Midterm Exam IV

☞ **Due Sunday, November 15th at 12:00PM noon Pacific.** ☜

Instructions

You have 48 hours to complete this exam. Take as much of that time as you need. We've designed the exam with the expectation that it will take you around three hours to finish.

Please type your answers the way you type up your problem sets. There's a LaTeX template available on Canvas if you'd like to use it, though it's not required. Once you're finished, submit your answers on Gradescope. Please leave appropriate buffer time to ensure your submission comes in by the deadline. As with the problem sets, we'll grade the last version you submit before the deadline, so feel free to periodically submit what you have just in case something comes up.

Honor Code Policies

You are required to abide by the Honor Code policies outlined in the Honor Code Policies handout available on Canvas. We'd like to call particular attention to the following rules.

This midterm exam must be completed individually. It is a violation of the Stanford Honor Code to communicate with any other humans about this exam, to solicit solutions to this exam, or to share your solutions with others.

This exam is open-book, so you are free to make use of all course materials on Canvas. You are also permitted to search online for conceptual information (for example, by visiting Wikipedia). However, you are not permitted to communicate with other humans about the exam or to solicit help from others. For example, you ***must not*** communicate with other students in the course about the exam, and you ***must not*** ask questions on sites like Chegg or Stack Overflow. (You may ask questions to the course staff on Ed; if you do, you must post your questions privately.)

All work done with the assistance of any material in any way (other than provided CS103 course materials) must include a detailed citation (e.g., "I visited the Wikipedia page for *X* on Problem 1 and made use of insights *A*, *B*, and *C*"). ***Copying solutions is never acceptable***, even with citation, and is always a violation of the Honor Code. If by chance you encounter solutions to a problem, navigate away from that page before you feel tempted to copy. ***Because of the revise-and-resubmit policy, there is no reason to violate your conscience to complete a take-home exam.***

If you become aware of any Honor Code violations by any student in the class, your commitments under the Stanford Honor Code obligate you to inform course staff.

Grading

To earn a satisfactory grade on this exam, you need to earn a raw score of **90%** or above. If your score is lower than this, you will be asked to revise your answers and resubmit by the following Sunday at 12:00PM noon Stanford time. Course staff will be available to coach you on understanding where your work needs improvement and how to proceed.

<u>You can do this.</u> Best of luck on the exam!

Problem One: La La Land (4 Points)

This problem asks you to write regular expressions for a collection of languages. To do so, download the starter files for Midterm Exam 4 and extract them somewhere convenient. Write your regular expressions in the file `res/LaLaLand.regexes`, and use the provided program to test your regular expressions before submitting online via Gradescope.

In what follows, let $\Sigma = \{\texttt{L}, \texttt{A}\}$.

i. **(2 Points)** Write a regular expression for the language $\{ w \in \Sigma^* \mid w$ has an even number of copies of the substring `LA` $\}$. For example, the string `LALA` is in this language, as are the strings ε and `LLLAAALLLAAA`, but the string `ALA` is not in the language.

ii. **(2 Points)** Write a regular expression for the language $\{ w \in \Sigma^* \mid w$ has at most three copies of the substring `LA` $\}$. For example, `LALALA` is in this language, as are `ALA` and ε, but `LALALALALA` is not.

Problem Two: Powers and Subsets (8 Points)

Let A and B be languages over some alphabet Σ. Prove by induction that if $A \subseteq B$, then $A^n \subseteq B^n$ for all $n \in \mathbb{N}$. (As a reminder, we consider zero to be a natural number.) In your proof, please refer back to the formal definitions of language exponentiation and concatenation, as you saw in Problem Set Six:

$$L^0 = \{\varepsilon\} \qquad L^{n+1} = LL^n$$

$$L_1L_2 = \{\ w \in \Sigma^* \mid \exists x \in L_1.\ \exists y \in L_2.\ w = xy\ \}$$

Problem Three: Nonregular Languages (6 Points)

Below is a list of languages. For each of the indicated languages, tell us whether it's regular. If it's not regular, give an example of an infinite distinguishing set for the language, and tell us what strings you'd append to two arbitrarily-chosen distinct strings from the language to prove those strings are distinguishable. No justification is required. We've filled in one of these for you as an example.

{ $\mathtt{a}^n\mathtt{b}^n \mid n \in \mathbb{N}$ }

☐ This language is regular.

✓ This language is not regular. In particular:

S = { $\mathtt{a}^n \mid n \in \mathbb{N}$ } is an infinite distinguishing set for the language, because

if we pick two distinct strings $\mathtt{a}^n$ and $\mathtt{a}^m$ arbitrarily from S, we can append

$\mathtt{b}^m$ to distinguish them.

i. **(2 Points)** { $w \in \{\mathtt{a}, \mathtt{b}\}^* \mid w$ has the same number of as and bs }

☐ This language is regular.

☐ This language is not regular. In particular:

S = _____ is an infinite distinguishing set for the language, because

if we pick two distinct strings _____ and _____ arbitrarily from S, we can append

_____ to distinguish them.

ii. **(2 Points)** { $w \in \{\mathtt{a}, \mathtt{b}\}^*$ | no two consecutive characters in w are the same, and w has the same number of as and bs }

☐ This language is regular.

☐ This language is not regular. In particular:

S = _____ is an infinite distinguishing set for the language, because

if we pick two distinct strings _____ and _____ arbitrarily from S, we can append

_____ to distinguish them.

iii. **(2 Points)** { $w \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^*$ | no two consecutive characters in w are the same, and w has the same number of as and bs }

☐ This language is regular.

☐ This language is not regular. In particular:

S = _____ is an infinite distinguishing set for the language, because

if we pick two distinct strings _____ and _____ arbitrarily from S, we can append

_____ to distinguish them.