

Datenbanken II

Übungsblatt 3 – WiSe 2019/20

9. Gegeben ist die Relation R (siehe Tabelle 1). Zeichnen Sie einen statischen Hash Index für das Attribut `Dept` dieser Relation. Der Hash Index besteht aus $b = 6$ Buckets, wobei jedes Bucket maximal 3 Einträge beinhalten kann (d.h. es können insgesamt $3 \cdot 6 = 18$ Einträge gespeichert werden - regulär). Allerdings kann es vorkommen, dass mehr als 3 Einträge in dasselbe Bucket gehasht werden. In einem solchen Fall spricht man von *Bucket Overflow*. In dieser Aufgabe wird *Overflow Chaining (Closed Addressing)* verwendet um mit solchen Fällen umzugehen (siehe Vorlesungsfolien zum Thema *Indexstrukturen* S. 65/90). Ein Zeiger auf ein Overflow Bucket benötigt wiederum einen Eintrag im Bucket, d.h. sobald ein Zeiger auf ein Overflow Bucket gespeichert wird, können nur mehr 2 Einträge für Schlüssel verwendet werden und im 3. Eintrag wird der Zeiger auf das Overflow Bucket gespeichert.

Duplikate im Hash-Index werden durch mehrfache Indexeinträge aufgelöst (siehe Vorlesungsfolien zum Thema *Indexstrukturen* S. 14/90).

Als Hash-Funktion für einen Eintrag $x \in R$ wird $h(x) = (s[0] + s[1]) \bmod b$ verwendet, wobei $s[i]$ der numerische Wert des Buchstabens an $(i + 1)$ -ter Position ist (siehe Tabelle 2). Beispielsweise berechnet sich der Hashwert für **MS** wie folgt: $(13 + 19) \bmod 6 = 32 \bmod 6 = 2$.

RegNo	Name	Dept
01700012	Abraham	EE
01700006	Alan	GG
01700001	Arthur	LS
01700019	Berta	AS
01700002	Charlie	MS
01700004	Claire	CL
01700015	Daryl	CS
01700009	Debra	LL
01700000	Dexter	NP
01700005	Doug	BE
01700011	Ephraim	GI
01700018	Gustavo	DS
01700020	Jake	AA
01700017	Jesse	PS
01700007	Joe	CL
01700014	Lucille	PL
01700013	Rick	CS
01700010	Rita	LL
01700003	Ryan	BA
01700008	Spencer	CP
01700016	Walter	BC

Tabelle 1: Relation R für Aufgabe 9.

Buchstabe	num. Wert
Aa	1
Bb	2
Cc	3
Dd	4
Ee	5
Ff	6
Gg	7
Hh	8
Ii	9
Jj	10
Kk	11
Ll	12
Mm	13
Nn	14
Oo	15
Pp	16
Qq	17
Rr	18
Ss	19
Tt	20
Uu	21
Vv	22
Ww	23
Xx	24
Yy	25
Zz	26

Tabelle 2: Numerische Werte für Aufgabe 9.

10. Verwenden Sie *erweiterbares Hashing* um die folgenden Suchschlüssel zu speichern. Die Schlüssel werden in folgender Reihenfolge eingefügt:

23, 25, 15, 7, 5, 8, 9, 10, 11, 12, 13, 26, 27, 28, 1, 2, 3

Als Hash-Wert wird der fünfstellige Binärwert des Suchschlüssels verwendet. In jedes Bucket passen 3 Suchschlüssel.

Zeichnen Sie den Index nach wichtigen Zwischenschritten (nach Vergrößerung des Verzeichnisses bzw. nach Teilung von Buckets).

11. Gegeben sei der Index in Abbildung 1, der per *erweiterbaren Hashing* erstellt wurde.

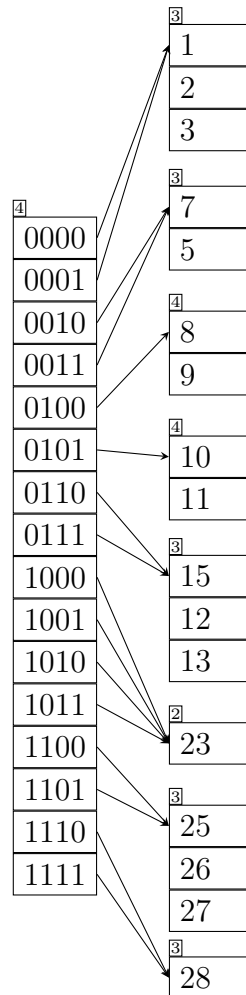


Abbildung 1: Index für Aufgabe 11.

Als Hash-Wert wird der fünfstellige Binärwert des Suchschlüssels verwendet. In jedes Bucket passen 3 Suchschlüssel.

Löschen Sie die folgenden Suchschlüssel (in dieser Reihenfolge) und zeichnen Sie den Index nach wichtigen Zwischenschritten (nach Verkleinerung des Verzeichnisses bzw. nach Zusammenführung von Buckets).

5, 7, 8, 12, 13, 9, 25