



# IIT Madras

ONLINE DEGREE

**Computational Thinking**  
**Professor Madhavan Mukund**  
**Department of Computer Science**  
**Chennai Mathematical Institute**  
**Professor G. Venkatesh**

**Indian Institute of Technology Madras**

**Resolve pronoun with its equivalent matching noun using dictionary and ordered lists**

Professor G. Venkatesh: So, we saw this idea of dictionary last time where we were trying to group together or index things by some name or category like say the date of birth.

Professor Madhavan Mukund: Date of birth.

Professor G. Venkatesh: So we were trying to group the students by date of birth and then we found that if two students went into the same group that meant that they had same birth date and then also we were doing the shops if you wanted to collect the total bill for example by customer then we found or the maximum purchases of some particular type by customer then instead of having a separate variable for each customer.

Professor Madhavan Mukund: We said variable.

Professor G. Venkatesh: Because, we said we have only three shops we can keep three variable but customers are many.

Professor Madhavan Mukund: Many.

Professor G. Venkatesh: And new ones may come if we get new bills.

Professor Madhavan Mukund: Right.

Professor G. Venkatesh: So, we is better to have one collection where we index by the name of the customer and we keep for each customer, we keep the value rather than creating a new variable for every customer so, this was our new.

Professor Madhavan Mukund: Dictionary.

Professor G. Venkatesh: New way of collecting information and then we have also said that with the dictionary we could have list like the list of students born on a particular day.

Professor Madhavan Mukund: Or even list of bills.

Professor G. Venkatesh: List of bills for a given customer and so on, so is here anything interested we can do with this word data set using dictionaries.

Professor Madhavan Mukund: See whole problem which we try to do noun, pronoun, resolution may be there is way we do it with list because you know specifically this category every word has a unique category.

Professor G. Venkatesh: So, you want to resolve these personal pronouns so, we know the category of word is a pronoun.

Professor Madhavan Mukund: Pronoun.

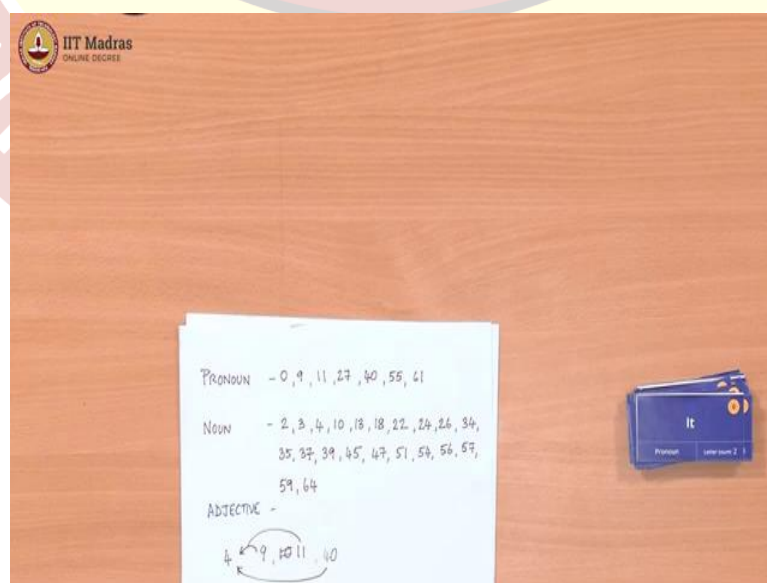
Professor G. Venkatesh: And we are looking for a matching noun.

Professor Madhavan Mukund: So why cannot we just take out all the cards and make a dictionary by category so pronoun dictionary which basically has categories as the entry index. So we can basically make a list of all pronouns, list of all the nouns, list of all the verbs. They may be useful because we may want to do something with dissolving word with find the subject to the verb, find object to the verb or something like that all these kind of interesting questions right now we are doing pronoun noun but, for all of that we will need to make a dictionary some kind of dictionary where we can keep together.

Professor G. Venkatesh: So, let us try and do that. So maybe we can try to do that, so let us try to make a dictionary shall we do it for all types or only for pronouns.

Professor Madhavan Mukund: Dictionary can be for all types but we will only keep track pronoun and noun right now.

(Refer Slide Time: 02:41)



Professor G. Venkatesh: So, we will start with so we have so we will just keep maybe the word position instead of writing the word.

