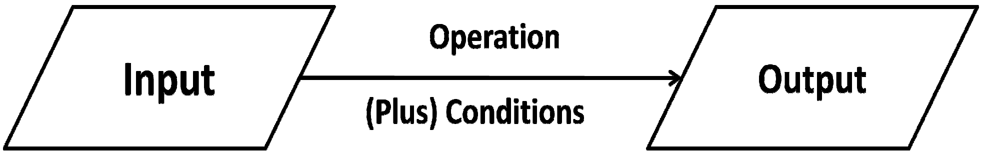
❖ Introduction :-

- The analysis subsystem is the **heart of the GIS.**
- It is also the most **abused subsystem of the GIS**.
- Much of the abuse of the analytical subsystem result from a **lack of understanding the nature of the spatial data** contain in the system.
- **GIS is the incomplete set of spatial analytical tools**. In many cases the user is obliged to combine GIS tools with other computational tools in order to accomplish his modeling task.
- Most GIS packages are heavily oriented toward **ability to overlay two or more maps to analyze** corresponding patterns or to operate on digital data obtain to remote sensing systems.
- In this chapter the framework of geographic ideas is limited to those that **deals fundamentally and explicitly with maps**.

❖ Preliminary Notes About Flowcharting :-

- Many software packages now incorporate some form of flowcharting utility for example, ESRI's Model Builder.
- **Advantage of Flowcharting**:
  - ✓ The GIS software vendors are placing on flowcharting, and their reorganization of its ability to systemize our thinking and documenting our procedure, we should being thinking about them immediately.



- This is the general form of most GIS flowcharts.
- Although the symbols will change, virtually every GIS model begins with **input** perform some **actions** and produce some **output**.

1. **Implementation Flowchart** :-
   - As you can see in diagram, we have an input (typically maps), an operation (a search, a measurement, or other GIS tasks), and an output of some kind. For the sake of simplicity it is assumed that the output is map in figure.
   - Notice how the flowchart shows both the objects operation and the flow direction. This type of flowchart is called an *Implementation Flowchart* **because it shows how the model actually works**.

2. **Formulation Flowchart** :-
   - As our modeling expertise grows, we will need to create a flowchart that flows in the other direction, so we can move from expected outcomes to the basic data elements on which the model is built.
   - This type of flowchart is called *Formulation Flowchart* **because it is use to design the modeling task.**

❖ GIS Data Query :-

- Each digital data base typically contain many maps, each processing a separate theme.
- For focusing the field study, you are likely to be trying to isolated numerous features for study.
- **You will often want to know:**
  - Which of the selected features occur most often?
  - How often they are occur?
  - Where they are located?

- **Identifying the absolute location of point data in Raster GIS:**
  - Though you will forgive a pun (શ્લેષ), you can easily identify the grid cell locations, most often by using an on screen courser device that allow you to point to each grid cell individually.
  - This usually result in a readout giving the row and column numbers, as well as the attribute code for the selected item.

- **In Raster GIS,**
  - ***lines*** are merely collection of grid cells that touch one another either along each side or diagonally.
  - ***Polygons/Region*** are group of grid cells that are connected in much the same way, or share attribute values.
  - To identify as entities you must identify them by attribute.
  - Tabulation of the result will again show **the amount of each category** and the **percentage**.

- **Process of finding point in Vector GIS:**
  - This process is performed by displaying all points in the cartographic portion of the database.
  - If you need to find a point of particular kind (e.g. Telephone Pole, Bridge Nesting Site), you will need to access the attribute database and its tables.
  - Most often you will perform a search that identify all table locations that shows the appropriate code for the object you wish to see.
  - Because attributes are linked to the entities, you will be able to selectively isolate these specific type of point object.
  - Because vector data contain explicit spatial information, you can easily obtain the exact coordinate in projected space.
  - These can be produce as tabular output or viewed directly on the screen by means of a pointing operation.
  - Because points are also used to represent line and lines are also used to represent polygon in vector data model, you can also obtain entity information about lines and polygons by selecting them individually.
  - **For Example**, you can identify all points in the coverage by simply displaying only points or you can obtain a list by accessing attribute table specific for point.

- This is true for lines of all kinds as well as for polygon.
- *To identify the absolute location for any two dimensional object*, you will need multiple sets of coordinate pair for each of the point that are used to identify all the line segments with which they are defined.

