Attribute Data Review

- Bits & Bytes
- Data types
  - Number (int, float, double, signed, unsigned…)
  - Text (string, character)
- Data structure (tables and fields)
  - Keys
  - Relating tables (Inner/outer join)
- GIS data models

Is Everything Best Represented as an Object? (Bian 2007)

- Environmental phenomena:
  - Objects (e.g., buildings)
  - Regions (e.g., campus)
  - Fields (e.g., urban residents)

- OO Paradigm
  - Encapsulation (identify, properties, behavior)
  - Composition (inheritance, aggregation, association)

- Identification of spatial objects
  - Scale (point, line, polygon, region, individual, mass)
  - Boundary (physical, precise vs. perceived, vague)
  - Attributes (homogeneous, discrete vs. heterogeneous, continuous)
  - Process (temporal change – mobile vs. sedentary)
  - Mobility (spatial change – solid vs. fluid)

Types of Environmental Phenomena
(Bian 2007)

Category	*Type*	*Example*
Object	Mobil individuals	Individual or groups of animals
Object	Sedentary individuals	Plants or bodies of water
Field	Masses of individuals	Vegetation
Region	Regions of individuals	Landscape patch
Field	Continuous solid mass	Land surface
Field	Continuous fluid mass	Water, air
Region	Sedentary regions in mass	Watershed
Region	Mobile regions in mass	Pollution plumes

GIS Data Models for Areal Fields
(Bian 2007)

Field Models	*GIS Model*	*Attribute Assoc*
Polygons	Vector	Piecewise
Contours	Vector	Sampled
TINs	Vector	Piecewise
Cell-grids	Raster	Sampled
Point-grids	Raster	Sampled
Irregular points	Vector	Sampled

Conceptual vs implementation OO models (Bian 2007)

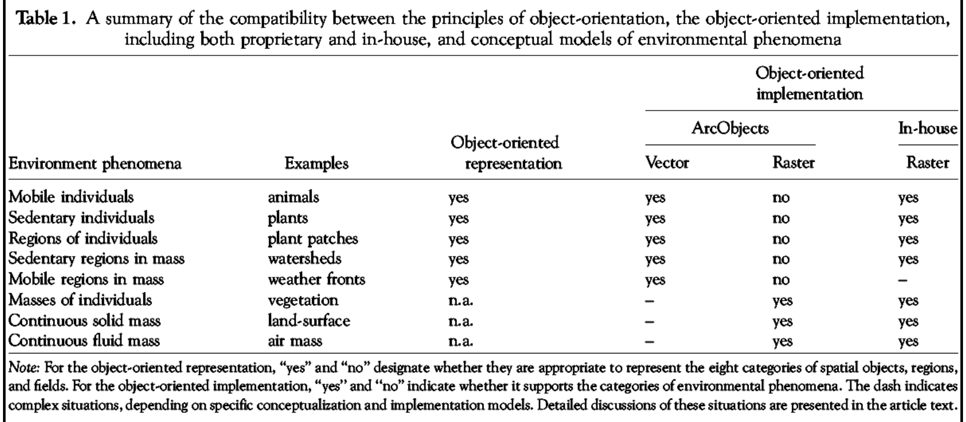


Table 1. A summary of the compatibility between the principles of object-orientation, the object-oriented implementation, including both proprietary and in-house, and conceptual models of environmental phenomena

			Object-oriented implementation		
			ArcObjects		In-house
Environment phenomena	Examples	Object-oriented representation	Vector	Raster	Raster
Mobile individuals	animals	yes	yes	no	yes
Sedentary individuals	plants	yes	yes	no	yes
Regions of individuals	plant patches	yes	yes	no	yes
Sedentary regions in mass	watersheds	yes	yes	no	yes
Mobile regions in mass	weather fronts	yes	yes	no	–
Masses of individuals	vegetation	n.a.	–	yes	yes
Continuous solid mass	land-surface	n.a.	–	yes	yes
Continuous fluid mass	air mass	n.a.	–	yes	yes

Note: For the object-oriented representation, "yes" and "no" designate whether they are appropriate to represent the eight categories of spatial objects, regions, and fields. For the object-oriented implementation, "yes" and "no" indicate whether it supports the categories of environmental phenomena. The dash indicates complex situations, depending on specific conceptualization and implementation models. Detailed discussions of these situations are presented in the article text.

