

Algorithmen auf Sequenzen

Übung – Blatt 4

Ausgabe: 28. April, **Besprechung:** 03.05. 9:00 Uhr; 05.05. 14:00 Uhr

Aufgabe 4.1

Führe auf dem Text AACGTAAGTACGAGAGTACG den BNDM-Algorithmus mit dem Muster AGTACGAG aus. Welche Fenster werden betrachtet? Welche Zustände sind nach jedem Schritt jeweils aktiv? Wie viele Zeichenvergleiche werden insgesamt benötigt?

Präsentationsaufgabe 4.2

Erstelle einen Benchmark, der den naiven Pattern-Matching-Algorithmus, den BNDM, den KMP-Algorithmus und den Shift-And Algorithmus (und ggf. einen EOSPM¹-Algorithmus deiner Wahl) testet. Nutze als Suchtext folgende Dateien:

<http://ls11-www.cs.tu-dortmund.de/people/kopczyns/Lehre/AAS/>

- [alice.txt](#)²
- [chromosom.txt](#)³
- [qt-source.txt](#)⁴

Die Muster sind dabei frei wählbar.

Anmerkung: Diese Aufgabe wird erst in der übernächsten Übungsstunde besprochen.

Aufgabe 4.3

Erstelle dem Horspool-Algorithmus entsprechend eine Sprungtabelle für das Muster AGTACGAG. Führe den Algorithmus auf dem Text AACGTAAGTACGAGAGTACG aus.

Aufgabe 4.4

Erstelle für das erweiterte Muster $b-x(1,3)-a-b-x(1,1)-a$ den entsprechenden NFA. Gib die Masken für a, b, sowie die Masken I und F an. Führe auf dem Text bbabbabbaaababbabab den erweiterten Shift-And Algorithmus aus. Welche Zustände sind nach jedem Schritt jeweils aktiv?

Aufgabe 4.5

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 ACGT von GTAC.

¹Exact-Online-Single-Pattern-Matching

²Original: <http://www.gutenberg.org/ebooks/19778>

³Original: <http://www.gutenberg.org/browse/authors/h#a856>

⁴Original: <http://qt.nokia.com/downloads/linux-x11-cpp>