

Compiler Construction

Lecture 7: Syntactic Analysis II (Top-Down Parsing)

Thomas Noll

Lehrstuhl für Informatik 2
(Software Modeling and Verification)

RWTH Aachen University

`noll@cs.rwth-aachen.de`

`http://www-i2.informatik.rwth-aachen.de/i2/cc10/`

Winter semester 2010/11

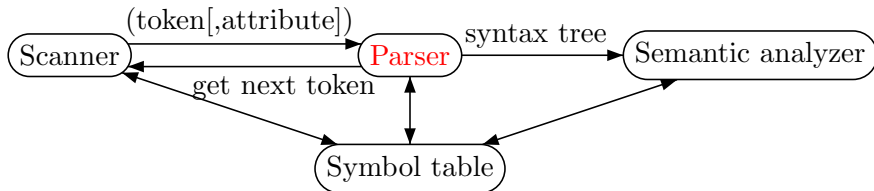
- 1 Repetition: Context-Free Grammars and Languages
- 2 Parsing Context-Free Languages
- 3 Nondeterministic Top-Down Parsing

Syntactic Analysis

Definition

The goal of **syntactic analysis** is to determine the syntactic structure of a program, given by a token sequence, according to a context-free grammar.

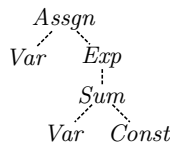
The corresponding program is called a **parser**:



Example: ... `x1` := `y2` + `1` ; ...

↓ Scanner

... (id, p_1)(gets,)(id, p_2)(plus,)(int, 1)(sem,) ... $\xrightarrow{\text{Parser}}$



Definition (Syntax of context-free grammars)

A **context-free grammar (CFG)** (over Σ) is a quadruple

$$G = \langle N, \Sigma, P, S \rangle$$

where

- N is a finite set of **nonterminal symbols**,
- Σ is a (finite) alphabet of **terminal symbols** (disjoint from N),
- P is a finite set of **production rules** of the form $A \rightarrow \alpha$ where $A \in N$ and $\alpha \in X^*$ for $X := N \cup \Sigma$, and
- $S \in N$ is a **start symbol**.

The set of all context-free grammars over Σ is denoted by CFG_{Σ} .

Remarks: as denotations we generally use

- $A, B, C, \dots \in N$ for nonterminal symbols
- $a, b, c, \dots \in \Sigma$ for terminal symbols
- $u, v, w, \dots \in \Sigma^*$ for terminal words
- $\alpha, \beta, \gamma, \dots \in X^*$ for **sentences**

Context-Free Grammars II

Context-free grammars generate context-free languages:

Definition (Semantics of context-free grammars)

Let $G = \langle N, \Sigma, P, S \rangle$ be a context-free grammar.

- The **derivation relation** $\Rightarrow \subseteq X^* \times X^*$ of G is defined by
$$\alpha \Rightarrow \beta \text{ iff there exist } \alpha_1, \alpha_2 \in X^*, A \rightarrow \gamma \in P$$
$$\text{such that } \alpha = \alpha_1 A \alpha_2 \text{ and } \beta = \alpha_1 \gamma \alpha_2.$$
- If in addition $\alpha_1 \in \Sigma^*$ or $\alpha_2 \in \Sigma^*$, then we write $\alpha \Rightarrow_l \beta$ or $\alpha \Rightarrow_r \beta$, respectively (**leftmost/rightmost** derivation).
- The **language generated by G** is given by
$$L(G) := \{w \in \Sigma^* \mid S \Rightarrow^* w\}.$$
- If a language $L \subseteq \Sigma^*$ is generated by some $G \in CFG_\Sigma$, then L is called **context free**. The set of all **context-free languages** over Σ is denoted by CFL_Σ .

