

Einführung in Geographische Informationssysteme

Einführungsvorlesung zur Ausbildung
“Anwendung geographischer Informationssysteme in Ökologie und Umweltschutz”

Universität Ulm, WS 1995/1996

Wolf-Fritz Riekert, FAW Ulm

Inhalt

- ◆ Geoinformation / Geodaten
- ◆ Geoinformationssysteme (GIS)
- ◆ Verfahren
- ◆ Anwendungen

Geoinformation

- ◆ Arten (thematisch, geometrisch / topologisch, kartographisch)
- ◆ Rasterdaten / Vektordaten
- ◆ Geoobjekte
- ◆ Objektarten
- ◆ Datenrepräsentation (z.B. in Datenbank)

Definitionen: Geoinformation / Geodaten

*Geoinformation =
Information mit Raumbezug
(und Zeitbezug)*

*Geodaten =
Sachdaten + Geometriedaten
(+ Chronometriedaten)*

Information vs. Daten

Daten

- alles, was auf dem Computer gespeichert werden kann
- können (müssen) interpretiert werden

Information

- besitzt Bedeutung für Menschen (“interpretierte Daten”)
- hat Nutzungsaspekt für Menschen

Raumbezug (1)

- ◆ Geodaten / Geoinformationen beziehen sich auf Orte oder Bereiche der Erde
- ◆ Erde als zweidimensionales oder dreidimensionales Gebilde (Erdoberfläche bzw. Erdkörper)
- ◆ Raumbezug durch Koordinaten
- ◆ oder symbolisch durch Namen, Nummern (z.B. Postleitzahlen)

Raumbezug (2)

- ◆ Erde als
 - Kugel
 - Rotationsellipsoid
 - Geoid
- ◆ Darstellbar durch Koordinaten
 - Geographische Koordinaten (Länge / Breite)
 - Ebene Koordinaten (“Gauß-Krüger”)
 - Geozentrische Koordinaten (globales System)

Dimensionalität von Geodaten

- 3D: Erdkörper als dreidimensionales Objekt
- 2D: Erdoberfläche als zweidimensionales Objekt
- 2.5D: Erdoberfläche als zweidim. Objekt + Höhe für jeden Punkt auf Erdoberfläche (Höhe wird zum Sachdatum)

Geodaten

Geodaten = Sachdaten + Geometriedaten
(+ Chronometriedaten)

Math. Funktion: Sachdaten = $f(\text{Ort}, \text{Zeit})$

Beispiel:

Temperatur(Feldberg, 21.Jan.) = - 10 Grad

Tabellenmetapher für Geodaten

Sachdaten

Geometriedaten

Höhe	Landnutzung	Gemeinde	X	Y
400	Wald	Ulm	400	900
390	Grünland	Ulm	420	910
350	Baggersee	Neu-Ulm	450	880
...

Was ist das Kernproblem der Geodatenverarbeitung?

Frage:

“Warum bereiten Geodaten besondere Probleme für die Informationstechnik?”

Viel gehörte Antwort:

“Weil Geodaten mehrdimensional sind.”

Stimmt die Antwort?

Das Kernproblem der Geodatenverarbeitung

*Es gibt unendlich viele Orte
(und Zeitpunkte) !*

Jeder dieser unendlich vielen Orte kann prinzipiell unterschiedliche Merkmale tragen

Lösungen des Kernproblems

- ◆ Vergrößerung des Raumbezugs:
⇒ Rasterdaten
- ◆ Vergrößerung (Klassifizierung) der Sachinformation:
⇒ Vektordaten

Rasterdaten (1)

- ◆ Matrix aus Merkmalswerten
- ◆ Jeder Merkmalswert ist einer Rasterzelle zugeordnet
- Raumbezug wird vergrößert auf die Rastergröße
- + Feine Auflösung der Merkmalswerte möglich

Rasterdaten (2)

0	20	20	40	40	0	0	0
0	20	10	40	40	40	40	0
20	20	10	40	20	10	10	0
20	10	10	20	20	10	10	0
0	10	10	10	10	10	10	0
0	0	0	0	10	10	0	0

Interpretation
z.B.:

