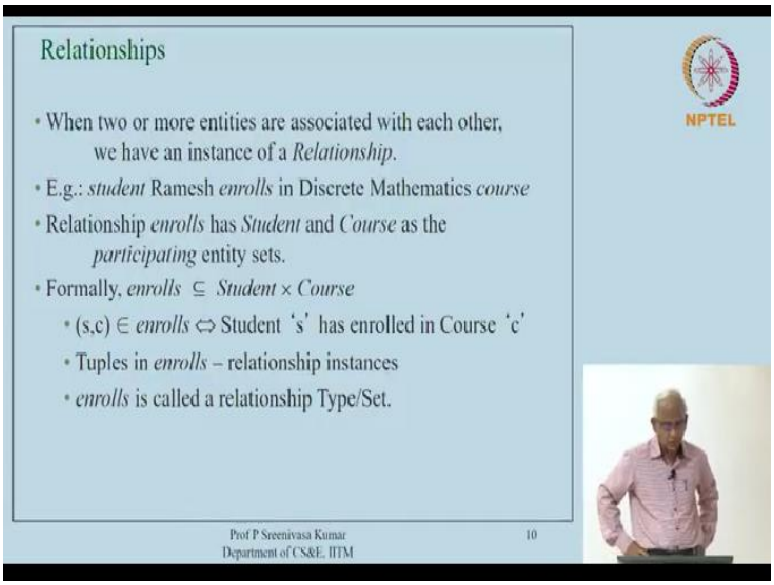
Database Systems
Prof. Sreenivasa Kumar
Computer Science and Engineering Department
Indian Institute of Technology-Madras

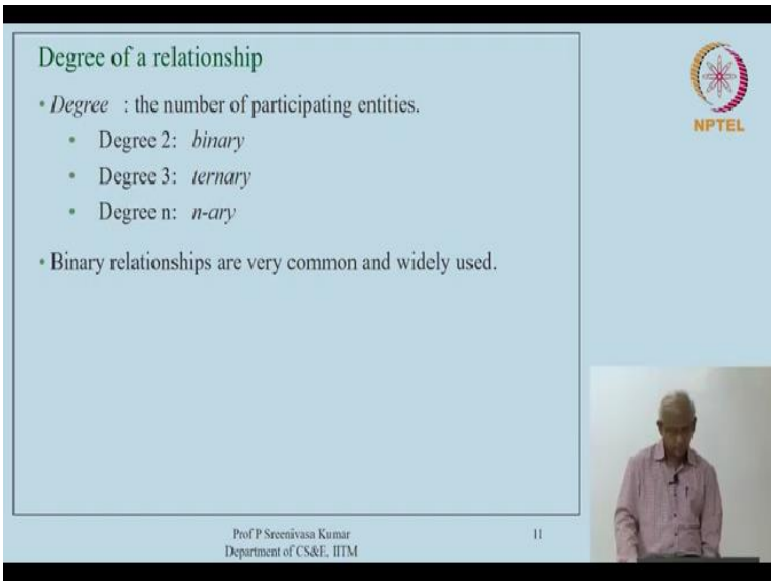
Lecture-05
Entities and Relationships

(Refer Slide Time: 00:20)



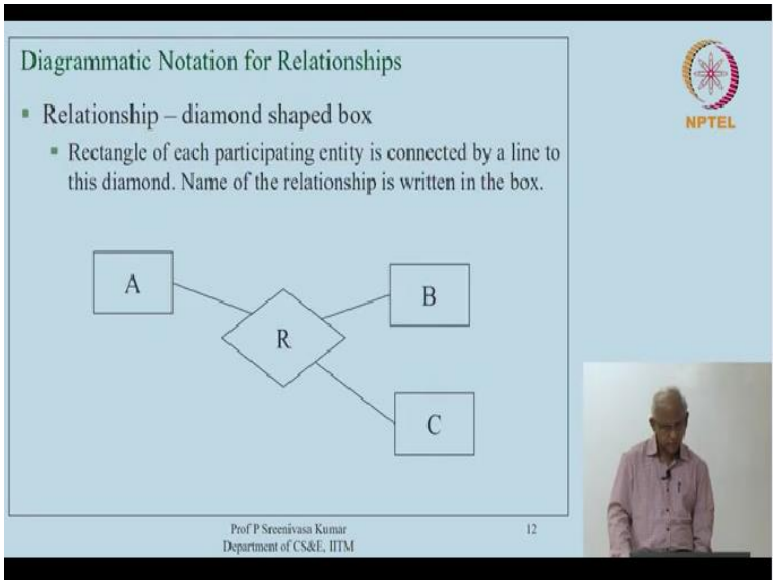
We started talking about this concept of relationships when 2 or more entities are associated with each other, we would like to capture that information into our model and so we have an instance of what we call as a relationship. So, we give an example it turns out to be a binary relationship, but in general there can be binary or ternary relationships.

(Refer Slide Time: 00:40)



So, we also talked about what is called a degree of a relationship which is if it is 2 participating entities then we have a binary relationship. Otherwise, if it is 3 ternary then otherwise, if it is greater than 3 then we call it as n-ary relationship. Binary relationships of course are.

