

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2013/2014
Prof. Dr. Jens Stoye · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2013winter/SequenzAnalyse>

Übungsblatt 10 vom 07.01.2014
Abgabe in einer Woche vor Beginn der Vorlesung.

Aufgabe 1 (Burrows-Wheeler Transformation)

(5 Punkte)

Gegeben sei der String $t = s\$ = \text{TGATGATGATGAGTAG\$}$.

1. Berechne die Burrows-Wheeler Transformierte $bwt(t)$ für t .
2. Schreibe $rle(bwt(t))$ als komprimierten String mit Hilfe von *run-length encoding* auf. Fasse dabei nur Buchstaben zusammen, die mindestens 3 mal hintereinander vorkommen.
3. Überprüfe, ob das Muster $p = \text{TGAG}$ im Text t vorkommt. Beschreibe den String-Matching-Algorithmus unter Verwendung der $bwt(t)$ beispielhaft an der Suche von p .

Aufgabe 2 (Suffixbaum und Suffixarrays)

(5 Punkte)

Gegeben sei der String $t = s\$ = \text{TGATGATGATGAGTAG\$}$.

1. Zeiche den Suffixbaum von t , sortiere die Bätter lexikographisch.
2. Beschrifte die Blätter mit dem Start-Index des zugehörigen Suffixes in t . Die Indizierung beginnt bei 0.
3. Finde den kürzesten eindeutigen Teilstring in t und gib ihn an.
4. Konstruiere die Suffixarrays pos , $rank$ und lcp .

Aufgabe 3 (Klausuraufgabe)

(4 Punkte)

Was für Aufgaben könnten wir wohl in der Klausur stellen? Denke dir eine Aufgabe mit Unterpunkten neu aus, von der du glaubst, sie könnte in der Klausur vorkommen. Gib auch die Lösung zu deiner Aufgabe mit an.