Chomsky normal form — CNF — format for rules of the grammar $\quad \epsilon \quad \in$

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

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

(2) $A \to a$ $\qquad a \in \Sigma$

(3) $A \to BC$ $\qquad A, B, C \in V$

(1) Only start symbol is allowed to produce empty strings

(2) $\underbrace{\text{Any var., including the start symbol}}$, produces/yields 1 terminal.

(3) " " " produces 2 variables.

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

$\{ S \to a$ ✓

$S \to AB$ ✓

$A \to a$ ✓

$B \to b \}$ ✓

ϵ

G_1, $w = ab$

$S \to AB \to aB \to ab$

parsing $w = ab$

CNF ✓

$S \to \epsilon$

$A \to \epsilon$ ✗

G is not in CNF

$\{$ ① $S \to \epsilon$ ✓

② $S \to AB$ ✓

③ $A \to aB$ ✗

④ $B \to b$ ✓

⑤ $A \to a \}$

$V = \{S, A, B\}$

$\Sigma = \{a, b\}$

start $= S$

All vars are denoted by capital letters

All terminals are denoted by small letters.

$S \to \epsilon$ ⇒

$A \to a$ ✓

$A \to BC$

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

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

$$S \rightarrow \cdots\cdots \rightarrow w$$

Let w has n characters/terminals $\Rightarrow w = w_1 w_2 \ldots\ldots w_n$ ✓

Let $w = ab$, $n = 2$, $w_1 = a$, $w_2 = b = w_n$

$$w = \overset{1}{a}\overset{2}{a}\overset{3}{b}\overset{4}{b}\overset{5}{a}\overset{6}{a}\overset{7}{c}\overset{8}{c}, \quad n = 8, \quad w_1 = a, \; w_2 = a \ldots\ldots w_8 = w_n = c$$

$$\boxed{w = w_1 \ldots\ldots \underbrace{w_i w_{i+1}} \ldots\ldots w_n}$$

2 to 7

$i = 4$

$w_i = b$, $w_{i+1} = a$

Case1:- $n = 0 \Rightarrow w = \epsilon \Rightarrow$ if "$S \rightarrow \epsilon$" is a producⁿ in G, then yes else no.

Language → $\overset{\text{target}}{\text{Strings}}$ are represented in terms of terminal symbols.

eng. lang terminal symbols — english alphabet (A - Z
a - z
0 - 9
⋮
)

"Algorithms"
"Sec K5th"
"ITER"

{ S → a ✓
S → aA ✓
A → a }

S, A ∈ V (variable set) ✓

a ∈ Σ ✓ (terminal set)

[a , aa]

w = aa

S → a X

S → aA → aa → parsing.

Parsing prob

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

yes/no   S → ······ → w

$w = w_1 \cdots w_i \; w_{i+1} \cdots w_n$

Case 1 $n = 0 \Rightarrow$ $w = \epsilon$ (empty string)

Case 2 $n = 1 \Rightarrow$ $w = t$ ($t \in \Sigma$)

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

$\underbrace{w_1 \cdots w_i}_{B} \cdot \underbrace{w_{i+1} \cdots w_n}_{C}$

logical and

✓ $A \to BC$

CNF

"$S \to \epsilon$" ✓ yes / × no

"$A \to t$" ✓ yes / × no

$\boxed{\begin{matrix} S \to \epsilon \\ A \to a \end{matrix}}$ ✓ $\boxed{A \to BC}$ ✓

can't give you strings of length > 1

ϵ, a

$$w = \underbrace{w_1 \cdots\cdots w_i}_{B} \cdot \underbrace{w_{i+1} \cdots w_n}_{C}$$

smaller recursive problems of the original problem?

yes

$$\begin{array}{ccc} & S \to w & \\ \swarrow & & \searrow \\ \boxed{B \to w_{1 \cdots i}} & & \boxed{C \to w_{i+1 \cdots n}} \end{array}$$

$\underbrace{w_1 \ \ldots \ w_i}_{B}$ $\underbrace{w_{i+1} \ \ldots \ w_n}_{C}$ $\rightarrow$ all possible splits

w_1 w_2 w_3 w_4 w_5 $\ldots$ w_n $\Rightarrow$

$A \rightarrow a$

$S \to \epsilon$

$S \to aAb$

$A \to ab$

$A \to aAb$

$w = aabb$

$S \to aAb \to aabb$ ✓ all possible splits

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

$S \to a$

$A \to a$

$S \to aA$

list which var produces which terminal?

$[(S, a), (A, a)]$ ✓

Table	split1	split2	......
$X \to YZ$			
$X_1 \to Y_1 Z_1$			
$X_2 \to Y_2 Z_2$			
⋮			

$w = aabb$

$\frac{n(n+1)}{2}$

$\frac{A+3}{4}$

	ϵ	a	abb	aa	bb	aab	b	a	ab	b	$aabb$
$S \to \epsilon$	T	F	F	F	F	F	F				
$S \to aAb$									F		
									T		
$A \to ab$			F	T	F						T
$A \to aAb$			T								

$X \to YZ$

$S \to aAb$

$M[i, j, \cancel{X}] = T \iff \underline{X}$ var can produce $\underline{w_{i \dots j}}$

$i = 1$
$j = n$
$\cancel{X} = S$

$M[1, n, S] \to$ whether G can produce $w_{1 \dots n}$

$X \to \underline{Y}\,\underline{Z}$

$w_{i \dots j}$ $\quad$ $w_{j+1} \dots w_k$

$$\bigvee_{i=k}^{j} M[i,k,Y]\cdot M[k+1,j,Z]$$

$X \to YZ$

$Y \to w_{i \cdots k}$

$Z \to w_{k+1 \cdots j}$

$w_{i \cdots j}$

$\vee = +$

$\cdot = *$

$w = w_{i \cdots j}$

$w_1 \mid w_2 \cdots w_n$ — split 1

$k=2$

$w_1 w_2 \mid w_3 \cdots w_n$ — split 2

$k=3$

$\overbrace{w_1 w_2 w_3}^{w_{i \cdots k}} \mid \overbrace{w_4 \cdots w_n}^{w_{k+1 \cdots j}}$ — split 3

$k=4$

$w_1 \cdots w_{n-1} \mid w_n$ — split 4

$k=n$

T logical and T

T logical or F

$\boxed{X \to YZ}$

$w_{i \cdots j}$

True/False

True/False

$$\bigvee_{i=k}^{j} M[i, k, Y] \cdot M[k+1, j, Z]$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$Y \to w_{i \cdots k} \qquad Z \to w_{k+1 \cdots j}$$

$\bigvee$