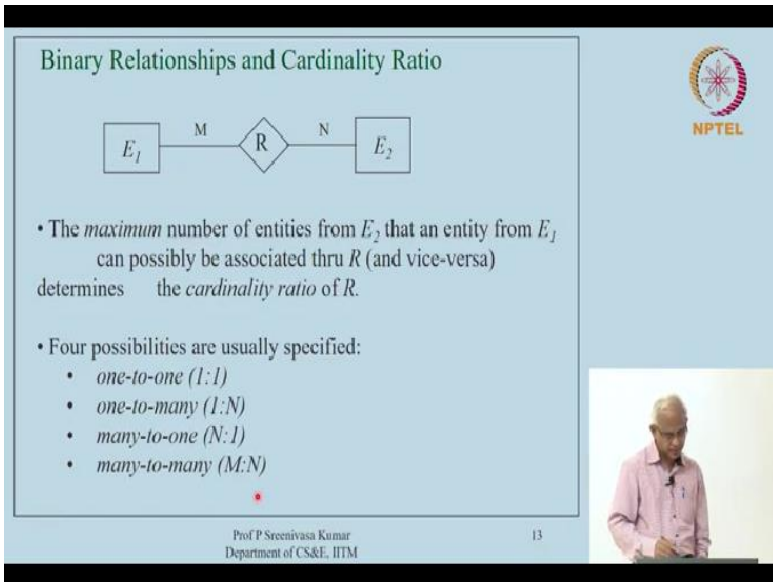
(Refer Slide Time: 01:05)



And we also talked about the diagrammatic notation for representing relationships; we represent the relationship as a diamond box, and then draws all these participating entities around it and then connect the participating entities to the box by lines. That is how we use this in diagram. Now let us move on I want to will there is a lot more information about these relationships that we would like to capture in general.

And so we will focus on binary relationships because they are the ones that occur more frequently than other kinds of relationships.

(Refer Slide Time: 01:50)

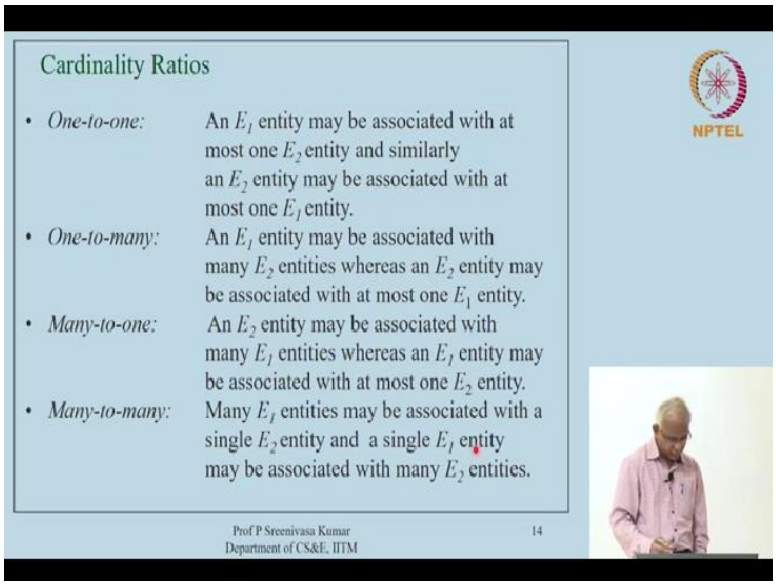


So, when we talk about binary relationships, one of the important things that we talk about is what is called a cardinality ratio of this binary relationship. So, it is denoted by having these kinds of numbers on these lines that are connecting the entities to the relationship. So, this is binary because there are 2 entities that are connected with the relationship or relationship type. So, I can use this relationship type and relationship interchangeably.

But we often be talking about relationship types okay. So, what is the maximum number of entities from E 2 side that are associated with E 1 and the vice versa. Kind of determines what is the cardinality ratio, we call it as a cardinality ratio of this relationship R okay and usually 4 different kinds of cardinality ratios are used to kind of model the different possibilities of these associations.

So, these are the 4 possibilities are usually specified like this, we call them as one to one, like written as 1 : 1 one to many 1 : N, N is generally stands for many, so and many to one and many to many. So, we will define these things in the next slide.

(Refer Slide Time: 03:33)

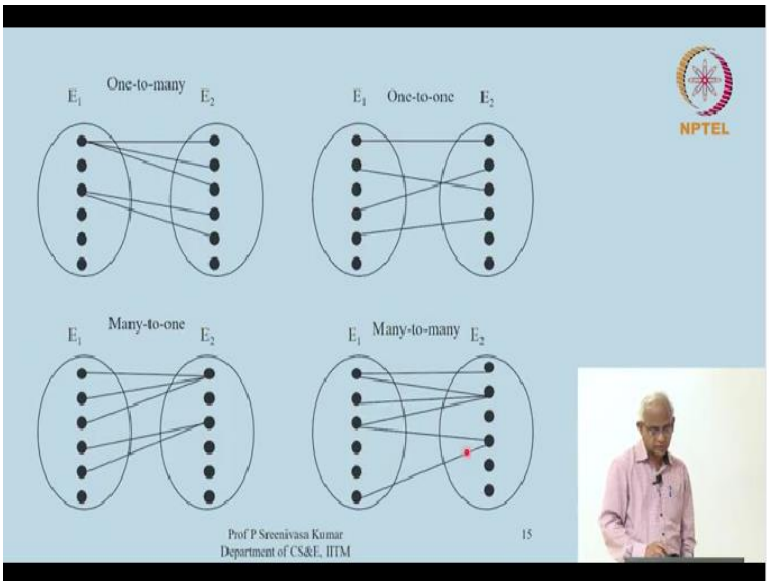


So, it is one to one means an entity E 1 may be associated with the utmost one entity. So remember that we are saying that it is the maximum number of entities that it might be associated with. So, an entity even may be associated with that most one entity E 2. And similarly, an E 2 entity may be associated with that most one even entity, if that is the case, we call that kind of relationship as a one to one relationship.