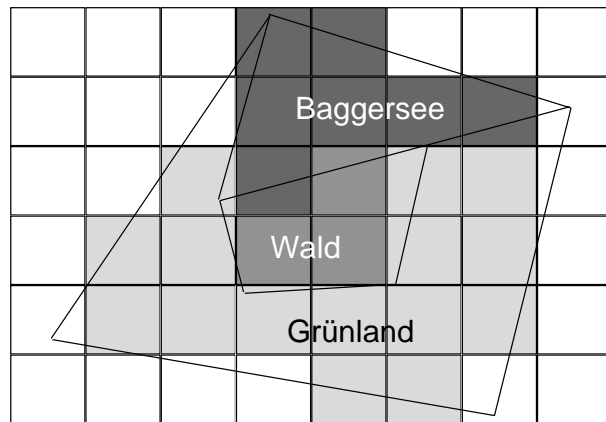
- Bilddaten
- Temperatur
- Höhe
- Lärm
- Verschmutzung
- Landnutzung
- Eigentümer
- ...

Rasterdaten (3)

Verschiedene Arten von Rasterdaten

- ◆ Kontinuierliche Werte (z.B. Höhen, Grauwertbilder)
- ◆ Diskrete Werte (z.B. Klassifikationsergebnisse)
- ◆ Binäre Rasterdaten (nur 0 oder 1 als Werte)

Rasterdaten (4)



Rasterdaten: Anwendung

- ◆ Bilddaten, Sensordaten
 - Satellitenbilddaten, Luftbilder
 - gescannte gedruckte Karten
- ◆ Umwelt, natürliche Phänomene
- ◆ Gut geeignet für kontinuierliche Verläufe von Merkmalen
- ◆ Aber: Nur Merkmale, keine Objekte

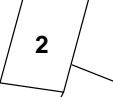
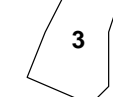
Rasterdaten: Datenstrukturen

- ◆ triviale Matrix-Darstellung
- ◆ Lauflängenkodierung (Run Length Code)
- ◆ Quadtree

Rasterdaten: Datenaustauschformate

- ◆ GIF (CompuServe, Unix, WWW)
- ◆ GRID (Arc/Info)
- ◆ TIFF (MacIntosh)
- ◆ PCX, BMP (PC)
- ◆ JPEG (WWW)
- ◆ Photo CD (Kodak)
- ◆ XWD (X Window System)
- ◆ andere

Geoobjekte

Objekt-ID	Höhe	Landnutzung	Gemeinde	Geometrie (Vektordaten)
1	400	Wald	Ulm	
2	390	Grünland	Ulm	
3	350	Baggersee	Neu-Ulm	

Objektbildung

- ◆ Merkmale werden vergrößert (klassifiziert)
- ◆ Bereiche mit homogenen Merkmalswerten (ggf. nach Klassifizierung)
 - ⇒ Geoobjekte
- ◆ Räumliche Ausdehnung der Bereiche
 - ⇒ Vektordaten
- ◆ Beispiele: Satellitenbildklassifikation, Höhenlinien

Vektordaten

- ◆ definieren Geometrielemente:
 - Punkt (0D), Linien (1D), Regionen (2D), Volumina (3D)
- ◆ werden benötigt zur Darstellung der Geometrie von Geoobjekten
- ◆ sehr genauer Raumbezug möglich
- ◆ "Topologie" kann repräsentiert werden
- ◆ aber: Vergrößerung der Merkmalswerte

Geometrielemente

sind definiert durch

- ◆ Koordinaten (x, y, z)
- ◆ Beziehungen zu anderen Geometrielementen (Topologie):
 - *Linie* enthält *Punkt*
 - *Region* ist begrenzt durch *Linie*
 - usw.

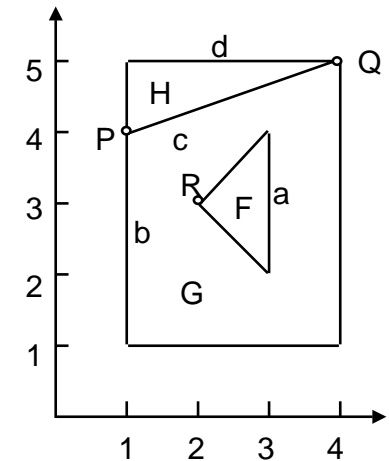
Vektordaten: Datenstrukturen

- ◆ Triviale Darstellung: "Spaghettidaten"
- ◆ Arc/Node-Repräsentation
- ◆ Hierarchisch

Vektordaten: "Spaghettidaten" (1)

- ◆ Triviale Darstellung
- ◆ Geometrieelemente definiert durch Koordinaten(listen)

Punkt P: 1|4
 Linie b: 1|4, 1|1, 4|1, 4|5
 Region F: 2|3, 3|2, 3|4
 Region G: 1|4, 1|1, 4|1, 4|5
 usw.

