

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2013/2014

Prof. Dr. Jens Stoye · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2013winter/SequenzAnalyse>

Übungsblatt 3 vom 05.11.2013

Abgabe in einer Woche vor Beginn der Vorlesung.

Aufgabe 1 (q -gram Distanz)

(3 Punkte)

Gegeben sind die Worte $x = \text{banenenbabane}$ und $y = \text{nebenanbenbananen}$.

1. Gib die vollständigen q -gram Profile für x und y für $q = 2$ über dem Alphabet $\Sigma = \{\mathbf{a, b, e, n}\}$ an.
2. Wie hoch ist die q -gram Distanz zwischen den Worten?

Aufgabe 2 (Worte mit dem gleichen q -gram Profil)

(4 Punkte)

Gegeben sind $q = 3$, $\Sigma = \{\mathbf{a, c, g}\}$ und $x = \text{gcaggagacaggac}$.

1. Gib das q -gram Profil von x an.
2. Zeichne den De-Bruijn Graphen zu x . Die Knoten des Graphen entsprechen den $q - 1$ -grams von x , die Kanten repräsentieren die q -grams.
3. Finde mit Hilfe des De-Bruijn Graphen ein Wort $y \neq x$ mit $d_q(x, y) = 0$.

Aufgabe 3 (Partitionierungen)

(3 Punkte)

Gib jeweils die Links-nach-Rechts und die Rechts-nach-Links Partitionierung von x bezüglich y und deren Länge an.

1. $x = \text{babbabbabaaa}$ und $y = \text{bab}$
2. $x = \text{bb}$ und $y = \text{cbbba}$
3. $x = \text{dca}$ und $y = \text{baab}$

Aufgabe 4 (Maximal-Matches Distanz)

(4 Punkte)

Berechne δ und gib jeweils die Links-nach-Rechts und Rechts-nach-Links Partitionierungen an.

1. $\delta(\text{xylophon} \parallel \text{colophonon})$
2. $\delta(\epsilon \parallel \text{bbbb})$
3. $\delta(\text{bbbb} \parallel \epsilon)$

Welche Beobachtung machst du bei 2. und 3.? Was sagt das über die Maximal-Matches Distanz aus?