

# Übung zu Algorithmen auf Sequenzen

## Blatt 4

Ausgabe: 15.11.2018    Besprechung: 22.11.2018

### Aufgabe 4.1

Erstellen Sie für den Aho-Corasick-Algorithmus einen Trie samt lps-Kanten und Ausgabemengen für die Mustermenge {**he**, **she**, **his**, **her**, **hers**}. Zeigen Sie für den Text **hishershes**, in welchem Zustand sich der Automat bei der Mustersuche nach jedem Zustandsübergang befindet.

### Aufgabe 4.2

Beschreiben Sie im Detail verschiedene Möglichkeiten, die Ausgabemengen für jeden Zustand des Aho-Corasick-Automaten zu implementieren. Analysieren Sie jeweils Speicherbedarf und Auswirkungen auf Konstruktionszeit des Automaten und Laufzeit der Mustersuche.

### Aufgabe 4.3

Erstellen Sie für das erweiterte Muster  $a\#ab\#(1,3)ab\#(1,2)b$  über dem Alphabet  $\Sigma = \{a, b\}$  den entsprechenden NFA. Geben Sie die Masken für alle Zeichen aus dem Alphabet, sowie für  $I$  und  $F$  an. Führen Sie eine Mustersuche auf dem Text **aabababbbababbaababaa** durch. An welchen Positionen wird das Muster erkannt?

### Aufgabe 4.4

Erstellen Sie für das erweiterte Muster  $aa?b?a?bb?aaa?b$  über dem Alphabet  $\Sigma = \{a, b\}$  den dazugehörigen NFA. Suchen Sie das Muster im Text **aabbbaabaab** mit Hilfe des Shift-And-Algorithmus. Erstellen Sie dazu alle nötigen Masken.