

Übung zu Algorithmen auf Sequenzen

Blatt 5

Ausgabe: 17.11.2016 Besprechung: 24.11.2016

Aufgabe 5.1

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 5.2

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 5.3

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 5.4

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.