CSE 2001 Fall 2012 Second test - version 1
Solutions

1. (3 points) Prove that the following language is not regular using the Pumping Lemma. You must include all the steps of the proof to get full credit. Assume that $\Sigma = \{a, b\}$.

$$\{a^n b^{2n} | n \geq 0\}$$

Solution: Assume that the given language (call it L) is regular. Therefore there exists a DFA for it. Let p be the pumping length. We choose string $s = a^p b^{2p} = xyz$ for pumping. From the third condition we know that $|xy| \leq p$ so y consists of only a's. Pumping down we get the string $s' = a^{p-k} b^{2p} \in L$ where $k > 0$. This is a contradiction because s' does not satisfy the criterion in the definition of L. Therefore L is not regular.

2. (3 points) Prove that non-regular languages are closed under complementation.

Solution: Let L be a non-regular language. If L^c is regular then its complement is a regular language since regular languages are closed under complementation. But the complement of L^c is L. Therefore L must be regular as well. Thus the complement of every non-regular language is non-regular, i.e., non-regular languages are closed under complementation.

3. (2 points) Define the Chomsky normal form for context-free languages.

Solution: The criteria for CNF are as follows (from page 109, third edition of Sipser):

   1. Every rule is of the form $A \to BC$ or $A \to a$, where A is any variable, B, C are any variables that are not the start variable and a is a terminal.
   2. The only rule that can have an ϵ on the right is $S \to \epsilon$.

4. (3 points) Is the following language context-free? First write yes or no and then provide a CFG/PDA if it is context-free or a Pumping Lemma proof if it is not.

$$\{a^m b^m a^n b^n | m \geq 0, n \geq 0\}$$

Solution: The language is context-free. A CFG for this language is

   1. $S \to UV$
   2. $U \to aUb|\epsilon$
   3. $V \to aVb|\epsilon$

The variables U, V independently generate strings of the form $a^k b^k$. The first rule concatenates the blocks produced by variables U, V.

5. (3 points) Construct a CFG for the following language $\{a^i b^j | 2i = 3j + 1\}$. Give a very brief explanation of why it is correct.

Solution: Note that j cannot be even, because if it is then $3j + 1$ is odd. Listing the first few values of j, we see that the corresponding strings are aab, a^5b^3, a^8b^5 and so on. The general formula is $a^{3k+2}b^{2k+1}$. So the CFG is
$S \to aab|aaaSbb$.

6. (3 points) Recall that an integer $p > 1$ is prime if and only if the only positive integer factors of p are 1 and p. Prove that the following language is not context-free using the Pumping Lemma. You must include all the steps of the proof to get full credit.

$$L = \{a^p | p \text{ is a prime number}\}$$

Solution: Assume that the given language L is context-free. Therefore there exists a PDA for it. Let p be the pumping length. We choose string $s = a^p = uvxyz$ for pumping. From the third condition we know that $|vxy| \leq p$. Obviously vxy consists of only a's. Suppose $|vy| = k$. Pumping up p times we get the string $s' = a^{p+kp} \in L$ where $k > 0$. We see that s' does not satisfy the criterion in the definition of L – observe that $p + kp = (1+k)p$ and $k > 0$. Thus $(1+k)p$ cannot be a prime number. This is a contradiction. Therefore L is not context-free.

7. (3 points) In class, it was proved that the language $REP = \{ww | w \in \Sigma^*\}$ is not context-free. Point out the major flaw in the following (erroneous) proof. You must explain why one of the statements is incorrect, and not just refer to the statement.

   1. Assume L is a context-free language.
   2. The concatenation of two context-free languages is context-free.
   3. REP is the concatenation of L and L.
   4. Therefore REP is context-free.

   Solution: Step 3 is the erroneous step.

   Note that $L \cdot L = \{wv | w \in L, v \in L\}$. So, $REP \subset L \cdot L$, but $REP \neq L \cdot L$.