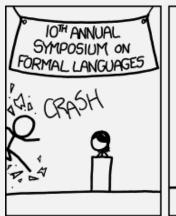
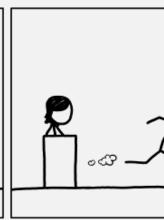
#### **UMBCS622**

### Pushdown Automata (PDAs)

Monday, October 4, 2021





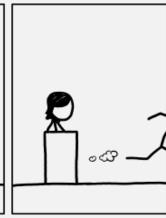


#### Announcements

- No class next Monday 10/11
- HW4 released
  - Due Sun 10/18 11:59pm EST
  - Note: this is a 2 week assignment!







### Last Time:

Regular Languages	Context-Free Languages (CFLs)
Regular Expression	Context-Free Grammar (CFG)
A Reg expr <u>describes</u> a Regular lang	A CFG <u>describes</u> a CFL

# Today:

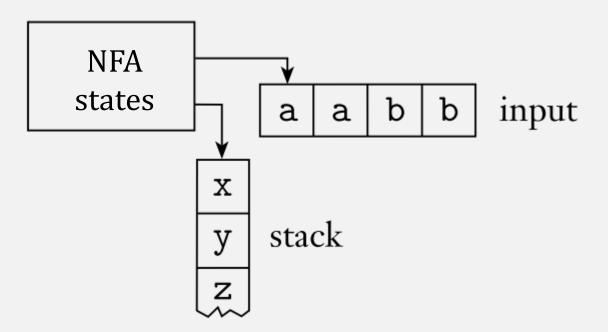
Regular Languages	Context-Free Languages (CFLs)
Regular Expression	Context-Free Grammar (CFG)
A Reg expr <u>describes</u> a Regular lang	A CFG <u>describes</u> a CFL
	TODAY:
Finite automaton (FSM)	Push-down automaton (PDA)
An FSM <u>recognizes</u> a Regular lang	A PDA <u>recognizes</u> a CFL

# Today:

Regular Languages	Context-Free Languages (CFLs)
Regular Expression	Context-Free Grammar (CFG)
A Reg expr <u>describes</u> a Regular lang	A CFG <u>describes</u> a CFL
	TODAY:
Finite automaton (FSM)	Push-down automaton (PDA)
An FSM <u>recognizes</u> a Regular lang	A PDA <u>recognizes</u> a CFL
DIFFERENCE:	DIFFERENCE:
A Regular lang is <u>defined</u> with a FSM	A CFL is <u>defined</u> with a CFG
<i>Must prove</i> : Reg expr ⇔ Reg lang	Must prove: PDA ⇔ CFL

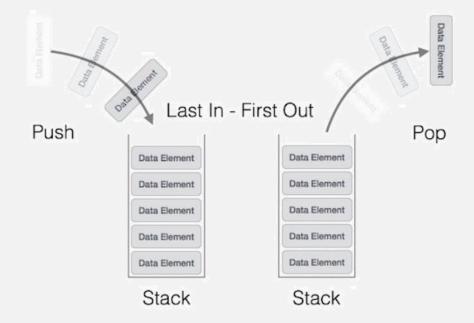
### Pushdown Automata (PDA)

• PDA = NFA + a stack



#### What is a Stack?

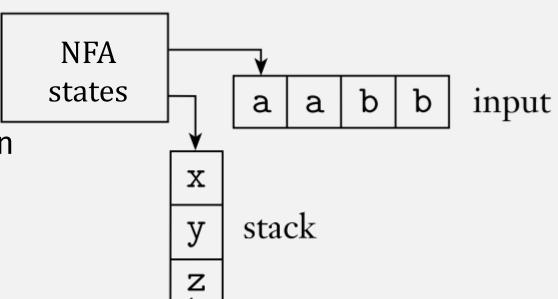
- Access to top element only
- 2 Operations: push, pop





### Pushdown Automata (PDA)

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



#### $\{0^n 1^n | n \ge 0\}$

### An Example PDA

(\$ = special symbol, indicating empty stack) Read Push Pop input read 0, no pop, push 0 0, $\varepsilon \rightarrow$ 0 (and repeat) arepsilon , arepsilon o \$  $q_2$ when machine starts: read 1, pop 0, no push - don't read input, 1,0 $\rightarrow \varepsilon$ (and repeat) - don't pop anything, - push empty stack symbol 1,0ightarrow arepsilon $q_3$ arepsilon,\$ 
ightarrow arepsilonaccept only when stack is empty

### Formal Definition of PDA

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.  $\Sigma$  is the input alphabet,
- 3.  $\Gamma$  is the stack alphabet,

Stack alphabet can have special stack symbols, e.g., \$

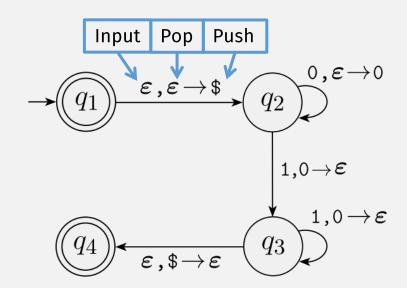
- 4.  $\delta: Q \times \Sigma_{\varepsilon} \times \Gamma_{\varepsilon} \longrightarrow \mathcal{P}(Q \times \Gamma_{\varepsilon})$  is the transition function, 5.  $q_0 \in \mathbb{Q}$  Input Popart state, and Push
- **6.**  $F \subseteq Q$  is the set of accept states.

