Interchanging Graphs Associated with Sorting by Transpositions

Pinglong You[a], Wai Chee Shiu[b,*], Wai Hong Chan[b,*], An Chang[c,†]

[a]School of Computer Science and Technology, Fujian Agriculture and Forestry University, Fuzhou, Fujian, P.R. China.

[b]Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, P.R. China.

[c]Software College/Center of Discrete Mathematics, Fuzhou University, Fuzhou, Fujian, P.R. China.

Abstract. Sorting a permutation by transpositions is one of important methods of sequence comparison in computational molecular biology for deriving evolutionary and functional relationships between genes. In the paper, we first associate the problem of sorting by transpositions by a new kind of interchanging graph, and then give such graph a nice characterization. Based on this characterization, we also give the connectivity of the interchanging graphs.

Keywords: permutation; transposition; interchanging graph; blow up; connectivity.

AMS Subject classifications: 05C75; 05E99

1 Introduction

Many combinatorial problems on the rearrangements of permutations have been studied intensively in last decade due to that they are important methods to analyze genome rearrangements in molecular biology([1]-[8]). Among them, sorting by reversals and sorting by transpositions are two fundamental types of rearrangement events occurring in DNA. It would

*The work is partially supported by the Faculty Research Grant, Hong Kong Baptist University.

†The work was supported by the National Natural Science Foundation of China (No. 10431020) and NSFFJ.

be desirable with an efficient algorithm for finding the minimal number of such events between two genomes. Caprara([6]) has shown that sorting by reversals is NP-hard. It is seems that sorting by transpositions is less well understood than sorting by reversals, and in particular, the complexity of sorting by transpositions still remains open. However, Bafna and Pevzner([3]) have designed a 3/2-approximation algorithm for the problem.

Our motivation on this problem stem a game with cards. Suppose that you have five cards numbered from 1 to 5 in decreasing order. The game requires that you should find the minimum step number of reversing the order of five cards in the way of exactly two adjacent blocks (here block means several adjacent cards) interchanged at one step. Of course, you can reverse the order of five cards in 4 steps just one by one. However, we can finish the game in 3 steps illustrated as the following Figure 1.

$$5\underline{4}\underline{3}21 \longrightarrow 52\underline{1}\underline{43} \longrightarrow 1\underline{4}\underline{5}23 \longrightarrow 12345$$

Figure 1: Transpositions of sorting 54321.

In fact, you can find more different ways to finish the game. A slight generalization of this small game is known as sorting a bridge hand([1]). Let $\pi = (a_1 a_2 \dots a_n)$ be a permutation of $\{1, 2, 3, \dots, n\}$. For a permutation π, a transposition $T(i, j, k)$ of π is represented by the following procedure.

$$\begin{aligned} T(i,j,k)\pi &= T(i,j,k)(a_1 a_2 \dots a_{i-1} \boxed{a_i \dots a_j}\,\boxed{a_{j+1} \dots a_k}\, a_{k+1} \dots a_n) \\ &= (a_1 a_2 \dots a_{i-1} \boxed{a_{j+1} \dots a_k}\,\boxed{a_i \dots a_j}\, a_{k+1} \dots a_n) \end{aligned}$$

Obviously, $T(i, j, k)$ has the effect on the permutation π such that two blocks $\boxed{a_i \dots a_j}$ and $\boxed{a_{j+1} \dots a_k}$ exchanged in π, where $1 \le i \le j < k \le n$. Moreover, transpositions on a permutation π generate the symmetric group $\mathcal{S}_n$. Let π, σ be two permutations in $\mathcal{S}_n$. The transposition distance between π and σ is the minimum number s of transpositions $T_1, T_2, \dots, T_s$ such that $T_s \dots T_2 T_1 \pi = \sigma$. Sorting a permutation π by transpositions is the problem of finding transposition distance $d(\pi)$ between π and ι, where ι is the identity permutation. Let $d(n) = \max_{\pi \in \mathcal{S}_n} \{d(\pi)\}$, i.e., the number of transpositions needed, in the worst case, to sort a permutation of length n. Bafna and Pevzner([3]) proved that $\lceil (n-1)/2 \rceil \le d(n) \le \lfloor (3n)/4 \rfloor$. H.Eriksson et al.([1]) improved the upper bound to $d(n) \le \lfloor \frac{2n-2}{3} \rfloor$ for $n \ge 9$ and showed that the minimum number of transpositions of sorting the permutation $n(n-1) \dots 321$ was $\lceil \frac{n+1}{2} \rceil$ when $n \ge 3$. H.Eriksson et al.([1]) also determined the exact values of $d(n)$ for $n \le 15$ by computations and theoretical arguments (Table 1). This is the answer to the game at the beginning.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
d(n)	0	1	2	3	3	4	4	5	5	6	6	7	8	8	9

