Shorter strings containing all k-element permutations

Eugen Zălinescu

Computer Science Department, ETH Zurich, Switzerland

Abstract

We consider the problem of finding short strings that contain all permutations of order k over an alphabet of size n, with $k \leq n$. We show constructively that $k(n-2)+3$ is an upper bound on the length of shortest such strings, for $n \geq k \geq 10$. Consequently, for $n \geq 10$, the shortest strings that contain all permutations of order n have length at most $n^2 - 2n + 3$. These two new upper bounds improve with one unit the previous known upper bounds.

Keywords: Combinatorial problems, Permutations, Subsequences

1. Introduction

The problem of finding (shortest) strings that contain, as subsequences, all permutations of order n was identified by R. Karp in 1971, and stated in [3]. In the sequel, it has been investigated in [7, 1, 5, 4, 6], which show constructively that $n^2 - 2n + 4$ is an upper bound on the length of shortest such strings, each work presenting a different construction. Moreover, M. Newey [7] showed that this bound is tight for $n \leq 7$. However, while it is tempting to conjecture tightness also for $n \geq 8$, counterexamples of length 83 and 102 have been recently and anonymously identified [2] for $n = 10$ and respectively $n = 11$. Motivated by these counterexamples, we present a construction of strings of length $n^2 - 2n + 3$ that contain all permutations of order n, for any $n \geq 10$.

A generalization of the above problem consists of finding (shortest) strings that contain all permutations of order k over an alphabet of size n. (For $k = n$ we find the original problem.) M. Newey [7] and C. Savage [8] present different constructions which show that $k(n-2)+4$ is an upper bound on the length of shortest such strings. Following C. Savage's approach, we use the construction from the particular case to build strings of length $k(n-2)+3$ that contain all permutations of order k over an alphabet of size n. We thus obtain the new upper bound $k(n-2)+3$ in the general case, for $n \geq k \geq 10$.

2. Preliminaries

We mainly use L. Adleman's [1] notations.

When a string σ is a subsequence of a string γ, we say that γ *contains* σ, or σ is contained in γ. For instance, abc contains ac. Given a set of strings Q, the string γ is *Q-complete* if γ contains each string in Q. For example, aba is $\{ab, ba\}$-complete.

For an alphabet Σ, we let P_Σ denote the set of all permutations of Σ, and L^k_Σ denote the set of all strings in Σ^* of length k in which no letter occurs twice.

For any $n \geq 1$, Σ_n denotes an alphabet of size n. We assume without loss of generality that $\Sigma_n = \{1, \ldots, n\}$. P_n abbreviates P_{Σ_n} and L^k_n abbreviates $L^k_{\Sigma_n}$. Note that $L^1_n = \Sigma_n$, $L^n_n = P_n$, and any element of P_n is L^1_n-complete. For instance, $P_3 = \{123, 132, 213, 231, 312, 321\}$, $L^2_3 = \{12, 13, 21, 23, 31, 32\}$, the string 1213121 is P_3-complete, and the string 1122332211 is L^2_3-complete.

Concatenation of strings is written as juxtaposition. The notation is extended to sets of strings as expected. For example, $\gamma Q := \{\gamma\sigma \mid \sigma \in Q\}$.

A string $\sigma = s_1 s_2 \ldots s_k \in \Sigma^*_n$ with $s_i \in \Sigma_n$ is an *R_n-string* if $s_{i+1} = (s_i \bmod n) + 1$ for all $1 \leq i \leq k-1$. We denote by $R^\ell_n(a)$ the R_n-string of length ℓ that starts with the letter a. For instance, $R^5_3(2) = 23123$.

3. Two previous constructions

Before presenting the details of our construction, we sketch the ones of [1] and [8].

