CS 6041 Theory of Computation

Review 2

Kun Suo

Computer Science, Kennesaw State University

https://kevinsuo.github.io/

Exam 2

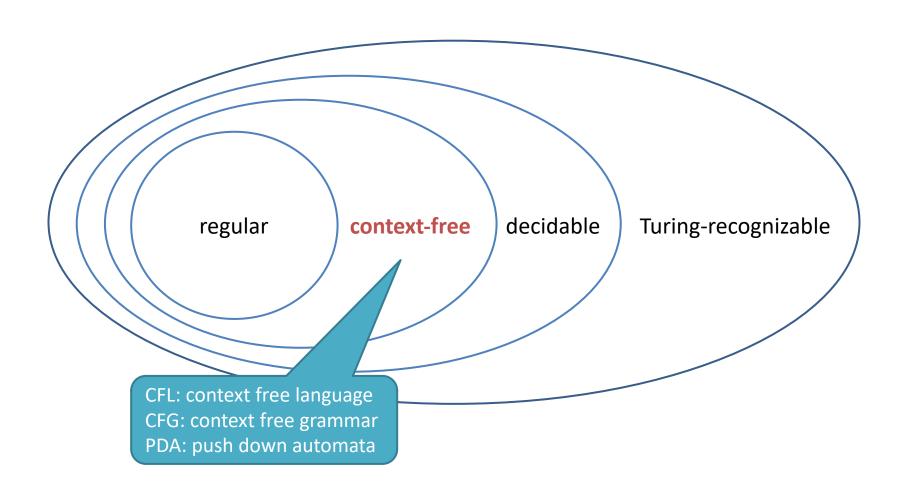
- similar as Exam 1

- 10 True/False question
 - 2 points each

- 4 short answer question
 - 20 points each

• 100 = 2*10 + 4*20

Context-free language



Context Free Grammar

• Example, G₁

Variable: A, B

Start variable:

A

3 substitution rules (productions)

$$A \rightarrow 0A1$$

$$A \rightarrow B$$

$$B \rightarrow \#$$

Terminals: 0, 1, #

$$A \Rightarrow 0A1$$

$$\Rightarrow$$
 00A11

$$\Rightarrow$$
 000A111

The language of grammar

Grammar G₁:

$$A \rightarrow 0A1$$

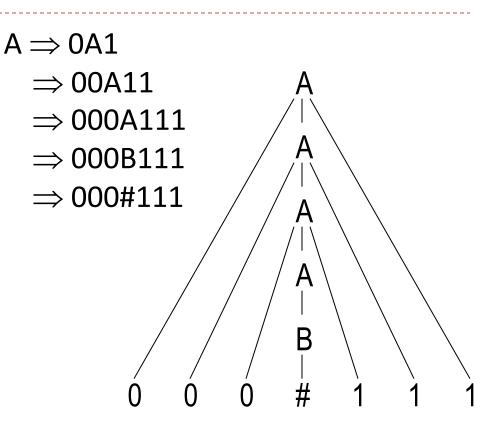
$$A \rightarrow B$$

$$B \rightarrow \#$$

The language of G₁:

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

- Context-free language
 - Languages generated by contextfree grammars



000#111

Definition of context-free grammar

- Context-free grammar is a 4-tuple $G=(V,\Sigma,R,S)$,
 - 1) V: finite variable set

2) Σ : finite terminal set

3) R: finite rule set $(A \rightarrow w, w \in (V \cup \Sigma)^*)$

4) S∈V: start variable

Design CFG is much difficult than designing an automata for language

Basic idea:

- 1. divide CFL into small parts
- 2. design CFG for each small part
- 3. combine them together

• Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$

- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - o Design CFG for $\{w \mid w=0^n1^n, n \ge 0\}$
 - Design CFG for $\{w \mid w=1^n0^n, n\geq 0\}$

- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - Design CFG for $\{w \mid w=0^n1^n, n\geq 0\}$

Generating same number of 0 and 1 Generating 0 before 1

01 0011 000111 00..011..1

- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - o Design CFG for $\{w \mid w=0^n1^n, n \ge 0\}$

Generating same number of 0 and 1 Generating 0 before 1

01 0011 000111 00..011..1

- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - Design CFG for $\{w \mid w=0^n1^n, n\geq 0\}$

