

Algorithmen und Datenstrukturen in der Bioinformatik

Fünftes Übungsblatt WS 14/15

Abgabe Donnerstag, 12:00

Niveau I

Aufgabe 1: Alignment mit linearem Speicherplatz

Für die Worte $W_1=HOWTO$ und $W_2=GOTOH$ soll mit linearem Speicherplatz ein Alignment erstellt werden. ($match=2$, $mismatch=-1$, $gap(d)=2$)

Ermitteln sie für den ersten Durchlauf die Tracebackpfeile r , die auf die Zellen der mittleren Spalte zeigen.

Aufgabe 2: Sequence alignment mit GOTOH

Gegeben seien die Sequenzen $S_1=GATTACA$ und $S_2=GTA$.

Erstellen Sie ein globales Alignment mit affinen Gapkosten:

$match=2$, $mismatch=-3$, $gap\ opening(d)=6$, $gap\ extension(e)=1$.

Verwenden Sie hierzu den GOTOH-Algorithmus aus der Vorlesung und notieren Sie ein optimales Alignment.

Niveau II

Aufgabe 3: Overlap Alignment

Ein *Overlap Alignment* ist das optimale Alignment zwischen dem Ende der einen und dem Anfang der anderen Sequenz, d. h. es werden jeweils für eine Sequenz keine Startgaps und für die andere keine Endgaps gewertet.

- Wie kann man den Algorithmus von Needleman/Wunsch modifizieren, so dass er ein overlap alignment ausgibt?

Motivation: Beim *shotgun sequencing* von Genomen müssen die erhaltenen Fragmente anhand von überlappenden gemeinsamen Sequenzen wieder zusammengesetzt werden.