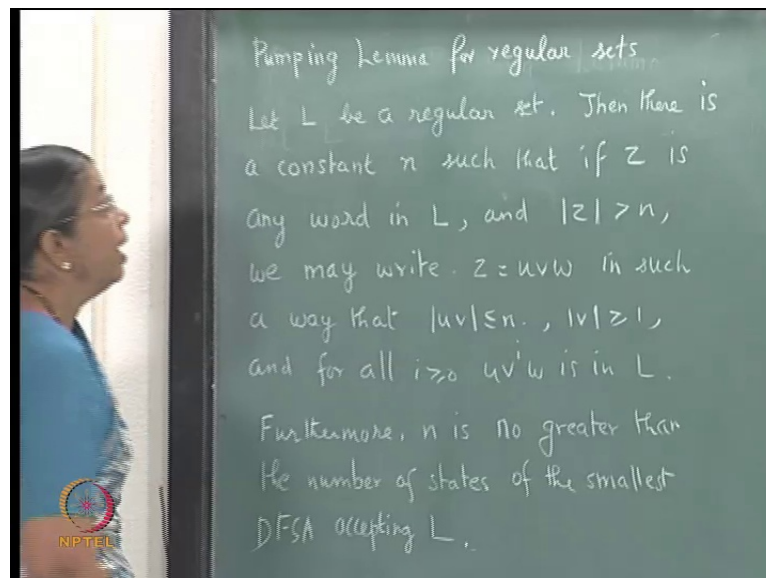


Theory of Computation
Prof. Kamala Krithivasan
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture No. #16
Pumping Lemmas For Regular Sets And CFL

So, today we shall consider pumping lemma for regular sets. This is very useful in proving certain sets are not regular. For example, you know that $a^n b^n$ is not regular. You cannot generate it with a right linear grammar or you cannot accept it with a DFSA. Why? The finite state automaton has a finite amount of memory, and if you want to accept $a^n b^n$, it should somewhere remember n , but it cannot remember infinite number of values. So, it is not possible intuitively, this is the argument. But formally, we can use what is known as the pumping lemma? And prove it.

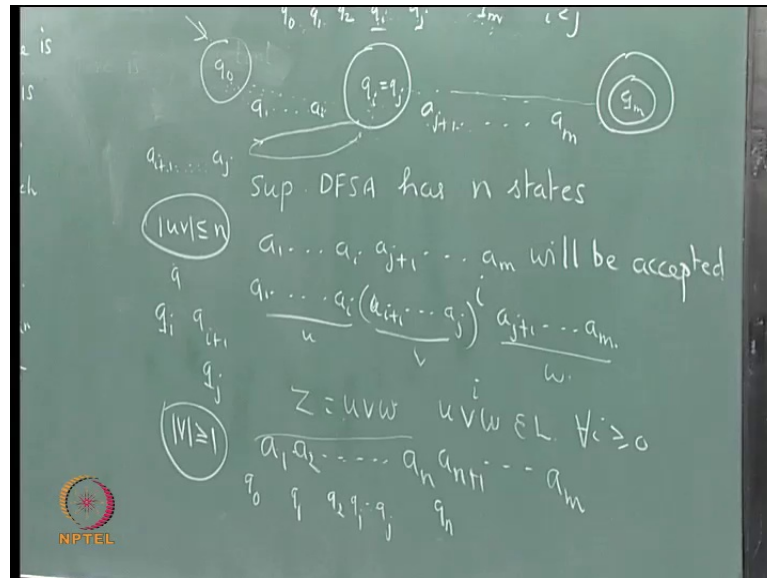
(Refer Slide Time: 01:02)



Now, what is the pumping lemma? The pumping lemma is like this, let L be a regular set, then there is a constant n such that if Z is a word in L and $|Z| > n$, the length of Z is greater than n . Then, you can write Z in the form uvw , such a way that the length of uv alone is less than or equal to n , and v is non empty, its length is greater than or equal to 1, such that

for all i greater than or equal to 0, $u v^i w$ is in L . This is called the pumping lemma for regular sets.

(Refer Slide Time: 01:50)



The idea is very simple, look at this structure of a word Z accepted. Suppose Z is a_1, a_2, \dots, a_m , and m is greater than n . You start with q_0 and after reading some portion, you reach a state q_i , and then you reach a final state q_m . Now, initially start with q_0 , after reading a_1 you go to some key q_1 , after reading a_2 you go to q_2 , and soon. After reading a_m you go to the final state and a_1, a_2, \dots, a_m is accepted by the automata.

Now, if m is greater than n , how many states are there? There are $m+1$ states here. Suppose the DFSA has n states. Suppose f DFSA has n states, the n corresponds really to the number of states. Furthermore, n is no greater than the number of states of the smallest DFSA accepting L . So, L is accepted by a machine having n states and Z is accepted by going from q_0 to a final state q_m .

Now, you see that $q_0, q_1, q_2, \dots, q_m$ there are $m+1$ of them and you are having only n states. So, by pigeonhole principle, some q_i and some q_j will be equal. They will be the same; you cannot have $m+1$ different states. So, two of them will be equal. So, from say q_i is equal to q_j . Then starting from q_0 and i is less than j say i is less than j , then starting from q_0 after reading a_1, a_2, \dots, a_i you go to q_i .

Then, after reading a_1 to a_j , you go to the same state q_j is the same as q_i . And, from a_{j+1} to a_m . You go from this to the final state you have a string $Z = a_1 a_2 \dots a_m$ accepted by the machine and the length of that is greater than n the number of states of the deterministic automata. Then there is a sequence of states which leads you to acceptance and these states all cannot be different from q_{naught} . After reading a_1 it goes to q_1 , after reading a_2 it goes to q_2 and so on.

There are $m+1$ states in this sequence and total you have only n different states. So, two of them have to be equal by pigeonhole principle. So, q_i is equal to q_j , say where i is less than j , then starting from q_{naught} , after reading $a_1 a_2 \dots a_i$, you go to this state. And after reading $a_{i+1} a_{i+2} \dots a_j$ again you go to this state, and a_j after from here, after reading $a_{j+1} \dots a_m$, you go the final state.

Now, you can see that $a_1 a_2 \dots a_j a_{j+1} \dots a_m$ will be accepted. Is not it from here to here, then here to here. It will be accepted and then a_1 to a_i then a_{i+1} to a_j . Starting from here, you can go here and you can traverse this path any number of times you want. And, then go to the final state, this will also be accepted by the same DFSA. It is a very simple concept. So, Z you can write this, you can call as u and this you can call as v and this you can call as w .

So, Z , you can write as $u v^i w$, such that $u v^i w$ belongs to L for i greater than or equal to 0. When i is 0, it is $u w$ like this, when i is 1. Whatever you had earlier $u v w$ i is 2 is $(()) u v v w$ and. Soon. So, the idea is here is very simple, now look at the statement of the lemma, when the length is greater than n . It may be written in this form, such that $|u| \leq n$ and $|v| \geq 1$ that is v is not empty. We are considering it deterministic automata length of $u v$ is less than or equal to n . Why, because if you see $a_1 a_2 \dots a_n$ plus, you are taking a string like this.

