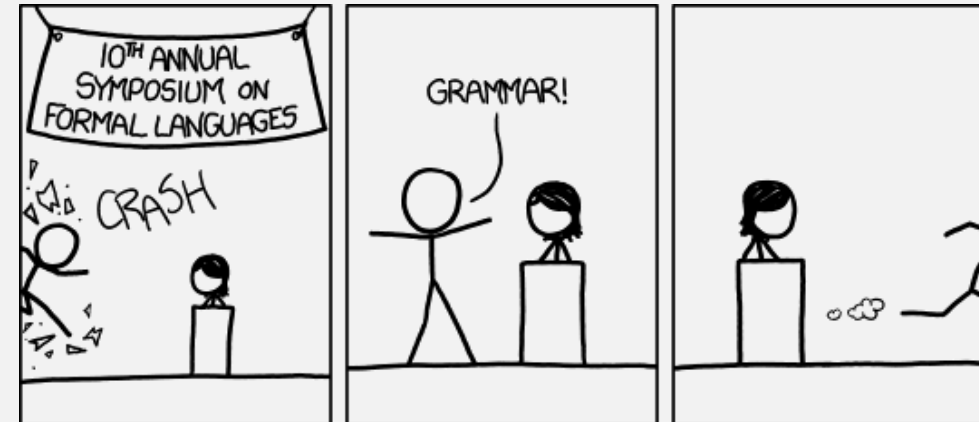


UMB CS 622

Pushdown Automata (PDAs)

Wednesday, March 20, 2024



Announcements

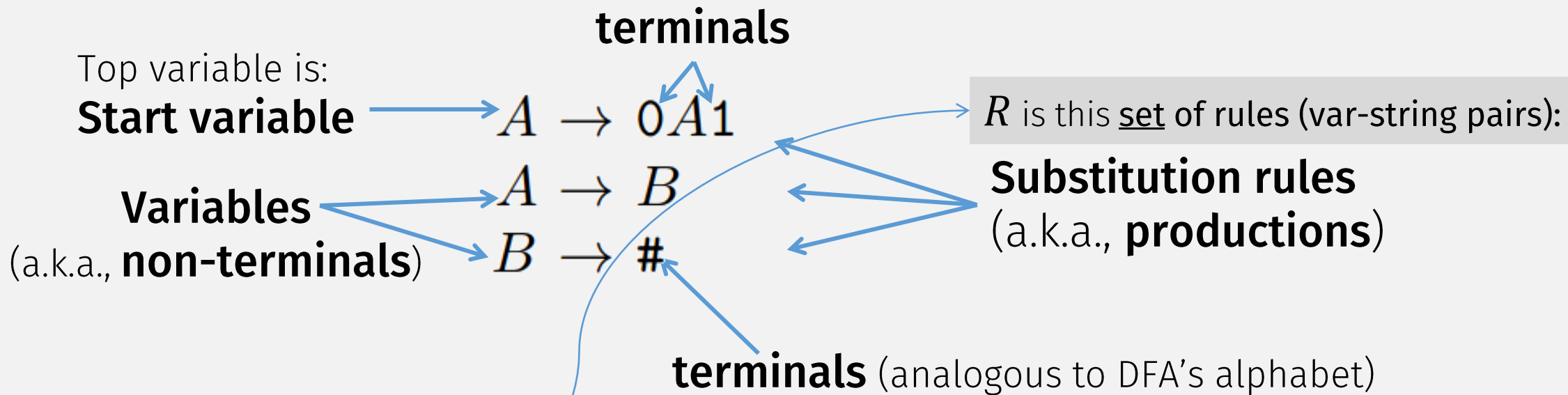
- HW 5 out
 - Due Mon 3/25 12pm noon



Last Time:

Context-Free Grammar (CFG)

Grammar $G_1 = (V, \Sigma, R, S)$



A *context-free grammar* is a 4-tuple (V, Σ, R, S) where

1. V is a finite set called the *variables*,
2. Σ is a finite set, disjoint from V , called the *terminals*,
3. R is a finite set of *rules*, with each rule being a variable and a string of variables and terminals, and
4. $S \in V$ is the start variable.

$$V = \{A, B\},$$

$$\Sigma = \{0, 1, \#\},$$

$$S = A,$$

Last Time:

Generating Strings with a CFG

Grammar $G_1 = (V, \Sigma, R, S)$

$$A \rightarrow 0A1$$

$$A \rightarrow B$$

$$B \rightarrow \#$$

Strings in CFG's language
= all possible **generated** / **derived** strings

$$L(G_1) \text{ is } \{0^n \# 1^n \mid n \geq 0\}$$

A CFG **generates** a string, by repeatedly applying substitution rules:

Example:

$$A \Rightarrow 0A1 \Rightarrow 00A11 \Rightarrow 000A111 \Rightarrow 000B111 \Rightarrow 000\#111$$

This sequence of steps is called a **derivation**

Last Time:

Derivations: Formally

Let $G = (V, \Sigma, R, S)$

Single-step

$$\alpha A \beta \xRightarrow{G} \alpha \gamma \beta$$

Where:

$\alpha, \beta \in (V \cup \Sigma)^*$ ← sequence of terminals or variables

$A \in V$ ← Variable

$A \rightarrow \gamma \in R$ ← Rule

A *context-free grammar* is a 4-tuple (V, Σ, R, S) , where

1. V is a finite set called the *variables*,
2. Σ is a finite set, disjoint from V , called the *terminals*,
3. R is a finite set of *rules*, with each rule being a variable and a string of variables and terminals, and
4. $S \in V$ is the start variable.

Last Time:

Derivations: Formally

Let $G = (V, \Sigma, R, S)$

Single-step

$$\alpha A \beta \xRightarrow{G} \alpha \gamma \beta$$

Where:

$$\alpha, \beta \in (V \cup \Sigma)^* \leftarrow \begin{array}{l} \text{sequence of} \\ \text{terminals or variables} \end{array}$$

$$A \in V \leftarrow \begin{array}{l} \text{Variable} \end{array}$$

$$A \rightarrow \gamma \in R \leftarrow \begin{array}{l} \text{Rule} \end{array}$$

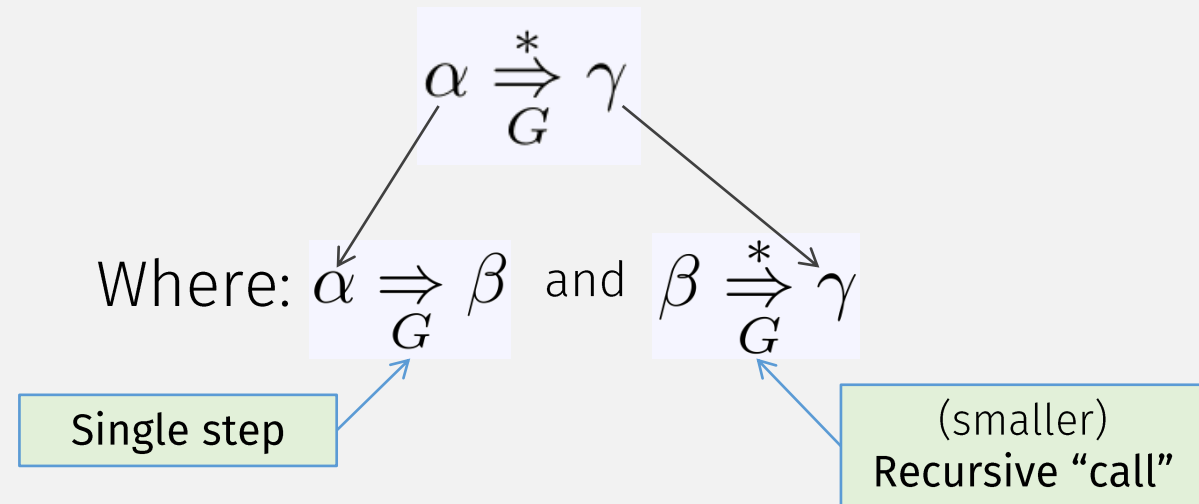
A *context-free grammar* is a 4-tuple (V, Σ, R, S) , where

1. V is a finite set called the *variables*,
2. Σ is a finite set, disjoint from V , called the *terminals*,
3. R is a finite set of *rules*, with each rule being a variable and a string of variables and terminals, and
4. $S \in V$ is the start variable.

