



Working with the Geodatabase: Powerful Multiuser Editing and Sophisticated Data Integrity

An ESRI® White Paper • February 2002

Copyright © 2002 ESRI
All rights reserved.
Printed in the United States of America.

The information contained in this document is the exclusive property of ESRI. This work is protected under United States copyright law and other international copyright treaties and conventions. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system, except as expressly permitted in writing by ESRI. All requests should be sent to Attention: Contracts Manager, ESRI, 380 New York Street, Redlands, CA 92373-8100, USA.

The information contained in this document is subject to change without notice.

U.S. GOVERNMENT RESTRICTED/LIMITED RIGHTS

Any software, documentation, and/or data delivered hereunder is subject to the terms of the License Agreement. In no event shall the U.S. Government acquire greater than RESTRICTED/LIMITED RIGHTS. At a minimum, use, duplication, or disclosure by the U.S. Government is subject to restrictions as set forth in FAR §52.227-14 Alternates I, II, and III (JUN 1987); FAR §52.227-19 (JUN 1987) and/or FAR §12.211/12.212 (Commercial Technical Data/Computer Software); and DFARS §252.227-7015 (NOV 1995) (Technical Data) and/or DFARS §227.7202 (Computer Software), as applicable. Contractor/Manufacturer is ESRI, 380 New York Street, Redlands, CA 92373-8100, USA.

@esri.com, 3D Analyst, ADF, AML, ARC/INFO, ArcAtlas, ArcCAD, ArcCatalog, ArcCOGO, ArcData, ArcDoc, ArcEdit, ArcEditor, ArcEurope, ArcExplorer, ArcExpress, ArcFM, ArcGIS, ArcGrid, ArcIMS, ArcInfo Librarian, ArcInfo, ArcInfo—Professional GIS, ArcInfo—The World's GIS, ArcLogistics, ArcMap, ArcNetwork, *ArcNews*, ArcObjects, ArcOpen, ArcPad, ArcPlot, ArcPress, ArcQuest, ArcReader, ArcScan, ArcScene, ArcSchool, ArcSDE, ArcSdl, ArcStorm, ArcSurvey, ArcTIN, ArcToolbox, ArcTools, ArcUSA, *ArcUser*, ArcView, ArcVoyager, *ArcWatch*, ArcWeb, ArcWorld, Atlas GIS, AtlasWare, Avenue, *BusinessMAP*, Database Integrator, DBI Kit, ESRI, ESRI—Team GIS, ESRI—The GIS People, FormEdit, Geographic Design System, Geography Matters, Geography Network, GIS by ESRI, GIS Day, GIS for Everyone, GISData Server, *InsiteMAP*, MapBeans, MapCafé, MapObjects, ModelBuilder, MOLE, NetEngine, PC ARC/INFO, PC ARCPLOT, PC ARCSHELL, PC DATA CONVERSION, PC STARTER KIT, PC TABLES, PC ARCEdit, PC NETWORK, PC OVERLAY, Rent-a-Tech, *RouteMAP*, SDE, SML, Spatial Database Engine, StreetEditor, StreetMap, TABLES, the ARC/INFO logo, the ArcAtlas logo, the ArcCAD logo, the ArcCAD WorkBench logo, the ArcCOGO logo, the ArcData logo, the ArcData Online logo, the ArcEdit logo, the ArcEurope logo, the ArcExplorer logo, the ArcExpress logo, the ArcFM logo, the ArcFM Viewer logo, the ArcGIS logo, the ArcGrid logo, the ArcIMS logo, the ArcInfo logo, the ArcLogistics Route logo, the ArcNetwork logo, the ArcPad logo, the ArcPlot logo, the ArcPress for ArcView logo, the ArcPress logo, the ArcScan logo, the ArcScene logo, the ArcSDE CAD Client logo, the ArcSDE logo, the ArcStorm logo, the ArcTIN logo, the ArcTools logo, the ArcUSA logo, the ArcView 3D Analyst logo, the ArcView Business Analyst logo, the ArcView Data Publisher logo, the ArcView GIS logo, the ArcView Image Analysis logo, the ArcView Internet Map Server logo, the ArcView logo, the ArcView Network Analyst logo, the ArcView Spatial Analyst logo, the ArcView StreetMap 2000 logo, the ArcView StreetMap logo, the ArcView Tracking Analyst logo, the ArcWorld logo, the Atlas GIS logo, the Avenue logo, the *BusinessMAP* logo, the Data Automation Kit logo, the Digital Chart of the World logo, the ESRI Data logo, the ESRI globe logo, the ESRI Press logo, the Geography Network logo, the MapCafé logo, the MapObjects Internet Map Server logo, the MapObjects logo, the MOLE logo, the NetEngine logo, the PC ARC/INFO logo, the Production Line Tool Set logo, the *RouteMAP* IMS logo, the *RouteMAP* logo, the SDE logo, The World's Leading Desktop GIS, *Water Writes*, www.esri.com, www.geographynetwork.com, www.gisday.com, and Your Personal Geographic Information System are trademarks, registered trademarks, or service marks of ESRI in the United States, the European Community, or certain other jurisdictions.