Next generation GIS data model?

Vector Data Model

- Represent discrete geometric objects
  - Isolated objects and connected coverages
- Points, lines, & areas (nodes, chains, polygons)
  - Scale
- Topology (geometric rules)
  - Categorical coverage
  - Planar vs non-planar network

Vector Data Model (cont.)

- Composite geometric objects
  - TIN, regions, dynamic segmentation (routes)
- Object-based vector data model
- Non-topological vector data
  - CAD .dxf
  - Arcview shape file
  - Data conversion issues

Topology

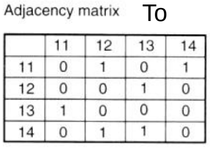
- Why topology
  - Enforce geometric rules for spatial representation and maintain data integrity (having implications in data interoperability)
  - Reduce data redundancy
  - Improve data access/update efficiency

Data Structures for Implementing Topology

- Coordinates (x, y)
- Digraph: adjacency and incidence matrices
- Line data model: arc-node list, arc-coordinate list
- Area data model: left/right list, polygon/arc list
- TIN: points, edges, & triangles list, adjacency matrix
- Region: region-arc and region-polygon lists
- Dynamic segmentation: section, routes, events tables
- OO data model

Adjacency & incidence matrices

Nodes: 11, 12, 13, 14
Arcs: 1, 2, 3, 4, 5, 6
Incidence: -1: end node, 1: start node
Adjacency: 1: Yes, 0: No

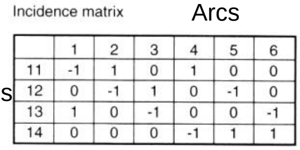


Adjacency matrix

To

From

	11	12	13	14
11	0	1	0	1
12	0	0	1	0
13	1	0	0	0
14	0	1	1	0



Incidence matrix

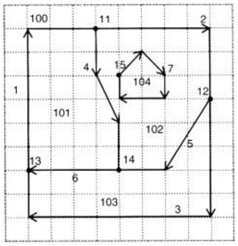
Arcs

Nodes

	1	2	3	4	5	6
11	-1	1	0	1	0	0
12	0	-1	1	0	-1	0
13	1	0	-1	0	0	-1
14	0	0	0	-1	1	1

ESRI’s Coverage Topology

- Connectivity: arcs - nodes
- Area definition: polygons – arcs
- Contiguity: arc directions, left/right polygon



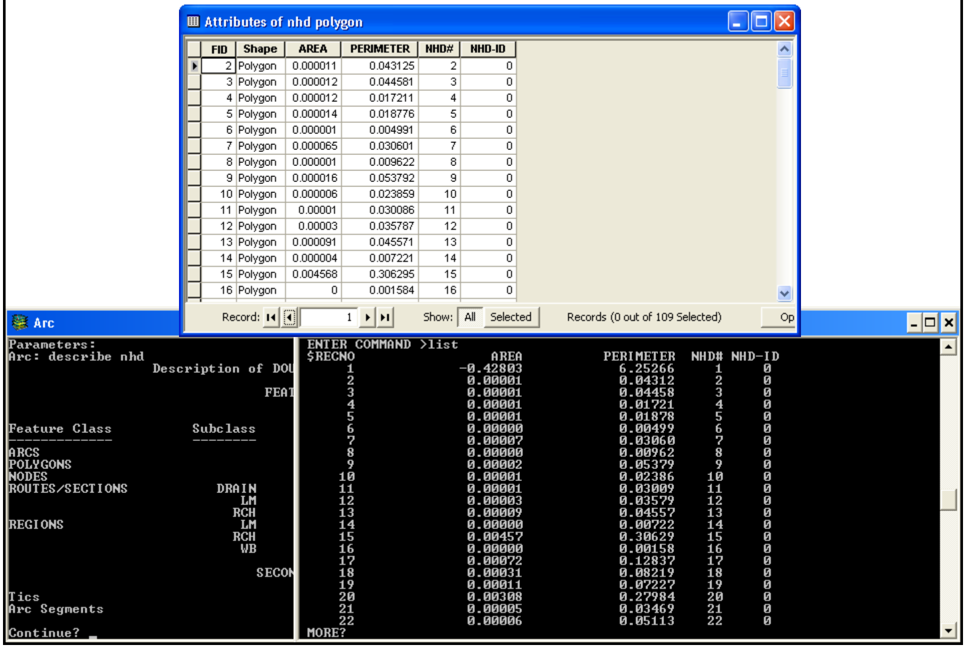
Left/right list

Arc#	L-poly	R-poly
1	100	101
2	100	102
3	100	103
4	102	101
5	103	102
6	103	101
7	102	104

