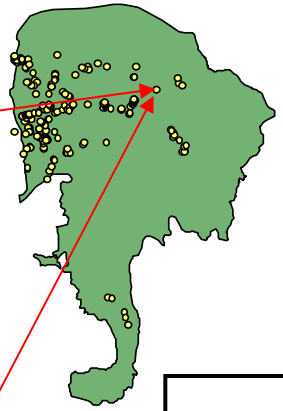
ESRI Object Models; Data Capture

Attributes of Condrey_samples_NAD27

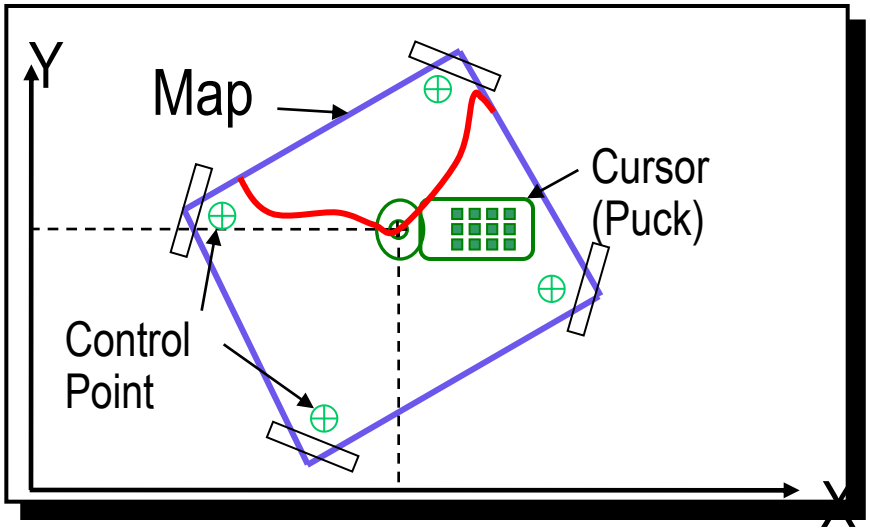
FID	Shape*	Id	Number	NAD27_East	NAD27North
0	Point	0	230	490921	4636832
1	Point	0	123	491124	4637700
2	Point	0	124	491375	4638149
3	Point	0	125	491522	4638902
4	Point	0	126	491653	4638751
5	Point	0	127	492872	4639489
6	Point	0	128	492963	4639515

Feature Class (spatial table)



Number	Age_Ma	1_sigma	Rx_Type	Size_kg
123	142	1.5	B_schist	3.4
124	136	2.0	G_schist	1.3
125				

Object Class (nonspatial table)



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

ESRI = Environmental Systems Research Institute, Inc.

Some ESRI History…

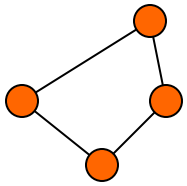
ESRI	Arc/Info	ArcView	ArcGIS
Date	1980-1999	1993-1999	2000 - present
Versions	1-7	1–3.2	8.0 – 10.0
Data Model	Coverage	Shapefile	Geodatabase
O.S.	Unix, PC DOS	Windows	Windows
Scripting Language	Arc Macro Language (AML)	Avenue Scripting	Vis. Basic for Appl. (VBA), Python
Database Software	Proprietary; Arc Tables	DBase	M.S. Access; ArcSDE for Oracle, etc.

ESRI Data Models

- **Topologic:**
  - ArcInfo - Coverage
  - ArcInfo ".EOO" – export format for coverage
  - ArcGIS - Geodatabase
- **Non-Topologic:**
  - ArcView (legacy) - Shapefile

