Lecture 8: PDAs, CFLs, CNFs, and CYK

Date: September 14, 2023.

Proposition 1 (Context-freeness). *Let $G = (N, \Sigma, P, S)$ be a CFG. For any $A \in N$ and $\alpha, \beta, \gamma \in (N \cup \Sigma)^*$, if $A \xrightarrow[G]{*} \beta$ then $\alpha A \gamma \xrightarrow[G]{*} \alpha\beta\gamma$.*

Let $M = (Q, \Sigma, \Gamma, \delta, s, \perp, F)$ be a PDA. For any $p, q \in Q$, $\alpha, \beta, \gamma \in \Gamma^*$, $x, y, z \in \Sigma^*$, if $(p, x, \alpha) \xrightarrow[M]{*} (q, y, \beta)$ then $(p, xz, \alpha\gamma) \xrightarrow[M]{*} (q, yz, \beta\gamma)$.

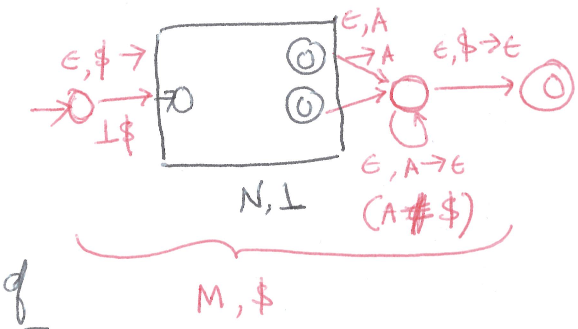
Theorem 2. *If L is a context-free language then there is a PDA M such that $\mathbf{L}(M) = L$.*

Theorem 3. *For any PDA M, $\mathbf{L}(M)$ is context-free.*

$M = (Q, \Sigma, \Gamma, \delta, s, \perp, F)$

$\forall x, q \in F$, if $(s, x, \perp) \xrightarrow[M]{*} (q, \epsilon, \gamma)$

then $\gamma = \epsilon$.



Intuition: Grammar G has non-terminals of the form $[pAq]$ where $p, q \in Q$, $A \in \Gamma$.

$[pAq] \xrightarrow[G]{*} w \iff (p, w, A) \xrightarrow[M]{*} (q, \epsilon, \epsilon)$

$w \in \Sigma^*$

$L(M) = \{ w \in \Sigma^* \mid [s \perp q] \xrightarrow[G]{*} w,\ q \in F \}$.

$G = (N, \Sigma, P, \mathcal{S})$

- $N = \{ [pAq] \mid p, q \in Q, A \in \Gamma \} \cup \{ \mathcal{S} \}$.

$\mathcal{S} \to [s \perp q]$ where $q \in F$.

$[pAq] \to a [q_1 B_1 q_2][q_2 B_2 q_3] \cdots [q_k B_k q]$ if $(p, \underbrace{a}_{\in \Sigma \cup \{\epsilon\}}, A, q_1, B_1 B_2 \ldots B_k) \in \delta$.

and $q_2, q_3 \ldots q_k \in Q$.

Transition Relation: $\delta \underset{\text{finite}}{\subseteq} Q \times (\Sigma \cup \{\epsilon\}) \times \Gamma \times Q \times \Gamma^*$

Context-free languages are not closed under intersection

Proposition 4. *If A is a context-free and B is regular then $A \cap B$ is context-free.*

$A \cap B$: Given w, determine if $w \in A$ and $w \in B$.

A is CFL : $\exists$ PDA $M = (Q, \Sigma, \Gamma, \delta, s, \perp, F)$ s.t $L(M) = A$.

B is regular : $\exists$ DFA $N = (Q', \Sigma, \delta', s', F')$ s.t $L(N) = B$.

Algo for $A \cap B$: Check if M accepts w & Check if N accepts w.] Simultaneously.

$$M_1 = ((Q \times Q'), \Sigma, \Gamma, \delta_1, (s, s'), \perp, F \times F')$$

$$\forall p_2 \in Q' \quad ((p_1, p_2), \epsilon, A, (q_1, p_2), B_1 \ldots B_k) \in \delta_1 \Leftarrow (p_1, \epsilon, A, q_1, B_1 \ldots B_k) \in \delta.$$

$$((p_1, p_2), a, A, (q_1, q_2), B_1 \ldots B_k) \in \delta_1 \Leftarrow (p_1, a, A, q_1, B_1 \ldots B_k) \in \delta \text{ and } \delta'(p_2, a) = q_2$$

If A is CFL, $h : \Sigma^* \to \Gamma^*$, then $h(A)$ is CFL.

Proposition 5. *If $A \subseteq \Gamma^*$ is context-free and $h : \Sigma^* \to \Gamma^*$ then $h^{-1}(A)$ is context-free.*

Problem defined $h^{-1}(A)$: Given w, determine $w \in h^{-1}(A)$.
determine $h(w) \in A$.

For each symbol a read from input
Run machine for A on $h(a)$ $\to$ finite.
PDA -

New PDA for $h^{-1}(A)$: Will an input buffer in its state.

$(Q \times \Gamma^*)$

G for A. $\quad B \to \underline{u}\, C_1 C_2 \ldots$

G for $h^{-1}(A)$. $\quad B \to \underline{w}\, C_1 C_2 \ldots \quad$ where $h(w) = u$.

Chomsky Normal Form: A CFG $G = (N, \Sigma, P, S)$ is said to be in *Chomsky Normal Form* if all production rules in P are of the form $A \to a$ or $A \to BC$, where $a \in \Sigma$ and $A, B, C \in N$.

Greibach Normal Form: A CFG $G = (N, \Sigma, P, S)$ is said to be in *Greibach Normal Form* if all production rules in P are of the form $A \to a\alpha$, where $a \in \Sigma$ and $\alpha \in N^*$.

Proposition 6. *For any CFG G, there is a CFG G_{CNF} in Chomsky Normal Form, and a CFG G_{GNF} in Greibach Normal Form, such that*

$$\mathbf{L}(G_{\mathsf{CNF}}) = \mathbf{L}(G_{\mathsf{GNF}}) = \mathbf{L}(G) \setminus \{\epsilon\}.$$

Cocke-Younger-Kasami (CYK) Algorithm: Given a CFG $G = (N, \Sigma, P, S)$ and an input string $w \in \Sigma^*$, the algorithm determines if $w \in \mathbf{L}(G)$.

When G is GNF, PDA recognizing L(G) can be constructed in such a way that it has no ϵ-transitions.