Fundamental Concept of Spatial Database

Unit: V

Semester: II

Paper Code: GIS 07

Name of Paper: Geographic Information System

PG Diploma in RS & GIS

Dr. SHYAMA PRASAD MUKHERJEE UNIVERSITY, RANCHI

4.1 Fundamental Concept of Spatial Database

A spatial database is defined as a collection of inter-related geodspatial data, that can handle and maintain a large amout of data which is shareable between different GIS applications.

Required functions of a spatial database are as follows.

- ✓ consistency with little or no redundancy
  maintenance of data quality including updating
  self-descriptive with metadata
  high performance by database mangement system with database langauage
  security including access control

In 1980's, GIS institution was centralized with a centralized spatial database. But in 1990's, the network concept arose, whice will be more convenient to meet user needs with distributed databases as shown in Figure 4.1. Such distributed databases in a network structure have the following benefits.

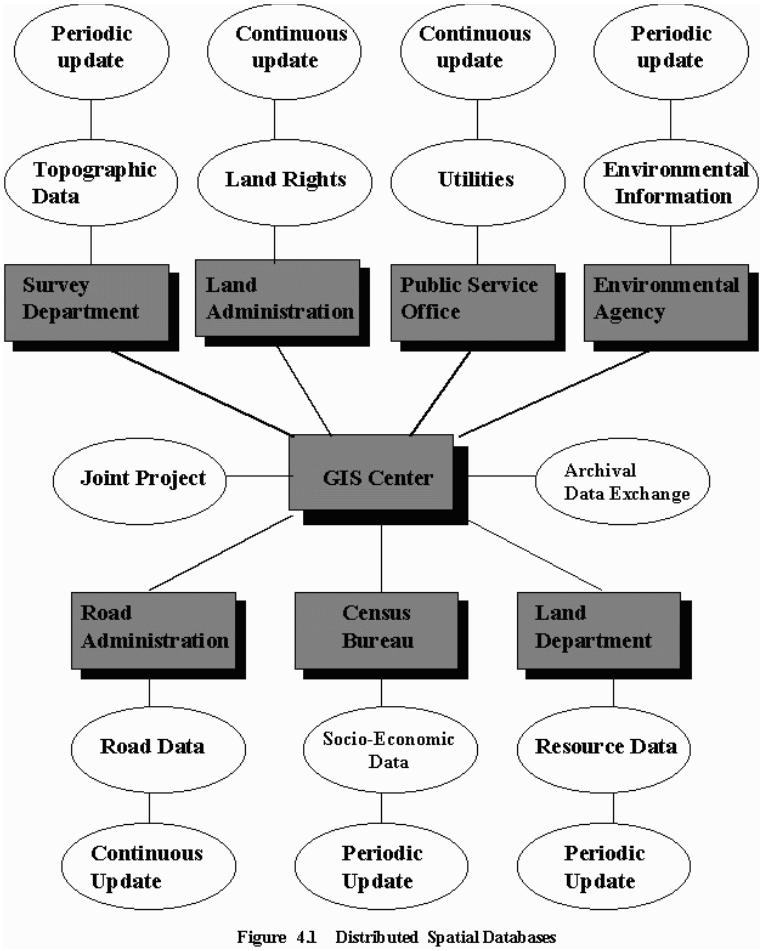


Figure 4.1 Distributed Spatial Databases

- ✓ better data storage and updating
- ✓ more efficient retrieval
- ✓ more efficient output

4.2 Design of Spatial Database

The design of spatial database will be made by the database manager who is responsible for the following issues.

- ✓ definition of database contents
- ✓ selection of database structure
- ✓ data distribution to users
- ✓ mainteneance and updating control
- ✓ day-to-day operation

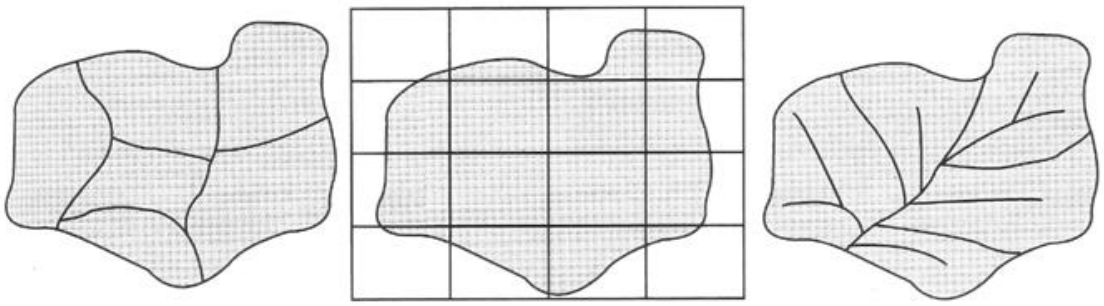
For the design of detail items, the following parameters should be well designed.