Table 1. Known values of $d(n)$

Since the symmetric group S_n is generated from a permutation by transpositions, one can associate naturally sorting by transpositions with a graph TG_n as the following. The graph TG_n has a vertex set $V(TG_n) = S_n$ and a edge set $E(TG_n) = \{(\pi, \sigma) | \text{there exists a transposition } T \text{ of } \pi \text{ such that } T\pi = \sigma\}$. Every transposition $T(i,j,k)$ has a inverse transposition $T^{-1}(i,j,k) = T(i,r,k)$, where $r = i + k - j - 1$. So the graph TG_n can be regarded as a undirected graph. Moreover, it is clear that TG_n is a Cayley graph([1]). When $n = 2,3$, TG_2 and TG_3(TG_1 is a trivial graph) are displayed in Figure 2, respectively.

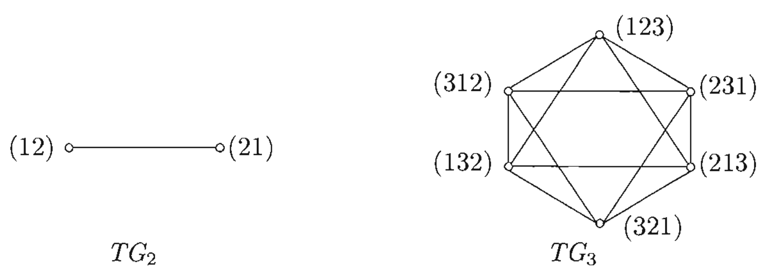

Figure 2: TG_2 and TG_3.

In the paper, we will study some properties of the graph TG_n in more details. In particular, a structural characterization of TG_n will be given, and by this characterization, the connectivity of interchanging graphs is obtained. For some undefined terminology and notations in the paper we refer to [9, 10].

2 Main Results

We will start with some lemmas.

Lemma 2.1 *The graph TG_n is vertex-transitive and $\binom{n+1}{3}$-regular.*

Proof Since all Cayley graphs are vertex-transitive, so is the graph TG_n. Thus TG_n is regular. Let $\pi \in V(TG_n)$. The set of neighbours of π,

$N(\pi)$ can be partitioned into two subsets, i.e. $N(\pi) = \{\sigma \in V(TG_n) | \sigma = T(i,j,k)\pi \text{ and } i < j < k\} \cup \{\sigma \in V(TG_n) | \sigma = T(i,j,k)\pi \text{ and } i = j < k\}$. Hence, $|N(\pi)| = \binom{n}{3} + \binom{n}{2} = \binom{n+1}{3}$. □

The subgraph of TG_n induced by the vertex set $\{(a_i a_{i+1} \ldots a_n a_1 a_2 \ldots a_{i-1}) \mid 1 \le i \le n\}$ in TG_n is denoted by $H(a_1 a_2 \ldots a_n)$. It is clear that $|V(H(a_1 a_2 \ldots a_n))| = n$.

Lemma 2.2 *The graph $H(a_1 a_2 \ldots a_n)$ is a complete graph.*

Proof Let $\pi_1 = (a_m a_{m+1} \ldots a_n a_1 \ldots a_{m-1})$, $\pi_2 = (a_l a_{l+1} \ldots a_n a_1 \ldots a_{l-1}) \in V(H)$ $(m < l)$ be two vertices of $H(a_1 a_2 \ldots a_n)$. Obviously, there exits a transposition $T(1, l-m, n)$ such that

$$T(1, l-m, n)\pi_1 = \pi_2$$

Thus we have $(\pi_1, \pi_2) \in E(H(a_1 a_2 \ldots a_n))$. So the result follows. □

If $\pi = (a_1 a_2 \ldots a_{n-1}) \in V(TG_{n-1})$ is a vertex of the graph TG_{n-1}, we denote the vertex $(a_1 a_2 \ldots a_{n-1} n) \in V(TG_n)$ by πn. And if the first element in the sequence of a vertex in $V(H(\pi n))$ is x , then we denote the vertex by $\pi_{(x)}$.