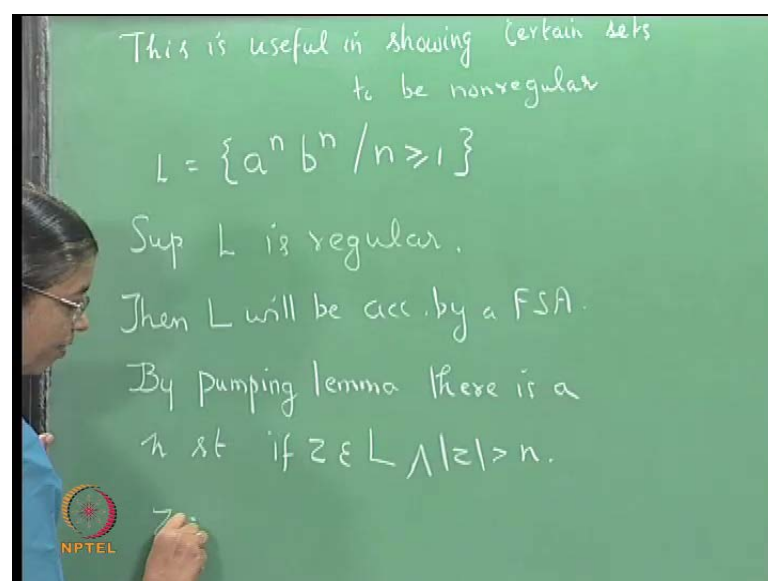
So, start from q_{naught} , after reading a_1 you go to q_1 , after reading a_2 you go to q_2 after reading a_n you go to q_n . So, among the $q_{naught} q_1 q_2 \dots q_n$ there are $n+1$ states. Here itself something has to repeat, you have only n states in the DFSA and in this sequence you are having $n+1$ states. So, by the pigeonhole principle here itself some q_i and some q_j will be equal. So, the $u v$ portion will occur within this. So, the rest of the portion, you can consider as I mean after the pump, you can get rest of the portion as w .

So, uv portion will occur in this. So, that is why you say length of uv is less than or equal to n , length of uv is less than or equal to n and two of them are repeated even if it is a single symbol. $q_i q_j$ even, if q_j is $q_i + 1$. You would have read one symbol at least is not it two of them are equal. Even, if you take the worst case that is successive states are identical, going from q_i to $q_i + 1$. You would have read one symbol. So, the length of v is length of v , it is not empty, it is at least one symbol, it can be more than that. So, you have these two conditions length of uv is less than or equal to n and length of v is greater than or equal to 1.

So, that is all the proof. So, let us look into the statement again. Let L be a regular set, then there is a constant n , such that constant is really the number of states of the DFSA. If Z is a word in L and the length of Z is greater than n , that is if you have a sufficiently long string accepted by the machine, then you can pump the middle portion uv power i Z . you can write as $u v w$ such that $u v^i w$ belongs to L .

Note that you can also have u empty v cannot be empty, but u can be empty. The pump can occur in the beginning itself, it is possible, see from here itself the pump can occur. So, because you are putting more and more v it is called pumping lemma. You are pumping the middle portion, is not it? vv squared v cubed like, That is why? the name pumping lemma for this

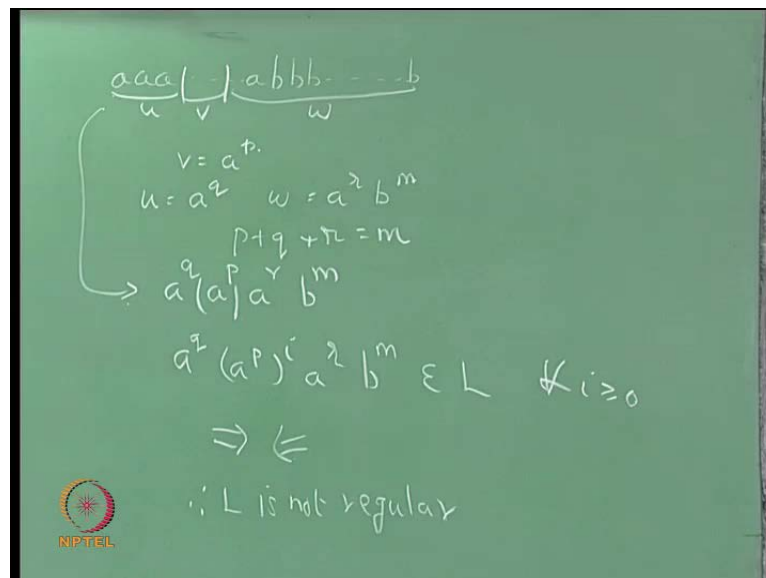
(Refer Slide Time: 11:17)



Now this is useful in showing that certain sets are not regular. This is useful in showing certain sets to be non-regular; you can show some sets are not regular using this lemma. Let us say, take the simple example a power n b power n greater than or equal to 1. This is not regular intuitively, the argument is you cannot. A finite automaton cannot remember everything if it has to accept a power n. You have it has to remember every n it is not possible with a finite amount of memory, even with the case of a carry in the case of a binary adder which you considered earlier.

It just distinguishes between two classes of input histories, one is that which produce the carry and that with a which produce one with carry. And another without carry. It distinguishes only between them, you cannot know each and every possible input history. So, if it has to accept a power n and b power n, it has to remember after reading the entire a sequence of a. It has to remember how many a(s) it has read. That is not possible by finite state automata. This is the intuitive argument tell us use the pumping lemma. Suppose the argument you have to write like this, suppose L is regular then L will be accepted by a FSA by pumping lemma. There is a n such that if Z belongs to L and length of Z greater than n.

(Refer Slide Time: 14:37)



Z can be written in the form u v w such that u v powers i w belongs to L. This is the pumping lemma; now consider a power m b power m. Where m is greater than n. So, you have a string aa followed by an equal number of b's. Now, using pumping lemma, this can

be written in the form $u v w$. And the other condition is length of $u v$ is less than or equal to n and v is greater than or equal to 1. So, the pump $u v$ portion will occur here.

