What is Where?

Getting Started With Geographic Information Systems

Chapter 5

Dursun Z. ªeker

Contents

- Basic Database Management
- Searches by Attribute
- Searches by Geography
- The Query Interface

Basic Database Management

A GIS can answer the two questions "*what*" and "*where*?" It also can answer the question "*what is where?*". The *where* components relates to the map behind all GIS activities. The *what* relates to the features, their size, geographical properties and attributes.

***Data retrieval* is the ability to gain access to a record and its attributes on demand in data organisation. The attribute and the map data have different means of access. At the most simplest level, the GIS is a computer program that accessed data stored in files.**

A GIS can answer the question: What is where?

WHAT: Characteristics of attributes or features.

WHERE: In geographic space.

Flat File Database

	Attribute	Attribute	Attribute
Record	Value	Value	Value
Record	Value	Value	Value
Record	Value	Value	Value

Arc/node map data structure with files

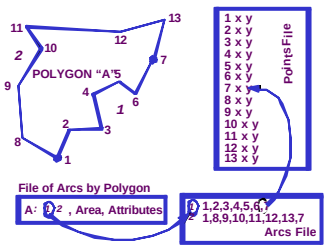


Figure 3.4 Arc/Node Map Data Structure with Files.

A GIS links attribute and spatial data

- Attribute Data
  - Flat File
  - Relations
- Map Data
  - Point File
  - Line File
  - Area File
  - Topology
  - Theme

Basic Database Management

At the logical level, access requires a *data model*, a theoretical construct that becomes the key for unlocking the the data's door. *A data model is a logical construct for the storage and retrieval of information.* It is computer's way of memorizing all the GIS that we need to use. GIS, must have at least two data models, and that two must have a bridge or link between them to tie the attributes and geography together. These are the map data model and attribute data model.

In the previous section, storage of map data has been considered, in this section storage of attributes will be mentioned.

What is a Data Model?

A logical construct for the storage and retrieval of information.

GIS map data structures are map data models.

Attribute data models are needed for the DBMS.

The origin of DBMS data models is in computer science.

- Data structure deal primarily with how the data are physically stored in files on the computer system. (digitally encoding, raster-vector)
- Data model is a logical means of organisation of data for use in an information system.

map data model

- **reality**
- **digital landscape model**
- **digital cartographic model**
- **map**
- **mental map**

attribute data model

- **DBMS**
- **data definition language is part of the DBMS:** types, lengths or numerical range of each attribute
- **data dictionary is a catalogue of all of the attributes**
- **every DBMS has the ability to examine the data dictionary.**
- **The data dictionary is a critical piece of metadata (data about data)**

Basic Database Management

The earliest database manegement systems (DBMS), date from the efforts of the early 1970s when large mainframe computers were used. This technology was called as automatic data processing. Database manegement went through its own revolution due to technological trends. Database manegement is effected from the hardware and interactive and graphical user interfaces. Whis these factors intellectual breakthroughs led to radical changes in the way that attribute data can be stored in files. Another revolution the object-oriented data base system will be mentioned later.

Computers and Databases

Databases are packages design to create, edit, manipulate and analyze data. To be suitable for a database, the data must consist of records, which provide information on individual cases, people, places, features etc. Each record may contain several fields each of which contains one item of information. There can be several classes of records in a database. Database creation involves several stages.

- Input of the spatial data,
- Input of the attribute data,
- Linking spatial and attribute data.

Computers and Databases

Spatial data is entered via digitised points and lines, scanned and vectorised lines or directly from other digital sources. For example, an airline reservation database might have the following classes of records and associated items:

@ ***Passengers:*** **name, phone, flight numbers.**

@ ***Aircraft:*** **type, registration number, number of seats.**

@ ***Crew:*** **names of pilot, copilot, cabin crew, home city.**

@ ***Flight:*** **number, departure and arrival times, aircraft.**

A DBMS contains:

– **data definition**
– **data dictionary**
– **data -entry module**
– **data update module**
– **report generator**
– **query language**

Data Base Management functions

- data entry
- verification
- modification (updating)
- query language
- Search

DBMS can be used to perform its more **advanced functions**

Sorting .
Reordering .
Subsetting .
Searching .