Generating same number of 0 and 1 Generating 0 before 1

 $01 \quad 0011 \quad 000111 \quad 00..011..1$

- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - Design CFG for $\{w \mid w=0^n1^n, n\geq 0\}$

Generating same number of 0 and 1 Generating 0 before 1

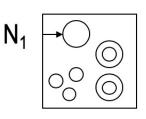
01 0011 000111 00..01..11

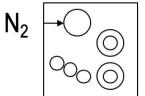
 $S \rightarrow 0S1$

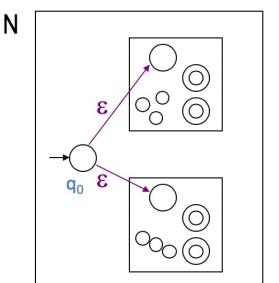
- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - o Design CFG for $\{w \mid w=0^n1^n, n \ge 0\}$
 - ▶ $G_1 = (\{S\}, \{0,1\}, \{S \rightarrow 0S1, S \rightarrow \varepsilon\}, S)$
 - Design CFG for $\{w \mid w=1^n0^n, n \ge 0\}$
 - G₂=({S},{0,1}, {S→1S0, S→ε}, S)

- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - Design CFG for {w|w=0ⁿ1ⁿ,n≥0}
 - $G_1 = (\{S_1\}, \{0,1\}, \{S_1 \to 0S_11, S_1 \to \varepsilon\}, S_1)$
 - Design CFG for $\{w \mid w=1^n0^n, n \ge 0\}$
 - ► $G_2 = (\{S_2\}, \{0,1\}, \{S_2 \rightarrow 1S_2 0, S_2 \rightarrow \epsilon\}, S_2)$

- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - o Design CFG for $\{w \mid w=0^n1^n, n \ge 0\}$
 - \rightarrow $G_1 = (\{S_1\}, \{0,1\}, \{S_1 \rightarrow 0S_11, S_1 \rightarrow \epsilon\}, S_1)$
 - Design CFG for $\{w \mid w=1^n0^n, n \ge 0\}$
 - ► $G_2 = (\{S_2\}, \{0,1\}, \{S_2 \rightarrow 1S_20, S_2 \rightarrow \epsilon\}, S_2)$







- Design CFG for $\{w \mid w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
 - o Design CFG for $\{w \mid w=0^n1^n, n \ge 0\}$
 - $G_1 = (\{S_1\}, \{0,1\}, \{S_1 \rightarrow 0S_11, S_1 \rightarrow \epsilon\}, S_1)$
 - Design CFG for $\{w \mid w=1^n0^n, n \ge 0\}$
 - ► $G_2 = (\{S_2\}, \{0,1\}, \{S_2 \rightarrow 1S_2 0, S_2 \rightarrow \epsilon\}, S_2)$

o G=({S,S₁,S₂},{0,1}, {S \rightarrow S₁, S \rightarrow S₂, S₁ \rightarrow 0S₁1, S₁ \rightarrow ϵ , S₂ \rightarrow 1S₂0, S₂ \rightarrow ϵ }, S)

Design CFG is much difficult than designing an automata for language

Other ideas:

- 1. Simulate the regular expressions
- 2. Look for a pattern from example strings
- 3. ...

