IIT Madras

ONLINE DEGREE

Computational Thinking
Professor Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute
Professor G. Venkatesh
Indian Institute of Technology, Madras
Concept of Dictionary to Solve Birthday Paradox Problem

Professor Madhavan Mukund: So let us go back to this problem which we had seen to try and identify whether two students in this class are born on the same day and same month.

Professor G Venkatesh: Okay.

Professor Madhavan Mukund: So we have seen ofcourse one way to do it which is to...

Professor G Venkatesh: Sort it.

Professor Madhavan Mukund: Yes. Before that we did this nested loop and then we made this nested loop efficient by doing some binning and looking only within a month...

Professor G Venkatesh: Binning, Yes, we did an interested thing within a month.

Professor Madhavan Mukund: Yes, so we said okay, we do not need to compare everything across all the cards but within March, within April and so on.

Professor G Venkatesh: Correct.

Professor Madhavan Mukund: And then we said that if we instead arranged it in order of birthday by...

Professor G Venkatesh: But that, both of them are costly, I mean business of comparing every card is costly but binning reduces it somehow.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Sorting is itself costly.

Professor Madhavan Mukund: Once you sort it, then the two birthdays that are identical will come next to each other.

Professor G Venkatesh: So sorting, without sorting can it do, if you have a large number of cards.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Sorting can take...

Professor Madhavan Mukund: yes, so now we have been doing this new thing of indexing by some (category) some name.

Professor G Venkatesh: Dictionary.

Professor Madhavan Mukund: So we have this dictionary. So supposing we maintain a dictionary in which the index value is the birthday.

Professor G Venkatesh: Full day, full birthday.

Professor Madhavan Mukund: Full birthday, that is not only...

Professor G Venkatesh: Not only the month...

Professor Madhavan Mukund: Yes. Month and date in that.

Professor G Venkatesh: 3rd June, say 3rd June becomes a dictionary.

Professor Madhavan Mukund: And we count...

Professor G Venkatesh: Count how many people have 3rd June.

Professor Madhavan Mukund: Yes, so we keep a count of how many people are born on each day.

Professor G Venkatesh: It is like we did shopping category like that.

Professor Madhavan Mukund: Yes. Like we were counting how many food items each person had.

Professor G Venkatesh: So this is cool. So then basically because the dictionary is efficient, why we know dictionary is efficient? Because computer tells you how to go to that...

Professor Madhavan Mukund: You can quickly update something or look it up or see if it is not there.

Professor G Venkatesh: See if it is there or not.

Professor Madhavan Mukund: so we can do that, right?

Professor G Venkatesh: So that must be, then that is really fast.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Then you do not have to compare any card with every card...

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: So we will just once...

Professor Madhavan Mukund: We can build this dictionary in one scan and then check whether there is any item which appears...

Professor G Venkatesh: which has a count of 2.

Professor Madhavan Mukund: Yes at least 2, right? That means there are at least 2 people born on the same date.

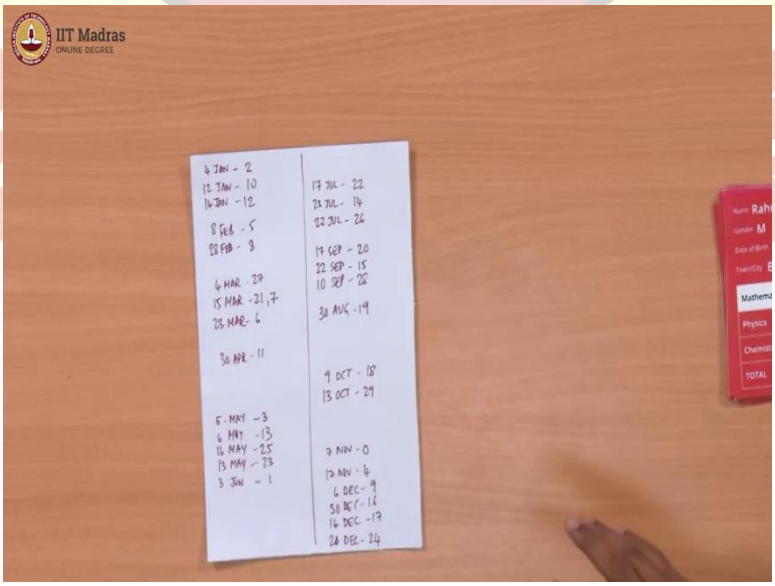
Professor G Venkatesh: Let us keep the card IDs.

Professor Madhavan Mukund: Yes let us keep the card IDs.

Professor G Venkatesh: Because then we will know who...

Professor Madhavan Mukund: So we will know who it is.

(Refer Slide Time: 02:18)



Professor G Venkatesh: Alright, so what are we doing? We are maintaining a dictionary.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Where the first column is, is the; what is it? It is the date of birth.

Professor Madhavan Mukund: Yes it is the day and month.

Professor G Venkatesh: Day and month, so 3rd June we will keep.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: And the second is the card ID.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: And then what we want to do at the end of this procedure, iteration is we want to see whether or not this dictionary contains any element which has two entries.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Right?

Professor Madhavan Mukund: Correct.

Professor G Venkatesh: It has two cards in it. Okay. Alright, let us go.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: So the first is 3rd June.

Professor Madhavan Mukund: Okay so just to simulate that we are going to find it fast. I am going to because I know, because we know that these are birthdays. Days of the month that June is roughly midway.

