IIT Madras

ONLINE DEGREE

Computational Thinking
Professor Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute
Professor G. Venkatesh
Indian Institute of Technology, Madras
Insertion sort and ordered list

Professor Madhavan Mukund: So, all along we have been using these cards as they are without having to worry about whether they are in any particular order. So, each time for instance when we are looking for the maximum or we were trying to add up the marks we were just scan them and whatever order they had. But I think some of these things might have become easier if we had some order in them.

For instance, if you wanted to find the maximum marks and it been ordered by marks then we know that the maximum is either in the beginning or in the end depending on whether it is increasing order, so…

Professor G. Venkatesh: Same goes with date of birth we said, find two people with the same date of birth, so if you had ordered it by date of birth…

Professor Madhavan Mukund: Yes, then they will come in sequence and so if two people have the same birth date they will come next to each other, so we just have to, we do not have to go across the whole thing we just have to see if there is neighboring cards.

Professor G. Venkatesh: So, arranging in some order…

Professor Madhavan Mukund: Some order, it seems to be useful thing.

Professor G. Venkatesh: Very useful thing, but if given cards like this, suppose I want to arrange it some order let us say I want to arrange it in the increasing order of this sequence number of the card, how do I do that? So, I know that this is not in sequence, so let us say I want to go in increasing order of sequence, so which means 0 is at the beginning at then, okay then this is the end. So, I know that this is 14 and this is not in correct order, so I have to exchange it, I have to put 4 before 14.

Professor Madhavan Mukund: Yes, but then will have to exchange something else, so that is not, so I have a…

Professor G. Venkatesh: Systematic way of.

Professor Madhavan Mukund: Yes, so I think normally when we say for example we actually play cards or something then normally what we do is we pick up a card and we put it in correct way, so we can do the same thing. So, we can pick up this first card 14, so ofcourse we have only 1 card it is in order then we pick up the 4 and now we decide whether the 4 must go before or after.

Professor G. Venkatesh: Should go before.

Professor Madhavan Mukund: Now we pick up this 22

Professor G. Venkatesh: Should go after 14.

Professor Madhavan Mukund: Right, so 22 must go after 14, so now we pick up this 11 and then 11 must be…

Professor G. Venkatesh: After 4.

Professor Madhavan Mukund: So, we find the place to insert it, so we insert it in the correct place so that, so what we are doing is we are building up a sequence which is in the correct order by taking each new card and putting it in the right place in this sequence.

Professor G. Venkatesh: So, when you take this new card you have to compare it with all the cards.

Professor Madhavan Mukund: Yes, so we are basically this is the systematic way to start here and say okay goes after 4, it goes after 11, it goes after 14, it goes after 22, goes in the end. Then we take 7 and will say it goes after 4 but 11 is bigger, so we put it between 4 and 11. So we are inserting a new card in the right place exactly. So, this is actually…

Professor G. Venkatesh: Is there a name?

Professor Madhavan Mukund: It is called insertion sort.

Professor G. Venkatesh: It is called insertion sort because we are inserting it in a right place.

Professor Madhavan Mukund: And we are finding a sorted sequence.

Professor G. Venkatesh: Net result is sorted.

Professor Madhavan Mukund: So now you take a 0 and I start and it is already smaller than 4, so I put it here and I take 16 and I go down and I find the place to put it , it goes after 14. So each time I am not doing anything very clever I am just going from the beginning, so I am saying 0, 4, 7, 11, 14, 16 and 22 is bigger so I put 21 before that and if I see 21 and 22 and 23 I can actually just to collapse nothing go in between, so this is to save space I will just do that.

Now I have 6, so 6 goes 0, 4 goes here, 13 goes obviously after 11 and above 14, 2 goes between 0 and 4. 29 goes all the way at the end, 9 comes before 11, 28 goes before so I put that.

Professor G. Venkatesh: 18?

Professor Madhavan Mukund: 18 comes here, so I am actually cheating by pretending but, so I will start everything I have to start from here and go through the whole list and find the…

Professor G. Venkatesh: It is a nested iteration.

Professor Madhavan Mukund: Yes, it is nested iteration. For every card here I will go through all the cards I have already seen and find the stop at the right place.

Professor G. Venkatesh: So, it is going to be 30 into 30 type of thing.

Professor Madhavan Mukund: Yes, no stopping, no avoiding that unfortunately.

Professor G. Venkatesh: Initially you are comparing with one card then two cards, 1 plus 2 plus 3…

Professor Madhavan Mukund: So as we had seen that would be something like N into N minus 1 by 2.

Professor G. Venkatesh: Yes, correct something like that n times minus 1 by 2, right?

