PDA design

Example:

Consider $L = \{s0s^R \mid s \in \{1,2\}^*\}$, s^R is s in reversed order.

Design a PDA to recognize L.

Equivalence of PDA and CFL

Theorem. Let $L \subseteq \Sigma^*$ be a language.
L is a CFL iff there is a PDA $\mathbf{G}$ s.t. $L_a(\mathbf{G}) = L$

Conclusion:
CFL $\Leftrightarrow$ CFG $\Leftrightarrow$ PDA

CFL — RL
PDA — FA

Equivalence of PDA and CFL

Let's prove: if $L \subseteq \Sigma^*$ is a CFL,
then there is a PDA $\mathbf{G}$ s.t. $L_a(\mathbf{G}) = L$

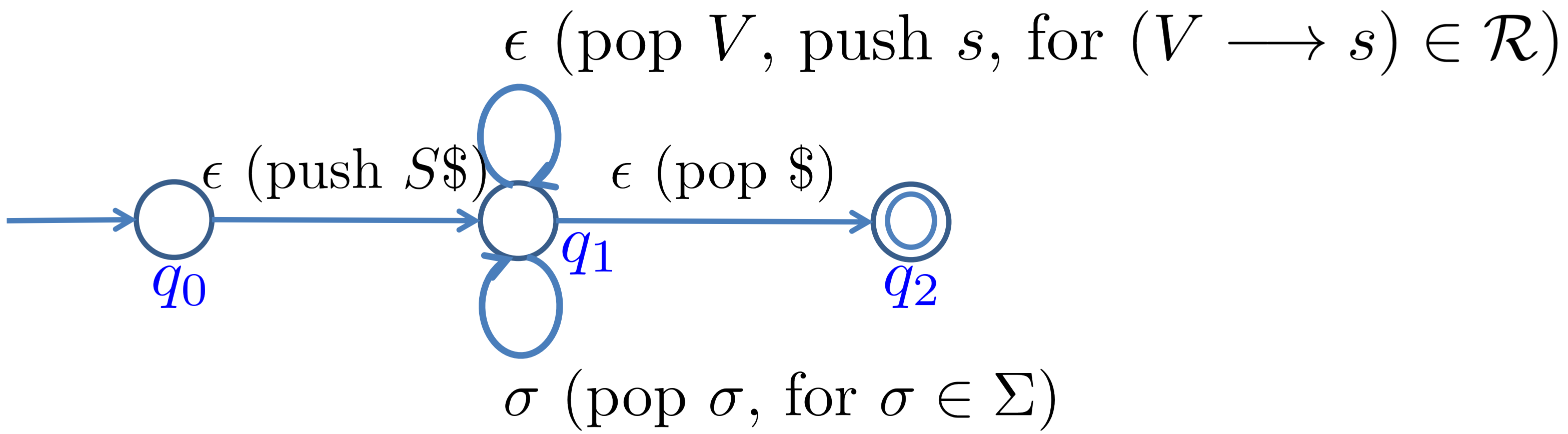
Proof: Let $L \subseteq \Sigma^*$ be a CFL.

Then by defn. there is a CFG $G = (\mathcal{V}, \Sigma, \mathcal{R}, S)$ s.t. $L(G) = L$

