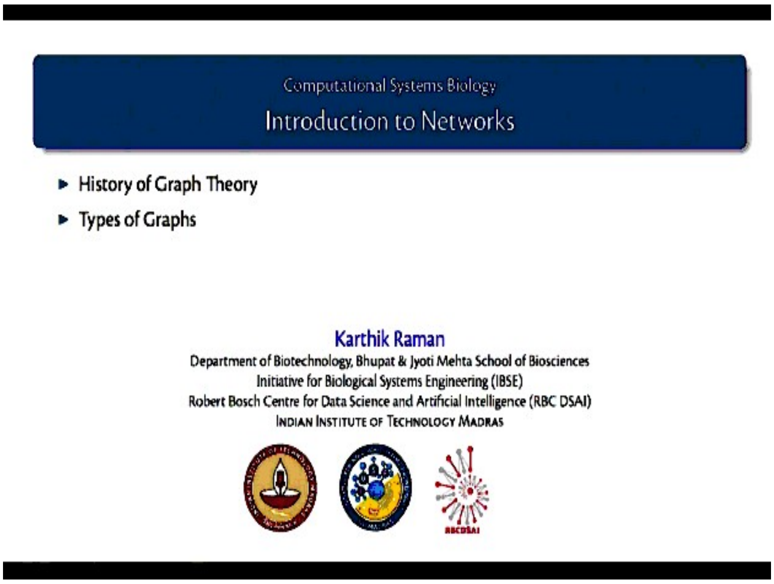
Computational Systems Biology
Karthik Raman
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture - 13
Introduction to Networks

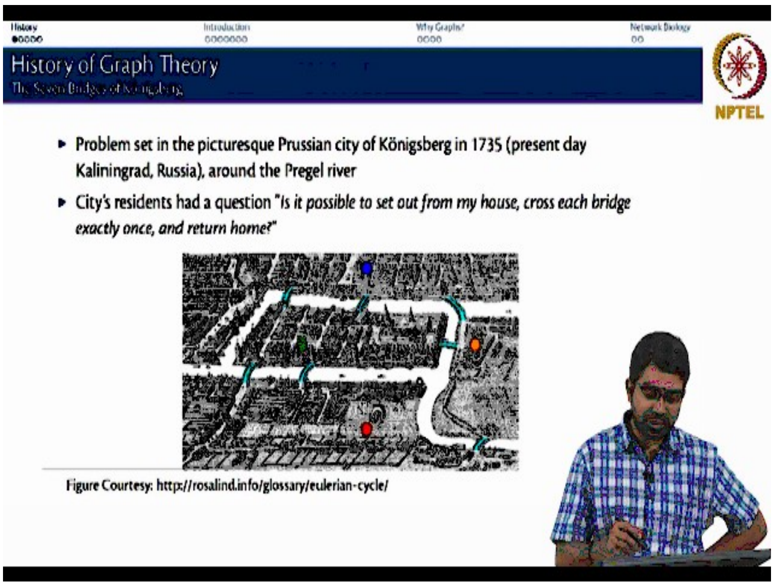
In this video we will look at graph theory.

(Refer Slide Time: 00:12)



We will start with a brief introduction to graph theory particularly the history of graph theory which is a very interesting anecdote where in you know you look at how Euler sort of created graph theory to solve a very simple problem. And then we will look at the different types of graphs and we will continue with the different types of graphs in the next lecture as well.

