Groundrules

- Homeworks will generally consist of *exercises*, easier problems designed to give you practice, and *problems*, that may be harder, trickier, and/or somewhat open-ended. You should do the exercises by yourself, but you may work with a friend on the harder problems if you want. One exception: no fair working with someone who has already figured out (or already knows) the answer. If you work with a friend, then write down who you are working with.
- If you've seen a problem before (sometimes we'll give problems that are "famous"), then say that in your solution (it won't affect your score, we just want to know). Also, if you use any sources other than the textbook, write that down too (it's fine to look up a complicated sum or inequality or whatever, but don't look up an entire solution).

Reading: Chapters 1 and 2 of Motwani & Raghavan.

Exercises

1. Consider the standard random walk on the path $\{0, 1, \ldots, n\}$ (which goes left/right w.p. 1/2 each). Let E_x be the expected time to hit $\{0, n\}$ starting from x. What would be the time to hit $\{0, n\}$ if we stayed put w.p. p, went left/right w.p. $(1-p)/2$ instead?
2. We know that 2-SAT can be solved in linear time. However, consider the following algorithm for 2-SAT:

   Pick an arbitrary starting assignment. If it is not a satisfying assignment, pick an arbitrary unsatisfied clause. Pick a random one of the two variables in it, and flip its value.
   Repeat this until you find a satisfying assignment (in which case output it), or you've done this for $2n^2$ steps (in which case say that the formula is not satisfiable).

   Show that the algorithm answers correctly with probability at least 1/2.
3. Prove that $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$ for independent discrete random variables X and Y.

Problems

1. **Coin Flipping.** In this course we will routinely assume that given any $p \in [0, 1]$ we can flip a coin with bias p. I.e., we will assume a procedure $\mathsf{coin}(p)$ that outputs 1 (or "heads") with probability p, and 0 (or "tails") with probability $1 - p$.
   (a) Suppose you just have a fair coin ($p = 1/2$), but want to implement $\mathsf{coin}(p)$ for any p. Say you have access to the binary representation of p. Give an algorithm for implementing $\mathsf{coin}(p)$ with expected *constant* running time (independent of p) as measured by the number of flips of your fair coin. What is the constant?
   (b) Suppose we have a source for randomness that outputs 1 with some **fixed but unknown** probability $q \in (0, 1)$ (and 0 with probability $1 - q$). Show how we can use this to implement $\mathsf{coin}(1/2)$. What is the expected number of random bits you use to output one fair bit?
2. **A Better Analysis for 3-SAT.** In lecture we saw an analysis for the randomized 3-SAT algorithm with success probability at least $(2/3)^n$. We now show that if we perform Step #2 for $3n$ iterations instead of just n, we get a success probability of $\approx (3/4)^n$.

(a) Clearly argue why the probability of success is at least $\sum_{k=0}^{n} \binom{n}{k} 2^{-n} \binom{3k}{k} (1/3)^{2k} (2/3)^k$. Be clear about where each term in this formula is coming from.

(b) Now use Stirling's formula to show that $\binom{3k}{k} (1/3)^{2k} (2/3)^k$ is at least $\frac{1}{\sqrt{10k}} \cdot 2^{-k}$ for $k \geq 1$. (You can use $n! = \sqrt{2\pi n} (\frac{n}{e})^n e^{\lambda_n}$, for $\lambda_n \in [\frac{1}{12n+1}, \frac{1}{12n}]$.)

(c) Complete the argument to show that the probability of success is at least $\Omega(\frac{1}{\sqrt{n}} \cdot (3/4)^n)$.

3. **Finding Long Paths in Graphs.** Given a graph $G = (V, E)$, you want to find long simple paths in the graph in polynomial time.

(a) (Algorithm 1: Dead easy.) Show that you can find a path of length k (if such a path exists) in time $n\Delta^k$, where Δ is the maximum degree of G.

(b) (Easy.) If the graph were directed (and a DAG), then show that you can deterministically find the longest path in G in time $O(m + n)$. Here, and in general, $m = |E|$ and $n = |V|$.

(c) (Algorithm 2:) Consider running the following procedure n times, and outputing the longest path found in these n tries.

> Take a random permutation of the vertices, and direct each edge from the lower endpoint to the higher endpoint to create a DAG $\vec{G}$. Find a longest path in $\vec{G}$.

Show that for $k = c \frac{\log n}{\log \log n}$ for some constant $c > 0$, Algorithm 2 will find a path of length k (if it exists) with probability at least 1/2.

(d) Now, consider a slight extension of this idea. Suppose you have a graph G, and you color the vertices using k colors (neighbors need not have different color). A path is called *polychromatic* if has $\ell \leq k$ vertices, and all the ℓ vertices have different colors.

i. Show that you can find a polychromatic path of length k in time that is $\text{poly}(n, k) 2^k$. (So, this is polynomial time for $k = O(\log n)$).

ii. (Algorithm 3:) Consider running the following procedure n times, and outputing the longest path found in these n tries.

> Take a random coloring of the vertices using k colors, and find the polychromatic path of length at most k in G.

Show that for $k = c \log n$ for some constant $c > 0$, Algorithm 3 will find a path of length k (if it exists) with probability at least 1/2. (Hint: Use Stirling's approximation.)

Note: the current best guarantee known is $2^{\sqrt{\log n}}$.

4. (Due to T. Cover and M. Rabin). Consider the following game. A friend writes down two numbers on two slips of paper and then randomly puts one in one hand and the other in the other hand. You get to pick a hand and see the number in it. You then can either keep the number you saw or else return it and get the other number. Say you end up with the number x and the other number was y. Then, your gain is $x - y$.

For a given (possibly randomized) strategy S, let $\mathbb{E}_{x,y}(S)$ denote its expected gain, given that the two numbers are x and y. For instance, if S is a deterministic strategy, then $\mathbb{E}_{x,y}(S) = \frac{1}{2} \cdot$(gain of S given that it is initially shown x and the other number is y) + $\frac{1}{2} \cdot$(gain of S given that it is initially shown y and the other number is x).

(a) Consider the strategy S = "if the first number I see is ≥ -17, then I keep it, else I switch." What is $\mathbb{E}_{x,y}(S)$ in terms of x and y?

(b) Give a randomized strategy S such that $\mathbb{E}_{x,y}(S) > 0$ for all $x \neq y$.