

Advanced Algorithms in Bioinformatics (P4)

Sequence and Structure Analysis

Freie Universität Berlin, Institut für Informatik
Knut Reinert, Sandro Andreotti
Sommersemester 2012

6. Exercise sheet, 19. June 2012

Discussion: 27. June 2012

Exercise 1.

BWT

- For the text **tacaacaatacaagag** construct the BWT and the arrays C and OCC . Use them to search for the pattern **aca**.

Exercise 2.

lcp-intervals:

- For the text **tacaacaatacag**, construct the lcp-table and apply the bottom-up algorithm to compute all lcp intervals.

Exercise 3.

child table:

- For the same text as in Exercise 2, compute the child table (up , $down$, $nextlIndex$). Verify if your results agree with the intervals computed in Exercise 2.

Exercise 4.

Top down traversal:

- Use the enhanced suffix array constructed in Exercises 2 + 3 and apply algorithm **answering_decision_queries** from the script to search for the pattern **caa**