# Evaluating Personal Geodatabases in Arc8

Eric Compas
October 20, 2003

## *Overview*

The purpose of this evaluation is to assess whether personal geodatabases are a better way of storing and managing the spatial and non-spatial NPS Inventorying & Monitoring data. The potential is to hold all data and metadata within one single database that would better maintain linkages and error checking between all components. This evaluation is to ask several specific questions about the capabilities of personal geodatabases within ArcGIS8 in managing this data. Additionally, several general geodatabase issues were examined. Terms that appear in **bold** below are specific ArcGIS terms and are defined in the glossary. Answers to questions are addressed in two ways: 1) how well the question can be answered given out-of-the-box functionality of software, and 2) how well the question can be addressed given additional coding or support tools. All testing was done with ArcGIS Workstation--personal geodatabases are severely restricted within ArcGIS Desktop.

## *Summary of conclusions*

- There were no insurmountable hurdles identified in combining a personal geodatabase with an I&M database.
- ArcGIS does recognize and honor the referential integrity setting made within Access on a relationship.
- ArcGIS does honor the validation rules that are specified within Access on a field.
- Geodatabases can be corrupted by making changes within Access. Some minor changes (hiding tables) can be made to minimize the chance of this.
- Access does not recognize **relationship classes** or **domains** defined in ArcGIS.
- Using **relationship classes** requires adding an ObjectID field to any table that participates in a relationship.
- Learning how to create, add data to, and manipulate personal geodatabases is non-trivial and does take some time to learn.

### Suggestions

These are specific to using personal geodatabases for I&M data.

- Further testing of spatial and non-spatial data combined in single personal geodatabase.
- Use ArcMap for spatial functions (editing features, topology, spatial analysis), but use Access for most of the data entry, validation of attributes (business rules), and table relationships. Access contains more sophisticated tools to ensure valid data and table relationships and to show errors when conflicts occur.
- Avoid use of **relationship class** and **domains** within ArcGIS. Use joins with referential integrity and validation rules within Access instead. Use ad-hoc joins and related within ArcGIS to display or analyze information in related tables.

# *Questions*

## Can personal geodatabases be opened (reading, editing) by more than one user in more than one application?

Yes, the personal geodatabase can be opened by:

- More than one user across a network in same or different applications, e.g. ArcMap, Access.
- More than one application on a machine.

Once a database is opened or viewed in ArcCatalog, the program places a lock on the database until ArcCatalog is closed (even if the user "closes" the geodatabase in the tree view). This prevents the opening of the database by Access in exclusive mode, i.e. required to make changes to forms, reports, or modules. Changes to tables (including adding fields or records) can be done in Access while ArcCatalog has this lock. However, these changes may not be visible to ArcCatalog until Access or ArcCatalog is closed (not sure where the edited and un-edited versions of the database are stored when this occurs). Conversely, if database is opened exclusively in Access, ArcCatalog can't even view the personal geodatabase--displays error message that can't create lock because of existing lock.

If two users are accessing the same personal geodatabase in ArcMap and one initiates an edit session, the other is still able to view the data but not enter edit mode.

From ArcGIS help: "Schema locking in personal geodatabases works in much the same way, except the locks are database-wide. Once an exclusive or shared lock is acquired on an item in a personal geodatabase, that lock applies to all items in the geodatabase." (under *Schema locking* heading)

The Microsoft Knowledge Base article 285822 describes how to list current users (reading the lock *.ldb databse). However, the results of the tool don't show which application have

what type of lock. The older application LDBView also doesn't give any more information.

Needed:

- Better way of figuring out which application has what type of lock on database.
- Tool to have ArcCatalog release lock when no longer using personal geodatabase.

## Can we maintain referential integrity between sample features and the non-spatial locations table (tbl_Locations)?

Referential integrity is a complex topic with many ways in which integrity between tables can be maintained. Below is a first look that is not intended to be conclusive.

In out-of-the-box ArcGIS, the answer is a qualitative yes.

- ArcGIS honors its own **relationship classes**.
    - I created a composite **relationship class** in ArcCatalog linking sample locations (tbl_Locations) to a point **feature class** representing these points (test_points). When trying to add an additional point (to test_points) using ArcMap, the editor would not let me commit the edits because the tbl_Locations did not have a related record. ArcMap, however, would let me delete points from test_points.
- ArcGIS recognizes joins defined within Access with **referential integrity** (this is noteworthy!).
    - When ArcMap allows editing on a linked table (I created a test table that was linked to tbl_Locations on LocationID), it won't let you modify records that would break the relationship. For example, I could only add records to the test table when the LocationID matched an existing one in tbl_Locations. ArcMap wouldn't always allow you to edit a table. For example, I could not edit records within the tbl_Locations table. Oddly enough, I could delete any fields in the table that don't participate in a relationship. I haven't figured out what triggers this.
    - If I edited a point feature which contained a join with referential integrity. I could add points with <null> values for the linked field (LocationID). If I edited this field, however, ArcMap would only accept values that matched those that existed within tbl_LocationID. Looks like Access does not mind <null> values as part of a relationship (haven't played with "Accept NULL value" settings).

