Middle East Technical University Department of Computer Engineering

Formal Languages and Abstract Machines
Take Home Exam 2

Mert AKÇA
2171163

1 Context-Free Grammars (10 pts)

a) Give the rules of the Context-Free Grammars to recognize strings in the given languages where $\Sigma = \{a, b\}$ and S is the start symbol.

$L(G) = \{w \mid \ w \in \Sigma^*;\ |w| \geq 3;$ the first and the second from the last symbols of w are the same$\}$ (2/10 pts)

$S \rightarrow Aaa,$
$S \rightarrow Abb,$
$A \rightarrow Ba,$
$A \rightarrow Bb,$
$B \rightarrow Bb \mid Ba \mid e$

$L(G) = \{w \mid \ w \in \Sigma^*;$ the length of w is odd$\}$ (2/10 pts)

$S \rightarrow AaA \mid AbA$
$A \rightarrow aAa \mid aAb \mid bAa \mid bAb \mid e$

$L(G) = \{w \mid \ w \in \Sigma^*;\ n(w, a) = 2 \cdot n(w, b)\}$ where $n(w, x)$ is the number of x symbols in w (3/10 pts)

$S \rightarrow A \mid e$
$A \rightarrow aAaAb \mid aAbAa \mid bAaAa \mid e$

b) Find the set of strings recognized by the CFG rules given below: (3/10 pts)

$S \to X \mid Y$
$X \to aXb \mid A \mid B$
$A \to aA \mid a$
$B \to Bb \mid b$
$Y \to CbaC$
$C \to CC \mid a \mid b \mid \varepsilon$

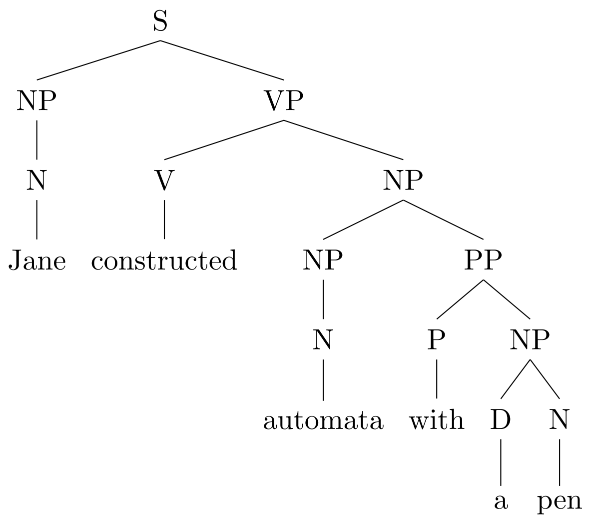
$\boxed{\text{L(G)} = \{a^n(a^+ \cup b^+)b^n, (a \cup b)^*ba(a \cup b)^*\}}$

2 Parse Trees and Derivations (20 pts)

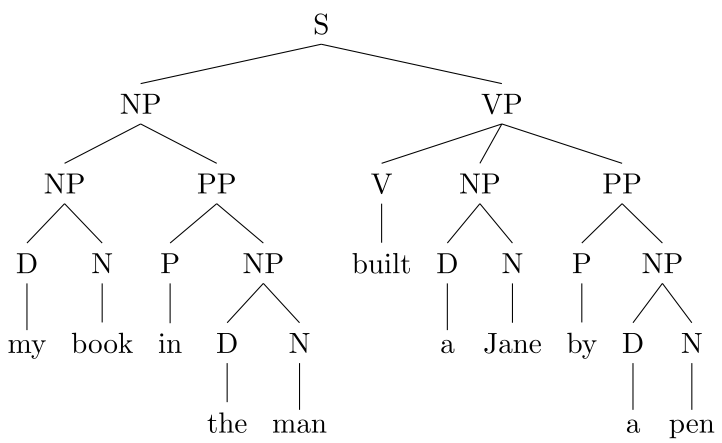
Given the CFG below, provide parse trees for given sentences in **a** and **b**.

```
S   → NP VP
VP  → V NP | V NP PP
PP  → P NP
NP  → N | D N | NP PP
V   → wrote | built | constructed
D   → a | an | the | my
N   → John | Mary | Jane | man | book | automata | pen | class
P   → in | on | by | with
```

a) Jane constructed automata with a pen (4/20 pts)



b) my book in the man built a Jane by a pen (4/20 pts)

