Homework 2 **CSE 591 Fall 2008**

Due October 2

1

Suppose we ask an Oracle a question $6n$ times. Each time we ask the same question with a YES/NO answer and each time (independently) we have a 3/4 chance of getting the correct answer. If we take the majority of the $6n$ answers, show that the probability we get the wrong answer is no more than $1/2^n$.

Prove this in two ways: first using the Chernoff bound, then using only basic arithmetic and the fact that the number of subsets of a set of size $6n$ is 2^{6n}.

2

Let K_n be the complete graph on n vertices: every vertex is adjacent to every other vertex. Suppose each of the $\binom{n}{2}$ edges is colored red or blue. Let $k \leq n$. This problem concerns the question of whether it is possible to color all the edges in such a way that no set of k vertices has all of its $\binom{k}{2}$ internal edges colored with the same color. For large n (compared to k) the answer is no, and the smallest such n is called the Ramsey number $R(k,k)$. The existence of Ramsey numbers follows from a theorem proved by Ramsey in 1929. You don't have to prove the existence of Ramsey numbers for this question. Instead, you will prove a lower bound on $R(k,k)$ using the probabilistic method.

Suppose you select the color of each edge randomly: red or blue, each with probability 1/2. Consider a set of k vertices and compute the probability that all of its edges are colored the same. Then compute the probability that among all the subsets of at least 2 vertices, every one had all its edges colored the same. Use the union bound.

How large does n have to be compared to k, to ensure that the probability that every set of (exactly) k vertices has edges of a unique color is strictly less than 1? (In this case the probability that every set sees both colors is positive.)

3

Consider the following two games. In game A, you repeatedly flip a biased coin (coin a) that comes up heads with probability $p_a < 1/2$, and tails with probability $1 - p_a$. You win a dollar if the coin comes up heads and lose a dollar if it comes up tails. (Clearly, this is a losing game for you.) In game B, you also repeatedly flip coins, but the coin that is flipped depends on how you have been doing so far in the game. Let w be the number of your wins so far and l the number of your losses. Each round you bet one dollar and so $w - l$ is exactly equal to your winnings so far. This game B uses two biased coins, coin b and coin c If your winnings in dollars are a multiple of 3, then you flip coin b, which comes up heads with probability p_b, and tails with probability $1 - p_b$. Otherwise you flip coin c, which comes up heads with probability p_c, and tails with probability $1 - p_c$. Again, you win a dollar if the coin you flipped comes up heads, and lose a dollar if it comes up tails.

(a) Suppose $p_a = 0.49$, $p_b = 0.09$ and $p_c = 0.74$. What is the expected gain (or loss) per round in playing game A for many rounds? What is the expected gain (or loss) per round in playing game B for many rounds?
(b) Consider now game C, in which you first toss a fair coin d. Then, if d comes up heads, you play game A, and if it comes up tails, you play game B. What is the expected gain (or loss) per round in playing game C for many rounds?

4

The NP-hard *set cover* problem is: given subsets $S_1, S_2, \ldots, S_m$ of $V = \{1, 2, \ldots, n\}$, cover V with as few of the S_is as possible (in other words, select as few S_is as possible so that their union contains V).

(a) Suppose we are given an instance of set cover whose optimal solution has size p (it uses p sets). Show that the simple greedy algorithm ("pick the largest subset; remove the elements of V it covers; repeat") has the property that after finding p subsets it has covered at least $(1 - 1/e)n$ elements of V.

(b) Suppose now that V contains some unknown "hidden" set H "of important points", and that we would like to find p subsets that cover a *constant fraction* of H, rather than just cover a constant fraction of V. Describe a randomized algorithm to do this. It should find p subsets, and have the property that for any $H \subseteq V$, the expected number of points in H covered is at least $|H|(1 - 1/e)$.

5

Consider the following special case of the set cover problem: given a graph $G = (V, E)$, the set of elements X is equal to the set E of edges in G, and the family of subsets $\mathcal{S}$ is exactly the set of "edge-neighborhoods": there is a set S_e in $\mathcal{S}$ for each edge $e \in E$, and the elements of S_e are the edges adjacent to e and e itself. All weights are equal to 1.

Give a 2-approximation for the minimum set cover in this case. (Hint: show that the minimum set cover has the same size as the minimum-size maximal matching—that is, a matching of minimum size among all those that cannot be enlarged by simply adding an edge.) Does your algorithm work when the weights are not all 1?