Chapter 2:- Geographic information and Spatial data types

2.1 Models and representations of the real world

Modelling is the process of producing an abstraction of the 'real world' sothat some part of it can be more easily handled.
Given the complexity of real world phenomena, our models can by definition never be perfect. We have limitations on the amount of data that we can store,limits on the amount of detail we can capture, and (usually) limits on the time we have available for a project. It is therefore possible that some facts or relationships that exist in the real world may not be discovered through our 'models'.
Any geographic phenomenon can usually be represented in various ways; the choice of which representation is best depends mostly on two issues. Firstly, what original, raw data (from sensors or otherwise) is available, and secondly, what sort of data manipulation is required or will be undertaken.

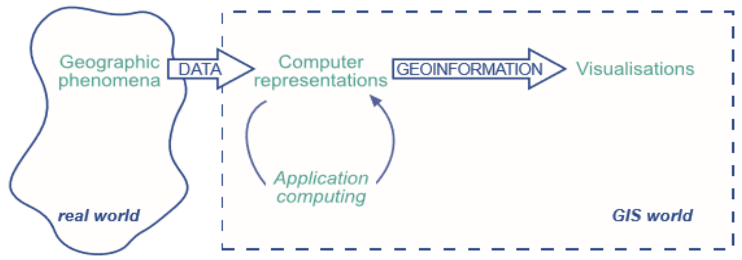


Figure 2.1: Representing relevant aspects of real-world phenomena inside a GIS to build models or simulations.

2.2 Geographic phenomena

2.2.1 Defining geographic phenomena

A GIS operates under the assumption that the relevant spatial phenomena occur in a two- or three-dimensional Euclidean space, unless otherwise specified.
Euclidean space can be informally defined as a model of space in which locations are represented by coordinates—(x, y) in 2D; (x, y, z) in 3D—and distance and direction can defined with geometric formulas. In the 2D case, this is known as the Euclidean plane, which is the most common Euclidean space in GIS use.

In order to be able to represent relevant aspects real world phenomena inside a GIS, we first need to define what it is we are referring to. We might define a geographic phenomenon as a manifestation of an entity or process of interest that:

- Can be named or described,
- Can be georeferenced, and
- Can be assigned a time (interval) at which it is/was present.

Referring back to the El Nino example discussed, one could say that there are at least three geographic phenomena of interest there. One is the Sea Surface Temperature, and another is the Wind Speed in various places. A third geographic phenomenon in that application is the array of monitoring buoys.

Eg :- To study water management:- objects of study are ground water levels, irrigation levels, meterological data, water budgets etc. (Do these follow the above properties?)

2.2.2 Types of geographic phenomena

Firstly, In order to be able to represent a phenomenon in a GIS requires us to state what it is, and where it is. We must provide a description—or at least a name—on the one hand, and a georeference on the other hand.

Secondly, some phenomena manifest themselves essentially everywhere in the study area, while others only do so in certain localities. If we define our study area as the equatorial Pacific Ocean, we can say that Sea Surface Temperature Fields can be measured anywhere in the study area. Therefore, it is a typical example of a (geographic) field.

A (geographic) field is a geographic phenomenon for which, for every point in the study area, a value can be determined

Some common examples of geographic fields are air temperature, barometric pressure and elevation. These fields are in fact continuous in nature. Examplesof discrete fields are land use and soil classifications.

Many other phenomena do not manifest themselves everywhere in the study area, but only in certain localities. The array of buoys of the previous chapter is a good example: there is a fixed number of buoys, and for each we know exactly where it is located. The buoys are typical examples of (geographic) objects.

(Geographic) objects populate the study area, and are usually well- distinguished, discrete, and bounded entities. The space between them is potentially ‘empty’ or undetermined

A simple rule-of-thumb is that natural geographic phenomena are usually fields, and man-made phenomena are usually objects.

