11.2 Spatial Data

Geographic information systems (*GIS*) and *spatial databases* provide powerful mechanisms to store and retrieve location data [164]. Such systems primarily concentrate on accessing large amounts of spatial data. In the Nimbus approach, the amount of data per server is low compared to geographic information systems. However, the characteristics of data and the required operations are similar.

According to Breunig [18], geo data can be divided into a number of *geo-objects*, each of them representing real geographic entities such as rivers or abstract entities such as city borders. A geo-object can be described by the following information:

- an identifier (e.g. a unique name),
- structure information (e.g. the relations between objects),
- the geometry,
- and a thematic description by attributes (containing e.g. the type of the object).

Sometimes, information about relations of objects is solely stored by their geometry. In contrast, some systems (e.g. street navigation systems) rely on topological relations between objects. From the geo scientist point of view, the thematic description plays an important role. It describes e.g.

- the classification of the object (e.g. topographic, land register, mineral resources, land development, trade and industry),
- the sub-classification (e.g. river, lake or source for topographic entities)
- statistical data (e.g. number of cows per square kilometre, coal-mining capacity in tons per year),
- owner of the corresponding area (for land registers),
- data about the acquisition process (e.g. age of the database entry, operator who enters the data, or accuracy information).

To organize thematic attributes, either *layers* or *classes* can be used. Using layers is the older approach where layers were represented by transparent foils, each representing a special type of geo-object. Maps were constructed by piling these foils and copying the results. The modern approach relies on class hierarchies, similar to class hierarchies in object-oriented software development. Classes of geo-objects can be subclassed in a similar way as software objects.

From this point of view, Nimbus domains are typical geo-objects, as they have a name (*d.name*), relations to other objects defined by the logical links and a geometric description (*d.c*). Up to now, domains do not contain explicit thematic attributes, but get their classification from the hierarchy (e.g. `geo`). As a result of the import of real land survey data (see below), thematic attributes are attached to the domain objects.

Properties of Geo-Spatial Data

From the resolution algorithms' point of view, the geometry is important. The geometric representation of geo-objects can be classified as follows:

- *accurate* vs. *approximated* representation,
- *raster*, *vector*, or *analytical* representation and
- the number of represented *dimensions*.

These classes are not completely independent of each other. Accurate representations represent the geometry exactly. Typically, only simple geometric shapes such as circles and rectangles can be represented accurately, thus typical spatial databases approximate shapes. Especially natural objects such as rivers cannot be represented exactly with a finite storage space.

A geometric representation can further be divided into a *raster*, *vector* or *analytical* representation (fig. 11-1). A raster represents a grid of elements, each of them either belongs to the geo-object or not. Multiple grids can represent different objects. As an advantage, some required operations such as the intersection of two objects have a simple implementation. However, more accurate rasters require huge storage spaces.

The vector representation uses points and lines to represent the geometry of geo-objects. The accuracy depends on the accuracy of a point representation and the number of lines spent to represent curved borders. To achieve a high accuracy, vector representations usually have a lower storage space requirement than comparable rasters, but the implementation of the required operations is more complex.

As a last representation, the geometry of a geo-object can be stored analytically, i.e. with the help of geometric primitives such as circles and rectangles. Some shapes can be exactly specified with a low amount of storage space, but there is no easy way to represent arbitrary shapes as required to represent geographic objects. Analytical representations do not play an important role for geo information systems.