Multi-step (recursively defined)

Base case: $\alpha \xRightarrow{*}_G \alpha$ (0 steps)

Recursive case: (> 0 steps)



Last Time:

Formal Definition of a CFL

A *context-free grammar* is a 4-tuple (V, Σ, R, S) , where

1. V is a finite set called the *variables*,
2. Σ is a finite set, disjoint from V , called the *terminals*,
3. R is a finite set of *rules*, with each rule being a variable and a string of variables and terminals, and
4. $S \in V$ is the start variable.

$$G = (V, \Sigma, R, S)$$

“the **language** of a grammar G is ...”

“... all possible sequences of terminal symbols (i.e., **strings**) ...”

“... that can be **generated** with rules of grammar G ”

$$L(G) = \left\{ w \in \Sigma^* \mid S \xRightarrow[G]{*} w \right\}$$

If a **CFG** generates all strings in a language L , then L is a **context-free language** (CFL)

Designing Grammars : Basics

1. Think about what you want to “link” together

- E.g., $0^n 1^n$
 - $A \rightarrow 0A1$
 - # 0s and # 1s are “linked”
- E.g., XML
 - ELEMENT \rightarrow \langle TAG \rangle CONTENT \langle /TAG \rangle
 - Start and end tags are “linked”

2. Start with small grammars and then combine

- just like with FSMs, and programming!

Designing Grammars: Building Up

- Start with small grammars and then combine (just like FSMs)
 - To create a grammar for the language $\{0^n 1^n \mid n \geq 0\} \cup \{1^n 0^n \mid n \geq 0\}$
 - First create grammar for lang $\{0^n 1^n \mid n \geq 0\}$:
$$S_1 \rightarrow 0S_1 1 \mid \epsilon$$
 - Then create grammar for lang $\{1^n 0^n \mid n \geq 0\}$:
$$S_2 \rightarrow 1S_2 0 \mid \epsilon$$
 - Then combine: $S \rightarrow S_1 \mid S_2$
$$S_1 \rightarrow 0S_1 1 \mid \epsilon$$
$$S_2 \rightarrow 1S_2 0 \mid \epsilon$$
 - ← New start variable and rule combines two smaller grammars
 - “|” = “or” = union (combines 2 rules with same left side)

(Closed) Operations for CFLs?

- Start with small grammars and then combine (just like FSMs)

- “Or”:

$$S \rightarrow S_1 \mid S_2$$

- “Concatenate”: $S \rightarrow S_1 S_2$

- “Repetition”: $S' \rightarrow S' S_1 \mid \epsilon$

Could you write out
the full proof of
these closure
operations?

Example: Creating CFG

alphabet Σ is $\{0,1\}$

$\{w \mid w \text{ starts and ends with the same symbol}\}$

1) come up with examples: In the language: **010, 101, 11011** **1, 0 ?**
Not in the language: **10, 01, 110** $\epsilon ?$

2) Create CFG:

Needed Rules:

$S \rightarrow \mathbf{0M0} \mid \mathbf{1M1} \mid \mathbf{0} \mid \mathbf{1}$ “start/end symbol are “linked” (ie, same); middle can be anything”

$M \rightarrow MT \mid \epsilon$ “middle: all possible terminals, repeated (ie, all possible strings)”

$T \rightarrow \mathbf{0} \mid \mathbf{1}$ “all possible terminals”

3) Check CFG: generates examples in the language; does not generate examples not in language

Regular Language vs CFL Comparison

Regular Languages	Context-Free Languages (CFLs)
Regular Expression	Context-Free Grammar (CFG)
<u>describes</u> a Regular Lang	<u>describes</u> a CFL

Regular Language vs CFL Comparison

Regular Languages	Context-Free Languages (CFLs)
Regular Expression	Context-Free Grammar (CFG)
<u>describes</u> a Regular Lang	<u>describes</u> a CFL
Finite State Automaton (FSM)	???
<u>recognizes</u> a Regular Lang	<u>recognizes</u> a CFL

Regular Language vs CFL Comparison

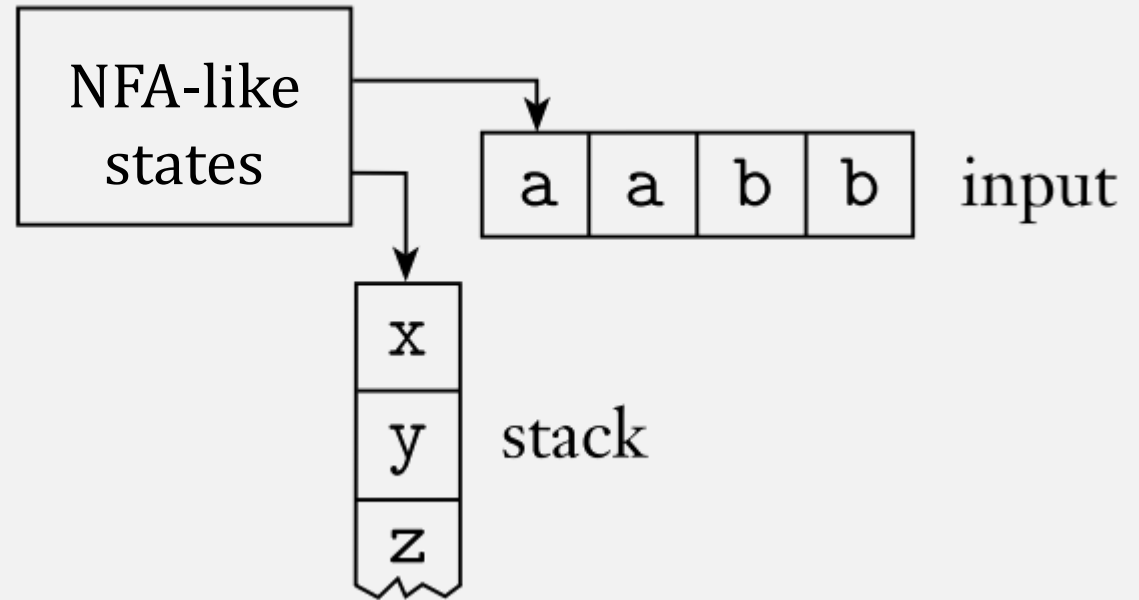
	Regular Languages	Context-Free Languages (CFLs)	
thm	Regular Expression <u>describes</u> a Regular Lang	Context-Free Grammar (CFG) <u>describes</u> a CFL	def
def	Finite State Automaton (FSM) <u>recognizes</u> a Regular Lang	Push-down Automata (PDA) <u>recognizes</u> a CFL	thm

Regular Language vs CFL Comparison

	Regular Languages	Context-Free Languages (CFLs)	
thm	Regular Expression <u>describes</u> a Regular Lang	Context-Free Grammar (CFG) <u>describes</u> a CFL	def
def	Finite State Automaton (FSM) <u>recognizes</u> a Regular Lang	Push-down Automata (PDA) <u>recognizes</u> a CFL	thm
	<u>Proved:</u>	<u>Proved:</u>	
	Regular Lang \Leftrightarrow Regular Expr	CFL \Leftrightarrow PDA	

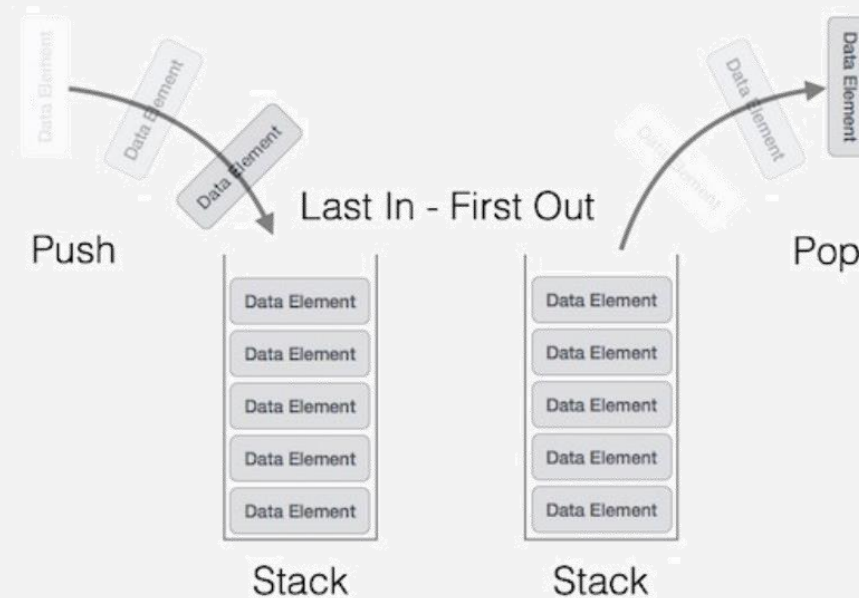
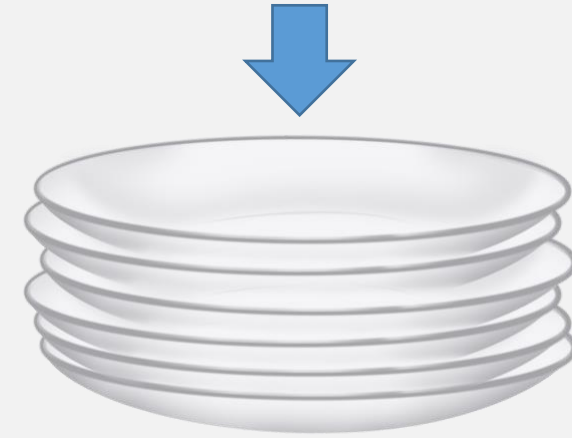
Pushdown Automata (PDA)