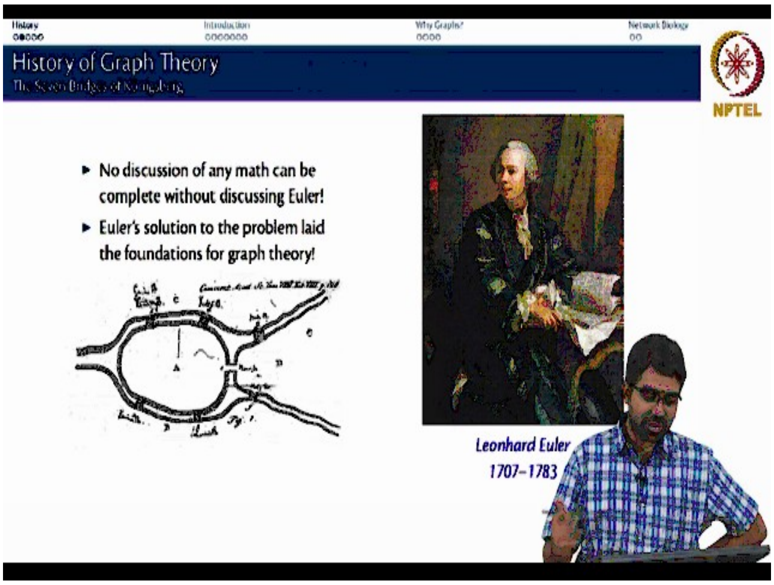
(Refer Slide Time: 00:44)



I think this is a very famous problem, which is well studied or well quoted in literature. It starts back with, way back in 1735 when recreational mathematics was a big thing. People used to do a lot of math for fun and a fun question that some people asked was they lived in the city of Konigsberg. It is presently in Russia around the Pregel river. I think you can see the river out there.

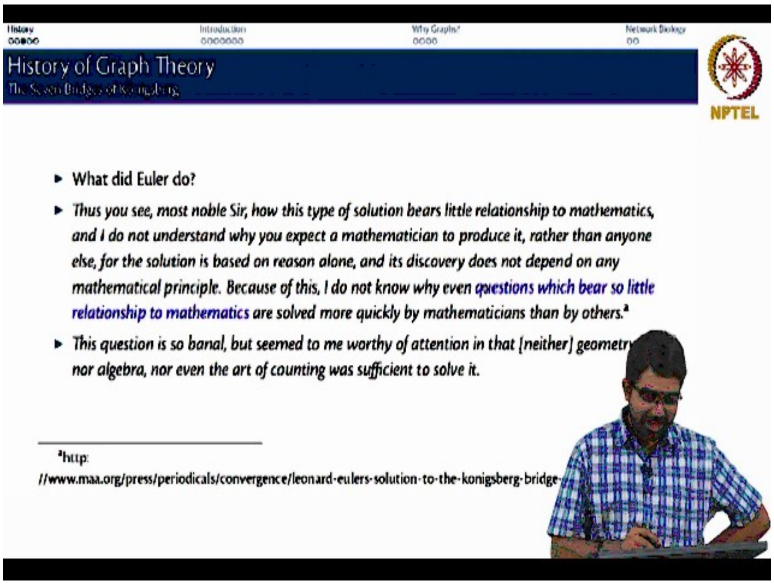
And the question the city's residents had was, is it possible to set out from my house, which could be anywhere here and take each of these bridges once and come back to my house? You start from anywhere in the city, trace each of these bridges and come back to my house right.

(Refer Slide Time: 01:29)



So, as always it was Euler who came up with the solution. We will look at Euler I think once again when we look at dynamical systems because he is contributed to several different branches. So, his solution essentially laid the foundation of graph theory.

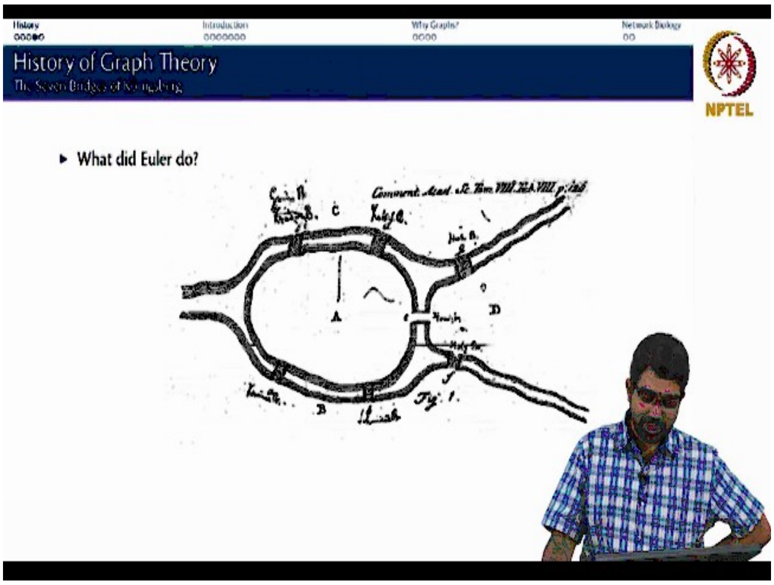
(Refer Slide Time: 01:48)