Lemma 2.3

$$V(TG_n) = \bigcup_{\pi \in V(TG_{n-1})} V(H(\pi n)).$$

Proof It is clear that elements in each $V(H(\pi n))$ are permutations of length n. Conversely, let $\pi' = b_1 b_2 \ldots b_i n b_{i+1} \ldots b_{n-1} \in V(TG_n)$. Then $\pi' \in V(H(b_{i+1} \ldots b_{n-1} b_1 b_2 \ldots b_i n))$. □

Remark: Since

$$\left| \bigcup_{\pi \in V(TG_{n-1})} V(H(\pi n)) \right| = |V(TG_n)| = n! = (n-1)!n = \sum_{\pi \in V(TG_{n-1})} |V(H(\pi n))|,$$

we have $V(H(\pi n)) \bigcap V(H(\alpha n)) = \emptyset$, for all $\pi, \alpha \in V(TG_{n-1})$ with $\pi \neq \alpha$.

Lemma 2.4 *Suppose $\pi, \sigma \in V(TG_{n-1})$ and $(\pi, \sigma) \in E(TG_{n-1})$, i.e., there exists a transposition $T(i,j,k)$ such that $T(i,j,k)\pi = \sigma$. Let $\pi = (a_1 a_2 \dots a_{n-1})$ and $E_{(\pi,\sigma)}$ be the set*

$$\{(\pi_{(x)},\ \sigma_{(x)}) | x = 1, 2, \dots, n\} \bigcup \{(\pi_{(a_i)},\ \sigma_{(a_{j+1})}),\ (\pi_{(a_{j+1})},\ \sigma_{(a_{k+1})}),\ (\pi_{(a_{k+1})},\ \sigma_{(a_i)})\}.$$

Then $E_{(\pi,\sigma)} \subseteq E(TG_n)$ (as shown in Figure 3).

Proof Since $e = (\pi, \sigma) \in E(TG_{n-1})$, i.e., there exists a transposition $T(i,j,k)$ such that $\sigma = T(i,j,k)\pi = (a_1 a_2 \dots a_{i-1} a_{j+1} \dots a_k a_i \dots a_j a_{k+1} \dots a_{n-1})$.

1. For any $x \in \{1, 2, \dots, n\}$, there exists m such that $a_m = x$. Take the triple (i,j,k) in the transposition $T(i,j,k)$ as the following way

   $(i_0, j_0, k_0) =$

   $$\begin{cases} (i - m + 1, k - m + i - j, k - m + 1) & \text{for } 1 \le m \le i - 1, \\ (j - m + 2, n - m + j - k + i, n - m + i) & \text{for } i \le m \le j, \\ (k - m + 2, k - m + j - i + 2, n - m + j + 1) & \text{for } j + 1 \le m \le k, \\ (n - m + i + 1, n - m + i + k - j, n - m + k + 1) & \text{for } m \ge k + 1. \end{cases}$$

   It is easy to verify that $T(i_0, j_0, k_0)\sigma_{(a_m)} = \pi_{(a_m)}$, that is, $(\pi_{(x)}, \sigma_{(x)}) \in E(TG_n)$, $x = 1, 2, \dots, n$.

2. Take $(i_0, j_0, k_0) = (1, j - i + 1, n - k + j)$ as the triple (i,j,k) in the transposition $T(i,j,k)$. Then we have

   $$T(i_0, j_0, k_0)\sigma_{(a_i)} = \pi_{(a_{k+1})}$$

   Hence $(\pi_{(a_{k+1})},\ \sigma_{(a_i)}) \in E(TG_n)$.

