

Aufgabe 5 [Hashing]

- (a) Nennen Sie zwei wünschenswerte Eigenschaften von Hashfunktionen.

surjektiv Die Abbildung soll surjektiv sein, d. h. jeder Index soll berechnet werden können.

gleichverteilt Durch die Hashfunktion soll möglichst eine Gleichverteilung auf die Buckets (Indexliste) erfolgen.

effizient Zudem sollte die Verteilung mittels Hashfunktion möglichst effizient gewählt werden.

- (b) Wie viele Elemente können bei Verkettung und wie viele Elemente können bei offener Adressierung in einer Hashtabelle mit m Zeilen gespeichert werden?

Verkettung Es darf mehr als ein Element pro Bucket enthalten sein, deswegen können beliebig viele Element gespeichert werden.

offene Adressierung (normalerweise) ein Element pro Bucket, deshalb ist die Anzahl der speicherbaren Elemente höchstens m . Können in einem Bucket k Elemente gespeichert werden, dann beträgt die Anzahl der speicherbaren Elemente $k \cdot m$.

- (c) Angenommen, in einer Hashtabelle der Größe m sind alle Einträge (mit mindestens einem Wert) belegt und insgesamt n Werte abgespeichert.

Geben Sie in Abhängigkeit von m und n an, wie viele Elemente bei der Suche nach einem nicht enthaltenen Wert besucht werden müssen. Sie dürfen annehmen, dass jeder Wert mit gleicher Wahrscheinlichkeit und unabhängig von anderen Werten auf jeden der m Plätze abgebildet wird (einfaches gleichmäßiges Hashing).

$\frac{n}{m}$ Beispiel: 10 Buckets, 30 Elemente: $\frac{30}{10} = 3$ Elemente im Bucket, die man durchsuchen muss.

- (d) Betrachten Sie die folgende Hashtabelle mit der Hashfunktion $h(x) = x \bmod 11$. Hierbei steht \emptyset für eine Zelle, in der kein Wert hinterlegt ist.

Index	0	1	2	3	4	5	6	7	8	9	10
Wert	11	\emptyset	\emptyset	3	\emptyset	16	28	18	\emptyset	\emptyset	32

Führen Sie nun die folgenden Operationen mit offener Adressierung mit linearem Sondieren aus und geben Sie den Zustand der Datenstruktur nach jedem Schritt an. Werden für eine Operation mehrere Zellen betrachtet, aber nicht modifiziert, so geben Sie deren Indizes in der betrachteten Reihenfolge an.

- (i) Insert 7

Index	0	1	2	3	4	5	6	7	8	9	10
Wert	11	\emptyset	\emptyset	3	\emptyset	16	28	18	7 ₂	\emptyset	32

- (ii) Insert 20

Index	0	1	2	3	4	5	6	7	8	9	10
Wert	11	\emptyset	\emptyset	3	\emptyset	16	28	18	7 ₂	20 ₁	32

- (iii) Delete 18

Index	0	1	2	3	4	5	6	7	8	9	10
Wert	11	\emptyset	\emptyset	3	\emptyset	16	28	del	7_2	20_1	32

del ist eine Marke, die anzeigt, dass gelöscht wurde und der Bucket nicht leer ist.

(iv) Search 7

Index	0	1	2	3	4	5	6	7	8	9	10
Wert	11	\emptyset	\emptyset	3	\emptyset	16	28	del	7_2	20_1	32

$$h(7) = 7 \bmod 11 = 7$$

7 (Index) \rightarrow del lineares sondieren \rightarrow 8 (Index) \rightarrow gefunden

(v) Insert 5

Index	0	1	2	3	4	5	6	7	8	9	10
Wert	11	\emptyset	\emptyset	3	\emptyset	16	28	5_3	7_2	20_1	32

Github: Staatsexamen/66115/2021/03/Thema-2/Teilaufgabe-2/Aufgabe-5.tex