Chapter II
Geographic Information Systems: Data Modeling

80%
“80 % of all data throughout the world have some geographic background”

Application Domains
- Urban planning
- Land use planning
- Rural and forestry planning
- Sea and Marine Applications
- Transports
- Natural resources (mines) planning and management
- Earth sciences
- Archaeology
- Big real estate planning and management
- etc

GIS: Data Modeling
- 2.1 – Geographic Data Modeling
- 2.2 – Data Acquisition
- 2.3 – Output Devices
- 2.4 – Metadata
- 2.5 – Spatial Data Consistency
- 2.6 – Extensions of XML
- 2.7 – Conclusions
2.1 – Geographic Data Modeling

- Discrete Objects
  - Generally modeled by their boundaries
  - What models to use for points, lines, areas and volumes?
- Attribute modeling
- Continuous phenomena (ex. Temperature)
  - Modeled as continuous fields

Earth Positioning

- Geodesy
- Coordinates
- Projections of the sphere

Geodesy

- The Earth is not exactly a sphere
- ellipsoid
- geoid
- altitude

Coordinates
Chapter II: GIS: Data Modeling

Coordinate center

Projections of a sphere

Globe Projections

Different projections
Deformations according to projections

Minimum path

Global curvilinear distance between two points

For instance in France (Lambert System)

\[ d(P_1, P_2) = R \times \arccos(\sin(LA_1) \times \sin(LA_2) \times \cos(LO_2 - LO_1)) \]

\[ R = \text{Mean Earth Radius} \]

\[ R = 6,378.135 \text{ Km (Equatorial Radius)} \]

\[ R = 6,356.766 \text{ Km (Polar Radius)} \]
Definitions of altitudes

Consequences

Fragment of the Dutch topographic map showing the border of Belgium and the Netherlands. The Mean Sea Level of Belgium differs -2.34m from the MSL of The Netherlands. As a result, contour lines are abruptly ending at the border.

Geographic data layers

Chapter II: GIS: Data Modeling
Real world models

Geographic Formats

Vector Format

Raster Format

Formalism with spatial pictograms and icons

- Small drawing (graphic symbol) representing some geometric type

- Spatial Pictograms

<table>
<thead>
<tr>
<th>Point</th>
<th>Line</th>
<th>Area</th>
<th>Volume</th>
</tr>
</thead>
</table>

- Usage

Other examples

Parking

Simple Pictogram

Town

Alternative Pictogram
Pictogram and icons

Conceptual model with pictograms and icons

OpenGIS Model

- Consortium of companies, research centers and administrations
- Interoperability of geographic applications
- Propositions of standards

http://www.opengis.org
Continuous Phenomena

- Theory of continuous fields
  - Scalar fields
  - Vector fields
- Applications
  - Meteorology
  - Sea studies
  - Terrains, soils
  - Etc.

Examples

Continuous Field Modeling

- Impossible to know the function everywhere
- Necessity of sampling points
- Necessity of interpolating functions
- Modeling (two levels)
  - Field as object (ex Temperature in a region)
  - Field as abstract data type (ex value of temperature in some point)

2.2 – Data Acquisition Modes

- Surveys
- Digitizing
- Aerial photos
- Satellite images
- Laser
- GPS
- Sensors
Theodolite

Aerial photos

Example of framework for aerial photo
Aerial photos
Characteristics

- Altitude: from 500 to 3,000 meters
- Format: 23 cm × 23 cm
- Scale from 1:3,000 to 1:25,000
- Photos pair → relief
- Parallaxes → determination of altitudes
- Photo-interpretation
- Orthophotos (mosaicking)

Distortions

- Camera
- Photo plan
- Irregular soil

Roof and soil coordinates

- Roof coordinates
- Soil coordinates
Realization of orthophotos

- Overlap: 60% longitudinal
- 25% lateral
- Selecting control points
- Elastic transformation (rubber sheeting)
- Corrections of distortions
- Cutting along roads or rivers

