

Prof. Dr.-Ing. Dr. h. c. T. Härder
 Fachbereich Informatik
 Arbeitsgruppe Datenbanken und Informationssysteme
 Universität Kaiserslautern

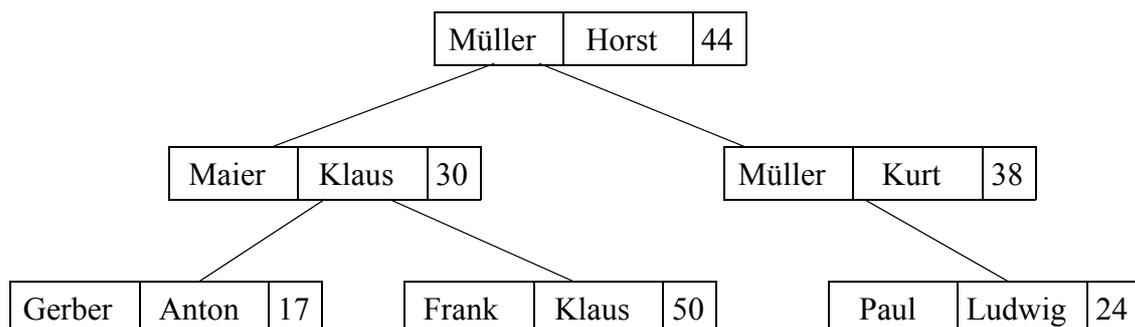
Übungsblatt 8 – Lösungsvorschläge

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

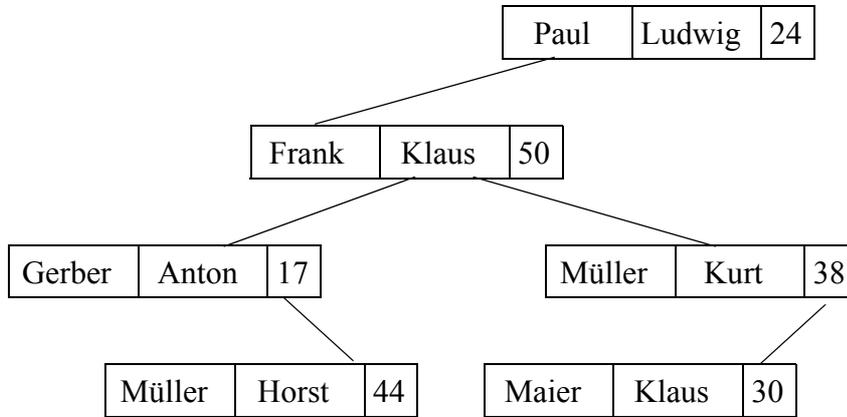
Aufgabe 1: k-d-Bauma)

a) Speichern Sie folgende Tupel in einem 3-d-Baum (k-d-Baum mit $k = 3$) mit den Schlüsselteilen „Nachname“, „Vorname“ und „Alter“ ab.

	Nachname	Vorname	Alter
1.	Müller	Horst	44
2.	Maier	Klaus	30
3.	Müller	Kurt	38
4.	Gerber	Anton	17
5.	Frank	Klaus	50
6.	Paul	Ludwig	24



b) Speichern Sie zum Vergleich die Tupel in umgekehrter Reihenfolge (6 -> 1) ab.



c) Suchen Sie in den sich aus a) und b) ergebenden Bäumen Tupel, deren Schlüssel die folgenden Werte haben:

1. Müller
2. Gerber, Anton

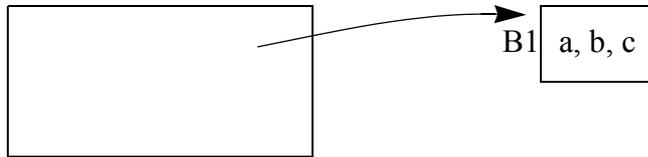
Aufgabe 2: Grid-File

Machen Sie sich die dynamischen Reorganisationen im GRID-File bei Einfügungen und Löschungen anhand des folgenden Beispiels klar. Ein Bucket kann jeweils 3 Sätze aufnehmen, das GRID-File ist anfangs leer. Von der Relation AUTO (KFZ-NR, MARKE, FARBE) sollen die beiden Attribute MARKE und FARBE mit der GRID-File-Technik abgespeichert werden.

