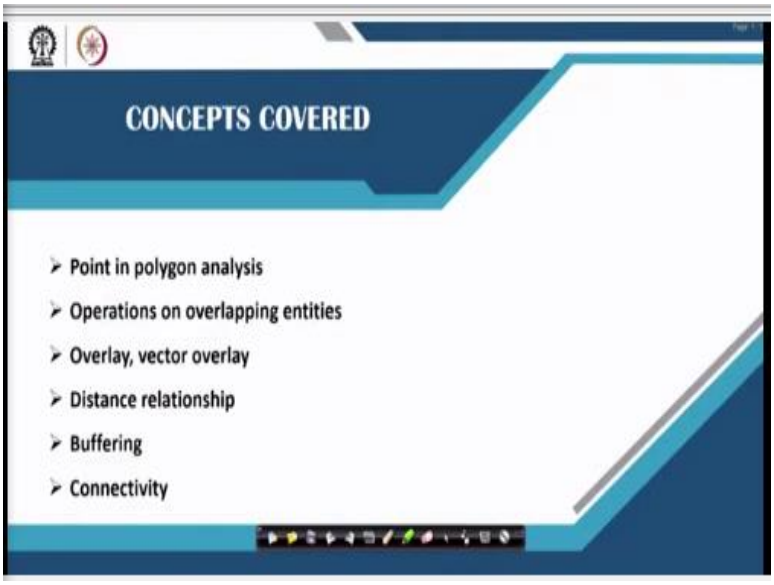
Geographic Information Systems
Prof. Bharath H Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology-Kharagpur

Module-09
Lecture-44
Basic Spatial Analysis (continued)

Hello namaste, welcome back to the course on geographic information system. In my previous class we had started looking at what do you mean by an overlay, what do you mean by overlay on a point on a polygon basis or overlay on vectors. Now as in such we continue with the same mode we will again go forward and understand with certain examples. In the last class I had given you only a theoretical overview of how this overlay analysis and how do you use it in your analysis or a buffering analysis how do you do it.

But now let us see in as a practical part of it how different things are applied at different context ok. This specific class is a extension class of the previous session whatever I have spoken a previous session and this is a exactly the thing that theoretically they had given I would put it in more in a practical purpose.

(Refer Slide Time: 01:17)



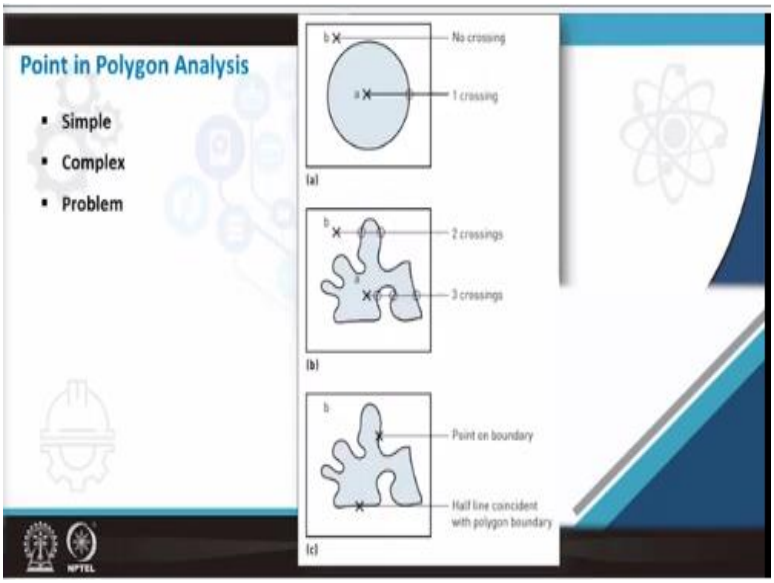
So ok, so I would speak today on point in polygon analysis, I spoke about how vector on vector data is possible, I have spoke about polygon and points I have spoke about polygon, polygon

overlay, point in polygon overlay. And also I spoke about how you can utilize it for various analysis. So I would take it forward, I would look at the same aspects as I spoke in the previous class.

