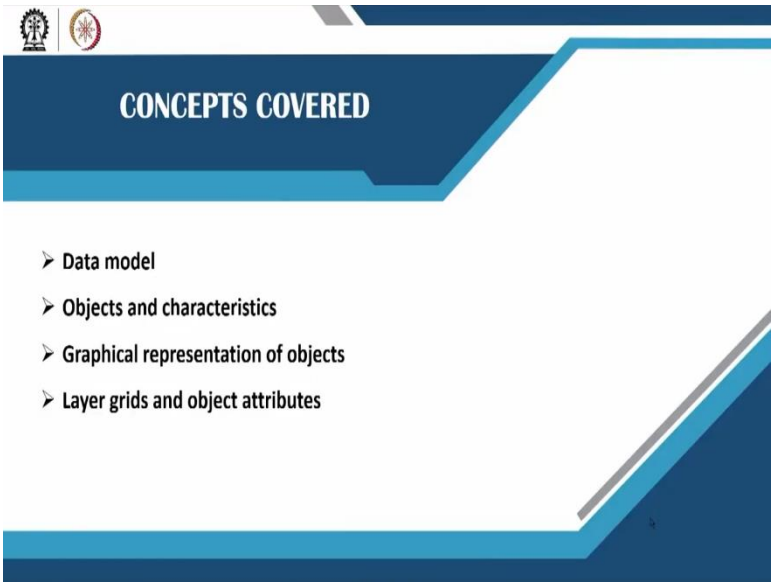
Geographic Information Systems
Prof. Bharath H Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology – Kharagpur

Module No # 02
Lecture No # 07
Real World to Digital World Through GIS (Continued)

Hello Namaste I am back with the module two the lecture two which is where we would be trying to understand again the real world to a digital world. So we will look at models here specifically, how are the models are build. Why there is a complexity in real world representation and how do we actually represent the real world. It is not as I said in my previous lecture.

It is not so easy to represent the real-world phenomena in terms of a digital world. So how do we convert real world phenomena? How the real-world phenomena is actually approximated in a digital world phenomena? So that is what we would understand in this particular lecture.

(Refer Slide Time 01:07)



So the concept that I would cover in this particular lecture is what do you mean by data module. I would revise what do you understand by a data module basically. Then we would also look at objects and characteristics which is how it is actually represented. Then we will also look at how we graphically represent an object. So whether it is in the form of grid, whether it is in form of

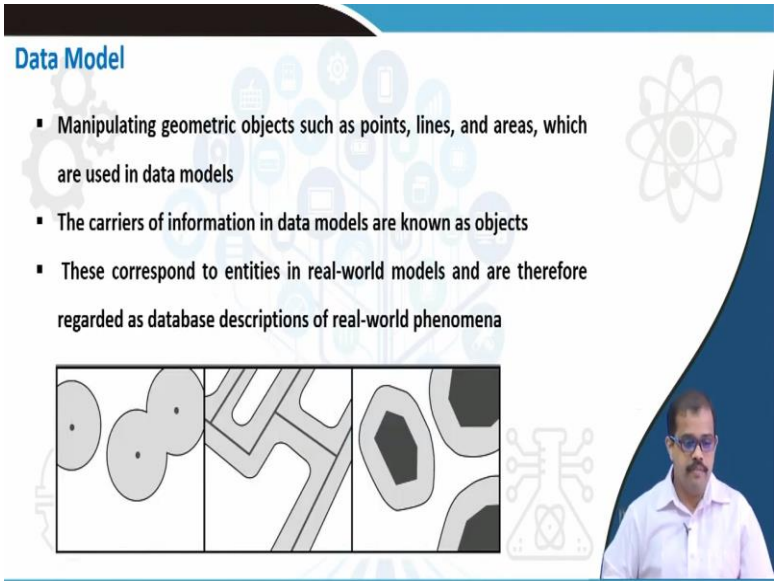
point line polygons. So how do we normally represent an object? Then we would look at what are different of layer grids and object attributes. What are different attributes that are there in this particular transformation of the real-world data into a digital world.

