Problem Set 6

Electronic submission via Gradescope due **11:59pm Tuesday 11/5**. You are strongly encouraged to submit a homework with a partner—that is, submit one homework with both of your names.

[You may discuss these problems with classmates. Feel free to look at wikipedia, course notes, etc. for reference material, but do not try to specifically search online for solutions to the problems. ***Your submission must be the original work of you and your partner, and you must understand everything that is written on your submission.*** *We strongly suggest that you write solutions using LaTex—see the course website for a latex solution template.]*

1. (Probabilistic Method Practice) Suppose we are investigating the social habits in a group of n chimpanzees, and after months of observations, for some pairs of chimpanzees A and B, we know whether A has spent more time grooming B or whether B has spent more time grooming A. We wish to aggregate these results into a single ranking of the 'altruism' of each chimpanzee that minimizes the number of "inconsistent pairs", where a pair A, B is inconsistent/violated if A is above B in the ranking but has spent less time grooming B than B spent grooming A.

   (a) (4 points) Prove that there exists a ranking that violates at most half the pairwise relationships.

   (b) (4 points) Prove that for sufficiently large n, there exists a set of grooming habits such that for *every* ranking, at least 49% of the pairwise relationships would be violated. [Hint: probabilistic method—choose a distribution over the grooming habits, then argue that the probability a random ranking violates significantly less than half the pairwise relationships is so small, that there is a good chance you picked a set of grooming habits with the property that no 'good' ranking exists.]

   (c) (1 point) Prove that there exists a ranking that violates strictly less than half the pairwise relationships.

2. An *edge coloring* of an (undirected) graph $G = (V, E)$ assigns exactly one color to each edge of the graph. We say that a colored path in the graph is *symmetric* if the path has an even number of edges, and the second half of the path is colored identically to the first half of the path (i.e. the sequence of colors in the second half of the path is the same sequence as in the first half). [Throughout this problem, by "path" we refer only to simple paths—ie paths that do not re-use any edges.]

   (a) (4 points) Prove that for any graph whose maximum degree is d, there exists a coloring using $10 \cdot d^2$ colors such that there are no "symmetric" paths of length 4 (i.e. no repeating paths consisting of 4 distinct edges). [Hint: Lovasz Local Lemma!]

   (b) (4 points) Given the setup in the previous part, give an algorithm that will find such a coloring in expected time polynomial in the size of the graph, and justify the runtime.

(c) (4 points) Prove that there is some constant C such that for any graph whose maximum degree is d, there exists a coloring using $C \cdot d^2$ colors such that there are no "symmetric" paths (of any length).

3. Consider a set of equations over variables $x_1, \ldots, x_n$, where each equation has the form $a_1 x_{i_1} + a_2 x_{i_2} + \ldots + a_r x_{i_r} \equiv a_{r+1} \mod 17$, for some r (that might vary from equation to equation) and set of coefficients $a_1, \ldots, a_r \in \{1, 2, \ldots, 16\}$, and $a_{r+1} \in \{0, \ldots, 16\}$. Additionally, suppose that each variable, x_i, occurs in at most 4 equations.

   (a) (6 points) Prove that there exists an assignment to the variables such the *none* of the equations are satisfied.

   (b) (2 points) Would your proof above continue to hold if the equations were modulo 18 rather than modulo 17? Explain why or why not?

4. Tightness of the Lovasz Local Lemma: One version of the LLL that we saw asserts that for any set of events $A_1, \ldots, A_n$, such that for each i, A_i is mutually independent of all but at most d events, then as long as $Pr[A_i] \leq \frac{1}{e(d+1)}$, then there is a nonzero chance of all events being simultaneously avoided.

   (a) (4 points) Define a set of events over a probability space such that each event is mutually independent of all but at most d other events, and $Pr[A_i] \leq 1/(d+1)$ for all i, but the probability of simultaneously avoiding all events A_i is 0. This shows that the constant e in the statement of the LLL cannot be replaced by 1.

   (b) (4 bonus points) For some constant $c > 1$, prove that the constant e in the LLL cannot be replaced by c.