On the other hand one to many is that an E 1 entity may be associated with many E 2 entities. Whereas, the E 2 entity may be associated with the utmost one E 1 entity okay and inverse of this yeah situation is many to one. That means an E 2 entity may be associated with many E 1 entities, whereas, an even entity may be associated utmost E 2 entity. I will give you examples of these different kinds in the next few slides.

Finally, we have many to many cardinality ratio in which many E 1 entities may be associated with a single E 2 entity and a single E 1 entity may be in fact associated with many E 2 entity. So we will look at some examples to make these things clear.

(Refer Slide Time: 05:14)

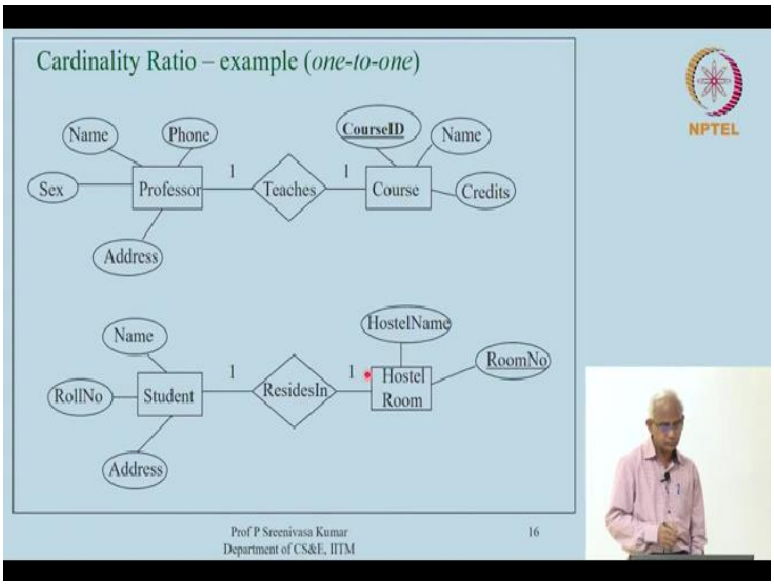


Before that I have a picture, which kind of you know, gives you clarity about this. So, this is one to one is here. So, you can see that, remember there is maximum, so this is some of the entities might not be associated with any entities on the other side. But here because it is one to one, if it all it is associated is a maximum of one. So, it is like that, here is one to many. So, each of these things may be associated with many E 2 entities.

And so that is the max. So, max is many it may not be associated with anybody else. Whereas here these E 2 entities associated with utmost one E 1 entity, and the reverse is the kind of situation in many to many, many to one relationship, as many to many, many E 1 entities may be associated for example, all these three associated with this and then you know these 2 entities are related with that.

So, I will show you examples of this in practical, we will take this the educational situation model that we are talking about and in that we will discuss various examples okay.

(Refer Slide Time: 06:40)



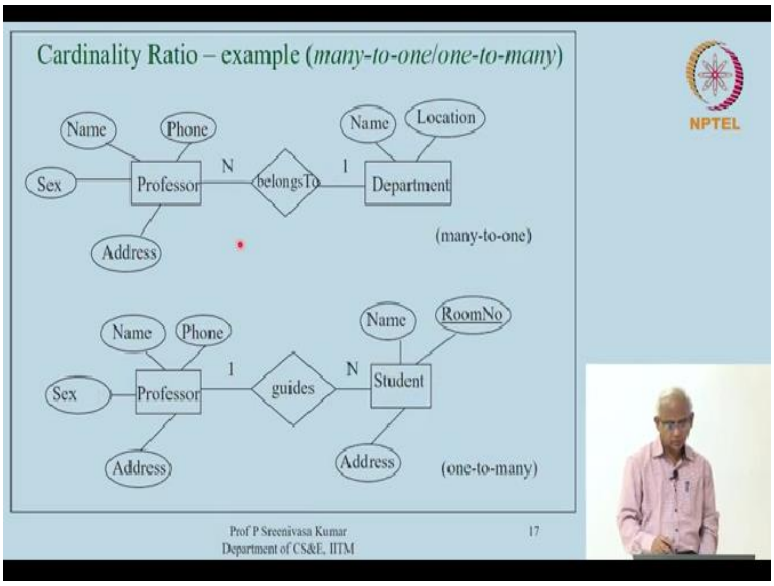
Let us begin with one to one. Again the whether a particular relationship type satisfies these you know kind constraints depend on the modeling the situation at hand situation at ground right. So, for example, if there is a policy in the institute saying that each professor, teachers, you know one course. So, no 2; professors do not simultaneously teach the course okay.

So, in such case professor teach a course association called the teachers is a one to one thing, in a similar way suppose there is a policy in the institute that the students have to be allotted single rooms, so to say I mean each person has to be associated exactly 1 room. So, besides in association between students and hostel rooms is a one to one relationship right. So, you actually; if you say that these are maximums.

