


Higher Engineering mathematics.
Professor P. N. Agrawal.
Department of mathematics.
Indian Institute of Technology, Roorkee.
Lecture 7: Language and Grammars-I.

Hello friends, welcome to my lecture on Languages and Grammars. We will have 3 lectures on this topic, so let us begin with the 1st lecture on this topic. See, in our daily life we come across 2 types of languages, one is the spoken language and then other one is the formal language. The spoken language is the one which you see in villages or in cities, that language is, I mean does not follow any rule. So, it is very complicated, it does not have any grammar or a set of rules which it follows. So, the other language which is the formal language, that language follows certain set of rules, it is scientifically defined, so that language is the one which we are going to discuss.

(Refer Slide Time: 1:18)



A language is called formal if it is provided with a mathematically rigorous representation of its alphabet of symbols, and formation rules specifying which strings of symbols count as well formed. Formal languages are studied in linguistics and computer science.
Formal language is a set of words, i.e. finite strings of symbols taken from the alphabet over which the language is defined.

A language is called formal if it is provided with a mathematically rigorous representation of its alphabet of symbols and formation rules specifying which strings of symbols count as well formed. Formal languages are studied in linguistics and computer science. Formal language is set of words that is finite strings of symbols taken from the alphabet over which the language is defined.

(Refer Slide Time: 1:52)

A grammar is a set of rules to define a valid sentence in any language. It does not describe the meaning of the sentences or what can be done with them in whatever context but only their form.



A grammar is a set of rules to define a valid sentence in any language. It does not describe the meaning of the sentences or what can be done with them in whatever context but only their form.

(Refer Slide Time: 2:08)

Alphabet: It is defined as a finite non-empty set of elements or symbols.
Example $\Sigma = \{0, 1\}$, $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $\Sigma = \{A, B, C, \dots, Z\}$ and $\Sigma = \{a, b, c, \dots, z\}$ where Σ is a representation for the alphabet.



Now, let us define an alphabet. An alphabet is defined as a finite non-empty set of elements or symbols, for example, $\Sigma = \{0, 1\}$ or Σ the set containing 0, 1, 2, 3 and so on up to 9 numbers and $\Sigma = a, b, c$ and so on up to Z, the upper case letters of English alphabet and Σ can also be taken as the set consisting of small case letters of the English alphabet a, b, c and so on up to z. Σ is the representation for the alphabet.

(Refer Slide Time: 2:43)

String A string or word on an alphabet Σ is defined as a finite sequence of symbols in Σ .

Example If $\Sigma = \{a, b\}$ then $u = abab$ and $v = aaabb$ are strings on Σ .



Now, let us define a string, a string or word, a string or word on an alphabet Σ is defined as a finite sequence of symbols in Σ . Say, for example you take Σ to be the set consisting of the letters a and b, then you can take the string or the word u as abab, it is consisting of the letters a and b, which are there in the set Σ . Okay. You can also take another example where we are taking aaabb, so it is a word or a string which is defined over Σ . So, u and v are both a finite sequence of symbols in Σ .

(Refer Slide Time: 3:31)

Definition

a^i is used to denote a sequence of i , a 's i.e. a^2 for aa , a^3 for aaa . Thus $\tilde{u} = aaabb = a^3b^2$. The length of a string u , denoted by $\|u\|$, is the number of symbols in the string, for example if $u = abab$ and $v = a^3b^2$ then $\|u\| = 4$ and $\|v\| = 5$.

The empty sequence of letters, denoted by λ or ϵ is also considered to be a string on Σ , called the empty string and $\|\lambda\| = 0$.



Now a^i , okay, a^i to the power i is used to denote a sequence of i a's, that is a^2 , we can write for a, a^3 we can write for aaa. So, if there are I a's, then I a's can be written as a^i . Now, if you have $u = aaabb$, then we can write it as $a^3 b^2$. Okay, the length of a string u , the length of a string u is denoted by this symbol, okay. $\|u\|$, you can call it is the number of, it is the number of symbols in the string.

So, for example in u there are 5 symbols $aaabb$, there are 5 symbols, so $\|u\| = 5$ here. Okay, now here we are taking u to be $abab$, so $\|u\| = 4$, there are 4 symbols $abab$, okay. And in v we have $a^3 b^2$, that is a is occurring 3 times aaa and b is occurring 2 times bb . So, $\|v\|$ will be $= 5$. Okay, The empty sequence of letters is denoted by λ or it is also denoted by ϵ , it is considered to be a string of Σ and we call it as empty string, okay, the $\|\lambda\|$ is defined as 0. So the empty sequence of letters is denoted by λ or ϵ and its norm is defined, length is defined as to be $= 0$.

