Environmental Data Analysis

GIS

Leibniz Universität Hannover

General

- Lectures: 29.11; 13.12; 10.01; 24.01 (GIS-Lab)
- Practice: 06.12; 20.12; 17.01; 24.01

- One home assignment (60%)
- Final written examination (40%) – open book (17.02.12 ?!)

- sagi.dalyot@ikg.uni-hannover.de (room 603, tel. 2472)
- david.siriba@ikg.uni-hannover.de (room 611, tel. 5255)
- Lijuan.Zhang@ikg.uni-hannover.de (room 616, tel. 19437)

2

General

Introducing the underlying principles and methods involved with working with Geographical Information Systems (GIS), and familiarizing and understanding related contents: advance technologic skill developing for (2D and 3D) analysis, modeling and visualization of geospatial data as a decision-making, planning and environmental tool.

The course concentrates in learning the GIS work-environment. As such, subjects that relate to general knowledge in geodesy and geoinformation will also be given, together with topics related to implementations and technological aspects of GIS.

Various subjects and topics will be implemented in ArcGIS 10 as exercises and assignments, providing with opportunities for development of practical skills in geo-processing data.

An introductory course to *GIS in Hydrology and Water Management*.

3

Textbooks on GIS

- Worboys, M.F. and Duckham, M.: *GIS: A Computing Perspective, Second Edition*, CRC Press, 2004.
- R. Laurini and D. Thompson: Fundamentals of Spatial Information Systems, Academic Press, 1992.
- Longley, Goodchild, Maguire, Rhind: *Geographic Information Systems and Science,Wiley, 2001.*
- Keith C. Clarke. *Getting Started with Geographic Information Systems. 2nd ed.,* Prentice Hall, 1999.
- Michael Zeiler: *Modeling Our World: The ESRI Guide to Geodatabase Design*, ESRI Press, 1999.

ESRI ArcGIS

- Booth Bob, *Getting started with ArcGis, GIS by ESRI, Redlands, 2001.*
- Minami Michael, *Using ArcMap, GIS by ESRI, Redlands, 2000.*
- Shaner Jeff, *Editing in ArcMap, GIS by ESRI, Redlands, 2000.*

Web resource

http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html

4

Topics Outline

- General Introduction – what is GIS? (spatial data, tasks,…)
- Data Modeling
  - Geometric- , thematic- , topologic- modeling
  - Data structures (types) – vector, raster
  - Datums, Projections, coordinate systems, geo-referencing
- Data Analysis (and geo-processing)
  - Thematic- , geometric- analysis
  - Queries
- Cartography
  - Graphical variables
  - Generalization
  - Data Presentation

5

Topics Outline (continued)

- Data Capture (incl. Topography)
  - Data acquisition and representation
  - Data interpolation and analysis (introduction)
  - DEMs (Introduction)
- Visualization (introduction)
  - 2D and 3D
  - ArcScene

6

Environmental Data Analysis - GIS

Introduction

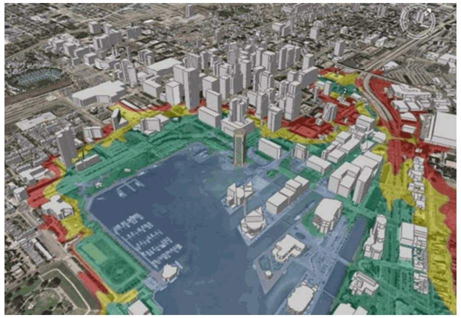
Sagi Dalyot
Liuan Zhang, David Siriba
Institute of Cartography and Geoinformatics
Leibniz Universität Hannover, Germany

How does the “machine” knows where I am?!

State, city, street, street number, distance to, direction to, traffic jams, …

8

How can the “machine” make us more aware?!

