

Indexierung

Ungeordnete Anordnung in der Tabelle:

PUNKTNUMMER	LAGE	HÖHE	...
...
P-128
P-129
P-130
...

Geordneter Zugriff trotz ungeordneter Anordnung in der Tabelle:

PUNKTNUMMER	INDEX
...	...
P-128	I2
P-129	I1
P-130	I5
...	...

INDEX	LAGE	HÖHE	...
I1
I2
I3
I4
I5

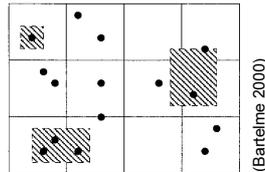
(Bartelme 2000)

Räumliche Indexierung

- Zweck: Ermöglichen einen schnellen Zugriff auf die Daten
- Index-Files mit 2 Spalten: Key-Value und Adresse der Seite im Datenfile
- Index-Files sind i.d.R. klein und deshalb schnell zu durchsuchen
- I.d.R. sortiert (Raum-füllende Kurven) oder eine bestimmte Struktur aufweisend, z.B. Gridfile, Suchbäume

Raumaufteilung (Punkte)

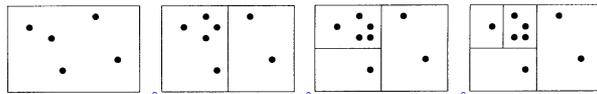
- Grobe Einteilung des (2D-)Raums zur Beschleunigung des Zugriffs
- Starres Schema:



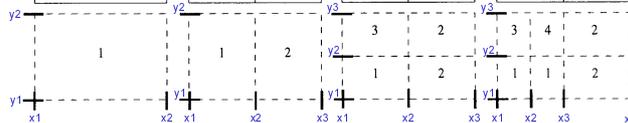
- Dynamisches Schema:
 - Gridfile
 - Quadtree
 - B-Tree
 - R-Tree

Gridfile

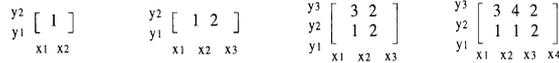
Datenraum:



Adressraum:



Adressfeld:



Griddirectory:

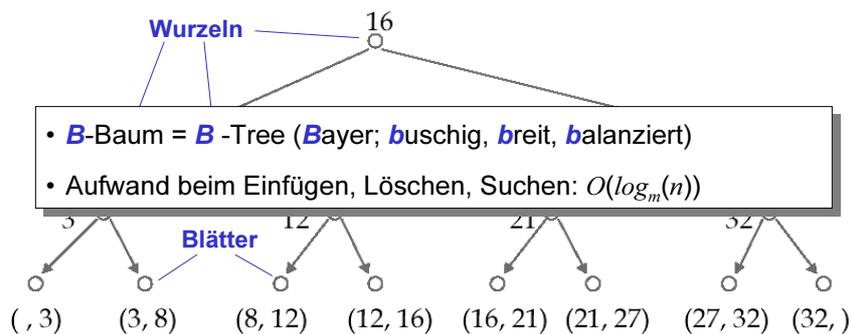
	$x_1 \leq x < x_2$	$x_2 \leq x < x_3$	$x_3 \leq x < x_4$
$y_2 \leq y < y_3$	3	4	2
$y_1 \leq y < y_2$	1	1	2

(Bartelme 2000)

Gridfile

- Aufteilung eines n -dimensionalen Datensatzes in rechteckige Blöcke
- Speicherung der Daten eines Blocks in "Buckets" fester Größe mit ≥ 1 (benachbarten) Blöcken pro Bucket (i.d.R. 10 bis 1000 Datensätze pro Bucket)
- Blöcke können unterschiedliche Größe haben
- bei Änderung der Datenmenge pro Bucket können Buckets aufgeteilt oder benachbarte Buckets zusammengefasst werden; jedoch müssen Blöcke konvex (rechteckig) sein
- Griddirectory: Zuordnung von Blöcken zu Buckets

B-Baum: Prinzip

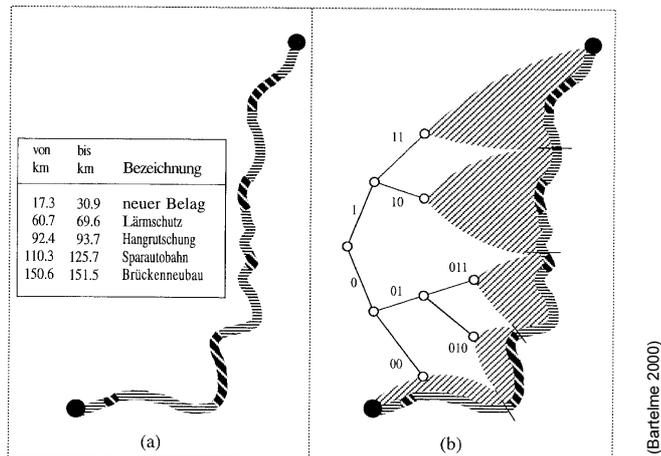


Blätter = Speicherblätter = Pages = Datenbuckets
= zusammenhängende Bereiche im Speicher

