Grammars & Languages

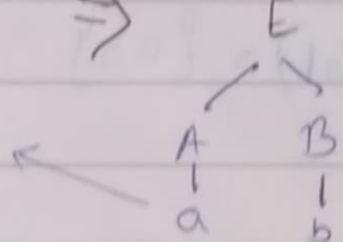
- In English, rules to produce/generate correct langu...
- In TOC, production rules that generate words in a lang...

Ex:

G: $E \to AB$	Variables: $\{E, A, B\}$	⇒ often represented using capital letters
$A \to a$	Terminals (symbols): $\{a, b\}$	⇒ represented by sma... letters
$B \to b$	Starting variable: E	

E says generates AB
A " a
B " b

Lang generated by G → $L(G) = \{ab\}$

⇒ 

Mathematically, any grammar G can be represented by 4 tuples - (V, T, P, S)

- V: Finite Non-empty set of variables (non-terminals or auxiliary symbol...
- T: Finite set of terminals.
- P: Finite non-empty set of production rules.
- S: start symbol (Symbol from where we start producing our sentences or strings)
  unique

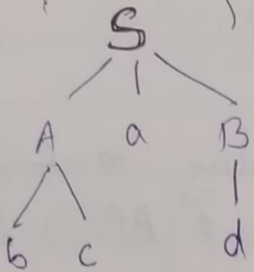
Parse Tree.
(Derivation tree)
(Syntax tree)

Ex:

1) $S \rightarrow AaB$

$A \rightarrow b|c$

$B \rightarrow d$



$\Rightarrow L(G) = \{bad, cad\}$

We can also represent it as

$$S \rightarrow \boxed{AaB} \rightarrow \boxed{baB} \rightarrow bad$$
$$\phantom{S \rightarrow \boxed{AaB}} \rightarrow \boxed{caB} \rightarrow cad$$

~~seque~~ sentential forms (All the strings in between start & end word. It ~~always~~ mostly contains both terminals & variables)

2)

G1: $S \rightarrow AB$
$A \rightarrow a$
$B \rightarrow b$

$L(G1) = \{ab\}$

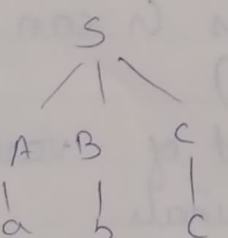
G2: $S \rightarrow Ab$
$A \rightarrow a$

$L(G2) = \{ab\}$

$\Rightarrow$ $G1 = G2$ iff $L(G1) = L(G2)$

There can be more than one grammar for a Language but for every grammar there'll be only 1 language.

3) G: $S \rightarrow ABC$
$A \rightarrow a$
$B \rightarrow b$
$C \rightarrow c$



$L(G) = \{abc\}$

* Backus - Naur Form (BNF) :

It says if $A \rightarrow \alpha_1$, $A \rightarrow \alpha_2$, $A \rightarrow \alpha_3$ then we can write it as $A \rightarrow \alpha_1 | \alpha_2 | \alpha_3$

Ex:

1) $S \rightarrow aS$, $S \rightarrow b$ $\Rightarrow$ $S \rightarrow aS | b$

2) $A \rightarrow a$, $A \rightarrow b$, $A \rightarrow \epsilon$ $\Rightarrow$ $A \rightarrow a | b | \epsilon$

3) $S \rightarrow AaB$

$\left.\begin{array}{l} A \rightarrow aA \\ A \rightarrow b \end{array}\right\} \Rightarrow A \rightarrow aA | b$

$\left.\begin{array}{l} B \rightarrow d \\ B \rightarrow e \end{array}\right\} \Rightarrow B \rightarrow d | e$

* Recursive production.

$A \rightarrow aA$ (Right)

$A \rightarrow Aa$ (Left)

$A \rightarrow aAb$ (General)

$A \rightarrow AaA$ (L & R)

$\left.\begin{array}{l} A \rightarrow aB \\ B \rightarrow bA \end{array}\right\} \Rightarrow A \rightarrow abA$ (Indirect)

Grammar consisting of recursive production rules $\Rightarrow$ recursive grammar

* Recursive & Non-recursive Grammar

- Recursive → Recursive production rules
- Non-recursive → No recursive production rules

<u>Recursive grammar</u> always generates <u>∞-lang.</u>

<u>Non-recursive grammar</u> always generate <u>a finite lang.</u>

Ex:

