Information Theory Exercise Sheet #5

University of Amsterdam, Master of Logic, Spring 2014
Coordinator: Christian Schaffner
Guest Lecturer: Teresa Piovesan
TA: Hoang Cuong

Out: Thu, 6 March 2014
Due: Thu, 13 March 2014, 11:00

To be solved in Class

1. *Quantifiying the loss when using the wrong code.* Prove that when designing a code with length $\ell(X)$, believing that the distribution is Q_X when the true distribution is P_X incurs a penality of $D(P_X||Q_X)$ in the average description length.

   More formally, when we think that the true distribution is Q_X, we set the code lengths to $\ell_x = \lceil \log \frac{1}{Q_X(x)} \rceil$. However, the true distribution turns out to be P_X and hence, the expected codeword length is $\mathbb{E}_{P_X}[\ell(X)] = \sum_x P_X(x)\ell_x$. Prove that

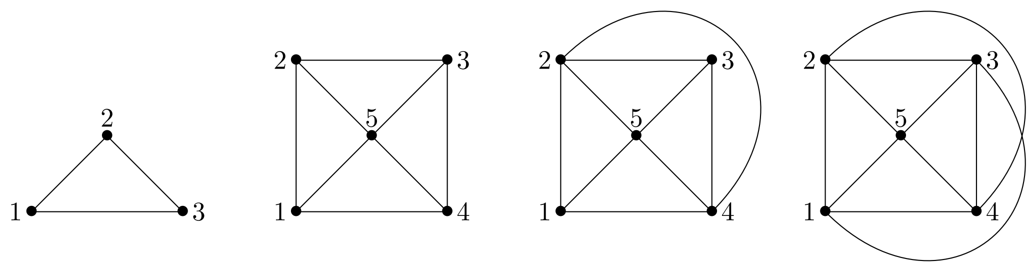
$$H(P_X) + D(P_X||Q_X) \leq \mathbb{E}_{P_X}[\ell(X)] \leq H(P_X) + D(P_X||Q_X) + 1$$

   Last time, we had trouble to show the lower bound as we probably did not consider explicitly that we set the ℓ_x to $\lceil \log \frac{1}{Q_X(x)} \rceil$. Let us try again with that extra clarification.

Homework

1. For each of the channels below, give the corresponding confusability graph.
   (a) [1 point] $\mathcal{X} = \{1,2,3,4,5\}$, $\mathcal{Y} = \{a,b,c\}$, $p_{Y|X}(a|1) = p_{Y|X}(b|1) = p_{Y|X}(a|2) = p_{Y|X}(b|2) = \frac{1}{2}$, $p_{Y|X}(b|3) = \frac{1}{3}$, $p_{Y|X}(c|3) = \frac{2}{3}$, $p_{Y|X}(c|4) = p_{Y|X}(c|5) = 1$.
   (b) [1 point] $\mathcal{X} = \{1,2,3,4,5\}$, $\mathcal{Y} = \{a,b,c,d\}$, $p_{Y|X}(a|2) = p_{Y|X}(b|2) = p_{Y|X}(c|2) = p_{Y|X}(a|4) = p_{Y|X}(c|4) = p_{Y|X}(d|4) = \frac{1}{3}$, $p_{Y|X}(b|3) = p_{Y|X}(c|3) = \frac{1}{2}$, $p_{Y|X}(a|1) = p_{Y|X}(d|5) = 1$.

2. For each of the confusability graphs below, describe one of the possible corresponding channels. Try to minimize number of output symbols you are using.



(a) (b) (c) (d)

(a) [1 point]

(b) [1 point]

(c) [1 point]

(d) [1 point]

(e) [2 points] Can you argue that you reached the minimal number of outputs in (a), (b), (c), (d) ?

(f) [1 point] Show that for any confusability graph G **with no isolated vertices**, there exists a corresponding channel with $|E(G)|$ output symbols.

3. *Shannon capacity of the complete graph.* A graph G with n vertices $V(G) = \{1, 2, \ldots, n\}$ is called *complete* if it has edges between any two vertices, i.e. $\forall i \neq j : ij \in E(G)$.

   (a) [2 points] Compute $\alpha(K_n)$, the independence number of the complete graph.

   (b) [2 points] Show that $K_n \boxtimes K_n = K_{n^2}$.

   (c) [2 points] Use (a) and (b) to prove that the Shannon capacity of K_n is 0. Note that this result formally confirms the intuition that channels whose confusability graphs are complete are useless for zero-error communication, because all symbols can possibly be confused with each other.

4. *Disjoint graphs.* For two graphs G and H, the graph $G + H$ is defined as the disjoint union of the two graphs[1]. Formally, assuming without loss of generality that $V(G) \cap V(H) = \emptyset$, then $V(G + H) = V(G) \cup V(H)$ and $E(G + H) = E(G) \cup E(H)$.

   For a graph G, the disjoint union of t copies of G is denoted as G^{+t}. Similarly, we write $G^{\boxtimes t}$ for the t-time strong product of G with itself.

   (a) [2 points] Prove that $\alpha(G + H) = \alpha(G) + \alpha(H)$.

   (b) [Bonus: +4 points] Prove that for any three graphs G, H, L, it holds that

   $$(G + H) \boxtimes L = (G \boxtimes L) + (H \boxtimes L)$$

   and for the same reason, it also holds that

   $$G \boxtimes (H + L) = (G \boxtimes H) + (G \boxtimes L)$$

   .

   (c) [4 points] Use (b) to derive that for any natural number $k \in \mathbb{N}$, $(G + G)^{\boxtimes k} = (G^{\boxtimes k})^{+2^k}$.

5. Let $\mathcal{X} = \mathcal{Y} = \{1, 2, 3, 4, 5, 6\}$. In this exercise, we compute the zero-error Shannon capacity of the noisy channel with transition probabilities $P_{Y|X}(y|x) = 1/3$ if and only if $x \equiv y \mod 2$.

   (a) [2 points] Give the confusability graph G of the noisy channel $P_{Y|X}$ described above.

   (b) [4 points] Use 4.(c) and 3.(a) and 3.(b) to show that the Shannon capacity of G is 1.

[1] You can think of $G + H$ as G and H "next to each other".