2.2.3 Geographic Fields

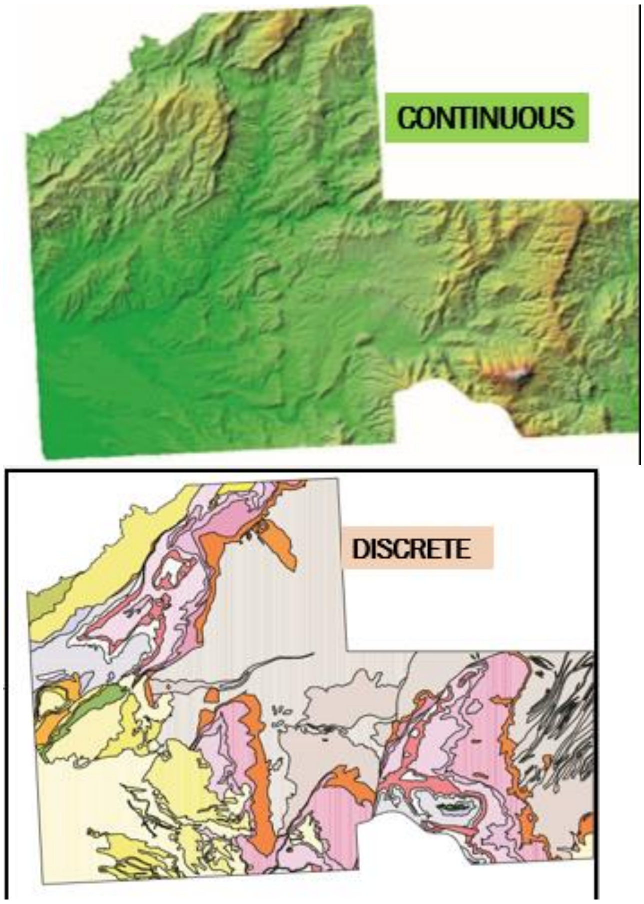
- Has a value everywhere in the study area.
- Field is a mathematical function that associates a particular value with any position in the study area. f(x,y) – f is the value, x,y is the position

Continuous:-

- Field values change gradually
- Field is differentiable; we can determine a measure of change in the field value per unit of distance everywhere and in any direction.
- Floating point values
- Example:- If Field is Elevation, measure is slope i.e change of elevation per metre distance, If Field is Soil Salinity, measure is salinity gradient i.e change of salinity per metre distance

Discrete:-

- Divides study space into mutually exclusive bounded parts with all locations in one part having same field value.
- Makes use of bounded features
- Assigns a value to every location in the study area.
- Integer values



Data types and values

1. Nominal data values are values that provide a name or identifier so that we can discriminate between different values, but that is about all we can do. Specifically, we cannot do true computations with these values. An example are the names of geological units. This kind of data value is called categorical data when the values assigned are sorted according to some set of non-overlapping categories. For example, we might identify the soil type of a given area to belong to a certain (pre-defined) category.

2. Ordinal data values are data values that can be put in some natural sequence but that do not allow any other type of computation. Household income, for instance, could be classified as being either 'low', 'average' or 'high'. Clearly this is their natural sequence, but this is all we can say—we can not say that a high income is twice as high as an average income.

3. Interval data values are quantitative, in that they allow simple forms of com- putation like addition and subtraction. However, interval data has no arithmetic zero value, and does not support multiplication or division. For instance, a temperature of 20 ◦C is not twice as warm as 10 ◦C, and thus centigrade temperatures are interval data values, not ratio data values.

4. Ratio data values allow most, if not all, forms of arithmetic computation**.** Rational data have a natural zero value, and multiplication and division of values are possible operators

 Source:- Principles of Geographic Information Systems,Otto Huisman and Rolf A.de By

(distances measured in metres are an example). Continuous fields can be expected to have ratio data values, and hence we can interpolate them.

We usually refer to nominal and categorical data values as 'qualitative' data, because we are limited in terms of the computations we can do on this type of data. Interval and ratio data is known as 'quantitative' data, as it refers to quantities. However, ordinal data does not seem to fit either of these data types. Often, ordinal data refers to a ranking scheme or some kind of hierarchical phenomena.

2.2.4 Geographic Objects

When a geographic phenomenon is not present everywhere in the study area, but somehow 'sparsely' populates it, we look at it as a collection of geographic objects. Such objects are usually easily distinguished and named, and their po- sition in space is determined by a combination of one or more of the following parameters:

- Location (where is it?),
- Shape (what form is it?),
- Size (how big is it?), and
- Orientation (in which direction is it facing?).

How we want to use the information about a geographic object determines which of the four above parameters is required to represent it.

For instance, in an in-car navigation system, all that matters about geographic objects like petrol stations is where they are. In the same system, however, roads are important objects, and for these some notion of location (where does it begin and end), shape (how many lanes does it have), size (how far can one travel on it) and orientation (in which direction can one travel on it) seem to be relevant information components.

Shape is usually important because one of its factors is dimension. This relates to whether an object is perceived as a point feature, or a linear, area or volume feature. The petrol stations mentioned above apparently are zero-dimensional, i.e. they are perceived as points in space; roads are one-dimensional, as they are considered to be lines in space. In another use of road information—for instance, in multi-purpose cadastre systems where precise location of sewers and manhole covers matters—roads might well be considered to be two-dimensional entities, i.e. areas within which a manhole cover may fall.

