Due: September 24, 2009 in class or 1:00 pm (via email to cs361-hw@cs.williams.edu)

1. DFA (20 points). Draw the fully-specified state diagram of a DFA that recognizes the following language $L$ over the alphabet $\Sigma = \{0, 1\}$:

   $L = \{ w \mid w$ begins with 0 and does not contain the substring 010$\}$.

2. Regexp (20 points). Give a regular expression describing the following language $L$ over the alphabet $\Sigma = \{0, 1\}$:

   $L = \{ w \mid w$ contains exactly two 0s separated by an even number of 1s$\}$.

3. Regexp from DFA (30 points). Give a regular expression describing the language recognized by the following DFA over the alphabet $\Sigma = \{0, 1\}$:

   ![State Diagram]

4. egrep (30 points). Give egrep commands that will exactly match each of the following languages (hint: type man egrep at a Linux or Mac terminal prompt):

   a) $L = \{ w \mid w$ is a Visa, MasterCard, or Discover credit card number$\}$ over the alphabet $\Sigma = \{0, 1, \ldots, 9\}$. (Visa card numbers start with 4 and have either 13 or 16 digits; MasterCard numbers start with the numbers 51 through 55 and have 16 digits; Discover card numbers begin with 6011 or 65 and have 16 digits.)

   b) $L = \{ w \mid w$ is a date in YYYYMMDD format between 19000101 to 20991231$\}$ over the alphabet $\Sigma = \{0, 1, \ldots, 9\}$. (You may assume that February always has 28 days.)

   c) $L = \{ w \mid w$ is an email address$\}$ over the alphabet $\Sigma = \{0, 1, \ldots, 9, a, b, \ldots, z, A, B, \ldots, Z, @, \}$ (You may assume that the last field of the domain name of the address is 2 or 3 letters long, such as com or uk.)