3. Sorting by Signed Reversals

Literature:

Bergeron, A., Mixtacki, J., & Stoye, J. (2005). *Chapter 10: The Inversion Distance Problem.* In: Gascuel, O. (ed.) Mathematics of Evolution and Phylogeny. Oxford University Press, pp. 262-290.

Tannier, E., Bergeron, A., & Sagot, M.-F. (2007). *Advances on sorting by reversals.* Discrete Applied Mathematics, Vol. 155(6-7), pp. 881–888.

Problem 2 (Sorting by Signed Reversals)**.** *Given two signed permutations π and σ, find a series of reversals $\rho_1, \dots, \rho_d$, such that $\pi \circ \rho_1 \circ \dots \circ \rho_d = \sigma$ and $srd(\pi, \sigma) = d$.*

It turns out that finding an actual sorting scenario is more complicated than computing the reversal distance.

3.1. An algorithm for computing a sorting scenario

The algorithm has two steps:

1. Merge or cut components to transform all unoriented into oriented components.
2. Apply reversals of type I to break cycles into trivial cycles (adjacencies).

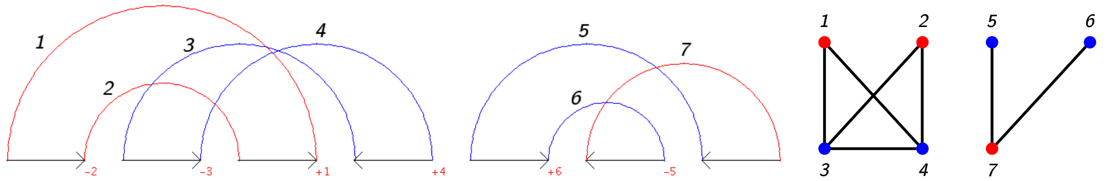
Step 1 can be performed in $O(n)$ time, see Bader, Moret, and Yan (2001), whereas step 2 requires $O(n^{\frac{3}{2}})$ time by the best, known algorithm by Tannier and Sagot (2005) and Han (2006). The remainder of this chapter focuses on step 2, discussing a less sophisticated algorithm that achieves this task only in quadratic time. Its underlying strategy is to identify so-called *save reversals*.

A reversal always *acts* on two (reality) edges, but it can affect other edges. To study these effects, we need a more appropriate data structure! Meet the *overlap graph*:

Definition 10. *The* overlap graph $OV(\pi)$ *of a permutation π is the graph whose vertices are the $n+1$ desire edges (arcs) of $BG(\pi)$ and whose edges correspond to crossings between them.*

Example 7. *The edge-labelled breakpoint graph and the overlap graph of* $\pi^1 = (0\ -2\ -3\ 1\ 4\ 6\ 5\ 7)$ *is shown in Figure 3.1.*

Observation 4. *Isolated vertices correspond to adjacencies of the permutation.*



(a) (b)

Figure 3.1.: (a) The breakpoint graph with labelled desire edges of permutation π^1; (b) The overlap graph of π with vertices corresponding to the edge label of its breakpoint graph.

Observation 5. *A component of π is a connected component of $OV(\pi)$.*

What happens to $OV(\pi)$ when we apply a reversal? To understand the intricacy of the reversal operation, it is helpful to look at vertex-induced subgraphs in $OV(\pi)$:

Definition 11 ((Vertex-) Induced Subgraph)**.** *Given a set of vertices $S \subseteq V$ of a graph $G = (V, E)$, the S-induced subgraph of G is the graph $G' = (S, E')$, where $E' = \{(u, v) \in E \mid u, v \in S\}$.*

Definition 12 (Local Complementation)**.** *Let G_v be the induced subgraph of a vertex v and its adjacent vertices (the neighborhood of v), the* local complementation *of v, denoted G/v, is the operation that complements all (i) edges and (ii) colors of vertices of subgraph G_v.*

