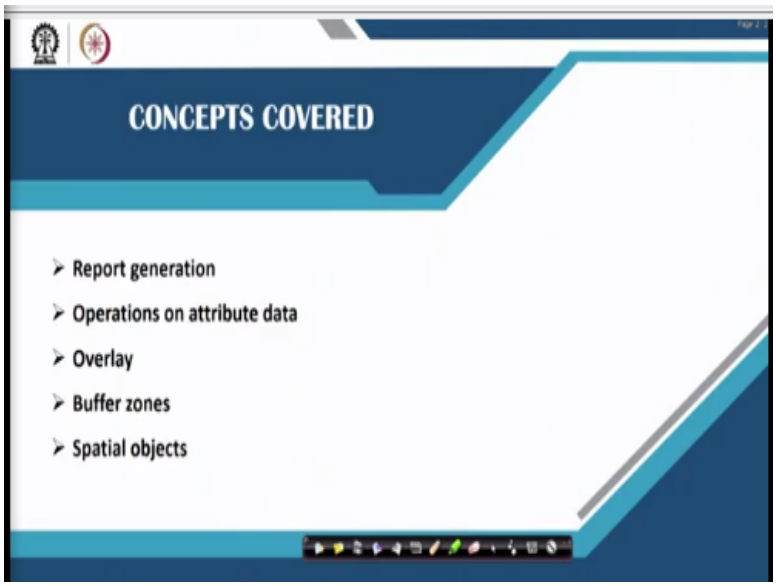
Geographic Information Systems
Prof. Bharath H Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology-Kharagpur

Module-09
Lecture-43
Basic Spatial Analysis (continued)

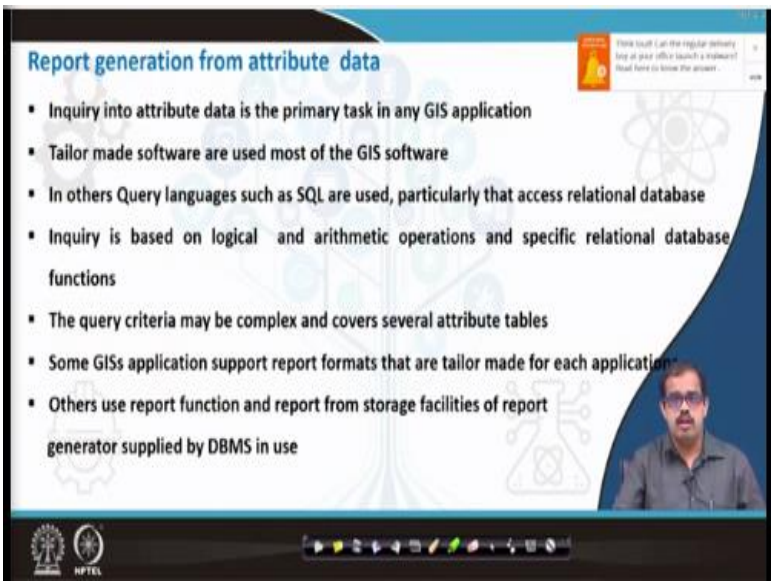
Hello namaste welcome back to the course on geographic information system. We were speaking about basics spatial analysis in the previous class. Let us consider the same thing in this class we will look into more of basics spatial analysis like overlay, buffering etc.

(Refer Slide Time: 00:41)



And proceed with more into advanced aspect. The first thing that I would like to look at how do you do a report generation. Then what are the different operations on attribute data, how do you do an operations on attribute data, some aspects of overlay, buffers and spatial objects. So if you are looking at any context overlay and buffers forms a very basic part of spatial operations, so let us look at it in detail and finally I would end this particular class with spatial objects.

(Refer Slide Time: 01:14)



And when you are looking at report generations from attribute data, when I say report generation you are actually providing an information to the user right. So when I am looking at attribute data, it is an inquiry into an attribute data is a very primary task. Whenever you are looking at GIS operations, querying any data any aspect of data is the very primary task. So this is applied as a GIS operations, when I am looking at tailor made softwares of most of the GIS or what are the GIS softwares are available.

This is a very inbuilt part of any of those softwares whether it is QGIS whether it is ArcGIS or any of those software which you are trying to use. So it is a very intrinsic part of any of the softwares ok, report generation is very essential in terms of processing the data as information. So in maybe using SQL's or any of the query languages particular that accesses relation database can be mostly used.

