Geodatabase Essentials – Part One
An Introduction to the Geodatabase

Jonathan Murphy
Colin Zwicker
Session Path

• The Geodatabase
  – What is it?
  – Why use it?
  – What types are there?

• Inside the Geodatabase

• Advanced Behavior

• Geodatabase Potpourri
What is the Geodatabase?

• Core ArcGIS data model
  – A comprehensive model for representing and managing GIS data

• A physical store of geographic data
  – Scalable storage model supported on different platforms

• A transactional model for managing GIS workflows

• Set of COM components for accessing data
Geodatabase Data Management Approach

• The geodatabase is built on an extended relational database.
  – Base relational model
  – Relational integrity
  – Base short transaction model
  – Reliability, Flexibility, Scalability
  – Supports continuous, large datasets

• Built on the simple feature model
  – Open access (OGC, C, COM, SQL)
Geodatabase Data Management Approach ...

- **Simple features + logic**
  - All geographic data stored as tables in a DBMS
  - Extend functionality and data integrity
  - Functionality is consistent across DBMS’

- **Application logic (software)**
  - Works on standard DBMS tables
  - Implements GIS integrity and behavior
  - Business rules, topology, networks
Geodatabase Data Management Approach ...

- Editing and data compilation
  - Rich set of editing tools
  - Maintain spatial and attribute integrity
  - Undo and redo edits
  - Multiple users editing the same data

- Versioning work flows
  - Multiple users editing over long periods of time
  - Archiving
  - Distributed data management

- Robust, customizable framework
  - Build and manage your own specific GIS solution
3 Types of Geodatabases

• **Personal Geodatabase**
  - Single user editing
  - Stored in MS Access
  - Size limit of 2 GB

• **File Geodatabase**
  - 1 TB per table
  - Reduced storage requirements

• **ArcSDE Geodatabase**
  - Stored in an enterprise DBMS
  - Supports multiuser editing via versioning
  - Requires ArcEditor or ArcInfo to edit
### 3 Types of Geodatabases...

<table>
<thead>
<tr>
<th></th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage format</strong></td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td><strong>Storage capacity</strong></td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Supported O/S platform</strong></td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Number of users</strong></td>
<td>Single editor</td>
<td>Single editor</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td></td>
<td>Multiple readers</td>
<td>Multiple readers</td>
<td></td>
</tr>
<tr>
<td><strong>Distributed GDB functionality</strong></td>
<td>Check out/check in and One-way replication</td>
<td>Check out/check in and One-way replication</td>
<td>Replication (all types) &amp; versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
### 3 Types of Geodatabases...

<table>
<thead>
<tr>
<th></th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage format</strong></td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td><strong>Storage capacity</strong></td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Supported O/S platform</strong></td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Number of users</strong></td>
<td>Single editor</td>
<td>Single editor</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td></td>
<td>Multiple readers</td>
<td>Multiple readers</td>
<td></td>
</tr>
<tr>
<td><strong>Distributed GDB</strong></td>
<td>Check out/check in</td>
<td>Check out/check in</td>
<td>Replication (all types) &amp;</td>
</tr>
<tr>
<td></td>
<td>and One-way replication</td>
<td>and One-way replication</td>
<td>versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
### 3 Types of Geodatabases…

<table>
<thead>
<tr>
<th></th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage format</strong></td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td><strong>Storage capacity</strong></td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Supported O/S platform</strong></td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Number of users</strong></td>
<td>Single editor</td>
<td>Single editor</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td></td>
<td>Multiple readers</td>
<td>Multiple readers</td>
<td></td>
</tr>
<tr>
<td><strong>Distributed GDB functionality</strong></td>
<td>Check out/check in and One-way replication</td>
<td>Check out/check in and One-way replication</td>
<td>Replication (all types) &amp; versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
Geodatabase Data Management

- Schema is defined in ArcCatalog
  - Define feature classes, datasets, relationships, etc

- Import and convert data from other formats
  - Shapefile
  - Coverage
  - CAD
  - Raster

- Copy and Paste

- Geodatabase XML Export / Import
  - For transferring Schema or Features and Schema

- Use an ESRI Data Model
  - Industry specific data models available
  - Copy geodatabase template
Editing Geodatabases

• ArcGIS datasets stored in the geodatabase are editable
  – Modify building footprints in parcel management
  – Add water mains to a water network
  – Update land owners information stored in a table
  – Etc…

• Transaction model for editing in ArcGIS
  – Edits are performed in an edit session
    • Open session – edit – save edits / don’t save edits
  – A series of edit operations constitutes a transaction
    • The transaction is either committed or rolled back
Editing Geodatabases...