Other companies and products mentioned herein are trademarks or registered trademarks of their respective trademark owners.

Working with the Geodatabase: Powerful Multiuser Editing and Sophisticated Data Integrity

An ESRI White Paper

Contents	Page
Overview	1
Why Use the Geodatabase?	2
Personal and Multiuser Geodatabases	2
Centralized GIS Data Management	2
Multiuser Editing.....	3
Support for Intelligent Features, Rules, and Relationships	3
Feature-Linked Annotation	3
Edit Topologically Integrated Features	4
Integrity Constraints for More Efficient Data Entry.....	4
Subtypes	4
Domains	4
Geometric Networks	5
Sources and Sinks	6
Dimensions	6
Relationships	6
Validation Rules	7
Attribution Rules	7
Connectivity Rules	7
Relationship Rules	8
Custom Rules	8
Data Models	8
Accessible Through Custom Applications.....	8
Comparing Functionality Between ArcView, ArcEditor, and ArcInfo.....	8
Conclusion.....	9

Working with the Geodatabase: Powerful Multiuser Editing and Sophisticated Data Integrity

Overview ArcGIS™ software is fundamentally based on a geodatabase (short for geographic database) data model. The geodatabase is a unique data format that is similar in structure to the coverage data model and also includes the functionality for multiuser editing. A geodatabase is a storage mechanism for spatial and attribute data that contains specific storage structures for features, collections of features, attributes, relationships between attributes, and relationships between features. The geodatabase has two major concepts: first, a geodatabase is a physical store of geographic information inside a database management system (DBMS); second, a geodatabase has a data model that supports transactional views of the database (versioning) that also supports objects with attributes and behavior. Behavior describes how an object (feature) can be edited and displayed. Behavior includes, but is not limited to, relationships, validation rules, subtypes, and default values. With associated behaviors data entry can be regulated more efficiently, and data contamination issues can be avoided.

Some of the major benefits acquired by migrating to the geodatabase are

- Centralized management of a wide variety of geographic information in a DBMS.
- Large data management in a continuous integrated environment (raster/vector).
- Full support for multiuser editing in an advanced versioning environment.
- Support for advanced data types and geometry (i.e., the ability to store three-dimensional coordinates, measures, and true curves).
- Faster and more efficient data entry with data rules and relationships.
- Create and edit feature-linked annotation.
- Create and edit geometric networks.
- Relationships with tabular data sources stored in any industry-standard DBMS.

- Create and edit topologically integrated features.
- Create subtypes and domains for maintaining database integrity.
- Support for the development of industry-standard data models.

Why Use the Geodatabase?