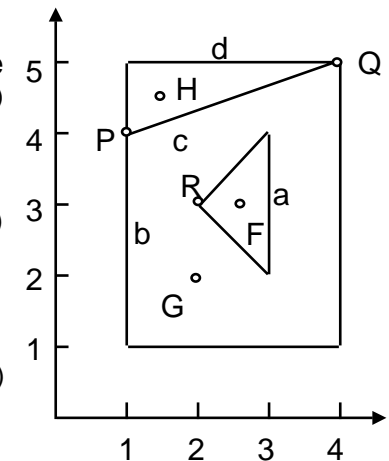
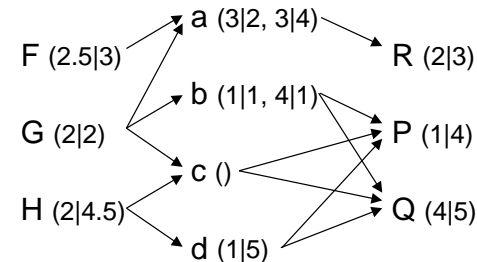


Vektordaten: "Spaghettidaten" (2)

- ◆ Sehr einfache Darstellung: Koordinatenlisten
- ◆ Gut geeignet als Datenaustauschformat
- ◆ Nachteil: Keine topologische Information vorhanden, Topologieaufbau erforderlich
- ◆ Problem: Inselflächen

Vektordaten: Hierarchische Darstellung (1)

Regionen (+ Inpunkte) Linien (+ Knickpunkte) Punkte (= "Knoten")



Vektordaten: Hierarchische Darstellung (2)

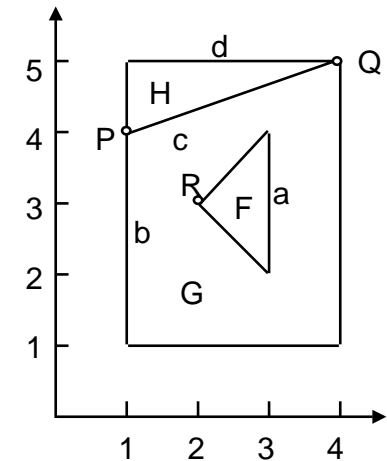
- ◆ Geometrieelemente sind definiert durch Koordinaten und Bestandteile
- ◆ Topologie repräsentiert über Bestandteilhierarchie
- ◆ redundanzfreie Darstellung
- ◆ Beispiel GIS SICAD open

Vektordaten: Arc-Node-Repräsentation (1)

Arc Table

Arc	Left	Right	Start	End
a	F	G	R	R
b	G	-	P	Q
c	G	H	Q	P
d	H	-	Q	P

- ◆ Eine Arc Table definiert die Topologie: Sie enthält Linien (Arcs) sowie anliegende Knotenpunkte (Nodes) und Regionen (Polygons)
- ◆ Weitere Datenstrukturen enthalten die Koordinaten



Vektordaten: Arc-Node-Repräsentation (2)

- ◆ spezialisiert auf "Netztopologien" (d.h. an jedem Ort gibt es maximal ein Geometrieelement)
- ◆ Topologie explizit dargestellt in Tabelle ("Arc Table")
- ◆ redundanzfreie Darstellung
- ◆ Beispiel: GIS Arc/Info

Vektordaten: Anwendung

- ◆ Gut anwendbar für Artefakte (d.h. von Menschen gemachte oder erdachte Objekte):
 - Topographie (z.B. Straßen, Gewässer, Gebäude)
 - Versorgung (z.B. Elektrizitätsleitungen)
 - Grundbesitz (Flurkarten)
- ◆ Schwieriger anwendbar für Umweltobjekte:
 - kontinuierliche Verläufe
 - Objektbildungsregeln oft unklar
 - Geometrien oft unscharf

Geoobjekte

- ◆ gehören einer Objektart an (z.B. Biotop)
- ◆ haben eindeutige ID (Name oder Nummer)
- ◆ besitzen eine Geometrie, dargestellt durch Geometrielement (Region, Linie, Punkt)
- ◆ besitzen Attribute (z.B. Schutzklasse = 2)
- ◆ stehen in Beziehung zu anderen Geoobjekten (z.B. Landkreis = Alb-Donau)

Objektart

- ◆ auch Objektklasse, Objekttyp genannt
- ◆ Objektartenkatalog enthält Beschreibungen aller Objektarten
- ◆ Durch Objektart sind mögliche Attribute und Beziehungen zu anderen Objekten gegeben
- ◆ Objektart legt Typ der Geometrie des Geoobjekts fest: Punkt, Linie, Region...

