

# Übungen zur Vorlesung Sequenzanalyse I

Universität Bielefeld, Wintersemester 2010/2011  
Dipl.-Inform Peter Husemann · Dr. Roland Wittler

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2010winter/SequenzAnalyse>

**Blatt 3 vom 29.10.2010**

**Abgabe in einer Woche vor Beginn der Vorlesung.**

## **Aufgabe 1** Hamming-Distanz

(3 Punkte)

Gegeben sind die Sequenzen  $s_1 = qppqppqppqpqpq$ ,  $s_2 = qppqpqqppqppq$  und  $s_3 = qppqpqqppqpq$ .

1. Berechne paarweise die Hamming-Distanzen zwischen diesen Sequenzen. Schreibe dazu eine Java-Funktion `public static int calcHammingDistance(String a, String b)`, die für Dich diese Aufgabe übernimmt.
2. Was gibt Dein Programm für die Sequenzen  $s_4 = ppppppqqppqq$  und  $s_5 = pppqqppqpq$  aus?

## **Aufgabe 2** Edit-Funktion

(2 Punkte)

Berechne von Hand die folgenden Ausdrücke und gib dabei *alle* Zwischenschritte an.

1.  $E(\text{ACAT}, \mathcal{I}_G \mathcal{C} \mathcal{I}_T \mathcal{I}_T \mathcal{I}_A \mathcal{C} \mathcal{C} \mathcal{D})$
2.  $E(\text{TAGCAT}, \mathcal{D} \mathcal{C} \mathcal{D} \mathcal{C} \mathcal{C} \mathcal{C} \mathcal{C})$

## **Aufgabe 3** Edit-Sequenz und Edit-Distanz

(2 Punkte)

Überführe die Sequenz  $x = \text{GAGTAG}$  in den String  $y = \text{AGCAT}$  durch Anwendung der *Edit*-Funktion.

1. Gib eine möglichst kurze Edit-Sequenz an, die  $x$  in  $y$  überführt. Wie sind die Kosten?
2. Gib eine der längsten Edit-Sequenzen an, die  $x$  in  $y$  überführt. Wie sind die Kosten?

## **Aufgabe 4** Edit-Distanz und reverses Komplement

(2 Punkte)

Um bei DNA Sequenzen das reverse Komplement zu erhalten, wird die Sequenz umgedreht und jeweils A und T vertauscht, sowie C und G. Das reverse Komplement von  $s = \text{AGCTTA}$  ist also  $rc(s) = \text{TAAGCT}$ . Zeige formal, dass sich die Edit-Distanz nicht ändert, wenn beide Sequenzen revers komplementiert werden:  $d(x, y) = d(rc(x), rc(y))$ .