The chief purpose of the geodatabase data model is to make it easier and more intuitive to use geographic information system (GIS) software across many applications. As such, the geodatabase is sometimes referred to as the next generation coverage. It is that and much more. The geodatabase data model centralizes data management and opens up the use of GIS to applications that were not feasible before. The geodatabase includes a variety of sophisticated editing features that allow the GIS analysis to be more efficient and less time-consuming.



Sophisticated Feature Editing with Rules and Relationships

Personal and Multiuser Geodatabases

The geodatabase comes in two varieties—personal and multiuser. Personal geodatabase support is built into ArcGIS and implemented using standard desktop tools and the Microsoft Access file format for project-level GIS. Multiuser databases are deployed using ArcSDE™ software and require a DBMS such as IBM® DB2®, Informix®, Oracle®, or Microsoft® SQL Server™. ArcEditor™ 8.1 or ArcInfo™ 8.1 software can create and edit personal and multiuser geodatabases, coverages, shapefiles, CAD, and a variety of other data formats. ArcView® 8.1 software works with personal geodatabases and shapefiles.

The geodatabase, whether personal or multiuser, can store tables, feature classes, feature data sets, and functionality such as rules, relationships, and geometric networks. Using a geodatabase is a simple and straightforward process. Geodatabases can be created, accessed, and managed through the standard menus and tools in ArcCatalog™, ArcToolbox™, and ArcMap™ software applications. The geodatabase model supports intelligent features, rules, and relationships that users can employ in GIS applications.

Centralized GIS Data Management

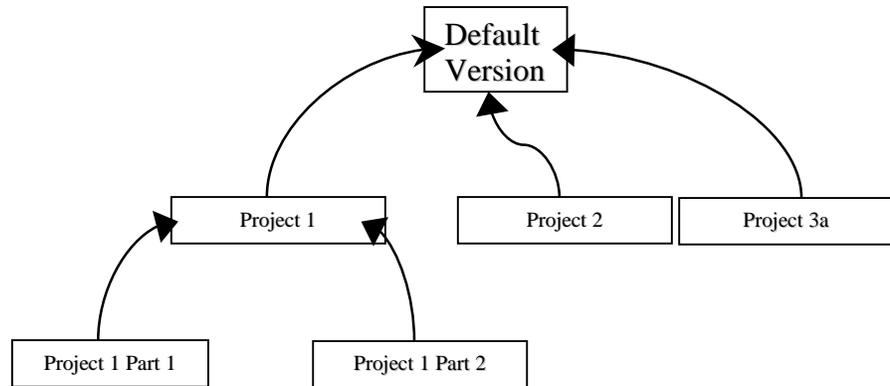
A key feature of the geodatabase is that all data (vector, raster, address, measures, CAD, etc.) is stored in a commercial off-the-shelf DBMS. The geodatabase is the natural solution for handling the growing amount of digital spatial data and at the same time

meets the need to leverage this data among the growing numbers of GIS users. This gives organizations the power to have an integrated data management policy covering all data, which can significantly simplify support and maintenance as well as reduce costs. The geodatabase supports multiple formats of spatial data including

- Simple features such as shapefiles
- Custom features with business logic and editing rules
- Attribute data
- Metadata
- Images
- Raster/Grid data
- CAD data

Multiuser Editing

The geodatabase can be used for multiuser access to databases storing continuous (nontiled) spatial data through the use of versioning and long transactions. Versioning allows simultaneous editing by multiple users (subject to DBMS permissions). This framework lets the analyst create versions of a geodatabase for the stages of a project, reconcile differences between versions, and update the master version of a geodatabase with the design as-built. Versioning is not supported by the personal geodatabase. For more information on versioning, download the white papers titled *Managing Workflow with Versions* and *Multiuser Geographical Information Systems with ArcInfo 8* from the ESRI Web site at arconline.esri.com.



