International Journal of Management, IT & Engineering
Vol.14 Issue 9, September 2024,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Spatial Data and Spatial Databases

Prashant Nadagoud

Abstract

Spatial data and spatial databases are fundamental to representing and analyzing geographic information in digital systems. This article explores the nature of spatial data, which describes the position, shape, and orientation of objects in space using geometric approximations. It outlines the hierarchical spatial data model consisting of elements, geometries, and layers, and discusses various coordinate systems used to reference spatial data. The article compares raster and vector data models for representing spatial information and introduces the concept of spatial databases, which are optimized for storing and querying spatial data. Key features of spatial databases, including spatial indexing and support for specialized spatial operations, are highlighted. The two-tier query model used in spatial databases is explained, emphasizing the balance between efficiency and accuracy in spatial data processing. This comprehensive overview provides insights into the structure, management, and application of spatial data in geographic information systems and related fields.

Database

It is an organized group of data in the digital form, for one or more purposes. The data are typically organized to model relevant aspects of reality, for example, the availability of rooms at a particular motel at a specific date.

Database Management System

It is a software package with computer programs that control the creation, maintenance, and use of a database. It is a collection of data records, files, and other objects.

Spatial Data

Spatial data describes the position, shape, and orientation of objects in space. The objects may be physical things, such as an office building or a mountain or a street, or abstract ones, such as an imaginary line marking the political boundary between countries.

Uses of Spatial Data

- Analyzing regional, national, or international sales trends.
- Taking up a decision such as where to place a new store based on proximity to customers and competitors.
- Navigating using a Global Positioning System (GPS) device.
- Tracking the delivery of a parcel by the customers.
- Populating geographic based information on a map rather than in a tabular or chart format.

Representing features on the Earth

In the real world, objects on the earth frequently have complex, uneven shapes. It would be very hard for any item of spatial data to define the exact shape of these objects. So, spatial data represents these objects by using simple, geometrical shapes that approximate their actual shape and position. These shapes are called geometries.

The below diagram shows the different types of geometries that can be used to represent the features on the earth.

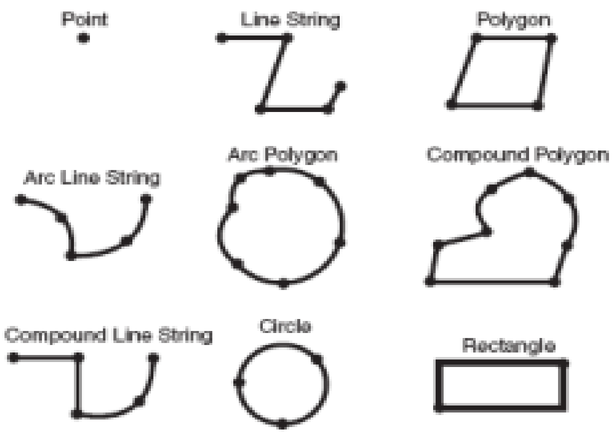

fig 1.1 Geometry Types

Spatial data types are organized in a hierarchical manner. Each sub-type inherits the properties and the behavior (methods or functions) of its super-type. The following figure represents the hierarchy of the spatial data types. Geometry being the super class.

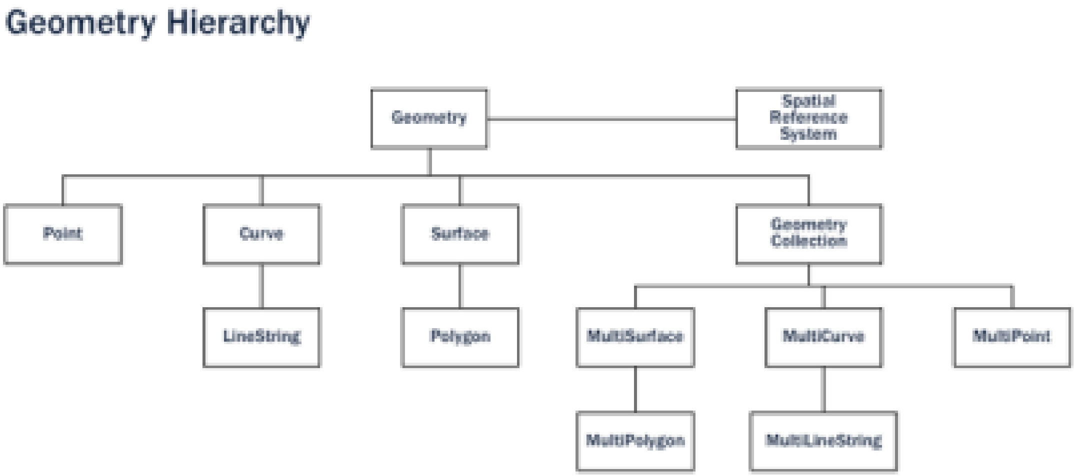

fig 1.2 Geometry Hierarchy

Spatial Data Model

It is a hierarchical structure consisting of *elements*, *geometries*, and *layers*, which correspond to representations of spatial data. Layers are composed of geometries, which in turn are made up of elements.

