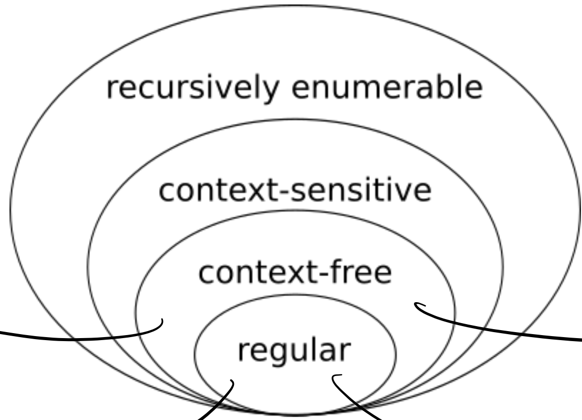
CSCI 301 - Lecture 23: Context-Free Grammars

Chomsky Hierarchy



③ Context-free Grammars

④ Pushdown automata

② Reg. Exp.

DFA ≡ NFA ①

Context-free Grammar

$$
\boxed{\begin{aligned}
S &\to AB \\
A &\to a \\
A &\to aA \\
B &\to b \\
B &\to bB
\end{aligned}}
$$

Derivation

$$
\begin{aligned}
S &\Rightarrow AB \\
&\Rightarrow aAB \\
&\Rightarrow aAbB \\
&\Rightarrow aaAbB \\
&\Rightarrow aaabB \\
&\Rightarrow aaabb
\end{aligned}
$$

Definition

A **context-free grammar** is a 4-tuple (V, Σ, R, S), where:

- V is a set of **variables** (aka **nonterminals**)
- Σ is an **alphabet** (of terminals), where $\Sigma \cap V = \emptyset$
- R is a set of **rules**, of the form

  $A \to w$, with $A \in V$, $w \in (\Sigma \cup V)^*$

  string
- $S \in V$ is a **start symbol**

Definitions

Let $G = (V, \Sigma, R, S)$ be a CFG. Let $A \in V$, and $u, v, w \in (\Sigma \cup V)^*$, and suppose $A \to w$ is a rule in R.

We say that uwv **can be derived in one step** from uAv.

We write this $uAv \Rightarrow uwv$.

Examples:

$aaAb \overset{?}{\Rightarrow} aaab$

$aaAb \overset{?}{\Rightarrow} aaaAb$

$aaAb \overset{?}{\not\Rightarrow} aabb$

$$\boxed{\begin{array}{l} S \to AB \\ A \to a \\ A \to aA \\ B \to b \\ B \to bB \end{array}}$$

This generalizes to a notion of can be derived from (in any number of steps), which we write $u \overset{*}{\Rightarrow} v$.

$$A \overset{?}{\Rightarrow^*} aaA$$

$$B \overset{?}{\Rightarrow^*} bbbbb$$

Definitions The language of a grammar G, $L(G)$, is the set of all strings in Σ^* that can be derived from S.

$$L(G) = \{ w \in \Sigma^* : S \overset{*}{\Rightarrow} w \}$$

A language A is context-free if there exists a context-free grammar G such that $L(G) = A$.

baabba

$S \to aSbS$

$S \Rightarrow bSaS$

$b\boxed{aSbS}aS$

$baaSbSbSaS$

$baabba$

Ex. PDA

$S \to bSaS$

Context-free but not regular:

$$\{a^n b^n : n \in \{0,1,2\ldots\}\}$$

$S \to \varepsilon$

$S \to aSb$

Notational aside:

$S \to \varepsilon$

$\quad | \, aSb$

Parse Trees

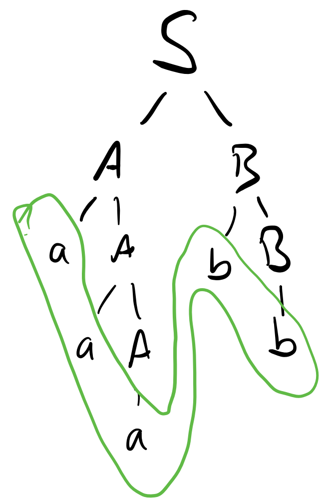
Grammar:

$$\boxed{\begin{aligned} S &\to AB \\ A &\to a \\ A &\to aA \\ B &\to b \\ B &\to bB \end{aligned}}$$

Derivation:

$$\begin{aligned} S &\Rightarrow AB \\ &\Rightarrow aAB \\ &\Rightarrow aAbB \\ &\Rightarrow aaAbB \\ &\Rightarrow aaabB \\ &\Rightarrow aaabb \end{aligned}$$

Parse Tree:



Ex. P+B

Derive $1 + 1 * 4$:

$$\begin{aligned} E &\to E + E \\ E &\to E - E \\ E &\to E * E \\ E &\to E \,/\, E \\ E &\to (\,E\,) \\ E &\to 0 \mid 1 \mid 2 \mid \cdots \mid 9 \end{aligned}$$

Def: A grammar is ________ if some string w has more than one parse tree.

Equivalently: A grammar is ________ if there is more than one left-most derivation for some string.