(Refer Slide Time 01:49)

So when we look at the data model as I said a real world data model facilitates the study of selected area of application by reducing the number of complexities considered. Which means to say that it is actually helping us to understand how the real world phenomena is there, which is quite complex which is quite dynamic, by converting it a data model and querying it for any instance of usage that we may require into. So it is facilitating any area of study which we would have collected and stored it as a database at any point of time.

If a real-world model is always to be used in any use it must be in a form of database. That is what I meant all along when I am actually representing the entire real world to a data world data model. So when we are looking at database data model always it is all about storing your data in a database. Converting your data into data model and storing it as database. That is what is extremely important the way you think about a data model.

(Refer Slide Time 03:01)

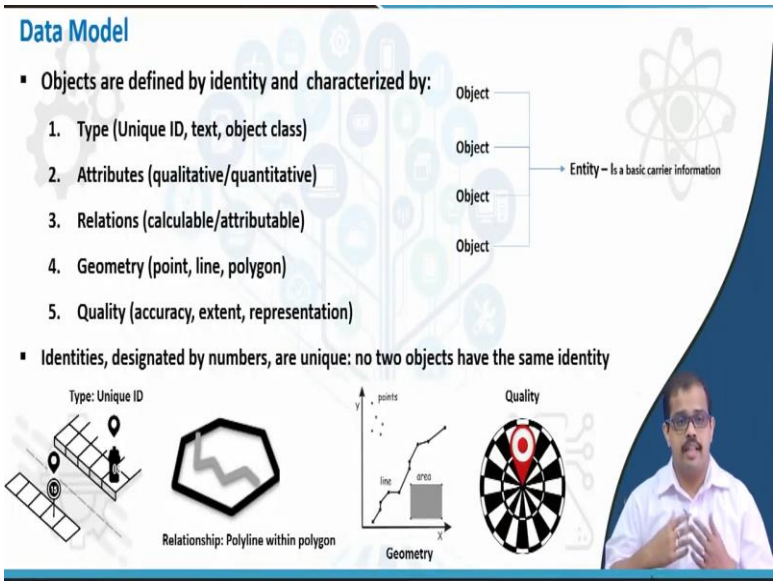


So when we look at the entire data model when we look at the geometric objects in the real world. For example you as I said there may be a number of trees around this particular building, there may be a number of roads that are actually may be connected to this particular area and there may be number of building in this particular area. So each of these are represented as point, lines and polygon in a vector data ok.

So when we look at this data huge number of geometric objects are there. So manipulating of this geometric object, so these are use in the form of a data models. So when we are looking at data models each of this objects are the are nothing but your information. These are the ones which actually carry your information. So we call it as a carrier of information in a data model is nothing but the objects. So when we are looking at this these correspond to the entities of the real world models and are there therefore regarded as database descriptors.

So when you look at the entire real world phenomena these becomes data at descriptors also as the carrier of the information of the real world in for the data model ok. So the point line and the polygon in a vector model now called as a descriptors, real world descriptors or your information notes.

(Refer Slide Time 04:32)

