Theory of Computation

L.EIC, 2nd Year

João M. P. Cardoso

Email: jmpc@acm.org



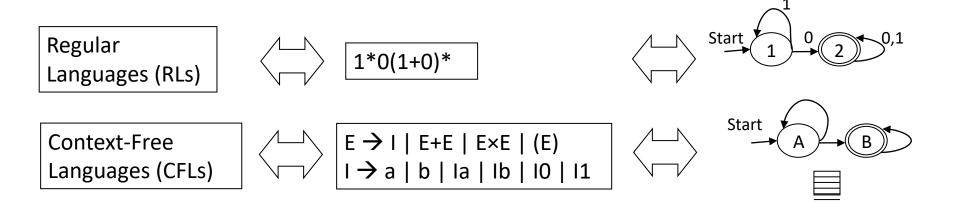


Outline

- ► Push Down Automata (PDAs)
- From CFGs to PDAs
- From PDAs to CFGs

Automata with Stack

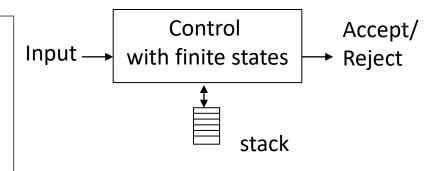
► The automata with stack are to the CFLs as the DFA, NFA, and ϵ -NFA automata are to RLs



Idea

- ► The pushdown automata (PDA) is a ε -NFA with a stack of symbols
 - ▶ Add the possibility to memorize a non-limited quantity of information
 - ▶ The PDA only has access to the top of the stack (LIFO) [no random access]

- How it works:
 - The control unit reads and consumes the input symbols
 - Transition to a new state based on the current state, input symbol, and symbol in the top of the stack
 - Spontaneous transitions with ϵ
 - Top of the stack substituted by symbols



Formal Definition of PDA

- ▶ Pushdown Automaton (PDA) P= (Q, Σ , Γ , δ , q_0 , Z_0 , F)
 - Q: finite set of states
 - $\triangleright \Sigma$: finite set of input symbols
 - $ightharpoonup \Gamma$: finite alphabet of the Stack
 - $\triangleright \delta$: transition function $\delta(q, a, X) = \{(p_1, \gamma_1), ...\}$ finite
 - \triangleright q is a state, a is an input symbol or ε , X is a stack symbol
 - \triangleright p₁ is a new state, γ_1 is the sequence of symbols that substituted X in the top of the stack
 - \triangleright γ₁= ε pop of the top of the stack
 - $\triangleright \gamma_1 = X$ stack is not changed
 - \triangleright γ_1 = YZ X substituted by Z followed by a push of Y
 - \triangleright q₀: initial state
 - ► Z₀: initial symbol in the stack (initial content of the stack)
 - ► F: set of accept/final states

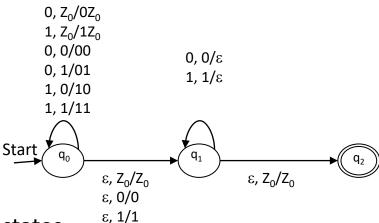
Example of the Palindromes

- $ightharpoonup L_{wwr} = \{ww^R \mid w \in (0+1)^*\}$ palindromes of even length
- CFG of the Language:
 - ▶ P \rightarrow ϵ | 0P0 | 1P1
- Build a PDA
 - ▶ Initial state q₀ means that the PDA didn't reach the middle of ww^R; store in stack the symbols of w
 - At every moment we assume that we may have reached the middle (end of w) and we follow a transition ε to q_1 ; the stack contains w, beginning in the bottom and finishing in the top; the non-determinism is simulated by the maintenance of the two states
 - ▶ In q₁ the PDA compares the input symbol with the top of the stack; if there is not a match, the prediction was wrong and this computing branch dies; other branch might succeed
 - ▶ If the stack gets empty (and the input is finished) the PDA discovered w and w^R