Some of the database functions

- **Creating and editing records, using customised screen**
- **Printing reports, using customised report forms, including subtotals and totals**
- **Selecting records based on user specified rules**
- **Updating records based on new information**
- **Linking records**

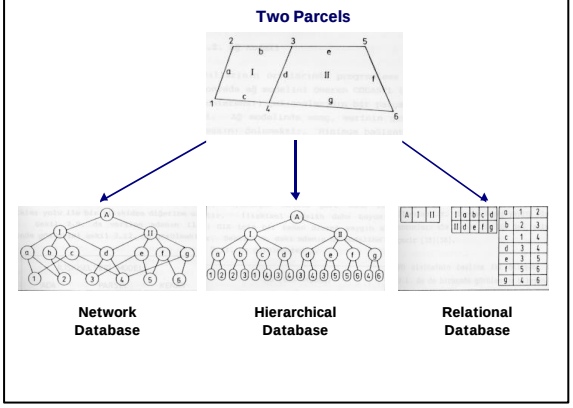
Types of Database

There are three (+1) types of database. These are:

- ✓ **Network**
- ✓ **Hierarchical**
- ✓ **Relational**

+ **Object - oriented**

They have different ways of modeling data within a database. Although all three are used the relational model has been most successful within GIS. Some of the well-known RDBMS are dBase, Oracle and Info. Many database use the same language, SQL (Standard Query Language) for formulating queries.



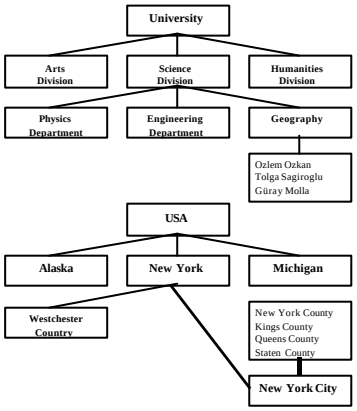
DBMS: hierarchical structure

-University
-Science Division
-Faculty
-Department
-Division

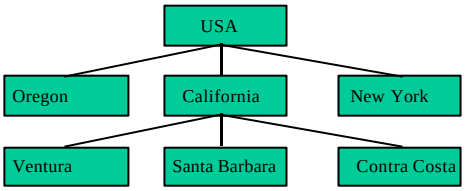
-state
-county ?
-city ?
-district

- multiple membership: (for a single house)
- a fire district
- a police district
- a school district
- a voting district
- a census district

Hierarchical structures for information. The university example is a simple hierarchy. The nation/state/county/city example is an "inverted" hierarchy. The hierarchical file organization or data model would not work in this case.



Historically, databases were structured hierarchically in files...



GIS and DBMS

Ability of the DBMS or GIS to get back on demand data that were previously stored.

Geographic search is the secret to GIS data retrieval.

Many forms of data organization are incapable of geographic search.

GI systems have embedded DBMSs, or link to a commercial DBMS.

Examples: Dbase, ORACLE, Excel, Paradox

Searches by Geography

When the attribute search is considered, *search and retrieval commands* such as show attribute, show records, generate a report, find, browse, sort, recode, restrict and compute mentioned.

In the map database records are features. There are some special attributes specific to the special data, and those relate to the coordinates and their measures, plus the charateristics of the lines and polygons.

Searches by Geography

- show all records in a spatial sense becomes either show all attributes or displays all features on a map.
- many GIS packages have *identify* or *locate* operations
- after selection successfully can then retrieve that feature's attributes from the nonmap database

DBMS: relational structure

- consist of several flat files
- each of them can contain different attributes associated with a record
- every record a unique identifier
- "key" attribute can then serve as a link between the flat files
- each linked files, have data entered, be edited and updated and searched separately and without affecting the others.

