

## Algorithms for Genome Rearrangement

Summer 2017

### Exercises

#### Exercise 09, 16.06.2017

1. Given a gene similarity graph  $G = (V_S, V_T, E)$  of two genomes  $S$  and  $T$ , recall the following measures that are functions of a matching  $\mathcal{M} \subseteq E$ : (4 P)

$$adj_{ST}(\mathcal{M}) = \sum_{\substack{e, f \in \mathcal{M} \\ e, f \text{ form cons. adj. in } S^{\mathcal{M}}, T^{\mathcal{M}}}} \sqrt{w(e) \cdot w(f)} \quad (1)$$

$$edg(\mathcal{M}) = \sum_{e \in \mathcal{M}} w(e) \quad (2)$$

$$F_{\alpha}(\mathcal{M}) = \alpha \cdot adj_{ST}(\mathcal{M}) + (1 - \alpha) \cdot edg(\mathcal{M}) \quad (3)$$

- (a) show that for any matching  $\mathcal{M}$  of  $G$  holds true that  $adj_{ST}(\mathcal{M}) \leq edg(\mathcal{M})$ .  
(b) Let  $\mathcal{M}'$  be a maximum weight matching in  $G$ . Show that for any matching  $\mathcal{M}$  of  $G$  that is a solution to problem FF-Adjacencies holds true that

$$(1 - \alpha) \cdot edg(\mathcal{M}') \leq F_{\alpha}(\mathcal{M}) \leq edg(\mathcal{M}) \leq edg(\mathcal{M}').$$

2. Devise an algorithm that enumerates all conserved adjacencies of two genomes  $S$  and  $T$  *without a matching*, that is (4 P)
- (a) a gene can participate in multiple conserved adjacencies, but  
(b) a conserved adjacency can only be established between pairs of genes that are directly adjacent, i.e no other gene is allowed to sit in between.
3. Read the paper of Martinez *et al.* "On the family-free DCJ distance and similarity" *Algorithms for Molecular Biology* 10(1), 13, 2015 and participate in the discussion in the tutorial on Friday, June 23. (4 P)

Hand in solutions before the tutorial on 23.06.2017