

Algorithmen zur Sequenzanalyse

Wintersemester 2019/2020
 Besprechung am 14.01.2020

Übungsblatt 6

Prof. Dr. E. Ohlebusch,
 Institut für Theoretische Informatik

Aufgabe 6.1.

Geben Sie einen Algorithmus zur Konstruktion der Bitvektoren des Wavelet-Baums für den Fall $\sigma = 2^h$ in Pseudocode an. Analysieren Sie die Laufzeit und den Speicherbedarf Ihres Algorithmus.

Hinweis: Nummerieren Sie die Buchstaben des Alphabets von 0 bis $\sigma - 1$ und benutzen Sie die Binärdarstellungen der Nummern.

Aufgabe 6.2.

Die Burrows-Wheeler Transformierte des Strings $S = in.ulm.um.ulm.und.um.ulm\$$ ist $BWT = mnmndmmn\$uuuluulliu.....$. Beantworten Sie mit Hilfe des Wavelet-Baums (siehe Abbildung 1) die folgenden Anfragen:

- $rank_i(BWT, 20)$
- $select_m(BWT, 4)$
- $BWT[17]$

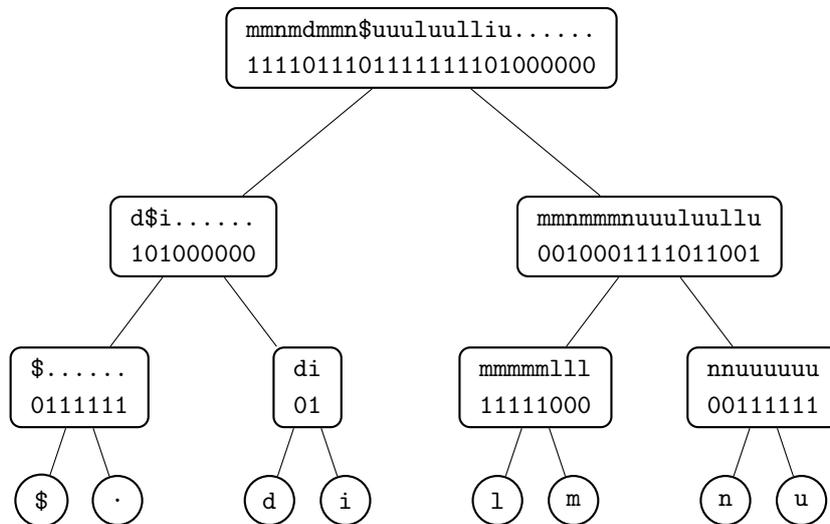


Abbildung 1: Wavelet-Baum des Strings $mnmndmmn\$uuuluulliu.....$

Aufgabe 6.3.

Berechnen Sie mit Hilfe von Algorithmus 27 im Skript das LCP-Array für den String $S = annasanannas\$$. Die Burrows-Wheeler-Transformierte von S ist $BWT = ssn\$nnannaaaa$.