Programming Languages: Syntax

Sudarshan S. Chawathe

2025-09-19

School of Computing and Information Science
& Climate Change Institute
University of Maine

Announcements and Reminders

- Sound and visuals check.
- Main online resource: Class Web site:
  - http://chaw.eip10.org/cos301/
  - also linked from my Web page, etc.
  - Includes (is) syllabus.
- Brightspace for some things only.
  - discussion forum.
  - homework and other submissions.
- Homework HW01 due tonight.
- Rules for use of Gen AI.
- PLY: Install, study examples, modify, experiment, etc.
  - Pointer in syllabus.

Plan for today

- Material from Chapter 2 of the textbook.
  - (Continuing) Theory of CFGs etc.
  - (review) Derivations (2.4).
  - *Parse trees* (2.5) and *abstract syntax trees (AST)* (2.6).
- (Review) Bigger picture question (related to homework):
  - How to implement a simple language like:
    - x = 5 + 3
    - y = 48 / (4 * 4)
    - z = x + 2 * y
    - etc.
- Class Exercise 2

(Review) CFG formally

- $G = (N, T, P, S)$
  - N: a set of symbols (*nonterminals*)
  - T: another set of symbols (*terminals*)
    - $N \cap T = \emptyset$
  - P: set of *productions*
    - each of form $n \rightarrow \alpha$ where
    - where $n \in N$ and $\alpha \in \{N \cup T\}^*$
  - $S \in N$: special nonterminal called *start*

(Review) CFG for infix expressions

- example from the textbook

(Review) Derivations

- Main question: Can a given string of tokens (terminals) be generated by a given grammar?
  - If so, how? Show the steps starting with the start symbol.
  - Is the derivation unique?
  - Is the leftmost derivation unique?
- Switch to example and practice in textbook.

Parse Trees

- Informally, a tree that has S as root and the children of each node are the items on the RHS of the rule that was used to replace the corresponding nonterminal. (Leaves correspond to terminals.)
- Example for infix expressions.
- Switch to practice problems.

Abstract Syntax Trees

- closely related to, but different from, parse trees.
- abstract away unimportant details such as order in which a sequence of nonterminals is expanded.
- Two main changes:
  - Nonterminal nodes are replaced by *corresponding* parts of the input sentence.
  - Unit productions are collapsed.
- Example for infix expressions.
- Practice problems in textbook.

Summary

- (E)BNF, CFGs, derivations, parse trees, ASTs.
- CFG formal definition.
- Practical examples of CFGs.
- Bigger picture:
  - Process source code into AST
  - then we can interpret or compile it etc.
- Before next class:
  - Read at least material due (syllabus).
  - Don't fall behind!
  - Experiment with PLY (Python Lex-Yacc):
    - http://www.dabeaz.com/ply/
- Class Web site: http://chaw.eip10.org/cos301/