

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2025/26

Prof. Dr. Jens Stoye

Leonard Bohnenkämper

Tutorien: Lennart Finke, Sofie Jans

<https://gi.cebitc.uni-bielefeld.de/teaching/2025winter/sai>

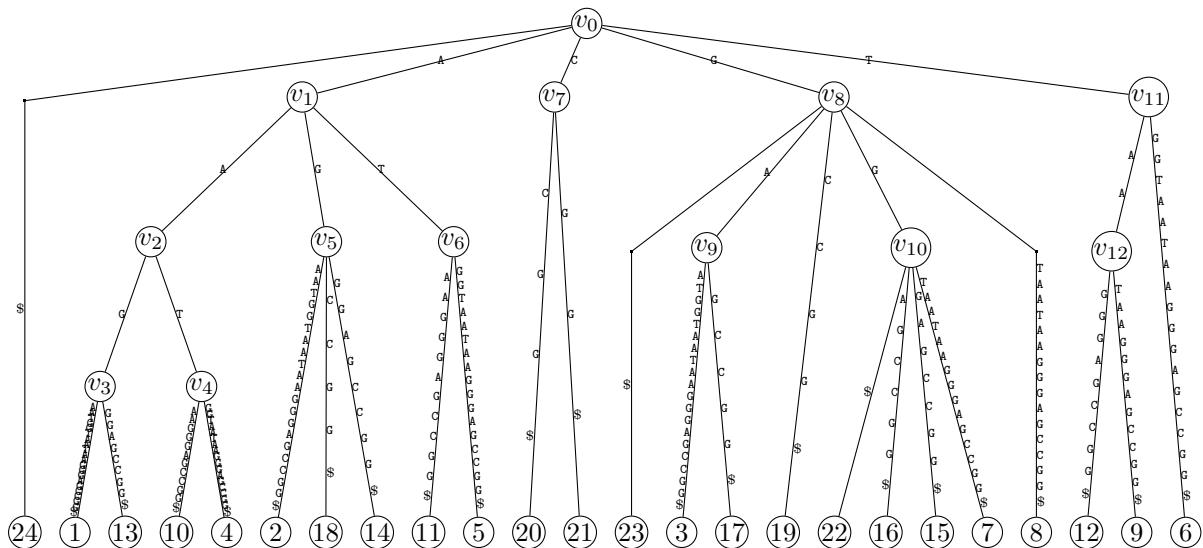
Übungsblatt 10 vom 15.01.2026

Abgabe bis 22.01.2026 bis 9:30 Uhr per Mail an den Tutor/die Tutorin

Aufgabe 1 (Suffix-Bäume und Suffix-Arrays)

(3 Punkte)

Betrachte folgenden Suffix-Baum einer Sequenz $s\$$.



1. Gib das Suffix-Array $\text{pos}(s\$)$ an.
2. Berechne das Array $\text{rank}(s\$)$.
3. Begründe kurz, warum dein Vorgehen aus den vorangegangenen Teilaufgaben korrekt ist.

Aufgabe 2 (Implementierung von rank und lcp)

(6 Punkte)

Vervollständige die Python-Implementierung im Übungsrepository¹.

1. Implementiere die Konstruktion des `rank`-Arrays aus `pos` in `construct_rank`.
2. Implementiere `longest_common_prefix_extension`, d.h. die Funktion die den längsten gemeinsamen Präfix von p_1 und p_2 berechnet, gegeben, dass die ersten l Zeichen bereits gleich sind.
3. Implementiere die Konstruktion des `lcp`-Arrays in `construct_lcp`.

Aufgabe 3 (Manber-Myers)

(8 Punkte)

Gegeben sei die Sequenz $s = \text{ATTATT}$.

1. Erstelle das Suffix-Array $\text{pos}(s\$)$ unter Verwendung des Manber-Myers-Algorithmus. Gib das Zwischenergebnis nach jeder Phase des Algorithmus' an.
2. Berechne die Arrays $\text{rank}(s\$)$ und $\text{lcp}(s\$)$.
3. Verwende Binärsuche im Suffix-Array $\text{pos}(s\$)$, um alle Vorkommen des Musters $p = \text{AT}$ in s zu finden. Gib alle Zwischenschritte an.

Aufgabe 4 (Manber-Myers – Laufzeit)

(3 Punkte)

Wie viele Phasen benötigt (im worst case) der Manber-Myers-Algorithmus für einen String der Länge n ?
Gib eine Beispieldsequenz an, bei der diese Anzahl Phasen erreicht wird.

¹https://gitlab.ub.uni-bielefeld.de/lbohnenkaemper/sqa-ex/-/tree/main/w12?ref_type=heads