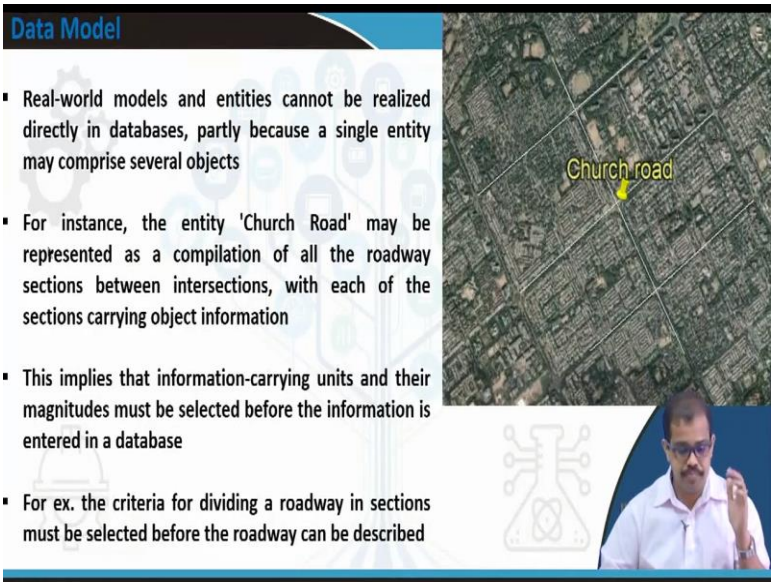


So when we are looking at the entire data model you have the objects that are defined by identity. So you have various objects in that particular data model that are representing in the form of identity and are characterized by you may have type, you may have for example I did given a example of a road, then you have attributes that maybe qualitative, quantitative it has always a relationships, then you have a geometry which is connected to it. So point, line and a polygon and you have a quality that is the accuracy and extent and representation.

So how do you understand what are things that are there in this particular data. So you should remember that I have already spoken you have you will always have a metadata. May having a metadata is extremely important in order to understand all of these aspects when you are looking at the entire data model. So identities designation by the number are extremely unique.

No two objects will have the same identity. If no two objects have a same identity then it means to say that it is actually giving you a bad result. Because your database would not be able to handle both with the same identity. And finally, you would get a you may get a wrong result or you may not even get a result. So that is how a data database works. So you should always have a separate identity. So as I said an object ID should be completely different.

(Refer Slide Time 06:04)



So when we are representing a real world now let us go back to the model to the real world to the digital world model. So when we are actually representing the real world, real world models and entity is normally cannot be easily put in to the data model because these are not, these are a single models entity but may have several objects several hundred, thousands, lakhs of objects.

So representing that is why extremely difficult. For example, you have a particular region here if I have to represent all of these building in that particular database. So representing this collecting this information is has to be extremely accurate. And all the attributes that are there for each of these which for different application is needed as you determine so these has to be done more accurately.

So determining this as extremely difficult but if I considered for example let us take this particular road which is called a church road ok, so if we take consider this it has different sections of roads, it has an intersection then each of these have information right. So when we are looking at this information when I say information the size of the road, length of the road, then on the side of the road you have number of the buildings which again represented as a entities.

So each of the building size is either the building information either what kind of building commercial, industrial or residential so that kind of information is associated with it. So

representing this is what makes it more complex. So this complex handling that can be done with GIS but the data that is collected must be handled in a very cautious manner.

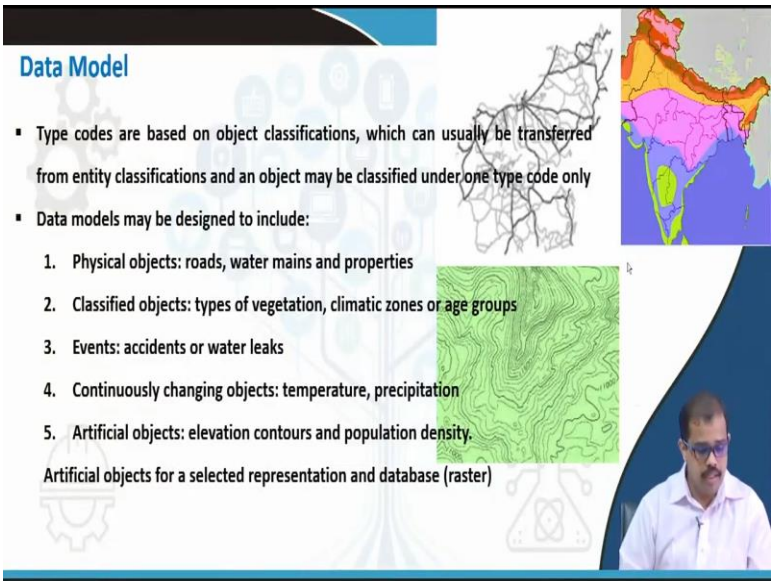
And most I mean most importantly it has to be a good data. The data collections should be extremely data collection is what extremely complex. And when we look at this, this implies the information carrying units and their magnitudes must be selected. That is what I meant. So before you go to the ground, before you understand on the ground what kind of information you need.

