

GEODATENBANKEN

Geo-Informationssysteme

"Ein Geoinformationssystem ist ein rechnergestütztes System, das aus Hardware, Software, Daten und den Anwendungen besteht. Mit ihm können raumbezogene Daten digital erfasst, gespeichert, verwaltet, aktualisiert, analysiert und modelliert sowie alphanumerisch und graphisch präsentiert werden. " [2]

Synonyme:

geographisches Informationssystem

Geoinformationssystem

GIS

Geographic bzw. Geographical Information System

Speicherung der Geodaten innerhalb eines GIS in Geodatenbank (es gibt auch andere Ansätze für die Speicherung).

Geodatenbanken - Vertreter

- „Oracle Spatial“, „Oracle Locator“
- IBM DB2 DataExtender - „Spatial Extender“
- MSSQL Server ab 2008
- Sybase - „SQL Anywhere for Spatial Data“
- PostgreSQL - „PostGIS“
- MySQL „Spatial Extensions“

Geodaten

Eigenschaften

- geometrisch
 - Ausdehnung und Lage im Raum, Form
- topologisch
 - relative räumliche Beziehungen von Geoobjekten zueinander
- thematisch
 - nominal (Namen, Kodierungen)
 - quantitativ (endliche Aufzählungstypen, z.B. Tag der Müllabfuhr)
 - qualitativ (Mess- und Berechnungswerte aus im Prinzip unendlichen Zahlenbereich, z.B. Einwohnerdichte)
- Zeitbezug
 - Gültigkeitszeitpunkte oder -räume
 - Dynamik des Geoobjektes
 - „moving objects“
 - spatio-temporale Datenmodelle

Metadaten

- semantisch
 - fachliche Aspekte
 - erlauben Beurteilung der Datenqualität
 - z.B. Titel, Inhaltsangabe, Aufnahmegeräte, Ortsbezug, Erhebungszweck, Urheber, Erhebungsdatum
- syntaktisch
 - Aufbau der Daten
 - Aussagen zur Qualität des Datenmodells
 - für den Datenaustausch erforderlich

- z.B. Datenmodell, Datenformat
- „weitere“
 - z.B. rechtliche Nutzungsaspekte, Referenzprojekte

Standardisierung von Geodatenmodellen

Simple Feature Modell

- ISO 19125: Simple Feature Access
 - Teil 1 - Common Architecture
 - Teil 2 - SQL Option
- Simple Features
 - maximal zwei Dimensionen
 - geradlinig verbunden
- definiert räumliche Operationen für Zugriff, Anfrage und Verarbeitung

„SQL/MM -Spatial“

- SQL/MM - ISO 13249
- Anwendungsgebiete - SQL-Erweiterungen
- allgemeiner Teil mit generellen Konzepten + vier Unterbereiche
 - „SQL/MM Full-Text“ - Suche in Dokumenten
 - „SQL/MM Still Image“ - Speicherung von Rasterbildern
 - „SQL/MM Data Mining“ - Data-Warehouse-Anwendungen
 - „SQL/MM Spatial“ für Geodaten
- „SQL/MM Spatial“ für Geodaten ähnlich OGC Simple Feature Model
- hier auch Datentypen für Kurven

Funktionalität von Geodatenbanken - Vergleich von Geometrien

9-Intersection-Modell

Das Simple Feature Model definiert in der Klasse Geometry eine Reihe von Methoden, die prüfen ob eine topologische Beziehung zwischen zwei Geometrien vorliegt.

Es erfolgt eine Zerlegung der Geometrie in drei Bereiche, ein Punkt im Datenraum liegt in genau einem der drei Bereiche.