• L={w| w has at least three 1s}, $\Sigma = \{0,1\}$

$$\Sigma^* 1 \Sigma^* 1 \Sigma^* 1 \Sigma^*$$

• L={w| w has at least three 1s}, $\Sigma = \{0,1\}$

$$\Sigma^* 1 \Sigma^* 1 \Sigma^* 1 \Sigma^*$$

$$S \rightarrow R1R1R1R$$

• L={w| w has at least three 1s}, $\Sigma = \{0,1\}$

$$\Sigma^* 1 \Sigma^* 1 \Sigma^* 1 \Sigma^*$$

 $S \rightarrow R1R1R1R$

 $R \rightarrow OR$

 $R \rightarrow 1R$

 $R \rightarrow \epsilon$

• L={w | w has odd length}, $\Sigma = \{0,1\}$

$$\Sigma(\Sigma \Sigma)^*$$

• L={w | w has odd length}, $\Sigma = \{0,1\}$

$$\Sigma(\Sigma \Sigma)^*$$

$$\begin{array}{c}
S \to 0 \\
S \to 1
\end{array}$$

• L={w | w has odd length}, $\Sigma = \{0,1\}$

$$\Sigma(\Sigma \Sigma)^*$$

 $S \rightarrow 0$

 $S \rightarrow 1$

 $S \rightarrow S00$

 $S \rightarrow S01$

 $S \rightarrow S10$

 $S \rightarrow S11$

• L={w | w has odd length}, $\Sigma = \{0,1\}$

$$\Sigma(\Sigma \Sigma)^*$$

 $S \rightarrow 0$

 $S \rightarrow 1$

 $S \rightarrow S00$

 $S \rightarrow S01$

 $S \rightarrow S10$

 $S \rightarrow S11$

• L={w| w has odd length and the middle symbol is 0}, $\Sigma = \{0,1\}$

0

000

001

100

101

00011

• • •

• L={w| w has odd length and the middle symbol is 0}, $\Sigma = \{0,1\}$

 $S \rightarrow 0$

 $S \rightarrow 0S0$

 $S \rightarrow 0S1$

 $S \rightarrow 1S0$

 $S \rightarrow 1S1$

0

000

001

100

101

00011

• • •

• L =
$$\{0^n1^n \mid n \ge 0\}$$
. $\Sigma = \{0,1\}$

$$S \rightarrow 0S1 \mid \epsilon$$

• L =
$$\{0^n1^{2n} \mid n \ge 0\}$$
. $\Sigma = \{0,1\}$

$$S \rightarrow 0S11 \mid \epsilon$$

• L =
$$\{00^*11^*\}$$
. $\Sigma = \{0,1\}$

01, 011, 0011, ...

How to design 00*

How to design 11*

• L =
$$\{00^*11^*\}$$
. $\Sigma = \{0,1\}$

How to design 00*

$$C \rightarrow 0$$

$$C \rightarrow 0C$$

• L =
$$\{00^*11^*\}$$
. $\Sigma = \{0,1\}$

How to design 11*

$$D \rightarrow 1$$

$$D \rightarrow 1D$$

• L =
$$\{00^*11^*\}$$
. $\Sigma = \{0,1\}$

How to design 00*

 $C \rightarrow 0$

 $C \rightarrow 0C$

How to design 00*11*

$$S \rightarrow CD$$

$$C \rightarrow 0C \mid 0$$

$$D \rightarrow 1D \mid 1$$

How to design 11*

$$D \rightarrow 1$$

$$D \rightarrow 1D$$

Ambiguity

- If a grammar generates the same string in several different ways, we say that the string is derived ambiguously in that grammar.
- If a grammar generates some string ambiguously, we say that the grammar is ambiguous.

```
• G_5: E \rightarrow
E+E \mid
E \times E \mid
(E) \mid a
```

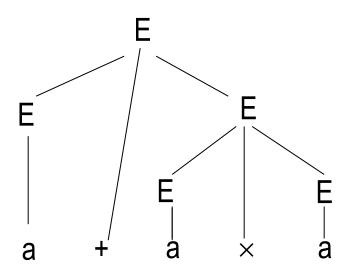
Ambiguity

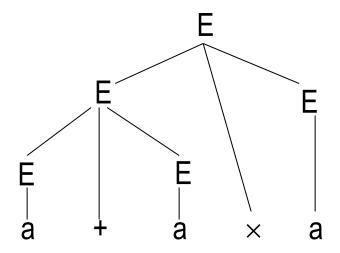
•
$$G_5$$
: $E \rightarrow$

$$E+E \mid$$

$$E\times E \mid$$

$$(E) \mid a$$





Leftmost derivation

A derivation of a string w in a grammar G is a
 leftmost derivation if at every step the *leftmost* remaining variable is the one replaced

•
$$E \Rightarrow E + E$$

$$\Rightarrow$$
 a+E

$$\Rightarrow$$
 a+E×E

$$\Rightarrow$$
 a+a×E \Rightarrow a+a×a

•
$$G_5$$
: $E \rightarrow$

Two different leftmost derivation

- E
 - \Rightarrow E+E
 - \Rightarrow a+E
 - \Rightarrow a+E×E
 - \Rightarrow a+a×E
 - ⇒ a+a×a
- E
 - $\Rightarrow \mathsf{E} \times \mathsf{E}$
 - \Rightarrow E+E \times E
 - \Rightarrow a+E×E
 - \Rightarrow a+a \times E
 - \Rightarrow a+a \times a