Geoinformationssysteme (GIS)

- ◆ anderer Name: Raumbezogene Informationssysteme
- ◆ Aufgabe: Erfassung, Verwaltung, Analyse und Präsentation ("EVAP") von raumbezogener Information
- ◆ sind mehr als nur Graphiksysteme
- ◆ können mehr als nur Kartographie

GIS-Historie

- ◆ Graphiksysteme zur Kartenproduktion (ab 70er Jahre)
- ◆ Anfügung von Sachdaten an Graphikelemente
- ◆ Objektmodelle für Geodaten (ab 80er Jahre)
- ◆ Integration mit relationalen Datenbanken (ab 90er Jahre)
- ◆ Objektorientierte Techniken

GIS-Architektur

- ◆ Datenhaltung
- ◆ Verarbeitungsteil
 - Grundsystem
 - Fachschalen
- ◆ Benutzeroberfläche

Geodatenhaltung

- ◆ Proprietäre Datenhaltung
 - Ablage der Geodaten im Dateisystem
 - Spezielle Ablagestrukturen und Zugriffstechniken
- ◆ Marktgängige Datenbanksysteme
 - Problem: Nicht für Geodaten konzipiert
 - Vorteil: GIS-Entwickler brauchen sich nicht um Probleme der Datenhaltung kümmern

Datenbanken

- ◆ persistente Speicherung von Daten (über Programmende hinaus)
- ◆ Abfragesprache für Daten
- ◆ Mehrbenutzerbetrieb ("Transaktionen")
- ◆ Wiederanlauf bei Systemabstürzen

Transaktionen

- ◆ ermöglichen Mehrbenutzerbetrieb
- ◆ sorgen für ununterbrochene Ausführung zusammengehöriger Operationen
- ◆ Beispiel Flugreservierung:
 - “Wenn Sitzplatz vorhanden (1), dann reserviere Sitzplatz (2)”
 - Überprüfung (1) darf nicht von Reservierung (2) getrennt werden

Datenbanksysteme

- ◆ “Auslaufmodelle”:
 - Hierarchische Datenbanken
 - Netzwerkdatenbanken
- ◆ Stand der Technik:
 - **Relationale Datenbanksysteme**
- ◆ Künftig (?):
 - Objektorientierte Datenbanksysteme

Relationale Datenbanksysteme

Flüsse			Arten			Vorkommen	
FlussID	Name	Wasserqualität	ArtID	Name	Schutzstufe	FlussID	ArtID
1	Donau	3	20	Eisvogel	1	2	40
2	Iller	2	30	F-Reiher	2	3	20
3	Blau	2	40	Forelle	2	1	30
						2	30

- ◆ Information wird in Tabellen dargestellt
- ◆ Normalisierung: Separate Tabellen für
 - mehrfach auftretende Information (Beispiel: Tabelle Arten)
 - n:n-Beziehungen (Beispiel: Tabelle Vorkommen)

SQL

SQL (Structured Query Language):
Abfragesprache für relationale
Datenbanksysteme

Wichtigste Elemente:

- ◆ Selektion = Auswahl bestimmter Zeilen
- ◆ Projektion = Auswahl bestimmter Spalten
- ◆ Join = Zusammenfügung von Tabellen

SQL: Selektion, Projektion

Flüsse			Arten			Vorkommen	
FlussID	Name	Wasserqualität	ArtID	Name	Schutzstufe	FlussID	ArtID
1	Donau	3	20	Eisvogel	1	2	40
2	Iller	2	30	F-Reiher	2	3	20
3	Blau	2	40	Forelle	2	1	30
						2	30

Gesucht sind FlussID und Name (*Projektion*)
aller Flüsse mit Wasserqualität < 3:

```
SELECT FlussID, Name
FROM Flüsse
WHERE Wasserqualität < 3
```