Back to the Palindrome Example

- ► PDA of L_{wwr} P = ({q₀,q₁,q₂}, {0,1}, {0,1,Z₀}, δ , q₀, Z₀, {q₂})
 - $ightharpoonup Z_0$ is used to mark the bottom of the stack and allows in the end of the Reading of ww^R to move the PDA to the accept state q_2
 - $\delta(q_0,0,Z_0) = \{(q_0,0Z_0)\} \in \delta(q_0,1,Z_0) = \{(q_0,1Z_0)\} \text{ top of the stack on the right}$
 - ► $\delta(q_0,0,0) = \{(q_0,00)\}, \ \delta(q_0,0,1) = \{(q_0,01)\}, \ \delta(q_0,1,0) = \{(q_0,10)\}, \ \delta(q_0,1,1) = \{(q_0,11)\}$
 - $\delta(q_0, \varepsilon, Z_0) = \{(q_1, Z_0)\}, \ \delta(q_0, \varepsilon, 0) = \{(q_1, 0)\}, \ \delta(q_0, \varepsilon, 1) = \{(q_1, 1)\}$
 - $\delta(q_1,0,0) = \{(q_1, \epsilon)\}$ and $\delta(q_1,1,1) = \{(q_1, \epsilon)\}$
 - $\delta(q_1, \varepsilon, Z_0) = \{(q_2, Z_0)\}$

Transition Diagram



- Nodes are states
- Arrow Start indicates the initial state
- Edges represent to transitions
 - Label a,X/ α from q to p means that δ (q,a,X) contains (p, α)
 - ► The edge indicate the input and the top of the stack before and after the transition

Instantaneous Description (ID)

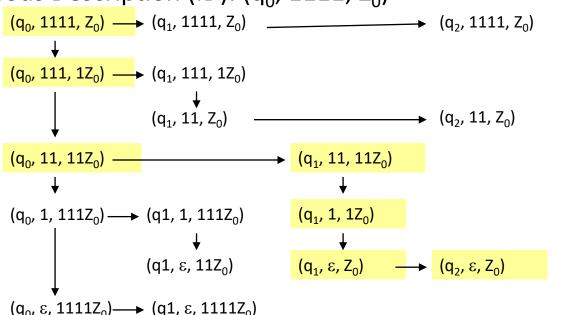
- Computation of a PDA
 - \blacktriangleright Evolves from configuration to configuration, in response to input symbols (or ϵ) and modifying the stack
 - ▶ In a DFA: all the Information in the state;
 - ► In a PDA: state + stack
- Instantaneous description (q,w,γ)
 - q: state
 - w: input reminiscent
 - γ: stack content (top at the left)
- ► Step of a PDA (Q, Σ , Γ , δ , q_0 , Z_0 , F)
 - ► If $\delta(q,a,X)$ contains (p,α) , for all strings w in Σ^* and β in Γ^* \bullet $(q, aw, X\beta)$ $\vdash_{P} (p,w,\alpha\beta)$
 - ► We use |-* for zero or more steps (computation)

The Palindrome Example

ε, 1/1

0, Z₀/0Z₀ 1, Z₀/1Z₀ 0, 0/00

- ► Input w=1111
- lnitial Instantaneous Description (ID): $(q_0, 1111, Z_0)$



- ► Given the PDA P= ({q,p}, {0,1}, {Z₀,X}, δ , q, Z₀, {p}) with
 - $\triangleright \delta(q,0,Z_0) = \{(q,XZ_0)\}$
 - $\delta(q,0,X) = \{(q,XX)\}$
 - $\delta(q,1,X) = \{(q,X)\}$

 - $\triangleright \delta(p,1,X) = \{(p,XX)\}\$
 - ► $\delta(p,1,Z_0) = \{(p, \epsilon)\}$
- Starting with the initial instantaneous description (ID), (q,w,Z_0) , show all the IDs reachable when the input is:
 - ►a) 01 b) 0011 c) 010