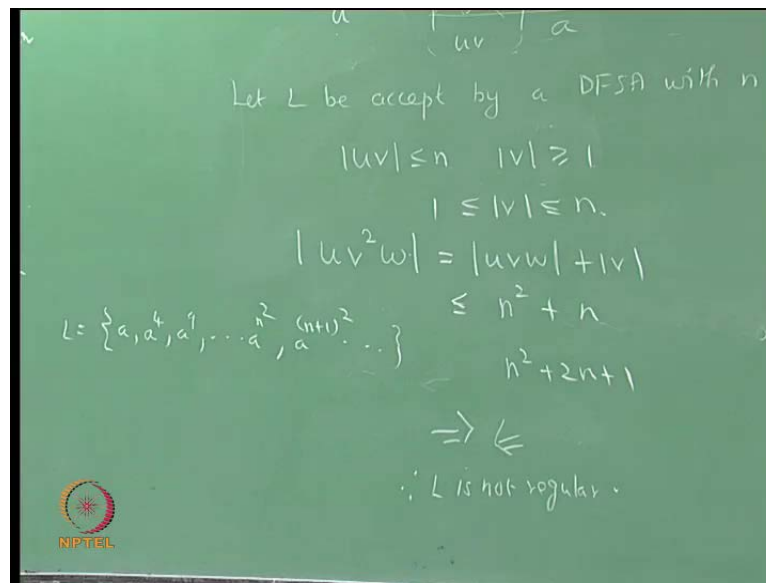
So, suppose something like this is there, this is v and this is w . We will not fall into $b u v$ will fall within the portion, you have chosen that way. You have taken one string $a^m b^m$, where m is greater than n . So, the $u v$ portion will occur within this. Suppose v is some a^p , then what happens u is some a^q w is $a^r b^m$. Where, what do you have? $p + q + r$ is equal to m . Is not it. So, you have a $a^q a^p a^r b^m$, the string is of this form $p + q + r$ is equal to m and you can pump this portion.

So, you will get a $a^q a^p a^r b^m$ belongs to L for all i greater than or equal to 0. That means, you can get a $a^q a^p a^r b^m$ belonging to L you can get a a^i . So, many strings you can get where the number of a 's will not be equal to the number of b 's. If it is a regular set, it has to satisfy the pumping lemma and if it satisfies the pumping lemma, you come to the conclusion that there are strings of the form $a^n b^m$ something like that.

Where the number of a 's is not equal to the number of b 's and such a string will belong to the language. You come to that conclusion, but that is not correct, the language L has strings of the form, where you have only equal number of a 's and equal number of b 's. String of a 's followed by an equal number of b 's. So, here you are arriving at a contradiction. Therefore, L is not regular, any questions? Let me work out one more example, the argument is similar, instead of this, I will consider a power n squared.

You want to show that a power n squared is not regular. Suppose L is regular, then L will be accepted by a FSA. And the pumping lemma, there is an n such that if the length of n^2 is greater than n . Then Z you can write in this form, such that $u v^i$ belongs to L

(Refer Slide Time: 18:37)



Now, here consider a power n square, it is a string of a 's and the first uv . Let L be accepted by a DFSA with n states, that is the n is for the pumping lemma. So, if you take a power n squared, you can write it like this. And the uv portion will be within the first n symbols.

So, suppose you consider length of uv is less than or equal to n , and length of v is greater than or equal to 1. So, length of v will be between 1 and n . And this is some portion is u some portion is v , and then it is w . So, if you consider uv squared w . What can you say about the length? The length of this will be length of uvw plus length of v . But what will be the length of uv^2w ? You have taken a power n squared. So, it will be n squared plus whatever it is, but it is less than or equal to n is not it length of v is less than or equal to n .

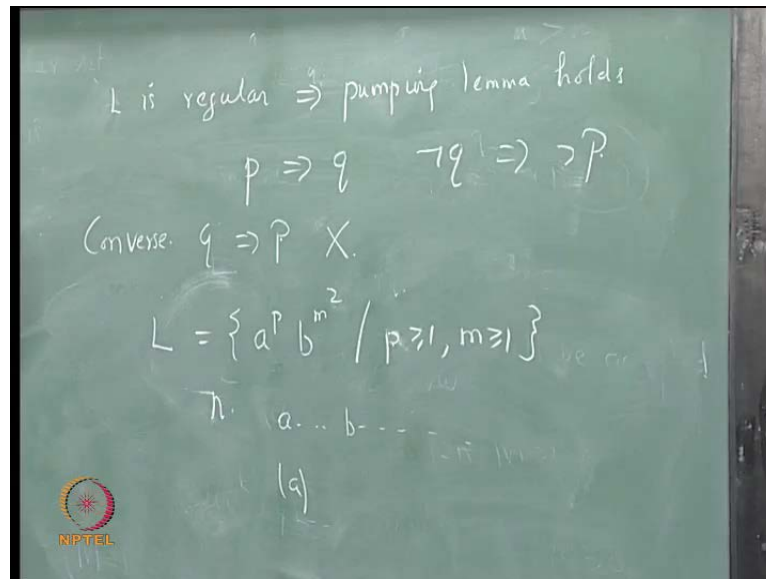
So, you get something a power some p or something whose value p value of p is n squared plus n . It is less than or equal to n squared plus n , it will be greater than n squared. So, after n squared, the next string you get is a power n plus 1 squared. Is not it? If you write a language L will be $a^4 a^9$. Then a power n squared the next string will be a power n plus 1 squared and soon. Is not it? and n plus 1 squared is n squared plus $2n$ plus 1.

So, by this argument you are getting a string which is between, whose length is between this and this, which is not possible. The language, the string whose length is next longer

string than a power n squared is a power $n + 1$ squared. But by this argument, you are getting a string whose length is in between n squared and $n + 1$ square which is not correct. So, you are arriving at a contradiction, therefore, L is not regular.

Many other languages, you can show to be non regular using this argument.

(Refer Slide Time: 22:22)



Now, pumping lemma says that L is regular, pumping lemma holds. The argument the way we have done is pumping lemma does not hold. So, L is not regular. So, what we are proving the **(())**.

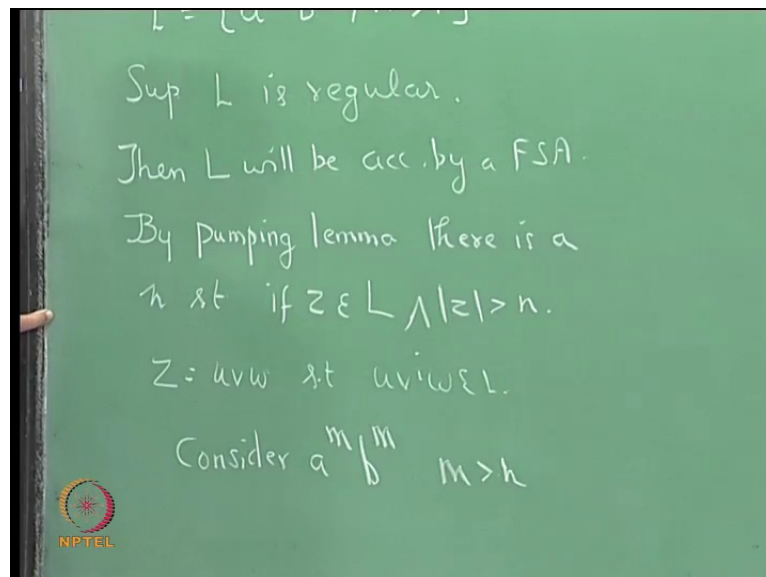
Contrapositive, we are proving the contrapositive. This is also from p implies q what we are showing is $\neg q$ implies $\neg p$, to show that this the argument we are giving to show that something is not regular. And it is like anyone can look at it as a sort of a game between two people. You are saying that you are choosing here, you want to show L is not regular. The adversary chooses n , he can choose any n , but once he chooses some n he keeps **(())**. Then you choose Z , the Z way you choose Z . You may choose depending upon n that is what you have done.