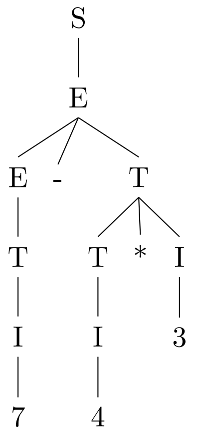


Given the CFG below, answer **c**, **d** and **e**

```
S  → E
E  → E + T | E - T | T
T  → T * I | T / I | I
I  → 0 | 1 | 2 | 3 | 4 | 6 | 7 | 8 | 9
```

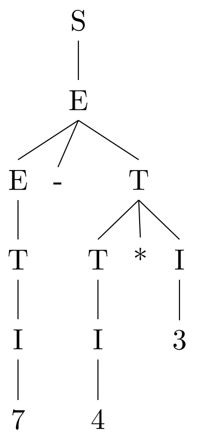
c) Provide the left-most derivation of 7 - 4 * 3 step-by-step and plot the final parse tree matching that derivation (4/20 pts)

$D = S \to E \to E - T \to T - T \to I - T \to 7 - T \to 7 - T * I \to 7 - I * I \to 7 - 4 * I \to 7 - 4 * 3$



d) Provide the right-most derivation of 7 - 4 * 3 step-by-step and plot the final parse tree matching that derivation (4/20 pts)

$D = S \to E \to E - T \to E - T * I \to E - T * 3 \to E - I * 3 \to E - 4 * 3 \to T - 4 * 3 \to I - 4 * 3 \to 7 - 4 * 3$

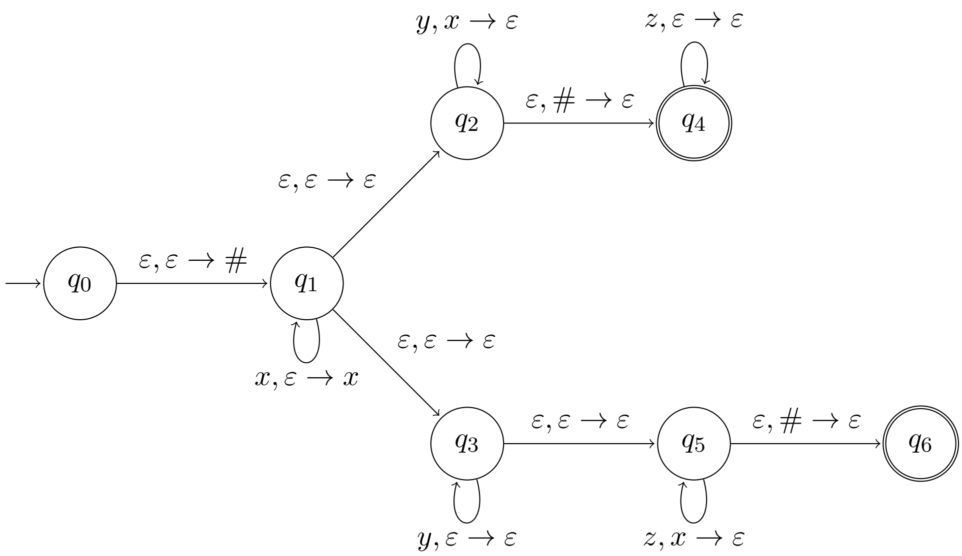


e) Are the derivations in **c** and **d** in the same similarity class? (4/20 pts)

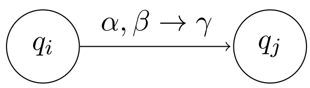
Yes, their final parse trees are the same

3 Pushdown Automata (30 pts)

a) Find the language recognized by the PDA given below (5/30 pts)

