

# Lecture 17 – Deterministic Pushdown Automata (DPDA)

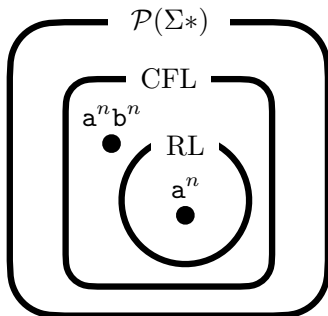
## COSE215: Theory of Computation

Jihyeok Park



2023 Spring

- A **pushdown automaton (PDA)** is an extension of  $\epsilon$ -NFA with a **stack**. Thus, PDA is **non-deterministic**.
  - Acceptance by **final states**
  - Acceptance by **empty stacks**
- Then, how about **deterministic PDA (DPDA)**?
- What is the **language class** of DPDA? Still, CFL?



$\text{PDA}_{\text{FS}}$   
(by final states)

||

$\text{PDA}_{\text{ES}}$   
(by empty stacks)

||

CFG

## 1. Deterministic Pushdown Automata (DPDA)

## 2. Deterministic Context-Free Languages (DCFLs)

Fact 1:  $\text{DCFL} \subset \text{CFL}$

Fact 2:  $\text{RL} \subset \text{DCFL}$

## 3. Languages Accepted by Empty Stacks of DPDA ( $\text{DCFL}_{\text{ES}}$ )

Fact 3:  $\text{DCFL}_{\text{ES}} \subset \text{DCFL}$

Fact 4:  $\text{DCFL}_{\text{ES}} = \text{DCFL having Prefix Property}$

Fact 5:  $\text{RL} \not\subset \text{DCFL}_{\text{ES}}$

## 4. Inherent Ambiguity of DCFLs

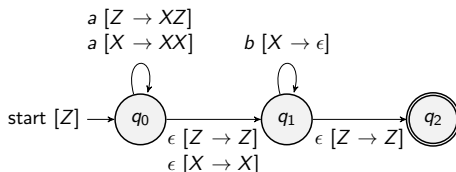
Fact 6:  $\text{DCFL} \subset \text{Non Inherently Ambiguous Languages}$

## Definition (Deterministic Pushdown Automata)

A PDA is **deterministic** if there is at most one-step move ( $\vdash$ ) from any configuration and we call it a **deterministic pushdown automaton (DPDA)**. In other words, a DPDA satisfies the following conditions:

- $|\delta(q, a, X)| \leq 1$  for all  $q \in Q$ ,  $a \in \Sigma \cup \{\epsilon\}$ , and  $X \in \Gamma$ .
- If  $\delta(q, \epsilon, X) \neq \emptyset$ , then  $\delta(q, a, X) = \emptyset$  for all  $a \in \Sigma$ .

For example, the following PDA is **NOT** deterministic:

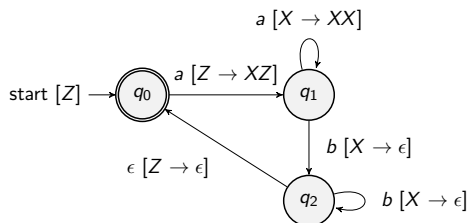


## Definition (Deterministic Pushdown Automata)

A PDA is **deterministic** if there is at most one-step move ( $\vdash$ ) from any configuration and we call it a **deterministic pushdown automaton (DPDA)**. In other words, a DPDA satisfies the following conditions:

- $|\delta(q, a, X)| \leq 1$  for all  $q \in Q$ ,  $a \in \Sigma \cup \{\epsilon\}$ , and  $X \in \Gamma$ .
- If  $\delta(q, \epsilon, X) \neq \emptyset$ , then  $\delta(q, a, X) = \emptyset$  for all  $a \in \Sigma$ .

