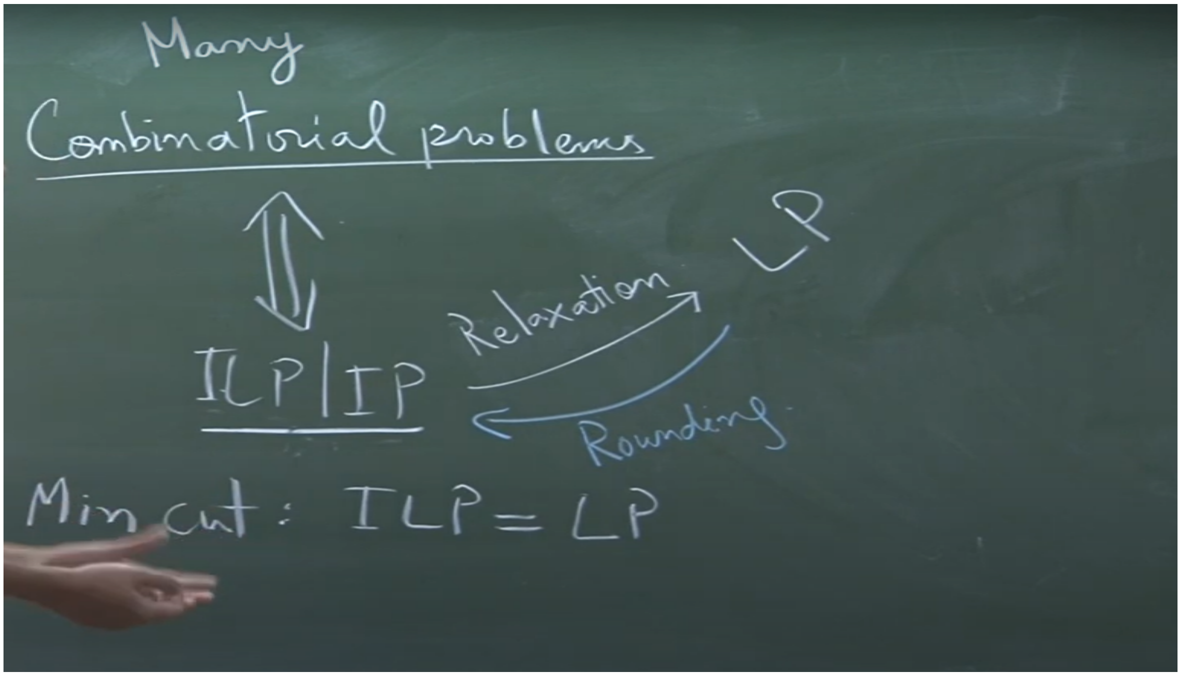
Linear Programming and its Applications to Computer Science
Prof. Rajat Mittal
Department of Computer Science and Engineering
Indian Institute Of Technology, Kanpur

Lecture – 44
Set cover problem

Welcome to the class on linear programming. I am really excited today, because as I said before we are going to talk about this very nice concept of relaxation and rounding right.



So, we have seen many examples of what we call combinatorial problems right covers, pecking, max flow, min cut. We will see another example set cover many of these combinatorial problems can be written as what we call linear program, but with few constraints saying that my variables are going to be integers. So, this is what we know as integer linear programs or integer programs correct. Remember why we started learning linear programs we said, because many problems can be phrased as linear programs.

So, the whole bunch of them can be solved using the same techniques great many of them can be phrased as linear program, but what about more. It turns out many problems are not linear programs, but something close to it linear programs with integer constraints. And question is can we still use our linear programming tools to say something about it right. And so there is an exact formulation in terms of integer linear program for these combinatorial problems.

I am not saying for all combinatorial problems for many combinatorial problems they can be exactly phrased as an integer linear program right. You saw it in your exam, you saw it as an example min cut right there. So, max cut had an LP, but min cut the clear LP had an integer constraint. And then because we do not know how to solve it and we believe that we cannot solve it efficiently do you remember why. Because NP hard problems you think which are very hard they can also be expressed as this.

