INDIAN INSTITUTE OF TECHNOLOGY MADRAS

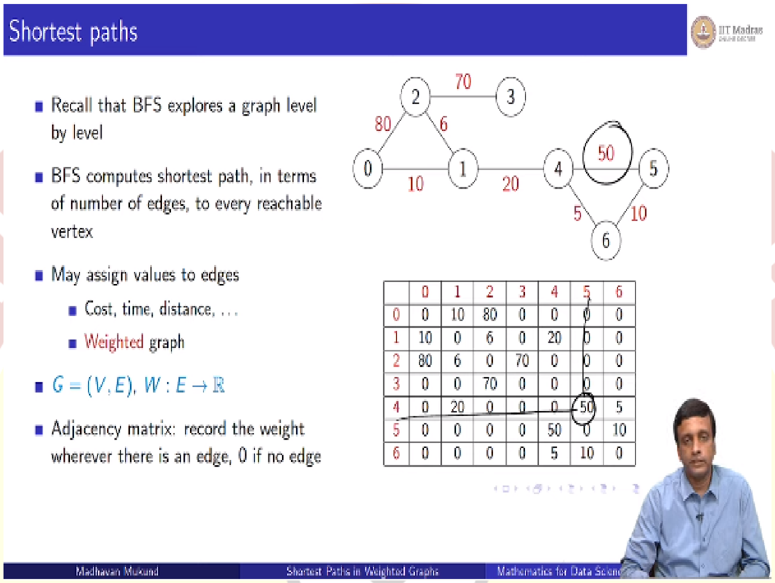
सिद्धिर्भवति कर्मजा

IIT Madras

ONLINE DEGREE

Mathematics for Data Science 1
Professor. Madhavan Mukund
Chennai Mathematical Institute
Lecture No. 12.1
Shortest Paths in Weighted Graphs

(Refer Slide Time: 0:18)



Shortest paths

- Recall that BFS explores a graph level by level
- BFS computes shortest path, in terms of number of edges, to every reachable vertex
- May assign values to edges
  - Cost, time, distance, ...
  - Weighted graph
- $G = (V, E)$, $W : E \to \mathbb{R}$
- Adjacency matrix: record the weight wherever there is an edge, 0 if no edge

	0	1	2	3	4	5	6
0	0	10	80	0	0	0	0
1	10	0	6	0	20	0	0
2	80	6	0	70	0	0	0
3	0	0	70	0	0	0	0
4	0	20	0	0	0	50	5
5	0	0	0	0	50	0	10
6	0	0	0	0	5	10	0

Madhavan Mukund | Shortest Paths in Weighted Graphs | Mathematics for Data Science

So now, let us look at a new type of graph called a weighted graph. So, remember that in a graph like this we have seen that a systematic way to explore this graph is breadth first search and breadth first search explores this graph level by level and therefore because it does it level by level, it discovers the vertices reachable from the starting vertex at successively longer distances and therefore, the BFS computes the shortest path in terms of number of edges to every reachable vertex.

Now in practice we often assign some values to the edges, so if you look at for example, a road map then you might see some numbers against each section of road representing the length of the road. Similarly, if you look at say a railway map or a airline map you might see the time it takes to do a segment of a journey or it could be the distance or it could be even the cost you know how much does this ticket cost.

So, these numbers indicate some more abstract information about the length of an edge than just the fact that I take one edge, so it is not enough for me to say that I took two flights, I need to know how long these flights are. So, if I take a hopping flight in a short distance say I go from Chennai to Bangalore to Mangalore, right it is not the same as taking a hopping flight which takes me from Chennai to Delhi to San Francisco. So, the two are still two length travels but one is enormously longer than the other in time.

