Handout 8: Context-free Closure Properties

8.1 Closure under union. The context-free languages are closed under union. Given two CFGs $G_1 = (V_1, \Sigma, R_1, S_1)$ and $G_2 = (V_2, \Sigma, R_2, S_2)$, we can construct a CFG $G = (V, \Sigma, R, S)$ such that $L(G) = L(G_1) \cup L(G_2)$:

$$V = V_1 \cup V_2 \cup \{S\}$$
$$R = R_1 \cup R_2 \cup \{S \rightarrow S_1 \mid S_2\}$$

Given two PDAs M_1 and M_2, we can construct a PDA M such that $L(M) = L(M_1) \cup L(M_2)$ as we did for NFAs.

This result can be useful to show that a language is context-free. For example, the language

$$L_7 \;=\; \{0^i 1^j 2^k \mid i \neq j \text{ or } j \neq k\}$$

is context-free, because it is the union of the two context-free languages $\{0^i 1^j 2^k \mid i \neq j\}$ and $\{0^i 1^j 2^k \mid j \neq k\}$.

8.2 Closure under concatenation. The context-free languages are closed under concatenation. Given two CFGs $G_1 = (V_1, \Sigma, R_1, S_1)$ and $G_2 = (V_2, \Sigma, R_2, S_2)$, we can construct a CFG $G = (V, \Sigma, R, S)$ such that $L(G) = L(G_1) \circ L(G_2)$:

$$V = V_1 \cup V_2 \cup \{S\}$$
$$R = R_1 \cup R_2 \cup \{S \rightarrow S_1 S_2\}$$

Given two PDAs M_1 and M_2, we can construct a PDA M such that $L(M) = L(M_1) \circ L(M_2)$ as we did for NFAs.

This result can be useful to show that a language is context-free. For example, the language

$$L_8 \;=\; \{0^i 1^j 2^k \mid j = i + k\}$$

is context-free, because it is the concatenation of the two context-free languages $\{0^i 1^i\}$ and $\{1^k 2^k\}$.

8.3 Closure under iteration. The context-free languages are closed under iteration. Given a CFGs $G = (V, \Sigma, R, S)$, we can construct a CFG $G' = (V', \Sigma, R', S')$ such that $L(G') = L(G)^*$:

$$V' = V \cup \{S'\}$$
$$R = R \cup \{S' \rightarrow \varepsilon \mid SS'\}$$

Given a PDA M, we can construct a PDA M' such that $L(M') = L(M)^*$ as we did for NFAs.

8.4 Closure under intersection with a regular language. If A is a context-free language and B is regular, then $A \cap B$ is context-free. Given a PDA $M_1 = (Q_1, \Sigma, \Gamma, \delta_1, q_{01}, F_1)$ and an NFA $M_2 = (Q_2, \Sigma, \delta_2, q_{02}, F_2)$, using the product construction we can construct a PDA $M = (Q, \Sigma, \Gamma, \delta, q_0, F)$ such that $L(M) = L(M_1) \cap L(M_2)$:

$$Q = Q_1 \times Q_2$$
$$((p_1, p_2), \gamma') \in \delta((q_1, q_2), \sigma, \gamma) \text{ iff } (p_1, \gamma') \in \delta_1(q_1, \sigma, \gamma) \text{ and } p_2 \in \delta_2(q_2, \sigma)$$
$$q_0 = (q_{01}, q_{02})$$
$$F = F_1 \times F_2$$

This result can be useful to show that a language is not context-free. For example, the language

$$L_8 = \{w \in \{0,1,2\}^* \mid \#(w,0) = \#(w,1) = \#(w,2)\}$$

is not context-free, because the non-context-free language is L_5 is the intersection of L_8 and the regular language $0^*1^*2^*$.

8.5 Nonclosure under intersection. The product of two PDAs would result in two stacks. Indeed, there are context-free languages whose intersection is not context-free. For example, L_5 is the intersection of the two context-free languages $\{0^i1^j2^k \mid i = j\}$ and $\{0^i1^j2^k \mid j = k\}$.

8.6 Nonclosure under complementation. Using de Morgan's law, from Sections 8.1 and 8.5 it follows that there are context-free languages whose complement is not context-free. For example, the complement of L_7 is not context-free, because $L_5 = \overline{L_7} \cap 0^*1^*2^*$.

A PDA $M = (Q, \Sigma, \Gamma, \delta, q_0, F)$ is *deterministic* if for all $q \in Q$, $\sigma \in \Sigma$, and $\gamma \in \Gamma$,

$$|\delta(q,\sigma,\gamma) \cup \delta(q,\sigma,\varepsilon) \cup \delta(q,\varepsilon,\gamma) \cup \delta(q,\varepsilon,\varepsilon)| = 1;$$

that is, in all situations M can proceed in exactly one way. If M is deterministic and $M' = (Q, \Sigma, \Gamma, \delta, q_0, Q \backslash F)$, then $L(M') = \overline{L(M)}$. It follows that context-free language L_7 cannot be recognized by a deterministic PDA; that is, unlike for finite automata, the deterministic variety of push-down automata is strictly less powerful than the nondeterministic variety.