

## Physics 239 Spring 2016 Assignment 3

Due 11am Tuesday, April 26, 2016

1. **Strong subadditivity, the classical case.** [From Barnett] Prove *strong subadditivity* of the Shannon entropy: for any distribution on three random variables,

$$H(ABC) + H(B) \leq H(AB) + H(BC).$$

(The corresponding statement about the von Neumann entropy is not so easy to show.)

Hint:  $q(a, b, c) \equiv \frac{p(a,b)p(b,c)}{p(b)}$  is a perfectly cromulent probability distribution on  $ABC$ .

What is the name for the situation when equality holds? Write the condition for equality in terms of the conditional mutual information  $I(A : C|B)$ .

2. **Symbol coding problem.**

You are a mad scientist, but a sloppy one. You have 127 identical-looking jars of liquid, and you have forgotten which one is the poison one. You have at your disposal 7 rats on whom your poor moral compass will allow you to test the liquids. However (the rats have a strong social network and excellent spies) you only get one shot: the rats must drink all at once (or they will catch on to what is happening and revolt). You may mix the liquids in separate containers; any rat that drinks any amount of poison will turn bright orange. Design a protocol to uniquely identify the poison jar.

3. **Huffman code.**

Make the Huffman code for the probability distribution  $p(x) = .5, .2, .15, .1, .05$ .

Compare the average word length to the Shannon entropy.

Bonus: what property of the distribution determines the deviation from optimality?

4. **Huffman code decryption problem.**

0000110111111000100101100101001100 1111100011  
11001011010101101011011100000011 01011010000011011 00101000 00000011

```

010000010111100101011100001101 10000101011011110100001001101
01011010000011011, 1000111001011000010111 010110100000110111000
(10111000001011000100110000010101 01011110000000001011110111001
10000010000100100) 00001001011 101000011111110111001 111110000000011
00000011010101 10000101011011110100001001101 01011010000011011
101000011010110000101101000011 1111100011010101 00001001011
1011101001001011 10101110101100010010000101010001101011.
1001011100001100001001010111100 00010110010100101000110101110111:
101010010001100011 00000011 00001101010100010010010010111110010111
1010101001001 100001001101000100000100011010100
000100101110110000000111111011010111 1011100000011010101
101000010000011010101 10000101011000111011101111000
0010001110000001010010100101100011 101000010101
000010010010111110010111011111001001011 010110100000110111000
0000001100011 1101011101110001010111101-0001001010100100001000
010110100000110111000.

```

Hint: I used the letter frequencies from *The Origin of Species*.

You might want to use `Mathematica` to do this problem!

5. **Analogy with strong-disorder RG.** [open ended, more optional question]

Test or decide the following consequence suggested by the analogy between Huffman coding and strong-disorder RG: The optimality of the Huffman code is better when the distribution is broader. A special case is the claim that the Huffman code is worst when all the probabilities are the same. Note that the outcome of the Huffman algorithm in this case depends on the number of elements of the alphabet.

Measure the optimality by  $\langle \ell \rangle - H[p]$  (or maybe  $\frac{\langle \ell \rangle - H[p]}{H[p]}$ ?).

6. **Binary symmetric channel.**

For the binary symmetric channel  $ABE$  defined in lecture, with  $p(a) = (p, 1 - p)_a$ ,  $p(e) = (q, 1 - q)_e$  and  $b = (a + e)_2$ , find all the quantities  $p(a, b)$ ,  $p(b)$ ,  $p(b|a)$ ,  $p(a|b)$  and  $H(B)$ ,  $H(B|A)$ ,  $I(B : A)$ ,  $I(B : A|E)$ .