

Algorithmen auf Sequenzen

Übung – Blatt 2 (2. Version)

Ausgabe: 14. April, **Besprechung:** 19.04. 9:00 Uhr; 21.04. 14:00 Uhr

Aufgabe 2.1

Finde heraus, wie die Zeichen Ä, Ö, Ü, ä, ö, ü und ß in UTF-8 binär und hexadezimal enkodiert sind. Wie sehen die Zeichen aus, wenn sie mit ISO-8859-15 wieder dekodiert werden? Erstelle auch eine analoge Tabelle dieser Zeichen, enkodiert in ISO-8859-15.

Aufgabe 2.2

Berechne für jedes DNA-Codon die Anzahl der synonymen und nichtsynonymen Stellen (Standard-Code). Können Aminosäuren auch durch 2 Basen kodiert werden? Begründe deine Antwort.

Präsentationsaufgabe 2.3

Erstelle sowohl einen NFA, als auch einen DFA für das Muster **AGTACGAG**. Zeige in welchen Zuständen sich der NFA während der Durchführung auf dem Text **AACGTAAGTACGAGAGTACG** befindet.

Aufgabe 2.4

Erkläre warum ein DFA eines gegebenen Musters genauso viele Zustände wie ein äquivalenter NFA hat.

Aufgabe 2.5

Analysiere die erwartete Laufzeit des naiven PM-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 $\sum_{i=1}^k p_i = 1$.