Theory of Computation

Fall 2012, Midterm Exam.

Due: Nov. 12, 2012

1. (20 pts) For each of the following languages L, state whether it is **(1) regular**, **(2) context-free but not regular**, or **(3) not context-free**. Prove your answer. Make sure, if you say that a language is context free, that you show that it is not also regular.

   (a) $\{w \in \{0,1\}^* \mid \exists k \geq 0$ and w is a binary encoding (leading zeros allowed) of $2^k + 1\}$. (E.g., $001001 \in L$, for 001001 is the binary encoding of $2^3 + 1$.)
   Soultion: (1) Regular. $L = 0^*(10 \cup 10^*1)$.

   (b) $\{0^p 1^q \mid 0 \leq p \leq q\}$.
   Solution: (2) Context-free but not regular.

   (c) $\{(a^m b^m)^n a^n b^n \mid m, n > 0\}$.
   Solution: (3) Not context-free.

   (d) $\{a^i b^j c^j d^i \mid i, j \geq 0\}$
   Solution: (2) Context-free but not regular.

2. (20 pts) True or False? Give a convincing argument. No penalties for wrong answers.

   (a) Let $L_4 = L_1 L_2 L_3$. If L_1 and L_3 are not regular and L_2 is regular, it is possible that L_4 is regular.
   Solution: True. Take $L_2 = \emptyset$. Then $L_4 = \emptyset$.

   (b) Every subset of a context-free language is context-free.
   Solution: False. $\{a^n b^n c^n \mid n \geq 0\} \subseteq a^*b^*c^*$ and $a^*b^*c^*$ is context-free.

   (c) It is possible that the intersection of an infinite number of regular languages is not regular.
   Solution: True. Let $S_i = (\bigcup_{1 \leq j \leq i} \{a^{p_j}\}) \cup \{a^k \mid k \geq p_i\}$, where p_j is the j-th prime number. E.g., $S_4 = \{\mathbf{a^2}, \mathbf{a^3}, \mathbf{a^5}, \mathbf{a^7}, a^8 . a^9, a^{10} ...\}$. Clearly, each S_i is regular. However, $\bigcap S_i = \{a^p \mid p$ is a prime $\}$ – not regular.

   (d) Given two alphabets Σ and Γ, and a language L over Σ. Let h be a homomorphism $h : \Sigma^* \to \Gamma^*$. Then $h^{-1}(h(L)) = L$, where h^{-1} is the inverse homomorphism of h.
   Solution: False. Let $\Sigma = \{a, b, c\}$ and $\Gamma = \{0, 1\}$, and $h(a) = h(b) = h(c) = 0$. Let $L = \{a\}$. Then $h(L) = \{0\}$. However, $h^{-1}(h(L)) = h^{-1}(\{0\}) = \{a, b, c\}$.

   (e) Let $M = (Q, \Sigma, \delta, q_0, F)$ be the minimal DFA recognizing the language $L(M)$ (i.e., the number of states cannot be reduced). Suppose M' is same as M except the initial state is changed to $q (\neq q_0)$, for some $q \in Q$, i.e., $M' = (Q, \Sigma, \delta, q, F)$. Assuming all states in Q are reachable from q, then M' is the minimal DFA recognizing $L(M')$.
   Solution: True. Because any two states of M (as well as M') are pairwise distinguishable.

3. (10 pts) Define $D(L) = \{s_1 s_2 \mid s_1 a s_2 \in L, s_1, s_2 \in \Sigma^*, a \in \Sigma\}$. That is, $D(L)$ is the language of strings that can be obtained by deleting exactly one symbol from some string in L. Prove in detail that if L is regular then $D(L)$ is also regular.
   Solution: Since L is regular, there exists a DFA

   $$A = (Q, \Sigma, \delta, q_0, F)$$

   that accepts L. Construct a new NFA $A' = (Q \times \{0,1\}, \Sigma, \delta', (q_0, 0), F \times \{1\})$ where

   $$\delta = \{(((q,i),\sigma),(q',i)),\ (((q,0),\epsilon),(q',1)) \mid ((q,\sigma),q') \in \delta, i \in \{0.1\}$$

   NFA A' essentially consists of two copies of A with ϵ-edges from the first copy to the second. The start state is in the first copy and the final states are all in the second, so every accepting path of A' includes exactly one ϵ-edge. Each ϵ-edge serves to delete exactly one symbol from a string in L; therefore A' accepts exactly language $D(L)$. We conclude that $D(L)$ is regular.

