You have 90 minutes to complete this exam. You may assume without proof any statement proved in class.

1. Find a regular expression for each of the following languages over \{0, 1\}:

   (1 pt) a. strings containing exactly two 0s;
   (1 pt) b. strings that begin or end with 00 or 11;
   (2 pts) c. strings containing both 11 and 010 as substrings;
   (2 pts) d. strings not containing the substring 110.
2 Prove or disprove:

(2 pts) a. if $L$ is regular and $L'$ is nonregular, then $L \cup L'$ is nonregular;

(3 pts) b. if $L$ is nonregular and both of $L'$ and $L \cap L'$ are regular, then $L \cup L'$ is nonregular.
Let $L_1$ and $L_2$ be languages recognized by DFAs with $n_1$ and $n_2$ states, respectively. Define $L = L_1 \cup L_2$. Prove that if $L$ is nonempty, then $L$ contains a string of length less than $\max(n_1, n_2)$.

Let $L$ be a nonempty language in which the shortest string has length $k$. Prove that $L$ cannot be recognized by a DFA with fewer than $k + 1$ states.
For each of the following languages over \{0, 1\}, determine whether it is regular, and prove your answer:

(a) nonpalindromes;  
(b) odd-length strings with middle symbol 0;  
(c) strings that contain a substring of the form \(wuw\) where \(u \in \{0, 1\}^*\) and \(w \in \{0, 1\}^+\);  
(d) strings with the property that in every prefix, the number of 0s and the number of 1s differ by at most 2.
SOLUTIONS
You have 90 minutes to complete this exam. You may assume without proof any statement proved in class.

1 Find a regular expression for each of the following languages over \{0, 1\}:

(1 pt) a. strings containing exactly two 0s;
(1 pt) b. strings that begin or end with 00 or 11;
(2 pts) c. strings containing both 11 and 010 as substrings;
(2 pts) d. strings not containing the substring 110.

**Solution.**

a. \(1^*01^*01^*\)

b. \((00 \cup 11) \Sigma^* \cup \Sigma^*(00 \cup 11)\)

c. \(\Sigma^*11\Sigma^*010\Sigma^* \cup \Sigma^*010\Sigma^*11\Sigma^*\)

d. \(0^*(10^+)^*1^*\)
Prove or disprove:

(2 pts)   a. if $L$ is regular and $L'$ is nonregular, then $L \cup L'$ is nonregular;

(3 pts)   b. if $L$ is nonregular and both of $L'$ and $L \cap L'$ are regular, then $L \cup L'$ is nonregular.

Solution.

a. False. Take $L = \Sigma^*$ and let $L'$ be any nonregular language. Then $L \cup L' = \Sigma^*$ is regular.

b. True. We will prove the following contrapositive form: if $L'$, $L \cap L'$, and $L \cup L'$ are regular, then $L$ is also regular. This claim follows from the closure of regular languages under difference because

$$L = (L \cup L') \setminus (L' \setminus (L \cap L')).$$
Let $L_1$ and $L_2$ be languages recognized by DFAs with $n_1$ and $n_2$ states, respectively. Define $L = L_1 \cup L_2$. Prove that if $L$ is nonempty, then $L$ contains a string of length less than $\max(n_1, n_2)$.

Solution. By the argument in the pumping lemma, any DFA with $n$ states that accepts some string accepts some string of length less than $n$. This is because any accepted string of length at least $n$ must visit some state more than once and can therefore be shortened.

Now, let $L = L_1 \cup L_2$ be nonempty. Consider two cases.

- If $L_1$ is nonempty, then by the first paragraph $L_1$ contains a string of length less than $n_1$, the number of states in its DFA.
- Analogously, if $L_2$ is nonempty, then it contains a string of length less than $n_2$.

Either way, $L$ contains a string of length less than $\max(n_1, n_2)$.

Let $L$ be a nonempty language in which the shortest string has length $k$. Prove that $L$ cannot be recognized by a DFA with fewer than $k + 1$ states.

Solution. Take any string $w = w_1 w_2 \ldots w_k$ in $L$. We claim that the $k + 1$ strings

$$
\varepsilon, \quad w_1, \quad w_1 w_2, \quad w_1 w_2 w_3, \quad \ldots, \quad w_1 w_2 w_3 \ldots w_k
$$

are each in a different equivalence class of $\equiv_L$. Indeed, for any $i < j$,

$$
(w_1 w_2 \ldots w_i)(w_{j+1} \ldots w_k) \notin L,
$$

$$
(w_1 w_2 \ldots w_j)(w_{j+1} \ldots w_k) \in L,
$$

where the first line holds because that string is shorter than $k$.

Since $\equiv_L$ has at least $k + 1$ equivalence classes, the Myhill-Nerode theorem implies that any DFA for $L$ must have at least $k + 1$ states.
For each of the following languages over \( \{0, 1\} \), determine whether it is regular, and prove your answer:

(a) nonpalindromes; 
(b) odd-length strings with middle symbol 0; 
(c) strings that contain a substring of the form \( uwu \) where \( u \in \{0, 1\}^* \) and \( w \in \{0, 1\}^+ \); 
(d) strings with the property that in every prefix, the number of 0s and the number of 1s differ by at most 2.

**Solution.**

In each part, \( L \) stands for the language in question.

(a) Nonregular. We showed in class that palindromes are a nonregular language. Since nonregular languages are closed under complement, nonpalindromes are also a nonregular language.

(b) Nonregular. We claim that the infinite collection of strings \( 1^{2i-1}, i = 1, 2, 3, \ldots \) are each in a different equivalence class of \( \equiv_L \). Indeed, for \( i < j \), the language contains \( 1^{2i-1}0^{2i} \) but not \( 1^{2j-1}0^{2j} \). By the Myhill-Nerode theorem, \( L \) is nonregular.

(c) Regular. The property of containing \( uwu \) for some nonempty \( w \) is equivalent to the property of having some symbol occur at least twice. Thus \( L \) is given by the regular expression \( L = \Sigma^*0\Sigma^*0\Sigma^* \cup \Sigma^*1\Sigma^*1\Sigma^* \), making it regular.

(d) Regular. The language is given by the following DFA:

![DFA Diagram]