

Algorithmen auf Sequenzen

Übungsblatt 10

Abgabetermin: Dienstag, 17.01.2017, 10 Uhr
(Via email an sophie.friedl@bio.ifi.lmu.de oder persönlich)

1. Aufgabe (Extended Suffix-Arrays)

Konstruieren Sie für das Wort $t = aaabaabab$ den zugehörige Extended Suffix-Array, d.h. den Suffix-Array A und die LCP-Tabelle L , und zeichnen Sie den zugehörigen LCP-Intervall-Baum für t (es gilt $a < b$).

2. Aufgabe (Maximale Repeats)

Entwerfen Sie einen möglichst effizienten Algorithmus zur Erkennung maximaler Repeats mit Hilfe eines Extended Suffix-Array. Dieser Algorithmus soll nur die LCP-Tabelle L zu dem zugehörigen Suffix-Array A verwenden. Range-Minimum-Queries und der LCP-Intervall-Baum dürfen hierbei nicht verwendet werden.

3. Aufgabe (Extended Suffix-Arrays)

Zeigen Sie, wie man die Tabellen P und N für den Extended Suffix-Array in Linearzeit konstruieren kann.

Hinweis: Der Algorithmus in Abbildung 4.8 im Skript zur Bestimmung von L^B kann hilfreich sein.