What did Euler do? He wrote this very nice letter, he basically said something like "This type of solution bears little relationship to mathematics and I don't understand why you want a mathematician to produce a solution and its discovery is based on reason, but not any mathematical principle". But he says, "I do not know why even questions which have so little relationship to math are still solved better by mathematicians".

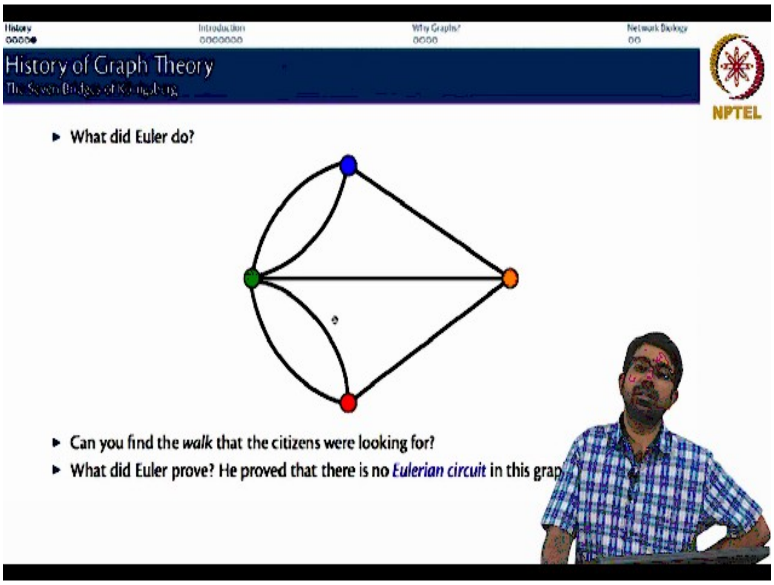
And he says it's very banal, the question. But it was worthy of attention because he could not use just geometry or just algebra or the principles of counting to solve it. He kind of needed to draw on multiple concepts here.

(Refer Slide Time: 02:31)



So, what did Euler do? I think you may have read about it. He took this map and converted it into the first ever graph, right.

(Refer Slide Time: 02:39)

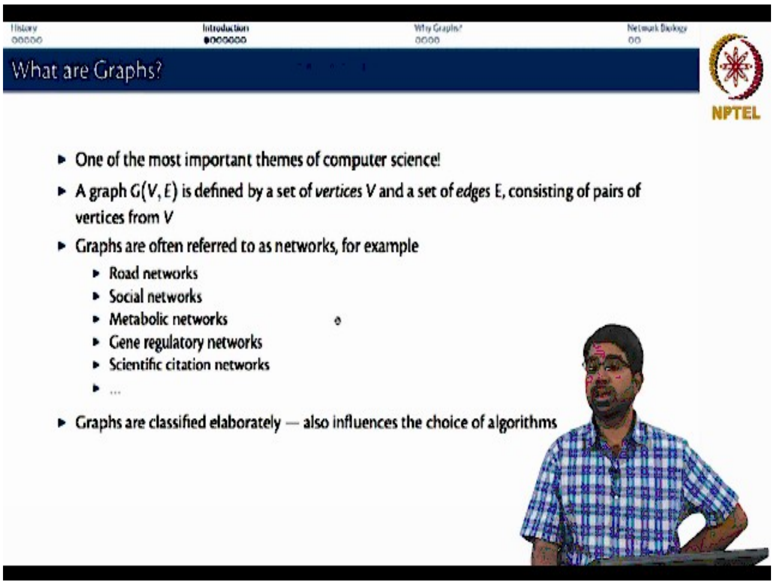


What is this graph? A graph essentially represents relationships between entities and their relationships. So, all the circles that you see here are nodes and the edges represent some sort of relationships. So, in this case your nodes are? There is only one city. Yeah, the islands, sort of islands in the city right and the edges represents your seven bridges. So, can you actually trace all these bridges and come back home starting with any point.