You have to make a list of information that you may need when you go and collect a data. Information that a particular person needs, and other person needs may not be the same. So what kind of information for your analysis is required is has to be first and foremost data mind. What should the data mind information and the data need, you go the ground and then look at what are the different data that you have to collect.

It become extremely easier for you and the data quality would be much better otherwise you will be scrabbling the information, some data would be actually missing in certain where the entities and that will lead to issues spread to your databases ok. For example as I gave an example of this road, if you want to divide this roadway into different sections it would have selected before that roadway how this particular road ways describe.

Can the roadway be describe in the proper manner if it is described in proper manner then it can be divided into different sections ok? So the first thing is look at the road understand the characteristic of the road and understand what is the headway of that particular road. So then you will able to understand what are the different sections. How can we define this section? What are the different information that has to be collected for each of the sections?

(Refer Slide Time 09:56)



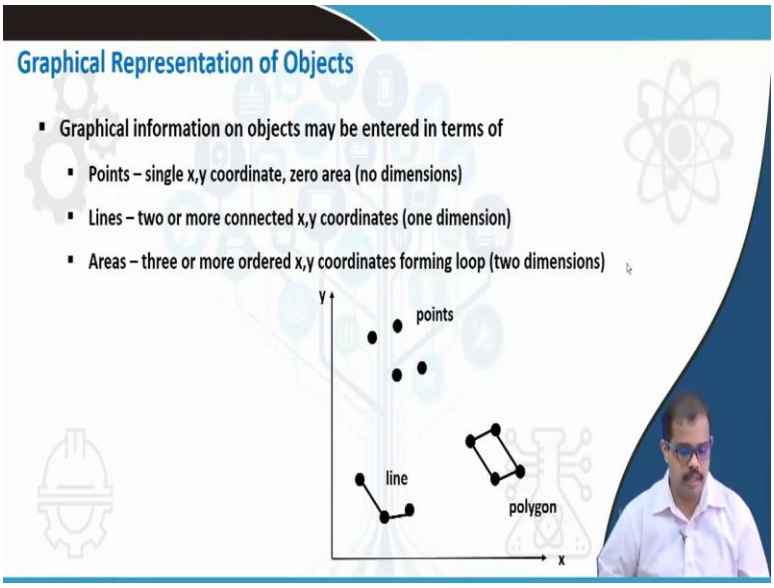
Then when we look at the type codes is the thing that is actually another point is extremely important I spoke about a the ID's before so these are based on object classifications which are usually transferred from entity classifications. I will show you when I am explaining the database how this object classification are done using the type codes. And then object maybe classified under one type code only.

So this is what is may you may call it as limitation but it is also an advantage. So it can be only one type code. If you have multiple type code then you are actually losing it on multiple front because you may not get the exact data that you may be looking into fit into the data model. So it has to be only one unique type code for every classification or every entity. So data models may be designed to include a physical object, you can have physical object as I previously said in a form of a geometry.

So it is road, water mains and it is properties classified objects it can be types of vegetation, climate zone, age groups. So these are the different types of classification events whether it is accident or water leak for example. Then continuously changing object for example temperature, precipitation if you have a real time sensing it can be easily fed into the database without much hassle in today's scenario. Then you have artificial object like elevation contours, population density. So these are just selected representation and of a data and database (()) (11:33) raster model.

So you can have a have any kind of data that is connected any kind of attribute that is connected, any kind of geometry data that is connected with a relationship. So this data model can have all of these but please remember that the type codes of unique kind is must be completely unique. That is why it is called as unique diaries. So you should be complete it should be completely unique so that you get some valuable outputs from that database.