(Refer Slide Time: 5:11)

Power of an alphabet:

If Σ is an alphabet, then Σ^k is defined as the set of strings of length k , each of whose symbols is in Σ . For example, if $\Sigma = \{0, 1\}$, then
 $\Sigma^1 = \{0, 1\}$, $\Sigma^2 = \{00, 01, 10, 11\}$, $\Sigma^3 = \{000, 001, 010, 011, 110, 101, 100, 111\}$
Kleene Star(Closure): Let Σ be an alphabet. The Kleene star Σ^* is defined as the set of all strings (including λ , the empty string) over the alphabet Σ . If $\Sigma = \{0, 1\}$, then
 $\Sigma^0 = \{\lambda\}$, $\Sigma^1 = \{0, 1\}$, $\Sigma^2 = \{00, 01, 10, 11\}$
 $\Sigma^3 = \Sigma^3 = \{000, 001, 010, 011, 110, 101, 100, 111\}$
 $\Sigma^* = \{\lambda, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 110, 101, 100, 111, \dots\}$
 $= \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \Sigma^4 \dots$
 $= \bigcup_{i=0}^{\infty} \Sigma^i$

7

Now, if Σ is an alphabet, then Σ^k is defined as the set of strings of length K , each of whose symbols is in Σ . So, if $\Sigma = \{0, 1\}$, then $\Sigma^1 = \{0, 1\}$, Σ^2 will be consisting of the strings of length 2 made up from the elements of Σ , so 01 are the elements of Σ , so we will have $00, 01, 10, 11$.

In Σ^3 we will have strings of length 3, okay, made up from the elements 0 and 1 so we can have $000, 001, 010, 011, 110, 101, 100, 111$, you can see each of them has length 3, okay, so they are elements of Σ^3 . Now, let us define closure, we also call it Kleene $*$. So, let Σ be an alphabet, then the Kleene $*$ Σ^* is defined as the set of all strings including λ , the empty

string. So, λ is also there in Σ^* . So, Σ^* is defined as the set of all the strings including λ , the empty string over the alphabet Σ .

So, if $\Sigma =$ the set consisting of 0 and 1, then $\Sigma^0, \Sigma^1 =$ the set consisting of λ only, λ is the empty string, $\Sigma^1 = \Sigma$, which is the set consisting of 0, 1. Σ^2 consists of elements, each one of length 2 and made up from the elements of Σ , so 00, 01, 10, 11. Σ^3 is the set made up from the elements 0 and 1 having length 3 each, so 000, 001, 010, 011, 110, 101, 100, 111, okay. Σ^* , now $\Sigma^* = \bigcup_{i=0}^{\infty} \Sigma^i$ where i varies from 0 to ∞ . $\Sigma^0 \Sigma^1 \Sigma^2$ and so on, okay.

So, when you consider the \cup of these sets $\Sigma^0, \Sigma^1, \Sigma^2, \Sigma^3, \Sigma^4$ and so on, you get Σ^* , so Σ^* is consisting of λ , okay, λ 01, then 00, then 01, 10, 11 and then the elements containing strings of length 3 each, okay, so 000, 001, 010, 011, then 110 and so on. So, that is Σ^* .

(Refer Slide Time: 8:05)

The set of non-empty strings from alphabet Σ is denoted by Σ^+ . Hence, $\Sigma^+ = \Sigma^* - \{\lambda\}$. $\Sigma^+ = \Sigma^*$



Operations on strings: Let u and v be two strings in Σ^* . The concatenation of u and v written as uv , is the string obtained by writing down the letters of v to the right end of u , for example, if $u = abab$ and $v = aaabb$ then $uv = ababaaabb$. $uv = ababaaabb$

Now, let us define Σ^{+} . Σ^{+} is defined as a set of all nonempty strings, okay. So, that means we have to exclude empty string from Σ^* to arrive at Σ^{+} . So, $\Sigma^{+} = \Sigma^* - \{\lambda\}$, λ we have seen, λ is $\Sigma_0, \Sigma_0 = \lambda$. So, we removed the empty string λ from the set Σ^* to get the set of nonempty strings that is Σ^{+} . Now, let us define operations on string. So, let u and v be strings in Σ^* , then concatenation of u and v written as uv is the string obtained by writing down the letters of v to the right end of u .

