Homework 1 – Due Thursday, February 4, 2021 at 11:59 PM

Reminder Collaboration is permitted, but you must write the solutions *by yourself without assistance*, and be ready to explain them orally to the course staff if asked. You must also identify your collaborators and write "Collaborators: none" if you worked by yourself. Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden.

Problems There are 7 required problems and one bonus problem. Problems 1-5 and 7 are to be turned in via Gradescope. Problem 6 will be autograded using AutomataTutor.

1. For each of the following languages, (i) give a plain English description of the language, and (ii) give two examples of strings in the language and two examples of strings outside the language.

   (a) $L_1 = \{w \in \{0,1\}^* \mid w = w^R\}$

   (b) $L_2 = \{x\mathsf{aa}y \mid x, y \in \{\mathsf{a},\mathsf{b}\}^*\} \cup \overline{\{x\mathsf{bb}y \mid x, y \in \{\mathsf{a},\mathsf{b}\}^*\}}$

2. For each of the following languages, (i) describe the language using set-builder notation and union/intersection/complement/reverse/concatenation operations (the notation used in Problem 1), and (ii) give two examples of strings in the language and two examples of strings outside the language.

   (a) L_3 = the set of all strings over alphabet $\{\mathsf{a},\mathsf{b},\mathsf{c}\}$ that have length at least 3 and have b as their third symbol.

   (b) L_4 = the set of all strings over alphabet $\{0,1\}$ that either start with 0 and have odd length, or start with 1 and have even length.

3. Which of the following statements are true or false, for all alphabets Σ? For each, provide either a proof or a counterexample.

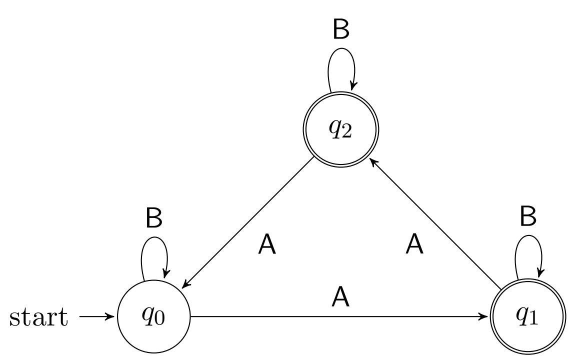
   (a) For all strings $x, y, z \in \Sigma^*$, we have $|x \circ (yz)^R| = |x| + |y| + |z|$. (Recall that $\circ$ denotes concatenation.)

   (b) For all languages $L_1, L_2 \subseteq \Sigma^*$, we have $(L_1 \circ L_2)^R = L_2^R \circ L_1^R$.

   (c) For all languages $L \subseteq \Sigma^*$, we have $L \circ \{\varepsilon\} = L \circ \emptyset$.

   (d) For all languages $L_1, L_2, L_3 \subseteq \Sigma^*$, we have $L_1 \circ (L_2 \cap L_3) = (L_1 \circ L_2) \cap (L_1 \circ L_3)$.

4. Consider the following state diagram of a DFA M using alphabet $\Sigma = \{\mathsf{A},\mathsf{B}\}$.



(a) What is the start state of M?

(b) What is the set of accept states of M?

(c) Give a formal description of the machine M (i.e., as a 5-tuple).

(d) What sequence of states does the machine go through on input ABBAB?

(e) Does the machine accept the string ABBAA?

(f) Does the machine accept the string ABABAA?

(g) What is the language recognized by M? (Hint: It has a simple English description using modular arithmetic.)

5. No Problem 5 – this was about the formal definition of NFAs, which we did not get to in class, so it will be assigned on HW 2.

6. This problem will be autograded using AutomataTutor. Visit `http://automatatutor.com/` and click on "Sign Up." Make an account using your name and `@bu.edu` email address (it is important for recording grades that the information for your account match the information on the course list provided by the university). We'll provide more specific information about how to register for this course on Piazza.

   Give state diagrams of DFAs with as few states as you can recognizing the following languages. You may assume that the alphabet in each case is $\Sigma = \{0, 1\}$.

   (a) $L_1 = \{w \mid w \text{ begins with 1 and ends with 00}\}$.

   (b) $L_2 = \{w \mid w \text{ contains at most three 1's}\}$.

   (c) $L_3 = \{w \mid w \text{ contains the substring 010}\}$.

   Give state diagrams of NFAs with as few states as you can recognizing the following languages. You may assume that the alphabet in each case is $\Sigma = \{0, 1\}$.

   (d) $L_4 = \{w \mid w \text{ contains an even number of occurences of the substring 01}\}$.

   (e) $L_5 = \{w \mid w \text{ contains substrings 100 and 10 which do not overlap}\}$.

7. Draw (and include in the PDF you submit to Gradescope) state diagrams of DFAs with as few states as you can recognizing the following languages. You may assume that the alphabet in each case is $\Sigma = \{0, 1\}$.

   (a) $L_6 = \{w \mid w$ is a string of the form $x_1y_1x_2y_2 \ldots x_ny_n$ for some natural number n such that $x_i, y_i \in \{0, 1\}$ and $x_i = y_i$ for all $i\}$.

   (b) $L_7 = \{w \mid w$ represents a binary number that is congruent to 1 modulo 3$\}$. In other words, this number minus 1 is divisible by 3. The number is presented starting from the most significant bit and can have leading 0s.

Bonus Problem

(a) Give a state diagram of a DFA recognizing the language $ADD = \{w \mid w$ is a string of the form $x_1y_1z_1x_2y_2z_2 \ldots x_ny_nz_n$ for some natural number n such that $x+y = z$ as binary numbers$\}$. Here, x, y, z are presented starting from the least significant bit and can have trailing 0's.

(b) Show that for any natural number n, the language $MOD_n = \{w \mid w$ represents a binary number that is divisible by $n\}$ is regular. The number is presented starting from the most significant bit and can have leading 0's.