Collections of geographic objects can be interesting phenomena at a higher ag- gregation level: forest plots form forests, groups of parcels form suburbs, streams, brooks and rivers form a river drainage system, roads form a road network, and SST buoys form an SST sensor network.

2.2.5 Boundaries

Where shape and/or size of contiguous areas matter, the notion of boundary comes into play. Location, shape and size are fully determined if we know an area's boundary, so the boundary is a good candidate for representing it. This is especially true for areas that have naturally crisp boundaries. A crisp boundary is one that can be determined with almost

arbitrary precision, dependent only on the data ac- quisition technique applied. Fuzzy boundaries contrast with crisp boundaries in that the boundary is not a precise line, but rather itself an area of transition.
As a general rule-of-thumb, crisp boundaries are more common in man-made phenomena, whereas fuzzy boundaries are more common with natural phenomena.

2.3 Computer representations of geographic information

We have seen that various geographic phenomena have the characteris- tics of continuous functions over space. Elevation, for instance, can be measured at many locations, even within one's own backyard, and each location may give a different value. In order to represent such a phenomenon faithfully in com- puter memory, we could either:

1. Try to store as many (location, elevation) observation pairs as possible, or
2. Try to find a symbolic representation of the elevation field function, as a formula in x and y—like ($3.0678x^2 + 20.08x\ 7.34y$) or so—which can be evaluated to give us the elevation at any given (x, y) location.

Both of these approaches have their drawbacks. The first suffers from the fact that we will never be able to store all elevation values for all locations; after all, there are infinitely many locations. The second approach suffers from the fact that we do not know just what this function should look like, and that it would be extremely difficult to derive such a function for larger areas.

We can use an interpolation function that allows us to infer a reasonable elevation value for locations that are not stored. A simple and commonly used interpolation function takesthe elevation value of the nearest location that is stored.
Interpolation is made possible by a principle called spatial autocorrelation. This is a fundamental principle which refers to the fact that locations that are closer together are more likely to have similar values than locations that are far apart—commonly referred to as 'Tobler's first law of Geography'. An obvious example of a phenomenon which exhibits this property is sea-surface temperature, where one might expect a high degree of correlation between measures taken close together.
To represent fields , tessellation approach is used and to represent objects , vector approach is used.

2.3.1 Regular Tessellations

A tessellation (or tiling) is a partitioning of space into mutually exclusive cells that together make up the complete study space.
With each cell, some (thematic) value is associated to characterize that part of space.
In a regular tessellation, the cells are the same shape and size. The simplest example is a rectangular raster of unit squares, represented in a computer in the 2D case as an array of n m elements
Three regular tessellation types are illustrated below

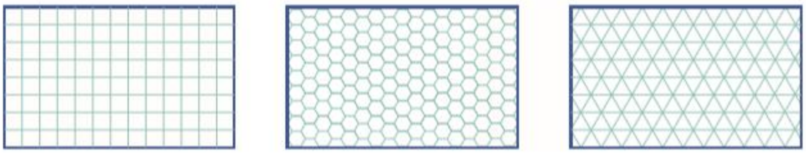


Figure 2.5: The three most common regular tessellation types: square cells, hexagonal cells, and triangular cells.

In all regular tessellations, the cells are of the same shape and size, and the field attribute value assigned to a cell is associated with the entire area occupied by the cell. The square cell tessellation is by far the most commonly used, mainly because georeferencing a cell is so straightforward. These tessellations are known under various names in different GIS packages, but most frequently as rasters.

A raster is a set of regularly spaced (and contiguous) cells with associated (field) values. The associated values represent cell values, not point values. This means that the value for a cell is assumed to be valid for all locations within the cell.

The size of the area that a single raster cell represents is called the raster's resolution. Sometimes, the word grid is also used, but strictly speaking, a grid refers to values at the intersections of a network of regularly spaced horizontal and perpendicular lines (see Figure 2.6). Grids are often used for discrete measurements that occur at regular intervals. Grid values are often considered synonymous with raster cells, although they are not.