1) $A \to aA \mid b \Rightarrow$ $A \to aA \to aaA \to aaaA \to \dots$; $A \to b$, ab, aab, $aaab \dots \Rightarrow a^*b$

2) $A \to Aa \mid b \Rightarrow$ $A \to Aa \to Aaa \to Aaaa \to \dots$; $A \to b$, ba, baa, $baaa \dots \Rightarrow ba^*$

3) $A \to aA \mid \epsilon \Rightarrow$ $A \to aA \to aaA \to aaaA \to \dots$; $A \to \epsilon$, a, aa, $aaa \dots \Rightarrow a^*$

4) $A \to Aa \mid \epsilon \Rightarrow a^*$

5) $A \to aA \mid a \Rightarrow$ $A \to aA \to aaA \to aaaA \to \dots$; $A \to a$, aa, aaa, $aaaa \dots \Rightarrow a^*a \Rightarrow a^+$

6) $A \to Aa \mid a \Rightarrow aa^* \Rightarrow a^+$

7) $A \to aA \mid bA \mid \epsilon \Rightarrow$ $A \to \epsilon$; $A \to aA \to a,\ aaA,\ abA \dots$; $A \to bA \to b,\ baA,\ bbA \dots \Rightarrow (a+b)^*$

8) $A \to aA \mid bA \mid a \mid b \Rightarrow (a+b)^+$

* Find languages generated by

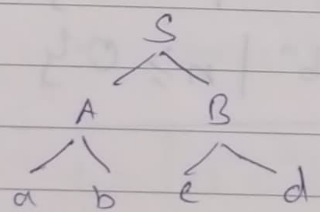
1) $S \to AB$
$A \to a|b$
$B \to c$

$L(G) = \{ac, bc\}$

2) $S \to AaB$
$A \to b$
$B \to c|d$

$S \to AaB \to baB \to bac,\ bad$ $\Rightarrow L(G) = \{bac, bad\}$

3) $S \to AB$
$A \to a|b$
$B \to c|d$



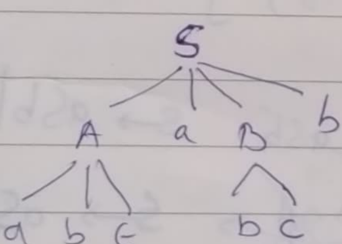
$L(G) = \{ac, ad, bc, bd\}$

4) $S \to AB|BA$
$A \to a|b$
$B \to c|d$

$S \to AB \to aB \to ac,\ ad$
$AB \to bB \to bc,\ bd$
$S \to BA \to cA \to ca,\ cb$
$BA \to dA \to da,\ db$

$\Rightarrow L(G) = \{ac, ad, bc, bd, ca, cb, da, db\}$

5) $S \to AaBb$
$A \to a|b|\epsilon$
$B \to b|c$



$\Rightarrow L(G) = \{aabb, aacb, babb, bacb, abb, acb\}$

6) $S \to Ab$
$A \to bB|b$
$B \to c|d|\epsilon$

$S \to Ab$
$\downarrow$
$bB|b$
$\downarrow$
$c|d|\epsilon$

$\Rightarrow L(G) = \{bb, bcb, bdb\}$

7) $S \to aSb \mid \epsilon$

$S \to aSb \to aaSbb \to aaaSbbb \ldots$

$\to \epsilon \quad ab \quad aabb \quad aaabbb \ldots$

$ab \quad a^2b^2 \quad a^3b^3 \ldots$

$\therefore L(G) = \{a^n b^n \mid n \geq 0\}$ There is no RE corresponds to this lang.

