

Prof. Dr. T. Härder  
 Fachbereich Informatik  
 AG Datenbanken und Informationssysteme  
 Universität Kaiserslautern

## *Übungsblatt 7*

Unterlagen zur Vorlesung: „[www.dvs.informatik.uni-kl.de/courses/DBSREAL/](http://www.dvs.informatik.uni-kl.de/courses/DBSREAL/)“

### Aufgabe 1: Externes Hashing mit Separatoren

192

Gegeben seien 5 Buckets mit einer Kapazität von 3 Sätzen. In der folgenden Tabelle seine Zuordnungen von Schlüssel zu Signaturfolgen bzw. Sondierungen gegeben:

Schlüssel	h	s
a	( 1 3 5 )	( 1110 0100 0001 )
b	( 2 4 1 )	( 1011 1100 0101 )
c	( 4 )	( 0000 )
d	( 3 )	( 0001 )
e	( 3 5 )	( 0011 0011 )
f	( 3 )	( 0010 )
g	( 3 2 4 )	( 0011 1011 1000 )
h	( 5 3 )	( 1010 0001 )
i	( 2 4 5 )	( 1011 0111 0100 )
j	( 3 2 1 )	( 1110 1011 0001 )
k	( 5 )	( 0101 )
l	( 3 1 )	( 0011 1010 )

Nach der Initialisierung seien alle Buckets leer. Anschließend werden folgende Operationen ausgeführt:

- a) Einfügen der Sätze a, b, c, d, e, f, g, h, i, j
- b) Löschen der Sätze d, e, f
- c) Einfügen der Sätze k, l

Machen Sie sich anhand dieser Operationen die Funktionsweise des Externen Hashing mit Separatoren klar.

Welche Probleme treten auf, wenn bei einer gleichmäßigen Verteilung der Separatoren sehr viele Einfüge- und Löschoptionen ausgeführt werden?

**Aufgabe 2: Lineares Hashing**

Gegeben sei eine Datei mit 3 Buckets mit einer Kapazität von je 3 Datensätzen. Für die Folge von Hashfunktionen gelten folgende Funktionen:

$$h_i(K) = K \bmod (3 * 2^i) \quad \text{mit } i \in \{0, 1, 2, \dots, n\}$$

Zu Beginn seien alle Buckets leer. Fügen Sie als Schlüssel folgende Zahlen in der gegebenen Reihenfolge mittels linearem Hashing ein:

7, 3, 2, 6, 10, 12, 11, 8, 14, 9, 23, 19, 24, 20, 25, 30

Zur Kontrolle der Speicherplatzbelegung verwenden Sie folgende Verfahren:

a) Unkontrolliertes Splitting

b) Kontrolliertes Splitting mit  $\beta = 0,7$

Für beide Verfahren geben Sie jeweils den Zustand der Datei bzw. Buckets vor und nach jedem Split-Vorgang eines Buckets an.

**Aufgabe 3: Verweislisten und Bitlisten** Gegeben seien folgende Datensätze von Bestellungsinformationen, welche die Adressen  $Z_n$  mit

$n \in \{1, 2, \dots, 10\}$  besitzen:

Adressen	Bestellnummer	Bestelldatum	Kundennummer
Z1	711	01.02.1999	100
Z2	600	03.01.1999	102
Z3	801	01.02.1999	141
Z4	505	02.01.1999	100
Z5	475	02.01.1999	102
Z6	730	01.02.1999	102
Z7	621	05.01.1999	141
Z8	699	09.01.1999	102
Z9	670	09.01.1999	100
Z10	515	02.01.1999	180

Invertieren Sie die oben vorgegebenen Datensätze jeweils nach den Attributen Bestelldatum und Kundennummer mit folgenden Verfahren:

a) Verweislisten

b) Bitlisten

**Aufgabe 4: Bitlistenkomprimierung**

Gegeben sei eine Bitliste der Länge 200 mit folgenden Eigenschaften:

An den Positionen 1, 2, 3, 73, 76, 90, 119, 135, 136, 161 tritt der Wert '1' und sonst '0' auf.

Führen Sie für die oben angegebene Bitliste eine Komprimierung nach jedem der folgenden Verfahren:

- a) *Laufkomprimierung* mit  $k = 6$
- b) *Nullfolgenkomprimierung* mit Codiereinheiten fester Länge mit  $k = 6$
- c) *Nullfolgenkomprimierung* mit Codiereinheiten variabler Länge mit einer festen Länge des Längenfeldes  $l = 3$
- d) *Nullfolgenkomprimierung* unter der Anwendung der *Golomb-Codierung* mit  $m = 4$