

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 11 vom 06.07.2007

Abgabe am 13.07.2007 vor der Vorlesung um 8:30 in H14

Die folgenden Aufgaben können als Probeklausuraufgaben aufgefasst werden.

Aufgabe 1 Gegeben sei ein kurzer String (z.B., $s = bcabcb$). Visualisiere den WOTD-Algorithmus an diesem Beispiel. Wie wird der Baum im Speicher dargestellt?

Aufgabe 2 Gegeben sei ein kurzer String (z.B., $s = bcabcb$). Visualisiere die einzelnen Phasen des Manber-Myers Algorithmus.

Aufgabe 3 Welche RNA-Sekundärstrukturelemente gibt es? Zeichne jeweils eine Skizze.

Aufgabe 4 Gib die Rekurrenz des Nussinov-Algorithmus an. Finde eine Struktur mit den meisten Basenpaaren von ACCUUAGAGUG (mit mindestens 3 ungepaarten Basen in einer Hairpin-Loop).

Aufgabe 5 Gegeben seien zwei kurze Sequenzen, z.B. $s = ACGTT$ und $t = CATGT$. Berechne die Edit-Matrix vorwärts und rückwärts und ermittle so die zusätzliche Kostenmatrix.

Aufgabe 6 Wozu dient parametrisches Alignment?

Aufgabe 7 Beschreibe den ray search Algorithmus (Text, kein Pseudocode).

Aufgabe 8 Begründe, warum die Region des Parameterraumes, in dem ein bestimmtes Alignment optimal ist, ein konvexes Polygon ist.

Aufgabe 9 Wozu dient längennormalisiertes lokales Alignment? Warum ist die Forderung nach einer Mindestlänge, bzw. einer zusätzlichen virtuellen Länge L , sinnvoll?

Aufgabe 10 Gegeben sei ein Alignment und eine Scoring-Funktion. Berechne den sum-of-pairs Score dieses Alignments.

Aufgabe 11 Gegeben sei ein Baum und ein Alignment, inkl. Sequenzen an den inneren Knoten. Berechne den tree-score dieses Alignments.

Aufgabe 12 Gegeben seien 3 oder 4 kurze Sequenzen. Berechne alle optimalen paarweisen Alignments. Gegeben sei ferner ein (suboptimales) multiples Alignment A . Berechne die Carillo-Lipman-Schranken für die paarweisen Alignments und visualisiere in jeder paarweisen Projektionsebene die zu berechnende Region.

Aufgabe 13 Zeige: Wenn alle optimalen paarweisen Alignments miteinander kompatibel sind, ist das optimale sum-of-pairs multiple Alignment genau das, das man aus den optimalen paarweisen Alignments zusammensetzen kann.

Aufgabe 14 Gegeben seien 3 bis 4 Sequenzen, z.B. $s_1 = ACGT$, $s_2 = CGT$, $s_3 = ACT$, $s_4 = ACTG$. Welche Sequenz ist bei der center star Methode das Zentrum? Wie sieht das center star Alignment aus? Findest du ein besseres (bzgl. sum-of-pair scoring)?

Aufgabe 15 Erkläre zwei grundsätzlich verschiedene Möglichkeiten, beim divide and conquer Alignment die Schnittpunkte zu bestimmen.

Aufgabe 16 Warum wählt man beim divide and conquer Alignment die Schnittpunkte häufig so, dass die erste Sequenz in zwei gleich große Hälften geteilt wird?

Aufgabe 17 Wie bestimmt man die Reihenfolge, in der bei progressivem Alignment die Sequenzen bzw. Teilalignments aligniert werden?

Aufgabe 18 Beschreibe das Konsistenz-Problem bei segmentbasiertem multiplen Alignment.

Aufgabe 19 Laufzeit und Speicherbedarf der in der Vorlesung behandelten Algorithmen?