

Distances between Signed Permutation

Huei-Min Wu

Department of Applied Mathematics
National Chiao Tung University, Taiwan

2012/11/05

Signed permutations allow taking the relative orientation of homologous markers into account and constitute a more biologically relevant model for genomes; for instance, reversals always change the strand of the reversed segment, so that the orientation of a gene inside the segment is changed as well. Therefore, whenever the orientation of genes is known and mutations that affect orientations are considered, it is better to use signed permutations as a model. Remarkably, some problems that are intractable for unsigned permutations become tractable for signed ones. Such a notable example is the reversal distance.

Recall: signed permutation

- Definition 2.8

A **signed permutation** on $\{1, 2, \dots, n\}$ is a permutation of the set $\{-n, \dots, -2, -1, 1, 2, \dots, n\}$ that satisfies $\pi_{-i} = -\pi_i$.

- Example.

$$\begin{pmatrix} -4 & -3 & -2 & -1 & 1 & 2 & 3 & 4 \\ 3 & -1 & 4 & 2 & \underline{-2} & -4 & 1 & -3 \end{pmatrix}$$

is simply denoted by $(-2 \ -4 \ 1 \ -3)$.

Signed Reversal

- Definition 4.3

For any permutation $\pi \in S_n^\pm$, the **signed reversal** $\rho(i,j)$ with $1 \leq i \leq j \leq n$ applied to π reverses the closed interval determined by i and j and reverses the sign of all its elements, transforming π into $\pi \circ \rho(i,j)$. Therefore, $\rho(i,j)$ is the following permutation:

$$\begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & \cdots & i-1 & -j & -(j-1) & \cdots & -(i+1) & -i & j+1 & \ddots & n \end{pmatrix}$$

Breakpoint Graph of a Signed Permutation

Definition 4.4

The **breakpoint graph** of a signed permutation π is the graph

$BG(\pi) = (V, E)$:

► vertex set V :

for each integer i in $\{1, \dots, n\}$, let i^- and i^+ be two vertices in V ;
add to V the two vertices 0^+ and $(n+1)^-$.

► edge set E :

the union of two perfect matchings of V , respectively denoted by **D** , the **desire edges**, and **R** , the **reality edges**:

D contains the edges $i^+(i+1)^-$ for all i in $\{1, \dots, n\}$.

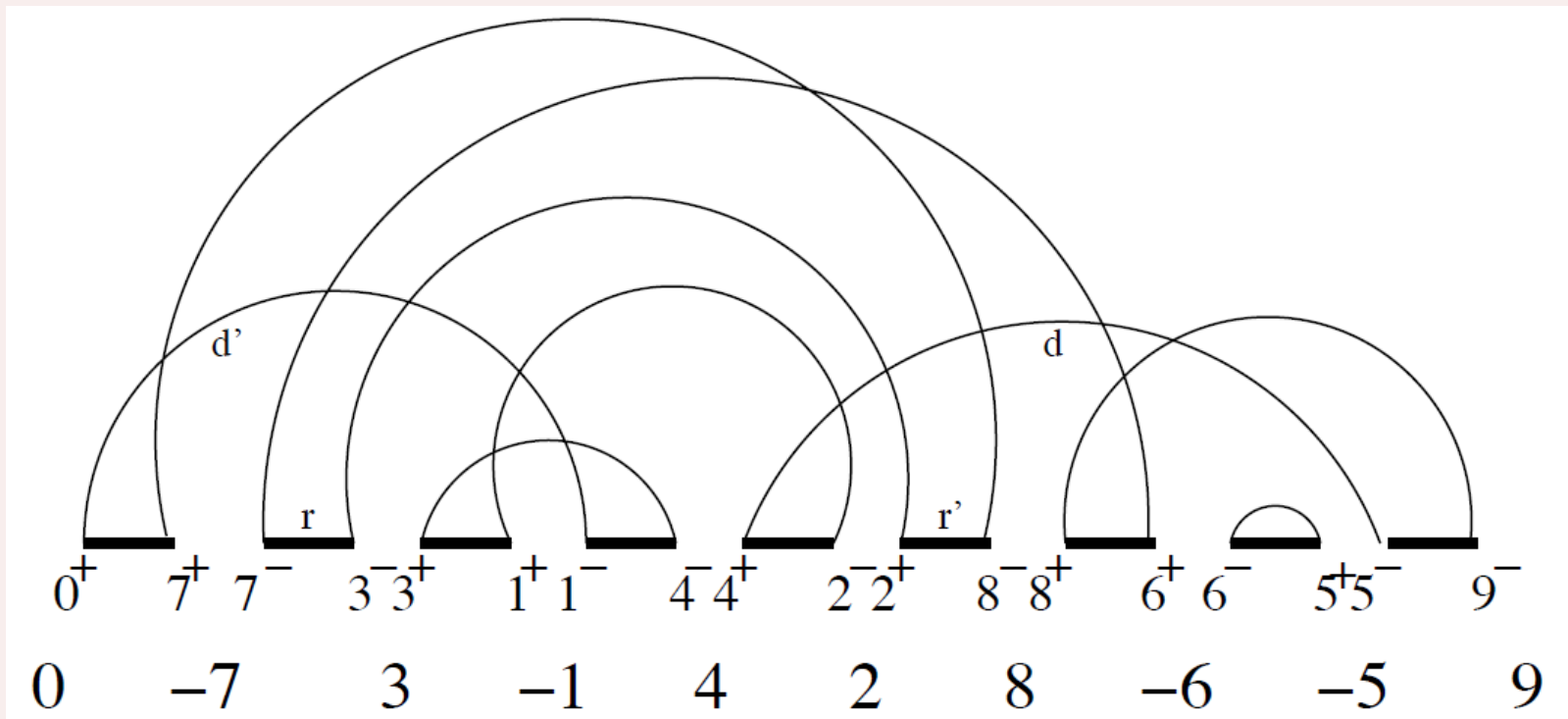
R contains an edge for all i in $\{1, \dots, n\}$,

