

Advanced Algorithms in Bioinformatics (P4)

Sequence and Structure Analysis

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

5. Exercise sheet, 16. May 2009
Discussion: 22. May 2009

Exercise 1.

1. Sketch the Function $U(n, q, \epsilon)$ for increasing values of n . Let $q = 7$, $\epsilon = 0.1$ and $n_0 = 30$. Draw it for the intervall $n = n_0 \dots n_0 + 35$.
2. How many q -hits do ϵ -matches of length $n \geq n_0$ at least have?

Exercise 2.

The Myers bitvector algorithm searches for approximate matches. For each position of the text it returns the edit distance of the best match ending at this position. How can we obtain the starting positions with another run of Myers algorithm?

Exercise 3.

Given the text catcabcb\$ construct the suffix array using the Manber Myers algorithm.

Exercise 4.

Work out the details of an algorithm to build a suffix array in linear time from a suffix tree.