CPS 230 Homework-6

Write the solution to each problem on a single page of a separate sheet of paper. The deadline for handing in solutions is November 25th.

1. **String Matching Problem:** (20 points)
   Give a linear-time algorithm to determine if a text $T$ is a cyclic rotation of another string $T'$. For example, *arc* and *car* are cyclic rotation of each other.

2. **String Matching Problems 2:** (20 = 12 + 8 points)
   A scattered substring of a string $y = b_1b_2\ldots b_n$ is a string $x = a_1a_2\ldots a_m$ with $m \leq n$ such that $a_i = b_j$, for some $1 \leq j_1 < j_2 < \ldots < j_m \leq n$. For example, 12345 is a scattered substring of the string 1ds2j34muy5dy.
   
   (a) Give an algorithm that decides whether or not $x$ is a scattered substring of $y$ in time $O(n)$.
   
   (b) Modify your algorithm to find the scattered substring that minimized $j_m - j_1$ in time $O(n^2)$.

3. **Pattern Matching Problems:** (20 = 12 + 8 points)
   Remember that $I$, $V$, $X$, $L$ are the symbols for 1, 5, 10, 50 in roman numbers. We can construct roman numbers between 1 and 88 using only these four symbols.
   
   (a) Give a regular expression $A$ for roman numbers between 1 and 88.
   
   (b) Draw the directed graph whose language is the set of strings defined by $A$.

4. **Pattern Matching Problems 2:** (20 = 10 + 10 points)
   How would you express the following additional symbols in the specification of a regular expression?
   
   (a) The wild card ‘?’ that indicates a position of the pattern that can be occupied by any character.
   
   (b) The negation $\bar{a}$ that prevents a position to be occupied by $a$, but any other character is allowed.

5. **Pattern Matching Problem 3:** (20 = 5 \times 4 points)
   Write the regular expression for the following string patterns, the alphabet $\Sigma = \{0, 1\}$:
   
   (a) The string contains 0101 as its substring.
   
   (b) The length of the string is at least three (less than or equal to three), and the symbol in the third position is 0.
   
   (c) The string contains at least two 0’s, and at most one 1’s (it contains no 1 or exactly one 1).
   
   (d) The string contains even number of 0’s, or it contains exactly two 1’s.