CHENNAI MATHEMATICAL INSTITUTE

M.Sc. / Ph.D. Programme in Computer Science

Entrance Examination, 27 May 2010

This question paper has 5 printed sides. Part A has 10 questions of 3 marks each. Part B has 7 questions of 10 marks each. The total marks are 100.

Part A

1. Over the alphabet $\{0, 1\}$, consider the language

$$L = \{w \mid w \text{ does not contain the substring } 0011\}.$$

Which of the following is true about L.

(a) L is not context free

(b) L is regular

(c) L is not regular but it is context free

(d) L is context free but not recursively enumerable

2. We need to choose a team of 11 from a pool of 15 players and also select a captain. The number of different ways this can be done is:

(a) $\binom{15}{11}$

(b) $11 \cdot \binom{15}{11}$

(c) $15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5$

(d) $(15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5) \cdot 11$

3. The area of the largest square that can be drawn inside a circle with unit radius is:

(a) $\sqrt{2}$ (b) 2 (c) 1 (d) None of the above

4. Let m and n range over natural numbers and let $Prime(n)$ be true if n is a prime number. Which of the following formulas expresses the fact that the set of prime numbers is infinite?

(a) $(\forall m)(\exists n)\ (n > m)$ implies $Prime(n)$

(b) $(\exists n)(\forall m)\ (n > m)$ implies $Prime(n)$

(c) $(\forall m)(\exists n)\ (n > m) \wedge Prime(n)$

(d) $(\exists n)(\forall m)\ (n > m) \wedge Prime(n)$

5. You have two normal, fair, dice, with faces labelled 1,2,. . . ,6. If you throw both dice, which of the following is true about the total value shown by the dice?

(a) The probability that the total is 6 is less than the probability that the total is 9.

(b) The probability that the total is 6 is equal to the probability that the total is 9.

(c) The probability that the total is 6 is greater than the probability that the total is 9.

(d) None of the above.

6. A simple graph is one with no self-loops or multiple edges. Among the simple graphs with n vertices and at most $20n - 3$ edges:

   (a) There is always a graph with all vertices connected to at least 42 other vertices.

   (b) For all such graphs the number of vertices connected to at least 42 other vertices is at most cn for some constant $c < 1$.

   (c) There are no graphs with each vertex connected to at most 38 other vertices.

   (d) None of the above.

7. For integer values of n, the expression $\dfrac{n(5n+1)(10n+1)}{6}$

   (a) Is always divisible by 5.

   (b) Is always divisible by 3.

   (c) Is always an integer.

   (d) None of the above.

8. In programming language terminology, *call by value* refers to the fact that:

   (a) A function call can return a value.

   (b) When a function is called, arguments are copied into local storage.

   (c) Functions can indirectly modify the value of external variables.

   (d) Every argument passed to a function must have a value.

9. Consider the following functions `f()` and `g()`.

```
f(){                      g(){
  w = 3;                     z = w;
  w = 4;                     z = z + 2*w;
}                            print(z);
                          }
```

We start with `w` set to `0` and execute `f()` and `g()` in parallel—that is, at each step we either execute one statement from `f()` or one statement from `g()`. What is the set of possible values printed by `g()`?

(a) `0,9,12` (b) `0,8,9,12` (c) `0,6,8,9,11,12` (d) `0,4,6,9,10,12`

10. Consider the following statements.

    (A) NP-complete problems are those that we know we can never solve efficiently.

    (B) If we find an efficient algorithm for one NP-complete problem, then we can solve all NP-complete problems efficiently.

    (C) Checking whether a number is a prime is an NP-complete problem.

    Then:

    (a) A and B are true but C is false.

    (b) A and C are false but B is true.

    (c) B and C are true but A is false.

    (d) All three statements are false.

Part B

1. An international cellphone company provides service on 7 different frequencies. They wish to set up business in Tamil Nadu and have fixed the locations of 100 towers for their new service. The company has to ensure that two towers broadcasting on the same frequency are at least 100 km apart, so that there is no interference of signals.

    (i) Model this problems using graphs.

    (ii) Describe an algorithm which will answer the question "Is it feasible to set up towers at the given locations and provide service on 7 different frequencies?". Your algorithm should say "feasible" if it is feasible, otherwise output the minimum number of frequencies needed to utilise all 100 towers.

2. Let G be a graph in which each vertex has degree at least k. Show that there is a path of length k in G—that is, a sequence of $k+1$ *distinct* vertices $v_0, v_1, ..., v_k$ such that for $0 \leq i < k$, v_i is connected to v_{i+1} in G.

