COL 352 Intro to Formal Languages and Theory of Comput, Tutorial Sheet 2

1. Give context-free grammars generating the following sets.

   (a) The set of all strings of balanced parentheses, i.e, each left parenthesis has a matching right parentheses and pairs of matching parentheses are properly nested.

   (b) The set of all strings over alphabet {a,b} with exactly twice as many a's as b's.

   (c) The set of all strings over alphabet {a,b,·,+,*,( ),∈,ϕ } that are well-formed regular expression over alphabet {a,b}. Note that we must distinguish between ϵ as the empty string and as a symbol in the regular expression. We use ∈ in the latter case.

   (d) The set of all strings over alphabet {a,b} not of the form ww for some string w.

2. Suppose G is a CFG with m variables and no right side of production longer than l. Show that if $A \Rightarrow_G^* \epsilon$, then there is a derivation of no more than $\frac{l^m - 1}{l - 1}$ steps by which A derives ϵ. How close to this bound can you actually come?

3. Suppose G is a CFG and w, of length l, is in $L(G)$. How long is a derivation of w in G if

   (a) G is in CNF

   (b) G is in GNF

4. Show that every CFL without ϵ is generated by a CFG all of whose productions are of the form $A \to a$, $A \to aB$, and $A \to aBC$.

5. A language L is said to have the $prefix\ property$ if no word in L is a proper prefix of another word in L. Show that L is $N(M)$ for DPDA M, then L has the prefix property. Is the foregoing necessarily true if L is $N(M)$ for a nondeterministic PDA M?

6. Show that the following are not context free languages.

   (a) $\{a^i b^j c^k | i < j < k\}$

   (b) $\{a^i b^j | j = i^2\}$

   (c) $\{a^i | i$ is a prime $\}$

   (d) the set of strings of a's, b's and c's with an equal number of each

   (e) $\{a^n b^n c^m | n \leq m \leq 2n\}$

7. Which of the following are CFL's?

   (a) $\{a^i b^j | i \neq j$ and $i \neq 2j\}$

   (b) **(a+b)*** - $\{(a^n b^n)^n \ | n \geq 1\}$

   (c) $\{ww^R w | w$ is in **(a+b)***$\}$

   (d) $\{b_i \# b_{i+1} | b_i$ is i in binary, $i \geq 1\}$

   (e) $\{wxw | x$ are in **(a+b)***, w is in $(a+b)^+\}$

   (f) **(a+b)*** - $\{(a^n b)^n \ | n \geq 1\}$

8. Show that if L is a CFL over a one-symbol alphabet, then L is regular. [Hint: Let n be the pumping lemma constant for L and let $L \subseteq 0^*$. Show that for every word of length n or more, say 0^m, there are p and q no greater than n such that 0^{p+iq} is in L for all $i \geq 0$. Then show show that L consists of perhaps some words of length less than n plus a finite number of linear sets, i.e., sets of the form $\{0^{p+iq} | i \geq 0\}$ for fixed p and q, $q \leq n$. You may want to bound the number of parse trees of a certain depth, call them the *base trees* so that every other larger parse tree can be obtained by pumping some portions of one of the base trees.]