from π_i^+ if π_i is nonnegative, and from π_i^- otherwise,
to π_{i+1}^- if π_{i+1} is nonnegative, and to π_{i+1}^+ otherwise.

Breakpoint Graph of a Signed Permutation

Example.

The breakpoint graph of the permutation $(-7 \ 3 \ -1 \ 4 \ 2 \ 8 \ -6 \ -5)$.
The desire edges are represented by thin lines, and the reality edges are represented in bold.



Breakpoint Graph of a Signed Permutation

- The breakpoint graph of a signed permutation is exactly the breakpoint graph of the unsigned permutation obtained by replacing every signed element with two unsigned elements, except that we omit edges of the form i^-i^+ .
- It is easy to check that every vertex of $BG(\pi)$ has degree 2 (it has one incident edge in R and one in D), so the breakpoint graph is a set of disjoint alternating cycles, which makes the disjoint cycle decomposition trivial, as is the case for the cycle graph.

Recall: common interval

- Definition.

An interval $I = \{|\pi_a|, \dots, |\pi_b|\}$ of a permutation π is a **common interval** if it is an interval of the identity permutation.

Let m_I and M_I denote, respectively, the smallest and largest elements of an interval I , that is,

$$m_I = \min_{i \in [a,b]} |\pi_i| \text{ and } M_I = \max_{i \in [a,b]} |\pi_i|.$$

- Example.

3 5 -4 1 -2

$$m_I = 3, M_I = 5$$

Conserved Interval

- Definition 4.1

A common interval $I = \{|\pi_a|, \dots, |\pi_b|\}$ is **conserved** if either $\pi_a = m_I$ and $\pi_b = M_I$, or $\pi_a = -M_I$ and $\pi_b = -m_I$.

- Example.



Component

- Definition 4.5

A **component** of a permutation is a conserved interval which is not the union of other conserved intervals.

- Example.



Component

- Definition 4.5

A point (π_i, π_{i+1}) belongs to a component C if both π_i and π_{i+1} are in C and C is the smallest component that contains them.

- Definition 2.13

The linear extension of π is the permutation defined by $\pi^l = (0 \ \pi_1 \ \dots \ \pi_n \ n+1)$.

- Definition 2.14

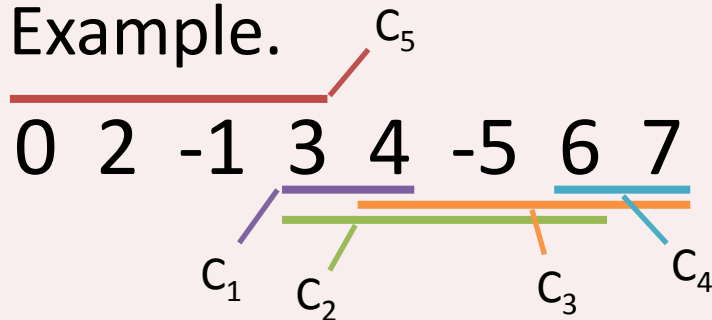
A point of π is an ordered pair (π_i^l, π_{i+1}^l) for $0 \leq i \leq n$.