- Zeichnen Sie alle wesentlichen Strukturen (*Directory*, Suchraum, *Buckets*) nach jeder der unten angegebenen Einfügungen.
- Löschen Sie danach die ersten 4 Einträge und zeichnen Sie die obigen Strukturen auf.

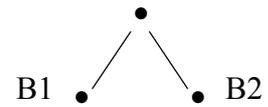
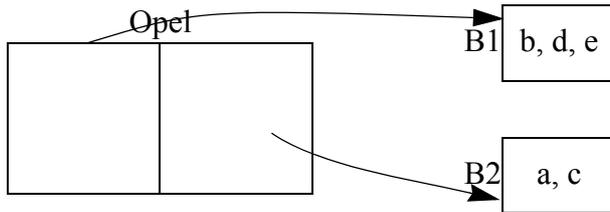
	<u>KFZ-NR.</u>	MARKE,	FARBE
a)	KL-PP 1	OPEL	GELB
b)	PS-A17	FORD	BLAU
c)	KL-CX 33	OPEL	BLAU
d)	KIB-AM 13	BMW	ROT
e)	SB-F16	AUDI	GRÜN
f)	KL-DZ 12	ALFA	SCHWARZ
g)	ZW-AL 43	VW	WEISS
h)	HOM-C 1	FIAT	VIOLETT
i)	SLS-AF 47	AUDI	AZUR
j)	KL-DM 31	ALFA	BEIGE
k)	VW-FS 40	SAAB	BRAUN

Eintragung a, b, c

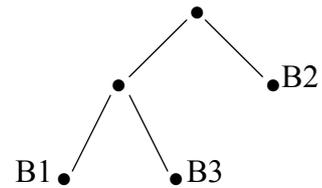
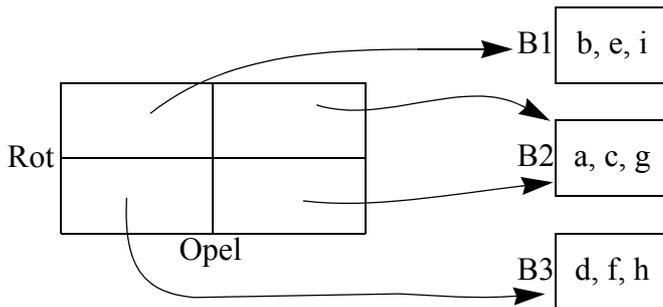


B1 •

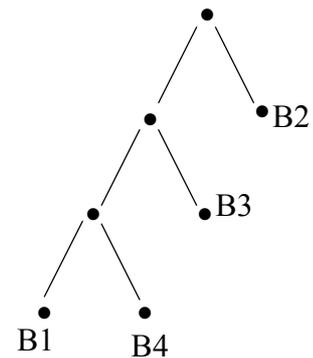
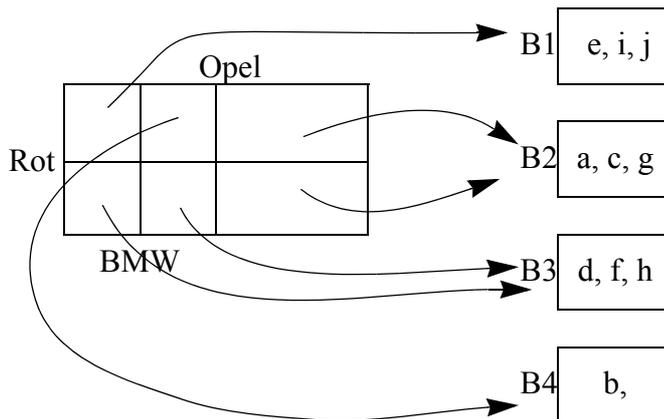
Eintragung von d führt zu Split bei „>= Opel“,
Eintragung e



