1. (20 pts) Finite Automata
   (a) Prove, by induction on the length of a string \( x \), that in the construction of a DFA from an NFA, \( \delta(q_0, x) = \{a, b, c, \ldots\} \) for the NFA if and only if \( \delta(q_0, x) = \{a, b, c, \ldots\} \) for the DFA.
   (b) Write an NFA for the set of all integers (the set of strings with alphabet \( \Sigma = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 0\} \)) which start with a different numeral than they end with.
   (c) Modulus calculator: We are used to the binary system, but other numeral systems exist, for instance, the "tertiary", or "base-3" system. Write your own modulus-5 calculator by creating an NFA which accepts the set of all ternary strings which do not start with a 0 and which are divisible by 5. (For example, your NFA should not accept "1210" because in decimal it is 93, but should accept the strings "101" and "120", which are 10 and 15, respectively, in the decimal system.)

2. (20 pts) Write regular expressions for:
   (a) The set of strings over alphabet \( \{a, b, c\} \) containing at least one \( a \) and at least two \( b \)'s.
   (b) The set of strings of \( 0 \)'s and \( 1 \)'s whose tenth symbol from the right is a 1.
   (c) The set of strings of \( 0 \)'s and \( 1 \)'s with at most one pair of consecutive \( 1 \)'s.
   (d) All strings of \( a \)'s and \( b \)'s such that at every point in the string, the number of \( b \)'s minus the number of \( a \)'s is zero, one, or two.
   (e) The set of all strings of \( 0 \)'s and \( 1 \)'s where the number of \( 0 \)'s is divisible by 5.
   (f) The set of strings of \( 0 \)'s and \( 1 \)'s where the number of \( 0 \)'s is divisible by 5 and the number of \( 1 \)'s is even.

3. (10 pts) Protein Synthesis: RNA, the DNA-based chains that encode templates for protein synthesis, can be viewed as very long strings formed from a four-nucleotide alphabet. Each RNA strand is a series of codons, which are "mini-strings" of length 3. Each codon has a specific meaning. Most represent amino acids, but some are "stop codons", which indicate the end of a protein template. One such stop codon is \( UAA \).
   A valid protein template over the alphabet \( \Sigma \) is a string of codons (the length is a multiple of 3) and in which the stop codon \( UAA \) occurs exactly once, as the last codon in the string. To simplify the problems, define the alphabet for RNA as a length-3 alphabet consisting of \( \Sigma = \{U, G, A\} \). (If you want, you can also use \( n \) to represent the expression \( (U + G + A); \text{e.g. } nn^* = (U + G + A)(U + G + A)^* \).
   (a) Write a regular expression which evaluates to exactly the set of valid protein templates.
   (b) Write a regular expression which represents the set of all strings over the alphabet \( \Sigma = \{U, G, A\} \) which are NOT valid protein templates.

4. (10 pts) Interweaving is an important technique that we will be using extensively. In this technique, a set of "weak" constraints are combined so that mutually, they enforce a much stronger constraint.
   (a) Given the set \( L = \{a^kba^{k+1}b | k \geq 1\} \), use copies of \( L \) modified by concatenation, union, intersection, and/or Kleene closure with other modified "copies" of \( L \) or with regular expressions, to produce a language which represents the set of all strings of \( a \)'s and \( b \)'s such that blocks of \( a \)'s are separated by exactly one \( b \), there are an even number of blocks of \( a \)'s, and each block of \( a \)'s has exactly one more \( a \) than the last; that is, of the form \( \{abaaba^3b...a^n b | n \text{ even and } n \geq 0\} \). (for example, your set should contain \( e, abaab, abaaabab, \) etc.)
   (b) Use the example strings \( s_1 = abaaaba^3b, s_2 = abaabaab, \) and \( s_3 = abaaabaab \) and show how the elements of your construction enforce different constraints so that \( s_1 \) is accepted but \( s_2 \) and \( s_3 \) are not. (This does NOT need to be a formal proof; just give intuitive explanations of how the elements of your construction work.)