L. Adleman's construction [1] starts from the string $R^{(n-2)(n-1)+2}_{n-1}(1)$. Next, the string is split into $n-1$ strings ($n-3$ strings of length $n-2$, first and last string of length $n-1$); we call these strings α-blocks. Finally, the letter n is inserted between the blocks, before the first block, and after the last block. The string thus obtained is P_n-complete. The proof is based on the following key property: any R_n-string of length $k(n-1)+1$ is L^k_n-complete.

C. Savage [8] generalizes L. Adleman's construction by starting instead from the string $R^{(k-2)(k-1)+2}_{k-1}(1)$, splitting it into $k-1$ blocks, and inserting the string[1] $k(k+1)\ldots n$ between the blocks, before the first block, and after the last block. The obtained string is L^k_n-complete and the proof is based on the same key property.

Our approach is similar. We start directly from $k-1$ so-called β-blocks, which are build from α-blocks by inserting the letter $(k-1)$ in appropriate positions. We then apply C. Savage's construction on the β-blocks, inserting the string $k(k+1)\ldots n$ between, before, and after the β-blocks, to obtain an L^k_n-complete string. However, as we do not have an analogue of the key property for β-blocks, our proof is completely different from the ones in [1] and [8].

4. A construction of L^k_n-complete strings of length $k(n-2)+3$

From now on, we fix two integers n and k, with $n \geq k \geq 10$. Also, we denote the letter $(k-1)$ by A.

[1] Actually, instead of the string $k(k+1)\ldots n$, any permutation of $\{k, k+1, \ldots, n\}$ would do.

The construction makes use of the so-called α-blocks and β-blocks. The *α-blocks* $\alpha_1, \ldots, \alpha_{k-1}$ are the following strings:

$$\begin{aligned}
&\alpha_1 := \alpha_2 := R_{k-2}^{k-2}(1) = 12\ldots(k-2),\\
&\alpha_j := R_{k-2}^{k-3}(k-j+2), \text{ for } 3 \le j \le k-3,\\
&\alpha_{k-1} := \alpha_{k-2} := R_{k-2}^{k-2}(4) = 45\ldots(k-2)123.
\end{aligned}$$

The *β-blocks* $\beta_1, \ldots, \beta_{k-1}$ are the following strings:

$$\begin{aligned}
&\beta_1 := \alpha_1 A, \quad \beta_2 := \alpha_2, \quad \beta_3 := A\alpha_3,\\
&\beta_j := a_1 a_2 \ldots a_{k-4} A a_{k-3}, \text{ for } 4 \le j \le \lfloor \tfrac{k}{2} \rfloor - 1,\\
&\beta_j := \alpha_j, \text{ for } \lfloor \tfrac{k}{2} \rfloor \le j \le \lceil \tfrac{k}{2} \rceil\\
&\beta_j := a_1 A a_2 \ldots a_{k-4} a_{k-3}, \text{ for } \lceil \tfrac{k}{2} \rceil + 1 \le j \le k-4,\\
&\beta_{k-3} := \alpha_{k-3} A, \quad \beta_{k-2} := \alpha_{k-2}, \quad \beta_{k-1} := A\alpha_{k-1},
\end{aligned}$$

where $\alpha_j = a_1 \ldots a_{k-3}$ for $4 \le j \le k-4$. Note that $\lfloor \frac{k}{2} \rfloor \le j \le \lceil \frac{k}{2} \rceil$ is equivalent with $j = \ell$ when $k = 2\ell$, and with $j \in \{\ell, \ell+1\}$ when $k = 2\ell+1$.

Table 1 details the β-blocks in all cases. A letter $b \in \Sigma_{k-1}$ is missing from β_j if b does not occur in β_j.