Early ESRI Data Models

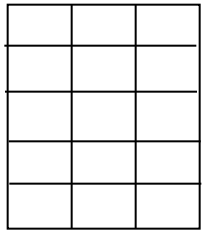
Spatial Data



Arc

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

Aspatial Data



Info

- 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

Slide courtesy of D. Maidment

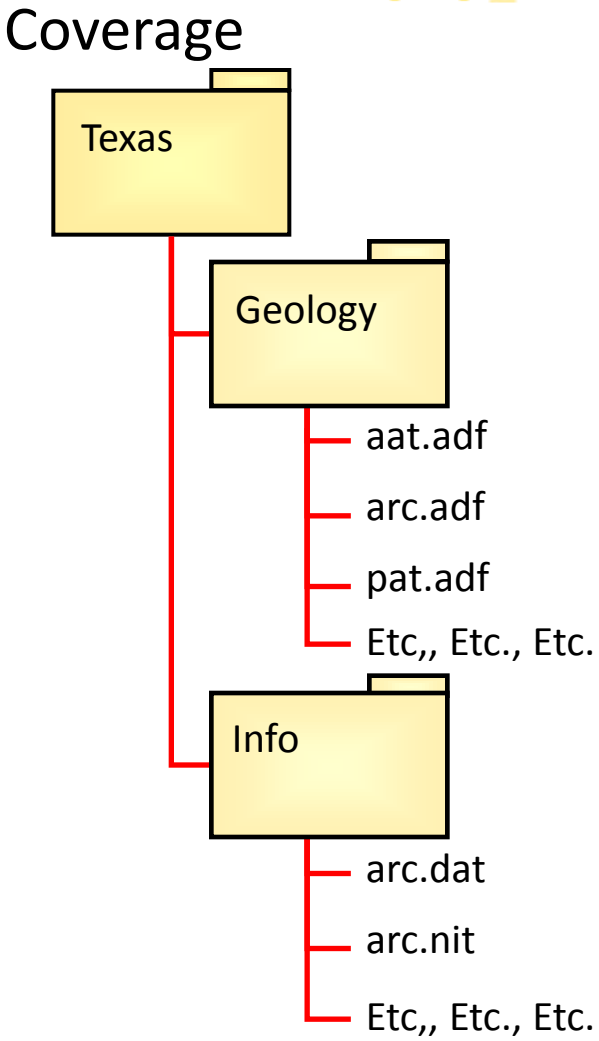
Data Organization

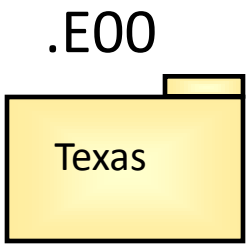
- Coverage
  - Data split between coverage and INFO *folders*
  - Common boundaries between polygons stored once
  - Topology explicitly stored
    - Planar graph maintained