In out-of-the-box Access, Access does not recognize **relationship classes** defined with ArcMap. Tools could be built to read the GDB_RelClasses table (what holds these classes) and define them as Access relationships. It does recognize it own defined joins with

referential integrity.

Tools needed:

- Way to analyze database to find referential integrity errors, e.g. spatial feature missing for sample location.

## Can we apply business rules (validation rules) to fields within the personal geodatabase? Same procedure for Access and ArcGIS?

In Access, you define business rules as validation rules. In ArcGIS, they are defined as **domains**. The validation rules in Access have much more flexibility than those of **domains**.

When editing tables in ArcMap, validation rules are checked before changes are committed (when you try to save edits). The error message does not specify which record (s) failed to pass the validation rules.

Domains, however, are not checked (odd). In fact, I can only figure out how to check domains when using a **feature class** and not just a table.

**Domains** within Access are ignored. Again, a tool could be created that examined defined domains and translates them to Access validation rules.

## Can feature datasets hold any geometry types (unlike coverages)?

Yes, they can hold any combination of point, multipoints, lines, and polygon **feature datasets**.

## What changes "break" or corrupt a personal geodatabase?

What does:

- Changing the table name of a **feature class**. ArcCatalog can't find it.
- Deleting one of the geodatabase, e.g. GDB_<name>, tables. For example, if the GDB_ObjectClasses database is deleted, ArcCatalog can no longer locate any of the **feature classes**.

What doesn't

- Adding new fields or records to the **feature classes's** table.

These list obviously aren't exhaustive. I'm sure there are many more ways to corrupt a

geodatabase.

# Other observations

## General disadvantages of geodatabases

- Complex management.
- New set of terminology (much of it redundant and unnecessary).
- Error messages when problem encountered in ArcGIS uninformative, e.g. importing errors in ArcCatalog.
- Have to add field to non-spatial tables to be "seen" by the geodatabase.
- Slower at drawing and selecting than shapefiles.
- Larger file sizes than shapefiles. Example: 53,000 polygons, shapefile 61.6 MB, geodatabase 65.4 MB.
- Can't natively store **layers** within geodatabase (can with ESRI sample).
- Different versions of personal geodatabases given software releases. What changes will ESRI make in the future that might mess up our coding?

## Joins/relates/relationship classes

- In ArcMap, even if a relationship class is established between a feature class and a table, you still have to create a join in order to expose this relationship for symbolizing, browsing table data, etc. The only exception is using the identify tool which does show field in the related table by default.
- In ArcMap, have no way of knowing that a relationship class has been defined for a feature class or table.
- Advantage of relationship classes may exist when editing. For example, if one feature is deleted, the related tabular data is also deleted.

# Potential tools

The following are suggestions for tools with in ArcCatalog or ArcMap that would help in creating and managing personal geodatabases:

- Batch import **feature classes** into **feature datasets**.
- Change the color or symbol for personal geodatabase tables within Access to mark them for users.
- Tools to convert **relationship classes** and **domains** to Access equivalents.
- View metadata within Access. Way to view XML metadata stored within a geodatabase within access so that ArcCatalog isn't required merely to browse metadata.
- View data within Access. Way to view features within a geodatabase within a

form or report (using ESRI's mapcontrol). Looks like there are still some problems with doing this).

# *Glossary*

Domain

The values or limits of an attribute used to validate attribute values. For numeric fields, can limit minimum and maximum values. For text fields, can limit values to a particular set of **coded values**. To view or edit, right click the personal geodatabase in ArcCatalog and choose properties.

Feature class

Actual collections of spatial features. Analogous to shapefiles or the components (points, lines, or polygons) of a coverage.

Feature dataset

A grouping of feature classes that share a similar spatial reference (projection and x/y domains). Feature datasets can be used to group features thematically. (Why use them? Ensure that new/imported feature lie within a particular extent? Because that's where you can define topology rules. Can't do that for a whole geodatabase or individual feature class.)

Relationship classes

Defined relationship (similar to relates or joins in ArcView) that tie a field in one table to that of another. Tables have to be registered with the geodatabase to create relationship (this requires adding an ObjectID field to your table).