Non-deterministic: produces a set of (STATE, STACK CHAR) pairs

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

# PDA Formal (b) efinition Example

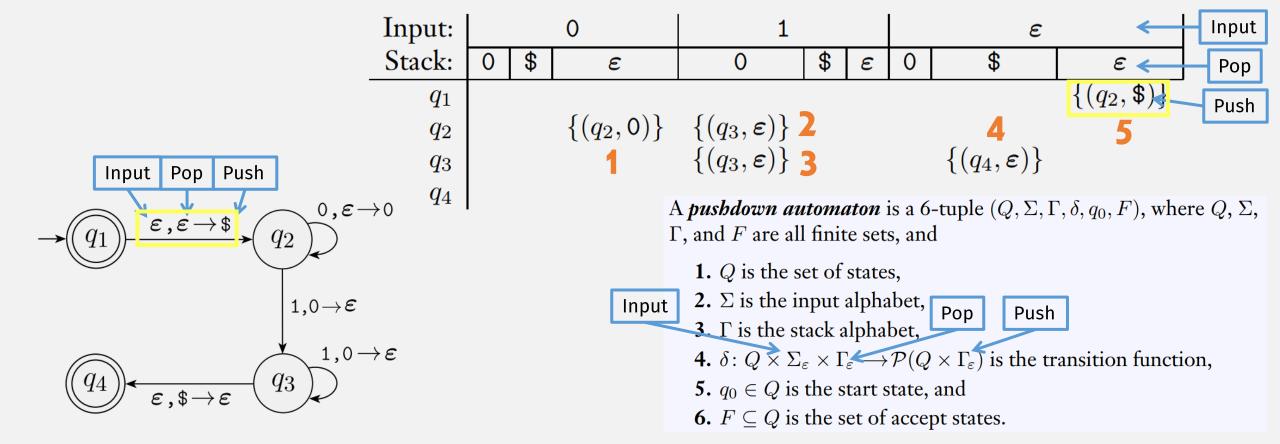
$$F = \{q_1, q_4\},\$$



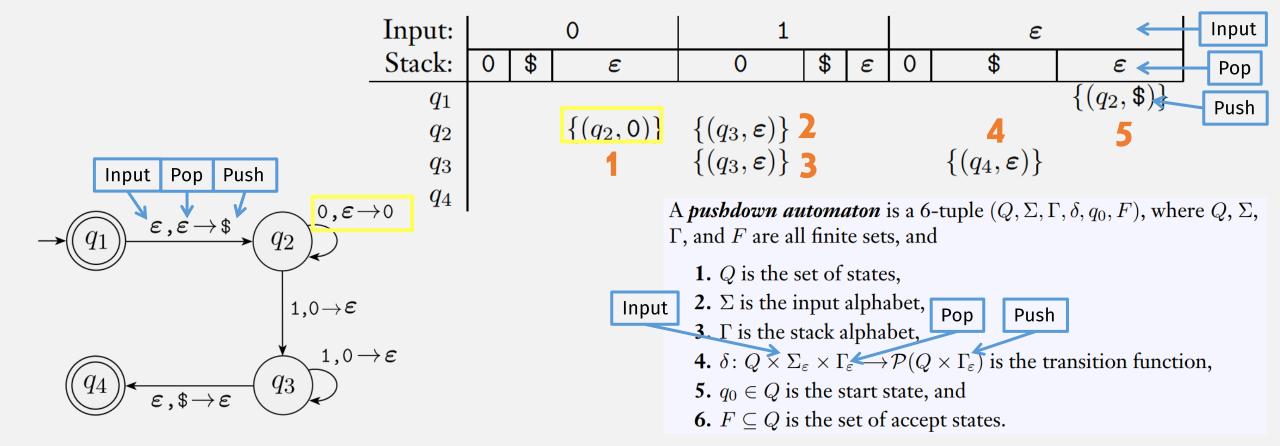
A pushdown automaton is a 6-tuple  $(Q, \Sigma, \Gamma, \delta, q_0, F)$ , where  $Q, \Sigma$ ,  $\Gamma$ , and F are all finite sets, and