So, if we can generically solve this. So, one way to prove P is equal to NP when millions of dollars is to show that you can solve integer linear programs right you can try. But for models like us we would say that oh this is this has an extra constraint what if I make it a linear program does it give us at least some information about my combinatorial problem right. So, I take by linear program by integer linear program convert it into an LP this step is called a relaxation. Mostly it is what we are doing is we are increasing the feasible set.

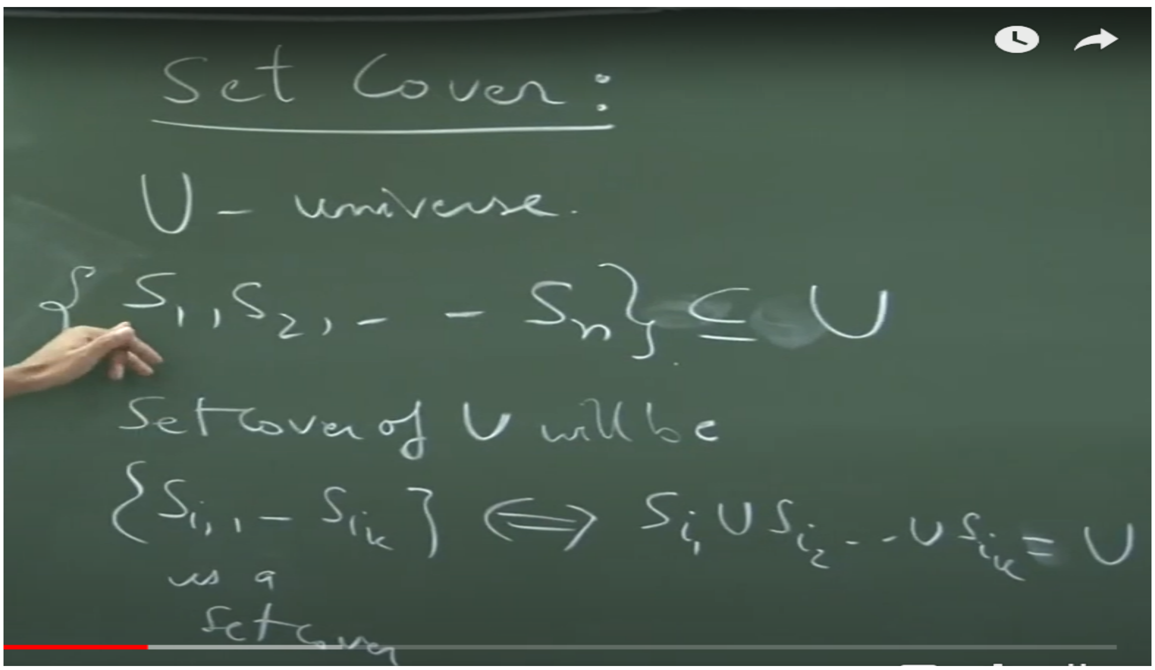
We are saying oh let us assume that our variables need not be 0 and 1 it could be anything between 0 and 1. It is not that our vectors are you know either in x axis or y axis, but you will say any vector which has norm 1 is allowed which has length 1 is allowed something like that. So, we change our domain to be continuous and in that way we are increasing our feasible set. So, now when we increase our feasible set for a minimization problem our value will be lower right. So, if it is a minimization problem we increase the feasible set the value will be lower in the case of maximization it will be higher.

The question we have to ask is does it become too high does it become close or something right. And in one of the cases especially in the case of min cut it turned out that these were instantly equal. There are many ways to see that we were actually lucky in this case right. So, the matrix of special kind here you can if you want to you can show that your all your points all your vertices of that feasible set are going to be integers. And now you know if you have a linear program whose vertices are going to be integers then your solutions are also going to be optimal solutions are going to be integers.

But that is not something which is going to hold for a general combinatorial problem. If you are if you want an exact description if the constraint matrix is unimodular something called unimodular then you can always do this. But we will not go into that we will say what if this is not equal and we expect it not to be equal many a times specifically because this can create NP hard this can capture NP hard problems. If ILP was equal to LP in all the cases we would get P is equal to NP right. So, now, let us take an example of a problem where this is not going to happen and see at least what we can do about it.