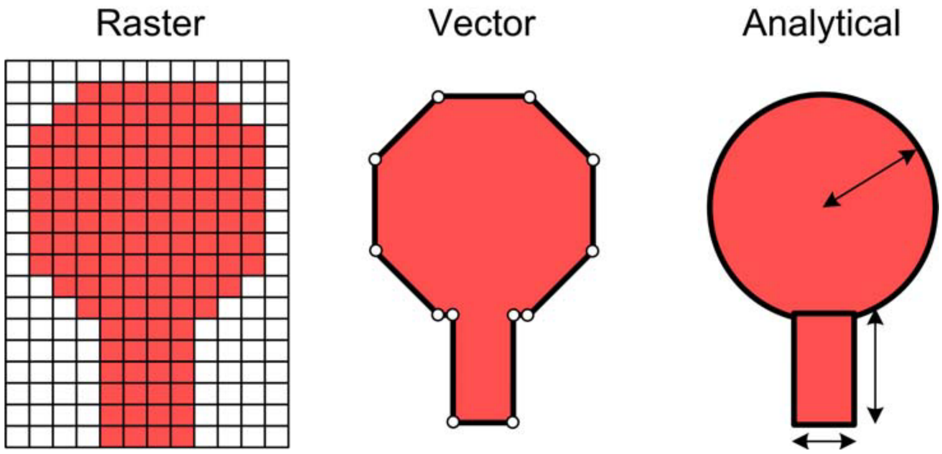

Fig. 11-1. Different geometric representations (2D)

A last classification considers the *dimension* of the represented geometric object (fig. 11-2):

- 2D representations project all objects to the earth's surface and only store projected, i.e. two-dimensional data. Even though a lot of information is lost, this representation is suitable for many scenarios. E.g. city maps usually only contain 2D data.
- 2+1D representations contain height data which is stored independently of the 2D data. Height information can be viewed as an independent layer of information.
- 2.5D representations store height information as an *attribute* of the geometric description (i.e. attached to a polygon point in a vector representation). As the height is only an attribute, it does not actually belong to the geometric description. Each point is still identifiable by its x-y-coordinates.
- 3D representations fully store geo-objects in three dimensions.

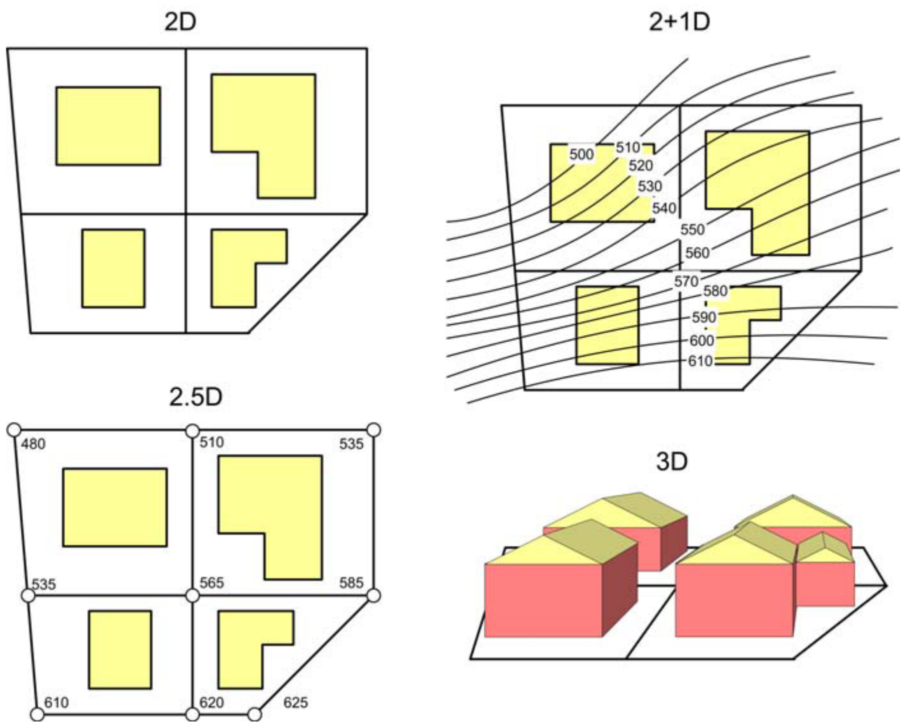

Fig. 11-2. Dimensions of geo data

In addition to the space dimension, often the time is added as a representation dimension. Spatio-temporal representations (3D space, one time dimension) consider the movement of objects over time. With this information, queries like "where was the course of the river in the 19th century?" can be answered.