Homework 4: Non-regular languages

CSE 30151 Spring 2016

Due 2016/02/16 at 11:55pm

Instructions

Please note that you will **lose one point** if you don't follow these instructions.

- You can prepare your solutions however you like, but you must submit them as a single PDF file.
- Please name your PDF as follows:
  - If you're making a complete submission, name your PDF file `netid-hw4.pdf`, where `netid` is replaced with your NetID.
  - If you're submitting some problems now and want to submit other problems later, name your PDF file `netid-hw4-1234.pdf`, where `1234` is replaced with the problems you are submitting at this time.
  - If you use the same name twice, only the most recent version will be graded!
- Submit your PDF file in Sakai. Don't forget to click the Submit (or Resubmit) button!

Problems

Each problem is worth 7 points. An additional one point is for legibility, and one point for following the submission instructions.

1. **Quotient languages.** The quotient L_1/L_2 of languages L_1 and L_2 is defined as:

$$L_1/L_2 = \{w \mid \exists u \in L_2 : wu \in L_1\}$$

   (a) If $L_1 = ab^*c$ and $L_2 = b^*c$, what is L_1/L_2?

   (b) Prove that if L_1 is a regular language and L_2 is a *finite* language, then L_1/L_2 is regular (by showing how to construct an automaton for L_1/L_2).

(c) [Problem 1.45] Prove that if L_1 is a regular language and L_2 is *any* language, then L_1/L_2 is regular.

Note: you only have to show that the automaton for L_1/L_2 exists, not how to actually construct it. Your proof should even work for such inscrutable languages as:

$$L_2 = \left\{ w \;\middle|\; \begin{array}{l} w \text{ is a program that prints the scores of all past} \\ \text{and future Notre Dame football games} \end{array} \right\}.$$

2. **Binary addition.**

(a) [Problem 1.32] Let

$$\Sigma_3 = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\},$$

so that a string over Σ_3 gives three rows of binary digits. Show that the following is regular:

$$B = \{w \in \Sigma_3^* \mid \text{the bottom row of } w \text{ is the sum of the top two rows}\}.$$

Hint: Design an automaton for B^R first, then convert it into an automaton for B.

(b) [Problem 1.53] Let $\Sigma = \{\mathtt{0}, \mathtt{1}, +, =\}$, and show that the following is not regular:

$$ADD = \{x{=}y{+}z \mid x, y, z \text{ are binary natural numbers, and } x = y + z \text{ is true}\}.$$

3. **Two similar but different languages** [Problem 1.49].

(a) Let $B = \{1^k y \mid y \in \{0,1\}^* \text{ and } y \text{ contains at least } k \text{ 1s, for some } k \geq 1\}$. Show that B is a regular language.

Hint: Try out some strings to see what does and doesn't belong to B, in order to find another simpler way of thinking about B.

(b) Let $C = \{1^k y \mid y \in \{0,1\}^* \text{ and } y \text{ contains at most } k \text{ 1s, for some } k \geq 1\}$. Show that C is not a regular language.

4. **Real-world languages.** Choose **one** of the following problems. You're welcome to do both, but will only get credit for the first one you do.

(a) Prove that C (the programming language) is not a regular language, using an alphabet of all ASCII characters. Don't forget that comments, strings, and `char` literals can contain arbitrary characters!

(b) Prove that English is not a regular language, assuming that the alphabet is the set of all English words in the dictionary. Consider sentences of the form

- the motorcycle rusted

- the motorcycle the guy rode rusted
- the motorcycle the guy my sister married rode rusted

Assume that this pattern of sentences goes on forever, that is:

$$\{(\langle\mathsf{det}\rangle\langle\mathsf{noun}\rangle)^n\langle\mathsf{verb}\rangle^n \mid n \geq 1\} \subseteq \text{English}$$
$$\{(\langle\mathsf{det}\rangle\langle\mathsf{noun}\rangle)^m\langle\mathsf{verb}\rangle^n \mid m \neq n\} \cap \text{English} = \emptyset.$$

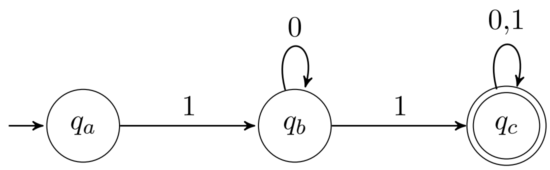
where $\langle\mathsf{det}\rangle$ is any determiner (the, my), $\langle\mathsf{noun}\rangle$ is any noun (motorcycle, guy, sister), and $\langle\mathsf{verb}\rangle$ is any verb (rusted, rode, married). Assume that all three word classes are finite and disjoint.

Solutions

1.

2. Problem 1.49

   - First note that the case $k = 1$ is a superset of all the other cases. The regular expression $10^*1(0\cup1)^*$ fully describes this language. Alternatively, just build it:

   - Here we apply the pumping lemma: Let pumping length be p. Consider the string $1^{p+1}01^{p+1}$. We know that this can be written as xy^iz with $|xy| \leq p$. Thus we have that $y = 1^j$ for some $j \geq 1$. Then $xy^0z = 1^{p+1-j}01^{p+1}$ and this has more 1's in z than in x, a contradiction.

DC: OK to eliminate case 1 by changing p to $p + 1$?