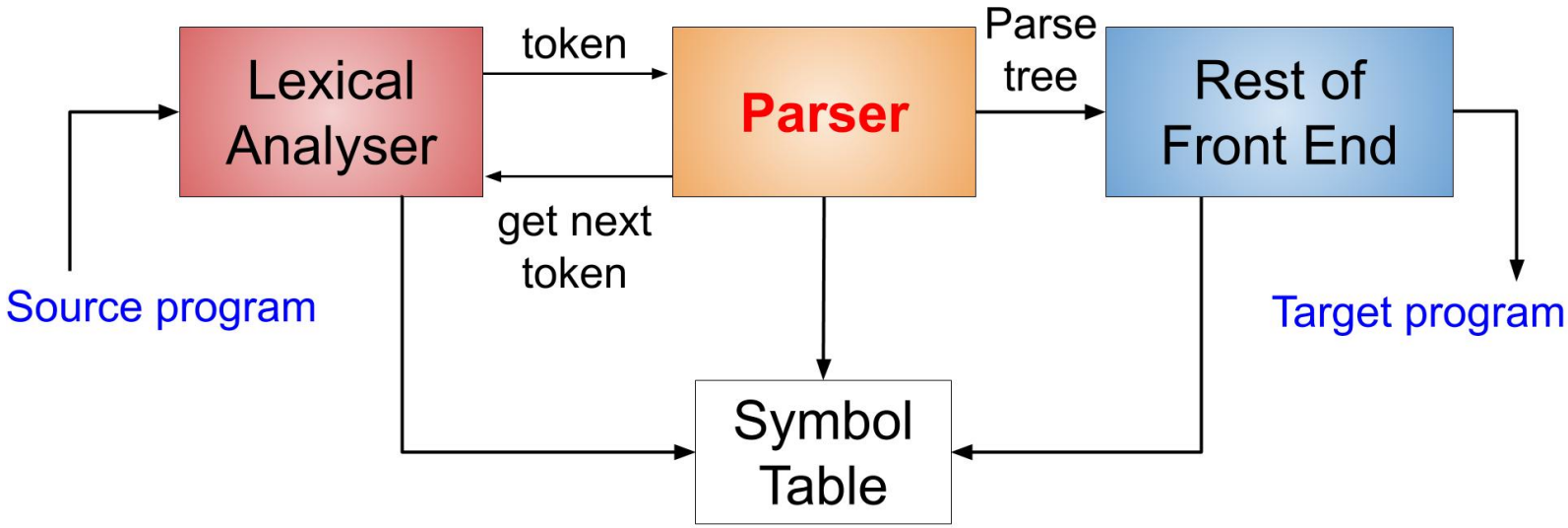
CSEN 1003: Compilers

Tutorial 3 - Context Free Grammars

16/2/2020 - 19/2/2020

Function of the Syntax Analyser



1. Determines if the input program is grammatical.
2. Generate error messages if the input is not grammatical.
3. Generate a parse tree.

For the Upcoming 5 Tutorials: Parsers!

Today's Plan

1. Context Free Grammars

Context Free Grammars

Definition

A context free grammar is a quadruple $\langle V, \Sigma, R, S \rangle$ where:

- V is a set of variables.
- Σ is the alphabet (terminals).
- R is a set of rules.
- S is the start variable.

Example

$$E \rightarrow E + E \mid E * E \mid T$$

$$T \rightarrow \texttt{id} \mid \texttt{num}$$

.

Derivations and Parse Trees

Example

Consider the following context-free grammar with the alphabet $\Sigma = \{+, *, a\}$:

$$S \quad \rightarrow \quad SS+ \mid SS* \mid \mathrm{a}.$$

Show that S derives `aa+a*`.

Derivations and Parse Trees

Example

Consider the following context-free grammar with the alphabet $\Sigma = \{+, *, a\}$:

$$S \quad \rightarrow \quad SS+ \mid SS* \mid \mathrm{a}.$$

Show that S derives aa+a*.

- The language of a Grammar G is the set:

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

.

Derivations and Parse Trees

Example

Consider the following context-free grammar with the alphabet $\Sigma = \{+, *, a\}$:

$$S \quad \rightarrow \quad SS+ \mid SS* \mid \mathrm{a}.$$

Show that S derives aa+a*.

- The language of a Grammar G is the set:

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

.

- A parser checks whether a program P is in the language of the grammar.

Grammar Ambiguity

Definition

A grammar G is ambiguous if there is some $w \in L(G)$ with more than one parse tree.

Example

Is the following grammar ambiguous?

$$S \rightarrow 1S0 \mid 1S \mid \varepsilon$$

Example

Is the following grammar ambiguous?

$$S \rightarrow aSbS \mid aS \mid \varepsilon$$

CFG Design

Example

Give a context-free grammar (CFG) for each of the following languages:

1. $L = \{\mathtt{a}^m\mathtt{b}^n\mathtt{c}^k \mid k = m + n \text{ and } m, n, k \geq 0\}$ over the alphabet $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$.
2. $L = \{\mathtt{a}^m\mathtt{b}^n \mid n \neq m\}$ over the alphabet $\Sigma = \{\mathtt{a}, \mathtt{b}\}$.
3. $L = \{w \mid w \text{ is a palindrome } \}$ over the alphabet $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$.

Grammar Correctness

Example

Consider the following grammar G:

$$\begin{aligned} S &\rightarrow \mathtt{a}S\mathtt{c} \mid T \\ T &\rightarrow \mathtt{b}T\mathtt{c} \mid \varepsilon \end{aligned}$$

and the language $L = \{\mathtt{a}^m\mathtt{b}^n\mathtt{c}^k \mid k = m + n \text{ and } m, n, k \geq 0\}$.
Prove that G is a correct grammar for L.