

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2024

Prof. Dr. Jens Stoye · Leonard Bohnenkämper

<https://gi.cebitec.uni-bielefeld.de/teaching/2024winter/sa>

Übungsblatt 12 vom 16.01.2025

Abgabe am 22.01.2025 bis 10:00 Uhr (morgens)

Aufgabe 1 (Center-Star-Approximation)

(5 Punkte)

Gegeben sind die Sequenzen $s_1 = \text{TAGC}$, $s_2 = \text{TTGC}$ und $s_3 = \text{AGC}$. Benutze für deine folgenden Berechnungen Einheitskosten.

1. Berechne die *Center-Sequenz* s_c .
2. Erstelle das multiple Alignment A_c und gib seine Sum-of-Pairs-Kosten an.
3. Was kannst du mithilfe deiner Lösung über die optimalen Alignment-Kosten sagen?
4. Beschreibe in eigenen Worten die Laufzeit- und Speicherplatzkomplexität der Center-Star-Approximation. Unterscheide dabei zuerst die einzelnen Phasen und erkläre dann das Gesamtergebnis.

Aufgabe 2 (Divide-and-Conquer-Alignment)

(5 Punkte)

Gegeben sind die Sequenzen $s_1 = \text{GAG}$, $s_2 = \text{TG}$ und $s_3 = \text{TA}$. Benutze für deine Berechnungen Einheitskosten.

1. Was ist der Unterschied zwischen einem optimalen und einem C-optimalen Schnitt?
2. Erstelle die Zusatzkostenmatrizen für die drei Sequenzen.
3. Gib alle C-optimalen Schnitte an.
4. Nenne alle möglichen optimalen multiplen Alignments, die dir die C-optimalen Schnitte angeben. Gib auch ihre Kosten an.
5. Wieso verkleinert sich der Suchraum von $\mathcal{O}(n^k 2^k)$ beim Sum-of-Pairs-optimalen multiplen Alignment auf $\mathcal{O}(n^{k-1})$ beim Divide-and-Conquer-Alignment?

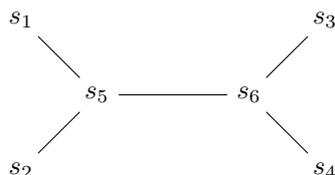
Aufgabe 3 (Sum-of-Pairs und Baumalignment)

(5 Punkte)

Gegeben ist folgendes Alignment

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

Berechne die Sum-Of-Pairs Kosten von s_1, s_2, s_3, s_4 dieses Alignments und die Baumalignment-Kosten von $s_1, s_2, s_3, s_4, s_5, s_6$ dieses Alignments unter dem unten dargestellten Baum. Nutze jeweils Einheitskosten.



Finde ein Alignment von s_1, s_2, s_3, s_4 mit geringeren Sum-Of-Pairs-Kosten.

Aufgabe 4 (Fitch-Algorithmus)

(5 Punkte)

Für die innere Minimierung der Formel für das Baumalignment kann der Fitch-Algorithmus verwendet werden. Finde eine optimale Beschriftung der inneren Knoten des Baums auf der Rückseite unter Einheitskosten. Trage auch Zwischenergebnisse in den Baum ein.

