

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2025/26

Prof. Dr. Jens Stoye

Leonard Bohnenkämper

Tutorien: Lennart Finke, Sofie Jans

<https://gi.cebitec.uni-bielefeld.de/teaching/2025winter/sa1>

Übungsblatt 12 vom 29.01.2026

Abgabe bis 05.02.2026 bis 9:30 Uhr per Mail an den Tutor/die Tutorin

Aufgabe 1 (MOVE it!)

(5 Punkte)

Unten findest du die Burrows-Wheeler Transformation vom letzten Zettel.

1. Gib die Zugehörige MOVE Datenstruktur an.
2. Führe eine exakte Textsuche in der MOVE Datenstruktur nach dem Pattern **GENI** durch. Wie oft kommt dieses im ursprünglichen Text vor? Gib MOVE und fast-forward Operationen explizit an.

\$	E
A	I
E	I
E	R
E	G
E	G
E	G
E	N
E	I
G	E
G	\$
G	R
H	I
I	N
I	N
I	R
I	T
L	A
N	E
N	E
N	E
R	L
R	H
R	E
R	E
T	R

Bonuspunkte auf der Rückseite.

Folgende Aufgaben sind freiwillig (z.B. zur Klausurvorbereitung) und zählen nicht in die Gesamtwertung, geben aber Bonuspunkte.

Aufgabe 2 (Wiederholung – Berechnung von Alignments)

(5 Bonuspunkte)

Gegeben seien die Sequenzen $x = \text{TACOCAT}$ und $y = \text{ACTA}$ sowie die Kosten-Funktion: Match = 0, Mismatch = 1, Indel = 2. Berechne die Alignments, die unter dem gegebenen Kostenschema sinnvoll sind, und begründe bei den übrigen, warum sie nicht sinnvoll sind.

1. Ein optimales globales Alignment von x und y .
2. Ein optimales semi-globales Alignment von x und y .
3. Ein optimales *free-end gap* Alignment von x und y .
4. Ein optimales lokales Alignment von x und y .

Aufgabe 3 (Wiederholung – Suffixbäume)

(5 Bonuspunkte)

Gegeben sei die Sequenzen $s = \text{BARBARBAR}$.

1. Berechne den Suffixbaum von s mit dem WOTD-Algorithmus.
2. Finde alle kürzesten einzigartigen Substrings von s mithilfe des Suffixbaums.
3. Finde alle maximalen Repeats von s mithilfe des Suffixbaums.
4. Gib das Suffixarray von s an.

Aufgabe 4 (Wiederholung – Multiple Alignments)

(5 Bonuspunkte)

Nutze in dieser Aufgabe das folgende Score-Schema: Matchscore 1, Mismatchscore -1 sowie Gapscore -1 . Betrachte folgendes Alignment:

$$A = \begin{pmatrix} \text{T} & \text{C} & \text{T} & \text{A} & \text{G} \\ \text{T} & - & \text{T} & - & \text{C} \\ - & \text{C} & \text{C} & - & - \end{pmatrix}$$

1. Berechne den Sum-Of-Pairs Score von A . Gib deinen Rechenweg mit an.
2. Betrachte das folgende progressive Alignment-Schema und berechne das Alignment B für $s_4 = \text{CAG}$. Was kannst du über die Optimalität deines Alignments sagen?

