} = \{(\tau_{\dots y^{2i}}, [\tau_{y^{2i} \dots y^{2i+1}}], \tau_{y^{2i+1} \dots})\}$, $1 \leq i < n/3$.

If we delete y from the sequences in L_{2i} (for $1 < i < n/3-1$) we obtain the sequences

$$\begin{aligned}& (x, \sigma_1, x, \sigma_2, x, \dots, \\ & \sigma_{3i-1}, x, \text{el}_1(\sigma_{3i}), [A \cup \{x\}], \text{el}_{-1}(\sigma_{3i+1}), x, \sigma_{3i+2}, \dots, \\ & x, \sigma_{n-1}, x, \sigma_n, x, \sigma_{n+1}, x),\end{aligned}$$

all of which are complete over $A \cup \{x\}$ by Theorem 8. For $i = 1$, the element $\text{el}_1(\sigma_3)$ does not appear. In this case, the sequences in L_2 are complete over $A \cup \{x\}$ by Theorem 8 because σ_2 contains all elements of A . Since τ is quasi-palindromic, the sequences in $L_{2/3(n+1)-2}$ are complete over $A \cup \{x\}$.

3. odd to even appearance of y .

Let $L_{2i-1} = \{(\tau_{\dots y^{2i-1}}, [\tau_{y^{2i-1} \dots y^{2i}}], \tau_{y^{2i} \dots})\}$, $1 < i < n/3$.

Let σ'_i denote the sequence σ_i without the last element and let $'\sigma_i$ denote the sequence σ_i without the first element.

By Corollary 7, the sequences $\lambda_1 = (x, \sigma_1, x, \sigma_2, x, \sigma'_3)$ and

$$\lambda_j = (\text{el}_{-1}(\sigma'_{3j-3}), \text{el}_{-1}(\sigma_{3j-3}), x, \sigma_{3j-2}, x, \sigma_{3j-1}, x, \sigma'_{3j}),$$

for $1 < j \leq (n+1)/3$, are 3-complete over $A \cup \{x\}$.

By quasi-palindromicity, the sequence $\rho_{(n+1)/3} = (' \sigma_{n-1}, x, \sigma_n, x, \sigma_{n+1}, x)$ and the sequences

$$\rho_j = (' \sigma_{3j-2}, x, \sigma_{3j-1}, x, \sigma_{3j}, x, \text{el}_1(\sigma_{3j+1}), \text{el}_1(' \sigma_{3j+1})),$$

for $1 \leq j < (n+1)/3$, are 3-complete, too.

Finally, for $1 < j < (n+1)/3$, each of the sequences in

$$M_j = \{(\text{el}_{-1}(\sigma'_{3j-3}), \text{el}_{-1}(\sigma_{3j-3}), x, \sigma'_{3j-2}, [\tau_{y^{2j-1} \dots y^{2j}}], \\ ' \sigma_{3j}, x, \text{el}_1(\sigma_{3j+1}), \text{el}_1(' \sigma_{3j+1}))\}$$

is 3-complete by Lemma 6.

Observe that each of the sequences in L_{2i-1} can be written as the simplification of $(\lambda_1, \dots, \lambda_{i-1}, \mu_i, \rho_{i+1}, \dots, \rho_{(n+1)/3})$ for some $\mu_i \in M_i$. Thus each sequence in L_{2i-1} is the concatenation of $(n+1)/3$ sequences which are 3-complete. Since there are $n+1$ elements in $A \cup \{x\}$, it follows that each sequence in L_{2i-1} is complete over $A \cup \{x\}$. $\square$

Corollary 10. *Let A be a set of $m > 4$ elements. Then there is a complete sequence of length $\lceil m^2 - \frac{7}{3}m + \frac{19}{3} \rceil$ over A .*

Proof. By the Main Theorem, we can construct sequences of length $\lceil m^2 - \frac{7}{3}m + \frac{19}{3} \rceil$ over a set of m elements whenever 3 divides $m-1$.

If $m-1 \equiv 1$ or $2 \pmod{3}$, then we first construct a sequence τ over a set with $l = m+2$ or $m+1$ elements, respectively, according to the Main Theorem. From such a sequence, we can obtain sequences over a smaller set by removing $c = \text{el}_{-1}(\sigma_1)$ from τ , as well as all the elements before the first appearance of c . The resulting sequence, $\text{del}_c(\tau_{c\dots})$, is complete by Lemma 1. Since the removed element appears $l-2$ times and there are $l-2$ elements preceding it, the resulting sequence has $\lceil l^2 - \frac{7}{3}l + \frac{19}{3} \rceil - 2l + 4 = \lceil (l-1)^2 - \frac{7}{3}(l-1) + \frac{19}{3} \rceil$ elements.

If $m-1 \equiv 1 \pmod{3}$, we can repeat this process once more with $\text{el}_{-1}(\sigma'_2)$. This is the last element to appear for the first time in the sequence $\text{del}_c(\tau_{c\dots})$ which resulted from the previous step. Removing all appearances of $\text{el}_{-1}(\sigma'_2)$ as well as all elements which precede its first appearance removes $2l-5$ elements from the sequence, thus we obtain a sequence of $\lceil (l-2)^2 - \frac{7}{3}(l-2) + \frac{19}{3} \rceil$ elements. $\square$

References

- [1] L. Adleman. Short permutation strings. *Discrete Math.*, 10:197–200, 1974.
- [2] V. Chvátal, D. A. Klarner, and D. E. Knuth. Selected combinatorial research problems. Technical Report STAN-CS-72-292, Stanford University, Department of Computer Science, Stanford, CA, USA, June 1972. infolab.stanford.edu/pub/cstr/reports/cs/tr/72/292/CS-TR-72-292.pdf.
- [3] G. Galbiati and F. P. Preparata. On permutation-embedding sequences. *SIAM Journal on Applied Mathematics*, 30(3):421–423, May 1976.
- [4] D. J. Kleitman and D. J. Kwiatkowski. A lower bound on the length of a sequence containing all permutations as subsequences. *J. Combinatorial Theory Ser. A*, 21(2):129–136, 1976.
- [5] P. J. Koutas and T. C. Hu. Shortest string containing all permutations. *Discrete Math.*, 11:125–132, 1975.
- [6] S. P. Mohanty. Shortest string containing all permutations. *Discrete Math.*, 31:91–95, 1980.
- [7] M. C. Newey. Notes on a problem involving permutations as subsequences. Technical Report STAN-CS-73-340, Stanford University, Department of Computer Science, Stanford, CA, USA, March 1973. infolab.stanford.edu/pub/cstr/reports/cs/tr/73/340/CS-TR-73-340.pdf.
- [8] E. Zălinescu. Shorter strings containing all k -element permutations. *Information Processing Letters*, 111(12):605–608, 2011.