Maps as Numbers: Data Models

The Task

- An accurate, registered, digital map that can be queried and analyzed...

Translate:

- Real World Locations, Paper Map ➔ Computer Files
- Spatial Data Models, Topology
- Entity Info. ➔ Queriable Database Files
- Relational or Object-Oriented Databases
- Relate Spatial Coordinates to Entity Info.

“Spatial DBMS” software = GIS software!

Data Models

- How is reality abstracted and codified?

<table>
<thead>
<tr>
<th>Reality</th>
<th>Conceptual Models</th>
<th>Logical Models</th>
<th>Physical Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wells</td>
<td>Produce from rocks that contain oil and gas.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wells are points, rock units are polygons (both are objects).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well A penetrates Fm. 1; produces oil. Well B penetrates Fm. 3; produces gas. Fm 3 overlies Fm. 1.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Store well locations with a particular file structure, relationships in a dBase table. Associate table with location.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Conceptual Models

Characterized all features or phenomenon as:

- **Discrete objects**: e.g. wells, roads, rock bodies, etc.
- **Object-based models**
- **Continuous phenomena**: e.g. gravity, magnetic intensity, topography, temperature, snowfall, soil pH, etc. = “fields” of values
- **Field-based models**
- **Organize objects and fields by a common theme**: e.g. geology, hydrography, transportation
- **Thematic layers**
Logical Models

- **VECTOR MODEL**
  - Discrete objects are represented by points and vectors, continuous fields by irregular tessellation of triangles (A Triangulated Irregular Network: “TIN”)

- **RASTER MODEL**
  - Discrete objects and continuous fields are represented by an array of square cells (pixels)

Logical Models

- How should discrete objects be coded?

   - **Raster Model**
   - **Vector Model**

Logical Models

Continuous Phenomena As Surfaces

- **Raster Topography**
  - Regular tessellations, e.g. DEM, DTM

- **Vector Topography**
  - Irregular tessellations, e.g. T.I.N.
Simple Vector Data Structure

Vector Line

<table>
<thead>
<tr>
<th>ID</th>
<th>X</th>
<th>Y</th>
<th>Node</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>503200</td>
<td>3200522</td>
<td>From</td>
</tr>
<tr>
<td>P2</td>
<td>503250</td>
<td>3200522</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>503300</td>
<td>3200460</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>503245</td>
<td>3200410</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>503350</td>
<td>3200410</td>
<td>To</td>
</tr>
</tbody>
</table>

(in UTM coordinates)

Simple Raster Data Structure:

Equivalent Binary “Flat File”
(Plus “Header” with Raster dimension, resolution and location)

```
1 1 0 0 0 0
0 0 1 0 0 0
0 0 0 1 0 0
0 0 1 0 0 0
0 1 1 1 1 1
```

Physical Models

Vector Models (Raster Next Time...)

- “Graphical”
- Topologic/georelational
- T.I.N.
- Network

“Graphical” Vector Model

- Lines have arbitrary beginning and end, like spaghetti on a plate
- Common lines between adjacent polygons duplicated
- Can lead to “slivers” of unassigned area = “sliver polygons”
“Graphical” Vector Model

- Shortcomings for maps:
  - No real world coordinates required
  - No identification of individual objects; no way to attach attributes
  - Details of relationships among objects (e.g. what’s adjacent) not stored, but needed for spatial analysis

Graphical Vector Structure

- Contains no explicit information about adjacency, containment or contiguity i.e.
  - Which polygons are adjacent?
  - Which polygons are contained within other polygons?
  - Which lines are connected? Where are they connected? Where do lines begin and end?

= “Spaghetti Data Model”

Topological Vector Model

- Store pts. as x,y geographic coordinates
- Store lines as arcs of connected pts.
- Store polygons as closed paths
  
  Also explicitly store ....

  - Where lines start and end (connectivity)
  - Which polygons are to the right and left side of a common line (adjacency)

Topology

- The geometric relationship(s) between entities (e.g. points, lines, areas); where is one thing with respect to another?
Topological Properties

- Spatial characteristics that are unchanged by transformations like scaling, rotation and translation are topologic.
- Non-topological: x, y coordinates, area, distance, orientation
- Topological:
  - Contiguity – what’s adjacent
  - Connectivity – what’s connected
  - Containment – what’s inside or outside of a region

Maintaining Topology: Planar Enforcement

- One and only one feature at every x, y location
- Lines cross at nodes; polygons space-filling, exhaustive, mutually exclusive (no overlaps or gaps)
- Sum of the area of all individual polygons equals the area of extent of all polygons
- Common boundaries stored only once
- A PLANAR GRAPH meets these conditions
- Allows spatial queries for adjacency, containment and rapid what-is-where
- (All raster data is of this sort)
Lines: Graphic vs. Topologic

- Graphic (Spaghetti)
  - vertices
  - Overshoot ("dangle")
  - Table of (x,y) coordinates