Principles Related to IDs

- ▶ If a sequence IDs (computations) is legal for a PDA P then the computations that result of adding any string w to the input in each ID is also legal
- ▶ If a computation is legal for a PDA P then the computations that result of adding any set of symbols below the bottom of the stack in each ID is also legal
 - ► Theorem 1: If (q,x,α) $\vdash^* (p,y,\beta)$ then $(q,xw,\alpha\gamma)$ $\vdash^* (p,yw,\beta\gamma)$
- ▶ If a computation is legal for a PDA P and a tail of the input is not consumed, then the computation that results of removing that tail from the input in each ID is also legal
 - ► Theorem 2: If (q,xw,α) $\vdash^* (p,yw,\beta)$ then (q,x,α) $\vdash^* (p,y,\beta)$

Comments

- Symbols for which P never looks cannot affect its computations
- Similar concept to the notion of context-free language:
 - ▶ What is in the sides does not affect the computation
- ► Theorem 2 is not the inverse of 1 because what is in the stack can influence the computation even if it is not discarded
 - ► For example, it can remove from the stack one symbol in each step and in the last step to add everything that was removed

Language of a PDA

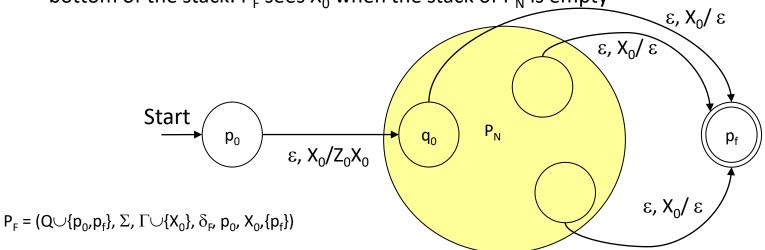
- Accepting by final state
 - ► Given the PDA P = (Q, Σ , Γ , δ , q_0 , Z_0 , F)
 - Language of P accepted by final state
 - ► L(P) = {w | (q_0, w, Z_0) | $+^* (q, \varepsilon, \alpha)$ } and $q \in F$
 - Final content of the stack is irrelevant
- Example:
 - (q_0, ww^R, Z_0) $+ (q_0, w^R, w^RZ_0)$ $+ (q_1, w^R, w^RZ_0)$ $+ (q_1, \varepsilon, Z_0)$ $+ (q_2, \varepsilon, Z_0)$
- Accepting by empty stack

 - Language accepted by empty stack, set of input w consumed by P emptying at the same time the stack (N(P) = stack null)
- ► Same example: modifications to empty the stack and to obtain $N(P_N)=L(P)$
 - $\delta(q_1, \varepsilon, Z_0) = \{(q_2, Z_0)\} \text{ becomes } \delta(q_1, \varepsilon, Z_0) = \{(q_2, \varepsilon)\}$
 - (q_0, ww^R, Z_0) | (q_0, w^R, w^RZ_0) | (q_1, w^R, w^RZ_0) | (q_1, ε, Z_0) | $(q_2, \varepsilon, \varepsilon)$

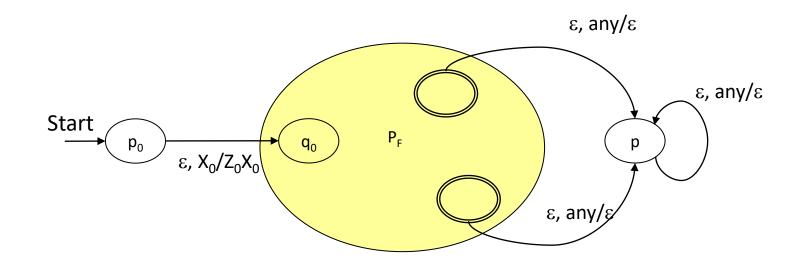
From the Empty Stack to the Final State

- ► Theorem: If L = N(P_N) for a PDA P_N = (Q, Σ , Γ , δ _N, q₀, Z₀) then there exists a PDA P_F such that L = L(P_F)
 - Two equivalent methods for accepting an input
 - ▶ While for a PDA P we can have $L(P) \neq N(P)$

Starting from P_N , use a new $X_0 \notin \Gamma$ as initial symbol of P_F and as a mark of the bottom of the stack: P_F sees X_0 when the stack of P_N is empty