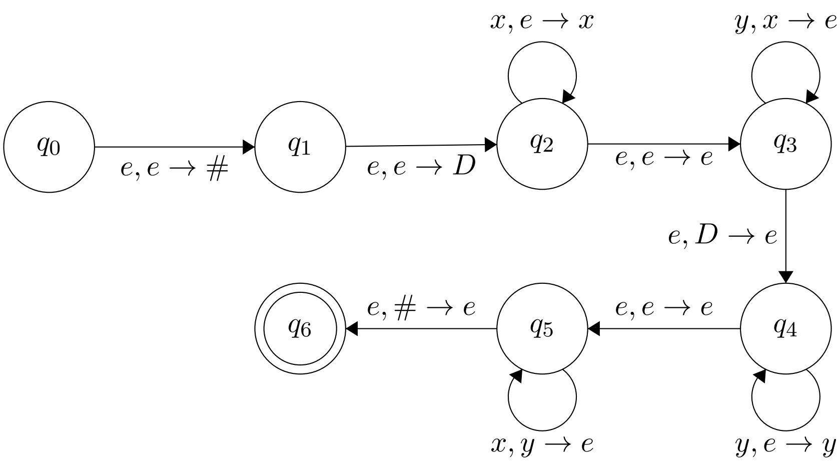


where the transition $((q_i, \alpha, \beta), (q_j, \gamma))$ is represented as:



L = $x^n y^n z^*$ or $x^n y^* z^n$, where $n \in N$

b) Design a PDA to recognize language $L = \{x^n y^{m+n} x^m \mid\ n, m \geq 0;\ n, m \in \mathbb{N}\}$ (5/30 pts)



$\to q_0$ is the start state

c) Design a PDA to recognize language $L = \{x^n y^m \mid n < m \leq 2n;\ n, m \in \mathbb{N}^+\}$ (10/30 pts)
Do not use multi-symbol push/pop operations in your transitions.
Simulate the PDA on strings xxy (with only one rejecting derivation) and $xxyyyy$ (accepting derivation) with transition tables.

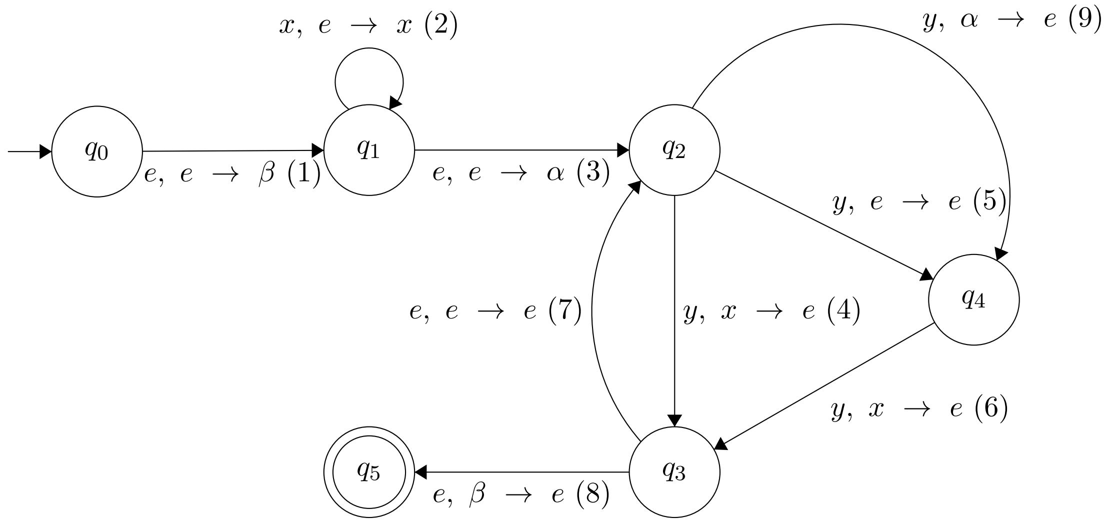


Table 1: Transition Table for xxy

State	Unread Input	Stack	Transition
q_0	xxy	e	-
q_1	xxy	β	1
q_1	xy	$x\beta$	2
q_1	y	$xx\beta$	2
q_2	y	$\alpha xx\beta$	3
q_3	e	$xx\beta$	9