Then, you know then you might actually wonder that is there a course that is can there be a course without any teacher. So, we do not normally expect that in a academic in this kind of a model, if a course is there and it is to be taught by somebody right. So, we will also be have to be bothered about the minimums right whether it is associated with at least one professor or not, we cannot allow a free floating course without nobody coming in to teach that course right.

So, we are to bother about mediums also we will come to that a little later okay. So, that is how some of these relationships will naturally be one to one.

(Refer Slide Time: 08:44)



Moving on here are one too many and many to many, many to one kind of relationships. Let us say we have a policy in the institute, that each professor has to belong to exactly one department, again, this is a ground reality a policy in the institute it is the, the institute might as well you know, adopt a different policy in which it allows professors to be associated with multiple departments as well okay.

But let us assume that in this situation each professor administratively belongs to one department, okay. In which case and also many professors will belong to one department. So it is a many to one association. Also, you might in this situation; it is also natural to assume that they cannot be free standing professor without belonging to any department. So, every professor must belong to one department. So that is a minimum associations that we are.

So in the association we want to also say that the professor has to belong to any part. Here we are now saying that he is utmost one department whereas a department should have at least one professor and can have several departments several professors okay. So like that we then this is a many to one relationship. So right now we are only bothered modeling the maximum. So we will also like see later how exactly to take care of the minimum participation also.

Now, let us look at these professor guides a student. Okay, here is a one to many relationship. So each professor can guide multiple students and student is always guided by utmost one professor.

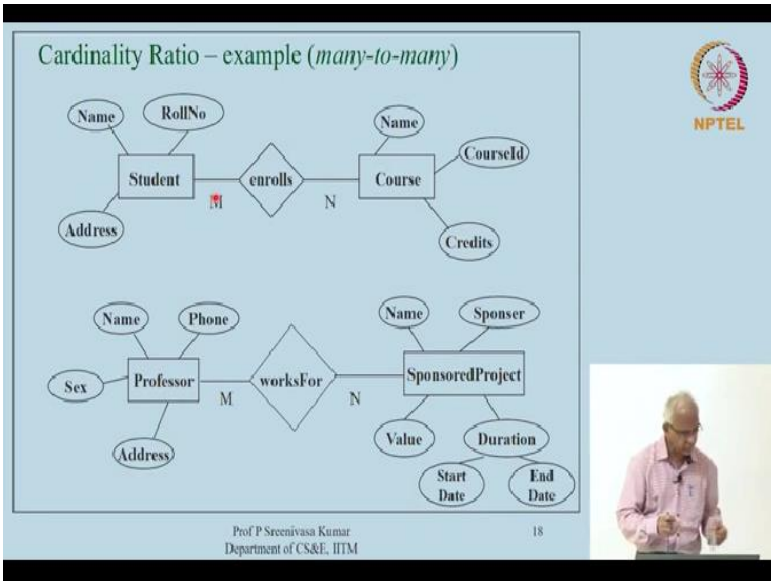
Let us say there is also a policy marriage. It is not sacrosanct. But you know let us say at the institute has adopted a policy saying that each student needs to be guided by one professor. So that he is not too confused about what to let us say.

From multiple professors guide the student, maybe there is a lot of confusion as to what to do right. So, let us say the institute has adopted a policy thing that a student has to be guided by one professor. And in order to get it too many students the institute obviously says that you should guide many students okay. So each professor should guide many students. So, the association between professor and student captured by this guides is the one to typical example of a one to many relationship okay.

So, are there any questions we are so these additional information as to whether an association is a one to one, one to many, many to one, many to many is in fact telling us some semantic information about the domain and we like to capture that and record it in the when we have in this notation or in a similar note other alternative notations up there. So, somehow we need to keep track of this information that.

Because that that is one the modelers will understand the domain in a better way. And later, of course, we will see how exactly this entire thing will influence the design of the actual design of the relational database system.

(Refer Slide Time: 13:03)



Now, we also have an example of some examples of this many to many relationships. So students, typically, so if you look at this enrollment relationship between students and courses, it is evidential many to many relationship. So a student enrolls for in a semester, many number of courses, the student is associated with many courses okay, and the course has many students enrolled for it.

We have 52 students in this class, like that many students are enrolled for 1 course, and a student enrolls for something like 5, 6 courses right. So, the enrollment relationship between students and courses if you see international, many to many relationship. Many entities on one side are associated with a single entity on the side and similarly many entities in this side are associated with a single entity on the other side.

Now, I have also given some okay one more example here is the works for relationship between a professor and as sponsored project. A sponsored project is a you know that is funded by some sponsor with certain amount to kind of carry out certain research work during a particular duration so like a 3 year project from funded by department of science and technology we have with value of 25 lakhs is given to a professor to kind of investigate some issues.

So, like that is sponsored projects entity can be imagined. And then professors are associated with sponsored projects. And typically the sponsors might say that at least 2 professors have to

be associated with a project because they cannot rely on I mean, they do not want to administratively make exactly one person responsible what if that person particular person, you know, they will have some health issue.

