Due: Monday October 7, in class (3pm).

1. The goal of this exercise is to work out an axiomatic definition of entropy. In other words, we start with some desirable properties of the entropy function, and show that $H(\cdot)$ is the only function satisfying these properties — and hence the formula for $H(\cdot)$ follows from the specified axioms. Let f be a function that takes a random variable X on finite support and outputs a real number with the following properties:

   - $f(X)$ only depends on the frequencies of the different values of X.
   - If X is a uniformly random point from a set of size M, and Y is a uniformly random point from a set of size $M' > M$, then $f(M') > f(M)$.
   - If X, Y are independent, then $f(X, Y) = f(X) + f(Y)$.
   - If B_q is such that $\Pr[B_q = 1] = q$ and $\Pr[B_q = 0] = 1 - q$, then $f(B_q)$ is a continuous function of q.
   - If B is a random variable taking $0/1$ values, and X is another random variable, then $f(BX) = f(B) + \Pr[B = 1] \cdot f(X|B = 1) + \Pr[B = 0] \cdot f(X|B = 0)$.
   - $f(B_{1/2}) = 1$.

   Show that $f(X) = H(X)$ for all finitely supported X. *Hint:* start with uniform X's, then proceed to $f(B_q)$.

2. (Problem 2.25 in the CT book). There isn't really a notion of mutual information common to three random variables. Here is one attempt at a definition: Using Venn diagrams, we can see that the mutual information common to three random variables X, Y, and Z can be defined by

   $$I(X; Y; Z) = I(X; Y) - I(X; Y|Z).$$

   The quantity is symmetric in X, Y, and Z despite the preceding asymmetric definition (this fact will follow from the identities below). Unfortunately, $I(X; Y; Z)$ is not necessarily non-negative. Find X, Y, and Z such that $I(X; Y; Z) < 0$, and prove the following two identities.

   (a) I(X;Y;Z)=H(XYZ)-H(X)-H(Y)-H(Z)+I(X;Y)+I(Y;Z)+I(Z;X).
   (b) I(X;Y;Z) = H(XYZ)-H(XY)-H(YZ)-H(ZX)+H(X)+H(Y)+H(Z).

   The first identity can be understood using the Venn diagram analogy for entropy and mutual information. The second identity follows easily from the first.

3. (Problem 2.39 in the CT book, extended). Let X, Y, Z be three Bernoulli(1/2) random variables that are pairwise independent: $I(X; Y) = I(Y; Z) = I(Z; X) = 0$.

   (a) Under this constraint, what is the minimum value for $H(XYZ)$?
   (b) Give an example achieving this minimum.
   (c) What would the minimum value for $H(XYZ)$ be if the constraint above was replaced with $I(X; Y) = I(Y; Z) = I(Z; X) = \alpha$, for some $0 \le \alpha \le 1$?
   (d) Show (by giving an example, or otherwise) that your bound from the previous part of the question is tight.

4. You are given a coin C, and know the following fact: with probability $1/2$ it is a fair coin (i.e. its outcomes are distributed as i.i.d Bernoulli $B_{1/2}$), and with probability $1/2$ it is ε-biased (i.e. its outcomes are distributed as i.i.d $B_{1/2+\varepsilon}$). Your goal is to determine which is the case with confidence of $3/4$. Let $C \in \{F, B\}$ be the random variable representing coin type. Let $T_1, T_2, \ldots, T_n$ represent tosses of the coin.

(a) Calculate $I(T_1; C)$ up to constant multiplicative terms (i.e. using $\Theta(\cdot)$ notation) in terms of ε;

(b) Prove that determining C with confidence of $3/4$ would require $\Omega(1/\varepsilon^2)$ coin tosses. You should only use part (a) and information-theoretic formalism.