3. The Income-Tax Department had prepared a list D of names of defaulters on March 31. However, the government extended the deadline to pay taxes till April 15.

    The IT department has now received two additional lists of names: a list B of names of people who have paid their taxes between April 1 and April 15 at a bank, and a list O of names of people have paid their taxes during this period online. Some people have paid part of their taxes at a bank and part online, so there may be names that appear in both B and O.

    From the lists D, B and O, the IT department wants to prepare a revised list of defaulters. The names in the original list D are sorted in alphabetical order while the names in B and O are listed according to the date on which they paid their taxes. Fortunately, no two people have the same name.

    Describe an efficient algorithm to compute the revised list of defaulters. Assume that the size of D, B and O is n, m and k respectively and that $n > m > k$. Describe the running time of your algorithm in terms of these parameters.

4. Indicate whether the following statements are true or false, providing a short explanation to substantiate your answers.

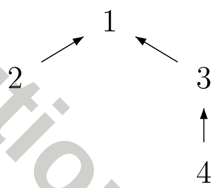
   (i) A DFA with n states must accept at least one string of length greater than n.

   (ii) A DFA that has n states and accepts an infinite language must accept at least one string x such that $2n < |x| < 3n$, where $|x|$ denotes the length of x.

   (iii) If a language L is accepted by an NFA with n states then there is a DFA with no more than 2^n states accepting L.

5. Sales have slumped at the Siruseri noodle factory and the management may need to terminate the contracts of some employees. Every employee has one immediate boss. The seniormost person in the company is the president, who has no boss. For legal reasons, if an employee's contract is not terminated, then his boss's contract cannot be terminated either. For how many different sets of employees can the management legally terminate contracts? Note that one possibility that has to be counted explicitly is that no employees' contracts are terminated (that is, the set of employees whose contract is terminated is the empty set).

   For example, suppose there are four employees, organised as follows. Each arrow points from an employee to his or her boss.



   Here, there are 7 different ways to terminate contracts for a set of employees, as follows:

$$\Big[\{1,2,3,4\},\{\},\{4\},\{2\},\{3,4\},\{2,4\},\{2,3,4\}\Big]$$

   Describe an algorithm to solve this problem efficiently.

6. You are given a list of positive integers along with a sequence of operations from the set {*,+}. You construct expressions from these two lists so that:

   - The numbers in the expression are drawn from the first list, without repetition and without altering their order.
   - All the operators in the second list are used in the expression, in the same order.

   For example, if the two lists are [1,3,2,1,4] and ['*','+'], the set of possible expressions you can form are

```
{1*3+2, 1*3+1, 1*3+4, 1*2+1, 1*2+4, ..., 2*1+4}
```

   For each expression, the value is computed by bracketing operators from the right. That is, the expressions above are evaluated as

```
{1*(3+2), 1*(3+1), 1*(3+4), 1*(2+1), 1*(2+4), ..., 2*(1+4)}
```

The aim is to determine maximum value among these expressions. In this example, the maximum value is 18, from the expression 3*2+4, which is evaluated as 3*(2+4).

You may assume that the length of the first list is more than the length of the second list.

Describe an algorithm to solve this problem.

7. A finite sequence of bits is represented as a list with values from the set {0,1}—for example, [0,1,0], [1,0,1,1], ...

   [] denotes the empty list, and [b] is the list consisting of one bit b. The function length(l) returns the length (number of bits) in the list l. For a nonempty list l, head(l) returns the first element of l, and tail(l) returns the list obtained by removing the first element from l. The operator ++ denotes list concatenation.

   For example:

   - head([0,1,0]) = 0, tail([0,1,0]) = [1,0],
   - head([1]) = 1, tail([1]) = [], and
   - [0,1,0]++[1] = [0,1,0,1].

   Conside the following functions:

   - op takes takes as input two bits and returns a bit.

     ```
     op(a,b)
       if (a = b) return(0)
       else return(1)
       endif
     ```

   - mystery1 takes as input two lists and returns a list.

     ```
     mystery1(s,t)
       if (length(s) != length(t)) then return(t)
       else if (length(s) = 0) then return(s)
            else return([op(head(s),head(t))] ++ mystery1(tail(s),tail(t)))
            endif
       endif
     ```

   - mystery2 takes as input two lists and outputs a list.

     ```
     mystery2(s,t)
       if (length(t) = 0) then return(s)
       else return( mystery1(mystery2(s,tail(t)),mystery2(s,tail(t))))
       endif
     ```

   (i) What is the value of length(mystery2(s,t)) in terms of length(s) and length(t)?

   (ii) Suppose s = t = 110100100. What are the first two bits of mystery2(s,t)?