

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2018

Dr. Daniel Dörr

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

Übungsblatt 4 vom 03.05.2018

Abgabe bis spätestens Freitag, 11.05.2018 23:59 GMT+2 via Email oder Postfach
(U10-151)

Aufgabe 1 (Globales Alignment)

(4 Punkte)

Wir definieren die folgenden Scores für die Editoperationen Insertion und Deletion: $\mathcal{I} = \mathcal{D} = -1$. Für Copy (\mathcal{C}) und Substitution (\mathcal{S}) gilt die Scorematrix \mathcal{S} :

	A	B	C
A	+2	-2	-2
B	-2	+2	-1
C	-2	-1	+2

1. Zeichne den globalen Alignmentgraphen für $x = \text{AB}$ und $y = \text{ACB}$ mit *allen* Kanten und Knoten.
2. Füge deiner Zeichnung die Kantenbeschriftungen λ und die Kantengewichte w hinzu.
3. Berechne die Knotenwerte $S(v)$.
4. Gib ein optimales globales Alignment an und markiere dessen Pfad im Graphen.

Aufgabe 2 (Affine Gapkosten)

(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 3 (Lokale und globale Alignment-Algorithmen)

(6 Punkte)

1. Erkläre den Unterschied zwischen einem lokalen und einem globalen Alignment. Nenne jeweils einen Algorithmus, der die Alignments berechnet.
2. Warum verwendet man einen eigenen Algorithmus zur Berechnung von lokalen Alignments, anstatt die globalen Alignments aller Paare von Substrings zu berechnen und sich das beste Alignment auszusuchen?
3. Berechne das lokale Alignment der beiden Strings $x = \text{ATGGCAG}$ und $y = \text{GTA}$. Verwende folgende Scores: $\mathcal{I}_c = \mathcal{D}_c = -2$, $\mathcal{C} = 2$ und $\mathcal{S} = -1$. Gib das Alignment an.
4. Für welche biologischen Fragestellungen würdest Du welchen Algorithmus anwenden?

Aufgabe 4 (Lokales und free-end-gap-Alignment)

(2 Punkte)

Beschreibe mit eigenen Worten, wie sich die Rekursionsschemata vom lokalen und vom free-end-gap-Alignment voneinander unterscheiden. Was bedeuten die Unterschiede bei der Berechnung des Alignments?