❖ Storage media

Volume, access speed and on line service should be considered. Table 4.1 shows the different types of storage media.

❖ Partition of data

Choice of administrative boundaries, map sheets, watersheds etc. will be made in consideration of GIS applications (see Figure 4.2)



(a) Administrative boundaries (b) Map sheets (c) Watersheds

Figure 4.2 Partition of Spatial Data

- **Standards**

Format, accuracy and quality should be standardized.

- **Change and updating**

Add, delete, edit and update should be well controlled by the database manager.

- **Scheduling**

Data availability, priorities, data acquisition etc. should be well scheduled.

- **Security**

Copyright, backup system and responsibilities should be well managed.

4.3 Database Management System

A database management system (DBMS) provides a number of functions to create, edit, manipulate and analyse spatial and non-spatial data in the applications of a GIS.

Major functions of a database are as follows:

- ✓ creating records of various data types; integer, real, character, data, image etc.
- ✓ operations ; sort, delete, edit, select etc.
- ✓ manipulation ; input, analysis, output, reformatting etc.
- ✓ query ; will be made by a standardized language such as SQL (Satandard Query Language)
- ✓ programming ; will be useful for application programs
- ✓ documentation ; metadata or description of the contents of the database should be complied.

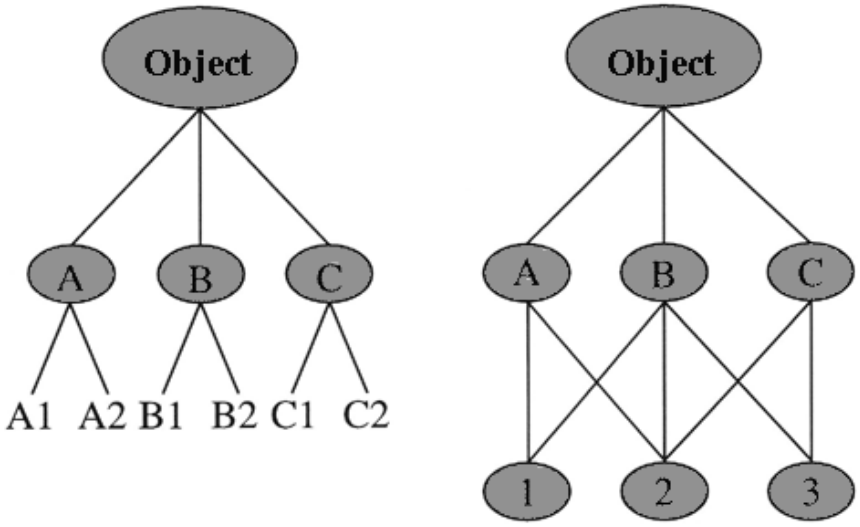
There are four types of database models :

- ✓ hierarchical model
- ✓ network model
- ✓ relational model
- ✓ object oriented model

