25th K-L

CFGs — TOC

{ Grammar
Parsing
C.F. ? }

Parsing CFG ? ⇒ DP

Eng lang
grammar

Hindi lang
grammar
meaningful → grammar

use this lang ⇒ make words / sentences / paragraphs / files in this lang

⇒ for communication

A set of rules for producing meaningful items of communication

"I text a you sent"

↓ gramatically wrong

"I sent you a text"

CFG

$G = \{ V, \Sigma, R, S \}$

- V: helps you for yielding target strings → denoted by capital letters
- Σ: terminal set (target string symbols) → denoted by small letters
- R: set of rules
- $S \in V$ ← belongs to

$$\{ S \to a \quad A \to a \quad S \to \epsilon \}$$

"I sent you a text"

(w) ↑ target string

LHS → can substitute/replace → RHS

- LHS: only variables
- RHS: any mix of terminals + variables
- target string: only terminal symbols

$S \longrightarrow \cdots \longrightarrow w$

$S \longrightarrow w$

parsing / generating / yielding

R:- $\{ S \to a$ (G1)
$A \to a \}$

start :- S

$V = \{S, A\}$

$\Sigma = \{a\}$

$w = a$

$S \to a$ ✓

R:- $\{ S \to a$ (1)
$S \to aA$ (2) ✓
$A \to a \}$

start = S

$V = \{S, A\}$

$\Sigma = \{a\}$

$w = aa$

$S \to aA$

$\to aa$

R:- $\{ S \to \epsilon$ (1)
$A \to a \}$ (2)

start = S

$V = \{S, A\}$

$\Sigma = \{a\}$

$w = \epsilon$ ← empty string

$S \to \epsilon$ ✓

$w = a$

$S \to X$

A is unreachable from S

Compilers

→ compiled code
interpreted

Python Interpreter

↓
whether the entire prg is a legal terminal string in the Python language?

R:- { LHS — RHS
:
:
} → Interpreter

↑ grammar size ↑ Parser

Parsing source bytes

Syntax errors

" Parsing Problem. "

Given a CF grammar G and a target string w, can G generate w?

↓ is in CNF

Chomsky Normal Form

(1) $S \to \epsilon$

(2) $A \to a$

(3) $A \to BC$

start = S

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

$\Sigma = \{a\}$

(1) Only start symbol gen. empty string

(2) Any variable including the start symbol can generate 1 terminal symbol

(3) " " " " " " " " 2 variables.

Note :- Any grammar can be converted to CNF.

$\{S \to a$ ✓

$A \to a$ ✓

$S \to AB$ ✓

$B \to b\}$ ✓

— CNF ? yes

$S \to \epsilon$ ✓

$A \to \epsilon$ ✗

$B \to \epsilon$ ✗

$\{S \to ab$ ✗

$A \to a$

$A \to AB$

$B \to b\}$

$S \to a\ b$

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

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

CNF = ? no

convert to CNF →

$S \to aB$ ✗

a

AB

$S \to \epsilon$

$S \to ab$ ✓

$\{S \to AB,$

$A \to a$

$A \to AB$

$B \to b\}$

CNF? yes

G $V = \{S\}, \Sigma = \phi, R = \phi$

start = S

↓

CNF ? yes

Q Given a CFG in CNF G, and a target string w, can G gen. w? yes/no

$w = w_1 w_2 \cdots w_n$ (n terminals in w) $\Leftrightarrow |w| = n$

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

$= w_{1 \cdots i} \; w_{i+1 \cdots n}$

Case1: if $n = 0 \Rightarrow w = \epsilon$ how to determine if G gen. w?