• **Personal Geodatabases**
  – Mainly single user editing on small datasets
  – Multiple readers
  – Editing locks at geodatabase level
    • Two editors cannot edit within the same geodatabase at same time

• **File Geodatabase**
  – Mainly single user editing small to very large datasets
  – Multiple readers
  – Editing locks at the dataset level
    • Multiple editors cannot edit the same table or stand-alone feature class at the same time
    • Multiple editors cannot edit feature classes in the same feature dataset at the same time
Editing Geodatabases...

• ArcSDE Geodatabases
  – Extend the transaction model with Versions
  – Multiuser editing without locking
    • Unique isolated view of the geodatabase

• Benefits of versioned editing
  – Multiple editors, editing over long periods of time
  – Undo / Redo
  – Archiving
  – Replication
Session Path

• The Geodatabase

• Inside the Geodatabase
  – Object class, Feature class, Raster dataset
  – Feature datasets
  – Validation rules
  – Domains, Subtypes, Relationship classes
  – Annotation, Dimensions
  – Exploring a Geodatabase DEMO

• Advanced Behavior

• Geodatabase Potpourri
Inside the Geodatabase

- A geodatabase contains datasets
- Datasets represent collections of information with a real-world interpretation
- Types of geographic datasets:
  - Tables
  - Object classes, feature classes, relationship classes
  - Feature datasets
  - Networks, Topologies, Raster and cadastral datasets
- Datasets have associated information to help manage integrity, behavior, and interpretation
  - Domains, Relational integrity, Topology, Metadata
Inside the Geodatabase…

Feature dataset
- Feature class
  - Polygon
  - Line
  - Point
  - Annotation
  - Dimension
  - Route

Relationship class
- Topology
- Geometric network
- Network dataset
- Terrain

Raster dataset
- Raster catalog

Survey dataset
- Project folder
- Project

Schematic dataset
- Toolbox
  - Tool
  - Model
  - Script

Behavior
- Attribute domains
- Attribute defaults
- Split/merge policy
- Relationship rules
- Connectivity rules
- Topology rules
Objects and Object Classes

- Objects are entities with properties and behavior
- An object is an instance of an object class
- All objects in an object class have the same properties and behavior
- An object can be related to other objects via relationships

A row stores an Object

A table stores an ObjectClass
Features and Feature Classes

- Builds on the Relational Model
- A feature is a spatial object
- A feature is an instance of a feature class
- Extended the relational model with
  - Geometry attribute types

A feature class is a table of rows, where each row has a geographic column
Geodatabase Supports Advanced Geometry

- Points, lines, polygons
  - Single and multipart features
- Text and surfaces
- Flexible coordinates
  - XY, Z, M

Feature with many parts

One record in feature class table
Geodatabase Raster Data

- **Support for many formats**
  - tiff, bmp, GRID, among others...

- **Attribute field in a table**

- **Raster dataset**
  - Separate rasters
  - Mosaicking

- **Raster catalog**
  - A collection of raster datasets
  - Accessed as one entity
  - Each can be accessed as a raster dataset
  - Each member can have its own storage properties
Feature Datasets

- A container object for other datasets
  - Same spatial reference
- Analogous to a coverage
  - Less restrictive
- Contain geometric networks, topologies, terrains, etc…
  - Optionally relationship classes

Subdivision
  - ParcelCorner
  - Parcel
  - ParcelAnno
  - LotLines
  - Parcel_Topo
  - LotDimensions
  - BoundryLines
Validation Rules

• Store attribute, connectivity, and relationship rules on objects as part of the geodatabase

• Predefined, parameter driven
  – Attribute range rule
  – Attribute set rule
  – Connectivity rule

• Perform custom validation by writing code
Domains

• Describe the legal values of a field type
  – Used to ensure attribute integrity

• Defined at the geodatabase level

• Types of domains:
  – Range
    • Valid values between a min / max range
    • A tree can have a height between 0 and 300 feet
    • A road can have between one and eight lanes
  – Coded Value
    • Valid values chosen from a set list
    • A tree can be of type oak, redwood, or palm
    • A road can be made of dirt, asphalt, or concrete
Subtypes

