

# CSC 341 - Fall 2022

## Problem Set 3 Solutions

**Problem 1.** Define the alphabet  $\Sigma = \mathbb{N}_2$ . For each of the following, give a regular expression which fully characterizes the language.

- a) Any string except 101010.
- b) Any string ending in 110.
- c) Any string which can be broken into a sequence of the following strings: 000, 111, 101, or 010.
- d) Any string of even length starting with 0 or any string of odd length starting with 1.
- e) Any string not containing the substrings 11 or 000.

**Solution 1.**

- a)  $\epsilon|0|1|00|01|10|11|000|\dots|101001|101011|101100|\dots|111111|(0|1)^7(0|1)^*$
- b)  $(0|1)^*110$
- c)  $(000|111|101|010)^*$
- d)  $\epsilon|0(0|1)((0|1)(0|1))^*|1((0|1)(0|1))^*$
- e)  $\epsilon|0|00|(1|01|001)(01|001)^*|(\epsilon|0|00)$

**Problem 2.** For a string  $\omega \in \Sigma^*$ , we define the *reverse* of  $\omega$ , written  $\omega^R$ , to be  $\omega$  with its sequence reversed. In other words, if

$$\omega = a_1 a_2 \dots a_n$$

for  $a_1, a_2, \dots, a_n \in \Sigma$ , then

$$\omega^R = a_n \dots a_2 a_1.$$

For a language  $A \subseteq \Sigma^*$ , we can define its reverse

$$A^R = \{\omega^R \mid \omega \in A\}$$

to be the reverse of every string in  $A$ .

Prove that the class of regular languages is closed under the reverse operator.

**Solution 2.** Let  $A$  be a regular language. Then there is some DFA  $D = (Q, \Sigma, \delta, q_0, F)$  with  $L(D) = A$ . We construct a new NFA  $N = (Q \cup \{q_s\}, \Sigma, \delta_R, q_s, \{q_0\})$  where on input  $q \in Q \cup \{q_s\}$  and  $c \in \Sigma_\epsilon$ ,

$$\delta_R(q, c) = \begin{cases} F & q = q_s \wedge c = \epsilon \\ \{q'\} & \exists q' \in Q : \delta(q', c) = q \\ \perp & \text{otherwise,} \end{cases}$$

where  $\perp$  indicates that  $N$  halts and rejects (i.e.  $N$  goes to some nonaccepting absorbing state). In other words  $\delta_R$  reverses the transitions of  $\delta$  and provides an epsilon transition from the new start state  $q_s$  to each of the old accepting states  $F$ . It remains to show that  $L(N) = A^R$ .

Let  $\omega \in L(N)$ . Then there exists  $q \in F$  such that  $\delta_R^*(q_s, \epsilon\omega) \ni \delta_R^*(q, \omega) \ni \{q_0\}$ . It then follows that there is a (necessarily deterministic by construction) path  $p_0p_1 \dots p_{|\omega|}$  with  $p_0 = q$  and  $p_{|\omega|} = q_0$  such that for each  $i$ ,  $\delta^R(p_i, \omega_i) = \{p_{i+1}\}$ . By construction, it must then be the case that for each  $i$ ,  $\delta(p_i, \omega_i) = p_{i-1}$ . But then  $p_{|\omega|} \dots p_1p_0$  is an accepting path of  $D$  on input  $\omega^R$ . Therefore  $\omega^R \in L(D)$ , hence  $\omega^R \in A$  and thus  $\omega \in A^R$ .

Now suppose  $\omega \in A^R$ . Then since  $L(D) = A$ , there is an accepting path  $p_0p_1 \dots p_{|\omega^R|}$  in  $D$  such that  $p_0 = q_0$ ,  $p_n \in F$ , and for each  $i$ ,  $\delta(p_i, \omega_i^R) = p_{i+1}$ . By construction, for each  $i$ ,  $\delta_R(p_i, \omega_i^R) = \{p_{i-1}\}$ . Thus  $p_{|\omega^R|} \dots p_1p_0$  is an accepting path of  $N$  on input  $\epsilon(\omega^R)^R = \epsilon\omega = \omega$ . Therefore  $\omega \in L(N)$ .  $\square$

**Problem 3.** Let  $A$  be language of all even length strings of  $\{0, 1\}^*$ . Note that  $A$  is regular.

- Prove or disprove that  $A \circ A^R$  is regular.
- Prove or disprove that  $\hat{A} = \{\omega\omega^R \mid \omega \in A\}$  is regular.
- Give a brief intuitive explanation in your own words of why one is regular while the other is not.

**Solution 3.**

- Let  $B$  be any regular language. Since the class of regular languages is closed under concatenation and reverse,  $B \circ B^R$  is regular. Since  $A$  is regular, we're done.  $\square$
- Suppose that  $\hat{A}$  is regular. Then it must satisfy the pumping lemma. Let  $p$  thus be the pumping length.