❖ Locating and Identifying Spatial Objects :-

- *GIS can be subjected to find or locate query*.
- **Question:** Why we need to be able to find and locate object in our GIS database?
- **Answer:** Because it is a fundamental process comparable to traditional map reading and because of its relation is more complex calculation.

1. **In first case** :
   - ✓ It is important to be able to **isolate**, **count** and **locate** object because these activities gives us an understanding of the overall complexity of our maps.
   - ✓ The objects in this map represent features on the ground.
   - ✓ One ***primary purpose of map*** is to display these objects, to facilitate the identification of the spatial relationships between and among themselves as well as to other objects on the landscapes.
   - ✓ **For Example**, it is important to know whether there many houses or just a few in a particular region. Or the number of road network might provide information that is impotent for routing sales staff or for truck traffic.

2. **In Second case :**
   - ✓ A quantitative measure of the amount of these objects will allow us to make direct analytical comparison to other coverage or other variable in same variable.
   - ✓ ***Comparison*** to other features on the landscape might be used to determine causes for distributional patterns, or at least to suggest the strong spatial relationship among cartographic variables.
   - ✓ **For Example**, information of number of houses may be useful for examine their relationship to the amount of land available for new houses.
   - ✓ Simple enumeration of object and their locations also lets you examine their relationships to more prominent (મુખ્ય) objects in the same map.
   - ✓ Most important aspect of being able to find and locate objects on the map is the ability to make further **measurement** and **comparison**.

❖ Defining Spatial Characteristics :-

- **Introduction** :-
  - ✓ Simply being able to find a point, line and area entity in a map is of little value.
  - ✓ Most of the object analysis find, count and locate in coordinate space are selected not on the basis of entity type alone but rather on what they represent in the real world.
  - ✓ **For the same reason, we most often search for objects, count their numbers, and note their locations on the basis of attributes.**

1. Point Attributes :-

- ✓ Point objects, like all objects, <u>differ not only in their locations but also in their attribute characteristics</u>.
- ✓ These differences provide us with often related spatial patterns of each group of objects.
- ✓ *<u>Point objects can differ by nominal type</u>*: For Example, we can isolate patients in the region that share diseases related to specific organ groups.
- ✓ *<u>These types may also be ranked within types</u>*: For Example, digestive diseases could be ranked as mild (કોમળ), moderate (સાધારણ), and severe (સખત).
- ✓ This would allow the user to identify patterns of distribution of such phenomena and to determine whether there is some reasonable cause for this pattern.
- ✓ If we can **<u>classify point object</u>** in **nominal** and **ordinal categories**, we can also separate out these objects by value in interval or ratio scales as well as by types.
- ✓ **For Example**, houses could be selected on the basis of market value; houses that are bellow 50,000, between 50,001 and 100,000 and so on.
- ✓ ***<u>These values would have to be store in database to permit retrieval of corresponding points</u>***.
- ✓ *<u>Quantitative measurements will allow us to select a wide range of groups or classes of each objects, depending on our needs</u>*.

- ✓ **<u>What is isolation</u>?**
    - o ***<u>Functioning GIS</u>*** is being able to define each group or category separately and tabulate the results of a search.
    - o *In other words, we need to be able to isolate these groups of point objects from those that are not important to our output or analysis objectives.*
    - o Because GIS works in **holistic cartography** (સર્વગ્રાહી નકશા દોરવાની વિદ્યા) program, we should be able to retrieve all the row data and then group the result in any fashion that suits out purpose.

- ✓ **<u>What is Comparison</u>?**
    - o *Beyond simply isolating categories of data, the GIS must allow us to locate each item individually for each class of data and **compare** it with other of its own kind.*
    - o That is, we need to be able to show the spatial relationship between individual items of a selected point object in a class of object so that we can later perform analytical operations to quantify these relationships.
    - o **For Example**, if we have isolate the houses that are coasting under 50,000 further comparison is:
    - o Located within a given distance of one another?
    - o Are they distributed evenly or randomly?
    - o Are they clustered or dispersed in their distribution?

2. Line Attributes :-

- ✓ *Line objects are **one-dimensional** entities defined by two or more points with corresponding coordinate pairs*.
- ✓ **Where to define points on a Line Object?**
  - Either the beginning or the ending of a line.
  - Some change in attribute along the line.
  - To display some place on the line attribute.
- ✓ **Examples of line item types** include railroads, streets, fault lines, fencerows, or streams.
- ✓ **How to Identify a Line?**
  - To identify the lines, it is necessary to identify all the coordinate pairs that make up the line in vector, or all the grid cell column and row values in the raster. This adds three other factors: Length, Azimuthal Direction and its Shape.

