

4. EXERCISE "BIOINFORMATICS", SS 17

Aufgabe 1: (7 Credits)

Prove the overlap lemma (see slide 13, `12-shotgun.pdf`).

Aufgabe 2: (3 Credits)

For every $m > 0$ find a substring-free set P such that the algorithm **MGreedy** returns a superstring at least $m - 2$ characters longer than the shortest possible $S^*(P)$. Explain your results.

Aufgabe 3: (5 Credits)

For a given string T and the pattern P , the worst-case time complexity of the brute-force pattern matching algorithm (see slide 3, `13-alignment.pdf`) is in $O(|P|(|T| - |P| + 1))$. Give an example of a string T and a pattern P such that the brute-force pattern matching algorithm indeed performs $|P|(|T| - |P| + 1)$ comparisons of characters to find *one* occurrence of P in T .

This illustrates that the bound is actually tight.

Aufgabe 4: (5 Credits)

Use the existence of a linear-time exact matching algorithm to solve the following problem in linear time. Given two strings α und β , determine if α is a cyclic (or circular) shift of β .

Deadline: Tuesday - May 2, 2017