1. Provide a regular expression to describe the language of the following NFA

\[ x^*y((xUy)y^*x(xy^*)^*)^* \]

2. For each of the two regular expressions below, (i) write two strings that are members of the language described by that expression and (ii) draw an NFA or DFA (your choice) to accept that string. Use the language \( \Sigma = \{4,5,6\} \)

a. \( 5^*(64^* U 5)6 \)

\[ 56, 66, 556, 566, 646, 5556, 5646, 6446, ... \]

b. \( (5 U 64)^* 456^* \)

\[ 45, 456, 545, 6454, 5545, 55545, 54566, 544545, 555545, 5555456, 554566, 545666, ... \]
3 Use the pumping lemma to show that the following languages are not regular.

a. \( A_1 = \{ x^{k+5}y^k \mid k \geq 1 \} \)

Assume \( A_1 \) is regular.

\[ w \in A_1 \text{ and } |w| > p : w = x^{p+5}y^px^5 \]

Divide \( w \) into \( xyz \): \( x = x^n \ y = x^m \ z = x^{p+5-(n+m)}y^px^p \) \( m > 0, n+m \leq p \)

Try pumping middle string: \( xy^2z = x^n x^{2m} x^{p+5-(n+m)}y^px^p = x^{p+5+m}y^px^p \)

First set of \( x \)'s has \( p+5+m \) characters, second set has \( p \) characters. Difference in number of \( x \)'s is: \( p+5+m - p = 5+m \). Since \( m > 0 \), difference is MORE than \( 5+m \), so \( xy^2z \) is not in language, thus a string \( w \) greater than the pumping length is not pumpable, thus language is not regular.

b. \( A_2 = \{ b^{n!} \mid n \geq 1 \} \) (Note \( b^1 = b \), \( b^2 = bb \), \( b^3 = bbbbb \), \( b^4 = b^{12} \) ...)

Assume \( A_2 \) is regular.

\[ w \in A_2 \text{ and } |w| > p : w = b^{p!} \]

Divide \( w \) into \( xyz \): \( x = b^n \ y = b^m \ z = b^{p!-(n+m)}y^pb^m \) \( m > 0, n+m \leq p \)

Try pumping middle string: \( xy^2z = b^n b^{2m} b^{p!-(n+m)}y^pb^m = b^{p!+m}y^pb^m \)

Number of \( b \)'s in pumped \( w \) is \( p!+m \) \( (p+1)! = (p+1) \times p! = p \times p! + p! \) \( \text{ (presuming } p > 1, \ p! > 1 \) \)

Since \( 0 < m \leq p \), \( p! + m \leq p! + p < p! + p! \)

Thus, number of \( b \)'s is smaller than the next factorial number of \( b \)'s. So \( xy^2z \) is not in language, thus a string \( w \) greater than the pumping length is not pumpable, thus language is not regular.

c. \( A_3 = \{ a^nb^mc^o \mid o, n, m \geq 0 \text{ and } o = n-m \} \)

Assume \( A_3 \) is regular.

\[ w \in A_3 \text{ and } |w| > p : w = a^p b^m c^{p-m} \text{ for any value of } m \text{ and pumping length } p \]

Divide \( w \) into \( xyz \): \( x = a^d \ y = a^f \ z = a^{p-(d+f)}b^m c^{p-m} \) \( f > 0, d+f < p \)

Try pumping middle string: \( xy^2z = a^d a^{2f} a^{p-(d+f)}b^m c^{p-m} = a^{p+f}b^m c^{p-m} \)

Number of \( a \)'s is \( p+f \), number of \( b \)'s is \( m \), number of \( c \)'s is \( p-m \)

Number of \( c \)'s is supposed to be \( p+f-m \). \( p+f-m < p-m \), so \( xy^2z \) is not in language.

Thus a string \( w \) greater than the pumping length is not pumpable, thus language is not regular.

d. \( A_4 = \{ 0110^R \mid w \in \{0,1\}^* \} \) Note, \( w^R \) is the reverse of \( w \), if \( w = 01001 \), \( w^R = 10010 \)
Incorrect question – this language has no w in its definition. I meant to write: $w1w^R$ for this question.

