

Prüfung zu Lehrveranstaltung 708.031 Datenstrukturen und Algorithmen

Es sind keinerlei Unterlagen oder Hilfsmittel erlaubt. Es dürfen nur einzelne, lose Blätter verwendet werden! Auf jedem Blatt muss der Name und die Matrikelnummer angegeben werden! Reine Arbeitszeit beträgt 90 Minuten.

1. Asymptotische Schranken (10 Punkte)

- a.) Beweisen oder widerlegen Sie folgende Aussagen (Achtung: Antworten ohne richtige Begründung erhalten KEINE Punkte!)
 - i.) Es existiert kein Suchalgorithmus welcher eine Laufzeit von $O(n^2 \log n)$ besitzt.
 - ii.) Jeder Algorithmus, welcher einen Speicher von $\Theta(n)$ benötigt, ist schneller als ein Algorithmus, welcher einen Speicher von $\Omega(n^3)$ benötigt.
 - iii.) Es existiert ein Algorithmus welcher eine Laufzeit von $O(n \cdot \log n)$ und einen Speicher von $\Omega(n^2)$ benötigt.
 - iv.) Jeder Algorithmus mit einer Laufzeit $T(n) = \Theta(n^c)$ mit $c > 2$ ist auch $T(n) = \Omega(\log_c n)$.
- b.) Lösen Sie folgende Rekursionen (ohne die Verwendung der *Master Methode*):
 - (i) $T(n) = O(n) + 3T(\frac{n}{3}), T(1) = O(1)$
 - (ii) $T(n) = O(\sqrt{n}) + T(\sqrt{n}), T(1) = O(1)$

2. Optimales Kodieren (10 Punkte)

- a.) Erklären Sie den Begriff der **Entropie**.
- b.) Folgender Text ist gegeben: "HALLO_IN_HONOLULU" - Ermitteln Sie dazu eine Häufigkeitstabelle (Achtung: auch das Leerzeichen " " ist hier ein Zeichen !) und konstruieren Sie daraus **Schritt für Schritt** den optimalen Codierungsbaum (nach Huffman). Wieviele Bits können bei diesem Text gegenüber einem Code mit je 3 Bits/Zeichen eingespart werden?
- c.) Schreiben Sie eine Pseudocode für die Implementierung der Huffman-Kodierung mit Halden. Wie lautet die Laufzeit $T(n)$?

3 Gestreute Speicherung(10 Punkte)

- a.) Erklären Sie das Grundprinzip (mit Skizze) der gestreuten Speicherung. Wann kommt es zu Kollisionen und wie werden dieses behandelt (2 Methoden).
- b.) Erklären Sie die Divisions- und die Multiplikationsmethode und definieren Sie die ideale Hashfunktion.
- c.) **Leiten** Sie den erwarteten Suchaufwand (Laufzeit $T(n)$) in einer Hashtabelle T mit Überläuferlisten **ab**. Betrachten Sie dabei getrennt die beiden Fälle (i) $w \notin T$ und (ii) $w \in T$.

4. Richtig oder Falsch (10 Punkte)

Stimmen folgenden Aussagen? Beachten Sie, dass es Punkte nur bei richtiger Antwort **MIT** richtiger Begründung gibt.

- a.) Es existiert ein optimaler, präfix-freier Binärcode mit den Kodewortelängen 2,2,3,3,3,4,5,5.
- b.) Der randomisierte Quicksort ist für alle möglichen Inputs (Annahme: ganzzahlig und es kommen keine Werte mehrfach vor) immer schneller als der nicht randomisierte Quicksort.
- c.) Eine ideale Hashfunktionen liefert immer einen Index zurück, welcher zu keiner Kollisionen führt.
- d.) Die Interpolationssuche ist immer schneller als die Binärsuche bei der Suche in sortierten linearen Feldern.
- e.) Es existiert ein Suchalgorithmus, welcher in vorsortierten linearen Feldern einen Speicherbedarf von $S(n) = O(n^2)$ besitzt.