

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

Übungsblatt 4

für die freiwillige Übung

Unterlagen zur Vorlesung: „www.dvs.informatik.uni-kl.de/courses/DBSREAL/“

Aufgabe 1: *Pointer-Swizzling* bei Bäumen

190

Gegeben ist ein vollständig belegter B*-Baum der Klasse $T(k, k^*, h)$, der in seinen Blattseiten auf Sätze verweist, die keine weiteren Verweise enthalten.

Die B*-Baum-Seiten besitzen folgendes Format:

Innerer Knoten	Z_0	R_1	Z_1	R_2	Z_2	\dots	R_p	Z_p	freier Platz
----------------	-------	-------	-------	-------	-------	---------	-------	-------	--------------

R_i = Referenzschlüssel (Wegweiser), $k \leq p \leq 2k$

Z_i = Verweis auf Sohnseite

Blattknoten	P	S_1	D_1	S_2	D_2	\dots	S_j	D_j	freier Platz	N
-------------	---	-------	-------	-------	-------	---------	-------	-------	--------------	---

P = PRIOR-Zeiger, N = NEXT-Zeiger, $k^* \leq j \leq 2k^*$

S_i = Satzschlüssel

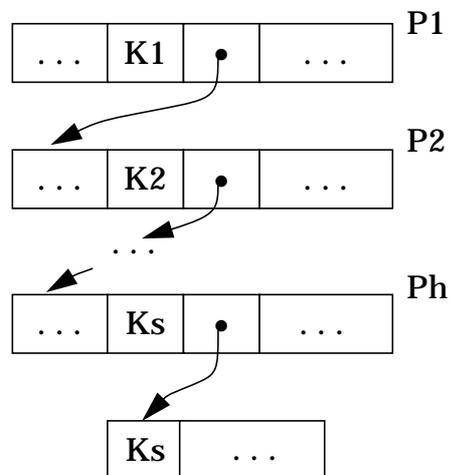
D_i = Verweis auf Satz (referenzierte Speicherung)

Eine innere Seite (inkl. der Wurzel) habe also bei voller Belegung $2k+1$ Verweise auf Seiten und eine Blattseite auf der Stufe h habe $2k^*$ Verweise auf Sätze.

Anfangs seien keine Seiten des B*-Baumes und keine der referenzierten Sätze im Hauptspeicher. Im folgenden soll Copy-Swizzling analysiert werden.

Skizzieren Sie für die verschiedenen Swizzling-Varianten die einzelnen Aktionen, die aus einem Suchvorgang für den Satz mit Schlüssel K_s resultieren, wenn dabei die Wurzelseite P_1 , auf Baumebene 2 die Seite P_2 , ... und auf Baumebene h die Seite P_h aufgesucht werden. Der Suchpfad ergebe sich durch die Schlüssel (Wegweiser) K_1 in

P1, K2 in P2, ..., Ks in Ph:



Wie viele Seiten, Deskriptoren und Sätze sind jeweils einzulagern und aufzusuchen, wenn als Verfahren

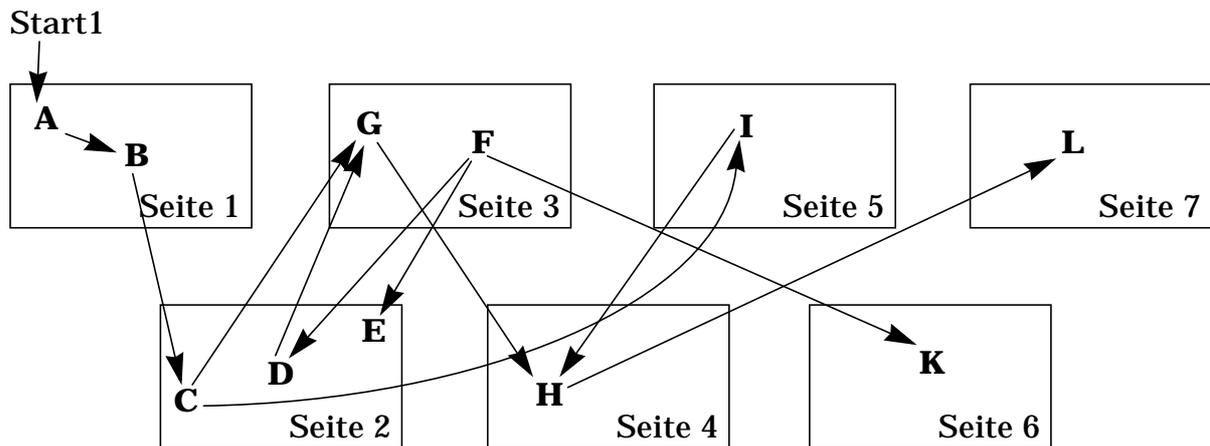
- a) eager, direct
- b) eager, indirect
- c) lazy, direct
- d) lazy, indirect

gewählt wird? „lazy“ bedeutet hier, daß bei erster Referenz ein Swizzling erfolgt.

