

You have 120 minutes to answer five questions.

Write your answers in the separate answer booklet.

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

1. For each statement below, there are two boxes in the answer booklet labeled “Yes” and “No”. Check “Yes” if the statement is **always** true and “No” otherwise, and give a **brief** (at most one short sentence) explanation of your answer. For example:

• $x + y = 5$

Yes

No

Suppose $x = 3$ and $y = 4$.

- If $7 = 5$, then Emily has been to the moon.

Yes

No

The hypothesis is false, so the implication is true.

Read each statement *very* carefully; some of them are deliberately subtle!

- The regular expression $(0 + 1 + \varepsilon)(00 + 11)^*$ represents a language with an infinite fooling set.
 - The regular expression 0^*0^* represents the language $\{0^n \mid n \geq 0 \text{ and } n \text{ is even}\}$.
 - For any two languages L_1 and L_2 , if $L_1 \cap L_2$ is regular, then at least one of L_1 and L_2 is regular.
 - For any two languages L_1 and L_2 , if $L_1 \cap L_2$ is nonregular, then at least one of L_1 and L_2 is nonregular.
 - For every regular language L , there is a DFA accepting L whose start state is accepting.
 - For every regular language L , there is an NFA with ε -transitions accepting L whose start state is accepting.
 - For every regular language L , the language $(L \bullet L \bullet L)^R$ (i.e., the reversal of concatenating three copies of L) is regular.
 - Every context-free grammar has at most two non-terminals.
 - For any two regular language L_1 and L_2 , their union $L_1 \cup L_2$ is context-free.
 - The language $\{0^n 1^m 0^n \mid n, m \geq 1\}$ is context-free.
2. Let $L = \{0^a 1^b 2^c 3^d \mid b, c, d \geq 0 \text{ and } a > d \text{ and } b = c\}$ be a language over $\{0, 1, 2, 3\}$. For example, L contains the strings 0 , 00011223 , and 003 , but L does not contain the strings ε , 02 , or 0123 .
- Prove** that L is *not* a regular language.
 - Describe a context-free grammar for L . You do not need to prove your context-free grammar is correct.

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

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

$$\text{bubble}(w) := \begin{cases} \varepsilon & \text{if } w = \varepsilon \\ 0 \cdot \text{bubble}(x) & \text{if } w = 0 \cdot x \text{ for some string } x \\ 010 \cdot \text{bubble}(x) & \text{if } w = 1 \cdot x \text{ for some string } x \end{cases}$$

For example,

$$\text{bubble}(00\underline{1}0\underline{1}0) = 00\underline{01}0\underline{001}00$$

- (a) **Prove** that $\text{bubble}(x \bullet y) = \text{bubble}(x) \bullet \text{bubble}(y)$ for all strings x and y over $\{0, 1\}$.
 (b) **Prove** that $\#(0, \text{bubble}(w)) \geq 2 \cdot \#(1, w)$ for all strings w over $\{0, 1\}$.

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

4. For any string $w \in \{0, 1\}^*$, let $\text{Delete2nd}(w)$ denote the string obtained by deleting the second symbol in w . If $|w| < 2$, then $\text{Delete2nd}(w) = w$. For example,

$$\text{Delete2nd}(1) = 1, \quad \text{Delete2nd}(0\underline{0}110) = 0110, \quad \text{Delete2nd}(0\underline{1}0101) = 00101.$$

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

- (a) $\text{INSERTED2ND}(L) = \{w \in \Sigma^* \mid \text{Delete2nd}(w) \in L\}$
 (b) $\text{DELETED2ND}(L) = \{\text{Delete2nd}(w) \mid w \in L\}$

5. For each of the following languages L over the alphabet $\{A, B, C\}$, describe a DFA that accepts L and give a regular expression that represents L . **Give your DFA states mnemonic names.** You do not need to do more to prove your answers are correct.

- (a) All strings in which every instance of the substring **AB** is immediately followed by a **C**. This language contains the strings ε , **CBA**, and **ABCBCACAABC**, but it does not contain the strings **ABABC** nor **ABCAB**.
 (b) All strings in which every **B** is immediately **preceded** by a **A** and immediately **followed** by a **C** (i.e., every **B** belongs to an instance of the substring **ABC**.) This language contains the strings ε , **ABC**, and **ACAABCC**, but it does not contain the strings **ABCBCAAABC** nor **CAAAB**.