CS208 Homework #4

Due Date: Mar 30, 2012

- ***Be brief, complete and stick to what has been asked. If needed, you may cite results/proofs covered in class without reproducing them.***
- ***Do not copy solutions from others***
- ***Penalty for copying: FR grade***

1. *[5+5 marks]* For each of the following languages, give a context-free grammar that generates the language. Specifically, you must indicate the set of non-terminals, the set of terminals, the set of production rules and the start symbol in your grammar.
   Important: In addition to the grammar, you must give a description (using a couple of sentences) of the set of strings represented by each non-terminal symbol in your grammar. You must also explain in a couple of sentences the justification behind each production rule in your grammar. Answers without these justifications will fetch 0 marks. You are, however, not required to give a formal proof that your grammar generates the language under consideration.

   In the following, $\#0(w)$ and $\#1(w)$ denote the number of 0s and 1s in a word $w \in \{0,1\}^*$.

   (a) $\{0^i 1^j 0^k \mid j + k = i, \text{ and } i, j, k > 0\}$.
   (b) $\{w \mid w \in \{0,1\}^*, \#0(w) = 2 \times \#1(w)\}$.

2. *[5 + 5 marks]* Show that each of the following languages is not context-free by showing that they violate the pumping lemma for context-free languages.

   (a) $\{w \circ w^R \circ w \mid w \in \{0,1\}^*\}$. In this question, w^R denotes the reverse of a string. Thus, $011^R = 110$.
   [Please note that this sub-question was earlier $\{w \circ w \mid w \in \{0,1\}^*\}$. That sub-question is now withdrawn since a solved version appears in one of the textbooks.]

(b) $\{0^n 1^n 2^i \mid 0 \leq i \leq n\}$

3. *[10 marks]* Consider the alphabet $\Sigma = \{a_1, \ldots a_N, b_1, \ldots b_N\}$ and the language $L = \bigcup_{i=1}^{M} \{a_i^j b_i^j \mid j \geq 0\}$. It is not hard to show that L is context-free. Show that it is not possible to have a pushdown automaton (PDA) $(Q, \Sigma, \Gamma, Q_0, X_0, \delta, F)$ that recognizes L by the empty stack if $|Q| < \sqrt{\frac{N}{|\Gamma|}}$. Recall that Q is the set of states of the PDA and Γ is its stack alphabet.

4. *[Bonus 10 marks (optional):]* For every natural number n, let $enc(n)$ denote the binary encoding of n with the most significant bit being 1. Thus, $enc(5) = 101$ and $enc(24) = 11000$. Show that $\{w \mid w = enc(p)$ for some prime no. $p\}$ is not context free. For purposes of this problem, you may use the following result (also called Fermat's theorem):
For every prime number $p > 2$, there exists a natural number $k > 1$ such that $2^{p-1} = k.p + 1$.