Problem Set 3

This problem set is due on **Wednesday February 14, by 5:00pm.**

Use the CS172 drop box.

Write **your name and your student ID number** on your solution. Write legibly. The description of your proofs should be as *clear* as possible (which does not mean *long* – in fact, typically, good clear explanations are also short.) Be sure to be familiar with the collaboration policy, and read the overview in the class homepage `www.cs.berkeley.edu/~luca/cs172`.

---

1. Define C to be all strings consisting of some positive number of 0's, followed by some string twice, followed again by some positive number of 0. For example 1100 is not in C, since it does not start with at least one 0. However 0001011010000000 is in C since it is three 0's, followed by 101 twice, followed by seven 0's. Prove that C is not regular.

2. Let A be the set of all binary strings which, when interpreted as a number with the most significant bit on the left, are divisible by 5. We know the language is regular from a previous homework. Construct an optimal DFA for A and prove its optimality by giving pairwise distinguishable strings, equal in number to the number of states in your DFA.

3. Consider the language $F = \{a^i b^j c^k \mid i, j, k \geq 0 \text{ and if } i = 1 \text{ then } j = k\}$.
   (a) Show that F acts like a regular language in the pumping lemma i.e. give a pumping length p and show that F satisfies the conditions of the lemma for this p.
   (b) Show that F is not regular.
   (c) Why is this not a contradiction?

4. Show that for any positive integer m, there exists a language A_m such that:
   (a) There is a DFA with m states which recognizes A_m.
   (b) No DFA with less than m states recognizes A_m.