For example, $u = abab$ and $v = aabb$, then $uv = ababaabb$ that is u , then we write the string of v to the right end of u , okay, so, v is $aabb$, okay, so $uv = ababaabb$. Okay. So, that is how we define the concatenation of the 2 strings u and v .

(Refer Slide Time: 9:36)

Properties of Concatenation:
 If u is a string, then $u^n = \underbrace{uu\dots u}_{n \text{ times}}$. As a special case $u^0 = \lambda$.
Associative law: For each u, v, w in Σ
 $u(vw) = (uv)w$
Identity element: The empty string λ is an identity element for concatenation i.e.
 $\lambda u = u\lambda = u$, for all u in Σ .
Right cancellation: or u, v, w in Σ
 $uw = vw \rightarrow u = v$.
Left cancellation: For u, v, w in Σ
 $wu = wv \rightarrow u = v$.
 For u, v in Σ , $\|uv\| = \|u\| + \|v\|$,
 i.e the length of concatenation of the strings is the sum of the individual lengths.



9

Now, properties, there are certain properties which concatenation follows. So, the concatenation follows certain properties, the properties are. If u is a string, then $u^n = \underbrace{uu\dots u}_n$ number of times, okay. And as a special case we can say that $u = \lambda$. Now, it also follows the associative law, if $u, v, w \in \Sigma$, then concatenation of V and W , when we consider its concatenation with U is same as concatenation of U and V with concatenation of W .

So the identity element is also there, the empty string λ is an identity element for concatenation because $\lambda U = U \lambda = U$ for all U in Σ . Now, it follows right cancellation law, whenever $U, V, W \in \Sigma$, then $WU = WV \implies U = V$. And it also follows left cancellation, whenever $U, V, W \in \Sigma$, then $UW = VW \implies U = V$.

And if U and $V \in \Sigma$, $\|U+V\| = \|U\| + \|V\|$. The length of concatenation of these strings, you can define this equality, it says that the length of concatenation of the strings = the sum of individual lengths. So, whenever we consider the length of the concatenation of UV , what we get is, it is equal to the sum of the lengths of the individual strings U and V .

(Refer Slide Time: 11:27)

$w = \lambda \lambda w$
 $w = \lambda a b c d$

$U_1 = \lambda$
 $U_2 = b c d$

$w = \lambda \lambda w$
 $w = \lambda a b c d$
 $w = \lambda a b c d$
 $w = \lambda a b c d$

$w = \lambda a b c d \lambda$

Substrings or subwords: A word u is called a substring or subword of w if $w = v_1 u v_2$. If $v_1 = \lambda$, then u is an initial segment. Note that λ is a subword of w since $w = \lambda \lambda w$. Also, w is a subword of w because $w = \lambda w \lambda$. For example, let $w = a b c d$ then the subwords of w are $\lambda, a, b, c, d, a b, b c, c d, a b c, b c d, a b c d$. The initial segments are $\lambda, a, a b, a b c, a b c d$. It is to be noted that $a c$ or $a c d$ are not subwords of w although their letters all belong to w .

$w = \lambda a b c d$ $U_1 = \lambda$ $U_2 = b c d$
 $w = \lambda a b c d$ $U_1 = \lambda$ $U_2 = c d$

$U_1 = \lambda, U_2 = w$
 - Hence $w = U_1 U_2 = \lambda \lambda w$
 $w = \lambda w \lambda$

$w = a b c d$
 $U_1 = a$
 $U_2 = b c d$

IIT ROORKEE
NPTL ONLINE
CERTIFICATION COURSE

10

Now, substrings or subwords. A word W is called a substrings or subword of W if we can find V_1 and $V_2 : W = V_1 U V_2$. If $V_1 = \lambda$, in particular if $V_1 = \lambda$, then U is defined as an initial segment. Okay, note that λ is a subword of W , why, because $W = \lambda W$, okay. So, λ is always a subword of W because we can find V_1 and V_2 , when we can take $V_1 = \lambda, V_2 = W$, such that $W = V_1 \lambda V_2$, where $V_1 = \lambda$, okay, let us take $V_1 = \lambda$ and $V_2 = W$.

Then W we can write as $V_1 U V_2$, okay. So, this equal $V_1 = \lambda$, U is we are taking as V , $U V$ are taking as λ and then $V_2 = W$. So, $W = \lambda W \lambda$, so λ is always a subword of each word W . Now, also W is a subword of W . Why, because they can take V_1 and V_2 part to be $= \lambda$, okay. $W =$, it can be written as $\lambda W \lambda$, so λ is always a subword of W .

