Theory of Computation

Spring 2020, Midterm Exam (Solutions)

Due: April 21, 2020

1. (10 pts) True or False? Prove your answer.

   (a) If $A \cap B$, $B \cap C$, and $A \cap C$ are regular languages, then $A \cup B \cup C$ is regular as well.

   (b) If $L \cdot L$ is a nonregular language, then L^* is nonregular as well.

   Solution

   (a) False. Let A be any nonregular language, and let $B = C = \emptyset$.

   (b) False. Consider the language $L = \{0, 1\} \cup \{0^n 1^n : n > 0\}$. Then $L^* = \{0, 1\}^*$

2. (20 pts) For each of the following languages $L_1, ..., L_{10}$, determine whether it is regular or not. No explanations are needed, however, $Score = \max\{0,\ \text{right-}\frac{1}{2}\ \text{wrong}\}$.

   (a) Let L be a given regular language. Let L_1 be the set of strings w such that $w^R w \in L$, where w^R is the reversal of w, e.g., $0010^R = 0100$. Is L_1 always regular?
   Solution: Regular. Pick a state q; Simulate $q \xrightarrow{w} q_f$ (in a forward direction) and $q \xrightarrow{w} q_0$ (in a backward direction, i.e., to check $q_0 \xrightarrow{w^R} q$) in parallel, where q_0 and q_f are the initial state and the final state, respectively.

   (b) L_2 is the set of strings of the form uv, where $u, v \in \{0, 1\}^*$ are palindromes. Note that a palindrome is a string x such that $x = x^R$, e.g., 001100 is a palindrome.
   Solution: Not Regular. $0^i 110^i 1 = (0^i 110^i)1 \in L_2$ but $0^j 110^i 1 \notin L_2$, for $i \neq j$. Hence, $0, 0^2, 0^3, 0^4$.. are in different equivalence classes.

   (c) L_3 is the set of binary strings in which the number of 0s and the number of 1s differ by an integer multiple of 17.
   Solution: Regular. Use states to keep track of the difference of 0s and 1s modulo 17.

   (d) $L_4 = \{x\#y \mid x, y \in \{0, 1\}^*,\ \text{when viewed as binary numbers, } x + y = 3y\}$. For instance, $1000\#100 \in L$, as 1000=8, 100=4, 8+4 = 3*4.
   Solution: Not regular, which we show using the Pumping lemma. We must start by choosing a string that is in fact in L_4. Let $w = 100^k\#10^k$.

   (e) $L_5 = \{w \mid w = xyzy,\ x, y, z \in \{0, 1\}^*\}$.
   Solution: Regular. The key to why this is so is to observe that y can be ϵ.

   (f) We define $max\text{-}string(L) = \{w \mid w \in L, \forall z \in \Sigma^*(z \neq \epsilon \Rightarrow xz \notin L)\}$. Suppose L is regular, is $L_6 = max\text{-}string(L)$ always regular?
   Solution: Regular. For each final state q_f of the original FA, if another final state q'_f is reachable from q_f, then remove q_f from the list of final states.

   (g) $L_7 = \{w \in \{a, b\}^* \mid$ the first, middle, and last characters of w are identical $\}$. For example, **a**bb**a**ab**a** $\in L_7$.
   Solution: Not regular, consider $L \cap ab^*ab^*a = \{ab^n ab^n a \mid n \geq 0\}$

   (h) L_8 is the set of strings that contain a substring of the form wuw where $u, w \in \{0, 1\}^*$. Note that x is a substring of y if there exist $s, t \in \Sigma^*$ such that $y = sxt$.
   Solution: Regular. w can be ϵ.

   (i) L_9 is the set of odd-length strings with middle symbol 0 (over alphabet $\{0, 1\}$).
   Solution: Not regular. Intersect the language with $1^*01^* = \{1^n 01^n\}$

   (j) $L_{10} = \{1^k y \mid y \in \{0, 1\}^*, y$ contains at most k 1s$, k \geq 1\}$. For instance, 11101 is in the language, as choosing $y = 01$ meets the requirement.
   Solution: Not regular. $L_{10} \cap 1^*01^* = \{1^n 01^m \mid m \leq n\}$.