And when you are looking at inquiry into based on logical arithmetic operations, a specific relational database of functions are used ok. The query criteria may be complex and cover several attributes it can be on single attribute, it can be on multiple attribute, it can be on all attributes in that particular table ok. And there is no limitations for any queries for example let us say that you are whatever the attribute table that you have created for a particular vector data has 10 attributes ok.

Your query can be on only one attribute, let us say only on area, you have created a number of polygons with areas. So let us say you are only querying on area, it can be on area, it is walled, it is probably state finally the country ok, that can be your query or on all aspects that are actually there on the it can be either arithmetic either Boolean or statistical whatever operations it is, it can be query on all of these.

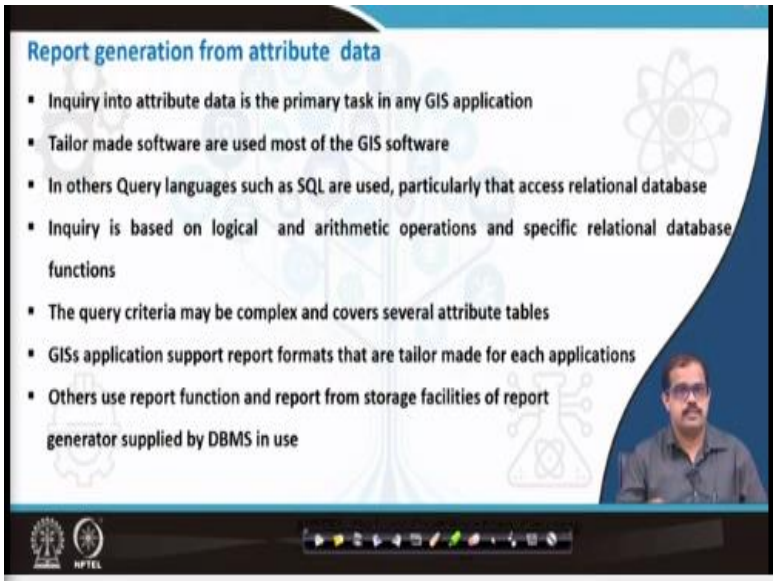
So your query is may not very limited, your query can be anything on that particular attribute. Some GIS application support report formats that are tailor made for each applications. So there are a lot of customized softwares for example, ArcGIS has a lot of flavors, when you have want to generate a report for a specific application.

If you are reporting for an urban application, your report generation is completely different when you are actually reporting it for an hydrological application. So softwares have their own way of representations. So normally, a paid software very heavily paid softwares have huge number of flavors of this representations, open Source does not have much of classification in this terms but report generations can be made ok, simple reports can be easily done.

Some GIS applications of some applications may have also support different thematic representations. So that is based on the user, you can if you understand the programming, whatever is done in a sophisticated way can also be done by you. It is just that you should understand how you code ok. Others use report function and report from storage facilities of the report generator supplied by DBMS in use.

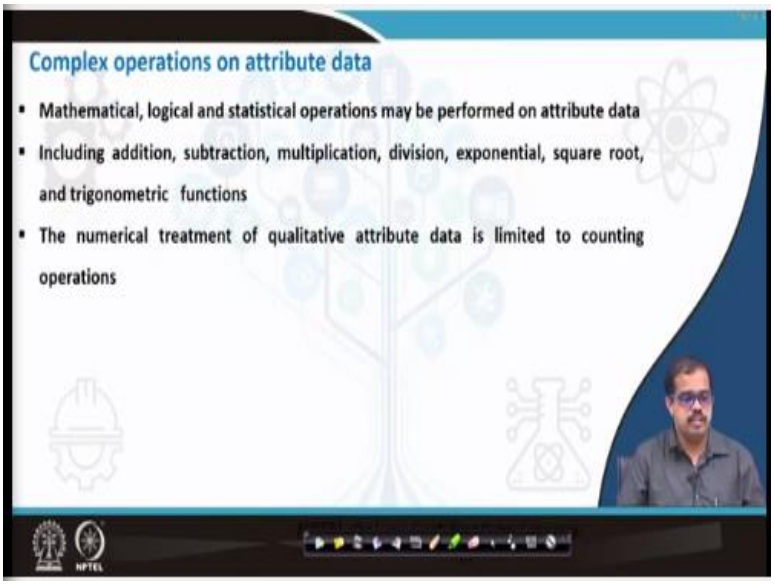
So by default there would be a report generator by a database management system, so that the same thing is used most of the open source have that specific attribute. But when you look at QGIS today, QGIS evolved a over a period of time and this evolving has bought it huge amount of flavors along with it. So now today the QGIS is also in terms of how well this particular thing can be done or how well is this more sophisticated as in terms of any of the paid softwares as you even it is you can compare it with ArcGIS the functions of ArcGIS. Most of things can be done using your open source softwares ok.