For example, a point might represent a building location, a line string might represent a road or flight path, and a polygon might represent a state, city, zoning district, or city block.

Element

An *element* is the basic building block of geometry. The supported spatial element types are *points*, *line strings*, and *polygons*. For example, elements might model star constellations (point clusters), roads (line strings), and county boundaries (polygons). Each coordinate in an element is stored as an X,Y pair. The exterior ring and the interior ring of a polygon with holes are considered as two distinct elements that together make up a complex polygon.

Point data consists of one coordinate.

Line data consists of two coordinates representing a line segment of the element.

Polygon data consists of coordinate pair values, one vertex pair for each line segment of the polygon. Coordinates are defined in order around the polygon (counterclockwise for an exterior polygon ring, clockwise for an interior polygon ring).

Geometry

A *geometry* (or geometry object) is the demonstration of a spatial feature, modeled as a structured set of primitive elements. It can consist of a single element, which is an instance of one of the supported primitive types, or a homogeneous or heterogeneous collection of elements. A multipolygon, such as one used to represent a set of islands, is a homogeneous collection. A heterogeneous collection is one in which the elements are of different types, for example, a point and a polygon.

Layer

A *layer* is a collection of geometries having identical feature set. For example, one layer in a GIS might include topographical features, while another describes population density, and a third describes the network of roads and bridges in the area (lines and points). Each layer's geometries and associated spatial index are stored in the database in standard tables.

Coordinate System

A *coordinate system* (also called a *spatial reference system*) is a way of assigning coordinates to a location and creating relationships between sets of such coordinates.

It enables the interpretation of a set of coordinates as a representation of a position in a real world space.

Each spatial data has a coordinate system associated with it. The coordinate system can be of type *georeferenced* (related to a specific representation of the Earth) or *not georeferenced* (that is, Cartesian, and not related to a specific representation of the Earth).

Spatial data can be associated with a Cartesian, geodetic (geographical), projected, or local coordinate system:

Cartesian coordinates are coordinates that measure the position of a point from a defined origin along axes that are perpendicular in the represented two dimensional or three-dimensional space.

If a coordinate system is not explicitly associated with a geometry, a Cartesian coordinate system is assumed.

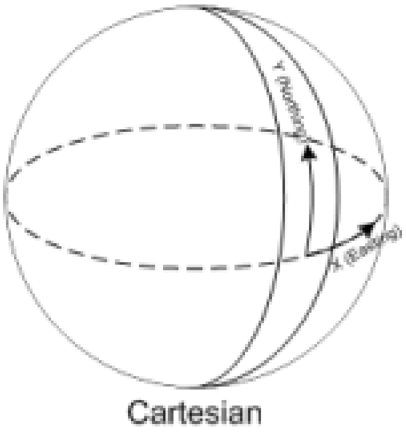

fig 1.3 Cartesian Coordinate System

Geodetic coordinates (sometimes called *geographic coordinates*) are angular coordinates (longitude and latitude), closely related to spherical polar coordinates, and are defined relative to a particular Earth geodetic datum. (A geodetic datum is a means of representing the figure of the Earth and is the reference for the system of geodetic coordinates.)

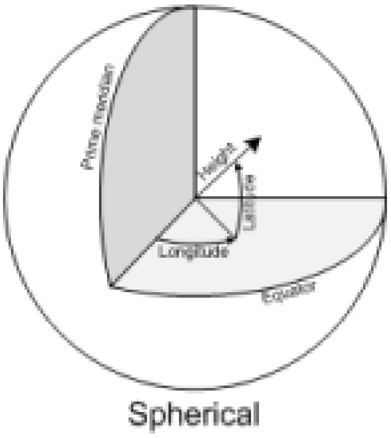

fig 1.4 Geographic Coordinate System

Projected coordinates are planar Cartesian coordinates that result from performing a mathematical mapping from a point on the Earth's surface to a plane.

