Carnegie Mellon University

15-859: Information Theory and Applications in TCS | Spring 2013

PROBLEM SET 4

Due by Tuesday, April 30

INSTRUCTIONS

- You are allowed to collaborate with up to two other students taking the class in solving problem sets. But here are some rules concerning such collaboration:
  1. You should think about *each problem* by yourself for at least 30 minutes before commencing any collaboration.
  2. Collaboration is defined as discussion of the lecture material and solution approaches to the problems. Please note that *you are not allowed to share any written material and you must write up solutions on your own without any "collaboration notes" as an aid.*
  3. You must clearly acknowledge your collaborator(s) in the write-up of your solutions.
  4. Of course, if you prefer, you can also work alone (see the last bullet item for some "credit" for doing so).
- Solutions typeset in LaTeX are strongly preferred.
- You should *not* search for solutions on the web. More generally, you are urged to try and solve the problems without consulting *any* reference material other than the course notes and what we cover in class. If for some reason you feel the need to consult some source, *please acknowledge the source* and try to articulate the difficulty you couldn't overcome before consulting the source and how it helped you overcome that difficulty. Alternatively, before turning to any such material, we encourage you to ask us for hints or clarifications.
- Please start work on the problem set early. The problem set has **six** problems and is worth a total of 100 points. As a rather rough guess/estimate, *scoring around* 85% *of the points, or* 75% *of the points if you work by yourself,* might suffice for an A on this problem set.

---

1. (15 points) Let $S \subset \mathbb{R}^n$ be a set of points and $T_1, \ldots, T_m$ be subsets of $[n]$ such that, for every $i \in [m]$, the set S projected to the coordinate positions picked by T_i has at least n_i distinct elements. Suppose each position $j \in [n]$ is included in at least k of the sets $T_1, \ldots, T_m$. Show that
$$|S|^k \le \prod_{i \in [m]} n_i.$$
(*Hint:* Shearer's lemma.)

2. (15 points) Consider the hypercube graph H with vertex set $\{0,1\}^m$ for some positive integer m, and edges $(x, y) \in E$ for $x, y \in \{0,1\}^n$ that differ in exactly 1 bit position. Let $S \subset \{0,1\}^m$.

   (a) Prove, using an entropy based argument, that the number of edges $(x, y) \in E$ with both $x, y \in S$ is at most $(|S| \log |S|)/2$.
   Hint: Consider the random variable $(X_1, X_2, \ldots, X_m)$ uniformly distributed on S. What's the value of $H(X_1|X_2, X_3, \ldots, X_m)$?

(b) Can you improve the above upper bound when $|S| \leq 2^{m/2}$?

3. (10 points) Let G be a graph with chromatic number $\chi(G)$ and a distribution P on its vertices. Show that

$$H(G, P) \leq \log_2 \chi(G).$$

4. (25 points) (Yet another characterization of graph entropy) The *vertex packing polytope* of a graph G with vertex set $V = [n]$, denoted by $\mathrm{VP}(G)$, is defined as the convex hull of characteristic vectors of the independent sets of G. That is,

$$\mathrm{VP}(G) := \{ \sum_{S \in I(G)} \alpha_S \cdot \chi(S) \colon (\forall S) \alpha_S \geq 0, \sum_{S \in I(G)} \alpha_S = 1 \},$$

where $I(G) \subset V$ is the set of all independent sets of G and $\chi(S) := (s_1, \ldots, s_n) \in \{0,1\}^n$ such that $s_i = 1 \Leftrightarrow i \in S$. The goal of this exercise is to show that the entropy of a graph G under distribution $P = (p_1, \ldots, p_n)$ is determined by

$$H(G, P) = \min_{\substack{(a_1, \ldots, a_n) \in \mathrm{VP}(G) \\ a_i > 0}} \sum_{i \in [n]} p_i \log_2 \big(\frac{1}{a_i}\big). \tag{1}$$

(a) First, prove that

$$\min_{\substack{(a_1, \ldots, a_n) \in \mathrm{VP}(G) \\ a_i > 0}} \sum_{i \in [n]} p_i \log_2 \big(\frac{1}{a_i}\big) \leq \min_{\substack{(X,Y) \\ X \sim P, X \in Y, Y \in I(G)}} I(X;Y).$$

In order to do so, let (X, Y) be the minimizer of the right hand side and show that the convex combination of the independent sets defined by the distribution of Y is a feasible solution for the optimization on the left hand side. *Hint:* You may want to use concavity of the log function and Jensen's inequality to simplify the calculations.

(b) Prove the reverse inequality

$$\min_{\substack{(a_1, \ldots, a_n) \in \mathrm{VP}(G) \\ a_i > 0}} \sum_{i \in [n]} p_i \log_2 \big(\frac{1}{a_i}\big) \geq \min_{\substack{(X,Y) \\ X \sim P, X \in Y, Y \in I(G)}} I(X;Y)$$

by letting $\sum_{S \in I(G)} \alpha_S \cdot \chi(S)$ to be the minimizer of the left hand side and naturally defining the joint distribution (X, Y) from the given distribution P and the coefficients α_S. *Hint:* Again, you may find concavity of the log function useful.

(c) Show that the minimum in (**??**) is attained by $\sum_{S \in I(G)} \alpha_S \cdot \chi(S)$ where the coefficients α_S are only supported on maximal independent sets (that is, any independent set S that is strictly contained in a larger one gets $\alpha_S = 0$).

5. (20 points) In this exercise, you will revisit the rectangle size bound and prove a version (in a similar spirit to the discrepancy method from class) that is useful for lower bounding *randomized* communication complexity.

(a) Consider the usual setup of a Boolean function $f : X \times Y \to \{0,1\}$. Let $\alpha, \rho \in (0,1)$. Suppose we have a distribution μ on $X \times Y$ such that for every large rectangle $R = S \times T \subseteq X \times Y$ with $\mu(R) \geq \rho$, we have a sizeable fraction of 1s in the rectangle; formally

$$\mu(R \cap f^{-1}(1)) > \alpha \cdot \mu(R \cap f^{-1}(0)) .$$

Prove that in this case

$$2^{R(f)} \geq \frac{1}{\rho}\left(\mu(f^{-1}(0)) - \frac{1}{\alpha}\right) .$$

(b) Using the above or otherwise prove that $R(\mathrm{IP}) \geq n - O(1)$.

6. (15 points) Consider the following variant of the indexing problem. Alice holds a string $x = (x_1, \ldots, x_n) \in \{0,1\}^n$ and Bob holds an index $i \in \{1, 2, \ldots, n\}$ *and* the bits $x_1, x_2, \ldots, x_{i-1}$. The goal is for Alice to send one message which will enable Bob to determine x_i with probability at least $2/3$. Prove that the randomized one-way communication complexity of this problem is still lower bounded by $\Omega(n)$ bits.