GEOG 413/613

LECTURE 3

1

Geospatial Databases

- A database can be thought of as an integrated set of data on a particular subject
  - Data are related and represent a specific aspect of the world (sometimes referred to as a miniworld)
  - Data are for a specific purpose (users and applications) to describe an organization or domain

2

Geospatial Databases

- Databases can be physically stored in files or by using specialist software programs called Database Management Systems (DBMS)
- Geospatial database is a collection of geographic (or spatial) data
  - Has entities (house, river, lake, road…)
  - Has attribute of these entities (location, size, type, name…)
  - Has spatial relationships (distances between entities, adjacency…)

3

Merits of Geospatial Databases

- Data stored at a single location reduces redundancy
  - Consider cadastral data needed by different levels of govt or departments
- Maintenance costs decrease
- Multiple applications and users can use the same data
  - Data are not dependent on software
- Data sharing is easier
  - Multiple interfaces and operations
- Data security and standards

4

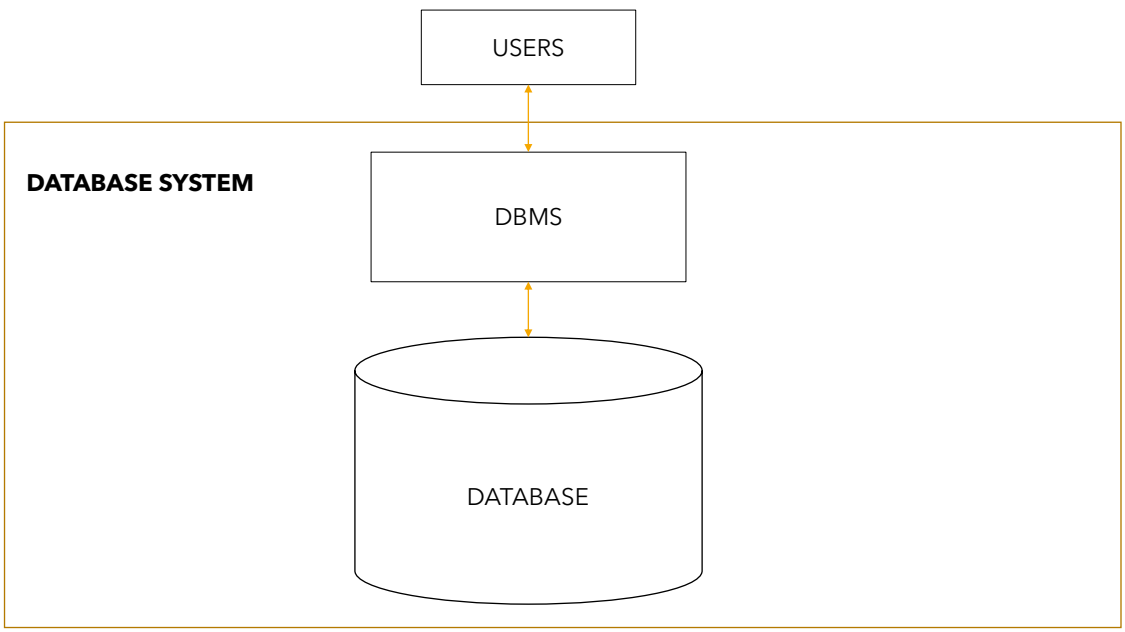
4

Disadvantages Geospatial Databases

- The cost of acquiring DBMS software can be quite high
- A DBMS can add unnecessary complexities for data management in small projects
- Single-user GIS will often be better for files rather DBs

5

5



6

6

Database Management Systems

- Small DBs for a handful of users can be stored on a hard disk
- However, large, complex databases with hundreds of users require specialist DBMS software to ensure data integrity and longevity.
  - A DBMS is a software application designed to organize the efficient and effective storage and access of data. It is used to define, construct, and manipulate a database.

7

7

Functions of a DBMS

- A data model:
  - a mechanism used to represent real-world objects digitally in a computer system, including standard data models suitable for representing several object types (e.g., integer and floating-point numbers, dates, and text)
  - Data model describes the structure of the DB
    - the data types
    - the relations between the data items
    - the constraints that should hold on the data
  - Geospatial DBMSs support geographic (spatial) object types

8

8

Functions of a DBMS

- Data input capability
  - tools to load data into databases in well-structured formats
- Indexing
  - An index is a data structuring used to speed up searching. All databases include tools to index standard database data types
- A query language
  - a standard data query/manipulation language called SQL (Structured/Standard Query Language).

