Automata and Grammars

SS 2018

Assignment 8: Solutions to Selected Problems

Problem 8.1. [Context-Free Grammars]

Let $G = (\{S, A, B, C\}, \{a, b, c\}, P, S)$ be the context-free grammar with the following set of productions:

$P = \{S \to aSb, S \to aAbb, S \to \varepsilon, A \to aAB, A \to bB, B \to aAb, B \to CC, C \to ba, C \to cS\}$.

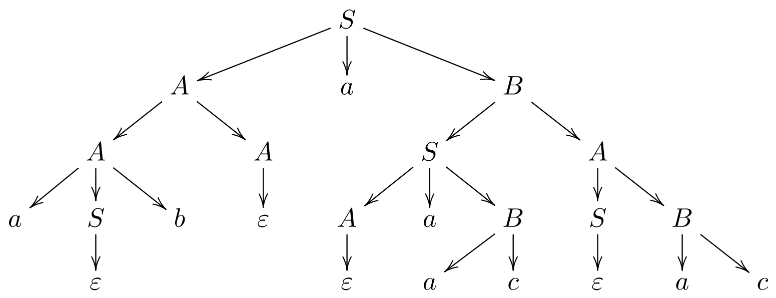
(a) Give a leftmost derivation for the word $w = aaababbababbababbb$.

(b) Give a right derivation for the word $w = aabbababacbb$.

(c) Give syntax trees for your derivations in (a) and in (b).

Problem 8.2. [Context-Free Grammars]

The following figure shows a syntax tree for some context-free grammar G and a word w:



(a) Determine the word w generated by this syntax tree.

(b) Give a leftmost derivation for this word.

(c) Give a list of all productions that are used in this syntax tree.

(d) What can be said about the ambiguity of the underlying grammar?

Solution. (a) $w = abaaacac$.

(b) The corresponding leftmost derivation is as follows:

$$\begin{array}{llllll} S & \to AaB & \to & AAaB & \to & aSbAaB \\ & \to abAaB & \to & abaB & \to & abaSA \\ & \to abaAaBA & \to & abaaBA & \to & abaaacA \\ & \to abaaacSB & \to & abaaacB & \to & abaaacac. \end{array}$$

(c) The following productions are used in the syntax tree:

$$P = \{S \to AaB, S \to \varepsilon, A \to AA, A \to aSb, A \to SB, A \to \varepsilon, B \to SA, B \to ac\}.$$

(d) The grammar G is ambiguous. Consider the leftmost derivations

$$\begin{array}{llllll} S & \to AaB & \to & AAaB & \to & AAAaB \\ & \to aSbAAaB & \to & abAAaB & \to & abaSbAaB \\ & \to ababAaB & \to & ababaSbaB & \to & abababaB \\ & \to abababaac & & & & \end{array}$$

and

$$\begin{array}{llllll} S & \to & AaB & \to & AAaB & \to & aSbAaB \\ & \to & abAaB & \to & abAAaB & \to & abaSbAaB \\ & \to & ababAaB & \to & ababaSbaB & \to & abababaB \\ & \to & abababaac. & & & & \end{array}$$

In the third step these leftmost derivations differ, but they both yields the word $abababaac$. This shows that the grammar G is ambiguous. □

Problem 8.3. [Context-Free Grammars]

Provide context-free grammars for the following languages:

$$\begin{array}{lll} \text{(a)} & L_1 & = & \{\, wcw^R \mid w \in \{a,b\}^* \,\}, \\ \text{(b)} & L_2 & = & \{\, a^k b^m c^n \mid k = m \text{ or } m = n \,\}, \\ \text{(c)} & L_3 & = & \{\, uv \mid u, v \in \{a,b\}^+, |u| = |v|, \text{ but } u \neq v \,\}. \end{array}$$

Solution. (a) Let $G_1 = (\{S\}, \{a,b,c\}, P, S)$, where P is defined as follows:

$$P = \{S \to aSa, S \to bSb, S \to c\}.$$

Then it is easily seen that $L(G_1) = L_1$.

(b) Let $G_2 = (\{S, A, B, C, D\}, \{a,b,c\}, P, S)$, where P is defined as follows:

$$P = \{S \to AB, S \to CD, A \to aAb, A \to \varepsilon, B \to cB, B \to \varepsilon, C \to aC, C \to \varepsilon, D \to bDc, D \to \varepsilon\}.$$

Then it is easily seen that $L(G_2) = L_2$.

(c) Let $G_3 = (\{S, A, B, C, D, E\}, \{a,b\}, P, S)$, where P is defined as follows:

$$\begin{array}{lll} P & = & \{S \to AB, S \to CD, \\ & & A \to EAE, A \to a, B \to EBE, B \to b, C \to ECE, C \to b, D \to EDE, D \to a, \\ & & E \to a, E \to b\}. \end{array}$$

Then $S \to AB \to^* x_1ax_2y_1by_2$, where $|x_1| = |x_2|$ and $|y_1| = |y_2|$. Hence, $|x_1| + |y_2| = |x_2| + |y_1|$, which means that the distinguished occurrences of a and b are at the same position in the first half u and in the second half v of the word $uv = x_1ax_2y_1by_2$. Hence, it follows that $u \neq v$. The case that the derivation starts with $S \to CD$ is analogous.

