CS648 : Randomized Algorithms
Semester I, 2011-12, CSE, IIT Kanpur

Assignment - 4 (due on 11th November : 9AM)

Note: Give complete details of the analysis of your solution. Be very rigorous in providing any mathematical detail in support of your arguments. Also mention any Lemma/Theorem you use.

1. **Method of bounded difference**

   This exercise is on the applications of method of bounded difference.

   (a) *Rumor Spreading*

   There were n persons in a city, with one of them knowing a rumor. On the morning of the first day, he selects one person from the city randomly uniformly, calls him and conveys the rum-our to him. Each person who comes to know this rumor continues to spread the rumor in the same way as the first person. So on the morning of any day, each person who knows the rum-our selects a person randomly uniformly from the city, calls him, and conveys the rumor to him/her. Show that the expected number of days until the entire city comes to know about the rumor is $O(\log n)$.

   This problem was discussed briefly in a lecture class last week. One has to pursue the approach of partitioning the experiment in to stages suitably. There are two stages as follows. In the first stage, the number of people who know the rumor is less than $n/2$. In the second stage, the number if people who know the rumor is more than $n/2$ but less than n. One has to show that the expected number of days spent in each of the stages is $O(\log n)$. In the first stage, one will have to show that in expected $O(1)$ rounds the number of persons knowing the rumor increases by some constant factor. For this purpose, it will be helpful to explore the relationship with the ball-bin problem. You will have to use method of bounded difference (MOBD) here. You are encouraged to provide a concise and formal analysis of the entire experiment.

   (b) *Triangles in a random graph*

   A random graph $G(n, p)$ is a graph on n vertices, where each edge is present in the graph independently with probability p. Let X be the random variable for the number of triangles present in $G(n, p)$. Fix $p = 1/2$ henceforth.

   i. Calculate $\mathbf{E}[X]$ and $\mathbf{E}[X^2]$. Give simple arguments to show that we can't use Chernoff bound to show that X is concentrated around $\mathbf{E}[X]$.

   ii. Chebyshev's inequality is a very simple and useful tool for getting a bound on the deviation of X from $\mathbf{E}[X]$. You may study it from any book/website. Use Chebyshev's inequality to get a bound on $\mathbf{P}[X > 5/4\mathbf{E}[X]]$.

   iii. Use the method of bounded difference to derive the bound on $\mathbf{P}[X > 5/4\mathbf{E}[X]]$. Compare it with the bound obtained using Chebyshev's inequality.

2. **Principle of deferred decision**

   Let K_n be a complete graph on n vertices where each edge is assigned a weight uniformly and randomly in the range $[0, 1]$. Show that the expected weight of MST of K_n is $O(1)$. For this purpose, analyse the Kruskal's algorithm on K_n using the principle of deferred decision.

3. **Delay sequences**

We shall describe one more tool for analysing the randomized algorithms in a lecture class in near future. This tool is called *delay sequences*. We shall demonstrate the power of this tool by solving the following problem in the class.

There are n counters numbered C_0 to C_{n-1} arranged on a circular loop. So C_i has two neighbors $C_{(i-1)\ mod\ n}$ and $C_{(i+1)\ mod\ n}$. Each counter C_i stores a count-value, called $\textsc{val}(c_i)$ which is initialized to 0 for all counters in the beginning. Each counter is also equipped with a fair coin as well. Let us now describe the experiment.

In the beginning of every minute, each counter C_i does the following: It tosses its coin. If the outcome is HEADS and none of its neighbors have count-value less than $\textsc{val}(C_i)$, then C_i increments $\textsc{val}(C_i)$ by 1.

Interestingly, with high probability, count-value of each counter will reach $\log n$ within $O(\log n)$ rounds. Can you prove it using the tools we have discussed ? There is a very insightful for this claim based on *delay sequences*.

The above problem of counters, in addition to being a very interesting problem, also proves a very surprising and useful result: A distributed computational environment in presence of failures works *almost* as good as a distributed computational environment without failures. We shall provide more details of this result during the lecture.

As an exercise of this assignment, you will have to solve the following generalization of the above problem. Suppose the counters are located on the vertices of a d-regular graph, probability of getting HEADS is p and the maximum permissible difference in the count-value of two neighboring counters is b.