PDA = NFA + a stack



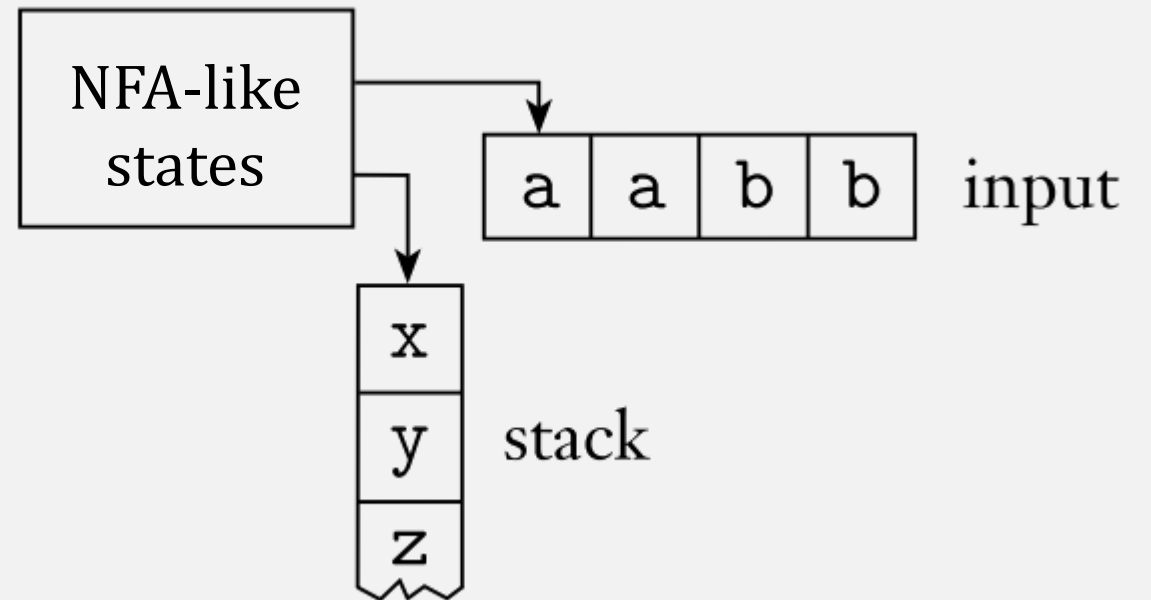
What is a Stack?

- A restricted kind of (infinite!) memory
- Access to top element only
- 2 Operations only: push, pop



Pushdown Automata (PDA)

- **PDA = NFA + a stack**
 - Infinite memory
 - read/write top location only
 - Push/pop



$$\{0^n 1^n \mid n \geq 0\}$$

An Example PDA

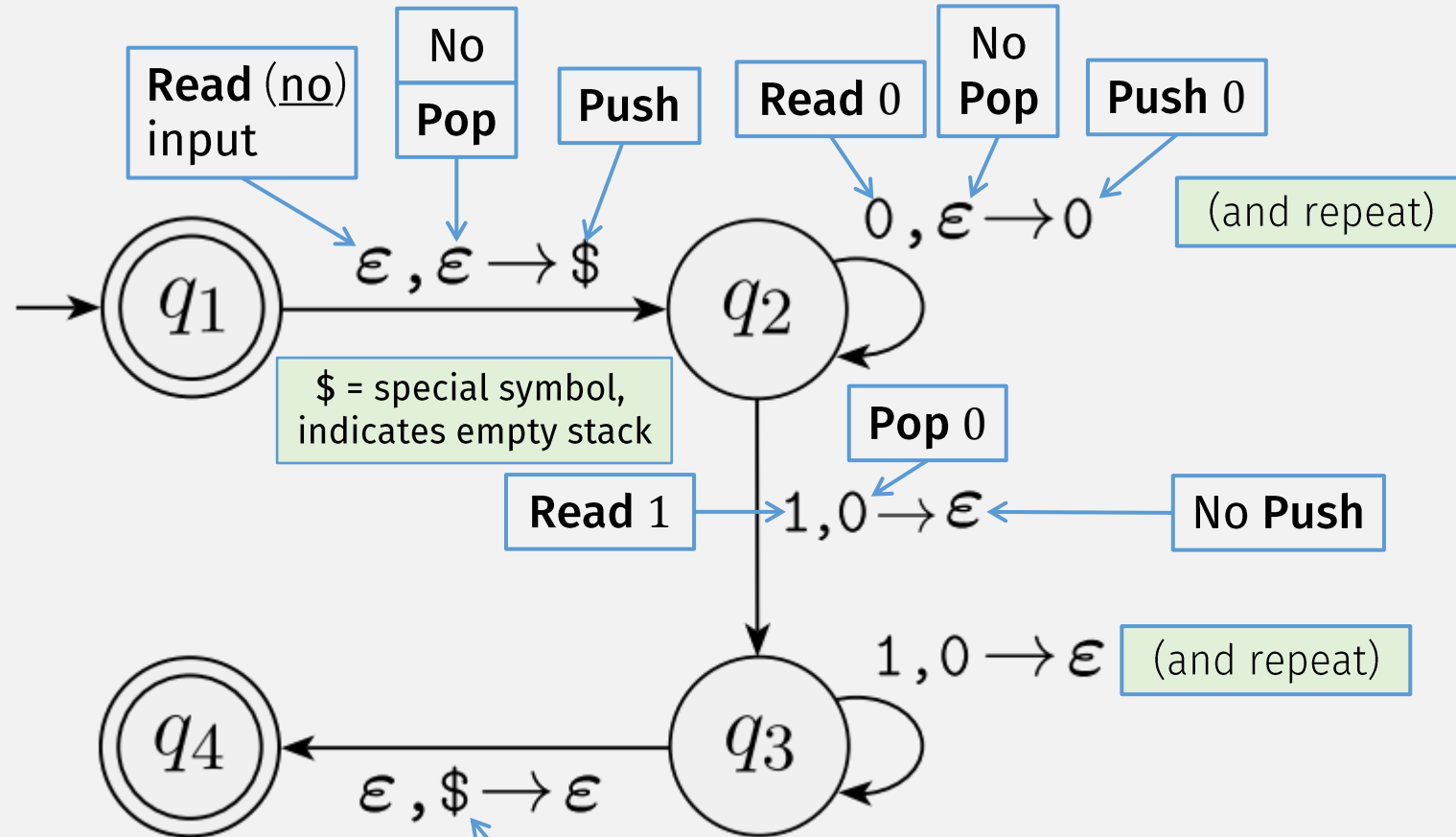
A PDA transition has 3 parts:

- **Read**

- **Pop**

- **Push**

- **Push**



This machine can only **pop \$** (and **accept**) when **stack is empty**, i.e., when $\# 0s = \# 1s$

Formal Definition of PDA

A *pushdown automaton* is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$, where Q, Σ, Γ , and F are all finite sets, and

1. Q is the set of states,
2. Σ is the input alphabet,
3. Γ is the stack alphabet,
4. $\delta: Q \times \Sigma_\epsilon \times \Gamma_\epsilon \longrightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ is the transition function,
5. $q_0 \in Q$ is the start state, and
6. $F \subseteq Q$ is the set of accept states.

