Middle East Technical University | Department of Computer Engineering

Formal Languages and Abstract Machines
Take Home Exam 2

Ahmet Dara VEFA
2237899

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)

$S- > aKaZ \mid bKbZ$
$Z- > a \mid b$
$K- > aK \mid bK \mid e$

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

$S- > aC \mid bC$
$C- > aaC \mid abC \mid bbC \mid baC \mid e$

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

$S- > e \mid SaSaSbS \mid SaSbSaS \mid SbSaSaS$

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

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

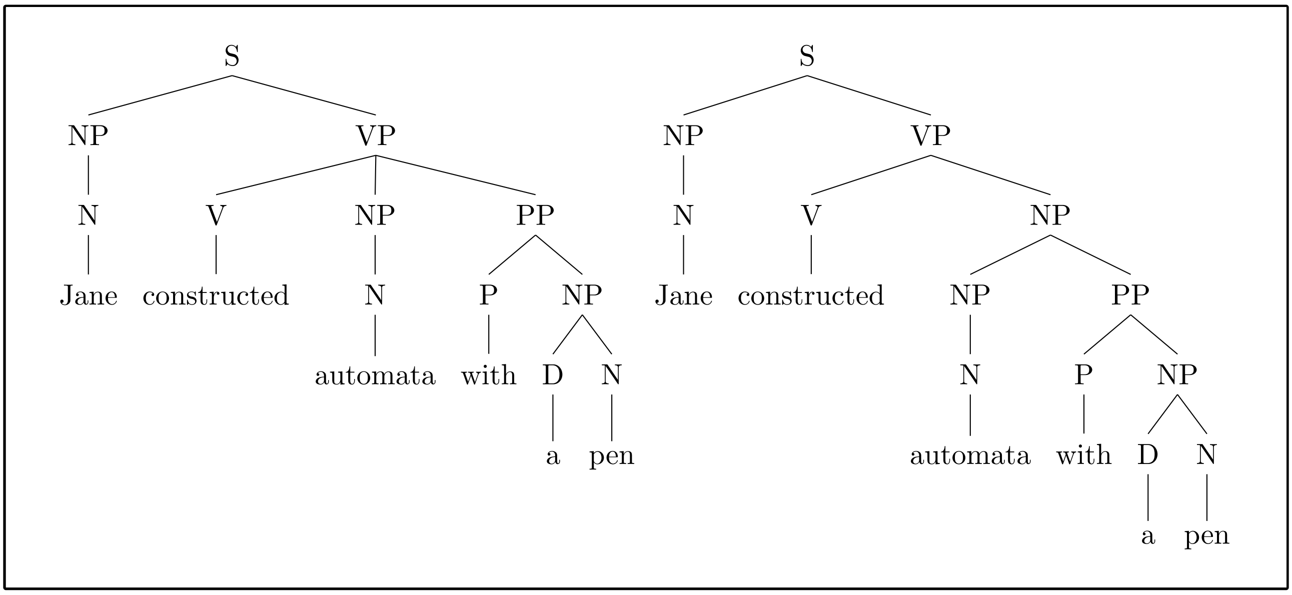
$$\boxed{L(G) = \{w | (w \in (a^k b^l,\ k \neq l)) \cup (w \in \Sigma^* ba \Sigma^*)\}}$$