Polygon/arc list

Polygon #	Arc#
101	1,4,6
102	4,2,5,0,7
103	6,5,3
104	7

Universal Polygon in a Polygon Coverage



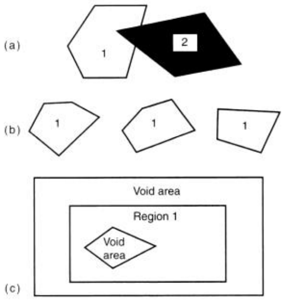
FID	Shape	AREA	PERIMETER	NHD#	NHD-ID
2	Polygon	0.000011	0.043125	2	0
3	Polygon	0.000012	0.044581	3	0
4	Polygon	0.000012	0.017211	4	0
5	Polygon	0.000014	0.018776	5	0
6	Polygon	0.000001	0.004991	6	0
7	Polygon	0.000065	0.030601	7	0
8	Polygon	0.000001	0.009622	8	0
9	Polygon	0.000016	0.053792	9	0
10	Polygon	0.000006	0.023859	10	0
11	Polygon	0.00001	0.030086	11	0
12	Polygon	0.00003	0.035787	12	0
13	Polygon	0.000091	0.045571	13	0
14	Polygon	0.000004	0.007221	14	0
15	Polygon	0.004568	0.306295	15	0
16	Polygon	0	0.001584	16	0

Record: 1 Show: All Selected Records (0 out of 109 Selected) Op

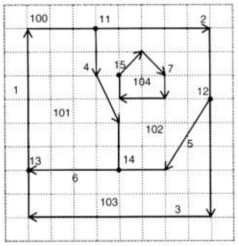
```
Parameters:
Arc: describe nhd
                    Description of DOU

                                  FEAT


Feature Class            Subclass
-------------            --------
ARCS
POLYGONS
NODES
ROUTES/SECTIONS             DRAIN
                               LM
                              RCH
REGIONS                        LM
                              RCH
                               WB

                                SECON

Tics
Arc Segments

Continue? _
```

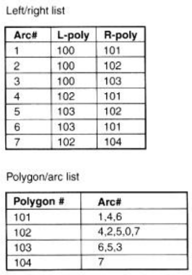
```
ENTER COMMAND >list
$RECNO                    AREA              PERIMETER  NHD# NHD-ID
       1              -0.42803                6.25266     1      0
       2               0.00001                0.04312     2      0
       3               0.00001                0.04458     3      0
       4               0.00001                0.01721     4      0
       5               0.00001                0.01878     5      0
       6               0.00000                0.00499     6      0
       7               0.00007                0.03060     7      0
       8               0.00000                0.00962     8      0
       9               0.00002                0.05379     9      0
      10               0.00001                0.02386    10      0
      11               0.00001                0.03009    11      0
      12               0.00003                0.03579    12      0
      13               0.00009                0.04557    13      0
      14               0.00000                0.00722    14      0
      15               0.00457                0.30629    15      0
      16               0.00000                0.00158    16      0
      17               0.00072                0.12837    17      0
      18               0.00031                0.08219    18      0
      19               0.00011                0.07227    19      0
      20               0.00308                0.27984    20      0
      21               0.00005                0.03469    21      0
      22               0.00006                0.05113    22      0
MORE?
```

Regions



Polygon Topology





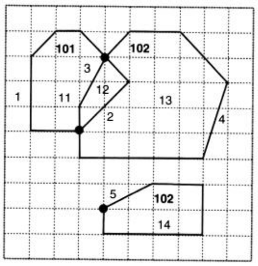
Left/right list

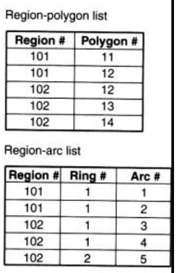
Arc#	L-poly	R-poly
1	100	101
2	100	102
3	100	103
4	102	101
5	103	102
6	103	101
7	102	104

