CSCI 338: Assignment 3

Nathan Stouffer

Problem 1

Design context-free grammars for the following languages

1.1 $A = \{a^n b^m | n \neq 2m\}$.

1.2 $B = \{a^i b^j c^k | i, j, k \geq 0$ and either $i = j$ or $j = k\}$.

1.3 $C = \{a^n b^m | n = 3m\}$.

1.4 $D = \{a^n b^m | n \leq m + 3\}$.

Proof: We now give the context grammars for each of the above languages.

1.1
$$S \longrightarrow aaSb \mid aB \mid A \mid B$$
$$A \longrightarrow Aa \mid a$$
$$B \longrightarrow Bb \mid b$$

1.2
$$S \longrightarrow IC \mid AK$$
$$I \longrightarrow aIb \mid \epsilon$$
$$K \longrightarrow bKc \mid \epsilon$$
$$C \longrightarrow Cc \mid \epsilon$$
$$A \longrightarrow Aa \mid \epsilon$$

1.3
$$S \longrightarrow \epsilon \mid aaaSb$$

1.4
$$S \longrightarrow X \mid aX \mid aaX \mid aaaX$$
$$X \longrightarrow aXb \mid B$$
$$B \longrightarrow Bb \mid \epsilon$$

$\square$

Problem 2

Decide whether the following grammar is ambiguous.

$$S \to AB|aaB$$
$$A \to a|Aa$$
$$B \to b$$

Proof: To show that the above grammar is ambiguous, we must show that there exist two leftmost derivations for some string in the grammar. Consider the following two derivations for aab.

Derivation 1: $S \overset{*}{\Longrightarrow} AB \overset{*}{\Longrightarrow} AaB \overset{*}{\Longrightarrow} aaB \overset{*}{\Longrightarrow} aab$
Derivation 2: $S \overset{*}{\Longrightarrow} aaB \overset{*}{\Longrightarrow} aab$

The above derivations are both leftmost and distinct, yet they derive the same string. So the given grammar must be ambiguous.

$\square$

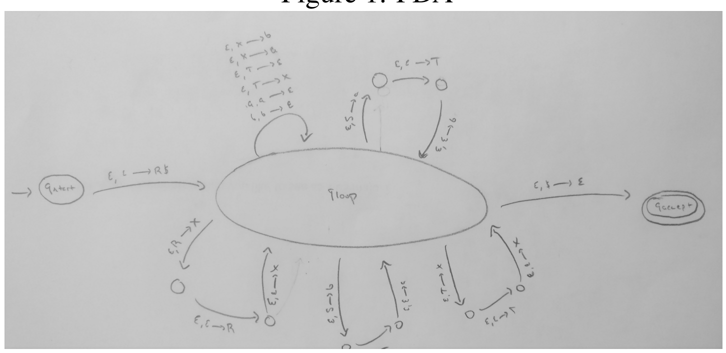
Problem 3

Convert the following CFG G to an equivalent PDA.

$R \to XRX|S$
$S \to aTb|bTa$
$T \to XTX|X|\epsilon$
$X \to a|b$

Proof: We show the equivalent PDA below.

Figure 1: PDA



∎

Problem 4

Let $G = (V, \Sigma, R, S)$ be the following grammar. $V = \{S, T, U\}$; $\Sigma = \{0, \#\}$; and R is the set of rules:

$S \to TT|U$
$T \to 0T|T0|\#$
$U \to 0U00|\#$

4.1 Describe $L(G)$ in English.

Proof: In English, $L(G)$ consists of the union of two sets. The first is the set of all strings that include two $\#$ symbols with any number of zeros on either side and in between the $\#$ symbols. The second is the set consisting of all strings beginning with n zeros, then a $\#$ symbol, and then $2n$ zeros.

□

4.2 Prove that $L(G)$ is not regular.

Proof: We will prove that $L(G)$ is not regular by contradiction. That is, we will assume that $L(G)$ is regular. Since $L(G)$ is regular, we can apply the pumping lemma for regular languages. Take p to be the pumping length of $L(G)$. Then, for any string $s \in L(G)$ where $|s| \geq p$ (we know s exists since p is finite and we construct an element of $L(G)$ to be as large as needed), the following must be true:

1. $xy^iz \in L(G)\ \forall i \geq 0$
2. $|y| > 0$
3. $|xy| \leq p$

We now choose $s = 0^p\#0^{2p} \in L(G)$. Since $L(G)$ is regular, we can decompose $s = xyz$. Because of condition 3, it must be the case that xy consists of no more than the first p letters of s, meaning that $y = 0^a$ where $1 \leq a \leq p$. We then choose $i = 0$ and wonder if it is true that $xz = 0^{p-a}\#0^{2p} \in L(G)$.

This rests on $2(p - a) = 2p \iff 2p - 2a = 2p \iff -2a = 0 \iff a = 0$, which is certainly not true since a cannot be 0. So $xz \notin L(G)$ and s cannot be pumped for $i = 0$.

Since s cannot by pumped for $i = 0$, the pumping lemma does not apply to $L(G)$. Since the pumping lemma does not apply to $L(G)$, it must be the case that $L(G)$ is non-regular.

