Lecture 17 – Deterministic Pushdown Automata (DPDA)

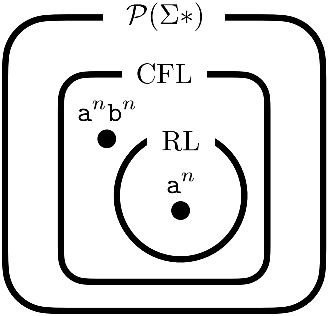
COSE215: Theory of Computation

Jihyeok Park

PLRG

2023 Spring

- A **pushdown automaton (PDA)** is an extension of ϵ-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?



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

$\parallel$

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

$\parallel$

CFG

Contents

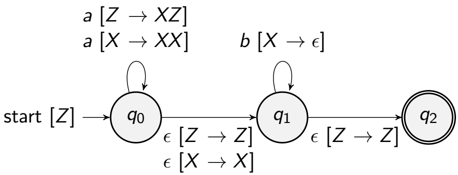
Definition of Deterministic Pushdown Automata

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 $\delta(q, \epsilon, X) \neq \varnothing$, then $\delta(q, a, X) = \varnothing$ for all $a \in \Sigma$.

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



Definition of Deterministic Pushdown Automata

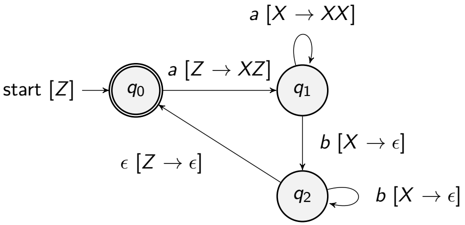
PLRG

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 $\delta(q, \epsilon, X) \neq \varnothing$, then $\delta(q, a, X) = \varnothing$ for all $a \in \Sigma$.

For example, the following PDA is deterministic:



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

Deterministic Context-Free Languages (DCFLs)

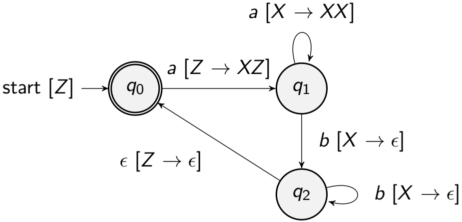
Definition (Deterministic Context-Free Languages (DCFLs))

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

Fact 2: RL $\subset$ DCFL

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

- $\boxed{\text{RL} \subseteq \text{DCFL}}$: 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,

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

- $\boxed{\text{DCFL} \setminus \text{RL} \neq \varnothing}$: We already know that the following language is a DCFL but not an RL:

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

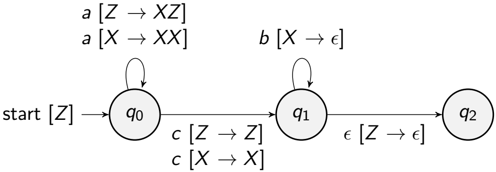
Definition (DCFL_{ES})

A language L is a **deterministic context-free language by empty stacks (DCFL_{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 DCFL_{ES}:

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

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

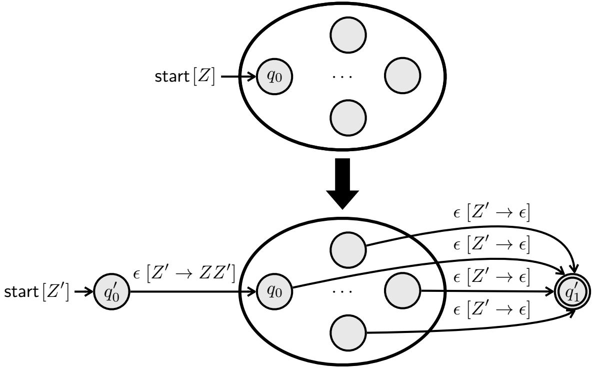


Fact 3: $\text{DCFL}_{\text{ES}} \subset \text{DCFL}$

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

All DCFL_{ES}s are DCFLs **BUT** there exists a DCFL that is not a DCFL_{ES}.

- $\boxed{\text{DCFL}_{\text{ES}} \subseteq \text{DCFL}}$: For a given DCFL_{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:



$\mathrm{DCFL}_{\mathrm{ES}} \subset \mathrm{DCFL}$

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

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

- $\boxed{\mathrm{DCFL} \setminus \mathrm{DCFL}_{\mathrm{ES}} \neq \varnothing}$: The following language is a DCFL but not a $\mathrm{DCFL}_{\mathrm{ES}}$:

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

Why?

