Automata Theory and Computability

Assignment 2

(Due on Fri 20 Sep 2013)

1. This question shows that the pumping lemma for regular languages is *not* a sufficient condition for regularity. Consider the language L below

$$(a+b+c)^*cc(a+b+c)^* + \bigcup_{n\geq 0}(a^+c)^n(b^+c)^n.$$

   (a) Show that L is not regular.

   (b) Show that L nonetheless satisfies the conditions of the pumping lemma. That is, the adversary has a winning strategy in the game over L. In other words, show that there exists an $n > 0$ such that for any string xyz in L, $|y| \geq n$, there exists strings u, v, w such that $y = uvw$, $v \neq \epsilon$, and for all i, the string $xuv^iwz \in L$.

2. We saw in class that ultimate periodicity of lengths of strings in a language was a *necessary* condition for regularity of the language. Argue that over a single-letter alphabet, the condition of ultimate periodicity is also *sufficient* for regularity.

3. Show that the following definitions of ultimate periodicity for a set of natural numbers X are equivalent:

   (a) There exist nonzero natural numbers n and p such that for all $m \geq n$, $m \in X$ iff $m + p \in X$.

   (b) There exist nonzero natural numbers n and p such that for all $m \geq n$, $m \in X$ implies $m + p \in X$.

   Use the second definition to give a simple solution to the problem given in the assignment: If L is a subset of $\{a\}^*$, then L^* is regular.

4. Construct a language $L \subseteq \{a, b\}^*$ with the property that neither L nor its complement contains an infinite regular set.

5. For a set of natural numbers A, define *binary*(A) to be the set of binary representations of numbers in A. Similarly define *unary*(A) to be the set of "unary" representations of numbers in A: *unary*$(A) = \{1^n \mid n \in A\}$. Thus for $A = \{2, 3, 6\}$, *binary*$(A) = \{10, 11, 110\}$ and *unary*$(A) = \{11, 111, 111111\}$.

   Consider the two propositions below:

   (a) For all A, if *binary*(A) is regular then so is *unary*(A).

   (b) For all A, if *unary*(A) is regular then so is *binary*(A).

One of the statements above is true and the other is false. Which is true and which is false?

6. Describe the equivalence classes of the canonical Myhill-Nerode relation for the language of equal number of a's and b's over the alphabet $\{a, b\}$. Draw a representation of the resulting canonical deterministic automaton for the language.

7. Give a regular expression which describes the language defined by the MSO sentence below over the alphabet $\{a, b\}$.

$$\begin{aligned}
&\forall x \forall y ((Q_b(x) \wedge succ(x,y)) \implies Q_a(y)) \wedge \\
&\exists X ((\exists x (zero(x) \wedge x \in X)) \wedge \\
&\quad (\exists x (last(x) \wedge \neg x \in X)) \wedge \\
&\quad (\forall x \forall y (succ(x,y) \implies (x \in X \Leftrightarrow \neg y \in X))) \wedge \\
&\quad \forall x (x \in X \implies Q_b(x))).
\end{aligned}$$

8. Give an MSO sentence describing the language $(ab^*a)^*$. Try to give a sentence as simple (logically understandable) as possible. In particular don't follow the route of constructing an automaton and then converting it to a sentence.

9. Let L be a regular language. Show that the language

$$\{x \mid \exists y \ |y| = 2^{|x|}, \text{ and } xy \in L\}.$$

is also regular. Give an explicit construction.

10. Professor Jones proposes the following procedure for finding the canonical automaton for a language L starting from a DFA $\mathcal{A}$ for L.

    (a) Reverse the transitions in $\mathcal{A}$, make the initial state final, and the final states initial, to obtain an NFA $\mathcal{B}$ for $rev(L)$.

    (b) Determinize $\mathcal{B}$ using the subset construction to get $\mathcal{C}$.

    (c) Repeat the above two steps, starting from $\mathcal{C}$, to get respectively an NFA $\mathcal{D}$, and DFA $\mathcal{E}$.

    (d) Return $\mathcal{E}$ as the canonical automaton for $\mathcal{A}$.

    Is Jones' procedure correct? Justify your answer.

Questions 9 and 10 may be discussed with your classmates or TA's. Hand in written answers to all questions.