Stack alphabet has special stack symbols, e.g., \$

Input Pop Push

Non-deterministic!
Result of a step is **set** of (STATE, STACK CHAR) pairs

$$Q = \{q_1, q_2, q_3, q_4\},$$

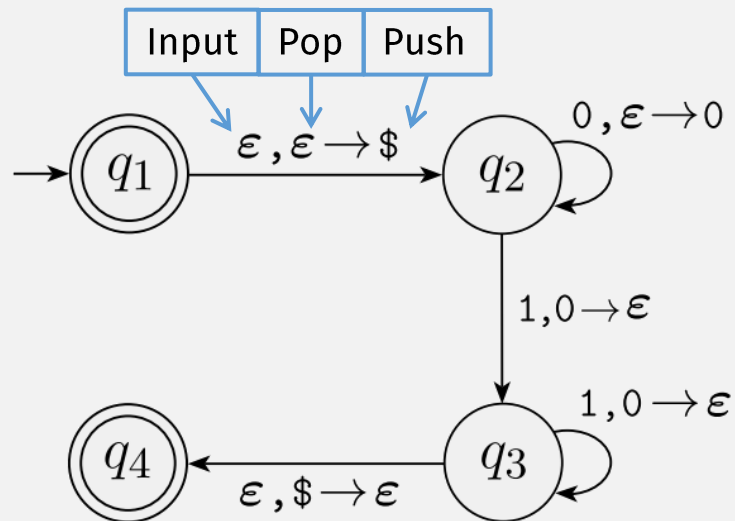
PDA Formal Definition Example

$$\Sigma = \{0, 1\},$$

$$\Gamma = \{0, \$\},$$

Stack alphabet has special stack symbol \$

$$F = \{q_1, q_4\},$$



A *pushdown automaton* is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$, where Q , Σ , Γ , and F are all finite sets, and

1. Q is the set of states,
2. Σ is the input alphabet,
3. Γ is the stack alphabet,
4. $\delta: Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ is the transition function,
5. $q_0 \in Q$ is the start state, and
6. $F \subseteq Q$ is the set of accept states.

Input

Pop

Push

$$Q = \{q_1, q_2, q_3, q_4\},$$

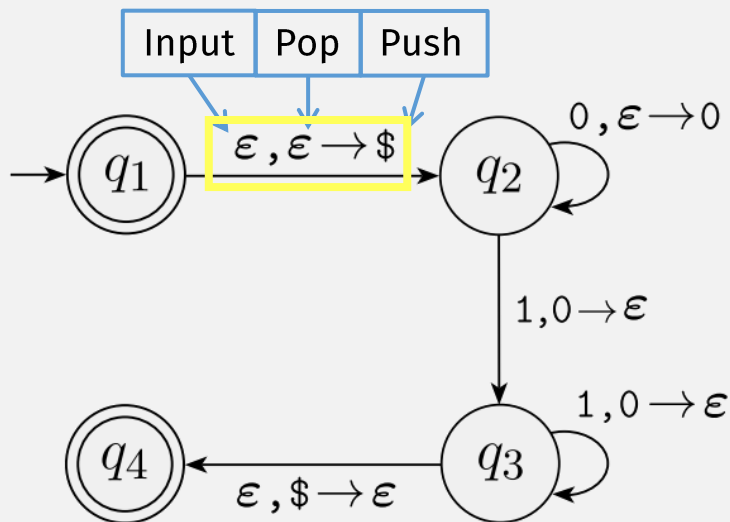
$$\Sigma = \{0, 1\},$$

$$\Gamma = \{0, \$\},$$

$$F = \{q_1, q_4\}, \text{ and}$$

δ is given by the following table, wherein blank entries signify \emptyset .

Input:	0			1			ϵ		
Stack:	0	\$	ϵ	0	\$	ϵ	0	\$	ϵ
q_1									$\{(q_2, \$)\}$
q_2			$\{(q_2, 0)\}$			$\{(q_3, \epsilon)\}$			
q_3			1			$\{(q_3, \epsilon)\}$			
q_4								4	$\{(q_4, \epsilon)\}$



A **pushdown automaton** is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$, where Q, Σ, Γ , and F are all finite sets, and

- Q is the set of states,
- Σ is the input alphabet,
- Γ is the stack alphabet,
- $\delta: Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ is the transition function,
- $q_0 \in Q$ is the start state, and
- $F \subseteq Q$ is the set of accept states.

$$Q = \{q_1, q_2, q_3, q_4\},$$

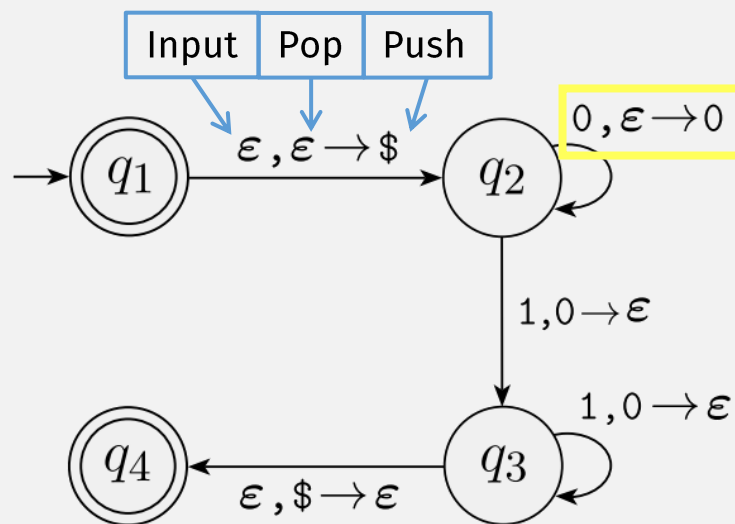
$$\Sigma = \{0, 1\},$$

$$\Gamma = \{0, \$\},$$

$$F = \{q_1, q_4\}, \text{ and}$$

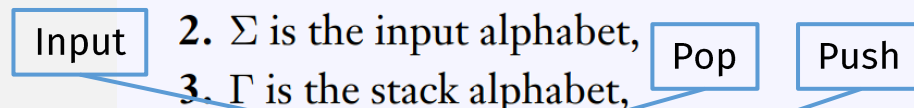
δ is given by the following table, wherein blank entries signify \emptyset .

Input:	0			1			ϵ		
Stack:	0	\$	ϵ	0	\$	ϵ	0	\$	ϵ
q_1									$\{(q_2, \$)\}$
q_2	$\{(q_2, 0)\}$			$\{(q_3, \epsilon)\}$					
q_3				$\{(q_3, \epsilon)\}$					$\{(q_4, \epsilon)\}$
q_4									



A **pushdown automaton** is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$, where $Q, \Sigma, \Gamma,$ and F are all finite sets, and

- Q is the set of states,
- Σ is the input alphabet,
- Γ is the stack alphabet,
- $\delta: Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ is the transition function,
- $q_0 \in Q$ is the start state, and
- $F \subseteq Q$ is the set of accept states.



$$Q = \{q_1, q_2, q_3, q_4\},$$

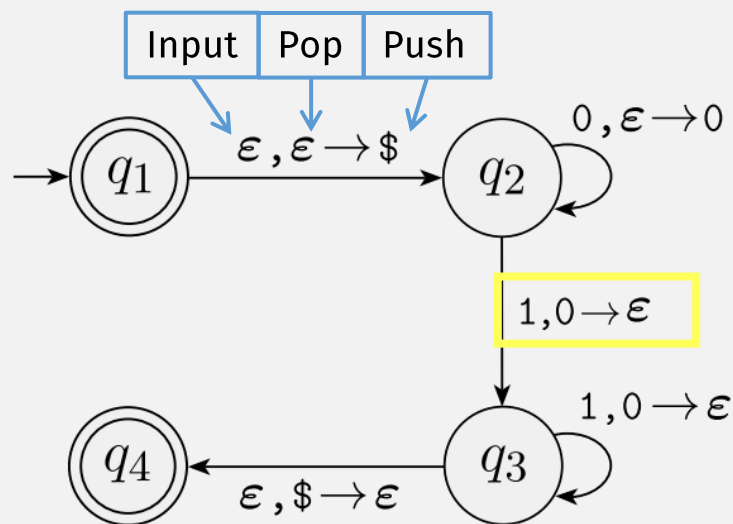
$$\Sigma = \{0, 1\},$$

$$\Gamma = \{0, \$\},$$

$$F = \{q_1, q_4\}, \text{ and}$$

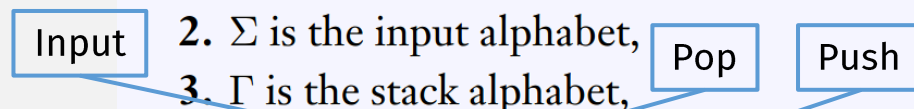
δ is given by the following table, wherein blank entries signify \emptyset .