8) $S \to aaSb \mid \epsilon$

$S \to aaSb \to aaaaSbb \to aaaaaaSbbb$

$\to \epsilon \quad aab \quad aaaabb \quad aaaaaabbb$

$(a^2)b^1 \quad (a^2)^2 b^2 \quad (a^2)^3 b^3$

$\therefore L(G) = \{(aa)^n b^n \mid n \geq 0\}$

$L(G) = \{a^{2n} b^n \mid n \geq 0\}$

9) $S \to aSa \mid bSb \mid \epsilon$

$\{\epsilon, aa, bb, abba, baab, abaaba, \ldots\}$

abba same

reverse

$L = \{ww^R \mid w \in (a+b)^*\}$

10) $S \to aSb \mid aAb$

$A \to aA \mid \epsilon$

a^*

$S \to aSb$, $S \to aA$; $aA \to aaA$, a

$S \to aSb \mid aa^*b$

$S \to aSb \mid a^+b$

Now

S
aSb
$aaSbb$
$aaaSbbb$
$a^nSb^n \quad n \geq 0$

Now, we can replace S with a^+b

$a^n a^+ b b^n$

$a^n a a^* b b^n$

$\Leftarrow a^{n+1} a^* b^{n+1}$

$L = \{a^m b^n \mid m \geq n, n \geq 1\}$

$\Leftarrow$ #a's is always $\geq$ #b's

1) $S \to aSb \mid aBb$ $\Rightarrow$ $S \to aSb \mid ab^*b$

$B \to bB \mid \epsilon$ $\qquad S \to aSb \mid ab^+$

$\Downarrow$

b^*

S

aSb

$aaSbb$

$aaaSbbb$

$\Downarrow$

We can replace S at any moment with ab^+ $\qquad a^nSb^n,\ n \geq 0$

$\therefore\ a^n ab^+ b^n \Rightarrow a^{n+1} b^* b b^n$

$a^{n+1} b^* b^{n+1}$

$\therefore\ L = \{a^m b^n \mid m \leq n,\ m \geq 1\}$

$\#a\text{'s} \leq \#b\text{'s}$

2) $S \to aSb \mid aAb$ $\Rightarrow$ $S \to aSb \mid ac^+b$

$A \to cA \mid c$

$\Downarrow$

$A \to c^+$

We can't write $A \to$ reg. E. but for our understanding

$\therefore\ L = \{a^n c^m b^n \mid m \geq 1;\ n \geq 1\}$

S

aSb

$aaSbb$

$aaaSbbb$

$\Downarrow$

We can replace S with ac^*b $\Leftarrow a^nSb^n$

$\Rightarrow a^n ac^+ b b^n$

$a^{n+1} c^+ b^{n+1}$

3) $S \to aSa \mid bSb \mid c$

Q.9

$S \to c$

$S \to aSa$

$S \to bSb$

$\Rightarrow \{c, aca, bcb, abcba, aacaa, bacab, baacaab \ldots\}$

$\Downarrow$

$L = \{wcw^R \mid w \in (a+b)^*\} \Leftarrow wcw^R$

4) $S \to AB$

$A \to aAb \mid \epsilon \Rightarrow (a^nb^n)\ n \geq 0$

$B \to cB \mid \epsilon \Rightarrow c^*$

$\Rightarrow$ ~~$S \to (ab)^* c^*$~~

$S \to a^n b^n c^*$

~~$L = \{a^n b^n c^m \mid m \geq 1$~~

$L = \{a^n b^n c^m \mid n \geq 0,\ m \geq 0\}$

15) $S \rightarrow SaSbS \mid SbSaS \mid \epsilon \Rightarrow \{\epsilon, ab, ba, abba, baab, baabba, \ldots\}$

$SaSbS$: $S \rightarrow 1a1b$, $S \rightarrow 1a1b$, $S \rightarrow \epsilon$ $\Downarrow$ #a's = #b's

$SbSaS$: $S \rightarrow \epsilon$, $S \rightarrow 1a1b$, $S \rightarrow 1a1b$; #a's = #b's

$\Rightarrow$ Language generated by this will contain equal no. of a's & b's

$$L = \{ w \in (a+b)^* \mid |w_a| = |w_b| \}$$

16) $S \rightarrow aSb \mid aAb \longrightarrow \} \Rightarrow S \rightarrow aSb \mid ac^n d^n b,\ n \geq 1$

$A \rightarrow cAd \mid cd \Rightarrow c^n d^n,\ n \geq 1$

S

aSb

$aaSbb$

$\Downarrow$

$a^m S b^m,\ m \geq 0$

$\Leftarrow$

$a^m a c^n d^n b b^m$ ($n \geq 1$, $m \geq 0$)

$a^{m+1} c^n d^n b^{m+1}$, $n \geq 1$, $m \geq 0$

$\Leftarrow$

$L = \{ a^m c^n d^n b^m \mid m \geq 1;\ n \geq 1 \}$

* Construct a grammar that generates

1) a) all strings using $\Sigma = \{a, b\}$ including ϵ $\rightarrow (a+b)^*$

$S \rightarrow aS \mid bS \mid \epsilon$

b) excluding ϵ

$S \rightarrow aS \mid bS \mid a \mid b$

2) a) strings that starts with 'b' $b(a+b)^*$

$S \rightarrow bA$

$A \rightarrow \epsilon \mid aA \mid bA$

b) Starts with 'ab'