Strategy: construct a PDA $\mathbf{G}$ s.t. $L_a(\mathbf{G}) = L(G)$

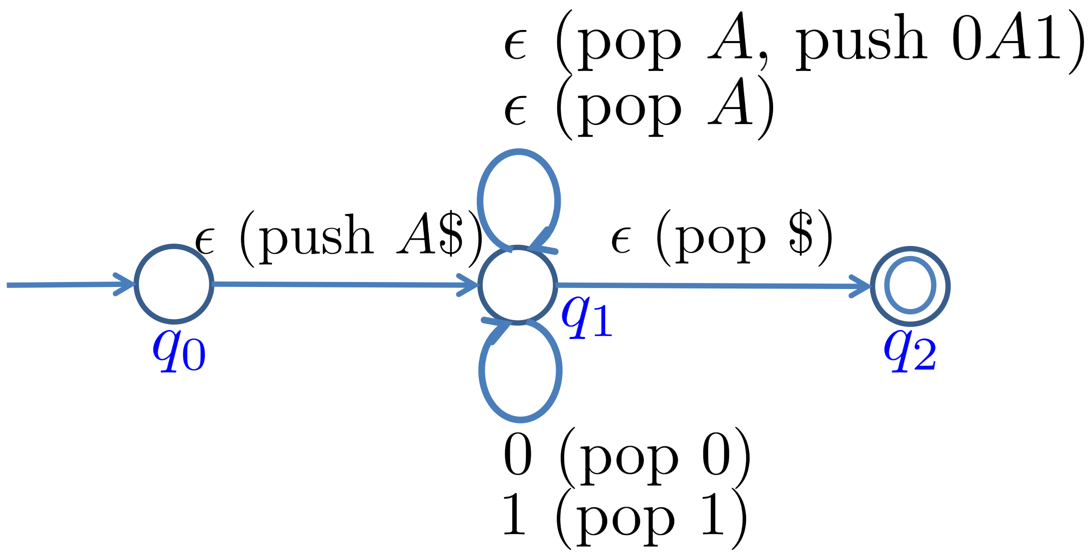
Equivalence of PDA and CFL

Given CFG $G = (\mathcal{V}, \Sigma, \mathcal{R}, S)$, construct PDA:



$Q = \{q_0, q_1, q_2\}$,
Σ,
$\Gamma = \{\$\} \cup \Sigma \cup \mathcal{V}$,
$\delta : Q \times (\Sigma \cup \{\epsilon\}) \times (\Gamma \cup \{\epsilon\}) \rightarrow Pwr(Q \times (\Gamma \cup \{\epsilon\}))$
q_0,
$Q_a = \{q_2\}$

