

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

Wintersemester 2020/2021

Prof. Peter Dittrich, Prof. Sebastian Böcker, Emanuel Barth, Marcus Ludwig

Ausgabe: 13. Januar 2021,  
Abgabe: 20. Januar 2021 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 generalisierten<sup>1</sup> 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?

---

<sup>1</sup>Der *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.