Middle East Technical University Department of Computer Engineering

Formal Languages and Abstract Machines
Take Home Exam 2

Kürşat Öztürk
2171874

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)

$S \rightarrow aAaa|aAab|bAba|bAbb$
$A \rightarrow a|b|AA|e$

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

$S \rightarrow aA|bA$
$A \rightarrow aa|ab|ba|bb|AA|e$

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

$S \rightarrow aaA|aAa|Aaa|e$
$A \rightarrow bS|Sb|b$

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

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

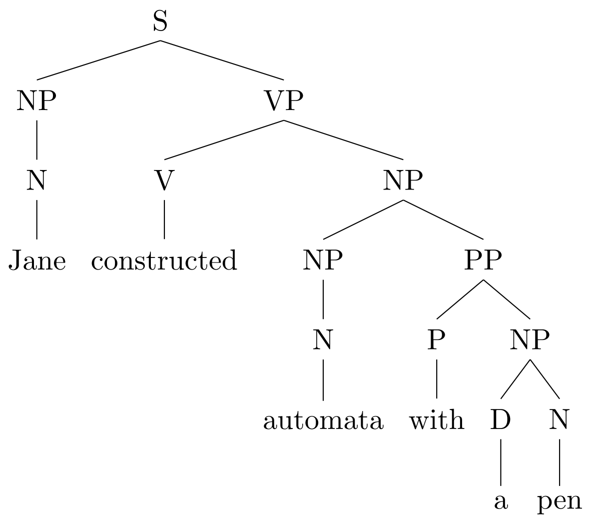
$L(G) = \{w | w \in \{a, b\}^* - \{ab\}; |w| > 0\}$

2 Parse Trees and Derivations (20 pts)

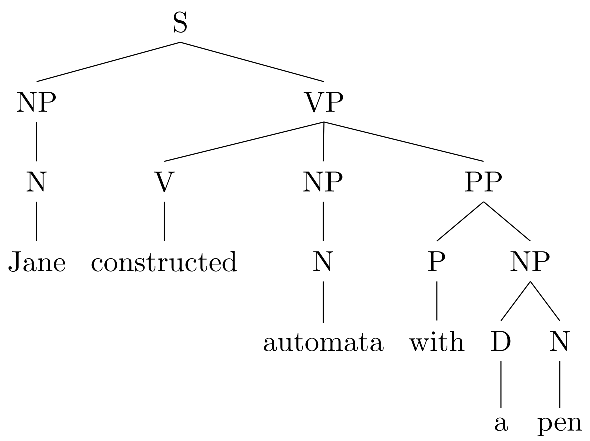
Given the CFG below, provide parse trees for given sentences in **a** and **b**.

```
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
```

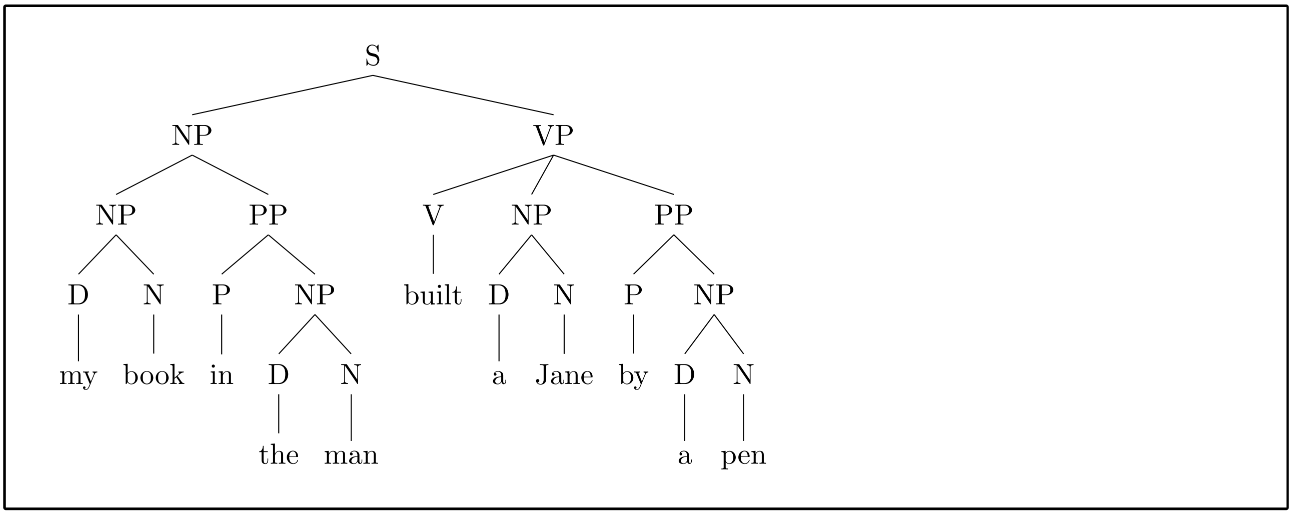
a) Jane constructed automata with a pen (4/20 pts)