Ergebnis

FlussID	Name
2	Iller
3	Blau

SQL: Join

Flüsse			Arten			Vorkommen	
FlussID	Name	Wasser-qualität	ArtID	Name	Schutz-stufe	FlussID	ArtID
1	Donau	3	20	Eisvogel	1	2	40
2	Iller	2	30	F-Reiher	2	3	20
3	Blau	2	40	Forelle	2	1	30
						2	30

Gesucht sind die Flüsse mit den in ihnen vorkommenden Arten mit Schutzstufe =2:

```
SELECT Flüsse.Name, Arten.Name
FROM Flüsse, Arten, Vorkommen
WHERE Flüsse.FlussID = Vorkommen.FlussID
AND Arten.ArtID = Vorkommen.ArtID
AND Schutzstufe = 2
```

Ergebnis

Flüsse. Name	Arten. Name
Donau	F-Reiher
Iller	F-Reiher
Iller	Forelle

Relationale Datenbank als "GIS"

- ◆ Geodaten werden in Tabellen abgelegt
- ◆ Gut geeignet für Sachdatenteil
- ◆ Gut geeignet für Topologie (Arc-Node-Modell)
- ◆ Probleme bei Geodaten: Datentyp "Liste (von Koordinaten)" fehlt.
- ◆ Moderne GIS-Architekturen erweitern relationale Datenbank zur vollständigen GIS-Funktionalität

Verfahren

- ◆ Erfassung
- ◆ Verwaltung
- ◆ Analyse
- ◆ Präsentation / Kartographie

Erfassung von Vektordaten

- ◆ Erfassung vor Ort ("in situ")
 - Feldvermessung
 - Global Positioning System (GPS)
- ◆ Häusliche Digitalisierung aus Landkarten und Luftbildern
 - Digitalisiertisch, -tablett
 - Photogrammetrische Auswertegeräte

Feldvermessung

- ◆ **Winkelmessung**
 - Theodolit
 - Tachymeter (auch für Streckenmessung)
- ◆ **Streckenmessung**
 - mechanisch
 - optisch
 - elektrisch

Global Positioning System

- ◆ 18 Satelliten in 20000 km Höhe
- ◆ Umlaufzeit ca. 12 Std.
- ◆ stets 4 Satelliten über Horizont
- ◆ strahlen Signale aus
- ◆ Empfänger vergleicht Laufzeiten
- ◆ Microcomputer berechnet Position

Digitalisiertisch, Digitalisiertablett

- ◆ **Ausrüstung**
 - Tisch bzw. Tablett
 - Lupe (eine Art Maus mit Fadenkreuz)
 - Ortungseinrichtung für Lupe
 - Serielle Computerschnittstelle
- ◆ Klick auf Lupe überträgt Koordinaten des Fadenkreuzes an Computer
- ◆ Koordinatenumrechnung durch Paßpunktentzerrung

Erfassung von Rasterdaten

- ◆ Scanner (für Landkarten und Luftbilder)
- ◆ Satellitensensoren (Fernerkundung)

Geokodierung (Entzerrung) (1)

- ◆ Transformation der erfaßten Koordinaten auf Koordinatensystem des zugrundeliegenden GIS
- ◆ i.d.R. erforderlich für Rasterdaten und digitalisierte Vektordaten
- ◆ Lösungsansatz durch Abbildungsgleichung
- ◆ Bestimmung der Abbildungsparameter über Paßpunktkoordinaten

Geokodierung (Entzerrung) (2)

Beispiel: Affine Abbildung

$$y_1 = a_{11}x_1 + a_{12}x_2 + a_1$$

$$y_2 = a_{21}x_1 + a_{22}x_2 + a_2$$

Paßpunktpaare ($x_1|x_2, y_1|y_2$)
(mind. 6 erforderlich)

einsetzen in Gleichung und Gleichung auflösen nach $a_{11}, a_{12}, a_{21}, a_{22}, a_1, a_2$

Anschließend Abbildungsgleichung auf erfaßte Daten anwenden.

Erfassung aus Luftbildern

- ◆ Quelle: Schwarz-weiss, Farb- oder Infrarotbilder
- ◆ Orthophoto: Verzerrungseffekte durch unterschiedliche Geländehöhen werden (vom Computer) eliminiert. (Erfordert Kenntnis des Höhenmodells.)
- ◆ Stereophotogrammetrie: Es wird 3D mit Hilfe eines Stereobetrachters gemessen,
- ◆ Zusätzlich Entzerrung mit Paßpunkten.