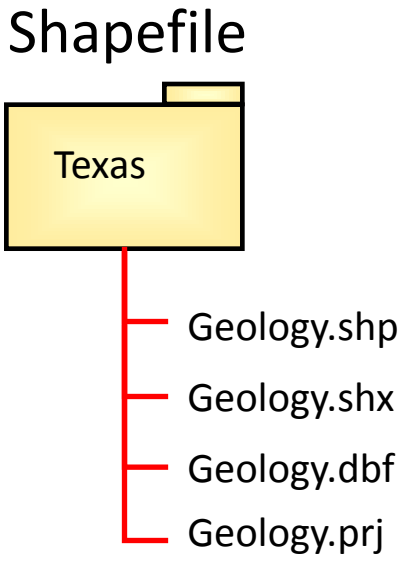
As in previous lecture

- 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



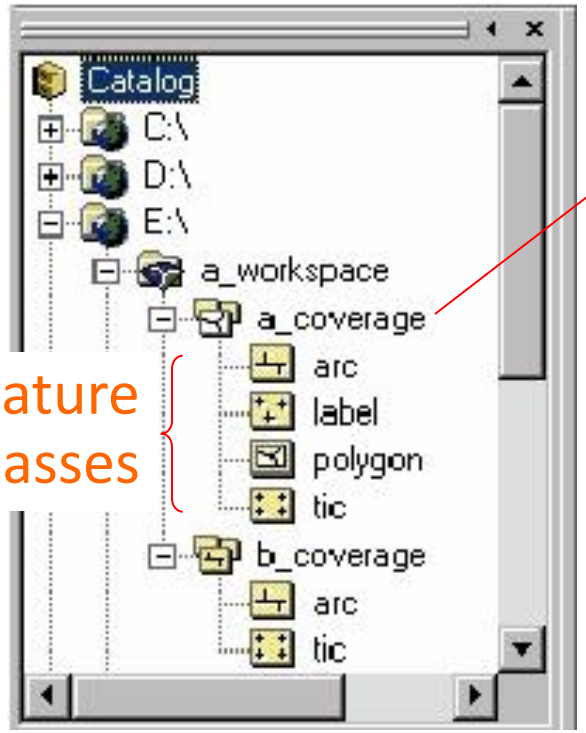




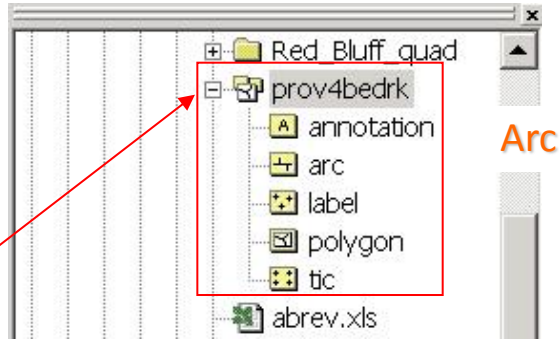
- One object shape (as points OR lines OR polygons) per file = "SHAPEFILE"
- Many related objects (as points AND lines AND polygons) per file = "COVERAGE"

Data Organization: Coverage in Windows Explorer and ArcCatalog

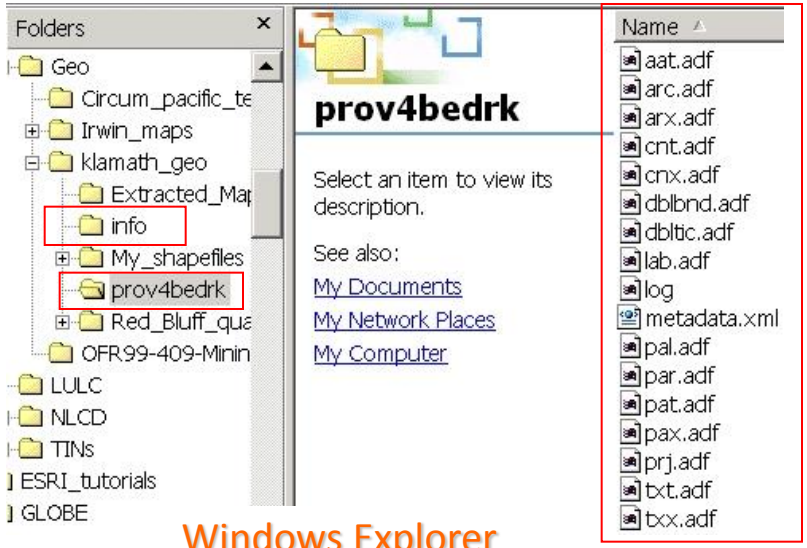
- ArcCatalog:
  Workspace/Coverage/
  Feature Class



Feature Classes



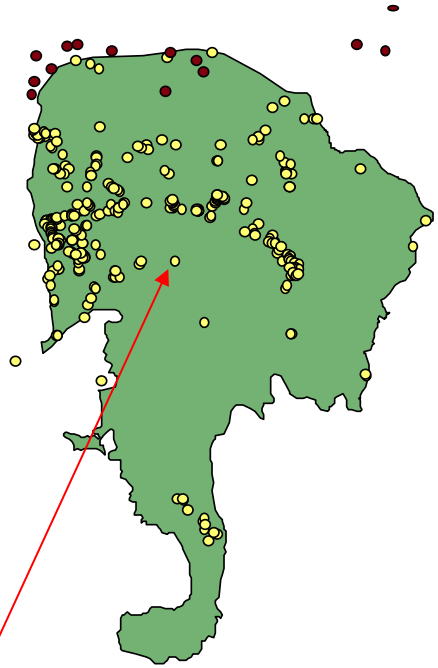
Arc Catalog



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



Attributes of Condrey_samples_NAD27

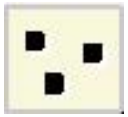
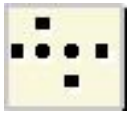
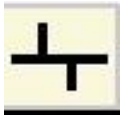
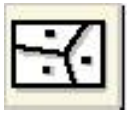
FID	Shape*	Id	Number	NAD27_East	NAD27North
0	Point	0	230	490921	4636832
1	Point	0	123	491124	4637700
2	Point	0	124	491375	4638149
3	Point	0	125	491522	4638902
4	Point	0	126	491653	4638751
5	Point	0	127	492872	4639489
6	Point	0	128	492963	4639515