- ✓ **What is isolation?**

  - o Like point objects, lines should be separable based on ordinal rankings or some measure of magnitude.
  - o Highway types such as single lane, double lane, three lane, and interstate freeway are examples of line objects organized by ordinal ranks.

- ✓ **What is Comparison?**

  - o These distinctly different road types can be compared only across the single spectrum of highway types; quantitative comparison to other non-highway line features is not permitted.
  - o In some cases, a single line may experience a change in attribute type, rank, or magnitude along its overall length. **For example**, a road may change from single to double lane.
  - o Another attribute characteristic involves not just the attributes of the line itself but also a comparison of what falls on either side of line attribute.
  - o **For Example**, we may want to define a hedgerow or fencerow not on the basis of type of vegetation but rather a comparison of the land covers that fall on either side of the Line entity.
  - o A search could be performed to identify each attribute for each bordering polygon. That's why, however, the relationship between the line and its neighboring polygons are explicitly encoded into the databases during input.

3. Area Attributes :-

- ✓ As before, each of these attributes needs to be explicitly stored in the database, whether as grid cell attributes or as vector polygon attributes. Isolation and retrieval are performed in exactly the same manner as for points and lines.
- ✓ Among the more useful attributes for polygon entities would be a measure of their shapes.

- ✓ Euclidean in that it could be measured as a deviation from some known geometric shape with predefined properties, such as a circle or a square.
- ✓ Another group of shape measures uses a set of mathematics called ***fractal geometry***, **in which the irregularity of the outside of a polygon is measured**. Related to shape is a measure of a polygon's elongation, or the **ratio between its long and short axes**.
- ✓ Another somewhat more commonly applied measurable attribute of polygons is their size.
  - ***In raster***, size is determined by counting the number of grid cells of a given category. Often, a reclassification process will be necessary to isolate these polygons, and then their area can more readily be found by a simple count of grid cells.
  - Measures of size in raster are not going to be particularly accurate because of the quantized nature of the grid cell.
  - ***In vector***, the perimeter of the polygonal lines is easily calculated by adding the line lengths, whereas the area is calculated much as would be done manually, by a series of length-times-width calculation for portion of each polygon.
  - The more complex the polygon, the more calculation must be performed.
- ✓ ***Two other measures regarding areas need to be mentioned as well:***
  - Contiguity
  - Homogeneity

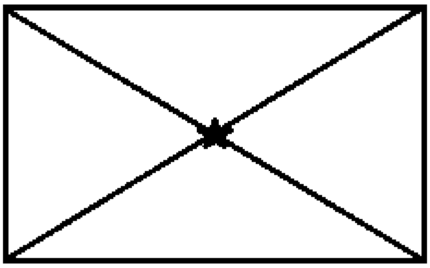
1. **Contiguity :-**
   - The first is **contiguity,** a measure of the wholeness or amount of perforation of a polygon. A large polygon that contains many holes (small polygons contained within the larger polygon) demonstrates much less contiguity than one with only a few holes or none at all.
2. **Homogeneity :-**
   - The second additional measurable attribute that could prove useful for polygonal features is the **homogeneity** of an area, **not necessarily defined as a single polygon**. Homogeneity is a measure of how much area of given portion of a map is directly in contact with polygonal features sharing the same attributes.

❖ Defining Spatial Characteristics :-

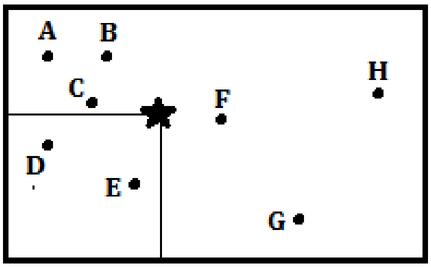
Thus far, we have worked with points, lines, and areas, based either on readily available descriptive characteristics or on attributes that are measurable through, sometimes simple, sometimes complex methods. All of these entities, however, have certain attributes that set them apart from the rest. Some attributes are a result of the methods of encoding-for example, nodes that are encoded during the digitizing process. Others may have to be determined-for example, centroids indicating the center of an area. We call these "higher-level objects" for lack of a better term, simply because of their uniqueness and usefulness for later GIS analysis. We will separate these higher-level objects into point, lines, and areas and examine each individually.