Professor G Venkatesh: So you are going to write in the middle.

Professor Madhavan Mukund: So i will writer, so this is January to June and this is going to be July to December. So I will write it somewhere here. So I will write somewhere here that 3rd June and the number is 1.

Professor G Venkatesh: 1, card number is 1. Then 5th May.

Professor Madhavan Mukund: 5th May

Professor G Venkatesh: 3

Professor Madhavan Mukund: Card number is 3.

Professor G Venkatesh: 4th Jan, so write on top, I guess 4th Jan. Card number is 2.

Professor Madhavan Mukund: So this is not the count but the card number.

Professor G Venkatesh: Card number. Yes and if there are more than 1 card, we will put a comma like a list...

Professor Madhavan Mukund: Yes we will, we are collecting a list of cards.

Professor G Venkatesh: List of cards. 22nd September.

Professor Madhavan Mukund: So it is somewhere here.

Professor G Venkatesh: 15, 17 July.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 22

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 23rd July.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 14, 28th February, 8.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 14th January, 12.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 17th November.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 4

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 7th November.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: So because of your technique here, you know whether it is there already or not quickly.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Probably the computer is doing something.

Professor Madhavan Mukund: Computer is doing something like that.

Professor G Venkatesh: 17th September.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 20.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 15th March.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 21, 22nd July

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 26, 12th January; 10, 26th December; 24.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 6th May.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 13.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 16th May

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 25, so far we have not seen anything, right?

Professor Madhavan Mukund: No.

Professor G Venkatesh: 4th March.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 27.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 30th August.

Professor Madhavan Mukund: Okay, so I will write it below September because I have run out of space.

Professor G Venkatesh: 19.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 13th October.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 29.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 14th December

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 17. 23rd March.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 6. 10th September.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 28. 30th December.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 16. 6th December, 9. 13th May.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 23. 15th March.

Professor Madhavan Mukund: We have a 15th March already.

Professor G Venkatesh: Oh You have? Wonderful! So we found the second.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: 7, 8th February.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 5. 9th October.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 18.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 30th April, 11. So this not only told us that we have one item with two, list with two elements but it also told us that for all others there are not any.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: So there is only one...

Professor Madhavan Mukund: Correct. So there is only one birthday which is in common.

Professor G Venkatesh: Birthday in common.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: And we did this very efficiently because we use a dictionary.

Professor Madhavan Mukund: Correct.

Professor G Venkatesh: Straight away go there and update.

Professor Madhavan Mukund: So we kept a count. Implicitly we kept a count for everyday in the year but we only actually use the count or the list for those days which appeared. So this is another thing. By looking at whichever names appear in this dictionary we actually know which days are actually birthdays of somebody in the class.

Professor G Venkatesh: Okay, that is another thing we can get.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Same effect.

Professor Madhavan Mukund: Yes, so we can actually get from these...

Professor G Venkatesh: You can get all the birthdays.

Professor Madhavan Mukund: Yes from the indices, we know which values are...

Professor G Venkatesh: how many are there? Can you count? How many birthdays are there? Just curiosity.

Professor Madhavan Mukund: It should be 29.

Professor G Venkatesh: It should be 29, Correct. Correct.

Professor Madhavan Mukund: 1, 2, 3, 4, 5...

Professor G Venkatesh: So 30 minus 1, right? Because 1 is gone…

Professor Madhavan Mukund: Yes, 13, 16, 20, 22, yes 29.

Professor G Venkatesh: 29. So we have 29 birthdays and 30 students

Professor Madhavan Mukund: Yes. But this idea you can actually keep track of the values also by how many values have come, is also interesting. So it is an interesting extra benefit that you get from a dictionary that to know whether something is, if you ask the following question, when I get a new birthday, does somebody has this birthday already or not. So that is part of the process of looking up the dictionary.

Professor G Venkatesh: Correct. Correct. Correct

Professor Madhavan Mukund: So we can also use it to keep track of how many different birthdays...

Professor G Venkatesh: Birthdays.

Professor Madhavan Mukund: actually have come and so in the previous thing for instance, when we were looking at shops and customers if you are doing it, you can also see who are all the customers whose bills have come. So whenever you use names as indices then...

Professor G Venkatesh: So we have seen some very interesting things, right? So we have started with lists, simple lists then we sorted the lists, ordered the lists. Some use of it. Now we have some very interesting things, it is not sorted actually.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Though you have written it like this.

Professor Madhavan Mukund: Even, here is not sorted, because I wrote August after September.

Professor G Venkatesh: So it is some jumble of things, but you have a way of randomly, directly going...

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: to there.

Professor Madhavan Mukund: Yes

Professor G Venkatesh: Dictionary element and doing some update on it and then what we saw basically is that inside that dictionary element itself, you can have a list.

Professor Madhavan Mukund: Yes. Like here, these are lists.

Professor G Venkatesh: That is a list.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: So, now we have a combination of a dictionary and a list.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: That should be a quite a powerful...

Professor Madhavan Mukund: Yes. So basically in a variable you could have a dictionary or a list but what is there inside a list is also a or a dictionary is also like a variable. So it could also be a list.

Professor G Venkatesh: So you could have a very powerful, it is a very-very powerful way of organising.

Professor Madhavan Mukund: yes. So we have very complex ways now of arranging...

Professor G Venkatesh: Arranging things. Hopefully we can use it, right?

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: We can use it to solve some...

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Some different types of problems.

Professor Madhavan Mukund: Yes.