

Spatial Databases as Models of Reality



presented by:

Tim Haithcoat

**University of Missouri
Columbia**

Compiled from:

Timothy L. Nyerges

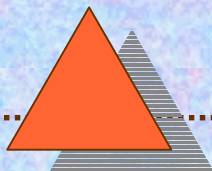
University of Washington





Introduction

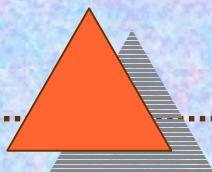
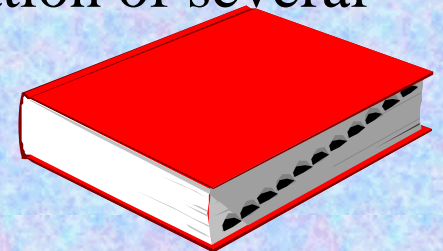
- ◆ The real world is too complex for our immediate and direct understanding
- ◆ We create “models” of reality that are intended to have some similarity with selected aspects of the real world
- ◆ Databases are created from these “models” as a fundamental step in coming to know the nature and status of that reality





Definition

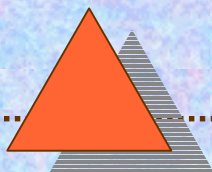
- ◆ **Spatial Database:** a collection of spatially referenced data that acts as a model of reality
 - A database is a model of reality in the sense that the database represents a selected set or approximation of phenomena
 - These selected phenomena are deemed important enough to represent in digital form
 - The digital representation might be for some past, present or future time period (or contain some combination of several time periods in an organized fashion)





Standards

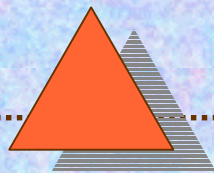
- ◆ Many of the definitions in this presentation were standardized by the US National Digital Cartographic Standard (DCDSTF, 1998)
 - These and many other standards have been developed to provide a nationally uniform means for portraying and exchanging digital cartographic data
 - These cartographic standards will form part of a larger standard being developed for the digital representation of all earth science information





Database Content & an Organization's Mission ORGANIZATION MANDATES

- ◆ Organizations have mandates to perform certain tasks that carry out their missions
 - Mandates are the reasons they exist as organizations
- ◆ Organizations have different needs for data depending on their mandates and the activities required to carry out these mandates
 - Mandates often help identify & define entities of interest, requiring a certain view of the world
 - What might seem at first glance to be the same data need in two different organizations can actually be quite different when looked at in more detail





Database Content & an Organization's Mission

DATABASE CONTENTS

Example: Transportation

- ◆ Highway data from different points of view:

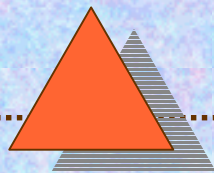
A natural resource organization might only need logging roads and the connecting access to state highways.



A transportation organization's main interest is in identifying highways used by the public. (The database might also be used to store detailed highway condition and maintenance information.)



- ◆ We would expect transportation's need for highway data to be more detailed than would the natural resource organization's





Database Content & an Organization's Mission

DATABASE CONTENTS

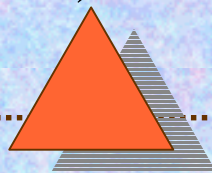
Example: Wetlands

- ◆ Wetland data from the different points of view:

An ecological organization might define wetlands as a natural resource to be preserved and restricted from developed. (That perspective might require considerable detail for describing the area's biology & physical resources.)



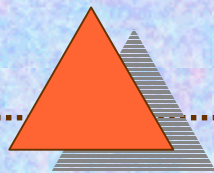
A taxing authority might define a wetland to be a “wasteland” and of very little value to society. (That description might require only the boundary of the wetland in the database.)





Database Content & an Organization's Mission DATABASE DESIGN

- ◆ In each organization only certain phenomena are important enough to collect and represent in a database
 - The data collection process involves a sampling of geographic reality to determine the status of that reality (whether past, present or future)
- ◆ Identifying the phenomena and then choosing an appropriate data representation for them is part of a process called database design.





Fundamental Database Elements

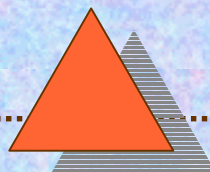
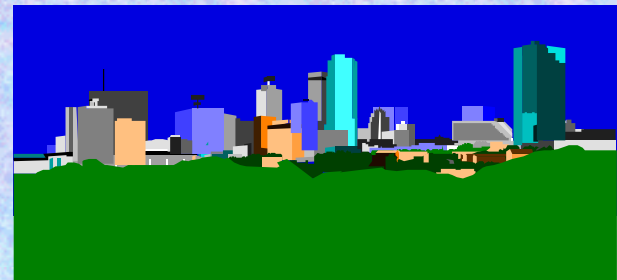
- ◆ Elements of reality modeled in a GIS database have two identities:
 - **Entity:** the element in reality
 - **Object:** the element as it is represented in the database
- ◆ A third identity that is important in cartographic applications is the **symbol** that is used to depict the object/entity as a feature on a map or other graphic display