4. Provide a regular expressions generating each of the languages below, presuming the alphabet $\Sigma = \{+, -, <, >\}$
   a. \{w \mid w \text{ contains less than 2 } - \text{ symbols} \} \\
      (+U<U>\)* U ((+U<U>\)*-(+U<U>\)*)

   b. \{w \mid w \text{ does not contain the substring } \rightarrow \} \\
      ((+U<U>\)* U (+U-U<)* U (-(<U+>))*\)*

   c. \{w \mid \text{each < is followed by + in w} \} \\
      ((-U+U>\)*(<+))*

   d. \{w \mid \text{the length of w is exactly 6} \} \\
      $\Sigma\Sigma\Sigma\Sigma\Sigma\Sigma$

5. The pumping lemma says that every regular language has a pumping length $p$, such the every string in the language can be pumped if it has length $p$ or more. If $p$ is a pumping length for language $A$, so is any length $p' \geq p$ that is a pumping length for $A$. For Example, if $A=01^*$, the minimum pumping length is 2. The reason is that the string $s=0$ is in $A$ and has length 1 yet $s$ cannot be pumped; but any string in $A$ of length 2 or more contains a 1 and hence can be pumped by dividing it so that $x=0$, $y=1$, and $z$ is the rest. For each of the following languages, give the minimum pumping length and justify you answer.

   a. $(010)^*01$ 
      01001 is the smallest string that can be pumped.

   b. $01^*$ 
      2 01 is the smallest string that can be pumped.

   c. $11^*(00^*)1$ 
      3 111 is the smallest string that can be pumped.

   d. 10011
6. Only one string in language. Specify pumping length to be greater than that string, so all strings in language \( \geq p \) can be pumped (no such strings exist, so this statement is vacuously true).

6. Convert the following grammar to be in Chomsky Normal Form.

\[
S \rightarrow Boo \mid SC \\
B \rightarrow mB \mid \varepsilon \\
C \rightarrow oCo \mid n
\]

<table>
<thead>
<tr>
<th>( S_0 \rightarrow S )</th>
<th>( S \rightarrow Boo \mid SC )</th>
<th>( S_0 \rightarrow S \rightarrow Boo \mid SC \mid oo )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S \rightarrow mB \mid \varepsilon )</td>
<td>( B \rightarrow mB \mid m )</td>
<td>( B \rightarrow mB \mid m )</td>
</tr>
<tr>
<td>( C \rightarrow oCo \mid n )</td>
<td>( C \rightarrow oCo \mid n )</td>
<td>( C \rightarrow oCo \mid n )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( S_0 \rightarrow B )</th>
<th>( S \rightarrow Boo \mid SC \mid U_0U_0 )</th>
<th>( S_0 \rightarrow B )</th>
<th>( S \rightarrow Boo \mid SC \mid U_0U_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S \rightarrow B ) ( U_0 ) ( U_0 \mid SC \mid U_0U_0 )</td>
<td>( S_0 \rightarrow Boo \mid SC \mid U_0U_0 )</td>
<td>( S_0 \rightarrow Boo \mid SC \mid U_0U_0 )</td>
<td>( S \rightarrow Boo \mid SC \mid U_0U_0 )</td>
</tr>
<tr>
<td>( B \rightarrow mB \mid m )</td>
<td>( B \rightarrow mB \mid m )</td>
<td>( B \rightarrow mB \mid m )</td>
<td>( B \rightarrow mB \mid m )</td>
</tr>
<tr>
<td>( C \rightarrow oCo \mid n )</td>
<td>( C \rightarrow oCo \mid n )</td>
<td>( C \rightarrow oCo \mid n )</td>
<td>( C \rightarrow oCo \mid n )</td>
</tr>
<tr>
<td>( U_0 \rightarrow o )</td>
<td>( U_0 \rightarrow o )</td>
<td>( U_0 \rightarrow o )</td>
<td>( U_0 \rightarrow o )</td>
</tr>
<tr>
<td>( U_m \rightarrow m )</td>
<td>( U_m \rightarrow m )</td>
<td>( U_m \rightarrow m )</td>
<td>( U_m \rightarrow m )</td>
</tr>
<tr>
<td>( D \rightarrow U_0U_0 )</td>
<td>( D \rightarrow U_0U_0 )</td>
<td>( D \rightarrow U_0U_0 )</td>
<td>( D \rightarrow U_0U_0 )</td>
</tr>
<tr>
<td>( E \rightarrow CU_0 )</td>
<td>( E \rightarrow CU_0 )</td>
<td>( E \rightarrow CU_0 )</td>
<td>( E \rightarrow CU_0 )</td>
</tr>
</tbody>
</table>

Final answer:

\[
S_0 \rightarrow B \ D \mid SC \mid U_0U_0 \\
S \rightarrow B \ D \mid SC \mid U_0U_0 \\
B \rightarrow U_mB \mid m \\
C \rightarrow U_0E \mid n \\
U_0 \rightarrow o \\
U_m \rightarrow m \\
D \rightarrow U_0U_0 \\
E \rightarrow CU_0
\]