COMP 335 – Introduction to Theoretical Computer Science

Solution to Assignment 3

1. Let L be any CFL. Then, we know that there is a context-free grammar $G = (V, \Sigma, S, P)$ that generates L. Using G, we obtain another CFG $G' = (V, \Sigma, S, P')$ such that if P includes the production $X \rightarrow$"whatever", then P' includes the corresponding production for variable X in which the right-hand side is the reverse of "whatever". It is easy to see that G' generates L^R, which means the language L^R is context-free. For this, let w be any string in L. Since $L(G) = L$, we know that there is a sequence of k derivations ($S \Rightarrow^{i_1} \Rightarrow \cdots \Rightarrow^{i_k} w$) in G that generates w. This implies that the string w^R could be dreived by the corresponding k sequence of derivations in G' (that is, $S \Rightarrow^{i_1} \Rightarrow \cdots \Rightarrow^{i_k} w^R$), which was to be shown.

2. This is yet another normal form for context-free gramamrs which can be obtained from Chomsky normal form. First note that productions of the form $A \rightarrow BC$ in Chomsky normal form (CNF) are also allowed in G. Also, productions of the form $A \rightarrow a$ in CNF can be replaced by the three productions: (1) $A \rightarrow aV_1V_2$, (2) $V_1 \rightarrow \lambda$, and (3) $V_2 \rightarrow \lambda$, in which V_1 and V_2 are new variables.

3. We will show that the languages (1), (2) and (5) are context-free, while languages defined in aprts (3) and (4) are not.

   (a).
   $$S \rightarrow AB$$
   $$B \rightarrow aCa \mid bCb$$
   $$C \rightarrow aCa \mid bCb \mid A$$
   $$A \rightarrow aa \mid ab \mid ba \mid bb$$

   (b).
   $$S \rightarrow aSa \mid bSb \mid a \mid b \mid \lambda$$

   (c). Suppose L_3 is CF. Then since it is infinite, we can apply the pumping lemma for CFL's. Let m be the integer in the P.L. Consider the string $w = a^{m+2}b^{m+1}c^m$ in L, whose size $|w| = 3m+3 \geq m$. We have the following possible scenarios consider for substrings u, v, x, y, z in the P.L.

```
                  <- m+2  -><- m+1 -><- m ->
              w= a........ab.......bc.....c
1                   v y
2                           v y
3                                   v y
4                   v       y
5                           v       y
6                       vy=a^k b^j
7                              vy=b^i c^p
```

Case 1. $v = a^{k_1}$ and $y = a^{k_2}$, for $1 \leq k_1 + k_2 \leq m$. We take $i = 0$. This gives $w_0 \notin L_3$, since the number of a's in w_0 is not more than number of b's: i.e., $n_a(w_0) = m+2-k_1-k_2 \leq m+1 = n_b(w_0)$.
Case 2. $v = b^{k_1}$ and $y = b^{k_2}$, for $1 \leq k_1 + k_2 \leq m$. We take $i = 0$. This gives $w_0 \notin L_3$, since $n_b(w_0) = m+1-k_1-k_2 \not> n_c(w_0) = m$.
Case 3. $v = c^{k_1}$ and $y = c^{k_2}$, for $1 \leq k_1 + k_2 \leq m$. We take $i = 2$. This yields $w_2 \notin L_3$, since $n_c(w_2) = m + k_1 + k_2 \geq n_c(w_2) = m+1$.
Case 4. $v = a^p$ and $y = b^q$, where $p, q \geq 1$. Taking $i = 0$, the number of a's in w_0 (also the number of b's) is not more than $n_c(w_0)$ (the same problem with the number of b's: $n_b(w_0) \leq n_c(w_0)$.
Case 5. Similar to case 4; We have that $v = b^p$ and $y = c^q$, for p, q 0. Taking $i = 2$, we can see that $n_a(w_2) \leq n_b(w_2)$.
Case 6: $v = a^p b^q$, for p, q 0. That is, v includes symbols a and b. In this case, taking $i = 2$, the string w_2 is not in L_3 since it does not have the proper pattern as v^2 includes both mixed a's and b's.
Case 7. $v = b^p c^q$, for $p, q > 0$. This case is similar to case 6 but b's and c's are mixed in w_2, and hence $w_2 \notin L_3$.
We considered all possible cased of decomposition of w and obtained a contradiction, and hence L_3 is not CF.