Polygon/arc list

Polygon #	Arc#
101	1,4,6
102	4,2,5,0,7
103	6,5,3
104	7

Region Topology





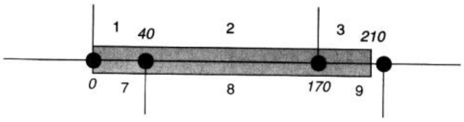
Region-polygon list

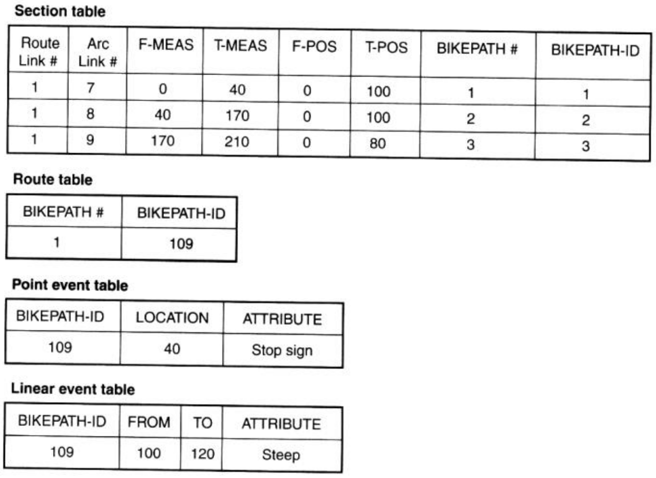
Region #	Polygon #
101	11
101	12
102	12
102	13
102	14

Region-arc list

Region #	Ring #	Arc #
101	1	1
101	1	2
102	1	3
102	1	4
102	2	5

Dynamic segmentation (To be covered in Week 8)





Section table

Route Link #	Arc Link #	F-MEAS	T-MEAS	F-POS	T-POS	BIKEPATH #	BIKEPATH-ID
1	7	0	40	0	100	1	1
1	8	40	170	0	100	2	2
1	9	170	210	0	80	3	3

Route table

BIKEPATH #	BIKEPATH-ID
1	109

Point event table

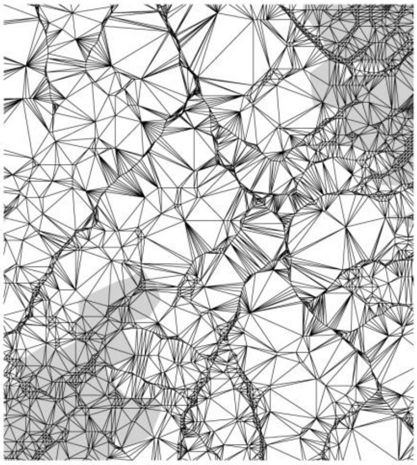
BIKEPATH-ID	LOCATION	ATTRIBUTE
109	40	Stop sign

Linear event table

BIKEPATH-ID	FROM	TO	ATTRIBUTE
109	100	120	Steep

Triangulated Irregular Network (TIN)

- Point (x, y, z)
- Edge
- Triangles
- Topology
  - –The triangle number
  - –The numbers of each adjacent triangle
  - –The three nodes defining the triangle
  - –The x, y coordinates of each node
  - –The surface z value of each node
  - –The edge type of each triangle edge (hard or soft)



Generating Topology

- Ways of generating topology in vector data model
  - – Automated methods
    - Delaunay triangulation, Arcinfo BUILD, Arcinfo REGIONQUERY
  - – Manual methods

