

Algorithmen zur Sequenzanalyse

Wintersemester 2015/2016
Besprechung am 11.12.2015

Übungsblatt 4

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

Aufgabe 4.1.

Entwerfen Sie einen linearen Algorithmus (ohne den Trick mit der Variablen *lastdiff*), der den *lcp*-Intervallbaum einmal bottom-up durchläuft und dabei alle maximalen Repeats ausgibt.

Aufgabe 4.2.

Entwerfen Sie einen linearen Algorithmus, der den *lcp*-Intervallbaum einmal bottom-up durchläuft und dabei simultan alle maximalen und supermaximalen Repeats ausgibt.

Aufgabe 4.3.

Ein String ω ist ein *prefix tandem repeat* des Strings S wenn gilt:

- ω ist ein Präfix von S
- es gibt einen String u mit $uu = \omega$

Geben Sie einen linearen Algorithmus an, welcher das längste *prefix tandem repeat* von S findet.

Aufgabe 4.4.

Sei ω ein längstes Repeat. Zeigen Sie, dass ω ein supermaximales Repeat (und damit auch ein maximales Repeat) sein muss.

Aufgabe 4.5.

Zeichnen Sie den de Bruijn Graph für $k = 2$ über dem Alphabet $\Sigma = \{a, c, g, t\}$. Finden Sie einen Eulerkreis in dem Graph und geben Sie die korrespondierende nicht-zyklische de Bruijn Sequenz an. Wie viele supermaximale Repeats enthält die Sequenz?