

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2012/2013

Prof. Dr. Jens Stoye · Nina Luhmann · Linda Sundermann

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

Übungsblatt 10 vom 18.01.2013

Abgabe in einer Woche vor Beginn der Vorlesung.

Aufgabe 1 (Laufzeit-Abschätzung)

(6 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 solcher Sequenzen kann man mit einem direkten multiplen Alignment ($\mathcal{O}(2^k \cdot n^k)$) in dieser Zeit alignieren?
3. Erkläre die Laufzeit des direkten multiplen Alignments.

Aufgabe 2 (Center-Star Approximation)

(5 Punkte)

Gegeben sind die Sequenzen $s_1 = GACTTGC$, $s_2 = ACTGG$ und $s_3 = GTTGC$. Benutze für deine folgenden Berechnungen Einheitskosten.

1. Berechne die *Center Sequenz* s_c . Du musst die paarweisen Alignments nicht explizit mit Matrizen berechnen.
2. Erstelle das multiple Alignment A_c und gebe seine Kosten an.
3. Beschreibe in eigenen Worten die Laufzeit- und Speicherkomplexität der Center-Star Approximation. Unterscheide dabei zuerst die einzelnen Phasen und erkläre dann das Gesamtergebnis.

Aufgabe 3 (Divide-and-Conquer Alignment)

(3 Punkte)

Gegeben sind die Sequenzen $s_1 = CTC$, $s_2 = AC$ und $s_3 = AT$. Benutze für deine folgenden Berechnungen Einheitskosten.

1. Was ist der Unterschied zwischen einem optimalen und einem C-optimalen Schnitt?
2. Erstelle die Zusatzkostenmatrizen für die drei Sequenzen.

Folgende Aufgabenstellungen werden aus der Wertung genommen:

3. Gib alle C-optimalen Schnitte an.
4. Nenne alle möglichen optimalen multiplen Alignments, die dir die C-optimalen Schnitte angeben. Gib auch ihre Kosten an.
5. Wieso verkleinert sich der Suchraum von $\mathcal{O}(n^k 2^k)$ beim Sum-of-Pairs-optimalen multiplen Alignment auf $\mathcal{O}(n^{k-1})$ beim Divide-and-Conquer Alignment?

Aufgabe 4 (Klausurfragen)

(2 Punkte)

Formuliere zwei Fragen zu bisherigen Themen, die dir nicht ganz klar sind. (Diese Fragen könnten im letzten Tutorium besprochen werden. Nutzt es als Klausurvorbereitung. (Wenn du keine Fragen mehr hast, schreibe zwei Fragen auf, die du dir gut in der Klausur vorstellen könntest.))