Sample location (points) Feature Class

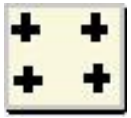
ArcInfo Coverage

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

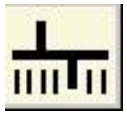
ArcInfo Coverages can contain:

- 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:

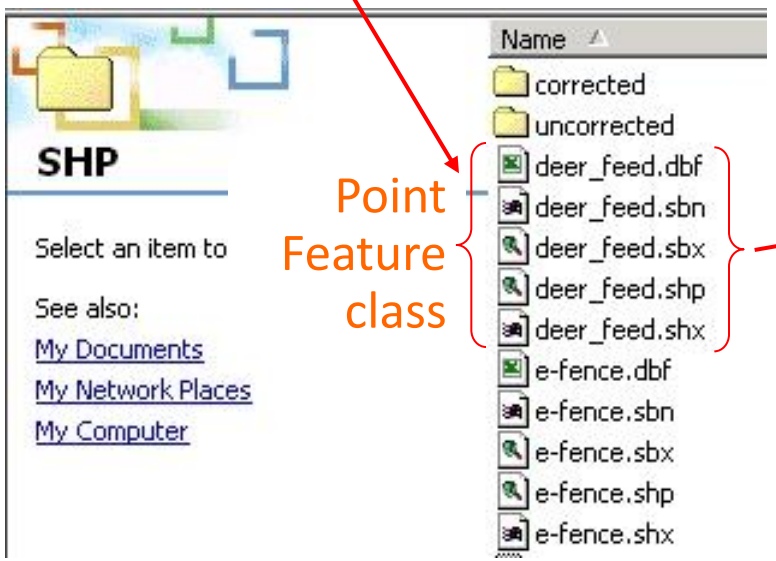
- Composite features:
  -  Routes – collections of Arcs with measurement system
  -  Regions – collections of polygons; adjacent, noncontiguous or overlapping

Shapefile format

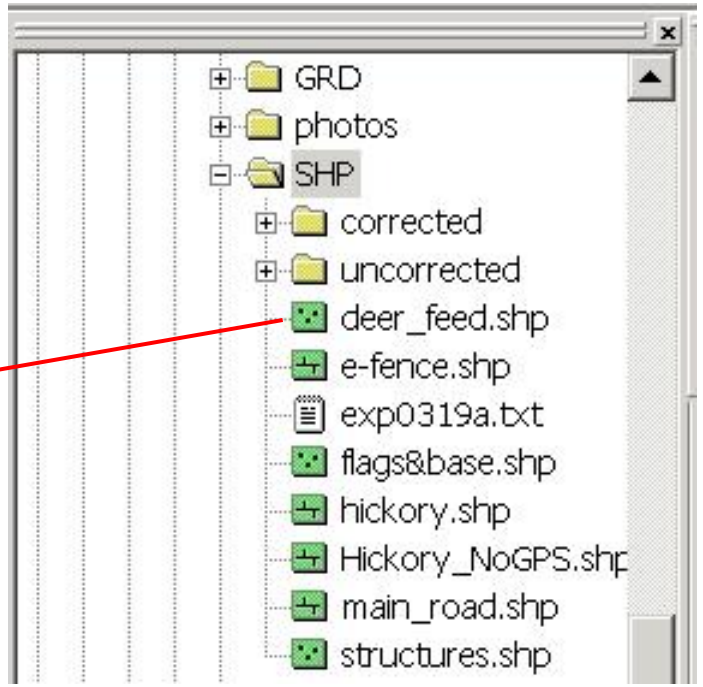
- Simpler than coverage; doesn't store topology
- Feature classes stored independently i.e. points, lines and polys. stored in physically separated files (e.g. no shared INFO table)
- For each type, spatial data stored in a .shp file, attribute data in a .dbf table.
- "Null" or "No Data" numerical values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

Shapefiles in ArcCatalog/Explorer

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



Window Explorer

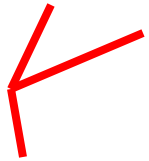


ArcCatalog

Shapefile feature class types:

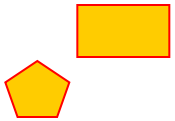
- Point, Multipoint

- Polyline (line with several paths)



- Polygon

Ring – closed, nonintersecting path – simple poly.