Professor Madhavan Mukund: So, this 27 again I sort all the way then finally it goes above the 28. 5 I come here and now I have this 4, 5, 6, 7 so let me collapse that. So, slowly this thing is collapsing, so this 25 comes before 27, 1 comes here, so 0, 1, 2 is finished. 26 again I go all the way I find it comes in between this 25 and this 26, 8 comes here I have between I mean know

that I have 4, 5, 6, 7 so it comes before the 9. 9, 10, 11 I have. So, 24 so I am just going a little faster, so 16 comes between 17, 16, 17, 18. 12 comes before 13 and finally 19 comes between 18 and 20. 3 comes right in front above 4, between 3 and 4 and 15 should come here. So, now in this process…

Professor G. Venkatesh: We have now turned this into a…

Professor Madhavan Mukund: All things are ordered, it should be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. So, we have done this ordering by systematically building up an ordered list one element at a time by inserting it into the correct place, so this is insertion. So, this is one way of doing it.

Professor G. Venkatesh: I guess there is other ways of going it but this one is enough for us.

Professor Madhavan Mukund: Yes, so at least we know one way of doing it. Later on, it is subject that will come up it is a very important process this business of sorting actually. So, as we not in this course but in more advance courses this will be one of the basic things how to do this kind of sorting as efficiently as possible because this n squared is a problem actually. So, if you take this n square and if you want to say for example sort, say all the names of say forget about people in India, even people in the big city with a few lakh people will take forever.

Professor G. Venkatesh: Will take for long time.

Professor Madhavan Mukund: Yes, because you will have 5 lakh into 5 lakh that many cannot sort.

Professor G. Venkatesh: Cannot sort. Okay we will discuss that later but right now I have a sorted list, not a problem, we sorted by this. I want to sort by marks.

Professor Madhavan Mukund: Yes, so for example we want to find the highest mark.

(Refer Slide Time: 07:20)
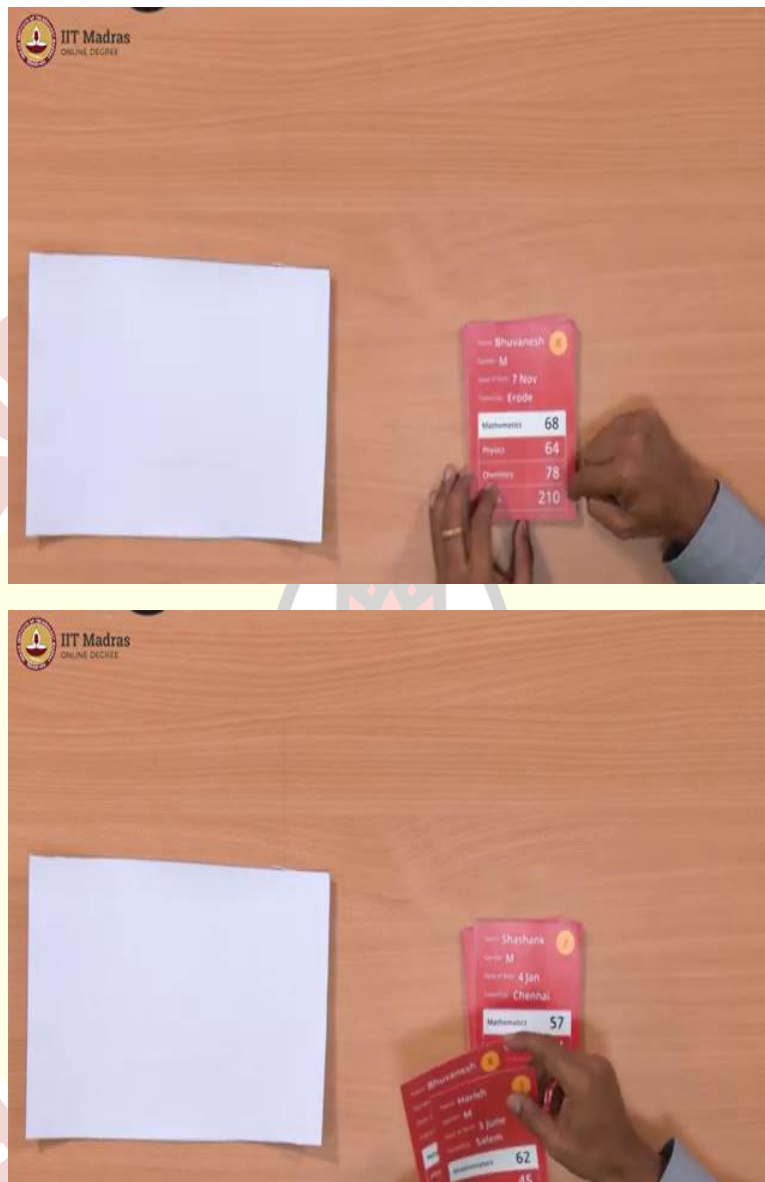
Professor G. Venkatesh: But this is not helping me because, I have to sort in by marks.