Remark: obviously,

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

- 1 Repetition: Context-Free Grammars and Languages
- 2 Parsing Context-Free Languages
- 3 Nondeterministic Top-Down Parsing

The Word Problem for Context-Free Languages

Problem 7.1 (Word problem for context-free languages)

Given $G \in CFG_{\Sigma}$ and $w \in \Sigma^$, decide whether $w \in L(G)$
(and determine a corresponding syntax tree).*

Problem 7.1 (Word problem for context-free languages)

Given $G \in CFG_\Sigma$ and $w \in \Sigma^$, decide whether $w \in L(G)$
(and determine a corresponding syntax tree).*

This problem is **decidable** for arbitrary CFGs:

- (for CFGs in Chomsky Normal Form)
Using the **tabular method by Cocke, Younger, and Kasami**
("CYK Algorithm"; time complexity $\mathcal{O}(|w|^3)$ ($\mathcal{O}(|w|^2)$)))

Problem 7.1 (Word problem for context-free languages)

Given $G \in CFG_{\Sigma}$ and $w \in \Sigma^$, decide whether $w \in L(G)$ (and determine a corresponding syntax tree).*

This problem is **decidable** for arbitrary CFGs:

- (for CFGs in Chomsky Normal Form)
Using the **tabular method by Cocke, Younger, and Kasami** (“CYK Algorithm”; time complexity $\mathcal{O}(|w|^3)$ ($\mathcal{O}(|w|^2)$))
- Using the **predecessor method**:

$$w \in L(G) \iff S \in \text{pre}^*({w})$$

where $\text{pre}^*(M) := \{\alpha \in X^* \mid \alpha \Rightarrow^* \beta \text{ for some } \beta \in M\}$
(polynomial [non-linear] time complexity)

Goal: exploit the special syntactic structures as present in programming languages (usually: no ambiguities) to devise parsing methods which are based on **deterministic pushdown automata with linear space and time complexity**

Goal: exploit the special syntactic structures as present in programming languages (usually: no ambiguities) to devise parsing methods which are based on **deterministic pushdown automata with linear space and time complexity**

Two approaches:

Top-down parsing: construction of syntax tree from the **root towards the leafs**, representation as **leftmost derivation**

Bottom-up parsing: construction of syntax tree from the **leafs towards the root**, representation as (reversed) **rightmost derivation**

Goal: compact representation of left-/rightmost derivations by index sequences

Definition 7.2 (Leftmost/rightmost analysis)

Let $G = \langle N, \Sigma, P, S \rangle \in CFG_{\Sigma}$ where $P = \{\pi_1, \dots, \pi_p\}$.

- If $i \in [p]$, $\pi_i = A \rightarrow \gamma$, $w \in \Sigma^*$, and $\alpha \in X^*$, then we write

$$wA\alpha \xRightarrow{i}_l w\gamma\alpha \quad \text{and} \quad \alpha Aw \xRightarrow{i}_r \alpha\gamma w.$$

Leftmost/Rightmost Analysis I

Goal: compact representation of left-/rightmost derivations by index sequences

Definition 7.2 (Leftmost/rightmost analysis)

Let $G = \langle N, \Sigma, P, S \rangle \in CFG_{\Sigma}$ where $P = \{\pi_1, \dots, \pi_p\}$.

- If $i \in [p]$, $\pi_i = A \rightarrow \gamma$, $w \in \Sigma^*$, and $\alpha \in X^*$, then we write

$$wA\alpha \xRightarrow{i}_l w\gamma\alpha \quad \text{and} \quad \alpha Aw \xRightarrow{i}_r \alpha\gamma w.$$

- If $z = i_1 \dots i_n \in [p]^*$, we write $\alpha \xRightarrow{z}_l \beta$ if there exist

$\alpha_0, \dots, \alpha_n \in X^*$ such that $\alpha_0 = \alpha$, $\alpha_n = \beta$, and $\alpha_{j-1} \xRightarrow{i_j}_l \alpha_j$ for every $j \in [n]$ (analogously for $\xRightarrow{z}_r$).

Leftmost/Rightmost Analysis I

Goal: compact representation of left-/rightmost derivations by index sequences

Definition 7.2 (Leftmost/rightmost analysis)

Let $G = \langle N, \Sigma, P, S \rangle \in CFG_{\Sigma}$ where $P = \{\pi_1, \dots, \pi_p\}$.