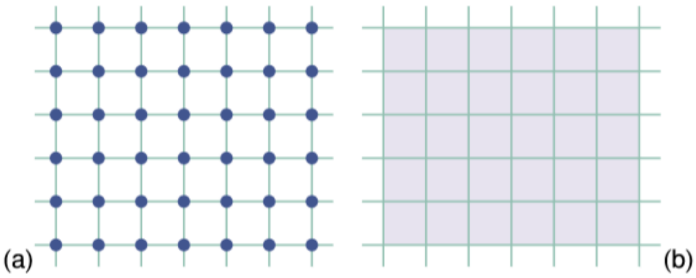


Figure 2.6: A grid (a) is a collection of regularly spaced (field) values, while a raster (b) is composed of cells. The associated values with each grid point or raster cell are not illustrated.

There are some issues related to cell-based partitioning of the study space. The field value of a cell can be interpreted as one for the complete tessellation cell, in which case the field is discrete, not continuous or even differentiable. Some convention is needed to state which value prevails on cell boundaries; with square cells, this convention often says that lower and left boundaries belong to the cell. To improve on this continuity issue, we can do two things:

- Make the cell size smaller, so as to make the 'continuity gaps' between the cells smaller, and/or
- Assume that a cell value only represents elevation for one specific loca- tion in the cell, and to provide a good interpolation function for all other locations that has the continuity characteristic.

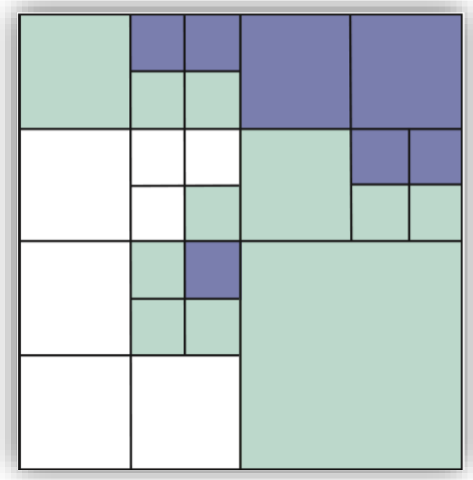
An important advantage of regular tessellations is that we know how they parti- tion space, and we can make our computations specific to this partitioning. This leads to fast algorithms. An obvious disadvantage is that they are not adaptive to the spatial phenomenon we want to represent. The cell boundaries are both artificial and fixed: they may or may not coincide with the boundaries of the phenomena of interest.

2.3.2 Irregular Tessellations

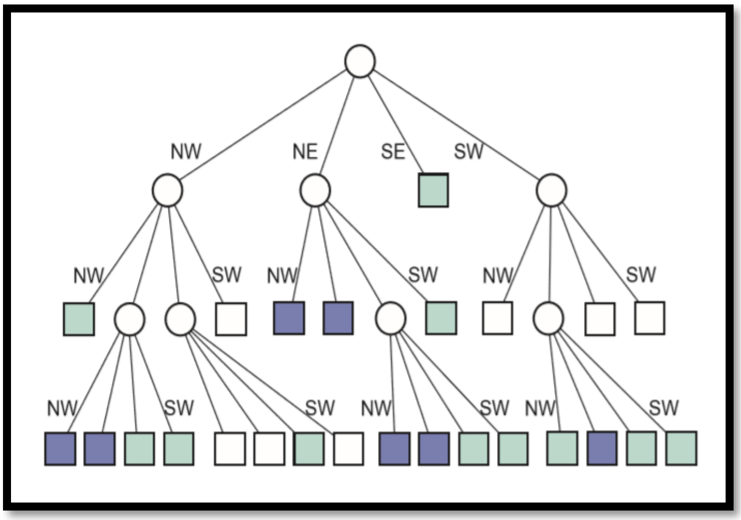
These are partitions of space into mutually disjoint cells, but now the cells may vary in size and shape, allowing them to adapt to the spatial phenomena that they represent.

The structure used to represent irregular tessellations is a region quadtree. It is based on a regular tessellation of square cells, but takes advantage of cases where neighbouring cells have the same field value, so that they can together be represented as one bigger cell.

A simple illustration is provided in Figure 2.7. It shows a small 8X8 raster with three possible field values: white, green and blue. The quadtree that represents this raster is constructed by repeatedly splitting up the area into four quadrants, which are called NW, NE, SE, SW for obvious reasons. This procedure stops when all the cells in a quadrant have the same field value. The procedure produces an upside-down, tree-like structure, known as a quadtree. In main memory, the nodes of a quadtree (both circles and squares in the figure below) are represented as records. The links between them are pointers, a programming technique to address (i.e. to point to) other records.



