

9. Übung zur Vorlesung “Einführung in die Bioinformatik I, 1. Teil”

Wintersemester 2019/2020

Prof. Peter Dittrich, Emanuel Barth, Marcus Ludwig

Ausgabe: 18. Dezember 2019,
Abgabe: 08. Januar 2020 zu Beginn der Übung

Suffixbäume

Aufgabe 1 (5 Punkte):

1. Bestimmen Sie für den String `abrakadabra` und `abaababab` den jeweiligen komprimierten Suffixbaum. Ist der komprimierte Suffixbaum nicht eindeutig, geben Sie bitte alle an.
Hinweis: Es ist ausreichend, eine alternative Beschriftung in anderer Farbe anzubringen oder separat alternative Beschriftungen zu kennzeichnen.
2. Bestimmen Sie für die Strings `AGTT`, `TAC` und `GTAG` den komprimierten generalisierten¹ Suffixbaum.

Aufgabe 2 (5 Punkte): Wie kann man mit Hilfe eines Suffixbaumes in einem Text der Länge n den längsten Teilstring finden, der in diesem Text auch rückwärts vorkommt? In welcher asymptotischen Laufzeit ist dies möglich?

¹Der *generalisierte* Suffixbaum für k Strings S_1, \dots, S_k ist der Suffixbaum für $S_1\$1S_2\$2 \dots S_k\$k$, wobei $\$1, \$2, \dots, \$k$ paarweise verschiedene Zeichen sind, die nicht im verwendeten Alphabet vorkommen.