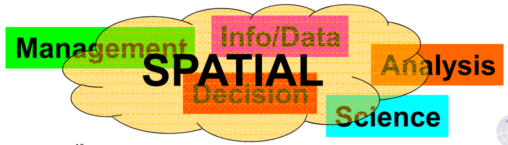


Decision making, different scenarios, 4D, different scales of analysis, model is ‘flexible’…

9

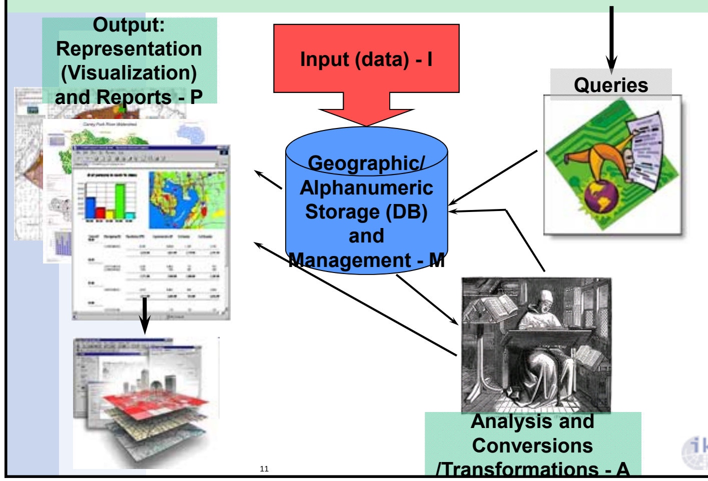
Definition(s) GIS

- “A GIS serves for acquisition, storage, analysis and presentation of all data, which describe a part of the earth’s surface, relating technical and administrative facilities as well as geo-scientific, economical and ecological structural conditions” [N. Bartelme, 1995]
- “A powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world for a particular set of purposes” [P.A. Burrough, 1986]
- “A system for supporting decision finding, which integrates spatial data in a problem solving environment’’ [D.J. Cowen, 1988]



10

Definition/scheme of GIS - IMAP



11

Geographic Information Systems – GIS

- Geographic Information Systems are systems for:
  - **I** – Capture and update (input)
  - **M** – Storage and management (DB)
  - **A** – Analysis and simulation
  - **P** – (Re)Presentation / Visualization

  of Spatial data [Bill and Fritsch, 1991]

- Physical and methodological separation of storage, analysis and visualization of spatial data
- Thus:
  - Flexible analysis
  - Flexible presentation
  - Flexible data combination
  - Integration of new data types (e.g. 3D, time, CAD …)

12

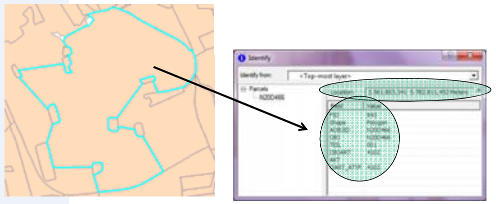
Geographic Information Systems – IMAP

- I: Data capture:
  - Data sources: images, raster maps, geodetic measurements, vector data, digital topographic landscape models, statistics, …
- M: Data management:
  - Data modelling and data structures (databases)
  - Goal: rich data structures that support later analysis processes
- A: Data analysis:
  - Thematic queries (SQL – Structured Query Language for Database Management Systems)
  - Geometric analysis (intersection, buffering, overlaying, …)
- P: Data presentation:
  - 2D-Visualization, maps, diagrams
  - 3D-Visualization
  - 4D-Visualization
  - Simulation and VR

13

Geographic Information Systems

- In general, spatial data (information/data in geographic/geospatial) that systems store/relate to:
  - **Position** - Information regarding certain positions/locations on ''earth'' relating to specific phenomena (features), and their geometry (shape), and
  - **Semantic data** (thematic/attributes) – Information regarding certain aspects of 'that' data: what? when? instances? status?..



14

Geographic Information Systems

- Large volumes of heterogeneous sets of spatial features (points, lines, areas, mixed) can be managed simultaneously
- Query of spatial database according to existence, position and characteristics of a very large number of features (phenomena) - management
- Ability for (human) interaction (queries, reports, …)
- System fits to various requirements of different users -> generic functionality

15

GIS vs. Mapping

- digital cartography / mapping deals with the automation of cartographic representation and production methods
  - primarily visualization and 'measuring' of spatial phenomena
  - visual analysis of data
  - Map as stored data (static)!
- research aspects in digital cartography
  - Interactive maps
  - Adaptive maps
  - Visualization of temporal phenomena
  - Automation in generalization
- Many commonalities between GIS and Cartography!