8 X 8 three valued raster (3 colours used)



Region Quadtree

Quadtrees are adaptive because they apply the spatial autocorrelation principle,
i.e. that locations that are near in space are likely to have similar field values. When a conglomerate of cells has the same value, they are represented together in the quadtree, provided boundaries coincide with the predefined quadrant boundaries. This is why we can also state that a quadtree provides a nested tessellation: quadrants are only split if they have two or more values. The square nodes at the same level represent equal area sizes, allowing quick computation of the area associated with some field value. The top node of the tree represents the complete raster.

2.3.3 Vector representations

Tessellations do not explicitly store georeferences of the phenomena they represent. Instead, they provide a georeference of the lower left corner of the raster, for instance, plus an indicator of the raster's resolution, thereby implicitly providing georeferences for all cells in the raster. In vector representations, an attempt is made to explicitly associate georeferences with the geographic phenomena. A georeference is a coordinate pair from some geographic space, and is also known as a vector. TIN is a representation for geographic fields.

Triangulated Irregular Networks

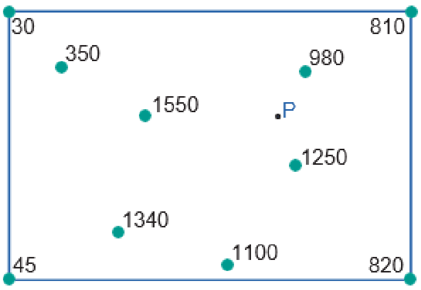
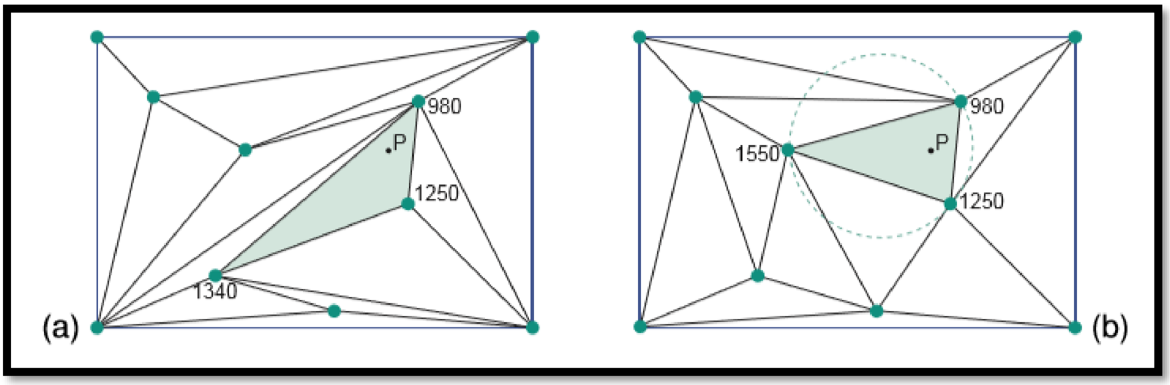


Figure 2.8: Input locations and their (elevation) values for a TIN construction. The location *P* is an arbitrary location that has no associated elevation measurement.

A commonly used data structure in GIS software is the triangulated irregular network, or TIN. It is one of the standard implementation techniques for digital terrain models, but it can be used to represent any continuous field. The principles behind a TIN are simple. It is built from a set of locations for which we have a measurement, for instance an elevation. The locations can be arbitrarily scattered in space, and are usually not on a nice regular grid. Any location together with its elevation value can be viewed as a point in three-dimensional space. This is illustrated in Figure 2.8. From these 3D points, we can construct an irregular tessellation made of triangles. Two such tessellations are illustrated in Figure 2.9.

Figure 2.9: Two triangulations based on the input locations of Figure 2.8. (a) one with many 'stretched' triangles; (b) the triangles are more equilateral; this is a *Delaunay triangulation*.



If we restrict the use of a plane to the area between its three anchor points, we ob- tain a triangular tessellation of the complete study space. Unfortunately, there are many different tessellations for a given input set of anchor points, as Figure 2.9 demonstrates with two of them. Some tessellations are better than others, in the sense that they make smaller errors of elevation approximation. For instance, if we base our elevation computation for location P on the left hand shaded tri- angle, we will get another value than from the right hand shaded triangle. The second will provide a better approximation because the average distance from P to the three triangle anchors is smaller.

The triangulation of Figure 2.9(b) happens to be a Delaunay triangulation, which in a sense is an optimal triangulation. It has two important properties. The first is that the triangles are as equilateral ('equal-sided') as they can be, given the set of anchor points. The second property is that for each triangle, the circumcircle through its three anchor points does not contain any other anchor point. One such circumcircle is depicted on the right of Figure 2.9(b).