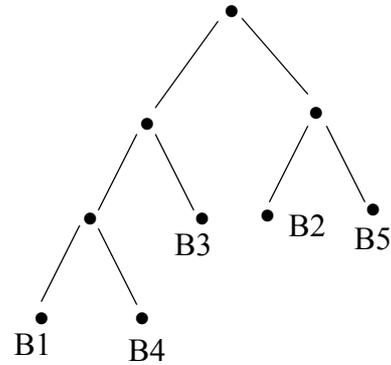
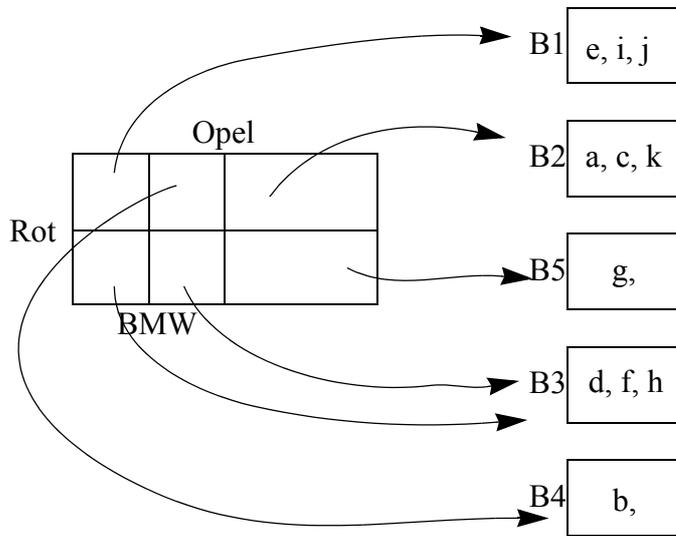
Eintragung f führt zu Split bei „>= Rot“,
Eintragung g, h, i



Eintragung j führt zu Split bei „>= BMW“,

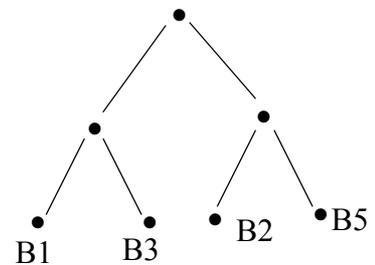
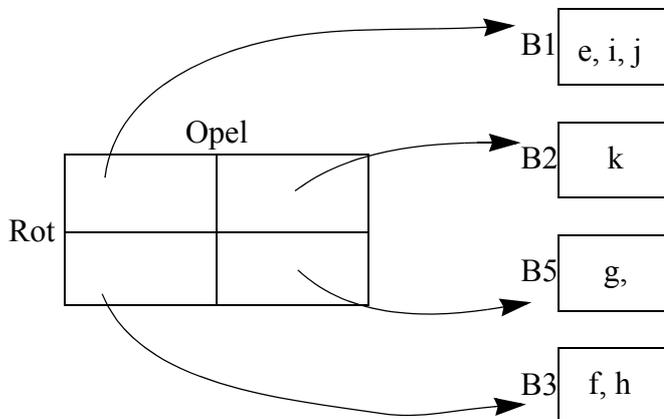


Eintragung von k führt zu Split bei „>= Rot“



Nach dem Löschen der ersten 4 Einträge sieht die Struktur wie folgt aus:

B2 und B5 werden nicht verschmolzen, da bei einer erneuten Einfügung dieselbe Aufteilung wieder durchgeführt werden müßte. B2 und B5 werden nur dann verschmolzen, wenn gleichzeitig auch B1 und B3 verschmolzen werden kann.



Nur verschmelzen, wenn:

1. die Buckets Zwillinge im Baum sind und
2. ein Bucket leer ist oder eine Trennlinie in einer Dimension verschwindet(, was aber sehr schwer zu überprüfen ist).

Die Suche im R-Baum liefert uns als mögliche Kandidaten die Figuren F3 sowie F5. Da aber die umfassenden Rechtecke nicht exakt der Figurkontur entsprechen muß, muß für die gefundenen Figuren einzeln das Enthaltensein des Punktes überprüft werden, so daß letztendlich lediglich F3 übrigbleibt.

b) Welche Figuren sind in dem durch die Punkte $(3.5, 0)$, $(8, 5.5)$ bestimmten Rechteck (Fenster) vollständig enthalten?