16

GIS vs. CAD

- CAD (Computer-Aided Design) – primarily interactive geometric modeling of 2D/3D objects (internal)
- very comfortable, high degree of interaction
- provision of plenty of construction aids, like geometric primitives, constraints (e.g. right angle, parallelism, ...)
- powerful visualization components (rendering, 3D visualization)
- main difference to GIS:
  - Limited semantic and thematic data
  - Limited analysis functionality (especially geometric analysis)
- Systems which integrate GIS and CAD, e.g. CAD-Systems with mapping extensions, e.g. MGE (Microstation+Intergraph GIS), Geographics, AutoCad Map

17

Relationship of GIS and other disciplines

Primary roots of GIS (after Jones 1997):

- Computer Science
  - databases
  - computational geometry – algorithms for processing and manipulation of geometric data (triangulation, convex hulls, intersection, etc.)
  - computer vision
  - information processing
  - cognitive science
  - artificial intelligence
- Geography, Geosciences
  - description, modelling, analysis
  - spatial processes

18

Relationship of GIS and other disciplines

- Mathematics
  - statistics
  - operations research
- Surveying/Mapping
  - geodetic data capture
  - photogrammetry and remote sensing
  - cartography
  - visualization

19

GIS Applications

- Land Information System (LIS)
  - real estate map, real estate attributes, real estate processes
  - topographical surveying
  - municipal surveying (city maps)
  - engineering surveying (documentation)
- Utility Information Systems (UIS)
  - automated mapping of facilities (electricity, cabling, sewer pipes, ...)
  - utility management
- Spatial Information Systems (SIS)
  - land use
  - region developing maps
  - land use supervising map
  - building planning

20

GIS Applications

- Environment Information Systems (EIS)
  - environmental goodness-of-fit
  - radioactivity
  - documentation of protected species
  - planning and simulation of agricultural facilities
  - documentation of air, water, soil (contaminated areas, resources, etc.)
- Special Purpose Information Systems (SPIS)
  - traffic navigation (cars, ships, planes)
  - transport management
  - military services
- And more…

21

Physical Components, Platforms

b) platforms

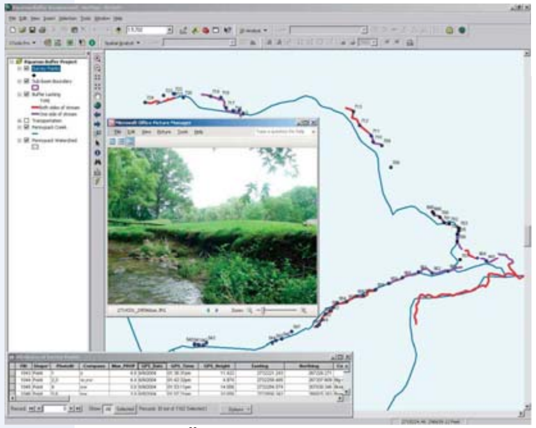
- Hardware (H) ~ 2-3 years
- Software (S) ~ 1 year
- Data (D) ~ 5-100 years
- Applications (A) ~ 10-50 years

b) platforms

- Web-GIS ⇨ passive (… and active – Web2.0)
- Workstation-GIS ⇨ active
- PC-GIS ⇨ passive and active