See if we chose a power m b power m . Where m is greater than n or a power n squared is not it. So, the adversary chooses n and you are choosing Z , then the adversary can write Z in the form uv for some uv satisfying the condition uv is less than or equal to n and v greater than or equal to 1. Then you show that uv power k does not belong to L for

some k . You show that this is not possible, that is why? It is not regular. This is the way it goes.

So, here if you look at the statement for all L , L is a regular set there exist n for all Z . Such that Z of Z is Z belongs to L and length of Z is greater than n . There is a way to write it in the form Z is equal to uvw . The statement of the pumping lemma, you look at it, then for all i Z $uv^i w$ belongs to L . This is the way the pumping lemma it is sort of you use in alternate manner for any regular set, for all regular of any regular set there a n for all Z belonging to L , such that length of Z greater than n .

(Refer Slide Time: 24:03)



z can be written in the form uvw , there is a way to write Z in the form, such that $uv^i w$ belongs to L . And you are corresponding to the for all quantifier, the adversary is that there exist quantifier. So, you first choose L then the adversary chooses n then you choose Z . Depending on n , you can also fix the Z once you know what is n ? Then adversary can write it in the form uvw . So, you show that for some k , $uv^k w$ does not belong to L . This is the sort of argument; we give for proving some sets are not regular.

Now, p implies q this is of the form, and what we have used is? Naught q implies naught p , contrapositive. What is a converse? What is a converse to pumping lemma? It will be of the form q inverse p , is not it? Converse is this, if you make a statement, it will be like this. If there is n such that for any Z belonging to L and L is some language, you start with L is some language. Then if there is a n such that for any Z of length greater than

n. You can write Z in the form $uv^i w$ such that $uv^i w$ belongs to L for all i then L is a regular set. Is not it? The converse will be of that form. Is it true? Is the converse true? Do you think it will be true?

How many of you think it is true? Converse will be true? How many of you think it is not true? How many of you think that converse will be true? You cannot say converse is not true. For that you consider that is the pumping lemma may hold L may not be regular. If the pumping lemma does not hold L is not regular, but if the pumping lemma holds you cannot conclude L is regular L may be regular or non regular you cannot say. So, take L to be a power n b power m squared n greater than or equal to 1 m greater than or equal to 1. Consider this is a regular set. This is not a regular set number, you can have a sequence of a 's, but the sequence of b (s) is m squared.

So, this is not a regular set, but if I choose n instead of n . See I do not want to use a same n , may be I will say a power p , p greater than or equal to 1 and m greater than or equal to 1. Suppose this satisfies the condition for all there is a n , then you can choose a string see the p is greater than or equal to 1 there will be at least one a . So, the string will be of the form ab .

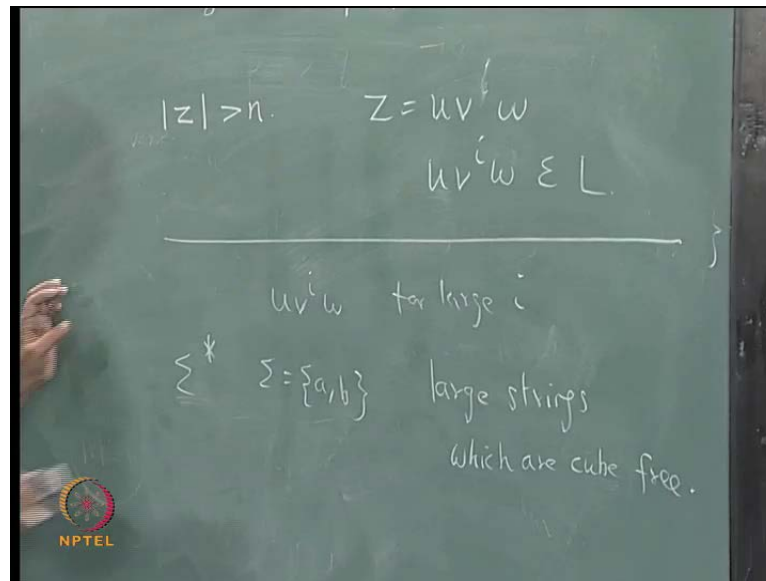
If you take a string of length n , whatever it is, the first symbol has to be a . It cannot begin with a b and the pump you can just pump at the first symbol alone you can pump. Then the resultant string will be in the language. Is not it? This union you have to also have b power q q greater than or equal to 1. Some sequence of a (s) non empty string of a (s) followed by a sequence of b 's, sequence of b (s) will be of the form m squared otherwise just sequence of b (s) alone. The reason for that, if it is if you take i to be 0 and you start with the string say $abbb$ something. If you pump a i is equal to 0 you will just get a string of b 's.

This you can pump, i times i , if you take to be 0 you have to get a string of b 's. Sequence of $(())$ that is why this is regular, anyway this is not regular. You can see that this is not regular because of this portion you have a sequence of a (s) followed by the sequence of b (s) which is not of a form, it is of this form. So, it not regular, but you can have any sequence of this without a 's. You can have any sequence of it this portion is regular does not matter, but the whole thing is not regular.

So, the converse to pumping lemma is not true, there is something like a converse to a pumping lemma which is a very involved theorem. It is not given in any book I think it is

only in the paper. There is a result by Ehrenfeucht, Parikh, and Rosenberg which is something like a converse to pumping lemma. Which assumes more condition, and then you show that under these conditions, it is not only pumping lemma. You identify certain positions and say that certain things can happen, occur and soon.

(Refer Slide Time: 33:05)



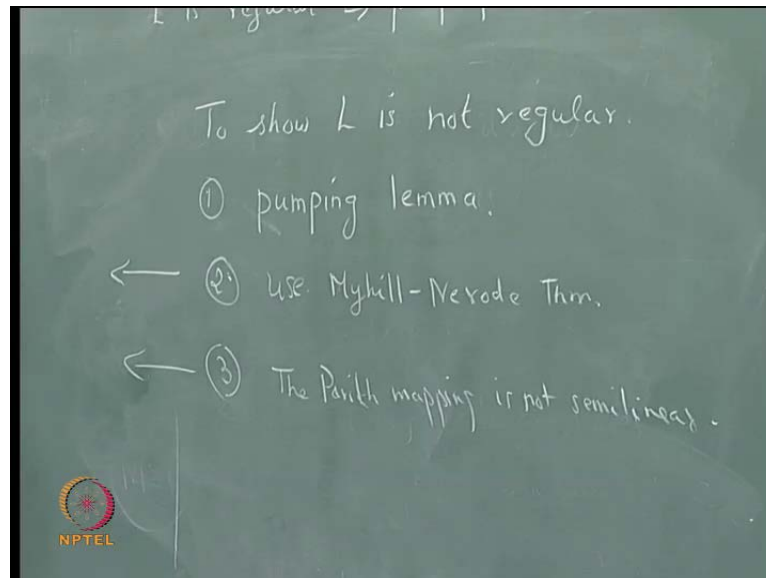
Then L will be a regular set something like that with this condition alone, if the converse is not true, moreover you must realise that. What pumping lemma says is? if you have a string, whose length is sufficiently large. Then Z can be written in the form $u v^i w$ such that $u v^i w$ belongs to L. You can keep on incrementing i, i can be one two three, it can go to infinity.