DBMS: relational structure

- Actions by key and links :

  relate, to select from two flat files that have a common key attribute,

  join, to take the relate output and merge them into a single database.

Most current GIS DBM is by relational databases.

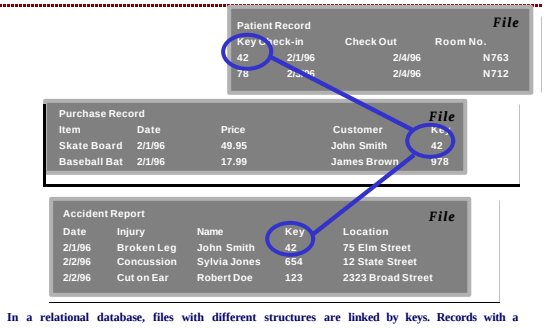
Based on multiple flat files for records

Dissimilar attribute structures

Connected by a common key attribute.

Key is a UNIQUE identifier at the "atomic" level for every record.

Relational Data Bases



In a relational database, files with different structures are linked by keys. Records with a common key relate to one feature within the GIS.

Retrieval Operations

- **Searches by attribute: find and browse.**
- **Data reorganization: select, renumber, and sort.**
- **Compute allows the creation of new attributes based on calculated values.**

Searches by Attibute

- All DBMSs include functions for basic data display, that is, show all attributes in a database, show all records with their attributes, and show all existing databases.
- The DBMS must support functions that fall into the category of query. A DBMS, should allow sufficient data query that any record can be isolated and any subset required for mapping found easily.

DBMS queries via the query language

sort

renumber

subset

search

Searches by Attibute

- Find : by search or by browse
- Sorting : alphabetically or numerically
- Restrict operation : selected attributes according to the their limits
- Compute operation : mathematical operations between attributes
- Renumber or recode

Command line attribute query

```
find in states where state_name = 'California'
<1 record in result>

use states

calculate in states population_density =
population / area
<50 records in result>

restrict in states where population_density >
1000
<20 records selected in result>
```


The Retrieval User Interface

GIS query is usually by command line, batch, menu or macro.

Most GIS packages use the GUI of the computer's operating system to support both a menu-type query interface and a macro or programming language.

SQL is a standard interface to relational databases and is supported by many GISs.

THE QUERY INTERFACE

Both database manegement and geographic information manegement share the fact that the user must somehow interact with the data in an appropriate way.

Macros, are files containing commands to be executed one at a time. If an error is detected in a macro, the execution can be stopped and the file modified to correct the mistake. In the typical form, a command consist of a keyword for the operation such as *IMPORT*, *OVERLAY*, *SELECT*, and so on, and a set of optional or required parameters.

THE QUERY INTERFACE

- command line, batch, or macro
- most GIS packages now are fully integrated with the WIMP (windows, icons, menus, and pointers) interface specified by the operating system (Windows or X-Windows), choises are now most commonly made by menu.
- most GIS packages use the GUI of the computer's operating system to support both a menu-type query interface and a macro or programming language (Arc View's Avenue, MapInfo's MapBasic, and Arc/Info's AML).

THE QUERY INTERFACE

The Structured Query Language (SQL), has accepted as a much used tool in regular database manegement. It is a standard interface to relational databases and is supported by many GISs.

Spatial Retrieval Operations

Attribute queries are not very useful for geographic search.

In a map database the records are features or themes

The spatial equivalent of a find is locate, the GIS highlights the result.

Spatial equivalents of the DBMS queries result in locating sets of features or building new GIS layers.