Disjointed Rings – multiple polygons define feature

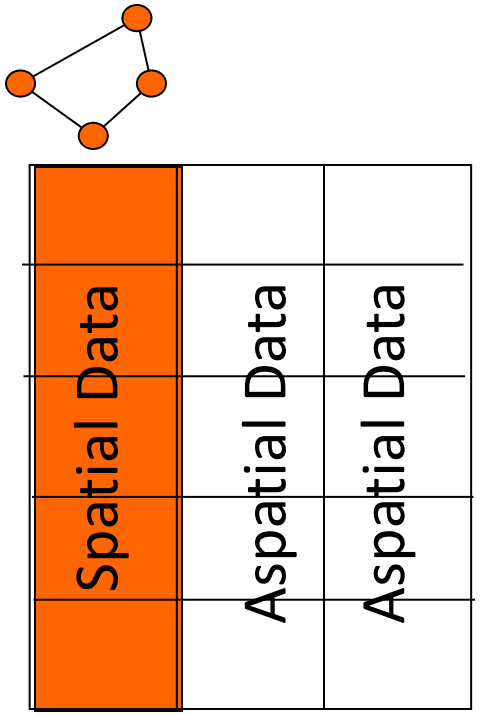
Nested Rings – “Island” or “Atoll” polygons

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).