1. Higher-Level Point Objects :-

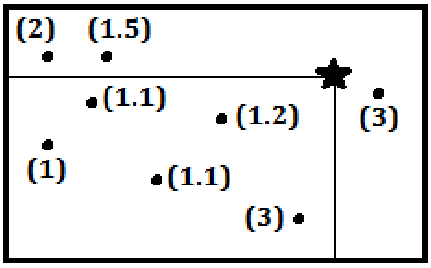
Two primary types of higher-level point objects are **Centroids** and **Nodes**. A centroid is most commonly defined as the point that occurs at the exact geographic center of an area or polygon. Its calculation is simple for simple polygonal shapes such as rectangles; when the polygons become quite complex, the complexity of the calculations needed to perform the task increases proportionality.



SIMPLE CENTROID



CENTER-OF-GRAVITY / MEAN CENTER



WEIGHTED MEAN CENTER

- Raster GIS is not well suited for this procedure.
- In, many cases, even vector GIS does not include this calculation as a stand-alone function. Simple or geographic centroids in vectors are calculated by a rule called the **trapezoidal rule,** which separates the polygon into a number of overlapping trapezoids.
- Then each trapezoidal centroid, or central coordinate, is calculated and its weighted average calculated.
- ***Centroids can also be placed at the center of a distribution of some phenomena, instead of at the absolute geographic center of a polygon***.
- The centroid would act as a point location, as if the data were actually calculated at this point. Then, through interpolation, the isolines or surfaces could be calculated on the basis of these point locations.

- **CENTER OF GRAVITY :-**

Point	X	Y
A	0.5	4.5
B	1.0	4.5
C	1.0	3.5
D	0.5	2.5
E	2.0	2.0
F	3.0	3.0
G	5.0	1.0
H	7.0	3.5
Total	**20.0**	**24.5**
Mean Center	**2.5**	**3.0625**

- Although we could still use the centroid of the county as the estimate of its population center, it would be more accurate if we placed it closer to the center of the distribution. This location, called the **mean center** or the **center of gravity**, requires us to separately average the x and the y coordinates for all points in the map.
- The final result would be a single pair of values representing the central point of the distribution of points.

$$\text{Mean Center X} = \text{X value for Center Point} = \frac{\textit{Sum of All X values}}{\textit{No.of Points}} = \frac{20.0}{8.0} = \mathbf{2.5}$$

$$\text{Mean Center Y} = \text{Y value for Center Point} = \frac{\textit{Sum of All Y values}}{\textit{No.of Points}} = \frac{24.5}{8.0} = \mathbf{3.0625}$$

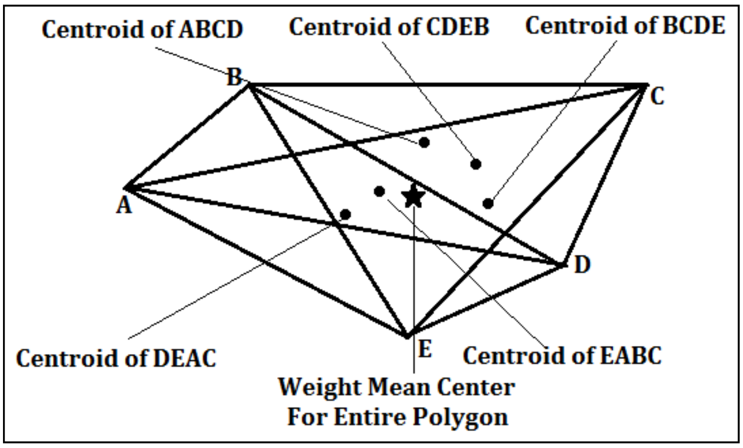
- **<u>WEIGHTED MEAN CENTER</u> :-**

Point	X	Y	f	fX	fY
A	0.5	4.5	2.0	2.25	9.0
B	1.0	4.5	1.5	1.5	6.75
C	1.0	3.5	1.1	1.1	3.85
D	0.5	2.5	1.0	0.5	2.5
E	2.0	2.0	1.1	2.2	2.0
F	3.0	3.0	1.2	3.6	3.6
G	5.0	1.0	3.0	15.0	3.0
H	7.0	3.5	3.0	21.0	10.5
			Total	**47.15**	**41.2**
			Weighted Mean Center	**5.894**	**5.15**

$$\text{Weighted Mean Center X} = \text{X value for Center Point} = \frac{\textit{Sum of All fX values}}{\textit{No.of Points}} = \frac{47.15}{8.0} = \mathbf{5.894}$$

$$\text{Weighted Mean Center Y} = \text{Y value for Center Point} = \frac{\textit{Sum of All fY values}}{\textit{No.of Points}} = \frac{41.2}{8.0} = \mathbf{5.15}$$

