Write your answers in the separate answer booklet.

You have 120 minutes (after you get the answer booklet) to answer five questions.

Please return this question sheet and your cheat sheet with your answers.

1. For each statement below, check "Yes" if the statement is ***always*** true and check "No" otherwise, and give a brief (one short sentence) explanation of your answer. Read these statements very carefully—small details matter!

   For any string $w \in \{0,1\}^*$, let w^C denote the *bitwise complement* of w, obtained by replacing every 0 in w with a 1 and vice versa, and let w^R denote the *reversal* of w, obtained by writing the symbols of w in reverse order. For example, $\varepsilon^C = \varepsilon^R = \varepsilon$ and $000110^C = 111001$ and $000110^R = 011000$.

   (a) Every language is either regular or context-free.

   (b) Every regular expression that does not contain a Kleene star (or Kleene plus) represents a finite language.

   (c) For every regular language $L \subseteq \{0,1\}^*$, the language $\{w^C \mid w^R \in L\}$ also regular.

   (d) If L has a fooling set of size 374, then every DFA for L has at most 374 states.

   (e) $(100)^*$ is a fooling set for the language $\{w \in \{0,1\}^* \mid \#(0,w) = \#(1,w)\}$.

   (f) For any language L and any finite language L', the language $L \cap L'$ is regular.

   (g) The regular expressions $(00+11)^*$ and $(00)^*(11)^*$ represent the same language.

   (h) The language $\left\{0^a 0^b \mid a > b\right\}$ is regular.

   (i) If F is a fooling set for some irregular language, then F^* contains infinitely many strings.

   (j) For every context-free language L, the language L^* is also context-free.

2. Recall that a *run* in a string w is a maximal non-empty substring of w in which all symbols are equal. For any non-empty string $w \in \{0,1\}^*$ let $\mathsf{Delete1st}(w)$ denote the string obtained by deleting the first run in w. For example,

   $$\mathsf{Delete1st}(\underline{111111}) = \varepsilon, \qquad \mathsf{Delete1st}(\underline{000}110) = 110, \qquad \mathsf{Delete1st}(\underline{1}00110) = 00110.$$

   Let L be an arbitrary regular language over the alphabet $\Sigma = \{0,1\}$. ***Prove*** that the following languages are also regular.

   (a) $\textsc{Insert1st}(L) = \left\{w \in \Sigma^* \;\middle|\; w \neq \varepsilon \text{ and } \mathsf{Delete1st}(w) \in L\right\}$

   (b) $\textsc{Delete1st}(L) = \left\{\mathsf{Delete1st}(w) \;\middle|\; w \neq \varepsilon \text{ and } w \in L\right\}$

Questions 3, 4, and 5 are on the back of this page.

3. For each of the following languages over the alphabet {A, B, N}, describe both a regular expression that matches the language and a DFA that accepts the language. You do not need to prove that your answers are correct.

   (a) All strings in $\{A, B, N\}^*$ where every N is directly between two As.

   This language contains the strings BANANA and ABAAANANABBB but does not contain the strings ANBAN or BAANNANA or or KEVIN.

   (b) All strings in which B and N are never adjacent, and every run has length at most 2.

   This language contains the strings BAANNANA and NANAANNABANANABBABANANANAB,[1] but does not contain the strings ANBAN or ABAAANANABBB or STUART.

4. Recall that $\#(a, w)$ denotes the number of occurrences of symbol a in string w.

   Consider the following recursively defined function on strings over {0, 1}:

$$\text{grow}(w) = \begin{cases} \varepsilon & \text{if } w = \varepsilon \\ 1 \cdot \text{grow}(x) & \text{if } w = 0x \text{ for some string } x \\ 01 \bullet \text{grow}(x) & \text{if } w = 1x \text{ for some string } x \end{cases}$$

   For example, grow(10001) = 0111101.

   Prove that the following identity holds for all strings $w \in \{0, 1\}^*$:

$$|\text{grow}(\text{grow}(w))| \;=\; 2 \cdot \#(0, w) + 3 \cdot \#(1, w).$$

   As usual, you may assume any result proved in class, in the lecture notes, in labs, in lab solutions, or in homework solutions. In particular, you may use the fact that $|xy| = |x| + |y|$ and $\#(a, xy) = \#(a, x) + \#(a, y)$ for every symbol a and all strings x and y.

5. Let L be the set of all strings $w \in (\text{BAN} + \text{ANA})^*$ in which the substring BAN and ANA appear the same number of times. For example, L contains the strings BANANA and ANAANABANBAN and the empty string, but L does not contain the strings BANANABAN or BANA or BOB.

   (a) ***Prove*** that L is not a regular language.

   (b) Describe a context-free grammar for L.

[1] "Nana Anna, ban an ABBA banana nab", meaning "Hey, Grandmother Anna! Don't let that Swedish pop group steal plantains!"