ESRI* Object Models; Data Capture

* Environmental Systems Research Institute

Conceptual Models

Characterized all features or phenomena as:

- **Object-based models**
- **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></th>
<th>Arc/Info</th>
<th>ArcView</th>
<th>ArcGIS Desktop</th>
<th>ArcGIS Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Versions</td>
<td>1-7</td>
<td>1–3.2</td>
<td>8.0 – 10.8</td>
<td>1.0-2.8</td>
</tr>
<tr>
<td>Data Model</td>
<td>Coverage</td>
<td>Shapefile</td>
<td>Geodatabase</td>
<td>Cloud, Geodatabase</td>
</tr>
<tr>
<td>O.S.</td>
<td>Unix, PC DOS</td>
<td>Windows</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>
<td>Python</td>
</tr>
<tr>
<td>Database Software</td>
<td>Proprietary; Arc Tables</td>
<td>DBase</td>
<td>M.S. Access, ArcSDE for Oracle, etc.</td>
<td>M.S. SQL, Oracle, PostgreSQL, etc.</td>
</tr>
</tbody>
</table>
 ESRI Object Models and Data Capture

Geo327G/386G: GIS & GPS Applications in Earth Sciences
Jackson School of Geosciences, University of Texas at Austin

ArcGIS Desktop Levels
(Licensing Levels)

- Basic – Entry level; make maps, do queries, some spatial analysis, some editing (shapefiles, personal geodatabases) – included with GTK ArcGIS Desktop
- Standard – midlevel; advanced editing, multi-user geodatabases; more tools in toolbox
- Advanced – full functionality; control of “all aspects of data building, modeling, analysis and map display

UT D.G.S. licenses

ArcGIS Extensions

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

ArcGIS Spatial Analyst
- ArcScene™ real-time interactive three-dimensional scenes
- Scenes 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

Geostatistical Analyst
- Advanced raster modeling
- Arc GRID calculator with Arc GRID algebra
- VBA for raster analysis

ArcInfo only
- ArcGIS GRID program in ArcInfo broker
- ArcGIS commands in Arc program

ArcGIS Licensing Levels

- Basic – Entry level; make maps, do queries, some spatial analysis, some editing (shapefiles, personal geodatabases) – included with GTK ArcGIS Desktop
- Standard – midlevel; advanced editing, multi-user geodatabases; more tools in toolbox
- Advanced – full functionality; control of “all aspects of data building, modeling, analysis and map display

UT D.G.S. licenses

ArcGIS Extensions

Others available: Network, Tracking, Survey, Maplex, (ERDAS Image Analyst)

Licensing and “Floating Seats”

This Lab (40 floating seats)

- Limit use to number of floating licenses
- License Manager
- License file with keycodes
- Arc Editor stand-alone licenses (field tablet computers)
- Floating seat
- Floating seat
- ArcView single-use seat
- Floating seat

(a COLA Server)
ESRI Data Models

- **Topologic:**
  - ArcInfo - Coverage
  - ArcInfo “.EOO” – export format for coverage
  - ArcGIS - Geodatabase

- **Non-Topologic:**
  - ArcView (legacy) - Shapefile

Early ESRI Data Models

- **Spatial Data**
  - Geographic coordinates and attributes are stored in separate but linked files

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

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

Data Organization

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

- **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”

- **Shapefile**
  - Many related features (as points AND lines AND polygons) per file = “COVERAGE”
Data Organization: Coverage in Windows Explorer and ArcCatalog

- ArcCatalog: Workspace>Coverage> Feature Class

- Feature Classes

Windows Explorer

Feature Class

- A collection of geographic objects with the same geometry (i.e. 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
Geodatabases in ArcCatalog/Windows Explorer

- Geodatabase/Feature Dataset/Feature Class

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
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>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>

Feature Class (spatial 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 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)
Digitizing strategies governed by:

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

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