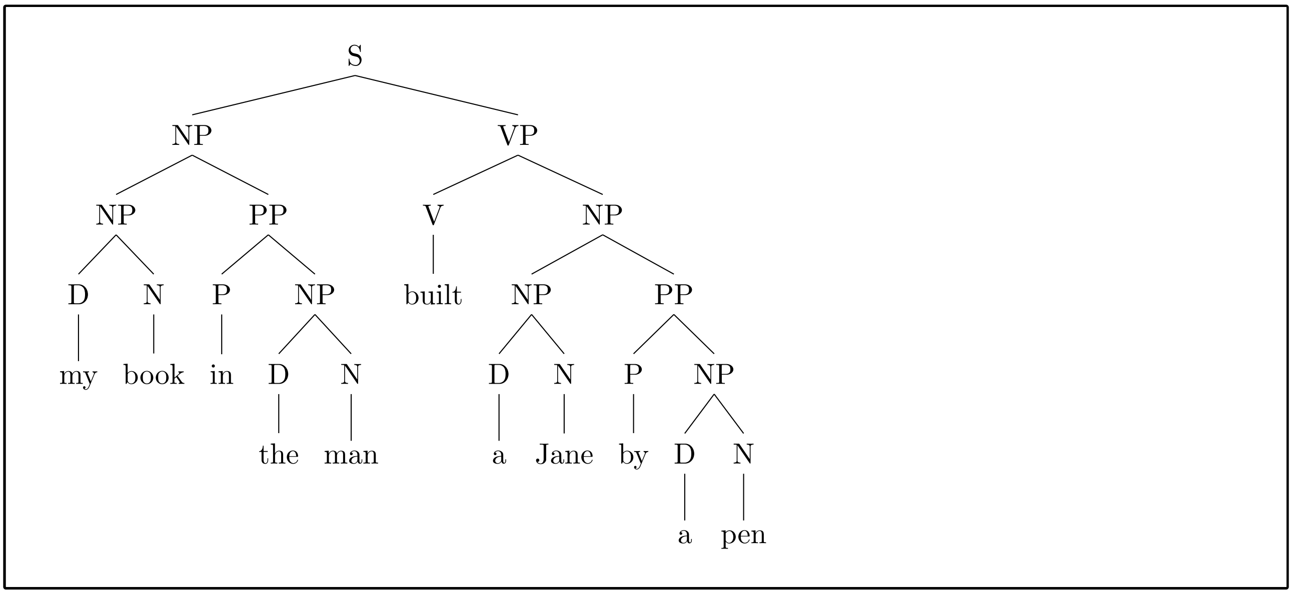
or



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



or

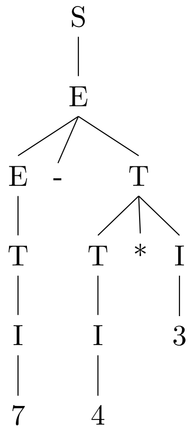


Given the CFG below, answer **c**, **d** and **e**

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

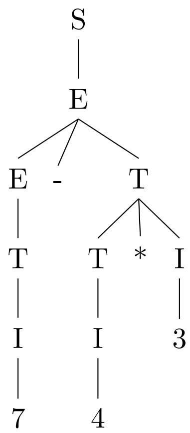
c) Provide the left-most derivation of 7 - 4 * 3 step-by-step and plot the final parse tree matching that derivation (4/20 pts)

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



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)

