Languages and Computation (COMP 2049) Lab 07

Closure Properties of Context-Free Languages

Through the following exercises, we will learn that the class of context-free languages (CFLs) is closed under union, concatenation, and star-closure. We may prove these closure properties by simple manipulation of grammars:

Example 1 (Union). Assume that the languages L_1 and L_2 are generated by context-free grammars (CFGs) $G_1 = (V_1, T_1, S_1, P_1)$ and $G_2 = (V_2, T_2, S_2, P_2)$, respectively. Without loss of generality, we may assume that there are no common variables between G_1 and G_2, i. e., $V_1 \cap V_2 = \emptyset$. In particular, this implies that $S_1 \neq S_2$.

- Question: Why can we assume this without loss of generality?

Solution

The answer is very easy. We can freely rename the variables of a CFG without changing the language that it generates.

We create a new variable S_3 which is not in $V_1 \cup V_2$. Then, we construct a new grammar $G_3 = (V_3, T_3, S_3, P_3)$ by assigning:

- $V_3 := V_1 \cup V_2 \cup \{S_3\}$
- $T_3 := T_1 \cup T_2$
- $P_3 := P_1 \cup P_2 \cup \{S_3 \to S_1 \mid S_2\}$

Clearly, G_3 is context-free and $L(G_3) = L_1 \cup L_2$. Hence, the class of CFLs is closed under union.

Using an argument similar to that of Example 1, prove that the class of CFLs is closed under:

(a) Concatenation;

(b) Star-closure.

Solution

(a) Assume that the languages L_1 and L_2 are generated by CFGs $G_1 = (V_1, T_1, S_1, P_1)$ and $G_2 = (V_2, T_2, S_2, P_2)$, respectively. Without loss of generality, we may assume that there are no common variables between G_1 and G_2, i. e., $V_1 \cap V_2 = \emptyset$. We create a new variable S_4 which is not in $V_1 \cup V_2$. Then, we construct a new grammar $G_4 = (V_4, T_4, S_4, P_4)$ by assigning:

   - $V_4 := V_1 \cup V_2 \cup \{S_4\}$
   - $T_4 := T_1 \cup T_2$
   - $P_4 := P_1 \cup P_2 \cup \{S_4 \to S_1 S_2\}$

   The grammar G_4 is context-free and $L(G_4) = L_1 L_2$. Hence, the class of CFLs is closed under concatenation.

(b) Assume that the languages L_1 is generated by CFG $G_1 = (V_1, T_1, S_1, P_1)$. We create a new variable S which is not in V_1. Then, we construct a new grammar $G = (V, T, S, P)$ by assigning:

   - $V := V_1 \cup \{S\}$
   - $T := T_1$
   - $P := P_1 \cup \{S \to S_1 S \mid \lambda\}$

   The grammar G is context-free and $L(G) = L_1^*$. Hence, the class of CFLs is closed under star-closure.