CS/ECE 374 A (Spring 2020)
Homework 4 (due Feb 20 Thursday at 10am)

Instructions: As in previous homeworks.

Problem 4.1: Prove that the following languages are not regular by using the fooling set method:

(a) $\{ww^Rw \mid w \in \{0,1\}^*\}$
(b) $\{(01)^{k^2+k} \mid k \geq 0\}$
(c) {all strings in $\{0,1\}^*$ that contain a palindrome of length at least 6 as a *prefix*}
(For example, **110011**01 is in the language, but 1110001101 is not.)

Problem 4.2:

(a) Prove that the following language is regular:

{all strings in $\{0,1\}^*$ that do not contain any palindrome of length at least 6 as a *substring*}

Don't give a DFA or NFA. Rather, use closure properties and basic facts about regular languages (for example, all finite languages are regular).
[Hint: if a string contains a long palindrome, wouldn't it contain a shorter one?]

(b) Prove that the following language is not regular:

{all imbalanced strings in $\{0,1\}^*$ that contains a balanced substring of length at least 6}

(Recall that a string is *balanced* iff it has the same number of 0's and 1's. A string is *imbalanced* iff it is not balanced.)
Don't use the fooling set method. Rather, use closure properties and a proof by contradiction. You may use the result from Problem 2.3. And you may also use the standard fact that $\{w \in \{0,1\}^* \mid \#_0(w) = \#_1(w)\}$ is not regular.

Problem 4.3: Give a context-free grammar (CFG) for each of the following languages in (a), (b), and (d). You must provide explanation for how your grammar works, by describing in English what is generated by each non-terminal. (Formal proofs of correctness are not required.)

(a) $\{w \in \{0,1\}^* \mid w \text{ is a palindrome of length at least } 6\}$.
(b) $\{0^i1^j0^k \mid k = 2i + 3j,\ i, j, k \geq 0\}$.
(c) Convert your grammar from (b) into Chomsky normal form.[1]
(d) $\{w \in \{0,1\}^* \mid w \text{ contains at least } |w|/2 \text{ consecutive 0's}\}$ (i.e., all strings of the form $x0^iy$ with $i \geq |x| + |y|$).

[1] See Section 5.8 of Jeff's book for the precise definition. To summarize, all production rules must be of the form $A \to BC$ or $A \to a$. For the start symbol S, the rule $S \to \varepsilon$ is allowed, but S should not appear on the right-hand side of any rule.