Can you start at say the blue point and like come back by tracing all of these? So, you could take this edge, this edge, this edge, this edge, but you would have missed a couple. You will see that you cannot find the walk that the citizens were looking for. In fact, it is named in Euler's honor. He proved that there is no Eulerian circuit in this graph. So, what is an Eulerian circuit?

It is one which starts at a particular node with its every edge in the graph and also every node and comes back to the original node. If it doesn't return to the original node it is called an Eulerian path.

(Refer Slide Time: 04:03)

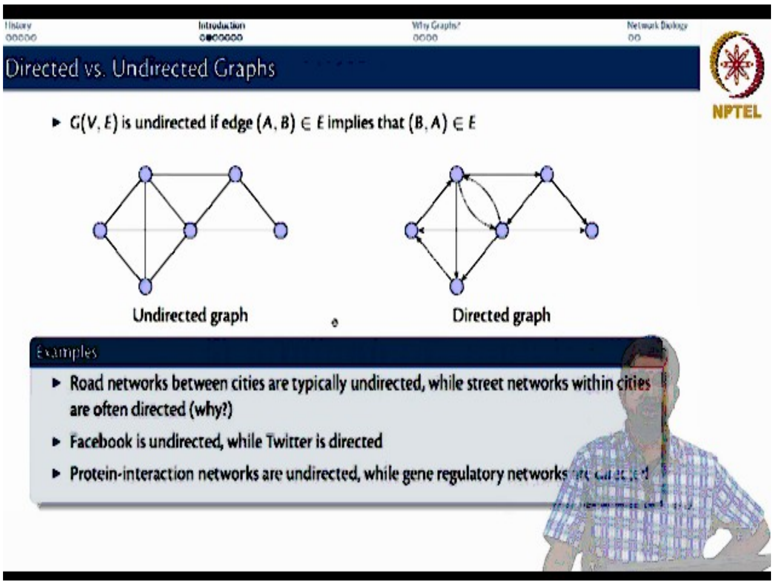


So, graphs are in fact one of the most important themes in computer science. I am sure you are familiar with Facebook graph and Facebook graph search and so on. So, graphs are basically all pervasive in computer science. Graph is defined as G (V, E) which is a set of vertices and edges. So, these edges are essentially a subset of VxV. So, every pair of nodes can be connected, you will find that in any graph some set of those nodes will be connected.

And you also typically call them as networks and you have different types of networks, you have road networks, social networks you are all familiar with, metabolic networks, gene regulatory networks, protein interaction networks, scientific co-authorship networks, scientific citation networks and so on and so forth. Food chain can also be viewed as a network.

And these graphs are also classified elaborately which influences what kind of algorithms you can run on them.

(Refer Slide Time: 05:11)



