

# Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2015/2016

Prof. Dr. Jens Stoye · M.Sc. Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2015winter/SequenzAnalyse>

## Übungsblatt 4 vom 17.11.2015

Abgabe in *zwei* Wochen vor Beginn der Vorlesung

### Aufgabe 1 (Berechnung von Alignments)

(6 Punkte)

Gegeben seien die Sequenzen  $x = \text{AGTCCCTA}$  und  $y = \text{CGTACCGGTAA}$  sowie die Score-Funktion: Match = 4, Mismatch = -3, Indel = -2.

1. Berechne alle optimalen globalen Alignments von  $x$  und  $y$ .
2. Berechne alle optimalen *free-end gap* Alignments von  $x$  und  $y$ .
3. Berechne alle optimalen lokalen Alignments von  $x$  und  $y$ .

### Aufgabe 2 (Lokale und globale Alignment-Algorithmen)

(4 Punkte)

1. Erkläre den Unterschied zwischen einem lokalen und einem globalen Alignment. Nenne jeweils einen Algorithmus, der die Alignments berechnet.
2. Für welche biologischen Fragestellungen würdest Du welchen Algorithmus anwenden?
3. 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?

### Aufgabe 3 (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?

### Aufgabe 4 (Affine Gapkosten)

(5 Punkte)

In der Vorlesung haben wir affine Gapkosten kennengelernt.

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  $x = \text{GTCC}$  und  $y = \text{GTTGCC}$  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 5 (Suboptimale Alignments)

(4 Punkte)

1. Beschreibe in eigenen Worten, was man unter *überlappenden* Alignments versteht.
2. Warum ist man bei der Bestimmung von suboptimalen Alignments in erster Linie an *nichtüberlappenden* Alignments interessiert?
3. Beschreibe kurz die Funktionsweise des Waterman-Eggert-Algorithmus. Durch welchen Trick kann man in der Praxis die Laufzeit verkürzen?