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

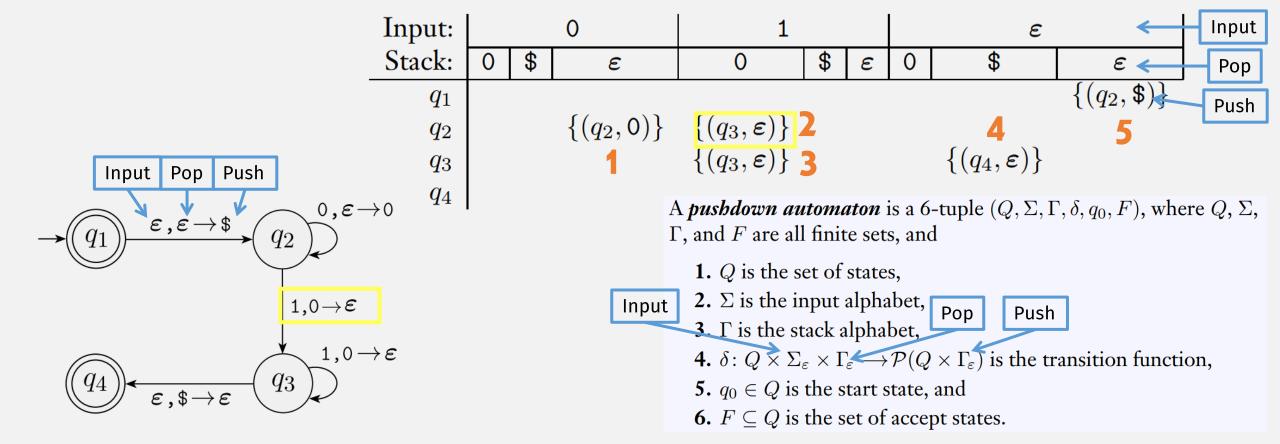
$$Q = \{q_1, q_2, q_3, q_4\},$$
  
 $\Sigma = \{0,1\},$   
 $\Gamma = \{0,\$\},$   
 $F = \{q_1, q_4\},$  and



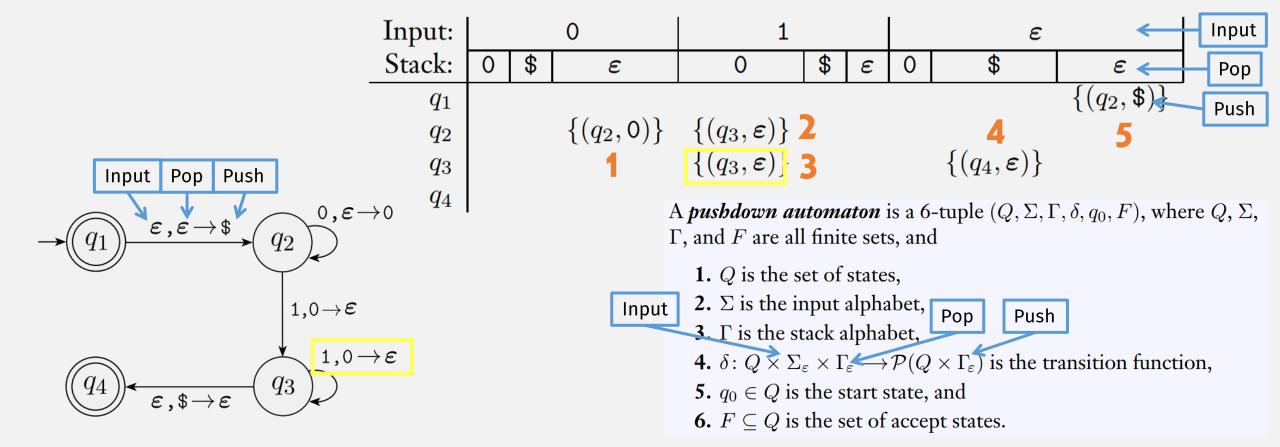
$$Q = \{q_1, q_2, q_3, q_4\},$$
  
 $\Sigma = \{0,1\},$   
 $\Gamma = \{0,\$\},$   
 $F = \{q_1, q_4\},$  and



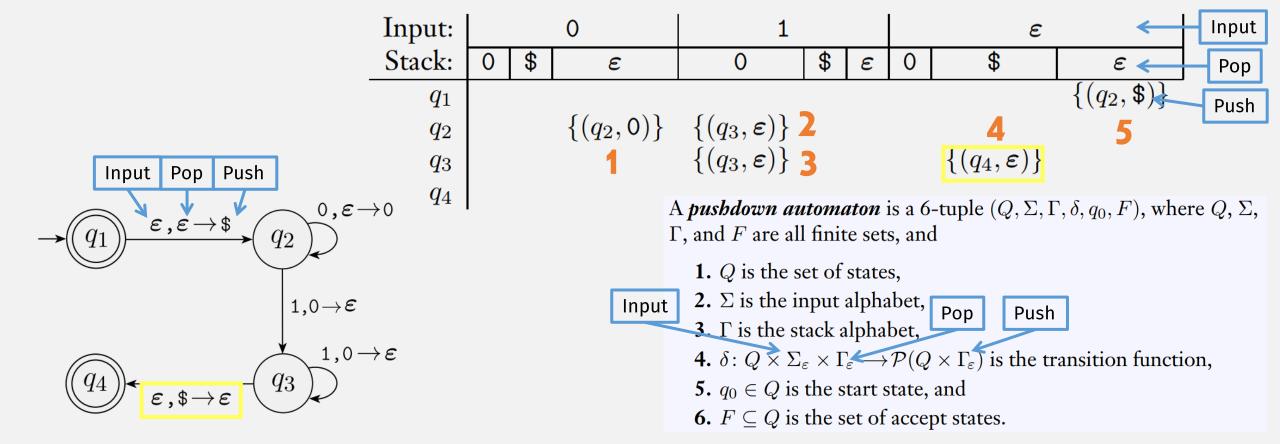
$$Q = \{q_1, q_2, q_3, q_4\},$$
  
 $\Sigma = \{0,1\},$   
 $\Gamma = \{0,\$\},$   
 $F = \{q_1, q_4\},$  and



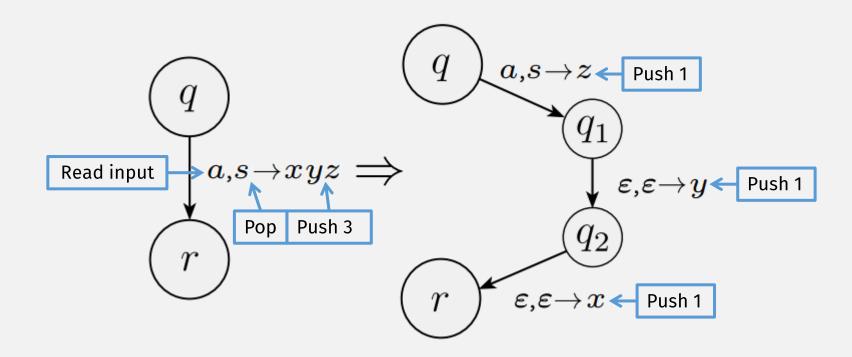
$$Q = \{q_1, q_2, q_3, q_4\},$$
  
 $\Sigma = \{0,1\},$   
 $\Gamma = \{0,\$\},$   
 $F = \{q_1, q_4\},$  and



$$Q = \{q_1, q_2, q_3, q_4\},$$
  
 $\Sigma = \{0,1\},$   
 $\Gamma = \{0,\$\},$   
 $F = \{q_1, q_4\},$  and



## Multi-Symbol Stack Pushes



Note the reverse order of pushes

## PDA Configurations (IDs)

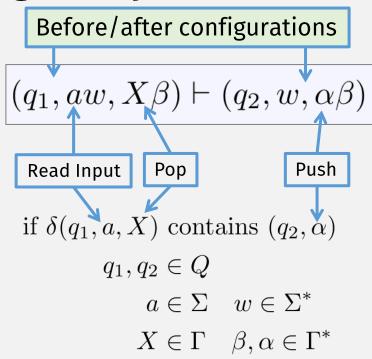
 When "running" an input string on a PDA, a configuration (or ID) is a snapshot of some point in the computation

```
• A configuration (or ID) (q, w, \gamma) has three components q = the current state w = the remaining input string \gamma = the stack contents
```

