
Algorithmische Bioinformatik II

Abgabetermin: Freitag, den 5. Dezember, 9⁰⁰ Uhr in Moodle

Hausaufgabe 1

Verwende den Algorithmus von Carrillo und Lipman zur Berechnung eines Sequenzen-Alignments zwischen zwei Sequenzen $s = ACAT$ und $t = TAGAT$. Hierzu sind für das Distanzmaß die **Gap-Kosten** von 3 und **Mismatch-Kosten** von 2 zu verwenden. Die **globale obere Schranke** für die Distanz von s und t ist mit 8 vorgegeben.

Hinweis: In der Vorlesung wurde dies für 3 oder mehr Sequenzen erläutert, natürlich funktioniert das Verfahren auch mit nur 2 Sequenzen.

Gib die Präfix-Matrix P und die Suffix-Matrix S sowie die kombinierte **Präfix-/Suffix-Matrix** $P + S$ an und **markiere alle Zellen**, die in den **Heap** aufgenommen wurden. Gib dabei ebenfalls die Berechnung der verwendeten **obere Schranke** im Relevanz-Test für das Sequenzpaar (s, t) an.

Hausaufgabe 2

Betrachte die Sequenzen $s_1 = AGACA$, $s_2 = GACAC$ und $s_3 = ACCA$. Berechne die C -optimalen Schnittpositionen mit Respekt zu $c_1 = 1$ sowie zu $c_1 = 2$ und die daraus resultierenden mehrfachen Alignments gemäß des Divide-and-Conquer-Alignment-Algorithmus, wobei nach der ersten Rekursion bereits jeweils ein optimales Alignment für die jeweiligen Präfixe bzw. Suffixe berechnet wird. Die zugrunde liegende Kostenfunktion für das SP-Distanzmaß sei $w(a, a) = 0$ und $w(a, b) = 1$ für alle $a \neq b \in \bar{\Sigma}$.

Tutoraufgabe 3 (Vorbereitung bis zum 03.12.25)

Beweise oder widerlege: Ein optimales mehrfaches Sequenzen-Alignment bezüglich des Sum-of-Pairs-Maßes induziert mindestens ein optimales paarweises Sequenzen-Alignment.