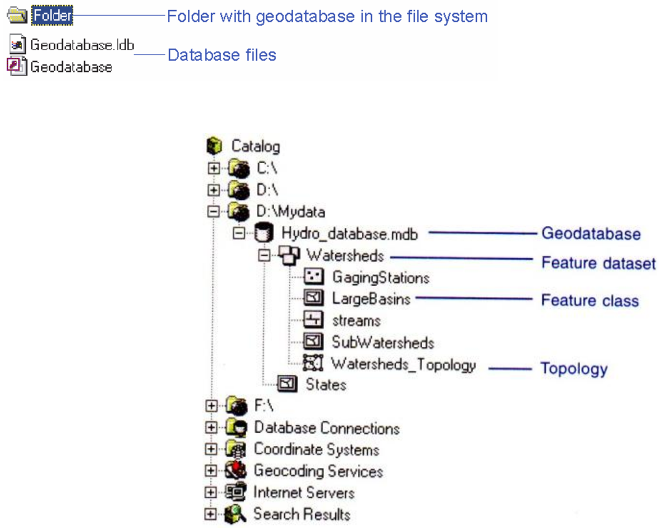
Object-Based Vector Data Model

- Georelational vector data model
- Object-oriented model
  - – Class and instance
  - – Properties and methods
  - – Interface:
    - Inheritance, encapsulation, polymorphism

ArcGIS Geodatabase Structure

- Spatial features: point, polyline, polygon
- Feature class
- Feature dataset
- Validation rules
  - – Attribute domain
  - – Relationship
  - – Connectivity
- User specified topology

ArcGIS Geodatabase Data Structure



Folder — Folder with geodatabase in the file system

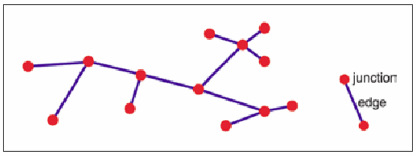
Geodatabase.ldb
Geodatabase — Database files

Catalog
C:\
D:\
D:\Mydata
Hydro_database.mdb — Geodatabase
Watersheds — Feature dataset
GagingStations
LargeBasins — Feature class
streams
SubWatersheds
Watersheds_Topology — Topology
States
F:\
Database Connections
Coordinate Systems
Geocoding Services
Internet Servers
Search Results

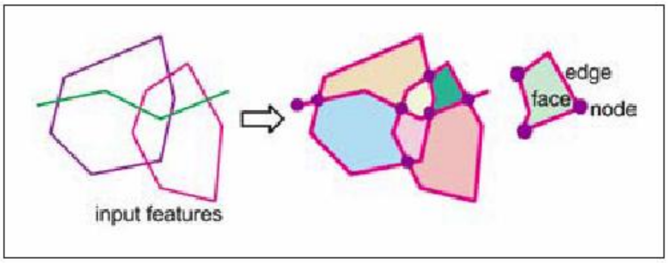
Features / Feature Classes

- All point, line, and polygon features can:
  - Be multipart
  - Have x,y; x,y,z; or x,y,z,m coordinates
    - (m-coordinates store distance measurement values, a line with m-coordinates becomes a route)
  - Be stored as continuous layers instead of tiled

- Network
  - Junctions, edges



Geodatabase Topology



- Details will be covered in Week 5.

ESRI's GIS Data Structures

Data Structure	***Type***	***Topology***	***Portability***	***Spatial Integrity***
Coverage	File-based	Required	Low	High
Shapefile	File-based	None	High	None
Personal Geodatabase	DBMS	Optional	High	High (if topology rules are defined)
File Geodatabase	File-based	Optional	High	High (if topology rules are defined)

Spatial Reference

- Prior to ArcGIS 9.2, spatial references were low precision. Each **integer coordinate** was allotted 31 bits rather than the 53 bits provided by high resolution spatial references created and maintained with 9.2 or above.

- With low precision, you have to specify a domain extent as well as a resolution (precision) value. A large domain extent is only possible with low resolution values.

- Therefore, when working with low-precision spatial references, you must carefully balance the trade-off between domain extent and the resolution or precision values.

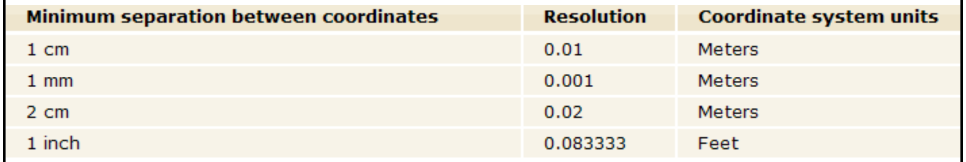
About x,y Resolution and Domain

- A low-precision geodatabase stores coordinates as positive 4-byte integers that have a maximum value of 2,147,483,647.