### "Running" an Input String on a PDA

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

#### Single-step



#### **Extended**

Base Case

$$I \stackrel{*}{\vdash} I$$
 for any ID  $I$ 

Recursive Case

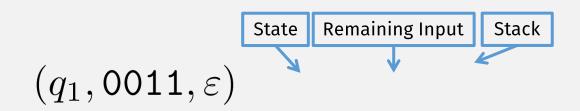
$$I \stackrel{*}{\vdash} J$$
 if there exists some ID  $K$  such that  $I \vdash K$  and  $K \stackrel{*}{\vdash} J$ 

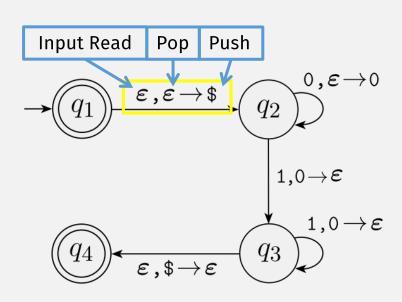
A configuration  $(q, w, \gamma)$  has three components q = the current state w = the remaining input string  $\gamma$  = the stack contents

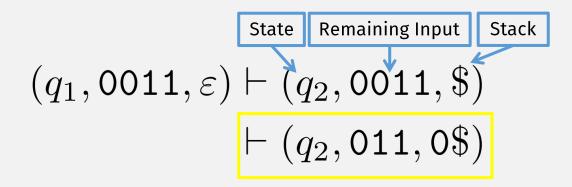
### Language of a PDA

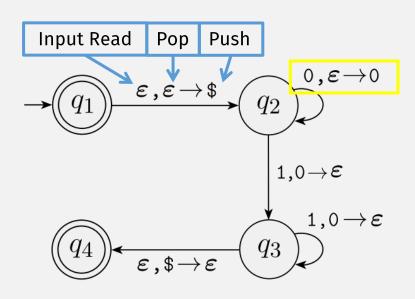
$$P = (Q, \Sigma, \Gamma, \delta, q_0, F)$$

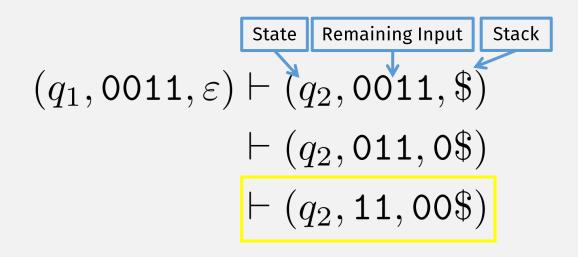
$$L(P) = \{ w \mid (q_0, w, \varepsilon) \vdash^* (q, \varepsilon, \alpha) \} \text{ where } q \in F$$