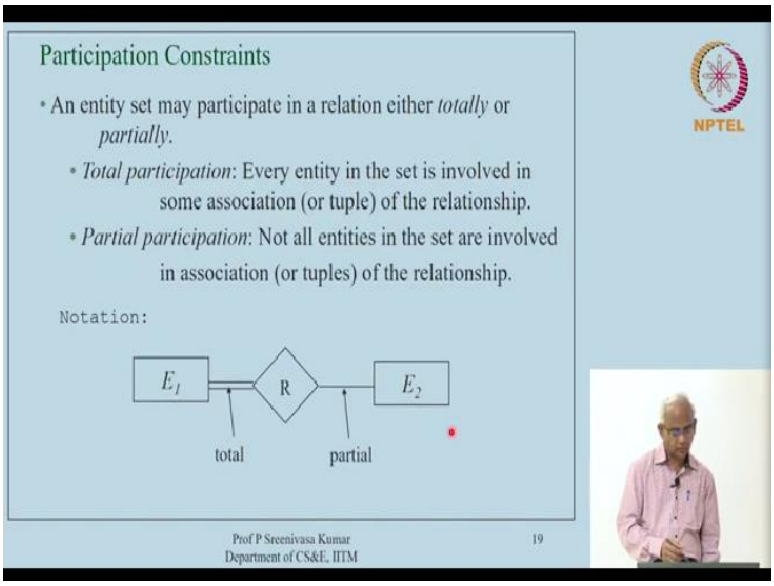
And then he has who will take care of the project. So at least 2 or more 3 professors associated with a sponsored project and a professor might have multiple research interests and might make multiple professors to multiple funding agencies and then obtain multiple projects in which he or she is playing a role. So, naturally professor might work for several sponsored p projects.

So, we have a situation which is an example of a many to many relationship here that professors work for many sponsored projects and each sponsored project has many professors working for it okay. Now, so at this stage of the development of the information system, we will use all these concepts, the concepts of entities and relationships and capture the domain that we are modeling.

And then later when we talk about relational data model, we will see how exactly we will translate this into relational data model. One suggestion that I have for you is that you know you must while you are using one data model, just stick to the concepts and the concepts that are tools that are provided in that model. So you should put that cap so to say, you must put the ER model cap and then think about how to model a situation.

Do not keep on thinking as to how we are going to, you know, translate this into a relational data model at this stage itself. We will have to talk about it in detail later okay. Now, this is one important aspect of relationships, typically, the binary relationships.

(Refer Slide Time: 17:51)



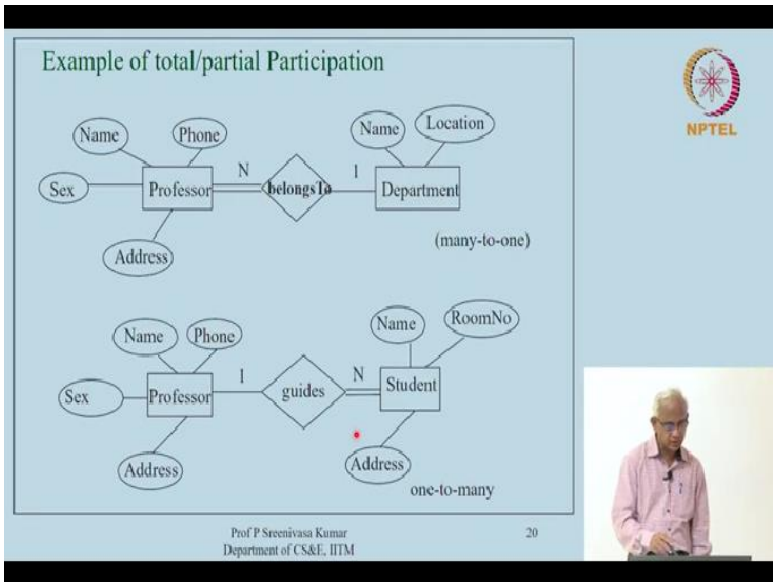
On the other hand, we also have what are called participation constraints. This will talk about the minimums as I was mentioning sometime earlier. So, an entity set may actually participate in a relation either totally or partially. What do we mean by total participation is that every entity in the set that entity set, for example, E 1 is involved in some association or a tuple of this relationship.

That means it appears at least in 1 tuple. So, that is the minimum right whereas the partial participation says that not all entities in this set are involved in the tuples in the relationship okay. So, the question that we are asking is very entity that is there in this entity set necessarily participate in the association in the relationship and so appear in some tuple or rather is that the case or it is not the case.

So, that is the thing. So, if it is not the case will say it is partial participation if it is indeed it is the case that every entity as you know does participate in some association with the entity on the other side then we call it as a total participation okay and the diagrammatic notation for that is that if it is total participation, then we basically put 2 lines that will connect the entity to the relationship okay.

Now, let me give you some examples of this, if it is just a single line, then it is a partial participation.

(Refer Slide Time: 19:51)



We can revisit some of the earlier examples and then see exactly how. So, for example, if you take this belongs to association between professor and department. We earlier said that they should not be any free standing professors who are not associated with any department okay. So, in our institute we have a policy that even professor must be administratively belong to some department okay.