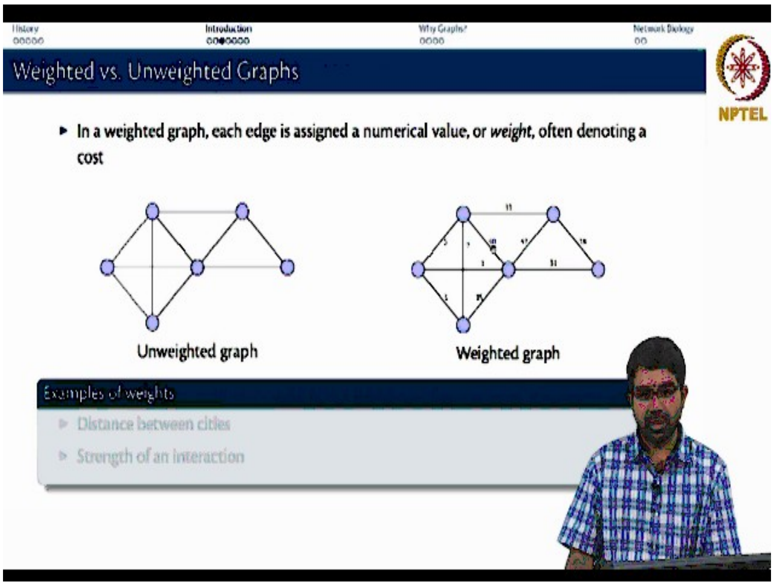
So, the most classic bifurcation of graphs is into directed versus undirected graphs. A directed graph is one in which there is a notion of direction, like A is connected to B, but it doesn't mean that B is connected to A. In an undirected graph, there is a path that can go both ways. A is connected to B as well as B is connected to A. So, a classic example of an undirected graph would be a social network where A is B's friend.

It also automatically means that B is A's friend. A and B are friends. Or in a protein interaction network you say A and B interact. There is no direction to that relationship. Where as in a gene regulatory network you would say that A regulates B. A represses B, B represses C, C activates A. So, all these will have some sort of directionality and the same thing for something like a twitter social network.

In Twitter, A follows B but B may not follow A back, right, so there is some sort of direction in the interactions. Similarly, road networks within a city are directed whereas road networks between cities are directed. You will find that road networks between cities are essentially highway so they basically go in both directions, right.

You will never have only highway that goes from Chennai to Bangalore, but not from Bangalore to Chennai. Whereas within a city you will have a number of one-way roads, one-way streets and so on, so directionality becomes important. As I mentioned, protein networks are undirected whereas gene regulatory networks are directed.

(Refer Slide Time: 07:07)



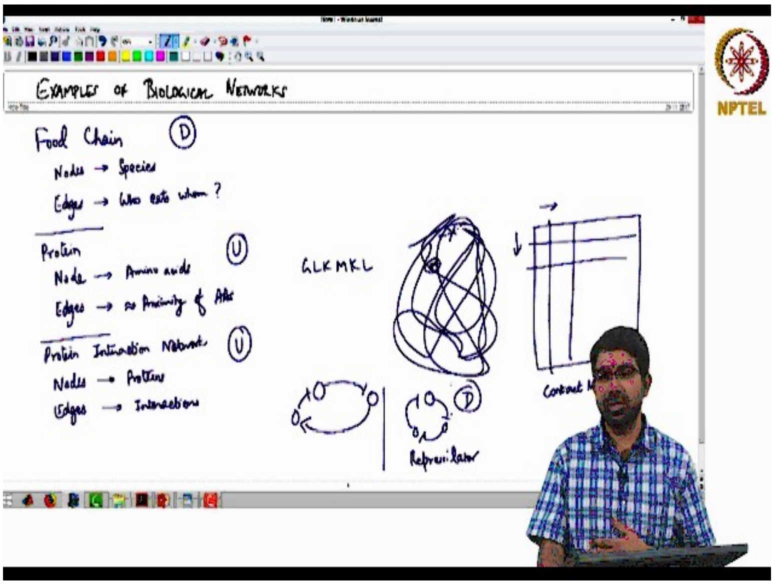
The next important concept is that of weights in graphs. How do you assign weights is an important thing? We will look at it when we study about reconstruction of these kinds of graphs. The graph on the left is an unweighted graph, the graph on the right is a weighted graph, what you see here is that there is, this is usually in computer science, this is usually a cost, there is a cost of going from node A to node B.

It costs you 5 rupees to go from node A to node B, right or it takes 5 hours to go from node A to node B. This is what happens whenever you query for route on Google Maps, right. It looks at a directed network and tries to find the shortest path that is available on the directed network, but the costs are usually either time or distance, depending upon what you want.