Professor Madhavan Mukund: So, if you sort y marks then…

Professor G. Venkatesh: Insertions are I start to 10, 198 then the 198 in front of 200. So, I can only maintain this way it looks like I maintain this sorted these cards in a sorted way only by one way of sorting, which is the way of sorting? I have to disturb the cards again, is there a way to fix this?

Professor Madhavan Mukund: So, I would suggest that we do not like we did for finding those sets we sort by keeping these positions of the cards and then depending on what value or what field we are looking at we move those positions, so we sort set.

Professor G. Venkatesh: Sort the bookmark.

Professor Madhavan Mukund: Sort the bookmarks, so we sort the bookmarks and then so we then we can pick up the cards and they order the bookmarks and that will be sorted for our criterion but we are not disturbing the cards.

Professor G. Venkatesh: We do not, okay.

Professor Madhavan Mukund: So, we can do a insertion sort like that.

Professor G. Venkatesh: Shall I try them, I want to, I did not want to disturb the set, I want only, so I want to insertion sort on the total marks.

(Refer Slide Time: 08:36)

Professor Madhavan Mukund: So, what I will need to do is keep both so that I know where we are sorting, so I will say that, so I will write it just for conveniences, so I will write the index so as 0 and 210, so this is my first card in some sense but I am sorting on the 210 I am not sorting on the 0. Now we move to the second card. So, this is 198, should go before this, so I would, so

we know let us, so I would now I will just write it again. So, I 198 and 0 to 10, so this is what I have, so now this is my sequence.

Professor G. Venkatesh: Sequence we are making.

Professor Madhavan Mukund: Yes, so the sequence is…

Professor G. Venkatesh: We had this set which is basically a, so the sequence when we had the earlier the set which we wrote down all these sets which we wrote, the order did not matter but here this is not like that, here the order is important actually.

Professor Madhavan Mukund: Yes, so in some sense this order is out of order with respect to the way the cards are arranged, this second order is the order that we are interested in so we want to find the smallest of the two that we have seen so far we know it is card 1, so it is an ordered set.

Professor G. Venkatesh: Something like an ordered, anything like an order set, a set is numbered. So, this is an ordered list.

Professor Madhavan Mukund: So, I think it is better to call it a list.

Professor G. Venkatesh: It is an ordered list, so this is 198, so 188…

Professor Madhavan Mukund: So now 188 will come before this, so I will just write it here for now because I can do that.

Professor G. Venkatesh: 2 and 188, 173.

Professor Madhavan Mukund: So, 173 is still before that, so now I have, so what is the card number; 3. So 3, 173; 2, 188 so maybe we should just do the first 10 cards because it will become a bit tedious. So now I have up to, okay.

Professor G. Venkatesh: 4, 240.

Professor Madhavan Mukund: So now 4, 240 will come back at the end because it is the biggest one.

Professor G. Venkatesh: 5, 247

Professor Madhavan Mukund: Unfortunately continues, so goes 5, 247.

Professor G. Venkatesh: 6, 250.

Professor Madhavan Mukund: Continuous.

Professor G. Venkatesh: 7, 252 again continuous, lucky, very lucky. 8, 189

Professor Madhavan Mukund: Sorry 7 was 252, so now 189, so I will insert it here so this is what insertion means so I will have to put it here. So, I will have 8, 189.

Professor G. Venkatesh: 9, 242.

Professor Madhavan Mukund: So, now 9 comes here, so I think that is the 10 card, because we started with 0.

Professor G. Venkatesh: 10, 204.

Professor Madhavan Mukund: Okay 10, 204 so that comes here, okay I think…

Professor G. Venkatesh: Okay I think this way we can do.

Professor Madhavan Mukund: So, after sorting 10 cards we just write it out neatly, we have 3, 173; 2, 188.

Professor G. Venkatesh: So, these 10 cards that we have 10 or 11 whatever it is, these are going to be these are still in that same order 0, 1, 2, 3, 4.

Professor G. Venkatesh: Yes, so physically they are in the same sequence.

Professor G. Venkatesh: Same sequence nothing has not affected it but if you go down this list.

Professor Madhavan Mukund: If you go down this list we can actually find out with respect to the marks therein increasing order.

Professor G. Venkatesh: 173, 188, 189, 198, 204, 210, 224, 240, 247, 250

Professor Madhavan Mukund: and 252 finally.