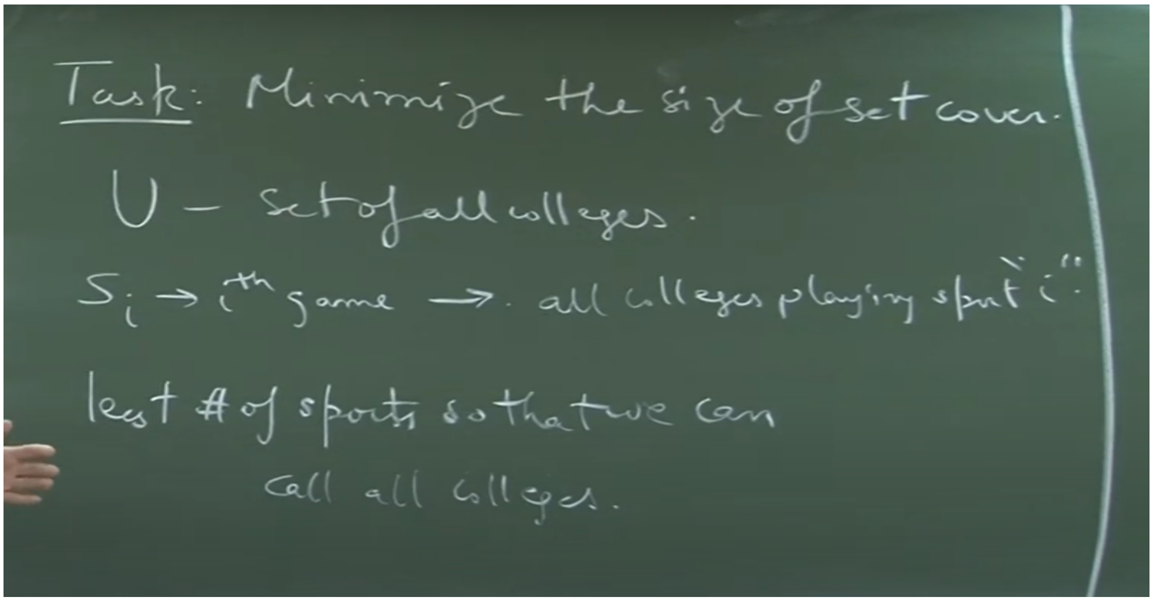
The idea is we will show that this LP value is close to the ILP value or using LP we can find an approximate solution of this integer linear program or my combinatorial problem. This step is called routing intuitively it makes sense we have a continuous solution you want to convert it into a integer solution or a discrete solution right. So, in some sense it is kind of a rounding off right and then in that case what we have to show is that even when we round it we do not lose much in terms of the objective value. Our rounded solution has decent enough objective value ok. So, to describe this process the problem the first problem we will consider is called a set cover problem.

So, first we will understand the context what kind of algorithms you want to give and then we will see how to apply this relaxation rounding technique. So, set cover problem is easy to describe we have a universe these are the sorry subsets of u. Notice that the size of u is not n I am denoting number of subsets to be n most of the game will be played with these subsets. Now, a set of set cover of u will be what do you think should be a set cover.



So, this is a set cover if other way to say it is that every element in the universe is contained in at least one set in my set cover and clearly if S 1 S 2 union S n is not u then we do not have a set cover. If it is u then S 1 to S n is the set cover. So, the non trivial problem is minimizing the size of the set cover ok. And this might seem like an abstract problem, but like most the problems in complex theory and mathematics there are many many problems in real life which can be framed as these problems that is why we study them in bulk. We said oh we study ten problems many a times your problem can be converted into one of those ten searching sorting satisfiability graph coloring or something and then you know how what to do it we have already solved those problems right.

So, then if I want to take an example for this let us talk about the Ghosh just got finished right. So, suppose you are the coordinator for Ghosh u is the set of all colleges and do I want it to be set of all colleges yes. And now S i corresponds to i th game and this what does S i contain all the colleges which play that sport right. And like a typical sports festival organizer you do not care about sports what you care about is that most of the colleges show up in your in your festival right. So, then you want to make sure that all the colleges show up in your festival, but since you do not have a big budget or you might have a big budget, but you want to spend it on you know getting the biggest celebrity you want to make sure that you organize least number of sports right.

