

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2017

Prof. Dr. Jens Stoye · M.Sc. Tizian Schulz

<https://gi.cebitec.uni-bielefeld.de/teaching/2017summer/sa>

Übungsblatt 4 vom 23.05.2017

Abgabe in einer Woche vor Beginn der Vorlesung.

Aufgabe 1 (Dynamische Programmierung)

(5 Punkte)

Schreibe ein Programm, das als Eingabe zwei Strings x und y bekommt und deren Edit-Distanz ausgibt. Die Edit-Distanz soll mittels dynamischer Programmierung bestimmt werden. Sprich mit deinem Tutor ab, welche Programmiersprachen du hierfür verwenden darfst.

Aufgabe 2 (Die Maximal-Matches-Distanz)

(6 Punkte)

Gegeben seien die beiden Strings $x = \text{ATATCGGTAATC}$ und $y = \text{CGATAGTAT}$.

1. Berechne die Links-Rechts- und die Rechts-Links-Partition von x in Bezug zu y und gib jeweils ihren Wert an.
2. Wie ist die Maximal-Matches-Distanz von x zu y definiert?
3. Zeige an einem von dir ausgedachten Beispiel, dass die Maximal-Matches-Distanz keine Metrik ist.

Aufgabe 3 (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	G	C
A	+2	-2	-2
G	-2	+2	-1
C	-2	-1	+2

1. Zeichne den globalen Alignmentgraphen für $x = \text{CG}$ und $y = \text{CA}$ 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.