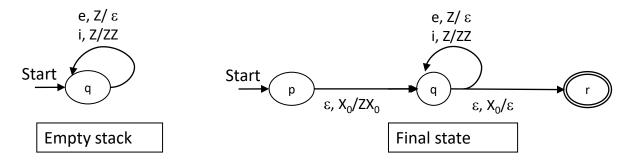


From the Final State to the Empty Stack



Example of Conversion

- ▶ Define a PDA which processes sequences formed with "i" and "e", meaning the *if* and *else*, constructions present in many programming languages, detecting invalid sequences (i.e., sequences with more "e's" than "i's" in a prefix)
 - ▶ Initial symbol: Z; stack with Z^n means that no. of i's no. e's = n-1
 - Accept by empty stack
 - Conversion to final state



Equivalence between PDAs and CFGs

- It is proved that the CFLs defined by a CFG are the languages accepted by a PDA by empty stack and thus also accepted by a PDA by final state
- ▶ Idea: given a CFG G build a PDA that simulates the leftmost derivations of G
 - \triangleright Any left syntax form non-terminal can be written as $xA\alpha$,
 - ▶ Where A if the leftmost variable,
 - x are the terminals in the left of A,
 - \blacktriangleright and α is the sequence of terminals and variables in the right of A.
 - \triangleright A α is named tail
 - \triangleright CFG G = (V,T,Q,S)
 - ▶ PDA that accepts L(G) by empty stack: P = ({q}, T, V \cup T, δ , q, S)
 - For each variable A:
 - $\triangleright \delta(q,\varepsilon,A)=\{(q,\beta) \mid A \rightarrow \beta \text{ is a production in G}\}$
 - For each terminal a:
 - \triangleright $\delta(q,a,a)=\{(q,\epsilon)\}$

From CFGs to PDAs

Given the CFG

```
E → I | E+E | E×E | (E)
I → a | b | Ia | Ib | IO | I1
```

- ▶ Obtain a PDA for accepting the same language by empty stack
 - $ightharpoonup P_N = (\{q\}, \{a,b,0,1,(,),+,\times\}, \{a,b,0,1,(,),+,\times,E,I\}, \delta, q, E)$
 - $\delta(q,\epsilon,I) = \{(q,a), (q,b), (q,Ia), (q,Ib), (q,I0), (q,I1)\}$

 - ▶ $\delta(q,a,a) = \{(q,\epsilon)\}; \delta(q,b,b) = \{(q,\epsilon)\}, \delta(q,0,0) = \{(q,\epsilon)\}; \delta(q,1,1) = \{(q,\epsilon)\}; \delta(q,(,() = \{(q,\epsilon)\}; \delta(q,),)) = \{(q,\epsilon)\}; \delta(q,+,+) = \{(q,\epsilon)\}; \delta(q,\times,\times) = \{(q,\epsilon)\}$
- Only one state
- ▶ Processing of variables is spontaneous
- Only the terminals consume inputs

- Using a CFG and the PDA for the language of expressions
 - a) Obtain a leftmost derivation for a×(a+b00)
 - b) Obtain the computing trace for the PDA, i.e., the sequence of the instantaneous descriptions

- Convert the CFG below to a PDA:
- 1. $S \rightarrow aAA$
- 2. $A \rightarrow aS/bS/a$

- Convert to PDA, the CFG with the following productions:
- 1. $A \rightarrow aAA$
- 2. $A \rightarrow aS \mid bS \mid a$
- 3. $S \rightarrow SS \mid (S) \mid \epsilon$
- 4. $S \rightarrow aAS \mid bAB \mid aB$
- 5. $A \rightarrow bBB \mid aS \mid a$
- 6. $B \rightarrow bA \mid a$

From PDAs to CFGs

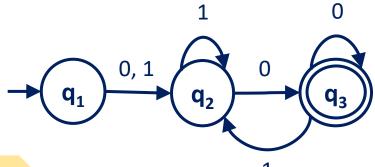
Let's Start by Converting FAs to CFGs

- ► One CFG variable for each FA state (e.g., q_i represented by L_i)
- ▶ One CFG rule for each transition
- For each accept state q_i we include $L_i \rightarrow \epsilon$
- Example:

$$L_1 \rightarrow 0L_2 \mid 1L_2$$

$$L_2 \rightarrow 1L_2 \mid 0L_3$$

$$L_3 \rightarrow 1L_2 \mid 0L_3 \mid \epsilon$$



The CFG obtained is usually identified as right-linear of right regular grammars (https://en.wikipedia.org/wiki/Linear grammar)

