

Chomsky
normal
form - CNF - format for rules of the grammar

ϵ \in

A grammar G is said to be in CNF if all the rules/productions are in :-

(1) $S \rightarrow \epsilon$ (\leftarrow empty string)

(2) $A \rightarrow a$

$a \in \Sigma$

(3) $A \rightarrow BC$

$A, B, C \in V$

- (1) Only start symbol is allowed to produce empty strings
- (2) Any var., including the start symbol, produces/yields 1 terminal.
- (3) " " " " produces 2 variables.

Note:- we can convert any CFG into CNF form.

$\{ S \rightarrow a \checkmark$
 $S \rightarrow AB \checkmark$
 $A \rightarrow a \checkmark$
 $B \rightarrow b \checkmark \}$

ϵ

CNF \checkmark

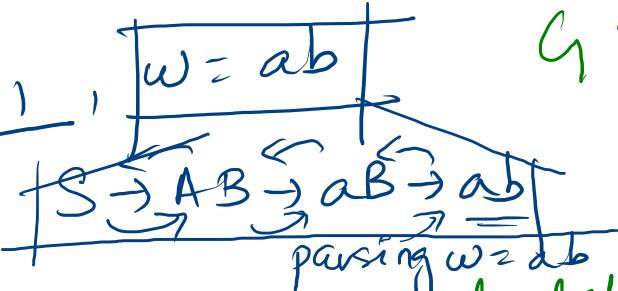
$S \rightarrow \epsilon$
 $A \rightarrow \epsilon \times$

$\{ ① S \rightarrow \epsilon \checkmark$
 $② S \rightarrow AB \checkmark$
 $\Rightarrow ③ A \rightarrow ab \times$
 $④ B \rightarrow b \checkmark$
 $⑤ A \rightarrow a \checkmark \}$

$V = \{ S, AB \}$

$Z = \{ a, b \}$

start = S



G is
not in
CNF

All vars are denoted by capital letters
All terminals are denoted by small letters.

$S \rightarrow \epsilon \Rightarrow$
 $A \rightarrow a \checkmark$
 $A \rightarrow BC$

Parsing - Process of creating a string's productions back to start symbol.

Q Given a grammar in CNF, a target string ω , can G gen. ω ?

Let ω has n characters
terminates $\Rightarrow \omega = \underline{w_1} \underline{w_2} \dots \underline{w_n}$ ✓

Let $\omega = \underline{ab}$, $n = 2$, $w_1 = a, w_2 = b = w_n$



~~$\omega = \underline{aabbaacc}$~~ , $n = 8$, $w_1 = a, w_2 = a \dots w_8 = w_n = ?$

$\boxed{\omega = \underline{w_1} \dots \underline{w_i} \underbrace{w_i^i w_{i+1}^i} \dots \underline{w_n}}$

2 to 2

$$\frac{i=4}{w_i^i=b, w_{i+1}^i=a}$$

Case:- $n=0 \Rightarrow \omega=\epsilon \Rightarrow$ if " $S \rightarrow \epsilon$ " is a product in G , then yes
else no.

Language \rightarrow ^{target} Strings are represented in terms of terminal symbols.

eng. lang terminal symbols — english alphabet (A - Z
a - z
0 - 9
:
"Algorithm"
"SICKS"
"ITER")

$\{ S \rightarrow a \}$

$S \rightarrow A \in V$ (Variable set)

$S \rightarrow aA$

$a \in T$ (Terminal set)

$A \rightarrow aB$

[a, aa]

~~Parsing prob~~

G(CNF), w , can G gen w ?

