Homework 6

Name: Yiqiao Jin
UID: 305107551

1

Below is a CFG for this language

$S \to A$

$A \to aAb|Ab|\varepsilon$

S is the start variable.

A is the variable for generating a at the beginning and b at the end of the new string. It either generates a pair of a and b, or only b. This ensures that $i \leq j$, the number of a's generated is always less than or equal to that of b's. When A finishes its generation, it generates ε.

An equivalent unambiguous language is:

$S \to AB$

$A \to aAb|\varepsilon$

$B \to Bb|\varepsilon$

This language also ensures that $i \leq j$. This language is unambiguous because every a must be paired with a b when A is generating. For example, for abb, the prefix ab is generated by A, and suffix b is generated by B.

Example strings that can be accepted by L: $abb, ab, b, \varepsilon, aaabbb$

2

Note: I use w instead of x in the original question so that it won't be confused with the x in the Pumping Lemma.

We assume that L_2 is a **CFL** and obtain a contradiction. Let p be the pumping length given by the pumping lemma. We can use the string $w = 0^p 1^p$ and let

$$s = ww^R w = \underbrace{0^p 1^p}_{w} \underbrace{1^p 0^p}_{w^R} \underbrace{0^p 1^p}_{w} = 0^p 1^{2p} 0^{2p} 1^p \in L$$

According to the pumping lemma for **CFL**, since $|s| \geq p$, s may be divided into five pieces $s = uvxyz$ satisfying the conditions

1. for each $i \geq 0$, $uv^i xy^i z \in L_2$
2. $|vy| \geq 1$
3. $|vxy| \leq p$

We consider the following 4 cases:

1) The substrings v and y contain some symbols from the 0^p in the first half of s

Since $|vxy| \leq p$, v and y CANNOT contain 0's from the 0^{2p} in the second half of s

Consider $s' = uxz = 0^i 1^j 0^{2p} 1^p$, where $i < p, j \leq 2p$. Since $|vxy| \leq p$, pumping down on s will still yield $|s'| \geq 5p$.

Since $s' = uxz \in L$, it must be of the form $tt^R t$. The first t must begin with the block of $i < p$ 0's followed by some number of 1's. Thus, $t^R t$ must contain a block of at most $2i < 2p$ 0's.

However, uxz contains 0^{2p} (a continuous block of 0's with length $2p$), a contradiction.

2) The substrings v and y contain some symbols from the 1^{2p} in the first half of s

Since $|vxy| \leq p$, v and y CANNOT contain 1's from the 1^p in the second half of s

Consider $s' = uxz = 0^i 1^j 0^k 1^p$, where $i \leq p, j < 2p, k \leq 2p$. Since $s' = uxz \in L$, it must be of the form $tt^R t$.

Since $|vxy| \leq p$, pumping down on s will still yield $|s'| \geq 5p$. The last t must end with the block of 1^p preceded by some number of 0's. $t = 0..0\underbrace{1^p}_{p}$

So $tt^R = 0..0\underbrace{1^p 1^p}_{2p}0..0$

Thus, tt^R must contain a continuous block of $2p$ 1's, whereas uxz contains a block of only $j < 2p$ 1's, a contradiction.

3) The substrings v and y contain some symbols from 0^{2p} in the second half of s

Since $|vxy| \leq p$, v and y CANNOT contain 0's from the 0^p in the first half of s

Consider $s' = uxz = 0^p 1^i 0^j 1^k$, where $i \leq 2p, j < 2p, k \leq p$. Assume $s' \in L$, s' must be of the form $tt^R t$. Since $|vxy| \leq p$, pumping down on s will still yield $|s'| \geq 5p$. This means $t^R t$ must contain a continuous block of 0^{2p}.

However, $s' = uxz$ contains a block of only $j < 2p$ 0's, a contradiction.

4) The substrings v and y contain some symbols from 1^p in the second half of s

Since $|vxy| \leq p$, v and y CANNOT contain 1's from the 1^{2p} in the first half of s

Consider $s' = uxz = 0^p 1^{2p} 0^i 1^j$, where $i \leq 2p, j < p$. Assume $s' \in L$, s' must be of the form $tt^R t$. Since $|vxy| \leq p$, pumping down on s will yield $|s'| \geq 5p$. The second t must end with the block of $1^j, j < p$ preceded by some number of 0's:

Thus, tt^R must contain a continuous block of at most $2j < 2p$ 1's, whereas uxz actually contains a block of $2p$ 1's, a contradiction.

From 1-4, the Pumping Lemma does NOT apply under all these cases. So we prove that this language L_2 is NOT a **CFL**.

3

1) $R \cap L$

This can be (1) (2) (3) (4). R is an FSL means R is also a CFL. Since L is non-CFG, no closure can be applied.

For example, let $\Sigma = \{0, 1\}$, $R = (00)^*$ and $L = \{ww | w \in \{0, 1\}^*\}$. $R \cap L = R$, a regular language (1)

Let $\Sigma = \{a, b, c\}$, $R = \{b, c\}^*$ and $L = \{a^i b^j c^k | 0 \leq i \leq j \leq k\}$. $R \cap L = \{b^j c^k | 0 \leq j \leq k\}$, which is (2) CFL but not FSL . If $R = \{a, b, c\}^*$, then $R \cap L = L\}$. This is (4) non-CFL.

2) A

A can be (1) FSL, OR (2) CFL and NOT FSL, OR (3) Inherently Ambiguous CFL.

A can be (1) since finite state language is a subset of CFL.

For example, the language $S \to Sa | aS | a | \varepsilon$ is an ambiguous CFL. But it represents a^*

A can be (2). If a language is not inherently ambiguous, we can eliminate the ambiguity in the language by substituting rules.

For example, the language $S \rightarrow 0S1|00S11|\varepsilon$ is an ambiguous CFL. It simply represents 0^n1^n , a CFL. We can eliminate the ambiguity in this case.

A can also be (3) if it is inherently ambiguous.

3) $\bar{C}$, the complement of C with respect to Σ

$\bar{C}$ can be (1) (2) (3) (4). A language that can be represented by a PDA is a CFL. But CFL is NOT closed under complement.

4) G

G can be (1) FSL, or (2) CFL.

The class of FSL is a subclass of CFG and every regular language is naturally context-free. Since G can be represented by an unambiguous CFG, it can either be a FSL (CFG can simulate any FSL), or it can be a CFG that is NOT an FSL.

5) $\bar{L}$, the complement of L with respect to Σ

$\bar{L}$ can be (1) (2) (3) (4). Since L is non-CFG, no closure can be applied to determine the class of $\bar{L}$.

6) X

X must be (4). FSL and CFL are closed under union. If X were CFL, then $L = X \cup S$ can fall into (1) (2) (3). But since L is non-CFL, X cannot be in (1) (2) (3).