

Übungen zur Vorlesung Sequenzanalyse II

Universität Bielefeld, SoSe 2012

Dr. Alexander Sczyrba · Nina Luhmann · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2012summer/SequenzAnalyse>

Übungsblatt 10 vom 15.06.2012

Abgabe in einer Woche vor Beginn der Vorlesung.

Bitte gib auch den Namen deines Tutors an.

Aufgabe 1 (Multiples Alignment von drei Sequenzen)

(6 Punkte)

Beschreibe, wie der Needleman-Wunsch-Algorithmus für paarweises Alignment verallgemeinert wird, um ein direktes Alignment von drei Sequenzen zu berechnen. Verwende die Standard-Editdistanz und

1. gib die Rekursion explizit an,
2. implementiere den Algorithmus in einer Programmiersprache deiner Wahl,
3. gib die Kosten eines optimalen Alignments von $a = \text{CTGT}$, $b = \text{AGT}$ und $c = \text{CAG}$ an.

Aufgabe 2 (Laufzeit-Abschätzung)

(4 Punkte)

Gegeben sind k Sequenzen der Länge $n = 1000$. Die Berechnung eines exakten paarweisen Alignments mittels *dynamic programming* dauert ca. zwei Sekunden. Ein Tag hat 86400 Sekunden und ein Jahr im Durchschnitt 365,2425 Tage.

1. Wie viele paarweise Alignments dieser Form kann man berechnen, bevor in 5 Milliarden Jahren die Sonne zu einem roten Riesen wird?
2. Wie viele Sequenzen kann man mit einem direkten multiplen Alignment ($\mathcal{O}(2^k \cdot n^k)$) in dieser Zeit alignieren?

Aufgabe 3 (Altschul-Gapkosten)

(3 Punkte)

1. Beschreibe die von Altschul vorgeschlagene Approximation affiner Gapkosten durch quasi-natürliche (Altschul-) Gapkosten.
2. Gib ein beispielhaftes multiples Sequenzalignment an, bei welchem sich die natürlichen Gapkosten und die Altschul-Gapkosten unterscheiden.