(Refer Slide Time: 05:28)



So this is about the report generations.

(Refer Slide Time: 05:32)



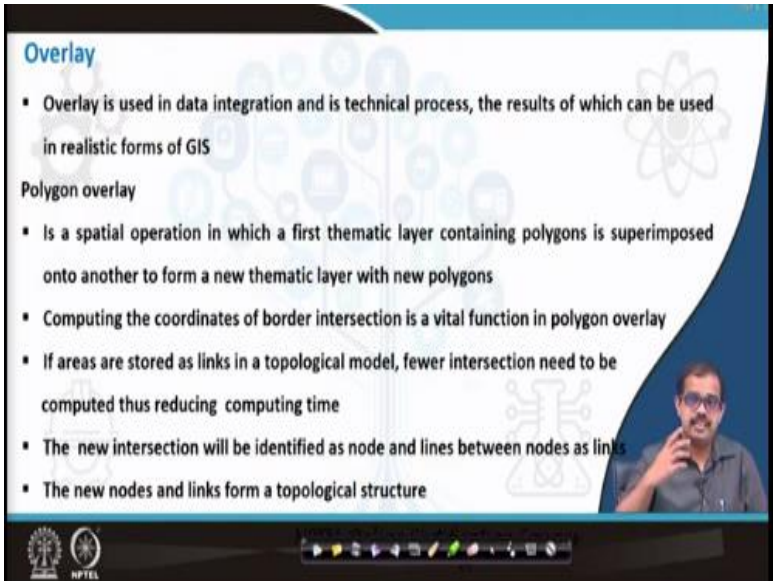
And when you are looking at complex operations on attribute data, so it can be mathematical operation, it can be logical operation, it can be a statistical operations. So this can be performed on any attribute data but only thing is that what kind of you cannot have an attribute data of where it is mentioned the name of people ok. And say that I want to do an mathematical operations it which cannot be possible.

So look at the logic behind that particular attribute, why that attribute is placed there based on that you can do any kind of applications any kind of operations. So including addition,

subtraction, multiplication, division, exponential analysis, square root trigonometric functions, so any of these applications can be done. So these are actually categorized into an complex operations.

The numerical treatment of a qualitative attribute data is limited to only counting operations, so that is the only limitations in terms of qualitative attribute data.

(Refer Slide Time: 06:31)



The next aspect is overlay, why overlaying is extremely important, why am I actually speaking it again and again. It is because overlaying is an critical part of analyzing different datasets that you may have. There may be certain datasets in a raster data, some dataset in a vector data and overlaying is actually give you the exact representation of the data or data or the real world ok. So overlaying is such data integrations and a technical process which can be used to get a realistic form of the earth surface in your GIS system ok.

If you want to really get an exclusively same representation of the real world into your GIS system, then you have to look at something like an overlay. Create different layers for different process and when you are actually comparing it or when you are visualizing the entire process. So if you can do an overlay, then it gives you an exact representation of what actually the real process is.

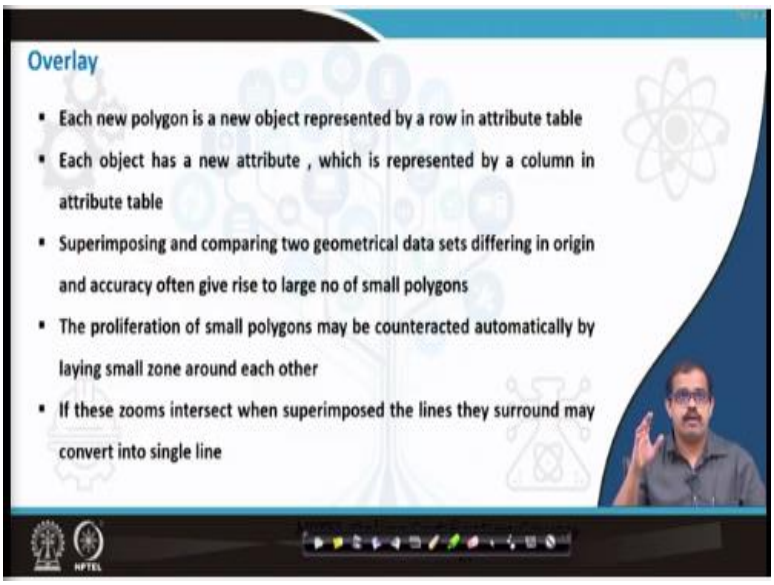
The first point of overlay is polygon overlay ok. When you are looking at this, it is let us say I have a thematic layer that is containing polygons is superimposed on another thematic layer, which is again having a number of polygons which means there is a thematic layer. Let us say I have created I am trying to analyze the urban situation or urban floods ok, now I have a thematic layer where I have digitized all the buildings fine.