$w = \underline{\underline{aa}}$
 $S \rightarrow a X$

$S \rightarrow aA \rightarrow \underline{\underline{aa}}$ \Rightarrow Parsing.

yes/no $S \rightarrow \dots \rightarrow w$

$$\underline{\omega = \omega_1 \dots \omega_i} \quad \underline{\omega_{i+1} \dots \omega_m}$$

Case 1

$$n=0 \Rightarrow \underline{\omega = \epsilon} \text{ (empty string)}$$

Case 2

$$n=1 \Rightarrow \underline{\omega = t} \quad (\underline{t \in \Sigma})$$

Case 3 $n \geq 2 \Rightarrow \omega$ has ≥ 2 terminals

$$\frac{\underline{\omega_1 \dots \omega_i}}{B} \cdot \frac{\underline{\omega_{i+1} \dots \omega_n}}{C}$$

logical
and

$$\checkmark A \rightarrow BC$$

$$\frac{\text{CNF}}{\underline{\underline{S \rightarrow \epsilon}}} \quad \begin{array}{ll} \checkmark & \text{yes} \\ \times & \text{no} \end{array}$$

$$\frac{\text{"A} \rightarrow t\text{"}}{\underline{\underline{A \rightarrow t}}} \quad \begin{array}{ll} \checkmark & \text{yes} \\ \times & \text{no} \end{array}$$

$$\frac{\underline{\underline{S \rightarrow \epsilon}}}{\underline{\underline{A \rightarrow a}}}$$

$$\frac{\checkmark \boxed{A \rightarrow BC}}{\checkmark}$$

can't give you
strings of length > 1



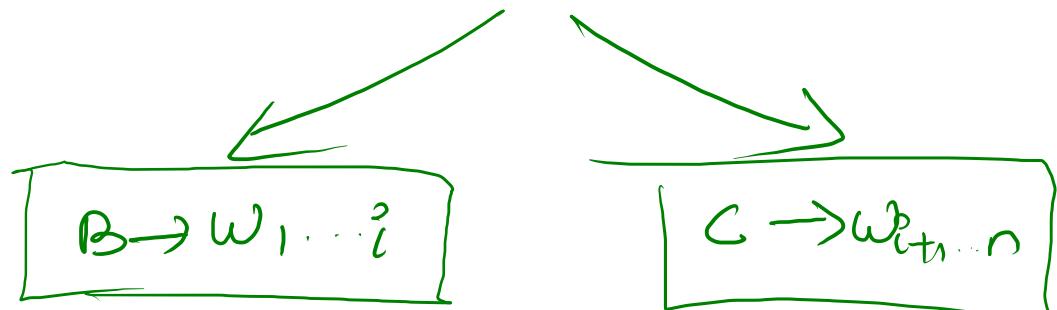
ϵ, a

$$w = \frac{w_1 \dots w_i}{B}, \quad \frac{w_{i+1} \dots w_n}{C}$$

smaller recursive problems of the original problem?

yes

$S \rightarrow w$



$w_1 \dots w_i$

B

$w_{i+1}^o \dots w_n$

C

→ all possible splits

⋮

⋮

w_1 w_2 ✓ w_3 ✓ w_4 w_5 ... w_n ⇒

A → a

$S \rightarrow \epsilon$
 $S \rightarrow aAb$
 $A \rightarrow ab$
 $A \rightarrow aAb$