Let  $\omega \in A$  and pick  $\omega = 0^p110^p \in \hat{A}$ . However we choose to write  $\omega$  as  $\omega = xyz$  for the pumping lemma, it must be the case that  $y = 0^k$  for some  $1 \leq k \leq p$  since  $|xy| \leq p$ . The pumping lemma then tells us that  $xy^2z \in \hat{A}$ . On the other hand,  $xy^2z = 0^{p+k}110^p$ , which is not of the form  $\nu\nu^R$  for any  $\nu$ , hence  $xy^2z \notin \hat{A}$ . This is a contradiction, so  $\hat{A}$  must not be regular.  $\square$

- $A \circ A^R$  only requires that we have a string from  $A$  and then a reversed string from  $A$ . Recognizing either is not hard. On the other hand,  $\hat{A}$  requires that we remember a string of unbounded length so we can recognize its reversed form. Finite automata obviously can't hold more than a finite amount of memory, so this is impossible.

**Problem 4.** Let  $\Sigma$  be an alphabet. Recall that the set of all languages over  $\Sigma$  is the family of all possible selections/exclusions of strings. In other words, it is the power set of the set of all strings,  $\mathcal{P}(\Sigma^*)$ . Notice that the set of all languages is uncountable even when  $|\Sigma| = 1$ . Also observe that the set of all regular languages  $\mathcal{R}$  over  $\Sigma$  is a subset of  $\mathcal{P}(\Sigma^*)$ . Prove that  $\mathcal{R}$  is countable. (Hint: you do not need to use diagonalization.)

**Solution 4.** Fix an alphabet  $\Sigma$ . Since a language is regular iff a DFA accepts it, then if the set of all DFAs is countable, then so is the class of regular languages. It suffices to show that the set of all DFAs is enumerable. We can enumerate the set of all DFAs as follows.

1. Let  $n = 1$
2. Repeat forever
  - (a) Let  $Q$  be the set of  $n$  states (this is unique up to renaming)
  - (b) There is a finite set of possible transition functions  $\Delta$  over  $Q$  and  $\Sigma$  since both  $Q$  and  $\Sigma$  are finite
  - (c) For each  $\delta \in \Delta$ 
    - i. For each  $q_0 \in Q$ 
      - A. For each  $F \in \mathcal{P}(Q)$ 
        - Output/Count the DFA  $(Q, \Sigma, \delta, q_0, F)$
  - (d) Increment  $n$

Thus the set of all DFAs is countable. □

**Problem 5.** We characterized a regular expression  $R$  over an alphabet  $\Sigma$  as matching one of the following (sometimes recursive) conditions:

1.  $R = a$  for some  $a \in \Sigma$
2.  $R = \epsilon$
3.  $R = \emptyset$
4.  $R = R_1|R_2$ , where  $R_1$  and  $R_2$  are both regular expressions
5.  $R = R_1 \circ R_2$ , where  $R_1$  and  $R_2$  are both regular expressions
6.  $R = R_1^*$ , where  $R_1$  is a regular expression.

Other expressions can be characterized in a similar manner. One major example is the arithmetic expressions. For example, if  $E$  is an arithmetic expression, we would expect  $(E)$  to also be an arithmetic expression but not  $(E$  or  $E)$ .

Define the alphabet  $\Sigma = \mathbb{Z}_{10} \cup \{+, -, *, /, (, )\}$  ( $-$  here may be binary or unary).

- a) Give a characterization similar to the characterization of regular expressions above for the arithmetic expressions.
- b) Prove that the arithmetic expressions are not regular.

**Solution 5.**

a) We characterize the arithmetic expressions as follows. We say that  $A$  is an arithmetic expression if  $A$  is

- 1) some  $\omega \in 0|(\mathbb{Z}_{10} \setminus \{0\})\mathbb{Z}_{10}^*$ ,
- 2)  $A_1 + A_2$  where  $A_1$  and  $A_2$  are arithmetic expressions,
- 3)  $A_1 - A_2$  where  $A_1$  and  $A_2$  are arithmetic expressions,
- 4)  $A_1 * A_2$  where  $A_1$  and  $A_2$  are arithmetic expressions,
- 5)  $A_1/A_2$  where  $A_1$  and  $A_2$  are arithmetic expressions,
- 6)  $(A_1)$  where  $A_1$  is an arithmetic expression, or
- 7)  $-A_1$  where  $A_1$  is an arithmetic expression.

b) Let  $\mathcal{A}$  be the set of all arithmetic expressions. Suppose that  $\mathcal{A}$  is regular. Then it must satisfy the pumping lemma. Let  $p$  be the pumping length, and pick  $\omega = ({}^p0)^p \in \mathcal{A}$ . Then however we write  $\omega = xyz$ , since  $|xy| \leq p$ , it follows that  $y = ({}^k$  for some  $1 \leq k \leq p$ . The pumping lemma insists that  $xz \in \mathcal{A}$ . However, if we remove  $k$  left parenthesis, then the result is clearly not an arithmetic expression since we would then have too many right parenthesis, hence  $xz \notin \mathcal{A}$ . This is a contradiction, therefore  $\mathcal{A}$  cannot be regular.  $\square$