Let us take $W = a b c d$, and let us see what are subwords of W , okay. So, 1st subword is λ because we have seen λ is always a subword of W , okay. Then a is subword of W , how a is this subword of W ? See, you can take $V_1 = \lambda$ I read the blue $= \lambda a b c d$, okay, so I am taking $V_1 = \lambda$ and $V_2 = b c d$, okay. So, then I am able to write $\lambda =, W = \lambda a b c d$. So, here V_1 is λ and V_2 is $b c d$.

So, I am able to write $W = V_1 U V_2$. So, a is a subword of W , similarly you can write b is a subword of W , and c is a subword of W , T is a subword of W . Now, $a c$, how is a subword of W ? So, $W =$, I write $\lambda a b$ and then I take $c d$ as $c d$, I take $V_2 = c d$, okay. So, here I am taking $V_1 = \lambda$ and $V_2 = c d$, okay. So, $a b$ is a subword of W , similarly $b c$ is a subword of W , we can write $W = a b c d$, where I am taking $V_1 = n$ and $V_2 = d$. So, I am able to write $W = \lambda, W = V_1 U V_2$, okay.

V_1 is a, V_2 is d. So, bc is a subword of W. Similarly cd is also a subword of W. Now, abc, abc is a subword of W, why, because I can write $W = \lambda abc$ and then d, okay. So, I am taking $V_1 = \lambda$ and $V_2 = d$, okay. So, abc is a subword of W, similarly bcd is a subword of W, here I will take $V_1 = a$ and $V_2 = \lambda$, okay. And abc is a subword of W because every word W is a subword of itself, so abcd is a subword of W.

Now, initial segments, let us see what are the initial segments here. Initial segment is λ , λ is always an initial segment, why, because I can write, see, let us look at the definition of initial segment. $W = V_1 UV_2$, okay, then we say that U is a subword of W. If $V_1 = \lambda$, okay, $V_1 = \lambda$, the U is an initial segment of W. So, I can say that λ is an initial segment of W because I can write $W = \lambda \lambda W$, okay, I can write $W = \dots$. So, here this V_1 is λ , U_2 is λ and V_2 I am taking as W.

So V_1 , when V_1 is λ , then this U is called if the initial segment, so λ will be an initial segment of W. Now, similarly a b, okay, how a b are initial segments, W can be written as λa and then V_2 I will take as bcd, okay. So, take $V_1 = \lambda$ and $V_2 = bcd$, then a will be an initial segment of W, okay. Now, let us look at the initial segments, okay, of the word W. So, λ is an initial segment of W because I can write $W = \lambda \lambda W$.

So, here V_1 is λ , okay. So, if V_1 is λ , then this is U, this is U, it is called initial segment of W. So, λ will be called as initial segment of W. , Now a will be initial segment of W because I can write $W = \lambda a$ and then bcd, okay. V_2 I will take as bcd and V_1 is λ , so I am able to write W as $V_1 \cup V_2$, okay. So, a is an initial segment of W. Okay, then a b is an initial segment of W. I can write $W = \lambda a b$ and then cd, so I will take $V_1 = \lambda$ and $V_1 = cd$.

And similarly abcd will be initial segment of λ , I will take $W = \lambda abc$ and then I will take $V_2 = d$, okay. And abc, abcd is always initial segment of W because then I will write $W = \lambda abcd$. So, I will take $V_1 = \lambda$ and $V_2 = \lambda$. So, the initial segments of W are λa , a b, abc, abcd. So, ac or a cd, okay, so let us look at a c or acd, ac or a cd are not subwords of W, although their letters $\in W$, you can see, their letters abcd, they $\in W$.

But, see, W is abcd, okay, W is abcd, so these letters ac, acd $\in W$ but ac is not a subword of W, why it is not a subword of W, because it will be called, it will be called subword of W, only when you have to, you can find V_1 and V_2 such that $W = V_1 \cup V_2$.

But $W = abcd$, so if I want to check whether ac is a subword of W or not, then I need to find V_1 and V_2 , such that $W = V_1 U V_2$, okay. So, ac , this is my U , okay, now I cannot find V_1 and V_2 such that I am able to get $abcd$. Okay, because b lies between a and c , in the word W , we have $abcd$, so b lies between a and c , so b cannot be inserted here, between a and c . V_1 is to the left of letter a and V_2 is to the right of the letter c . So, we can never get $abcd$. We will never get any V_1 and V_2 such that $W = V_1 ac V_2$, so ac is not a subword of W . Similarly, acd is also not a subword of W , we cannot find V_1, V_2 such that $W = V_1 acd V_2$.