Orthophoto principles

Orthophoto (result)

Color balancing

Before | After
Image deformation

Before

After

Orthophoto mosaicking

Covering a DTM

Principle of Laser Scanning
Laser range scanning

Satellite Images

Brazil: Satellite image organization
Chapter II: GIS: Data Modeling

Visual Information Systems

Pr. Robert Laurini

Spectroscopic scale:

- Wavelengths:
  - 30 cm, 1 GHz
  - 3 cm, 10 GHz
  - 0.3 cm, 100 GHz
  - 300 µm, 10^12 Hz
  - 30 µm, 10^14 Hz
  - 3 µm, 10^16 Hz
  - 0.3 µm, 10^18 Hz
  - 300 Å, 10^18 Hz
  - 30 Å, 10^18 Hz
  - 3 Å, 10^18 Hz
  - 0.3 Å, 10^18 Hz
  - 0.003 Å, 10^18 Hz

- Hyperfrequencies:

Visible range:
- 0.7 µm
- 0.4 µm

Spectral regions:
- Infrared
- Ultraviolet
- Gamma-Rays

Satellite applications:
- METEOSAT
- NOAA
- LANDSAT
- SPOT

Remote sensing principles:
- Filtering
- Proper emissions
- Reflected solar radiations

Satellites and usage:
- Geo-stationary orbit
- Polar orbit
- Phases
- Tracking

Temporal scale:
- Minutes
- Hours
- Days
- Months
- Years
- Centuries

Spatial Amplitude:
- 1000 km
- 100 km
- 10 km
- 1 km
- 1 dam
- 1 m

Geographical applications:
- Microclimate
- Biotopes
- Mesoclimates
- Soil
- Volcanism
- Pedogenesis
- Pedogenesis
- Glaciations
- Agriculture
- Urban planning
- Environment
- Military Satellites
- Aerial photos
- Satellite and usage
- Agriculture
- Humidity

Remote sensing principles:
- Solar radiations
- Reflected solar radiations

Additional resources:

Subsequent passages:

Spectral signature

http://www.rsacl.co.uk/remote_sensing/main.htm

Global Positioning System

Principles of GPS

Ikonos

Visual Information Systems

Chapter II: GIS: Data Modeling
Differential GPS

Measuring with GPS

Voice Technology

- Provided by Datria / Stantec
- GPS-positioned messages are stored into computers
- Interesting for example to describe certain situations

Measures with fixed sensors
Chapter II: GIS: Data Modeling

Acquisition with vehicle

DIGITIZING

Common errors in digitizing

Rubber-sheeting

Necessity of sharing geometry and topology

Necessity of snaping nodes on arcs

Undershoot

Overshoot

Initial map

New map

Control points

To move

Fixed points
2.3 – Output devices

- Various devices
- Interactivity level
- Level of Graphic Semiology

Map scanning

Original segments

After scanning

After skeletonizing

Choosing scales

(a) Range of acquisition scales and precision

(b) Range of utilization scales and precision

Flatbed plotter
2.4 – Metadata

• Data about data
  – Origin
  – Quality
  – Consistency
  – Completeness
  – Updating

• Standards
  – CEN
  – FGDC

FGDC Metadata

Norm ISO19115
2.5 – Consistency of Spatial Data

- Spatial integrity constraints
- Semantics of spatial integrity constraints
- Towards Quality control
- Conclusion

Quality Control

- New databases
  - At creation, powerful controls
  - Then, checking after each updating, insertion and deletion
- Old databases
  - Powerful quality control procedures
  - Correction of erroneous objects
  - Then, checking after each updating, insertion and deletion

Spatial integrity constraints

- Elements about integrity constraints in databases
- Consistency and precision
- Semantics and data structures
- Definition of spatial integrity constraints (SIC)
- Examples for terrains
- Constraints and derived spatial data
Elements about integrity constraints in databases

