

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2022

Prof. Dr. Jens Stoye · Dr. Marília D. V. Braga

<https://gi.cebitec.uni-bielefeld.de/teaching/2022summer/sa>

Übungsblatt 5 vom 5.5.2022

Abgabe am 12.5.2022 bis 12:00 Uhr (mittags)

Aufgabe 1 (Berechnung von Alignments)

(6 Punkte)

Gegeben seien die Sequenzen $x = \text{ATCCGGA}$ und $y = \text{TCGGTCTCG}$ sowie die folgende Score-Funktion:

$$\text{MATCH} = 4 ; \text{ MISMATCH} = -3 ; \text{ INDEL} = -2$$

1. Berechne die Matrix für *free-end gap* Alignment und gib alle optimalen Alignments an.
2. Berechne die Matrix für lokales Alignment und gib alle optimalen Alignments an.

Aufgabe 2 (Suboptimale lokale Alignments)

(6 Punkte)

1. Warum ist man bei der Bestimmung von suboptimalen Alignments in erster Linie an *nichtüberlappenden* Alignments interessiert? Welche Probleme möchte man vermeiden?
2. Gegeben seien die Sequenzen $x = \text{AACTAGAG}$ und $y = \text{TCTGCAAT}$ sowie die folgende Score-Funktion:

$$\text{MATCH} = 3 ; \text{ MISMATCH} = -2 ; \text{ INDEL} = -2$$

- (a) Berechne die Alignment-Matrix für lokales Alignment mit dem Smith-Waterman-Algorithmus und gib ein optimales Alignment an.
- (b) Berechne die Matrix für nichtüberlappende, suboptimale lokale Alignments nach *Waterman-Eggert* und gib das erste suboptimale Alignment an.
- (c) Aktualisiere die Matrix und gib das nächste suboptimale Alignment an.

Aufgabe 3 (Affine Gapkosten)

(8 Punkte)

1. Warum sollten bei affinen Gapkosten die Gap-open-Kosten d nicht niedriger gewählt als die Gap-extension-Kosten e , also $d \geq e$?
2. Zeige, dass affine Gapkosten subadditiv sind, also dass gilt: $g(\ell_1 + \ell_2) \leq g(\ell_1) + g(\ell_2)$
3. Berechne die Matrizen \mathbf{S} , \mathbf{V} und \mathbf{H} der Sequenzen $x = \text{ACTTT}$ und $y = \text{AAGGCT}$ für globales Alignment mit affinen Gapkosten mit Hilfe des Gotoh-Algorithmus.

Verwende dabei

$$\begin{cases} \text{Score für MATCH} & = 3 \\ \text{Score für MISMATCH} & = -2 \end{cases} \quad \begin{cases} \text{Kosten für GAP-OPEN :} & d = 3 \\ \text{Kosten für GAP-EXTENSION :} & e = 1 \end{cases}$$

Gib alle optimalen Alignments an.

Du kannst diese Seite für deine Lösung verwenden:

Aufgabe 3: $\left\{ \begin{array}{l} \text{Score: MATCH} = 3, \text{ MISMATCH} = -2 \\ \text{Kosten: GAP-OPEN} : d = 3, \text{ GAP-EXTENSION} : e = 1 \end{array} \right.$

V	ε	A	A	G	G	C	T
ε							
A							
C							
T							
T							
T							

S	ε	A	A	G	G	C	T
ε							
A							
C							
T							
T							
T							

H	ε	A	A	G	G	C	T
ε							
A							
C							
T							
T							
T							