(Refer Slide Time: 20:07)

Prefix and Suffix:

Let x and y be two words over an alphabet Σ . Then x is said to be equal to y , written $x = y$, if they have the same length and the same symbols at the same positions. x is called a prefix of y if there exists a word z over Σ such that $xz = y$. For example, if $x = in$ and $y = indian$, then x is a prefix of y , since $xz = y$, where $z = dian$. On the other hand, if $x = can$ and $y = scan$ then x is not a prefix of y and y is not a prefix of x .

$y = indian$
 $x = in$
 $z = dian$
 $xz = in \cdot dian = indian = y$

$y = scan$
 $x = can$
 $z = ?$
 $xz = can \cdot ? \neq scan = y$

Now, similarly let us consider prefix and suffix. Let X and Y be 2 words over and alphabets Σ , okay. So, let X, Y be 2 words over and alphabets Σ , then X is called $= Y$, written as $X = Y$, if X and Y both have the same length and they have same symbols at the same positions. X will be called a prefix of Y if I can find a word Z over Σ such that $XZ = Y$, okay. For example, if suppose $X = in$ and $Y = indian$, $Y = Indian$, okay.

Then this is my X , okay, then \exists the word Z , okay, : $Y = XZ$, so Z I will take as $dian$. So, in this case where X is in and Y is $indian$, okay, we are able to find Z which is $dian$, such that $Y = XZ$, so X is a prefix of Y . Now, if you look at this example, $X = can$, $Y = scan$, then x is not a prefix of Y because we are not able to find any Z such that $Y = X, Z$, okay, $Y = XZ$, $X scan$. Okay, so $X scan$ and Z is some word from the alphabets Σ , so Z cannot be found such that you get $Y scan$, okay.

Because S occurs to the left of can , so we will not be able to get any Z , such that $scan = can Z$, okay. There does not exist any Z , which can give you $can Z = scan$. So, X is not prefix of Y , similarly you can see Y is not prefix of X , if Y is to be a prefix of X , then X should be $= YZ$ for some Z , some word Z . Okay, and $X = can$ and $Y = scan$, okay, $scan$. So, you cannot get can , okay, you cannot get any Z such that $scan Z = can$. So, Y is also not a prefix of X .

(Refer Slide Time: 22:57)

x is called a proper prefix of y if $x \neq \lambda$, $x \neq y$, and x is a prefix of y . Since $x = \lambda x$, λ is always a prefix of any word x including self. Further $x = x\lambda$, hence x is always prefix of itself.

Similarly, we define suffix and proper suffix. For example, if $x = \text{happy}$ and $y = \text{unhappy}$ then x is a suffix of y and in fact a proper suffix of y .

$x = z\lambda$
 $\lambda = \lambda\lambda$
 $x = \lambda x$

$\lambda = \lambda\lambda$
 x is called a suffix of y if \exists a word z such that $y = zx$
 $y = zx$
proper suffix x is called a proper suffix of y if $z \neq \lambda$, $z \neq y$ and x is a suffix of y .

Now, X is called a proper prefix of Y , now defined proper prefix, proper prefix of Y if $X \neq \lambda$ and $X \neq Y$, also X is a prefix of Y . So, it will be called proper prefix of Y if $X \neq \lambda$, $X \neq Y$ and X is a prefix of Y . Now, since $X = \lambda X$, λ is always a prefix of any word, including itself because I can write $\lambda = \lambda \lambda$. So, λ is always a prefix of any word that follows from here and it is also a prefix of itself. Now, further $X = X \lambda$, so X is always a prefix of itself, okay, these are simple things. Now, let us define, similarly we can define suffix and proper suffix.

Okay, a word X will be called, X is called a suffix of Y if \exists a word Z , such that $Y = ZX$, then X will be called a suffix of Y , okay. And we can also define proper suffix, proper suffix means, okay, X is called a proper suffix of Y if $X \neq \lambda$, $X \neq Y$, okay, and X is a suffix of Y , X is a suffix of Y , okay. So, like we have seen here, that λ is always a prefix of any word X , we can also say that $X = X \lambda$, okay.

So, λ is a suffix of any word X and also $\lambda = \lambda \lambda$, okay. So, λ is a suffix of itself, okay. And also $X = \lambda X$, okay, $X = \lambda X$, so X is a suffix of itself. Now, here, we have an example, so let us say take $X = \text{happy}$, okay, $X = \text{happy}$ and $Y = \text{unhappy}$, okay, unhappy , then X is a suffix of Y or we can check that it is a proper suffix of Y because $X \neq \lambda$, $X \neq Y$ and X is a suffix of Y .

