

Übungen zur Vorlesung Sequenzanalyse II

Universität Bielefeld, SoSe 2007

Dr. Sven Rahmann · Dipl.-Inf. Peter Husemann · Dipl.-Biol. Constantin Bannert

<http://gi.cebitec.uni-bielefeld.de/teaching/2007summer/sequenzanalyse/>

Blatt 9 vom 15.06.2007

Abgabe am 22.06.2007 vor der Vorlesung um 8:30 in H14

Aufgabe 1 Gib einen effizienten Algorithmus an, mit dem man ein optimales paarweises Alignment (bzw. nur seine Kosten oder Score) finden kann, das durch einen vorgegebenen Punkt (i, j) im Alignmentgraphen führt. Nenne zwei bis drei Anwendungen dafür.

Aufgabe 2 Wähle zwei kurze Sequenzen und eine einfache Kostenfunktion (z.B. Edit-Distanz). Berechne die optimale paarweisen Alignmentkosten der Sequenzen. Berechne dann für jeden Punkt (i, j) des Alignmentgraphen die zusätzlichen Kosten des optimalen Alignments durch (i, j) .

Aufgabe 3 Diskutiere die folgenden zwei Standpunkte:

1. Multiples Alignment mit der sum-of-pairs score ist exakt nicht lösbar. Die sum-of-pairs score hat andererseits keine gesicherte biologische Grundlage. Also ist es kein Fehler, eine biologisch inspirierte schnelle Heuristik zu verwenden.
2. Um überhaupt herauszufinden, wie biologisch relevant eine bestimmte scoring-Funktion ist, muss man erst in der Lage sein, diese exakt zu optimieren.

Aufgabe 4 Informiere dich über BALiBASE 3. Was enthält es, wofür ist das gut, etc.?