For example, the following PDA is deterministic:



$$\begin{aligned}
 (q_0, aabb, Z) &\vdash (q_1, abb, XZ) \\
 &\vdash (q_1, bb, XXZ) \\
 &\vdash (q_2, b, XZ) \\
 &\vdash (q_2, \epsilon, Z) \\
 &\vdash (q_0, \epsilon, \epsilon)
 \end{aligned}$$

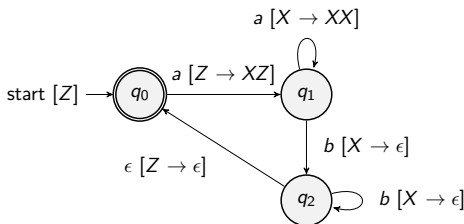
## Definition (Deterministic Context-Free Languages (DCFLs))

A language  $L$  is a **deterministic context-free language (DCFL)** if and only if there exists a DPDA  $P$  such that  $L = L_F(P)$  where  $L_F(P)$  is the language accepted by **final states** of  $P$ .

For example, the following language is a DCFL:

$$L = \{a^n b^n \mid n \geq 0\}$$

because it is accepted by final states of the following DPDA:



## Fact 1: DCFL $\subset$ CFL

All DCFLs are CFLs **BUT** there exists a CFL that is not a DCFL.

For example, the following language is a CFL but not a DCFL:

$$L = \{ww^R \mid w \in \{a, b\}^*\}$$

Why?

Fact 2:  $RL \subset DCFL$ 

All RLs are DCFLs **BUT** there exists a DCFL that is not an RL.

- $RL \subseteq DCFL$ : For a given RL  $L$  and its corresponding DFA

$$D = (Q, \Sigma, \delta, q_0, F)$$

we can construct a DPDA  $P$  that accepts  $L$  as follows:

$$P = (Q, \Sigma, \{Z\}, \delta_P, q_0, Z, F)$$

where  $\forall q \in Q. \forall a \in \Sigma. \delta_P(q, a, Z) = \{(\delta(q, a), Z)\}$ . Then,

$$(q_0, w, Z) \vdash^* (q, \epsilon, Z) \iff \delta^*(q_0, w) = q \quad \square$$

- $DCFL \setminus RL \neq \emptyset$ : We already know that the following language is a DCFL but not an RL:

$$L = \{a^n b^n \mid n \geq 0\} \in DCFL \setminus RL$$



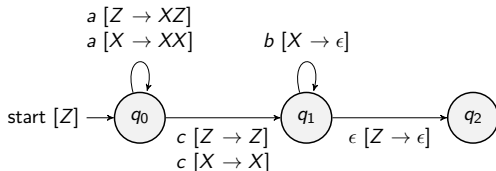
Definition ( $\text{DCFL}_{\text{ES}}$ )

A language  $L$  is a **deterministic context-free language by empty stacks** ( $\text{DCFL}_{\text{ES}}$ ) if and only if there exists a DPDA  $P$  such that  $L = L_E(P)$  where  $L_E(P)$  is the language accepted by **empty stacks** of  $P$ .

For example, the following language is a  $\text{DCFL}_{\text{ES}}$ :

$$L = \{a^n cb^n \mid n \geq 0\}$$

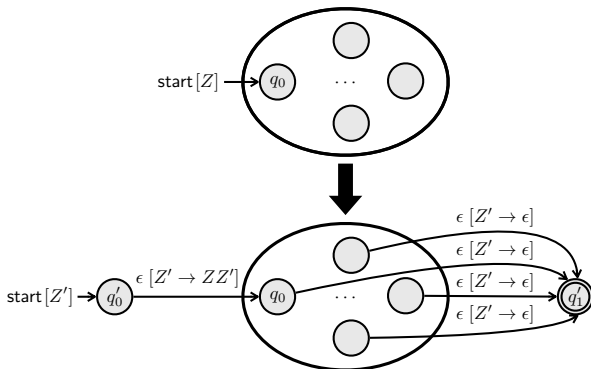
because it is accepted by empty stacks of the following DPDA:



Fact 3:  $\text{DCFL}_{\text{ES}} \subset \text{DCFL}$ 

All  $\text{DCFL}_{\text{ES}}$ s are DCFLs **BUT** there exists a DCFL that is not a  $\text{DCFL}_{\text{ES}}$ .