2 Parse Trees and Derivations (20 pts)

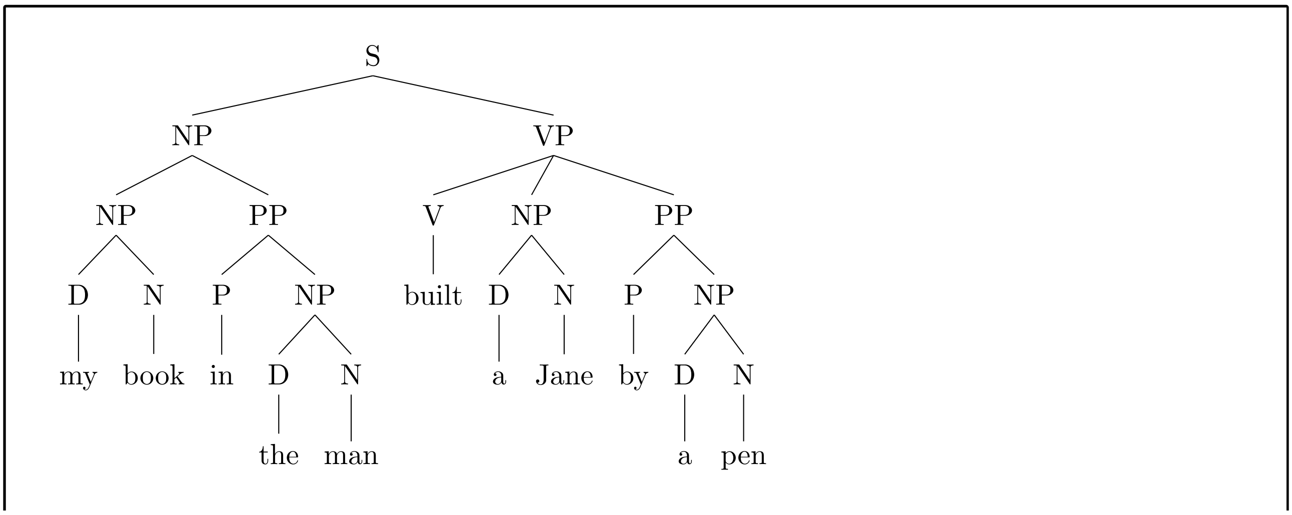
Given the CFG below, provide parse trees for given sentences in **a** and **b**.

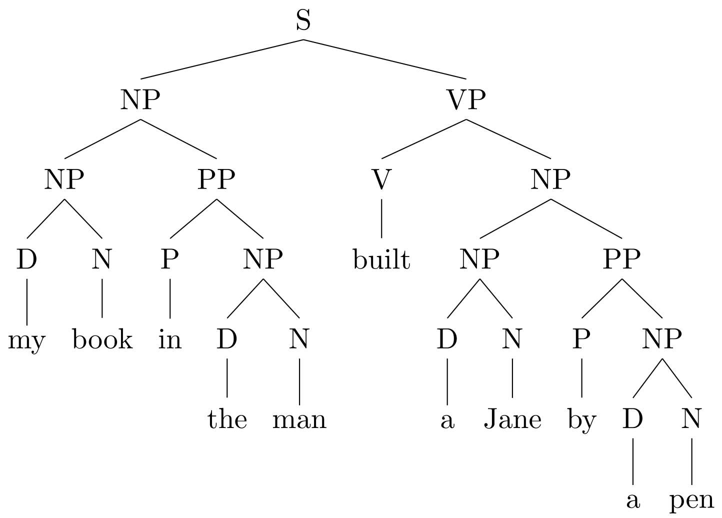
```
S   → NP VP
VP  → V NP | V NP PP
PP  → P NP
NP  → N | D N | NP PP
V   → wrote | built | constructed
D   → a | an | the | my
N   → John | Mary | Jane | man | book | automata | pen | class
P   → in | on | by | with
```

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)



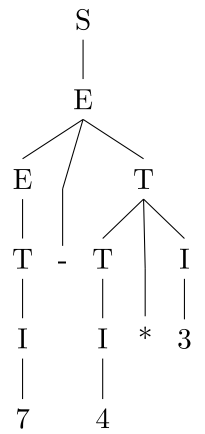

S
NP VP
NP PP V NP
D N P NP built NP PP
my book in D N D N P NP
the man a Jane by D N
a pen