(Refer Slide Time: 26:04)

Reversal of a string: The reversal of a string is obtained by flipping over the last symbol. For example, if $u = xyz$ then $u^R = zyx$, where u^R is the reversal of u .

Palindrom: A palindrom over an alphabet Σ is a string that reads the same both forwards and the backwards. For example, $u = abba$ is a palindrom because $u = u^R$.

Example: If $S = \{aa, b\}$ then the Kleene star $S^* = \bigcup_{i=0}^{\infty} S^i$.
 $S^0 = \{\lambda\}$, $S^1 = \{b\}$, $S^2 = \{aa, bb\}$, $S^3 = \{aab, bbb, baa\}$, $S^4 = \{aaaa, aabb, bbba, bbbb\}$

Then $S^* = \{\lambda, b, aa, bb, aab, bbb, baa, aaaa, aabb, bbba, bbbb, \dots\}$

If $S = \{a, b, c\}$ then $S^0 = \{\lambda\}$, $S^1 = \{a, b, c\}$ hence

$S^* = \{\lambda, a, b, c, aa, ab, ac, bb, ba, bc, cc, ca, cb, \dots\}$

Example: If $\Sigma = \{0, 1\}$, $P = \{01, 011, 101\}$, $Q = \{1, 11\}$ then

$PQ = \{011, 0111, 0111, 01111, 1011, 10111\}$

and $QP = \{101, 1011, 1101, 1101, 11011, 11101\}$.

Now, let us define reversal of a string. The reversal of a string is obtained by flipping over the last symbol. Okay, for example if $U = XYZ$, then U^R , U^R means its reversal of U , U^R will be $= ZYX$, $U^R = ZYX$, U^R is the reversal of U . What is the palindrom, a palindrom over and alphabets Σ is a string that treats the same both forwards and backwards. For example, if $U = abba$, then it is a palindrom because U^R is also $= abba$, okay.

So, U is same as U^R , so it is a palindrom. Now, if $S = ab$, okay, in the set S , let us consider the set S , $S = ab$, then the Kleene $*$, Kleene $*$ we have earlier defined, so here the Σ set, Σ set is the set S , it is consisting of a and b , so $S^i = \bigcup_{i=0}^{\infty} S^i$. okay. Now, S^0 , S^0 is the set containing λ , S^1 is the set containing length, elements of length 1, okay, so we have b here. S^2 , because if you remember Σ^k , was the set of all strings of length k , okay, so here Σ is S , so S^1 means a string of length 1 and there are 2 elements here, one is b and the other one is a .

b has length 1, a has length 2, so only b will be there in S^1 . S^2 can have all elements of length 2, okay, so a will be there and from b we can have bb , so bb will be there. S^3 , okay, S^3 can have elements of length 3, strings of length 3, so we can form from a and b , we have to form strings of length 3. So, we can have aab , we can also have baa , this one and we can also have bbb , okay, nothing else is possible here, so that is S^3 . And S^4 , from S^4 , in S^4 we have to have strings of length 4, so from a we can write $aaaa$, it is a string of length 4.

And then from b , a and b , we can write $aabb$, okay, we can also write $bbba$, okay. So, $aaaa$, $aabb$, $bbba$ and we can then write $bbbb$, okay, b can be repeated 4 times, so $bbbb$, so these 4 elements will be there in S^4 . Now, $S^* = \bigcup_{i=0}^{\infty} S^i$, where $i = 0$ to ∞ , so S^i , so S^* is this much,

okay, this set. Now, if S is abc , the S^0 is λ , S^1 is the set of strings of length 1, so abc , all are there,

λ is then λabc , and then length S^2 , S^2 will contain elements made up from of length 2, made up from abc , so we will have $aa, ab, ac, bb, ba, bc, cc, ca, cb$, okay, this will be S^2 , and then similarly we can write S^3 .

So, $aa, ab, ac, bb, ba, bc, cc, ca, cb$, these are elements $\in S^2$. And then we will have elements of length 3 each made up from abc . Okay, now let us take $\Sigma = 01$, alphabet Σ we take as 01 , then let us also take $P =$ the set containing $01, 011, 101$ and Q as the set containing 11 . Then PQ , okay, this as PQ is defined as, now we write the elements of P , okay, into write-off P , we write the element from Q , okay. So, let us take the element from P to be 01 , okay, this one. Then PQ will give us to the right of 01 we write the elements from Q , so 1 , okay, this is one element.