Component

- Definition 4.5

A point (π_i, π_{i+1}) **belongs** to a component C if both π_i and π_{i+1} are in C and C is the smallest component that contains them.

- Example.



$(3,4)$ belongs to C_1 , not belongs to C_2 .

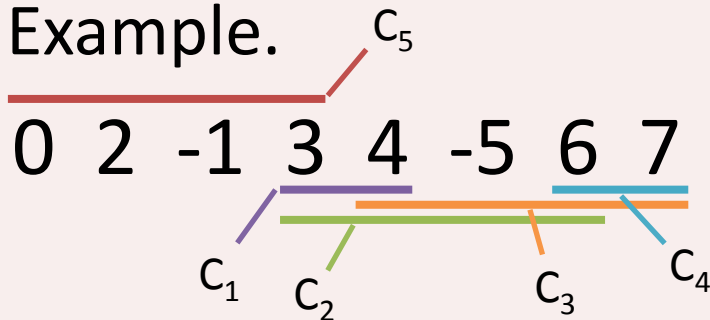
$(6,7)$ belongs to C_4 , not belongs to C_3 .

Component

- Definition 4.5

A point is **oriented** if its elements have opposite signs.
A component is said to be **oriented** if an oriented point belongs to it.

- Example.



(4,-5) belongs to C_2 , C_2 is oriented.
(2,-1) belongs to C_5 , C_5 is oriented.

The Structure of the Family of all Components

Property 4.1

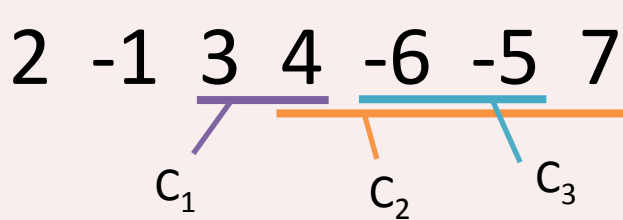
For two different components of a signed permutation, one of the following holds:

- They are disjoint.
- One is contained in the other and they have different extremities.
- They overlap on one element.

The Structure of the Family of all Components

- They are disjoint.
- One is contained in the other and they have different extremities.
- They overlap on one element.

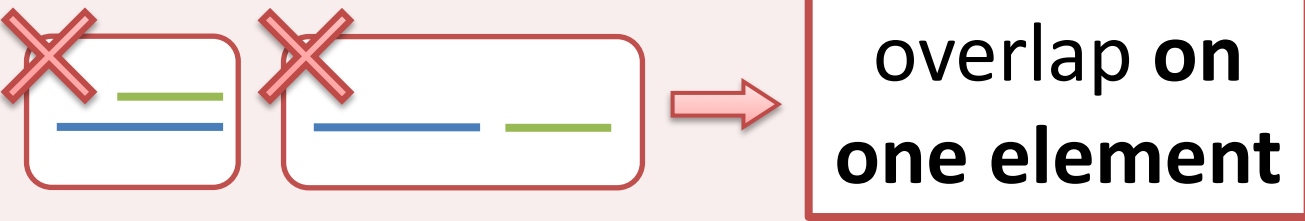
Example.



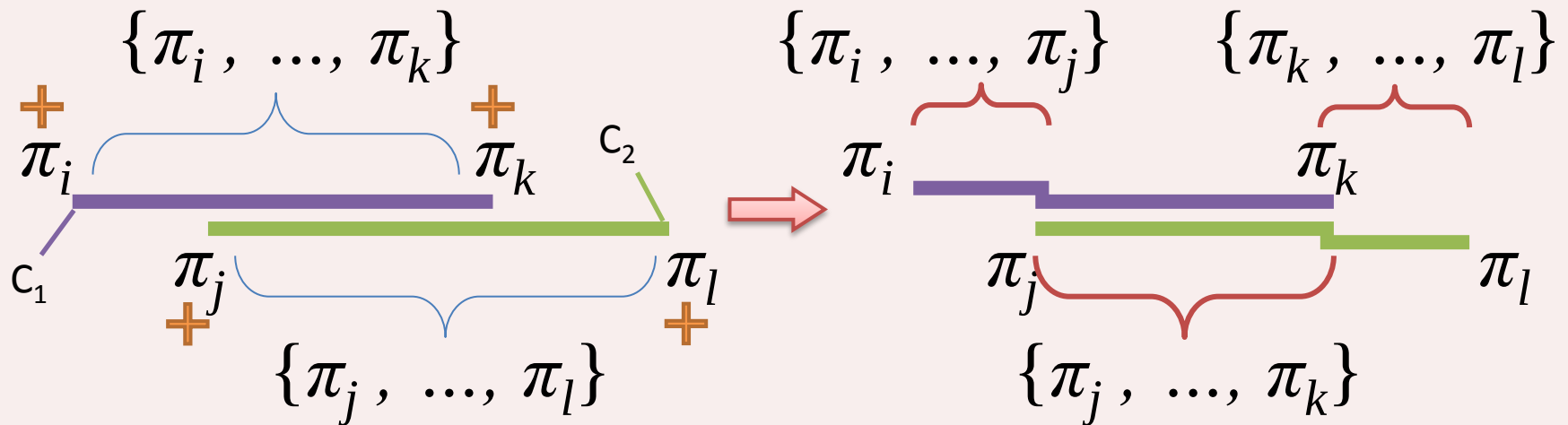
C_1 and C_2 overlap on one element.