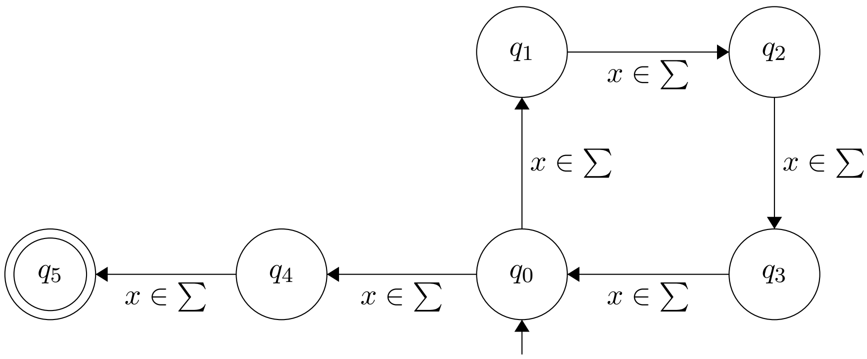
β *is not top of the stack so we cannot move final state*

Table 2: Transition table for xxyyyy

State	Unread Input	Stack	Transition
q_0	xxyyyy	e	-
q_1	xxyyyy	β	1
q_1	xyyyy	$x\beta$	2
q_1	yyyy	$xx\beta$	2
q_2	yyyy	$\alpha xx\beta$	3
q_4	yyy	$xx\beta$	9
q_3	yy	$x\beta$	6
q_2	yy	$x\beta$	7
q_4	y	$x\beta$	5
q_3	e	β	6
q_5	e	e	8

d) Given two languages L' and L as $L' = \{w \mid w \in L;\ |w| = 4n + 2\ for\ n \in \mathbb{N}\}$ (10/30 pts)
If L is a CFL, show that L' is also a CFL by constructing an automaton for L' in terms of another automaton that recognizes L.

Let L'' be a regular language $L'' = \{w \mid |w| = 4n + 2\ for\ n \in N\}$. This language has a deterministic automaton $M_1 = (K_1, \sum, \delta, s_1, F_1)$



As we can see L'' is an regular language because it's accepted by an finite automaton.

Intersection of language L and L'' (strings with length of 4n+2) is L'. L is a CFL and L" is a regular language, so L' must be also a CFL by Theorem 3.5.2 (textbook, p.144).

4 Closure Properties (20 pts)

Let L_1 and L_2 be context-free languages which are not regular, and let L_3 be a regular language. Determine whether the following languages are necessarily CFLs or not. If they need to be context-free, explain your reasoning. If not, give one example where the language is a CFL and a counter example where the language is not a CFL.

a) $L_4 = L_1 \cap (L_2 \setminus L_3)$ (10/20 pts)

$L_4 = L_1 \cap (L_2 \cap L_3')$
L_2 is a CFL and L_3' is a regular language (Complement of a regular language is also regular) so $(L_2 \cap L_3')$ is CFL
Therefore, L_4 is intersection of two CFL's, so we cannot decide whether it's a CFL or not.

b) $L_5 = (L_1 \cap L_3)^*$ (10/20 pts)

$(L_1 \cap L_3)$ is intersection of a regular language and context-free lang. so we can say that $(L_1 \cap L_3)$ is CFL
Because CFL's are closed under kleene star property $(L_1 \cap L_3)^*$ is also a CFL.

5 Pumping Theorem (20 pts)

a) Show that $L = \{a^n m^n t^i \mid n \leq i \leq 2n\}$ is not a Context Free Language (10/20 pts)
using Pumping Theorem for CFLs.