Then, 01 with 11 , so 0111 , okay. So we get 2 elements, okay, when we pick up an element 01 from P and take all elements from Q and then take the 2nd element from P 011 and then join this to the elements of Q . So, we write the elements of Q to the right of 011 . So, we have $0111, 01111$ and then 101 , that is the 3rd element of P . So, we had joined the element 101 of P the elements of Q , so 101 then 1 and then 101 and then 11 , okay. So, these are the elements of PQ . And similarly we write the elements of Q .

We pick up an element of Q and then write to the right of the element of Q all the elements of P . So, when we have 1 here, Q , we take the element 1 of Q and then write all the elements of P to the right of 1 . So, 01 , that is one element, then 1011 , then 1101 , okay. Now we have adjusted all elements of P , now let us take the 2nd element of Q , 11 . Okay, and then we write to the right of 11 , all elements of P , so $1101, 11011$ and then 11101 , okay. So, these are the elements Of QP .

(Refer Slide Time: 32:19)


Language: Any collection L of words from an alphabet Σ , i.e., any subset of Σ^* , is called a language over Σ .

Sentence: A string in a language is called a sentence of L .

Example: If $\Sigma = \{a, b\}$ then
 $\Sigma^* = \{\lambda, a, b, aa, ab, bb, ba, aaa, aab, abb, aba, bbb, baa, bab, bba, \dots\}$

Then $L_1 = \{a, ab, ab^2, \dots\}$ i.e. L_1 consists of all words which begin with an a followed by one or more b 's.
 $L_2 = \{a^m b^n; m > 0, n > 0\}$ i.e. L_2 consists of all words beginning with one or more a 's followed by one or more b 's.
 $L_3 = \{a^m b^m; m > 0\}$ i.e. L_3 consists of all words beginning with one or more a 's followed by same number of b 's.
 $L_4 = \{b^m a b^n; m > 0, n > 0\}$ i.e. L_4 consists of all words with exactly one a which is neither the first nor the last letter of the word i.e. there is one or more b 's before and after a .

$\Sigma^* = \bigcup_{k=0}^{\infty} \Sigma^k$
 $\Sigma^0 = \{\lambda\}, \Sigma^1 = \{a, b\}$
 $\Sigma^2 = \{aa, ab, ba, bb\}$



Now language, any collection of L of words from an alphabet Σ , that is any subset of Σ^* is called a language over Σ . We have already defined, $\Sigma^* = \bigcup_{k=0}^{\infty} \Sigma^k$, $\Sigma^0 = \{\lambda\}$, $\Sigma^1 = \Sigma$, that is Σ , Σ^2 is, we have earlier seen aa, a, b, ba, bb , that is Σ^2 .

And then similarly we can write Σ^3 and so on. So, then \cup of all these sets, that is Σ^* , so $\lambda, ab, aa, ab, bb, ba$, then aaa, aab and so on that is Σ^* . Now, let us take L_1 to be this set a, ab, ab^2 , okay, then L_1 is a subset of Σ^* , you can see, $L_1 \subseteq \Sigma^*$, so L_1 defines a language. And you can describe this language in words, and one consists of all words which begin with an a followed by one or more b 's. So, we can describe this $\subseteq \Sigma^*$ which defines a language. This language consists of all words which begin with an a followed by one or more b 's.

And then another set you can take L_2, L_2 of Σ^* . It is containing elements a^m, b^n , where m is a positive integer, n is a positive integer. So, this language L_2 again describes a language, so L_2 consists of all words beginning with one or more a 's followed by one or more b 's. So, here a 's and b 's, number of a 's and b 's can be different, okay. So, that is the language L_2 . And then the language L_3 , it consists of a^m, b^n , where m is a positive integer. So, L_3 consists of all words beginning with one or more a 's followed by same number of b 's.

And L to the power 4, $L_4, L_4 = b^m a b^n$. So here a lies in the middle, okay and on the left of a we have b's, okay, one or more b's, on the right of a also be one or more b's. So, L_4 consists of all words with exactly one a which is neither the 1st, nor the last letter of the word, that is there is one or more b before and after a, okay.