Einfache Kartierungen (z.B. Biotopkartierung)

- ◆ Orthophoto als Erfassungshilfe (vom Landesvermessungsamt)
- ◆ Begehung des Gebiets
- ◆ Einzeichnen der Objekte auf Orthophoto ("Filzstift")
- ◆ häusliche Digitalisierung, Entzerrung

Erfassung über Scanner

- ◆ Vorteil: kein Digitalisiertablett etc. erforderlich, aber: Rasterdaten
- ◆ Digitalisierung auf dem Bildschirm ("digitizing on screen")
- ◆ oder Mustererkennungstechniken
 - Erkennung von Schriften, Linien, Signaturen auf Landkarten, anschließend manuelle Nachbearbeitung der Geodaten
 - Spektralklassifikation (wie bei Satellitenbildern)

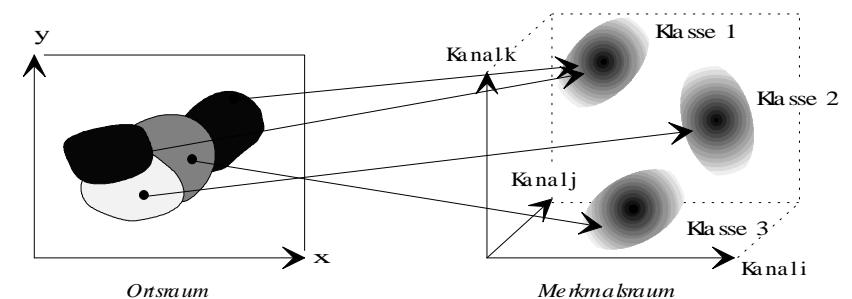
Erfassung aus Satellitenbilddaten

- ◆ Geokodierung ähnlich wie Luftbilder
- ◆ Klassifikation der Bilddaten = Vergrößerungen der Merkmale im Rasterbild
- ◆ Objektbildung (Objekte = zusammenhängende, homogene Bereiche im klassifizierten Bild)
- ◆ Raster/Vektor-Wandlung zur Erzeugung der Vektorgeometrien

Klassifikation von Satellitenbilddaten

- ◆ Unüberwachte Klassifizierung
 - Automatische Bildung von Merkmalsklassen ("Clusteranalyse")
 - Bedeutung der Klassen muß durch Experten analysiert werden
- ◆ Überwachte Klassifizierung
 - Experte gibt Trainingsgebiete an für gewünschte Zielklassen
 - Automatische Klassifikation der Bildpunkte nach diesen Klassen

Klassifikation von Satellitenbilddaten



Analyse von Vektordaten

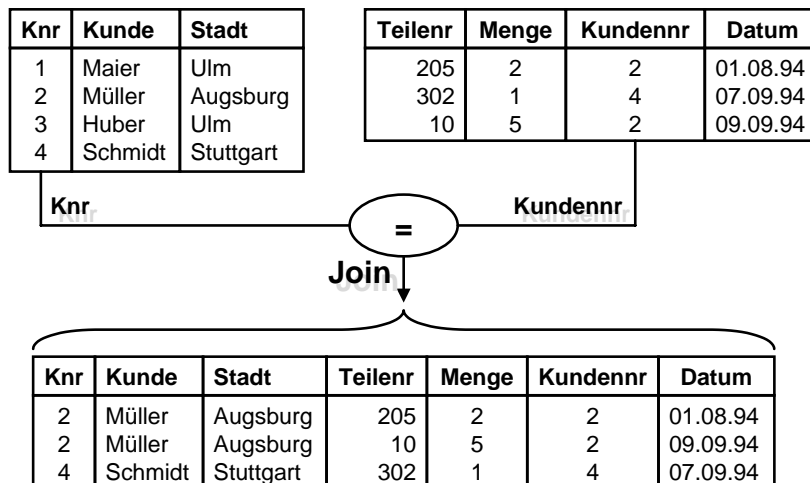
- ◆ Räumliche Abfragen
 - Verwendung räumlicher Prädikate
 - Räumliche SQL-Erweiterungen
- ◆ Generatoren
 - Verschneidung
 - Pufferbildung