- Categorize objects or features into groups
  - Share the same attributes

- Defined at the class level

- Select a field to base the subtype on
  - Short or long integer field
  - Can have different default values and domains for each field
  - Can define behavior rules between subtypes
Relationship Classes

- **Association between objects in one class and another**
  - A class may participate in multiple relationship classes
- **Simple relationships**
- **Composite relationships**
  - Related objects can message each other
  - Can trigger behavior (cascade delete, move to follow, custom, etc.)
- **Associate rules with relationship classes**
  - Each Parcel can have between 1 to 3 Buildings

<table>
<thead>
<tr>
<th>Parcel</th>
<th>Buildings</th>
</tr>
</thead>
<tbody>
<tr>
<td>OID</td>
<td>Zone</td>
</tr>
<tr>
<td>28</td>
<td>Commercial</td>
</tr>
<tr>
<td>794</td>
<td>Residential</td>
</tr>
<tr>
<td>858</td>
<td>Residential</td>
</tr>
</tbody>
</table>
**Relationship Classes**

- **Association between objects in one class and another**
  - A class may participate in multiple relationship classes
- **Simple relationships**
- **Composite relationships**
  - Related objects can message each other
  - Can trigger behavior (cascade delete, move to follow, custom, etc.)
- **Associate rules with relationship classes**
  - Each Parcel can have between 1 to 3 Buildings

### Parcel

<table>
<thead>
<tr>
<th>OID</th>
<th>Zone</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>Commercial</td>
<td>10000</td>
</tr>
<tr>
<td>794</td>
<td>Residential</td>
<td>5000</td>
</tr>
<tr>
<td>858</td>
<td>Residential</td>
<td>6050</td>
</tr>
</tbody>
</table>

### Buildings

<table>
<thead>
<tr>
<th>OID</th>
<th>Parcel_ID</th>
<th>Building Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28</td>
<td>Office Building</td>
</tr>
<tr>
<td>2</td>
<td>794</td>
<td>Townhouse</td>
</tr>
<tr>
<td>3</td>
<td>794</td>
<td>Townhouse</td>
</tr>
<tr>
<td>4</td>
<td>858</td>
<td>Condo</td>
</tr>
</tbody>
</table>
Annotation

- Annotation feature classes may be
  - Feature linked or Non-feature linked

- Composite relationship manages link

- Can store text as well as other graphics
  - Lines, arrows, boxes, etc...

![Diagram showing feature classes and relationship classes with annotated locations and relationships.](image-url)
Dimension Features

- Type of annotation that displays specific distances on a map

- Graphic features stored in a dimension feature class

- “Smart” feature
  - Special drawing
  - Special editing
Object Behavior

• You can:
  – Instantiate classes with predefined behavior. *(Dimensions and Annotation)*
  – Control the default value and acceptable values for any attribute in a class. *(Domains and Validation)*
  – Partition the objects in a class into like groups. *(Subtypes)*
  – Control the general and network relationships in which an object can participate. *(Relationship Classes)*

• Out of the Box in ArcGIS!
  – Configurable, no programming required
Exploring a Geodatabase Demo

• Explore a Geodatabase
  – Tables
  – Feature Classes
  – Subtypes
  – Domains
  – Relationship Classes
Session Path

• The Geodatabase
• Inside the Geodatabase

• Advanced Behavior
  – Geometric Networks
  – Network Datasets
  – Geodatabase Topology
  – Advanced behavior DEMO

• Geodatabase Potpourri
**Geometric Networks**

- Uses edges and junctions to model network systems
- Each feature class has a role in the network
  - A network may have multiple feature classes in the same role
- Connectivity relationships between feature classes
  - Based on geometric coincidence
  - Can associate connectivity rules with the network
  - Connectivity is maintained **on the fly**
Geometric Networks

- A geometric network is associated with a logical network
  - Each network feature is associated with one or more elements in the logical network

- Trace solvers on the logical network provide
  - Connectivity tracing, cycle detection, flow directions
  - Upstream/downstream tracing, Isolation tracing
Network Datasets

- Network designed for the transportation industry
- Does not replace the Geometric Network
- Multimodal scenarios
- Edges & Junctions
- Attributes
  - Properties to control traversability
    - Cost, restriction, descriptor
  - On-the-fly calculation of costs
  - Improves analysis
Network Dataset Functionality

• **Multimodal**
  – Points span multiple connectivity groups
  – used to create connectivity between lines in different groups

