

# Übungen zur Vorlesung Grundlagen der Sequenzanalyse

Universität Bielefeld, WS 2006/07

Dr. Sven Rahmann · Dipl.-Bioinf. Katharina Jahn

<http://gi.cebitec.uni-bielefeld.de/teaching/2006winter/sequenzanalyse/>

**Blatt 4 vom 09.11.2006**

**Abgabe am 16.11.2006 vor der Vorlesung um 8:30 in H3**

**Aufgabe 1** Gib zu  $q = 3$  und  $x = \text{cbabbcabcbabaa}$  das  $q$ -gram Profil an. Finde ein Wort  $y \neq x$  mit  $d_q(x, y) = 0$ .

**Aufgabe 2** Für das Alphabet  $\Sigma = \{0, 1, \dots, 7\}$  und Wortlänge  $q = 4$ , berechne den Rang (eine der beiden Möglichkeiten) des Wortes 2175 und danach (ohne vollständige Neuberechnung, sondern durch ein Update in konstanter Zeit) den Rang des Wortes 1753. Welches Wort hat den Rang 2175?

**Aufgabe 3** Begründe folgende Aussagen zu symmetrischen log-odds Scorematrizen:

1. Eine Scorematrix für die evolutionäre Zeitspanne der Länge  $t = 0$  hat positive Diagonalelemente; ansonsten sind alle Scores gleich  $-\infty$ .
2. Für  $t = \infty$  sind alle Scores gleich Null.

**Aufgabe 4** Formuliere den DP-Algorithmus zur Berechnung der Edit-Distanz so um, dass nun der Edit-Score berechnet wird. Berechne den Edit-Score von WRITERS und VINTNER mit Hilfe des Algorithmus. Berechne dann die Edit-Distanz aus den Wortlängen und dem Edit-Score.