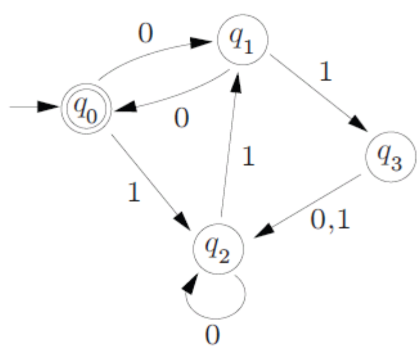
3. (10 pts) Use the pumping lemma to show that the following language is not regular. Show your steps in detail.

$$\{(ba)^n b^n \in \{a,b\}^* \mid n \geq 0\}$$

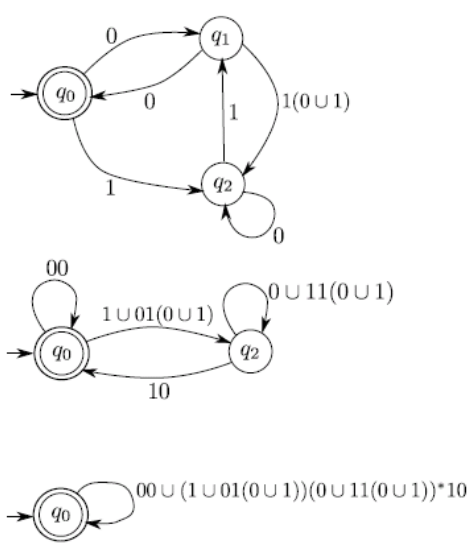
Solution:

Assume to the contrary that A is regular. Therefore it has a pumping length $p \geq 1$. Consider $s = (ba)^p b^p$, we observe that $s \in A$. Since $|s| \geq p$, every proper break down of it must be "pumpable". Let $s = xyz$ be a proper break down of string x according to pumping lemma, that is $|xy| \leq p$ and $y \neq \epsilon$. Since $|(ba)^p| = 2p$, the string xy is just a prefix of $(ba)^p$. Since $y \neq \epsilon$, when we pump down y, we will remove at least one a or one b from the $(ba)^p$ prefix of s while s will have still p trailing b's; this means that $xz \notin A$, which contradicts the pumping lemma. Thus A can not be regular.

4. (10 pts) Convert the following DFA into a regular expression that describes the same language. We eliminate states in the order q_3, q_1, q_2. We do not need to add dummy initial and final states here, since they don't play a role in this example (there is just one final state in the DFA). Show your work in sufficient detail.



Solution:



So our regular expression is $\left(00 \cup (1 \cup 01(0 \cup 1))(0 \cup 11(0 \cup 1))^* 10\right)^*$.

5. (15 pts) (Myhill-Nerode Theorem)

(a) (7 pts) Given a language L, consider the equivalence relation $\equiv_L$ on Σ^* ($\equiv_L \subseteq \Sigma^* \times \Sigma^*$) defined by: $x \equiv_L y$ if and only if for all $z \in \Sigma^*, xz \in L$ iff $yz \in L$. Write down the equivalence classes of $\equiv_L$ for the language $L = \{0^n1^n \mid n > 0\}$ over the binary alphabet $\{0, 1\}$.

Solution

- $\{0^i\}$ for $i = 0, 1, 2, ...$;
- $\{0^{n+i}1^n \mid n \geq 1\}$ for $i = 0, 1, 2, ...$;
- $\{0, 1\}^* - \{0^n1^m \mid n \geq m\}$

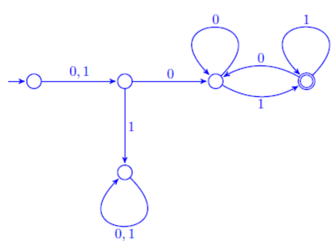
Any two classes in the first group can be distinguished from one another by a string of all 1s; likewise for the second group. Any class in the first group can be distinguished from any class in the second group by a string of the form 011...1. Finally, the strings in $\{0, 1\}^* - \{0^n1^m \mid n \geq m\}$ form a separate class because they are precisely those strings that cannot be made into a string of L by appending any suffix.

