Matemática Contemporânea, Vol 39, 77-84
https://doi.org/10.21711/231766362010/rmc399

Upper bounds and exact values on transposition distance of permutations *

Luís Felipe Ignácio Cunha **Luis Antonio B. Kowada**

Abstract

One of the main operations of genome rearrangement is the transposition (exchange of contiguous blocks). Recently, the problem of computing the minimum amount of operations of transpositions needed to transform one sequence into another (transposition distance) between two permutations was proved to be NP-hard [3]. The exact distance is known for few cases. We show how to sort lonely permutations of the family $u_{n,3}$ applying $\lfloor \frac{n}{2} \rfloor + 1$ transpositions. Thus, if 4 divides $n+1$ then the transposition distance is $d_t(u_{n,3}) = \lfloor \frac{n}{2} \rfloor + 1$, and if 4 does not divide $n+1$, we have that $\lfloor \frac{n}{2} \rfloor \leq d_t(u_{n,3}) \leq \lfloor \frac{n}{2} \rfloor + 1$.

1 Introduction

The transposition is one of the main operations of genome rearrangement that can be viewed as the swap of contiguous blocks. A possible biological explanation for this rearrangement is the duplication of a block of genes, followed by the deletion of the original block [2, 10].

Permutations can be joined into equivalence classes where for all elements in the same class the transposition distance are equal. However some classes have just one permutation, so they are called *unitary toric classes* and correspond to the *lonely permutations*, for more details see in [5, 7].

The problem of determining the transposition distance of permutations is NP-hard [3] and there are no tight bounds for the distance. We combine

*2000 *AMS Subject Classification.* 68R10, 05C75 05C85.
Key Words and Phrases. Genome rearrangement, transposition distance, lonely permutations.
*This research was supported by FAPERJ

the reality and desire diagram [1] and the unitary toric classes to obtain tight bounds for the transposition distance of lonely permutations.

2 Preliminaries

Definition 2.1. [1] *A* transposition, *denoted by* $t(i,j,k)$, *where* $1 \leq i < j < k \leq n+1$, *is defined as the permutation*

$$t(i,j,k) := [1\,2\ldots i{-}1\; j\; j{+}1 \ldots k{-}1\; i\; i{+}1 \ldots j{-}1\; k \ldots n].$$

The transposition $t(i,j,k)$ "cuts" the elements between the positions j and $k-1$ (both inclusive) and "pastes" them immediately before the i-th position. The permutation π is a bijective function of n elements into these n elements given by $\pi = [\pi_1\pi_2\ldots\pi_{i-1}\pi_i\ldots\pi_{j-1}\boxed{\pi_j\ldots\pi_{k-1}}\pi_k\ldots\pi_n]$, then

$$\pi \cdot t(i,j,k) = [\pi_1\pi_2\ldots\pi_{i-1}\boxed{\pi_j\ldots\pi_{k-1}}\pi_i\ldots\pi_{j-1}\pi_k\ldots\pi_n].$$

Definition 2.2. [1, 8] *The* transposition distance $d_t(\pi)$ *of a permutation* π *is the length* q *of the shortest sequence of transpositions* t_1, t_2, $\ldots, t_q$, *such that* $\pi t_1 t_2 \ldots t_q = [1\,2\,\ldots\,n]$. *If* $\pi = [1\,2\,\ldots\,n]$, *then* $d_t(\pi) = 0$.

In the study of determining the transposition distance, it is useful to give special names and symbols for some permutations. These are: the *identity permutation of* n *elements*, $\iota_{[n]} := [1\,2\,\ldots\,n]$; the *reverse permutation of* n *elements*, $\rho_{[n]} := [n\,n{-}1\,\ldots\,2\,1]$; and the *lonely permutation of* n *elements*, $u_{n,\ell} := [\overline{\ell}\,\overline{2\ell}\,\overline{3\ell}\,\ldots\,\overline{n\ell}]$, such that $\gcd(n+1,\ell) = 1$ and $\overline{x}$ denotes the remainder of the division of x by $n+1$.

Remark that $\iota_{[n]} = u_{n,1}$ and $\rho_{[n]} = u_{n,n}$.