Fact 4: DCFL_{ES} = DCFL having Prefix Property

Fact 4: DCFL_{ES} = DCFL having Prefix Property

A language L is a DCFL_{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:

$$\forall x, y \in \Sigma^*. \; ((xy \in L \land y \neq \epsilon) \implies x \notin L)$$

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 proper prefix of $\mathtt{ab} \in L$.
Thus, L is a **DCFL** but **NOT** a $\textbf{DCFL}_{\textbf{ES}}$.

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

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

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

- $\boxed{\text{RL}\setminus\text{DCFL}_{\text{ES}} \neq \varnothing}$: 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 proper prefix of aaaa $\in L$.
Thus, L is a **RL** but **NOT** a **DCFL$_{\text{ES}}$**.

$$L \in \text{RL} \setminus \text{DCFL}_{\text{ES}}$$

Inherent Ambiguity of DCFLs

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.

- $\boxed{\text{A DCFL}_{\text{ES}} \text{ has an unambiguous grammar}}$: For a given DCFL$_{\text{ES}}$ L and its corresponding PDA P, we can define a CFG for L as follows:
  - For all $0 \leq j < n$,
  $$S \to A^{Z}_{0,j}$$
  - 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 $0 \leq k_1, \cdots, k_m < n$. Then,
  $$A^{X}_{i,k_m} \to a\ A^{X_1}_{j,k_1}\ A^{X_2}_{k_1,k_2} \cdots A^{X_m}_{k_{m-1},k_m}$$

  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:
  $$A^{X}_{i,j} \Rightarrow^* w \qquad \text{if and only if} \qquad (q_i, w, X) \vdash^* (q_j, \epsilon, \epsilon)$$
  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.

- $\boxed{\text{A DCFL has an unambiguous grammar}}$: 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_{ES} 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:

$$S \rightarrow X\$ \qquad X \rightarrow \mathtt{a}X\mathtt{b} \mid \epsilon$$

Then, the unambiguous grammar G for L is:

$$S \rightarrow X\$ \qquad X \rightarrow \mathtt{a}X\mathtt{b} \mid \epsilon \qquad \$ \rightarrow \epsilon$$

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.

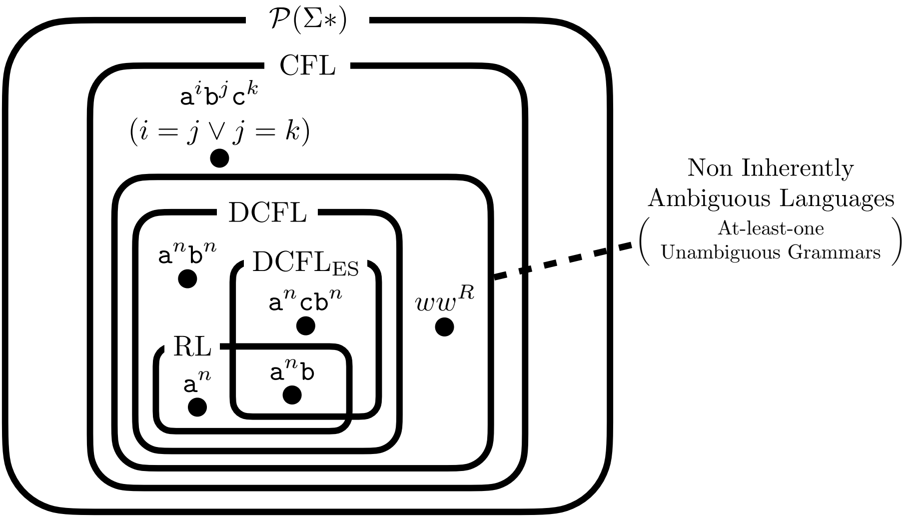
- $\boxed{\text{Non Inherently Ambiguous Languages} \setminus \text{DCFL} \neq \varnothing}$: For example, the following language is a non inherently ambiguous language but not a DCFL:

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

because the following unambiguous grammar G represents L:

$$S \to aSa \mid bSb \mid \epsilon$$

Summary



$\mathcal{P}(\Sigma*)$
CFL
$\mathtt{a}^i\mathtt{b}^j\mathtt{c}^k$
$(i = j \vee j = k)$
DCFL
$\mathtt{a}^n\mathtt{b}^n$
DCFL_{ES}
$\mathtt{a}^n\mathtt{c}\mathtt{b}^n$
RL
$\mathtt{a}^n$
$\mathtt{a}^n\mathtt{b}$
ww^R
Non Inherently
Ambiguous Languages
At-least-one
Unambiguous Grammars

Next Lecture

PLRG

- Normal Forms of Context-Free Grammars

Jihyeok Park
jihyeok_park@korea.ac.kr
https://plrg.korea.ac.kr