Equivalence of PDA and CFL



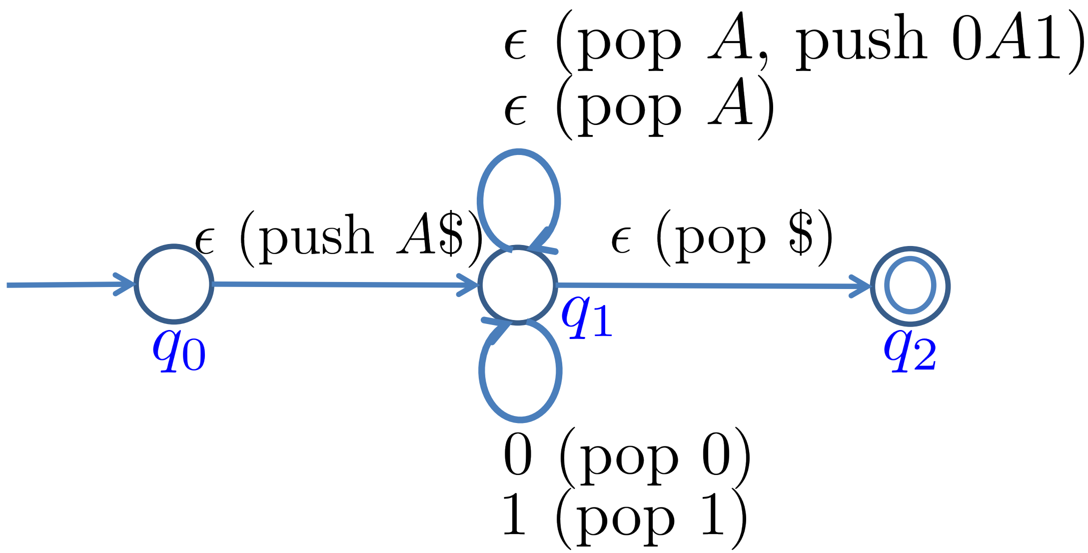
e.g. (CFG) G: $A \longrightarrow 0A1$

$A \longrightarrow \epsilon$

Consider a derivation

$A \longrightarrow^* s = 0011$

Equivalence of PDA and CFL



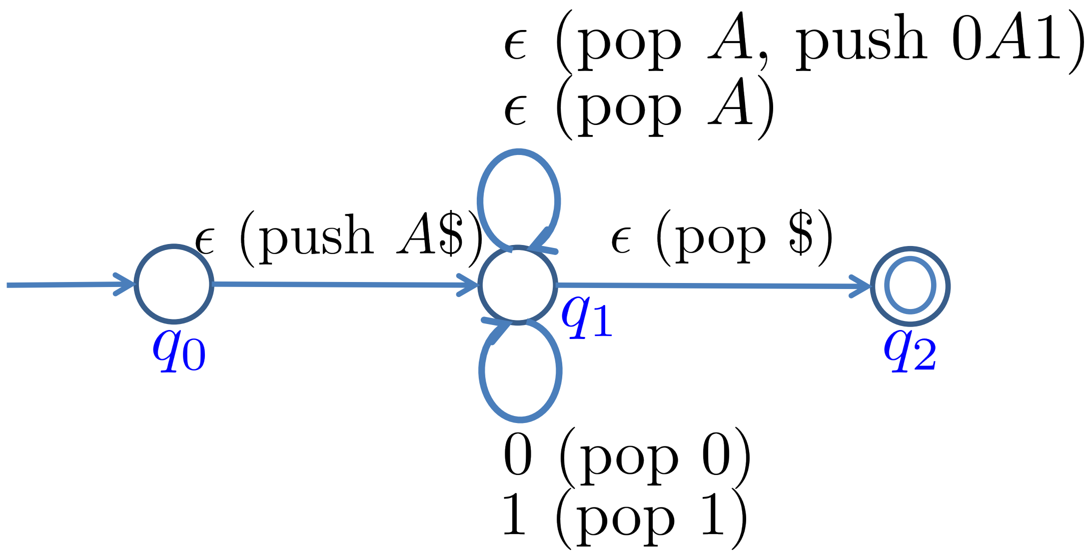
e.g. (CFG) G: $A \longrightarrow 0A1$

$A \longrightarrow \epsilon$

Consider $s = 01$ s.t.

$(q_2, \cdot) \in \delta(q_0, s, \cdot)$

Equivalence of PDA and CFL



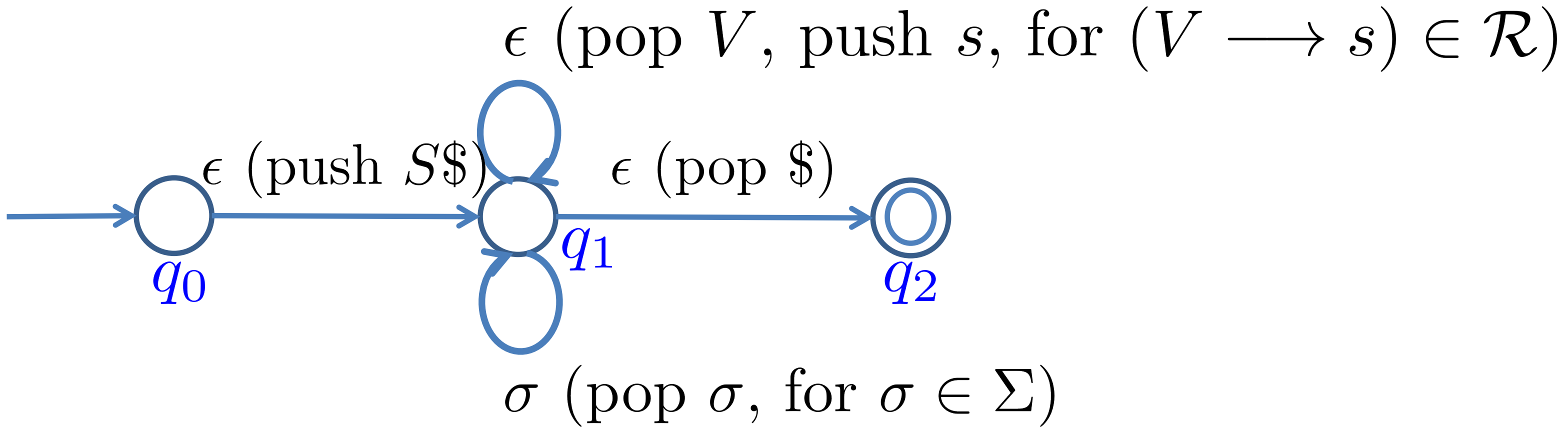
e.g. (CFG) G:

$$A \longrightarrow 0A1$$
$$A \longrightarrow \epsilon$$

In fact: $s \in L(G)$ iff $s \in L_a(\mathbf{G})$

Equivalence of PDA and CFL

Given CFG $G = (\mathcal{V}, \Sigma, \mathcal{R}, S)$, construct PDA:



PDA $\mathbf{G} = (Q, \Sigma, \Gamma, \delta, q_0, Q_a)$

Claim: in general $s \in L(G)$ iff $s \in L_a(\mathbf{G})$
(proof in Sipser's book, pp.118-119)

So $L(G) = L_a(\mathbf{G})$

Equivalence of PDA and CFL

Theorem. $L \subseteq \Sigma^*$ is a CFL iff
there is a PDA $\mathbf{G}$ s.t. $L_a(\mathbf{G}) = L$

So we have proved: if $L \subseteq \Sigma^*$ is a CFL,
then there is a PDA $\mathbf{G}$ s.t. $L_a(\mathbf{G}) = L$

The reverse direction: if $\mathbf{G}$ is a PDA,
then $L_a(\mathbf{G})$ is context-free.