- Rand
 - für Punkte: Rand leer
 - für Linien: Endpunkte der Linie
 - für Flächen: äußerer und innerer Ring
 - allgemein: Rand ist ein Geoobjekt der nächst-niedrigeren Dimension
- Inneres
 - für Punkte: Punkt selber
 - für Linien: Linie ohne Endpunkte
 - für Flächen: Punkte, die innerhalb des äußeren Rings, aber nicht innerhalb des inneren Rings liegen
 - allgemein: Inneres ist Geoobjekt ohne Rand
- Äußeres - Datenraum ohne Rand und Inneres

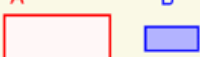
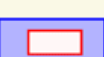


Die drei Bereiche werden paarweise darauf geprüft, ob sie sich in irgendeiner Weise schneiden.





$$A_x \cap B_y = \text{true} \Leftrightarrow A_x \cap B_y \neq \emptyset \Leftrightarrow \exists p \in A_x : p \in B_y$$

Das Ergebnis der Prüfung wird in der Intersection Matrix zusammengestellt.

$$A \otimes B = \begin{bmatrix} A_I \cap B_I & A_I \cap B_R & A_I \cap B_A \\ A_R \cap B_I & A_R \cap B_R & A_R \cap B_A \\ A_A \cap B_I & A_A \cap B_R & A_A \cap B_A \end{bmatrix}$$

Von rechnerisch 2^9 Kombinationen sind nur 8 sinnvoll.

A disjoint B	$\begin{bmatrix} \text{false} & \text{false} & \text{true} \\ \text{false} & \text{false} & \text{true} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	
A inside B	$\begin{bmatrix} \text{true} & \text{false} & \text{false} \\ \text{true} & \text{false} & \text{false} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	
A contains B	$\begin{bmatrix} \text{true} & \text{true} & \text{true} \\ \text{false} & \text{false} & \text{true} \\ \text{false} & \text{false} & \text{true} \end{bmatrix}$	
A equals B	$\begin{bmatrix} \text{true} & \text{false} & \text{false} \\ \text{false} & \text{true} & \text{false} \\ \text{false} & \text{false} & \text{true} \end{bmatrix}$	

A meets B	$\begin{bmatrix} \text{false} & \text{false} & \text{true} \\ \text{false} & \text{true} & \text{true} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	
A covers B	$\begin{bmatrix} \text{true} & \text{true} & \text{true} \\ \text{false} & \text{true} & \text{true} \\ \text{false} & \text{false} & \text{true} \end{bmatrix}$	
A coveredBy B	$\begin{bmatrix} \text{true} & \text{false} & \text{false} \\ \text{true} & \text{true} & \text{false} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	
A overlaps B	$\begin{bmatrix} \text{true} & \text{true} & \text{true} \\ \text{true} & \text{true} & \text{true} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	

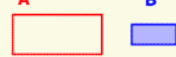
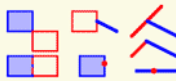


Es ergeben sich folgende Äquivalenzen

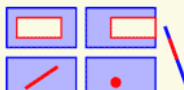


- A disjoint B \leftrightarrow B disjoint A
- A inside B \leftrightarrow B contains A
- A equals B \leftrightarrow B equals A
- A meets B \leftrightarrow B meets A
- A covers B \leftrightarrow B coveredBy A
- A overlaps B \leftrightarrow B overlaps A

DE-9-Intersection-Modell

Dieses Modell erlaubt differenziertere Aussagen als das 9-Intersection-Modell. In der Intersection Matrix werden Aussagen zur Dimension des Schnittes hinterlegt:

- Schnitt wird durch einzelnen Punkt beschrieben: 0
- Schnitt wird durch Linien beschrieben: 1
- Schnitt wird durch Flächen beschrieben: 2
- es liegt kein Schnitt vor: -1
- es liegt (irgendein) Schnitt vor: T
- es ist egal, ob ein Schnitt vorliegt („don't care“): *

A disjoint B	$\begin{bmatrix} \text{F} & \text{F} & \text{*} \\ \text{F} & \text{F} & \text{*} \\ \text{***} \end{bmatrix}$	
A touches B ($d(A) > 0 \vee d(B) > 0$)	$\begin{bmatrix} \text{F} & \text{T} & \text{*} \\ \text{***} \vee \begin{bmatrix} \text{F} & \text{***} \\ \text{T} & \text{***} \end{bmatrix} \vee \begin{bmatrix} \text{F} & \text{***} \\ \text{*} & \text{T} \end{bmatrix} \\ \text{***} \end{bmatrix}$	
A crosses B ($d(A) < d(B)$)	$\begin{bmatrix} \text{T} & \text{*} & \text{T} \\ \text{***} \\ \text{***} \end{bmatrix}$	
A crosses B ($d(A) = d(B) = 1$)	$\begin{bmatrix} \text{0} & \text{*} & \text{*} \\ \text{***} \\ \text{***} \end{bmatrix}$	

