

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2014/2015

Dr. Roland Wittler · Nina Luhmann · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2014winter/SequenzAnalyse>

Übungsblatt 3 vom 28.10.2014

Abgabe in einer Woche vor Beginn der Vorlesung.

Aufgabe 1 (Globales Alignment)

(4 Punkte)

1. Welche Zeichen dürfen in einer Alignmentsspalte bei gegebenem Alphabet Σ vorkommen? Welche Kombination ist nicht erlaubt?
2. Gib das Alignment A an, welches der Editsequenz $CCCDL_ZCCDDD$ bei gegebenem String $x = BRASILIEN$ entspricht.
3. Berechne die Projektion $y = \pi_{\{2\}}(A)$.
4. Kann es ein weiteres globales Alignment $A' \neq A$ der Sequenzen x und y geben? Wenn ja, gib eines an. Wenn nein, begründe warum es keines geben kann.

Aufgabe 2 (Alignment Graph)

(4 Punkte)

Wir definieren die folgenden Scores für die Editoperationen Insertion und Deletion: $\mathcal{I} = \mathcal{D} = -2$. Für Copy (\mathcal{C}) und Substitution (\mathcal{S}) gilt die Scorematrix $\mathcal{S}_{a,c}$:

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

1. Zeichne den globalen Alignmentgraphen für $x = ABC$ und $y = AB$ mit *allen* Knoten, Kanten, Kantenbeschriftungen λ und Kantengewichten w .
2. Berechne die Knotenwerte $S(v)$.
3. Gib ein optimales globales Alignment an und markiere dessen Pfad im Graphen.

Aufgabe 3 (Anzahl Alignments)

(4 Punkte)

Theorem B.1 auf Seite 139 im Skript nennt eine rekursive Berechnungsvorschrift für die Anzahl an möglichen globalen Alignments von zwei Sequenzen mit den Längen m und n :

$$N(m, n) = N(m - 1, n - 1) + N(m - 1, n) + N(m, n - 1)$$

1. Erläutere die Formel. (Tipp: Lies den Beweis des Theorems.)
2. Implementiere die Rekursion.
3. Berechne $N(n, n)$ für $n = 1, \dots, 20$ und stelle das Ergebnis als Graph dar – mit n auf der X-Achse und $N(n, n)$ auf der Y-Achse.