(Refer Slide Time 12:02)

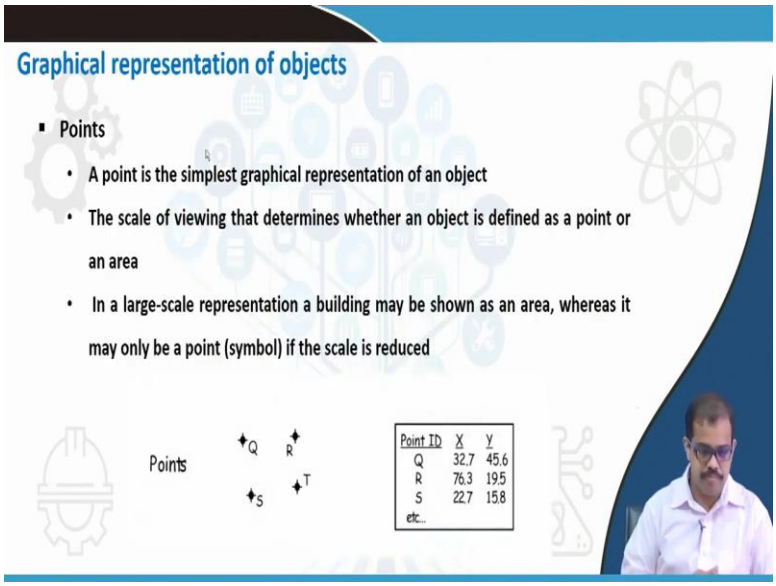


So when we are look looking at the graphical representation of this objects as I previously said this objects can be entered in terms of points where you have single x, y coordinates it has zero area. It is only representing that particular coordinate particular point if you for example when you look at you are going to one place another so the first thing is if you are putting your point in your google maps in your mobile so that point is representing your place of stay, that is a point.

So then you will give the point of the destination so that is nothing but a line. So two or more connected x y coordinates connected points becomes a line. So it is always one dimensional it gives you the route ok. If you are representing the entire dimension of where you would actually stay then it will become sub polygon. It your house is nothing but a polygon. Number of connected line is nothing but the polygon.

So points can number of points can be represented as line, number of line segments is then represented as a polygon. So connected line segment and close line segment is nothing but a polygon.

(Refer Slide Time 13:19)



So when we look at points, a point is the most simplest graphical representation of an object on the earth surface so be careful. You are representing an object ok. You are not representing a phenomenon or an area it is an object. So in a large scale representation a building maybe shown as an area, whereas it may be only a point. So this is where people make a mistake. So when you are representing a large scale you represented as area but it is a point only if the scale is actually reduced.

When you are actually looking at the entire image so buildings can be a point, can be a polygon. So it dependent if it is a very large scale region because for example if you are looking at the entire campus of the IIT Kharagpur then I focus only on a particular building and try to map it maybe 3, 4 building then it becomes a large scale maps. So you have every information about that those four buildings so this can be represented as a polygon.

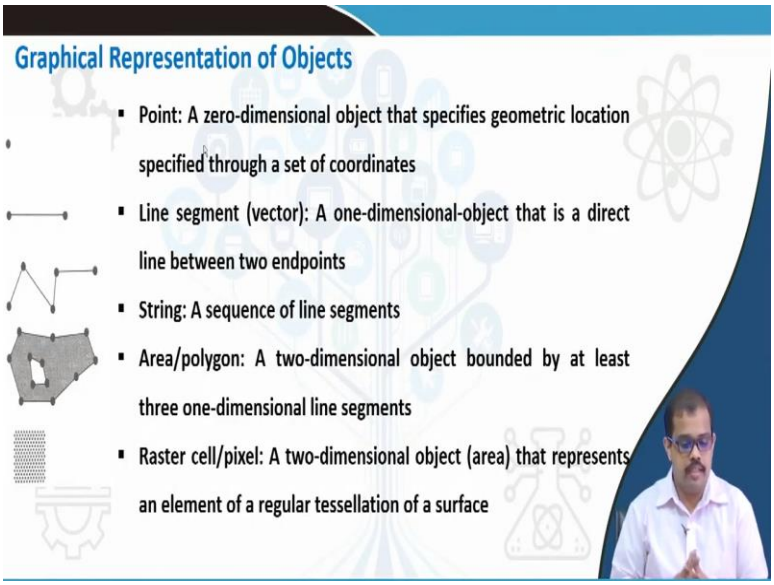
Whereas if you are looking at the entire IIT Kharagpur so each of this building is nothing but an point so that is represented as a point. So when we are looking at lines, line connect at least two points lines always has two points, so they connect at least two points. So and I used to represent objects that may be define in one dimensional. I gave you example of your house and your destination that you have reach to so that you used your Google map.