Multilevel Tree

Support for Intelligent Features, Rules, and Relationships

The geodatabase data model supports, as standard, a rich collection of objects (rows in a database table) and features (objects with geometry). The geodatabase supports advanced feature types such as geometric and logical networks, true curves, complex polylines, and user-defined features. Vector features can now have two, three, or four dimensions (x, y, z , and m). Users can define topological and associative relationships and rules that define how feature classes interact. All this can be done without programming using the menu-driven graphical user interface.

Feature-Linked Annotation

Once you have labeled your spatial features, you can then convert the labels into an annotation feature class. An option in this process is to make the annotation linked to the spatial features so that when a spatial feature is moved, the label will move with it.

Edit Topologically Integrated Features

Topologically associated feature classes in a geodatabase should be stored in an integrated feature data set. A feature data set is a collection of feature classes. The user can use the topological editing tools (such as the Shared-Edge edit tool) in ArcMap to maintain the topological associations between features in an integrated feature data set. Since it is common for vector data sets to have features that share common boundaries or corners, the ability to directly edit all of these features within one spatially common feature data set is a huge advantage for the GIS community.

Integrity Constraints for More Efficient Data Entry

By managing data in a geodatabase, the user can define rules about how the data can be edited. Using the geodatabase, any user can set up subtypes and domains to ensure the attribute integrity of the database.

Subtypes

With subtypes the analyst can define categories for his/her features, which allow the analyst to distinguish objects based on different behaviors, attributes, and access privileges. For example, if telephone poles only have three locations (highways, arterial roads, secondary roads), the user subtypes can be set up for each pole location so that during data entry, users can easily associate the proper attribute. Users can also set defaults for the attribute value that is most likely to be associated with telephone poles. For example, if 90 percent of telephone poles are located on highways, having this attribute set as a default value can save time. ArcCatalog lets the user browse through easily understood descriptions of any subtypes, so the user does not have to remember cryptic coding schemes.

Domains

Domains control the allowable set of values that may be input for any field in a geodatabase table. This limits the possibility of data entry error because the choices for each data value are limited within defined parameters. For example, telephone pole heights are between 20 and 50 feet; therefore, if a value outside this range were entered it would be invalid. There are two types of attribute domains.

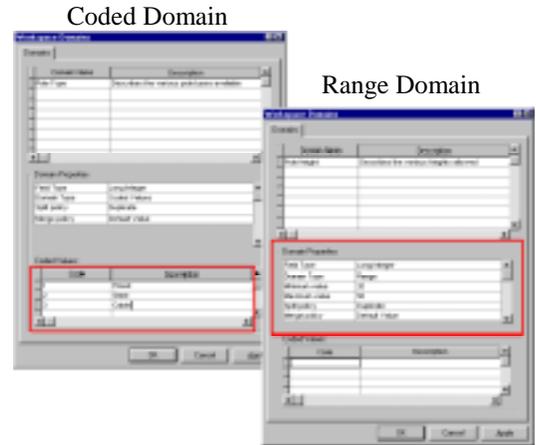
- Coded Values—Specify listings of acceptable codes. For example, the type of telephone poles could be coded as 1 (Wood), 2 (Steel), or 3 (Cable).
- Range—Specify acceptable starting and ending numeric values. For example, telephone pole heights could be between 20 and 50 feet.

To learn more about subtypes and domains, visit the *ArcUser* April 2000 Web page and look under "Learn More About Subtypes and Domains."

J-8783



Subtypes categorize attribute values with a single field in any geodatabase feature class or table.

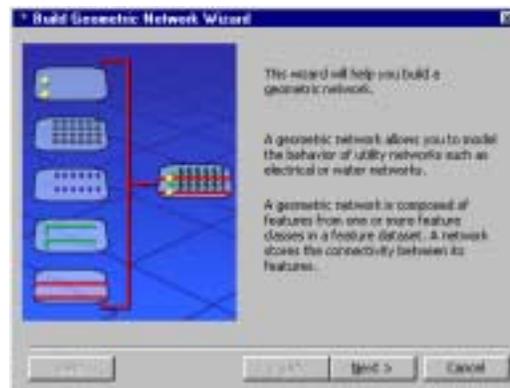


Domains control the allowable set of values for any field in a geodatabase.