Let $\Sigma_{k,n} := \{k, \ldots, n\}$ and $\Sigma_{A,n} := \Sigma_{k,n} \cup \{A\}$. The set of strings T_n^k is

$$\begin{aligned}
\big\{\tau_1\beta_1\tau_2\ldots\tau_{k-1}\beta_{k-1}\tau_k \mid \tau_j \in P_{\Sigma_{A,n}} &\text{ for } k \text{ odd and } j = \lceil \tfrac{k}{2} \rceil,\\
\tau_j \in P_{\Sigma_{k,n}} &\text{ otherwise}\big\}.
\end{aligned}$$

That is, strings in T_n^k are obtained by concatenating the β-blocks and inserting between them, before the first block, and after the last block, arbitrary permutations of the set $\Sigma_{k,n}$; with one exception: for k odd ($k = 2\ell+1$), the block inserted in the middle (that is, $\tau_{\ell+1}$) is an arbitrary permutation of $\Sigma_{A,n}$.

The set we have just built is the focus of this paper. As examples, the following two strings are in T_{10}^{10} and respectively in T_{13}^{11}:

- $a \cdot 12345678A \cdot a \cdot 12345678 \cdot a \cdot A1234567 \cdot a \cdot 812345A6 \cdot a \cdot 7812345 \cdot a \cdot 6A781234 \cdot a \cdot 5678123A \cdot a \cdot 45678123 \cdot a \cdot A45678123 \cdot a$ and
- $bcd \cdot 123456789A \cdot cbd \cdot 123456789 \cdot dbc \cdot A12345678 \cdot cbd \cdot 9123456A7 \cdot bcd \cdot 89123456 \cdot bAdc \cdot 78912345 \cdot dcb \cdot 6A7891234 \cdot dcb \cdot 56789123A \cdot dbc \cdot 456789123 \cdot bdc \cdot A456789123 \cdot bdc$,

where $\cdot$ denotes concatenation and delimits the β-blocks, and a, b, c, and d denote the letters 10, 11, 12, and respectively 13. We recall that $A = 9$ in the first string, and $A = 10$ in the second.

5. Main result

This section is devoted to the proof of the following theorem.

Theorem 1. *Any string in T_n^k is L_n^k-complete and has length $k(n-2)+3$.*

	j	β_j	$\lvert\beta_j\rvert$	missing
	1	$1\,2\,3\ldots(k-4)\,(k-3)\,(k-2)\,A$	$k-1$	-
	2	$1\,2\,3\ldots(k-4)\,(k-3)\,(k-2)$	$k-2$	A
	3	$A\,1\,2\ldots(k-5)\,(k-4)\,(k-3)$	$k-2$	$k-2$
	4	$(k-2)\,1\,2\ldots(k-5)\,A\,(k-4)$	$k-2$	$k-3$
	5	$(k-3)\,(k-2)\,1\ldots(k-6)\,A\,(k-5)$	$k-2$	$k-4$
		$\vdots$		
	j	$(k-j+2)\ldots(k-j-1)\,A\,(k-j)$	$k-2$	$k-j+1$
		$\vdots$		
$k=2\ell$	$\ell-1$	$(\ell+3)\,(\ell+4)\,(\ell+5)\ldots\ell\,A\,(\ell+1)$	$k-2$	$\ell+2$
$k=2\ell$	ℓ	$(\ell+2)\,(\ell+3)\,(\ell+4)\ldots(\ell-2)\,(\ell-1)\,\ell$	$k-3$	$\ell+1, A$
$k=2\ell$	$\ell+1$	$(\ell+1)\,A\,(\ell+2)\ldots(\ell-3)\,(\ell-2)\,(\ell-1)$	$k-2$	ℓ
		or		
$k=2\ell+1$	$\ell-1$	$(\ell+4)\,(\ell+5)\,(\ell+6)\ldots(\ell+1)\,A\,(\ell+2)$	$k-2$	$\ell+3$
$k=2\ell+1$	ℓ	$(\ell+3)\,(\ell+4)\,(\ell+5)\ldots(\ell-1)\,\ell\,(\ell+1)$	$k-2$	$\ell+2, A$
$k=2\ell+1$	$\ell+1$	$(\ell+2)\,(\ell+3)\,(\ell+4)\ldots(\ell-2)\,(\ell-1)\,\ell$	$k-3$	$\ell+1, A$
$k=2\ell+1$	$\ell+2$	$(\ell+1)\,A\,(\ell+2)\ldots(\ell-3)\,(\ell-2)\,(\ell-1)$	$k-2$	ℓ
		$\vdots$		
	$k-5$	$7\,A\,8\ldots3\,4\,5$	$k-2$	6
	$k-4$	$6\,A\,6\ldots2\,3\,4$	$k-2$	5
	$k-3$	$5\,6\,7\ldots2\,3\,A$	$k-2$	4
	$k-2$	$4\,5\,6\ldots1\,2\,3$	$k-2$	A
	$k-1$	$A\,4\,5\,6\ldots1\,2\,3$	$k-1$	-