Then you have properties boundaries are typical lines, as a electric power lines, telecommunication lines, cables etc., Roads and rivers on the other hand maybe either lines or areas. So you should be extremely careful. Roads can be lines in a very coarse resolution data. If you are if you have very good resolution data then roads can be even polygons ok. So depending on this scale that you have used then if you are representing polygons, polygons are used to represent objects defining in two dimensions, a lake, a area of a forest, a area of a building, a area of a town these are represented by areas or polygons.

Again when you look at the physical size in relation with the scale determines whether an object is represented by area or a point. If I have told you before when I am explaining the points so what kind of scale that you are actually representing whether you are representing the entire IIT Kharagpur campus or representing only the 4 buildings in the IIT Kharagpur campus. So if you are representing 4 IIT four buildings in the IIT Kharagpur of campus each of building becomes an entity and each of this building will be in a polygon shape.

But if you are trying to take the entire vast campus of IIT Kharagpur supposed to be the biggest campus in IIT's so each of this buildings will become a point. So please keep this classification in mind an area is delineated by at least 3 connecting lines you should have at least 3 connecting lines may get a polygon. Each of which comprise a number of points. As I said line point number of points form a line, number of line segment form a polygon. So this loop has to be understood properly.

(Refer Slide Time 16:46)

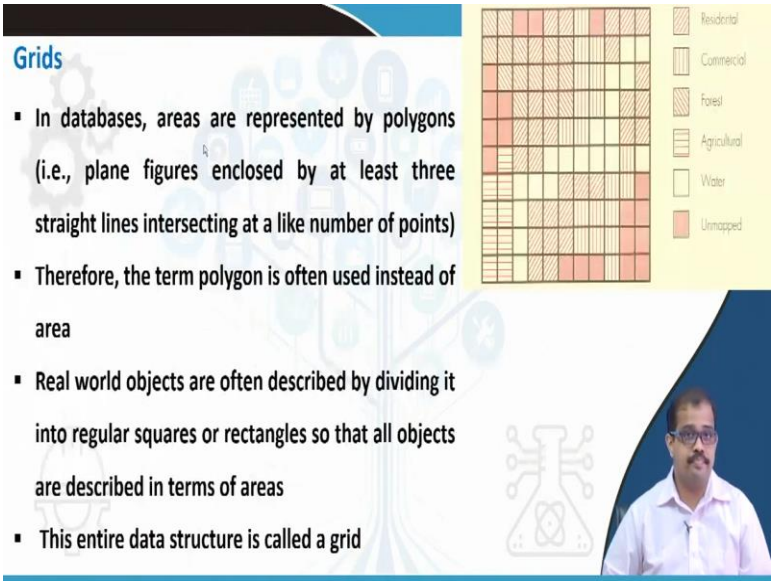


So when we are looking at the graphical representation of point it is just I am trying to revise what I spoke in last few slides. So points a zero-dimensional object that represents that gives you the specific geometric location through a set of coordinates. Then line segment a one-dimensional object that is directly line between 2 endpoints 2 points that are actually connecting. Then you have a string a sequence of line segments, so number of line segments. So that also can be also be representation of your route ok.

So if you are looking at a polygon a 2 dimensional object bounded by at least 3 dimensional line segments ok. So finally, I would like to introduce one more concept here is a raster. So what we are looking vector data model now if we look at the raster is number of cell or a pixel. This I have spoken in the previous lecture. So if you zoom in your photo you can see a completely a box for example you can just take a photo in you mobile phone and keep on zooming it until you find only boxes where your software will not be able to zoom in further ok.