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
```

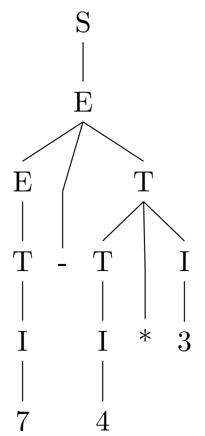
c) Provide the left-most derivation of 7 - 4 * 3 step-by-step and plot the final parse tree matching that derivation (4/20 pts)

$S => E => E{-}T => T{-}T => I{-}T => 7{-}T => 7{-}T{*}I => 7{-}I{*}I => 7{-}4{*}I => 7{-}4{*}3$



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)

$S => E => E - T => E - T * I => E - T * 3 => E - I * 3 => E - 4 * 3 => T - 4 * 3 => I - 4 * 3 => 7 - 4 * 3$

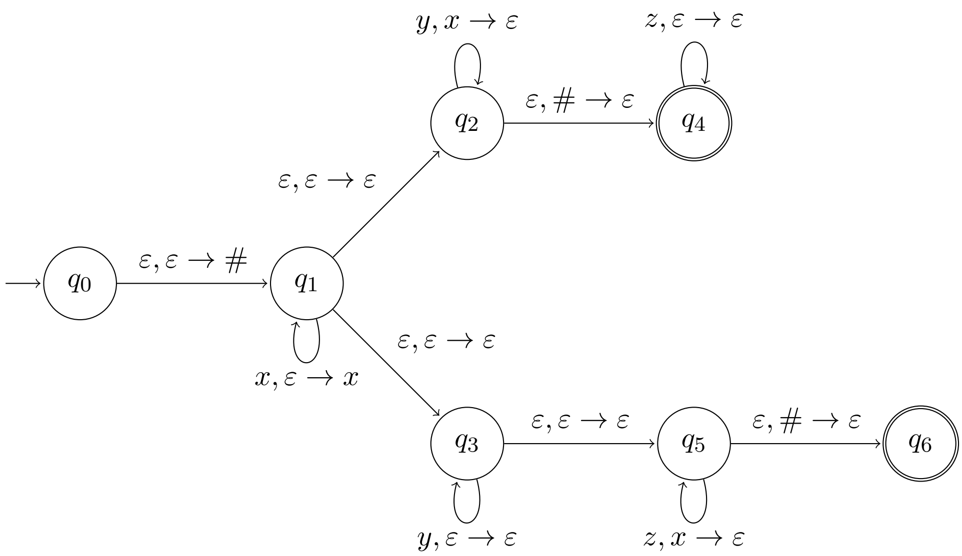


e) Are the derivations in **c** and **d** in the same similarity class? (4/20 pts)

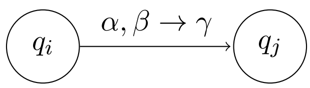
Yes, since they both give the same parse tree they are in the same similarity class.

3 Pushdown Automata (30 pts)

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

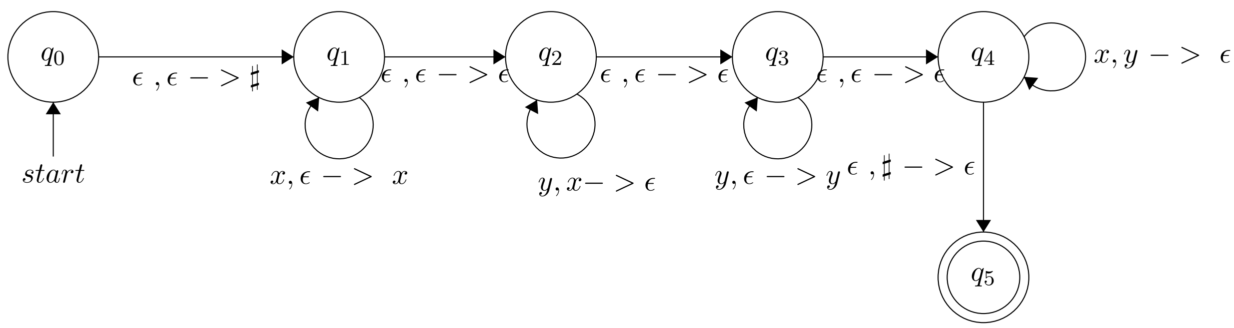


where the transition $((q_i, \alpha, \beta), (q_j, \gamma))$ is represented as:

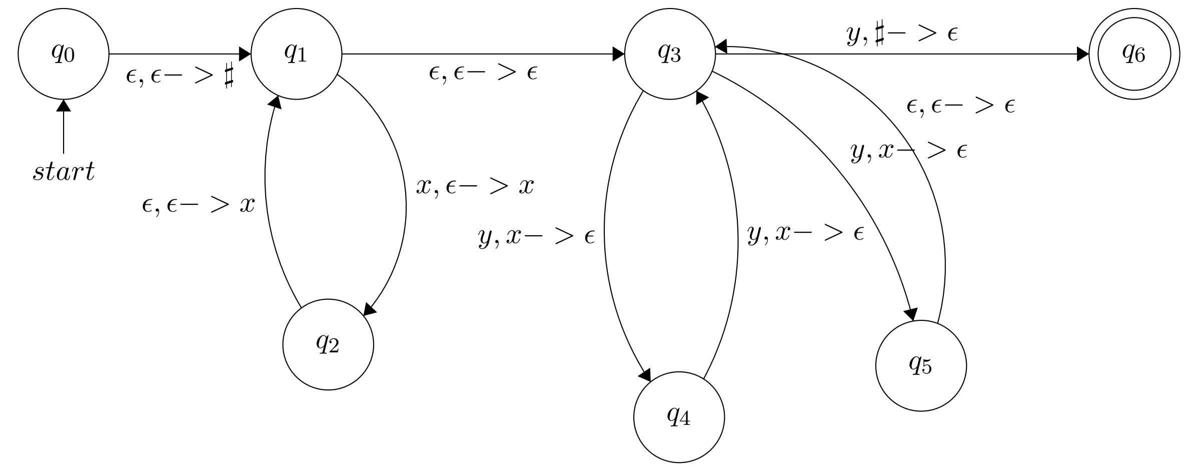


$L = \{w | w \in (x^n y^m z^n \cup x^m y^m z^n), (n, m \in N)\}$

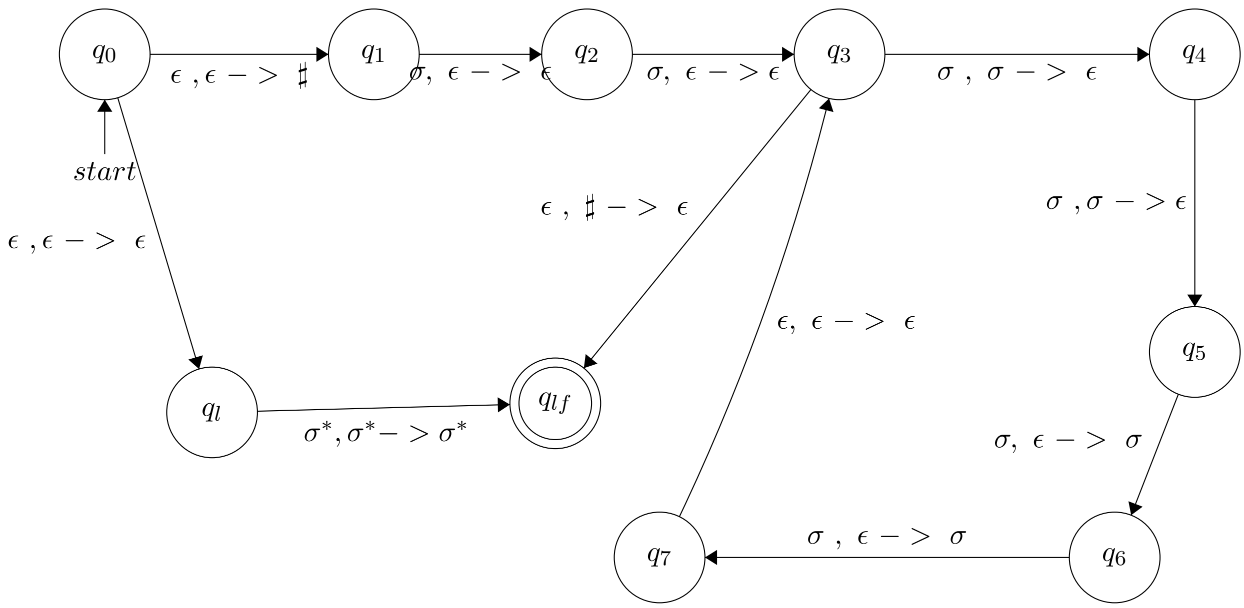
b) Design a PDA to recognize language $L = \{x^n y^{m+n} x^m \mid n, m \geq 0;\ n, m \in \mathbb{N}\}$ (5/30 pts)



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.



Where σ is the alphabet of L, q_l is the starting node of L, q_{lf} is the finish node of L

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)

Let$L_6 = L_2 - L_3 = L_2 \cap \bar{L}_3$
intersection of a non regular CF language and the complement of a regular language is CF language.
