

Lecture: Foundations of Sequence Analysis
Winter Semester 2003/2004

Exercises

Exercise 8, Discussion: 01/26/2004.

1. Given a string S .
 - (a) Give a formal specification of the set of all minimal unique substrings (prefix-minimal as well as suffix-minimal).
(A substring should be given by a pair (i, l) , where i is the starting position with respect to S and l is the length.)
 - (b) Give a formal specification of the set of all prefix-minimal unique substrings of S .
 - (c) Given the string $S = AAGAGTAGAGGT$ and its suffix tree. Find all prefix-minimal unique substrings of length at least 3.
2. Given the strings $S_1 = AACCATC$ $S_2 = ACTCATC$.
 - (a) Construct the generalized suffix tree for S_1 and S_2 .
 - (b) Calculate all maximal unique matches by using the generalized suffix tree.
3. Given a string S and the suffix tree for S .
 - (a) Describe an algorithm to find all maximal repeats.
 - (b) Calculate all maximal repeats for the string $S = AAGAGTAGAGGT$ by using the suffix tree for S .