GIS operations

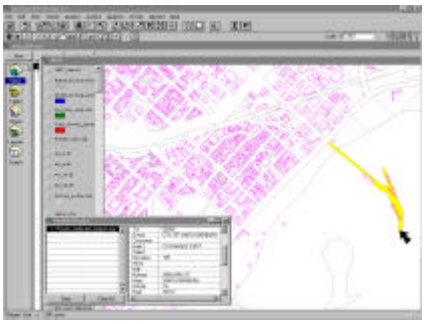
- select and join (change the map scale or extent)
- the buffer operation (within a certain distance of a point, a set of points, a line, or an area)
- a map *overlay* (merging attributes across flat files)
- use of multiple operations in sequence
- weighting

Spatial Search

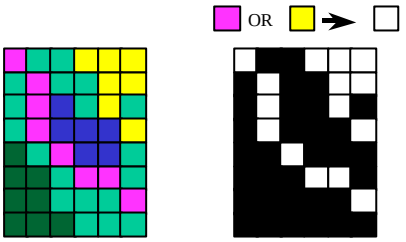
Buffering is a spatial retrieval around points, lines, or areas based on distance.

Overlay is a spatial retrieval operation that is equivalent to an attribute join.

identify

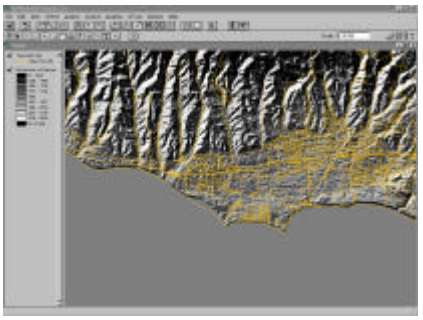


Recode



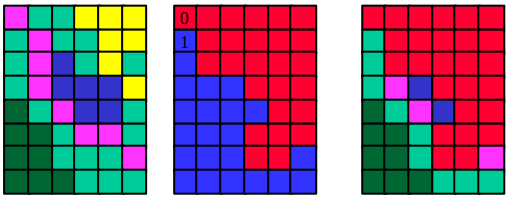
ITU Photogrammetry Division

Data overlay



ITU Photogrammetry Division

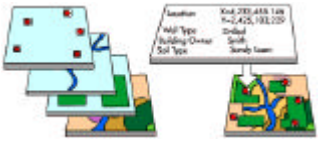
Overlay



ITU Photogrammetry Division

Overlay Analysis

The integration of different data layers involves a process called overlay. At its simplest, this could be a visual operation, but analytical operations require one or more data layers to be joined physically.



This overlay, or spatial join, can integrate data on soils, slope, and vegetation, or land ownership with tax assessment.

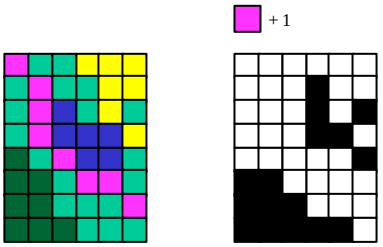
ITU Photogrammetry Division

Types of overlay operations

And
Or
Max
Min
Exhaustive set

ITU Photogrammetry Division

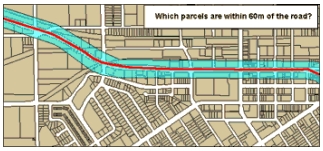
Buffer



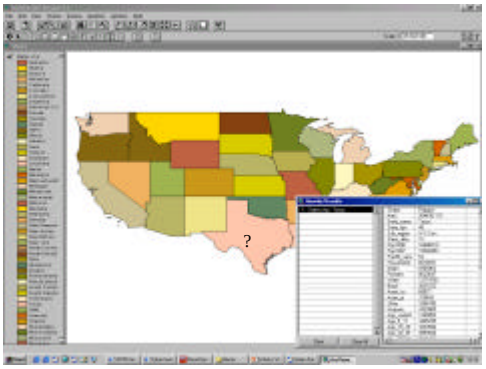
ITU Photogrammetry Division

Proximity Analysis – BUFFER operation

- How many houses lie within 100 m of this water main?
- What is the total number of customers within 10 km of this store?
- What proportion of the alfalfa crop is within 500 m of the well?

