Tutorial 6
Language and Automata Theory

Context-free Languages and Push-down Automata

1. What context-free languages will be generated by the following two (separate) context-free grammars, $G = (\{S, A, B\}, \{a, b, c\}, P, S)$, where P consists of the following production rules?

   (a)
   $$S \longrightarrow ASB \mid \varepsilon$$
   $$A \longrightarrow a$$
   $$B \longrightarrow bb \mid b$$

   (b)
   $$S \longrightarrow abScB \mid \varepsilon$$
   $$B \longrightarrow bB \mid b$$

2. Define the context-free grammars for the following context-free languages. Are your defined CFGs ambiguous / non-ambiguous?

   - $L_{2a} = \{a^i b^j c^k \mid i, j, k \geq 0,\ \text{and } i = j \text{ or } i = k\}$
   - $L_{2b} = \{a^i b^j c^k \mid i, j, k \geq 0,\ \text{and } i = j \text{ or } j = k\}$
   - $L_{2c} = \{a^i b^j c^k \mid i, j, k \geq 0,\ \text{and } i + j = k\}$
   - $L_{2d} = \{a^i b^j c^k \mid i, j, k \geq 0,\ \text{and } i + k = j\}$
   - $L_{2e} = \{a^i b^j c^k \mid i, j, k \geq 0,\ \text{and } i + k < j\}$
   - $L_{2f} = \{w \mid w \in a, b^* \text{ and } w \neq w^{rev}\} =$ all non-palindromes over $\{a, b\}$

   Now, design (separate) push-down automata which accepts all of the above CFL. State whether your PDA accepts by final state or empty stack.

3. Which of the following language(s) is/are context-free? Give justifications – For CFL, you need to give the CFG, otherwise you have to prove using Pumping Lemma for CFL.

   - $L_{3a} = \{a^m b^n \mid m, n \geq 0, m = 2n\}$
   - $L_{3b} = \{a^m b^n \mid m, n \geq 0, m \neq 2n\}$
   - $L_{3c} = \{a^m b^n \mid n \geq 0, 3n \leq m \leq 5n\}$
   - $L_{3d} = \{a^m b^n \mid m, n \geq 1 \text{ and } n \neq 2m\}$
   - $L_{3e} = \{a^k \mid k \text{ is a power of } 2\}$
   - $L_{3f} = \{x \in \{0, 1\}^* \mid \overline{x} = x^{\mathbf{R}}\}$.
   - $L_{3g} = \{a^m b^n c^{m+n} \mid m, n \geq 1\}$
   - $L_{3h} = \{a^m b^m c^{m+n} \mid m, n \geq 1\}$
   - $L_{3i} = \{a^l b^m c^n \mid l \geq 0, l < m \text{ and } l < n\}$
   - $L_{3j} = \{a^n b^m c^k \mid n, m, k \geq 1 \text{ and } n + k = m\}$
   - $L_{3k} = \{a^p \mid p \text{ is a prime}\}$
   - $L_{3l} = \{a^{q!} \mid q \geq 0\}$

   Here, $\overline{x}$ denotes the Boolean complement and $x^{\mathbf{R}}$ denotes the reverse of the string x.

4. A context-free grammar G is called *ambiguous* if some string in $\mathcal{L}(G)$ has two different derivation/parse trees. Answer the following.

   (a) Show that, the following context-free grammar, $G = (\{S\}, \{a, b, c\}, P, S)$, with the production rules (P) defined as: $S \longrightarrow aS \mid aSbS \mid c$, is ambiguous.

   (b) Construct a non-ambiguous grammar, G', that derives the same language.

   (c) Also prove $\mathcal{L}(G) = \mathcal{L}(G')$.

5. Consider the following context-free grammar $G = (\{S, A, B\}, \{a, b\}, P, S)$ with P containing the following production rules: $S \longrightarrow ABS \mid AB$, $A \longrightarrow aA \mid a$, $B \longrightarrow bA$. Which of the following strings are in $\mathcal{L}(G)$ and which are not?

   (a) $aabaab$, (b) $aaaaba$, (c) $aabbaa$, (d) $abaaba$.

   Provide derivations for those that are in $\mathcal{L}(G)$ and reasons for those that are not.

6. Given the following languages over the alphabet $\{a, b\}$, design one-state PDAs that accepts by empty stack (separate PDA for each one).

   (a) $L_{6a} = \mathcal{L}\Big((a+b)^* b\Big)$

   (b) $L_{6b} = \mathcal{L}\Big(a(a+b)^* b\Big)$

   (c) L_{6c} = set of all palindromes over $\{a, b\}$

   Additionally, explore the following:

   (a) Since L_{6a} and L_{6b} are regular languages, can you directly present the left linear and the right linear grammar for them and then formally derive the NFAs from these grammars?

   (b) Can you also develop two (separate) DFAs for L_{6a} and L_{6b} directly? Then, from those DFAs, can you again formally derive the grammars for the same?

7. Consider the two CFGs G and G' with the start symbols S and S' and with the only productions:

   Productions of G : $S \longrightarrow aS \mid B, \quad B \longrightarrow bB \mid b$

   Productions of G' : $S' \longrightarrow aA' \mid bB', \quad A' \longrightarrow aA' \mid B', \quad B' \longrightarrow bB' \mid \varepsilon$

   Prove that, $\mathcal{L}(G) \subset \mathcal{L}(G')$, i.e., $\mathcal{L}(G)$ is strictly contained in $\mathcal{L}(G')$.

8. Convert the following context-free grammars into equivalent CFG in Chomsky Normal Form. Then, proceed converting these Chomsky Normal Form CFGs into Greibach Normal Form CFGs.

   (a) $S \longrightarrow BSB \mid B \mid \varepsilon$
   $B \longrightarrow 00 \mid \varepsilon$

   (b) $S \longrightarrow ABa \mid \varepsilon$
   $A \longrightarrow aab$
   $B \longrightarrow Ac$

   (c) $S \longrightarrow AbA$
   $A \longrightarrow Aa \mid \varepsilon$

9. Provide grammars in Chomsky normal form for the following languages.

   (a) $L_{9a} = \left\{a^n b^{2n} c^k \mid k, n \geq 1\right\}$

   (b) $L_{9b} = \{a, b\}^* - \left\{x \in \{a, b\}^* \mid x = x^{\mathbf{R}}\right\}$, where $x^{\mathbf{R}}$ denotes the reverse of the string x.

   (c) $L_{9c} = \left\{a^n b^k c^n \mid k, n \geq 1\right\}$