(proof in Sipser's book, pp.121-122)

Pumping Lemma
(Context-Free Languages)

So far

We have studied context-free languages,
and the equivalent context-free grammar, PDA

Are there any languages that are not context-free?
Or is there a limitation of push-down automata?

Consider $L = \{\alpha^n \beta^n \gamma^n \mid n = 0, 1, \ldots\}$

Is it possible to design a PDA to recognize L?

Pumping lemma

If L is a context-free language, then there is a number p (pumping length) s.t. $(\forall s \in L)|s| \geq p \Rightarrow (s = uvxyz) \wedge$

$$((\forall i \geq 0)uv^i xy^i z \in L) \wedge$$
$$(|vy| > 0) \wedge$$
$$(|vxy| \leq p)$$

v^i: $v^0 = \epsilon$, $v^1 = v$, $v^2 = vv, \ldots$

y^i: $y^0 = \epsilon$, $y^1 = y$, $y^2 = yy, \ldots$

$|vy| > 0$: either $v \neq \epsilon$ or $y \neq \epsilon$

$|vxy| \leq p$: v, x, y together has length at most p

Pumping lemma is a *necessary* condition for CFL
(Proof p.127)

Pumping lemma

Now we use pumping lemma to show
$L = \{\alpha^n \beta^n \gamma^n \mid n = 0, 1, \ldots\}$ is not a context-free language

Assume on the contrary L *is* context-free.
Then by pumping lemma, there is a pumping length p s.t.

$$
\begin{aligned}
(\forall s \in L)|s| \geq p \Rightarrow (s = uvxyz)\wedge & \\
((\forall i \geq 0)uv^ixy^iz \in L)\wedge & \quad \cdots \text{ condition 1)} \\
(|vy| > 0)\wedge & \quad \cdots \text{ condition 2)} \\
(|vxy| \leq p) & \quad \cdots \text{ condition 3)}
\end{aligned}
$$

Consider the string $s = \alpha^p \beta^p \gamma^p$.
Since $s \in L$ and $|s| > p$, s can be split into u, v, x, y, z
satisfying the three conditions

We consider two cases to show this is impossible.

Pumping lemma

Case 1: Both v and y contain only one typle of symbol:
i.e. v and y do not contain both α and β, or both β and γ
Then string $uvvxyyz$ cannot contain equal numbers of α, β, γ
and $uv^2xy^2z \notin L$. This violates condition 1).

Case 2: Either v or y contains more than one typle of symbol:
i.e. v or y contains both α and β, or both β and γ
Then string $uvvxyyz$ has β before α, or γ before β
and $uv^2xy^2z \notin L$. This violates condition 1).

Conclusion: $L = \{\alpha^n\beta^n\gamma^n \mid n = 0, 1, \ldots\}$ is not context-free

Example

Use pumping lemma to show
$L = \{s0s \mid s \in \{1,2\}^*\}$ is not a context-free language

Assume on the contrary L *is* context-free.
Then by pumping lemma, there is a pumping length p s.t.

$$(\forall s \in L)|s| \geq p \Rightarrow (s = uvxyz) \wedge$$
$$((\forall i \geq 0) uv^i xy^i z \in L) \wedge \quad \cdots \text{ condition 1)}$$
$$(|vy| > 0) \wedge \quad \cdots \text{ condition 2)}$$
$$(|vxy| \leq p) \quad \cdots \text{ condition 3)}$$

Consider the string $s = 1^p 2^p 0 1^p 2^p$.
Since $s \in L$ and $|s| > p$, s can be split into u, v, x, y, z satisfying the three conditions

Pumping lemma

Consider substring vxy. By condition 3) the length is at most p.

Note: substring vxy must cross the midpoint 0:
i) If vxy occurs before 0 (i.e. 1^p2^p)
then in $uvvxyyz$ strings before and after 0 have different lengths and $uv^2xy^2z \notin L$. This violates condition 1).
ii) If vxy occurs after 0 (i.e. 1^p2^p)
then in $uvvxyyz$ strings before and after 0 have different lengths and $uv^2xy^2z \notin L$. This violates condition 1).

But if substring vxy crosses the midpoint 0,
then $uxz(i = 0)$ has the form $1^p2^j01^k2^p$ (or $1^p2^j1^k2^p$),
where j, k cannot both be p.
and $uv^0xy^0z \notin L$. This violates condition 1).

Conclusion: $L = \{s0s \mid s \in \{1, 2\}^*\}$ is not context-free