- $G_5: E \rightarrow$
 - E+E |
 - $E \times E$
 - (E) | a

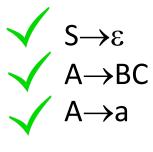
Chomsky normal form (CNF)

CNF: only allow CFG in the following forms

Only start variable S $S \rightarrow \epsilon$ can generate ε $A \rightarrow BC$

Variables can only generate:

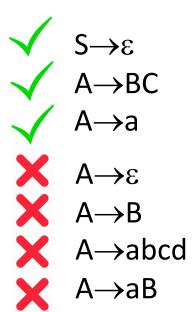
- 1, two variables
- 2, single terminal





Techniques for CNF

- Add new start variable if needed
- A $\rightarrow \epsilon$, merge above rules with A
- A→B, replace B with terminals or other rules
- A \rightarrow aB, replace with U \rightarrow a, A \rightarrow UB
- A \rightarrow abcd, replace with A \rightarrow aU₁, U₁ \rightarrow bU₂, U₂ \rightarrow cd
- A→BCD, similar as the above



Chomsky normal form example

G₆:
$$S \rightarrow ASA \mid aB$$
,
 $A \rightarrow B \mid S$
 $B \rightarrow b \mid \varepsilon$

Get the CNF for G₆



G₆:
$$S \rightarrow ASA \mid aB$$
,
 $A \rightarrow B \mid S$
 $B \rightarrow b \mid \varepsilon$
 $A \rightarrow B$
 $A \rightarrow B$
 $A \rightarrow B$
 $A \rightarrow B$
 $A \rightarrow aB$
 $A \rightarrow aB$

(2a)
$$S_0 \rightarrow S_X$$

$$S \rightarrow ASA \mid aB$$

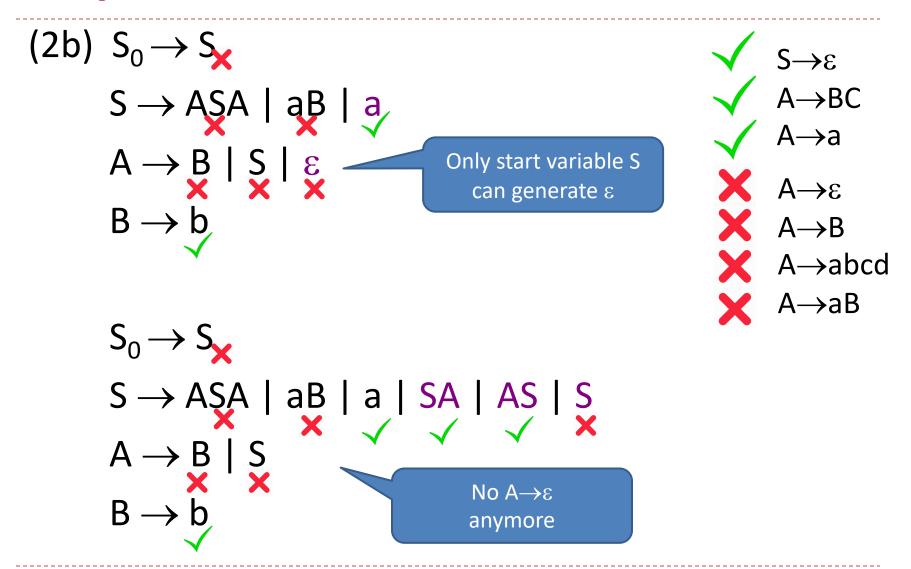
$$A \rightarrow B \mid S$$

$$B \rightarrow b \mid \varepsilon$$
Conly start variable S

$$Can generate \varepsilon$$

$$A \rightarrow B$$

$$A \rightarrow AB$$



(3a)
$$S_0 \rightarrow S_{\bullet}$$

 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS \mid S_{\bullet}$
 $A \rightarrow B \mid S_{\bullet}$
 $B \rightarrow b$
 $S_0 \rightarrow S_{\bullet}$
 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow B \mid S_{\bullet}$
 $B \rightarrow b$

(3b)
$$S_0 \rightarrow S_{\times}$$

 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow B \mid S_{\times}$
 $B \rightarrow b$