So look at that so those are called as pixels or cells. So each of so collection of the pixel in definite number of rows and columns it represents an raster image or a raster model ok. Any of your photos, satellite data, any of your images that have been considered is are raster data ok in a form a photo etc., is raster data ok. So this can be again associated with geometric object, it can be associated with geographical data, it can be associated then with the database. So it depends on how you are actually categorizing in this raster data.

(Refer Slide Time 18:43)



So when we look at the raster data the next concept that we have to understand is grids. So for example on your right hand side I have represented if I have consider this particular area in a particular region ok. So if I am representing this area and in different region I would if I classify it as a residential, commercial, forest, agricultural, water, unmapped. So what I do is I define this in equal number of grids ok.

So number of straight lines in intersecting at a like number of points. So it is equal number of square boxes fine. So you have equal number of square boxes. So each of this box representing a polygon ok. So always when we are looking at technically the vector model, so we or a raster model we use a terms a point, line and a polygon and a grids or pixels etcetera. Not we do not use terms such as area.

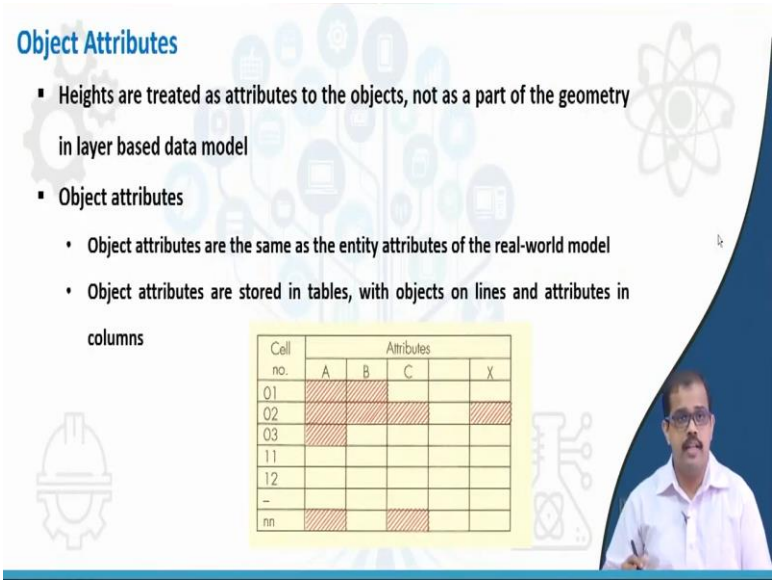
So when we are looking at this these are the number of polygons that are been represented. Each of polygons are represented by real world objects ok which are in regular square and have all terms of the area. These are representing in each area. So all of this polygons have each areas each for example this residential area is represented as this classification. So wherever you find of this particular classification symbol so that is nothing but the residential area in this region right.

So this is how the data is represented as grid. So the grid data has is a collection of cells, collection of equally space equal area cells or if I have to say technically number of polygons.

Each of this polygon have real world information. All real-world information is then coded in terms of certain classification procedure or classification types. So each of this classification types are is as mentioned here. These are the different codes. This is just an example.

So there are lot of what we call it as cartographic instrument the way you want to classify based on the instruments you can look at the presentation. But normally you have a certain standard of representing each and every object. So that is how it has to be represented. It cannot have your own way of representing every phenomenon on the earth surface.

(Refer Slide Time 21:25)



So finally, when we look at when we looked at grid we have something called as object attributes. So heights are treated as attributes to objects not as a part of geometry in the layer based model. For example, what I am speaking about is previous example which I am speaking about the building. So normally heights can be treated as attributes to objects ok not as a geometry. If you have a geometry of a building let say you as have given a example of IIT Kharagpur if they have collected information about 4 buildings.

