
Algorithmische Bioinformatik II

Abgabetermin: Freitag, den 9. Dezember, 9⁰⁰ Uhr in Moodle

Tutoraufgabe 1 (Vorbereitung bis zum 07.12.22)

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

Hausaufgabe 2

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 3

Sei $w : \bar{\Sigma}^2 \rightarrow \mathbb{R}_+$ eine Kostenfunktion für ein Distanzmaß mit $w(-, -) := 0$, die die Eigenschaften einer Metrik (Definitheit, Symmetrie und Dreiecksungleichung) erfüllt. Sei weiter $w' : \bar{\Sigma}_0^k \rightarrow \mathbb{R}_+$ die zugehörige Consensus-Kostenfunktion für ein Distanzmaß eines k -fachen Sequenzen-Alignments.

Zeige, welche Eigenschaften einer Metrik (Definitheit, Symmetrie und Dreiecksungleichung) von w' erfüllt sind und welche nicht (Gegenbeispiel).