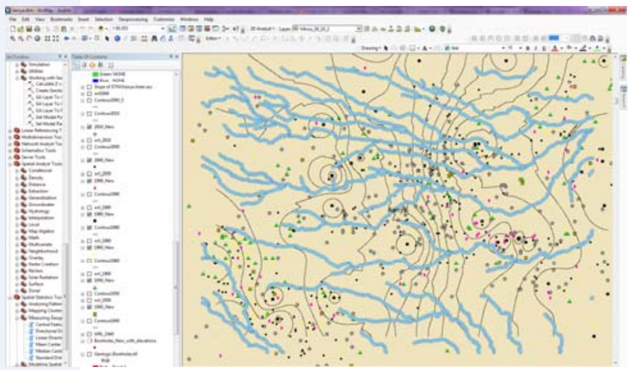
22

GIS – Geometry + Thematic Data



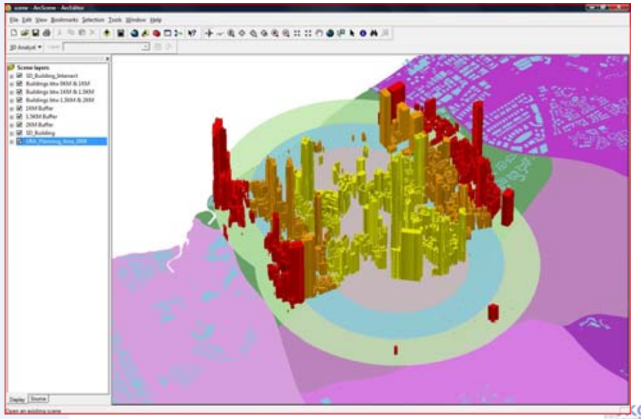
23

Spatial Query: Buffer-Operation



24

Integration of 3D in Standard-GIS



25

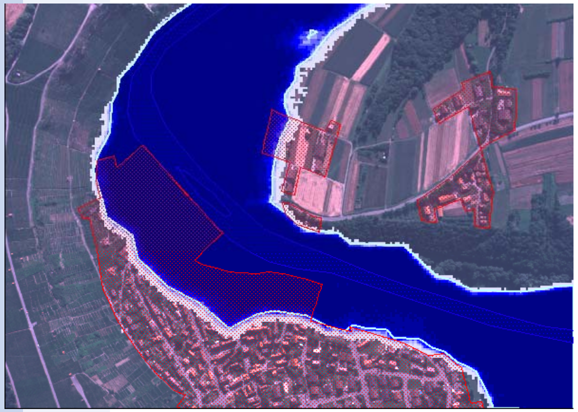
Integration of the 3rd Dimension

- 3D visualization, e.g. with ArcGIS 3D-Analyst/ArcScene
- however: no complex 3D-Analysis capabilities (yet)

26

Example: Flooding Prediction

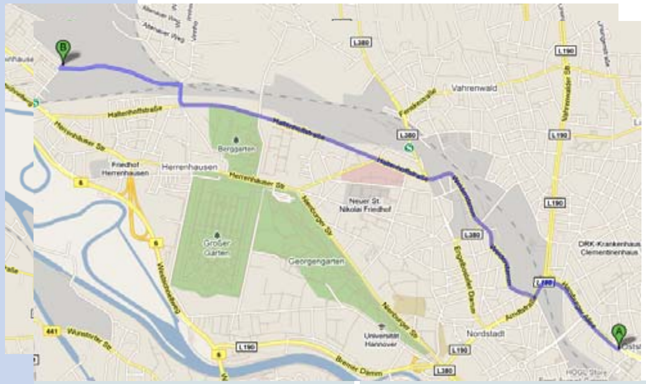
- Multi-temporal - probability of flooding



> 95 %
> 90 %
> 80 %
> 70 %
> 60 %
> 50 %
> 40 %
> 30 %
> 20 %
> 10 %
> 0 %

27

Route planning on the Internet



How do I get from the main station to the University?

By car… 6,7km and 11mins.

Walking… 5,5 km and 68 mins.

28

Location Based Services

- Information is provided depending on the location / position of the user
  - E.g. „where is nearest restaurant?“
- Components
  - Mobile device
    - PDA, Laptop, Mobile/smart phone
  - Positioning accuracy (planar)
    - Mobile phone cell: 200m (city), 3km (rural areas)
    - GPS (5-10m)
    - Manual input (Address, positioning in map - geocoding)
  - Communication: GSM, UMTS, WLAN
  - Server with underlying (geo-)data

29

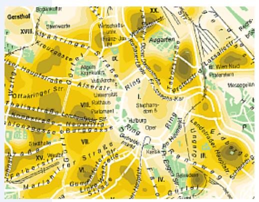
Location Based Services

- Applications:
  - Tourism
  - Navigation
  - Hiking
  - Museum guide
  - Traffic jams….

30

Geo-Marketing

- 80% of data in any organization have spatial connection and are spatial-based, e.g. addresses of customers, location of shops, routes (traveling-agent), ….
- Geo-marketing can analyze and visualize these connections
  - Integration of organization data with statistics about inhabitants, their income level, economy, ….
- Application: location-based planning



31

Web2.0 (future of GIS?!)

- Continuous updating of data
- More data types
- Cloud-data is used for various of applications and analysis
- And… ?!