Input:	0			1			ϵ		
Stack:	0	\$	ϵ	0	\$	ϵ	0	\$	ϵ
q_1									
q_2			$\{(q_2, 0)\}$			$\{(q_3, \epsilon)\}$			
q_3			1			$\{(q_3, \epsilon)\}$			
q_4									$\{(q_4, \epsilon)\}$



A **pushdown automaton** is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$, where Q, Σ, Γ , and F are all finite sets, and

- Q is the set of states,
- Σ is the input alphabet,
- Γ is the stack alphabet,
- $\delta: Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ is the transition function,
- $q_0 \in Q$ is the start state, and
- $F \subseteq Q$ is the set of accept states.



$$Q = \{q_1, q_2, q_3, q_4\},$$

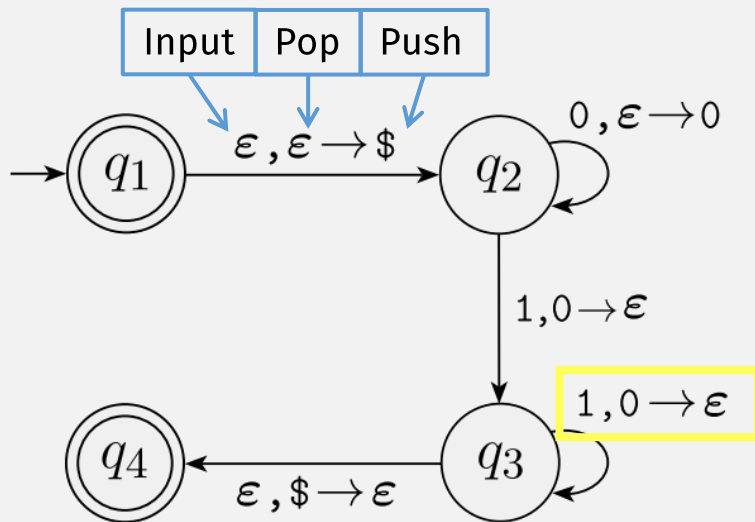
$$\Sigma = \{0,1\},$$

$$\Gamma = \{0, \$\},$$

$$F = \{q_1, q_4\}, \text{ and}$$

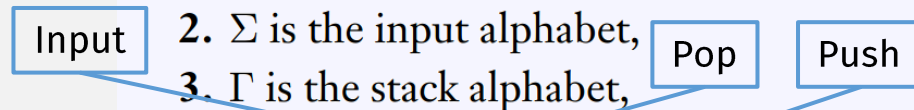
δ is given by the following table, wherein blank entries signify \emptyset .

Input:	0			1			ϵ		
Stack:	0	\$	ϵ	0	\$	ϵ	0	\$	ϵ
q_1									
q_2			$\{(q_2, 0)\}$			$\{(q_3, \epsilon)\}$			
q_3			1			$\{(q_3, \epsilon)\}$			
q_4									$\{(q_4, \epsilon)\}$



A **pushdown automaton** is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$, where $Q, \Sigma, \Gamma,$ and F are all finite sets, and

- Q is the set of states,
- Σ is the input alphabet,
- Γ is the stack alphabet,
- $\delta: Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ is the transition function,
- $q_0 \in Q$ is the start state, and
- $F \subseteq Q$ is the set of accept states.



$$Q = \{q_1, q_2, q_3, q_4\},$$

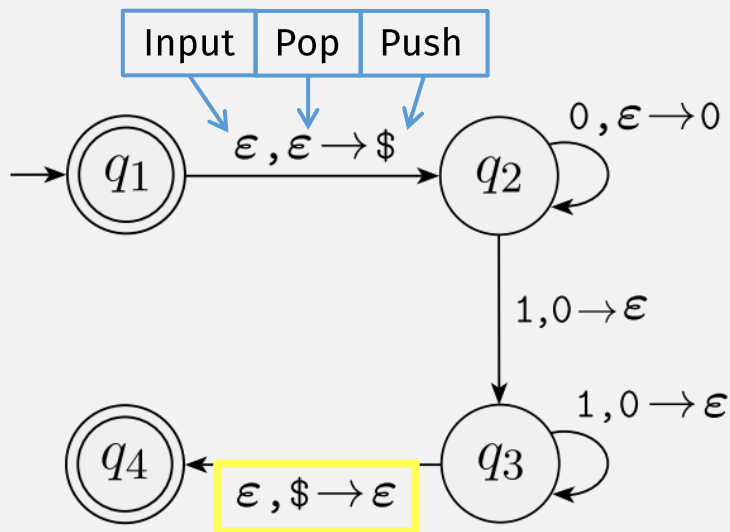
$$\Sigma = \{0, 1\},$$

$$\Gamma = \{0, \$\},$$

$$F = \{q_1, q_4\}, \text{ and}$$

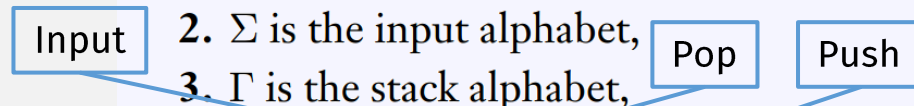
δ is given by the following table, wherein blank entries signify \emptyset .

Input:	0			1			ϵ		
Stack:	0	\$	ϵ	0	\$	ϵ	0	\$	ϵ
q_1									
q_2			$\{(q_2, 0)\}$			$\{(q_3, \epsilon)\}$			
q_3			1			$\{(q_3, \epsilon)\}$			
q_4									$\{(q_4, \epsilon)\}$



A **pushdown automaton** is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$, where Q, Σ, Γ , and F are all finite sets, and

- Q is the set of states,
- Σ is the input alphabet,
- Γ is the stack alphabet,
- $\delta: Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ is the transition function,
- $q_0 \in Q$ is the start state, and
- $F \subseteq Q$ is the set of accept states.



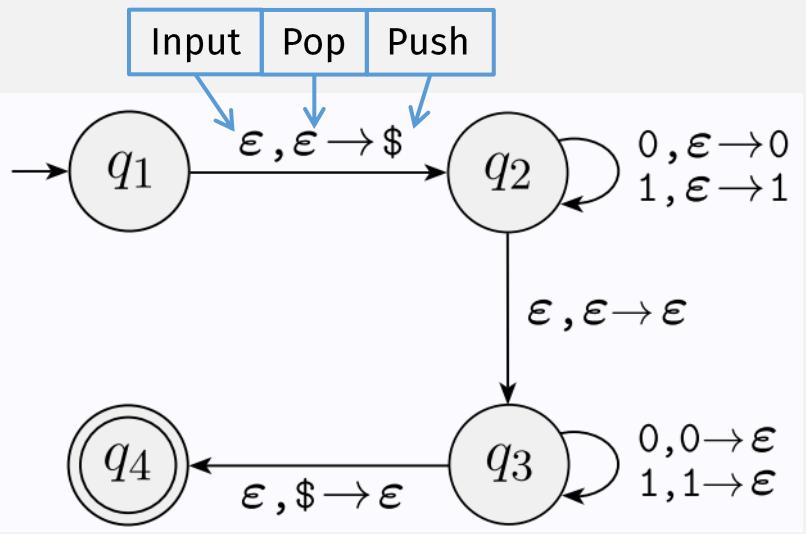
In-class exercise:
Fill in the blanks

$Q =$
 $\Sigma =$
 $\Gamma =$
 $F =$

δ is given by the following table, wherein blank entries signify \emptyset .

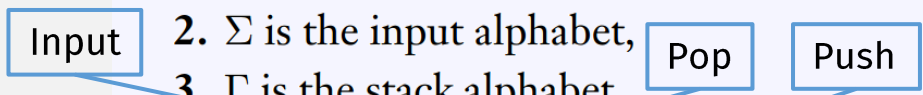
Input:	0			1			ϵ			←	Input
Stack:	0	\$	ϵ	0	\$	ϵ	0	\$	ϵ	←	Pop

PDA M_3 recognizing the language $\{ww^R \mid w \in \{0,1\}^*\}$



A *pushdown automaton* is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$, where $Q, \Sigma, \Gamma,$ and F are all finite sets, and

1. Q is the set of states,
2. Σ is the input alphabet,
3. Γ is the stack alphabet,
4. $\delta: Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ is the transition function,
5. $q_0 \in Q$ is the start state, and
6. $F \subseteq Q$ is the set of accept states.



