
Algorithmen auf Sequenzen

Abgabetermin: Donnerstag, den 1. Februar vor der Vorlesung

Aufgabe (Notenbonus) 1

Konstruiere für das Wort $t = aaabaabab$ das zugehörige Extended-Suffix-Array, d.h. das Suffix-Array A und die LCP-Tabelle L und zeichne den zugehörigen LCP-Intervall-Baum von t (es gilt $\$ < a < b$).

Aufgabe (Notenbonus) 2

Betrachte das Wort $t = \text{ABANANAANDANANANAS}$.

- Konstruiere die Burrows-Wheeler-Transformierte \hat{t} zu $t\$$.
- Gib die zugehörige LF-Funktion für \hat{t} an.
- Gib die Werte $C(\cdot)$ und $Occ(\cdot, \cdot)$ an.
- Suche nach $s = \text{NANA}$ im FM-Index für t mit Hilfe des in der Vorlesung angegebenen Algorithmus für die Rückwärtssuche im FM-Index.

Es gilt $\$ < A < B < D < N < S$.

Aufgabe 3

Sei $t \in \Sigma^n$ ein Text, $p \in \Sigma^m$ ein Muster und $k \in \mathbb{N}_0$ eine Konstante.

Beschreibe, wie in Zeit $O(kn + m)$ festgestellt werden kann, ob und ggf. wo p in t mit höchstens k Mismatches vorkommt.