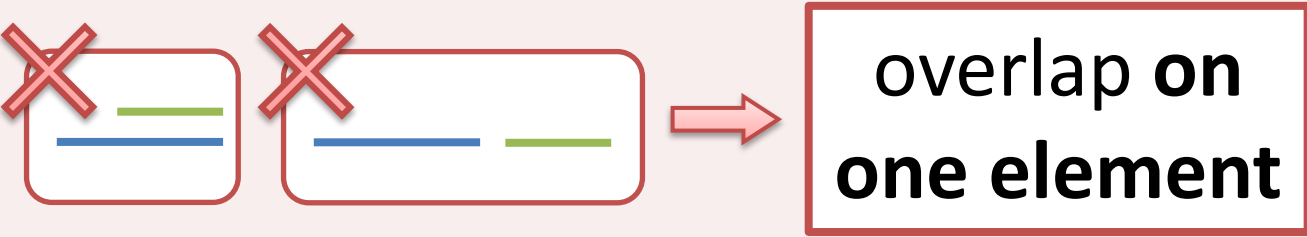
C_2 contains C_3 .

C_1 and C_3 are disjoint.

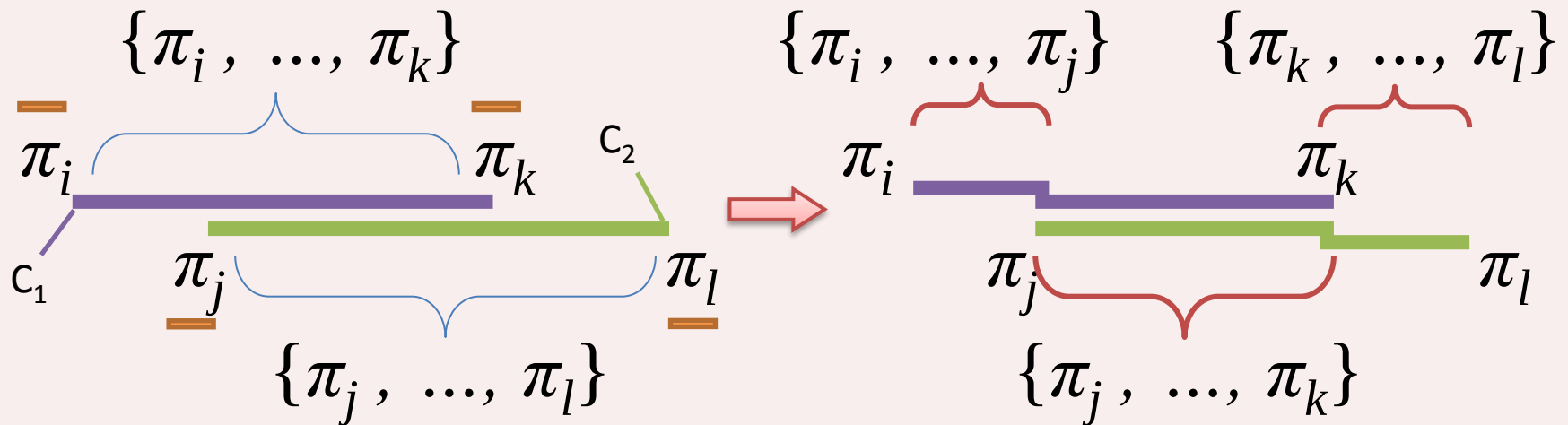
Proof:  overlap on one element

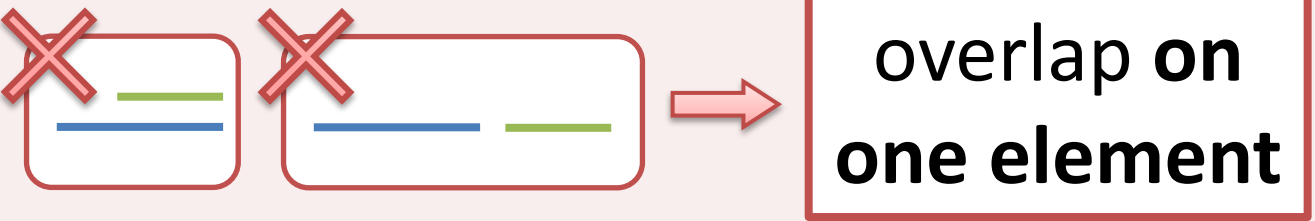
Case 1.