Professor G. Venkatesh: So, I can also now because I know the smallest in this is card index 3 I can because I have the bookmark, I can go straight to 3.

Professor Madhavan Mukund: And then you can find out what marks, what that names, names of the person anything that you…

Professor G. Venkatesh: Anything I want I can find out, Rida, Chennai whatever anything I want to find out. So, this is way to sort, you can sort now we have sorted by the sequence number, we have sorted marks, we can sort it also by maths marks you just give me another list.

Professor Madhavan Mukund: Will give you another list so then you can find…

Professor G. Venkatesh: Sort by physics marks you can even…

Professor G. Venkatesh: Even chemistry mark.

Professor Madhavan Mukund: You can even sort by birth date if you want starting from January to December.

Professor G. Venkatesh: So you can have any number of ordered list, this is an ordered list, any number of ordered list and each ordered list represents a way of sorting the set of cards. Now some of the problems that we did earlier I have been become much simpler, right because its ordered list.

Professor Madhavan Mukund: Correct, so we already saw that this keeping track of the top 3 prizes for instance if we had the sorted sequence we could have even done it in descending order instead of increasing order decreasing order, so that the top highest marks right to the beginning.

Professor G. Venkatesh: So, the top 3 first 3 marks will be the top 3.

Professor Madhavan Mukund: And we could keep the top 3, so we could keep 3 list; one for maths, one for physics, one for chemistry and then one for total. And then each total we just have to check whether they are in the early parts of the maths, physics, chemistry list.

Professor G. Venkatesh: And there we do not even see if it is not the first 3 you may want to go to the 4th, 5th you can do all that. If it is already sorted out.

Professor Madhavan Mukund: Exactly, so will be much simpler.

Professor G. Venkatesh: So much simpler to do all those exceptional cases that we discussed that there is no girl or boy, all those thing, so you can look at the top.

Professor Madhavan Mukund: So you can even find the first, if you have lot of boys or lot of girls right at the top you can find out where you have to go to find the first one of the different gender and then see what would happen, so many things you could do like that.

Professor G. Venkatesh: I also notice basically some other thing right the thing that we did much long back which is about giving grades, we try to give grades to all those students based on the marks. So we said let us find the minimum mark, let us find the maximum mark and divide this maximum minus minimum into some buckets, we just arbitrarily chose some buckets. And then the first bucket we gave A grade, B grade like that we did.

If it were a organize as a in descending order of marks let us say, we could look at the descending order go down the descending order and then first so many students we can give a grade.

Professor Madhavan Mukund: For instance you might decide that you want to give 20 percent of the students says, then you might look for the first 6 out of 30 give them A then you might say that you want to give so many Bs.

Professor G. Venkatesh: Say 35 percent you want to give B.

Professor Madhavan Mukund: So you can exactly do that.

Professor G. Venkatesh: You can do one third, you can find and then you can…so in order to divide equally. So, if it is sorted, if it is organized in a descending or ascending order many things that we did earlier may case turns out it is much easier. And as we discussed even that same birthday problem…

Professor G. Venkatesh: Yes, if it were sorted.

Professor Madhavan Mukund: For instance, if you want to find out for example in this case two students have got the same marks, instead of comparing students against students one by one you just first arrange them and then look they will come consecutively.

Professor G. Venkatesh: Consecutively they will come.

Professor Madhavan Mukund: Two people have got 240 they will appear in the same I mean one after the other in the list, so if you sort by birthday two people have the same day on which they are born then we will find…

Professor G. Venkatesh: Okay, let us take slightly more revolving right, this business of finding matching the student study pair. I want to take a student who is doing well in physics and pair that student with a student who is not doing well in physics but the other student must be doing well in some subject which this student is not doing, is there a something like that we can do this order, pair order?

Professor Madhavan Mukund: Well you could arrange it one order for physics and may be in the other order for maths and then try and see if they can be matched up in that order. You take the person who has done best in maths worst in physics and see if the corresponding thing works that they have done…

Professor G. Venkatesh: So, somebody in the top of this list, is in the bottom of this list mean if it is both in the…

Professor Madhavan Mukund: Yes, both in the same order, you wanted take one from the top of the list and one from the bottom of the list or the other way around and see if that is a good pairing. So that would be a just one good way to look for these because we did it kind of by inspection.

Professor G. Venkatesh: We did some cluster binning something, so we did use. But not a very systematic way of doing thing, but if you have ordered it potentially one could go down more systematically one could explore that space.

Professor Madhavan Mukund: Yes, that is right.

Professor G. Venkatesh: I think we should spend some time doing that later on, I mean that might be a good problem to work on pairing, okay.