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

   (2 pts)  
   a. strings of length at most 2015 over a given alphabet $\Sigma$;  

   (3 pts)  
   b. binary strings in which the number of 0s and the number of 1s are not both odd.

Solution.

   a. $(\Sigma \cup \epsilon)^{2015}$  
   b. $1^*(01*01^*)^* \cup 0^*(10*10^*)^*$. There is a hard way to solve this problem and an easy one.  
   The hard way is to construct an NFA and convert it to a regular expression. The easy way is to observe that this language is the union of two simpler languages: (i) binary strings with an even number of 0s, and (ii) binary strings with an even number of 1s.
(2 pts) 2 Simplify $(0 \cup 1)^*01(0 \cup 1)^* \cup 1^*0^*$ as much as possible, and explain your answer in detail.

**Solution.** The regular expression $(0 \cup 1)^*01(0 \cup 1)^*$ corresponds to binary strings that contain 01. The only strings that do not have this property are of the form $1^*0^*$. Thus, the union of $(0 \cup 1)^*01(0 \cup 1)^*$ and $1^*0^*$ simplifies to $(0 \cup 1)^*$.

(3 pts) 3 Prove or disprove: the Kleene star of every language is a regular language.

**Solution.** The claim is false. Consider $L = \{0^n1^n : n \geq 0\}$. Then $L^*$ is the set of strings of the form $0^{n_1}1^{n_1}0^{n_2}1^{n_2} \cdots 0^{n_k}1^{n_k}$ for some $n_1, n_2, \ldots, n_k$. Let $p$ be arbitrary and consider the string $w = 0^p1^p \in L^*$. If $x, y, z$ are strings such that $y$ is nonempty, $|xy| \leq p$, and $w = xyz$, then $xy^2z \notin L^*$. Therefore by the pumping lemma, $L^*$ is nonregular.
Let $L$ be the language of palindromes over $\{0, 1\}$. Determine the equivalence classes of $\equiv_L$.

**Solution.** Let $u, v$ be arbitrary strings with $u \neq v$. Then for $N$ sufficiently large,

$$u \ 01^N \ 0 u^R \in L,$$
$$v \ 01^N \ 0 u^R \notin L.$$

Thus, every string is in an equivalence class by itself.

*Note.* A common mistake is to claim that $uu^R \in L, v u^R \notin L$ for any pair of distinct strings $u, v$. This claim fails for many string pairs, e.g., $0, 00$ as well as $0, 01010101$.

Construct a DFA for $0^*1^*0^+$ with the smallest possible number of states. Prove that your DFA is the smallest possible.

**Solution.** The following DFA with five states recognizes $0^*1^*0^+$:

The five strings $\varepsilon, 0, 1, 10, 101$ are in pairwise distinct equivalence classes of $\equiv_{0^*1^*0^+}$, with distinguishing suffixes given by the following table:

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$0$</th>
<th>$1$</th>
<th>$10$</th>
<th>$101$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
<td>$010$</td>
<td>$010$</td>
<td></td>
</tr>
<tr>
<td>$0$</td>
<td>$\varepsilon$</td>
<td>$10$</td>
<td>$10$</td>
<td>$10$</td>
</tr>
<tr>
<td>$1$</td>
<td>$010$</td>
<td>$010$</td>
<td>$10$</td>
<td>$10$</td>
</tr>
<tr>
<td>$10$</td>
<td>$10$</td>
<td>$10$</td>
<td>$10$</td>
<td>$10$</td>
</tr>
<tr>
<td>$101$</td>
<td>$0$</td>
<td>$0$</td>
<td>$0$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

By the Myhill-Nerode theorem, we conclude that no smaller DFA exists.
For each of the following languages, determine whether it is regular, and prove your answer:

(2 pts) a. binary strings in which the number of 1s is a multiple of the number of 0s;
(2 pts) b. binary strings of the form $uvuw$, where $u, v, w$ are strings and $u$ is nonempty;
(2 pts) c. nonempty binary strings of even length with the two middle symbols unequal;
(2 pts) d. strings over the alphabet $\{a, b, c\}$ that contain each of the three alphabet symbols.

Solution.

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

a. Nonregular. For any positive integers $i < j$, we have $0^i 1^i \in L$ but $0^j 1^j \notin L$. Therefore, the strings $0, 00, 000, \ldots, 0^n, \ldots$ are each in a distinct equivalence class of $\equiv_L$. By the Myhill-Nerode theorem, $L$ is nonregular.

b. Regular. This language contains precisely those strings in which the first symbol occurs again, which corresponds to the regular expression $1 \Sigma^* 1 \Sigma^* \cup 0 \Sigma^* 0 \Sigma^*$.

c. Nonregular. For any even positive integers $i \neq j$, we have $0^i 1^i \in L$ but $0^j 1^j \notin L$. Therefore, the strings $0^2, 0^4, \ldots, 0^{2n}, \ldots$ are each in a distinct equivalence class of $\equiv_L$. By the Myhill-Nerode theorem, $L$ is nonregular.

d. Regular. The language is given by $\{a, b\}^* \cup \{a, c\}^* \cup \{b, c\}^*$. Thus, it is obtained by applying the operations of Kleene star, union, and complement to the regular languages $\{a, b\}, \{a, c\}, \{b, c\}$. Since regular languages are closed under these operations, the result is regular.