

# Übungen zur Vorlesung Algorithmen auf Sequenzen

TU Dortmund, SS 2008

Prof. Dr. Sven Rahmann

**Blatt 6 vom 19.05.2008**

**Abgabe am Fr 23.05.2008 in der Vorlesung**

**Aufgabe 1** Zeige, dass reguläre Sprachen bezüglich Durchschnitten abgeschlossen sind. Gibt es eine einfache Methode, aus den regulären Ausdrücken für Sprachen  $L_1, L_2$  den regulären Ausdruck für  $L_1 \cap L_2$  zu konstruieren?

**Aufgabe 2** Vollziehe den in der Vorlesung besprochenen Algorithmus zur Konstruktion des Syntaxbaums aus einem regulären Ausdruck am Beispiel  $((a^*b)^*|a((cbc|b^*)c^*))b$  nach.

**Aufgabe 3** Konstruiere den Thompson-NFA für dasselbe Beispiel.

**Aufgabe 4** Ein Abel'sches Muster ist eine Abbildung  $P : \Sigma \rightarrow \mathbb{N}_0$ , die jedem Symbol des Alphabets eine Häufigkeit zuordnet; es wird auch als symbolische Summe  $\sum_{c \in \Sigma} m_c c$  geschrieben mit  $m_c = P(c)$ . Den Faktor  $m_c = 1$  lässt man weg.

Ein Abel'sches Muster  $P$  passt auf einen String  $t$ , wenn die Symbol-Häufigkeiten in  $t$  mit den in  $P$  gegebenen übereinstimmen. Das Suchproblem für Abel'sche Muster besteht darin, in einem Text  $T$  alle Teilstrings zu finden, auf die  $P$  passt.

Beispiel:  $P = 2a + 3b + c$  passt auf alle Strings, die genau aus zwei a's, drei b's und einem c bestehen.

Gib einen effizienten Präfix-basierten und einen Suffix/Teilstring-basierten Algorithmus für das Suchproblem für Abel'sche Muster an.