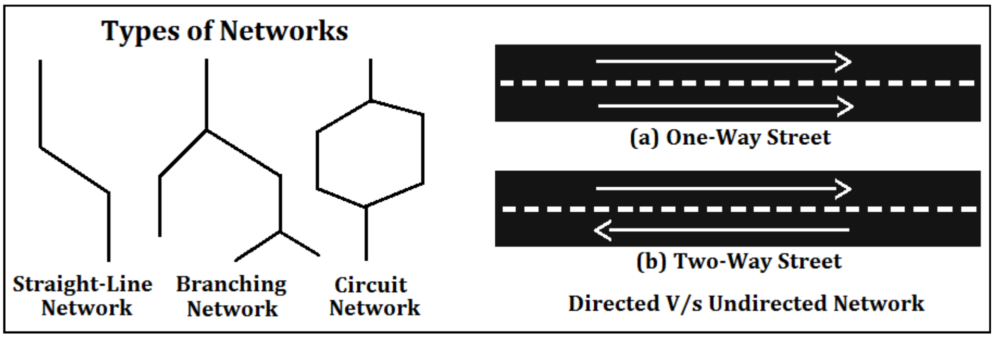
- In addition, if these points had additional weights – for example, if the point indicates both the locations and the amount of **shopping** done at each store- we could further place our center on the <u>basis of this additional weighting factor</u>. The procedure, called the **weighted mean center**, simply requires us to multiply each X and each y by a weighting factor (amount of shopping in this case).
- Diagram of the Trapezoids defined in an irregular polygon.
- <u>Each overlapping trapezoid has a calculated mean center from which the Weighted Mean Center can further be determined. And this weighted mean center is the center point for the entire polygon</u>.



- <u>Our second type of higher-level point object, **the node**</u>, was mentioned in Chapter 4. In this case the points are significant not as individuals but as specific locators along line and area entities. Nodes do not occur as specific objects in raster GIS, so there is no need to comment on their calculation in those systems. Because nodes carry with them an indicator of a change in attribute, the ability to identify them is vital to many attribute selection and manipulation procedures. Nodes are generally encoded explicitly during input and should be easily separated or identified through simple search procedures in the GIS.

2. Higher-Level Line Objects :-

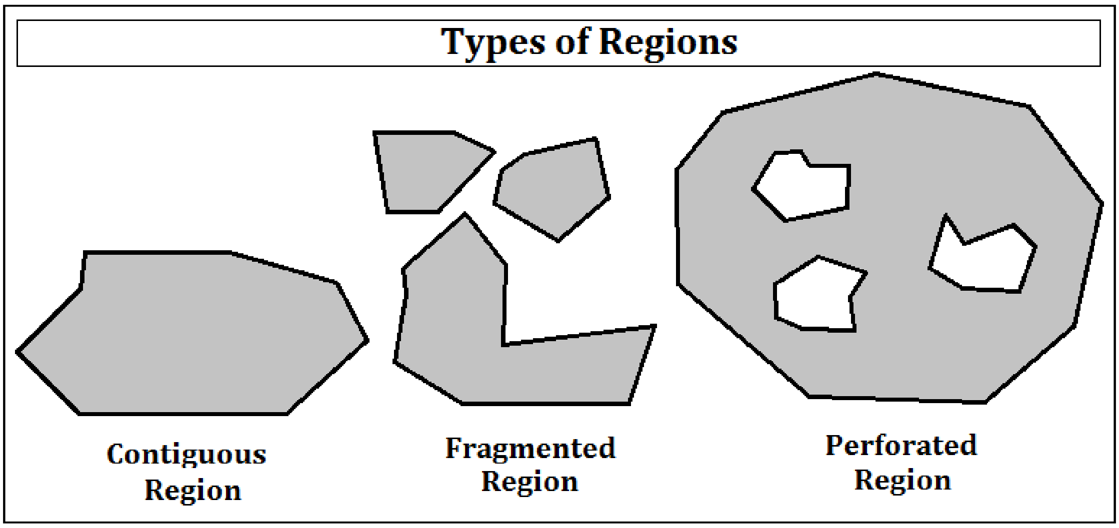
- Three <u>different types of lines</u> are particularly important and justify calling their objects higher level. We touched on the **first type** when we examined the <u>relationship between line attributes and associated attributes for the adjacent polygons</u>.
- These lines are most often called **borders**, and major change in attribute value or even collection of attributes is recognized or assumed as one moves across them.
- In other words, borders are particularly important because of their locations relative to adjacent polygons.
- **For example :-**
  - <u>Hedgerow</u> :-
    1. In our hedgerow example, borders need to have associated with them a set of attributes that easily allow portions of the map to be separated out because all the objects on one side are substantially different from all those on the other.
  - <u>National Border Between US & Canada</u> :-
    1. The line serving as the border between these two countries should allow one to identify all the states in the United States as fundamentally separate from all the provinces and territories of Canada on the north.
    2. If the line separating these two areas on the map is not provided with attribute information clearly indicating its status as an international border, you may have to force the GIS to separate its united states from Canada, rather than simply performing a search on the basis of the neighbors of the borderline.
- ***<u>Lines also can become higher-level objects</u>*** when they are located in a particular way relative to other lines rather than to polygons, as in our last example. <u>In many case, lines are not simply indications of locations of linear objects or of boundaries between polygons but rather than are connected by nodes to from</u> **networks**.
- *<u>"**Networks** can most appropriately be defined as a set of interconnected line entities whose attributes share some common theme primarily related to flow</u>."*