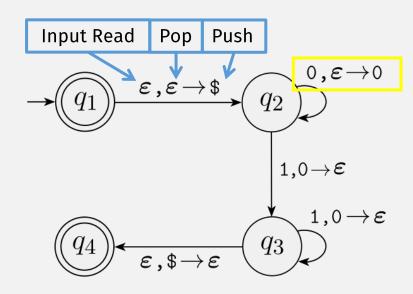


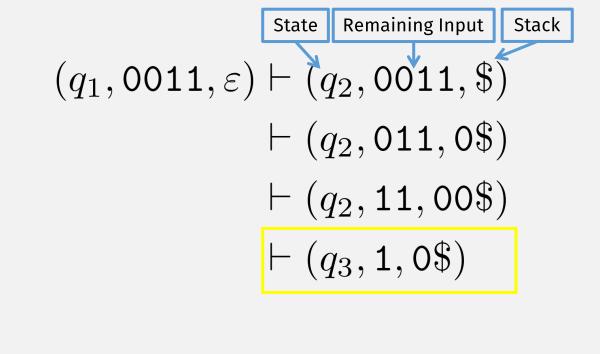


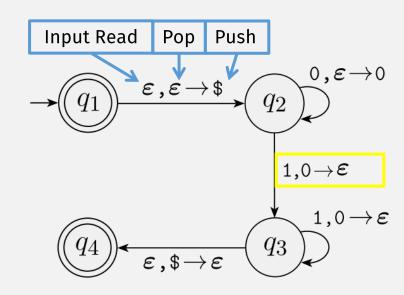


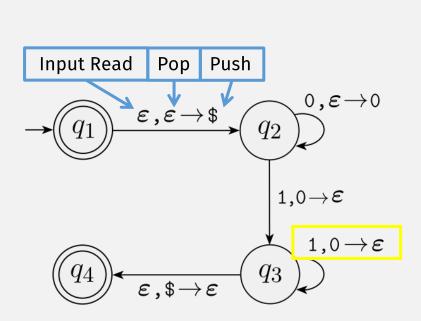


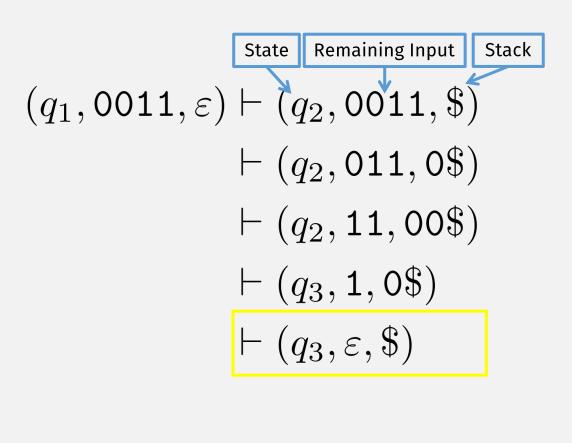


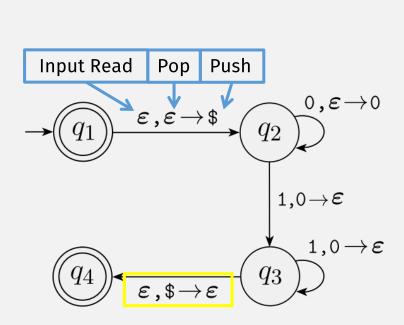


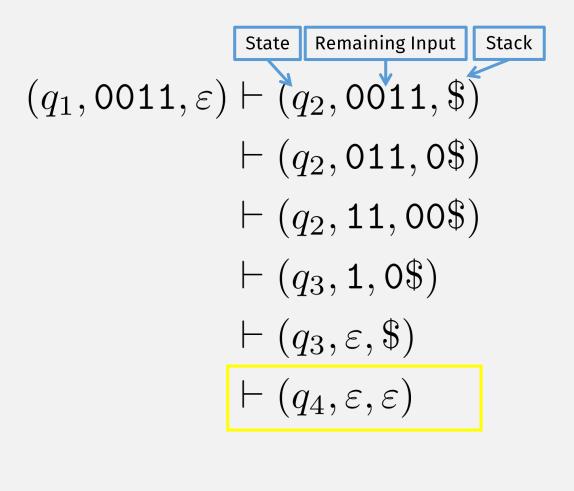












# If $(q_1,x,\alpha) \vdash^* (q_n,y,\beta)$ Assume is true then $(q_1,x\boldsymbol{w},\alpha\boldsymbol{\gamma}) \vdash^* (q_n,y\boldsymbol{w},\beta\boldsymbol{\gamma})$ Must prove

#### A PDA Theorem

**Proof:** (by induction on the number of steps in the sequence)

Adding to <u>end of input</u> or <u>bottom of stack</u> doesn't affect the computation

- <u>Base Case</u> (0 steps): If  $(q_1, x, \alpha) \vdash^* (q_1, x, \alpha)$  then  $(q_1, xw, \alpha\gamma) \vdash^* (q_1, xw, \alpha\gamma)$ 
  - TRUE, from definition of  $\vdash^*$ :  $I \vdash^* I$  for any ID I
- Inductive Case

Need to prove:

IH says: if this is true ... How do we know these steps are true?

$$\begin{array}{c|c} & \text{If } (q_1,x,\alpha) \overset{\checkmark}{\vdash^*} (\boldsymbol{q_{n-1}},\boldsymbol{x'},\boldsymbol{\alpha'}) & \vdash (q_n,y,\beta) & \leftarrow & \text{From the assumption!} \\ & \text{Then } (q_1,x\boldsymbol{w},\alpha\boldsymbol{\gamma}) \overset{\ast}{\vdash^*} (\boldsymbol{q_{n-1}},\boldsymbol{x'w},\boldsymbol{\alpha'\gamma}) & \vdash (q_n,y\boldsymbol{w},\beta\boldsymbol{\gamma}) & \leftarrow & \text{Left to prove} \\ \end{array}$$

... then this is true

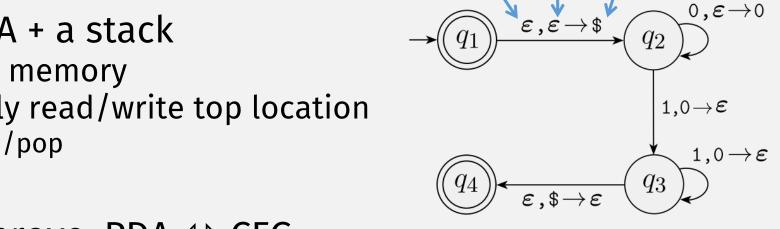
These steps must use the same  $\delta$  transition, why?

Same state, input char, and stack top!

### **CFL** ⇔ **PDA**

### Pushdown Automata (PDA)

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



Input

Pop

Push

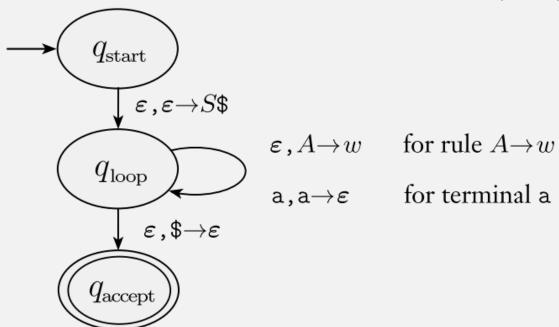
- Want to prove: PDA ⇔ CFG
- Then, to prove that a language is context-free, we can either:
  - Create a CFG, or
  - Create a PDA

### A lang is a CFL iff some PDA recognizes it

- ⇒ If a language is a CFL, then a PDA recognizes it
  - (Easier)
  - We know: A CFL has a CFG describing it (definition of CFL)
  - To prove forward dir: Convert CFG→PDA
- ← If a PDA recognizes a language, then it's a CFL

#### **CFG→PDA**

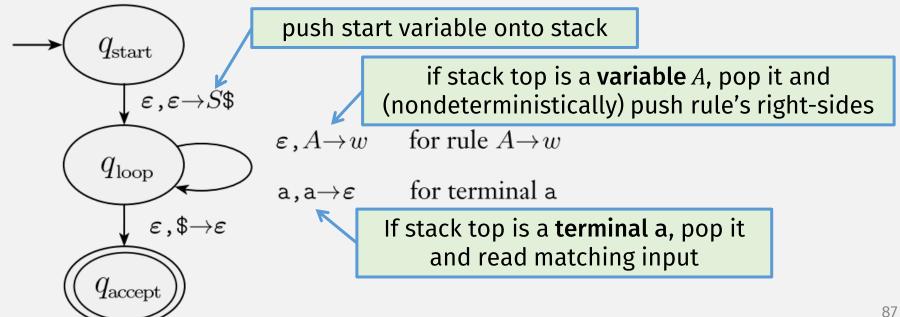
- Construct a PDA from CFG such that:
  - PDA accepts input string only if the CFG can generate that string
- Intuitively, PDA will <u>nondeterministically</u> try all rules



