To show that a set is not a CFL - use pumping lemma in the contrapositive form.

For all $k \geq 0$, $\exists\, z \in A$ s.t $|z| \geq k$ and for all split of z into substrings $z = uvwxy$ with $vx \neq \epsilon$ and $|vwx| \leq k$, there exists an $i \geq 0$ s.t $uv^iwx^iy \notin A$.

$$A = \{ w \in \{a,b,c\}^* \mid \#_a(w) = \#_b(w) = \#_c(w) \}$$

Claim. A is not a CFL.

Theorem. CFLs are closed under intersection with regular sets.

if $A \subseteq \Sigma^*$ is a CFL and $B \subseteq \Sigma^*$ is a regular set then $A \cap B$ is a CFL.

$$B = A \cap L(a^* b^* c^*) = \{ a^n b^n c^n \mid n \geq 0 \}$$

if A is a CFL then by Theorem B is a CFL which is a contradiction.

Example.

$A = \{ww \mid w \in \{a,b\}^*\}$.

Claim. A is not a CFL.

$A' = A \cap L(a^*b^*a^*b^*) = \{a^n b^m a^n b^m \mid n,m \geq 0\}$

Suffices to show that A' is not a CFL

Consider any k. Choose $z = a^k b^k a^k b^k$

We have $|z| \geq k$. No matter which way z is split

$z = uvwxy$ where $vx \neq \epsilon$ and $|vwx| \leq k$

with $i=2$ it can be shown that $uv^iwx^iy \notin A'$.

By pumping lemma, A' is not regular.
so A is not regular.

Deterministic PDA.

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

- $\dashv$ is a special symbol not in Σ. Right end marker

$$\delta \subseteq (Q \times (\Sigma \cup \{\dashv\} \cup \{\epsilon\}) \times \Gamma) \times (Q \times \Gamma^*)$$

- For any $p \in Q$, $a \in \Sigma \cup \{\dashv\}$, $A \in \Gamma$, δ contains exactly one transition of the form

$$((p, a, A), (q, \beta)) \text{ or } ((p, \epsilon, A), (q, \beta))$$

- $\perp$ is always at the bottom of the stack

All transitions involving $\perp$ should be $((p, a, \perp), (q, \beta\perp))$

Acceptance is by final state: $(s, x\dashv, \perp) \xrightarrow[M]{*} (q, \epsilon, \beta)$

Properties:

- DCFLs are closed under complementation.
- DCFLs are not closed under union, intersection.

DCFLs $\subsetneq$ CFL.

CFG to NPDA.

Suppose $G = (N, \Sigma, P, S)$. To construct NPDA M s.t $L(M) = L(G)$.

Assume that all productions of G are of the form

$$\underbrace{A \to c B_1 B_2 \cdots B_k}_{\text{Greibach Normal Form.}} \quad , c \in \Sigma \cup \{\epsilon\},\ k \geq 0.$$

Construct $M = (\{q\}, \Sigma, N, \delta, q, S, \phi)$

M has one state and accepts with empty stack

Σ - Set of terminal symbols in G : input alphabet in M

N - Set of nonterminals of G is the stack alphabet of M.

S - Start symbol of G, the initial stack symbol of M.

Definition of δ:

For each production $A \to c B_1 B_2 \cdots B_k$, add the following to δ.

$$((q, c, A), (q, B_1 B_2 \cdots B_k))$$

$M = (\{q\}, \Sigma, N, \delta, q, S, \phi)$

δ: For each $A \to cB_1 B_2 \cdots B_k$, $((q, c, A), (q, B_1 B_2 \cdots B_k)) \in \delta$.

Example. Balanced Parentheses.

Production Rules in G	Transitions in M.
1. $S \to [BS$	$((q, [, S), (q, BS))$.
2. $S \to [B$	$((q, [, S), (q, B))$
3. $S \to [SB$	$((q, [, S), (q, SB))$
4. $S \to [SBS$	$((q, [, S), (q, SBS))$
5. $B \to]$	$((q,], B), (q, \epsilon))$

Leftmost derivation - productions are always applied to the leftmost nonterminal.