Table 1: The β-blocks.

The length of a string in T_n^k is $\sum_{j=1}^{k} |\tau_j| + \sum_{j=1}^{k-1} |\beta_j| = k(n-k+1) + 2(k-1) + 1(k-3) + (k-1-3)(k-2) = k(n-2)+3$. When counting the length of $\beta_1\beta_2\ldots\beta_k$, we note that for k even, there are 2 blocks of length $k-1$, 1 block of length $k-3$, and the rest of the blocks has length $k-2$. For k odd the same counting holds by moving the letter A from the $\tau_{\ell+1}$-block to the β_ℓ-block.

We define recursively the following sets of strings:

$$S_1 := P_{\Sigma_{k,n}}\beta_1,$$

$$S_{j+1} := \begin{cases} S_j\, P_{\Sigma_{A,n}}\beta_{j+1} & \text{for } k \text{ odd and } j = \lceil \frac{k}{2} \rceil, \\ S_j\, P_{\Sigma_{k,n}}\beta_{j+1} & \text{otherwise,} \end{cases}$$

with $1 \le j < k-1$. Note that any string in S_j is a prefix of a string in T_n^k, and that $S_{k-1}P_{\Sigma_{k,n}} = T_n^k$.

The first point of the following lemma is key in establishing the theorem. However, in order to prove it, its statement needs to be strengthened.

For a string γ, we let γ^{-i} be γ without the last i letters.

Lemma 1. *The following statements hold:*

(1) *For any $1 \le j \le k-1$, any $\gamma \in S_j$, γ is $L^j_{\Sigma_n}$-complete.*

(2) *For any $1 \le j \le \lfloor \frac{k}{2} \rfloor - 1$, any $\gamma \in S_j$, γ^{-2} is $L^j_{\Sigma_n - \{k-j-1, A\}}$-complete.*

Proof. We will often use the following property: if $\gamma = a_m a_{m-1} \ldots a_1 \in \Sigma^*$ is L^j_Σ-complete, then γ^{-i} is $L^j_{\Sigma - \{a_1,\ldots,a_i\}}$-complete, where $i < m$ and $j \le |\Sigma| - i$. In what follows, γ_j denotes an arbitrary element of S_j.

We proceed by induction on j.

Base cases. Clearly, γ_1 is $L^1_{\Sigma_n}$-complete and γ_1^{-2} is $L^1_{\Sigma_n-\{k-2,A\}}$-complete. We also easily check that γ_2 is $L^2_{\Sigma_n}$-complete and γ_2^{-2} is $L^2_{\Sigma_n-\{k-3,A\}}$-complete.

Inductive case. Suppose now that $j \geq 3$. We let $\gamma_j = \gamma_{j-1}\tau_j\beta_j$ with $\gamma_{j-1} \in S_{j-1}$ and $\tau_j \in P_{\Sigma_{k,n}}$, and $\gamma_{j-1} = \gamma_{j-2}\tau_{j-1}\beta_{j-1}$, with $\gamma_{j-2} \in S_{j-2}$ and $\tau_{j-1} \in P_{\Sigma_{k,n}}$. By induction hypothesis, γ_{j-1} is $L^{j-1}_{\Sigma_n}$-complete, γ_{j-2} is $L^{j-2}_{\Sigma_n}$-complete, and, if $j \leq \lfloor \frac{k}{2} \rfloor - 1$ then γ_{j-1}^{-2} is $L^{j-1}_{\Sigma_n-\{k-j,A\}}$-complete and γ_{j-2}^{-2} is $L^{j-2}_{\Sigma_n-\{k-j+1,A\}}$-complete.