Räumliche Abfragen

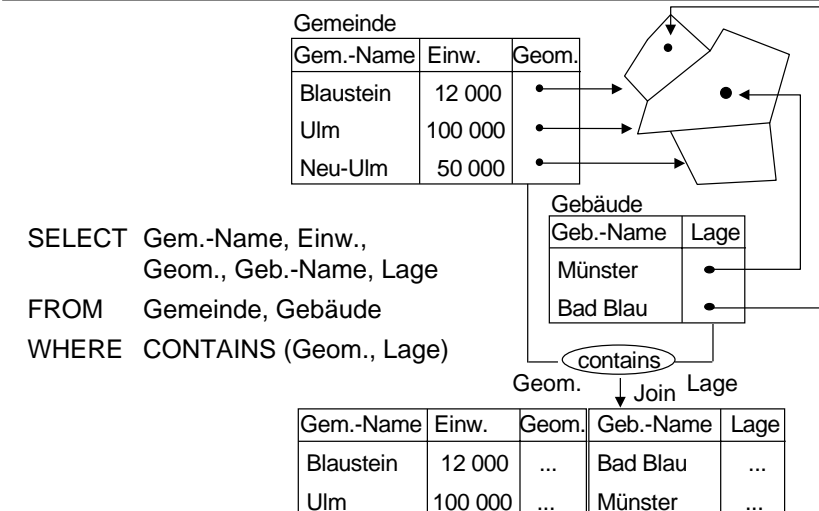
- ◆ räumliche Prädikate zur Selektion
 - INTERSECTS(X,Y)
 - CONTAINS(X,Y)
 - DISTANCE(X,Y) < D
 - ...
- ◆ Erweiterung der SQL-Syntax in WHERE-Klausel, z.B.:

WHERE DISTANCE(Flüsse.Lage, 5000|4000) < 120

Klassischer Equi-Join

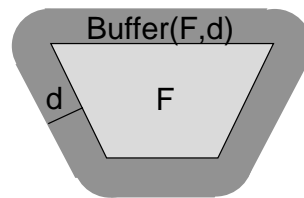


Spatial Join

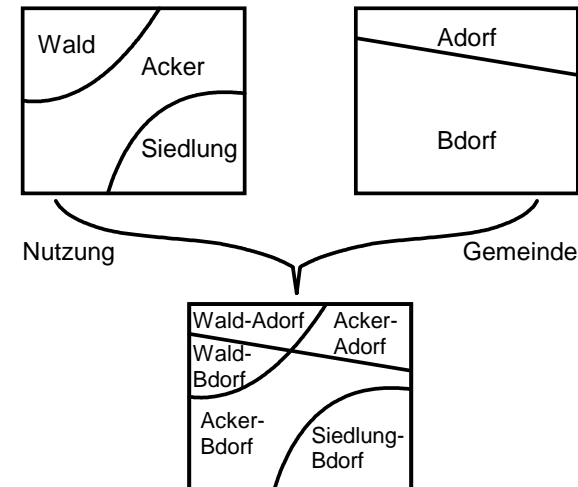


Pufferzonenbildung (Buffering)

- ◆ Pufferzone Buffer(F,d)
- ◆ schließt alle Punkte um das Geometrieelement F im Abstand d mit ein
- ◆ Anschließende Verschneidung liefert Objekt(teil)e im Abstand d
- ◆ Interpretation: Einflussgebiete, z.B. Lärmausbreitung, Verschmutzung