32

Environmental Data Analysis - GIS

Data Modeling

Sagi Dalyot
Liuan Zhang, David Siriba
Institute of Cartography and Geoinformatics
Leibniz Universität Hannover, Germany

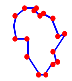
Data Modelling in GIS

Overview:

- conceptual model
  - geometric modelling
  - topologic modelling
  - thematic modelling
- data structures for
  - vector data
  - raster data
  - thematic data

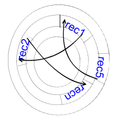
Reality

Data model



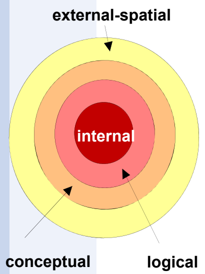
Data structure

File structure



Views (Schemes) in GIS

Differentiation according to four views:



- external-spatial: usage of GIS products, graphical user interface, system setup, “application view”
- conceptual: structure definition for themes and geometry
- logical: mapping of themes and geometry onto DBMS, determine data model
- internal: data access regulation (physical, administration)

35

External View

Reality

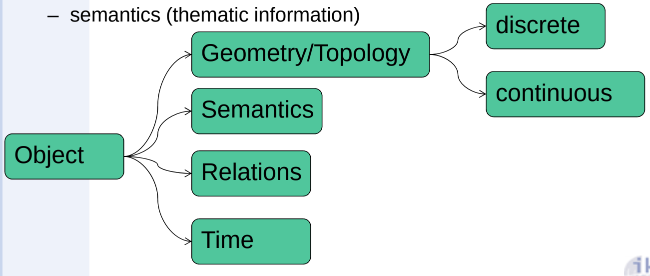
How to model?

GIS (Computer System, Knowledge Based System)

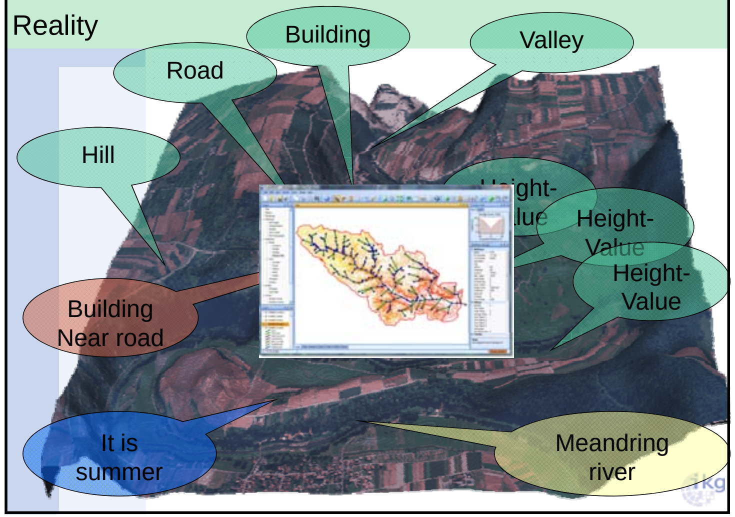
- transition from reality to computer system involves abstraction process; thus the model only reflects certain aspects of reality
- only what passes through "glasses of perception" (what fits to the predefined model) is stored in the computer system
- model has to represent entities (objects, phenomenon) and the relationships among them ‘as true as possible’
- the richness of these models determines the system's range of applications and capabilities
- specification of ‘extract of reality’ to be represented/modeled in the GIS - application dependent

Conceptual Model

- conceptual model describes transformation into given predefined sets of **data structures**
- objects are structured according to:
  - geometry
  - topology
  - semantics (thematic information)



Reality



Aspects of reality

- Objects
  - Phenomena
    - E.g. city, river, building
  - Facts
    - Abstract, e.g. population density, traffic load
  - Events, e.g. growth of city
- Properties of objects
  - Width of a river, unemployment rate in city, ...
- Relations between objects
  - Neighborhood, dependencies, size-relations
- Spatial arrangements
  - Cluster
  - Alignments
- Temporal attributes
  - Emergence, building of road, temporal parking

