CSE322: Formal Models in Computer Science Spring 2006

PROBLEM SET 3
Due Friday, April 21, 2006, in class

Reading Assignment: Sipser's book, Sections 1.3 and 1.4.

Instructions: The basic instructions are the same as in Problem Set 1.

There are **SIX** questions in this assignment. One of them is a bonus question. Recall that all bonus questions are for extra credit.

Of course you are not required to think about the problems in the order they appear here. In particular, the order of appearance may not have any correlation with the order of relative difficulty.

Again, it never hurts to start working on the problems early.

---

1. (10 points) Prove that the following language is regular:

   $$L_1 = \{w \mid w \in \{0,1\}^* \text{ and contains equal number of occurrences of the substrings 01 and 10}\}.$$

   So for example, $101 \in L_1$ as it contains one occurrence of 01 and 10 while $1010 \notin L_1$ as it contains two occurrences of 10 and one of 01.

   (*Hint :* There might be more than one way of describing the strings in a language.)

   Compare this language with $L_2 = \{w \mid w \in \{0,1\}^*$ and contains the same number of 0s and 1s $\}$, which we will show in class is not regular. Note that at first glance, both L_1 and L_2 seem to require remembering more than a constant number of things (which might lead us to wrongly conclude that *both* of them are not regular languages). Hence, this example shows that sometimes our intuition can be wrong.

2. (10 points) (**Bonus**) Prove that $\mathsf{half}(L)$ is regular, where the operation is defined as follows:

   $$\mathsf{half}(L) = \{x \mid \text{ for some } y \text{ such that } |x| = |y|, xy \text{ belongs to } L\}.$$

   (For this problem, presenting a *correct* construction of an NFA/DFA (which should include a formal description), along with an informal yet convincing explanation of why the construction works will be sufficient. Fully formal proof of correctness using induction is *not* required.)

3. ($2 \times 10 = 20$ points) Give regular expressions for the following languages.

   (a) L_3, which is the language of all valid comments in the C language. Assume for this problem that a valid comment in C starts with /# and ends with #/ with no intervening #/. Assume for simplicity that the alphabet for L_3 is $\{/, \#, a, b\}$.
   For example, $/\#ab\#/$ and $/\#a/\#b\#/$ are in L_3 while $/\#ab$ and $/\#a\#/b\#/$ are not.

   (b) $L_4 = \{w \mid w \in \{0,1\}^*$ and w starts with a 1 and has odd length or starts with a 010 and has length that is a multiple of 3$\}$.

4. (*) (5 points) Do the following pair of regular expressions describe the same language ?

$$(10 \cup 1)^*1 \text{ and } 1(01 \cup 1)^*$$

   Prove your assertion.

5. (*) ($2 \times 10 = 20$ points) Using the construction covered in class (Lemma 1.55 in Sipser's text) convert the following regular expressions into NFAs. Show all the steps of your construction (like Example 1.56 in Sipser's text).

   (a) $(a^*b(b^*a)a^*b)^*$.

   (b) $(0 \cup 01)^* \cup (1 \cup 01)^*$.

6. (*) ($5+10 = 15$ points) Consider the DFA $M = \langle \{q_0, q_1, q_2, q_3\}, \{0,1\}, \delta, q_0, \{q_1\} \rangle$, with the following transition function: for $i = 0, 1, 2, 3$; $\delta(q_i, 0) = q_{(2i \bmod 4)}$ and $\delta(q_i, 1) = q_{((2i+1) \bmod 4)}$.

   (a) Let L_5 be the language accepted by M. Give a simple description of L_5.

   (b) Using the NFA to regular expression conversion procedure we covered in class, obtain a regular expression that describes L_5. Show your steps.