From PDAs to CFGs

- ► Idea:
 - Recognizing that the main event of the PDA processing is the pop of a symbol while consuming the input
- ► Add variables to the grammar for
 - Each elimination of a stack symbol X
 - ► Each state transition from p to q eliminating X, represented by a compound symbol [pXq]
- From PDA P= (Q, Σ , Γ , δ _N, q₀, Z₀) build CFG G= (V, Σ , R, S)
 - ► Variables V: contain S and the compound symbols [pXq]

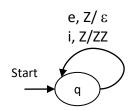
From PDAs to CFGs (cont.)

► Productions R:

- ► For all states p, G contains S \rightarrow [q₀Z₀p] (q₀ is the start state of the PDA)
 - Symbol $[q_0Z_0p]$ generates all the strings w that pop Z_0 from the stack while going from state q_0 to state p, (q_0, w, Z_0) $\vdash^* (p, \varepsilon, \varepsilon)$
 - ▶ Hence S generates all the strings w that empty the stack
- ▶ If $\delta(q,a,X)$ contains $(r, Y_1Y_2...Y_k)$, $k \ge 0$, $a \in \Sigma$ or $a = \varepsilon$, then for all the lists of states $r_1, r_2, ..., r_k$, G contains (when k = 0, the pair is (r, ε))

$$[qXr_k] \rightarrow a[rY_1r_1][r_1Y_2r_2]...[r_{k-1}Y_kr_k]$$

A way to pop X and to go from q to r_k is to read a (it can be ϵ) and use some input to pop Y_1 while going from r to r_1 , etc.



- ► Convert the PDA $P_N = (\{q\}, \{i,e\}, \{Z\}, \delta_N, q, Z)$ to a CFG
 - Accept strings that for the first time don't follow that each "e" needs to correspond to a previous "i"
- ► Solution:
 - Only a state q and a stack symbol Z
 - Two variables: S, start symbol; [qZq], unique symbol from the states and symbols of P_N
 - ▶ Production:
 - ► S \rightarrow [qZq] (if there were more states p and r we would have S \rightarrow [qZp] and S \rightarrow [qZr])
 - ► From $\delta_N(q,i,Z)=\{(q,ZZ)\}$ obtain $[qZq]\rightarrow i[qZq]$ (if there were more states p and r we would have $[qZp]\rightarrow i[qZr]$ [rZp]) ...
 - ► From $\delta_N(q,e,Z)=\{(q,\epsilon)\}$ obtain $[qZq]\rightarrow e$ (Z is substituted by nothing)
 - ▶ Naming A to [qZq] we obtain $S \rightarrow A e A \rightarrow iAA \mid e$

 $\triangleright P_N = (\{q,r\},\{0,1\},\{X,Z\},\delta_N,q,Z)$

1.
$$\delta(q, 0, Z) = \{(q, XZ)\}$$

2.
$$\delta(q, 0, X) = \{(q, XX)\}$$

3.
$$\delta(q, 1, X) = \{(r, \epsilon)\}$$

4.
$$\delta(r, 1, X) = \{(r, \varepsilon)\}$$

5.
$$\delta(r, \varepsilon, Z) = \{(r, \varepsilon)\}$$

State	input	stack	New state	stack
q	0	Z	q	XZ
q	0	Χ	q	XX
q	1	X	r	3
r	1	Χ	r	3
r	3	Z	r	3

- $ightharpoonup P_N = (\{q,r\}, \{0,1\}, \{X,Z\}, \delta_N, q, Z)$
- Possible V = {S, [qZq], [qZr], [qXq], [qXr], [rZq], [rZr], [rXq], [rXr]}
- ► [qZq], [qXq], [qXr], [rXr], [rZr]
- \triangleright S \rightarrow [qZq] | [qZr]
- $[qZq] \rightarrow 0[qXq][qZq] \mid 0 [qXr][rZq]$
- $[qXq] \rightarrow 0[qXq][qXq] | 0[qXr][rXq]$
- $[qXr] \rightarrow 0[qXq][qXr] | 0[qXr][rXr]$
- ightharpoonup [rZq] \rightarrow
- $[qZr] \rightarrow 0[qXr][rZr] \mid 0[qXq][qZr]$