(Refer Slide Time: 35:25)

Since the language is a set, all the operations like union, intersection and difference hold true for languages also.
 The complement of a language L , denoted by $\bar{L} = \Sigma^* - L$.
 The inverse of L , denoted by L^{-1} , is given by $L^{-1} = \{w^R; w \in L\}$
 The concatenation of L_1 and L_2 is given by
 $L_1 L_2 = \{uv; u \in L_1, v \in L_2\}$
 The union of L_1 and L_2 , denoted by $L_1 \cup L_2$, is given by
 $L_1 \cup L_2 = \{w; w \in L_1 \text{ or } w \in L_2\}$



Now, we have seen that languages are set, okay, language you set, so all the operations like \cup , \cap and difference hold true for languages. So, compliment of a language L , compliment of a language L is defined as $\bar{L} = \Sigma^* - L$, okay. And inverse of language L is defined as $L^{-1} = \{W^R, W \in L\}$ So, $\{W^R$ means reversal of W , so in L^{-1} , all the elements that $\in L$, we take their reversal, okay. So, those elements $\in L^{-1}$. Now we have concatenation of L_1 and L_2 . Okay, concatenation of the set L_1 and L_2 is given by $L_1 L_2$, it consists of the strings UV , where $u \in L_1$ and $v \in L_2$. The \cup of L_1 and L_2 denoted by $L_1 \cup L_2$ is the set containing W , where $W \in L_1$ or $W \in L_2$.

(Refer Slide Time: 36:28) 41:19

Now, let us define power of a language, okay, suppose L is the language, then L^0 consists of empty string λ , okay, $L_1 = L$, L^{m+1} is defined as $L^m L$, okay, there is concatenation of L^m with L , okay, where $m > 0$. For example suppose $L = ab, c$, L is consisting of 2 elements ab and c ,

okay, over this set Σ , okay. So, then L_0 is the set containing λ , $L_1 = L$, so it is consisting of ab and c , L_2, L_2 is LL , by this definition, by this definition it is LL .

So, concatenation of L with L , okay, concatenation of L with $LL = L = ab, c, ab$ and c , so L^2 will be LL . So, we write the elements of L to the right of the elements of L . So, ab we write and then to the right of this ab we write ab , okay, that is one thing. Then ab , to the right of ab we write another element c , okay. Then we can also have c to the right of c we can write a, b , okay. That is L^2 , and then we can also have cc , okay. So, $abab, abc, cab$ and cc , these are the elements of L^2 .

And then we have L^3, L^3 is L^2L , okay, L^2 is $abab, abc$, then we have c, a, b , then we have cc , okay and we have L , L is containing ab and c , okay, so concatenation of L^2 with L , okay. So, we have $ababab$, that is one element, then we have $abab$ and then c . So, corresponding to ab we have two elements, $ababab, ababc$ and then corresponding to abc again we have 2 elements $abcab$ and then $abcc$. cab gives $cabab, cabcc$. Then cc gives $ccab$ and then ccc . So, we have altogether, $2^3 = 8$, that is (i.e.) 8 elements, 1, 2, 3, 4, 5, 6, 7, 8.

Now negative power of a language is not defined, do not consider L^{-1} which we have defined here as the negative power of L , it is just a notation, it is the set which consists of the reversal of the words of W . It is not the inverse as we usually understand. Now, let us consider a to be the set a, b, c . Let us find L^* , okay. So, L^* when $L = b^2, L = \{a, b, c\} = abc$. L^* if you take, then, if you take $L = b^2$, then L^* consists of λ and then all elements made from b^2 .

So, $b^2, b^2b^2, b^2b^2b^2$ and so on. So, that means it is $\lambda, b^2, b^4, b^6, b^8$ and so on. So, it consists of empty string λ and all even integral powers of b , even positive integral powers of b . Okay, that is L^* , for the case b^2 . And $L = ab$. When you take $L = a, b$, then $L^*, L^* = \lambda$, then aa, ab, ba, bb , okay, and then we will have aaa okay, aaa then we have aab , then we can have $baba$, we can have bbb and so on, okay.

So, these will be the elements of length 3, okay. We can take a 3 times, we can take $2a$'s and we can take $1b$, we can take $1a$, we can also take abb , okay, that will also be an element, okay, abb . We can then take single b and $2a$'s, okay, single b and $2a$'s, baa , okay. And then we will have bba and bbb . So, those elements will be there of length 3 and so on we can have. So, L^* will consist of λ and all elements made up from a and b . Now $c = abc$, so let us find L^* here.

So, L' will be, this is L , $L = abc$, so $L = \lambda$, L' here will be λ , then we will have abc cube and then we will have $abac$ cube, okay, we will also have aa here, okay, then a b , then ac q , then bb , ba , bb , bc cube, okay, then ca , cb , cc cube and so on, okay. Sorry, ca , c cube a , c cube c and then c cube c cube. Okay. So, all these elements \in to L' , when $L = abc$ cube. That is all in this lecture, thank you very much for your attention.