$L_1 \cap L_6$ is the intersection of 2 CF languages. Since CF languages are not closed under intersection L_4 can be either non-CF:
let $L_1 = \{a^n b^n c^m : m, n \in N\}$
let $L_2 \cap \bar{L}_3 = \{a^m b^n c^n : m, n \in N\}$ L_4 is not-CF.
or it can be CF:
let $L_1 = \{a^n b^n : n \in N\}$
let $L_2 = \{a^n b^n : n \in N\}$
let $L_3 = \{\}$
so $L_2 \cap \bar{L}_3 = L_2 = L_1$, therefore $L_5 = L_1 \cap L_2 = \{a^n b^n : n \in N\}$.
So L_5 is CF.

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

$(L_1 \cap L_2)^*$ is always CF since $(L_1 \cap L_2)$ can be anything from $\{\}$ to a CF language, which makes it a CF language($RL \subseteq CFL$). CF languages are closed under kleene star(Theorem 3.5.1), so L_5 is context free.

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.

Assume L is a CFL. Then for some $w = a^k m^k t^{2k}$, where K is the pumping length. By the Pumping Theorem, there exists a split $w = uvxyz$ st. $\mid vxy \mid < K \ and \ \mid vy \mid \geq 1$.
There are 5 cases.
First one is where vxy is all taken from a^k .
Second one is where vxy is all taken from m^k .
Third one is where vxy is all taken from t^{2k} .
Fourth one is where vxy is taken from $a^k m^k$.
Fifth one is where vxy is taken from $m^k t^{2k}$.
For cases 1 and 2: vxy consists of a's or m's. If we pump vxy for i=2 we get $w' = a^l m^j t^{2k}$ where $l \neq j$. This is a contradiction with the $w = a^k m^k t^{2k}$.
