Boğaziçi University
Department of Physics

Phys 497 Spring 2008

Problem Set #3

Due in class Tuesday, 6 May 2008

Problem 1: (20 pts)
The average information per symbol from a source has been expressed as follows:

$$\sum_i P(A_i) \log_2 \frac{1}{P(A_i)}$$

If you consider the contribution from the term of a single event, say, A_k, it can be expressed as:

$$P(A_k) \log_2 \frac{1}{P(A_k)}$$

What should $P(A_k)$ be so that this partial contribution is the maximum possible contribution? (This simply amounts to finding the maximum value of $f(x) = x \log_2(1/x)$ between 0 and 1!)

Problem 2: (40 pts)
Consider a source which outputs on of three symbols, and each ouput is assumed to be independent. Let us call the symbols A, B, C.

a) With no knowledge about the probability of each symbol, we assume they are all equally likely. What is the entropy of the source in this case?

b) Using a fixed length code, how many bits does it take to encode one symbol on the average?

c) Will Huffman coding help in this case? Invent a codebook using the Huffman coding algorithm, and find out how many bits it takes to code one symbol in that case.

d) Now, we are given the following information: $P(A) = 7/8$, $P(B) = 1/16$, $P(C) = 1/16$. What is the entropy of the source now?

e) Generate a Huffman code for this case as well, and find how many bits it takes to encode one symbol in this case.

Note that the answer of part **d** is below 0.75, making this a quite low-entropy source. However, your answer to part **e** is above 1.0. In principle, you should be able to transmit more than one symbol per bit, but with this Huffman code, that does not seem possible...

f) Can you go over one symbol per bit on the average with any Huffman codebook? Why? / Why not?

g) Find a coding method so that you can transmit more than one symbol per bit on the average. (You need not worry about the end of the message, you can assume the symbols go on forever.)

h) **Optional challenge!** Find a coding method to use at most 0.75 bits per symbol on the average.

Problem 3: (40 pts)
Consider a binary channel, which is not symmetric. The input symbols are 0 and 1, and the associated events we call A_0 and A_1. The output symbols are the same, and the output events are called B_0 and B_1. The source produces twice as many ones as zeroes, thus we know that $P(A_0) = 1/3$ and $P(A_1) = 2/3$.

The channel is asymmetric, as mentioned. Thus, $P(B_0|A_0) = 4/5$, and $P(B_1|A_1) = 3/5$. (This is enough information about the channel!)

a) What is your uncertainty U before you observe an output event? That is, what is U_{before}?

b) Assuming that you have observed a zero at the output (that is, event B_0 has been observed), what is your uncertainty now ($U_{after}(B_0)$)?

c) Which is greater? Did your uncertainty increase or decrease by observing B_0? Try to explain what is going on in your own words.

d) Also calcualte $U_{after}(B_1)$ and M. Is M greater than (or equal to) zero as promised?

e) Now, let the *source* be symmetric, that is, let $P(A_0) = 1/2$ and $P(A_1) = 1/2$, and repeat parts **b** and **c.**.

Problem 4: (40 pts)
You have estimated that the probability of there being a fire in your home any given night is $p = 10^{-6}$. You decide to install a fire alarm system at home. You are given the following information about the alarm system:

During any night, given there is a fire, the alarm will sound with 99.99% probability. The alarm is very sensitive, so during any night, given there is no fire, there is a 1% probability that the alarm will sound anyway.

a) Before solving the rest of the problem, state whether or not you think this is a useful alarm system.

b) During a night, there is no alarm. What is the probability that you will burn to death anyway?

c) You wake up with the sound of the alarm, at night. What is the probability that there is a fire?

d) Now do you think this is a useful alarm system?

e) You move to a different home, which contains lots of combustible material, and now you estimate the probability that there is a fire at home to be 10^{-2}. Repeat parts **b**, **c**, and **d**.