

Übung zu Algorithmen auf Sequenzen

Blatt 9

Ausgabe: 10.01.2019 **Besprechung:** 17.01.2019

Gegeben sind Muster $P = \text{babb}$ und Text $T = \text{aababbaaaabab}$. Suchen Sie P in T mit maximal einem Fehler.

Aufgabe 9.1

Nutzen Sie den Algorithmus von Ukkonen und zeigen Sie genau, welche Felder der Matrix berechnet werden müssen.

Aufgabe 9.2

Nutzen Sie den fehlertoleranten Shift-And-Algorithmus.

Aufgabe 9.3

Bestimmen Sie die 1-Nachbarschaft von babba . (Hinweis: Sie können Die NFA-Topologie verwenden, um die Möglichkeiten systematisch aufzuzählen.)

Zusatzaufgabe 9.4

Bestimmen Sie durch Simulationen die erwartete Edit-Distanz $d(n, k)$ zweier zufälliger Sequenzen der Länge n über einem Alphabet der Größe k (Gleichverteilung). Betrachten Sie für festes k die Asymptotik $d_k := \lim_{n \rightarrow \infty} d(n, k)/n \in [0, 1]$ und geben Sie ungefähre Werte d_2, d_3, \dots, d_{10} an. Wie wären die entsprechenden Werte für die Hamming-Distanz?

Zusatzaufgabe 9.5

Wie können Sie effizient prüfen, ob ein String s eine zyklische Permutation eines anderen Strings t ist? (Hinweise: Sie dürfen $|s| = |t|$ voraussetzen. Denken Sie an alle Techniken, die Sie im Lauf der Vorlesung kennengelernt haben.)