A within B	$\begin{bmatrix} \text{T} & \text{*} & \text{F} \\ \text{***} & \text{F} \\ \text{***} \end{bmatrix}$	
A overlaps B ($d(A) = d(B)$, $d(A) \neq 1$, $d(B) \neq 1$)	$\begin{bmatrix} \text{T} & \text{*} & \text{T} \\ \text{***} \\ \text{T} & \text{*} & \text{*} \end{bmatrix}$	
A overlaps B ($d(A) = d(B) = 1$)	$\begin{bmatrix} \text{1} & \text{*} & \text{T} \\ \text{***} \\ \text{T} & \text{*} & \text{*} \end{bmatrix}$	

Äquivalenzen

- A contains B \leftrightarrow B within A
- A intersects B $\leftrightarrow \neg$ (A disjoint B)

Das Simple Feature Modell stellt Methoden zur Verfügung, mit denen geprüft werden kann, ob für zwei Geometrien eine bestimmte topologische Beziehung vorliegt.

Des Weiteren steht mit "relate" eine Methode zur Verfügung, der zwei Geometrien und eine beliebige DE-9-Intersection Matrix zur Prüfung übergeben werden können. Auch kann die DE9IM für zwei gegebene Geometrien ermittelt werden

Räumliche Datenbankankfragen

räumliche Basisanfragen

Enthält eine Anfrage eine oder mehrere Operationen mit Raumbezug, so wird die Anfrage auf eine oder eine Folge von räumlichen Basisanfragen zurückgeführt.

- Punktabfrage/point query
bestimmt für einen gegebenen Punkt p alle Geoobjekte, die p enthalten
- Rechteckanfrage/window query
bestimmt für gegebenes Rechteck r alle Geoobjekte, die r schneiden
- Regionsanfrage/region query
bestimmt für gegebenes Polygon alle Geoobjekte, die pol schneiden
- Abstandanfrage/distagepance query
bestimmt alle Geoobjekte, die sich in vorgegebener Entfernung $dist$ zu einem Abfragepunkt p befinden (entspricht kreisförmiger Regionsabfrage mit Radius $dist$)
- Bestimmung der nächstgelegenen Nachbarn/nearest neighbour query, NQQ
bestimmt zu gegebenem Geoobjekt die nächstgelegenen Objekte
- geometrischer Verbund/spatial join
Verbundoperation zwischen Relationen, die mindestens eine geometrische Verbundbedingung enthält

mehrstufige Anfragebearbeitung

- mehrere Filterschritte
 - Reduzierung der Menge potentieller Kandidaten
 - Identifizierung sicherer Treffer
- dabei Nutzung schnell ausführbarer Algorithmen und Datenrepräsentationen in den frühen Schritten
- ein finaler Verfeinerungsschritt mit Test der exakten Bedingung auf durch Filterschritte reduzierter Kandidatenmenge

Verwendung von Approximationen

Verwendung vereinfachter Geometrien

- minimal umgebendes Rechteck (MUR/Bounding Box)
- konvexe Hülle

Verwendung verschiedener Approximationsverfahren

- konservative/progressive Approximation
 - konservativ
 - enthalten Geometrie des approximierten Objektes vollständig
 - geeignet zum Eliminieren von falschen Treffern
 - progressiv
 - sind vollständig in Objektgeometrie enthalten
 - geeignet zum Identifizieren sicherer Treffer
- ein-/mehrelementige Approximation

Dabei gelten folgende Trade-Offs

1. Je kleiner die Fehlfläche ist, desto ...
 - ... häufiger können Anfragen ohne Betrachtung der exakten Objektgeometrie beantwortet werden.
 - ... besser ist die Approximationsgüte.