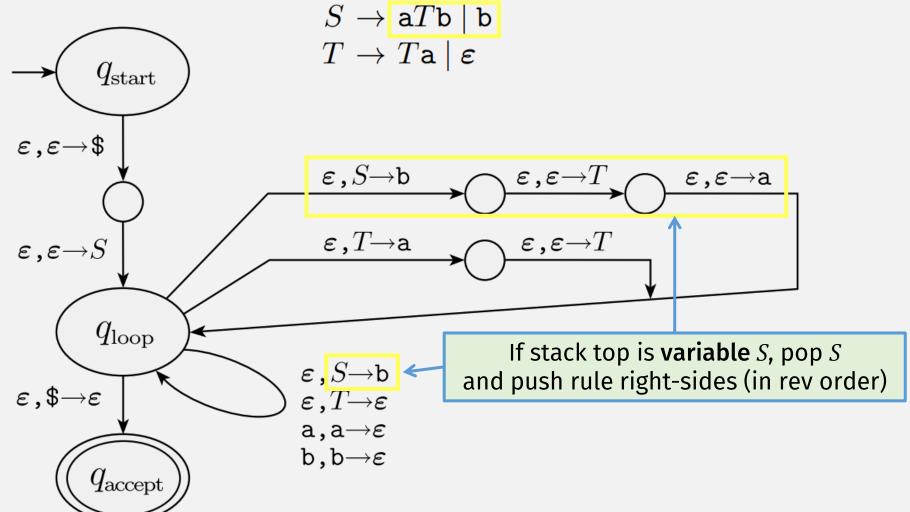
#### **CFG**→**PDA**

- Construct a PDA from CFG such that:
  - PDA accepts input string only if the CFG can generate that string

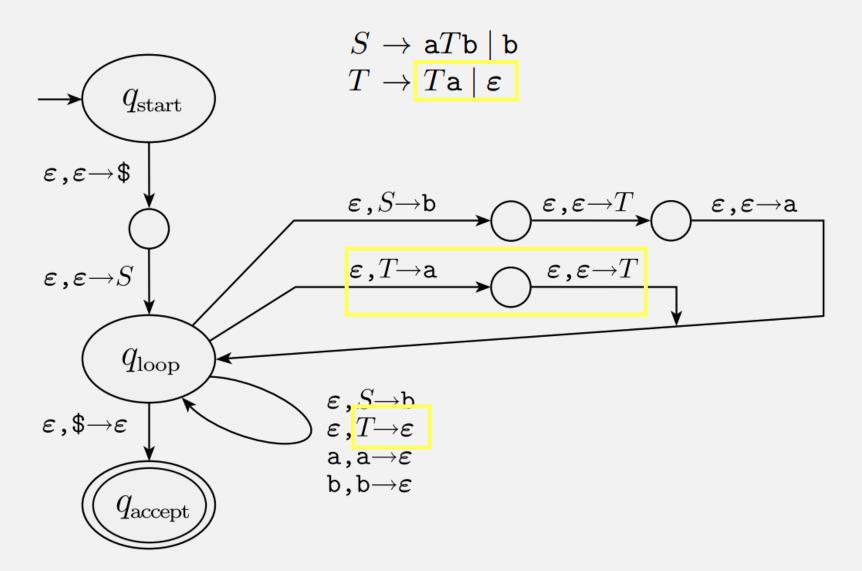
• Intuitively, PDA will <u>nondeterministically</u> try all rules



