

Text Indexing and Information Retrieval

Übungsblatt 7

Besprechung: 26.11.2018

Aufgabe 1 (Praxis)

- (a) Implementieren Sie den BWT-Algorithmus. Benutzen Sie hierzu bestehende Suffixarray-Konstruktionsalgorithmen (etwa aus vorherigen Übungsblättern oder öffentlichen Repositories).
- (b) Implementieren Sie die Rücktransformation, d. h., Ihr Algorithmus liest eine BWT ein und gibt den ursprünglichen Text aus. Die *einzige* Eingabe ist hierbei die BWT.
- (c) Komprimieren Sie die BWT mithilfe von (1) Run-length Kodierung und (2) Move-to-Front-Kodierung.

Aufgabe 2 (Theorie)

Erweitern Sie den Algorithmus für die Rücktransformation der BWT so, dass er auch das Suffix-Array für den Originaltext T berechnet. Kann dabei das LF-Array mit dem Suffix-Array überschrieben werden?

Aufgabe 3 (Theorie)

Beschreiben Sie einen Linearzeit-Algorithmus, der aus dem BW-transformierten Text den Originaltext *vorwärts* rekonstruiert.