Answers for Chapter 2: A SIMPLE SYNTAX-DIRECTED TRANSLATOR

hmsjwzb@zoho.com

June 8, 2025

2.2.7 Exercises for Section 2.2

Exercise 2.2.1 : Consider the context-free grammar

$$S \rightarrow S\,S+ \mid S\,S* \mid a$$

a) Show how the string $aa + a*$ can be generated by this grammar

b) Construct a parse tree for this string

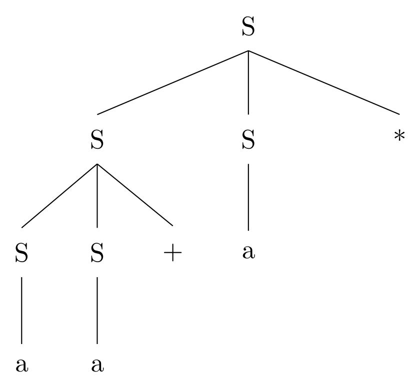
c) What language does this grammar generate? Justify your answer.

Answer:

a) Apply the following production in sequence.

$$\begin{aligned}
&S \rightarrow S_1\, S_2\, * \\
&S_1 \rightarrow S_3\, S_4 + \\
&S_3 \rightarrow a \\
&S_4 \rightarrow a \\
&S_2 \rightarrow a
\end{aligned}$$

b) The parse tree for it



c) This language generate post-fix expression with + and * as operator, a as operand

Exercise 2.2.2: What language is generated by the following grammars? In each case justify your answer.

a) $S \to 0\ S\ 1 \mid 0\ 1$

b) $S \to +\ SS \mid -\ SS \mid a$

c) $S \to S\ (\ S\ )\ S \mid \epsilon$

d) $S \to a\ S\ b\ S \mid b\ S\ a\ S \mid \epsilon$

e) $S \to a \mid S\ +\ S \mid S\ S \mid S\ * \mid (\ S\ )$

Answer:

a) A language contain first half as 0, second half as 1

b) A pre-fix expression with + and - as operator and a as operator

c) matched Parentheses

d) A language with same number of a and b

e) An infix expression with + and * as operator and a as operand

Exercise 2.2.3: Which of the grammars in Exercise 2.2.2 are ambiguous?
Answer:

c) $S \rightarrow S\ (\ S\ )\ S \mid \epsilon$ is ambiguous, the input ()() has two corresponding grammar tree.

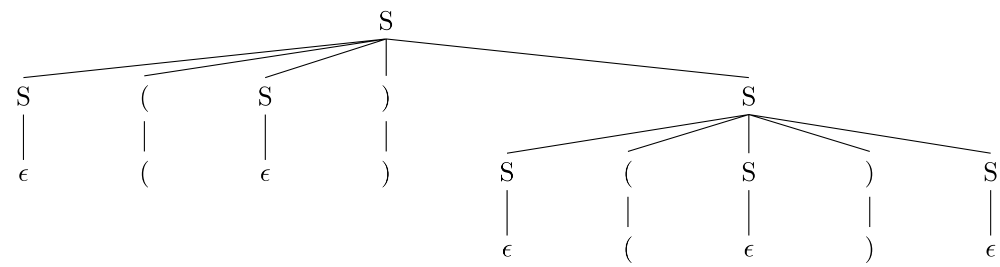


Figure 1: grammar tree 1

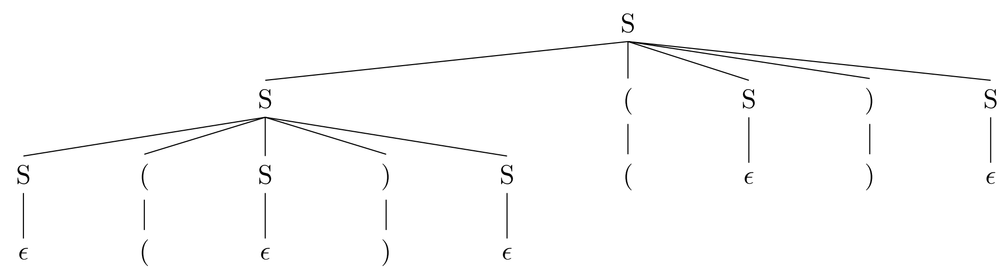


Figure 2: grammar tree 2

d) $S \rightarrow a\ S\ b\ S \mid b\ S\ a\ S \mid \epsilon$ is ambiguous

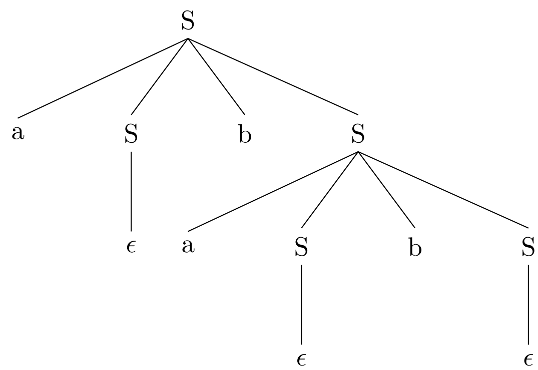


Figure 3: Derivation for abab

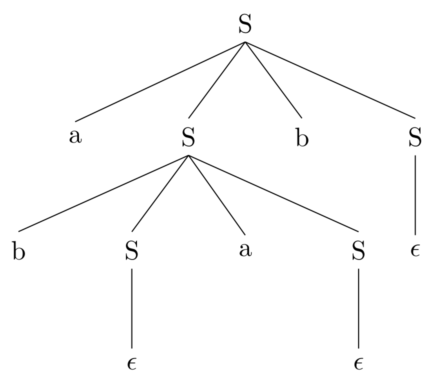


Figure 4: Derivation for abab

Exercise 2.2.4: Construct unambiguous context-free grammars for each of the following languages. In each case show that your grammar is correct.