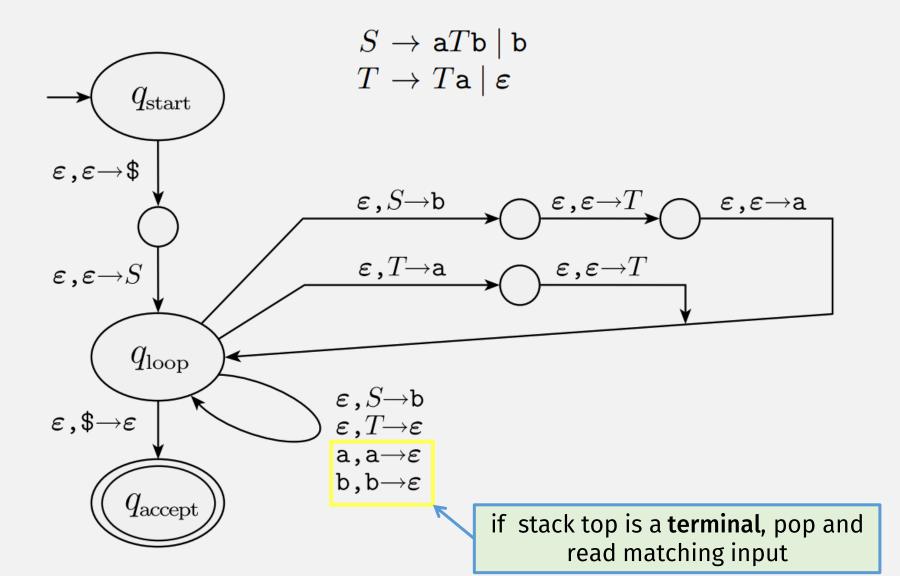
### Example **CFG→PDA**



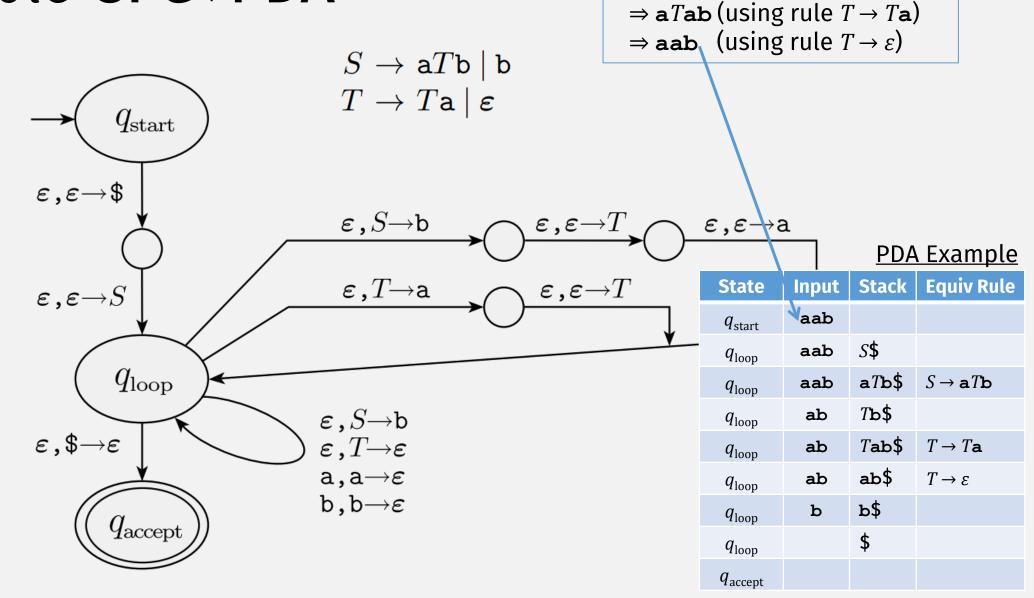
## Example CFG>PDA



### Example **CFG→PDA**



### Example CFG>PDA

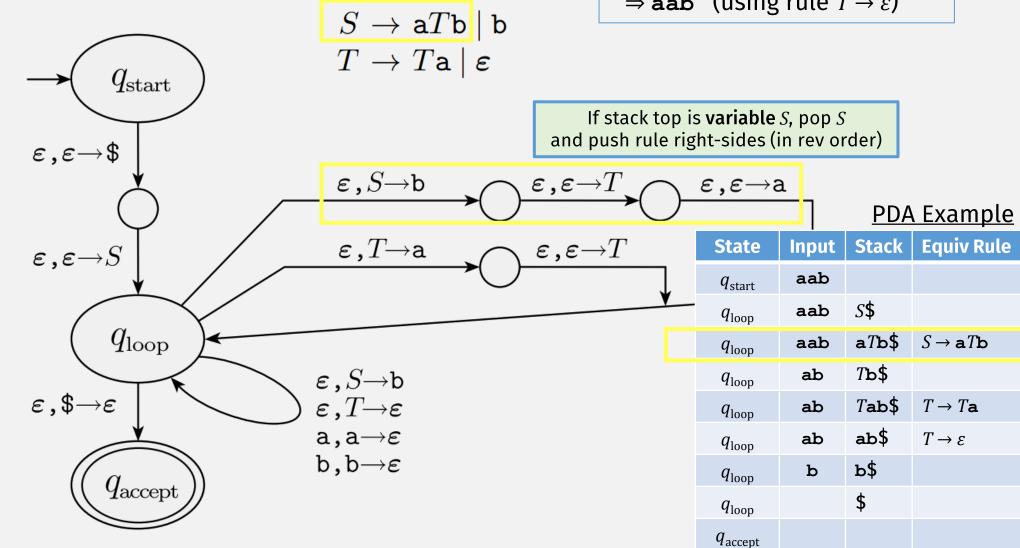


**Example Derivation using CFG:** 

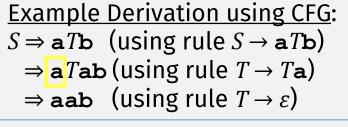
 $S \Rightarrow \mathbf{a} T \mathbf{b}$  (using rule  $S \rightarrow \mathbf{a} T \mathbf{b}$ )

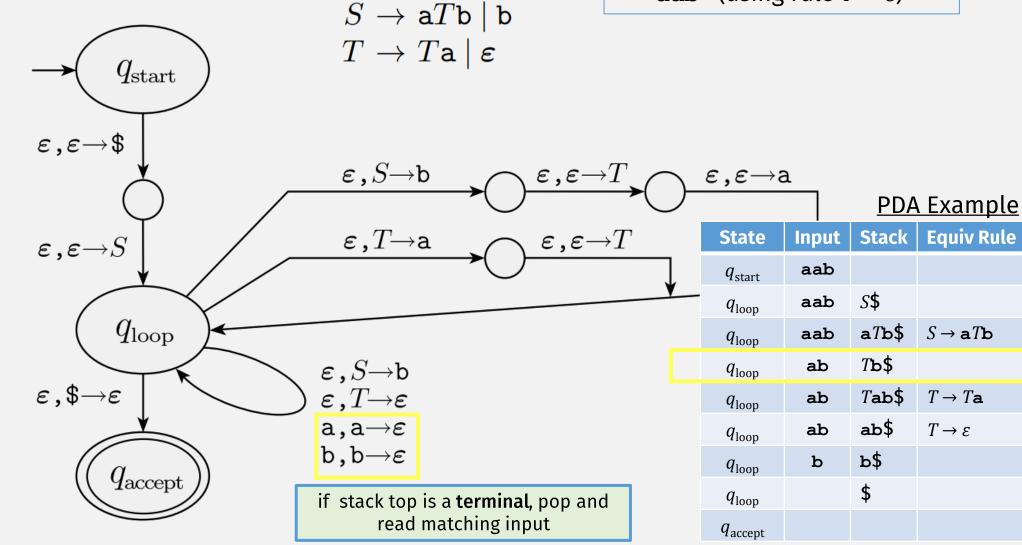
### Example CFG>PDA

Example Derivation using CFG:  $S \Rightarrow \mathbf{a}T\mathbf{b}$  (using rule  $S \rightarrow \mathbf{a}T\mathbf{b}$ )  $\Rightarrow \mathbf{a}T\mathbf{a}\mathbf{b}$  (using rule  $T \rightarrow T\mathbf{a}$ )  $\Rightarrow \mathbf{a}\mathbf{a}\mathbf{b}$  (using rule  $T \rightarrow \varepsilon$ )



