Kolmogorov Complexity – Assignment 4

1. The relative entropy (or Kullback-Leibler divergence) between two probability distributions p and q over a universe U is

$$D(p||q) = \sum_{x \in U} p(x) \log \frac{p(x)}{q(x)}.$$

   (a) Prove that for all distribution p and q, $D(p||q) \geq 0$ with equality only holding when p and q are identical distributions (i.e., $p(x) = q(x)$, for all x in U)

   (b) Given two random variables X and Y whose joint distribution is given by the distribution $p(x, y)$, show that the

$$I(X : Y) = D(p(X, Y)||p_X(Y) \cdot p_Y(X))$$

   where $p_x(y)$ denotes the conditional distribution on Y given $X = x$ (similarly $p_y(x)$).

   (c) Use (a) and (b) to conclude that for all random variables X and Y, $I(X : Y) \geq 0$. Furthermore, equality occurs if and only if X and Y are independent random variables.

   (d) Use (c) to conclude that conditioning always reduces entropy, i.e., for all random variables X and Y, $H(X) \geq H(X|Y)$.

   (e) (Subadditivity of entropy) Prove that for all random variables X and Y, $H(X, Y) \leq H(X) + H(Y)$.

   (f) For any random variable X defined over a universe U, show that $H(X) \leq \log |U|$ [Hint: Use (a)]

2. Consider linear inequalities of sizes of finite sets with unions and how that compares with Kolmogorov complexity. For example the statement $C(x, y) \leq C(x) + C(y) + O(\log n)$ corresponds to $|X \cup Y| \leq |X| + |Y|$.

   (a) Prove or disprove: Every linear inequality that holds for Kolmogorov complexity (up to additive logarithmic factors) also holds for sizes of sets with unions.

   (b) Prove or disprove: Every linear inequality that holds for sizes of sets over unions holds for Kolmogorov complexity (up to additive logarithmic factors).