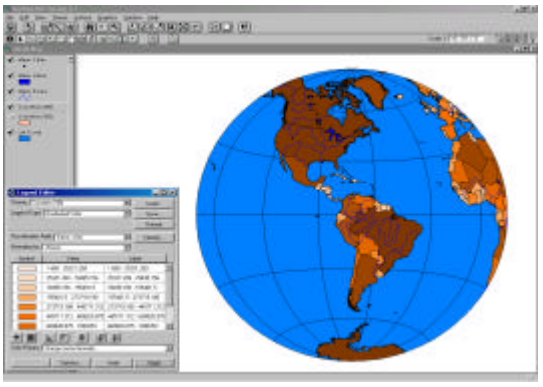


To answer such questions, GIS technology uses a process called buffering to determine the proximity relationship between features.

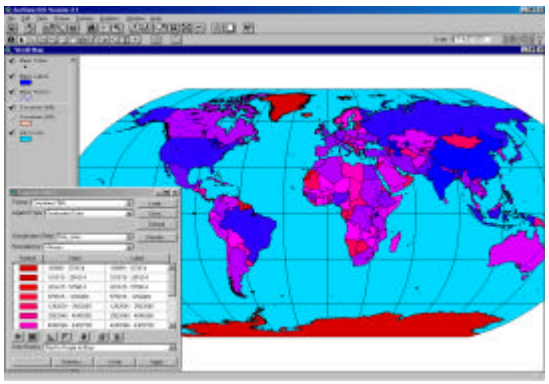


Special search in GIS using the *identify* command.

Mercator Projection

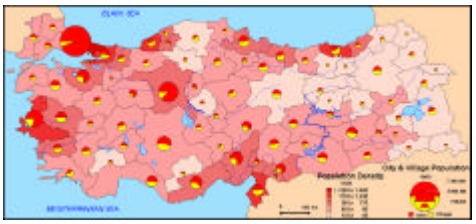


Lambert Equal Area Azimuthal Projection



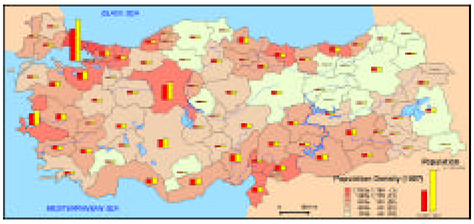
Robinson Projection

Searches by Geography



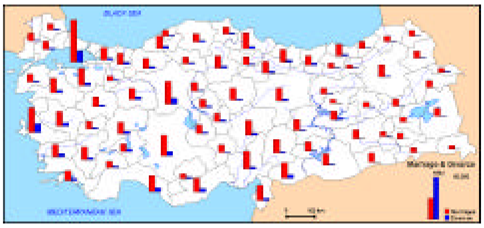
This figure has been taken from Electronic Atlas Project – ITU 1185

Searches by Geography



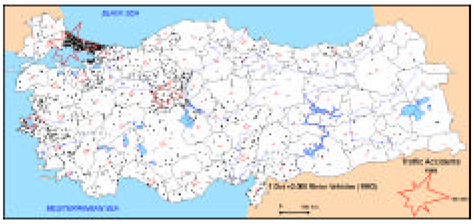
This figure has been taken from Electronic Atlas Project – ITU 1185

Searches by Geography



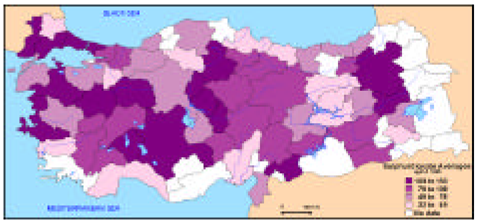
This figure has been taken from Electronic Atlas Project – ITU 1185

Searches by Geography



This figure has been taken from Electronic Atlas Project – ITU 1185

Searches by Geography



This figure has been taken from Electronic Atlas Project – ITU 1185

Complex Retrieval: Map Algebra

Combinations of spatial and attribute queries can build some complex and powerful GIS operations, such as weighting.

Weighted overlay analysis really just complex retrieval.

Coming next…

Why is it there?