Je komplexer die Approximation, desto ...

 - ... besser ist die Approximationsgüte und desto wirksamer ist der entsprechende Filterschritt.
 - ... aufwändiger ist ihre Berechnung und der Verarbeitungsaufwand und desto mehr Speicherplatz benötigt sie.
2. Je mehr Elemente verwendet werden, desto ...

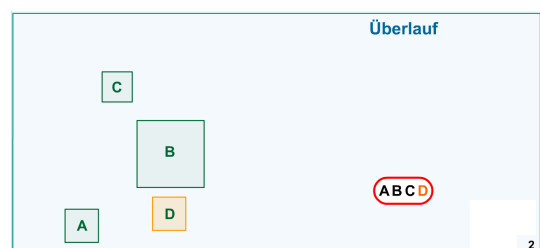
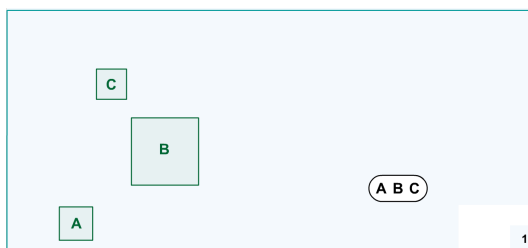
- ... besser ist die Approximationsgüte.
- ... wirksamer ist der entsprechende Filterschritt.
- ... höher ist der Verarbeitungsaufwand.
- ... mehr Speicherplatz benötigt die Approximation.