- If $i \in [p]$, $\pi_i = A \rightarrow \gamma$, $w \in \Sigma^*$, and $\alpha \in X^*$, then we write

$$wA\alpha \xRightarrow{i}_l w\gamma\alpha \quad \text{and} \quad \alpha Aw \xRightarrow{i}_r \alpha\gamma w.$$

- If $z = i_1 \dots i_n \in [p]^*$, we write $\alpha \xRightarrow{z}_l \beta$ if there exist $\alpha_0, \dots, \alpha_n \in X^*$ such that $\alpha_0 = \alpha$, $\alpha_n = \beta$, and $\alpha_{j-1} \xRightarrow{i_j}_l \alpha_j$ for every $j \in [n]$ (analogously for $\xRightarrow{z}_r$).
- An index sequence $z \in [p]^*$ is called a **leftmost analysis** (**rightmost analysis**) of α if $S \xRightarrow{z}_l \alpha$ ($S \xRightarrow{z}_r \alpha$), respectively.

Example 7.3

Grammar for arithmetic expressions:

$$\begin{aligned} G_{AE} : \quad E &\rightarrow E+T \mid T & (1, 2) \\ T &\rightarrow T*F \mid F & (3, 4) \\ F &\rightarrow (E) \mid a \mid b & (5, 6, 7) \end{aligned}$$

Leftmost/Rightmost Analysis

Example 7.3

Grammar for arithmetic expressions:

$$\begin{aligned} G_{AE} : \quad E &\rightarrow E+T \mid T & (1, 2) \\ T &\rightarrow T*F \mid F & (3, 4) \\ F &\rightarrow (E) \mid a \mid b & (5, 6, 7) \end{aligned}$$

Leftmost derivation of $(a)*b$:

$$\begin{array}{ccccccc} E & \xRightarrow{2}_l & T & \xRightarrow{3}_l & T*F & \xRightarrow{4}_l & F*F & \xRightarrow{5}_l & (E)*F \\ & \xRightarrow{2}_l & (T)*F & \xRightarrow{4}_l & (F)*F & \xRightarrow{6}_l & (a)*F & \xRightarrow{7}_l & (a)*b \end{array}$$

$\Rightarrow$ leftmost analysis: 23452467

Leftmost/Rightmost Analysis

Example 7.3

Grammar for arithmetic expressions:

$$\begin{aligned} G_{AE} : \quad E &\rightarrow E+T \mid T & (1, 2) \\ T &\rightarrow T*F \mid F & (3, 4) \\ F &\rightarrow (E) \mid a \mid b & (5, 6, 7) \end{aligned}$$

Leftmost derivation of $(a)*b$:

$$\begin{array}{ccccccc} E & \xRightarrow{2}_l & T & \xRightarrow{3}_l & T*F & \xRightarrow{4}_l & F*F & \xRightarrow{5}_l & (E)*F \\ & \xRightarrow{2}_l & (T)*F & \xRightarrow{4}_l & (F)*F & \xRightarrow{6}_l & (a)*F & \xRightarrow{7}_l & (a)*b \end{array}$$

$\Rightarrow$ leftmost analysis: 23452467

Rightmost derivation of $(a)*b$:

$$\begin{array}{ccccccc} E & \xRightarrow{2}_r & T & \xRightarrow{3}_r & T*F & \xRightarrow{7}_r & T*b & \xRightarrow{4}_r & F*b \\ & \xRightarrow{5}_r & (E)*b & \xRightarrow{2}_r & (T)*b & \xRightarrow{4}_r & (F)*b & \xRightarrow{6}_r & (a)*b \end{array}$$

$\Rightarrow$ rightmost analysis: 23745246

General assumption in the following: every grammar is reduced

Definition 7.4 (Reduced CFG)

A grammar $G = \langle N, \Sigma, P, S \rangle \in CFG_{\Sigma}$ is called **reduced** if for every $A \in N$ there exist $\alpha, \beta \in X^*$ and $w \in \Sigma^*$ such that

$$\begin{aligned} S &\Rightarrow^* \alpha A \beta \quad (A \text{ **reachable**}) \text{ and} \\ A &\Rightarrow^* w \quad (A \text{ **productive**}). \end{aligned}$$

- 1 Repetition: Context-Free Grammars and Languages
- 2 Parsing Context-Free Languages
- 3 Nondeterministic Top-Down Parsing

Approach:

- 1 Given $G \in CFG_{\Sigma}$, construct a **nondeterministic pushdown automaton** (PDA) which accepts $L(G)$ and which additionally computes corresponding leftmost derivations (similar to the proof of “ $L(CFG_{\Sigma}) \subseteq L(PDA_{\Sigma})$ ”)
 - input alphabet: Σ
 - pushdown alphabet: X
 - output alphabet: $[p]$
 - state set: omitted

Approach:

- ① Given $G \in CFG_\Sigma$, construct a **nondeterministic pushdown automaton** (PDA) which accepts $L(G)$ and which additionally computes corresponding leftmost derivations (similar to the proof of “ $L(CFG_\Sigma) \subseteq L(PDA_\Sigma)$ ”)
 - input alphabet: Σ
 - pushdown alphabet: X
 - output alphabet: $[p]$
 - state set: omitted
- ② **Remove nondeterminism** by allowing **lookahead** on the input:
 $G \in LL(k)$ iff $L(G)$ recognizable by deterministic PDA with lookahead of k symbols

The Nondeterministic Top-Down Automaton I

Definition 7.5 (Nondeterministic top-down parsing automaton)

Let $G = \langle N, \Sigma, P, S \rangle \in CFG_{\Sigma}$. The **nondeterministic top-down parsing automaton** of G , $NTA(G)$, is defined by the following components.

- **Input alphabet:** Σ
- **Pushdown alphabet:** X
- **Output alphabet:** $[p]$
- **Configurations:** $\Sigma^* \times X^* \times [p]^*$ (top of pushdown to the left)
- **Transitions** for $w \in \Sigma^*$, $\alpha \in X^*$, and $z \in [p]^*$:
 - expansion steps: if $\pi_i = A \rightarrow \beta$, then $(w, A\alpha, z) \vdash (w, \beta\alpha, zi)$
 - matching steps: for every $a \in \Sigma$, $(aw, a\alpha, z) \vdash (w, \alpha, z)$
- **Initial configuration** for $w \in \Sigma^*$: (w, S, ε)
- **Final configurations:** $\{\varepsilon\} \times \{\varepsilon\} \times [p]^*$

The Nondeterministic Top-Down Automaton I

Definition 7.5 (Nondeterministic top-down parsing automaton)

Let $G = \langle N, \Sigma, P, S \rangle \in CFG_{\Sigma}$. The **nondeterministic top-down parsing automaton** of G , $NTA(G)$, is defined by the following components.

- **Input alphabet:** Σ
- **Pushdown alphabet:** X
- **Output alphabet:** $[p]$
- **Configurations:** $\Sigma^* \times X^* \times [p]^*$ (top of pushdown to the left)
- **Transitions** for $w \in \Sigma^*$, $\alpha \in X^*$, and $z \in [p]^*$:
 - expansion steps: if $\pi_i = A \rightarrow \beta$, then $(w, A\alpha, z) \vdash (w, \beta\alpha, zi)$
 - matching steps: for every $a \in \Sigma$, $(aw, a\alpha, z) \vdash (w, \alpha, z)$
- **Initial configuration** for $w \in \Sigma^*$: (w, S, ε)
- **Final configurations:** $\{\varepsilon\} \times \{\varepsilon\} \times [p]^*$

Remark: $NTA(G)$ is nondeterministic iff G contains $A \rightarrow \beta \mid \gamma$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned} G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\ T &\rightarrow T*F \mid F & (3, 4) \\ F &\rightarrow (E) \mid a \mid b & (5, 6, 7) \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned} G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\ T &\rightarrow T*F \mid F & (3, 4) \\ F &\rightarrow (E) \mid a \mid b & (5, 6, 7) \end{aligned}$$

Leftmost analysis of $(a)*b$:

$((a)*b, E, \varepsilon)$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned} G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\ T &\rightarrow T * F \mid F & (3, 4) \\ F &\rightarrow (E) \mid a \mid b & (5, 6, 7) \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned} &((a)*b, E, \varepsilon) \\ \vdash &((a)*b, T, 2) \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T*F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 &((a)*b, E, \varepsilon) \\
 \vdash &((a)*b, T, 2) \\
 \vdash &((a)*b, T*F, 23)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned} G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\ T &\rightarrow T*F \mid F & (3, 4) \\ F &\rightarrow (E) \mid a \mid b & (5, 6, 7) \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned} &((a)*b, E, \varepsilon) \\ \vdash &((a)*b, T, 2) \\ \vdash &((a)*b, T*F, 23) \\ \vdash &((a)*b, F*F, 234) \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T*F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 &((a)*b, E, \varepsilon) \\
 \vdash &((a)*b, T, 2) \\
 \vdash &((a)*b, T*F, 23) \\
 \vdash &((a)*b, F*F, 234) \\
 \vdash &((a)*b, (E)*F, 2345)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : \quad & \textcolor{red}{E} \rightarrow E + T \mid \textcolor{red}{T} & (1, 2) \\
 & T \rightarrow T * F \mid F & (3, 4) \\
 & F \rightarrow (\textcolor{blue}{E}) \mid \textcolor{blue}{a} \mid \textcolor{blue}{b} & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(\textcolor{blue}{a}) * \textcolor{blue}{b}$:

$$\begin{aligned}
 & (\textcolor{blue}{(a)} * \textcolor{blue}{b}, E, \varepsilon) \\
 \vdash & (\textcolor{blue}{(a)} * \textcolor{blue}{b}, T, 2) \\
 \vdash & (\textcolor{blue}{(a)} * \textcolor{blue}{b}, T * F, 23) \\
 \vdash & (\textcolor{blue}{(a)} * \textcolor{blue}{b}, F * F, 234) \\
 \vdash & (\textcolor{blue}{(a)} * \textcolor{blue}{b}, (\textcolor{blue}{E}) * F, 2345) \\
 \vdash & (\textcolor{blue}{a} * \textcolor{blue}{b}, \textcolor{red}{E} * F, 2345)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T * F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 &((a)*b, E, \varepsilon) \\
 \vdash &((a)*b, T, 2) \\
 \vdash &((a)*b, T * F, 23) \\
 \vdash &((a)*b, F * F, 234) \\
 \vdash &((a)*b, (E) * F, 2345) \\
 \vdash &((a)*b, E * F, 2345) \\
 \vdash &((a)*b, T * F, 23452)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T*F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 &((a)*b, E, \varepsilon) \\
 \vdash &((a)*b, T, 2) \\
 \vdash &((a)*b, T*F, 23) \\
 \vdash &((a)*b, F*F, 234) \\
 \vdash &((a)*b, (E)*F, 2345) \\
 \vdash &((a)*b, E)*F, 2345) \\
 \vdash &((a)*b, T)*F, 23452) \\
 \vdash &((a)*b, F)*F, 234524)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T * F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 &((a)*b, E, \varepsilon) \\
 \vdash &((a)*b, T, 2) \\
 \vdash &((a)*b, T * F, 23) \\
 \vdash &((a)*b, F * F, 234) \\
 \vdash &((a)*b, (E) * F, 2345) \\
 \vdash &((a)*b, E) * F, 2345) \\
 \vdash &((a)*b, T) * F, 23452) \\
 \vdash &((a)*b, F) * F, 234524) \\
 \vdash &((a)*b, a) * F, 2345246)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T * F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 &((a)*b, E, \varepsilon) \\
 \vdash &((a)*b, T, 2) \\
 \vdash &((a)*b, T * F, 23) \\
 \vdash &((a)*b, F * F, 234) \\
 \vdash &((a)*b, (E) * F, 2345) \\
 \vdash &((a)*b, E) * F, 2345) \\
 \vdash &((a)*b, T) * F, 23452) \\
 \vdash &((a)*b, F) * F, 234524) \\
 \vdash &((a)*b, a) * F, 2345246) \\
 \vdash &(() * b,) * F, 2345246)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T * F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 & \left((a)*b, E, \varepsilon \right) \\
 \vdash & \left((a)*b, T, 2 \right) \\
 \vdash & \left((a)*b, T * F, 23 \right) \\
 \vdash & \left((a)*b, F * F, 234 \right) \\
 \vdash & \left((a)*b, (E) * F, 2345 \right) \\
 \vdash & \left(a * b, E) * F, 2345 \right) \\
 \vdash & \left(a * b, T) * F, 23452 \right) \\
 \vdash & \left(a * b, F) * F, 234524 \right) \\
 \vdash & \left(a * b, a) * F, 2345246 \right) \\
 \vdash & \left() * b,) * F, 2345246 \right) \\
 \vdash & \left(* b, * F, 2345246 \right)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T*F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 &((a)*b, E, \varepsilon) \\
 \vdash &((a)*b, T, 2) \\
 \vdash &((a)*b, T*F, 23) \\
 \vdash &((a)*b, F*F, 234) \\
 \vdash &((a)*b, (E)*F, 2345) \\
 \vdash &(a)*b, E)*F, 2345) \\
 \vdash &(a)*b, T)*F, 23452) \\
 \vdash &(a)*b, F)*F, 234524) \\
 \vdash &(a)*b, a)*F, 2345246) \\
 \vdash &())*b,)*F, 2345246) \\
 \vdash &()*b, *F, 2345246) \\
 \vdash &()b, F, 2345246)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T * F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 & \left((a)*b, E, \varepsilon \right) \\
 \vdash & \left((a)*b, T, 2 \right) \\
 \vdash & \left((a)*b, T * F, 23 \right) \\
 \vdash & \left((a)*b, F * F, 234 \right) \\
 \vdash & \left((a)*b, (E) * F, 2345 \right) \\
 \vdash & \left(a)*b, E) * F, 2345 \right) \\
 \vdash & \left(a)*b, T) * F, 23452 \right) \\
 \vdash & \left(a)*b, F) * F, 234524 \right) \\
 \vdash & \left(a)*b, a) * F, 2345246 \right) \\
 \vdash & \left() * b,) * F, 2345246 \right) \\
 \vdash & \left(* b, * F, 2345246 \right) \\
 \vdash & \left(b, F, 2345246 \right) \\
 \vdash & \left(b, b, 23452467 \right)
 \end{aligned}$$

