Problem Set 2: Context-Free Languages

Daniel Frishberg

Cal Poly CSC 445, Fall 2024

1 Instructions for Submission

The **regular deadline** for this problem set is **Thursday, November 7, at 11:59pm.** The **late deadline** for this problem set is **Friday, November 8, at 11:59pm.** Please upload a PDF or image of your work to Canvas. Legible handwriting or typing is fine. Please submit **just your answers, and not the problem set document.**

2 How to solve/use this problem set

Some questions are graded on effort. These require a credible, reasonable attempt for credit. Any such attempt will receive full credit. Other questions are graded on correctness. **Questions that are graded on correctness are indicated as such in bold.**

3 Required Problems

1. Give context-free grammars that generate the following languages:

   (a) (6 points) **(Graded on effort)**

   $$\{uu^R \mid u \in \Sigma^*\}$$

   where $\Sigma = \{0, 1\}$

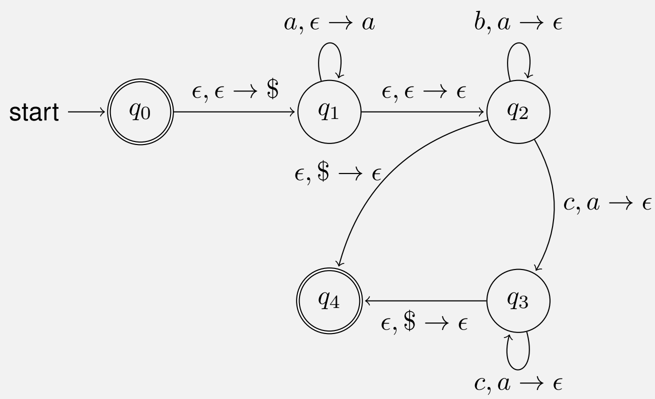
   $$S \to 0S0 \mid 1S1 \mid \varepsilon$$

   (b) (12 points) **(Graded on correctness)**

   $$\{a^i b^j c^k \mid i = j + k \text{ and } i, j, k \geq 0\}$$

$$
\begin{aligned}
S &\to aSc \mid T \mid \varepsilon \\
T &\to aTb \mid \varepsilon
\end{aligned}
$$

2. (16 points) **(Graded on correctness)** Construct a PDA that recognizes the language in Question 1(b).



3. (12 points) **(Graded on effort)** Convert the following grammar to Chomsky normal form. You do not have to follow the specific steps of the conversion algorithm; any valid CNF grammar that is equivalent to the original grammar is sufficient.

$$
\begin{aligned}
S &\to SUU \\
U &\to TTU \mid \varepsilon \\
T &\to aa \mid b
\end{aligned}
$$

$$
\begin{aligned}
S_0 &\to SX \mid SU \\
S &\to SX \mid SU \\
X &\to UU \\
Y &\to TT \\
U &\to YU \mid TT \\
T &\to AA \mid b \\
A &\to a
\end{aligned}
$$

4. Consider the following context-free grammar G in Chomsky normal form (CNF):

$$\begin{aligned}
S &\to AU \mid \varepsilon \\
T &\to AU \\
U &\to TB|b \\
A &\to a \\
B &\to b
\end{aligned}$$

(a) (6 points) **(Graded on correctness)** Write, in set-builder notation, the language this CFG generates.

$$L = \{a^n b^n \mid n \geq 0\}$$

(b) (14 points) **(Graded on correctness)** Use the CYK algorithm to test whether G generates the string $w = aabb$. (Just show the final table and say whether G generates w or not.)

$\{S, T\}$			
$\emptyset$	$\{U\}$		
$\emptyset$	$\{S, T\}$	$\emptyset$	
$\{A\}$	$\{A\}$	$\{U, B\}$	$\{U, B\}$
$\boldsymbol{a}$	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{b}$

G generates w.

5. (20 points) **(Graded on correctness)** Prove using the Pumping Lemma for context-free languages that the language

$$L_1 = \{\#u\#u \mid u \in \{0,1\}^*\}$$

is not context free where $\Sigma = \{0, 1, \#\}$.

This illustrates that the basic task of checking whether two strings are equal cannot be straightforwardly accomplished with a CFG or a PDA.

Suppose for a contradition that L is context free. Then there exists a pumping length p. Let $w = \#0^p1^p\#0^p1^p$. Then $w \in L$ and $|w| \geq p$. Therefore there exists u, v, x, y, z such that $w = uvxyz$, and:

(a) For all $i \geq 0$, $uv^ixy^iz \in L$

(b) $|vy| > 0$

(c) $|vxy| \leq p$

Let $i = 2$. Then $uv^2xy^2z = uvvxyyz$ and is in L by (a). Since $|vxy| \leq p$, vxy must be contained in the first half of w or the second half of w. After pumping, the string is now of the form $\#u^{'}\#u$ where $u^{'}$ is not equal to u, where the two halves of the string are not the same. Therefore $uv^2xy^2z \notin L$. This is a contradiction.

6. (14 points) **(Graded on correctness)** Closure (Sipser Ex. 3.12 p. 175) Consider the language

$$L = \{\#x_1\#x_2\# \cdots \#x_l \,|\forall i, x_i \in \{0,1\}^*, \\ \text{and } \exists i \neq j, x_i = x_j\}$$

where $\Sigma = \{0, 1, \#\}$

Intuitively, $w \in L$ consists of l different strings, where at least two of the strings are identical.

(By the way, this is the **complement** of the *element distinctness problem* in Sipser Ex. 3.12, p. 175.)

Prove that L is not context free using the closure properties of context-free languages and the fact that L_1 from the previous question is not context free.

Suppose for a contradiction that L is context free.
Let L_2 be the language $\#\{0,1\}^*\#\{0,1\}^*$ and is regular. Then by closure properties of context free languages, $L \cap L_2$ is context free. However, $L \cap L_2 = L_1$, which we know is not context free. This is a contradiction.