In an unweighted graph there is no notion of cost between 2 nodes. But the one difference in biological networks is that we usually do not have a notion of cost in many cases, there are exceptions, we will get to it, but in most cases your edge weights represents only a confidence. I can say with 40% confidence that node A and node B interact. I can say with 90% confidence that node A and node B interact.

So, this becomes a challenge, right because most algorithms you will see assume that the edge weight is actually a cost and not a probability or a strength or something like that. So, what is the strength of an interaction? That could be one thing. If you are talking about, you can have different kinds of networks, we will probably stop by that in a minute. So, in different networks you will have different description for the edges, right.

(Refer Slide Time: 08:55)



Let's maybe look at a few examples. So, do you know of any biological networks? Food chain. What are the nodes? Species. Edges? Right? Is this directed? Very much, right. Do you have bidirectional edges? Not sure. So, that is food chain. So, lets actually look at an unusual network. I am not talking about the protein network, I am talking about a protein right.

So, what is a node here, it could be atoms or better still amino acids. Edges? Well, amino acids are bonded to one another but then that would be a boring network, right. So if this where your peptide, you typically will have a structure like this. This is how a protein folds and now you have this protein coming close to this protein. So what you can do is, you can compute a matrix of distance, right?

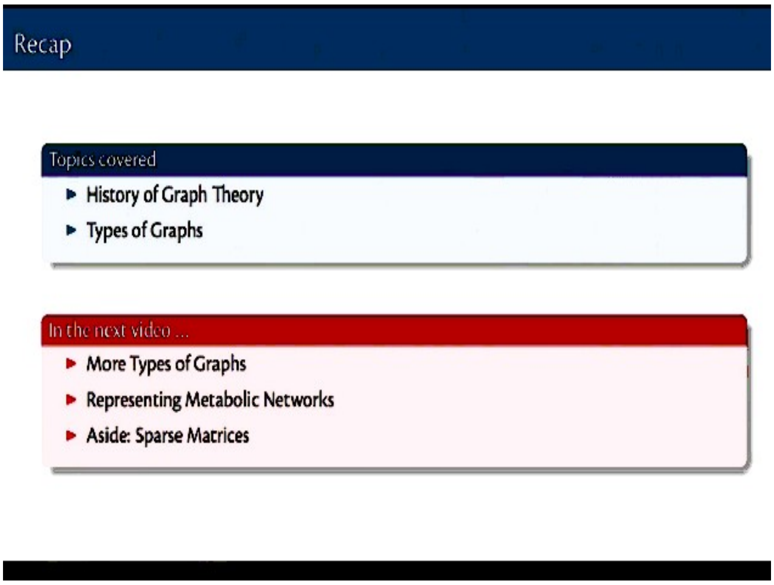
So, you have amino acids on either sides and you can compute a distance and if this distance is less than a threshold, you can put an edge, right. So you would say that these 2 nodes are connected to one another but this node and this node are not connected to one another. So, do read about contact maps. So here edges represent proximity or could you think of something else? Energy, right?

There is some energy of interaction, right. So, you can use a threshold, right if this interaction energy is very high you can have something, if its very low, you cannot have an edge and so on. A protein could be viewed as a network. Let's look at a simpler example now. Protein interaction network, this is much easier. So, every node, by the way, is this directed? The previous protein, is it a directed network or an undirected network? Undirected network.

A is close to B, B is close to A, so this is undirected. Food chain, directed. Protein interaction network, undirected. So, here nodes are proteins, edges, interactions, whereas in a gene regulatory network you would find that, this is a very classic network for those of you who have seen it before. Right A inhibits B, B inhibits C, C inhibits A. Right, this is of course a classic directed network.

So, these are some examples of networks and you have many, many more different types of networks and there are many properties that we wish to study for these networks. Let's quickly look at some more interesting types of networks.

(Refer Slide Time: 13:53)



So, I hope today you got an introduction to the history of graph theory and the basics of graphs and the different types of graphs that exist. We will continue with some more types of graphs in the next video and start looking at representing metabolic networks and also an interesting data structure for representing matrices on a computer.