

### 13. Übung „Algorithmen der Bioinformatik I“

1. Gegeben sei die Levensthein-Distanz als Kostenfunktion. Im Alignment-Graphen bezeichnen wir die Menge der Knoten  $(i, j)$  mit  $|i - j| \leq k$  als Schlauch der Breite  $k$  um die Hauptdiagonale.

Zeigen Sie:

- Für Strings  $S_1, S_2$  mit einer Distanz  $D(S_1, S_2) \leq k$  verbleibt der Pfad des optimalen Alignments im Schlauch der Breite  $k$  um die Hauptdiagonale. (2 Punkte)
- Für Strings  $S_1, S_2$  mit  $|S_1| = |S_2|$  und einer Distanz  $D(S_1, S_2) \leq 2k$  verbleibt der Pfad des optimalen Alignments im Schlauch der Breite  $k$  um die Hauptdiagonale. (2 Punkte)

Gegen sei die Levensthein-Distanz als Kostenfunktion. Ein globales Alignment zweier Strings werde mit der Technik des dynamischen Programmierens bestimmt. Zeigen Sie, dass für die Einträge  $D_{i,j}$  der dabei generierten Matrix gilt

$$\begin{aligned} |D_{i,j} - D_{i,j-1}| &\leq 1 && \text{für } i \geq 0, j \geq 1 \\ |D_{i,j} - D_{i-1,j}| &\leq 1 && \text{für } i \geq 1, j \geq 0 \\ D_{i,j} - D_{i-1,j-1} &\geq 0 && \text{für } i \geq 1, j \geq 1 \end{aligned}$$

(3 Punkte)

2. Unter einem approximativen maximalen Pair eines Strings  $S$  verstehen wir ein Alignment zweier Teilstrings  $S_1 = S[i_1..j_1]$  und  $S[i_2..j_2]$ , wobei ausgeschlossen sei, dass ein Zeichen (an der selben Position) mit sich selbst aligniert wird.

Beispiel: Für  $S = xaxb$  mit  $S_1 = xax$  und  $S_2 = xb$  kann man die Alignments

$$\begin{array}{cccc} x & a & x & - \\ - & - & x & b \end{array} \quad \text{und} \quad \begin{array}{ccc} x & a & x \\ x & b & - \end{array}$$

angeben, wobei das erste Alignment kein approximatives maximales Pair darstellt, da  $S(3)$  mit sich selbst aligniert wird, wohingegen das zweite Alignment ein approximatives maximales Pair ist.

Ein optimales approximatives maximales Pair ist ein approximatives maximales Pair mit maximaler Ähnlichkeit für das Alignment.

- Entwickeln Sie einen Algorithmus, der zu String  $S$  das optimale approximative maximale Pair bestimmt. Hinweis: Nutzen Sie die DP-Matrix für das lokale Alignment. Die Hauptdiagonale spielt dabei eine wichtige Rolle. (3 Punkte)
  - (Zusatzaufgabe) Welche Folgen hätte es, wenn man statt der maximalen Ähnlichkeit minimale Kosten für das optimale approximative maximale Pair verlangen würde? (2 Punkte)
3. Arbeiten Sie Ihre Aufzeichnungen zur Vorlesung durch und stellen Sie eventuelle Unklarheiten fest. Formulieren Sie eine Fragestellung, die Sie in der nächsten Übung behandelt wissen möchten. (2 Punkte)