So I have another thematic layers where I have digitized all the buildings that are closer to where you have water bodies ok. So now I have polygons as buildings in both the images, so now it is polygon on polygon overlay ok. So now I can look at either computing the coordinates of intersection for any of the polygon overlay, it can be looking at how this overlay and what kind of things are there.

If let us say I have a vector layer beneath it which is a drainage network, I can find out which are the drainage that are compromised for to develop to build that particular building. And because of which is there is urban flooding that is happening in a particular city right. So if the areas of whatever the polygons that you have created in each of these both images, so if it is actually linked in a topological model fewer intersections needed to be computed thus reducing a computing time.

Otherwise your computing time will be extremely high ok. Normally the new intersections are identified as nodes and lines between nodes are called as links in your polygon overlay. The new nodes and links form a topological structure ok when you are looking at the topological model it forms a topological structure.

(Refer Slide Time: 09:38)



So when you are looking at this is one form of overlay. So when your each new polygon in the new object representation by a row in attribute table. So now when we go into the attribute table, whatever you are trying to overlay both of these. So now you have 2 polygons, one I said is completely buildings, buildings that were there for few years. Other layer is the buildings that are very recently built with the lake bodies digitized.

Now when you are looking at both of these each new polygon that you superimpose, you have this something like both the images are like this. So when you are looking at this, so each object in attribute that are represented by a column, there is a new map that is formed. So whatever the objects are added to those maps have a new column at in an attribute data. So superimposing or comparing 2 geometry that sets differing in origin and accuracy may actually give you error.

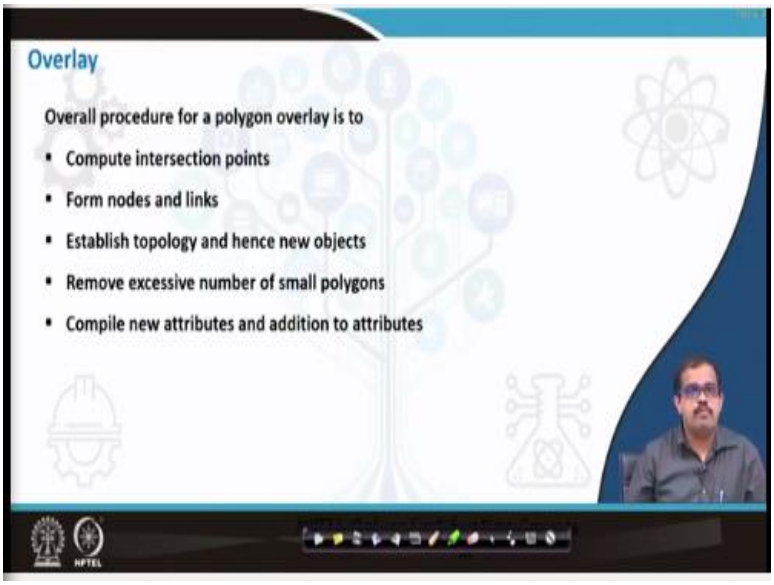
For example what I am meant to say here is you have created a map which is actually belonging to let us say an x city ok. And you whatever the coordinate system whatever the datum that you have used or whatever the geographical quantities that you have used, differs from the second image. Then you will not be able to actually give out or information that is in a true sense. So both of them have to maintain the same geographical quantities.

Once they maintain the both the same geographical quantities then only you will be able to superimpose them and compare 2 data sets otherwise it is impossible to actually get the real

world quantities ok. The proliferation of small polygons maybe contracted automatically by laying the smaller zones around each other ok. If these zones intersect when superimpose then the lines they surround may convert into a single line.