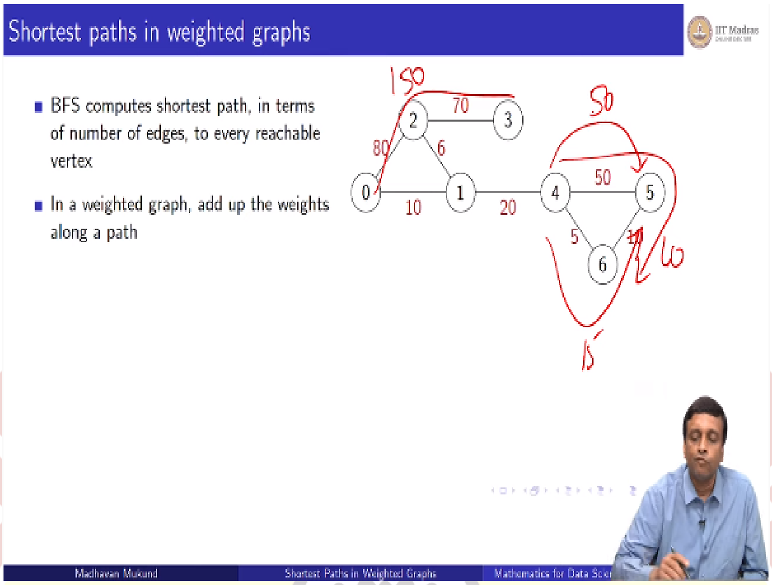
So, this is the kind of information that we want, so we want to take this kind of a weighted graph, so what is a weighted graph formally? A weighted graph is just a graph, so we have a set of vertices, we have a set of edges but we have weights, so weight formula is a function, right, a function which takes every edge that is present in the graph and assigns it some real number.

So, notice that we are not claiming at this moment that the real number is positive or negative, we will discuss what negative weights will mean but most of the interesting things that we can think of the weights will be 0 or more. So, we can think of a 0 cost edge sometimes the cost an edge which is not there has a 0 cost edge or sometimes it may be there but typically in any reasonable scenario weights are positive but we will see a situation where weights could actually be negative and make sense.

So, the first thing is we have to, we are going to work with these graphs and we have an adjacency matrix way of working with a graph in general which records the presence of an edge. So, how do we record a weighted graph, so what is our representation of weighted graphs? So, in adjacency matrix what we can do is that whenever we normally put a 1 saying there is an edge, instead of the 1 we can put the weight, so assuming that there are no 0 weights, then wherever there is a 0 there is no edge, wherever there is a 1 there is a weight.

So, if you look at this graph here for instance, if I look at the edge for instance 4 to 5, then I look at the entry 4, 5 and now I have a 50 there rather than just a 1. So, this is a very simple way to represent weighted graphs just take the adjacency matrix and at each entry i comma j put the weight of the edge i comma j and if there is no edge or if the weight is 0 put a 0.

(Refer Slide Time: 3:18)

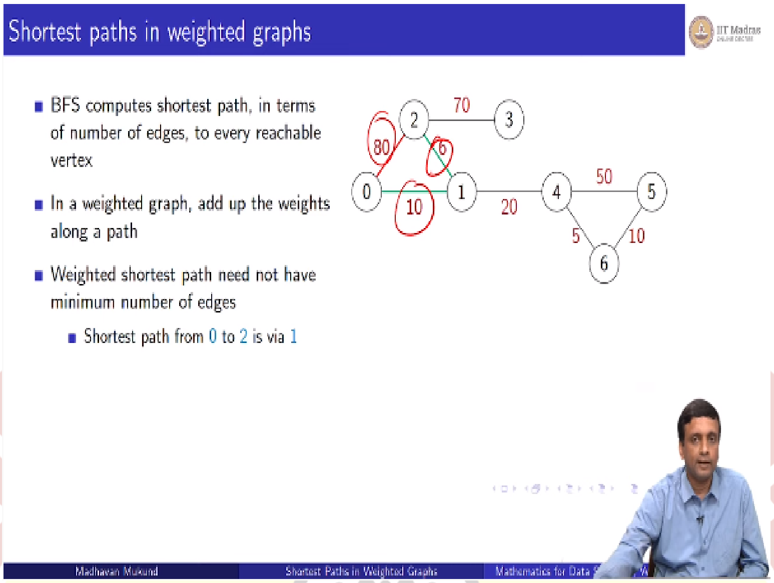


Shortest paths in weighted graphs

- BFS computes shortest path, in terms of number of edges, to every reachable vertex
- In a weighted graph, add up the weights along a path

So, our interest is initially to compute shortest paths in such graphs, so we have these weighted graphs, where we have some edge weights between edges and we want to find the shortest path, so what is the shortest path now? For us the shortest path will be the sum of the weight, so for example, if I take this path, then the weight of that path, the length of that path is 80 plus 70 is 150. If, I take this path for instance, the length of that path is 60. More interestingly if I take this path the length of that part is 15 whereas the direct path from here to here is just is 50. So, going from 4 to 5 via 6 is actually shorter than going from 4 to 5.

