Assignment 4: **CFG grammars and languages**

(200 points)

You are required to solve the following problems:

1. (20 points) Recall the CFG $G_4 = (\{E, T, F\}, \{+, x, (,), a\}, R, E)$ where R is the set of productions:

$$
\begin{aligned}
E &\longrightarrow E + T | T \\
T &\longrightarrow T * F | F \\
F &\longrightarrow (E) | a
\end{aligned}
$$

   Give the leftmost and the rightmost derivations for each of the following strings:

   (a) `a`

   (b) `a + a`

   (c) `a+a+a`

   (d) `((a))`

   You get 5 points for each correct leftmost and rightmost derivations.

2. (a) (10 points) Use the languages $A = \{a^m b^n c^n | m, n \geq 0\}$, $B = \{a^n b^n c^m | m, n \geq 0\}$ and $C = \{a^n b^n c^n | n \geq 0\}$ to show that class of context-free languages is not closed under intersection.

   (b) (10 points) Use part (1) above and DeMorgan's law to show that the class of context-free languages is not closed under complementation.

3. (20 points) Let $E = \{a^i b^j | i \neq j and 2i \neq j\}$. Show that E is a context-free language.

4. (35 points) Give context-free grammars that generate the following languages. In all parts the terminal set is $\{0, 1\}$. You get 5 points for each correct context-free grammar you have constructed.

   (a) $L_1 = \{w | w$ contains at least three `1`-s $\}$

   (b) $L_2 = \{w | w$ starts and ends with the same symbol $\}$

   (c) $L_3 = \{w |$ the length of w is odd $\}$

   (d) $L_4 = \{w |$ the length of w is odd and its middle symbol is 0 $\}$

   (e) $L_5 = \{w | w$ contains more `1`-s than `0`-s $\}$

   (f) $L_6 = \{w | w = w^{\mathcal{R}}$, that is w is a palindrome $\}$

   (g) $L_7 = \emptyset$

5. (15 points) Give a context-free grammar that generates the language:

$$A = \{a^i b^j c^k | i, j, k \geq 0 \wedge (i = j \vee j = k)\}$$

   Is your grammar ambiguous? Why or why not?

6. (30 points) Consider the CFG $G = (V, \Sigma, R, S)$, where $V = \{A, B\}$, $\Sigma = \{0\}$, $R = \{A \rightarrow BAB|B|\epsilon; B \rightarrow 00|\epsilon\}$, and $S = A$. You are required to convert G into an equivalent CFG G', where G' is in Chomsky normal form. Use the procedure given in Theorem 2.9.

7. (40 points) Use pumping lemma to prove that the following languages are not context-free. You get 10 points for each correct proof.
   (a) $L_1 = \{0^n 1^n 0^n 1^n | n \geq 0\}$
   (b) $L_2 = \{0^n \# 0^{2n} \# 0^{3n} | n \geq 0\}$
   (c) $L_3 = \{w\#x | w \text{ is a substring of } x,\ w, x \in \{a, b\}^*\}$
   (d) $L_4 = \{x_1 \# x_2 \# \ldots \# x_k | x_i \in \{a, b\}^* \text{ and for some } i \neq j,\ x_i = x_j\}$.

8. (20 points) Show that, if G is a CFG in Chomsky normal form, then for any string $w \in L(G)$ of length $n \geq 1$, exactly $2n - 1$ steps are required for any derivation of w.

9. (20 pints) Let $D = \{xy | x, y \in \{0, 1\}^* \wedge |x| = |y| \ but \ x \neq y\}$. Show that D is a context-free language by constructing a PDA that recognizes it.

10. (20 points) Show that $F = \{a^i b^j | i = kj \text{ for some } k > 0\}$ is not context-free.

Duedate: Monday 2 November 2009.