$w = aabt$

$S \rightarrow aAb \rightarrow aab\ b$

all possible splits

a	a	abb	-	✓/✗
aa	aa	bb	-	✓/✗
aab	aab	b	-	✓/✗
ab	aabb	c	-	✓/✗
bb	ε	aabb	-	✓/✗

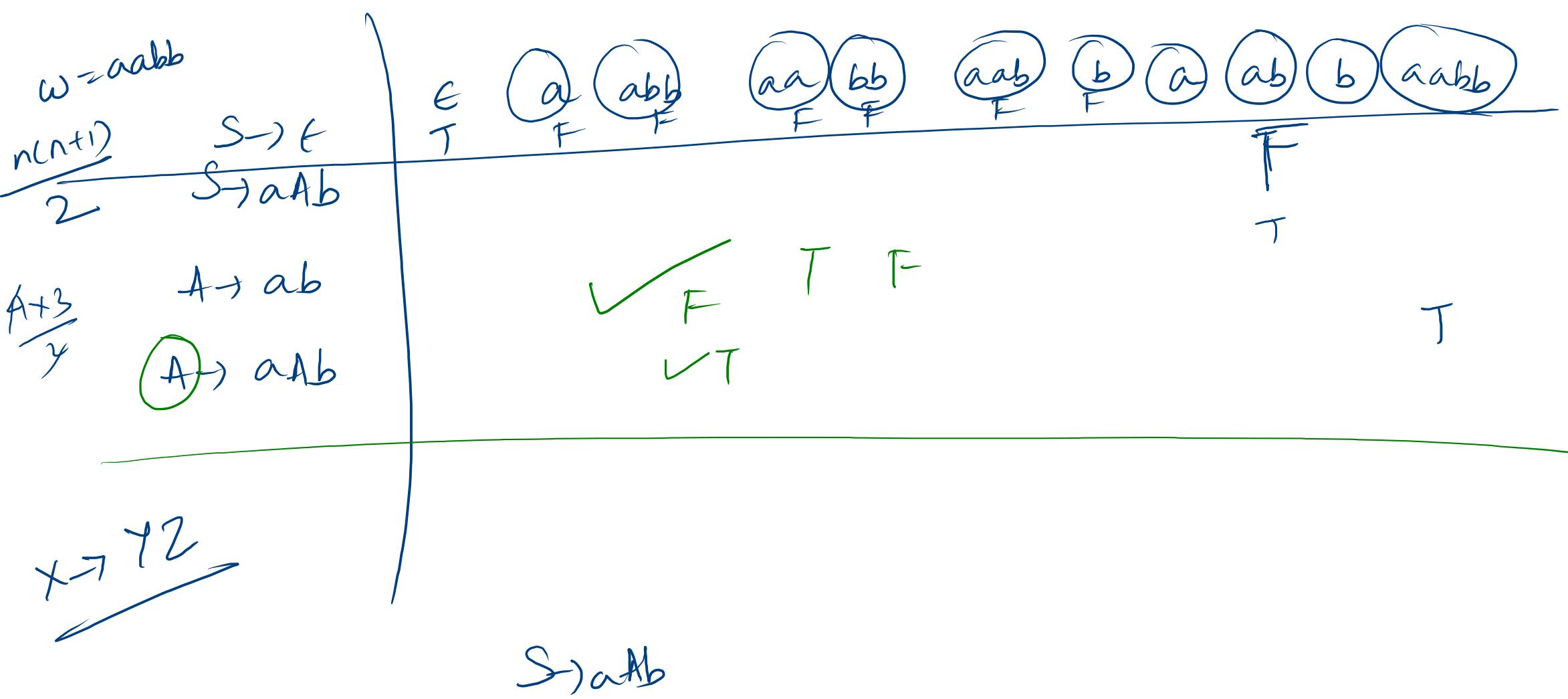
$S \rightarrow a$ $A \rightarrow a$ $S \rightarrow aA$

list which var produces which terminal?

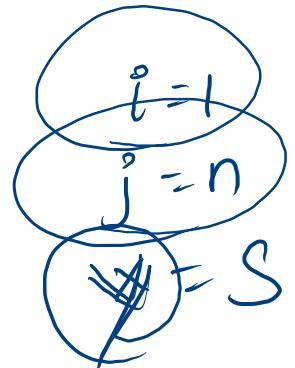
$(S, a), (A, a)$ ✓

Table

	split 1	split 2	-----
$X \rightarrow Y_1 Z$	$X_1 + Y_1 Z_1$		
	$X_2 \rightarrow Y_2 Z_2$		
	:		



$M[i, j, \cancel{x}] = T \Leftrightarrow$ var can produce $\underline{w_i \dots j}$



$M[1, n, \cancel{S}] \rightarrow$ whether G can produce $w_1 \dots n$

$X \rightarrow \underline{YZ}$
 $w_i \dots j$ $w_j \dots w_k$

$$\bigvee^j \frac{M[i, k, y] \cdot M[k+1, j, z]}{i=k}$$

$i=k$



$x \rightarrow \underline{y} \underline{z}$

$\checkmark Y \rightarrow \underline{\omega_i \dots k}$



$\checkmark Z \rightarrow \underline{\omega_{k+1} \dots j}$

$\underline{\omega_i \dots \underline{j}}$

$\vee = \neq$

$\cdot = *$

$$\omega = \omega_0 \dots j$$

