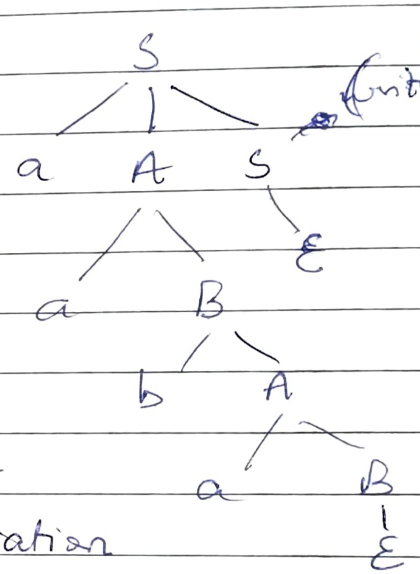
DERIVATION TREE

Also called parse tree of a grammar G.

Capital letters → variables
small letters → terminals
variables can also be called non-terminals.

Making a derivation tree.

$S \to \cancel{aS}\, aAS \mid \varepsilon$
$A \to aB$
$B \to \varepsilon$
$B \to bA$

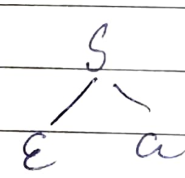


aAS
(writing in order)

So all the leaves in order ~~are~~ is called yield of derivation
Here yield is aaba.

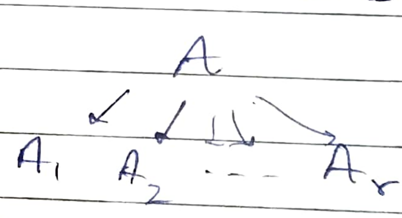
ε only occurs at leaves and it doesn't have any siblings

like eg.



$S \to \varepsilon a$ doesn't makes any sense

$A \to A_1 A_2 A_3 \cdots A_r$ ⟷

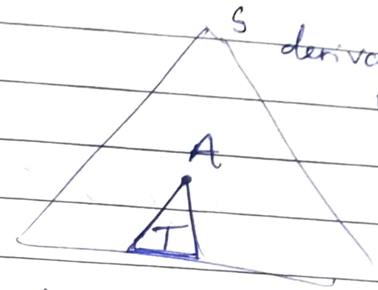


We can say one from other

• Theorem

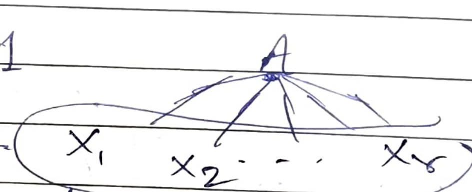
Let $G(V,T,P,S)$ be a context free grammar (CFG). Then for any $A \in V$, $A \overset{*}{\Rightarrow} \alpha$ iff there is a subtree ~~of~~ T of derivation tree with its root labelled A whose yield is α.

Theorem Means $A \overset{*}{\Rightarrow} \alpha \iff$



For iff. lets prove "$\Leftarrow$" part
induction on internal nodes of T

basis: $i=1$



Let A be just 1 internal node, and it has such tree

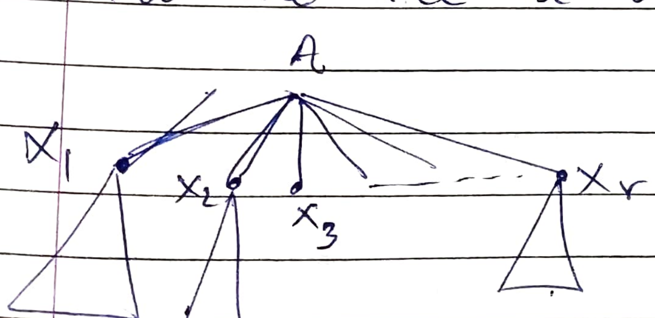
it can be said $A \overset{*}{\Rightarrow} X_1 X_2 \dots X_r$