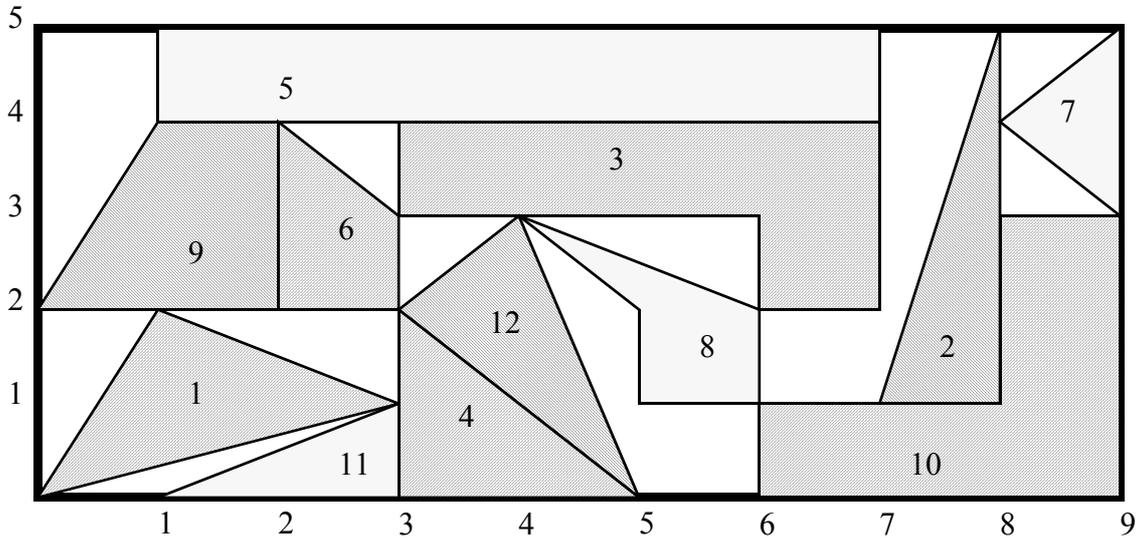
Diese Frage ist unter ausschließlicher Nutzung der R-Baum-Einträge zu beantworten und resultiert in F3, F4

c) Bestimmen Sie alle mit dem Fenster $(2, 2)$, $(4, 4)$ überlappenden Figuren.

Hier liefert uns der R-Baum wiederum nur mögliche Kandidaten, nämlich F1 und F3. Für diese muß wieder konkret überprüft werden, ob sie überhaupt in dem Suchfenster liegen, was hier der Fall ist.

Aufgabe 4: R-Baum - Suchverfahren

Gegeben sei folgende Anordnung von Flächen.

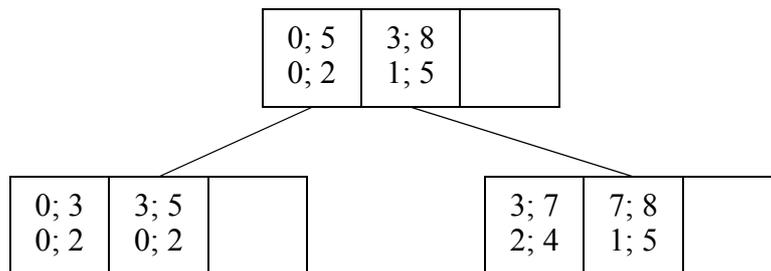


a.) Tragen Sie diese in der durchnummerierten Reihenfolge in einen R-Baum mit den Parametern $M=3$ (maximale Anzahl von Einträgen in einem Knoten) und $m=2$ (minimale Anzahl von Einträgen in einem Knoten) ein.