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

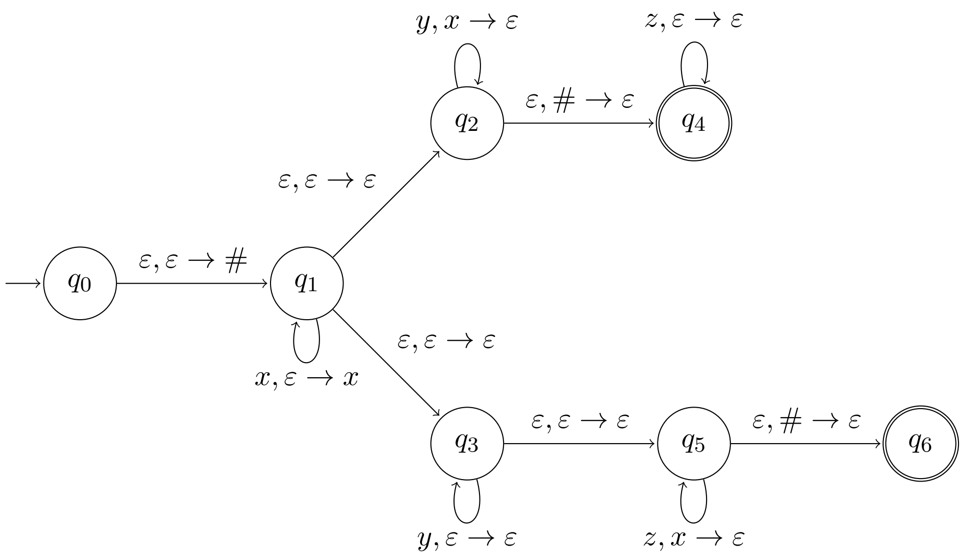


e) Are the derivations in **c** and **d** in the same similarity class? (4/20 pts)

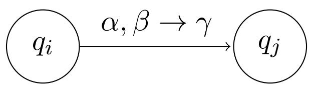
Yes. They both, have the same parse tree, so they in the same similarity class.

3 Pushdown Automata (30 pts)

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

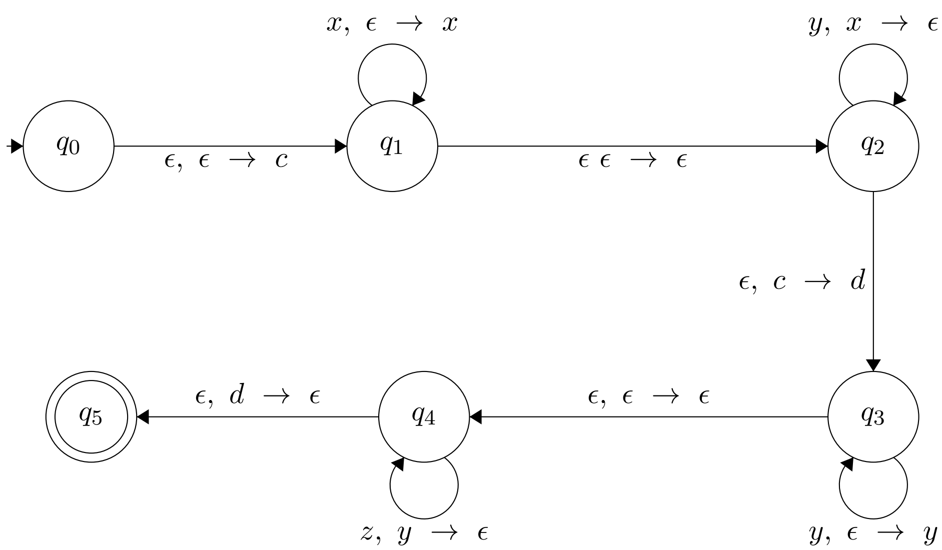


where the transition $((q_i, \alpha, \beta), (q_j, \gamma))$ is represented as:

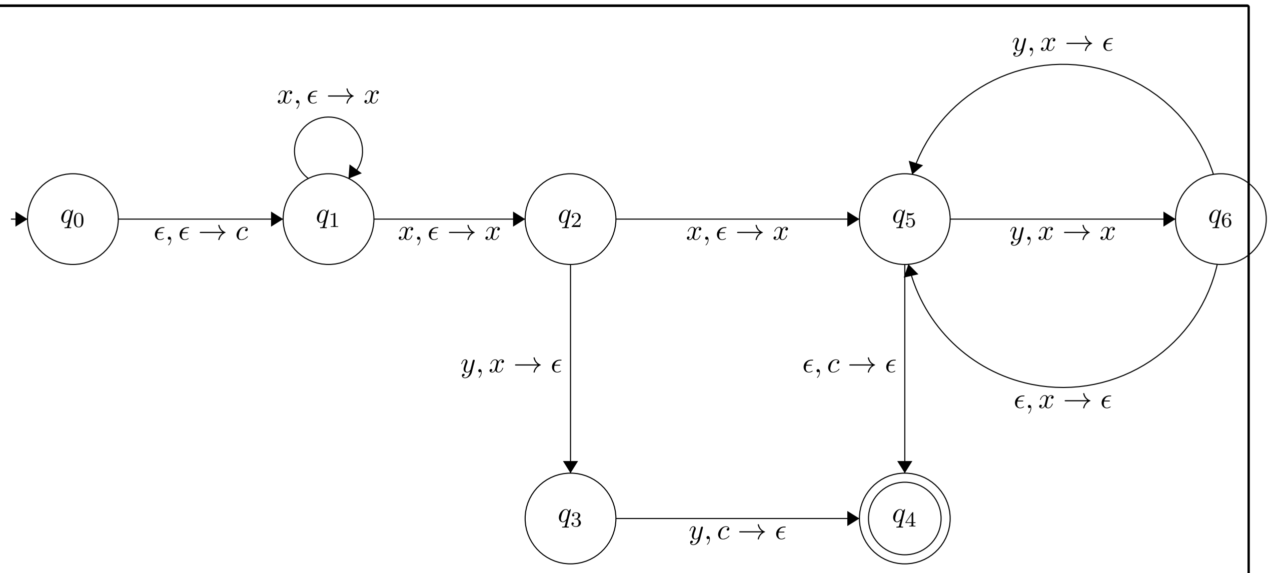


$L = \{x^n y^n z^m \text{ and } x^n y^m z^n | n, m \geq 0 \text{ and } n, m \in N\}$

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)



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.



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 $L_2 = \{w||w| = 4 * n + 2$ for $n \in N\}$
It is given that L is a CFL.
L_2 is a regular by the definition(it can be drawn that an NFA accepting L_2)
$L' = L \cap L_2$
By the theorem 3.5.2 Intersection of a CFL with a regular language is also CFL.
So, L' is also CFL.

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)

$L_2 \setminus L_3 = L_2 \cap \bar{L}_3$
since regular languages close under complementation, $\bar{L}_3$ is also a regular language.
Intersection of CFL's with a regular language is also a CFL so $L_2 \cap \bar{L}_3$ is a CFL.(Theorem 3.5.2 in textbook)
Since CFL's are not close under intersection (Theorem 3.5.4), $L_1 \cap (L_2 \cap \bar{L}_3)$ is not necessarily CFL.

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

Intersection of CFL's with a regular language is also a CFL so $L_1 \cap L_3$ is a CFL(Theorem 3.5.2 in textbook)
Since CFL's are closed under Kleene Star (Closure property) L_5 is necessarily 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 using Pumping Theorem for CFLs. (10/20 pts)

$w = a^p m^p t^{2p}$ and $w \in L$
$w = uvxyz$ such that
$|vxy| \leq p$
$|vy| > 0$
$uv^n xy^n z \in L$

case 1:

uvxy = a^{p-1}
z = $am^p t^{2p}$
if we take n = 0:
$uv^0 xy^0 z = a^{p-|vy|} m^p t^{2p} \notin L$

case 2:

uv = a^{p-1}
x = am
y = m^{p-2}
z = mt^{2p}
if we take n = 0:
$uv^0 xy^0 z = a^{p-|v|} m^{p-|y|} t^{2p} \notin L$

case 3:

u = $a^p m$
vxy = m^{p-2}
z = mt^{2p}
if we take n = 0:
$uv^0 xy^0 z = a^p m^{p-|vy|} t^{2p} \notin L$

case 4:

u = $a^p m$
v = m^{p-2}
x = mt
yz = t^{2p-1}
if we take n = 0:
$uv^0 xy^0 z = a^p m^{p-|v|} t^{2p-|y|} \notin L$

case 5:
u = $a^p m^p t$
vxyz = t^{2p-1}
if we take n = 0:

$uv^0xy^0z = a^pm^pt^{2p-|vy|} \notin L$

So, By Pumping Theorem, Language is not CFL.

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

$w = a^pb^{2p}a^p$ and $w \in L$
$w = uvxyz$ such that
$|vxy| \leq p$
$|vy| > 0$
$uv^nxy^nz \in L$

case 1:

uvxy = a^{p-1}
z = $ab^{2p}a^p$
if we take n = 0:
$uv^0xy^0z = a^{p-|vy|}b^{2p}a^p \notin L$

case 2:

uv = a^{p-1}
x = ab
y = b^{2p-2}
z = ba^p
if we take n = 0:
$uv^0xy^0z = a^{p-|v|}b^{2p-|y|}a^p \notin L$

case 3:

u = a^pb
vxy = b^{2p-2}
z = ba^p
if we take n = 0:
$uv^0xy^0z = a^pb^{2p-|vy|}a^p \notin L$

case 4:

u = a^pb
v = b^{2p-2}
x = ba
yz = a^{p-1}
if we take n = 0:

$uv^0xy^0z = a^pb^{2p-|v|}a^{p-|y|} \notin L$

case 5:
u = $a^pb^{2p}a$
vxyz = a^{p-1}
if we take n = 0:
$uv^0xy^0z = a^pb^{2p}a^{p-|vy|} \notin L$

So, By Pumping Theorem, Language is not CFL.

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) $a^*bc \cup a^nb^nc$

answer here ...

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

answer here ...