- $\boxed{\text{DCFL}_{\text{ES}} \subseteq \text{DCFL}}$ : For a given  $\text{DCFL}_{\text{ES}}$   $L$  and its corresponding DPDA  $P$  by **empty stacks**, we can always construct a DPDA  $P'$  that accepts  $L$  by **final states** as follows:



### Fact 3: $\text{DCFL}_{\text{ES}} \subset \text{DCFL}$

All  $\text{DCFL}_{\text{ES}}$ s are DCFLs **BUT** there exists a DCFL that is not a  $\text{DCFL}_{\text{ES}}$ .

- $\text{DCFL} \setminus \text{DCFL}_{\text{ES}} \neq \emptyset$ : The following language is a DCFL but not a  $\text{DCFL}_{\text{ES}}$ :

$$L = \{a^n b^n \mid n \geq 0\} \in \text{DCFL} \setminus \text{DCFL}_{\text{ES}}$$

Why?

**Fact 4:  $\text{DCFL}_{\text{ES}} = \text{DCFL}$  having Prefix Property**

A language  $L$  is a  $\text{DCFL}_{\text{ES}}$  if and only if  $L$  is a DCFL having the **prefix property**.

**Definition (Prefix Property)**

A language  $L$  has the **prefix property** if and only if for any word  $w \in L$ , any proper prefix of  $w$  is not in  $L$ :

$$\forall x, y \in \Sigma^*. (xy \in L \wedge y \neq \epsilon \Rightarrow x \notin L)$$

For example, the following language is a **DCFL** but does **NOT** have the **prefix property**:

$$L = \{a^n b^n \mid n \geq 0\}$$

because  $\epsilon \in L$  is a prefix of  $aabb \in L$ .

Thus,  $L$  is a **DCFL** but **NOT** a  $\text{DCFL}_{\text{ES}}$ .

Fact 5:  $RL \not\subseteq DCFL_{ES}$ 

There exists a  $RL$  that is not a  $DCFL_{ES}$ .

- $RL \setminus DCFL_{ES} \neq \emptyset$ : For example, the following language is a  $RL$  but does **NOT** have the **prefix property**:

$$L = \{a^n \mid n \geq 0\}$$

because  $aa \in L$  is a prefix of  $aaaa \in L$ .

Thus,  $L$  is a  $RL$  but **NOT** a  $DCFL_{ES}$ .

$$L \in RL \setminus DCFL_{ES}$$

## Definition (Inherent Ambiguity)

A language  $L$  is **inherently ambiguous** if all CFGs whose languages are  $L$  are ambiguous. (i.e. there is no unambiguous grammar for  $L$ )

- **Is there any DCFL that is inherently ambiguous?**  
(i.e., is there any DCFL always having ambiguous grammars?)
- **Is there any DCFL that is not inherently ambiguous?**  
(i.e., is there any DCFL having at least one unambiguous grammar?)
- **Is there any inherently ambiguous language that is not a DCFL?**  
(i.e., is there any unambiguous grammar whose language is not a DCFL?)

Fact 6: DCFL  $\subset$  Non Inherently Ambiguous Languages

All DCFLs have at least one corresponding unambiguous grammars **BUT** there exists a non inherently ambiguous language that is not a DCFL.

- A DCFL<sub>ES</sub> has an unambiguous grammar: For a given DCFL<sub>ES</sub>  $L$  and its corresponding PDA  $P$ , we can define a CFG for  $L$  as follows:
  - For all  $0 \leq j < n$ ,

$$S \rightarrow A_{0,j}^Z$$

- For all  $q_i \in Q$ ,  $a \in \Sigma \cup \{\epsilon\}$ , and  $X \in \Gamma$ , consider any  $(q_j, X_1 \cdots X_m) \in \delta(q_i, a, X)$  and  $0 \leq k_1, \dots, k_m < n$ . Then,

$$A_{i,k_m}^X \rightarrow a A_{j,k_1}^{X_1} A_{k_1,k_2}^{X_2} \cdots A_{k_{m-1},k_m}^{X_m}$$

For any word  $w \in L$ ,  $w$  has a unique moves from the initial configuration to the final configuration in  $P$ . And, we know that:

$$A_{i,j}^X \Rightarrow^* w \quad \text{if and only if} \quad (q_i, w, X) \vdash^* (q_j, \epsilon, \epsilon)$$

Thus, the above CFG is **unambiguous**.

