CS340: Theory of Computation — Sem I 2021-22

Lecture Notes 12: Properties of Context-free Languages

Raghunath Tewari — IIT Kanpur

1 Closure Properties

Let us study the various closure properties of CFLs.

1. **Union**

   Let L_1 and L_2 be two CFLs accepted by the CFGs $G_1 = (V_1, \Sigma, P_1, S_1)$ and $G_2 = (V_2, \Sigma, P_2, S_2)$ respectively. Then $L_1 \cup L_2$ will be accepted by the CFG $G = (V, \Sigma, P, S)$ where

   $$\begin{aligned} V &= V_1 \cup V_2 \cup \{S\}, \text{ and} \\ P &= P_1 \cup P_2 \cup \{S \longrightarrow S_1 \mid S_2\}. \end{aligned}$$

2. **Concatenation**

   Let L_1 and L_2 be two CFLs accepted by the CFGs $G_1 = (V_1, \Sigma, P_1, S_1)$ and $G_2 = (V_2, \Sigma, P_2, S_2)$ respectively. Then $L_1 \cdot L_2$ will be accepted by the CFG $G = (V, \Sigma, P, S)$ where

   $$\begin{aligned} V &= V_1 \cup V_2 \cup \{S\}, \text{ and} \\ P &= P_1 \cup P_2 \cup \{S \longrightarrow S_1 S_2\}. \end{aligned}$$

3. **Star**

   Let L_1 be a CFL accepted by the CFG $G_1 = (V_1, \Sigma, P_1, S_1)$. Then L_1^* will be accepted by the CFG $G = (V, \Sigma, P, S)$ where

   $$\begin{aligned} V &= V_1 \cup \{S\}, \text{ and} \\ P &= P_1 \cup \{S \longrightarrow S_1 S \mid \epsilon\}. \end{aligned}$$

4. **Reversal**

   Let L_1 be a CFL accepted by the CFG $G_1 = (V, \Sigma, P, S)$. Then $\text{rev}(L_1)$ will be accepted by the CFG $G = (V, \Sigma, P_R, S)$ where for every rule $A \longrightarrow X_1 X_2 \ldots X_k$ in P, add the rule

   $$A \longrightarrow X_k \ldots X_2 X_1$$

   to P_R.

5. **Homomorphism and Inverse Homomorphism**

   CFLs are also closed under homomorphism and inverse inverse homomorphism. The construction is left as an exercise.

Exercise 1. Show that CFLs are closed under homomorphism and inverse inverse homomorphism.

(*Hint*: For homomorphism start with a CFG and for inverse homomorphism start with a PDA.)

6. **Intersection with a Regular language** Let L_1 be a CFL and L_2 be a regular language, then $L_1 \cap L_2$ is a CFL.

   The idea is to take a PDA N for L_1 and a DFA M for L_2 and construct a PDA (say N') for $L_1 \cap L_2$. N' will be a "product automaton" of M and N, making a moving according to both M and N at each step simultaneously. In addition, N' will use its own stack to simulate the stack of N.

   Formally, let $N = (Q_N, \Sigma, \Gamma, \delta_N, q_{0_N}, F_N)$ and $M = (Q_M, \Sigma, \delta_M, q_{0_M}, F_M)$. We construct $N' = (Q, \Sigma, \Gamma, \delta, q_0, F)$ as follows:

   - $Q = Q_N \times Q_M$
   - $q_0 = (q_{0_N}, q_{0_M})$
   - $F = F_N \times F_M$
   - $((r, s), Y) \in \delta((p, q), a, X)$ if
     - $(r, Y) \in \delta_N(p, a, X)$, and
     - $s = \delta_M(q, a)$ if $a \in \Sigma$, and $s = q$ if $a = \epsilon$.

1.1 Non-closure under certain operations

What about other set operations such as intersection, complement and set difference?
It turns out that CFLs are *not* closed under these operations.

- Consider the two languages

$$\begin{aligned} L_1 &= \{a^n b^n c^m \mid n, m \geq 0\} \\ L_2 &= \{a^n b^m c^m \mid n, m \geq 0\} \end{aligned}$$

  Here is a CFG for L_1:

$$\begin{aligned} S &\longrightarrow S_1 C \\ S_1 &\longrightarrow a S_1 b \mid \epsilon \\ C &\longrightarrow cC \mid \epsilon \end{aligned}$$

  Similarly one can construct a CFG for L_2 as well. But now we can write our favourite non context-free language $L = \{a^n b^n c^n \mid n \geq 0\}$ as,

$$L = L_1 \cap L_2.$$

  Hence CFLs are not closed under intersection.

- If CFLs were closed under complement, then by DeMorgan's law they would be closed under intersection as well. Hence CFLs are not closed under complement.

- For a language $L \subseteq \Sigma^*$, $\overline{L} = \Sigma^* \setminus L$. This shows that CFLs are not closed under set difference as well.

1.2 Some Applications

1. Let
$$L_1 = \{w \in \{a,b,c\}^* \mid \#_a(w) = \#_b(w) = \#_c(w)\}.$$
Note that $L_1 \cap L(a^*b^*c^*) = \{a^n b^n c^n \mid n \geq 0\}$. Since $L(a^*b^*c^*)$ is a regular language and the language on the right hand side of the equal sign is not a CFL therefore L_1 is not a CFL.

2. Show that
$$L_2 = \{a^n b^n a^{2m} b^{2m} \mid n, m \geq 0\}$$
is context-free.

   We use the fact that $L' = \{a^n b^n \mid n \geq 0\}$ is context-free. Consider the homomorphism h defined as
$$\begin{aligned} h(a) &= aa \\ h(b) &= bb. \end{aligned}$$
Then $h(L') = \{a^{2n} b^{2n} \mid n \geq 0\}$. Now observe that $L_2 = L' \cdot h(L')$. Therefore L_2 is context-free.