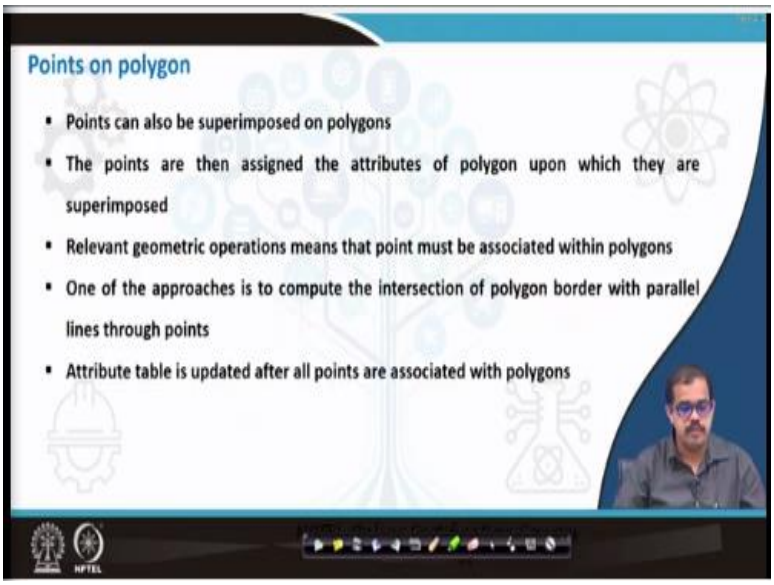
For example if there are a number of lines if they superimpose each other and if they can be converted as a single line when it is actually zoomed over.

(Refer Slide Time: 11:51)



Then overlay procedure of a polygon overlay is to one is to compute intersection points, either form nodes and links established topology and hence new objects. And remove excessive number of small polygons compiled new attributes and additions to attribute. So this is how the overlay operations is done.

(Refer Slide Time: 12:12)

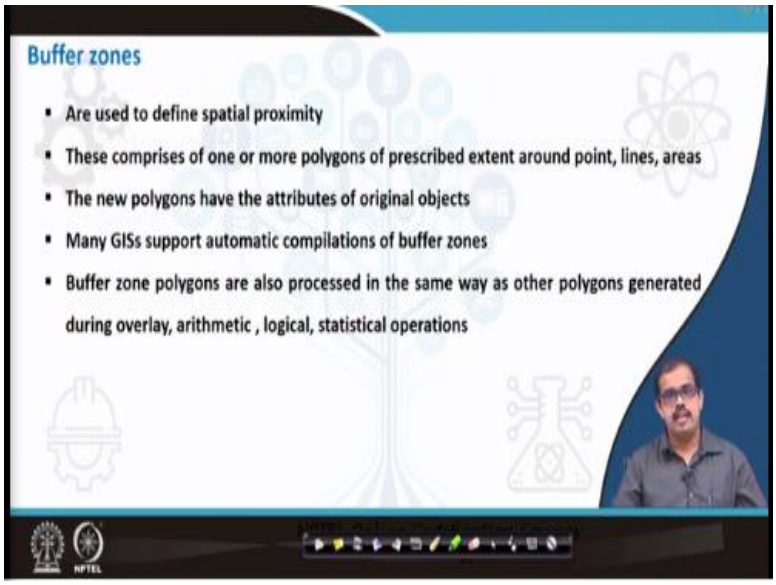


Then points on polygon, this is the second type of overly operations, points can be superimposed on polygons it is as simple as this you have number of polygons. Now, I want to just find out how many trees are there in a city. So I have the tree layer which is points ok, then I have the city boundary ok, I have buildings in that particular city. So buildings are nothing but polygons, so how many trees are around each building, if I have to find out.

So I have a points of those trees I superimpose on this. So that gives me a perspective of how a city has a number of trees and how actually the bylaws the building bylaws are respecting the environmental norms also ok or how it is actually followed by the citizens who are actually residing in those buildings or it is official area of that particular resident ok. So the points are then assigned to attributes of the polygon upon which they are superimposed ok, each of these points are added as attributes.

Relevant geometry operation means that points maybe associated with polygons or it is the one that is related to the polygons. So one of the approach is to compute the intersection of a polygon border ok or border with parallel lines through points, you use points to compute it. Or attribute table is updated after these points are associated with the polygon. So always any operations on any datasets it is the one that is updated with is that is the attribute data.

(Refer Slide Time: 13:53)



Then the next concept is buffer zones, so now if you want to have a buffer zones it is actually defined as a spatial proximity. Let us say that there is a city ok, there is one city x, now I want to find out if the city has grown over it is limit a threshold of the urban growth. Now I want to see how the city is actually spreading towards outskirts ok, what I will do is that, I will draw an buffer of 1 kilometer.

Now I will see how much the city has grown in this 1 kilometer, then I will understand ok if it has grown already grown in this 1 kilometer then I will look at 5 kilometer, then 10 kilometer, 15 kilometer. There may be at a certain point of time where there is no much urban growth in that region. So that analysis may help me to understand how the city has outgrown itself from the periphery towards outskirts, this is called urban sprawling ok.