A TIN clearly is a vector representation: each anchor point has a stored georeference. Yet, we might also call it an irregular tessellation, as the chosen triangulation provides a partitioning of the entire study space.

Point representations

Points are defined as single coordinate pairs (x, y) when we work in 2D, or co- ordinate triplets (x, y, z) when we work in 3D.

Points are used to represent objects that are best described as shape- and size- less, one-dimensional features. Whether this is the case really depends on the purposes of the spatial application and also on the spatial extent of the objects compared to the scale applied in the application. For a tourist city map, a park will not usually be considered a point feature, but perhaps a museum will, and certainly a public phone booth might be represented as a point. Besides the georeference, usually extra data is stored for each point object. This so-called attribute or thematic data, can capture anything that is considered rel- evant about the object.

Line representations

Line data are used to represent one-dimensional objects such as roads, railroads, canals, rivers and power lines. Again, there is an issue of relevance for the appli- cation and the scale that the application requires. For the example application of mapping tourist information, bus, subway and streetcar routes are likely to be relevant line features. Some cadastral systems, on the other hand, may consider roads to be two-dimensional features, i.e. having a width as well.

Above, we discussed the notion that arbitrary, continuous curvilinear features are as equally difficult to represent as continuous fields. GISs therefore approximate such features (finitely!) as lists of nodes. The two end nodes and zero or more internal nodes or vertices define a line. Other terms for 'line' that are commonly used in some GISs are polyline, arc or edge. A node or vertex is like a point (as discussed above) but it only serves to define the line, and provide shape in order to obtain a better approximation of the actual feature.

The straight parts of a line between two consecutive vertices or end nodes are called line segments. Many GISs store a line as a simple sequence of coordinates of its end nodes and vertices, assuming that all its segments are straight. This is usually good enough, as cases in which a single straight line segment is con- sidered an unsatisfactory representation can be dealt with by using multiple (smaller) line segments instead of only one.

Figure 2.10: A line is de- fined by its two end nodes and zero or more internal nodes, also known as ver- tices. This line represen- tation has three vertices, and therefore four line seg- ments.

Area representations

When area objects are stored using a vector approach, the usual technique is to apply a boundary model. This means that each area feature is represented by some arc/node structure that determines a polygon as the area's boundary. Common sense dictates that area features of the same kind are best stored in a single data layer, represented by mutually non-overlapping polygons. In essence, what we then get is an application-determined (i.e. adaptive) partition of space.

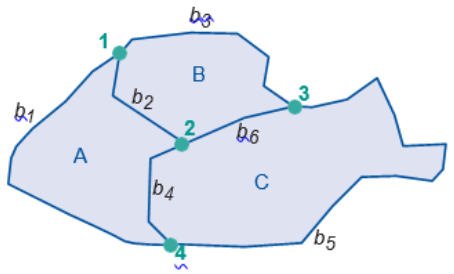


Figure 2.11: Areas as they are represented by their boundaries. Each boundary is a cyclic se- quence of line features; each line—as before—is a sequence of two end nodes, with in between, zero or more vertices.

A simple but naive representation of area features would be to list for each polygon simply the list of lines that describes its boundary. Each line in the list would, as before, be a sequence that starts with a node and ends with one, possibly with vertices in between. But this is far from optimal. To understand why this is the case, take a closer look at the shared boundary between the bottom left and right polygons in Figure 2.11. The line that makes up the boundary be- tween them is the same, which means that using the above representation the line would be stored twice, namely once for each polygon. This is a form of data duplication—known as data redundancy—which is (at least in theory,) unneces- sary, although it remains a feature of some systems.

line	*from*	*to*	*left*	*right*	*vertexlist*
b_1	4	1	W	A	. . .
b_2	1	2	B	A	. . .
b_3	1	3	W	B	. . .
b_4	2	4	C	A	. . .
b_5	3	4	W	C	. . .
b_6	3	2	C	B	. . .