Then operations on overlapping entities, how do you do an operation on overlapping entities. Then the vector overlays the same I would not stress much, then the distance relationship I have already spoken the previous class, I would not pick up much here. But that is one of those things that you have to see when you are looking at how the distance is measured and it is actually changing.

Then we look at the buffering part of it what are the different buffers, how buffering is an issue, how the buffering can be utilized for various analysis, then the connectivity.

(Refer Slide Time: 02:19)



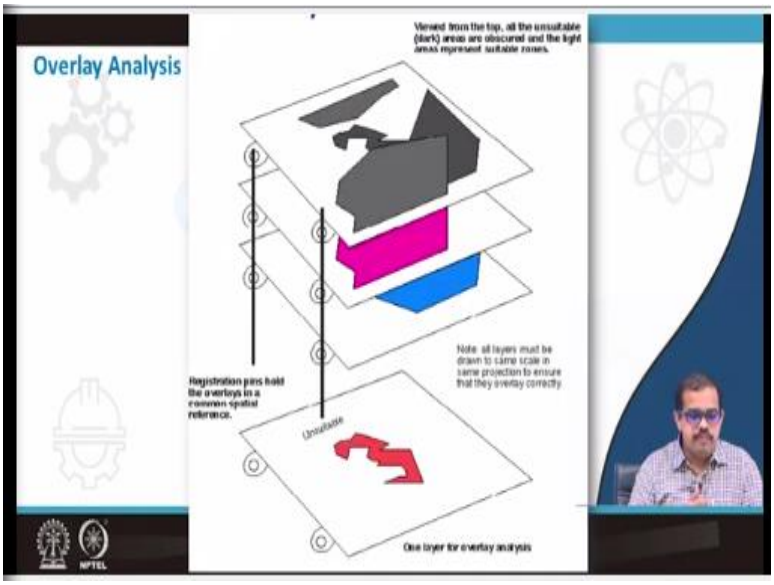
Now when we are looking at point in polygon analysis. For example, there is one problem here. So if you have this a polygon, this is a circular polygon, so you have 2 points that are here. But there is only one point that is actually crossing the polygon which means to say that this point will be highlighted in as a simple case of point in polygon analysis, which means it is either indicating a specific property of that particular polygon ok or it may be a qualitative or it may be a quantitative issue.

But when you are looking at any complex issues, you may see that there may be a lot of crossings for particular polygon. If you look at here, these are 2 points if we take a point b there is 2 crossing and where if you take a point a, it has 3 crossings. If it is done if you are trying to understand a cross sectional analysis in this particular study. So that is what is a point that is how a point in polygon analysis is done.

For example, if I draw a line like this and try to see that particular point how it has analyzed over a particular region ok. So if there are more crossings for particular polygon by a point then it becomes a complex problem. When it is only 1 or maybe even 2 becomes a more simpler problem which is much easier to solve in any of the polynomial equations ok. And the third one is the problematic issue where there is a point on the boundary or some point as outside the boundary and some point as inside the boundary.

So such issues are more of complex, so you have to either understand what percentage of that polygon is actually indicative of this particular point. So to find that it you have polygon you have to understand the features of your polygon, the shape of your polygon and then try to when way in how much what is the characteristics of that polygon. So these are certain issues that you may face when you are looking at the overlay analysis.

(Refer Slide Time: 04:31)



And when we look at the entire overlay for example, if you have number let us say as I previously said why I have shown you this specific image is that. Tell you that whatever their maybe a number of layers that you have. Let us say the summer is working on urban flood, so they have develop the drainage network. The urban area map from maybe 2 decades map or 3 decades map, they have digitize, they have done the classification etc. whatever it maybe.

Then they have the ward boundaries of each of the maps then they develop and natural drainage system, they develop the slope, the hill shade, so all of these quantities they have developed. Now that person has to look at what are the different quantities which are actually causing the urban floods. So now you overlay one on another for if you have all those different layers, so we overlay one on another using certain quantitative issues.