Professor Madhavan Mukund: Index should be good enough.

Professor G. Venkatesh: So, we had this.

Professor Madhavan Mukund: I go through I can say whether it is pronoun or noun I think it is easy and then you can just record it so, you make a list of so the dictionary is basically storing the list of cards.

Professor G. Venkatesh: List of cards of a given type.

Professor Madhavan Mukund: Of a given type do we turn it around or I will keep it like this or.

Professor G. Venkatesh: Turn it around so we keep the order I think.

Professor Madhavan Mukund: So, it is pronoun and word zero.

Professor G. Venkatesh: So, the first pronoun is zero this is verb I am skipping it.

Professor Madhavan Mukund: Monday is a noun 2, Morning is a noun 3, Swaminathan is a noun 4, this verb, adjective, preposition, verb, pronoun his is a pronoun 9, eyes is a noun 10, he is a pronoun 11, Monday is a noun 13, so each time you call noun or a pronoun I added to the list for the appropriate dictionary and term.

Professor G. Venkatesh: Because not all nouns are interested to us.

Professor Madhavan Mukund: Interested name nouns will see later specially is an adverb, unpleasant adjective, in is a preposition, the is an article, calendar is a noun 18.

Professor G. Venkatesh: 18.

Professor Madhavan Mukund: Preposition, article, adjective, freedom noun 22, preposition, Saturday noun 24, conjunction, noun Sunday 26, it is a pronoun 27.

Professor G. Venkatesh: 27.

Professor Madhavan Mukund: Verb, adjective, preposition, verb, preposition, article, Monday is a noun 34.

Professor G. Venkatesh: 34.

Professor Madhavan Mukund: So, we seen Monday already I think but you should keep both

Professor G. Venkatesh: So, this is a different copy of Monday.

Professor Madhavan Mukund: Even been something else Monday can only been one thing mood is a noun 35, preposition, work is a noun , see work for example could be verb. I work here, so it is worth keeping the card as this, work 37 noun, discipline noun 39, he is a pronoun 40.

Professor G. Venkatesh: Many more nouns and pronouns that looks like.

Professor Madhavan Mukund: So, resolution would be harder, shuddered 41.

Professor G. Venkatesh: That is a verb.

Professor Madhavan Mukund: Verb, preposition, article, adverb, thought noun.

Professor G. Venkatesh: 45.

Professor Madhavan Mukund: 45 preposition, noun school 47, the, dismal, adjective, building is a noun 51, article, adjective, Vedanayagam is a noun 54, his 55, class is a noun 56, teacher is a noun 57, conjunction, headmaster is a noun 59, preposition, his is a pronoun 61, thin, long, cane is a noun 64.

Professor G. Venkatesh: So, now we could have done this for every type of category but, we have already done it for pronouns and nouns.

Professor Madhavan Mukund: Now, we want to do some kind of resolve pronoun for noun resolve so, what we said is for every pronoun that we want to resolve we have to look for the nearest noun before that and say that this pronoun refers to that noun.

Professor G. Venkatesh: So, we cannot do it for all pronouns. For example it, 0, it does not make sense so next pronoun is at 9.

Professor Madhavan Mukund: 9 so I can go order nine directly.

Professor G. Venkatesh: So, since the cards were in sequence these lists were also in sequence.

Professor Madhavan Mukund: Also, in sequence.

Professor G. Venkatesh: It is ordered.

Professor Madhavan Mukund: It is ordered actually. It is a ordered list that makes very difference because you can search so 9 his now, what we said is we have to find a noun.

Professor G. Venkatesh: Which is before 9.

Professor Madhavan Mukund: Before 9.

Professor G. Venkatesh: So, now we know the noun so we know that the noun is before 9 so we started could be 2, 3, 4 and we wanted to look for the nearest one.

Professor Madhavan Mukund: But, the nearest one should also be a sensible one.

Professor G. Venkatesh: Let us see.

Professor Madhavan Mukund: Let see so 2 is 1, 2, 2 is Monday anyway that is not the nearest one, three is morning, four is Swaminathan.

Professor G. Venkatesh: So, 9 in some sense so, this is the pronoun and this is the noun so, by without going back and reading the whole thing just by looking at this because now are not interested in the adjectives, adverbs verbs so we can kind of and now we know that 4 is for 9 so, now let us see what is 11 the next pronoun.