Proof:  **overlap on one element**

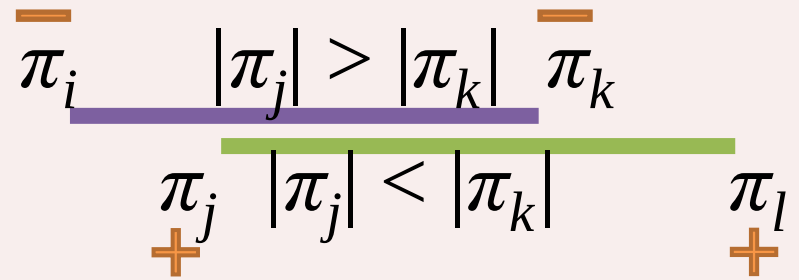
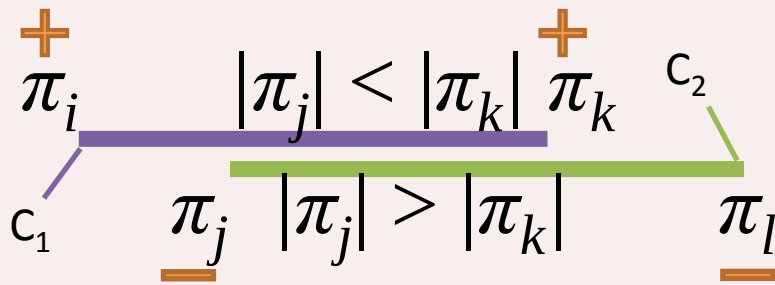
Case 2.



Proof:  overlap on one element

Case 3.

Case 4.



- **The Space of All Optimal Solutions**

Almost all studies on sorting sequences of reversals have been devoted to giving only one sequence, though it has been remarked that there are often plenty of them (it may be over several millions even for $n \leq 10$).

- **Experimental Results**

Bader et al. [26] provide some experimental tests of their algorithm to compute the reversal distance. Bergeron and Strasbourg [46] have performed some experiments on an $O(n^3)$ algorithm to sort by reversals.

Variants of Sorting by Reversals

- Prefix Reversals
- Reversals That Are Symmetric Around a Point
- Weighted Reversals
- Fixed-Length Reversals

Combined Operations

More work has been conducted on combined operations in the signed case than in the unsigned case, because these problems are attempts to come back to the biological application, and signed permutations seem to be a more realistic model.

- Reversals and Transpositions
- Reversals, Transpositions, Transreversals, Revrevs
Transreversal: a transposition followed by a reversal on the transposed segment.
Revrev: two reversals of contiguous segments.

References

- [29] V. Bafna and P. A. Pevzner, *Genome rearrangements and sorting by reversals*, in Proceedings of the Thirty-fourth Annual Symposium on Foundations of Computer Science (FOCS), IEEE Computer Society Press, 1993, pp. 148– 157.
- [44] A. Bergeron, *A very elementary presentation of the Hannenhalli-Pevzner theory*, Discrete Applied Mathematics, 146, no. 2 (2005), pp. 134– 145.
- [45] A. Bergeron and J. Stoye, *On the similarity of sets of permutations and its applications to genome comparison*, in Proceedings of the Ninth International Computing and Combinatorics Conference (COCOON), T. Warnow and B. Zhu, eds., vol. 2697 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 68– 79.
- [142] Y. Diekmann, M.-F. Sagot, and E. Tannier, *Evolution under reversals: Parsimony and conservation of common intervals*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 4, no. 2 (2007), pp. 301– 309.
- [199] S. Hannenhalli and P. A. Pevzner, *Transforming cabbage into turnip: Polynomial algorithm for sorting signed permutations by reversals*, Journal of the ACM, 46, no. 1 (1999), pp. 1– 27.
- [352] E. Tannier, A. Bergeron, and M.-F. Sagot, *Advances on sorting by reversals*, Discrete Applied Mathematics, 155, no. 6– 7 (2007), pp. 881– 888.



Thanks for your attention !

*You know you've read a good book
when you turn the last page and feel as if you've lost a friend.*