

Übungen zur Vorlesung Sequenzanalyse I

Universität Bielefeld, WiSe 2009/2010

Prof. Dr. Jens Stoye · Dipl.-Inform. Nils Hoffmann

<http://wiki.techfak.uni-bielefeld.de/gi/GILectures/2009winter/SequenzAnalyse>

Blatt 8 vom 3.12.2009

Abgabe in einer Woche vor Beginn der Vorlesung.

Aufgabe 1 Edit-Distanz und Alignment

(2 Punkte)

Gegeben sind die Sequenzen $x = \text{TTAGACA}$ und $y = \text{TATACG}$.

1. Berechne die Edit-Distanz für Einheitskosten von x und y mit Hilfe der Edit-Matrix.
2. Gib eine optimale Edit-Sequenz und das dazugehörige optimale Alignment an.

Aufgabe 2 Die Anzahl Globaler Alignments

(3 Punkte)

Erstelle ein Programm in einer Programmiersprache deiner Wahl, welches

1. $N(m, n)$ rekursiv berechnet,
2. $N(n, n)$ approximativ berechnet,
3. $N'(m, n)$ exakt berechnet und
4. $N'(m, n)$ approximativ berechnet.

Gib die Ergebnisse für $n = m \in \{1, 5, 10, 100\}$ grafisch an und vergleiche diese. Diskutiere den Unterschied zwischen $N(m, n)$ und $N'(m, n)$.

Aufgabe 3 Alignmentklassen

(2 Punkte)

Gegeben sind die Sequenzen $x = \text{AB}$ und $y = \text{CD}$.

1. Bestimme alle 13 möglichen Alignments von x und y .
2. Welche sechs davon sind *echt* verschieden?

Aufgabe 4 Alignmentgraph

(3 Punkte)

Wir definieren die folgenden Scores für die Editoperationen Insertion und Deletion: $\mathcal{I}_c = \mathcal{D}_c = 0$. Für Copy (\mathcal{C}) und Substitution (\mathcal{S}) gilt die Scorematrix $\mathcal{S}_{a,c}$:

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

1. Zeichne den Alignmentgraphen für $x = \text{AC}$ und $y = \text{AB}$ mit allen Kanten und Knoten.
2. Füge deiner Zeichnung die Kantenbeschriftungen λ und die Kantengewichte w hinzu.
3. Berechne die Knotenwerte $s(v)$ und zeichne den Backtracing-Pfad eines optimalen Alignments ein.