• **Turns**
  – Turns do not alter connectivity, but traversability (e.g. U-Turn restriction)
Geodatabase Topology

- A topology manages a set of simple feature classes that share geometry

- Topology is used to
  - Integrate feature geometry
  - Validate features
  - Control editing tools
  - Define relationships between features
  - Ensure the quality of your data
Topological Integrity

• Topology defines integrity rules for associated feature classes
  – Participating feature classes / subtypes
  – Cluster tolerance, ranks and rules
    • Cluster Tolerance for XY and Z

• Rules are evaluated during Validation
  – Define rules when creating the Topology

• Violations of these rules are expressed as error features managed in the database as a part of the topology
  – Error and Exceptions
  – Examine and Fix errors in ArcMap
Topology Error Examples

- Rules enforced to maintain topological integrity
  - 25+ topology rules in ArcGIS

**Must not overlap**
- Polygons must not overlap within a feature class or subtype.
  - Polygons can be disconnected or touch at a point or touch along an edge.

**Must be properly inside polygons**
- Points in one feature class or subtype must be inside polygons of another feature class or subtype.

**Must not have dangles**
- The end of a line must touch any part of one other line or any part of itself within a feature class or subtype.

- Point errors are created where the points are outside or touch the boundary of the polygons.

- Point errors are created at the end of a line that does not touch at least one other line or itself.
Editing with a Topology

• Editing creates a **dirty area**
  – Area has been edited and may contain errors
  – Can be symbolized

• Errors are found during **validation**
  – Errors have properties
    • What rule was violated
    • Which feature(s) created the error

• Your options:
  – Ignore the error
  – Mark as exception
  – Fix the error
Geodatabase Behavior Demo

• Explore a Geodatabase
  – Topology
  – Geometric Network
Session Path

• The Geodatabase

• Inside the Geodatabase

• Advanced Behavior

• Geodatabase Potpourri
  – Terrains
  – Cartographic representations
  – Cadastral
Terrains

• Massive point datasets in a multi-resolution, on-the-fly generated TIN
  – Dataset for modeling 3D surfaces
  – Modeled within a feature dataset
  – User defined terrain (pyramid) levels
    • Different resolutions & vertical tolerances

• Requires 3D Analyst
  – Extension to define & edit
  – No license needed to view
Representations

• Property of a feature class
  – Stores info about feature symbology

• One feature class - multiple representations

• Rules and overrides

• Representation Management Toolset
Cadastral Editor

- Solution for parcel data management
  - Survey Analyst extension

- Uses COGO attributes and survey control to improve spatial accuracy

- Cadastral editing
  - Cadastral editor toolbar
  - Cadastral fabrics
    - Group layer with sublayers
    - Jobs
Summary

• The Geodatabase
  – Data model, Storage, Transaction model, COM components

• Inside the Geodatabase
  – Datasets, Validation rules, data behavior and integrity

• Advanced Behavior
  – Geometric Networks, Network Datasets, and Topology

• Geodatabase Potpourri
  – Terrains, Representations, Cadastral
Other Sessions

- **Geodatabase Essentials Part 1**
  - Wed 1:30 pm Room 6C \ Fri 8:30 am Room 10

- **Geodatabase Essentials Part 2**
  - Wed 8:30 am Room 4 \ Thurs 8:30 am Room 3 \ Fri 9:00 am Room 4

- **Managing Distributed Data with Geodatabase Replication**
  - Tues 3:15 pm \ Thurs 10:15 am Room 6D

- **Topology in the Geodatabase**
  - Tues 1:30 pm \ Thurs 8:30 am Room 6C

- **Geometric Networks in the Geodatabase**
  - Wed 1:30 pm Room 3

- **Working with Raster Data in ArcGIS**
  - Wed 1:30 pm Room 6D

- **Editing with ArcGIS – Tips and Tricks**
  - Wed 8:30 am \ Thurs 10:15 am Room 3

- **Geodatabase Editing Workflows – An Introduction**
  - Wed 8:30 am \ Thurs 1:30 pm Room 6C

- **Geodatabase Editing Workflows – Advanced**
  - Wed 10:15 am \ Thurs 3:15 pm Room 6C
Other Geodatabase Resources

Geodatabase Resource Center - http://resources.esri.com/geodatabase/

Inside the Geodatabase Blog - www.esri.com/geodatabaseblog