ESRI Object Models and Data Capture

Conceptual Models

Characterized all features or phenomena as:

- Discrete objects; e.g. wells, roads, rock bodies, etc.
  - **Object-based models**
- Continuous phenomena; e.g. gravity, topography, temperature, snowfall, soil pH, etc.
  - **Field-based models**

Outline

- ESRI Software Family
- ESRI Object Data Models
  - History
  - Data Organization – Physical Models
    - Coverage
    - Shapefile
    - Geodatabase
- Data Capture
  - Digitizing
    - "Heads Down"
  - "Heads Up"
  - Building Topology

Some ESRI History...

<table>
<thead>
<tr>
<th>ESRI</th>
<th>Arc/Info</th>
<th>ArcView</th>
<th>ArcGIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Versions</td>
<td>1-7</td>
<td>1–3.2</td>
<td>8.0 – 10.2</td>
</tr>
<tr>
<td>Data Model</td>
<td>Coverage</td>
<td>Shapefile</td>
<td>Geodatabase</td>
</tr>
<tr>
<td>O.S.</td>
<td>Unix, PC DOS</td>
<td>Windows</td>
<td>Windows</td>
</tr>
<tr>
<td>Scripting Language</td>
<td>Arc Macro Language (AML)</td>
<td>Avenue Scripting</td>
<td>Vis. Basic for Appl. (VBA); Python</td>
</tr>
<tr>
<td>Database Software</td>
<td>Proprietary; Arc Tables</td>
<td>OBase</td>
<td>M.S. Access; ArcSDE for Oracle, etc.</td>
</tr>
</tbody>
</table>
ArcGIS Licensing Levels

- **ArcView** – Make maps, do queries, some spatial analysis, some editing (shapefiles, personal geodatabases) – included with GTK ArcGIS Desktop
- **ArcEditor** – plus edit multi-user geodatabases; more tools in toolbox
- **ArcInfo** – full functionality; comes with ArcInfo Workstation (i.e. “legacy” ArcInfo v. 7). *UT D.G.S. licenses*

ArcGIS Extensions

- **ArcView, ArcEditor, and ArcInfo**
  - Advanced raster modeling
  - Arc GRID calculator with ARC GRID algorithms
  - VBA for raster analysis

- **ArcGIS Spatial Analyst**
  - ArcScene™ real-time interactive three-dimensional scenes
  - Scenes views in ArcCatalog
  - Three-dimensional modeling tools
  - ARC TIN tools

- **ArcGIS 3D Analyst**
  - Advanced kriging and surface modeling
  - Exploratory spatial data analysis tools
  - Probability, threshold, and error mapping

- **ArcInfo only**
  - Arc GRID program in ArcInfo Workstation
  - Arc GRID commands in Arc program
  - ARC TIN™ commands in Arc program
  - Surfacescene command

ESRI Data Models

- **Topologic:**
  - ArcInfo - Coverage
  - ArcInfo "E00" – export format for coverage
  - ArcGIS - Geodatabase

- **Non-Topologic:**
  - ArcView (legacy) - Shapefile
Early ESRI Data Models

- **Coverages**
  - Developed for workstation Arc/Info ~ 1980
  - Complex structure, proprietary format
  - Attributes in Info tables

- **Shapefiles**
  - Developed for ArcView ~ 1993
  - Simpler structure in public domain
  - Attributes in dBase (.dbf) tables

Geographic coordinates and attributes are stored in separate but linked files.

Data Organization

- **Coverage**
  - Data split between coverage and INFO folders
  - Common boundaries between polygons stored once
  - Topology explicitly stored
  - Planar graph maintained
  - As in previous lecture

- **Shapefile**
  - Data divided among three or more files (.shp, .shx, .dbf, .sbx, .sbn, et al.)
  - Common boundaries between polygons stored twice
  - Topology created on-the-fly
  - Planar graph not required

Folder/File Organization

- **Coverage**
  - One feature shape (as points OR lines OR polygons) per file = "SHAPEFILE"
  - Many related features (as points AND lines AND polygons) per file = "COVERAGE"

- **Shapefile**
  - Geology.shp
  - Geology.shx
  - Geology.dbf
  - Geology.pj

ArcCatalog: Workspace>Coverage>Feature Class

ArcCatalog

Windows Explorer
Feature Class

- A collection of geographic objects with the same geometry (point, line, polygon) that share the same attributes.
- A shapefile contains one feature class
- A coverage can contain many feature classes

ArcInfo Coverage

- An integrated, homogeneous set of feature classes (pts., lines, polygons) stored together
- Feature classes unified by a theme, e.g. hydro
  - Spatial (coordinate) data stored in binary files;
  - Attributes and topologic data stored in INFO tables
  - Stored within a “Workspace”

ArcInfo Coverages can contain:

- Primary feature classes:
  - Points, with attributes in PAT (point attribute table)
  - Nodes, with attributes in NAT
  - Arcs, with attributes in AAT
  - Polygons, with interior label points and attributes in PAT

Coverages feature classes can contain:

- Secondary features:
  - Tics – registration points for digitized data
  - Annotations – text for map
  - Links – vectors used for adjusting local area to known locations
    (spatial adjustment)
Coverages can also contain:

- Composite features:
  - Routes – collections of Arcs with measurement system
  - Regions – collections of polygons; adjacent, noncontiguous or overlapping

Shapefile format

- Simpler than coverage; doesn’t store topology
- Feature classes stored independently i.e. points, lines and polys. stored in physically separated files (e.g. no shared INFO table)
- For each type, spatial data stored in a .shp file, attribute data in a .dbf table.
- "Null" or "No Data" numerical values not supported in attribute tables