$S \rightarrow abA$

$A \rightarrow \epsilon \mid aA \mid bA$

3) a) strings that end in 'a'

$S \rightarrow Aa$

$A \rightarrow \epsilon \mid aA \mid bA$

b) Strings that end in 'ab'

$S \rightarrow Aab$

$A \rightarrow \epsilon \mid aA \mid bA$

4) Strings that contain 'ab' as a substring

$S \rightarrow AabA$

$A \rightarrow \epsilon \mid aA \mid bA$

5) a) Strings that starts & ends in 'a'

$S \rightarrow aAa \mid a$

$A \rightarrow \epsilon \mid aA \mid bA$

5) b) Starts & ends in same symbol

$S \rightarrow a \mid b \mid aAa \mid bAb$

$A \rightarrow \epsilon \mid aA \mid bA$

c) Starts & ends in diff symbol

$S \rightarrow aAb \mid bAa$

$A \rightarrow \epsilon \mid aA \mid bA$

6) a) Third symbol from left is 'b'

$S \rightarrow AAbB$

$A \rightarrow a \mid b$

$B \rightarrow \epsilon \mid aB \mid bB$

or

$S \rightarrow AbB$

$A \rightarrow aa \mid ab \mid ba \mid bb$

$B \rightarrow \epsilon \mid aB \mid bB$

b) 4^{th} symbol from right is 'a'

$S \rightarrow AaBBB$

$A \rightarrow aA \mid bA \mid \epsilon$

$B \rightarrow a \mid b$

7) a) #a's = 2

$S \rightarrow AaAaA$

$A \rightarrow bA \mid \epsilon$

b) #a's ≤ 2

$S \rightarrow A \mid AaA \mid AaAaA$

$A \rightarrow bA \mid \epsilon$

or

$S \rightarrow BABAB$

$A \rightarrow a \mid \epsilon$

$B \rightarrow \epsilon \mid bB$

c) #a's ≥ 2

$S \rightarrow AaAaA$

$A \rightarrow \epsilon \mid aA \mid bA$

7) d) # a's is even
b^* or $(b^*ab^*ab^*)^*$

$S \to B|X$
$B \to \epsilon|bB$ (b^*)
$X \to \epsilon|AX$ $((b^*ab^*ab^*)^*)$
$A \to BaBaB$ $(b^*ab^*ab^*)$

e) # a's is odd
$\Rightarrow \underbrace{b^*ab^*}_{1} + \underbrace{(b^*ab^*ab^*)^*}_{\text{even}}$

$S \to BaBX$
$B \to \epsilon|bB$ → This will generate b^*
$X \to \epsilon|AX$ → $(b^*ab^*ab^*)^*$
$A \to BaBaB$ → this will generate $b^*ab^*ab^*$

f) # a's $\equiv 1 \pmod 3$
1, 4, 7, 10, 13, ...
1 + 0, 3, 6, 9, 12 ....
$b^*ab^*(b^*ab^*ab^*ab^*)^*$

$S \to BaBX$
$B \to \epsilon|bB$
$X \to \epsilon|AX$
$A \to BaBaBaB$

8) a) $|w| = 2$
$S \to ab|ba|aa|bb$ or $S \to AA$, $A \to a|b$

b) $|w| \le 2$
$S \to \epsilon|A|AA$
$A \to a|b$
or $S \to AA$, $A \to \epsilon|a|b$

c) $|w| \ge 2$
$S \to AAB$
$A \to a|b$
$B \to \epsilon|aB|bB$

d) $|w| \equiv 0 \pmod 2$
$((a+b)^2)^*$

$S \to X$
$A \to a|b$
$X \to \epsilon|AAX$

e) $|w| \equiv 1 \pmod 2$
$(a+b)((a+b)^2)^*$

$S \to AX$
$A \to a|b$
$X \to \epsilon|AAX$

f) $|w| \equiv 2 \pmod 3$
2, 5, 8, 11, 14, ...
$(a+b)^2((a+b)^3)^*$

$S \to AAX$
$A \to a|b$
$X \to \epsilon|BX$
$B \to AAA$

9) a) $\{a^m b^n \mid m, n \geq 0\}$

$S \to AB$

$A \to \epsilon \mid aA$

$B \to \epsilon \mid bB$

b) $\{a^m b^n \mid m, n \geq 1\}$

$S \to AB$

$A \to a \mid aA$

$B \to b \mid bB$

c) $\{a^m b^n \mid m \geq 1, n \geq 2\}$

$S \to AB$

$A \to aA \mid a$

$B \to bb \mid bB$

d) $\{a^m b^n \mid m+n \text{ is even}\}$

m : even | m : odd
n : even | n : odd

$\underbrace{(aa)^*}_{A}\underbrace{(bb)^*}_{B}$ $\quad a\underbrace{(aa)}_{A}{}^* b\underbrace{(bb)}_{B}{}^*$