For example, if I say that I will weigh each of these layers at a particular weight. So you will adopt maybe any (()) (05:38) analysis which is the advance spatial analysis or it may be your AHP. You try to say that this is a weightage of this particular layer or this is the importance of this particular layer in this particular analysis. So with that you will also find out where actually the drainage network has blocked and where the flooding is happened.

So if you can understand that the flooding is happening because of the blockage of drainage network then it is highly sensitive area. So that is how you can come up with a flood vulnerability map which is actually addressing the urban floods in that particular region. So that is one example that I have just mentioned, so you can have any number of layers, any number of quantities, but these layers have to be utilized in analysis.

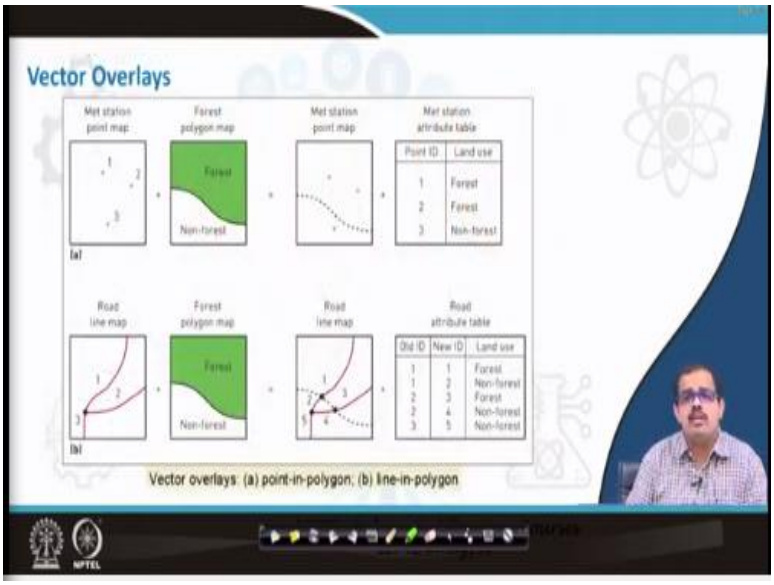
For example, if you are using an economic layer ok in your urban floods, so that gives you a different angle of analysis of this particular analysis in urban floods. But it may be useful for certain section of people who are actually looking at your model, it may not be useful for many sections or people who are looking at your model. So the amount of data that is necessary the same data you will use it for different kinds of analysis, that is why overlay analysis is extremely useful in GIS ok.

(Refer Slide Time: 06:57)

So just to give another example this is one of the examples from one of the research papers. So when you look at it I have ward level boundaries of a particular place ok, then I have another vector layer which is actually giving me the vulnerability region ok of anything. Now this is just an example ok, now I overlay on the other, so once I overlay one on another I know which area is actually vulnerable.

Based on this the output will be the vulnerability map of both the layers which are overlaid on each other ok. So that is exactly the example or application of why you do an overlay analysis. So if you have a physical layer ok, let us say you have 3, 4, 5 physical maps of all these layers, you take a printout, use it. But you will not be able to understand which region overlaps on each other. And what is a quantification of each of this layer, that can be done using GIS, that is why GIS is a powerful tool.

(Refer Slide Time: 08:07)



Then you have vector overlays, so when I say vector on vector overlays, it maybe both point in polygon and line in polygon ok, I have discussed already the point in polygon. This example gives you exactly details of a planning of how you can look at the planning map. Now there is metrological station point map in these 3 regions, now I want to know what is the rainfall in that particular region and what is it is corresponding clause in that particular region.

Now, these are the 3 metrological stations, so I have values attribute data that is attached to it, has the amount of rainfall in that particular region. Now, I have a map which is actually a forest map versus a non forest ok. Now I overlay one on the other, so let us based on these overlay I find out. In that map what is a particular land use and what is other attribute information such as it is rainfall etc. ok.