3. Similar to Case 2, we can show that $(\pi_{(a_i)},\ \sigma_{(a_{j+1})})$, $(\pi_{(a_{j+1})},\ \sigma_{(a_{k+1})}) \in E(TG_n)$ □

Lemma 2.5

$$E(TG_n) = \Big(\bigcup_{\pi \in V(TG_{n-1})} E(H(\pi n)) \Big) \bigcup \Big(\bigcup_{(\pi,\sigma) \in E(TG_{n-1})} E_{(\pi,\sigma)} \Big).$$

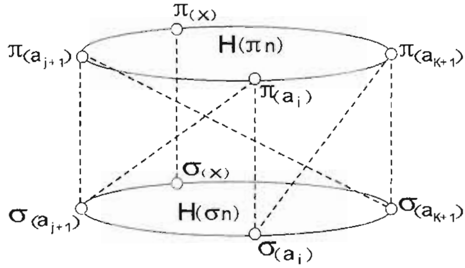

Figure 3: The edges (dotted lines) in $E_{(e)}$.

Proof

$$\begin{aligned} & \sum_{\pi \in V(TG_{n-1})} |E(H(\pi n)| + \sum_{(\pi,\sigma) \in E(TG_{n-1})} |E(\pi, \sigma)| \\ = & (n-1)!\binom{n}{2} + \binom{n}{3}\frac{(n-1)!}{2}(n+3) \\ = & \binom{n+1}{3}\frac{(n)!}{2} \\ = & |E(TG_n)|. \end{aligned}$$

Hence, if $(\pi', \sigma') \in E(TG_n)$, then either

1. $\pi', \sigma' \in V(H(\pi n))$ for some $\pi \in V(TG_{n-1})$, or
2. $\pi' \in V(H(\pi n))$ and $\sigma' \in V(H(\sigma n)$, for some $(\pi, \sigma) \in E(TG_{n-1})$.

$\square$

Then, we can easily find the way to construct TG_n by 'blowing up' TG_{n-1} as follows.

1. Extract each vertex $\pi \in V(TG_{n-1})$ to the complete graph $H(\pi n)$;

2. Add edges to connect the vertices $V(H(\pi n)$ to the "corresponding" vertices in $V(H(\sigma n)$ for all edges $(\pi, \sigma) \in E(TG_{n-1})$.

Example 2.1 *Figure 4 illustrates that how TG_{n-1} is blown up to TG_n ($n = 2, 3$) by the method mentioned above.*

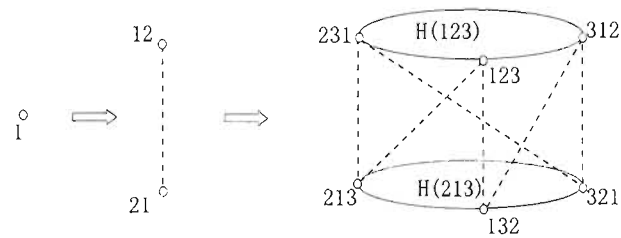

Figure 4: TG_{n-1} is blown up to TG_n with $n = 2, 3$.

References

[1] H. Eriksson, K. Eriksson, J. Karlander, L. Svensson and J. *Wästlund*, Sorting a bridge hand, Discrete Math., 241 (2001), 289-300.

[2] V. Bafna and P.A. Pevzner, Genome rearrangements and sorting by reversals, SIAM J. Comput. 25 (1996), 272-289.

[3] V. Bafna and P.A. Pevzner, Sorting by transpositions, SIAM J. Discrete Math., 11 (1998), 224-240.

[4] W.H. Gates and C.H. Papadimitriou, Bounds for sorting by prefix reversals, Discrete Math., 27 (1979), 47-57.

[5] A. Capraran, Sorting permutations by reversals and Eulerian cycle decompositions, SIAM J. Discrete Math., 12 (1999), 91-110.

[6] A. Capraran, Sorting by reversals is difficult, Proceedings of the RECOMB'97, ACM Press, New York, 1997.

[7] D.A. Christie and R.W. Irving, Sorting strings by reversals and by transpositions, SIAM J. Discrete Math., 14 (2001), No.2, 193-206.

[8] D.A. Christie, Sorting permutations by block-interchanges, Inform. Process. Lett., 60 (1996), 165-169.

[9] J.A. Bondy and U.S.R.Murty, Graph theory with applications, The Macmillan Press Ltd., 1976.

[10] C.D. Godsil, Algebraic combinatorics, Chapman and Hall, New York, 1993.