Definition 2.3. [1, 8] *Given a permutation* π *of* n *elements, the* Reality and Desire Diagram *is a graph on the following set of vertices:*

$$V(RD(\pi)) := \{0, -1, +1, -2, +2, \cdots, -n, +n, -(n+1)\}.$$

and whose set of edges is partitioned into two sets R *and* D, *respectively reality and desire edges, defined as*

$$\begin{aligned} R :=\ & \{(+\pi_i,\ -\pi_{i+1}\ |i = 1, \cdots, n-1\} \cup \{(0,\ -\pi_1),\ (+\pi_n,\ -(n+1))\}, \\ D :=\ & \{(+i,\ -(i+1)\ |i = 1, \cdots, n-1\} \cup \{(0,\ -1),\ (+n,\ -(n+1))\}. \end{aligned}$$

We say that a cycle in a permutation π *has length* k *if it has exactly* k *reality edges (or* k *desire edges) in a cycle of* $RD(\pi)$.

Definition 2.4. [4, 8] *The* reduced permutation *of π, $gl(\pi)$, is the permutation whose reality and desire diagram $RD(gl(\pi))$ is equal to $RD(\pi)$ without the cycles of length 1, keeping the order of the elements.*

Theorem 2.5. [1] *A lower bound of transpositions distance is*

$$d_t(\pi) \geq \left\lceil \frac{(n+1) - c_{odd}(\pi)}{2} \right\rceil,$$

where $c_{odd}(\pi)$ is the number of odd cycles in π.

Lemma 2.6. [8] *Given a lonely permutation $u_{n,\ell}$ with $\ell > 1$, the number of cycles in $RD(u_{n,\ell})$ is $gcd(n+1, \ell-1)$.*

Corollary 2.7. [8] *Given a lonely permutation $u_{n,\ell}$ with $\ell > 1$, the length of each cycle in $RD(u_{n,\ell})$ is $k = \frac{(n+1)}{gcd(n+1,\ell-1)}$.*

Proposition 2.8. [7] *If $\overline{\ell^{-1}}$ is the multiplicative inverse of ℓ modulo $(n+1)$, then $d_t(u_{n,\ell}) = d_t(u_{n,\overline{\ell^{-1}}})$.*

Proposition 2.9. *If $n = 3q$, then $2q+1 \equiv 3^{-1} \pmod{n+1}$ and if $n = 3q+1$, then $q+1 \equiv 3^{-1} \pmod{n+1}$.*

Proof. The hypothesis $n = 3q$ implies the computation of the inverse as follows:

$$3(2q+1) = 2(3q+1) + 1 \equiv 1 \pmod{n+1}.$$

Now, the hypothesis $n = 3q+1$ implies:

$$3(q+1) = 3q+1+2 = n+2 \equiv 1 \pmod{n+1}.$$

□

Definition 2.10. [6] *The* concatenation *of the sequence a with the sequence b, denoted by $a \odot b$, is the operation that joins both sequences. The* generalized concatenation *is denoted by $\bigodot_{i=j}^{z} f(i)$ and is defined by the concatenation of the sequences $f(i)$, with i ranging from j to z.*

Example 1. *The lonely permutation $u_{n,\ell}$ can be described as the following concatenation:*

$$u_{n,\ell} = \bigodot_{i=1}^{n} [\overline{i\ell}] = [\overline{\ell}\, \overline{2\ell}\, \overline{3\ell} \ldots \overline{n\ell}].$$

3 Upper bounds and exact values on transposition distance of lonely permutation $u_{n,3}$

Theorem 3.1 shows a tight bound on the transposition distance for a family of lonely permutations.

Theorem 3.1. *In a lonely permutation $u_{n,3}$ if 4 divides $n+1$, $d_t(u_{n,3}) = \lfloor \frac{n}{2} \rfloor + 1$ and, if 4 does not divide $n+1$, $\lfloor \frac{n}{2} \rfloor \leq d_t(u_{n,3}) \leq \lfloor \frac{n}{2} \rfloor + 1$.*

Proof. Let $n = 3q$. So $d_t(u_{n,3}) = d_t(u_{n,2q+1})$. Let $\pi(0) = u_{n,2q+1}$, thus

$$\pi(0) = \bigodot_{i=0}^{q-1} [(2q+1+i)\ (q+1+i)\ (i+1)].$$

The transpositions to apply are $t_i(i, 3i, 3i+2)$ with $i = 1, 2, \cdots, q-1$. After k transpositions, where $1 \leq k \leq q-1$, the permutation is:

$$\pi(k) = \bigodot_{i=1}^{k}[i] \odot \bigodot_{i=0}^{k}[2q+k+1-i] \odot \bigodot_{i=1}^{k+1}[q+i] \odot [k+1] \odot \bigodot_{i=k+1}^{q-1} [(2q+1+i)\ (q+1+i)\ (i+1)].$$