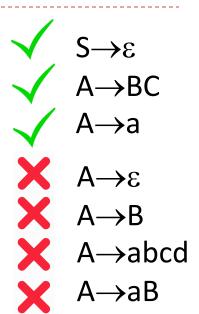
$$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow B \mid S$
 $B \rightarrow b$

∠ A→aB

(3c)
$$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow B \mid S$
 $B \rightarrow b$
 $S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid S$
 $A \rightarrow b \mid S$

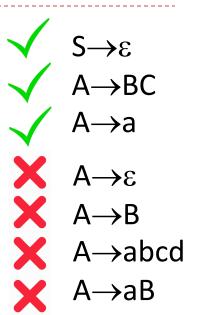


(3d)
$$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid S$
 $B \rightarrow b$

$$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid ASA \mid aB \mid a \mid SA \mid AS$
 $B \rightarrow b$

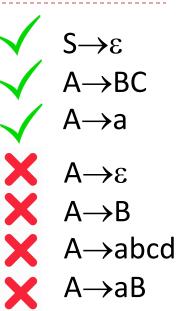


(4)
$$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid ASA \mid aB \mid a \mid SA \mid AS$
 $B \rightarrow b$

$$S_0 \rightarrow AA_1 \mid aB \mid a \mid SA \mid AS$$

 $S \rightarrow AA_1 \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid AA_1 \mid aB \mid a \mid SA \mid AS$
 $B \rightarrow b$
 $A \rightarrow SA$

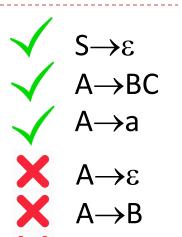


(5)
$$S_0 \rightarrow AA_1 \mid aB \mid a \mid SA \mid AS$$

 $S \rightarrow AA_1 \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid AA_1 \mid aB \mid a \mid SA \mid AS$
 $B \rightarrow b$
 $A_1 \rightarrow SA$

$$S_0 \rightarrow AA_1 \mid UB \mid a \mid SA \mid AS$$

 $S \rightarrow AA_1 \mid UB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid AA_1 \mid UB \mid a \mid SA \mid AS$
 $B \rightarrow b$
 $A_1 \rightarrow SA$
 $A \rightarrow AA_2 \mid AA_3 \mid AA_4 \mid AA_5 \mid AA$



A→abcd

(5)
$$S_0 \rightarrow AA_1 \mid UB \mid a \mid$$

$$SA \mid AS$$

$$S \rightarrow AA_1 \mid UB \mid a \mid$$

$$SA \mid AS$$

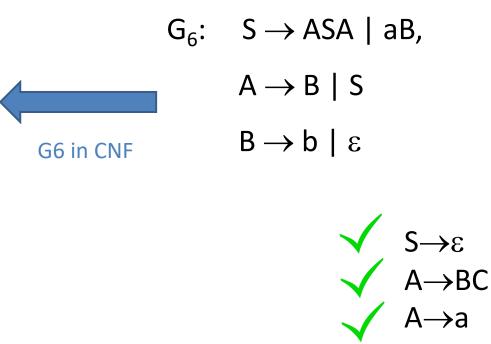
$$A \rightarrow b \mid AA_1 \mid UB \mid a \mid$$

$$SA \mid AS$$

$$B \rightarrow b$$

$$A_1 \rightarrow SA$$

$$U \rightarrow a$$



Χ Α→ε

 $X A \rightarrow B$

A→abcd

A→aB

What does PDA look like?