We consider the two statements in turn.

(1) Let $\sigma = s_1 \dots s_j \in L^j_{\Sigma_n}$.
Suppose first that s_j occurs in $\tau_j\beta_j$. As γ_{j-1} is $L^{j-1}_{\Sigma_n}$-complete, γ_{j-1} contains $s_1 \dots s_{j-1}$. Thus γ_j contains σ.
Suppose now that s_j does not occur in $\tau_j\beta_j$. Then $j \leq k-2$. Note that for k odd and $j = \lceil \frac{k}{2} \rceil$, the letter A is missing from β_j, but it is not missing from $\tau_j\beta_j$. Depending on the value of j, we can have one of the following cases:
   (a) $s_j = k-j+1$, for $3 \leq j \leq k-3$. We have that the last letter of γ_{j-1} is $(k-j+1)$. Then γ_{j-1}^{-1} is $L^{j-1}_{\Sigma_n-\{k-j+1\}}$-complete. Hence $s_1 \dots s_{j-1}$ is contained in γ_{j-1}^{-1}, and s_j equals the last letter of γ_{j-1}. Thus, σ is contained in γ_{j-1}, and moreover in γ_j.
   (b) $s_j = A$, for $j = k-2$. As γ_{j-1} is $L^{j-1}_{\Sigma_n}$-complete and the last letter of γ_{j-1} is A, we have that γ_{j-1}^{-1} is $L^{j-1}_{\Sigma_n-\{A\}}$-complete. Hence σ is contained in γ_{j-1}.
   (c) $s_j = A$, for $j = \frac{k}{2}$, k even, and for $j = \lfloor \frac{k}{2} \rfloor$, k odd. Note that $j \geq 5$ and that s_j equals the second last letter of β_{j-1}. We can have:
      (i) $s_{j-1} \notin \{k-j+1, k-j+2\}$. That is, s_{j-1} is neither the missing letter from β_{j-1}, nor the letter after A in β_{j-1} (i.e. the last letter of β_{j-1}). Then s_{j-1} is contained in $\tau_{j-1}\beta_{j-1}^{-2}$. As γ_{j-2} is $L^{j-2}_{\Sigma_n}$-complete, $s_1 \dots s_{j-2}$ is contained in γ_{j-2}. Hence σ is contained in γ_{j-1}.
      (ii) $s_{j-1} = k-j+2$ (the missing letter from β_{j-1}). As γ_{j-2} is $L^{j-2}_{\Sigma_n}$-complete and the last letter of β_{j-2} is $(k-j+2)$, we have that γ_{j-2}^{-1} is $L^{j-2}_{\Sigma_n-\{k-j+2\}}$-complete, hence also $L^{j-2}_{\Sigma_n-\{k-j+2,A\}}$-complete. Hence $s_1 \dots s_{j-2}$ is contained in γ_{j-2}^{-1}, and s_{j-1} equals the last letter of γ_{j-2}. Thus σ is contained in γ_{j-1}.
      (iii) $s_{j-1} = k-j+1$ (the last letter of β_{j-1}). We distinguish two cases:
         – For $j-2 \geq 4$, as γ_{j-2}^{-2} is $L^{j-2}_{\Sigma_n-\{k-j+1,A\}}$-complete and the last letter of γ_{j-2}^{-2} is $(k-j+1)$, it follows that γ_{j-2}^{-3} is $L^{j-2}_{\Sigma_n-\{k-j+1,A\}}$-complete. Thus $s_1 \dots s_{j-2}$ is contained in γ_{j-2}^{-3} and s_{j-1} equals the third last letter of γ_{j-2}. Hence σ is contained in γ_{j-2}.

