Computability - Exercise 4

All answers should be proved formally

Due Monday, June 16, 11:59 AM

1. (a) Specify the Myhill-Nerode equivalence classes of the language $L_1 = \{0^i 1^j \mid i > j\}$. (You do not have to prove that these are the equivalence classes, but do be careful not to miss anything). Is L_1 regular?

   (b) Is $L_2 = \{w \in \{0,1\}^* \mid$ the number of 0's in w is greater then the number of 1's in $w\}$ regular? (remember to prove your answer).
   Hint: You may use the language 0^*1^*.

2. Describe the (Myhill-Nerode) equivalence classes of the language $L = (0+1)^*010(0+1)^*$ and draw a deterministic finite automaton (a.k.a. DFA) for L, based on these classes. (No proof required)

3. (based on a question from last year's midterm exam)
   Let $C = \{L_1, L_2, \ldots, L_n\}$ be a finite set of regular languages over an alphabet Σ. Suppose that for each one of the languages in the set, the number of Myhill-Nerode equivalence classes is k. For $m \geq 0$, let L^m denote the language

   $$L^m = \{w \in \Sigma^* : w \text{ belongs to exactly } m \text{ languages from } C\}$$

   Give a tight bound for the number of Myhill-Nerode equivalence classes of L^m.
   A tight bound is a function $f : \mathbb{N} \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, such that for all n,k, and m, there is no set $C = \{L_1, \ldots, L_n\}$ for which the DFA for L^m needs more than $f(n, m, k)$ states (that is, f is an upper bound), and there exists a set $C = \{L_1, \ldots, L_n\}$ for which an automaton for L^m needs at least $f(n, k, m)$ states (that is, f is also a lower bound).
   Justify your answer (without a formal proof).

4. For each of the following languages over $\Sigma = \{0, 1\}$, write a context-free grammar with the **minimal** number of variables that generates the language (without further proof).

   (a) $\{w \mid w = w^R\}$ (w^R denotes the reverse of w).

   (b) $\{w \mid w \neq w^R\}$.

   (c) $\{w \mid$ the number of 0's in w equals to the number of 1's$\}$.

5. (optional)
   Let G be a context-free grammar in Chomsky normal form that contains k variables. Show that, if G generates some string using a derivation with at least 2^k steps, then $L(G)$ is infinite (that is, contains infinitely many words).