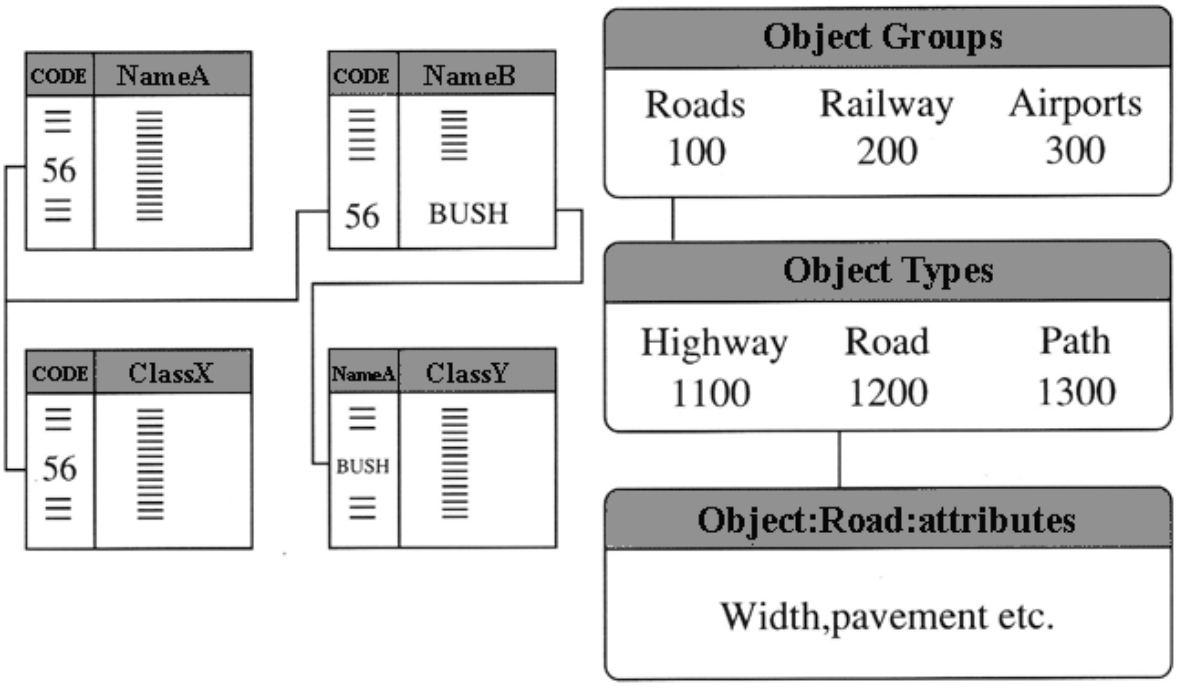
Although all four types are used, the relational model has been most successful in GIS. Well known relational databases include dBase, Oracle and Info. Object oriented model is a new concept that has been recently developed. There has been debate on which of the two-layers or object oriented is efficient in GIS. Layers may be efficient for natural resources management, for example with different layers of land use, soil, geology, agriculture, forests etc. On the other hand object orientation may be more convenient for facility management with grouped attributes. Figure 4.3 shows the concept of four types of database model.

4.4 Hierarchical Model

Several records or files are hierachically related with each other. For example, an organization has several departments, each of which has attributes such as name of director, number of staffs, annual products etc. Each department has several divisions with attributes of name of manager, number of staffs, annual products etc. Then each division has several sections with attributes such as name of head, number of staff, number of PCs etc. Hierachical model is a type of tree structure as shown already in Figure 4.3 (a). A set of links connect all record types in a tree structure. The advantages of hierachical model are high speed of access to large datasets and eases of updating. However, the disadvantage is that linkages are only possible vertically but not horizontally or diagonally, that means there is no relation between different trees at the same level unless they share the same parent. The Quadtree, that is used to access a small part of a large raster image or map area, is a type of hierachical model. Quadtree first divides a total map area into 4, 16, 32, ... step by step as shown in Figure 4.4 (a). Secondly a quadtree is built as shown in Figure 4.4 (b), that makes the access to a particular area at high speed. Numbering of 0, 1, 2 and 3 known as Morton order, makes effective coding of a block or a pixel in a raster model. For example the block of 211 in Figure 4.4 (a) can be expressed 100101 in a pair of base 2 digits, while the conventional block number (4, 3), line and row number needs more bits in a computer.