(Refer Slide Time: 4:02)



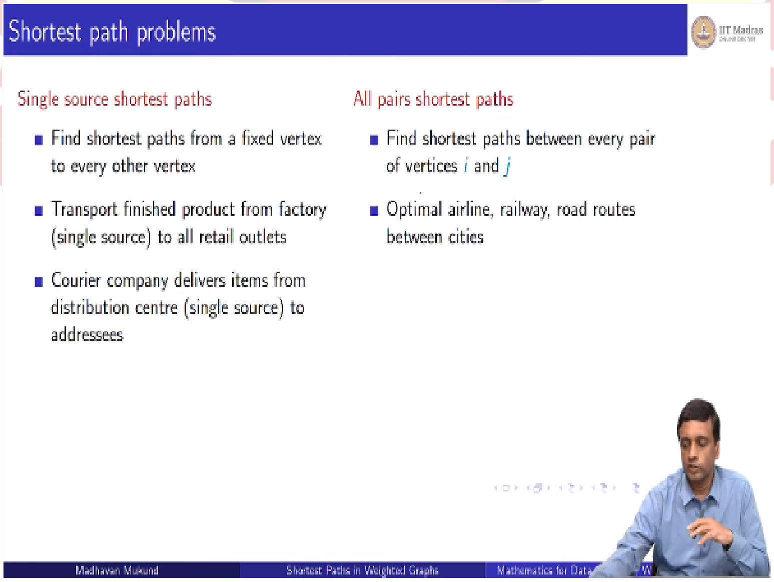
Shortest paths in weighted graphs

- BFS computes shortest path, in terms of number of edges, to every reachable vertex
- In a weighted graph, add up the weights along a path
- Weighted shortest path need not have minimum number of edges
  - Shortest path from 0 to 2 is via 1

Madhavan Mukund Shortest Paths in Weighted Graphs Mathematics for Data

So, in general this is the situation that in a weighted graph, the weighted shortest path does not need to have the minimum number of edges in the unweighted sense. So, if I look at the very beginning of this graph from 0 I can go from 0 to 2 in one step but then I pay a cost of 80, whereas if I go from 0 to 1 and 1 to 2, so I take two steps then actually I get a cost which is 10 plus 6 which is much less 16. So, the shortest path from 0 to 2 is actually via 1, an indirect path.

(Refer Slide Time: 4:30)



Shortest path problems

Single source shortest paths

- Find shortest paths from a fixed vertex to every other vertex
- Transport finished product from factory (single source) to all retail outlets
- Courier company delivers items from distribution centre (single source) to addressees

All pairs shortest paths

- Find shortest paths between every pair of vertices i and j
- Optimal airline, railway, road routes between cities

Madhavan Mukund Shortest Paths in Weighted Graphs Mathematics for Data

So now, what are we going to try and solve with these weighted graphs? Well there are more than one type of shortest path problem, so the first type of shortest path problem is one where we start at a fixed vertex and we want to find out the distance to every other vertex, so this is called the single source shortest path problem. So, we would start a fixed vertex and find out how long it is from this vertex to every other vertex, why is this an interesting problem?

Well, there are many applications where this is interesting for instance, suppose you are a manufacturer, so you have a factory and you make things and now you have to take things from your factory, your finished products and distribute them to the shops where they are sold, the retail outlets. So, you have a single source your factory and then you have to find the most efficient way, so the shortest path in terms of whatever you are measuring the cost of travelling or the time it takes to travel or the distance it takes to travel, whatever is the cost that you want to count towards the transportation cost, you would like to find the shortest transportation cost from your factory to every one of your retail outlets. So, this is a single source shortest path to every other vertex.

Alternatively for instance, you could be a courier company, so what happens in a courier company is that they have these flights between cities, so all the packages which go to say Delhi come from different destinations and they land in Delhi and they go to a centralized clearing facility in Delhi. So, overnight you might have flights coming from Calcutta, from Bombay, from Bangalore, from Chennai and all that all this information, all these packages come to Delhi and now they have to deliver them out.