– If $j-2 < 4$ then $j-2 = 3$ (since $j \geq 5$). The last letter of γ_{j-2}^{-1} is $(k-j+1)$. It follows that γ_{j-2}^{-1} is $L^{j-2}_{\Sigma_n - \{k-j+1,A\}}$-complete. Thus $s_1 \dots s_{j-2}$ is contained in γ_{j-2}^{-2} and s_{j-1} equals the second last letter of γ_{j-2}. Hence σ is contained in γ_{j-1}.

(2) Let $\sigma = s_1 \dots s_j \in L^j_{\Sigma_n - \{k-j-1,A\}}$. We can have:

(a) $s_j \notin \{k-j+1, k-j\}$. That is, s_j is neither the missing letter of β_j nor the last letter of β_j. Then s_j is contained in $\tau_j \beta_j^{-2}$. And as γ_{j-1} is $L^{j-1}_{\Sigma_n}$-complete, it follows that σ is contained in γ_j^{-2}.

(b) $s_j = k-j+1$. As γ_{j-1} is $L^{j-1}_{\Sigma_n}$-complete and $(k-j+1)$ is the last letter of γ_{j-1} we have that γ_{j-1}^{-1} is $L^{j-1}_{\Sigma_n - \{k-j+1\}}$-complete. It follows that γ_{j-1}^{-1} contains $s_1 \dots s_{j-1}$, and thus σ is contained in γ_{j-1}, thus in γ_j^{-2}.

(c) $s_j = k-j$. As γ_{j-1}^{-2} is $L^{j-1}_{\Sigma_n - \{k-j,A\}}$-complete, we have that γ_{j-1}^{-2} contains $s_1 \dots s_{j-1}$. For $j-1 \leq 3$, $(k-j)$ is the second last letter of γ_{j-1}, hence γ_{j-1}^{-1} contains σ. For $j-1 > 3$, $(k-j)$ is the third last letter of γ_{j-1}, hence γ_{j-1}^{-3} is $L^{j-1}_{\Sigma_n - \{k-j,A\}}$-complete, thus γ_{j-1}^{-2} contains σ.

□

The following lemma implies that the reverse of any string in T_n^k is isomorphic with a string in T_n^k. Let us denote by $\overline{\omega}$ the reverse of a string ω. Also, for $\gamma = \tau_1 \beta_1 \tau_2 \dots \tau_{k-1} \beta_{k-1} \tau_k \in T_n^k$, let $\gamma_j := \beta_j \tau_{j+1} \dots \beta_{k-1} \tau_k$ for any $1 \leq j \leq k-1$.

Let g be the bijection on Σ_n given by $g(a) := 1 + (k+1-a) \mod (k-2)$ for $1 \leq a \leq k-2$, and $g(a) := a$ for $k-1 \leq a \leq n$.

Lemma 2. *For any $\gamma \in T_n^k$, for any $1 \leq j \leq k-1$, we have $\overline{\gamma_j} \in g(S_{k-j})$.*

Proof. Clearly, $\overline{\tau} \in P_{\Sigma_{k,n}}$ for any $\tau \in P_{\Sigma_{k,n}}$.

We have $\overline{\alpha_j} = g(\alpha_{k-j})$ for any $1 \leq j \leq k-1$. Hence also $\overline{\beta_j} = g(\beta_{k-j})$ for any $1 \leq j \leq k-1$. It follows that $\overline{\gamma_j} = \overline{\beta_j \tau_{j+1} \dots \beta_{k-1} \tau_k} = \overline{\tau_k} \overline{\beta_{k-1}} \dots \overline{\tau_{j+1}} \overline{\beta_j} = \overline{\tau_k}\, g(\beta_1) \dots \overline{\tau_{j+1}}\, g(\beta_{k-j}) = g(\overline{\tau_k} \beta_1 \dots \overline{\tau_{j+1}} \beta_{k-j})$. As $\overline{\tau_k} \beta_1 \dots \overline{\tau_{j+1}} \beta_{k-j} \in S_{k-j}$, we have $\overline{\gamma_j} \in g(S_{k-j})$. □

We can now conclude the proof of Theorem 1.