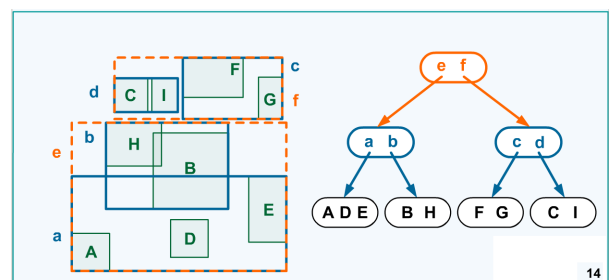
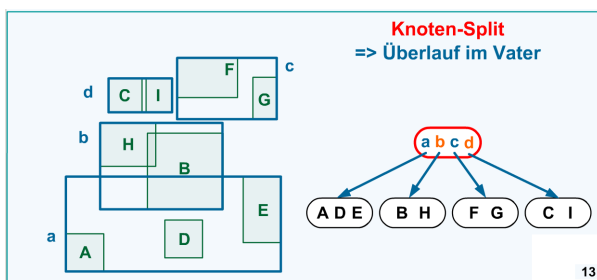
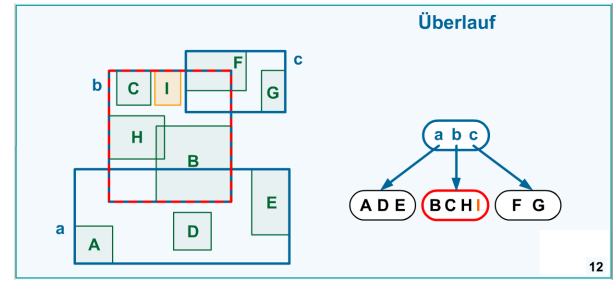
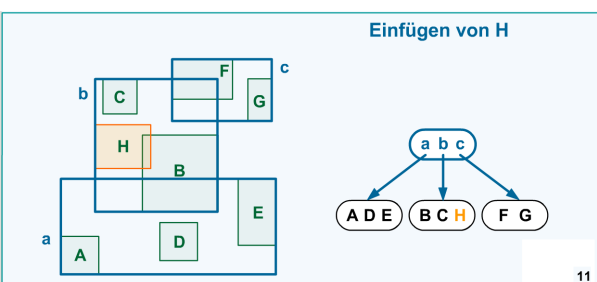
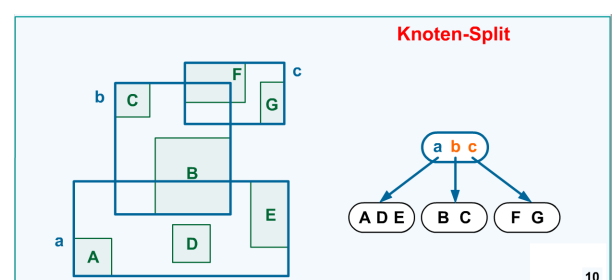
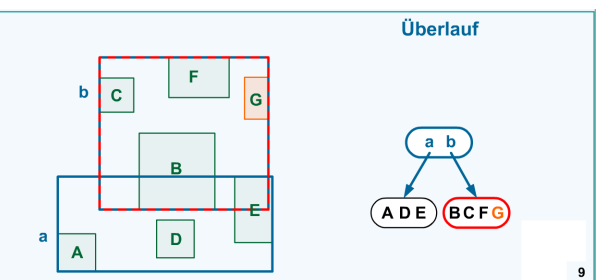
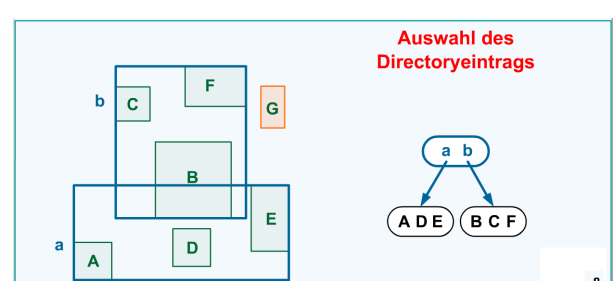
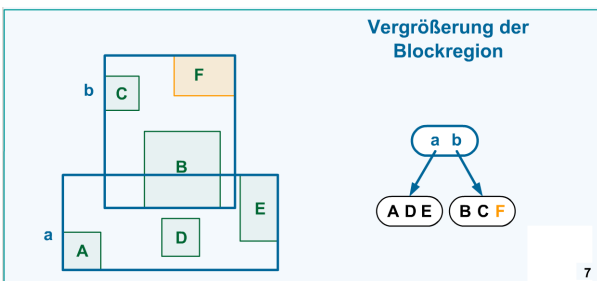
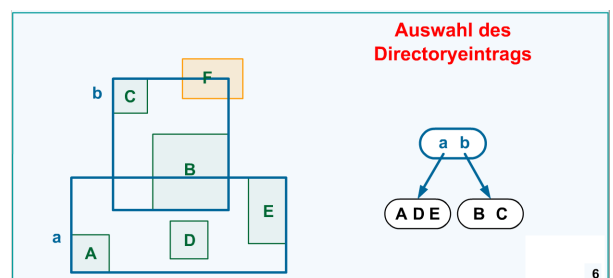
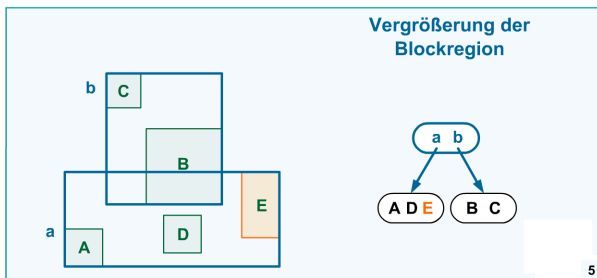
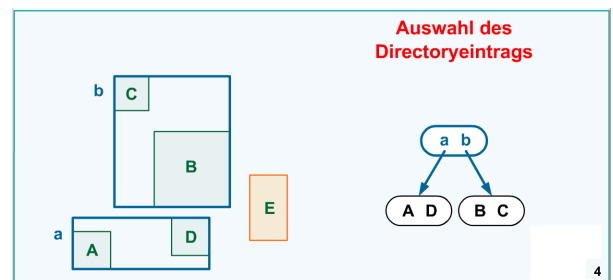
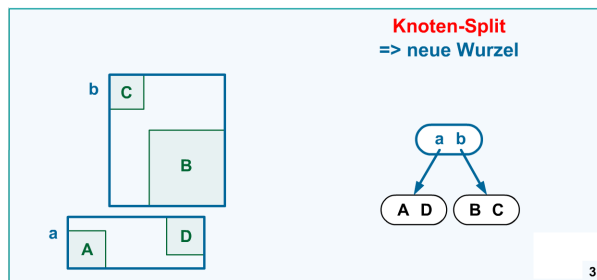
Nutzung räumlicher Indexstrukturen

