

Foundations of Sequence Analysis
Winter 2005/2006

Exercises

Übung 6, Besprechung am 12.12.2004 (14-16h in E0-160)
bzw. 08.12.2004 (14-16h in C01-148 und E01-108).

1. Lokales Alignment.

Gegeben sind zwei Sequenzen $\mathbf{u} = \text{TCCATACT}$ und $\mathbf{v} = \text{ACACAACACG}$. Berechnen Sie $loc_{\sigma}(\mathbf{u}, \mathbf{v})$ und das lokale optimale Alignment (mit Stift und Papier) mit Hilfe des Smith-Waterman Algorithmus, wobei die Score-Funktion wie folgt gegeben ist:

$$\sigma(\alpha \rightarrow \beta) = \begin{cases} +2 & \text{if } \alpha, \beta \in \mathcal{A} \text{ and } \alpha = \beta \\ -3 & \text{if } \alpha, \beta \in \mathcal{A} \text{ and } \alpha \neq \beta \\ -2 & \text{if otherwise} \end{cases}$$

2. Smith-Waterman Algorithmus.

- (a) Implementieren Sie den Smith-Waterman Algorithmus zur Berechnung von optimalen lokalen Alignments. Das Programm sollte vier Parameter haben:

int indelscore - Score für eine Insertion bzw. Deletion.

NNMatrix scorematrix - Eine symmetrische $n \times n$ -Matrix, die die Replacement-Scores enthält.

String u - Erste zu alignierende Sequenz.

String v - Zweite zu alignierende Sequenz.

- (b) Testen Sie den Algorithmus, indem Sie aus dem Internet eine BLOSUM-Matrix und zwei Proteinsequenzen herunterladen und den Algorithmus auf diese Daten anwenden.

3. Maximale Matches Modell.

- (a) Lesen und verstehen Sie Kapitel 3.5 (The Maximal Matches Model) des Skripts.

- (b) Gegeben seien zwei Sequenzen $\mathbf{u} = \text{TGAGCTCTCTAC}$ und $\mathbf{v} = \text{TACTCTCAATGC}$.

- i. Berechnen Sie die links-nach-rechts Partition von \mathbf{u} bezüglich \mathbf{v} .
- ii. Berechnen Sie die rechts-nach-links Partition von \mathbf{u} bezüglich \mathbf{v} .
- iii. Berechnen Sie die links-nach-rechts Partition von \mathbf{v} bezüglich \mathbf{u} .
- iv. Berechnen Sie die rechts-nach-links Partition von \mathbf{v} bezüglich \mathbf{u} .

- (c) Beweisen Sie, dass die links-nach-rechts Partition eine minimale Partition ist. (Hilfe: Beweis durch Widerspruch.)