To show. Leftmost derivation of x in G corresponds to an accepting computation of M on input x.

Example: Input $x = [[[]][]]$

Rule		Configurations of M
	S	$(q, [[[]][]], S)$
3	$[SB$	$(q, [[]][]], SB)$
4	$[[SBSB$	$(q, []][]], SBSB)$
2	$[[[BBSB$	$(q,]][]], BBSB)$
5	$[[[]BSB$	$(q,][]], BSB)$
5	$[[[]]SB$	$(q, []], SB)$
2	$[[[]][BB$	$(q,]], BB)$
5	$[[[]][]B$	$(q,], B)$
5	$[[[]][]]$	(q, ϵ, ϵ).

Lemma 1. For any $z, y \in \Sigma^*$, $\gamma \in N^*$ and $A \in N$,

$A \xrightarrow[G]{n} z\gamma$ by a leftmost derivation iff $(q, zy, A) \xrightarrow[M]{n} (q, y, \gamma)$

Proof. Induction on n.

Base case, $n = 0$: Straight forward.

Induction step.

Suppose $A \xrightarrow[G]{n+1} z\gamma$ using a leftmost derivation.

$c \in \Sigma \cup \{\epsilon\}, \beta \in N^*$

Suppose $B \to c\beta$ was the last production applied.

$$A \xrightarrow[G]{n} uB\alpha \xrightarrow[G]{1} uc\beta\alpha = z\gamma. \quad / \ z = uc \text{ and } \gamma = \beta\alpha$$

By induction hypothesis, $(q, ucy, A) \xrightarrow[M]{n} (q, cy, B\alpha)$.

By definition of M, $((q, c, B), (q, \beta)) \in \delta$.

therefore, $(q, cy, B\alpha) \xrightarrow[M]{1} (q, y, \beta\alpha)$

Thus we have

$$(q, zy, A) = (q, ucy, A) \xrightarrow[M]{n+1} (q, y, \beta\alpha) = (q, y, \gamma).$$

Conversly, Suppose $(q, zy, A) \xrightarrow[M]{n+1} (q, y, \gamma)$

let $((q, c, B), (q, \beta)) \in S$ be the last transition taken by M.

Then $z = uc$ for some $u \in \Sigma^*$, $\gamma = \beta\alpha$ for some $\alpha \in \Gamma^*$ and

$$(q, ucy, A) \xrightarrow[M]{n} (q, cy, B\alpha) \xrightarrow[M]{1} (q, y, \beta\alpha)$$

By induction hypothesis $A \xrightarrow[G]{n} uB\alpha$ by a leftmost derivation in G.

By definition of S in M, $B \rightarrow c\beta$ is a production of G.

Then, $A \xrightarrow[G]{n} uB\alpha \xrightarrow[G]{1} uc\beta\alpha = z\gamma$ by a leftmost derivation.

Theorem $L(G) = L(M)$.

Proof.

$x \in L(G)$ iff $S \xrightarrow[G]{*} x$ by a leftmost derivation [Defn of $L(G)$]

iff $(q, x, S) \xrightarrow[M]{*} (q, \epsilon, \epsilon)$ [Lemma 1]

iff $x \in L(M)$ [Definition of $L(M)$]

NPDAs accept only CFLs.

Step 1. Every PDA can be simulated by a PDA with one state.

Step 2. Every PDA with one state has an equivalent CFG.

→ "Invert" the construction in the previous lecture.

Suppose $M = (\{q\}, \Sigma, \Gamma, \delta, s, \perp, \phi)$

Define $G = (\Gamma, \Sigma, P, \perp)$ as follows.

For every transition $((q, c, A), (q, B_1 \cdots B_k)) \in \delta$,

add the production $A \to cB_1 B_2 \cdots B_k$ in P.

Lemma 1. For any $z, y \in \Sigma^*$, $\gamma \in N^*$ and $A \in N$,

$A \xrightarrow[G]{n} z\gamma$ by a leftmost derivation iff $(q, zy, A) \xrightarrow[M]{n} (q, y, \gamma)$

Theorem $L(G) = L(M)$.