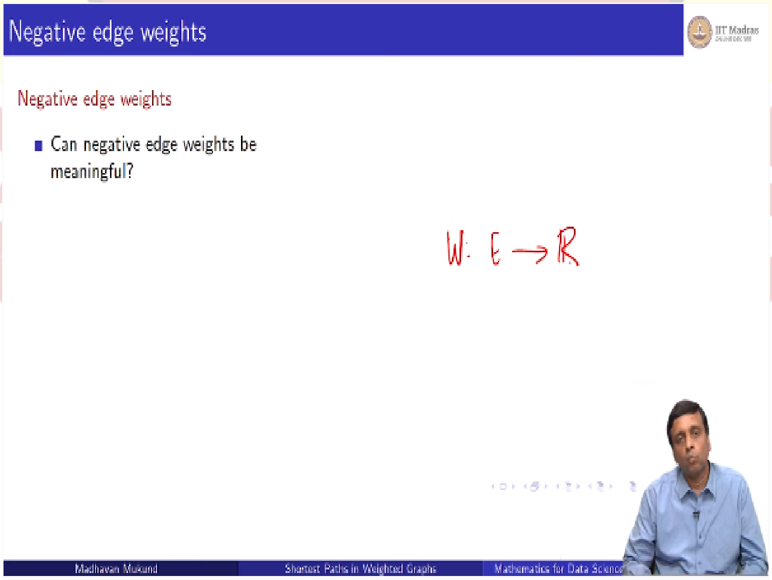
So, the starting point, the distribution center where all these things are initially brought in from the airport or air cargo wherever they come, that is a single source and now they have to now find the most efficient way to distribute it to all the destinations where these packages have to be. So, the single source shortest path probably has a number of applications and therefore it is an interesting version to solve.

On the other hand, sometimes we want to know something about every pair. Now, of course, you could take the single source thing and start from every vertex and find every other vertice distance and then you will get all pairs but generally there may be a better way to find the distance between all pairs. So, for every i and every j we want to find out the shortest distance from i to j. So, the single source is fixed and starting point, say fix vertex 7, and from 7 what is the shortest distance to everything?

Now, this is for every i and j, from not only from 7 to j, I want to know from 9 to j, from 11 to j, from 7 to 11, from 9 to 11 and so on and this is the kind of thing typically that say if you are managing a booking site and somebody says I want to go from city A to city B, then you have to be able to provide in terms of the cost or time or some metric the cheapest way to go from A to B.

So, somebody might say that I want to reach there as fast as possible, so what is the shortest flight? Some people might say I want the cheapest ticket. So, based on whatever is the cost that you are associating, then you will have to find and then you need to be able to do this for any pair because you have customers who can be going from anywhere to anywhere. So, this is the all pair shortest path problem, so these are the two problems that we will initially look at in the context of weighted graphs, single source shortest paths and all pair shortest paths.

(Refer Slide Time: 7:45)


Negative edge weights

IIT Madras

Negative edge weights

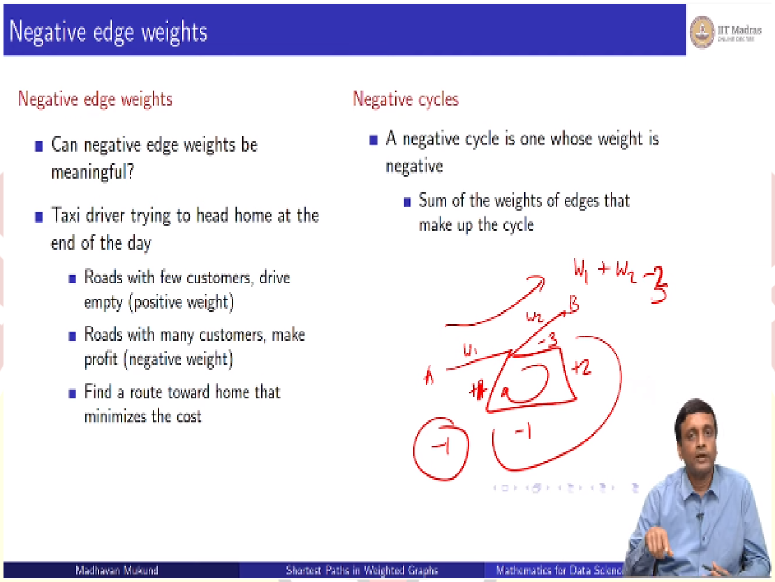
- Can negative edge weights be meaningful?

$W: E \rightarrow R$

Madhavan Mukund | Shortest Paths in Weighted Graphs | Mathematics for Data Science

Now at the beginning I alluded to the fact that we have this function, so we said that we have this function which takes every edge and gives us an arbitrary number and in principle there is nothing to prevent this number, this weight of an edge from being negative, so what if I am thinking of this cost, what would be a reasonable scenario where I could have negative edge weights?