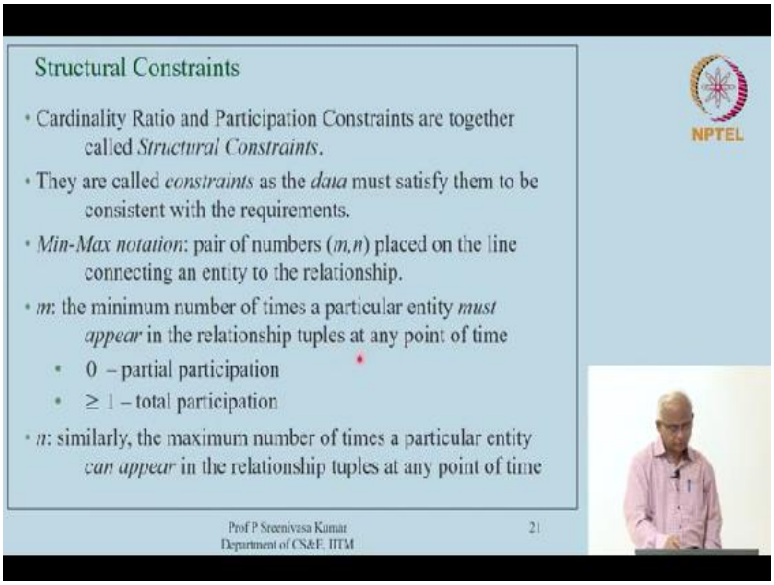
This is not a universal policy, but this is a policy of a particular institute, a different institute might allow some professor to be a some academic to be just be there in institute without really belonging to a particular department. It is possible right, but we here we are talking about a situation where every professor must belong to one department. So, this is a total participation. The relation the association between professor and belongs to relationship it is total that means every professor must appear in some tuple.

And earlier we have said that utmost one. So, now we are saying at least one earlier we have said utmost one. So together now, it becomes that he has to belong to one department and he has to become exactly one department, he or she. Now and if you know, for example, look at the guides relationship. It is not really compulsory for a professor to guide a student. A particular student, a particular professor might be unpopular.

And might not actually get; any racist student to guide, so we will allow that in the system whereas unfortunately the students do not have a choice. They have to have a guide. So, we will capture so if the institute policy is that every student must have a guide whether he like it or not will obviously then the student is having a total participation in the guides relationship. So, you can see that every student has to be in some tuple of other of this guides relationship.

Whereas the professor might not take guiding any particular any student at all, so the one here indicates that yeah, each student is guided by utmost one, utmost one professor and he should be guided by at least one professor is what total participation now says okay. So, together these 2 things, the cardinality ratio and the participation constraints will allow us to capture more detail information about the relationships that are existing between entities and we should use make use of these tools in order to capture that domain knowledge.

(Refer Slide Time: 23:28)



So, together they are called the structural constraints, so cardinality ratio and the participation constraints are together are called the structural constraints. Why do we call constraints is that we could probably call them as structural properties also, but then we will call them as constraints because later on we will see how, you know, the system, the information system that we are developing, might be actually asked to enforce this things, you know.

If there is a policy of this kind saying that, you know, a student has to be guided by at least one professor, and then we built the system based on these assumptions, and then if there is a student without any guide, then the system will refuse to you know, accept that situation and then prompt you saying that you must supply this information to me, because in your model every student must have at least one okay.

So that kind of so you can think of them as constraints also that is why they are called constraints, because the system can be held responsible for enforcing them. So that the information that we have correctly reflects our assumption support the window mind. Here is an alternate notation which nicely captures both these 2 things together you know. So, this is called min max notation.

A pair of numbers m and n where you know n at least has to be at least 1 is placed on the line connecting the entity to the relationship and what is the meaning is that m is the okay. The minimum number of times entity must appear in the relationship tuples at any point of time. And n similarly is the max. So min max right, this is the main and that is the max. So, minimum number of times a particular entity must appear in a relationship tuples at any point of time.

So, if it is 0 then it captures partial participation right. So, the student that particular entity may not appear in any. So, even 0 is allowed. So that basically takes care of partial participation and greater than equal to 1 some number, okay. You might want to actually if it is you know, you do not know the actual number and if there is no such okay, we will come to that.

So, it should be greater than equal to 1 for total participation in a similarly the n is the maximum number of times a particular entity can appear in the relation. So this is what we have done in cardinality ratio while talking about cardinality ratio, what is the max is right we talked about that. So at any point of time it can. So what the cardinality ratios basically said is that it is many without putting any actual number.

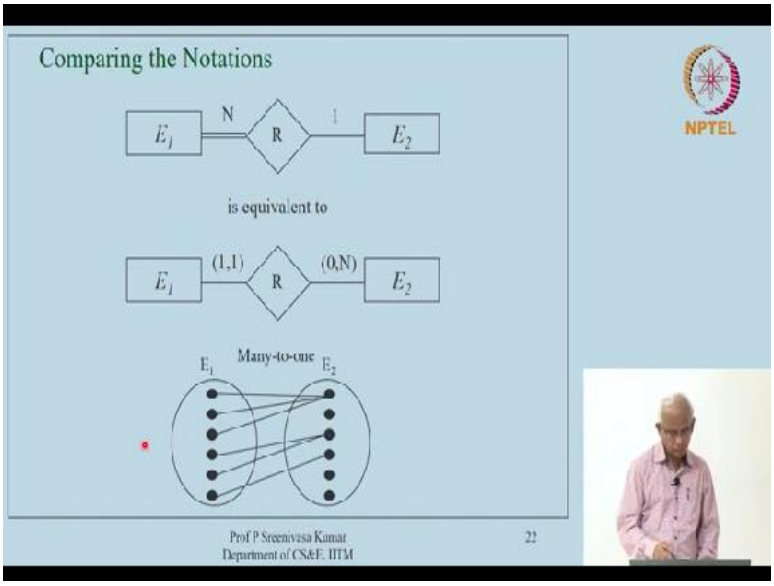
But now we have this facility to kind of even make the further precise and then put a number saying let us say the institute has a let us our department, for example has a policy saying thing

that each professor can guide utmost 4 students, some number. In order to ensure that there is a good distribution of students across professors, let us say we adopted a policy that each professor can guide utmost 4 PhD students.