Geometric Networks

The movement of people, the transportation and distribution of goods and services, the delivery of resources and energy, and the communication of information all occur through definable network systems. Line and point feature classes in a single feature data set can be organized into a geometric network. The network combines the line and point feature classes into a single database entity to model linear networks—for example, electrical networks—and maintains topological relationships between the participating feature classes. The geometric network automatically maintains the explicit topological relationships between all the features that participate in a geometric network. With ArcCatalog, ArcInfo and ArcEditor users can easily create a geometric network using the Geometric Network wizard. An entirely new and empty geometric network can be devised that a user can then populate. Alternatively, geometric networks can be assembled from existing features in a feature data set.

Geometric networks consist of edge network features and junction network features. An example of an edge feature is a water main, while a junction feature might be a valve.



Friendly Wizard Interface for Creating Geometric Networks

Sources and Sinks

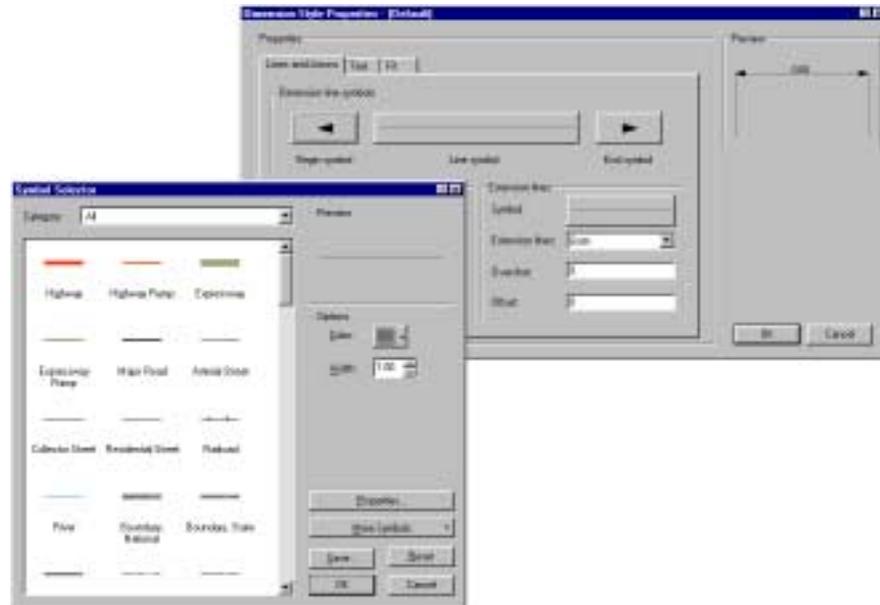
Networks are often used to model real-world systems in which the direction of movement through the network is well-defined. For example, the flow of electricity in an electrical network is from the power generation station to the customers. In a water network, the flow direction may not be as well-defined, but the flow of water may be from a pump station to a customer or from customers to a treatment plant.

Flow direction in a network is calculated from a set of sources and sinks. In the above examples, electricity and water flow are driven by sources and sinks. Flow is away from sources, such as the power generation station or a pump station, and toward sinks such as a water treatment plant.

Dimensions

Dimensions are a special kind of map annotation that show specific lengths or distances on a map. A dimension may indicate the length of a side of a building or land parcel, or it may indicate the distance between two features such as a fire hydrant and the corner of a building.

The geodatabase currently supports two types of dimensions: aligned and linear. Aligned dimensions run parallel to the baseline and represent the true distance between the beginning and ending dimension points. Linear dimensions do not represent the true distance between the beginning and ending dimension points, but aid in analysis and assessment of corresponding features.