9

9

Functions of a DBMS

- Security
  - DBMSs provide controlled access to data
  - data access rights (read, write, modify, delete) for users to parts or all of DB
- Backup and recovery
  - to protect system from failure and incorrect (accidental or deliberate) update
- Database administration tools
  - Setting the structure of a database (the schema), indexing, backup and recovery, access control

10

10

Functions of a DBMS

- Applications
  - General-purpose tools for creating, using, and maintaining databases (e.g., forms and reports)
- Application programming interfaces (APIs) for further customization

11

Types of DBMSs

- Three main types of DBMSs have been used in GI systems:
  - relational (RDBMS)
    - comprises a set of tables, each a two-dimensional list (or array) of records containing attributes about the objects
    - most of the DBMSs are built on relational DB concepts
  - object (ODBMS)
    - RDBMSs were focused primarily on business applications such as banking, human resource management, and stock control and inventory, they were never designed to deal with rich data types, such as geographic objects, sound, and video
    - ODBMSs can store objects persistently
  - object-relational (ORDBMS)
    - An ORDBMS is an RDBMS engine with some additional capabilities for dealing with objects.
      - Provide object description (attributes such as color, size, and age)
      - Provide object methods or functions such as drawing instructions, query interfaces, and interpolation algorithms)
      - Examples: IBM's Informix, Microsoft's SQL Server, Oracle's Oracle DBMS and PostgreSQL (open source development)

12

Geographic DBMS Extensions

- Several ORDBMS have spatial database extension to provide core support for geographic data types and functions
  - The open-source DBMS PostgreSQL has the PostGIS extension which supports spatial types and functions
- Typically spatial extensions provide
  - Indexing; Storage management; Transaction services; Query language; DB replication services; Query parser; Query optimizer

13

13

Storing Geospatial Data in DBMS Tables

- Databases users interact with an object class
  - The Object Class is also what is referred to as a layer or feature class (data on a particular these)
  - Object classes are stored in standard database tables
  - Database tables are designed along the following principles
    - There is only one value in each cell at the intersection of a row and column.
    - All values in a column are about the same subject.
    - Each row is unique
    - There is no significance to the sequence of columns
    - There is no significance to the sequence of rows

14

14

Storing Geospatial Data in DBMS Tables

- A database table is a two-dimensional array of rows and columns
- Each object class is stored as a single database table in a DBMS.
  - Table rows contain objects (instances of object classes, e.g., data for a single polygon)
  - Table columns contain object properties (the attributes)
  - Geographic database tables have a geometry column (sometimes called the shape column. The coordinate values may be stored in a highly compressed format.

