

7. Übung „Algorithmen der Bioinformatik I“

1. a) Konstruieren Sie den impliziten Suffixbaum für den String *abbcabbd* (mit Suffixlinks). Tragen Sie die bei der Konstruktion des Suffixbaumes angewandten Regeln (I, IIa, IIb oder III) in eine Tabelle ein, wobei die Zeilen den Phasen $i + 1 = 1, 2, \dots, n$ und die Spalten den Erweiterungen $j = 1, \dots, i + 1$ entsprechen. Tragen Sie dabei auch die implizit angewandten Regeln ein! (3 Punkte)
- b) Wir haben bereits gezeigt, daß die Anwendungen von Regel I, II, III in den Zeilen der Tabelle in dieser Reihenfolge auftreten. Formulieren Sie eine Vermutung über die Reihenfolge des Auftretens in den Spalten der Tabelle! Läßt sich über Regel II noch mehr sagen? (3 Punkte)
2. Konstruieren Sie für jedes $n \in \mathbb{N}$ eine Zeichenreihe $t \in \{a, b\}^n$, d.h. geben Sie eine Konstruktionsvorschrift an, so dass die Summe der Längen der Kantenmarkierungen als Zeichenreihen aus Σ^+ des zu t gehörigen Suffix-Baumes mindestens $O(n^2)$ ist. (3 Punkte)

Zusatzaufgabe zum Knobeln (Lösung derzeit unbekannt...):

3. Welcher Zusammenhang besteht zwischen dem Suffixbaum $ST(S)$ eines Strings S und dem Baum $ST(S^r)$ des reversen Strings S^r ? Die Suffixlinks in beiden Bäumen sollen bei der Lösung der Aufgabe hilfreich sein... (5 Zusatzpunkte)