- These definitions are based on those defined by the DCDSTF, 1998

Entity

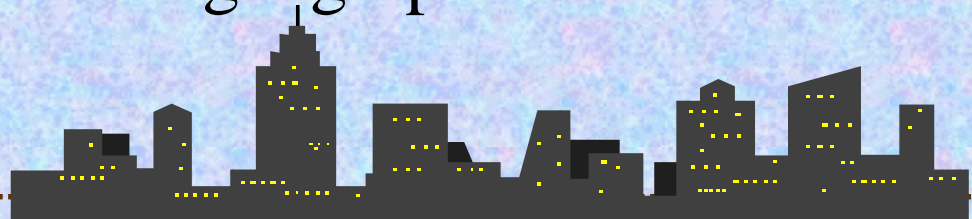
- ◆ A phenomenon of interest in reality that is not further subdivided into phenomena of the same kind
 - Example: a city could be considered an entity and subdivided into component parts, but these parts would NOT be called cities, they would be districts, neighborhoods or the like
 - Example: a forest could be subdivided into smaller forest compartments






Object

- ◆ A digital representation of all or part of an entity
- ◆ The method of digital representation of a phenomenon varies according to scale, purpose and other factors
 - Example: a city could be represented geographically as a point if the area under consideration were continental in scale
 - the same city could be geographically represented as an area if dealing with a geographic database for a state or a county

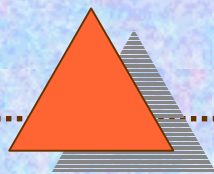




**Basic
0-Dimensional
Object Types**

● **Point**

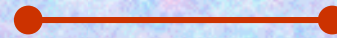
* **Node**



Basic 1-Dimensional Object Types



Line



Line Segment



String



Arc



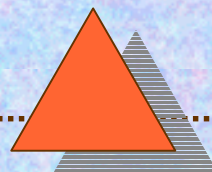
Link



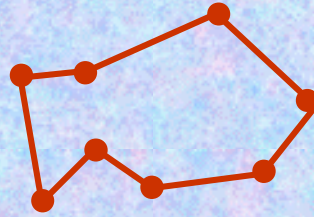
Direct link



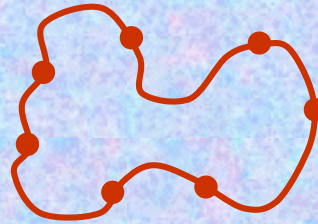
Chain



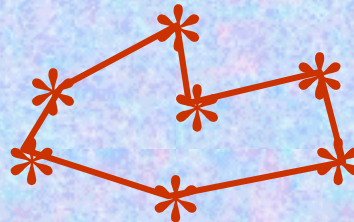
Basic 1-Dimensional Object Types (Continued)



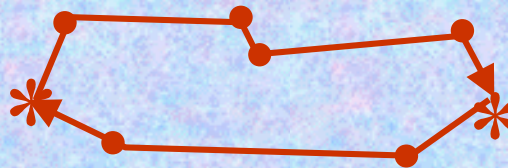
Ring created from
string(s)



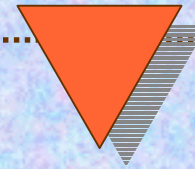
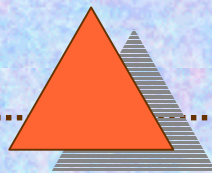
Ring created from
arc(s)



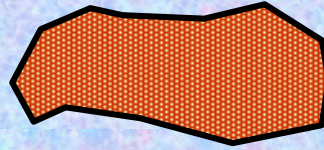
Ring created from
link(s)



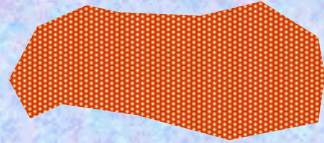
Ring created from
Chain(s)



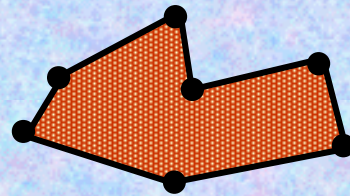
Basic 2-Dimensional Spatial Object Types



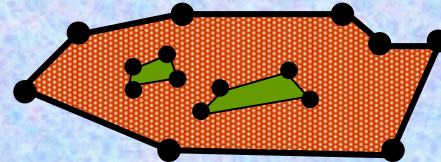
Area



Interior area



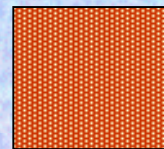
Simple Polygon



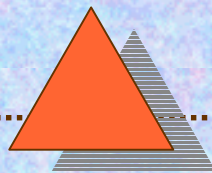
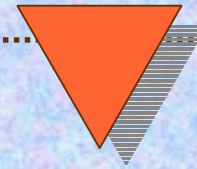
Complex Polygon



