

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

Wintersemester 2019/2020

Prof. Peter Dittrich, Emanuel Barth, Marcus Ludwig

Ausgabe: 11. Dezember 2019,
Abgabe: 18. Dezember 2019 zu Beginn der Übung

Suffixbäume

Aufgabe 1 (5 Punkte): Bestimmen Sie für die Strings `abrakadabra`, `abaababab` und `rhabarber` den jeweiligen Suffixbaum.

Aufgabe 2 (5 Punkte): Geben Sie einen String der Länge n an, der die Anzahl innerer Knoten eines Suffixbaumes minimiert, und einen String, der sie maximiert. Geben Sie die Anzahl innerer Knoten im jeweiligen Suffixbaum an.

Aufgabe 3 (5 Punkte): Geben Sie einen Algorithmus an, der für einen vorhandenen Suffixbaum für einen String S bei Eingabe eines Patterns P die Position jedes Vorkommens von P in S ausgibt und eine Laufzeit von $O(m + k)$ benötigt, wobei $m = |P|$ und k die Anzahl der Vorkommen von P in S ist. Begründen Sie Korrektheit und Laufzeit Ihres Algorithmus.

Aufgabe 4 (5 Punkte): Geben Sie einen Algorithmus an, der unter Verwendung eines Suffixbaumes für einen gegebenen String S der Länge n in einer Laufzeit von $O(n)$ den längsten Teilstring P findet, der mindestens zweimal in S vorkommt. Begründen Sie Korrektheit und Laufzeit des Algorithmus.