Semantics

Relations

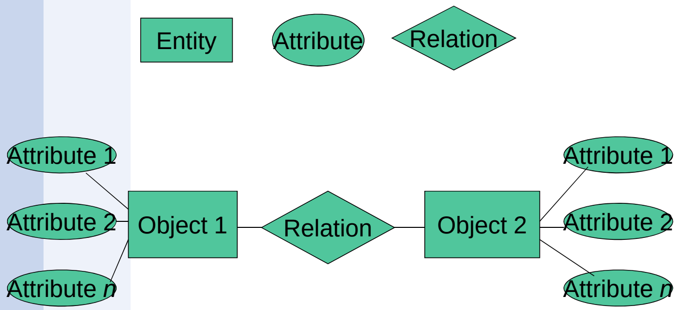
Time

Conceptual Modelling

- Describe <u>aspects</u> of reality that are relevant for a given task or application (sometimes derived by system limitations)
- Several aspects should be a-priori considered:
  - Completeness: modelling should aspire to be complete as possible
  - Accuracy: the quality grade in describing/modeling the object in regard to "it's" reality (in positioning, consistency, etc.)
  - Resolution: the size of the smallest object required for modelling (in raster = pixel size), e.g., desired scaling
- Only what is explicitly modelled can later be used for analysis and visualization processes
- Questions:
  - Which properties do objects have/store/present?
  - Which relations do objects share with other objects?
  - To what extent the modelling is essential?
  - What state/process is needed to be analyzed?

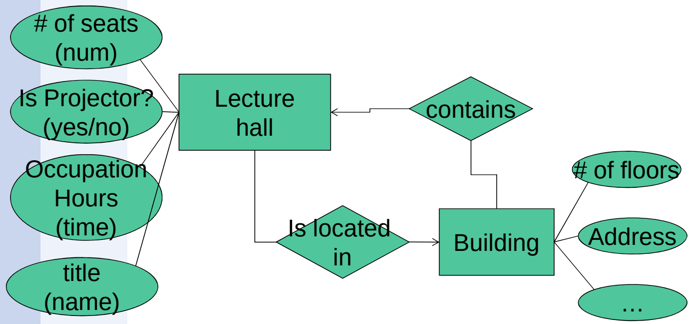
Entity-Relationship-Model (ER-Model)

- Describe (models) objects and their relationships



Entity-Relationship-Model (ER-Model)

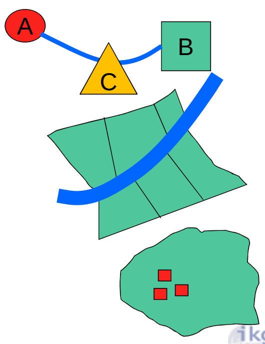
- Example: object modeled is a **lecture hall**



Spatial properties and relations

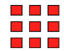
- Length and width (road, river, electricity-line)
- Area (building, city)
- Volume (3D) (building, mountain)
- Perimeter (lake, parcel)
- Form (circular, elongated, …)
- Orientation
- Inclination (3D)

- Connection (A to B via C)
- Neighborhood (parcels and river)
- Containment (settlements in forest)
- Direction (left, behind, on top)

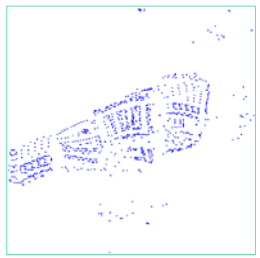


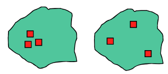
Spatial distributions

- Regular vs. irregular distributions
  - Building in city
  - Lakes, islands



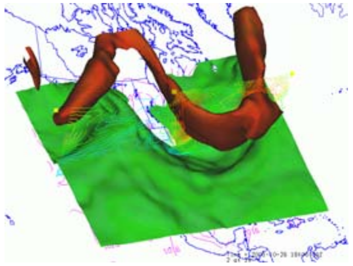
- Clustered vs. arbitrarily distributed





Spatio-temporal patterns

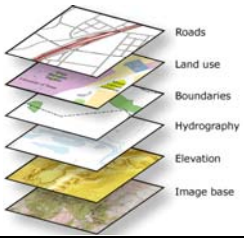
- Temporal changes
  - Change of landuse
  - Movement of tribes
  - Change in area size: e.g. growth of city
  - Movement of commuters
  - Change in water level



Thematic Modelling

Thematic Modelling

- layer principle: theme (phenomenon) = function of position
  - confer overlay of transparencies with different information, linked by common geometry (x,y-coordinates/position)
  - each layer contains different information (semantics, attributes,...)
  - simple model; easy superimposition
  - no hierarchy - all layers are of equal priority
  - Each layer can store a certain/specific geometry

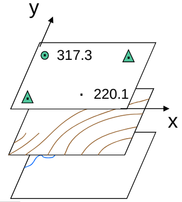


Thematic Modelling

The layer principle

Phenomenon x,y
- theme 1 ➡ layer 1 (geometry and attributes)
- theme 2 ➡ layer 2 (geometry and attributes)
- ⋮
- theme *n* ➡ layer *n* (...)

