

Übung zu Algorithmen auf Sequenzen Blatt 2

Ausgabe: 30. Oktober 2014 **Besprechung:** 13. November 2014

Aufgabe 2.1

Erstelle den NFA und den DFA für das Muster $P = \text{AGTACGAG}$. Zeige, in welchen Zuständen sich der NFA während der Durchführung auf dem Text $T = \text{AACGTAAGTACGAGAGTACG}$ befindet.

Stelle dann die lps-Funktion für P auf und führe den KMP-Algorithmus auf T aus.

Gib weiter die Masken zu P für den Shift-And-Algorithmus an.

Gib die Horspool-Sprungtabelle zu P an und führe den Horspool-Algorithmus auf T aus.

Aufgabe 2.2

Erkläre in eigenen Worten, warum der DFA eines gegebenen Musters genauso viele Zustände wie der äquivalente NFA hat.

Aufgabe 2.3

Analysiere die erwartete Laufzeit des naiven Pattern-Matching-Algorithmus, wenn die Buchstaben des Alphabets mit unterschiedlichen Wahrscheinlichkeiten vorkommen. Sei $\Sigma = \{\sigma_1, \dots, \sigma_k\}$. Die Wahrscheinlichkeit für den Buchstaben σ_i sei $p_i \geq 0$ an jeder Stelle, unabhängig von den anderen Stellen. Natürlich ist $\sum_{i=1}^k p_i = 1$.

Aufgabe 2.4

Wie kann man elegant und effizient feststellen, ob ein Muster P eine zyklische Permutation eines anderen Musters Q ist? Beispiel: 34567812 ist eine zyklische Permutation von 12345678, ebenso ATTATGC von ATGCATT.