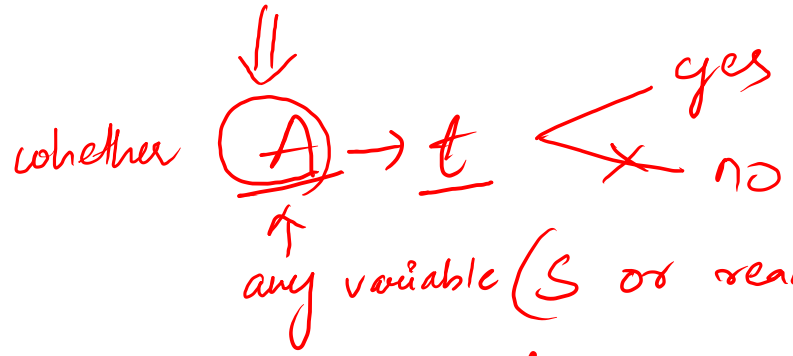
$\epsilon = \sim\sim$

$a \in X$

$\Downarrow$

$S \to \epsilon$? yes / no

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



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

$w = a\ b$

reachable from S

$A \to BC$

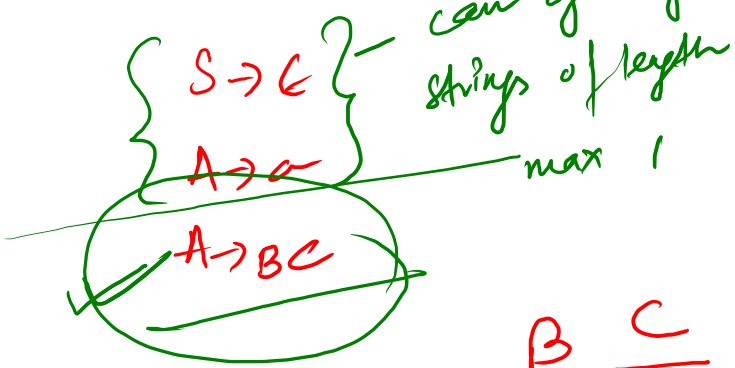
$B \to a$, $C \to b$

$w = abc$

$A \to BC$

① a, $bc \to abc$

② ab, $c \to abc$



splits

① $a|bc = \frac{B}{a}, \frac{C}{bc}$

② $ab|c = \frac{ab}{B}, \frac{c}{C}$

$S \xrightarrow{G_1} w$

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

$A \to BC$

B → can B gen. $w_1 \ldots ?$ and

C → can C gen $w_{i+1 \ldots n}$?

smaller recursive problems of the original prb.

consider all substrings of target string → determine all possible splits

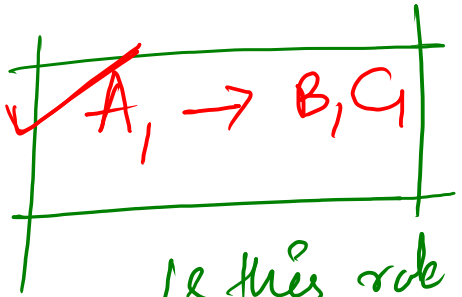
CNF

$A \to BB$

for each split,

1. $A \to BC$ → T/F (B) · T/F (C)
2. $A_1 \to B_1 C_1$
3. $A_2 \to B_2 C_2$

$A_3 \to B_3 C_3$



Is this rule able to gen. any possible split?

$A \to BC$

K splits $\to \omega$ $\quad M[i,k,Y] \cdot M[k+1,j,Z]$

1st split $\frac{T/F}{B_1} \cdot \frac{T/F}{C_1} \to IR_1\ (T/F)$ ✓

2nd split $\frac{T/F}{B_1} \cdot \frac{T/F}{C_1} \to IR_2\ (T/F)$ ✓

3rd split $\frac{T/F}{B_1} \cdot \frac{T/F}{C_1} \to IR_3\ (T/F)$ ✓

$\vdots$

K^{th} split $\frac{T/F}{B_1} \cdot \frac{T/F}{C_1} \to IR_K\ (T/F)$ ✓

$\to$ Ans ✓✓

for all rules of the form $A \to BC$, repeat.

rule $A_2 \to B_2C_2$ → ans

V

rule $A_3 \to B_3C_3$ → ans

⋮

w ? by any rule of $A \to BC$

$\{ S \to \epsilon$

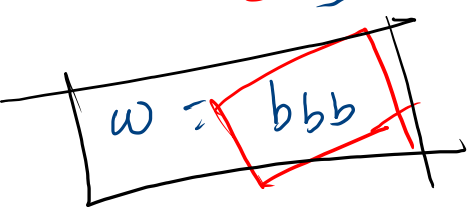
① $S \to AB$

② $A \to BB$

$A \to a$

$B \to b \}$

$A \to BC$



$S \to AB$

$\to BBB$

$\to bbb$

---

b | b b ⇒ b, bb

bb | b ⇒ bb, b

② $A \to BB$

b, bb

B B = F

T · F

bb, b

B B

F · T = F

V

F

① $S \to AB$; b, bb

A B

1 F · F 2 = F

$S \to AB$; bb, b

A B

can $A \to bb$?

$A \to BB \to bB \to bb$

T · T = T

T

= T (yes)

$M[i,k,Y] \cdot M[k+1,j,Z]$

$w_{i \cdots j} \Rightarrow w_{i \cdots k} w_{k+1 \cdots j}$

$i=4, j=n$

$k \geq i$

$X \to YZ \longrightarrow w_{k+1 \cdots j}$

$w_{i \cdots k}$

sp1 → if $Y \to w_{i \cdots k} \Rightarrow M[i,k,Y] = true$??

sp2 → if $Z \to w_{k+1 \cdots j} \Rightarrow M[k+1,j,Z] = true$??

Subproblems are independent → e.g. merge sort

Parsing CFG → subproblems are independent.

bbb b bb

bb b

in a string of length n, no. of substrings = ?

$\frac{n(n+1)}{2} \simeq O(n^2)$

K splits

M[i, j, X]

[w_i .... w_K w_{K+1} .... w_j]

$X \to YZ \Rightarrow$

$|w| = n$

i → n

j → n

K → n

n^3