(<http://www.informatik.uni-rostock.de/FB/Praktik/psue/Lehrer/files/plo/ien00-01.pdf>)

B-Baum für ein 1D-Objekt

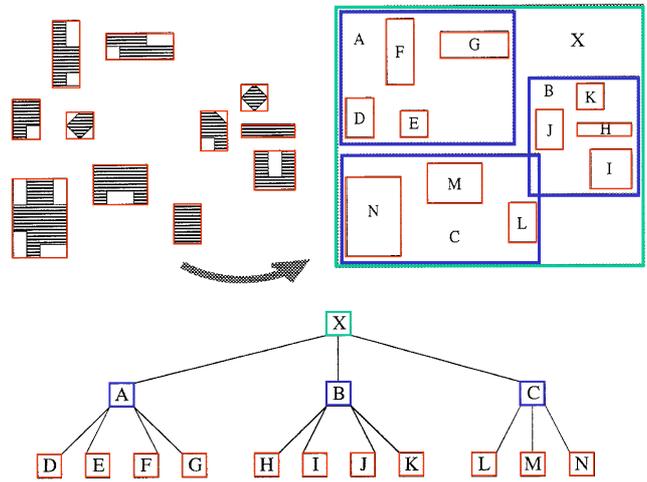
Bsp.: Straße



R-Baum

- Guttmann (1984)
- Erweiterung des B-Baums für k Dimensionen
- eine räumliche Zugriffsmethode, die den Raum mit hierarchisch genisteten, eventuell sich deckenden, Kästen aufspaltet
- Objekte werden durch die minimalen umgebenden Rechtecke (MBR = minimum bounding rectangle) repräsentiert
- Überlappungen der Regionen sind möglich

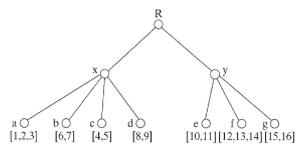
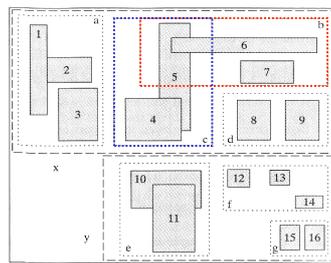
R-Baum für 2D-Objekte



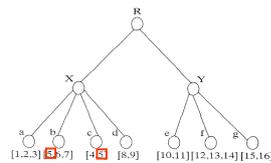
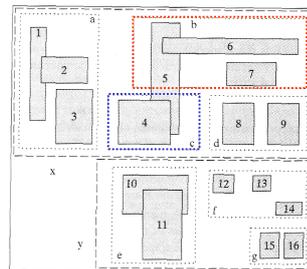
(Bartelme 2000)

Vergleich

R-Baum



R+-Baum



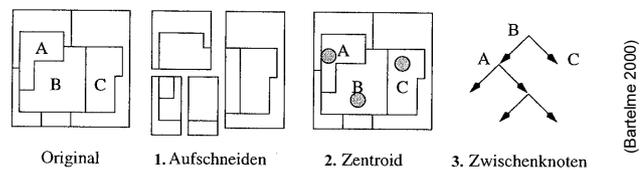
(Adam und Gangopadhyay 1997)

Raumaufteilung (Objekte)

Problem: Objekte werden von Grenzlinien der Blätter geschnitten

Verschiedene Ansätze:

1. „**Aufschneiden**“ der Objekte entlang der Blattgrenzen
2. Zuordnung der Objekte zu ihren **Zentroiden** (= Punkte)
3. Speicherung der Objekte in den **Zwischenknoten** eines Baums, dem sie gerade noch zerschneidungsfrei zugeordnet werden können



4. Darstellung der Objekte als Punkte in einem **höherdimensionalen Parameterraum**, z.B. in dem 4D-Raum $(x_{\min}, x_{\max}, y_{\min}, y_{\max})$

Raumaufteilung (Objekte)

4. Darstellung der Objekte als *Punkte* in einem **höherdimensionalen Parameterraum**, z.B. in dem 4D-Raum $(x_{\min}, x_{\max}, y_{\min}, y_{\max})$

Bsp.: 1D Datenraum \rightarrow 2D-Parameterraum

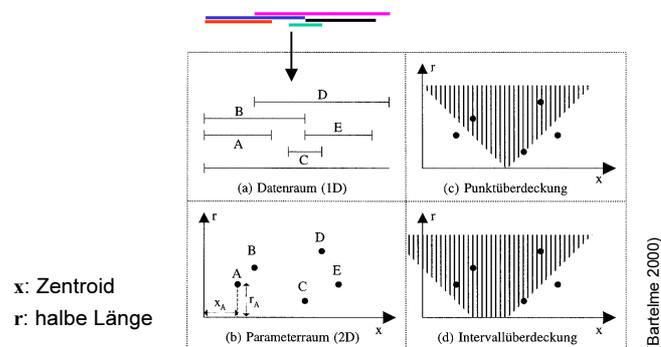


Abbildung 8.15: Datenraum (1D), Parameterraum (2D) und Suchbereiche