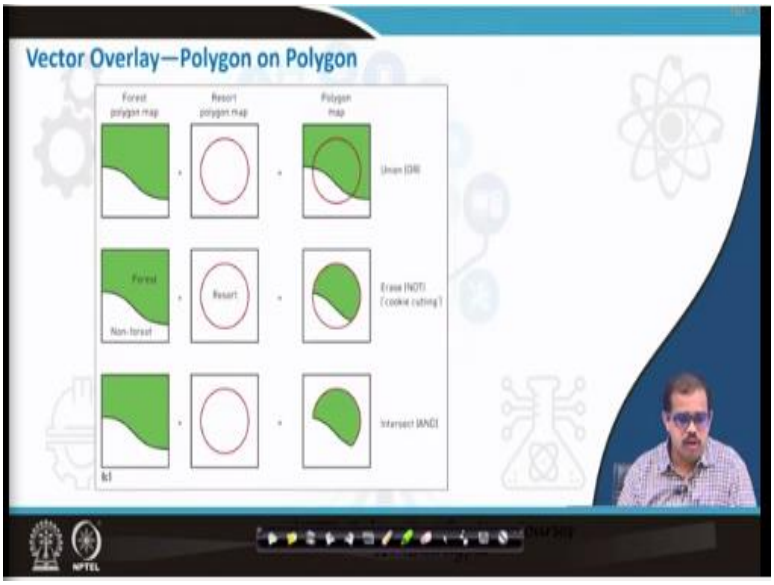
So now I will be able to find out whether on a broader level, it is not exactly on those details. But it is on a broader level you will say whether forest has certain influences on rainfall or rainfall pattern. The second example here is a line on polygon overlay, so for example here you have a line map ok this is a road line**.** Let us say that there is a road that has to be built and your road connecting to places.

Now you have to produce a environmental impact assessment report, so what you if the best thing is do is get an a forest non forest or vegetation map of India from various sources. It maybe

from Bhuvan, it maybe from MOF website or any of those sources none you have a roadmap ok. So now once you have a roadmap which is a line ok then you have a forest non forest area which maybe a probably most of it is a polygons, now I overlay one on other ok.

So now I know where actually it has crisscrossing each other though a new ID is allocated it tells me this particular line where it is crisscrossing is forest or non forest ok. So which means to say that I can do an analysis on understanding how it is actually environmentally impacting a particular region when such development projects happen. That is why you can use overlay tool in much bigger context ok.

(Refer Slide Time: 10:46)



So there for example, another example is you have a resort that is coming up in that region or it is proposed. So use simple tool is use a Boolean operation, Boolean union operations ok and I say union it is R. So once you do a Boolean operation you can find out where that particular project is proposed. Once you understand where the project was proposed you can take a decision or anyone who is considering a policy decision or who is a policy maker can easily say that say this particular region cannot be allotted for the resort.

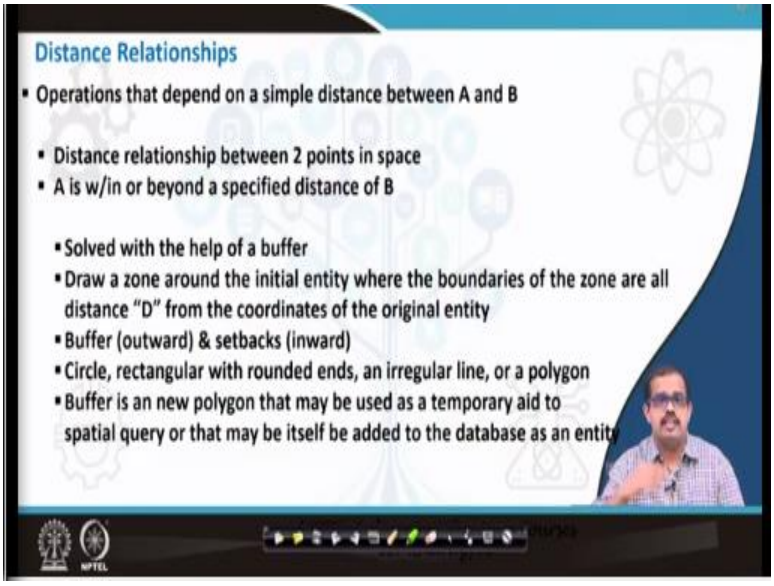
Because it is in the forested region and the forest region cannot be converted as a commercial activity ok or for example, if let us say I have a forest map, I have a resort map ok. So if I want to know that how much percentage of forest falls into this resort area, as simple as this user not

operation of Boolean. So what I am trying to imply is that any operator can be easily applied in any kind of analysis in GIS, only thing is that it has to be logical ok.