(b) (8 pts) Suppose the equivalence classes induced by $\equiv_L$ for a language L (over $\Sigma = \{0, 1\}$) accepted a DFA M has the following five equivalence classes. Furthermore, $001101 \in L$, and M has only one final state.

- $C_1 = \{\epsilon\}$
- $C_2 = \{0, 1\}$
- $C_3 = \{01, 11, 010, 011, ...\}$
- $C_4 = \{00, 000, 0010, 1010, ...\}$
- $C_5 = \{000001, 0011, 10011...\}$

Draw the DFA M.

Solution



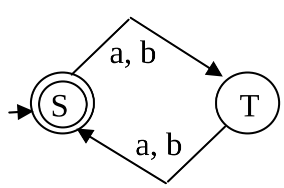
6. (5 pts) Consider the following right-linear grammar over alphabet $\{a, b\}$:

$$S \to aT \mid bT \mid \epsilon$$

$$T \to aS \mid bS$$

Draw a finite automaton to accept the language generated by the above grammar.

Solution:



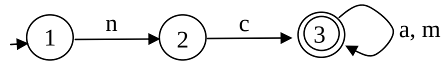
7. (10 pts) A set of pairs of strings $F = \{(x_i, y_i) \mid i = 1, 2, ..., n\}$ is called a *fooling set* for a language L if for each i, j in $\{1, 2, ..., n\}$,

   (a) $x_i y_i \in L$, and

   (b) if $i \neq j$, then $x_i y_j \notin L$ or $x_j y_i \notin L$

   There is a theorem saying that if F is a fooling set for a regular language L, then every NFA for L has at least $|F|$ states. Consider the following language: $L = nc(m+a)^*$, where $\Sigma = \{a, c, m, n\}$.

   (a) (5 pts) Give (draw) an NFA with the minimum number of states to accept L.
   Solution:



   (b) (5 pts) Find a fooling set to show that the NFA you design is minimum.
   Solution: $F = \{(\epsilon, ncma), (n, cma), (nc, ma)\}$

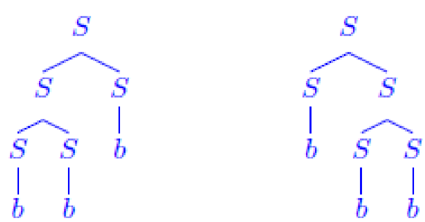
8. (12 pts) Consider the context-free grammar G: $S \to SS \mid aS \mid b$ (over alphabet $\Sigma = \{a, b\}$).

   (a) (4 pts) Prove that this grammar is ambiguous.

   (b) (4 pts) The language $L(G)$ is in fact regular. Write a regular expression as simple as possible to express $L(G)$.

   (c) (4 pts) Give an equivalent unambiguous grammar.

   Solution:



   (a)

   (b) $(a+b)^*b$

   (c) $S \to aS \mid bS \mid b$

9. (8 pts) We define $L_1/L_2 = \{u \in \Sigma^* \mid uv \in L_1, \exists v \in L_2\}$, and $L_1 \setminus L_2 = \{u \in \Sigma^* \mid vu \in L_1, \exists v \in L_2\}$. Suppose $L_1 = \{a^n b^n c^n \mid n \geq 0\}$ and $L_2 = \{b, c\}^*$. Answer the following questions:

   (a) $L_1/L_2 = ?$
   Solution: $\{a^m b^m c^n \mid m \geq n \geq 0\} \cup \{a^m b^n \mid m \geq n \geq 0\}$

   (b) $L_2/L_1 = ?$
   Solution: L_2, as $\epsilon \in L_1$ and $L_2 \cdot \{\epsilon\} = L_2$.

   (c) $L_1 \setminus L_2 = ?$
   Solution: L_1, as $\epsilon \in L_2$ and $\{\epsilon\} \cdot L_1 = L_1$.

   (d) $L_2 \setminus L_1 = ?$
   Solution: L_2, as $\epsilon \in L_1$ and $\{\epsilon\} \cdot L_2 = L_2$..