Middle East Technical University — Department of Computer Engineering

CENG 280

Formal Languages and Abstract Machines

Spring 2018-2019

Take Home Exam 2

Due date: April 26, 2019, Friday, 23:55

Objectives

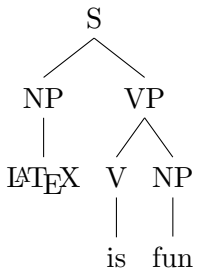
To familiarize with context-free languages, grammars for CFLs, parse trees and derivations, pushdown automata, CFL and PDA equivalence, closure properties of CFLs, pumping theorem for CFLs, Chomsky normal form, and Cocke-Younger-Kasami algorithm for parsing, Deterministic PDA.

Specifications

- You must adhere to the notation used in the textbook.
- Your solution should be delivered as a .tex file based on your modification of the provided template file. For convenience, a simple code for drawing a tree is included in the following. On the left-hand side you can see the code segment, and generated tree is placed on the right. You can also use the automata template given in THE2.

```
% preamble
\usepackage{tikz}
\usepackage{tikz-qtree}

% document
% use qtree
\Tree [.S [.NP $$\LaTeX$$ ] [.VP [.V is ] [.NP fun ] ] ]
% or tikz-qtree with possible tikz options|\\
\begin{tikzpicture}[scale=1] |\\
\Tree [.S [.NP $$\LaTeX$$ ] [.VP [.V is ] [.NP fun ] ] ]
\end{tikzpicture}
```



- The questions and submission regulations are included in subsequent sections. While designing your solutions to the tasks, explicitly state any assumptions you make and pay particular attention to the notation you use. Your proofs must be sound and complete. Grading will be heavily affected by the formalization of your solutions.

1 Context-Free Grammars (26 pts)

a. Formally construct CFGs that generate each of the following languages. (16 pts)

(i) $\{a^i\, ba^j\, ba^{i+j} \in \{a,b\}^* \mid i,j \in \mathbb{N}\}$

(ii) $\{w \in \{a,b\}^* \mid$ the first, last and middle character of w are the same, $|w| > 3$ and is odd$\}$.

(iii) $\{a^i\, b^j\, c^k \in \{a,b,c\}^* \mid i,j,k \geq 0$ and $i \neq j$ or $j \neq k\}$

(iv) $\{w_1 c w_2 c \ldots c w_k c c w_j^R \in \{a,b,c\}^* \mid k \geq 1,\ 1 \leq j \leq k,\ w_i \in \{a,b\}^+$ for $i = 1, \ldots, k\}$

b. Consider the CFG $G = (V, \Sigma, R, S)$, where (10 pts)

$$\begin{aligned} V &= \{a, b, S, A, B\}, \\ \Sigma &= \{a, b\}, \\ R &= \{S \to aB \mid bA, \\ &\qquad A \to a \mid aS \mid BAA, \\ &\qquad B \to b \mid bS \mid ABB\}. \end{aligned}$$

Prove that $L(G)$ is the set of all nonempty strings in $\{a,b\}$ that have equal numbers of occurrences of a and b.

2 Parse Trees and Derivations (10 pts)

Consider the CFG $G = (V, \Sigma, R, S)$ where

$$V = \{a, b, S, A, B\},\ \Sigma = \{a, b\},\ R = \{S \to Ab \mid aaB,\ A \to a \mid Aa,\ B \to b\}.$$

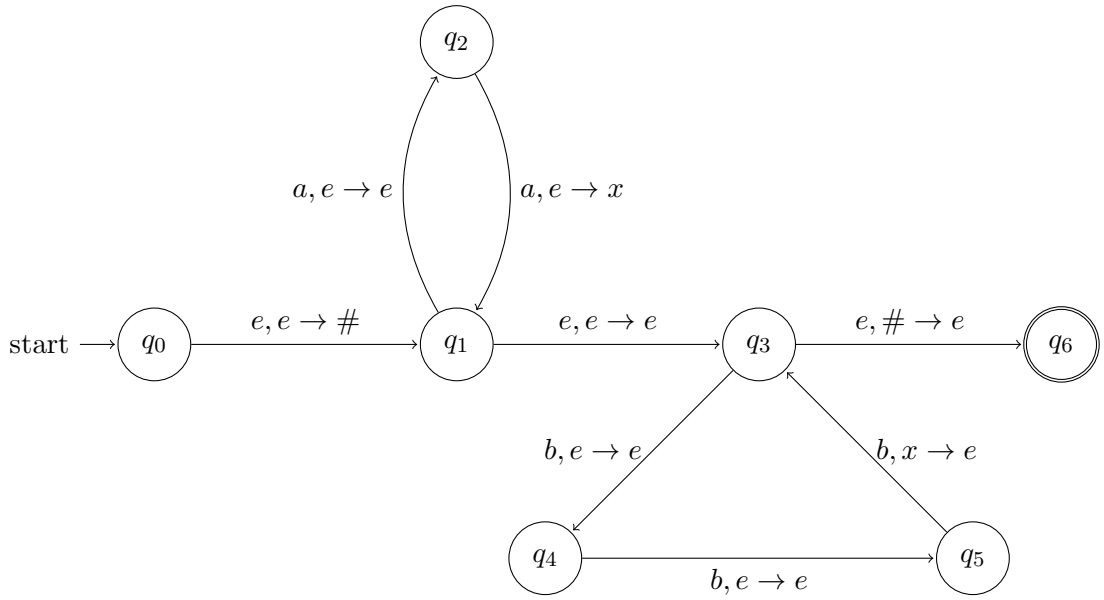
a. Find a string $s \in L(G)$ that has two leftmost derivations. Give these derivations and corresponding parse trees. (4 pts)

