Algorithms **Monsoon 2018**
Questions for Quiz 1 **28th August**

Your Quiz will have 3 questions, one from each of three sections given here.
Please come with sufficient papers to write your answers.
No books, smart phones are allowed. Maximum marks is 30.

Section A

1. (10 points) Consider the problem of determining a DFA and a regular expression are equivalent. Express the problem as a language and show that it is decidable.
2. (10 points) Let $ALL_{DFA} = \{\langle A \rangle |\ A \text{ is a DFA and } L(A) = \Sigma^*\}$. Show that ALL_{DFA} is decidable.
3. (10 points) Let $\mathbf{C}_{CFG} = \{\langle G, k \rangle |\ G \text{ is a CFG and } L(G) \text{ contains exactly } k \text{ strings where } k \geq 0 \text{ or } k = \infty\}$. Show that $\mathbf{C}_{CFG}$ is decidable.
4. (10 points) Let $A = \{\langle R, S \rangle | R \text{ and } S \text{ are regular expressions and } L(R) \subseteq L(S)\}$. Show that A is decidable.

Section B

1. (a) (5 points) Show that the solution of $T(n) = T(\lceil \frac{n}{2} \rceil) + 1$ is $O(\log n)$.
   (b) (5 points) Use master method to show that the solution to the binary-search recurrence $T(n) = T(\frac{n}{2}) + \Theta(1)$ is $T(n) = \Theta(\log n)$.
2. (a) (5 points) Show that the solution of $T(n) = 2T(\lfloor \frac{n}{2} \rfloor + 17) + n$ is $O(n \log n)$.
   (b) (5 points) Use a recursion tree to determine a good asymptotic upper bound on the recurrence $T(n) = T(n-1) + T(\frac{n}{2}) + n$. Use substitution method to verify your answer.
3. (a) (5 points) Use master method to give tight asymptotic bounds for $T(n) = 2T(\frac{n}{4}) + n^2$.
   (b) (5 points) Use a recursion tree to give an asymptotically tight solution to the recurrence $T(n) = T(n-a) + T(a) + cn$ where $a \geq 1$ and $c > 0$ are constants.

Section C

1. (a) (5 points) Generalize Huffman's algorithm to ternary codewords (i.e., codewords using the symbols 0,1 and 2), and prove that it yields optimal ternary codes.
   (b) (5 points) Show how to find the maximum spanning tree of a graph, that is, the spanning tree of largest total weight.
2. (a) (5 points) Let $G = (V, E)$ be a weighted graph with a distinguished vertex s and all edge weights are positve and distinct. Is it possible for a tree of shortest paths from s and a minimum spanning tree in G to not share any edges? If so, give an example. If not, give reason.

(b) (5 points) Suppose the symbols a, b, c, d, e occur with frequencies $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}$ respectively. What is the Huffman coding of the alphabet? If this encoding is applied to a file consisting of 1,000,000 characters with the given frequencies, what is the length of the encoded file in bits.

3. (a) (5 points) Suppose that a data file contains a sequence of 8-bit characters such that all 256 characters are about equally common: the maximum character frequency is less than twice of the minimum character frequency. Prove that Huffman coding in this case is no more efficient than using an ordinary 8-bit fixed-length code.

   (b) (5 points) Prove or disprove Prim's algorithm works correctly when there are negative edges.

Books for reference:

1. Introduction to Theory of Computation, by Micheal Sipser.
2. Introduction to Algorithms, by CLRS.
3. Algorithms, by S.Dasgupta, C.H.Papadimitriou, and U.V.Vazirani.