

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2019

Dr. Daniel Dörr · Michel T. Henrichs

<https://gi.cebitec.uni-bielefeld.de/teaching/2019summer/sa>

Übungsblatt 5 vom 6.5.2019

Abgabe am 13.5.2019 bis 12:00 Uhr (mittags)

Aufgabe 1 (Berechnung von Alignments)

(6 Punkte)

Gegeben seien die Sequenzen $x = \text{AGTCCCTA}$ und $y = \text{CGTACCGGTAA}$ sowie die Score-Funktion: Match = 4, Mismatch = -3, Indel = -2.

1. Berechne alle optimalen globalen Alignments von x und y .
2. Berechne alle optimalen *free-end gap* Alignments von x und y .
3. Berechne alle optimalen lokalen Alignments von x und y .

Aufgabe 2 (Gotoh)

(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 $x = \text{CTATGTGA}$ und $y = \text{CTGA}$ effizient mit Hilfe des Gotoh-Algorithmus und gib dessen Gesamtscore an. Verwende dabei: Score für Match = 2, Score für Mismatch = 0, Kosten für Gap-open $d = 2$, sowie Kosten für Gap-extension $e = 0.5$.

Aufgabe 3 (Suboptimale Alignments)

(5 Punkte)

1. Beschreibe in eigenen Worten, was man unter überlappenden Alignments versteht.
2. Warum ist man bei der Bestimmung von suboptimalen Alignments in erster Linie an nichtüberlappenden Alignments interessiert? Welche Probleme möchte man vermeiden?
3. Beschreibe kurz die Funktionsweise des Waterman-Eggert-Algorithmus. Durch welchen Trick kann man in der Praxis die Laufzeit verkürzen?
4. Gegeben seien $x = \text{GTAA}$ und $y = \text{GCTA}$. Berechne ein lokales Alignment mit dem Smith-Waterman-Algorithmus und das erste nicht überlappende, suboptimale Alignment nach Waterman-Eggert. Verwende dazu: Score für Match = 3, Score für Mismatch = 1, Kosten für Indels = 1.