The proof of the last expression, $\pi(k)$, is by induction. When $k = 1$, observe that the transposition $t_1(1,3,5)$ over $\pi(0)$ moves the elements 1 and $(2q+2)$ to the beginning of the permutation, thus:

$$\pi(1) = \bigodot_{i=1}^{1}[i] \odot \bigodot_{i=0}^{1}[2q+2-i] \odot \bigodot_{i=1}^{2}[q+i] \odot [2] \odot \bigodot_{i=2}^{q-1}[(2q+1+i)\ (q+1+i)\ (i+1)].$$

Suppose after $k-1$ transpositions that the permutation is:

$$\pi(k-1) = \bigodot_{i=1}^{k-1}[i] \odot \bigodot_{i=0}^{k-1}[2q+k-i] \odot \bigodot_{i=1}^{k}[q+i] \odot [k] \odot \bigodot_{i=k}^{q-1}[(2q+1+i)\ (q+1+i)\ (i+1)].$$

Now applying the transposition $t_k(k, 3k, 3k+2)$ over $\pi(k-1)$, observe that the elements k and $(2q+k+1)$ moves into the elements $(k-1)$ and $(2q+k)$:

$$\pi(k) = \bigodot_{i=1}^{k-1}[i] \odot [k\ (2q+k+1)] \odot \bigodot_{i=0}^{k-1}[2q+k-i] \odot \bigodot_{i=1}^{k}[q+i] \odot [k+1]$$
$$\odot \bigodot_{i=k+1}^{q-1} [(2q+1+i)\ (q+1+i)\ (i+1)].$$

And this permutation can be rewritten as:

$$\pi(k) = \bigodot_{i=1}^{k}[i] \odot \bigodot_{i=0}^{k}[2q+k+1-i] \odot \bigodot_{i=1}^{k+1}[q+i] \odot [k+1] \odot \bigodot_{i=k+1}^{q-1}[(2q+1+i)\ (q+1+i)\ (i+1)].$$

Remark that as we said $1 \le k \le q-1$, these transpositions are valid while the last generalized concatenation exists, i.e, if $k > q-1$ there is not the concatenation ranging from $k+1$ to $q-1$. Thus, after $q-1$ transpositions:

$$\pi(q-1) = \bigodot_{i=1}^{q-1}[i] \odot \bigodot_{i=0}^{q-1}[3q-i] \odot \bigodot_{i=1}^{q}[q+i] \odot [q].$$

The $(q-1)$ initial elements are ordered the way we want as in the identity, the elements from $2q$ to $(3q-1)$ are in the reversal way and from $(q+1)$ to $2q$ are ordered equal the identity but not in the position we intend.

Now, observe that $\rho_{[q+2]} = gl(\pi(q-1))$. As $d_t(\rho_{[q+2]}) = \left\lfloor \frac{q+2}{2} \right\rfloor + 1$ [9], in total we applied $(q-1) + \left\lfloor \frac{q+2}{2} \right\rfloor + 1 = \left\lfloor \frac{n}{2} \right\rfloor + 1$ transpositions.

Let $n = 3q+1$. So $d_t(u_{n,3}) = d_t(u_{n,q+1})$. Let $\pi(0) = u_{n,q+1}$, thus

$$\pi(0) = \bigodot_{i=0}^{q-1}[(q+1+i)\ (2q+2+i)\ (i+1)] \odot [(2q+1)].$$

The transpositions to apply are $t_i(i, 3i, 3i+2)$ with $i = 1, 2, \cdots, q$. After k transpositions, where $1 \le k \le q$, the permutation is:

$$\pi(k) = \bigodot_{i=1}^{k}[i] \odot \bigodot_{i=0}^{k}[q+k+1-i] \odot \bigodot_{i=2}^{k+2}[2q+i] \odot [k+1]$$
$$\odot \bigodot_{i=k+1}^{q-1}[(q+1+i)\ (2q+2+i)\ (i+1)] \odot [2q+1].$$

