DUE: Tuesday, September 3

1. (9 points) For each of the following languages over \{0,1\}, produce a DFA to accept it.
   
   a) \( L_0 = \{ z \mid |z| \leq 3 \} \).
   
   b) \( L_1 \) = the set of all strings with exactly one occurrence of 00. That is, 1011001101 \( \in L_1 \) but 10110001101 \( \notin L_1 \).
   
   c) \( L_0 \cap \sim L_1 \).

2. (10 points) For any language \( L \), define the language
   
   \[
   \text{NotPrefix}(L) = \left\{ w \mid (w \in L) \land (\exists x \, wx \in L) \rightarrow (x = \varepsilon) \right\}
   \]

   So if \( L = \{0^n \mid n \geq 0\} \), then \( \text{NotPrefix}(L) = \emptyset \), and if \( L = \{\varepsilon, 0101, 01, 11\} \), then
   \( \text{NotPrefix}(L) = \{0101, 11\} \). Show that if \( L \) is regular, then \( \text{NotPrefix}(L) \) must be regular.

3. (8 points) For each of the following conjectures, state whether it is true or false and justify your answer.
   
   a) **Conjecture**: For any languages \( L_0 \) and \( L_1 \), \( L_0 \cup L_1 \) if and only if \( L_0^* = L_1^* \).
   
   b) **Conjecture**: For any \( \Sigma \) and any \( L \in \Sigma^* \), if \( L \) is infinite then \( \sim L \) is finite.
1. a) $L_0$ is accepted by the DFA $M_0 = \left(\{q_0, q_1, q_2, q_{\text{fail}}\}, \{0, 1\}, \delta, q_0, \{q_0, q_1, q_2, q_{\text{fail}}\}\right)$, where being in state $q_i, 0 \leq i \leq 3$, "means" that $i$ input symbols have been read.

\[
\begin{array}{c|cc}
\delta & 0 & 1 \\
\hline
q_0 & q_1 & q_1 \\
q_1 & q_2 & q_2 \\
q_2 & q_3 & q_3 \\
q_3 & q_{\text{fail}} & q_{\text{fail}} \\
q_{\text{fail}} & q_{\text{fail}} & q_{\text{fail}} \\
\end{array}
\]

b) $z \in L_1$ if and only if one instance of 00 has been read and then every 0 that follows is either the last character in $z$ or is followed by a 1. In the following acceptor for $L_1$, states $q_0$ and $q_1$ "mean" that no instance of 00 has been read, and states $q_2$ and $q_3$ mean that exactly one instance of 00 has been read. If the last symbol read was 0 then the machine is in $q_4$ or $q_3$, and if the last symbol read was 1 then the machine is in $q_0$ or $q_2$. If more than one instance of 00 has been read, then the machine is in state $q_{\text{fail}}$.

\[
M_1 = \left(\{q_0, q_1, q_2, q_3, q_{\text{fail}}\}, \{0, 1\}, \delta, q_0, \{q_2, q_3\}\right)
\]

\[
\begin{array}{c|cc}
\delta & 0 & 1 \\
\hline
q_0 & q_1 & q_0 \\
q_1 & q_2 & q_0 \\
q_2 & q_3 & q_2 \\
q_3 & q_{\text{fail}} & q_2 \\
q_{\text{fail}} & q_{\text{fail}} & q_{\text{fail}} \\
\end{array}
\]

c) $L_0 \cap \sim L_1$ is the set of all strings of length at most 3 which have either no occurrences of 00 or at least one occurrence of 00. So $L_0 \cap \sim L_1 = \{\varepsilon, 0, 1, 01, 10, 11, 000, 010, 011, 101, 110, 111\}$

Since we don’t care about a minimal number of states, we can design the machine with state $q_z$ denoting that input string $z$ has been read, and state $q_{\text{fail}}$ denoting that the input which has been read is not a prefix of any string in our language. Language $L_0 \cap \sim L_1$ is accepted by the DFA $\left(\{z \mid |z| \leq 3\} \cup \{q_{\text{fail}}\}, \{0, 1\}, \delta, q_0, \{q_z \mid z \in L_0 \cap \sim L_1\}\right)$, where $L_0 \cap \sim L_1$ is as listed above, and with transition function $\delta(q_z, a) = \begin{cases} q_{z+a}, & \text{if } |z| < 3 \\ q_{\text{fail}}, & \text{otherwise} \end{cases}$ for $z \in \{0, 1\}^*$, $a \in \{0, 1\}$, and $\delta(q_{\text{fail}}, 0) = \delta(q_{\text{fail}}, 1) = q_{\text{fail}}$. 
2. If \( L \) is regular, then there is a DFA \( M = (Q, \Sigma, \delta, s, F) \) such that \( L(M) = L \). Construct \( F^* \) as the set of all states \( q \in F \) for which there is a path in \( M \) from \( q \) to a state of \( F \). Removing all states of \( F^* \) from \( F \) yields a machine which accepts \( \text{NotPrefix}(L) \).

Claim: \( M^* = (Q, \Sigma, \delta, s, \{q \in F \mid q \notin F^*\}) \) accepts \( \text{NotPrefix}(L) \).

Proof: We need to show that for each of the following three cases and for any \( z \in \Sigma^* \), \( (z \in \text{NotPrefix}(L)) \iff (z \in L(M^*)) \).

Case 1 \( \tilde{\delta}(s, z) \notin F \)

So \( z \notin \text{NotPrefix}(L) \). Since \( \{q \in F \mid q \notin F^*\} \subseteq F \), then \( \tilde{\delta}(s, z) \notin \{q \in F \mid q \notin F^*\} \).

Case 2 \( \tilde{\delta}(s, z) \in F^* \)

By the definition of \( F^* \), there must be a nonempty \( x \) leading from \( \tilde{\delta}(s, z) \) to a state of \( F \). That is \( \tilde{\delta}(\tilde{\delta}(s, z), x) \in F \) for some \( x \neq \epsilon \). So \( \tilde{\delta}(s, z) \notin \text{NotPrefix}(L) \), and \( z \notin L(M^*) \).

Case 3 \( \tilde{\delta}(s, z) \in \{q \in F \mid q \notin F^*\} \)

Clearly \( z \in L \), and by the definition of \( M^* \), \( z \in L(M^*) \). But since \( z \) is not a proper prefix of any string belonging to \( L \), it must follow that \( z \in \text{NotPrefix}(L) \).

3. a The conjecture is false. If \( L_0 = \{0\} \) and \( L_1 = \{0\}^* \), then \( L_0 \neq L_1 \) but \( L_0^* = L_1 = L_1^* \).

b The conjecture is false. Choose \( \Sigma = \{0, 1\} \) and \( L \) as the binary strings ending in 0, which contains the even natural numbers and hence is infinite. Language \( \sim L \) contains the odd binary numbers and hence is infinite also. The reason for the word “contains” above is that these languages also contain binary numbers with leading 0’s. But, clearly, any set containing an infinite set must be infinite.