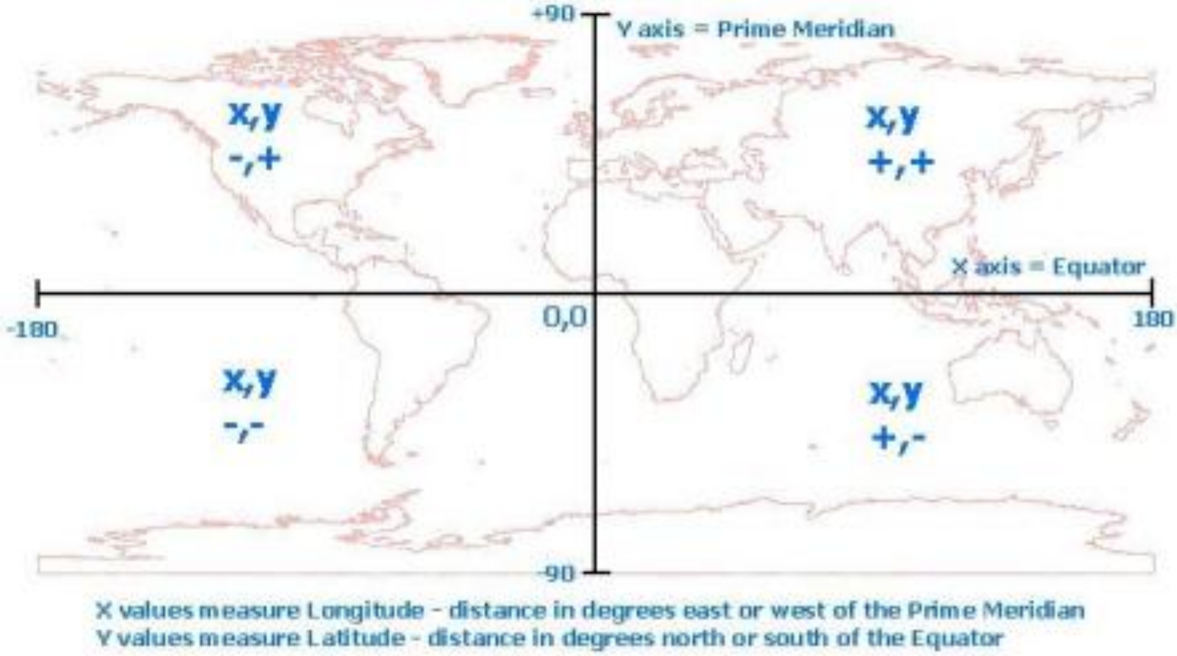

Fig 1.5 Projected Coordinate System

Local coordinates are Cartesian coordinates in a non-Earth (non-georeferenced) coordinate system. Local coordinate systems are often used for CAD applications and local surveys.

Types of Spatial Data Model

In GIS, there are 2 basic spatial data models representing the real world:

Raster Model

In the raster data model, land cover is represented as single square cells.

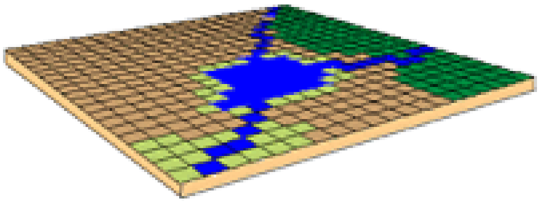

fig 1.6 Raster data model

Raster data are good at:

▪ Representing continuous data (e.g., slope, elevation)

▪ Representing multiple feature types (e.g., points, lines, and polygons) as single feature types (cells).

▪ Rapid computations ("map algebra") in which raster layers are treated as elements in mathematical expressions.

▪ Analysis of multi-layer or multivariate data (e.g., satellite image processing and analysis).

Vector Data

In the vector data model, features of the earth are represented as

▪ Points

▪ Lines/routes

▪ Polygons/regions

▪ Triangulated irregular Network

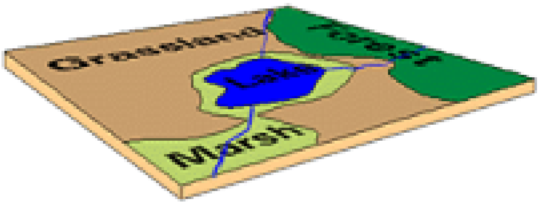

fig 1.7 Vector data model

Vector data are good at:

▪ Accurately representing true shape and size

▪ Representing non-continuous data (e.g., rivers, political boundaries, road lines, mountain peaks)

Overview of Spatial Database

Spatial Database

A *Spatial Database* is a database that is optimized to store and query data that is related to objects in space, including points, lines and polygons. While typical databases can understand various numeric and character types of data, further functionality needs to be added for databases to process spatial data types. These are typically called *geometry* or *features*. The Open Geospatial Consortium (OGC) created the simple features specification and sets standards for adding spatial functionality to database systems.