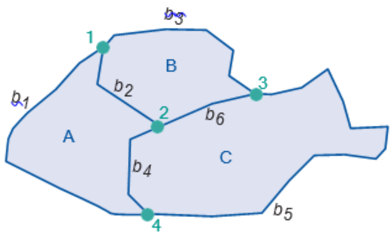
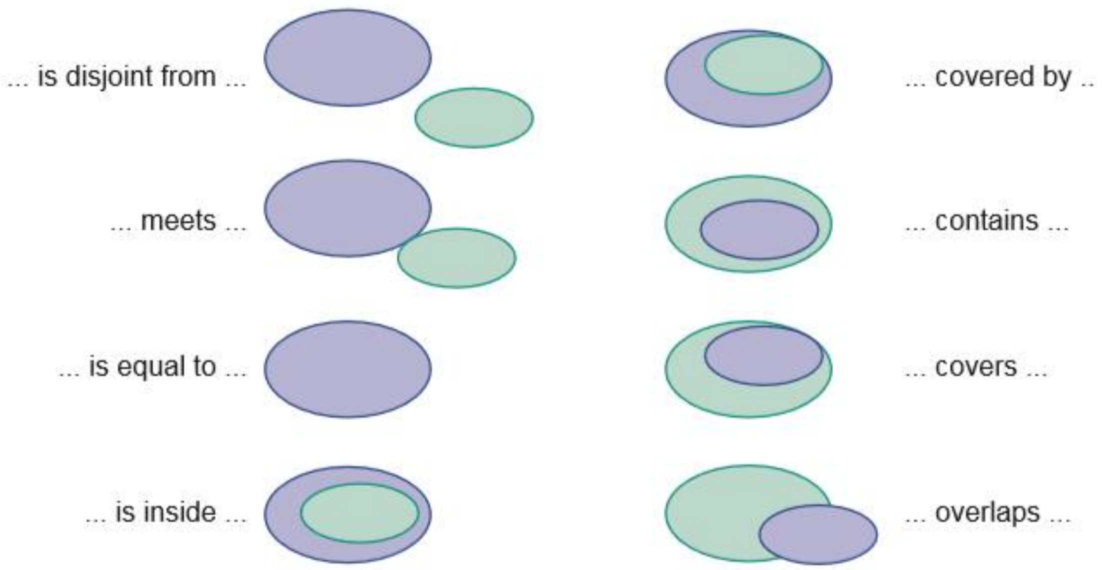


Figure 2.12: A simple boundary model for the polygons *A*, *B* and *C*. For each arc, we store the start and end node (as well as a vertex list, but these have been omitted from the table), its left and right polygon. The 'poly- gon' *W* denotes the out- side world polygon.

Topology of two dimensions

- 2 properties:- interior and boundary
- Interior of a region R – largest set of points for which we construct a disk-like environment around it that falls completely inside R.
- Boundary of a region R – set of points that belong to R; but not to the interior to R
- Eight spatial relationships



 Source:- Principles of Geographic Information Systems,Otto Huisman and Rolf A.de By

2.3.5 Scale and resolution

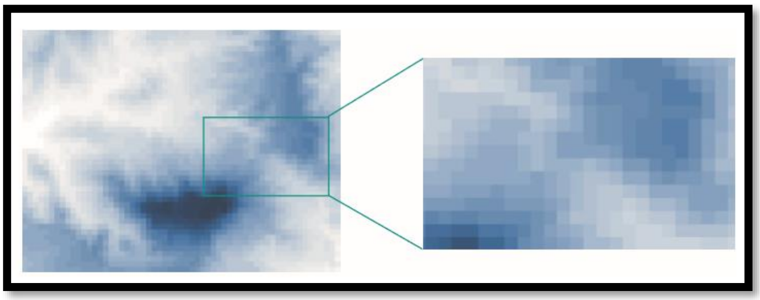
Map scale can be defined as the ratio between the distance on a paper map and the distance of the same stretch in the terrain. A 1:50,000 scale map means that 1 cm on the map represents 50,000 cm, i.e. 500 m, in the terrain.

'Large-scale' means that the ratio is large, so typically it means there is much detail, as in a 1:1,000 paper map. 'Small-scale' in contrast means a small ratio, hence less detail, as in a 1:2,500,000 paper map. When applied to spatial data, the term resolution is commonly associated with the cell width of the tessellation applied.

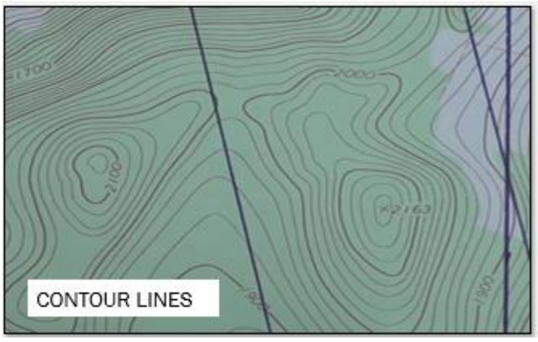
Digital spatial data, as stored in a GIS, is essentially without scale: scale is a ratio notion associated with visual output, like a map or on-screen display, not with the data that was used to produce the map.