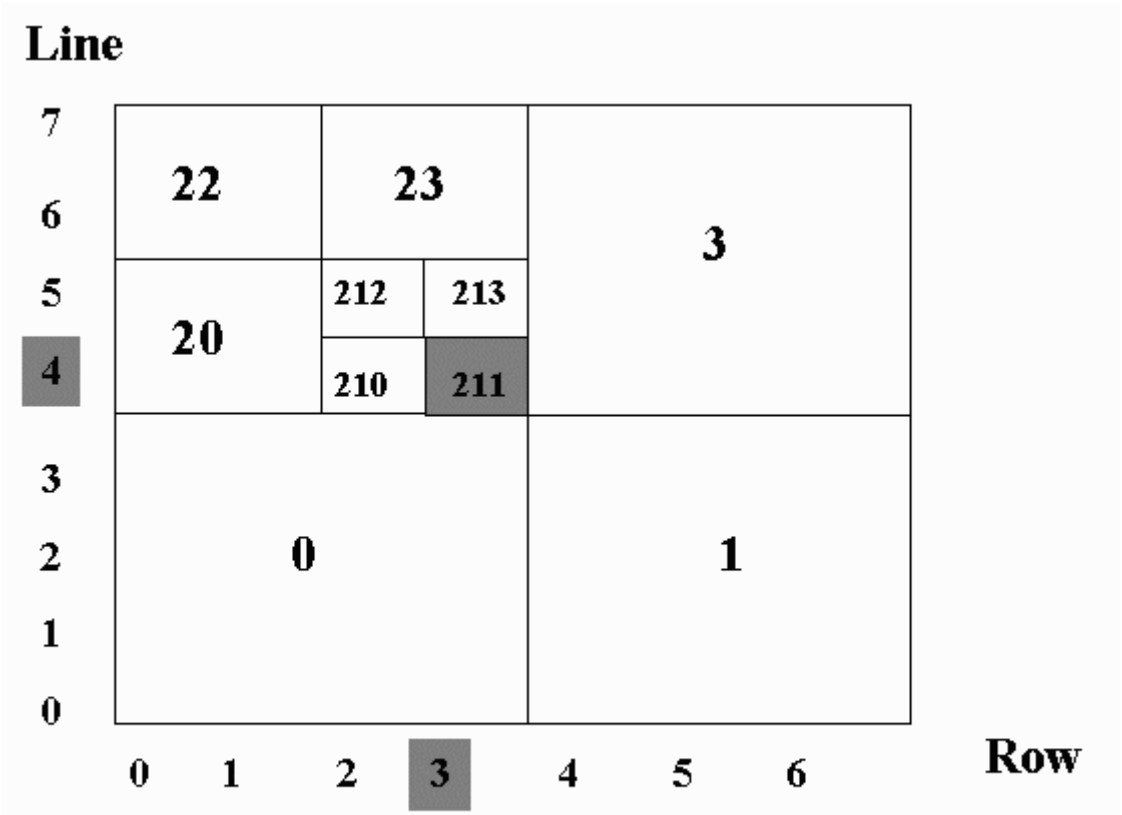


(a) Hierarchical model (b) Network model

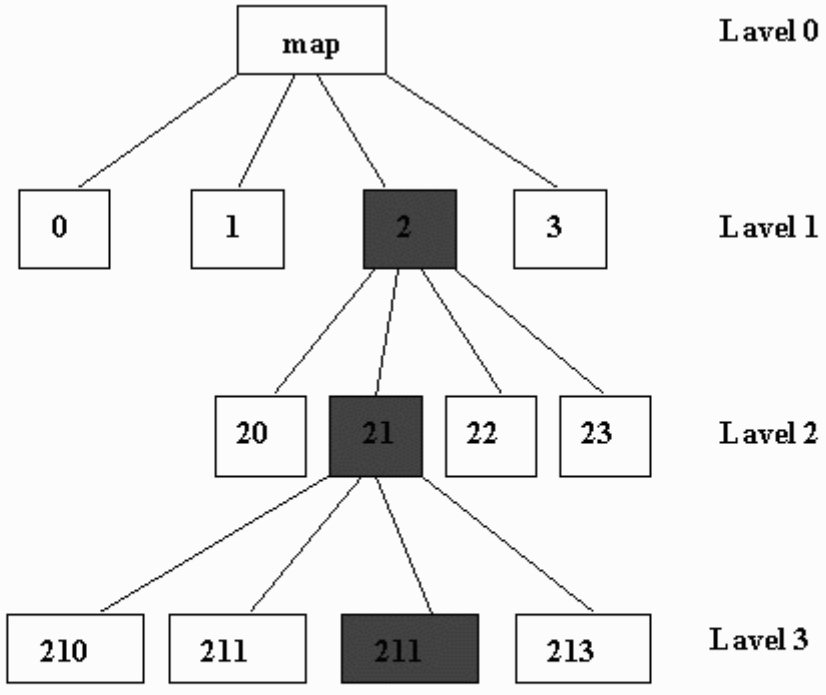


(c) Relation of model (d) Object oriented model

Figure 4.3 Concept of Database Models



(a) Partition of a map



(b) Quadtree

Figure 4.4 Quadtree Model

4.5 Relational Database

Relational database is the most popular model for GIS. For example, the following relational database software's are widely used.

- ✓ INFO in ARC/INFO
- ✓ DBASE III for several PC-based GIS
- ✓ ORACLE for several GIS uses

Relational database is a model to link the complex spatial relationships between objects.

The spatial objects are tabulated in tables consisting of records with a set of attributes as shown already in Figure 4.3 (c). Each table (called relation) consists of different number of attributes, which is called the degree. The degree of an attributes refers to n-ary (ex. unary, binary etc.) relation.