- Likelyhood control of value
- Existential integrity
- Referential integrity
- User-defined constraints

Consistency and precision

- Quality control
- Precision and accuracy

Common errors

Semantics and data structure

- A structure claiming "I am a square", is it really a square?
- Necessity of controls
- In some case, add complimentary data

- Example
- NOQUAD (#quad, #pt1, #pt2, #pt3, #pt4)
- PUNTO(#pt, x, y)
Standards

- ISO 19113
  - Geographic Information – Quality Principles
- ISO 19114
  - Geographic Information – Quality Evaluation Procedures
- ISO 19115
  - Geographic Information – Metadata

Dimensions of quality

- Geometric
  - Primitive → Objects → Class of objects → Set of data
- Semantic
  - Value → Domain → Attribute → Class of objects → Semantics
- Temporal
  - Time primitive → Complete history → All temporal aspects → Set of temporal data

Example of consequences of error

4-sided polygon and its area

There is no error

The road passes at the third floor of a building!!!
Representation of semantics of different spatial objects

<table>
<thead>
<tr>
<th>Set of points</th>
<th>Region: (#point)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No rule No SIC.</td>
</tr>
</tbody>
</table>

Closed polyline

<table>
<thead>
<tr>
<th>Region: (#point)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule: point in a polyline</td>
</tr>
<tr>
<td>SIC: 2 neighbouring points must differ</td>
</tr>
</tbody>
</table>

Polygon

<table>
<thead>
<tr>
<th>Region: (#point)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule: point in a polyline</td>
</tr>
<tr>
<td>SIC: closure and no degenerescence</td>
</tr>
</tbody>
</table>

Definition of spatial integrity constraints

- IC = Predicate against a database
- SIC = Predicate with geometric and/or topological conditions

Examples of a terrain model

- **R0 (#terrain, #triangle)**

- **R1 (#triangle, #segment1, #segment2, #segment3)**

- **R2 (#segment, #point1, #point2, #triangle1, #triangle2)**

- **R3 (#point, x, y, z)**

Example with tables
Examples of inconsistencies

R1 #triangle #segment1 #segment2 #segment3
123
ae
g
null
c
h
d
f
h
unknown segment2 modified segment3 triangle 4 missing

R2 #segment #point1 #point2 #triang1 #triang2
a
b
c
d
e
f
ij
AA
C
B
C
EB
DG
BC
F
CE
F
EB
H
1
null
2
null
43
null
null
44
null
7
double
null
point1 changed point2 changed same triangles same extrémities error in triangle2 segment en trop segment manquant

R3 #point x y z
A
B
CD
E
F
I
41
null
2
0
5
3
3
4
1
2
3
vertex within triangle 2 unknown coordinates coordinate outside isolated point

Triangles in the database
Segments in the database
Vertices in the database

Derived data structures

Example:

Model 1
R1 (#triangle, (#vertex)^2 )
R2 (#vertex, x, y ).

Model 2
R1bis (#triangle, (#angle)^3 )
R2bis (#angle, value-in-degrees).

Semantics of SIC

• Geometry and topology
• Example of chaining SIC
• Constraints and representation

Geometry and topology

• Usage of topology
• Usage of trigonometry
• Usage of some theorems
Examples of valid and degenerated polygons

Valid and non-valid tessellations

Example of chaining the checking of constraints

Let consider the following database:

\[ R1 \equiv \text{(#point, x, y)} \]
\[ R2 \equiv \text{(#segment, (#point)^2)} \]
\[ R3 \equiv \text{(#polygon, (#segment)^*)} \]
\[ R4 \equiv \text{(#tessellation, (#polygon)^*)} \]
Constraints and multiplicity of representations

Geometry and topology

- Validity of polygonal tessellations
- Validity of networks

Example of cadaster

Validity of polygonal tessellations

- 1 – validity of points
- 2 – validity of segments
- 3 – validity of polygons
- 4 – validity of tessellation
- 5 – formula of Euler-Poincaré: \( P + V = S + 1 \)
Example of network with errors