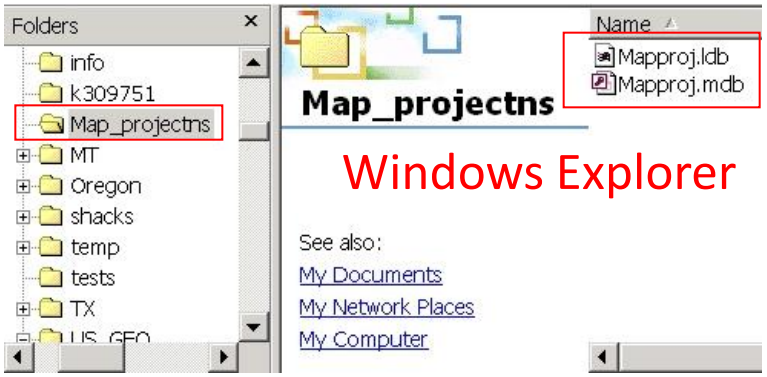
Slide courtesy of D. Maidment

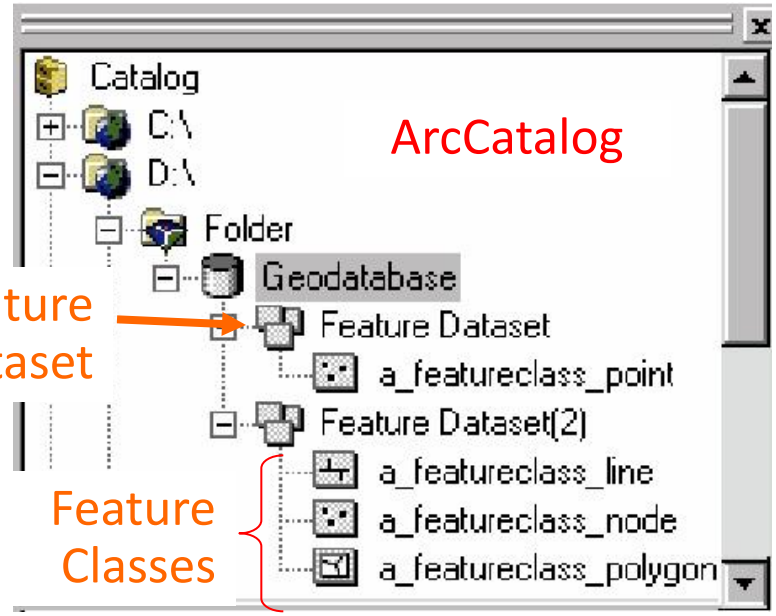
Geodatabase Model

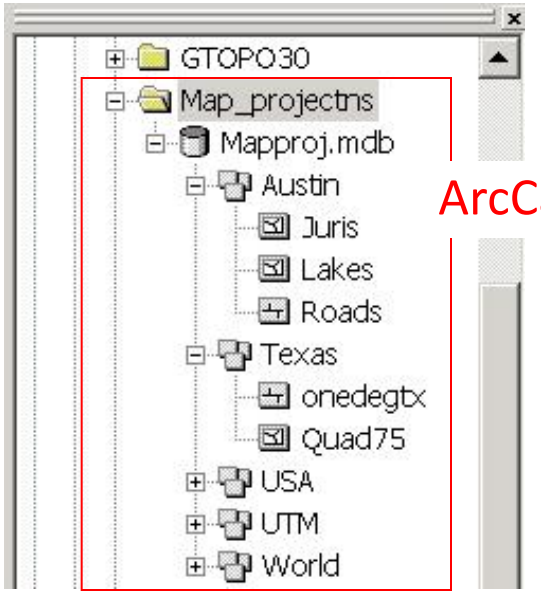
- Data structure capable of storing objects with behaviors and relationships, not merely graphical shapes with topology and attributes
- All spatial and attribute data for a feature are stored in a row of a single table
- A Geodatabase is a top-level container for feature classes, coverages, shapefiles,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

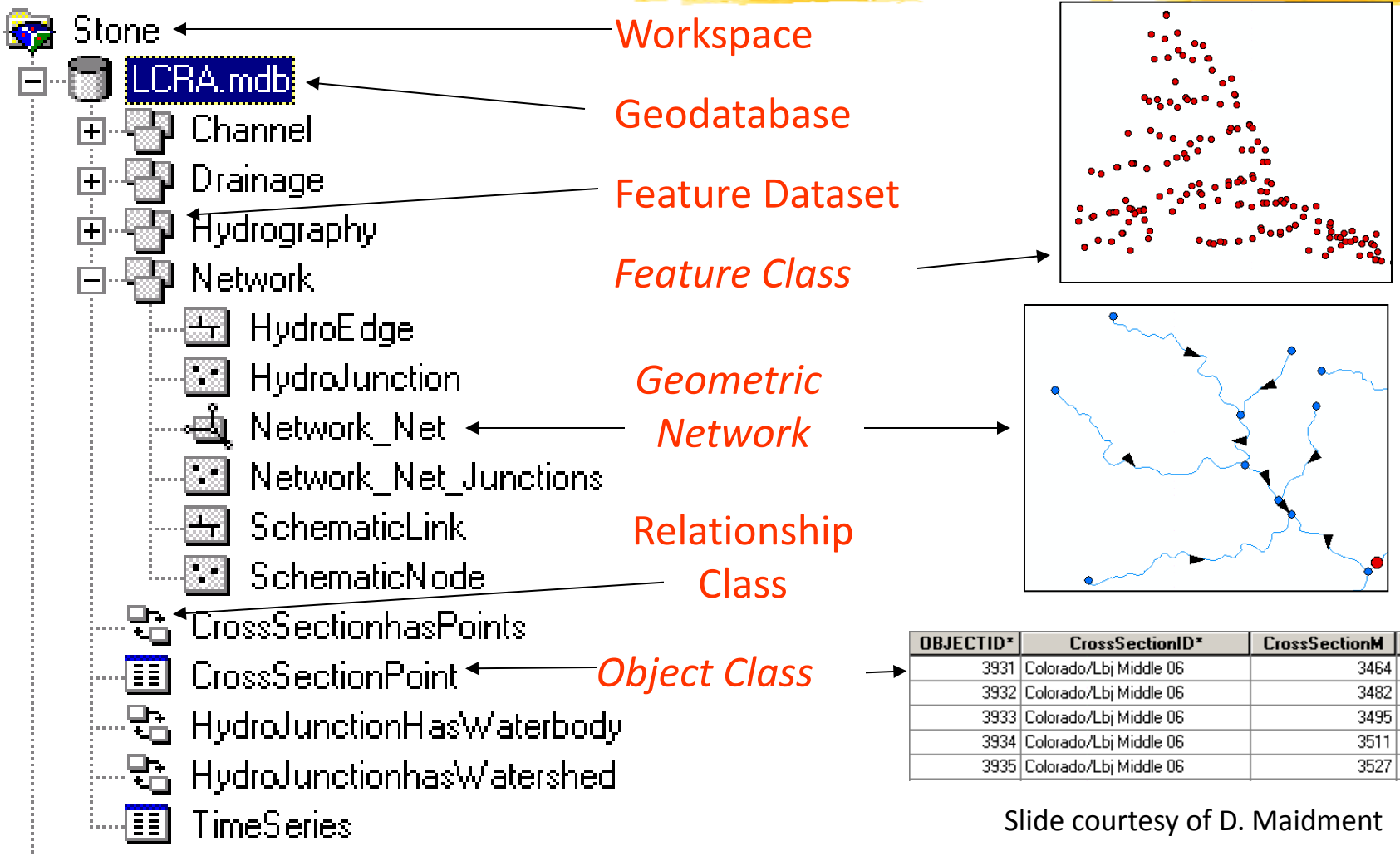
- Network Junctions (special Nodes)
- Network Edges