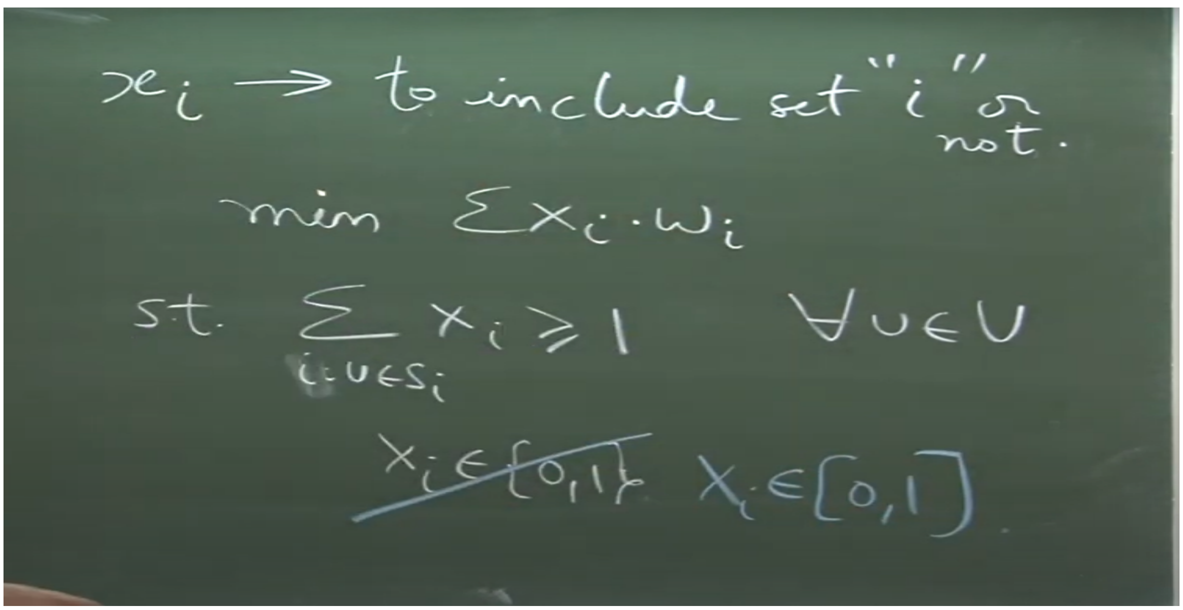


So, this will be your problem and then you do not have to worry about it because we as a complex theory people have solved the problem for you. We have not solved the problem we have told you that is very hard. So, do not solve it, but yes. So, and you know if given ten more minutes you can all of you can come up with an with a real life situation which formulates into a set curve problem. So, even though it might look like a very abstract thing this is really a an important problem ok.

I could have started with this and come up with set cover problem, but right. So, it is clear to you that this is this is a nice problem right many you want to buy groceries. So, that every spice is covered or something like that many right many cases can be converted this problem. So, finally, we want to solve this problem and now since you are taking my course every time you see a problem the thought is can I write a linear program for this. It is a very hard problem.

So, probably no, but though as we have pointed out we can create an integer linear program for it. So, take out your pen and paper and write an integer linear program for this. Like we have done in most of these cases we define this indicator variable which denotes whether I am taking a particular set in my set cover or not correct. Then what do I want to minimize? I want to minimize the number of sets ok.

Sorry. At least ok. This is a integer linear program because of this correct and now I can also introduce a weighted version of it where there is a cost associated with taking each set not percent each set right. This is corresponding to the probably for our example this is the cost of organizing that sport someone might say oh there is no point organizing cricket because it takes twice as much cost as football plus hockey or something right using football hockey I can cover right. So, then if I have a cost of w i with x i then I can just make it this would not change most of our discussion. So, if you want you can remove this it is up to you or if you want to keep it you can keep it just makes it a problem more general ok. So, this is the problem which you want to solve which is very hard.



So, we say let us convert it into an LP. Before we solve this problem and let us actually define when we are happy with the solution ok. We know that this problem is hard that means we do not we cannot give a exact solution believe we believe that we cannot give an exact solution.