Syntax and semantics
Context-Free Languages

1 Learning Objectives

1. Give a formal description of a CFG.
2. Construct a derivation of a given string.
3. Draw a parse tree of a derivation.
4. Define a CFG of a given context-free language.
5. Convert a DFA into an equivalent CFG.
6. Prove that a grammar is ambiguous.
7. Convert a CFG into an equivalent CFG in Chomsky normal form.
8. Construct a PDA that recognizes a language.
9. Construct a PDA equivalent to a given CFG.
10. Use Pumping Lemma to prove that a language is not context-free.

2 Readings

Sipser's book:
Part I – Automata and Languages, section 2. Context-Free Languages

3 Homework - Exercises

Exercise 1. Consider the languages

$$L_1 = \{a^i b^j c^k \mid i, j, k \geq 0, i = j \text{ eller } j = k\}\,.$$

$$L_2 = \{a^i b^j c^k \mid i, j, k \geq 0, i = j \text{ or } j = k\}$$

1. Define context-free grammars for each of the two languages. It is enough to set the rules of grammars.
2. Prove that the grammars you have constructed are ambiguous. Do not forget to provide an example. *Hint: aabbcc.*

Exercise 2. Consider the context-free grammar $G = (V, \Sigma, R, S)$:

$$\begin{aligned}
V &= \{S, A\} \\
\Sigma &= \{\mathtt{0}, \mathtt{1}\} \\
R &= \{S \to \mathtt{00}S, S \to A\mathtt{1}S, S \to A, A \to \varepsilon\}
\end{aligned}$$

1. Show that `001` $\in L(G)$ by constructing a derivation for this string.

2. Show that `1100` $\in L(G)$ by constructing a derivation for this string.

Exercise 3. Consider the following context-free grammar

$$\begin{aligned}
A &\to BAB \mid B \mid \varepsilon \\
B &\to \mathtt{00}\dagger \mid \varepsilon
\end{aligned}$$

Convert the grammar into an equivalent grammar in Chomsky-normal form.

Exercise 4. Consider the following context-free grammar with start variable S_1:

$$\begin{aligned}
S_1 &\to AS \mid AA \mid \varepsilon \mid AS_1 \\
S &\to ASA \mid bb \mid \varepsilon \\
A &\to AA \mid S \mid c
\end{aligned}$$

Is this grammar in Chomsky-normal form? If yes, explain why. If no, explain why it is not, and circle in the rules above all the parts where you think that there are problems.

Exercise 5. Consider the languages defined bellow

$$L_1 = \{a^i b^j \mid i \leq j\}$$

$$L_2 = \{w \in \{a, b\}^* \mid w \text{ contains the same number of a's and b's}\}$$

$$L_3 = \{a^k b^{2k} \mid k \geq 0\}$$

Prove that they are context-free.

Exercise 6. Consider the languages

$$L_1 = \{x \in \{(,)\}^* \mid x \text{ is a string of balanced parentheses }\}$$

$$L_2 = \{a^i b^i c^k \mid i, j, k \geq 0; k = i + j\}$$

Construct PDAs that recognize each of these languages.

Exercise 7. Consider the context-free grammar G:

$$\begin{aligned} S &\to \mathtt{11}S \mid \mathtt{0}A\mathtt{1}S\mathtt{00} \mid \mathtt{10} \\ A &\to \mathtt{1}S \mid AA \mid \mathtt{00} \end{aligned}$$

1. Find three strings in $L(G)$.
2. Define a pushdown automaton that is equivalent to G.

Exercise 8. Consider the language

$$L_1 = \{a^n \mid n \text{ is a prime number}\} .$$

Prove, by using the Pumping lemma, that L_1 is not context-free.

Exercise 9. Consider the language

$$L_2 = \{w \in \{a, b, c\}^* \mid w \text{ have equal number of occurrences of } a,\ b \text{ and } c\} .$$

Prove, by using the Pumping lemma, that L_2 is not context-free.