For geometric networks

Plus other classes

- Relationship classes
- Object Classes – tabular data without geography

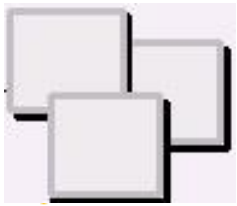
ArcGIS Geodatabase



OBJECTID*	CrossSectionID*	CrossSectionM	Elevation
3931	Colorado/Lbj Middle 06	3464	801
3932	Colorado/Lbj Middle 06	3482	800
3933	Colorado/Lbj Middle 06	3495	800
3934	Colorado/Lbj Middle 06	3511	799
3935	Colorado/Lbj Middle 06	3527	799

Slide courtesy of D. Maidment

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)

Number	Age_Ma	1_sigma	Rx_Type	Size_kg
123	142	1.5	B_schist	3.4
124	136	2.0	G_schist	1.3
125				

Object Class (nonspatial table)

Relationship

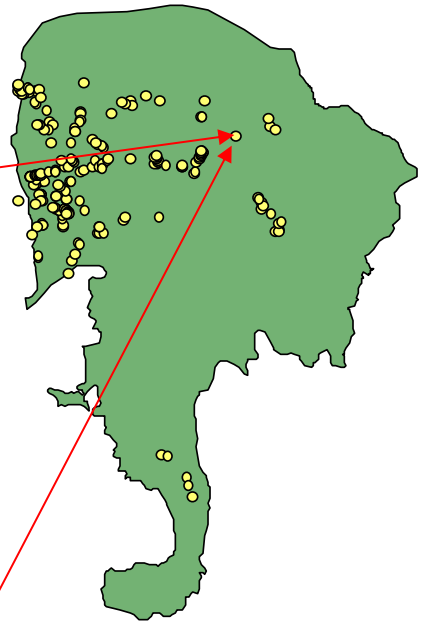
- A relationship is an association or link between two objects in a database.
- A relationship can exist between spatial objects (features in feature classes), non-spatial objects (objects in object classes), or between spatial and non-spatial objects.

Relationship class

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

Attributes of Condrey_samples_NAD27

FID	Shape*	Id	Number	NAD27_East	NAD27North
0	Point	0	230	490921	4636832
1	Point	0	123	491124	4637700
2	Point	0	124	491375	4638149
3	Point	0	125	491522	4638902
4	Point	0	126	491653	4638751
5	Point	0	127	492872	4639489
6	Point	0	128	492963	4639515



Feature Class (spatial table)

Number	Age_Ma	1_sigma	Rx_Type	Size_kg
123	142	1.5	B_schist	3.4
124	136	2.0	G_schist	1.3
125				

Object Class (nonspatial table)