In a relational model, the following two important concepts should be defined.

I. **Key of relation ;** a subset of attributes
II. **Unique identification ;** e.g. the key attributes is a phone directory in a set of last name, first name and address.
III. **non redundancy ;** any key attribute selected and tabulated should keep the key's uniqueness. e.g. address can not be dropped from telephone address, because there may be many with the same names.
IV. **Prime attribute :** an atribute listed in at least one key.

The most important point of the relational database design is to build a set of key attributes with a prime attribute, so as to allow dependence between attributes as well as to avoid loss of general information when records are inserted or deleted. Table 4.2 shows how to build the relational database by normalizing an unstructured table.

The advantages of the relational database are as follows.

- ✓ there is no redundancy.
- ✓ type of building of an owner can be changed without destroying the relation between type and rate.
- ✓ a new type of building for example "Clay" can be inserted.

4.6 Object Oriented Database

An Object-Oriented model uses functions to model spatial and non-spatial relationships of geographic objects and the attributes. An object is an encapsulated unit which is characterized by attributes, a set of orientations and rules.

An object oriented model has the following characteristics.

- ✓ **generic properties** : there should be an inheritance relationship.
- ✓ **abtraction :** objects, classes and super classes are to be generated by classification, **generalization,** association and aggregation.
- ✓ **adhoc queries** : users can order spatial operations to obtain spatial relationships of geographic objects using a special language.

For example, let us try to represent a thought : "Hawaii is an island that is a state of USA" in GIS. In this case, we don稚 mind the geographic location with latitude and longtitude in the conventional GIS model. This is not appropriate to use the layers. In an object oriented model, we are more careful with spatial relationships for example, "is a" (island is a land) and "part of" (state is a part of country). In addition Hawii (state) has Honolulu City and also is in Pacific Region. Figure 4.5 (a) shows "is a" inheritance for the super class of land, while Figure 4.5 (b) shows the spatial relationships for the object of state.

An object oriented database is based on a semantic model as shown in Figure 4.6, which is usually manged by a spacial language although the language has not yet been fully completed.

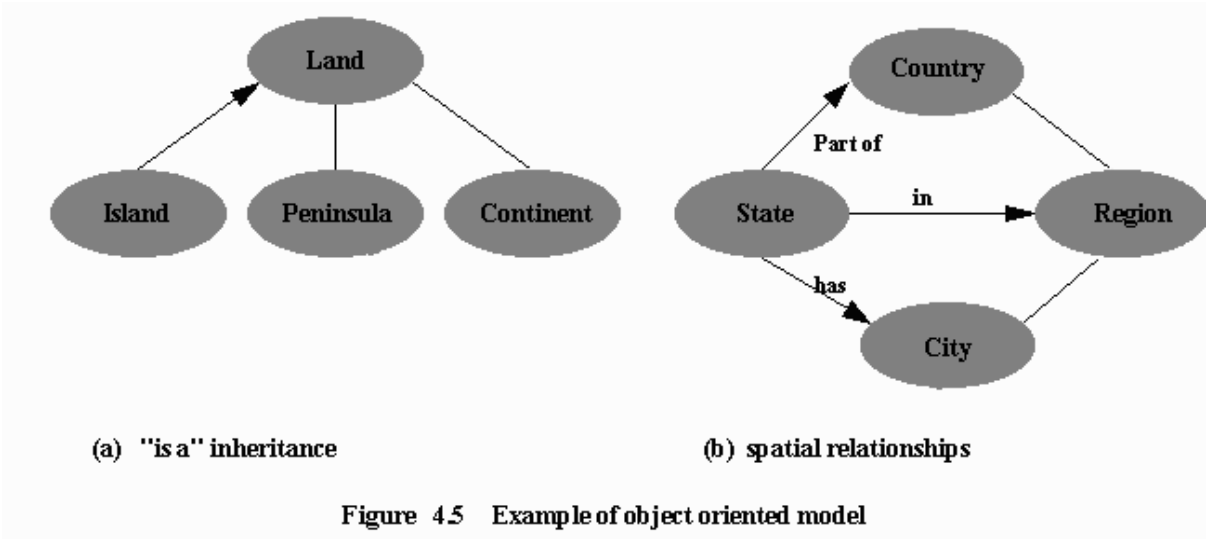


Figure 4.5 Example of object oriented model