Problem 3.1. (10 points) Give a regular expression for the complement of the language defined by $0^*(01 \cup 1)^*$. 

Hint: Translate the regular expression into/from a finite automata and use operations on finite automata.

Solution: 
First, we build an NFA that accepts the language. Then, we remove $\varepsilon$ transitions so as to have an automaton that we can determinize.

Applying the subset construction algorithm for the determinization, we get the following DFA. Complementing the DFA is simply done by swapping accepting and non-accepting states. Finally, we transform the automaton by adding a new initial state and a single accepting state.

We can now build the regular expression corresponding to the last automaton using the state elimination algorithm seen during the lecture.

Remove $q_{0,2}$

Remove $\emptyset$
Problem 3.2. (20 points) For each of the following three languages, either show that the language is regular by giving a finite automaton (or a regular expression), or prove that the language is not regular using the pumping lemma. In both cases you have the freedom to use the closure properties of regular languages. You may assume that the language of all words over the alphabet Σ = {0, 1} that contain the same number of 0's and 1's is nonregular (we proved this in class).

a. The set $A_1$ of all words over the alphabet Σ = {0, 1} that contain the same number of the substrings 00 and 11. For example, 100110000111 ∈ $A_1$.

b. The set $A_2$ of all words over the alphabet Σ = {0, 1} that contain the same number of the substrings 01 and 10. For example, 110101 ∈ $A_2$.

c. The set $A_3$ of all words over the alphabet Σ = {0, 1, #} that have the form

$$u_1u_2\ldots u_kw_1#v_1v_2\ldots v_m#w_1w_2\ldots w_n,$$

for binary numbers $u = u_1\ldots u_k$ and $v = v_1\ldots v_m$ and $w = w_1\ldots w_n$ such that $w = u + v$. For example, 010#0011#101 ∈ $A_3$.

Solution:

a. Not Regular

We propose two different ways to show that $A_1$ is nonregular. The first solution uses the pumping property. The second solution uses the closure property of regular languages.

- We show that language $A_1$ does not have the pumping property.

1. Consider any arbitrary $p > 0$.
2. Choose $w = 0^p1^{2p}$, $w \in L_1$.
3. Consider $w = xyz$ any pumping decomposition of $w$ with $|xy| \leq p$, $|y| > 0$.
   It is clear that $y$ must have the form $0^j$, $j > 0$.
4. Choose $i = 2$ (to pumping once).
5. We get $xy^2z = 0^{2p+j}1^{2p} \notin L_1$. Therefore $L_1$ is not regular.
• Let’s assume that $A_1$ is regular. We know that $0^*1^*$ is regular. Since regular languages are closed under intersection, $A_1 \cap 00^*11^*$ should be regular. However, $A_1 \cap 00^*11^* = \{0^n1^n \mid n \in \mathbb{N}\}$, which, as shown in class, is nonregular. We therefore get to a contradiction which forces us to reject the assumption that $A_1$ is regular.

b. : Regular. The following DFA accepts $A_2$:

![DFA Diagram]

The DFA is based on the fact that an accepted word has to start and end with the same symbol in order to have the same number of changes from 0 to 1 (resulting in a substring 01) and from 1 to 0 (resulting in a substring 10). Note that the number of substrings 10 and 01 may differ by at most one in any word whereas in the previous question where the difference between the number of substrings 00 and 11 could be anything.

c. : Not Regular.

We use the pumping lemma.

Consider any arbitrary $p \geq 1$ and choose $w = 1^p 0^p 1^p$, $w \in L_3$.

Let $w = xyz$ be any pumping decomposition of $w$ with $|xy| \leq p$, $|y| > 0$.

It is clear that $y$ must have the form $1^j$, $j > 0$. Pumping once, (choosing $i = 2$ in the pumping lemma) we get $1^{p+j} 0^p 1^p \notin L_3$. Thus, $L_3$ is not regular.