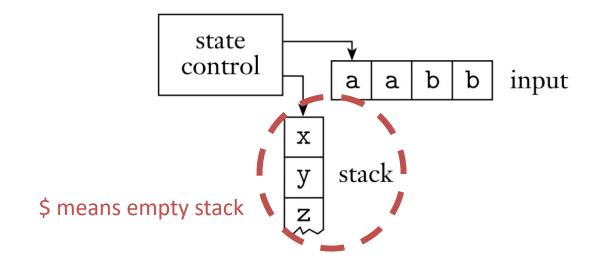
finite automaton

Memory = 1

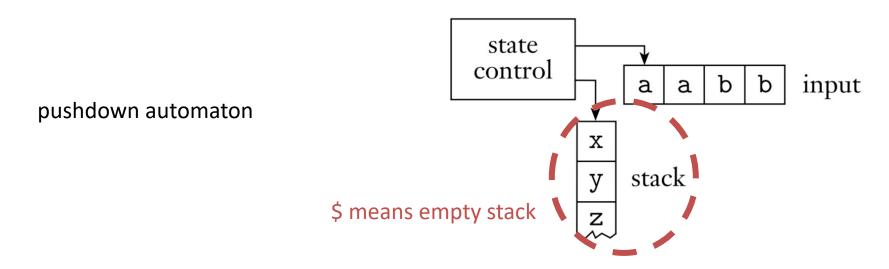
state control a a b b input

pushdown automaton

Memory = N



What does PDA looks like?

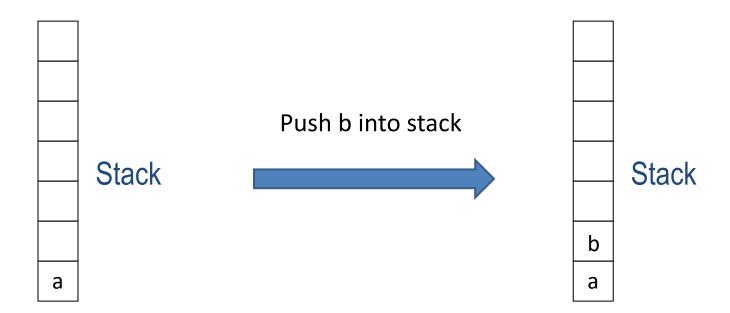


Pushdown automata has more memories than finite automata

PDA = finite automata + A stack (unlimited size)

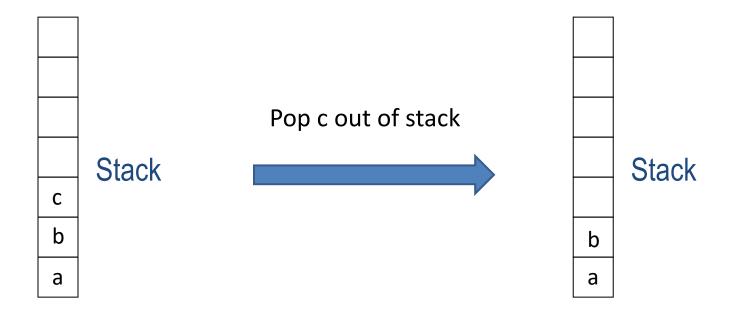
Stack operation

Push: add to the top of stack



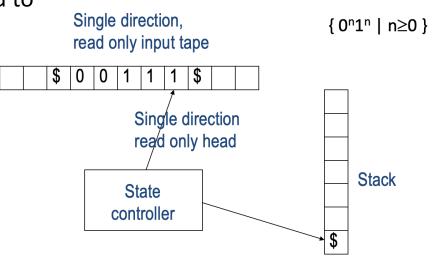
Stack operation

Pop: remove from the top of stack



Informal description for PDA to recognize some languages

- L = {w | w has some features}
- Read symbols from input
 - STEP1: regular?
 - If the language is regular, do not need to use stack; if not regular, define operations on stack
 - STEP2: define operations:
 - When to push
 - When to pop
 - STEP 3: determine accept/reject:
 - Under which cases, accept
 - Under which cases, reject



Definition of PDA (non-deterministic)

- PDA M=(Q, Σ , Γ , δ ,q₀,F), where
 - 1) Q: set of states
 - 2) Σ : input alphabet, $\Sigma_{\varepsilon} = \Sigma \cup \{\varepsilon\}$
 - 3) Γ : stack alphabet, $\Gamma_{\varepsilon} = \Gamma \cup \{\varepsilon\}$
 - 4) $\delta: Q \times \Sigma_{\varepsilon} \times \Gamma_{\varepsilon} \rightarrow P(Q \times \Gamma_{\varepsilon})$, transition function
 - 5) $q_0 \in \mathbb{Q}$: start state
 - 6) F⊆Q: accept state set

