

Übung zu Algorithmen auf Sequenzen Blatt 6

Ausgabe: 8. Januar 2015 **Besprechung:** 22. Januar 2015

Aufgabe 6.1

Sei F_6 der sechste Fibonacci-String (Länge 13). Berechne die BWT von F_6 .

Aufgabe 6.2

Transformiere den BWT-transformierten Text `rrt$bttooooo` zurück.

Aufgabe 6.3

Führe auf dem Text `ACATGATGATCTGATCA$` eine Rückwärtssuche (Backward Search) mit dem Muster `TGAT` durch. Erstelle dazu die BWT und alle weiteren notwendigen Hilfsarrays.

Aufgabe 6.4

Betrachte den Quellcode von bzip2 unter <http://www.bzip.org/downloads.html>. Echte Dateien, die man komprimieren möchte, haben nicht notwendigerweise ein Wächterzeichen am Ende. Wie wird dieses Problem von bzip2 gelöst, und welcher Algorithmus wird verwendet, um die BWT eines Blocks der Datei zu erzeugen?

Aufgabe 6.5

Da sich ein BWT-transformierter Text häufig besser komprimieren lässt, liegt es nahe, die Transformation mehrfach anzuwenden, also zu einem String s die Kompressionseigenschaften von s , $bwt(s)$, $bwt(bwt(s))$, usw. zu betrachten. Warum "bringt das nichts"?