Instructions

Homework should be done in groups of one to three people. You are free to change group members at any time throughout the quarter. Problems should be solved together, not divided up between partners. A single representative of your group should submit your work through Gradescope. Submissions must be received by 11:59pm on the due date, and there are no exceptions to this rule.

Homework solutions should be neatly written or typed and turned in through Gradescope by 11:59pm on the due date. No late homeworks will be accepted for any reason. You will be able to look at your scanned work before submitting it. Please ensure that your submission is legible (neatly written and not too faint) or your homework may not be graded. A typed PDF submission is recommended.

You may consult their textbook, class notes, lecture slides, instructors, TAs, and tutors for help with homework. You should not look for answers to homework problems in other texts or sources, including the internet. Only post about graded homework questions on Piazza if you suspect a typo in the assignment, or if you don’t understand what the question is asking you to do. Other questions are best addressed in office hours.

Your assignments in this class will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. You should always explain how you arrived at your conclusions, using mathematically sound reasoning. Whether you use formal proof techniques or write a more informal argument for why something is true, your answers should always be well-supported. Your goal should be to convince the reader that your results and methods are sound.

Reading Sipser Sections 1.2, 1.3, 1.4 (to page 79)

Key Concepts Regular expressions, equivalence between languages described by regular expressions and languages recognized by automata, non regular languages, pumping lemma.

1. (10 points) (Sipser 1.64) Let N be an NFA with k states that recognizes some language A
   a. Show that, if A is nonempty, A contains some string of length at most k.
   b. Show that, if \bar{A} is nonempty, \bar{A} contains some string of length at most 2^k.

   Note: By giving an example, can you show that part (a.) is not necessarily true when "A" is replaced by its complement, "\bar{A}". (Not for credit.)
2. (10 points) Write a regular expression describing each of the following sets.

a. \( \{w \in \{0, 1\}^* \mid w \text{ starts with a } 1 \text{ and has odd length, or starts with a } 0 \text{ and has even length}\} \)

b. \( \{w \in \{a, b\}^* \mid w \text{ is any string not in } a^*b^*\} \)

To ensure you receive full credit, please remember what we said at the beginning of this assignment: Your assignments in this class will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. **You should always explain how you arrived at your conclusions**, using mathematically sound reasoning. Whether you use formal proof techniques or write a more informal argument for why something is true, your answers should always be well-supported. Your goal should be to convince the reader that your results and methods are sound.

3. (10 points) Transform the following regular expression into an equivalent NFA using the procedure studied in class and described in the textbook

\[ ((a \cup ba)b)^* \]

Draw the resulting NFA using JFLAP. You can simplify your NFA to some extent, e.g., by removing some \( \varepsilon \)-transitions and unnecessary states. In your homework submission, include JFLAP images of state diagrams illustrating at least two intermediate steps in the transformation process.

**Note:** You can extend this problem to now transform the NFA you obtained into an equivalent 4-state DFA using the procedure studied in class and described in the textbook. (Not for credit.)

4. (10 points) Recall that a prime number is a positive integer greater than one whose only positive integer factors are one and itself. A composite integer is a positive integer greater than one which is not prime. Define the following languages over the unary alphabet \( \{1\} \)

\[ P = \{1^m \mid m \text{ is prime integer}\}, \]

\[ C = \{1^m \mid m \text{ is a composite integer}\}. \]

For example, the strings 1111, 1^6 and 1^8 are in \( C \); the strings \( \varepsilon, 1, 11, 1^3, 1^5 \) are not in \( C \); the strings \( \varepsilon, 1, 1111, 1^6 \) are not in \( P \). **Note:** the exponent here indicates concatenating the symbol 1 with itself a specified number of times; it is not algebraic exponentiation of the number 1.

Prove that the language \( C \) is not regular, using closure properties of the class of regular languages. As part of your proof, you may use the fact that the set \( P \) is not regular (you do not need to prove this fact; we will do this next week). You should not use the Pumping Lemma in your solution.

5. (10 points) A common misconception about regular languages is that if a language is regular, then its subsets or supersets must be too. In this question, we’ll show that this is false.

a. Give an example of a nonregular language \( A \) and a regular language \( B \) such that \( A \subseteq B \). Justify (either by proving or citing a fact proved in class or in the textbook) any claims you make about sets being regular or nonregular.

b. Give an example of a regular language \( A \) and a nonregular language \( B \) such that \( A \subseteq B \). Justify (either by proving or citing a fact proved in class or in the textbook) any claims you make about sets being regular or nonregular.