Paper Map ➡ Files Of Coordinates

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

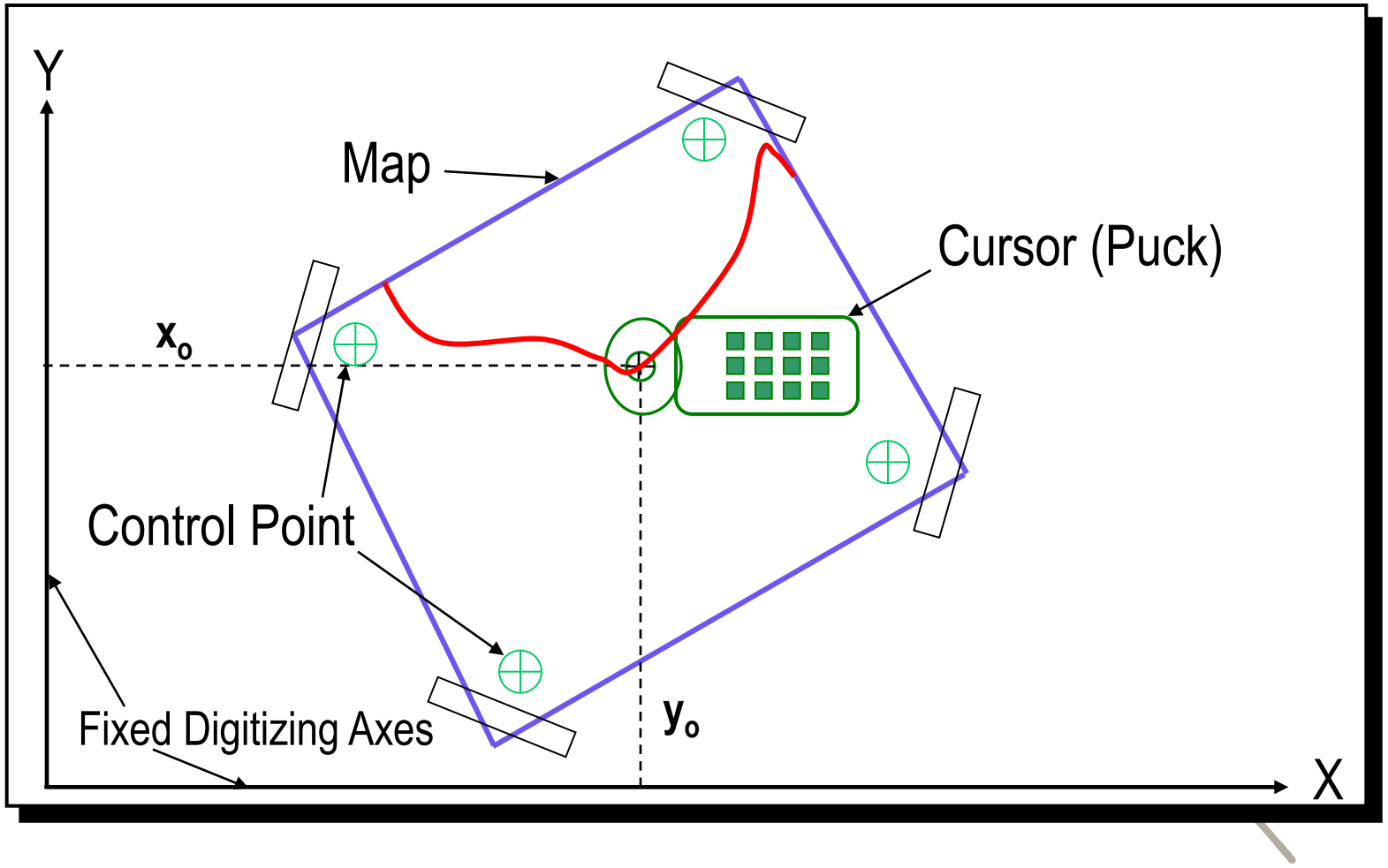
Digitizing is:

- 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 table once available in Rm. 6.202
- 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 table



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.)
- Separating features as point, line or polygon and tracing them to separate files (themes)

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?

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

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