

Sequence Analysis SS 2013

Freie Universität Berlin, Institut für Informatik

Knut Reinert, Sandro Andreotti

Sommersemester 2013

6. Exercise sheet, 26. May 2013

Discussion: 30. May 2013

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.

BWT - compressing L

- Let R be the MTF encoding of L and Y the corresponding list of characters. Give an algorithm in pseudocode to decode R into L

Exercise 3.

BWT - compressing pos

- Present an example that proves the following assumption stated in the script:
If we mark every η -th row in the matrix \mathcal{M} the worst case time of a pos query is $O(\frac{\eta-1}{\eta}n)$