The there are different types of sprawling but this is one of the types of urban sprawling. So it means to say that there is a sprawl across that particular sections of roads. So this is what the buffer zones can offer, it can say that how the urban area is actually even interacting with the regions just next to the urban or the peri-urban areas. And how this particular region is being affected, so such analysis can be made that is why you need a buffer zone.

And let us say you have a `water body and you want to find out you know what is the watershed of that particular region. So which means to say that watershed should be relatively with

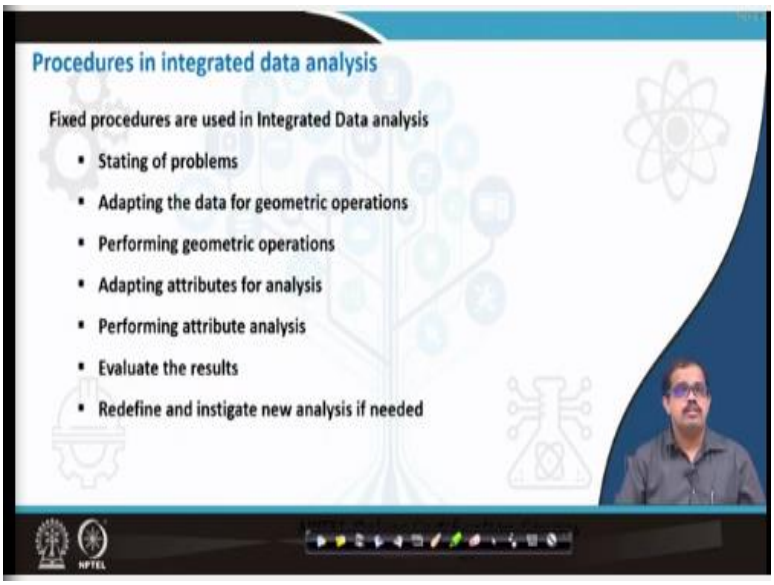
vegetation and you should to understand that, so what how much is the maybe let us say a 5 meter, 15 meters, 5 kilometer or 15 kilometer or 25 kilometer whatever is the rule you try to draw a buffer.

And see what kind of infringement has happened as far as that water body is concerned. Such kind of analysis can be done that is why you need to understand that there is a tool called as buffer zones which can be used as a spatial analytics for various kind of applications ok. These comprise of one or more polygons and of prescribed extent around a point, line or an area, it can be around a point, it can be around the line, it can be around an area.

So it depends on what kind of application it is ok. So when you are looking most of the softwares today have an application with the buffer zones, any of the software can actually provide you a tool where you can create a buffer. Buffer can be based on the way that particular phenomena if you have. Let us say I have an road which is actually changing its course. So if the course is something like this, you can draw buffer based on that course or let us say you want a uniform buffer of a circle.

So you can also draw a uniform buffer of a circle, so it depends on the user what kind of buffer and how the buffer has to be done. Buffer zones polygons are also processed in the same way other polygons are generated during overlay arithmetic logical or statistical operations. So the way it is applied is very different other than that there is exactly no difference one of these polygons in applications of in the any of the applications that we may use ok.

(Refer Slide Time: 17:23)



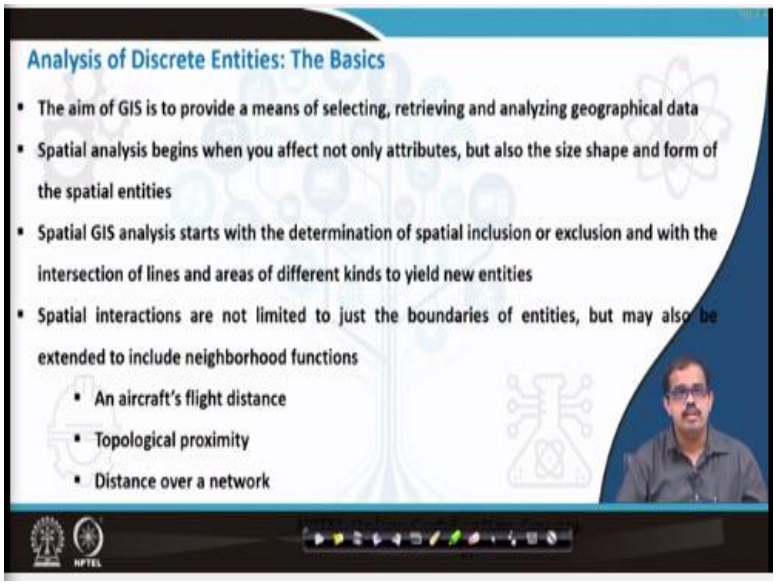
So when in case you want to do an integrated analysis, there are fixed procedures that are used in integrated data analysis one is stating of problems ok. First you have to state a what is the problem that you are trying to understand. If the problem is sensible to be understood, that is the first step that you have to always look at. Then adapting the data of for geometric operations, what kind of data, how the geometric operations are done, what kind of operations has to be done.