So, such information we can gather and use convey that using this max thing here. So then you can put the number 4 like for example there to say that this professor can appear in 4 number of this association tuples with 4 different students. So, that way this min max notation as far as binary as relationships are concerned will capture these cardinality ratio and participation constraints together.

And in fact, slightly generalize them to kind of make them even more precise, because we can convey the instead of simply saying many, we can now put some in bond on the put a contract number there. So, it is useful to have such kind of facility.

(Refer Slide Time: 28:44)

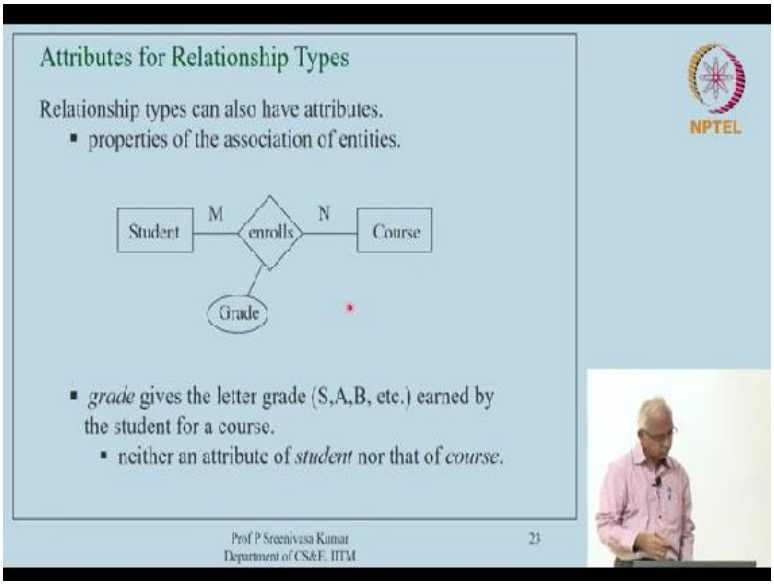


So, for example, such a many to one relationship will now be equal to this 1,1. So, it is a minimum of 1 and the maximum 1. So, for example, you can take faculty member belongs to department. So, you must belong under itself total participation. So, he must belong to at least 1 department and utmost 1 department and departments might not have professors can have as many professors as is possible.

So, there is no upper limit on the number of professors that they have. So, supposing in the institute says that your you know, the maximum number of faculty that you can have as determined by the student faculty ratio is some 40 or something like there is some number, then you can probably put that number there okay. So, that way these notations are very useful to capture the domain knowledge that we are modeling.

And they are not so difficult to kind of you know kind of convey to a person who does not know computer science. So, when we engage in discussions with the stakeholders for whom we are building the information system, these concepts are not too difficult to kind of explain to them. So, we will explain these concepts for them and talk in terms of these things and then gather the information that we are supposed to model. So, these are rigorous enough to capture the information, but simple enough to actually explain.

(Refer Slide Time: 30:45)



Now, moving on it is possible that any questions and whatever we have been talking about, it is possible just like you know additional information about entities was model using the notion of an attribute, you know, an attribute says entities have attributes like name of the student, roll number of the student and so on. We talked about attributes of students. In a similar way this relationship types also can have attributes okay.

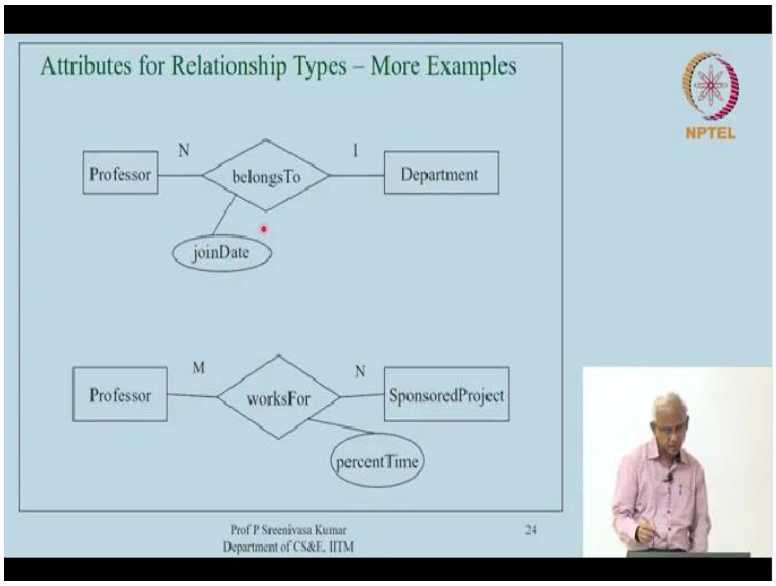
For example, this is a very typical relatable attribute right. So if a student enrolls for a course, at the end of it, he or she will get a grade letter grade saying that, yes you got the highest grade yes or the next highest grade A or B etc. your performance in this course is graded like this. So you get a letter grade. So this grade attribute is has values from these symbols that are used to represent cord and the performance shown by the student.