15

15

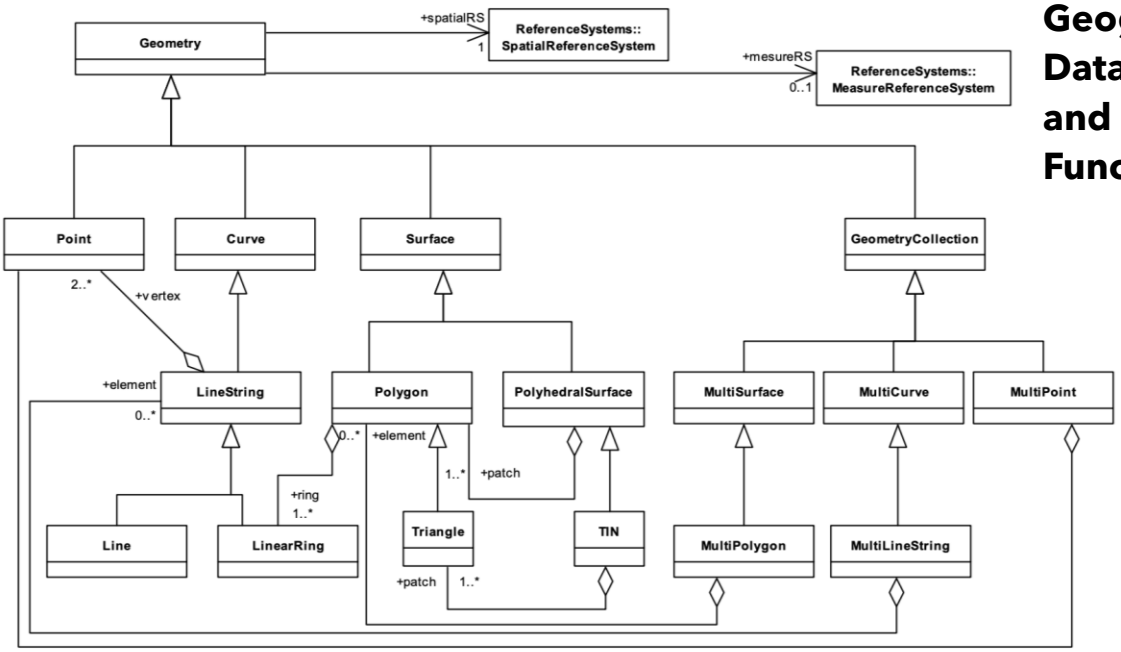
SQL

- SQL (Structured Query Language) programming language designed to retrieve sets (row and column combinations) of data from relational databases
- It is the standard database query language it has geographic capabilities
- Some DBMS can have proprietary SQL extensions that are usually only used on their system

16

16

Geographic Data Types and Functions

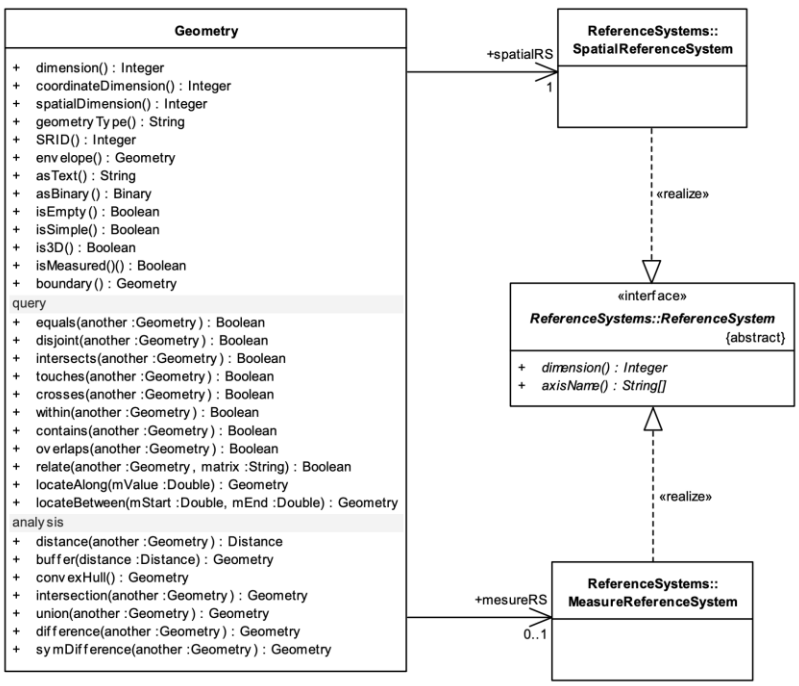


17

OGC Geometry Hierarchy

17

Geographic Data Types and Functions



18

OGC Geometry Class Operations

18

Geographic Database Design

- A GIS will normally be associated with a specific domain e.g. water, forestry, land management
- A database design is the process of modelling the data for this specific domain.
  - This design includes:
    - specification of all data types and relationships
    - the actual database configuration required to store them
- The database design process involves three stages: conceptual, logical, and physical

19

19

DB Design: Conceptual Model (high level)

1. Model the User's View
   - Identify organization functions
   - Determine data require for these functions
   - Organize the data into groups for management
- One can present this conceptual model as as report with accompanying tables

20

20

DB Design: Conceptual Model

2. Define the Objects and their Relationships
   - Specify object types (classes) and functions
   - Specify relationships between object types

- Object models and diagrams are used to describe a set of object classes and the relationships between them

21

DB Design: Conceptual Model

3. Select Geographic Representation
   - discrete object or continuous field
   - It is possible to change between representation type later but it can be complex and expensive with a risk of information loss

