

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2018

Dr. Daniel Dörr

<https://gi.cebitec.uni-bielefeld.de/teaching/2018summer/sa>

Übungsblatt 5 vom 08.05.2018

Abgabe 17.05.2018

Aufgabe 1 (Affine Gapkosten (Wdh. Aufgabe 4.2))

(5 Punkte)

1. Warum sollten die Kosten für eine *gap extension* nicht höher als die Kosten für ein *gap open* gewählt werden, also $e \leq d$?
2. Zeige, dass affine Gapkosten subadditiv sind.
3. Berechne ein optimales globales Alignment mit affinen Gapkosten von den Sequenzen $x = \text{AACTGG}$ und $y = \text{AAGG}$ effizient mit Hilfe des Gotoh-Algorithmus (berechne die Matrizen S , H und V) und gib dessen Gesamtscore an. Verwende dabei: Score für Match = 2, Score für Mismatch = 0, Kosten für Gap-open $d = 1$, sowie Kosten für Gap-extension $e = 0.5$.

Aufgabe 2 (Scores und Kosten)

(4 Punkte)

Gegeben seien Sequenzen $x = \text{BABABB}$ und $y = \text{BABBA}$.

1. Berechne alle optimalen globalen Alignments der Sequenzen x und y unter Standard-Editkosten ($\mathcal{S}_c = \mathcal{I} = \mathcal{D} = 1$ und $\mathcal{C} = 0$)
2. Berechne ein optimales globales Alignment der Sequenzen x und y mit Scores $\mathcal{S}_c = \mathcal{I} = \mathcal{D} = 0$ und $\mathcal{C} = 1$, das für die Kosten in 1.) nicht optimal wäre.
3. In welchen Fällen sollte man Scores statt Kosten verwenden?
4. In welchen Fällen könnte es notwendig sein, Kosten statt Scores zu verwenden? (1 Bonuspunkt)

Aufgabe 3 (Suboptimale Alignments)

(4 Punkte)

Gegeben seien Sequenzen $x = \text{GATC}$, $y = \text{ACGCTCC}$, berechne ein lokales Alignment mit dem Smith-Waterman-Algorithmus und das erste nicht überlappende, suboptimale Alignment nach *Waterman-Eggert*. Verwende dazu Scores mit $\mathcal{C} = 2$, $\mathcal{S}_c = 0$ und $\mathcal{I}_c = \mathcal{D}_c = -1$.