State	input	stack	New state	stack
q	0	Z	q	XZ
q	0	Χ	q	XX
q	1	Χ	r	3
r	1	Χ	r	3
r	3	Z	r	3

- ightharpoonup [rXq] \rightarrow
- ightharpoonup [rZr] ightharpoonup ϵ
- ightharpoonup [rXr] \rightarrow 1
- ightharpoonup [qXr] ightharpoonup 1

- Possible V = {S, [qZq], [qZr], [qXq], [qXr], [rZq], [rZr], [rXq], [rXr]}
- ► [qZq], [qXq], [qXr], [rXr], [rZr]
- \triangleright S \rightarrow [qZq] | [qZr]
- ightharpoonup [qZq] \rightarrow 0[qXq][qZq]
- \triangleright [qXq] \rightarrow 0[qXq][qXq]
- ightharpoonup [qXr] \rightarrow 0[qXq][qXr] | 0[qXr][rXr] | 1
- \triangleright [qZr] \rightarrow 0[qXr][rZr] | 0[qXq][qZr]

State	input	stack	New state	stack
q	0	Z	q	XZ
q	0	Χ	q	XX
q	1	Χ	r	3
r	1	Χ	r	3
r	3	Z	r	3

- ightharpoonup [rZr] $\rightarrow \varepsilon$
- ightharpoonup [rXr] ightharpoonup 1

- Possible V = {S, [qZq], [qZr], [qXq], [qXr], [rZq], [rZr], [rXq], [rXr]}
- ► [qZq], [qXq], [qXr], [rXr], [rZr]
- \triangleright S \rightarrow [qZq] | [qZr]
- \triangleright [qZq] \rightarrow 0[qXq][qZq]
- ightharpoonup [qXq] ightharpoonup 0[qXq][qXq]
- ightharpoonup [qXr] \rightarrow 0[qXq][qXr] | 0[qXr][rXr] | 1
- \triangleright [qZr] \rightarrow 0[qXr][rZr] | 0[qXq][qZr]

State	input	stack	New state	stack
q	0	Z	q	XZ
q	0	Χ	q	XX
q	1	X	r	3
r	1	Χ	r	3
r	3	Z	r	3

- ightharpoonup [rZr] $\rightarrow \varepsilon$
- ightharpoonup [rXr] ightharpoonup 1

State	input	stack	New state	stack
q	0	Z	q	XZ
q	0	Χ	q	XX
q	1	Χ	r	3
r	1	Χ	r	3
r	3	Z	r	3

- Possbile V = {S, [qZq], [qZr], [qXq], [qXr], [rZq], [rZr], [rXq], [rXr]}
- ► [qZq], [qXq], [qXr], [rXr], [rZr]

- ightharpoonup [rZr] $\rightarrow \varepsilon$
- ightharpoonup [rXr] \rightarrow 1

- \triangleright S \rightarrow [qZq] | [qZr]
- ightharpoonup [qXr] ightharpoonup 0[qXr][rXr] | 1
- \triangleright [qZr] \rightarrow 0[qXr][rZr] | 0[qXq][qZr]

- Possible V = {S, [qZq], [qZr], [qXq], [qXr], [rZq], [rZr], [rXq], [rXr]}
- ► [qZq], [qXq], [qXr], [rXr], [rZr]
- \triangleright S \rightarrow [qZr]
- ightharpoonup [qXr] \rightarrow 0[qXr][rXr] | 1
- \triangleright [qZr] \rightarrow 0[qXr][rZr]

State	input	stack	New state	stack
q	0	Z	q	XZ
q	0	X	q	XX
q	1	X	r	3
r	1	Χ	r	3
r	3	Z	r	3

- ightharpoonup [rZr] $\rightarrow \varepsilon$
- ightharpoonup [rXr] \rightarrow 1

