

Advanced Algorithms in Bioinformatics (P4)

Sequence and Structure Analysis

Freie Universität Berlin, Institut für Informatik
Prof. Dr. Knut Reinert, Sandro Andreotti
Sommersemester 2009

7. Exercise sheet, 6. June 2009

Discussion: 12. June 2009

Exercise 1.

Match Refinement Example

Given Sequences $A = \text{AAGCGCCCGCG}$ and $B = \text{AAGCGGGCCCGCG}$ and the projection maps (no indels):

$$\begin{array}{ll} \alpha_{S_1}[0, 5] \rightarrow [0, 5] & \beta_{S_1}[0, 5] \rightarrow [0, 5] \\ \alpha_{S_2}[2, 5] \rightarrow [11, 14] & \beta_{S_2}[11, 14] \rightarrow [2, 5] \\ \alpha_{S_3}[8, 11] \rightarrow [2, 5] & \beta_{S_3}[2, 5] \rightarrow [8, 11] \\ \alpha_{S_4}[5, 11] \rightarrow [8, 14] & \beta_{S_4}[8, 14] \rightarrow [5, 11] \end{array}$$

- Draw the two sequences and their corresponding segment matches.
- Compute the minimal resolved refinement by applying the algorithm from the lecture.

Exercise 2.

Prove the lemmata used for the Manhattan distance and the sum-of-pairs distance in the chaining problem, as discussed in the lecture (Lemma 1 and Lemma 3).

*Exercise 3.

Graph-theoretic formulation of the chaining problem:

Find an example where the size of the edge set is $\Omega(V^2)$ and removal of transitive edges does not improve the worst-case asymptotic.