Lemma 2. *For a permutation π and an oriented vertex v of the overlap graph, $OV(\pi \circ \rho(v)) = OV(\pi)/v$.*

Theorem 4. *If $OV(\pi)$ has no unoriented component, then it has an oriented vertex v such that $OV(\pi)/v$ has no unoriented component.*

The proof (see Bergeron *et al.*) makes use of the following definition and lemma:

Definition 13. *The score of a reversal is the number of oriented vertices in the resulting permutation.*

The score s of a reversal can be easily computed in the overlap graph: Let v be a vertex of $OV(.)$. Clearly the score of $\rho(v)$ is given by

$$s(\rho(v)) = T + U(v) - O(v) - 1\,,$$

where T is the total number of oriented vertices in the overlap graph and $U(v)$, $O(v)$ are the number of unoriented, respectively oriented vertices adjacent to v.

Example 7 (continued)**.** *Figure 3.2 shows for each vertex of $OV(\pi^1)$ its corresponding reversal score.*

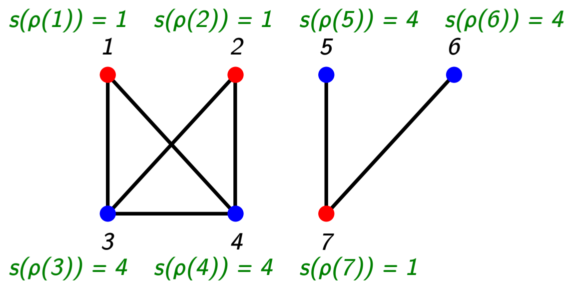


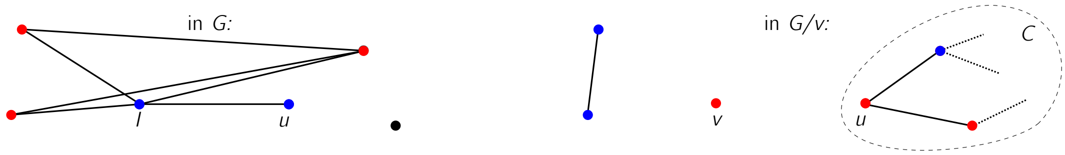
Figure 3.2.: Reversal scores corresponding to vertices of $OV(\pi^1)$.

Lemma 3. *A type I reversal of maximal score does not create new unoriented components.*

Proof. We now show that if an unoriented component would be created by the reversal with maximal score, then this unoriented component would contain a vertex whose corresponding reversal has an even higher score, contradicting the initial assumption.

Given an overlap graph $G = (V, E)$ and a vertex v with $s(\rho(v))$ being maximal over all vertices in V. Assume that there is an unoriented component C in G/v. Clearly, if C is oriented in G, then, in order to become unoriented, component C must contain at least one vertex u that is connected to v in G and thus is affected by the complementation.

Observe that all unoriented vertices of G that are connected to v must also be connected to u, otherwise C would not be unoriented in G/v. Thus, $U(u) \geq U(v)$.



Following the same logic, all oriented vertices connected to u in G must also be connected to v, otherwise C would not be unoriented in G/v. Thus, $O(u) \leq O(v)$ and we have $T + U(v) - O(v) - 1 \leq T + U(u) - O(u) - 1 \Leftrightarrow s(\rho(v)) \leq s(\rho(u))$.

Now, if $O(u) = O(v)$ and $U(u) = U(v)$, then v's and u's induced subgraph in G are identical, hence all vertices that are no longer connected to v in G/v are also not connected to u.

In other words, for u to be part of an unoriented component of G/v entails the neighborship of an additional unoriented vertex that is not connected to v in G. But this would mean that $s(\rho(u)) > s(\rho(v))$, a contradiction! □

Definition 14. *A reversal with maximal score is* safe.

Theorem 5. *If $\rho(v)$ is a safe reversal of permutation π, then $srd(\pi \circ \rho(v)) = srd(\pi) - 1$.*