Or what kind of weather it is polygon on polygon, polygon on point, point polygon on line. So what kind of analysis are you trying to do, what kind of geometric operations would be done. And adapting attributes for each analysis, what attributes you need and how it is actually reflecting on the information. Then performing attribute analysis once it is selected, finally evaluating the results that is very important.

Whether you have to evaluate and it is very important that you validate that result ok. If you have done it then whatever the data that you have generated is sensible otherwise the data that you have generated is really not working with the real world. Then redefine and instigate new analysis if needed ok. In case let us say the validation data is giving a wrong result which is not coinciding with the data that you have.

The thing is that you have to again reperform the entire analysis that you would have performed. It may be only for spatial analysis, it maybe the analysis from data acquisition to your spatial analysis, so it can be anything. So they are depending on what kind of output you get the validation has to be done. Once the validated you will understand where you have missed in case you have not missed you would have got good results ok.

(Refer Slide Time: 19:13)



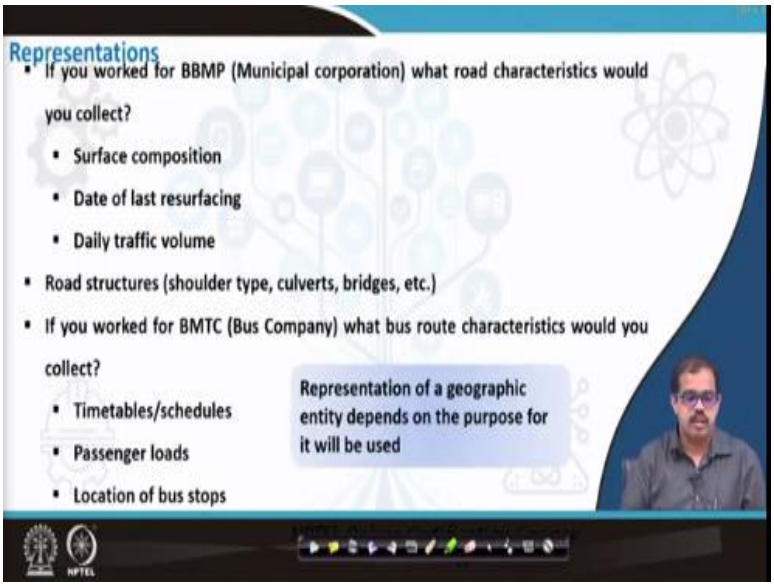
So, very important as look at very, very, very important part is how do you collect data. So once you have collected data extremely well then the other part will fall in line ok. So the next thing is analysis of discrete quantities. So if you are looking at discrete quantities is how do you actually select retrieve analyze geographical data as a discrete value. So when you are looking at this, when you are doing any analysis on discrete quantity.

Always there is an effect on attributes ok when most of the spatial analysis are done on it also affects the size, the shape and the form of the entities ok. So when I say discrete entities, it is discrete objects, these are entities ok. Each of the entities will be affected in terms of size, shape or form. When you are looking at spatial GIS, spatial GIS starts with determination of spatial inclusion or exclusion, I was speaking about in the previous class only about inclusion.

Now it is also with inclusion and exclusion, you can look at both of these when you are looking at interactions that are not limited to bound just the boundaries of the entities. But can also be

extended to include your neighborhood functions. For example, an aircraft distance ok, then topological proximity it can be distance over the network. So all of these can be analysis of a discrete entity ok.

(Refer Slide Time: 20:52)



If you have done analysis, how do you actually are present, it actually depends on what kind of analysis you are doing and for what aspect. For example let us say, if there is a work for BBMP**,** BBMP is a municipal corporation Bruhat Bengaluru Mahanagara Palike ok. So what road characteristic would you collect, if you are collecting a characteristics of a road, what is that you actually collect.