(Refer Slide Time: 8:12)



Negative edge weights

Negative edge weights

- Can negative edge weights be meaningful?
- Taxi driver trying to head home at the end of the day
  - Roads with few customers, drive empty (positive weight)
  - Roads with many customers, make profit (negative weight)
  - Find a route toward home that minimizes the cost

Negative cycles

- A negative cycle is one whose weight is negative
  - Sum of the weights of edges that make up the cycle

Madhavan Mukund | Shortest Paths in Weighted Graphs | Mathematics for Data Science

So now, imagine that you are an Uber driver or an Ola driver or any one of these cab companies, so you have a certain amount of hours when you drive your cab and at a certain point you want to start heading home and reach home roughly when you finish your work rather than eventually traveling across the entire city empty, so this is all, I mean many of us in days when we have taken Ubers, I found that towards the end of the day for example it is little difficult sometimes to get long distance things if the driver is not living in that direction and they would say ‘No, no sorry I am going in the opposite direction because I am heading home’.

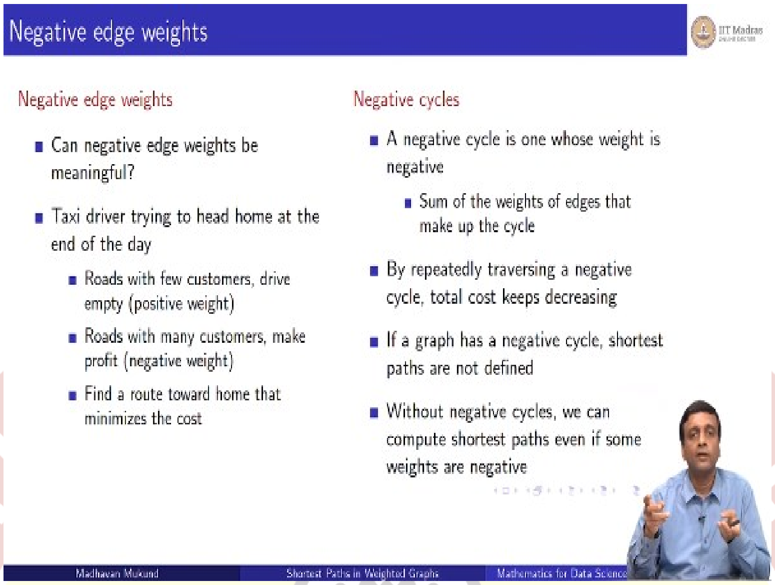
So, here is a taxi driver trying to head home at the end of the day, so maybe he has an hour of service left and he wants to try and optimize where he lands up at the end of this hour, so he has a minimum amount of time to travel home. So, now he has to take a choice, so he has to start looking for customers maybe right, so if he takes a road which has very few customers then he will be losing money, so there will be a cost that he is paying, so that is a positive cost.

On the other hand, if you travel on a road with many customers then you are likely to find somebody who will hail you for a ride, you might get a call so therefore you will earn money, so you have a negative cost, so where you are not taking customers, you are paying for driving the car, you are paying of petrol cost and other costs and therefore, you are losing money so that is a positive cost, where you are gaining money, where you are earning money it is a negative cost and you want to obviously get more money, so you want to reduce the cost you want to make the negatives more than the positive. So, you want to find actually a route towards home which minimizes the cost, so this is one example where negative edge weights make sense.

Now what happens to our shortest path problem in the presence of negative edge weights? So, the problem is not so much with negative edge weights but negative cycles. So, supposing I have somewhere in my graph, a cycle which has something like minus 3, plus 2, minus 1 and plus 2 for example, if I go around the cycle then what do I do? I add up the weight, so I get minus 3 plus 2 is minus 1, right, minus 1 plus, so maybe I should maybe make this also plus 1, so minus 3 plus 2 is minus 1, minus 1 minus 1 minus 2, minus 2 plus 1 is minus 1, so the total weight of the cycle is minus 1. In other words if I go around the cycle once I reduce my cost by 1.

So, if I am supposing I am going from A to B, so there could be some arbitrary weights here W1 and W2, so I do W1 plus W2 to go from here to there but if I want to reduce it I go around the cycle uselessly once and I get minus 1, I go around it again and I get minus 2, I go around it again and I get minus 3, so I can go around the cycle as many times as I want, and keep reducing my cost, I can make it as negative as I want and it does not make sense because I mean we are actually taking longer and longer paths but because the cycle has a negative cost we are able to do this, so this is clearly something which is, which makes the problem undefined.