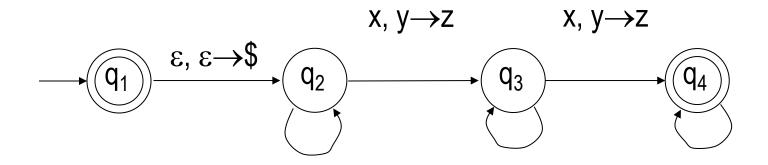
Design PDA

- $L(M_2)=\{a^nb^nc^m | m,n\geq 0\}$
 - Operation:
 - ☐ For an input a, and push a into stack
 - ☐ For an input b, pop one a from the stack
 - Determine accept/reject
 - □ If the stack is empty when finish reading b, then after reading all the cs, accept;
 - □ Otherwise, reject;

Design PDA

• $L(M_2)=\{a^nb^nc^m | m,n\geq 0\}$

x,y->z x: input y->z: the top of stack changes



- $L(M_2)=\{ a^nb^nc^m | m,n\geq 0 \}$
- $X, Y \rightarrow Z$

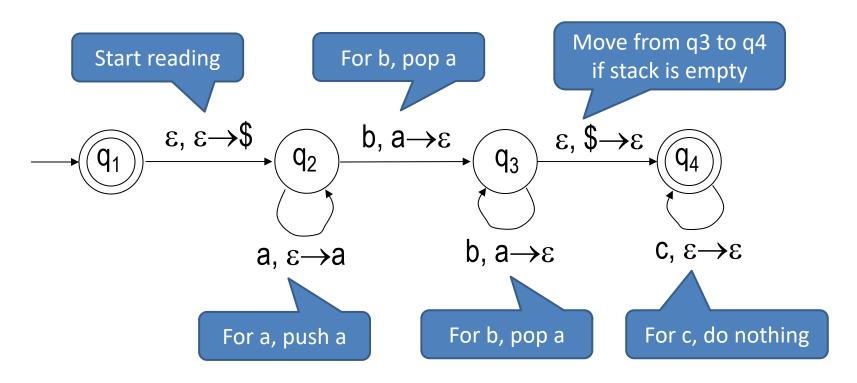
 $X, y \rightarrow Z$

 $X, y \rightarrow Z$

- Operation:
 - ☐ For an input a, and push a into stack
 - ☐ For an input b, pop one a from the stack
- Determine accept/reject
 - ☐ If the stack is empty when finish reading b, then after reading all the cs, accept;
 - Otherwise, reject;

Design PDA

• $L(M_2)=\{a^nb^nc^m | m,n\geq 0\}$



Pumping lemma

Suppose A is CFL,

then there exist a number p(the pumping length) where,

if $s \in A$ and $|s| \ge p$, then s = UVXYZ,

Satisfying the following

- 1) ∀i≥0, uvⁱxyⁱz∈A;
- 2) |vy|>0;
- 3) |vxy|≤p.

1) ∀i≥0, uvⁱxyⁱz∈A;

- 2) |vy|>0;
- 3) |vxy|≤p.

Example: $B=\{a^nb^nc^n \mid n\geq 0\}$

• Proof:

Suppose B is CFL, p is the pumping length,

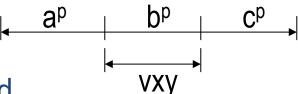
let s=apbpcp

Then s = uvxyz, that

∀i≥0, uvⁱxyⁱz∈B;

|vy|>0, v and y have at least one kind
of symbol;

|vxy|≤p, v and y have at most two
kinds of symbol;



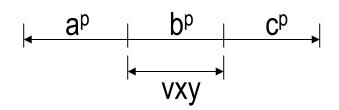
$$\begin{vmatrix} a^p & b^p & c^p \\ \hline \\ vxy & \end{vmatrix}$$

Example: B={ aⁿbⁿcⁿ | n≥0 }

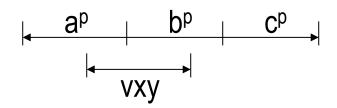
- 1) ∀i≥0, uvⁱxyⁱz∈A;
- 2) |vy|>0;
- 3) |vxy|≤p.

• Proof:

If v and y have one kind of symbol, then in uvⁱxyⁱz (i>1), a/b/c has different numbers;



If v and y have two kinds of symbol, then in uvⁱxyⁱz (i>1), a/b/c has different numbers;



Contradiction.

Exam 2

- 10 True/False question
 - 2 points each

- 4 short answer question
 - 20 points each

Time: check the time @D2L