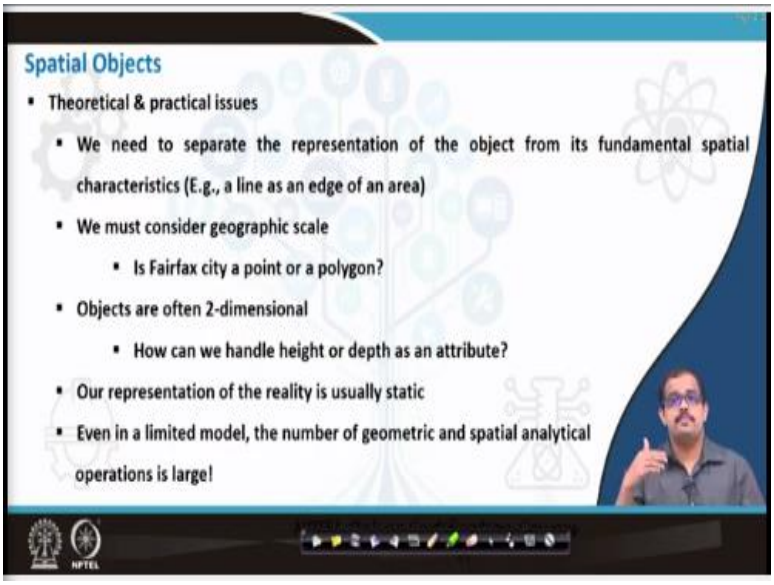
It maybe your surface composition, date of last resurfacing, it maybe daily volume traffic that is what is necessary for you to actually plan what kind of road has to be there and how it has to be there ok. So that is what means by representations or let us say if the same thing is applied in BMTC, BMTC is nothing but a bus company ok. So it is a transport corporation which is actually a bus company.

Now if that is a specific problem then probably I would like to know what is the timetable of those process or schedule of buses arriving and leaving ok, at a particular point. Then what is a passenger loads in every of the sectors, how is the location of the bus stop located and what are

the facilities that are there right. So these are 2 different cases of representing in a different way, so all the representations cannot be clubbed as a single rule.

So each of the representations have their own way of representing different quantities. So every representation of a geographic entity depends on what is the purpose it has been used or it will be used. So each user should himself or herself understand that how the representation has to be, what is the use that he or she is trying to do. And what is the final analysis output that will be done for a particular representation ok.

(Refer Slide Time: 22:55)



Now if it is a spatial object when you are looking at both theoretical and practical use, we need to separate the first thing is representation object from it is fundamental spatial characteristics ok. For example there is a line on an edge of an area or a polygon, remote, corrected ok, we must consider the geographical scale. The scale in 2 different vector models would be very different, so you have to look at what is the different geographical scale.

A Fairfax city is a point or a polygon, let us say if you are representing entire India. So each city is a point ok, if you are representing only a particular city of Kolkata that city is nothing is a polygon. So please be extremely careful on what geographical scale you are considering ok. Based on the geographical scale and the application you have to choose what kind of particular representation you are doing ok.

Objects are often 2 dimensional, we have to handle the height and depth as an attribute, please keep it in mind. Whatever you are trying to do will be a form of attribute ok, our representation of reality is usually stand the real word is static. When you are looking at a GIS model, it cannot be dynamic. Now, nowadays because of representation, because of the web evolvement of the web GIS in a larger form.

Now it is becoming more dynamic, but yet in today's context, in today's scenario it is yet static. Even in a limited model, the number of geometric and spatial analytical operations is quite large. Even if you do a very small model of a single road ok, the analytical and geometrical aspects that you have to consider is very large in terms of spatial objects. So be extremely careful when you are looking at different model and how the aspects of the model are being considered ok.

(Refer Slide Time: 25:01)

So this is about how you use as overlay a buffer zone and how the spatial objects are being represented. So what I am trying to make you understand is that how actually an analysis is done, what are the different tools of analysis that are being done. That is what I expressed in terms of overlay whether you are doing operations on attribute data or on buffer zones.

And now once we have understood each of these operations as different aspects. Now the final thing is how the spatial objects can be manipulated, how spatial objects can be used in terms of

analysis of data. So that is exactly the way of representations that you have to have. And please keep this in mind the representation of anything is very contextual or application based, it cannot be a similar for any all the applications ok.

So and user, it depends on the user, depends on the application and depends on what kind of data that particular phenomena or objective is actually handling it. So look at that aspect and look at the representation how it is done ok. So I will end this class by these spatial objects. In the next class, we will look at advanced analysis, next 2 classes we will look at what do you mean by advanced spatial analysis.

And how do we do an advanced spatial analysis, and we will end this particular week with that, so until then have a nice time and thank you.