(d). Suppose L_4 is CFL, and since it is infinite, we can apply the PL for CFLs. Let us consider the string $w = a^m b^m a^m b^m$ in L_4 whose length $|w| = 4m \geq m$. Then, by P.L., there are substrings u, v, x, y, z in Σ^* where $w = uvxyz$, $|vxy| \leq m$, $|vy| \geq 1$, such that $w_i = uv^ixy^iz \in L_4, \forall i \geq 0$.

```
         w= <- m  -><- m -><- m  -><- m ->
               a......ab.....ba......ab.....b
1                      v y
2                            v y
3                                    v y
4                                            v y
5                     v        y
6                              v       y
7                                      v       y
8                       vy=a^k b^j
9                             vy=b^k a^j
10                                   vy=a^k b^p
```

Case 1. v and y consist of only a's. That is, $u = a^p$, $v = a^{k_1}$, $x = a^q$, $y = a^{k_2}$, $z = b^{m-p-k_1-q-k_2}b^m a^m b^m$. where $k_1 + k_2 \geq 1$ Consider w_i for $i = 2$: that is $w_2 = a^{m+k_1+k_2}b^m a^m b^m$. Since $k_1 + k_2 \geq 1$, then w_2 has more a's at the beginning than it has in the second part of a's. Thus $w_2 \notin L_4$.

The treatment of cases 2 to 4 is similar to case 1.

Case 5. $v = a^p$ and $y = b^q$, where p, q 0. Taking $i = 0$, we get $w_0 = a^{m-p}b^{m-q}a^m b^m$ which is not in L_4.

Case 6. $v = b^p$ and $y = a^q$, where p, q 0. Taking $i = 0$, we get $w_0 = a^m b^{m-p}a^{m-q}b^m$ which is not in L_4.

Case 7. $v = a^p$ and $y = a^q$, where p, q 0. Taking $i = 0$, we get $w_0 = a^m b^m a^{m-p}b^{m-q} \notin L_4$.

Case 8. $v = a^p b^q$; that is v includes both symbols a and b. Taking $i = 2$, we get w_2 which is not in L_4 since it does not have the proper pattern as a's and b's are mixed in the first part of the string w_2.

The treatment of cases 9 and 10 are similar to Case 8, by taking $i = 2$.

Since for every decomposition of w, we obtained a string which was not in L_4, we can conclude that L_4 is not CF.

(e).

$$S \to AB \mid BA \mid A \mid B$$
$$A \to aAa \mid aAb \mid bAa \mid bAb \mid a$$
$$B \to aBa \mid aBb \mid bBa \mid bBb \mid b$$

4. A desired PDA for L_1 is given in Fig.1 and for L_2 in Fig. 2.

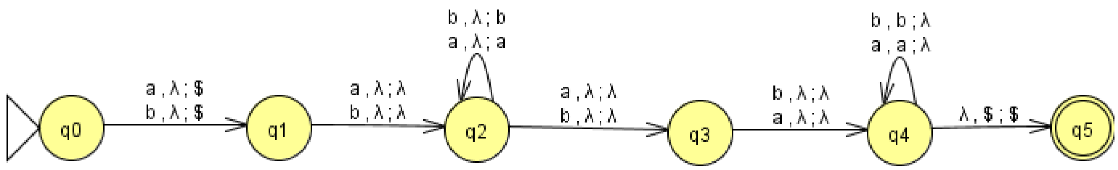

Fig. 1. A PDA for L_1

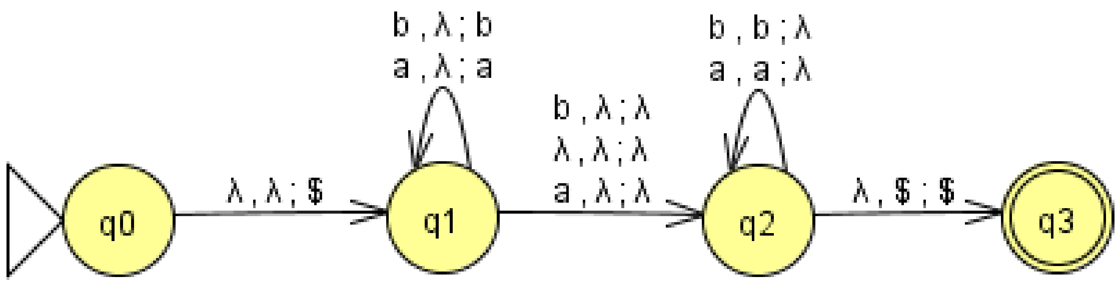

Fig. 2. A PDA for L_2