Example: Hydrology map
- E1 : wells (points)
- E2 : streams (lines)
- E3 : rivers (lines)
- ⋮
- E*n* : lakes (polygons)



All layers have to be defined in the **same coordinate system**

Thematic modelling – layer principle

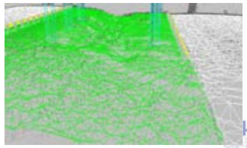
- Advantages:
  - very simple model
  - map is divided into several thematic layers
  - simple overlay by activation of different layers
- Disadvantages:
  - Not object-related
  - No explicit relations between objects, e.g.,
    - Part-of-relation of parcel and the buildings inside
    - Is-a-relation of road and traffic object

Representation of reality as Geo-objects

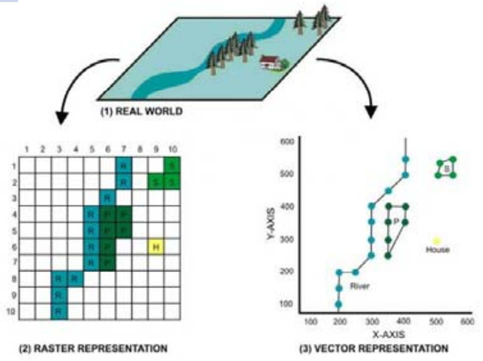
- Modelling with vector data

- Modelling with raster data

- (Modelling with triangulated data (vector-based representation))



Raster and Vector Model



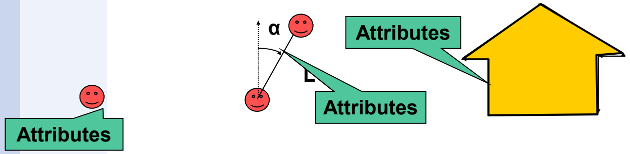
Aronoff (1989)

Raster and Vector Model

Vector Data – presentation via a set of vectors and integrated mathematical functions: start and end point (or start point and length and orientation); accompanied by certain attribute values: color, width, etc.

All vectors are grouped into 3 types of geometries:

- Point (node, vertex, intersection): trees, sewage, light-poles, etc.
- Line/arc (polyline): streets, roads, rivers, etc.
- Polygon (area-based): buildings, cities, lakes, etc.



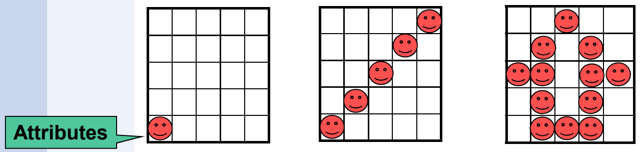
Raster and Vector Model

Raster Data – Bit-Map representation via a set of (equal-spaced) cells, i.e., pixels (picture element), where each has a certain numeric value, i.e., color, saturation.

The quality of the data is derived by the number of bits representing each cell (pixel) and by the raster's resolution – separation capability.

For example:

- 8-bit grayscale = 0-255 (256 values)
- 8-bit RGB = 16.7 million values

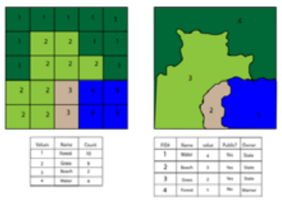


Raster and Vector Model

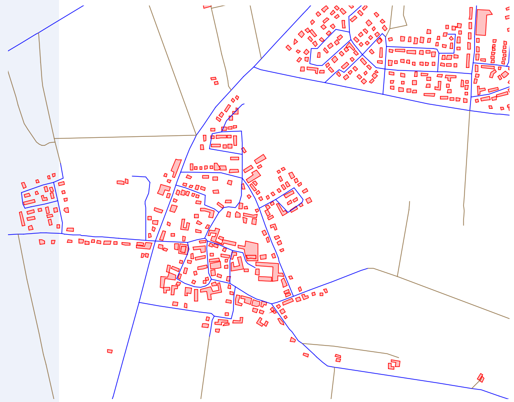
	Raster	Vector
Resolution	Derived by pixel-size	Coordinates
Characterization	Dense and heterogeneous	Sparse and homogenic
Data space	huge	small
Curves	Limited	Efficient and explicit
Zoom	Sensitive and distortions	Insensitive and undistortioned
Representation change	Simplicite mathemtical functions	Designated algorithms
Analysis	Relatively fast, data-overlay	Requires resources