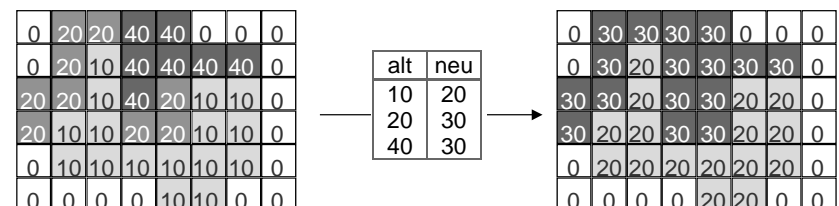
Vektorverschneidung



Analyse von Rasterdaten

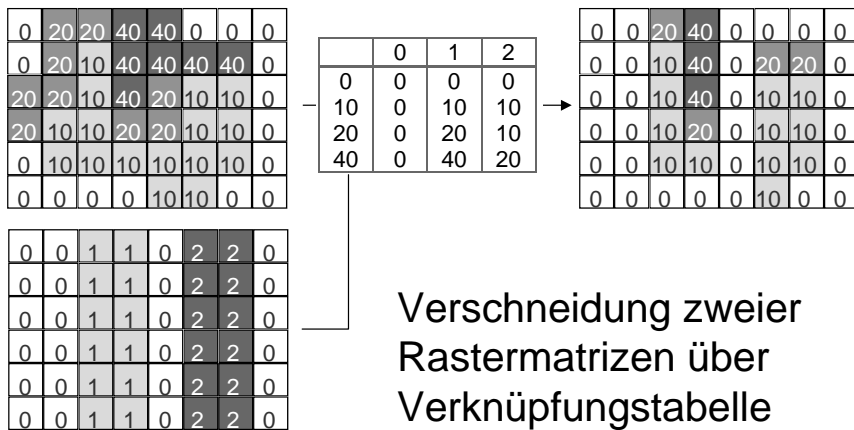
- ◆ Klassifizierung
- ◆ Umklassifizierung
- ◆ Verschneidung
- ◆ Umgebungsoperatoren
- ◆ Resampling (Übergang auf andere Rastergröße)
- ◆ Koordinatentransformation (z.B. Entzerrung)

Umklassifizierung



Die Werte der Rastermatrix werden gemäß Zuordnungstabelle ersetzt

Rasterverschneidung

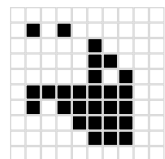


Umgebungsoperatoren

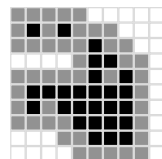
Wert eines Rasterelements bestimmt sich aus dem Wert seiner Nachbarn (z.B. 3 x 3-Umgebung, 5 x 5-Umgebung)

- ◆ Mittelwertbildung zur "Glättung" der Werte
- ◆ Konturbildung durch Differenzenbildung
- ◆ Expansion und Kontraktion
- ◆ ...

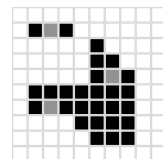
Beispiele für Umgebungsoperatoren Expansion und Kontraktion



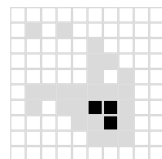
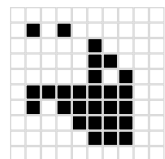
Ausgangsdaten



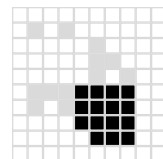
Expansion (Dilatation)



Kontraktion (Erosion)



Kontraktion



Expansion

Präsentation

- ◆ Graphische Darstellung von Geodaten
- ◆ Eng verknüpft mit Kartographie
- ◆ aber graphische Ausgestaltung änderbar
 - Einblenden / Ausblenden von "thematischen Ebenen" ("Folien, Layers")
 - Änderung von Symbolen, Farben, Strichstärken etc.
 - Hervorhebung einer Selektionsmenge (z.B. Ergebnis einer Anfrage)

Karten

Eine Karte besteht aus:

- ◆ Kartenrahmen, Koordinatengitter
- ◆ Legende
- ◆ eigentlicher Karteninhalt

Darstellungselemente

- ◆ Was man auf einer Karte sieht, sind graphische Darstellungen von Geoobjekten (nicht die Geoobjekte selbst)
- ◆ Darstellungselemente werden einem "Signaturkatalog" entnommen
- ◆ Verwendete Signaturen werden in Legende dokumentiert
- ◆ Die Darstellung ist gegenüber der Realität "generalisiert" (d.h. vergrößert)

Signaturen

- ◆ Punktsignaturen (z.B. für Kirche, Schloß)
 - Symbol, Farbe, Größe
- ◆ Liniensignaturen (z.B. Eisenbahn, Straße)
 - Linienart, Farbe, Breite
- ◆ Flächensignaturen (z.B. Wald, Gewässer)
 - Füllung (Farbe, Symbole), Umrandung wie Liniensignatur
- ◆ Textsignaturen (z.B. geogr. Namen)
 - Textfont, Größe, Attribute (z.B. kursiv)

Generalisierung

- ◆ Generalisierung
 - Ausgedehnte Objekte als Punkte oder Linien darstellen
 - Konturen "glätten", bei Strassen etc. Kurven auslassen
 - Verdrängung von Objekten, z.B. Gebäude links der Straße
- ◆ Doppelte Bedeutung des Maßstabs:
 - als Abbildungsmaßstab
 - als Maß für die angewandte Generalisierung

Anwendungen

- ◆ Topographie
- ◆ Geologie
- ◆ Biotopkartierung
- ◆ Versorgung
- ◆