

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 1 vom 22.10.2013

Abgabe in einer Woche vor Beginn der Vorlesung.

Organisatorisches

Wichtige Hinweise: Bitte unbedingt im eKVV für die Vorlesung und die Übungen registrieren. Auf der Homepage der Veranstaltung steht eine elektronische Version des Skripts zum Download bereit. Die gedruckte Version wird in der ersten Übung verteilt.

Abgabe der Übungszettel: Auch wenn die Bearbeitung der Aufgaben in einer Gruppe (maximal zwei Personen) erfolgt ist, so müssen die Lösungen zu den Aufgaben separat und unter eigenem Namen **vor** Beginn der Vorlesung am Dienstag abgegeben werden. Alternativ könnt ihr die Zettel auch euren Tutoren geben oder sie **vor** Beginn der Vorlesung am Dienstag in den Briefkasten bei Raum U10-151 werfen. (Die Abgabe im Briefkasten folgt auf eigene Gewähr.) Bitte den Tutorientermin bzw. den Tutor, sowie euren Namen und den eures Gruppenpartners deutlich auf der Abgabe vermerken.

Quellenangabe auf den Übungszetteln: Benutzt ihr bei der Bearbeitung der Übungsaufgaben andere Quellen als das Skript zur Vorlesung, so gebt diese bitte immer mit an. Ist nicht ersichtlich, woher ihr eure Informationen genommen habt, können wir euch leider keine Punkte geben.

Teilnahme an der Klausur: Die Klausur kann am Ende mitschreiben, wer bis dahin A&D, Grundlagen der Programmierung oder Ähnliches bestanden, mindestens 50% der Übungspunkte erreicht und mindestens zweimal in den Übungen eine Aufgabe vorgerechnet hat.

Aufgaben

Aufgabe 1 (Diskrete Metrik)

(3 Punkte)

Auf jeder Menge \mathcal{X} lässt sich die diskrete Metrik $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ definieren:

$$d(x, y) := \begin{cases} 0, & \text{wenn } x = y, \\ 1, & \text{wenn } x \neq y. \end{cases}$$

Weise nach, dass es sich bei d um eine Metrik auf \mathcal{X} handelt, indem Du die Definitionen (1) - (3) aus dem Skript (S. 12) überprüfst.

Aufgabe 2 (Die Hamming-Distanz)

(4 Punkte)

In der Vorlesung wurde die Hamming-Distanz eingeführt. Implementiere eine Funktion, die als Parameter zwei Strings a und b übergeben bekommt und ihre Hamming-Distanz ausgibt. Wenn die beiden Strings eine unterschiedliche Länge haben, soll eine Fehlermeldung ausgegeben werden. Sprich mit deinem Tutor ab, welche Programmiersprache du benutzen kannst.

Aufgabe 3 (*Edit*-Operationen)

(2 Punkte)

Berechne $E(\text{AGACCT}, \text{CS}_A\text{CDCS}_G)$ und gib alle Zwischenschritte an.

Aufgabe 4 (*Edit*-Sequenz und -Distanz)

(5 Punkte)

Überführe den String $x = \text{BAUSTELLE}$ in den String $y = \text{LAUBFALE}$ durch Anwendung der *Edit*-Funktion.

1. Gib eine möglichst kurze *Edit*-Sequenz an, die x in y überführt.
2. Gib drei der längsten *Edit*-Sequenzen an, die x in y überführen.
3. Wie hoch ist die *Edit*-Distanz bei deinen Lösungen?