Pixel



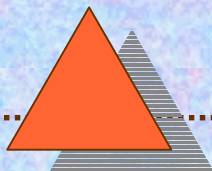
Grid Cell





Attributes

- ◆ An attribute is a characteristic of an entity selected for representation
- ◆ Usually non-spatial
 - Though some may be related to the spatial character of the phenomena under study
 - ▼ Examples: area, perimeter



Attribute Value



- ◆ The actual value of the attribute that has been measured (sampled) and stored in the database
- ◆ An entity type is almost always labeled and known by attributes
 - Example: a road usually has a name and is identified according to its class (i.e., alley, freeway)
- ◆ Attributes values often are conceptually organized in attribute tables which list individual entities in the rows and attributes in the column
 - Entries in each cell of the table represent the attribute value of a specific attribute for a specific entity

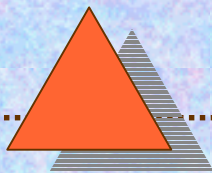


- Note: attribute table is not an official DCDSTF term



Database Model

- ◆ Is a conceptual description of a database defining entity type and associated attributes
 - Each entity type is represented by specific spatial objects
- ◆ Examples of database models can be grouped by application area
 - Example: transportation applications require different database models than to natural resource applications





Database Model ~ Continued

- ◆ After the database is constructed, the database mode is a view of the database which the system can present to the user
 - Other views can be presented, but this one is likely useful because it was important in the conceptual design
 - ▼ Example: the system can model the data in vector form but generate a raster for purpose of display to the user
 - Need not be related directly to the way the data are actually stored in the database
 - ▼ Example: census zones may be defined as being represented by polygons, but the program may actually represent the polygon as a series of line segments



Layers

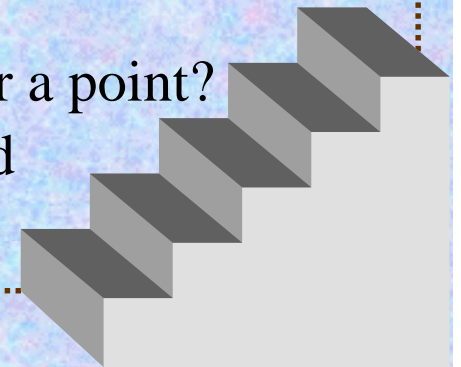
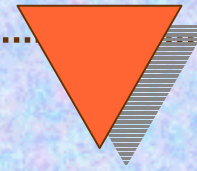
- ◆ Spatial objects can be grouped into layers, also called overlays, coverages or themes
- ◆ One layer may represent a single entity type or a group of conceptually related entity types
 - Example: a layer may have only stream segments or may have streams, lakes, coastline and swamps
 - Options depend on the system as well as the database model
 - Some spatial databases have been built by combining all entities into one layer



Steps in Database Design

◆ Conceptual

- Software and hardware independent
- Describes and defines included entities
- Identifies how entities will be represented in the database
 - ▼ i.e., selection of spatial objects - points, lines, areas, raster cells
- Requires decisions about how real-world dimensionality and relationships will be represented
 - ▼ These can be based on the processing that will be done on these objects
 - ▼ i.e., should a building be represented as an area or a point?
 - ▼ i.e., should highway segments be explicitly linked in the database?





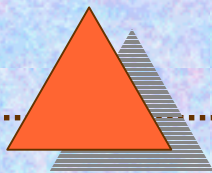
Steps in Database Design ~ Continued

◆ Logical

- Software specific but hardware independent
- Sets out the logical structure of the database elements, determined by the database management system used by the software

◆ Physical

- Both hardware and software specific
- Requires consideration of how files will be structured for access from the disk

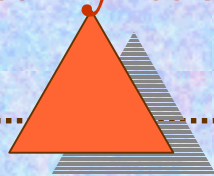




Desirable Database Characteristics

Database should be:

- ◆ **Contemporaneous:** should contain information of the same vintage for all its measured variables
- ◆ **As detailed as necessary** for the intended applications
 - The categories of information & subcategories within them should contain all of the data needed to analyze or model the behavior of the resource using conventional methods & models
- ◆ **Positionally accurate**

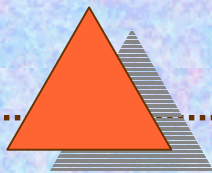




Desirable Database Characteristics ~ Continued

Database should be:

- ◆ **Exactly compatible** with other information that may be overlain with it
- ◆ **Internally accurate**, portraying the nature of phenomena without error: requires clear definitions of phenomena that are included
- ◆ **Readily updated** on a regular schedule
- ◆ **Accessible** to whoever needs it



Issues in Database Design

- ◆ Almost all entities of geographic reality have at least 3-dimensional spatial character, but not all dimensions may be needed
 - *Example:* highway pavement has a depth which might be important, but is not as important as the width, which is not as important as the length
- ◆ Representation should be based on types of manipulations that might be undertaken
- ◆ Map-scale of the source document is important in constraining the level of detail represented in a database
 - *Example:* on a 1:100,000 map individual houses or fields are not visible

