Homework 5

1. Give context-free grammars that generate the following languages.

   (a) $\{ w \in \{0,1\}^* \mid w \text{ contains at least three 1s} \}$

   (b) $\{ w \in \{0,1\}^* \mid w = w^{\mathcal{R}} \text{ and } |w| \text{ is even} \}$

   (c) $\{ w \in \{0,1\}^* \mid \text{the length of } w \text{ is odd and the middle symbol is } 0 \}$

   (d) $\{ a^i b^j c^k \mid i,j,k \geq 0, \text{ and } i = j \text{ or } i = k \}$

   (e) $\{ a^i b^j c^k \mid i,j,k \geq 0 \text{ and } i + j = k \}$

   (f) $\{ a^i b^j c^k \mid i,j,k \geq 0 \text{ and } i + k = j \}$. [Hint: use problem 3b.]

   (g) $\emptyset$

   (h) The language A of strings of properly balanced left and right brackets: every left bracket can be paired with a unique subsequent right bracket, and every right bracket can be paired with a unique preceding left bracket. Moreover, the string between any such pair has the same property. For example, $[\,]\,[\,[\,[\,]\,[\,]\,]\,[\,]\,] \in A$.

2. Let $T = \{ 0,\ 1,\ (,\ ),\ \cup,\ ^*,\ \emptyset,\ e \}$. We may think of T as the set of symbols used by regular expressions over the alphabet $\{0,1\}$; the only difference is that we use e for symbol ε, to avoid potential confusion in what follows.

   (a) Your task is to design a CFG G with set of terminals T that generates exactly the regular expressions with alphabet $\{0,1\}$.

   (b) Using your CFG G, give a derivation and the corresponding parse tree for the string $(0 \cup (10)^* 1)^*$.

3. (a) Suppose that language A_1 has a context-free grammar $G_1 = (V_1, \Sigma, R_1, S_1)$, and language A_2 has a context-free grammar $G_2 = (V_2, \Sigma, R_2, S_2)$, where, for $i = 1, 2$, V_i is the set of variables, R_i is the set of rules, and S_i is the start variable for CFG G_i. The CFGs have the same set of terminals Σ. Assume that $V_1 \cap V_2 = \emptyset$. Define another CFG $G_3 = (V_3, \Sigma, R_3, S_3)$ with $V_3 = V_1 \cup V_2 \cup \{S_3\}$, where $S_3 \notin V_1 \cup V_2$, and $R_3 = R_1 \cup R_2 \cup \{ S_3 \to S_1,\ S_3 \to S_2 \}$. Argue that G_3 generates the language $A_1 \cup A_2$. Thus, conclude that the class of context-free languages is closed under union.

   (b) Prove that the class of context-free languages is closed under concatenation.

   (c) Prove that the class of context-free languages is closed under Kleene-star.

4. Convert the following CFG into an equivalent CFG in Chomsky normal form, using the procedure given in Theorem 2.9.

$$\begin{aligned} S &\rightarrow BSB \mid B \mid \varepsilon \\ B &\rightarrow \mathsf{00} \mid \varepsilon \end{aligned}$$

5. Consider the CFG $G = (V, \Sigma, R, S)$, where $V = \{ S \}$ is the set of variables with S as the starting variable, alphabet $\Sigma = \{ +, -, \times, /, (,), \mathsf{0}, \mathsf{1}, \mathsf{2}, \ldots, \mathsf{9} \}$, and rules R as

$$S \rightarrow S + S \mid S - S \mid S \times S \mid S/S \mid (S) \mid -S \mid \mathsf{0} \mid \mathsf{1} \mid \cdots \mid \mathsf{9}$$

The CFG G generates the language $L(G)$ of some types of simple arithmetic expressions.

(a) Consider the strings $---\mathsf{5}$ and $\mathsf{2}+--\mathsf{4}$. Give derivations showing that each string belongs to $L(G)$.

(b) Suppose that we want to disallow such strings. Give another CFG that achieves this. More specifically, strings such as $\mathsf{2}-\mathsf{3}$, $\mathsf{2}+-\mathsf{3}$ and $\mathsf{2}--\mathsf{3}$ are allowed, but not $\mathsf{2}+--\mathsf{3}$ nor $\mathsf{2}---\mathsf{3}$.