Fact 6: DCFL  $\subset$  Non Inherently Ambiguous Languages

All DCFLs have at least one corresponding unambiguous grammars **BUT** there exists a non inherently ambiguous language that is not a DCFL.

- A DCFL has an unambiguous grammar: For a given DCFL  $L$ , we can define another DCFL  $L'$  with a special symbol  $\$$  as follows:

$$L' = L\$ = \{w\$ \mid w \in L\}$$

Then,  $L'$  is a DCFL<sub>ES</sub> because it has the prefix property. Thus,  $L'$  has an unambiguous grammar  $G'$ . Now, we can define an unambiguous grammar  $G$  for  $L$  by treating  $\$$  as a variable with the rule  $\$ \rightarrow \epsilon$ .

For example, if the given DCFL is  $L = \{a^n b^n \mid n \geq 0\}$ , then  $L' = \{a^n b^n \$ \mid n \geq 0\}$  and  $G'$  is:

$$S \rightarrow X\$ \quad X \rightarrow aXb \mid \epsilon$$

Then, the unambiguous grammar  $G$  for  $L$  is:

$$S \rightarrow X\$ \quad X \rightarrow aXb \mid \epsilon \quad \$ \rightarrow \epsilon$$



**Fact 6: DCFL  $\subset$  Non Inherently Ambiguous Languages**

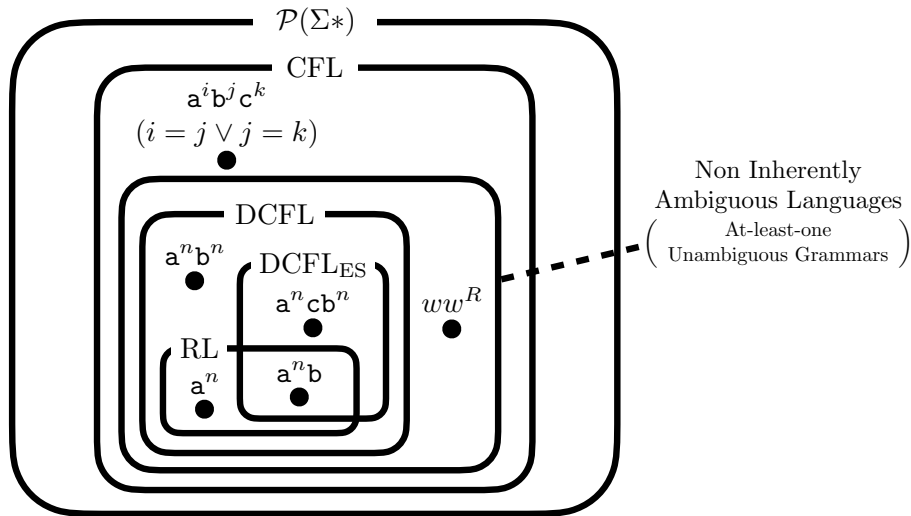
All DCFLs have at least one corresponding unambiguous grammars **BUT** there exists a non inherently ambiguous language that is not a DCFL.

- Non Inherently Ambiguous Languages \ DCFL  $\neq \emptyset$ : For example, the following language is a non inherently ambiguous language but not a DCFL:

$$L = \{ww^R \mid w \in \{a, b\}^*\}$$

because the following unambiguous grammar  $G$  represents  $L$ :

$$S \rightarrow aSa \mid bSb \mid \epsilon$$



- Normal Forms of Context-Free Grammars

Jihyeok Park

[jihyeok\\_park@korea.ac.kr](mailto:jihyeok_park@korea.ac.kr)

<https://plrg.korea.ac.kr>