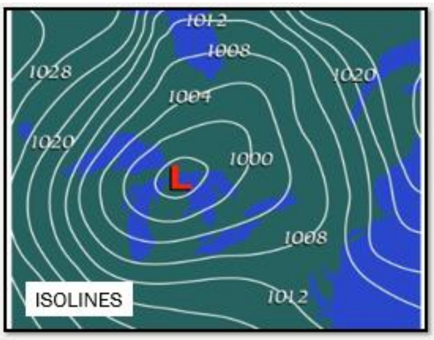
2.3.6 Representation of geographic fields

- Can be represented through tessellation (rasters),TIN or vector
- Raster representation is shown below



- Here, the raster represents a continuous field like elevation.
- Different shades of blue indicate different elevation values; darker blues – higher elevation.
- Each cell has a value; shown in a shade of blue ; so that it is more legible.
- Raster is a list of m X n values
- Vectors are represented as follows :-



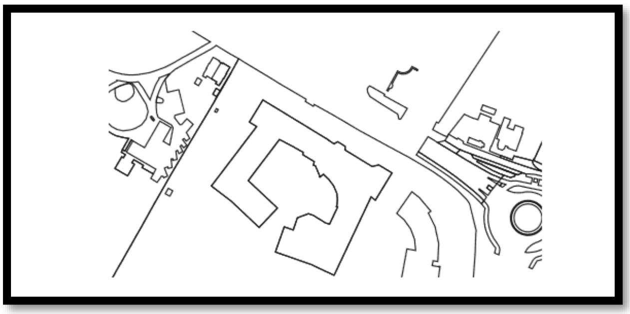


- isolines (a linear feature that connects the points with equal field value) and
- contour lines (line that joins points of equal elevation)

 Source:- Principles of Geographic Information Systems,Otto Huisman and Rolf A.de By

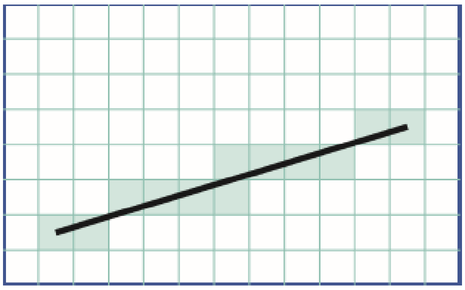
2.3.7 Representation of geographic objects (Using Tessellations)

Vectors are represented as follows:-



Various objects like buildings, bikes, road lanes are represented as area objects in a vector representation

Rasters are represented as follows:-



An actual straight line and its representation in a raster

2.4 Temporal Dimension

Examples of the kinds of questions involving time include:

- Where and when did something happen?
- How fast did this change occur?
- In which order did the changes happen?

Here we will present a brief examination of different 'concepts' of time.

1. **Discrete and continuous time:** Time can be measured along a discrete or continuous scale. Discrete time is composed of discrete elements (seconds, minutes, hours, days, months, or years). In continuous time, no such discrete elements exist, and for any two different points in time, there is always another point in between. We can also structure time by events (points in time) or periods (time intervals). When we represent time periods by a start and end event, we can derive temporal relationships between events and periods such as 'before', 'overlap', and 'after'.

2. **Valid time and transaction time:** Valid time (or world time) is the time when an event really happened, or a string of events took place. Transaction time (or database time) is the time when the event was stored in the database or GIS. Observe that the time at which we store something in the data- base/GIS typically is (much) later than when the related event took place.

3. **Linear, branching and cyclic time:** Time can be considered to be linear, extending from the past to the present ('now'), and into the future. This view gives a single time line. For

some types of temporal analysis, branching time—in which different time lines from a certain point in time onwards are possible—and cyclic time—in which repeating cycles such as seasons or days of a week are recognized, make more sense and can be useful.

4. **Time granularity:** When measuring time, we speak of granularity as the precision of a time value in a GIS or database (e.g. year, month, day, sec- ond, etc.). Different applications may obviously require different granu- larity. In cadastral applications, time granularity might well be a day, as the law requires deeds to be date-marked; in geological mapping applica- tions, time granularity is more likely in the order of thousands or millions of years.

5. **Absolute and relative time:** Time can be represented as absolute or relative. Absolute time marks a point on the time line where events happen (e.g. '6 July 1999 at 11:15 p.m.'). Relative time is indicated relative to other points in time (e.g. 'yesterday', 'last year', 'tomorrow', which are all relative to 'now', or 'two weeks later', which is relative to some other arbitrary point in time.).

****