4. (15 pts) Let middle be a function that maps from any language L over some alphabet Σ to a new language as follows:

$$middle(L) = \{x \mid \exists y, z \in \Sigma^*, (yxz \in L)\}.$$

   (a) (5 pts) Let $L = \{w \in \{a,b\}^* \mid \#_a(w) = \#_b(w)\}$. What is middle($L$)? (Here $\#_a(w)$ denotes the number of a's in w. E.g., $\#_a(ababa) = 3$.)
   Solution: $\{a,b\}^*$

   (b) (10 pts) Prove formally that, for any language L, if L is regular then middle(L) is also regular.
   Solution: Idea: Let M be an FA accepting L. The proof is by building two extra copies of M, both of which mimic all of M's transitions except they read no input. From each state in copy one, there is a transition labeled ϵ to the corresponding state in M, and from each state in M there is a transition labeled ϵ to the corresponding state in the second copy. The start state of M^* is the start state of copy 1. So M^* begins in the first copy, performing, without actually reading any input, whatever M could have performed while reading some initial input string y. At any point, it can guess that its skipped over all the characters in y. So it jumps to M and reads x. At any point, it can guess that its read all of x. Then it jumps to the second copy, in which it can do whatever M would have done on reading z. If it guesses to do that before it actually reads all of x, the path will fail to accept since it will not be possible to read the rest of the input.

5. (10 pts) Let L be the language over $\Sigma = \{a,b\}$ consisting of all words x for which the number of a's in x equals the number of b's in x.

$$L = \{x \in \Sigma^* \mid \#_a(x) = \#_b(x)\}$$

   Let R_L be the relation induced by L as discussed in class.

   (a) (4 pts) Is aaR_Laaa? Is $\epsilon R_L ab$? Why?
   Solution: (1) aa and aaa are not R_L related, because $aabb \in L$ but $aaabb \notin L$.
   (2) $\epsilon R_L ab$ is correct.

   (b) (6 pts) Use Myhill-Nerode Theorem to show that L is not regular.
   Solution: $\forall i,j \geq 0, i \neq j$, a^i and a^j are not R_L related. Hence R_L induces an infinite number of equivalence classes.

6. (10 pts) Let L be the language of the regular expression a^*b^*. Prove formally that any DFA accepting L must have at least two final states. (Hint: Proof by contradiction.)
Solution: Let $M = (Q, \Sigma, \delta, q_0, F)$ be a DFA accepting L. Since $\epsilon \in L$, $q_0 \in F$. Let $\hat{\delta}(q_0, ab) = q_1 \in F$. We claim that $q_0 \neq q_1$. If otherwise (i.e., $q_0 = q_1$), then consider $\hat{\delta}(q_0, abab) = \hat{\delta}(q_0, ab) = q_0$ – contradicting the fact that $abab \notin L$.

7. (15 pts) Answer the following questions:

   (a) (10 pts) Prove that $L = \{0^{(2n+1)^2} \mid n \geq 0\}$ is not regular using the Pumping Lemma.
   Solution: Idea: Let k be the pumping constant. Consider $x = 0^{(2k+1)^2} = u \cdot v \cdot w$, with $0 \leq |v| \leq k$. Then $|u \cdot v^2 \cdot w| = (2k+1)^2 + |v| \leq (2k+1)^2 + k < (2(k+1)+1)^2$. Hence, $u \cdot v^2 \cdot w \notin L$.

   (b) (5 pts) Use the above result to show that $L' = \{0^{n^2+n} \mid n \geq 0\}$ is not regular by closure properties. (Do not use Myhill-Nerode Theorem or the Pumping Lemma; use only (a) and closure properties of regular languages.)
   Solution: Define a homomorphism $h(0) = 0000$. Then by the closure properties under homomorphism and concatenation. $h(L') \cdot \{0\}$=$\{0^{4n^2+4n+1} \mid n \geq 0\}$=$L$. Hence, L' is not regular.