$\therefore$ $S \to XY \mid aXbY$

$X \to \epsilon \mid AX \quad (A^*)$

$A \to aa$

$B \to bb$

$Y \to \epsilon \mid BY \quad (B^*)$

e) $\{a^m b^n \mid m+n \text{ is odd}\}$

m : even | m : odd
n : odd | n : even

$(aa)^* b(bb)^* \quad a(aa)^*(bb)^*$

$S \to XbY \mid aXY$

$A \to aa$

$B \to bb$

$X \to \epsilon \mid AX$

$Y \to \epsilon \mid AY$

10) a) $\{a^m b^n \mid m = n\}$

$S \to aSb \mid \epsilon$

b) $\{a^m b^n \mid m = 2n\}$

$S \to aaSb \mid \epsilon$

c) $\{a^m b^n c^n \mid m, n \geq 1\}$

$S \to AB$

$A \to a \mid aA$

$B \to bc \mid bBc$

d) $\{a^n b^m c^n \mid m, n \geq 1\}$

$S \to aSc \mid aBc$

$B \to b \mid bB$

e) $\{a^m b^n c^p \mid n = m+p ;\ n, m, p \geq 0\}$

ϵ, abbc, ab, bc, ...

$m=p=0$, ϵ, $n=0$ — $p=0$ — $m=0$

$S \rightarrow AB$

$A \rightarrow \epsilon \mid aAb$

$B \rightarrow \epsilon \mid bBc$

$\underbrace{\overset{2}{aa}\ \overset{2}{bb}}_{\text{equal } a \,\&\, b} \mid \underbrace{\overset{1}{b}\ \overset{1}{c}}_{\text{equal } b \,\&\, c}$

11) a) $\{a^m b^n \mid m>n ;\ m,n \geq 1\}$

$S \rightarrow aSb \mid aAb$

$A \rightarrow a \mid aA$

11) b) $\{a^m b^n \mid m<n ;\ m,n \geq 1\}$

$S \rightarrow aSb \mid aAb$

$A \rightarrow b \mid bB$

c) $\{a^m b^n \mid m \neq n ;\ m,n \geq 1\}$ — $m<n$ or $m>n$

~~$S \rightarrow aSb \mid aAb \mid aBb$~~

~~$A \rightarrow a \mid aA$~~

~~$B \rightarrow b \mid bB$~~

$S \rightarrow S_1 \mid S_2$

$S_1 \rightarrow aS_1b \mid aAb$

$S_2 \rightarrow aS_2b \mid aBb$

$A \rightarrow a \mid aA$

$B \rightarrow b \mid bB$

12) a) $\{w \epsilon w^R \mid w, \epsilon \in \{a,b\}^*\}$ — ϵ is word

$S \rightarrow aSa \mid bSb \mid \epsilon$

$S \rightarrow aAa \mid bAb$

$A \rightarrow \epsilon \mid aA \mid bA$

b) $\{wcw^R \mid w \in (a+b)^*\}$ — c is symbol

$S \rightarrow aSa \mid bSb \mid c$

c) $\{ww^R \mid w \in (a+b)^*\}$

$S \rightarrow aSa \mid bSb \mid \epsilon$

13) a) $\{w \in \Sigma^* \mid |w_a| = |w_b|\}$

$S \rightarrow SaSbS \mid SbSaS \mid \epsilon$

_ a _ b _

_ b _ a _

S

b) $\{w \in \Sigma^* \mid |w_a| = 2|w_b|\}$

$S \rightarrow SaSaSbS \mid$

$SaSbSaS \mid$

$SbSaSaS \mid \epsilon$

_ a _ a _ b _

_ a _ b _ a _

_ b _ a _ a _

*14) $\{a^n b^n c^n \mid n \geq 1\}$

abc, aabbcc, aaabbbccc, ...

$S \rightarrow abc \mid aSAc$

$CA \rightarrow AC$

$bA \rightarrow bb$

S

aSAc

abc

a abcAC

aabACC

aabbcc