Let pumping length = 4, $|vxy| \leq 4$
S = aaaammmmtttt(tttt)

Case 1: $|vxy|$ is in the first boundary
$aa\underbrace{\text{aamm}}_{\text{vxy}}mmtttt(tttt)$ for $i = 2$;
aaaaaammmmmmtttt(tttt) : number of a's is not equal to number of m's.

Case 2: $|vxy|$ is not in boundary it's in a's or m's
$\underbrace{\text{aaaa}}_{\text{vxy}}mmmmtttt(tttt)$ for $i = 2$;
aaaaaammmmtttt(tttt) : number of a's is not equal to number of m's.
Also:
$aaaa\underbrace{\text{mmmm}}_{\text{vxy}}tttt(tttt)$ for $i = 2$;
aaaammmmmmtttt(tttt) : number of a's is not equal to number of m's.

Case 3: $|vxy|$ is in t's
$aaaammmm\underbrace{\text{tttt}}_{\text{vxy}}(tttt)$ for $i = 4$;
aaaammmmtttttttttt(tttt) : the mininmum number of t's is greater than 2n.

Therefore, the rule is not met in all cases for every $i \geq 0$, this language is not context-free language.

b) Show that $L = \{a^n b^{2n} a^n \mid n \in \mathbb{N}+\}$ is not a Context Free Language (10/20 pts)
using Pumping Theorem for CFLs.

Let pumping length = 3, $|vxy| \leq 3$
S = aaabbbbbbaaa

Case 1: $|vxy|$ is in the boundaries
$a\underbrace{aab}_{\text{vxy}}bbbbbaaa$ for $i = 2$;
aaaabbbbbbbaaa : the number of a's is not equal.
Also:
$aaabbbb\underbrace{bba}_{\text{vxy}}aa$ for $i = 2$;
aaabbbbbbbaaaa : the number of a's is not equal.

Case 2: $|vxy|$ is not in the boundaries

$aaa\underbrace{bbb}_{\text{vxy}}bbbaaa$ for $i = 2$;

aaabbbbbbbbbaaa : the number of b's is not double of number of a's.

Also:

$aaabbbbbb\underbrace{aaa}_{\text{vxy}}$ for $i = 2$;

aaabbbbbbaaaaaa : the number of a's is not equal.

Therefore, the rule is not met in all cases for every $i \geq 0$, this language is not context-free language.

6 CNF and CYK (not graded)

a) Convert the given context-free grammar to Chomsky Normal Form.

$S \to XSX \mid xY$
$X \to Y \mid S$
$Y \to z \mid \varepsilon$

answer here ...

b) Use the grammar below to parse the given sentence using Cocke–Younger–Kasami algorithm. Plot the parse trees.

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP
NP → I | she | me | Houston
NP → Det Nom
Nom → book | flight | meal | money
Nom → Nom Noun
Nom → Nom PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP
PP → Prep NP
Det → that | this | the | a
Noun → book | flight | meal | money
Verb → book | include | prefer
Aux → does
Prep → from | to | on | near | through

book the flight through Houston

Empty parse table:

1:5 → 1:1 2:5 1:5 → 1:2 3:5 1:5 → 1:3 4:5 1:5 → 1:4 5:5				
1:4 → 1:1 2:4 1:4 → 1:2 3:4 1:4 → 1:3 4:4	2:5 → 2:2 3:5 2:5 → 2:3 4:5 2:5 → 2:4 5:5			
1:3 → 1:1 2:3 1:3 → 1:2 3:3	2:4 → 2:2 3:4 2:4 → 2:3 4:4	3:5 → 3:3 4:5 3:5 → 3:4 5:5		
1:2 → 1:1 2:2	2:3 → 2:2 3:3	3:4 → 3:3 4:4	4:5 → 4:4 5:5	
1:1	2:2	3:3	4:4	5:5
book	the	flight	through	Houston

rest of the answer here ...

7 Deterministic Pushdown Automata (not graded)

Provide a DPDA to recognize the given languages, the DPDA must read its entire input and finish with an empty stack.

a) $a^*bc \cup a^n b^n c$

answer here ...

b) $(aa)^*c \cup a^n b^n c$

answer here ...