So in the third map if you see there is again a forest and non forest, now I am doing a intersection which means AND operator. So it is intersecting both of this map, when I intersect it means to say it gives me the exact area where the forest is there ok. I do not, I have just a forest map, I do not have a non forest map. Now, once I intersect non forest area is completely removed and I get only the forested.

Now I get decision maker can easily consider a decision of whether that land has to be allotted there or not as simple as that. So this is where overlay plays a very, very important role in giving us important decisions on how that particular operations can be used or utilized for providing direct informations ok.

(Refer Slide Time: 12:53)



So there is the next thing called as distance relationship, yesterday I have spoke about the distance relationship. But when you are looking at distance relationship, you have operations that depend on simple distances between A to B. Let us say tomorrow you decide for example, you will develop a map of your institute, once you develop a map of your instrument you should make it more dynamic.

So what you will try to do is that, if anyone who is entering the institute you will have to tell them what is the distance from point A to point B and how to reach them ok. So how do you do this, basically you will have to look at the shortest distance route or different routes that are there in that particular institute and this are then used by nodes to calculate the distances ok. The best way to do this is by using a buffer, help of a buffer ok, so how do you do this.

Basically, first I will draw a zone around the initial entity where the boundaries of the zone are all distance D from the coordinates of the original entity ok. Now buffer both buffer that is outside and the buffer that is inside, so when I say buffer that is outside, that just outside the entity that you are representing and buffer inside is the inner entity of that. If you take a 5 kilometer or for 2.5 is an outer entity and 2.5 is a inner entity, so that you will look at.

Then you will look at the circular, the rectangular with the rounded ends or irregular lines or a polygon, what kind of polygon can connect these. Once you have understood once those polygons have been drawn or lines have been drawn, then buffer to the newer polygon using a temporary aid can be done using a spatial query. So which means to say that once you start locating that particular region to reach that particular point.

Then you will say that this particular by at any query how do I connect these 2 polygons are these, 2 lines or these 2 points. So once you have connected those 2 points, then your route is actually established, that is how you do it ok.

(Refer Slide Time: 15:11)

So now for example this is exactly what I was speaking, so you have to reach this point from this point ok. So this is the distance that you have to calculate, now you draw a buffer ok. Now this part is a outer buffer and this part of a inner buffer for this region. Let us say from the point if this is the entity, this outer buffer and this part is a inner buffer.

Now with this outer buffer and a inner buffer what you try to do is that you will find out how the lines connect each of these points across those lines ok. If there is already a set of road, ok. Now, you will draw several buffers in this region ok, based on this particular point, you will then try to trace this particular line, so this line is actually connecting these 2 points.

So once you have found out that it is connecting those 2 points, then you calculate what is the shortest distance and other buffer regions, which is the shortest distance. Once that is calculate that is how Google gives you the shortest distance and maybe you have 2 or 3 more distances that is actually put out there. So it says that this is with tolls, without tolls, you have this distance. So those are based on the database is information that you have, when you are actually travelling.

Google would have collected all those database information either through surveys or either through it is manual inputs that it may have it in it is database ok, so this is one example of buffering.

(Refer Slide Time: 16:45)

So when you are looking at a buffering you can do a line buffering, a point buffering or it can be polygon buffering ok. So when you are looking at a line buffering, so you are actually looking at what is across the line, so the effect is across the line. Then if you are looking at a point buffering what is the effect around the point and when you are looking at a area you are actually looking at every point on the area what is the effect on every point of the area from the buffered region or the buffered source.

That is exactly why you need a buffer point line and feature or an area in order to look at it is analysis. So very good example is that, for example let us say I have planned a new metro line that is coming up. So now I want to understand what it may have an impact on next neighborhood ok, maybe because of if you raise an FAR how much is the economic impact that government may have.

