18.404/6.5400 Michael Sipser
Theory of Computation
Due via Gradescope at $\boxed{\text{noon}}$ sharp, Thursday, September 19, 2024 **Problem Set 1**

Instructions: Read the rules for collaboration, course bibles, etc., on the Course Information sheet (see the course homepage). Submit your solutions on Gradescope as described on the homepage. Usually, we will have presented all the material you need to solve each problem set by a week before its deadline.

Read all of Chapters 1 and 2 except Section 2.4.

0.1 Read and solve, <u>but do not turn in</u>: Book, 1.14 . [Swapping NFA accept/non-accept states]

0.2 Read and solve, <u>but do not turn in</u>: Book, 1.31 . [Closure under reversal]

0.3 Read and solve, <u>but do not turn in</u>: Book, 1.46b . [Pumping lemma]

0.4 Read and solve, <u>but do not turn in</u>: Book, 2.4 and 2.5 . [Practice with CFLs]
Note, problems marked with ᴬ have solutions at the end of the chapter.

You may assume the solutions to the above problems when solving the problems below.

1. For every $b \geq 2$, let $\Sigma_b = \{\texttt{0}, \texttt{1}, \ldots, b-1\}$ where each member of Σ_b is considered to be a single symbol. Consider a string $s \in \Sigma_b^*$ to be a number written in base b. The empty string represents the number 0. Let $MFIVE_b = \{s \in \Sigma_b^* |\ s \text{ represents some multiple of 5 in base } b\}$. For example, $\texttt{1010} \in MFIVE_2$ and $\texttt{120} \in MFIVE_3$ and $\texttt{45} \in MFIVE_{10}$.
   (a) Show that $MFIVE_3$ is regular by giving the state diagram of a DFA M_3 that recognizes it. (Hint: Simulate long division. You need only 5 states.)
   (b) Show that $MFIVE_d$ is regular for each $d \geq 2$ by giving the formal description of a DFA M_d that recognizes it.

2. The *Hamming distance* $H(x, y)$ between two strings x and y of equal length, is the number of corresponding symbols at which x and y differ. For example, $H(\texttt{1101111}, \texttt{0001111}) = 2$. For any language A, let $N_1(A) = \{w|\ H(w, x) \leq 1 \text{ for some } x \in A\}$.
Show that the class of regular languages is closed under the N_1 operation.

3. Let $D = \{w|\ w \in \{\texttt{0}, \texttt{1}\}^* \text{ is not a palindrome (i.e., } w \neq w^{\mathcal{R}})\}$. Prove that D is not regular.

4. Let $\Sigma = \{\texttt{0}, \texttt{1}\}$.
   (a) Let $TUT = \{tut|\ t, u \in \Sigma^*\}$. Show TUT is regular.
   (b) Let $TUTU = \{tutu|\ t, u \in \Sigma^*\}$. Show $TUTU$ is not regular.

5. Let x and y be strings over some alphabet Σ. Say x is a ***substring*** of y if $y \in \Sigma^* x\, \Sigma^*$.
Define the ***avoids*** operation for languages A and B to be

$$A \textit{ avoids } B = \{w|\ w \in A \text{ and } w \text{ doesn't contain any string in } B \text{ as a substring}\}.$$

Show that the class of regular languages is closed under the *avoids* operation.
(Hint: Theorems we've previously shown may be helpful.)

6. Let M_1 and M_2 be DFAs that have k_1 and k_2 states, respectively, where $L(M_1) \neq L(M_2)$. Using an argument similar to the proof of the pumping lemma, show that there is a string s where $|s| \leq k_1 k_2$ and where exactly one of M_1 and M_2 accepts s.

7.⋆ (⋆ means optional) Improve the bound in Problem 6 to show that such an s exists where $|s| \leq k_1 + k_2$.