Problem Set 1

This problem set is due on **Friday January 30, by 4:00pm.**

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. Prove that the following languages are regular, either by exhibiting a regular expression representing the language, or a DFA/NFA that recognizes the language:
   (a) the set of all words in the Oxford English dictionary, for $\Sigma = \{a, b, \ldots, z\}$
   (b) all strings that do not contain the substring aba, for $\Sigma = \{a, b\}$ (for instance, $aabaa$ contains the substring aba, whereas $abba$ does not)
   (c) all strings that do not contain 3 consecutive occurences of the same letter, for $\Sigma = \{a, b\}$
2. (Sipser, problem 1.24) For any string $w = w_1 w_2 \cdots w_n$, the reverse of w, written as w^R is the string w in reverse order, $w_n \cdots w_2 w_1$. For any language A, let $A^R = \{w^R \mid w \in A\}$. Show that if A is regular, so is A^R.
3. For any language A with alphabet Σ, let

   $$A^{sub} = \{w \in \Sigma^* \mid w \text{ is a substring of } x, \text{ for some } x \in A\}$$

   Show that if A is regular, so is A^{sub}.
4. Let k be a positive integer. Let $\Sigma = \{0, 1\}$, and L be the language consisting of all strings over $\{0, 1\}$ containing a 1 in the kth position from the end (in particular, all strings of length less than k are not in L).
   (a) Construct a DFA with exactly 2^k states that recognizes L.
   (b) Construct a NFA with exactly $k + 1$ states that recognizes L.