1. Find a regular expression for each of the following languages:

a. binary strings that contain both 01 and 10 as substrings; (2 pts)

b. strings over the alphabet \{a, b, c\} that do not contain ab as a substring; (2 pts)

c. the complement of the language 0^*1^*2^*...9^* over the alphabet \{0, 1, 2, ..., 9\}. (2 pts)

Solution:

a. \((0^+1^+0 \cup 1^+0^+1)\Sigma^*\)

b. \((a^*c \cup b)^*a^*\)

c. \(\bigcup_{i>j} \Sigma^*ij\Sigma^*\), where the union is over all \(i, j \in \{0, 1, 2, ..., 9\}\) with \(i > j\).

(1 pt) 2. Simplify the regular expression \((00 \cup 01 \cup 10 \cup 11)^*0^*1^*\) as much as possible, and justify your answer.

Solution:

\[(00 \cup 01 \cup 10 \cup 11)^*0^*1^* = (\Sigma\Sigma)^*0^*1^* \supseteq (\Sigma\Sigma)^*(\Sigma \cup \varepsilon) = \Sigma^*\.

So, the answer is \(\Sigma^*\).
3 Let $L$ be the set of odd-length binary strings in which the middle symbol also occurs elsewhere in the string. Use the pumping lemma to prove that $L$ is nonregular. You must not use the Myhill–Nerode theorem or closure properties.

**Solution.** Let $p > 1$ be arbitrary. Consider the string $w = 0^p10^{p^2+p} \in L$. Let $w = xyz$ be any decomposition such that $y$ is nonempty and occurs within the first $p$ zeroes. Then pumping in an additional $p!/[|y|]$ copies of $y$ shifts the “1” into the middle position. Thus, $xyy^{p!/|y|}z \notin L$. By the pumping lemma, $L$ is nonregular.

4 Determine the equivalence classes of $\equiv_L$ for the language $L = \{0^n1^n : n \geq 0\}$ over the binary alphabet.

**Solution:**

- $\{0^i\}$ for $i = 0, 1, 2, \ldots$;
- $\{0^{n+i}1^n : n \geq 1\}$ for $i = 0, 1, 2, \ldots$;
- all other strings are lumped together in one class, $\{0^n1^m : n \geq m\}$.

Any two classes in the first group can be distinguished from one another by a string of all 1s; likewise for the second group. Any class in the first group can be distinguished from any class in the second group by a string of the form 011\ldots1. Finally, the strings in $\{0^n1^m : n \geq m\}$ form a separate class because they are precisely those strings that cannot be made into a string of $L$ by appending any suffix.
5 Construct a DFA for $\Sigma 0\Sigma^*1$ with the smallest possible number of states, where $\Sigma = \{0, 1\}$. Prove that your DFA is the smallest possible.

**Solution.** The following DFA with five states recognizes $L = \Sigma 0\Sigma^*1$:

![DFA Diagram]

By the Myhill–Nerode theorem, no smaller DFA exists for $L$ because each of the five strings $\varepsilon, 0, 00, 01, 001$ is in a different equivalence class of $\equiv_L$. Their distinguishing suffixes are as follows:

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon$</th>
<th>0</th>
<th>00</th>
<th>01</th>
<th>001</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00</td>
<td>1 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>001 01 01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>001</td>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
</tr>
</tbody>
</table>

6 True or false? Prove your answer.

(2 pts) **a.** If $A \cap B$, $B \cap C$, and $A \cap C$ are regular languages, then $A \cup B \cup C$ is regular as well.

(2 pts) **b.** If $LL$ is a nonregular language, then $L^*$ is nonregular as well.

**Solution.**

**a.** False. Let $A$ be any nonregular language, and let $B = C = \emptyset$. Then the pairwise intersections $A \cap B = B \cap C = A \cap C = \emptyset$ are regular, but the union $A \cup B \cup C = A$ is nonregular.

**b.** False. Consider the language $L = \{0, 1\} \cup \{0^n1^n : n \geq 0\}$. Then $L^* = \{0, 1\}^*$, which is a regular language. We will now show that $LL$ is nonregular. For any positive integers $i \neq j$, we have $0^i1^i0^j1^j \in LL$ but $0^i1^i0^j1^i \notin LL$. Therefore, each of the strings $0, 0^2, 0^3, 0^4, \ldots$ is in a different equivalence class of $\equiv_{LL}$. Since there are infinitely many equivalence classes, $LL$ is nonregular by the Myhill–Nerode theorem.
For each of the following languages, determine whether it is regular and prove your answer:

(2 pts) a. strings of the form $0^n1^m0^n$ for some $n, m \geq 0$;
(2 pts) b. strings of the form $uv$, where $u, v \in \{0, 1\}^*$ are palindromes;
(2 pts) c. binary strings in which the number of 0s and the number of 1s differ by an integer multiple of 2017.

Solution.

In each part, $L$ stands for the language in question.

a. Nonregular. For any positive integers $i \neq j$, we have $0^i1^i0^i \in L$ but $0^j1^i0^i \notin L$. Therefore, each of the strings $0, 0^2, 0^3, 0^4, \ldots$ is in a different equivalence class of $\equiv_L$. Since there are infinitely many equivalence classes, $L$ is nonregular by the Myhill–Nerode theorem.

b. Nonregular. For any positive integers $i \neq j$, we have $0^j110^i1 = (0^i110^i)1 \in L$ but $0^j110^i1 \notin L$. Therefore, each of the strings $0, 0^2, 0^3, 0^4, \ldots$ is in a different equivalence class of $\equiv_L$. Since there are infinitely many equivalence classes, $L$ is nonregular by the Myhill–Nerode theorem.

c. Regular. A DFA for this language simply keeps track of the difference between the number of 0s and the number of 1s, performing all calculations modulo 2017. Formally, the DFA is $(\{0, 1, 2, \ldots, 2016\}, \{0, 1\}, \delta, 0, \{0\})$, where

$$
\delta(q, \sigma) = \begin{cases} 
(q + 1) \mod 2017 & \text{if } \sigma = 0, \\
(q - 1) \mod 2017 & \text{if } \sigma = 1.
\end{cases}
$$