**In-class exercise:
Fill in the blanks**

$$Q = \{q_1, q_2, q_3, q_4\},$$

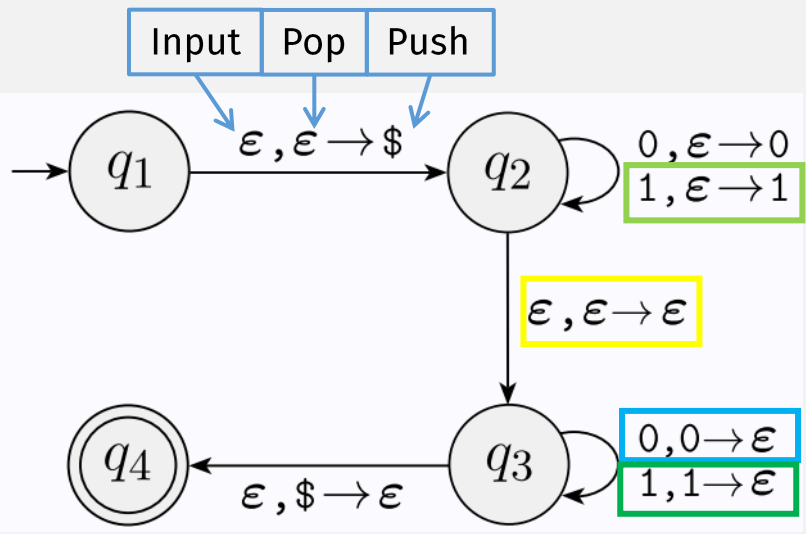
$$\Sigma = \{0,1\},$$

$$\Gamma = \{0,1, \$\},$$

$$F = \{q_4\}$$

δ is given by the following table, wherein blank entries signify \emptyset .

Input:	0			1				ϵ			
Stack:	0	\$	ϵ	0	1	\$	ϵ	0	\$	ϵ	
q_1											
q_2			$\{(q_2, 0)\}$			$\{(q_2, 1)\}$				$\{(q_2, \$)\}$	
q_3	$\{(q_3, \epsilon)\}$			$\{(q_3, \epsilon)\}$						$\{(q_3, \epsilon)\}$	
q_4								$\{(q_4, \epsilon)\}$			



PDA M_3 recognizing the language $\{ww^R \mid w \in \{0,1\}^*\}$

DFA Computation Rules

Informally

Given

- A DFA (~ a “Program”)
- and Input = string of chars, e.g. “1101”

A DFA computation (~ “Program run”):

- Start in *start state*
- Repeat:
 - Read 1 char from Input, and
 - Change state according to *transition rules*

Result of computation:

- Accept if last state is *Accept state*
- Reject otherwise

Formally (i.e., mathematically)

- $M = (Q, \Sigma, \delta, q_0, F)$
- $w = w_1w_2 \cdots w_n$

A DFA **computation** is a sequence of states:

- specified by $\hat{\delta}(q_0, w)$ where:

- M **accepts** w if $\hat{\delta}(q_0, w) \in F$
- M **rejects** otherwise

DFA Multi-step Transition Function

$$\hat{\delta}: Q \times \Sigma^* \rightarrow Q$$

- Domain (inputs):
 - state $q \in Q$
 - string $w = w_1 w_2 \cdots w_n$ where $w_i \in \Sigma$
- Range (output):
 - state $q \in Q$

A DFA **computation** is a sequence of states:

(Defined recursively)

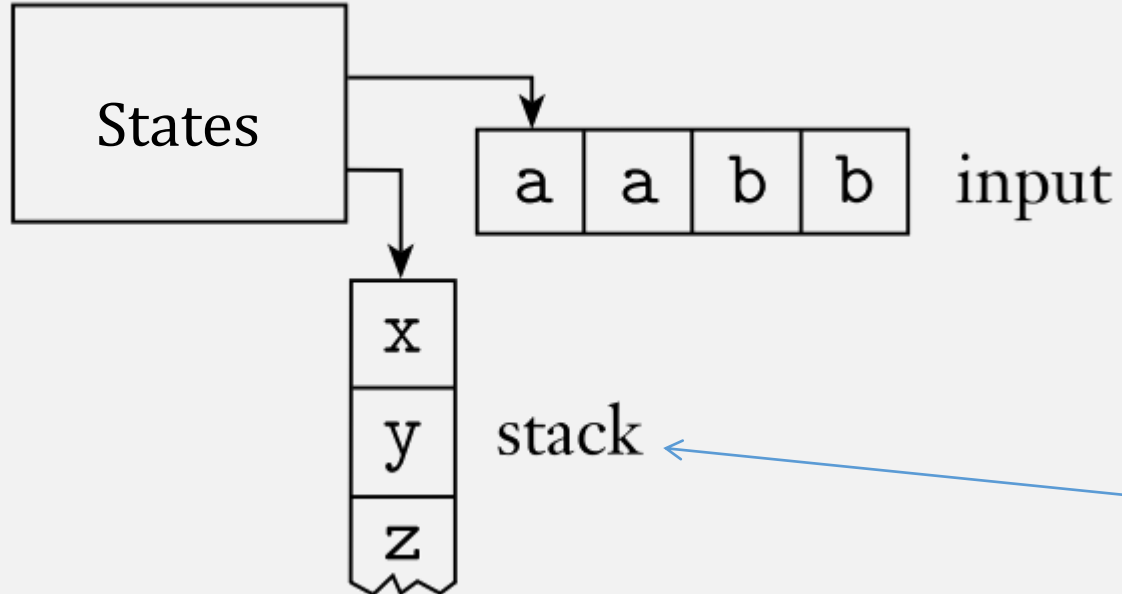
Base case $\hat{\delta}(q, \varepsilon) = q$

Recursive Case $\hat{\delta}(q, w'w_n) = \delta(\hat{\delta}(q, w'), w_n)$

where $w' = w_1 \cdots w_{n-1}$

PDA Computation?

- **PDA** = NFA + a stack
 - Infinite memory
 - Push/pop top location only



A DFA **computation** is a sequence of states ...

A PDA **computation** is a not just a sequence of states ...

... because the **stack contents** can change too!

PDA Configurations (IDs)

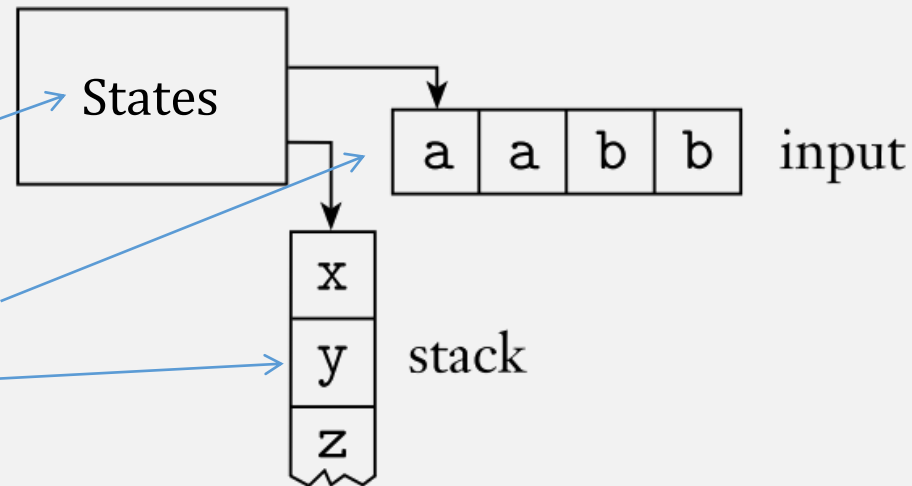
- A **configuration** (or **ID**) is a “snapshot” of a PDA’s computation

- 3 components (q, w, γ) :

q = the current state

w = the remaining input string

γ = the stack contents

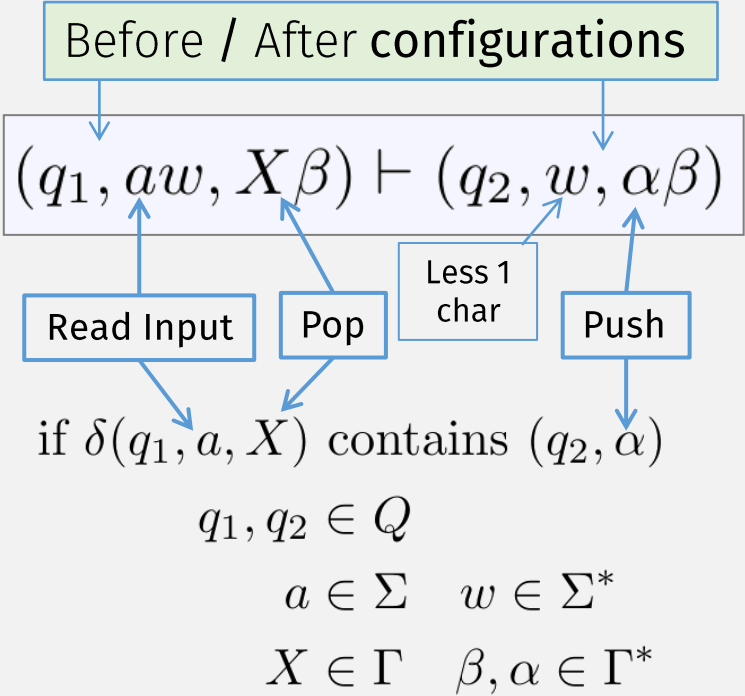


A **sequence of configurations** represents a **PDA** computation

PDA Computation, Formally

$$P = (Q, \Sigma, \Gamma, \delta, q_0, F)$$

Single-step



Multi-step

- Base Case 0 steps

$$I \vdash^* I \text{ for any ID } I$$

- Recursive Case > 0 steps

$$I \vdash^* J \text{ if there exists some ID } K \text{ such that } I \vdash K \text{ and } K \vdash^* J$$