Let us say I have first building is computer science, that other building is electrical engineering, third one is you have (()) (22:08) block where my particular department is there. Then you have maybe chemical engineering. So when we look at all of these 4 building ok if I am trying to give you a height of the building then it becomes an attribute of an object or an entity ok. Rather the geometry of a building for example what is the way of that building is represented whether it is

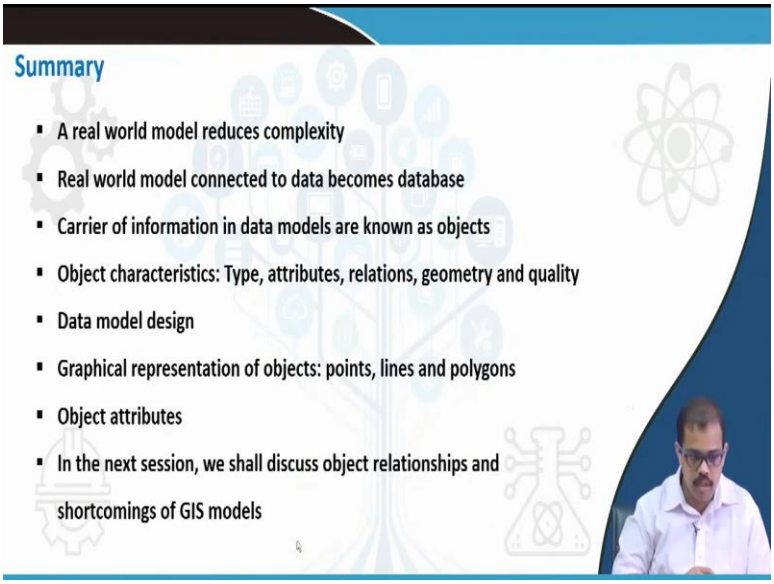
square, it is hexagonal or whether it has a random shape or whether it is rectangle or whether it is a circle.

So all of these becomes a geometrical shapes ok this becomes a geometrical way of representing that particular object, but height is always is an attribute. This is where student actually makes a mistake. Always they try to put height as a geometry, so ones you put height as a geometry then it is actually representing a 3D format of that particular data which is actually not representing the relations ok. But you can form relations from a 2D data to a 3D data ok.

So when you are looking at object attributes, object attributes are the same as entity attributes of the real world models ok. So that is what I said object and entity are go hand in hand. Object attributes are always same as the entity attributes. Object attributes are stored in tables with objects as lines. So this which is the thing that everyone has to understand objects are lines ok on lines and attributes in columns ok.

For example, I have table here which is actually collected as a database. So when you look at this these are different objects that are represented here cell no 1, 2, 3, 11, 12 to n. So these are nothing but the objects whereas these are those attributes that are stored in each of this cells. So that is where you have an object you have number of attributes that is actually connected to a cells.

(Refer Slide Time 24:15)



So to summarize this particular session we have a real world model that actually reduces that has complexity but you have to reduce the complexity in order to put it into a data model. So real world model connected to a data becomes a database. So ones I have represented in a data model so it becomes a database. The database always has certain geometric information these are nothing but your carrier of information's ok.

I spoke about point line and a polygon. These are represented different information, different size of information, different entity of information. So these are nothing but having the different kinds of information embedded. So we have object characteristics for example the type, the attributes, the relations, the geometry and the quality. So these are different object characteristic. Then you have a data model design.

So I will go to details of data model design but as of now ahh please understand it is in the form of grids, in the form of attribute, how the attribute the objects are actually put in the form of a gridded data. So then you have graphical representation of points that is in the form of point line in the polygon. So we have looked at object and attributes, how the objects are put in.

You always objects are in the form of number of rows and you have columns the each of these columns are represented in a form of as attributes. So object and attributes so these have relations. So as of both of this class so the very important concept that we have learnt is how the real world translate into a data model.

So now the next session of the lecture we would look at how the object relationships are maintain and what are the difference shortcoming of these models this GIS model and how do we look at GIS as a tool which can be a extremely useful in our analysis. That is what we would look at so the that that is it for today. Thank you very much.