- Topologic (with meatballs)
  - vertices
  - arcs
  - Table of (x,y) coordinates
  - Table of arcs with IDs, starting and ending nodes

Polygons: Polygon-Arc Topology

<table>
<thead>
<tr>
<th>Arc Table</th>
<th>Arc</th>
<th>ID</th>
<th>L.Poly</th>
<th>R.Poly</th>
<th>F.Node</th>
<th>T.Node</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>World</td>
<td>100</td>
<td>P1</td>
<td>N1</td>
<td>N2</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>P1</td>
<td>102</td>
<td>P2</td>
<td>N2</td>
<td>N1</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>P2</td>
<td>103</td>
<td>World</td>
<td>N2</td>
<td>N1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Polygon Table</th>
<th>Poly ID</th>
<th>FID</th>
<th>Arcc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>100</td>
<td>A1, A2</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>102</td>
<td>A2, A3</td>
<td></td>
</tr>
</tbody>
</table>

Arc Coordinates Table

<table>
<thead>
<tr>
<th>Arc</th>
<th>Start</th>
<th>Vertices</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>N1</td>
<td>v7, v1, v11, ...</td>
<td>N2</td>
</tr>
<tr>
<td>A2</td>
<td>N2</td>
<td>v8</td>
<td>N1</td>
</tr>
<tr>
<td>A3</td>
<td>N2</td>
<td>v1, v2, v6</td>
<td>N1</td>
</tr>
</tbody>
</table>

Lines: Arc-Node Topology

- Vertex Table (V)
  - Example values: ID, x, y

- Node Table (n)
  - Example values: ID, x, y

- Arc Table (A)
  - Table showing arcs with IDs, F, and T Node

- Poly ID | FID | Arcs |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>100</td>
<td>A1, A2</td>
</tr>
<tr>
<td>P2</td>
<td>102</td>
<td>A2, A3</td>
</tr>
</tbody>
</table>

Why Bother With Topology?

- Provides a way of error trapping and geometry validation after data entry
- All lines must meet at nodes, all polygons must close, polygons can't overlap, all lines in a network must join
- *Permits spatial queries, precise measurements*
What Kind of Queries Does Topology Permit?

- Connectivity
  - What is shortest path between features or locations? (networks, flow)
  - Find all fault trace intersections

- Contiguity
  - What’s adjacent: e.g. Show all granite/limestone contacts
  - Combine all contiguous units with a specific attribute (e.g. lithology) into a single unit

- Containment (= “Area Definition”)
  - What proportion of an area is underlain by a specific rock type?
  - What is spatial density of specific feature(s)?

Vector Models

- Graphical
- Topologic/"georelational"
- T.I.N.
- Network

Triangulated Irregular Network - TIN

- Topological 3-D model for representing continuous surfaces using a tessellation of triangles

Triangular Irregular Network

- Network ("tessellation") of edge-sharing triangles made from irregularly spaced points with x, y and z values
- Density of triangles varies with density of data points (e.g. spacing of contours) - c.f. raster with uniform data density - advantages for file size
- Triangle sides are constructed by connecting adjacent points so that the minimum angle of each triangle is maximized (see “Delaunay Triangulation” for details); i.e. a “fat” triangle, not a “sliver” triangle.
- Can render faces, calculate slope, aspect, surface shade, hidden-line removal, etc.
- Practical limit for computation on desktop is ~ 10-15 million nodes
How Are Triangle Created?
Find the Delaunay Triangulation

- Find the set of circumcircles such that no point lies within a circumcircle
- Circumcircle is the circle that passes through all 3 corners of a triangle
- For 4 or more points on the same circumcircle (e.g., a rectangle) the D. Triangulation is not unique
- For a set of points on a line, the D. Triangulation is degenerate (no triangle)
- D. Triangulation avoids sliver triangles — better represents average slopes and aspects

TIN Topology

<table>
<thead>
<tr>
<th>Node</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>3</td>
<td>44</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>7</td>
<td>50</td>
</tr>
</tbody>
</table>

Triangle | Node List | Neighbors
---|-----------|-----------
A | 1, 2, 6 | C, E
B | 2, 3, 4 | C
C | 2, 4, 6 | B, D, A
D | 4, 5, 6 | E, C
E | 5, 1, 6 | A, C, D

TIN for Seiad Valley, CA

- Triangle edges symbolized
- Faces symbolized for elevation & aspect

3-D TIN Scenes of Seiad Valley fault

After Zeiler, Modeling our World, p. 165
3-D TINS, Grand Canyon

- Bright Angel Trail
- Grand Canyon at Bright Angel Creek

Vector Models

- Graphical
- Topologic/"georelational"
- T.I.N.
- Network - not discussed, see Help files

Logical Models:

- Graphical
- Topologic/"georelational"
- T.I.N.
- Network - not discussed, see Help files