So, you will get an infinite number of strings beginning infinite number of strings which belong to L. It does it mean that if you have sufficiently large string belonging to L. You can write it in the form $u v^i w$ for large i. That is not true, what pumping lemma says is if you have a fairly large string, then you can write it in this form. And you can get infinite number of strings of this form it does not mean that if you have a very large string you can write it in this form $u v^i w$ for large i.

One example is simple example is Σ^* , where Σ is a b. You know that there are large strings which are cubed. We have seen that there are large strings which are cube, that is no sub string will occur three times consecutively. So, this is what it means you can

make use of this to show that other things like that $w^c w^r$ is not regular a power p p is a prime not regular a power two power n greater than or equal to zero not regular and soon.

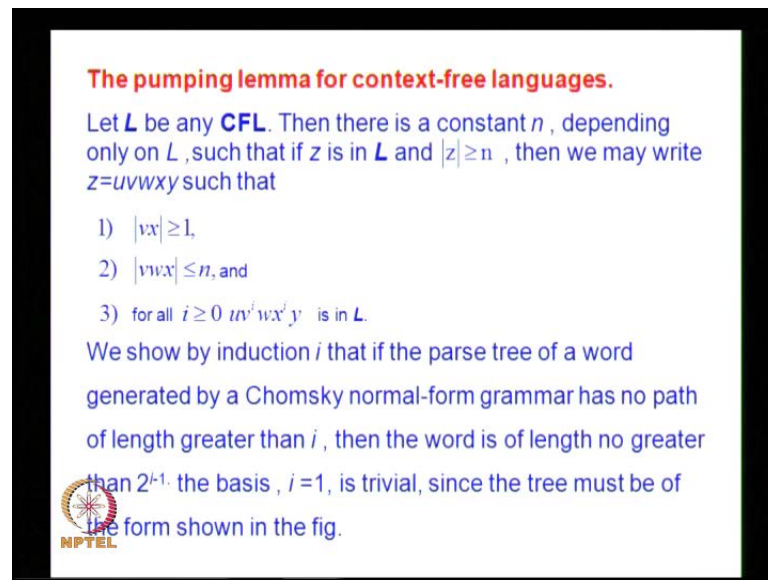
(Refer Slide Time: 35:43)



Actually, it is important to show that certain sets are not regular and there are three ways of doing that. One is use pumping lemma to show something in that to show L is not regular. You can use three methods, one is use pumping lemma. The next method is use myhill-nerode theorem, this we shall be considering next. Third the parikh mapping is not semilinear; this is another way you can show. These two things we shall consider later.

Now, we shall consider the pumping lemma for context free languages, for regular sets what we found is? If the string is sufficiently long then a portion of it can be pumped any number of times. For context free languages, what we defined is? Given a context free language, if you have a string which is sufficiently long, then two portions of the string can be pumped equal number of times. So, that we get an infinite number of strings.

(Refer Slide Time: 37:16)




The pumping lemma for context-free languages.

Let L be any CFL. Then there is a constant n , depending only on L , such that if z is in L and $|z| \geq n$, then we may write $z = uvwxy$ such that

- 1) $|vx| \geq 1$,
- 2) $|vwx| \leq n$, and
- 3) for all $i \geq 0$ uv^iwx^iy is in L .

We show by induction i that if the parse tree of a word generated by a Chomsky normal-form grammar has no path of length greater than i , then the word is of length no greater than 2^{i-1} . The basis, $i = 1$, is trivial, since the tree must be of the form shown in the fig.

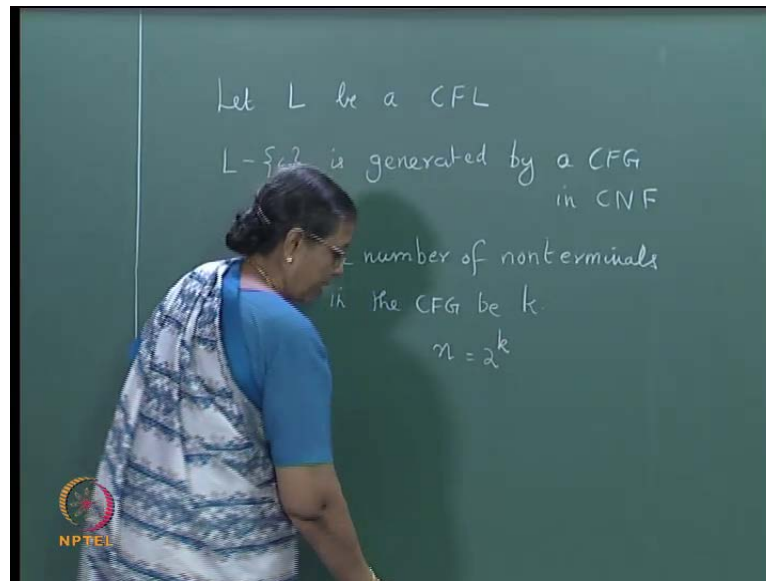


So, the statement can be like this, let L be any contextfree language. Then there is a constant n depending only on L such that if Z is in L and the length of Z is greater or equal to n . Then we may write Z as $u v w x y$ such that length of $v x$ is greater than or equal to one. What it means is both v and x cannot be simultaneously epsilon and second condition is length of $v w x$ is less than or equal to n . And for all i greater than or equal to zero $u v$ power $i w x$ power $i y$ is in L that is both the portions v and x can be simultaneously find equal number of times.

Now, you can see that, if you have a sufficiently long string in a language the corresponding parse tree should be, should have a path which is sufficiently long. In fact, we can show this we can show by induction on i that is the parse tree of a word generated by a chomsky normal form grammar has no path of length greater than i . Then the word is of length no greater than two i minus 1. So, in order to prove the pumping lemma first we shall take the contextfree grammar to be in chomsky normal form.