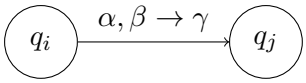
b. Find an equivalent unambiguous context-free grammar. (3 pts)

c. Give the unique leftmost derivation and parse tree for the string s from **a.** with respect to the grammar defined in **b.** (3 pts)

3 Pushdown Automata (23 pts)

a. Find the language generated by the PDA given below (5 pts)



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

b. Design a PDA that generates the complement of the language $\{ww \in \{a, b\}^* \mid w \in \{a, b\}^*\}$. (8 pts)

c.

(i) Design a PDA M that generates the language $\{a^i\, b^j\, c^k \in \{a, b, c\}^* \mid i, j, k \geq 0$, and $i = j$ or $j \neq k\}$. (5 pts)

(ii) Show that $aabcc \in L(M)$ and $bac \notin L(M)$ by tracing M on these strings. (5 pts)

4 PDA and CFGs (8 pts)

a. Consider the CFG $G = (V, \Sigma, R, E)$, where (4 pts)

$$
\begin{aligned}
V &= \{a, +, \times, (,), E, T, F\}, \\
\Sigma &= \{a, +, \times, (,)\}, \\
R &= \{E \to E + T \mid T, \\
&\qquad T \to T \times F \mid F, \\
&\qquad F \to (E) \mid a\}.
\end{aligned}
$$

Convert G to an equivalent PDA using Lemma 3.4.1.

b. Show that if $M = (K, \Sigma, \Gamma, \Delta, s, F)$ is a PDA, then there is another PDA $M' = (K', \Sigma, \Gamma, \Delta', s, F)$ such that $L(M') = L(M)$ and for all $((q_i, u, \beta), (q_j, \gamma)) \in \Delta'$, $|\beta| + |\gamma| \leq 1$. (4 pts)

5 Closure Properties and Pumping Theorem (29 pts)

a. Use closure properties for CFLs to prove that the following languages are context-free.

(i) $\{a^m\, b^{m+n}\, a^n \in \{a,b\}^* \,|\, m,n \in \mathbb{N}\}$ (5 pts)

(ii) $\{a,b\}^* - L$, where $L = \{babaabaaab \ldots ba^{n-1}ba^n \in \{a,b\}^* \,|\, n \geq 1\}$ (8 pts)

b. Use Pumping Theorem for CFLs to show that following languages are not context-free.

(i) $\{a^m b^n \in \{a,b\}^* \,|\, m,n \in \mathbb{N} \text{ and } m \leq n^2\}$ (8 pts)

(ii) $\{www \in \{a,b\}^* \,|\, w \in \{a,b\}^*\}$ (8 pts)

6 Closure Properties (4 pts)

(T/F) If L is a CFL, $\overline{L}$ is not a CFL since CFLs are not closed under complementation.

(T/F) There exists CFLs L_1 and L_2 such that $L_1 \cap L_2$ is a CFL.

(T/F) If L_1 is a CFL and L_2 is a regular language, $L_1 - L_2$ is a CFL.

(T/F) Every subset of a CFL is a CFL.

7 CNF and CYK (not graded)

a. Consider the CFG $G = (\{a,b,c,S,A,B,C\}, \{a,b,c\}, R, S)$, where

$$\begin{aligned} R = \{ & S \to aAB \,|\, bBA \\ & A \to BS \,|\, C \\ & B \to bA \\ & C \to c \,|\, e\}. \end{aligned}$$

Convert G into an equivalent CFG in Chomsky normal form.

b. Using CYK decide whether the following strings belong to $L(G)$.

(i) $w_1 = babcb$

(ii) $w_2 = acbbab$

8 Deterministic Pushdown Automata (not graded)

Construct a DPDA that generates the given languages.

a. $\{a^n b^m \in \{a,b\}^* \,|\, n,m \in \mathbb{N} \text{ and } m \geq n+2\}$

b. $\{w \in \{a,b\}^* \,|\, w$ starts and ends with the same symbol and have the same number of as and bs $\}$

c. $\{a^n b^m a^n \in \{a,b\}^* \,|\, m,n \in \mathbb{N}\}$

Submission

- You should submit your THE2 as a PDF file with the identifier `the2.pdf` on odtuclass. Please use the template provided on odtuclass with appropriate modifications.
- Soft-copies should be uploaded strictly by the deadline.
- **Late Submission:** You have two days in total for late submission with penalties of 20 points and 50 points reduction in your grade for the first and second day, respectively. No further submissions are accepted.

Regulations

1. **Cheating:** This take-home exam has to be completed and submitted **individually**. Teaming up, sharing solutions anywhere other parties might access and using work belonging to others as part or in whole are considered cheating. **We have zero tolerance policy for cheating**. People involved in cheating will be punished in accordance with the university regulations
2. **Newsgroup:** You must follow the newsgroup (cow.ceng.metu.edu.tr) for discussions and possible updates on a daily basis. You are advised to initiate discussions on COW so that most parties will benefit.