22

DB Design: Logical Model (medium level)

1. Match to Geographic Database Types
   - Match object types to specific data types (point, line, area, georaster, etc.)
     - This could be implemented in Oracle, PostGIS,…
2. Organize Geographic Database Structure
   - Define topological rules and relationships
   - Define coordinate systems

- Hides details of database structure but outlines how data is organised

23

23

DB Design: Physical Model (low level)

1. Final stage before the implementation of the DB
   - Details of how data is structured in a database to minimise storage requirements and improve performance
   - Choices on hardware would be made at this stage

24

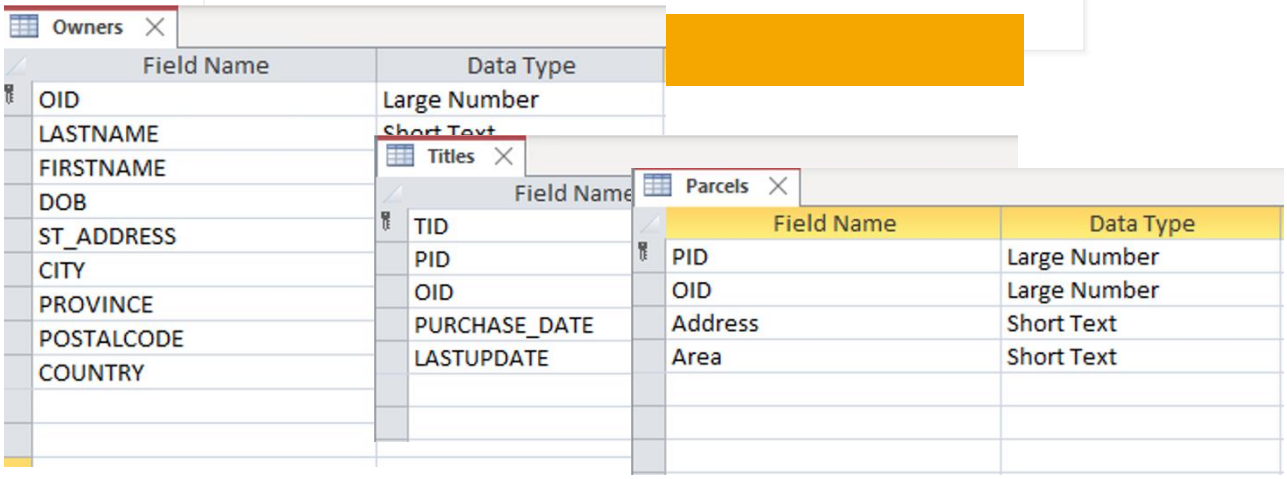
24

Database Schema and Instances

- A data models make a clear distinction between the content and its description
  - the description of the database is called *database schema*
  - the data items that reside in a database at a specific point in time form the *database instance*
  - The database schema is specified during the database design process
    - and is not expected to change frequently
    - each modification on the schema may affect the database relation

25

Database Schema



26

Database Schema

```
CREATE SCHEMA cadastre;
create table cadastre.parcels (
    PID BIGINT NOT NULL,
    OID BIGINT NOT NULL,
    ADDRESS VARCHAR (50) NOT NULL
    AREA AGE INT NOT NULL,
    PRIMARY KEY (PID)
);
```

27

Database Instance

OID	LASTNAME	FIRSTNAME	DOB	ST_ADDRESS	CITY	PROVINCE	POSTALCODE	COUNTRY
12345678	Abe	Ed	11/01/80	123 West St	Sunset	DD	VVVVVV	
23456789	Bud	Fiona	13/02/81	123 East Ave	Meadows	GG	YYYYYY	
34567890	Cage	Gayle	15/03/82	780 Run Cr	Springfield	EE	UUUUU	

TID	PID	OID	PURCHACE_DATE	LAT_UPDATE
2097541	5667890256	12345678	11/01/22	123 West St
5689014	2677908390	23456789	13/02/21	123 East Ave
2678902	2342567123	34567890	15/03/22	780 Run Cr

PID	OID	ADDRESS	AREA
5667890256	12345678	123 Pine St	0.3
2677908390	23456789	123 Queens Ave	0.23
2342567123	34567890	780 Mayor Cr	0.65

28