Because of you will have a lot of I mean once the FAR is raised, so you will have a lot of income in through building property issues etc. So when you are looking at that, so you can do such analysis when you have a buffer. So your first you will fix up a certain amount of buffer, so if you increase the buffer in this region this is amount of impact if you increase to this.

So that is how certain additions are considered when they are actually when such big infrastructures are coming up. If you are looking at a point, so let us say that if there is a new

infrastructure that is coming at a point or let us say a new flag post has been put up ok or if there is a new signal has been put up. So what they try to do is that, if that signal comes here, what maybe the effect on the across that entire region.

And how many signals are required at that particular point and how it is going to affect the traffic load in the entire region, so that is analysis can be easily done with just a buffer ok. So first they will fix up 1 buffer and see what is the load, then they fix up the next buffer and look at the load where there is a least, so they try to maintain that region as one signal corridor. So the next signal corridor is placed in the next region.

Similarly when you are looking at area I have given you an example yesterday about the urban area. For example let us say this is one of the city ok or let us say it is a lake whatever you may consider it to. If a city is there, so I want to understand what is a growth of that particular city over a period of time ok. If many cities like mega cities of India, it maybe Delhi, it maybe Mumbai or some very big cities like Chennai, Hyderabad or Bangalore.

So what they try to do, normally the phenomena is it starts at the center there is certain growth at the center and slowly stretch towards the out search. So now with different phenomenas because of different distance by the government, so there may be a spread in different regions. So then you had national highways and you have your state highways, so based on this there may be a spread outside the city boundary also.

So what you do you have a 2 kilometer buffer or a 5 kilometer buffer or a 10 kilometer buffer. Once you have understood that, then you can find out at every instant at every buffer that you have to consider. So you can find out exactly where the urban area is intensified. For example if you take Bangalore, so Bangalore if you try to analyze, so once you see the outskirts it may be towards the Bangalore south.

If you see there are lot of development that would have happened, then towards (()) (20:31) they would have happened there would have been a lot of developments because of the highway connectivity towards Bangalore. So that kind of analysis how actually the skirting of the urban

area happens how can be easily analyzed, it can be done using different metrics etc. for different analysis. So that is how the buffer is extremely useful in terms of analytics ok.

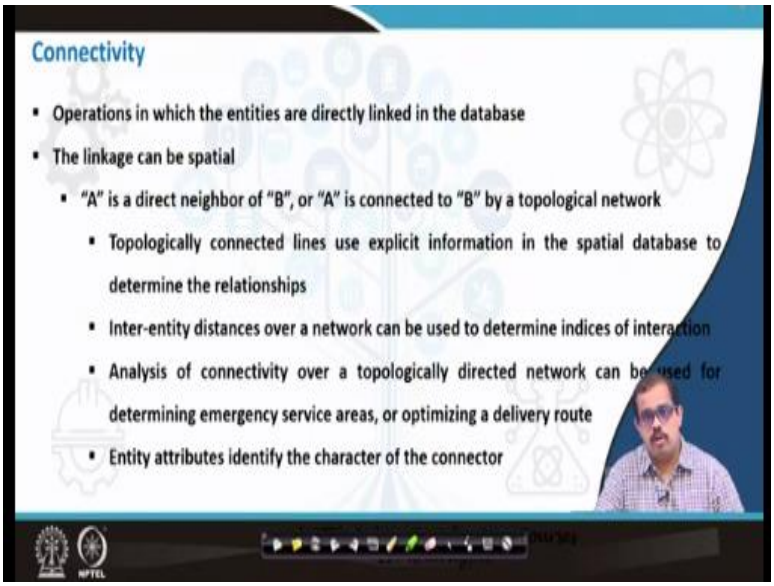
(Refer Slide Time: 20:57)

So then operations on a overlapping polygon, so this normally when you have polygons which overlaps they actually create issues that are in the digitized boundaries ok. It maybe because of this spurious polygons, you have one polygon like this and that polygon sitting like this. So it maybe because of certain digitization errors or it may be the errors because of different sources, that have been collected.