As in the preview case, the proof of the last expression is by induction. When $k = 1$, observe that the transposition $t_1(1,3,5)$ over $\pi(0)$ moves the elements 1 and $(q+2)$ to the beginning of the permutation, thus:

$$\pi(1) = \bigodot_{i=1}^{1}[i] \odot \bigodot_{i=0}^{1}[q+2-i] \odot \bigodot_{i=2}^{3}[2q+i] \odot [2] \odot \bigodot_{i=2}^{q-1}[(q+1+i)\ (2q+2+i)\ (i+1)] \odot [2q+1].$$

Suppose after $k-1$ transpositions that the permutation is:

$$\pi(k-1) = \bigodot_{i=1}^{k-1}[i] \odot \bigodot_{i=0}^{k-1}[q+k-i] \odot \bigodot_{i=2}^{k+1}[2q+i] \odot [k] \odot \bigodot_{i=k}^{q-1}[(q+1+i)\ (2q+2+i)\ (i+1)] \odot [2q+1].$$

Now applying the transposition $t_k(k, 3k, 3k+2)$ over $\pi(k-1)$, observe that the elements k and $(q+k+1)$ moves into the elements $(k-1)$ and $(q+k)$:

$$\pi(k) = \bigodot_{i=1}^{k-1}[i] \odot [k \ (q+k+1)] \odot \bigodot_{i=0}^{k-1}[q+k-i] \odot \bigodot_{i=2}^{k+1}[2q+i] \odot$$
$$[(2q+k+2) \ (k+1)] \odot \bigodot_{i=k+1}^{q-1}[(q+1+i) \ (2q+2+i) \ (i+1)] \odot [2q+1].$$

And this permutation can be rewritten as:

$$\pi(k) = \bigodot_{i=1}^{k}[i] \odot \bigodot_{i=0}^{k}[q+k+1-i] \odot \bigodot_{i=2}^{k+2}[2q+i] \odot [k+1]$$
$$\odot \bigodot_{i=k+1}^{q-1}[(q+1+i) \ (2q+2+i) \ (i+1)] \odot [2q+1].$$

Remark that as we said $1 \leq k \leq q$, these transpositions are valid while the last generalized concatenation exists, i.e, if $k > q$ there is not the concatenation ranging from $k+1$ to $q-1$. Thus, after q transpositions:

$$\pi(q) = \bigodot_{i=1}^{q}[i] \odot \bigodot_{i=-1}^{q-1}[2q-i] \odot \bigodot_{i=2}^{q+1}[2q+i].$$

The q initial elements are ordered the way we want as in the identity, the elements from $(q+1)$ to $(2q+1)$ are in the reversal way and from $(2q+2)$ to $(3q+1)$ are ordered equal the identity but not in the position we intend.

Now, observe that $\rho_{[q+1]} = gl(\pi(q))$. As $d_t(\rho_{[q+1]}) = \left\lfloor \frac{q+1}{2} \right\rfloor + 1$ [9], in total we applied $q + \lfloor \frac{q+1}{2} \rfloor + 1 = \lfloor \frac{n}{2} \rfloor + 1$ transpositions.

Let $n = 3q+2$. In this case does not exist lonely permutation because $gcd(n+1, 3) = 3 \neq 1$.

Until here, we observe the possibility to order $u_{n,3}$ applying $\lfloor \frac{n}{2} \rfloor + 1$ transpositions. Now, we will show when it is the exact value of the transposition distance and when is not.

- Suppose that 4 divides $(n+1)$.

By *corollary 2.7*, the length of each cycle is even, $\frac{n+1}{2}$. So by the *theorem 2.5*, $d_t(\pi) \geq \frac{(n+1)}{2} = \lfloor \frac{n}{2} \rfloor + 1$.

As we ordered applying $\lfloor \frac{n}{2} \rfloor + 1$ transpositions, we proved that the exact transposition distance is

$$d_t(u_{n,3}) = \left\lfloor \frac{n}{2} \right\rfloor + 1.$$

- Suppose that 4 does not divide $(n+1)$.

