Sample solution for homework 2

Question 1. It can be proved that for every word $w$, the following holds. $w \in L(G)$ if and only if in every prefix of $w$, the number of $a$’s is greater than or equal to the number of $b$’s.

In this case, the words in (a) and (c) do not belong to $L(G)$, whereas those in (b) and (d) do. The derivation trees for (b) and (d) are omitted here.

Question 2. In the following $X, S$ are variables and $S$ is the start variable.

(a) $L_1 = \{a^m b^n \mid m > n\}$ is generated by the following grammar:

$$
S \rightarrow aSb \mid aS \mid a
$$

(b) $L_2 = \{w^R \mid w \in \{a,b\}^*\}$ is generated by the following grammar:

$$
S \rightarrow X$

$$
X \rightarrow aXa \mid bXb \mid$

(c) $L_3 = \{w_1^R \cdot w_2^R \cdot \ldots \cdot w_k^R \mid \text{each } w_i \in \{a,b\}^* \text{ for some } k \geq 1\}$ is generated by the following grammar:

$$
S \rightarrow SX^R \mid X$

$$
X \rightarrow aXa \mid bXb \mid$

(d) $L_4$ is the complement of the language $\{a^m b^n \mid m \geq n \geq 0\}$ over the alphabet $\{a,b\}$. Note that $w \notin L_4$ if and only if one of the following holds.

- $w = a^m b^n$, where $m < n$.
- There is $a$ that appears after $b$.

Such words can be generated by the following grammar:

$$
S \rightarrow X \mid YbaY$

$$
X \rightarrow aXb \mid Xb \mid b$

$$
Y \rightarrow aY \mid bY \mid \epsilon$

$X$ generates the first case, whereas $YbaY$ generates the words in the second case.

Question 3.

(a) (Sketch) Suppose to the contrary that there is CFG $G = (\Sigma, V, R, S)$ that generates $L_5$.

Consider $w = a^n b^n$, where $n > 2M^{|R|} + 1$. By pumping lemma, the word $ww$ can be partitioned into $sx_1yzt$ such that $y < M^{|R|} + 1$, $|x| + |z| > 0$ and $sx_1yzt \in L(G)$, for every $i \geq 0$.

For $x$ and $z$ to be “pumped” for every $i \geq 0$ and still belong to $L_5$, $x$ must be in the first $w$, and $z$ in the second $w$. Moreover, the position of $x$ in the first $w$ must be the same as the position of $z$ in $w$. Since $n > 2M^{|R|} + 1$, this implies that $|y| > M^{|R|} + 1$, which contradicts the condition in pumping lemma. Thus, there is no such $G$ that generates $L_5$.  

[1]
(b) Suppose to the contrary that there is CFG $G = \langle \Sigma, V, R, S \rangle$ that generates $L_6$.

Consider $a^n$, where $n$ is a prime number bigger than $M^{R} + 1$. By pumping lemma, $a^n$ can be partitioned into $sxyzt$ such that $|x| + |z| > 0$ and $sx^iyz^it \in L(G)$, for every $i \geq 0$. If we pick $i = n + 1$, then:

$$
|sx^{n+1}yz^{n+1}| = |sx^nyz| + n(|x| + |z|)
= n + n(|x| + |z|)
= (n + 1)(|x| + |z|),
$$

which is not a prime and contradicts the assumption that $G$ generates $L_6$.

**Question 4 (2 points).** For (a), if we take $L = \{a^nb^n \mid n \geq 0\}$, which is CFL, and $K = \Sigma^*$, which is regular, then $L \cap K = L$ which has been shown to be not regular.

Similarly, for (b), if we take $L = \{a^nb^n \mid n \geq 0\}$, which is CFL, and $K = \emptyset$, which is regular, then $L \cup K = L$ which has been shown to be not regular.

**Question 5 (2 points).** As before, we simply write $\sim$ to denote $\sim_L$.

Now consider $L = \{a^nb^n \mid n \geq 0\}$, which is CFL. Note that if $m \neq n$, we have: $a^m \not\sim a^n$. This is because $a^mb^m \in L$ but $a^ma^n \notin L_0$ (due to the fact that $m \neq n$).

Therefore, we have infinitely many words $a, a^2, a^3, \ldots$ where each of them belongs to different equivalent classes of $\sim$. Thus, the index of a CFL may not necessarily be finite.