So there may be different solutions to look at it, the first solution is designate boundaries of one feature as a dominant boundary ok, which means to say if there is 2 polygon. If there is a polygon this is one polygon and this is another polygon let us say like the sitting like this ok, exactly like this. So if I designate this polygon as the major boundary, then this should sit something like this, that is what is one of the solutions.

Then otherwise, eliminate polygons that have areas smaller than the critical threshold, delete that area at once. So you can find out what is the threshold area then passes smoothing window over the polygon coordinates this can be used by fuzzy tolerance very easily done and most studies do this. They do a smoothing window, so that all spurious polygons are removed or adjusted in the way that it has to set on each other ok or sharing the same common boundary.

(Refer Slide Time: 22:23)



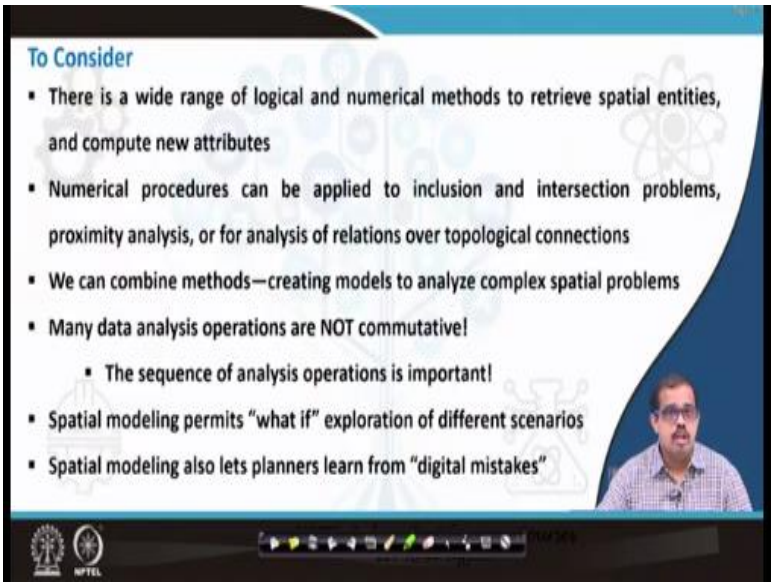
So that is about how the overlay in polygon is done. Now the next thing is connectivity, how do you find out the entities that are directly linked in a database. For example linkage can be spatial cannot be even spatial but it is attributed to be spatial. So how do we do this for example let us say A is a direct neighbor of B ok. So or A is connected to B that is how we speak, A is a direct neighbor of B or A is connected to B by a topological network ok.

If we consider a point A and a point B both are direct neighbors but they are connected via topological network. When I say topological network, the topologically connected lines use explicit information in a spatial database determine the relationships among them. Neighbors is a relationship ok or connected by such distances and have this such topological connections is again a relationship ok.

Always spatial database are related with relationship determined by the relationship. Inter-entity distances over a network can be used to determine indices of interaction. Analysis of connectivity over a topologically directed network can be used for determining emergency services or optimizing a delivery route. It can be used for any kind of analysis I am just giving some examples of where a particular analysis can be used.

Entity attributes identify the character of a connector, what is the different connectors, I have given you example of what is the connector, what is the node and how they are connected across each other.

(Refer Slide Time: 24:11)



So the very important thing that I have to point out here, before I finish up the basic part of the analysis is that people have to understand certain things when you are looking at GIS what are the things basically you look at is. There may be a lot of logical, there may be a lot of numerical methods to retrieve spatial entities or compute spatial entities or do any kind of analysis or any kind of operations, so you have to look at exactly what is valid ok.

So please look at the theoretical explanation of how you are applying a particular tool or why you are applying a particular tool. For example, some of those students who actually put in their resume as I know GIS ok. But when they come to the lab and so I asked them, so why have we used a particular tool, they would not be able to explain they would say that this particular software has this tool, so I have used this to do the analysis, no, that is not the right way.

So you should be able to understand and interpret why that particular tool is there and why it has to be use. It may not be the best tool that can be used but if you understand the science behind it, it is much easier to actually even the manipulate that tool for the analysis that you have to do.

Then numerical procedures can be applied to inclusion and intersection problems, proximity analysis or analysis of relations over topological connections.

