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.

Questions 2 and 4 were swapped in the question sheet—but not in the answer booklet(!)—that was distributed during the exam.

1. For each statement below, check "True" if the statement is always true and check "False" otherwise, and give a brief (one short sentence) explanation of your answer. Read these statements very carefully—small details matter!

   (a) Every irregular language is infinite.

   (b) The language $(0 + 1(01^*0)^*1)^*$ is context-free.

   (c) Every subset of a regular language is regular.

   (d) The language $\{0^a 1^b \mid a + b \text{ is divisible by } 374\}$ is regular.

   (e) If language L is regular, then L has a finite fooling set.

   (f) For every language L, if for every string $w \in L$ there is a DFA that accepts w, then L is regular.

   (g) If language L is accepted by an NFA with n states, then its complement $\Sigma^* \setminus L$ is also accepted by an NFA with n states.

   (h) 0^*1^* is a fooling set for the language $\left\{0^i 1^j 0^{i+j} \;\middle|\; i, j \geq 0\right\}$.

   (i) Every regular language is accepted by a DFA with an odd number of accepting states.

   (j) The context-free grammar $S \to 1T \mid T1 \mid \varepsilon;\; T \to 0S \mid S0$ generates all strings in which the number of 0s equals the number of 1s.

2. Recall that a *run* in a string w is a maximal non-empty substring of w in which all symbols are equal. For example, the string $\underbrace{0}\underbrace{111111}\underbrace{000}\underbrace{1}\underbrace{0000}$ consists of five runs.

   Let L be the set of all strings in $\{0, 1\}^*$ in which every run of 0s is followed immediately by a longer run of 1s. For example, the strings 000111111011 and 11101100011111 and 11111 are in L, but the strings 00000111 and 0111000000 are not.

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

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

3. For any string $w \in \{0,1\}^*$, let sortpairs(w) denote the string obtained by dividing w into pairs of symbols, sorting each pair into non-decreasing order, and leaving the last symbol if w has odd length. We can define sortpairs recursively as follows:

$$\text{sortpairs}(w) := \begin{cases} w & \text{if } w = \varepsilon \text{ or } w = 0 \text{ or } w = 1 \\ 01 \cdot \text{sortpairs}(x) & \text{if } w = 10x \text{ for some string } x \\ ab \cdot \text{sortpairs}(x) & \text{if } w = abx \text{ for some string } x \text{ and bits } a, b \text{ where } ab \neq 10 \end{cases}$$

For example,

$$\text{sortpairs}(\underbrace{00}\underbrace{10}\underbrace{11}\underbrace{10}\underbrace{01}1) = \underbrace{00}\underbrace{01}\underbrace{11}\underbrace{01}\underbrace{01}1$$

Recall that $\#(1, w)$ denotes the number 1s in the string w. For example, $\#(1, \varepsilon) = 0$ and $\#(1, 00101100011) = 5$.

(a) ***Prove*** that $\#(1, \text{sortpairs}(w)) = \#(1, w)$ for every string w.

(b) ***Prove*** that sortpairs(sortpairs(w)) = sortpairs(w) for every string w.

As usual, you can 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 $\#(1, xy) = \#(1, x) + \#(1, y)$ for all strings x and y.

4. Recall that a *run* in a string w is a maximal non-empty substring of w in which all symbols are equal. For example, the string $\underbrace{0}\underbrace{111111}\underbrace{000}\underbrace{1}\underbrace{0000}$ consists of five runs.

For any string $w \in \{0,1\}^*$, let compact(w) denote the string obtained by replacing each run with a single symbol from that run. For example, compact$(\varepsilon) = \varepsilon$ and

$$\text{compact}(011111100010000) = 01010.$$

Let L be an arbitrary regular language.

(a) ***Prove*** that the language $\textsc{Compact}(L) = \{\text{compact}(w) \mid w \in L\}$ is regular.

(b) ***Prove*** that the language $\textsc{Uncompact}(L) = \{w \in \Sigma^* \mid \text{compact}(w) \in L\}$ is regular.

5. For each of the following languages L over the alphabet $\Sigma = \{0,1\}$, describe a DFA that accepts L **and** give a regular expression that represents L. You do not need to justify your answers.

(a) Strings that do not contain the substring 01110.

(b) Strings that contain at least one odd-length run of 0s that is followed immediately by an even-length run of 1s.