

Präsenzübungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2015/2016

Prof. Dr. Jens Stoye · M.Sc. Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2015winter/SequenzAnalyse>

Präsenzübungsblatt 2, Woche 46/2015

Aufgabe 1 (Beziehung der Edit-Distanz zur q -gram Distanz)

Zwischen der Standard-Edit-Distanz mit Einheitskosten und der q -gram-Distanz besteht folgende Beziehung:

$$\frac{d_q(x, y)}{2q} \leq d(x, y)$$

Lies dir den Beweis dazu im Skript auf Seite 22 durch.

Berechne nun die Edit-Distanz (ohne Matrix) und die q -gram-Distanz ($q = 3$) von den beiden Strings $x = aaaaa$ und $y = aabaa$. Fertige dann dafür eine Skizze an, die die Beziehung verdeutlicht.

Aufgabe 2 (Laufzeit der q -gram-Distanz)

Begründe, warum die Berechnung der q -gram-Distanz von zwei Strings in linearer Zeit berechnet werden kann. Gehe dafür die Schritte durch, die notwendig sind, um die Distanz zu berechnen.