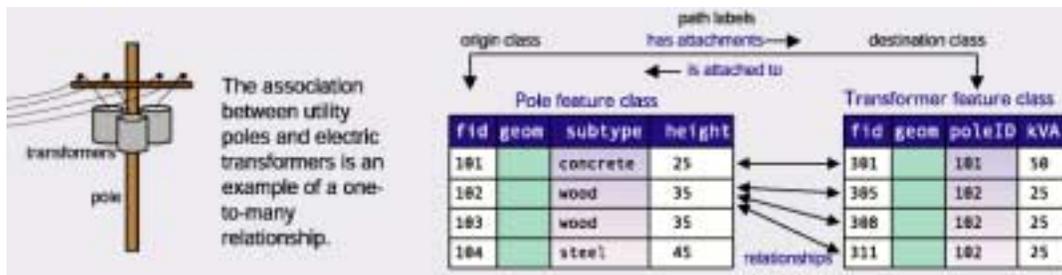
Choose from a variety of style options for customizing the look and feel of a dimension feature class.

Relationships

The geodatabase gives users the power to create advanced relationships among multiple tables in the DBMS. A relationship is an association between two objects in a feature data set. These objects can be nonspatial (objects) or spatial (features). Besides identifying the associated objects, relationships can have additional properties.

Relationships are organized into relationship classes. Each relationship in a relationship class has the same origin class and destinations class. Any object class may participate in many relationship classes. With relationships, the geodatabase ensures referential integrity. For example, a parcel data set may have multiple owners. By creating relationships between different owner databases (i.e., commercial, residential), the city planner can appropriately identify all owners of the parcel. Essentially, tabular relationships in a geodatabase give users the ability to simultaneously interact with related information for their specific analysis.

Pole-Transformer Relationship



The transformers are directly related to the type of pole they reside on.

Validation Rules The geodatabase supports four broad types of validation rules for maintaining database integrity. These include attribute validation rules, network connectivity rules, relationship rules, and custom rules.

Validation rules can ensure that the data entry clerk does not enter invalid entries for the data in question, for example, when modeling a system in which water mains can only be made of cast iron, ductile iron, or copper and there are four possible ground surface types. When the data entry changes values for an attribute, users might want to make sure that only legal or valid values are inserted into the attributes for that lateral.

Attribution Rules Attribute validation rules include subtypes for assigning data categorically as well as attribute domains for constraining the values allowed in any particular attribute for a table, feature class, or subtype. Essentially, attribute domains constrain the values that can be applied to the data so that the data entry clerk will be less likely to enter incorrect values. Subtypes make it easier for this clerk to assign the data into easily understandable categories.

Connectivity Rules In most networks, it may not be desirable to have all edge types to be able to logically connect to all junction types. Similarly, not all edge types can logically connect to all other edge types through all junction types. For example, in a water network, a hydrant can connect to a hydrant lateral but not to a service lateral. Similarly, in the same water network, a 10-inch transmission main can only connect to an 8-inch transmission main through a reducer. Network connectivity rules constrain the type of network features that may be connected to one another and the number of features of any particular type that can be connected to features of another type.

There are two types of connectivity rules: edge–junction rules and edge–edge rules. An edge–junction rule is a connectivity rule that establishes that an edge of type A may connect to a junction of type B. An edge–edge rule is a connectivity rule that establishes that an edge of type A may connect to an edge of type B through a set of junction types.

By establishing these rules, along with other rules such as attribute domains, the analyst can maintain the integrity of the network data in the database.

Relationship Rules

Relationship classes can have an associated set of relationship rules. Relationship rules control which object subtypes from the origin class can be related to which object subtypes in the destination class. They can also be used to specify a valid cardinality range for all permissible subtype pairs. Cardinality describes how many objects of type A are related to an object of type B. For example, the subtype wood pole may be able to support from zero to three transformers, whereas the subtype steel pole may support zero to five transformers. In the first case, the cardinality range would be 0–3; in the second case, it would be 0–5.

