

Special course in Computer Science:  
Advanced Text Algorithms  
Exercise set 2  
Due: 18.12.2018

1. Consider the following sequences:

S1: AATATTCGCATCCGGTA  
S2: TATATCGCTCCGTA  
S3: AATTTTCGCATCCGCCTA  
S4: AATAAATCGCTCCGTA  
S5: TTAATATGCATCCCGTA

Apply the Center Star Method to align the strings S1, S2, and S3. Extend the previously generated multiple alignment with the remaining two strings, by enlarging the previous star graph with two more branches, from the previously chosen centre to the new elements.

2. For two words  $u$  and  $v$ , Table  $SUF_u$  is defined for the positions  $1 \leq i \leq |v|$ . The value  $SUF_u[k]$  is the size of longest suffix of  $v[1..k]$  which is also a suffix of  $u$ .

Use the suffix tree method to provide an algorithm computing the  $SUF_u$  table in linear time in the size of  $v$  (independent of the size of  $u$ )

(Discuss in details the complexity of the algorithm)

3. The operation test is a composition of two smaller boolean functions: `righttest` and `lefttest` (see lecture slides from Lecture 10: Detecting text regularities):

- `righttest` searches for a square whose centre is in  $v$ ,
- `lefttest` searches for a square whose centre is in  $u$ .

Prove that (i.e., provide an algorithm) the Boolean value `lefttest(u,v)` can be computed in  $O(|u|)$  time (independent on  $|v|$ ).  
(you can use any of the functions previously defined and analyzed in the lectures)