

2. ÜBUNG "BIOINFORMATIK", SS 17

Aufgabe 1: (5 Credits)

Prove *Lemma 2* stated in the lecture:

Given a set of strings $P = \{S_1, \dots, S_r\}$ that is substring-free and a permutation $\Pi = \sigma_1 \dots \sigma_r$ of the integers $1, \dots, r$ that implies the ordered set $\{S_{\sigma_1}, \dots, S_{\sigma_r}\}$. Let

$$S(\Pi) = \text{pref}(S_{\sigma_1}, S_{\sigma_2}) \text{pref}(S_{\sigma_2}, S_{\sigma_3}) \dots \text{pref}(S_{\sigma_{r-1}}, S_{\sigma_r}) S_{\sigma_r}$$

Show that $S(\Pi)$ is a superstring $S(P)$.

Aufgabe 2: (7.5 Credits)

Implement an algorithm in C++ that determines the shortest common supersequence (SCS) of n user-defined strings.

Give the computed results and the intermediate steps of your algorithm on the strings $S_1 = \text{CCTT}$, $S_2 = \text{ACCCT}$, $S_3 = \text{TTC}$.

Determine the runtime of your algorithm depending on the number of the strings and their length.

Send the source-code via email to the tutor Nikolai Nøjgaard.

Aufgabe 3: (7.5 Credits)

Let X, Y, Z and Z' be distinct strings s.t. the set $\{X, Y, Z, Z'\}$ is substring-free.

Prove the following statement:

If $\text{ov}(X, Y) \geq \max\{\text{ov}(X, Z), \text{ov}(Z', Y)\}$, then $\text{ov}(X, Y) + \text{ov}(Z', Z) \geq \text{ov}(X, Z) + \text{ov}(Z', Y)$.

Deadline: Tuesday - April 18, 2017 - 8:30am