

Präsenzübungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2016/2017

Prof. Dr. Jens Stoye · M.Sc. Tizian Schulz

<https://gi.cebitec.uni-bielefeld.de/teaching/2016winter/sa>

Präsenzübungsblatt 1, Woche 43/2016

Aufgabe 1 (Grundlagen)

Diskutiert gemeinsam mit eurem Tutor die Inhalte des zweiten Kapitels im Skript.

Aufgabe 2 (Anzahl von Subsequenzen und Substrings)

Versuche, dir die folgenden Formeln jeweils anschaulich klar zu machen und fertige kleine Skizzen an.

1. Wie viele Subsequenzen der festen Länge k hat ein String der Länge n ?
2. Wie viele Subsequenzen mit den Längen $0 \leq k \leq n$ hat ein Wort der Länge n insgesamt?
3. Wie viele Substrings der festen Länge k hat ein String der Länge n ?
4. Wie viele Substrings mit den Längen $0 \leq k \leq n$ hat ein Wort der Länge n insgesamt?

Aufgabe 3 (Verschiedene Metriken)

Aus der Vorlesung kennst du bereits folgende Metriken mit ihren Definitionen:

$$d_1(x, y) := \sum_{i=1}^n |x_i - y_i| \quad (\text{Mannheim metric oder Manhattan distance})$$

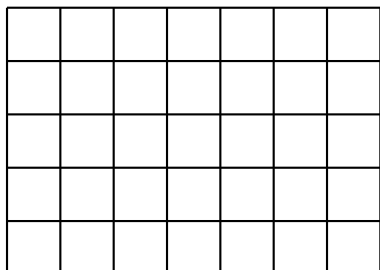
$$d_2(x, y) := \sqrt{\sum_{i=1}^n |x_i - y_i|^2} \quad (\text{Euclidean distance})$$

$$d_\infty(x, y) := \max_{i=1, \dots, n} |x_i - y_i| \quad (\text{maximum metric})$$

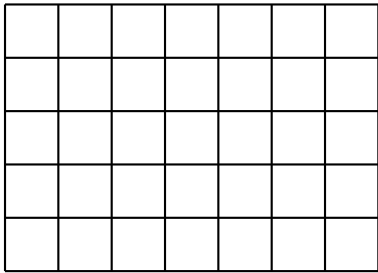
$$d_H(x, y) := \sum_{i=1}^n \mathbb{1}_{x_i \neq y_i} \quad (\text{Hamming distance})$$

Zeiche in die folgenden Koordinatensysteme die Punkte A und B jeweils so ein, dass sich die folgenden Distanzen ergeben:

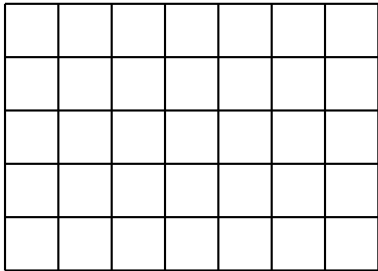
1. $d_1(A, B) = 7$



2. $d_2(A, B) = \sqrt{40}$



3. $d_\infty(A, B) = 5$



4. $d_H(A, B) = 1$

