“B problems” must be turned in.

Problem B1 (80 pts). This problem illustrates the power of the congruence version of Myhill-Nerode.

Recall that the reversal of a string, \( w \in \Sigma^* \), is defined inductively as follows:

\[
\epsilon^R = \epsilon \\
(ua)^R = au^R,
\]

for all \( u \in \Sigma^* \) and all \( a \in \Sigma \).

(1) Let \( \sim \) be a congruence (on \( \Sigma^* \)) and assume that \( \sim \) has \( n \) equivalence classes. Define \( \sim_R \) and \( \approx \) by

\[ u \sim_R v \iff u^R \sim v^R, \text{ for all } u, v \in \Sigma^* \text{ and } \approx = \sim \cap \sim_R. \]

Prove that the relation \( \approx \) is a congruence and that \( \approx \) has at most \( n^2 \) equivalence classes.

(2) Given any regular language \( L \) over \( \Sigma^* \) let

\[ L^{(1/2)} = \{ w \in \Sigma^* \mid ww^R \in L \}. \]

Prove that \( L^{(1/2)} \) is also regular using the relation \( \approx \) of part (1).

(3) Let \( L \) be any regular language over some alphabet \( \Sigma \). For any natural number \( k \geq 2 \), let

\[ L^{(1/k)} = \{ w \in \Sigma^* \mid (ww^R)^{k-1} \in L \} = \{ w \in \Sigma^* \mid \underbrace{ww^Rww^R \cdots ww^R}_{k-1} \in L \}. \]

Also define the languages

\[ L^{1/\infty} = \{ w \in \Sigma^* \mid (ww^R)^{k-1} \in L, \text{ for all } k \geq 2 \}, \text{ and } \]

\[ L^\infty = \{ w \in \Sigma^* \mid (ww^R)^{k-1} \in L, \text{ for some } k \geq 2 \}. \]
Prove that every language \( L^{(1/k)} \) is regular.

(4) Prove that there are only finitely many distinct languages of the form \( L^{(1/k)} \) (this means that the set of languages \( \{ L^{(1/k)} \}_{k \geq 2} \) is finite). Prove that \( L^{1/\infty} \) and \( L^\infty \) are regular.

**Problem B2 (100 pts).** Which of the following languages are regular? Justify each answer.

1. \( L_1 = \{ wcw \mid w \in \{a, b \}^* \} \). (here \( \Sigma = \{a, b, c\} \)).
2. \( L_2 = \{ xy \mid x, y \in \{a, b \}^* \text{ and } |x| = |y| \} \). (here \( \Sigma = \{a, b\} \)).
3. \( L_3 = \{ a^n \mid n \text{ is a prime number} \} \). (here \( \Sigma = \{a\} \)).
4. \( L_4 = \{ a^m b^n \mid gcd(m, n) = 23 \} \). (here \( \Sigma = \{a, b\} \)).
5. Consider the language \( L_5 = \{ a^{4n+3} \mid 4n + 3 \text{ is prime} \} \).

Assuming that \( L_5 \) is infinite, prove that \( L_5 \) is not regular.

6. Let \( F_n = 2^{2^n} + 1 \), for any integer \( n \geq 0 \), and let
\[
L_6 = \{ a^{F_n} \mid n \geq 0 \}.
\]

Here \( \Sigma = \{a\} \).

**Extra Credit (from 10 up to 100 pts).** Find explicitly what \( F_0, F_1, F_2, F_3 \) are, and check that they are prime. What about \( F_4 \)?

Is the language \( L_7 = \{ a^{F_n} \mid n \geq 0, F_n \text{ is prime} \} \) regular?

**Extra Credit (20 pts).** Prove that there are infinitely many primes of the form \( 4n + 3 \).

The list of such primes begins with
\[
3, 7, 11, 19, 23, 31, 43, \cdots
\]

Say we already have \( n + 1 \) of these primes, denoted by
\[
3, p_1, p_2, \cdots, p_n,
\]

where \( p_i > 3 \). Consider the number
\[
m = 4p_1 p_2 \cdots p_n + 3.
\]

If \( m = q_1 \cdots q_k \) is a prime factorization of \( m \), prove that \( q_j > 3 \) for \( j = 1, \ldots k \) and that no \( q_j \) is equal to any of the \( p_i \)'s. Prove that one of the \( q_j \)'s must be of the form \( 4n + 3 \), which
shows that there is a prime of the form $4n + 3$ greater than any of the previous primes of the same form.

**Problem B3 (80 pts).** The purpose of this problem is to get a fast algorithm for testing state equivalence in a DFA. Let $D = (Q, \Sigma, \delta, q_0, F)$ be a deterministic finite automaton. Recall that state equivalence is the equivalence relation $\equiv$ on $Q$, defined such that,

$$p \equiv q \iff \forall z \in \Sigma^* (\delta^*(p, z) \in F \iff \delta^*(q, z) \in F),$$

and that $i$-equivalence is the equivalence relation $\equiv_i$ on $Q$, defined such that,

$$p \equiv_i q \iff \forall z \in \Sigma^*, |z| \leq i (\delta^*(p, z) \in F \iff \delta^*(q, z) \in F).$$

A relation $S \subseteq Q \times Q$ is a forward closure iff it is an equivalence relation and whenever $(p, q) \in S$, then $(\delta(p, a), \delta(q, a)) \in S$, for all $a \in \Sigma$.

We say that a forward closure $S$ is good iff whenever $(p, q) \in S$, then $\text{good}(p, q)$, where $\text{good}(p, q)$ holds iff either both $p, q \in F$, or both $p, q \notin F$.

Given any relation $R \subseteq Q \times Q$, recall that the smallest equivalence relation $R_\approx$ containing $R$ is the relation $(R \cup R^{-1})^*$ (where $R^{-1} = \{(q, p) \mid (p, q) \in R\}$, and $(R \cup R^{-1})^*$ is the reflexive and transitive closure of $(R \cup R^{-1})$). We define the sequence of relations $R_i \subseteq Q \times Q$ as follows:

$$R_0 = R_\approx,$$

$$R_{i+1} = (R_i \cup \{(\delta(p, a), \delta(q, a)) \mid (p, q) \in R_i, a \in \Sigma\})_\approx.$$

(1) Prove that $R_{i_0 + 1} = R_{i_0}$ for some least $i_0$. Prove that $R_{i_0}$ is the smallest forward closure containing $R$.

*Hint.* First, prove that

$$R_i \subseteq R_{i+1}$$

for all $i \geq 0$. Next, prove that $R_{i_0}$ is forward closed.

If $\sim$ is any forward closure containing $R$, prove by induction that

$$R_i \subseteq \sim$$

for all $i \geq 0$.

We denote the smallest forward closure $R_{i_0}$ containing $R$ as $R^\dagger$, and call it the forward closure of $R$.

(2) Prove that $p \equiv q$ iff the forward closure $R^\dagger$ of the relation $R = \{(p, q)\}$ is good.
Hint. First, prove that if $R^i$ is good, then

$$R^i \subseteq \equiv .$$

For this, prove by induction that

$$R^i \subseteq \equiv_i$$

for all $i \geq 0$.

Then, prove that if $p \equiv q$, then

$$R^i \subseteq \equiv .$$

For this, prove that $\equiv$ is an equivalence relation containing $R = \{(p,q)\}$ and that $\equiv$ is forward closed.

**TOTAL: 260 points + 30 points**