ECE 499/599 Data Compression/Information Theory
Spring 06

Dr. Thinh Nguyen

Homework 1 Solution
Due 04/18/06 at the beginning of the class

Problem 1:

a) 2 bits.
b) 1.75 bits.
c) 1.739818 bits.

Problem 2: Do problem 4 in chapter 2 of the textbook.

First show that f(x) = -xlogx – (a-x)log(a-x) is maximum for x = a/2, then use this fact to prove whether H(Q) is greater or smaller than H(P). This can be done by taking first derivative to estabilish a local extreme point at a/2. Then, taking the second derivative showing f''(a/2) > 0, hence f(a/2) is maximum. We then proceed as follows.

$$\begin{aligned} H_Q - H_P &= -\sum_{i=1}^{m} q_i \log_2 q_i + \sum_{i=1}^{m} p_i \log_2 p_i \\ &= -q_{j-1} \log_2 q_{j-1} - q_j \log_2 q_j + p_{j-1} \log_2 p_{j-1} + p_j \log_2 p_j \end{aligned}$$

Given a function

$$f_a(x) = -x \log x - (a - x) \log(a - x)$$

we can easily show that $f_a(x)$ is maximum for $x = \frac{a}{2}$ Let

$$p_{j-1} + p_j = c$$

then

$$q_{j-1} = q_j = \frac{c}{2}$$

Then

$$\begin{aligned} H_Q - H_P &= -\frac{c}{2}\log_2\frac{c}{2} - \frac{c}{2}\log_2\frac{c}{2} + p_j\log_2 p_j + (c-p_j)\log_2(c-p_j) \\ &= f_c(\frac{c}{2}) - f_c(p_j) \\ &\geq 0 \end{aligned}$$

Therefore $H_Q \geq H_P$.

Another neat proof:

We know the highest entropy is obtained when all the probabilities are equal to each other. In this case, we only need to consider the amount of information of {qj-1,qj} and {pj-1, pj} since the probabilities of other symbols are identical for p and q. Since qj-1 = qj, {qj-1,qj} must be maximum. Since entropy is additive (with statistical weight) , we have H(Q) > H(P).

Note: I will give bonus points for anyone who comes up with elegant proof like one above.

Problem 3: For a certain exam, 75% of the participating students in the exam pass, 25% do not pass. Of the students who have passed, 10% own a car. Of the students who have failed, 50% own a car.

a) How much information does one receive if one is told the result of a student's exam?

There are four possible situations, passing, not passing, owning a car, not owning a car, are denoted by s, s', c and c', repsectively. When the result of an exam is announced, this delivers an amount of information

H(result) = -p(s)logp(s) – p(s')logp(s') = -(3/4)log(3/4) – (1/4) log(1/4) = 0.81 bit

b) How much information is contained in the announcement of a student who has passed that he does or does not have a car?

If a student, who has passed, announces whether or not he has a car, then there are two possibilities c and c' with the given probabilities. Thus,

H(car owner/passed) = -p(c|s)log[p(c|s)] – p(c'|s)log[p(c'|s)] =

-(1/10)log(1/10) –(9/10)log(9/10) = 0.47 bits

c) How much uncertainty remains concerning the car ownership of a student if he announces the result of his exam?

There are four possibilities in total. The corresponding probabilities are

p(s,c) = (3/4)(1/10) = 3/40
p(s,c') = (3/4)(9/10) = 27/40
p(s',c) = (1/4)(1/2)=1/8
p(s',c') = (1/4)(1/2) = 1/8

The amount of information that is delivered by announcing the result of the exam as well as possible car ownership is then

H(car ownership, result) = (-3/40)log(3/40) – (27/40)log(27/40) – 2(1/8)log(1/8) = 1.41 bits

The remaining uncertainty about car ownership, if the result of the exam is given, is then
H(car ownership/result) = H(char ownership, result) – H(result) = 1.41-.81 = 0.6 bits.

One can also obtain this result by calculating the conditional amount of information directly as per the definition for the conditional amount of information.