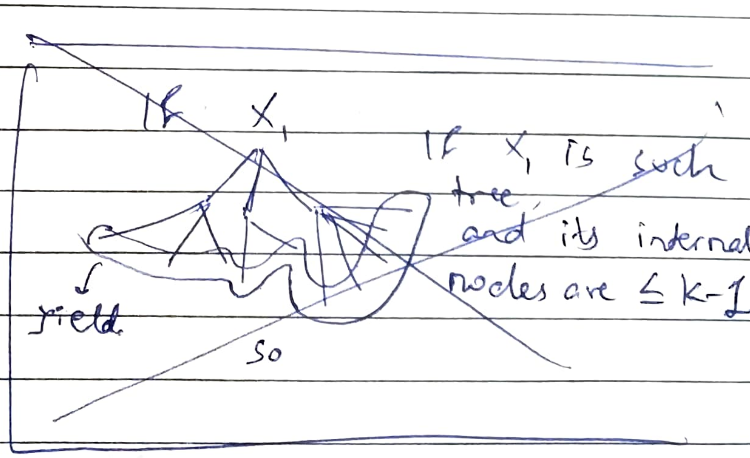
~~$X_1 X_2 \dots X_r$~~

suppose this yield is α so $A \overset{*}{\Rightarrow} \alpha$

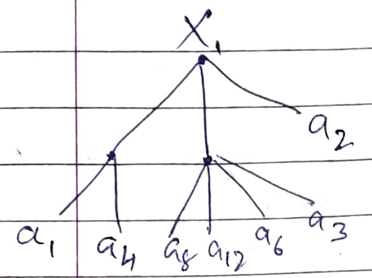
IH: for $i<k$, this is true.
IS: for $i=k$.

Let the tree be like





So apart from 'A', there are $k-1$ internal nodes.



Let X_1 be such a tree

So yield of X_1 is $\underbrace{a_1 a_4 a_8 a_{12} a_6 a_3 a_2}_{\alpha_1}$

~~yield of~~

internal nodes of $X_1 \leq k-1$

we can say, from IH

$X_1 \stackrel{*}{\Rightarrow} a_1 a_4 a_8 a_{12} a_6 a_3 a_2$

Similarly

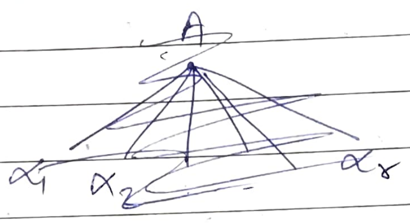
$X_2 \stackrel{*}{\Rightarrow} \alpha_2$

$\vdots$

$X_r \stackrel{*}{\Rightarrow} \alpha_r$

X_1 yield is α_1, then $X_1 \stackrel{*}{\Rightarrow} \alpha_1$

~~Let by~~ Doing this for all children of A



Production

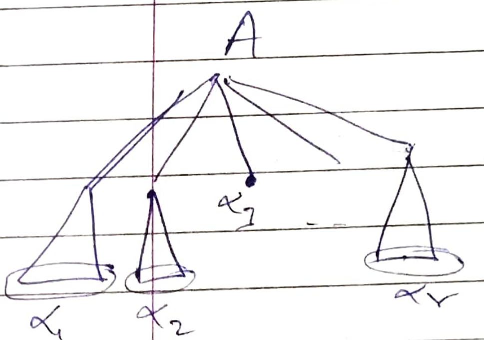
~~yield of A = $\alpha_1\alpha_2 \dots \alpha_r$~~

step 1 $A \Rightarrow X_1 X_2 \dots X_r$

(in 0 or more steps) $\stackrel{*}{\Rightarrow} \alpha_1 \alpha_2 \dots \alpha_r$

$\Rightarrow \alpha.$

(as at one time only one transition is done in sentential form)



So yield of A from graph $= \alpha_1 \alpha_2 \alpha_3 \dots \alpha_r$

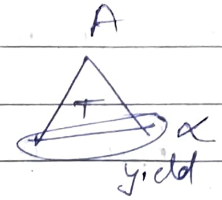
$= \alpha$

~~Hence Proved~~ //.

Part 2 "$\Rightarrow$"

We have to prove

$A \stackrel{*}{\Rightarrow} \alpha \Rightarrow$



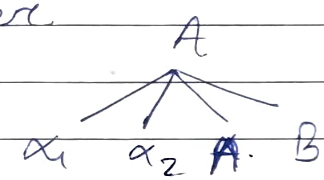
we need to focus on this and prove other

induction on number of steps of the derivation of $A \stackrel{*}{\Rightarrow} \alpha$

basis: In 1 step.

$$A \Rightarrow \underbrace{\alpha_1 \alpha_2 A B}_{\text{taken arbitrarily}}$$

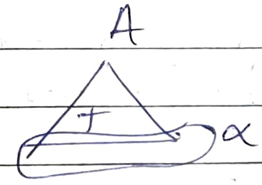
So $A \rightarrow \alpha_1 \alpha_2 AB$ must be there.
and thus we must have a tree $\longrightarrow$



For given production α, we found a tree with root A having yield α

IH: If no. of steps $< k$ then

$A \xRightarrow{(<k)} \alpha$



symbol for number here describes no. of steps for production

I.S: no. of steps is k.