Professor Madhavan Mukund: Next one is a he.

Professor G. Venkatesh: So, now we know that there is no except for 10 what is 10.

Professor Madhavan Mukund: Eyes.

Professor G. Venkatesh: Obviously eyes is not swaminathan noun others so, 10 is also sorry 11 is also back to the same 4 so, 27.

Professor Madhavan Mukund: Next is 27.

Professor G. Venkatesh: This is also good because we can quickly find the pronouns without having to read the whole.

Professor Madhavan Mukund: 27 is it.

Professor G. Venkatesh: So, it normally we do not think of is.

Professor Madhavan Mukund: We are thinking early personal pronouns refereeing to so, it is not good in fact find the noun object but anyway we are doing right now so leave it then we go to 40.



Professor G. Venkatesh: 40.

Professor Madhavan Mukund: 40 is he.

Professor G. Venkatesh: So, this should be something close to 40 if we go now so we have to go back and search so it is obviously beyond four because four is where we were we have to go and then 39 is the nearest one.

Professor Madhavan Mukund: 39 is discipline.

Professor G. Venkatesh: So, then 37.

Professor Madhavan Mukund: 37 is work.

Professor G. Venkatesh: Now, 35.

Professor Madhavan Mukund: 35 is mood.

Professor G. Venkatesh: Now, 34.

Professor Madhavan Mukund: 34 is Monday.

Professor G. Venkatesh: 26 you have to go back a bit.

Professor Madhavan Mukund: 26 Sunday.

Professor G. Venkatesh: 24.

Professor Madhavan Mukund: 24 is Saturday.

Professor G. Venkatesh: 22.

Professor Madhavan Mukund: 22 is freedom.

Professor G. Venkatesh: 18.

Professor Madhavan Mukund: 18 is calendar, I think we come back to swaminathan only. We did it last time.

Professor G. Venkatesh: 13.

Professor Madhavan Mukund: 13 is Monday.

Professor G. Venkatesh: 10.

Professor Madhavan Mukund: 10 is eyes.

Professor G. Venkatesh: So, basically even for 40.

Professor Madhavan Mukund: We came back to 4.

Professor G. Venkatesh: Finally, if we think about this then 40 also come back to same four.

Professor Madhavan Mukund: I say easy actually basically because it is sorted just have to find the one before it.

Professor G. Venkatesh: Correct and then go backwards.

Professor G. Venkatesh: Find the one which is of the right the right type of noun because it is a person it cannot be Monday or work or something like that.

Professor Madhavan Mukund: So, again here we saw what is seeing we are seeing that we have made a dictionary, dictionary basically is dictionary by category so you will have noun, pronouns, adjectives, adverbs, verbs, prepositions, article whatever it is and inside it we are storing the list of cards you could actually in this case you could actually put card also because.

Professor G. Venkatesh: Because, each card is of one has a unique category.

Professor Madhavan Mukund: Unique category.

Professor G. Venkatesh: And even you know the same card appears twice like you said work if it appears in the text somewhere else as verb and once as an noun then the copy there will be another number, different number, so this is actually separating out the words into these different categories.

Professor Madhavan Mukund: This partitioning.

Professor G. Venkatesh: Partitioning them so every word belongs to exactly one of these. To set is partitioned into these categories and so you could keep the card but you could also keep the index. We have kept the index because you may use you may have another dictionary for example this dictionary is making sure that everything is in one, it is useful actually everything goes into one bin because this is binning.

Professor G. Venkatesh: This is binning..

Professor Madhavan Mukund: This is what we discuss as binning.

Professor G. Venkatesh: This is dictionary which actually is binning.



Professor Madhavan Mukund: Binning which is partitioning.

Professor G. Venkatesh: Partitioning.

Professor Madhavan Mukund: But you may have multiple dictionaries so therefore it might be useful to keep the index.

Professor G. Venkatesh: And for instance if I mean just as an example I think if we had just say adjective where kept we did not record it but, we could have as we went along then you could also ask for reasons this adjective is refereeing to which noun so again you can do a similar thing.

Professor Madhavan Mukund: Similar thing find the earlier adjective or later.

Professor G. Venkatesh: Or the nearest noun either side.

Professor Madhavan Mukund: And decide.

Professor G. Venkatesh: Which one it is, so all kinds of analysis you can do just looking at this list and not going back and going through the whole paragraph again.