Custom Rules

By accessing the underlying ArcObjects™ software through any Component Object Model (COM)-compliant programming language, anyone can also create custom rules for a specific data model. See *Exploring ArcObjects* for more information. The *Exploring ArcObjects* books Volume 1 and Volume 2 are located in the ArcGIS Desktop Media Kit CD 3.

Data Models

The geodatabase data model allows for the creation of common or essential data models for specific industries and applications. ArcGIS data models provide ready-to-use nonproprietary frameworks for modeling and capturing the behavior of real-world objects in a geodatabase. They are built on the accepted standards in each field and provide data models that can be configured and customized using ArcEditor 8.1 (or ArcInfo 8.1). Data migration is the most time-consuming and most costly part of a GIS project. These data models can provide "quick start" solutions that optimize performance and establish industry standards that improve data sharing. For more information about each specific data model, please visit <http://arconline.esri.com/> and click on the data models section.

Accessible Through Custom Applications

Geodatabases can be accessed via the standard menus and tools in ArcCatalog, ArcMap, and ArcToolbox. Programmers can also use the ArcObjects, OLE DB, and SQL application programming interfaces (APIs) included in the software. These data access capabilities are built using industry-standard APIs.

Comparing Functionality Between ArcView, ArcEditor, and ArcInfo

ArcGIS provides a varying level of functionality between ArcView, ArcEditor, and ArcInfo. ArcView allows the user to create and edit personal geodatabases. However, the user cannot create or manage geometric networks and relationships. Basic functionality such as defining attribute domains is included. For a more robust geodatabase environment, ArcEditor or ArcInfo are recommended. ArcEditor and ArcInfo give the analyst the full functionality of the geodatabase. Personal and multiuser geodatabases can be created and edited. Features such as geometric networks, relationships, feature-linked annotation, multiuser editing, custom data models, and image storage are all incorporated within ArcEditor and ArcInfo.

Table 1
Geodatabase Functionality Comparison Between ArcView, ArcEditor, and ArcInfo

Features	Personal Geodatabase with ArcView	Personal Geodatabase with ArcEditor or ArcInfo	Multiuser Geodatabase with ArcEditor or ArcInfo
Editors	One	One	Many
Create and edit simple features (points, lines, areas, static annotation)	✓	✓	✓
Define and use attribute domains	✓	✓	✓
Set database schema	✓ ¹	✓	✓
Versioning (long transactions)			✓
Store images, raster/grid			✓
Create and edit features with subtypes or dimension features		✓	✓
Establish behavior (relationships, geometric networks, feature-linked annotation, etc.)		✓	✓
Create and edit custom features		✓	✓
Database size	≤250K features ²	≤250K features ²	Unlimited
Requires ArcSDE			✓
Supported databases	Microsoft Jet	Microsoft Jet	Oracle Microsoft SQL Server IBM DB2 Informix

¹ Limited to simple features in a personal geodatabase.

² This is an approximate limit affected by two factors—file size and computer memory. Microsoft Jet 4.0, used by the personal geodatabase, has a 2 GB file size limit. In addition, a personal geodatabase is a single file that is loaded into computer memory. Therefore, performance can become unacceptable even for file sizes less than 2 GB. The recommended 250,000 feature limit is based on ESRI's experience with typical GIS data sets stored in a personal geodatabase.

Conclusion The geodatabase integrates the vast flexibility of the ArcInfo coverage and incorporates a variety of new features, which ultimately make geospatial editing less time-consuming and more intuitive. With features such as subtypes and domains, feature-linked annotation, multiuser editing with versioning, and advanced geometric networks, GIS users can effectively take their geospatial analysis to the next level.

Find out more about this robust data format at
<http://www.esri.com/software/arcgis/geodatabase.html>.