Algorithms in Comparative Genomics

Bielefeld University, Summer Semester 2025
Dr. Marília D. V. Braga · Prof. Dr. Jens Stoye

https://gi.cebitec.uni-bielefeld.de/teaching/2025summer/cg

Exercise sheet 12, discussion: 18.07.2025

Exercise 1 (Median gene clusters)
Given two linear genomes (represented as sequences of gene families):

$$\begin{aligned} \mathbb{A} &= [1,4,7,6,5,4,5,1,4,3,2] \\ \mathbb{B} &= [1,2,3,4,3,4,5,7] \end{aligned}$$

1. Find all common intervals of $\mathbb{A}$ and $\mathbb{B}$ of size at least 2 and list all their occurrences in $\mathbb{A}$ and in $\mathbb{B}$. Which of these occurrences are maximal?

2. Find five maximal median gene clusters of $\mathbb{A}$ and $\mathbb{B}$ of size at least 3 with a symmetric set distance of at most 1.

Exercise 2 (Factoradic numbers)
Obviously there exist bijective mappings between the numbers $1, 2, \ldots, n!$ and the (unsigned) permutations over $\{1, 2, \ldots, n\}$. Find such a mapping that is computable in both directions in polynomial time.