- If you need to store meter precision, you have 2.14 billion meters to work with (approximately 53 times the circumference of the earth).

- If you need to store centimeters precision, in which case you would have 2.14 billion centimeters to work with (about one-half the circumference of the earth).

- Resolution values represent the minimum allowed separation between two coordinate values. Resolution values are used to convert decimal values to the integers stored in the geodatabase.



Minimum separation between coordinates	Resolution	Coordinate system units
1 cm	0.01	Meters
1 mm	0.001	Meters
2 cm	0.02	Meters
1 inch	0.083333	Feet

Advantages of ESRI Geodatabase

- Functionality (ArcObjects)
  - Attribute domain
  - Connectivity rules/geometric networks
  - Relationship class
  - Topology rules
- Web-based, versioned operations (ArcSDE)
- Portability (Personal Geodatabase)
- Integration with RDBMS
- Customized data models

Humans still do mapping

July 9, 2004

Navigation systems need them to plot streets

BY CHRISTOPHER JENSEN
Newhouse News Service

CLEVELAND – How the rewards of mapping have changed. Guys like Magellan got glory; modern-day mappers get fast food and a night at some motel.

Kind of unfair. After all, trundling around a new housing division near Cleveland, John Rhinerson and Brian Nash are doing their part to help every motorist who ever got lost or fretted about how to get from here to there.

Nash and Rhinerson work for Navteq Corp., a Chicago-based company that provides the geographic data for more than 90 percent of the navigation systems sold by automakers in the United States.

A variety of manufacturers, from Alpine to VDO, make in-car navigation systems. Usually offered as options, they allow a motorist to type in a destination and get turn-by-turn instructions. Many also offer information on everything from gas stations and restaurants to automobile dealerships.

Such assistance is possible thanks to a powerful computer, geographic data from Navteq and a global positioning system. The GPS takes information from several satellites to tell a vehicle's location and how quickly it is traveling. Make a wrong turn, and the navigation system will gently request a U-turn ("when safe") or recalculate how to get you to your destination.

How well and how easily each navigation system works varies greatly depending on the company that created the system, according to studies by J.D. Power and Associates, the market research firm. Some systems frustrate and others delight, depending on design factors such as how easy the system is to operate.

But none of them would be possible without detailed information on what is where, which is where Nash, who is based in Columbus, and Rhinerson, who is based in Youngstown, do their part.

JOSHUA GUNTER, NEWHOUSE NEWS SERVICE

John Rhinerson, right, gives directions to driver Brian Nash as they make their way through a neighborhood to plot new roads for vehicle navigation systems.

They are among the 250 Navteq field analysts in the United States and Canada who drive the nation's roads, big and little, collecting such information, said Kelly Smith, the vice president of marketing and communications for Navteq.

On a June afternoon, Rhinerson and Nash are in the full Navteq hunter-gatherer mode. Nash is driving a white Ford Taurus with a GPS antenna on the roof. Rhinerson has a huge laptop computer on his lap hooked into the GPS, which is taking signals from the satellites and recording the route they are driving. They are charting a new housing complex. As they drive, the GPS records a line showing their route on the laptop. Straights. Turns. Cul-de-sacs.

Nash is the driver and spotter. It is his job to note details such as street names ("Lakeview ... one word") and the house numbers at the beginning and end of each street. Nash seems quite fond of people who use large signs for their house numbers or, better yet, paint them on a boulder.

"Those are Navteq-friendly," he said.

Rhinerson has an electronic pad that allows him to write such information on the laptop's screen.

Some mapping chores are more challenging than others. Rural routes and subdivisions can have a more leisurely pace. Working in a city, dealing with traffic and trying to collect all the proper information is more frantic, but Rhinerson and Nash say they enjoy the hectic, go-go-go pace.

The information they gather goes into the Navteq database, which is then sold to companies that make navigation systems. There are different levels of detail. If you want tiny streets, it costs more.

More detail may also make a navigation system operate more slowly, taking longer to calculate a new route, which can upset consumers. That explains why some navigation systems have more detailed information than others.