Aufgabe 2: *Pointer-Swizzling (in-place, direct, lazy)*

187

Gegeben seien die Objekte A bis L, die auf externem Speicher in den Seiten 1 bis 7 wie unten abgebildet abgelegt sind. Die Pfeile repräsentieren dabei die Beziehungen der Objekte untereinander.



Es wird angenommen, daß diese Verweise als Externspeicheradressen vorliegen. Ein Transaktion referenziere die einzelnen Objekte in folgender Reihenfolge:

A B C G F E I H L K

Im folgenden werde immer von einem *Swizzling* mit den Eigenschaften *in-place*, *direct*, *lazy* ausgegangen, wobei bei der ersten Referenz auf eine Hauptspeicheradresse umgestellt wird.

a) Welche Auswirkungen haben die obigen Objektreferenzen auf die Hauptspeicherbelegung, wenn davon ausgegangen werden kann, daß alle Seiten in den Hauptspeicher passen?

b) Angenommen es stünden nur 5 Seiten im Hauptspeicher zur Verfügung. Beim Einlesen der Seite mit Objekt L in Seite 7 müßte dann Platz geschaffen werden. Welche Seite(n) könnte(n) verdrängt werden?

c) Wie wäre die Situation, wenn zwischenzeitlich in Seite 1 ein neues Objekt eingefügt worden wäre?

d) Welche Informationen für die einzelnen Seiten/Objekte werden beim *Pointer-Swizzling* benötigt?

Aufgabe 3: Abbildung von Datensätzen in Seiten

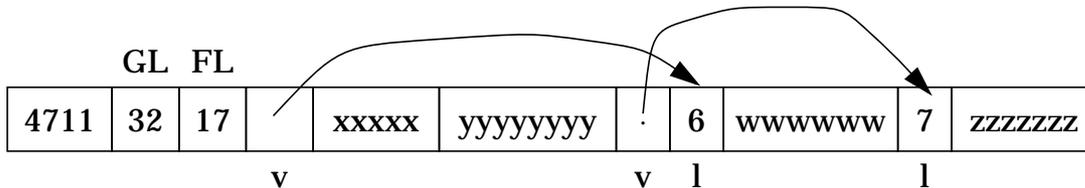
165

Jede Datenseite benötigt gewisse Informationen, die der Identifikation und der Selbstbeschreibung dienen. Sie werden im sogenannten *Page-Header* abgelegt. Überlegen Sie, welche Informationen zur Beschreibung einer Datenseite notwendig sind.

a) Entwerfen Sie für den *Page-Header* eine geeignete Datenstruktur und legen Sie den Typ der einzelnen Elemente dieser Struktur fest.

Die Sätze einer Tabelle TAB seien nach dem 4. Verfahren im Skript als Speicherstrukturen abgelegt. Anfangs haben sie 4 Felder und sind im Katalog durch v, f5, f8, v beschrieben.

Ein bestimmter Satz mit SKZ = 4711 habe folgende Repräsentation, wobei der Zeiger eines v-Feldes 2 Byte und die Längendarstellung eines v-Feldes (l) 1 Byte ausmachen.



Durch

Alter Table TAB (Add A-Feld var) und Alter Table TAB (Add B-Feld fixed (3))

werden dynamisch zwei Felder an TAB-Sätze angehängt, was zunächst nur Auswirkungen auf die Katalog-Information von TAB hat: v, f5, f8, v, v, f3. Die einzelnen Sätze von TAB werden erst bei ihrer Aktualisierung mit Werten für A-Feld und B-Feld modifiziert.

Welche Änderungen an der Speicherstruktur von Satz 4711 ergeben sich, wenn

- b) beide neuen Felder noch keinen Wert besitzen
- c) B-Feld den Wert bbb und A-Feld noch keinen Wert besitzt
- d) A-Feld den Wert 5aaaaa und B-Feld noch keinen Wert besitzt
- e) beide neuen Felder die obigen Werte haben?