- ***Three major forms of networks*:**
  - **Straight lines**, as one might find in a major interstate highway.
  - **Branching trees**, as one might when looking at stream networks.
  - **Circuits,** as one might find in street patterns whose lines lead directly back to the starting point.
- In addition, all these network types can be defined as **directed** or **undirected**.
  - *<u>In a directed network</u>*, the flows are allowed to move in a single direction only.
  - *<u>In an undirected network</u>*, the flows are allowed to move in both directions.

3. Higher-Level Area Objects :-

- As with points and lines, areas can also occur as higher-level objects. The polygons themselves can be used to define regions of similar geographic attributes.
- Among the more important aspects of geographic research today, as in the past, is the definition of **regions**: "*Areas of the earth that exemplify some unity of describable attributes*."
- **For example**, political regions are defined by national boundaries, ethnic region by a similarity of origin, and biogeography regions on the basis of organismic similarities.
- Inside the GIS, these region definitions can be based on the attributes defining each polygon or set of polygons.
- **For Example**, we might be able to define region within our GIS by selecting all the polygons that show forest as the major vegetation component. This will give us a "forest" region.
- Because defining regions is a major endeavor in itself, the chances are good that simply selecting the appropriate polygons or sets of grid cells will not suffice to produce definitions. ***Instead***, we will most likely have to combine several sets of attributes from different coverage to define our regions.



- **Types of Regions :-**
  - Regions differ not only in their attributes and in the way the attributes are manipulated to define them but also in the way they are configured in space.
  - There are three types, based on spatial configuration:
    1. Contiguous Regions.
    2. Fragmented Regions.
    3. Perforated region.

- **Contiguous regions** :-

  These are wholly contained in a single polygon. Although a contiguous region is wholly contained, its attributes can be defined as homogeneous or heterogeneous but with some similarity of heterogeneous mix, as we have seen.

- **Fragmented regions** :-

  These regions share some commonality of attributes. Again, they are either a homogeneous or a heterogeneous mix containing identifiable similarities but are composed of more than one polygon from separated by intervening space that does not share the same mix of attributes. For example, one could define as a forest region a number of polygons, scattered throughout the coverage but having the same mix of trees or tree types. There are no limitations to the distance between polygons as long as the similarity of attributes is maintained.

- **Perforated regions**,

  It defining criterion for the region remains the same-homogeneity of attributes or of the attribute mix. In perforated regions, however, the uniform polygon is interspersed with smaller polygons that do not share a mix of attributes with the surrounding polygon. As such, the region is defined as the surrounding matrix, whereas the smaller internal polygons are said to be the perforations. Clearly, there is a possible relationship between perforated and fragmented regions. If the smaller polygons contained in the perforated region are found to share common attributes, they too may be considered to be a region and can easily be separated out from the background region.

--X--X--X--X--X--X--