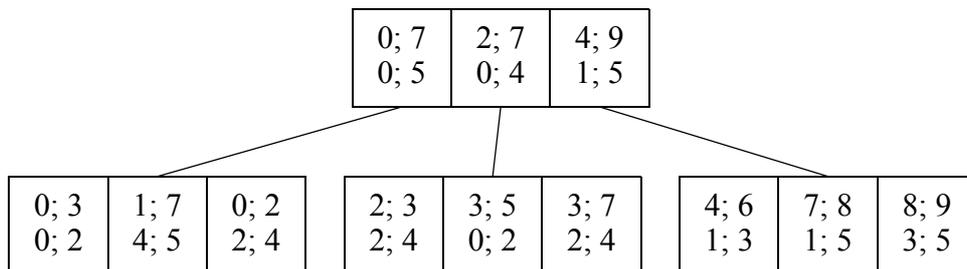
Nach dem Eintragen der Flächen 1, 2, 3

0; 3	7; 8	3; 7
0; 2	1; 5	2; 4

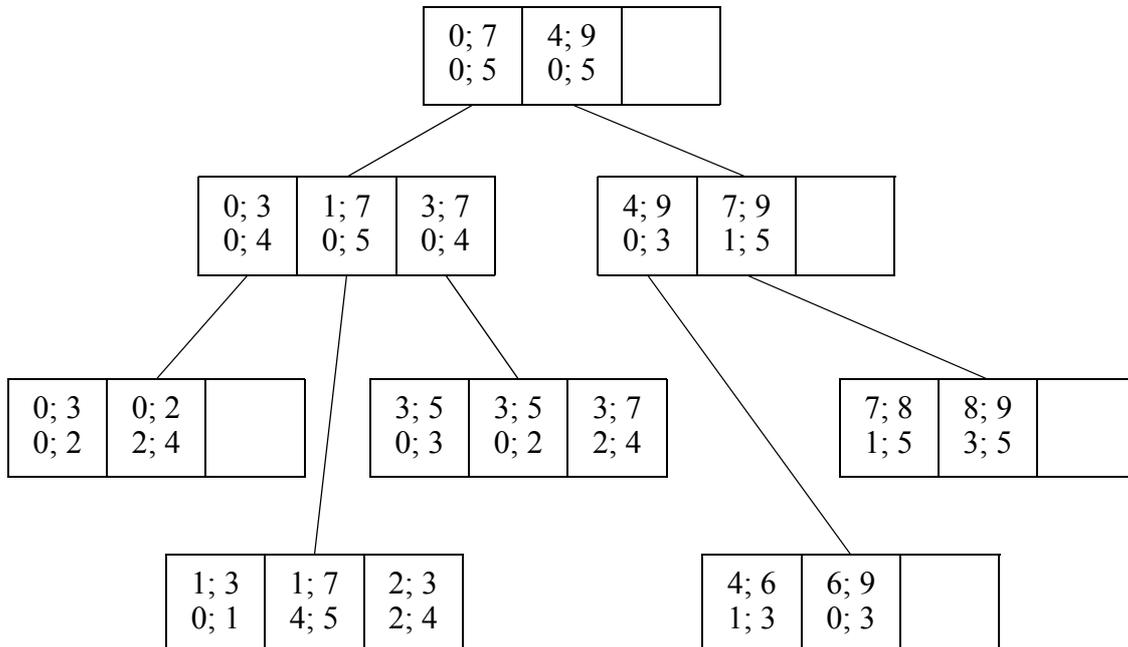
Nach dem Eintragen von Fläche 4



Nach dem Eintragen von Flächen 5, 6, 7, 8, 9
(Beim Eintragen von F8 wird F3 verschoben)



Nach dem Eintragen aller Flächen



b.) Schreiben Sie die Funktionen

`FINDE_ALLE_VOLLSTAENDIG_ENTHALTENEN_FLAECHEN (Rechteck)`

Funktion braucht nur auf dem R-Baum zu arbeiten, die konkreten Figuren interessieren nicht, da diese vollständig durch die umgebenden Rechtecke beschrieben werden.

`FINDE_ALLE_TEILWEISE_ENTHALTENEN_FLAECHEN (Rechteck)`

Der R-Baum kann uns hierbei nur die Flächen liefern, die möglicherweise das gegebene Rechteck schneiden. Für jede dieser Flächen ist eine konkrete Überprüfung notwendig!

die alle Flächen findet, die vollständig resp. teilweise in dem gegebenen Rechteck enthalten sind.

c.) Testen Sie die Funktionen mit dem Rechteck, daß durch die Eckpunkte (3,1) und (7,4) beschrieben wird.

Vollständig: 3, 8

Teilweise: 12, 4 (+ vollständig 3, 8)

d.) Welche Verbesserung des R-Baums erlaubt eine einfachere Suche?

Der R*-Baum, da der Suchraum durch überlegte Zerlegung in Teilflächen eingeschränkt wird. (Hier eine Zerlegung bei $x=3$ und $x=7$)

Beachte aber größeren Overhead in der Baumorganisation