CS648 : Randomized Algorithms
CSE, IIT Kanpur

Practice sheet
Hashing & Probabilistic methods

1. **Hashing with larger space for hash table**
   Let H be a universal hash family. Suppose we wish to build a single level hash table only. Fortunately, we are allowed to have a hash table of size larger than s. Suppose we wish to have a worst case search time $O(s^{1/2})$. What should be the minimum value of n (size of hash table) that would serve this purpose. Design a Las Vegas algorithm to build such a hash table.

2. **Hashing using matrices**
   Let the size of universe be 2^u, so each element in the universe can be represented by u bits only. The hash table size is 2^b. We wish to construct a universal hash family based on matrices. This will require you to interpret each binary number as a vector. Let $\mathcal{M}$ be the set of all b-by-u matrices consisting of 0/1 entries. For a matrix $M \in \mathcal{M}$ we define function h_M as follows.

   $$h_M(x) = M \cdot x \qquad \forall x \in U$$

   Here $M \cdot x$ is the product of M and vector x where we do addition modulo 2. Prove that the following hash-family is a universal hash family.

   $$H = \{h_M | M \in \mathcal{M}\}$$

3. **Random Permutation**
   Let n be a prime number and let $S = \{1, 2, ..., n-1\}$. We are given an array A storing a permutation of S.

   (a) We wish to permute A randomly such that the following condition is satisfied.

   $$\mathbf{P}(A[i] = j) = \frac{1}{n-1}$$

   How will you do it using $O(\log n)$ random bits only ?

   (b) Suppose, we wish to permute A such that the following condition (in addition to the one mentioned in part (a)) is also satisfied for any $i \neq j$ and $k \neq \ell$.

   $$\mathbf{P}(A[i] = k \text{ and } A[j] = \ell) = \frac{1}{n(n-1)}$$

   How will you do it using $O(\log n)$ random bits only ?

   Are you not amazed by such a small number of random bits used to generate random permutation ?

4. **Monochromatic Clique**
   Prove that, for every integer n, there exists a coloring of the edges of the complete graph K_n by two colors so that the total number of monochromatic copies of K_4 is at most $\binom{n}{4}2^{-5}$.

5. **Line intersecting a circle**
   There are several circles of total circumference 10 inside a square of side length 1. Prove that there is a line that intersects at least 4 of the circles.
   Hint : Choose any one side of the square. Select a random point on this side and draw a line perpendicular to the side. What will be expected number of circles it will intersect ?

6. **Large cut**
   In the lecture class, we showed that a graph having m edges has a cut of size at least $m/2$. In fact, this bound can be further improved slightly: If G has $2n+1$ vertices and m edges, then it has a cut of size at least $m(n+1)/(2n+1)$. Show this formally and rigorously. Then try to get a bound better than $m/2$ for a graph having even number of vertices ?
   Hint : Use another simple randomized algorithm to partition the vertices.

7. **Independent set** Let $G = (V, E)$ be an undirected graph on n nodes and m edges, where degree of each node is at most $d > 1$. A subset S of nodes is said to be an independent set if no edge exists between any pair of nodes in S. The following is a sketch of a distributed algorithm that computes an independent set S of *large* size for G in one round and using $O(m)$ messages.

   1 Each node i tosses a coin that gives heads with probability p;

   2 Each node i sends the outcome of its coin toss to all its neighbours and, in turn, receives their outcomes as well.

   3 Each node i adds itself to S if ....

   (a). Complete the algorithm by suitably writing the text of dotted line of Step 3 stated above.
   (b). Express the expected size of S in terms of n, d, and p. You must give proper justification for arriving at the expression as well.
   (c). What should be the value of p that maximizes expected size of the S?

8. **A glimpse of the area of streaming algorithms...**

   Note: This problem is just for fun and you may skip it for the exam.

   You enter a shopping mall. There is a row of apples and you are standing at one end of this row. The row is so long that you can not see the other end of the row and hence don't know the exact number of apples in the row. You have a fair coin in your pocket and there is a positive integer k. How will you get out of the mall with a uniformly random sample of k apples subject to the following constraints:

   - You have a bag that can accommodate at most k apples at any time.
   - You are allowed to make only a single pass over the row of apples. (Note that once you discard an apple, you can not place it back in the bag).

   What is the expected number of coin tosses that you will have to make ?