When n is even: In this case there exists one cycle and this is an odd cycle, by *lemma 2.6* and *corollary 2.7*. So, $d_t(\pi) \geq \left\lceil \frac{(n+1)-c_{odd}(\pi)}{2} \right\rceil = \frac{n}{2}$.
When n is odd: In this case there exist two cycles, $gcd(n+1, 2) = 2$, and these are odd cycles (4 does not divide $(n+1)$). So $d_t(\pi) \geq \left\lceil \frac{(n+1)-c_{odd}(\pi)}{2} \right\rceil = \frac{n-1}{2} = \lfloor \frac{n}{2} \rfloor$.

As we ordered applying $\lfloor \frac{n}{2} \rfloor + 1$ transpositions we conclude:

$$\left\lfloor \frac{n}{2} \right\rfloor \leq d_t(u_{n,3}) \leq \left\lfloor \frac{n}{2} \right\rfloor + 1.$$

□

4 Conclusion and future work

Table 1 shows the known results about the exact values on transposition distance for lonely permutations $u_{n,\ell}$. In these cases, we have the exact transposition distance, besides the $u_{n,\ell^{-1}}$, given by proposition 2.8. In this work, we have shown how to sort lonely permutations $u_{n,3}$, giving the exact transposition distance for some cases and upper bounds for other cases. A table in [8] is given with the exact values of transposition distance for $n, \ell \leq 18$. The present work agrees with those exact values, when 4 divides $n+1$. For the case 4 does not divide $n+1$, the present work agrees with some exact values and gives the upper bound of $\lfloor \frac{n}{2} \rfloor + 1$ for the other cases. We intend to investigate other families of lonely permutations, in order to obtain additional exact values for the transposition distance.

n	ℓ	$d_t(u_{n,\ell})$	Ref.	Year	n	ℓ	$d_t(u_{n,\ell})$	Ref.	Year
q	q	$\lfloor \frac{q}{2} \rfloor + 1$	[9]	1997	$4q-1$	$2q+1$	$2q$	[4]	1999
$6q$	$2q+1$	$3q+1$	[4]	1999	$2q$	2	q	[8]	2010
$6q-2$	$4q$	$3q$	[4]	1999	$4q-1$, $q>1$	3	$2q+1$	*	2011

Table 1: Exact values known for the transposition distance of lonely permutations. The $*$ means contribution of the present work.

References

[1] BAFNA, V.; PEVZNER, P.A.: Sorting by transpositions. SIAM J. Disc. Math. 11(2) (1998) 224–240.

[2] BOORE, J.L.: The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals. In Sankoff, D., Nadeau, J.H., eds.: Comparative Genomics. Kluwer Academic Publishers (2000) 133–148.

[3] BULTEAU, L.; FERTIN, G.; RUSU, I.: Sorting by Transpositions is Difficult arXiv/1011.1157 (2010).

[4] CHRISTIE, D.A.: Genome Rerrangement Problems. PhD thesis, University of Glasgow (1999).

[5] ERIKSSON,H.; ERIKSSON, K.; KARLANDER, J., SVENSSON, L.; WÄSTLUND, J.: Sorting a bridge hand. Discrete Mathematics. 241(1) (2001) 289–300.

[6] FORTUNA, V. J.: Distâncias de Transposição entre Genomas. Masther Thesis, Instituto de Computação - Unicamp, 2005.

[7] HAUSEN, R.; FARIA,L.; FIGUEIREDO, C.M.; KOWADA,L.A.: Unitary toric classes, the reality and desire diagram, and sorting by transpositions. SIAM J. Disc. Math. 24(3) (2010) 792–807.

[8] KOWADA, L. A.; HAUSEN, R.; FIGUEIREDO, C. M.: Bounds on the transposition distance for lonely permutations. BSB, Brazilian Symposium on Bioinformatics 6268 (2010) 35–46.

[9] MEIDANIS, J.; WALTER, M.E.M.T.; DIAS,Z.: Transposition distance between a permutations and its reverse. Proc. Fourth South Am. Workshop String Processing, (1997), 70–79.

[10] SANKOFF, D.; LEDUC, G.; ANTOINE, N.; PAQUIN, B.; LANG, B.F.; CEDERGREN, R.: Gene order comparisons for phylogenetic inference: evolution of the mitochondrial genome. Proceedings of the National Academy of Sciences of the United States of America 89(14) (1992) 6575–6579.

Instituto de Matemática
Universidade Federal Fluminense
S/N Mário Santos Braga Street.
CEP: 24.020-140, Valonguinho, Niterói, Brazil.
lfignaciocunha@id.uff.br, luis@vm.uff.br