- Possible V = {S, [qZq], [qZr], [qXq], [qXr], [rZq], [rZr], [rXq], [rXr]}
- ► [qZq], [qXq], [qXr], [rXr], [rZr]
- \triangleright S \rightarrow [qZr]
- ightharpoonup [qXr] \rightarrow 0[qXr][rXr] | 1
- ightharpoonup [qZr] \rightarrow 0[qXr][rZr]

State	input	stack	New state	stack
q	0	Z	q	XZ
q	0	Χ	q	XX
q	1	X	r	3
r	1	Χ	r	3
r	3	Z	r	3

- ightharpoonup [rZr] $\rightarrow \varepsilon$
- ightharpoonup [rXr] \rightarrow 1

CFG:
$$V = \{S, [qZr], \underline{[qXr]}, \}$$

- ►S \rightarrow [qZr]
- ightharpoonup [qXr] ightharpoonup 0[qXr] 1 | 1
- ightharpoonup [qZr] ightharpoonup 0[qXr]

Substituting [qZr] by A and [qXr] by B:

- $\triangleright S \rightarrow A$
- \triangleright B \rightarrow 0B1 | 1
- $\triangleright A \rightarrow OB$

State	input	stack	New state	stack
q	0	Z	q	XZ
q	0	Χ	q	XX
q	1	Χ	r	3
r	1	Χ	r	3
r	3	Z	r	3

- ► PDA P = ({p,q}, {0,1}, {X,Z}, δ , q, Z)), with the following transition function:
 - 1. $\delta(q, 1, Z) = \{(q, XZ)\}$
 - 2. $\delta(q, 1, X) = \{(q, XX)\}$
 - 3. $\delta(q, 0, X) = \{(p, X)\}$
 - 4. $\delta(q, \varepsilon, X) = \{(q, \varepsilon)\}$
 - 5. $\delta(p, 1, X) = \{(p, \varepsilon)\}$
 - 6. $\delta(p, 0, Z) = \{(q, Z)\}$
- Convert it to a CFG

- PDA P = ({p,q}, {0,1}, {X,Z}, δ, q, Z)), with the following transition function:
 - 1. $\delta(q, 1, Z) = \{(q, XZ)\}$
 - 2. $\delta(q, 1, X) = \{(q, XX)\}$
 - 3. $\delta(q, 0, X) = \{(p, X)\}$
 - 4. $\delta(q, \varepsilon, X) = \{(q, \varepsilon)\}$
 - 5. $\delta(p, 1, X) = \{(p, \varepsilon)\}$
 - 6. $\delta(p, 0, Z) = \{(q, Z)\}$
- Convert it to a CFG

S is the start symbol: From rule (3): $[qXq] \rightarrow 0[pXq]$ $S \rightarrow [qZq] \mid [qZp]$ From rule (1): $[qXp] \rightarrow 0[pXp]$ $[qZq] \rightarrow 1[qXq][qZq]$ From rule (4): $[qZq] \rightarrow 1[qXp][pZq]$ $[qXq] \rightarrow \varepsilon$ $[qZp] \rightarrow 1[qXq][qZp]$ From rule (5): $[qZp] \rightarrow 1[qXp][pZp]$ $[pXp] \rightarrow 1$ From rule (2): From rule (6): $[qXq] \rightarrow 1[qXq][qXq]$ $\lceil pZq \rceil -> 0 \lceil qZq \rceil$ $[qXq] \rightarrow 1[qXp][pXq]$ $[pZp] \rightarrow 0[qZp]$ $[qXp] \rightarrow 1[qXq][qXp]$

 $[qXp] \rightarrow 1[qXp][pXp]$

- ► PDA P= ({q,p}, {0,1}, {Z₀,X}, δ , q, Z₀, {p}) with transition function
 - $\delta(q,0,Z_0) = \{(q,XZ_0)\}$
 - $\delta(q,0,X) = \{(q,XX)\}$
 - ► $\delta(q,1,X) = \{(q,X)\}$

 - $\delta(p,1,X) = \{(p,XX)\}$
- Obtain a CFG