Shapefiles in ArcCatalog/Explorer

- Folder / Shapefile
- Three or more files per feature class

Shapefile feature class types:

- Point, Multipoint
- Polyline (line with several paths)
- Polygon

- Ring – closed, nonintersecting path – simple poly.
- Disjointed Rings – multiple polygons define feature
- Nested Rings – "Island" or "Atoll" polygons
Shapefile Topology

- Shapefiles don’t store information about adjacency
- Topology is generated on the fly — vertices stored in systematic fashion to deal with containment and adjacency
- **Planar enforcement can be broken by editing — not required in structure of shapefile**
- But...tools available to maintain planar enforcement when digitizing in heads-up mode

Geodatabase model

- Stores geographic coordinates as one of many attribute in a relational database table; no separation between aspatial and spatial data, as in earlier models
- Uses MS Access for “Personal Geodatabase” (single user)
- Uses Oracle, DB2 or other commercial relational databases for “Enterprise GIS” (many simultaneous users).

Geodatabase Model

- Data structure capable of storing objects with behaviors and relationships, not merely graphical shapes with topology and attributes
- All spatial and attribute data for a feature are stored in a row of a single table
- A Geodatabase is a top-level container for feature classes, coverages, shapefiles, rasters, et al. (more later) — ALL DATA CAN BE IN ONE CONTAINER AND ARE THUS PORTABLE
Feature classes in Geodatabase include:

- Points, Multipoints (groups of points)
- Lines
- Polygons

Plus ....
- Network Junctions (special Nodes)
- Network Edges

For geometric networks
- Plus other classes
- Relationship classes
- Object Classes – tabular data without geography

Geodatabase Feature Datasets

- Set of Feature Classes, some with topologies, that share the same spatial reference
- All feature classes with topologies must be stored within a Feature Dataset
- Analogous to coverage

Object Class

- A collection of nonspatial objects that share the same attributes and are stored in a table (i.e. a simple table)

<table>
<thead>
<tr>
<th>Number</th>
<th>Age_Ma</th>
<th>1_sigma</th>
<th>Rx_Type</th>
<th>Size_kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>142</td>
<td>1.5</td>
<td>B_schist</td>
<td>3.4</td>
</tr>
<tr>
<td>124</td>
<td>136</td>
<td>2.0</td>
<td>G_schist</td>
<td>1.3</td>
</tr>
<tr>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Object Class (nonspatial table)
Relationship

- A relationship is an association or link between two objects in a database.
- A relationship can exist between spatial objects (features in feature classes), non-spatial objects (objects in object classes), or between spatial and non-spatial objects.

Relationship class

E.g. relationship between spatial and non-spatial objects

<table>
<thead>
<tr>
<th>Number</th>
<th>Age_Ma</th>
<th>1_sigma</th>
<th>Rx_Type</th>
<th>Size_kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>142</td>
<td>1.5</td>
<td>B_schist</td>
<td>1.4</td>
</tr>
<tr>
<td>124</td>
<td>136</td>
<td>2.0</td>
<td>G_schist</td>
<td>1.3</td>
</tr>
<tr>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Paper Map ➔ Files Of Coordinates

- How are they organized?
  - Data Models, Topology
- How are they stored?
  - Data Organization
- How are coordinates captured?
  - Data Entry, Encoding

Digitizing is:

- Conversion of spatial data to digital form
  - Lines, points or polygons are traced to record coordinates of their locations
- Term conventionally used to denote the process of creating VECTOR data
  - Scanning produces raster data (“bit maps”)
  - But software exists to convert raster to vector so can digitize (“vectorize”) scanned images
Digitizing is accomplished via:

- Digitizing table or tablet
- "heads-down" digitizing
- Large digitizing table
- A mouse, on screen
- "heads-up" digitizing
- Aerial photos, other raster or vector sources as base to digitize from
- Software that converts raster to vector
  - Vectorization – batch or interactive modes, e.g. ArcScan extension

Digitizing strategies governed by:

- Will data be used for queries and analysis or just visual display?
  - i.e. Topology important or not?
  - "True" G.I.S. functionality or not?
- What are accuracy requirements and how much generalization is permitted?

Digitizing with a tablet involves:

- Digitize 3 reference points – define position of map w.r.t. digitizing table
- Establishing 4 or more control points - distinctive features at known locations that can be used to register the map to ground coordinates (e.g. UTM, lat./lon.) = “georeferencing”
- Separating features as point, line or polygon and tracing them to separate files (themes)
- (Heads-up digitizing starts with georeferencing)
Spaghetti vs. Topologic models

- Spaghetti: Points, lines, polygons and their attributes stored in tables
- Topological:
  - Same, but with corresponding tables of information about what’s adjacent or what’s within what

“Building Topology”

- Clean: Edit to ensure planar enforcement
  - Remove sliver polygons & gaps between polygons
  - Correct overshoots, undershoots, leaky polygons
- Build: Add topological attributes to spaghetti
  - Manual
  - Automatic
- Digitizing with topology performed in ArcInfo or with tools in ArcToolbox, ArcMap and ArcCatalog
- Changes to polygons or lines affect topological attributes – Strict rules for editing coverages in ArcMap (topology tools available)

Heads-up digitizing

- Decide whether new file will have planar enforcement
- Create new point, line or polygon feature class(es) in ArcCatalog
- Edit feature class(es) to add features and attributes
- Stop editing
- Save edits as part of new feature class