COMS 4995 (Randomized Algorithms): Exercise Set #11

For the weeks of November 25–December 4, 2019

Instructions:

(1) *Do not turn anything in.*

(2) The course staff is happy to discuss the solutions of these exercises with you in office hours or in the course discussion forum.

(3) While these exercises are certainly not trivial, you should be able to complete them on your own (perhaps after consulting with the course staff or a friend for hints).

Exercise 51

Recall the balls and bins setting and the power of two choices result from Lecture #21. In that proof, we defined the sequence

$$\beta_4 = \frac{n}{4}$$

and

$$\beta_{i+1} = 2\frac{\beta_i^2}{n}$$

for all $i \geq 4$.

(a) Prove that once $i \geq \log_2(\log_2 n) + O(1)$, $\beta_i < 1$.

(b) Prove that the largest i for which $\beta_{i+1} \geq 12 \ln n$ satisfies $i = \log_2(\log_2 n) \pm O(1)$.

Exercise 52

Prove the following fact that we used in our proof in Lecture #21. Let $Z_1, \ldots, Z_n$ be a sequence of arbitrary random variables and $Y_1, \ldots, Y_n$ a sequence of 0-1 random variables with the property that each Y_i is fully determined by $Z_1, \ldots, Z_i$. (In the application in Lecture #21, the Z_i's correspond to the bins that each ball wound up in, and the Y_i's correspond to the Y_t's from lecture.) Suppose that

$$\mathbf{Pr}[Y_i = 1 \mid Z_1, \ldots, Z_{i-1}] \leq p$$

with probability 1 (over $Z_1, \ldots, Z_{i-1}$). Let $X_1, \ldots, X_n$ denote i.i.d. Bernoulli random variables with parameter p. Prove that, for every threshold k,

$$\mathbf{Pr}\left[\sum_{i=1}^{n} Y_i > k\right] \leq \mathbf{Pr}\left[\sum_{i=1}^{n} X_i > k\right].$$

[Hint: Induction on n.]

Exercise 53

Continuing with Lecture #21, suppose we choose $d \geq 2$ bins uniformly at random (with replacement) rather than 2. (As usual, a ball is placed in the least loaded of the d randomly chosen bins, with ties broken randomly.) Generalize the heuristic analysis from Lecture #21 to argue that we might guess that the expected maximum load under this process is

$$\approx \log_d(\log_2 n) = \frac{\log_2(\log_2 n)}{\log_2 d}.$$

Exercise 54

Generalize the argument in Lecture #22 to prove that Moser's algorithm, when given a k-SAT formula satisfying $d \leq 2^{k-3}$, terminates (with a satisfying assignment) within $O(m \log m)$ calls to Fix with high probability (tending to 1 as $m \to \infty$). (Here m denotes the number of clauses in the given formula, and d equals one plus the maximum degree of the dependency graph in which vertices correspond to clauses and edge correspond to pairs of overlapping clauses.)

[Hint: Suppose the algorithm fails to terminate within T calls to Fix with probability at least $\frac{1}{2}$, where the probability is over the $n + kT$ random bits used in the initialization and the first (at most) T Fix calls. Following the argument from lecture, what can you say about T? What if the probability is at least $\frac{1}{4}$?]