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 at least two 0s;

   (1 pt) b. strings that do not end with 01;

   (2 pts) c. strings containing both 101 and 010 as substrings;

   (2 pts) d. strings not containing the substring 000.

Solution.

   a. \(\Sigma^*0\Sigma^*0\Sigma^*\)

   b. \(\varepsilon \cup 0 \cup 1 \cup \Sigma^*(00 \cup 10 \cup 11)\)

   c. \(\Sigma^*(101\Sigma^*010 \cup 010\Sigma^*101 \cup 1010 \cup 0101)\Sigma^*\)

   d. \(1^*(01^+ \cup 001^+)\)*(\(\varepsilon \cup 0 \cup 00)\)
2. Prove or disprove:

(2 pts) a. there exists a nonregular language $L$ such that $L^*$ is regular;

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

Solution.

a. True. We showed in class that the language $L = \{n^2 : n \geq 0\}$ over alphabet $\{0, 1\}$ is nonregular. Since $0, 1 \in L$, we have $L^* = \{0, 1\}^*$. Therefore, $L^*$ is regular.

b. False. Take any nonregular language $L'$, and let $L = \Sigma^*$. Then $L$ is regular, $L'$ and $L \cap L'$ are nonregular, but $L \cup L' = \Sigma^*$ is regular.
(3 pts) 3 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 $L \neq \Sigma^*$, then $L$ excludes some string of length less than $n_1n_2$.

**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, consider $L = L_1 \cup L_2$. By the Cartesian-product construction from class, $L$ has a DFA with $n_1n_2$ states. Swapping the accept and reject states in that DFA gives a DFA for $\overline{L}$ with $n_1n_2$ states. By the first paragraph, this means that $\overline{L}$ must contain a string of length less than $n_1n_2$ if $\overline{L}$ is nonempty.

(3 pts) 4 Let $L$ be any language. Prove that the equivalence class of $\equiv_L$ that contains $\varepsilon$ either contains no other strings, or contains infinitely many strings.

**Solution.** Assuming $w \equiv_L \varepsilon$ for some nonempty string $w$, we will show that the infinite collection of strings $\varepsilon, w, w^2, w^3, \ldots$ all belong to the same equivalence class.

Since $w \equiv_L \varepsilon$, we know the following for all $v$:

$$wv \in L \iff v \in L.$$

In particular, for every $i$ and every string $u$,

$$w(w^i u) \in L \iff w^i u \in L.$$

By definition, this means that $w^{i+1} \equiv_L w^i$ for all $i$. 
5 For each of the following languages over \{0, 1\}, determine whether it is regular, and prove your answer:

(a) strings in which the number of 0s and the number of 1s are both divisible by 5;
(b) even-length strings of length at least 2 with the two middle symbols equal;
(c) strings that contain a nonempty substring of the form \(ww\) for some \(w \in \{0, 1\}^*\);
(d) strings with the property that in some prefix, the number of 0s is 3 more than the number of 1s.

**Solution.**

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

(a) Regular. \(L\) is the intersection of the regular languages \((0^*10^*)^5\) and \(((1^*01^*)^5\)^*, and is therefore regular by the closure of regular languages under intersection.

(b) Nonregular. We claim that the infinite collection of strings \((01)^i\), \(i = 1, 2, 3, \ldots\) are each in a different equivalence class of \(\equiv_L\). Indeed, for \(i < j\), the language contains \((01)^i(00)^j\) but not \((01)^j(00)^j\). By the Myhill-Nerode theorem, \(L\) is nonregular.

(c) Regular. By direct inspection, \(L\) contains every string of length 4. Thus, \(L\) contains every string of length at least 4. Of strings shorter than 4, the language contains only 00, 11, \(\Sigma00, 00\Sigma\), \(\Sigma11, 11\Sigma\). In summary, \(L\) is given by the regular expression \(L = \Sigma^*(00 \cup 11)\Sigma^* \cup \Sigma^4\Sigma^*\) and is therefore regular.

(d) Nonregular. We claim that the infinite collection of strings \(1^i\), \(i = 1, 2, 3, \ldots\), are each in a different equivalence class of \(\equiv_L\). Indeed, for \(i < j\), the language contains \(1^i0^{i+3}\) but not \(1^j0^{i+3}\). By the Myhill-Nerode theorem, the language is nonregular.