And you can also now typically see that this particular attribute is in fact, the attribute of this association between the student and the course, it is not an attribute of the student alone, because the student typically takes several courses, student entity they cannot simply have a letter, you know just associated with that, because the student and the course combination is the one that has this letter great associated with it right.

When the student has taken that course, this is the grade that he or she is obtained, in a similar way it cannot be an attribute of the course, because what do we mean by giving a letter grade to a course right. So, when the student has taken a course, then that association is the one that has the grade attribute. So some many times it is natural to kind of you know, give this facility of having attributes to the relationship types also.

And so we do give that option and so attributes can be associated with relationship times. Of course in a okay, we will show you some.

(Refer Slide Time: 33:24)



Maybe some more examples of that, now, for example, professor belongs to department it is a many to one relationship that we looked at earlier. The join date could be a possible attribute as to when the professor joined that particular department, where is a natural attribute because he is associated with the department. If you observe everything keenly here, since each professor belongs to exactly one department right.

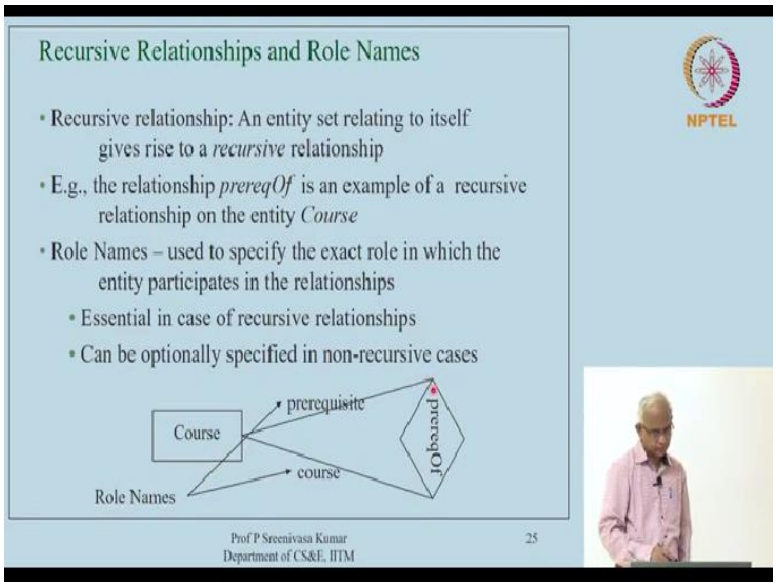
So maybe it is possible to migrate this attribute to the professor right. It is possible to kind of migrate this to the professor. But we could as well associated with the relationship. It cannot be migrated to the department. Obviously, the joint date whose joined date are you talking about right with the department cannot have a joint but it kind of naturally belongs to this association between professor and department.

It is possible that a professor has served one department for some time and served another department little later and so you might want to actually have this joined associated with the professor department combination okay. Now, here is another attribute. We talked about professors getting involved in sponsored projects. So, if a professor is involved in sponsored projects, this sponsoring agency might like to know what is the percentage of time that you are going to devote to this project okay.

You have certain working time; you have teaching and other commitments okay. So, but you know, we would like to know how much is you know is given to this person particular sponsor project. So, the percentage time given to a sponsored project is a possible attribute in this context. Recall that this is a many to many relationship and so, the person time is in fact is, you know is a attribute of the association between professor.

And a particular sponsored project. So a professor might have 3 sponsored projects, and then might say that okay, I am going to do 10% of my time to each of these 3 okay, so these are interesting scenarios, which we can capture using these conceptual tools that are available in any questions in what we have been talking about.

(Refer Slide Time: 36:30)



Here is another small interesting idea called recursive relationships and associated with that is roll names. We an entity set might reflect itself; and does kind of give rise to what is called a recursive relationship. I mean, this is like, you know, the in database palace they use this term. But you can it is not the same sense of you know, recursion as we see it in other contexts. Basically, what we have here is that there is a relationship type.

And in that relationship type a entity is participating more than once. So far, I said there are some participating entities. So, I never said that a particular entity should not participate to more than once in an entity in a relationship type right. I will show you an example of such a kind of a thing

right here and the prerequisite off. Prerequisite of is a relationship and you can see that the entity course parts but 2 times in that in the sense that.

One as a course, the other one as the prerequisite course for the course, so we have to then kind of distinguish as to how, in what role it is this particular entity participating in the association. And so we have to have what are called roll names associated with the participation of the entity in the relationship. So, that will disambiguate as what roll this particular part entities is taking up.

Now, so here is a picture. So, prerequisite of is a relationship and a course is participating in this relationship. So, basically what we are saying here is that certain courses are prerequisite of other courses. That means, unless you complete that course, you should not take the other course, prerequisite the meaning of prerequisite is very clear to you. So, unless you do data structures course you should not take databases course.

For example so, the data structures course is a prerequisite for the database systems course. So, the course entity which is basically consisting of all these courses is participating in this prerequisite off relationship once in the role of a course and once in the role of a man in the in the role of a prerequisite. So, these roll names become essential to kind of say that how the causes associated with the how entity in general is participating in the relationship okay.