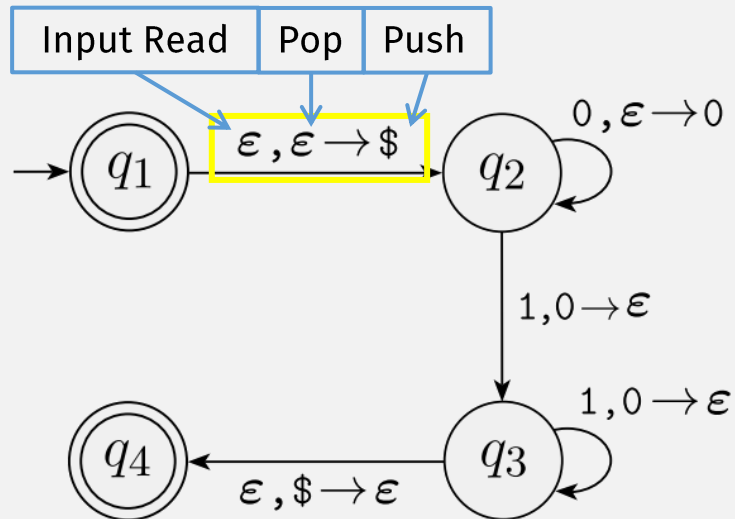
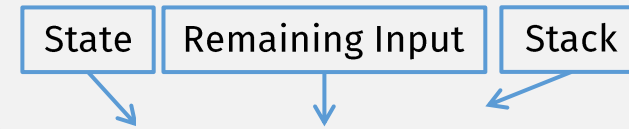
Single step Recursive "call"

A configuration (q, w, γ) has three components
 q = the current state
 w = the remaining input string
 γ = the stack contents

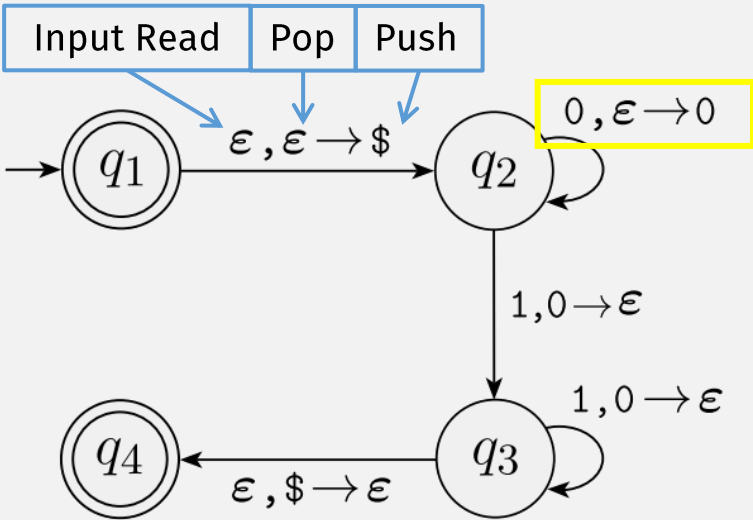
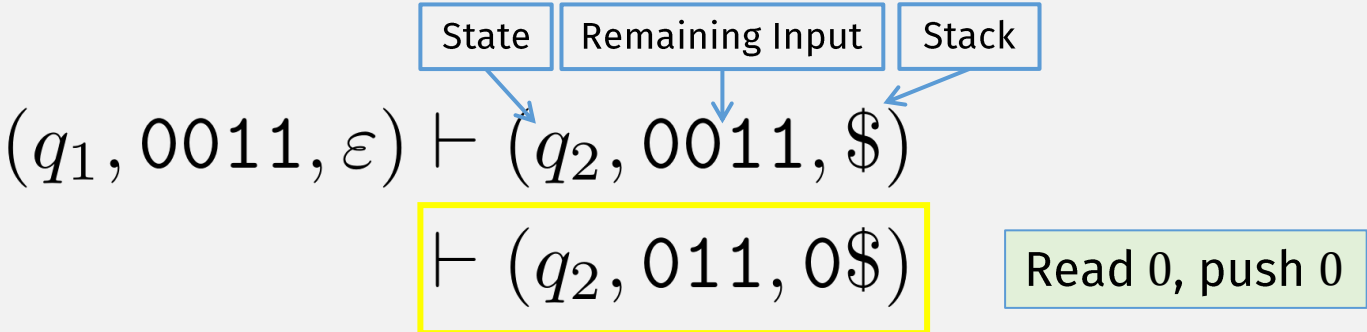
This specifies the **sequence of configurations** for a PDA computation

PDA Running Input String Example

$(q_1, 0011, \epsilon)$



PDA Running Input String Example



PDA Running Input String Example

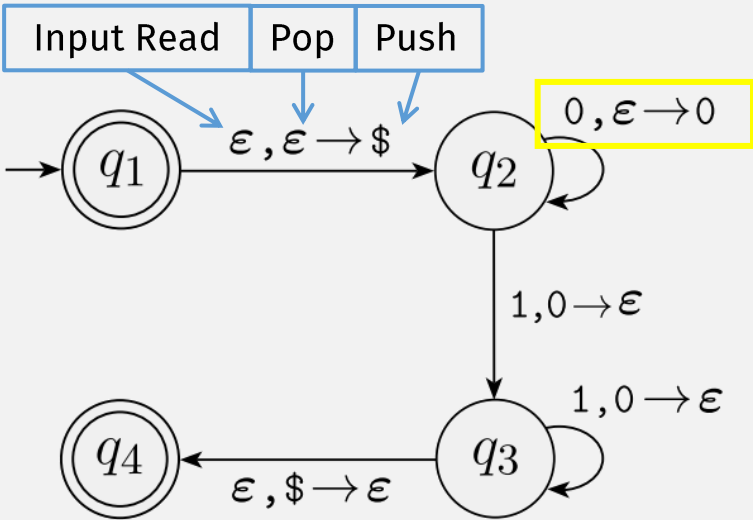
State	Remaining Input	Stack
-------	-----------------	-------

$(q_1, 0011, \epsilon) \vdash (q_2, 0011, \$)$

$\vdash (q_2, 011, 0\$)$

$\vdash (q_2, 11, 00\$)$

Read 0, push 0

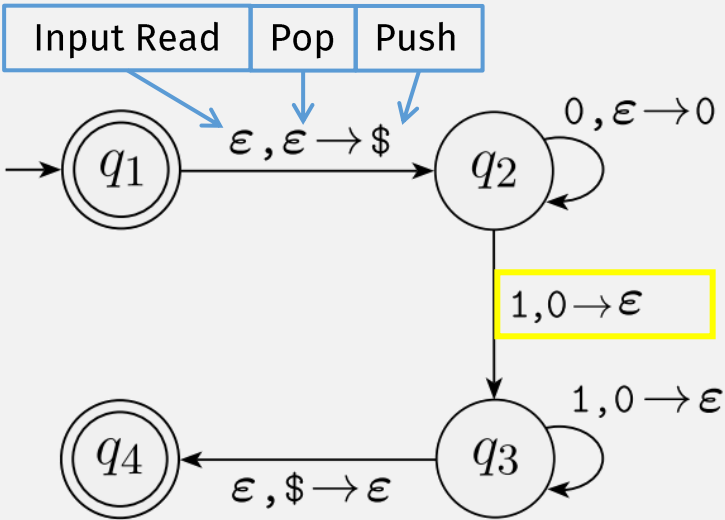


PDA Running Input String Example

State	Remaining Input	Stack
-------	-----------------	-------

$(q_1, 0011, \epsilon) \vdash (q_2, 0011, \$)$
 $\vdash (q_2, 011, 0\$)$
 $\vdash (q_2, 11, 00\$)$
 $\vdash (q_3, 1, 0\$)$

