

Praktikum Algorithmik SS 2005

Aufgabe 3: MERGESORT im Pseudocode

Aufgabe 4: Analyse von MERGESORT mit dem Master-Theorem

Aufgabe 5: Nicht-deterministischer Algorithmus RANDOMIZEDQUICKSELECT

Aufgabe 6: Programmierung und Analyse von RANDOMIZEDQUICKSELECT

Abgabe: Gruppen A u. D bis 20. KW, Gruppen B u. C bis 21. KW

Name: Matr-Nr:

Datum: Unterschrift des Dozenten (wenn bestanden):

Lernziele der Aufgaben 3 u. 4: Divide & Conquer – Prinzip, Anwendung des Master-Theorems zur Analyse eines D&C-Algorithmus

Aufgabe 3 (MERGESORT)

Formulieren Sie MERGESORT(A, p, r) rekursiv als Divide&Conquer-Algorithmus. Unterscheiden Sie in Ihrem Algorithmus deutlich die drei Schritte *Divide*, *Conquer* und *Combine*.

Hinweise: $A = (A[1], A[2], \dots, A[n])$ ist die zu sortierende Folge, p ist der Index der ersten Elements und r der Index des letzten Elements der Folge A .

Gegeben ist eine Prozedur MERGE(A, p, q, r), die zwei Folgen ($A[p], \dots, A[q]$) und ($A[q+1], \dots, A[r]$) zu einer sortierten Folge ($A[p], \dots, A[r]$) mischt.

Schreiben Sie MERGESORT im Pseudocode. Erklären Sie die Prozedur MERGE verbal oder mit Pseudocode.

Aufgabe 4 (Analyse von MERGESORT)

- Identifizieren Sie in Ihrem Programm genau die Schritte *Divide*, *Conquer* und *Combine* und überlegen Sie sich für jeden dieser Schritte die Zeitkomplexität.
- Stellen Sie mit den Überlegungen aus a) eine Differenzgleichung für MERGESORT auf.
- Lösen Sie die Differenzgleichung mit Hilfe des in der Vorlesung besprochenen Mastertheorems.

Lernziele der Aufgaben 5 u. 6: Am Beispiel des Algorithmus RANDOMIZEDQUICKSELECT soll eine Decrease-and-Conquer-Strategie (auch Prune-and-Conquer genannt) angewendet werden (→ Kapitel 1.3 Algorithmische Entwurfsmuster). Simulation eines schnellen nicht-deterministischen Algorithmus.

Aufgabe 5 (Verstehen von RANDOMIZEDQUICKSELECT)

Gegeben ist eine unsortierte Folge von n Elementen, auf der eine Ordnungsrelation definiert ist. Die Aufgabe besteht darin, das k -kleinste Element in der Folge zu finden. Ein zufallsgesteuerter Algorithmus RANDOMIZEDQUICKSELECT wird eine mittlere Laufzeit von $O(n)$ haben, wobei der Algorithmus mögliche Entscheidungen zufallsgesteuert trifft. Über die Verteilung der Eingabedaten muss keine statistische Annahme gemacht werden. Im Worst Case wird die Laufzeit $O(n^2)$ betragen. Mit dem zufallsgesteuerten Algorithmus RANDOMIZEDQUICKSELECT wird hier ein schneller „nicht-deterministischer“ Algorithmus simuliert.

Idee des Algorithmus

Gegeben sei eine unsortierte Folge S von n vergleichbaren Elementen und eine Integerzahl $k \in [1, n]$. Wähle zufällig ein Element x aus S und benutze dieses als Pivot-Element, um S in die drei Subsequenzen L , E und G zu teilen. L speichert alle Elemente von S , die kleiner als x , E alle die gleich x und G alle die größer als x sind. Dieser Schritt heißt **Decrease-** oder auch **Prune-**Schritt (prune = abschneiden). Anhand des Vergleichs von k mit der Kardinalität $|L|$, $|E|$ und $|G|$ der drei Mengen wird entschieden, mit welchen Parameterwerten für S und k der Algorithmus RANDOMIZEDQUICKSELECT rekursiv aufgerufen wird. Der Pseudocode der Algorithmus-Idee sieht wie folgt aus:

Algorithmus RANDOMIZEDQUICKSELECT(S , k):

Input: Sequenz S von n vergleichbaren Elementen und Integer $k \in [1, n]$

Output: Das k -kleinste Element von S

if $n = 1$ **then**

return das (erste) Element von S

 wähle ein Element x zufällig aus S aus

 entferne alle Elemente aus S und speichere sie in den Folgen:

- L , speichert alle Elemente aus S kleiner als x
- E , speichert alle Elemente aus S gleich x
- G , speichert alle Elemente aus S größer als x

if $k \leq |L|$ **then** ... ?

else if $k \leq |L| + |E|$ **then** ... ?

else RANDOMIZEDQUICKSELECT(G , $k - |L| - |E|$)

Bearbeiten Sie folgende Teilaufgaben:

- Erklären Sie, warum die Brute-Force-Methode zur Lösung des Select-Problems $O(n^2)$ Zeit benötigt!
- Erklären Sie genau die Idee des beschriebenen Algorithmus RANDOMIZEDQUICKSELECT, und vervollständigen Sie den Pseudocode an den mit Fragezeichen versehenen Stellen. Für jede der drei Fallunterscheidungen findet ein rekursiver Aufruf statt. Wie lauten die Parameter dieser Aufrufe?
- Worin besteht der nicht-deterministische Kern des Algorithmus RANDOMIZEDQUICKSELECT? An welchen Stellen trifft der Algorithmus „selbst“ Entscheidungen, die zu einer sehr schnellen Lösung des Select-Problems führen? Bitte genau erklären!!

- Führen Sie den Algorithmus von Hand auf Papier für die folgende Menge S und $k = 9$ aus. Untersuchen Sie jeweils für die Wahl von x die Fallunterscheidungen, und schreiben Sie jeden Rekursionsaufruf mit den geeigneten Parametern auf.

$S = \{ 3, 9, 15, 8, 5, 17, 10, 1, 13, 6, 20, 11, 15, 25, 29, 2, 3, 14, 4, 3 \}$

Zur Konstruktion der beiden rekursiven Aufrufe versuchen Sie, Antworten auf folgende zwei Fragen zu finden:

- Welcher Zusammenhang besteht zwischen der Größe der Menge L und dem k -kleinsten Element von S ?
- Was kann man über das k -kleinste Element aussagen, wenn $k = |L| + |E|$?

- Diskutieren Sie die Laufzeit-Performance des Algorithmus RANDOMIZEDQUICKSELECT.

Aufgabe 6 (Programmieren, Testen und Analysieren von RANDOMIZEDQUICKSELECT)

Programmieren Sie den Algorithmus RANDOMIZEDQUICKSELECT und geben Sie in jedem Rekursionsschritt die Mengen L , E und G sowie deren Kardinalität auf dem Bildschirm aus.

- Erzeugen Sie die Startmenge S mit Hilfe eines Zufallszahlengenerators. Folgende Parameter zur Erzeugung von S geben Sie interaktiv ein
 - die Größe von S
 - den Wertebereich von S , z.B. $[10, 8888]$
 - den Wert k
 - Geben Sie das Pivotelement $x \in S$ vor jeder Aufteilung S in die Mengen L , E und G , d.h. in jedem Rekursionsschritt interaktiv ein
- Der Ablauf des Algorithmus RANDOMIZEDQUICKSELECT soll Schritt für Schritt am Bildschirm verfolgt werden. Geben Sie dazu alle relevanten Daten in übersichtlicher Form aus. Halten Sie dazu den Algorithmus nach jedem Rekursionsschritt an und setzen Sie Verarbeitung durch eine Tastenbetätigung und nach Eingabe eines Wertes für x fort. Muss x ein Element aus S sein?
- Führen Sie einen experimentellen Ratio-Test und Power-Test für RANDOMIZEDQUICKSELECT durch und zeichnen Sie die Laufzeiten mit einem Plot-Programm. Erklären und dokumentieren Sie die Ergebnisse.