

# Übungen zur Vorlesung Sequenzanalyse II

Universität Bielefeld, SS 2011  
Prof. Dr. Jens Stoye · Daniel Dörr

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2011summer/SequenzAnalyse>

**Blatt 2 vom 06.05.2011**

**Bitte gib auch den Namen deiner Tutorin an.**

**Abgabe in einer Woche vor Beginn der Vorlesung.**

## **Aufgabe 1 (Verallgemeinerter Suffixbaum)**

(4 Punkte)

Zeige: Zwei Strings  $s$ ,  $t$  sind genau dann identisch ( $s = t$ ), wenn alle Blattkanten des gemeinsamen Suffixbaumes  $s\#t$  entweder mit  $\$$  oder  $\#$  beginnen.

## **Aufgabe 2 (Links-Rechts-Partition)**

(4 Punkte)

Die Links-Rechts-Partition  $P_{lr}(s, t)$  eines Strings  $s$  bezüglich eines Strings  $t$  (siehe Abschnitt 3.8 im Skript) kann mit Hilfe eines Suffixbaumes effizient berechnet werden.

- Überlege dir einen Algorithmus, der die Links-Rechts-Partition  $P_{lr}(s, t)$  in linearer Zeit berechnet.
- Verwende diesen Algorithmus, um  $P_{lr}(s, t)$  für  $s = \text{ABAAABABB}$  und  $t = \text{ABABABA}$  zu berechnen.

## **Aufgabe 3 (Maximale Repeats)**

(4 Punkte)

Lies für diese Aufgabe Abschnitt 8.7.4 im Skript, insbesondere den Algorithmus zum Finden maximaler Repeats.

Zeige mit Hilfe des Suffixbaumes: In jedem String der Länge  $n$  gibt es höchstens  $n$  maximale Repeats.

Bedenke: Es stimmt nicht, dass an jeder Position nur ein maximaler Repeat beginnen oder enden kann.