Validity of networks

- 1 – validity of nodes
- 2 – validity of edges/arcs
- 3 – network connexity
- 4 – orientation of the network (iff orientated graph)

Example for contours

Parcels and buildings

Parcels and water network
Introduction to Quality control

- New databases
  - at the DB inception
  - checking after each update/delete/insert
- Old databases
  - powerful control procedures and correction of erroneous objects
  - checking after each update/delete/insert

Evolution of quality

Data quality level

Time

Top quality level

Strong
Maintenance
Actions

Components of quality

- Lineage
- Accuracy
- Resolution
- Feature completeness
- Timeliness
- Consistency
- Quality of metadata

Cost Balance

Induced cost when no maintenance

Annual cost of quality maintenance

Evolution of data quality when no maintenance actions are made

Evolution of data quality when regular maintenance actions are made
Conclusion about consistency

- Importance of quality control
- Cost balance
- *What is the cost of an error??*
- Strong quality control at the creation
- Quality control and lifecycle

Quality Visualization

- Various modes
- Example in photogrammetry
- Metadata
- Application

Various modes

- With messages
  - Quality message
  - For the clicked objects, the quality indicator is:
  - -5 for extremities
  - -2 in the central part

- With colors
  - Increasing intensity

Other modes

- With noise
  - Nice music
  - Bad Noise

- With animation
  - Oscillation cycle in the uncertainty swath
  - Step 1, Step 2, Step 3, Step 4, Step 5
Quality indicator
(Traffic light metaphor)

- Confidence
- Data Quality
- Risk seriousness
- Descriptive
- Spatial
- Temporal
- Reliability

Geometric Precision with colors

Definition of a buffer for road extraction

- The buffer depends on:
  - positional accuracy
  - quality of path axes
  - and attribute quality

Road extraction

Extracted lines

Extracted roads
Results in a rural zone

Results:
83 % accepted
13 % non accepted
4 % indecision

Intervisibility with errors

Optimal path in a terrain taking quality into account

Metadata
Consequences of an error

At motorway construction, this building was not stored into the DB.

Conclusion about quality control

- Importance of quality control
- Cost of checking
- Cost of correction
- Cost in case of errors ???

2.6 – Extensions of XML

- Objective: processing geographic vector data on Internet
- Interest:
  - To reduce server loads
  - To alleviate interchanges between clients and servers
  - To allow queries at client side
  - To install local processing at client level

Extensions

- SVG (Scalable Vector Graphic)
- GML (Geography Markup Language)
- LandXML
SVG

• To increase graphic capabilities of XML
• Not originally planned for cartography
• The maps is seen as a drawing
• Possibility of interactivity
• Possibility to modify some drawing attributes

GML

• Really encoding geographic vector data
• Targeted applications: mapping and spatial analysis
• Creation of a small GIS at client level
• Capacity of using spatial and non-spatial attributes
• Opening towards interoperability
LandXML

- Format specification in civil engineering, surveying and architecture
- Transferring data between actors
- 2D and 3D

Comparison - usage

<table>
<thead>
<tr>
<th>Domains</th>
<th>GML</th>
<th>SVG</th>
<th>LandXML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban planning</td>
<td>X</td>
<td>X</td>
<td>XX</td>
</tr>
<tr>
<td>Environmental planning</td>
<td>X</td>
<td>X</td>
<td>XX</td>
</tr>
<tr>
<td>Cadaster</td>
<td>XX</td>
<td>X</td>
<td>XX</td>
</tr>
<tr>
<td>Statistical mapping</td>
<td>XX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

2.7 – Conclusions

- 80% of information throughout the world have some spatial component
- Geographic Databases among the biggest in the world
- Usage for other domains
  - Geomarketing
  - Real estate management
  - Location-Based Services
  - Pervasive Information Systems