Raster and Vector Model

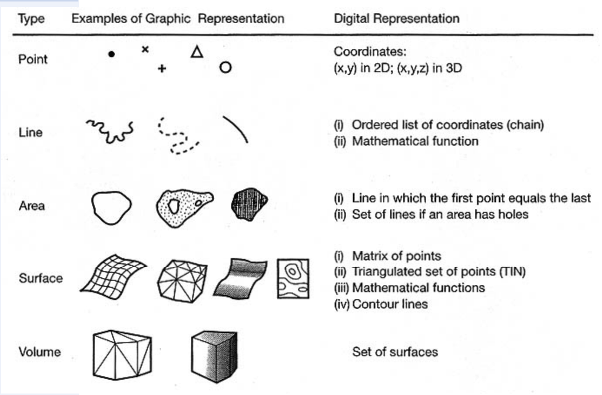
- **Point** – single positioning in space. Dimensionless, and has no size and orientation.
- **Line** – the shortest path between two points. 1D only (length). Arc/Polyline – a sequence of line segments or infinite points (minimum 2 positions).
- **Polygon** – bounded shape/geometry composed of a finite number of line-segments (edges) congregating at the vertices. Polygon is a 2D body and has area (minimum 3/4 points).
- **Raster** – bit-map image that can present any of the above – or their combination.



Modeling with vector data



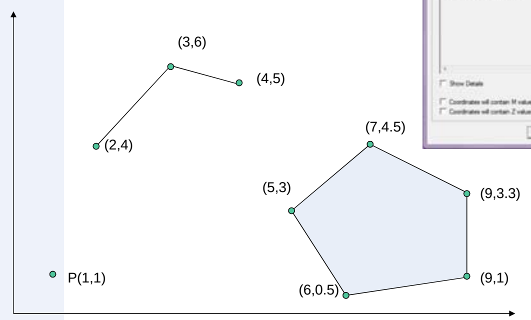
Geometric primitives



Type	Examples of Graphic Representation	Digital Representation
Point		Coordinates: (x,y) in 2D; (x,y,z) in 3D
Line		(i) Ordered list of coordinates (chain) (ii) Mathematical function
Area		(i) Line in which the first point equals the last (ii) Set of lines if an area has holes
Surface		(i) Matrix of points (ii) Triangulated set of points (TIN) (iii) Mathematical functions (iv) Contour lines
Volume		Set of surfaces

Jones (1997)

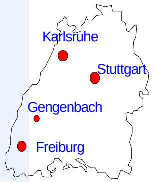
Points, lines, polygons



- Data capture: position / centre line / digitization of boundary

Geometric Modelling

point: elementary object; carries geometric information



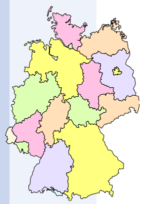
- identification (name / number)
- coordinates: x/y/(z) in given reference system
- type (e.g. trigonometric point, city center, ...)
- quality measures: point error, reliability, date of capture, ...

line: string of points (>= 2 points)

- identification (name / number)
- line points
- type (e.g. street, river, breakline, network, ..)
- type of connection between points: straight line, circular line, higher order polynomial, ...
- direction

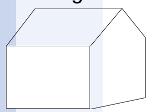
Geometric Modelling

polygon: closed string of points



- identification
- type (e.g. house, parcel of land, states of F.R. of Germany ...)
- sequence of points (or lines)
- size
- ...

volumes: modelling strategies from computer graphics and CAD. Modelling of 3D-bodies is of increasing importance in GIS.



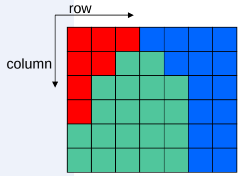
Modeling with raster data

Pixel (cell) – value only with no explicit attributes

Raster data

- Gray/RGB values represent
  - Light intensity
  - Derived values, e.g. land-use classification
  - Distribution of noise along a road
  - Terrain height
  - ....

Requires image analysis and computer vision techniques and algorithms (classification)



Value	Land use
1	Water
2	Settlement
3	Forest