

Algorithms in Comparative Genomics

Universität Bielefeld, WS 2020/2021

Dr. Marília D. V. Braga · Leonard Bohnenkämper

<https://gi.cebitec.uni-bielefeld.de/teaching/2020winter/cg>

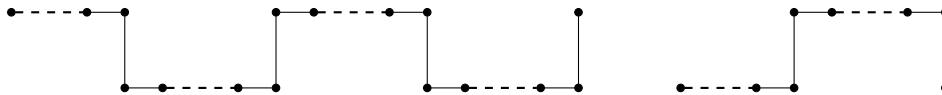
Exercise sheet 12, 28.1.2021

Exercise 1 (Deducting chains of recombinations and cappings)

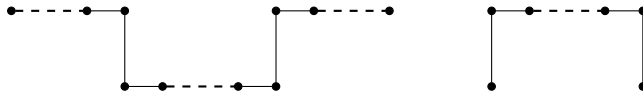
(8 pts)

Identify the path and indel-types of each of the sets of the following components. Which chain of deducting recombination covers all the components in the subtask? Give the capping associated with the corresponding chain.

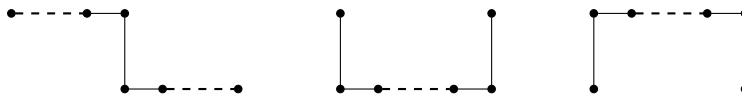
1.



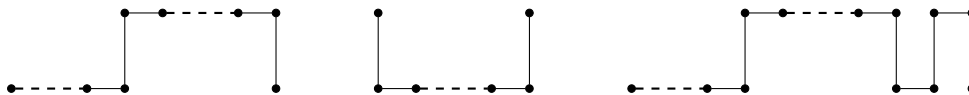
2.



3.



4.



Exercise 2 (Adapting the ILP to different matching models)

(1+2+3 pts)

Given the capped multi-relational graph that contains all possible indel-edges¹:

1. How can you enforce the maximal matching model using ILP constraints (Hint: You only need one extra constraint on the x -variable).
2. Adapt the ILP to the exemplary matching model (Hint: You need only need one extra constraint on the x -variable).
3. Adapt the ILP to the intermediate matching model (Hint: You need extra constraints on the x -variable, but also on the objective function²).

Exercise 3 (Proving important lemmata)

(1+1+2 pts)

1. Prove (e.g. via contradiction): There is always an even number of transitions in a cycle.
2. Prove (e.g. via induction): Prove that the formulas for computing the indel-potential via transitions and via runs are equivalent.
3. Prove that the DCJ-indel distance of natural genomes is NP-hard by giving a reduction from balanced genomes. For which matching model does it apply?

¹As opposed to removing some to enforce the maximal matching model

²Are n_* , p_* still constant? :)