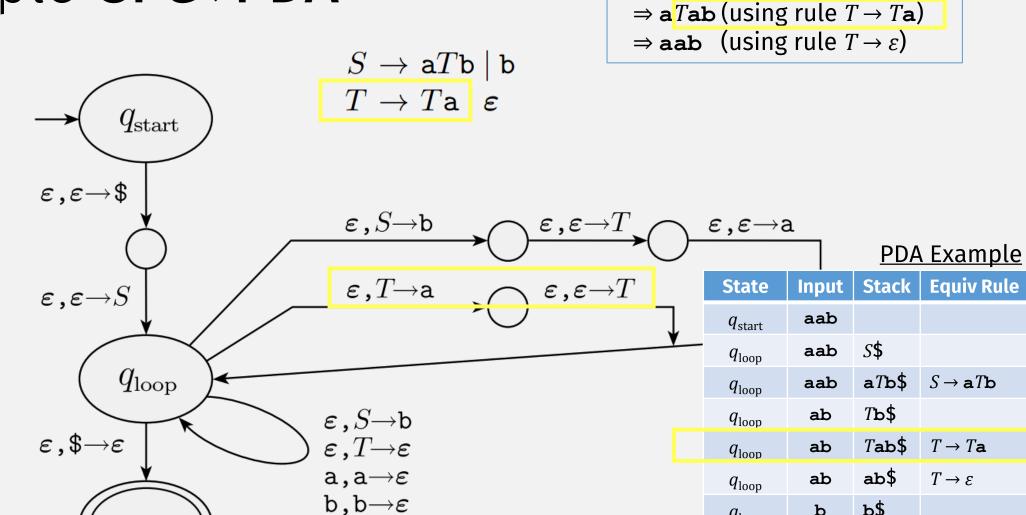
### Example **CFG→PDA**





### Example **CFG→PDA**

 $q_{
m accept}$ 



**Example Derivation using CFG:** 

b

 $q_{\mathrm{loop}}$ 

 $q_{\rm loop}$ 

 $q_{\rm accept}$ 

**b**\$

 $S \Rightarrow \mathbf{a} T \mathbf{b}$  (using rule  $S \rightarrow \mathbf{a} T \mathbf{b}$ )

### A lang is a CFL iff some PDA recognizes it

- $\boxed{\hspace{0.1cm}}$   $\Rightarrow$  If a language is a CFL, then a PDA recognizes it
  - Convert CFG→PDA

- ← If a PDA recognizes a language, then it's a CFL
  - (Harder)
  - Need to: Convert PDA→CFG

#### PDA→CFG: Prelims

#### Before converting PDA to CFG, modify it so:

- 1. It has a single accept state,  $q_{\text{accept}}$ .
- 2. It empties its stack before accepting.
- **3.** Each transition either pushes a symbol onto the stack (a *push* move) or pops one off the stack (a *pop* move), but it does not do both at the same time.

#### **Important:**

This doesn't change the language recognized by the PDA (confirm this to yourselves)

### $PDA P \rightarrow CFG G$ : Variables

$$P = (Q, \Sigma, \Gamma, \delta, q_0, \{q_{\text{accept}}\})$$
 variables of  $G$  are  $\{A_{pq} | p, q \in Q\}$ 

- Want: if P goes from state p to q reading input x, then some  $A_{pq}$  generates x
- So: For every pair of states p, q in P, add variable  $A_{pq}$  to G
- Then: connect the variables together by,
  - Add rules:  $A_{pq} \rightarrow A_{pr}A_{rq}$ , for each state r
  - These rules allow grammar to simulate every possible transition
  - (We haven't added input read/generated terminals yet)
- To add terminals: pair up stack pushes and pops (essence of a CFL)\*8

### PDA P -> CFG G: Generating Strings

$$P = (Q, \Sigma, \Gamma, \delta, q_0, \{q_{\text{accept}}\})$$
 variables of  $G$  are  $\{A_{pq} | p, q \in Q\}$ 

• The key: pair up stack pushes and pops (essence of a CFL)

if  $\delta(p, a, \varepsilon)$  contains (r, u) and  $\delta(s, b, u)$  contains  $(q, \varepsilon)$ ,

put the rule  $A_{pq} \rightarrow aA_{rs}b$  in G

### PDA P -> CFG G: Generating Strings

$$P = (Q, \Sigma, \Gamma, \delta, q_0, \{q_{\text{accept}}\})$$
 variables of  $G$  are  $\{A_{pq} | p, q \in Q\}$ 

• The key: pair up stack pushes and pops (essence of a CFL)

if  $\delta(p, a, \varepsilon)$  contains (r, u) and  $\delta(s, b, u)$  contains  $(q, \varepsilon)$ ,

put the rule  $A_{pq} \rightarrow aA_{rs}b$  in G

### PDA P -> CFG G: Generating Strings

$$P = (Q, \Sigma, \Gamma, \delta, q_0, \{q_{\text{accept}}\})$$
 variables of  $G$  are  $\{A_{pq} | p, q \in Q\}$ 

• The key: pair up stack pushes and pops (essence of a CFL)

if  $\delta(p, a, \varepsilon)$  contains (r, u) and  $\delta(s, b, u)$  contains  $(q, \varepsilon)$ ,

put the rule  $A_{pq} \rightarrow a A_{rs} b$  in G

## A language is a CFL $\Leftrightarrow$ A PDA recognizes it

- $| \longrightarrow |$  If a language is a CFL, then a PDA recognizes it
  - Convert CFG→PDA

- ✓ 

  ✓ If a PDA recognizes a language, then it's a CFL
  - Convert PDA→CFG

### Check-in Quiz 10/4

On Gradescope