Example 7.6

Grammar for
arithmetic expressions
(cf. Example 7.3):

$$\begin{aligned}
 G_{AE} : E &\rightarrow E+T \mid T & (1, 2) \\
 T &\rightarrow T * F \mid F & (3, 4) \\
 F &\rightarrow (E) \mid a \mid b & (5, 6, 7)
 \end{aligned}$$

Leftmost analysis of $(a)*b$:

$$\begin{aligned}
 & \left((a)*b, E, \varepsilon \right) \\
 \vdash & \left((a)*b, T, 2 \right) \\
 \vdash & \left((a)*b, T * F, 23 \right) \\
 \vdash & \left((a)*b, F * F, 234 \right) \\
 \vdash & \left((a)*b, (E) * F, 2345 \right) \\
 \vdash & \left(a * b, E) * F, 2345 \right) \\
 \vdash & \left(a * b, T) * F, 23452 \right) \\
 \vdash & \left(a * b, F) * F, 234524 \right) \\
 \vdash & \left(a * b, a) * F, 2345246 \right) \\
 \vdash & \left() * b,) * F, 2345246 \right) \\
 \vdash & \left(* b, * F, 2345246 \right) \\
 \vdash & \left(b, F, 2345246 \right) \\
 \vdash & \left(b, b, 23452467 \right) \\
 \vdash & \left(\varepsilon, \varepsilon, \textcolor{red}{23452467} \right)
 \end{aligned}$$

Theorem 7.7 (Correctness of $\text{NTA}(G)$)

Let $G = \langle N, \Sigma, P, S \rangle \in \text{CFG}_\Sigma$ and $\text{NTA}(G)$ as before. Then, for every $w \in \Sigma^$ and $z \in [p]^*$,*

$$(w, S, \varepsilon) \vdash^* (\varepsilon, \varepsilon, z) \quad \text{iff} \quad z \text{ is a leftmost analysis of } w$$

Theorem 7.7 (Correctness of $\text{NTA}(G)$)

Let $G = \langle N, \Sigma, P, S \rangle \in \text{CFG}_\Sigma$ and $\text{NTA}(G)$ as before. Then, for every $w \in \Sigma^*$ and $z \in [p]^*$,

$(w, S, \varepsilon) \vdash^* (\varepsilon, \varepsilon, z)$ iff z is a leftmost analysis of w

Proof.

$\implies$ (soundness): see exercises

$\impliedby$ (completeness): on the board

