BITS Pilani
Pilani Campus

Context-Free Grammar

Shashank Gupta
Assistant Professor
Department of Computer Science and Information Systems

Context-Free Grammar

Recognizes a larger class of languages than regular languages.

Finds applications in programming languages and compiler design.

Informal Description

Terminal symbols similar to the input alphabet symbols.

- **Variable symbols** can be replaced with a string of variables and terminals.

Production Rules: Dictates how variables get replaces

- **Start variable**: Starting point of the computation.
- Similar to the start state of FA.

How does Computation Proceeds?

1. Write down the start variable as the current string.

2. Pick a variable in the current string and replace it with one of its production rules.

3. Continue step 2 until no more variables are left.

Example

- Terminals: {0,1}
- Variables: {S}
- Production Rules
- $S \rightarrow 0S1$
- $S \rightarrow \epsilon$
- Start Symbol S

$S \rightarrow 0S1$
$S \rightarrow 00S11$
$S \rightarrow 000S111$
$S \rightarrow 000111$

Hence, this grammar accepts the language of the form $0^n1^n \mid n \geq 0$

Note: Variables are denoted by Capital Letters

Example

- Terminals: {0,1}
- Variables: {S, A, B}
- Production Rules
- $S \rightarrow ASB \mid \in$
- $A \rightarrow 0$
- $B \rightarrow 11$
- Start Symbol S

$S \rightarrow ASB$
$S \rightarrow 0SB$
$S \rightarrow 0ASBB$
$S \rightarrow 00BB$
$S \rightarrow 0011B$
$S \rightarrow 001111$

Hence, this grammar accepts the language of the form $0^n1^{2n} \mid n \geq 0$

Context-Free Grammar (CFG)

A context-free grammar is the 4-tuple (V, Σ, P, S), where

- V is a finite set called variables,
- Σ is a finite set disjoint from V called terminals,
- $P \subseteq V \times \{V \cup \Sigma\}^*$ is a finite set of production rules,
- $S \in V$ is the Start variable

More Terminologies about CFG

If $A \in V$, $u, v \in \{V \cup \Sigma\}^*$ and $A \rightarrow w$ is a rule, then we say that uAv yields uwv in one step. This is denoted as $uAv \Rightarrow uwv$.

The language of G, $L(G) = \{w \in \Sigma^* \mid S \overset{*}{\Rightarrow} w\}$

- A language $L \subseteq \Sigma^*$ is called a context-free language (CFL in short) if there is a CFG G, such that $L = L(G)$.

Examples of CFG

Construct a CFG that generates all strings of a's and b's excluding null string

- L = {a, b, ab, ba, aab, baaba, ------}

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

- Grammar including null string ?
- $S \rightarrow \in \mid aS \mid bS$

Examples of CFG

Construct a CFG that generates all strings of a's and b's where each string starts with a and ends with b

- L = {ab, aab, abb, abbabb, ------}

$S \rightarrow aAb$

$A \rightarrow \in \mid aA \mid bA$