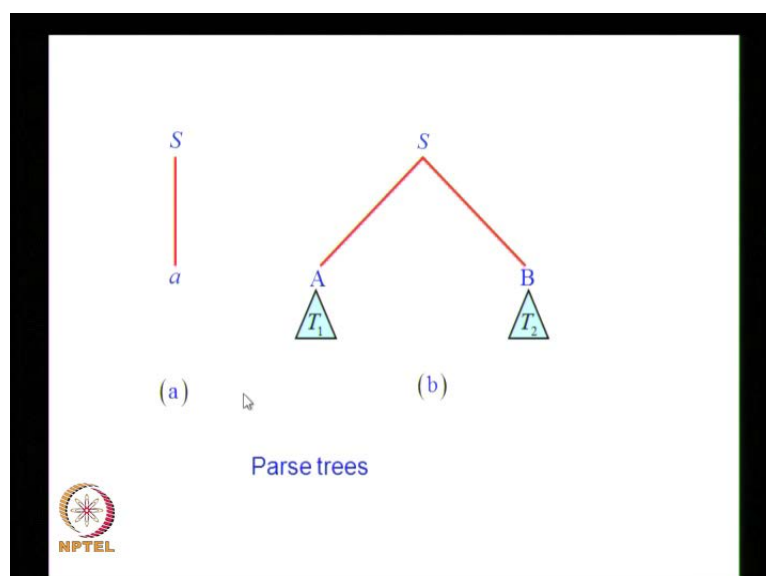
(No audio from 38:57 to 39:21)

(Refer Slide Time: 38:57)



Let the contextfree language minus epsilon, we are not very much bother about epsilon, because we are going to consider only strings of length greater than or equal to n for alarge n . So, let L minus epsilon be generated by a contextfree language in chomsky normal form. Let the number of non terminals in the c f g be k , and then we take n to be two power k . Now in this grammar we can show that if the parse tree of a word generated by the chomsky normal form grammar has no path of length greater than i , then the word is of length no greater than two power i minus 1, for the basis i one is trivial since the tree must be of the form shown in the figure.

(Refer Slide Time: 40:32)



If the path length is one, you have a situation like this the length of the string generated is one you see that i is 1. So, $2^1 - 1$ is 2^0 which is one. So, the base case holds.

Now, for an induction assume that the result is true up to path length $i - 1$ and show that the result is true for path length i . Now you consider this tree with path length maximum i , then the first step of the tree will be like this. And from this non-terminal there will be a subtree here. The maximum path length in t_1 or t_2 will be $i - 1$. So, the maximum length of the string generated in this portion will be $2^{i-1} - 1$. Similarly the maximum length of the string generated here is $2^{i-1} - 1$. So, the total length of the string generated by this tree will be $2^{i-1} - 1 + 2^{i-1} - 1$ which is $2^i - 2$. So, this proves the simple result.


(Refer Slide Time: 41:45)

Let G have k variables and let $n = 2^k$. If z is in $L(G)$ and $|z| \geq n$ then since $|z| > 2^{k-1}$ any parse tree for z must have a path of length at least $k+1$. But such a path has at least $k+2$ vertices, all but the last of which are labeled by variables. Thus there must be some variable that appears twice on the path.

Some variable must appear twice near the bottom of the path. In particular, let P be a path that is as long or longer than any path in the tree. Then there must be two vertices v_1 and v_2 on the path satisfying the following conditions.

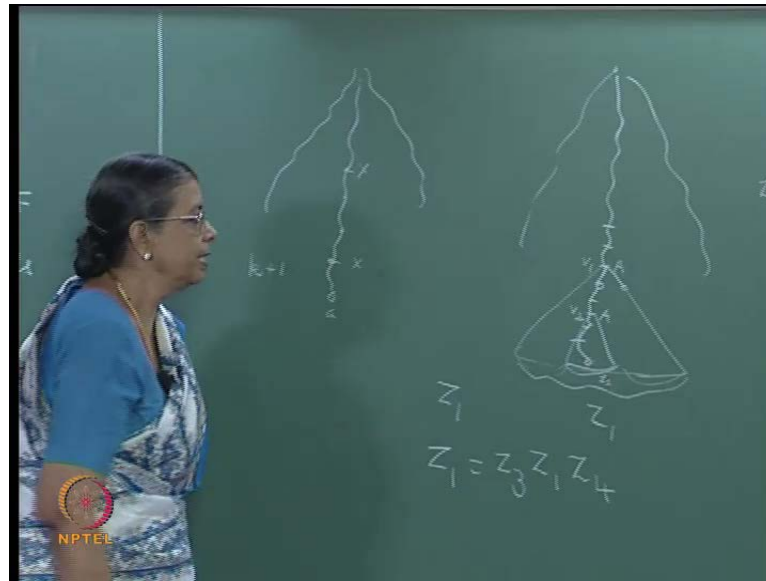
- 1) The vertices v_1 and v_2 both have the same label, say A .
- 2) Vertex v_1 is closer to the root than vertex v_2 .

3) The portion of the path from v_1 to the leaf is of length at most $k+1$.

 NPTEL

Now, let G have k variables or k non-terminals and we take n to be 2^k if Z is in $L(G)$ and the length of Z is greater than or equal to n . Then because the length of the string is greater than or equal to $2^k - 1$ any parse tree for Z must have a path of length at least $k + 1$. If it is less than that then the string will not be of this length, but such a path will have at least $k + 2$ vertices. And the last one will be a leaf rest of them will be non-terminals and there must be some variable that will appear in the path.

(Refer Slide Time: 42:36)



So, let us consider like this, the tree has a path of length $k+1$ or more. Now, let this be the longest path there may be one or two such paths, let us take one of them. You find that the last one will be a terminal symbol rest of the nodes in this path will be non-terminals. And there are only k non-terminals and there are $k+1$ nodes with labels which are non-terminals. By the pigeonhole principle at least one non-terminal will be repeated that is somewhere here, somewhere here, the same non-terminal will be repeated as the label of the internal node.

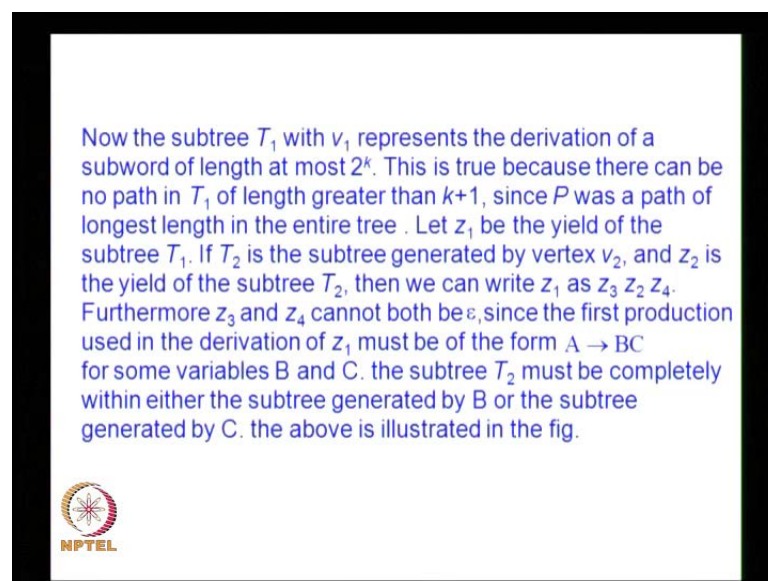
In fact, you can say something more than that; some variable must appear twice near the bottom of the path. This you can say, in fact, let p be a path that is as long or longer than any path in the tree. Then there must be two vertices v_1 and v_2 on the path satisfying the following conditions: the vertices v_1 and v_2 . The vertices v_1 and v_2 both have the same label a , the vertex v_1 is closer to the root than vertex v_2 . The portion of the path from v_1 to the leaf is of length at most $k+1$. How can you get this?