$A \xRightarrow{k} \alpha$ (given)

we split it into first step and rest $k-1$ steps.

$$A \Rightarrow A_1 A_2 \cdots A_r \xRightarrow{k-1} \alpha$$

Let α be concatenation of $\alpha_1 \alpha_2 \ldots \alpha_r$

$$\alpha_1 \alpha_2 \ldots \alpha_r = \alpha.$$

We can say

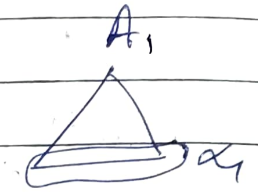
$$A \Rightarrow A_1 A_2 \ldots A_r \overset{k-1}{\Longrightarrow} \alpha_1 \alpha_2 \ldots \alpha_r$$

Let A_1 have a ~~y~~ production of α_1 and it will be in $(\leq k-1)$ steps

Remember, this is just ~~an~~ existential proof. We just need to prove there exists...

$$A_1 \overset{\leq k-1}{\Longrightarrow} \alpha_1 \text{, by IH}$$

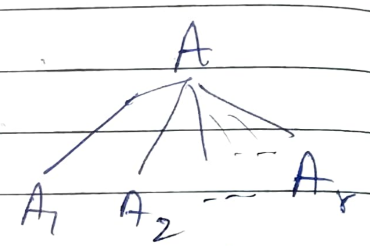
we can have a tree



Similarly we have for $A_2, A_3, \ldots, A_r$

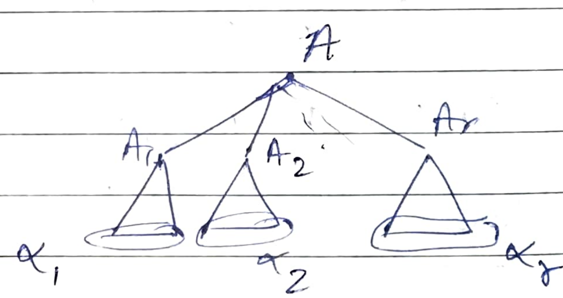
reviewing our first step.

$A \Rightarrow A_1 A_2 \ldots A_r$ corresponding tree is



and now substituting $A_1, A_2, \ldots A_r$ rooted subtrees

~~According to the ~~earlier~~ main statement~~



~~$A \Rightarrow \alpha = \alpha_1 \alpha_2 \ldots \alpha_r$~~

Yield of this tree is $\alpha_1 \alpha_2 \alpha_3 \ldots \alpha_r$ which is equal to α

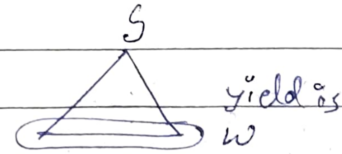
So ~~we~~ according to our production, we have a ~~deriv~~ subtree rooted at A, having yield α.

Hence Proved //.

Corollary

If $w \in L(G)$ for CFG, then w has atleast one parse tree

$S \overset{*}{\Rightarrow} w$



Parse tree defination : Ordered, rooted tree that represents a structure (here our branched tree structure) of a string according to some production of CFG.

WRONG ~~You may get confused between parse tree and derivation tree, but here in parse tree, S is not necessarily said to be a start point.~~

2 parse trees ~~having different structures (depending on order of derivation) are called different parse~~ trees, but both are derivation tree.

- Leftmost derivation
(substituting, expanding leftmost variable)

eg.

$S \rightarrow aAS \mid a$

$A \rightarrow SbA \mid SS \mid ba$

$S \Rightarrow a\underline{A}S \Rightarrow a\underline{S}bAS \Rightarrow aab\underline{A}S$
$\Rightarrow aabba\underline{S}$
$\Rightarrow aabbaa.$

Similar logic for rightmost derivation

$S \Rightarrow aA\underline{S} \Rightarrow aAaA\underline{S} \Rightarrow aAa\underline{A}a \Rightarrow a\underline{A}abaa$
$\Rightarrow abaabaa$

Third one is arbitrary production.

Corresponding to a parse tree, there is a unique leftmost derivation
(To have example, look at first one we took while defining derivation tree, 6 pages back, yield was aaba)

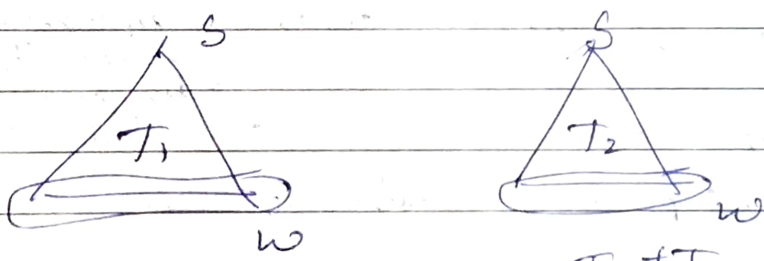
Similarly ~~there~~ corresponding to a parse tree there exists a unique rightmost derivation

~~AMBIGUO~~

Ambiguous grammar

$w \in L(G)$

There exist multiple derivations to same w.
So there may be different parse trees.



Both derived using G.

$T_1 \neq T_2$

As ~~So~~ w ~~can~~ has multiple derivations and multiple parse trees, it is said to have ambiguous grammar

Inherently ambiguous language: For a language L, it is not possible to come up with a non-ambiguous grammar