- übliche Indizierungsverfahren ungeeignet
 - Vergleichsoperation teuer
 - keine lineare Ordnung
- Index arbeitet auf MUR
- Anspruch an räumlichen Index
 - Approximationen verwalten
 - Basisanfragen effizient
 - dynamisches Einfügen, Löschen, Verändern
 - gute Speicherplatzausnutzung
 - Indexstruktur robust gegen Ungleichverteilung der Geoobjekte im Datenraum (Speicherplatzausnutzung, Bearbeitungszeit)
- verschiedene Verfahren werden kombiniert: Clipping, Punkttransformation, Einbettung, Überlappung, Mehrschichtentechnik

Räumliche Indexe

- Quadrantenbäume
 - räumliche Datenstruktur
 - teilen k-dimensionalen Datenraum in 2^k gleich große Zellen
 - können Punkte, Linien oder Flächen verwalten
- R-Bäume
 - balancierter Baum
 - R+-Baum, R*-Baum
 - wichtige Operationen beim Einfügen, bestimmen das Leitungsverhalten des Baums maßgeblich
 - "ChooseSubtree" - drei Fälle
 - innerhalb genau einer Blockregion des aktuell betrachteten Verzeichnisknotens
 - innerhalb mehrerer Blockregionen → wähle Blockregion mit der geringsten Fläche aus
 - in keiner der Blockregionen vollständig enthalten → wähle die Region, die den geringsten Flächenzuwachs erfährt
 - Überlauf - „Split“
 - bei Überlauf ist nicht eindeutig klar, welche Einträge wohin
 - beim Split-Algorithmus unterscheiden sich die R-Baum-Varianten
 - mögliche Ziele
 - Minimierung der Überlappung zwischen den Blockregionen
 - Minimierung der Fläche der Blockregionen
 - Minimierung des Rands der Blockregionen
 - Maximierung der Speicherplatzausnutzung
 - Einfügen im R-Baum





PostGIS

PosGIS Eckdaten

- Extension zum DBMS PostgreSQL
- Version 1.0 von 2004, aktuell 2.0 (2.1)
- standardkonform
 - Simple Feature Model
 - SQL/MM

Datentypen

- GEOMETRY, GEOMETRYCOLLECTION
- POINT, POINTZ, MULTIPOINT
- LINESTRING, MULTICURVE, MULTILINESTRING
- POLYGON, MULTISURFACE, MULTIPOLYGON
- CIRCULARSTRING, COMPOUNDCURVE
- CURVEPOLYGON
- GEOGRAPHY

Ausgabeformate

- WKT - well known text
- WKB - well known binary
- SVG - scalable vector graphics
- GML - geography markup language
- KML - keyhole markup language

Methoden

- Abfrage von Geometrieeigenschaften
- Validierung von Geometrien
- geometrische Funktionen
- metrische Funktionen
- Test topologischer Prädikate
- Verschneidung von Geometrien
- geometrische Aggregatfunktionen

Quellen und weiterführende Links

- [1] „Geodatenbanksysteme in Theorie und Praxis: Einführung in objektrelationale Geodatenbanken unter besonderer Berücksichtigung von Oracle Spatial“, Brinkhoff, T., 2013
- [2] „Geoinformatik in Theorie und Praxis“, De Lange, N., 2006
- [3] „OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 1: Common Access & Part 2: SQL option“, <http://www.opengeospatial.org/standards>
- [4] PostGIS Documentation, <http://postgis.net/docs/manual-2.1/>
- [5] „Fernstudienmaterialien Geoinformatik“: „Geodatenbanken“, „Geodatenstandards“, <http://www.fergi.uni-osnabrueck.de/moodle/>