So, in the derivation tree path Z take the longest path, if there are more than one take one of them start from the leaf go up the path up to a path length of $k+1$. You consider a path length of $k+1$ here. In that there will be $k+1$ non-terminal nodes that is nodes with labels as non-terminals. The last one will be a leaf among this $k+1$ by the pigeonhole principle two of them will have the same label k , say call this as v_1 , call this as v_2 .

So, v_1 is nearer to the root than v_2 and both of them have the same label. The path from v_1 to the leaf is at most $k + 1$, in the worst case this can be a . Otherwise somewhere a will be here. So, the path length from v_1 to the leaf will be at most $k + 1$, now if you consider this result of this tree, the string derived let it be Z_1 . This one is Z_1 and from this another tree will be subtree will be derived, call this as Z_2 . So, you can write Z_1 as $Z_3 Z_2 Z_4$, this portion is Z_3 , this portion is Z_4 .

You can see that both Z_3 and Z_4 cannot be ϵ ; simultaneously one can be ϵ possible. But both of them cannot be ϵ , the reason is from here, you use a Chomsky normal form grammar. So, it can be like this and this subtree will entirely lie within one of them. Either in this case, it lies on the left of its left subtree, sometimes it may lie on the right subtree, but whatever it is the other tree will be non null, and so, you have this feature.

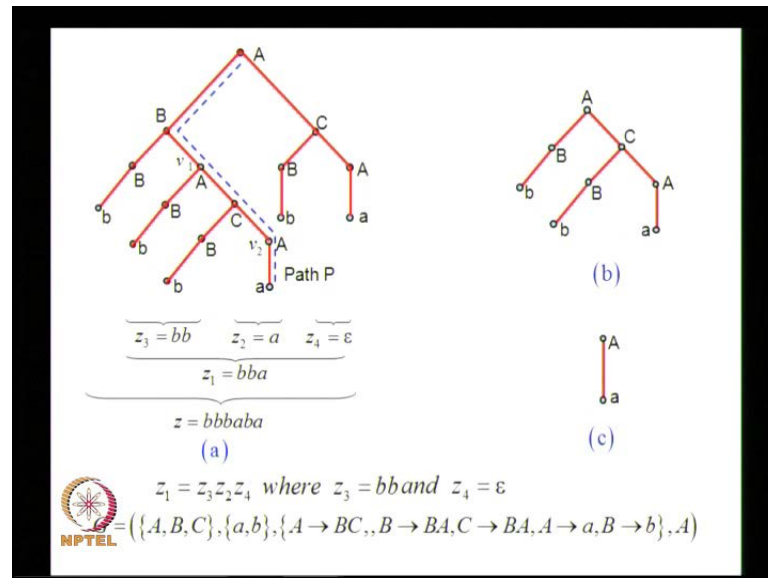
(Refer Slide Time: 46:50)



Thus you find that a tree T_1 with v_1 represents the derivation of a sub word of length at most two power k . And this is true, because there can be no path in T_1 of length greater than $k + 1$. So, in the figure what we saw is the length of the string derived from the first a , that is v_1 is at most two power k , that is why we have the first condition in the theorem or may be the second condition in the theorem length of $v w x$ is less than or equal to n , and p has the path of longest length in the entire tree.

Let Z_1 be the yield of the sub tree t_1 and t_2 is the sub tree generated with a vertex v_2 . And the string generated by that is Z_2 . So, you can write Z_1 as $Z_3 Z_2 Z_4$ both Z_3 and Z_4 cannot be epsilon. Since the first production used in the derivation must be of the form a goes to $b c$ for some variables b and c . And the sub tree t_2 must be completely within either the sub tree generated by b or the sub tree generated by c .

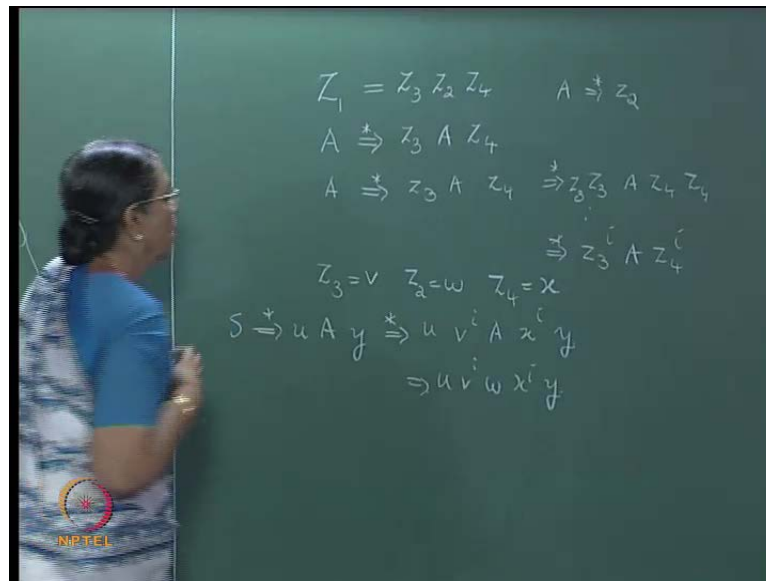
(Refer Slide Time: 48:13)



So, we have this result as an example you can consider this, you consider this example where the grammar is given by three non terminals, it is in Chomsky normal form rules are a goes to $B C$, B goes to $B A$, C goes to $B A$ and soon. a goes to a , b goes to b , you see that you have generation for this tree.

Now, if you consider a path, what is k here? k is three. So, consider a path of length four, start from this one two three four in this. You find that the non terminal a is repeated here. So, v_1 is nearer to the root than v_2 and $v_1 v_2$ have the same label the result of the sub tree with this as the root. You look at this, then this generates Z_1 bba is Z_1 what is Z_2 generated by this sub tree? That is just this a . So, is that one is $b b a$ Z_2 is a Z_3 is $b b$ this portion and Z_4 is epsilon because this entirely falls on one side.