Read 1, pop 0



PDA Running Input String Example

State	Remaining Input	Stack
-------	-----------------	-------

$(q_1, 0011, \epsilon) \vdash (q_2, 0011, \$)$

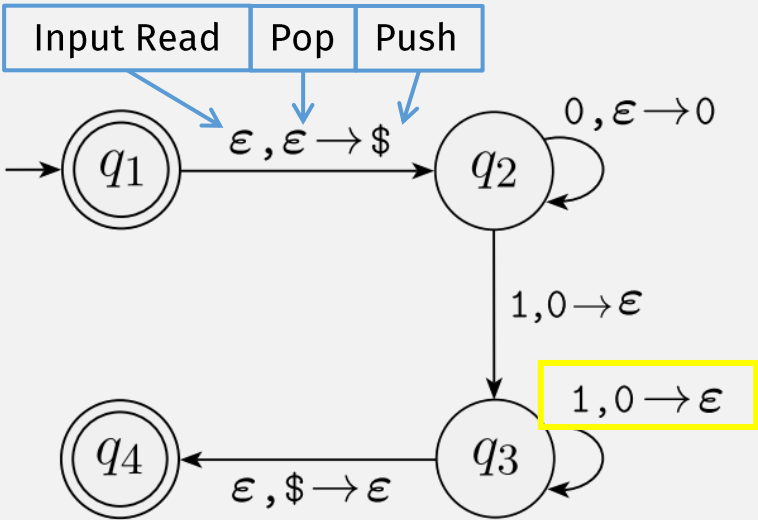
$\vdash (q_2, 011, 0\$)$

$\vdash (q_2, 11, 00\$)$

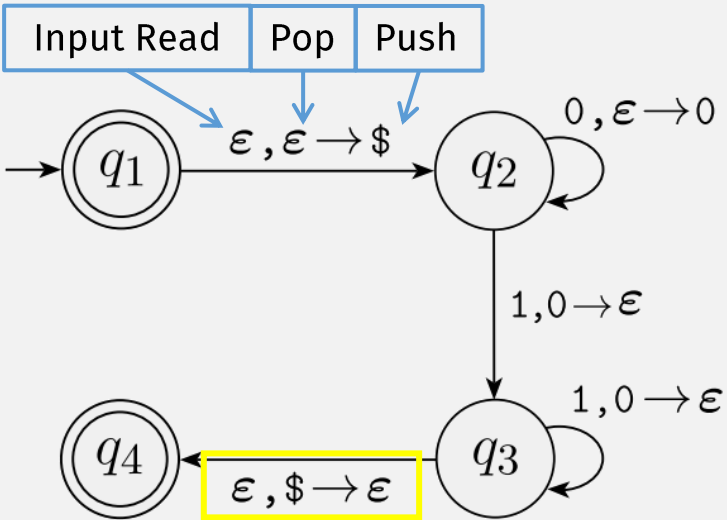
$\vdash (q_3, 1, 0\$)$

$\vdash (q_3, \epsilon, \$)$

Read 1, pop 0



PDA Running Input String Example



State | Remaining Input | Stack

- $(q_1, 0011, \epsilon) \vdash (q_2, 0011, \$)$
- $\vdash (q_2, 011, 0\$)$
- $\vdash (q_2, 11, 00\$)$
- $\vdash (q_3, 1, 0\$)$
- $\vdash (q_3, \epsilon, \$)$
- $\vdash (q_4, \epsilon, \epsilon)$

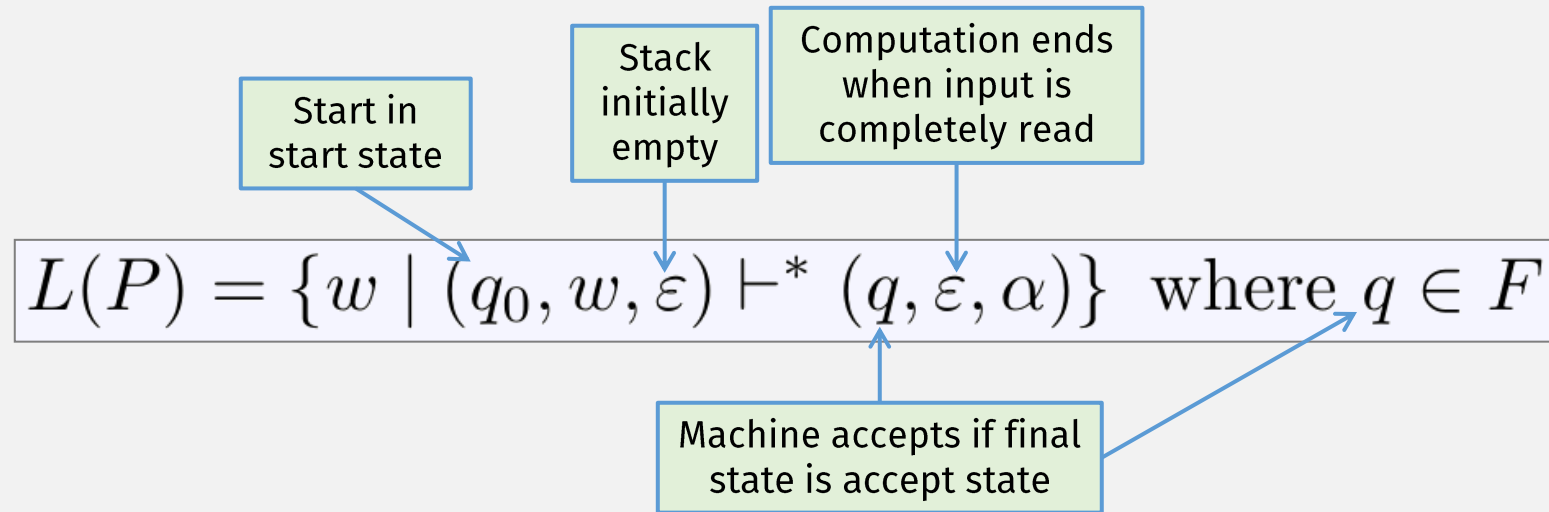
pop empty stack symbol

Flashback: Computation and Languages

- The **language** of a machine is the **set of all strings that it accepts**
- E.g., A DFA M **accepts** w if $\hat{\delta}(q_0, w) \in F$
- Language of $M = L(M) = \{ w \mid M \text{ accepts } w \}$

Language of a PDA

$$P = (Q, \Sigma, \Gamma, \delta, q_0, F)$$



A **configuration** (q, w, γ) has three components

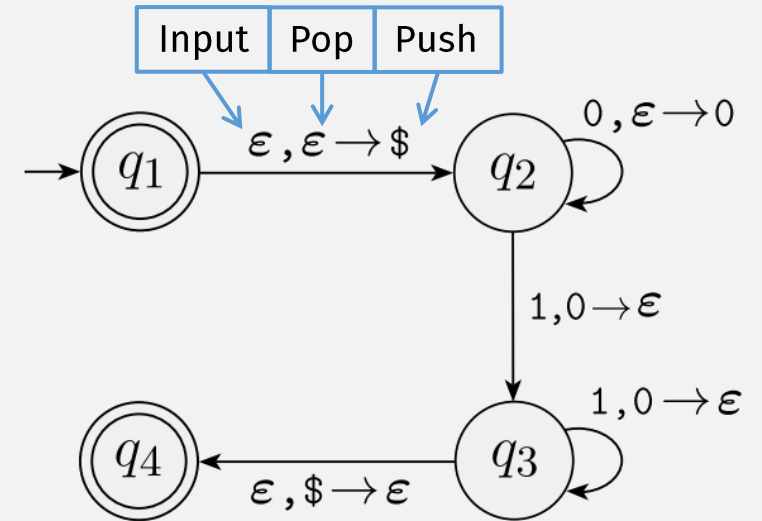
q = the current state

w = the remaining input string

γ = the stack contents

PDA's and CFL's?

- **PDA** = NFA + a stack
 - Infinite memory
 - Push/pop top location only
- Want to prove: PDA's represent CFL's!
- We know: a CFL, by definition, is a language that is generated by a CFG
- Need to show: PDA \Leftrightarrow CFG
- Then, to prove that a language is a CFL, we can either:
 - Create a CFG, or
 - Create a PDA



A lang is a CFL iff some PDA recognizes it

⇒ If a language is a **CFL**, then a PDA recognizes it

- We know: A CFL has a CFG describing it (definition of CFL)
- To prove this part: show the CFG has an equivalent PDA

⇐ If a PDA recognizes a language, then it's a CFL

Submit in-class work 3/20

On gradescope