On the other hand, if $w = uv \in L_3$, then $u = x_1ay_1$ and $v = x_2by_2$ or $u = x_1by_1$ and $v = x_2ay_2$ for some words satisfying $|x_1| = |x_2|$ and $|y_1| = |y_2|$. Let us write $y_1x_2 = x_3y_3$, where $|x_3| = |x_2| = |x_1|$ and $|y_3| = |y_1| = |y_2|$. Then $S \to AB \to^* x_1ax_3y_3by_2 = x_1ay_1x_2by_2 = uv = w$, that is, $w \in L(G_3)$. The case that $u = x_1by_1$ and $v = x_2ay_2$ is completely analogous. Hence, it follows that $L(G_3) = L_3$. □

Problem 8.4. [Simplifying Context-Free Grammars]

Let G be the following context-free grammar:

$$G = (\{S, A, B, C\}, \{a, c\}, \{S \to aACa, A \to B, A \to a, B \to C, B \to c, C \to cC, C \to \varepsilon\}, S).$$

(a) Determine a proper context-free grammar that is equivalent to G.

(b) Remove the ε-productions from G.

(c) Remove the chain productions from G. Is the resulting grammar G' proper? If not, then determine a proper context-free grammar that is equivalent to G' and that has neither ε-productions nor chain productions.

Solution. (a) $V_{\text{term}} = \{ X \in V \mid \exists w \in \{a,c\}^* : X \to_P^* w \} = \{S, A, B, C\}$ and $V_{\text{reach}} = \{ X \in V \mid \exists \alpha, \beta \in (V_{\text{term}} \cup T)^* : S \to_P^* \alpha X \beta \} = \{S, A, B, C\}$. Thus, all nonterminals of G are useful, that is, G is already a proper context-free grammar.

(b) $V_1 = \{ X \in V \mid X \to_P^* \varepsilon \} = \{A, B, C\}$. Now we remove the ε-production $C \to \varepsilon$ and add the following new productions $S \to aAa$, $S \to aCa, S \to aa$ and $C \to c$. In this way we obtain the context-free grammar $G_1 = (\{S, A, B, C\}, \{a, c\}, P_1, S)$, where P_1 is defined as follows:

$$\begin{aligned} P_1 \;=\; & \{S \to aACa, S \to aAa, S \to aCa, S \to aa, \\ & A \to B, A \to a, B \to C, B \to c, C \to cC, C \to c\}. \end{aligned}$$

(c) We have $A \to_{P_1} B$ and $B \to_{P_1} C$, which are the only chain productions. Thus, we see that no nonterminal is equivalent to any other nonterminal. Now by ordering the nonterminals as $S < A < B < C$, we obtain that $X < Y$ holds for each chain production $(X \to Y)$. In order to get rid off of the chain productions, we now proceed as follows:

- Remove the chain production $(B \to C)$ and add the production $(B \to cC)$.
- Remove the chain production $(A \to B)$ and add the productions $(A \to c)$ and $(A \to cC)$.

The resulting grammar is $G_2 = (\{S, A, B, C\}, \{a, c\}, P_2, S)$, where P_2 is defined as follows:

$$\begin{aligned} P_2 \;=\; & \{S \to aACa, S \to aAa, S \to aCa, S \to aa, \\ & A \to c, A \to cC, A \to a, B \to cC, B \to c, C \to cC, C \to c\}. \end{aligned}$$

$V_{\text{term}} = \{ X \in V \mid \exists w \in \{a,c\}^* : X \to_P^* w \} = \{S, A, B, C\}$ and $V_{\text{reach}} = \{ X \in V \mid \exists \alpha, \beta \in (V_{\text{term}} \cup T)^* : S \to_P^* \alpha X \beta \} = \{S, A, C\}$. Thus, the nonterminal B is not useful. By removing this nonterminal and all productions containing B we obtain the desired proper context-free grammar $G_3 = (\{S, A, C\}, \{a, c\}, P_3, S)$, where P_3 is defined as follows:

$$\begin{aligned} P_3 \;=\; & \{S \to aACa, S \to aAa, S \to aCa, S \to aa, \\ & A \to c, A \to cC, A \to a, C \to cC, C \to c\}. \end{aligned}$$

□

Problem 8.5. [Chomsky Normal Form]

Convert the following context-free grammar $G = (\{S, A\}, \{a, b, c\}, P, S)$ into an equivalent grammar that is in Chomsky Normal Form using the construction detailed in the proof of Theorem 3.9, where P_1 is defined as follows:

$$P = \{S \to cS, S \to aAb, S \to ab, A \to aAb, A \to ab, A \to cc\}.$$