

## Übungen zur Vorlesung Algorithmen zur Sequenzanalyse Sommersemester 2003

Die Aussagen der ersten beiden Aufgaben bilden die Grundlage für Ukkonens cutoff Algorithmus zur Lösung des  $k$ -Differenzen Problems (den wir später kennenlernen werden).

**Aufgabe 4.1** Es sei  $\delta$  die Einheitskostenfunktion. Zeigen Sie, dass die Tabelle  $E_\delta$  zur Berechnung der Levenshtein Distanz mittels Dynamic Programming folgende Eigenschaften hat:

1.  $E_\delta(i, j) - 1 \leq E_\delta(i + 1, j) \leq E_\delta(i, j) + 1, i \in [0, m - 1], j \in [0, n]$ .
2.  $E_\delta(i, j) \leq E_\delta(i + 1, j + 1) \leq E_\delta(i, j) + 1, i \in [0, m - 1], j \in [0, n - 1]$ .
3.  $E_\delta(i, j + 1) - 1 \leq E_\delta(i, j) \leq E_\delta(i, j + 1) + 1, i \in [0, m], j \in [0, n - 1]$ .

D.h. benachbarte Einträge in der Tabelle  $E_\delta$  unterscheiden sich um höchstens 1 und die Einträge in jeder Diagonalen bilden eine nicht absteigende Folge.

**Aufgabe 4.2** Zeigen Sie, dass für alle  $(i, j) \in [0, m - 1] \times [0, n - 1]$  die folgenden Eigenschaften gelten ( $\delta$  ist wieder die Einheitskostenfunktion):

1. Falls  $E_\delta(i, j) \leq E_\delta(i, j + 1)$  und  $E_\delta(i, j) \leq E_\delta(i + 1, j)$ , dann ist  $E_\delta(i, j) = E_\delta(i + 1, j + 1)$  gdw.  $u_{i+1} = v_{j+1}$ .
2. 
$$E_\delta(i + 1, j + 1) = \begin{cases} E_\delta(i, j) & \text{if } u_{i+1} = v_{j+1} \\ 1 + E_\delta(i, j + 1) & \text{else if } E_\delta(i, j + 1) < E_\delta(i, j) \\ 1 + \min\{E_\delta(i + 1, j), E_\delta(i, j)\} & \text{otherwise} \end{cases}$$

**Aufgabe 4.3** Bestimmen Sie die einfache Levenshtein-Distanz von  $u = FREIZEIT$  und  $v = ZEITGEIST$  mit Hilfe des vorgestellten Algorithmus zur schnellen Berechnung der einfachen Levenshtein-Distanz (Script S.20).

**Aufgabe 4.4** Seien  $u = acbacbdab$  und  $v = bbacbdbcab$ .

1. Bestimmen Sie die einfache Levenshtein-Distanz von  $u$  und  $v$ .
2. Bestimmen Sie die Levenshtein-Distanz von  $u$  und  $v$ .
3. Berechnen Sie  $mmdist(u, v)$  und  $mmdist(v, u)$ .
4. Bestimmen Sie den  $q$ -Wort Abstand von  $u$  und  $v$  für  $q = 2$  und  $q = 3$ .
5. Berechnen Sie  $score_{fasta}(u, v)$  für  $q = 2$  und  $q = 3$ .