a) Arithmetic expressions in postfix notation.

b) Left-associative lists of identifiers separated by commas.

c) Right-associative lists of identifiers separated by commas.

d) Arithmetic expressions of integers and identifiers with the four binary operators +,-,*,/.

!e) Add unary plus and minus to arithmetic operators of (d)

Answer:

a) Arithmetic expressions in postfix notation.

$$\begin{aligned} expr &\rightarrow expr\ expr\ operator \\ &\quad \mid operand\ operand\ operator \\ operator &\rightarrow +\ -\ *\ \div \\ operand &\rightarrow [1-9]\ remain \\ remain &\rightarrow [0-9]\ remain \end{aligned}$$

b) Left-associative lists of identifiers separated by commas.

$$list\ \rightarrow list,\ \texttt{ident}$$

c) Right-associative lists of identifiers separated by commas.

$$list\ \rightarrow \texttt{ident},\ list$$

d) Arithmetic expressions of integers and identifiers with four binary operator $+, -, *, /$

$$\begin{aligned} expr &\rightarrow expr\ operand\ integer \\ &\quad \mid integer \\ operand &\rightarrow + \mid - \mid * \mid\ / \\ integer &\rightarrow msd\ remain \\ msd &\rightarrow [1..9] \\ remain &\rightarrowtail [0..9]^*\ remain \end{aligned}$$

!e) Add unary plus and minus arithmetic operators of (d)

$$\begin{aligned} expr &\rightarrow expr\ operand\ unary \\ &\mid unary \\ unary &\rightarrow -\ integer \\ &\mid +\ integer \\ operand &\rightarrow + \mid - \mid * \mid\ / \\ integer &\rightarrow msd\ remain \\ msd &\rightarrow [1..9] \\ remain &\rightarrowtail [0..9]^*\ remain \end{aligned}$$

Exercise 2.2.5:

a) Show that all binary strings generated by the following grammar have values divisible by 3. *Hint.* Use induction on the number of nodes in parse tree.

$$num \ \rightarrow\ 11 \mid 1001 \mid num\ 0 \mid num\ num$$

b) Does the grammar generate all binary strings with values divisible by 3?

a) Show that all binary strings generated by the following grammar have values divisible by 3. *Hint.* Use induction on the number of nodes in parse tree.

$$num \ \rightarrow\ 11 \mid 1001 \mid num\ 0 \mid num\ num$$

(a) The binary strings are divisible by 3

Proof.

$$num \ \rightarrow\ 11 \mid 1001$$

The show terminal symbol generated by num are divisible by 3.
A sum of every digit of a number is divisible 3, then the number is divisible by 3.

As we can see from above grammar tree, the sum of all digits of num is divisible by 3. So it should be divisible by 3. ☐

(b) Does the grammar generate all binary strings with values divisible by 3?
Yes, it generate all binary strings with values divisible by 3.

Exercise 2.2.6: Construct a context-free grammar for roman numerals

- via wikipedia, we can categorize the single roman numerals into 4 groups:

$$I,\ II,\ III \mid I\ V \mid V,\ V\ I,\ V\ II,\ V\ III \mid I\ X$$

then get the production:

$$\begin{aligned} digit &\rightarrow smallDigit \mid I\ V \mid V\ smallDigit \mid I\ X \\ smallDigit &\rightarrow I \mid II \mid III \mid \epsilon \end{aligned}$$

- find a simple way to map roman to arabic numerals. For example

(a) $XII => X,\ II => 10 + 2 => 12$

(b) $CXCIX => C, XC, IX => 100 + 90 + 9 => 199$

(c) $MDCCCLXXX => M, DCCC, LXXX => 1000 + 800 + 80 => 1880$

- via the upper two rules, derive the production

$$\begin{aligned} romanNum &\rightarrow thround\ hundred\ ten\ digit \\ throusand &\rightarrow M \mid MM \mid MMM \mid \epsilon \\ hundred &\rightarrow smallHundred \mid C\ D \mid D\ smallHundred \mid C\ M \\ smallHundred &\rightarrow C \mid CC \mid CCC \mid \epsilon \\ ten &\rightarrow smallTen \mid X\ L \mid L\ smallTen \mid X\ C \\ smallTen &\rightarrow X \mid XX \mid XXX \mid \epsilon \\ digit &\rightarrow smallDight \mid I\ V \mid V\ smallDigit \mid I\ X \\ smallDigit &\rightarrow I \mid II \mid III \mid \epsilon \end{aligned}$$