□

Problem 5

Convert the following CFG into an equivalent CFG in Chomsky Normal Form

$A \rightarrow BAB|B|\epsilon$
$B \rightarrow 00|\epsilon$

Proof: We begin by adding a start rule $S \longrightarrow A$. This gives

$$S \longrightarrow A$$
$$A \longrightarrow BAB \mid B \mid \epsilon$$
$$B \longrightarrow 00 \mid \epsilon$$

We now remove the rule $B \longrightarrow \epsilon$

$$S \longrightarrow A$$
$$A \longrightarrow BAB \mid BA \mid AB \mid A \mid B \mid \epsilon$$
$$B \longrightarrow 00$$

And now $A \longrightarrow \epsilon$

$$S \longrightarrow A \mid \epsilon$$
$$A \longrightarrow BAB \mid BB \mid BA \mid AB \mid A \mid B$$
$$B \longrightarrow 00$$

The unit rule $A \longrightarrow A$ has no consequence so we remove it. In addition, we replace the rule $S \longrightarrow A$

$$S \longrightarrow BAB \mid BB \mid BA \mid AB \mid B \mid \epsilon$$
$$A \longrightarrow BAB \mid BB \mid BA \mid AB \mid B$$
$$B \longrightarrow 00$$

We now introduce the new rule $C \longrightarrow AB$

$$S \longrightarrow BC \mid BB \mid BA \mid AB \mid B \mid \epsilon$$
$$A \longrightarrow BC \mid BB \mid BA \mid AB \mid B$$
$$B \longrightarrow 00$$
$$C \longrightarrow AB$$

Finally, we create the new rule $D \longrightarrow 0$. This puts us in Chomsky Normal Form.

$$S \longrightarrow BC \mid BB \mid BA \mid AB \mid DD \mid \epsilon$$
$$A \longrightarrow BC \mid BB \mid BA \mid AB \mid DD$$
$$B \longrightarrow DD$$
$$C \longrightarrow AB$$
$$D \longrightarrow 0$$

□

Problem 6

Using pumping lemma to prove that the following languages are not context-free.

6.1 $L = \{a^n b^j c^k | k = nj\}$.

Proof: We will prove that L is not context free by contradiction. That is, we will assume that L is context free. Since L is context free, we can apply the pumping lemma. Take p to be the pumping length of L. Then, for any string $s \in L$ where $|s| \geq p$ (we know s exists since p is finite and we could construct an element of L to be as large as needed), the following must be true:

1. $uv^ixy^iz \in L \; \forall i \geq 0$
2. $|vy| > 0$
3. $|vxy| \leq p$

Choose $s = a^p b^p c^{p^2} \in L$. We can then decompose $s = uvxyz$. We now consider three cases for s.

The first case is when one of v, y contains more than one type of symbol. Pumping up results in a string not of the form $a^n b^j c^k$, since there will be some symbols in between (ie $aabaabbcccc$).

We now consider the case where $v = a^{m_1}$ and $y = b^{m_2}$. Pumping up when $i = 2$ results in a string $a^{p+m_1} b^{j+m_2} c^k$. So is $k = (p + m_1)(p + m_2) = p^2 + (m_1 + m_2)p + m_1 * m_2$? No, since $k = p^2 < p^2 + (m_1 + m_2)p + m_1 * m_2$ and one of m_1, m_2 is greater than 0.

Now we have the case $v = b^{m_1}$ and $y = c^{m_2}$. Pumping up for $i = 2$ results in the string $s = a^p b^{p+m_1} c^{p^2+m_2}$. So is $k = p^2 + m_2 = p(p + m_1) \iff p^2 + m_2 = p^2 + p * m_1 \iff m_2 = p * m_1$. So $s_i \in L$ only when $m_2 = p * m_1$. Can this ever be the case? Well, $m_1 = 0 \implies m_2 = 0$, but that cannot be the case since $|vy| > 0$. So it must be true that $p * m_1 \geq p \implies m_2 \geq p$, yet this also cannot be the case since $|vxy| \leq p$.

Thus a contradiction is found in all cases and L is not context free.

□

6.2 $\quad L = \{a^n b^j | n \geq (j-1)^3\}$.

Proof: We will prove that L is not context free by contradiction. That is, we will assume that L is context free. Since L is context free, we can apply the pumping lemma. Take p to be the pumping length of L. Then, for any string $s \in L$ where $|s| \geq p$ (we know s exists since p is finite and we could construct an element of L to be as large as needed), the following must be true:

1. $uv^i xy^i z \in L \; \forall i \geq 0$
2. $|vy| > 0$
3. $|vxy| \leq p$

Choose $s = a^{(p-1)^3} b^p \in L$. Then we can decompose $s = uvxyz$. Now consider the following four cases.

If one of v, y contains more than one type of symbol then pumping results in leaving the form $a^n b^j$ for any $n, j \in \mathbb{N}$.

Now consider the case where $vxy = a^m$ for $1 \leq m \leq p$. Then we can pump down for the case $i = 0$ so that $s = a^{(p-1)^3 - m} b^p$. Clearly $(p-1)^3 \neq (p-1)^3 - m$ since $m \neq 0$.

We now view the case where $vxy = b^m$ where $1 \leq m \leq p$. We again pump down ($i = 0$) and now $s = a^{(p-1)^3} b^{p-m}$. Yet it cannot be the case that $(p-1)^3 = (p-m-1)^3$ since $m \neq 0$.

We now consider the final case where $v = a^{m_1}$ and $y = b^{m_2}$ for both $m_1, m_2 \geq 1$. We then pump up for the case $i = 2$ and now $s = a^{(p-1)^3 + m_1} b^{p+m_2}$. This case rests on the fact that $(p-1)^3 + m_1 = (p + m_2 - 1)^3$. Yet, since $m_2 \geq 1$, it must be true that $(p + m_2 - 1)^3 \geq p^3$. But since $m_1 < p$, it must also be true that $(p-1)^3 + m_1 < (p-1)^3 + p < p^3$. So $(p-1)^3 + m_1 < (p + m_2 - 1)^3$.

So we have found a contradiction in all cases and L is not context free.

$\square$