

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2012/2013

Prof. Dr. Jens Stoye · Nina Luhmann · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2012winter/SequenzAnalyse>

Übungsblatt 1 vom 12.10.2012

Abgabe in zwei Wochen am 26.10. 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 Freitag abgegeben werden. Alternativ könnt ihr die Zettel auch euren Tutoren geben oder sie **vor** Beginn der Vorlesung am Freitag in den Briefkasten bei Raum U10-151 werfen. 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 oder Grundlagen der Programmierung bestanden, mindestens 50% der Übungspunkte erreicht und mindestens zweimal in den Übungen eine Aufgabe vorgerechnet hat.

Aufgaben

Aufgabe 1 (Textsuche) (6 Punkte)

Welche Algorithmen zur Textsuche kennst du bereits aus der Vorlesung *Algorithmen und Datenstrukturen* (es sollten vier Stück sein)? Beschreibe diese jeweils mit 2 bis 3 Sätzen. Welche Problemstellung lösen diese Algorithmen? Welche Problemstellungen gibt es darüber hinaus in der Textsuche?

Andere Quellen als das Skript zur Vorlesung und das zu *Algorithmen und Datenstrukturen* bitte angeben!

Aufgabe 2 (Komplexitätsklassen) (4 Punkte)

Gegeben sind zwei Algorithmen A_1 und A_2 . Bei einer Eingabgröße von n braucht A_1 $f_1 = n^4 + 2$ Rechenschritte, A_2 benötigt $f_2 = 2n \cdot \log n + 200$ Schritte.

1. Stelle die Funktionen grafisch in einem gemeinsamen Koordinatensystem dar. Es reicht, wenn du die Funktionen im Bereich $n = 0$ bis $n = 10$ skizzierst.
2. Ermittle die Komplexitätsklasse der Algorithmen in der \mathcal{O} -Notation.
3. Welcher Algorithmus hat die asymptotisch schlechtere Laufzeit? Ab welchem n_0 wird das deutlich, mit der Vorgabe, dass $c = 1$ ist?

Aufgabe 3 (Alphabete) (4 Punkte)

Gegeben sei das Alphabet $\Sigma = \{a, h\}$.

1. Wie ist Σ^5 definiert?
2. Schreibe alle Worte $w \in \Sigma^3$ auf.
3. Wie sind Σ^* und Σ^+ definiert und was ist der Unterschied zwischen ihnen?
4. Was ist der Unterschied zwischen ε , $\{\varepsilon\}$ und $\{\}$?

Aufgabe 4 (Substrings)

(4 Punkte)

Gegeben sei das Wort $s = \text{GIRAFFE}$.

1. Gib alle Substrings der Länge 4 und eine Subsequenz der Länge 4 von s an.
2. Gib alle Präfixe und Suffixe von s an.
3. Versuche, eine allgemeine Formel für die Anzahl der Substrings einer festen Länge k in einem beliebigen Wort der Länge n zu finden. Wie viele Substrings der Länge 5 enthält s ?

Aufgabe 5 (Metriken)

(3 Punkte)

Eine Metrik $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$, ist auf der Menge \mathcal{X} vollständig definiert durch:

$$d(x, y) = 0 \leftrightarrow x = y \tag{1}$$

$$d(x, y) = d(y, x), \forall x, y \in \mathcal{X} \tag{2}$$

$$d(x, y) \leq d(x, z) + d(z, y), \forall x, y, z \in \mathcal{X} \tag{3}$$

Beweise, dass aus diesen Definitionen die Nichtnegativität folgt:

$$d(x, y) \geq 0, \forall x, y \in \mathcal{X}$$

Aufgabe 6 (Hamming-Distanz)

(4 Punkte)

Gegeben sind die Sequenzen $s_1 = \text{HAHAHA}$, $s_2 = \text{AHAAAA}$, $s_3 = \text{HHMMMM}$ und $s_4 = \text{HAMM}$. Berechne mit einem selbst geschriebenen Programm in einer Programmiersprache deiner Wahl sämtliche paarweisen Hamming-Distanzen zwischen diesen Sequenzen und gib diese an. Gib das Programm bitte auch in elektronischer Form an deinen Tutor ab.