Features of Spatial Databases

Database systems use indexes to quickly look up values and the manner that most databases index data is not optimal for spatial queries. So, spatial databases use a spatial index to improve database operations.

In Addition to typical SQL queries such as SELECT statements, spatial databases can perform a wide variety of spatial operations. The following query types and many more are supported by the Open Geospatial Consortium:

I. Spatial Measurements: Finds the distance between points, polygon area, etc.

II. Spatial Functions: Modify existing features to create new ones, for example by providing a buffer around them, intersecting features, etc.

- Spatial Predicates: Allows true/false queries such as 'is there a residence located within a mile of the area we are planning to build the landfill?'

III. Constructor Functions: Creates new features with an SQL query specifying the vertices (points of nodes) which can make up lines. If the first and last vertex of a line are identical the feature can also be of the type polygon (a closed line).

IV. Observer Functions: Queries which return specific information about a feature such as the location of the center of a circle

Not all spatial databases support these query types.

Query Model

Spatial databases use a *two-tier* query model to resolve spatial queries and spatial joins. The term is used to indicate that two distinct operations are performed to resolve queries. The output of the two combined operations yields the exact result set.

The two operations are referred to as *primary* and *secondary* filter operations.

▪ The primary filter permits fast selection of candidate records to pass along to the secondary filter. The primary filter compares geometry approximations to reduce computation complexity and is considered a lower-cost filter. Because the primary filter compares geometric approximations, it returns a superset of the exact result set.

▪ The secondary filter applies exact computations to geometries that result from the primary filter. The secondary filter yields an accurate answer to a spatial query. The secondary filter operation is computationally expensive, but it is only applied to the primary filter results, not the entire data set.

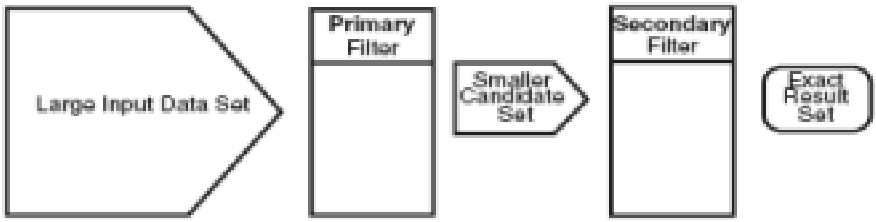

fig 1.8 Query Model

Different vendors of Spatial Database Systems

- Boeing's Spatial Query Server spatially enables Sybase ASE.
- Smallworld VMDS, the native GE Smallworld GIS database.
- SpatiaLite extends Sqlite with spatial datatypes, functions, and utilities.
- IBM DB2 Spatial Extender can be used to enable any edition of DB2, including the free DB2 Express-C, with support for spatial types.
- Oracle Spatial.
- Microsoft SQL Server has support for spatial types since version 2008.
- PostgreSQL DBMS uses the spatial extension PostGIS to implement the standardized datatype *geometry* and corresponding functions.
- MySQL DBMS implements the datatype *geometry* plus some spatial functions.
- Neo4j - Graph database that can build 1D and 2D indexes as Btree, Quadtree and Hilbert curve directly in the graph.

References:

- Wikipedia. Spatial Databases, http://en.wikipedia.org/wiki/Spatial_database.
- IBM. DB2 Database Downloads, http://www-01.ibm.com/support/docview.wss?uid=swg24023357.
- IBM. DB2 Spatial Extender, http://www14.software.ibm.com/webapp/download/product.jsp?s=p&id=ERAI-4JGW5S.
- Wikipedia. B-Tree, http://en.wikipedia.org/wiki/B-tree.
- Sharpgis. Shape2sql, http://www.sharpgis.net/page/SQL-Server-2008-Spatial-Tools.aspx
- National Weather Services. OST/SEC GIS Map Group, http://www.nws.noaa.gov/geodata/catalog/national/html/province.htm.