(Refer Slide Time: 50:51)

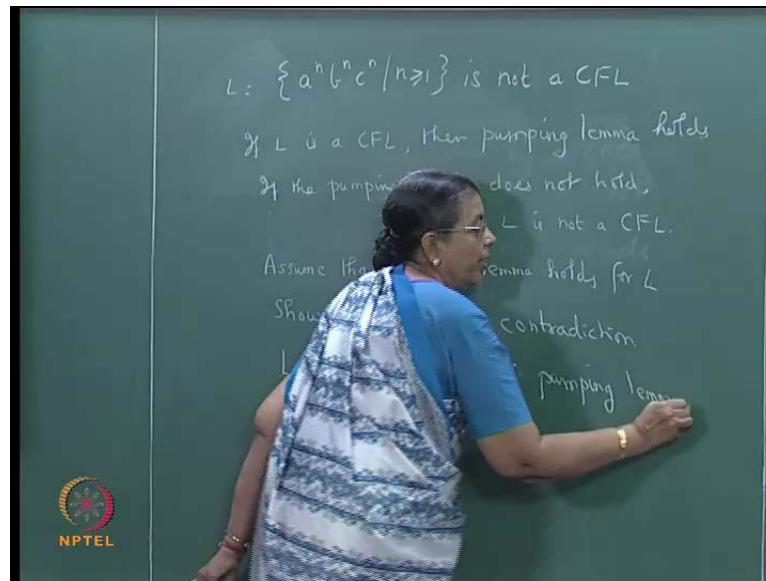


So, what you have is? Z_1 can be written in the form $Z_3 Z_2 Z_4$ where Z_3 is b^i and Z_4 is ϵ . So, what you find is from the non terminal A you are deriving Z_3 a Z_4 . Now this portion you can repeat any number of times actually if you repeat this any number of times Z_3 a Z_4 Z_3 a Z_3 Z_4 Z_4 and soon. You can pump Z_3 any number of times and the same number of times Z_4 also will be pumped.

Now, if in the theorem we take Z_3 to be v , Z_2 to be w , Z_4 to be x . Starting from S you will have a derivation $u A y$, if you look into this tree, some portion will be generated here, that is u . And some portion will be generated here, that is y . We are bothered about this portion. And here this can be these two portions can be pumped. So, from this you can generate Z_3 and Z_4 any number of times simultaneously. That is v can be generated any number of times, x can be generated any number of times. And then finally, from A you will derive Z_2 a also derives Z_2 . So, and Z_2 is nothing, but w . So, $u v^i w x^i y$. So, we get the pumping lemma for context free languages.

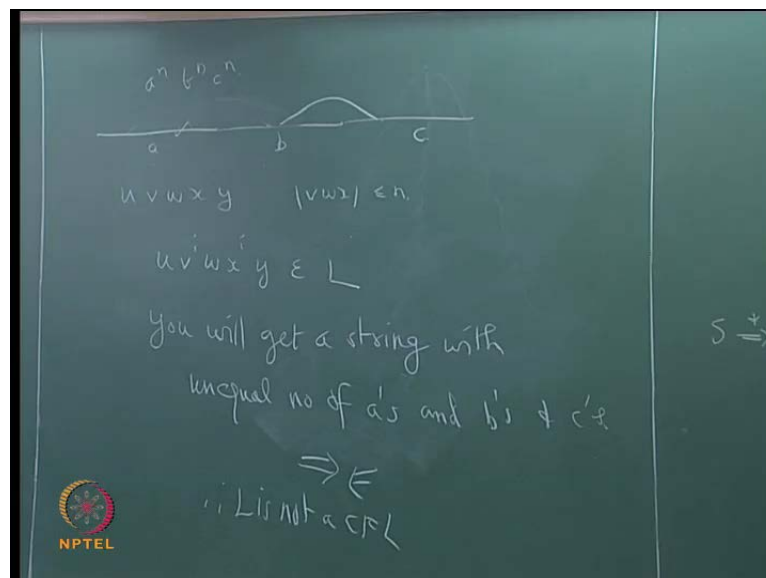
Let us see, how the pump occurs? You see that in this example which we consider L here, see look at this example, this is Z_1 . So, on the left you are getting a portion of the string which is b^i on the right you are getting a string ϵ .

(Refer Slide Time: 54:13)



For example, we can show that a power n b power n c power n greater than or equal to one is not a CFL, it is a context sensitive language. How can we show that this is not a CFL? Actually what we are going to prove is if L is a CFL then pumping lemma holds. The contra positive of it will be, if the pumping lemma does not hold L is not a CFL. So, in order to show that, assume that pumping lemma holds for L , this L then show you arrive at a contradiction.

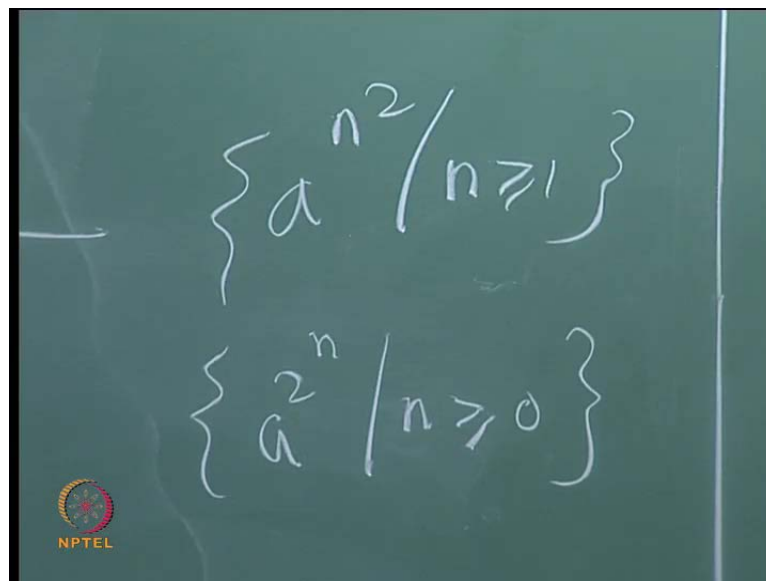
(Refer Slide Time: 56:27)



Let n be the constant of the pumping lemma, then consider the string $a^n b^n c^n$. So, you have a string of a 's equal number of b 's equal number of c 's. Now, this can be written as uv^iwx^iy . Where, the length of vwx is less than or equal to n . So, this vwx portion can occur within a or b or within c or like this. It cannot occur like this, it cannot include both a and c , it can include a and b or it can include b and c . It can include just one symbol that is also possible, but it cannot include a and c .

So, when you pump uv^iwx^iy will belong to L and you will find that when you pump. If the portion is like this, vwx portion is like this, the number of a 's will increase number of b 's will increase. But the number of c 's will remain the same if the portion is like this. Number of b 's and c 's will increase, but the number of a 's will remain the same. So, you will get a string which has unequal number of a 's b 's and c 's. You will get a string with unequal number of a 's and b 's and c 's which is a contradiction. Therefore, L is not a CFL. Thus we can use the pumping lemma for context free languages to show that certain languages are not context free.

(Refer Slide Time: 58:37)



We can use it for languages like a^{n^2} n greater than or equal to one or a^{2^n} n greater than or equal to zero, and soon. Thus you have seen the pumping lemma for regular sets, and also the pumping lemma for context free languages.

One thing you have to note is we use the automaton, the finite state automaton for regular sets. And, we use the context free grammars for proving the pumping lemma

for context free languages. Both are similar in the sense that you can pump a portion, if it is a regular set, you can pump two portions simultaneously for a context free language.