
Algorithmische Bioinformatik I

Abgabetermin: Freitag, der 15. Juli, 9⁰⁰ Uhr in Moodle

Hausaufgabe 1

Konstruiere den Suffix-Baum für $t = t_1 \cdots t_7 = aababaa$ mit Ukkonens Algorithmus aus der Vorlesung. Gib für jeden Präfix $t_1 \cdots t_i$ von t dabei alle ausgeführten Zwischenschritte an, markiere insbesondere die Position des aktiven Knotens und Endknotens im jeweiligen Suffix-Baum. Zeichne dabei nur die verwendeten und neu eingetragenen Suffix-Links mit jeweils einer anderen Farbe ein. Der Übersichtlichkeit wegen sollen hier als Kantenlabel die Zeichenreihen selbst anstelle der üblicherweise verwendeten Referenzen auf t verwendet werden.

Hausaufgabe 2

Sei $s = s_0 \cdots s_{m-1} \in \Sigma^m$.

- Was kann man über das Wort s aussagen, wenn alle seine Z-Werte positiv sind;
- Zeige, wie sich die Z-Werte Z_2, \dots, Z_{p+1} ohne Zeichenvergleiche bestimmen lassen, wenn $Z_1 = p > 0$ gilt.

Hinweis: Begründung nicht vergessen.

Tutoraufgabe 3 (Vorbereitung bis zum 13. Juli 2022)

Gegeben sei eine Menge $S = \{s_1, \dots, s_\ell\}$ von Zeichenreihen mit $n = \sum_{i=1}^{\ell} |s_i|$. Konstruiere einen Algorithmus, mit dem man in Zeit $O(n)$ alle Zeichenreihen $s_i \in S$ finden kann, die Teilwörter einer anderen Zeichenreihe $s_j \in S$ sind.

Hinweise: Ein Suffix-Baum kann helfen. Korrektheitsbeweis und Laufzeitanalyse nicht vergessen.