Let γ be an arbitrary string in T_n^k and $\sigma = s_1 \dots s_k$ be an arbitrary string in L_n^k. Let j be such that $s_j \in \Sigma_{k,n}$, with $1 \leq j \leq k$. There is such a j by the pigeonhole principle, as $|\Sigma_n - \Sigma_{k,n}| = k-1$ and the k letters of σ are distinct. Let $\gamma' = \tau_1 \beta_1 \tau_2 \dots \tau_{j-1} \beta_{j-1}$ and $\gamma'' = \beta_j \tau_{j+1} \beta_{j+1} \dots \beta_{k-1} \tau_k$. Then $\gamma = \gamma' \tau_j \gamma''$. (If $j = 1$ then $\gamma' = \varepsilon$ and $\gamma = \tau_1 \gamma''$, while if $j = k$ then $\gamma'' = \varepsilon$ and $\gamma = \gamma' \tau_k$, where ε is the empty string.)

As $\gamma' \in S_{j-1}$, by Lemma 1(1), we have that γ' contains $s_1 \dots s_{j-1}$. By Lemma 2, we have that $\overline{\gamma''} \in g(S_{k-j})$. Thus, again by Lemma 1(1), we have that $g^{-1}(\overline{\gamma''})$ contains $g^{-1}(\overline{s_k \dots s_{j+1}})$, and hence γ'' contains $s_{j+1} \dots s_k$. Clearly, τ_j contains s_j. Putting the three pieces together, we obtain that γ contains σ.

6. Conclusions

In this paper we have built a set T_n^k of strings, each string being of length $k(n-2)+3$ and containing all permutations of order k over an alphabet of size n, for $n \geq k \geq 10$. We thus improve by one unit the previous known upper bound on the length of the shortest such strings, which was $k(n-2)+4$. It remains open if further improvements are possible. We also do not know whether the set T_n^k is complete, that is, whether there exist other such strings, not isomorphic with the ones in T_n^k.

In the particular case when $k = n$, our construction shows that $f(n) \leq n^2-2n+3$, for $n \geq 10$, where $f(n)$ denotes the length of the shortest strings that contain all permutations of order n (over an alphabet of size n). As M. Newey [7] proved that $f(n) = n^2-2n+4$, for $3 \leq n \leq 7$, it is compelling to find the values of $f(8)$ and $f(9)$.

Acknowledgments. Mohammad Torabi Dashti introduced us to the problem. Lăcrămioara Aştefănoaei made useful comments. We warmly thank them both. We also thank the anonymous referees for their helpful suggestions.

[1] Leonard Adleman. Short permutation strings. *Discrete Mathematics*, 10:197–200, 1974.

[2] Open Garden Forum (`http://garden.irmacs.sfu.ca/?q=op/smallest_universal_supersequence`), 2009. Retrieved on March 22nd, 2011.

[3] Vašek Chvátal, David A. Klarner, and Donald E. Knuth. Selected combinatorial research problems. Technical Report 292, Stanford University, 1972.

[4] Giulia Galbiati and Franco P. Preparata. On permutation embedding sequences. *SIAM Journal of Applied Mathematics*, 30(3):421–423, 1976.

[5] P. J. Koutas and T. C. Hu. Shortest string containing all permutations. *Discrete Mathematics*, 11:125–132, 1975.

[6] S. P. Mohanty. Shortest string containing all permutations. *Discrete Mathematics*, 31:91–95, 1980.

[7] Malcolm Newey. Notes on a problem involving permutations as subsequences. Technical Report 340, Stanford University, 1973.

[8] Carla Savage. Short strings containing all k-element permutations. *Discrete Mathematics*, 42:281–285, 1982.