For case 3: vxy consists of t's. If we pump vxy for i=2 we get $w' = a^k m^k t^j$ where j¿2k.This is a contradiction with the $w = a^k m^k t^{2k}$.
For case 4: vxy consists of a's and m's. If we pump vxy for i=0, we get $w' = a^j m^p t^{2k}$ where either $j \neq p$ or 2j¡2k.This is a contradiction with the $w = a^k m^k t^{2k}$.
For case 5: vxy consists of m's and t's.If we pump vxy for i=2 we get $w' = a^k m^j t^l$ where either $k \neq j$ or $l > 2k$.This is a contradiction with the $w = a^k m^k t^{2k}$.
So by pumping theorem this language is not CF.

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.

Assume L is a CFLç Then for some $w = a^k b^{2k} a^k$, where k is the pumping length. By the pumping theorem, there exists a split$w = uvxyz$ st. $\mid vxy \mid < K \ and \ \mid vy \mid \geq 1$.
There are 5 cases.
First one is where vxy is all taken from a^k .
Second one is where vxy is all taken from b^{2k} .
Third one is where vxy is all taken from a^k .
Fourth one is where vxy is taken from $a^k b^{2k}$.
Fifth one is where vxy is taken from $b^{2k} a^k$.
For cases 1,2 and 3: vxy consists of a's or b's. If we pump vxy for i=0 we get $w' = a^p b^r a^s$ where either $p \neq s$ or $r \neq 2p$ or $r \neq 2s$. This is a contradiction with the $w = a^k b^{2k} a^k$.
For cases 4 and 5: vxy consists of a's and b's. If we pump vsy for i=0, we get $w' = a^p b^r a^s$ where either $p \neq s$ or $r \neq 2p$ or $r \neq 2s$. This is a contradiction with the $w = a^k b^{2k} a^k$.
So by pumping theorem this language is not CF.

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 ...