So when you are looking at this topological connectors, it is extremely important for you to analyze what kind of relations are there. So without relations, most of them try to utilize the way the analysis is done but they do not ever think about what is a topological connections that exist, it may be a non existence all throughout the period. So but look at the entire topological connections and then derive what are the spatial relationship is there.

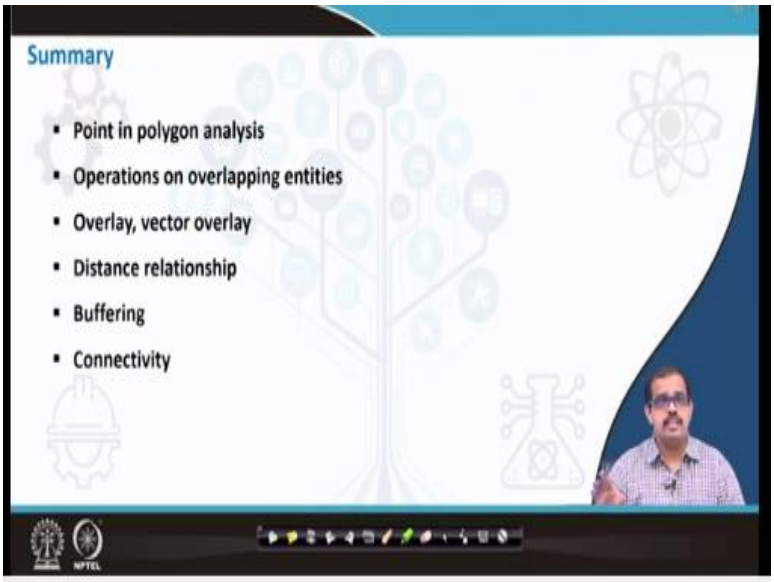
So please look at how topologically each of these are connected ok. We can combine methods, create models unless complex spatial problems. But when you are creating models be careful on what is the model behavior, why are you adopting that particular model ok. So and what kind of analysis you would be doing, so you have to plan it first and look at what kind of models you will be looking at.

So many data analysis operations are not commutative ok, so keep this in mind. The sequence of analysis operations is very important the way it is operated. So you cannot say that I have used this operator, so I have used this first and this next without any logic, you should have a certain logic of why it is being used. Then spatial modeling permits what if exploration of a different scenarios ok, this is very important in terms of spatial modeling.

Always look at scenarios which the models were in you create scenarios, so it could give you the more flavors of how the model behaves at different context. So your model will be more successful in case it can give you better results at different context when you use what if scenarios. If many brains work across, so those brains have various ways of thinking. So what if actually gives you the exact way of the replication of the real world scenario.

Spatial modeling also let us planners learned from digital mistakes, so that will help them in a larger way. But only thing is you should be good enough to understand which means that if you can understand the basics, then it is easier for you to interpret what do you mean by a digital mistake ok.

(Refer Slide Time: 27:54)



So these are the things that you have to consider. So as I said first we looked at point in polygon analysis wherein we looked at how, what is the simple problem, what is a complex problem and what is a problematic problem ok. So look at those issues when you are looking at point in polygon analysis. And we looked at how the overlapping entities can be operated.

For example, we looked at how the line entity is operated on a polygon or how a line entity is used for various analysis. Then we looked at vector overlay, buffering, buffering we discussed about how buffering can be done whether it is urban scenario or whether it is application to any natural resource management whatever you can do it.

Then distance relationship that is extremely important, how do you do a distance relationship probably one can understand it is applications in a very large context when you are developing the entire database for a particular region or city or a university. Then we looked at connectivity, so that is extremely important. And finally I ended this particular class with certain cautious things that you have to look at when you are looking at spatial analysis.

So in the next class, we would take up the last part of how do you do an analysis is the advanced analysis, what are different advanced tools do you have ok. So once we have understood this, if you have understood the advanced tools at the end I would also give you a certain ways of how,

what is the issues when you actually look at a GIS is a data ok as a tool, as a data. So that is how would be the next class so until then have a nice time, thank you very much.