(Refer Slide Time: 11:10)



Negative edge weights

Negative edge weights

- Can negative edge weights be meaningful?
- Taxi driver trying to head home at the end of the day
  - Roads with few customers, drive empty (positive weight)
  - Roads with many customers, make profit (negative weight)
  - Find a route toward home that minimizes the cost

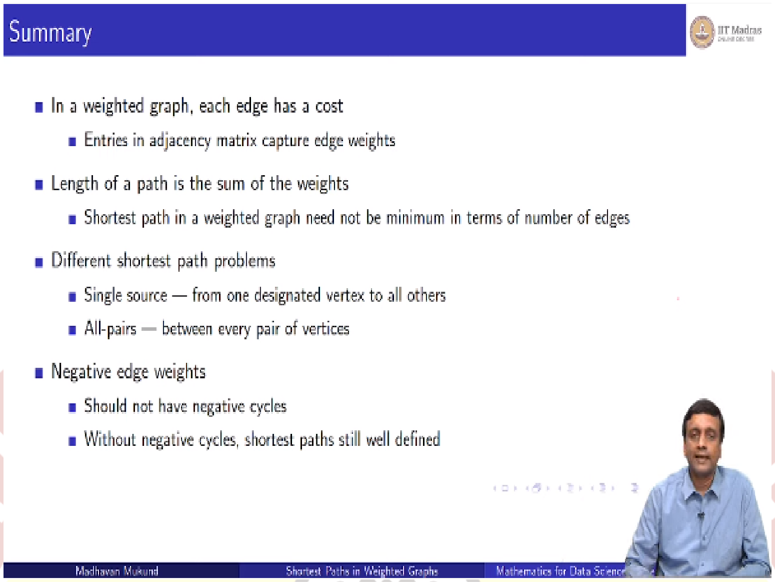
Negative cycles

- A negative cycle is one whose weight is negative
  - Sum of the weights of edges that make up the cycle
- By repeatedly traversing a negative cycle, total cost keeps decreasing
- If a graph has a negative cycle, shortest paths are not defined
- Without negative cycles, we can compute shortest paths even if some weights are negative

Madhavan Mukund Shortest Paths in Weighted Graphs Mathematics for Data Science

So, if we have a negative cycle we can just go round and round that cycle and then there are no shortest paths anymore on anything which goes past that cycle because every time I want to reduce the cost just go around cycle once. But if I do not have negative cycles then it is fine. So, when a graph has negative cycles, shortest paths are not defined but if you have negative edges you might have some edges which are negative but you do not have negative cycles, then it is fine. So, if you have negative edges but no negative cycles you can still do shortest paths but you have to then be careful that your algorithm does not depend on the edges being positive.

(Refer Slide Time: 11:45)



Summary

- In a weighted graph, each edge has a cost
  - Entries in adjacency matrix capture edge weights
- Length of a path is the sum of the weights
  - Shortest path in a weighted graph need not be minimum in terms of number of edges
- Different shortest path problems
  - Single source — from one designated vertex to all others
  - All-pairs — between every pair of vertices
- Negative edge weights
  - Should not have negative cycles
  - Without negative cycles, shortest paths still well defined

Madhavan Mukund | Shortest Paths in Weighted Graphs | Mathematics for Data Science

To, summarize we have looked at what is called a weighted graph, so in a weighted graph we attach to every edge a cost, a number, which we call the weight and these can be described for us in our adjacency matrix by entering the cost or the weight of every edge in place of a 1. So, then once we have these edge weights, then we can measure the length of a path in terms of the weight. So, not just how many steps I take in terms of edges but what is the total sum of the weights across these edges.

And so now I get a new notion of shortest path which is probably more natural from the way we think about graphs representing sort of spatial things at least, so we get the sum of all the edges that we traverse but the weights of the sums, not just the number of edges and we saw that this now will give us something which is not necessarily the same as the shortest path in terms of number of edges. We could have a shorter path which has a longer higher cost as compared to a longer path.

So, we said that there are two types of shortest path problems with at least two types which we will find interesting, one is the single source path where we start at a fixed vertex and we want to find out where we can, how fast we can go to every other vertex, so this is for example the delivery problem for a courier company or we have the all pairs problem which is typically the type of problem that you need to solve if you run a travel agency, you need to be able to tell somebody from anywhere to anywhere what is the best way to go.

And finally, we looked at this peculiar problem of negative edge weight, so we gave a justification that there can be reasonable situations which are modeled by negative edge weights and if we still want to be able to compute shortest paths in the presence of negative edge weights, what we need to ensure is that there are no negative cycles because if we have negative cycles, then the shortest path is not defined but we do not have negative cycles even if we have negative edge weights, we can hope to find shortest paths.