365 Written Assignment 2

Alen Mehmedbegovic

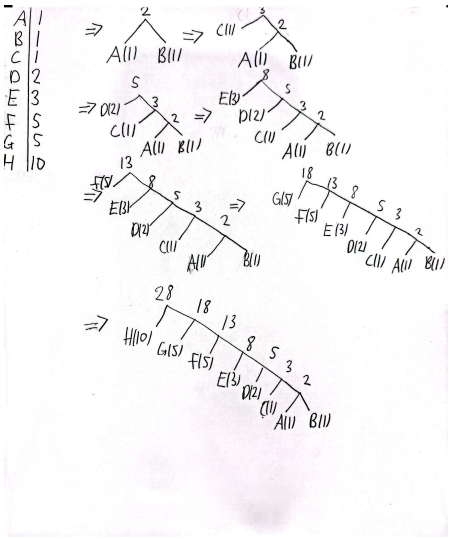
March 11th, 2024

1 7.8 Q1

50/50 chance for black or white.
prob(white) = 1/2, prob(black) = 1/2.
If our "checkboard" image is n by n, we would have n^2 entries in total. For either of the values, the probability distribution is the same. Meaning our entropy would be

$$\sum_{i=1}^{n^2} -\frac{1}{2} * log_2 \frac{1}{2} = n^2 * -\frac{1}{2} * log_2 \frac{1}{2} = n^2 * \frac{1}{2} * log_2 2 = n^2 * \frac{1}{2} = \frac{n^2}{2}$$

2 7.8 Q2



3 7.8 Q3

This is an 8x8 grid meaning 64 total entries.
8 99's, 8 20's, 16 50's, 32 0's.
pr(99) = 1/8, pr(20) = 1/8, pr(50) = 1/4, pr(0) = 1/2.

$$Entropy = -\frac{1}{8} * log_2\frac{1}{8} - \frac{1}{8} * log_2\frac{1}{8} - \frac{1}{4} * log_2\frac{1}{4} - \frac{1}{2} * log_2\frac{1}{2}$$

$$Entropy = \frac{1}{8} * log_2 8 + \frac{1}{8} * log_2 8 + \frac{1}{4} * log_2 4 + \frac{1}{2} * log_2 2$$

$$Entropy = \frac{1}{8} * 3 + \frac{1}{8} * 3 + \frac{1}{4} * 2 + \frac{1}{2} * 1$$

$$Entropy = \frac{3}{8} + \frac{3}{8} + \frac{2}{4} + \frac{1}{2}$$

$$Entropy = \frac{7}{4}$$

4 7.8 Q7b

(change ii to "If Arithmetic coding is used, what are the lower and higher bounds of the interval after encoding BBB and ABC, respectively")

Original bounds from 0.0 to 1.0 are A: [0,0.5), B: [0.5,0.9), C: [0.9,1)

BBB:
First B changes bounds to 0.5-0.9. Now the range is A: [0.5,0.7), B: [0.7,0.86), C: [0.86,0.90).
Second B changes bounds to 0.7-0.86. Now the range is A: [0.7,0.78), B: [0.78,0.844), C: [0.844,0.86).
Last B changes bounds to 0.78-0.844. As this is the last symbol and the question isn't asking the specific symbol ranges within the bound, we can answer with lowerbound: 0.78, higher bound: 0.844

ABC:
A changes bounds to 0-0.5. Now the range is A: [0,0.25), B: [0.25,0.45), C: [0.45,0.5).
B changes bounds to 0.25-0.45. Now the range is A: [0.25,0.35), B: [0.35,0.43), C: [0.43,0.45).
C changes bounds to 0.43-0.45. As this is the last symbol and the question isn't asking the specific symbol ranges within the bound, we can answer with lowerbound: 0.43, higher bound: 0.45

5 EX 1

Given sequence 00100011001011011101001110111011

1) String length: 32. Number of zeroes: 14, ones: 18. Pr(0) = 7/16, Pr(1) = 9/16.

$$Entropy = -\frac{7}{16} * log_2 \frac{7}{16} - \frac{9}{16} * log_2 \frac{9}{16}$$

2) Pr(0,0) = 49/256, Pr(0,1) = 63/256 , Pr(1,0) = 63/256, Pr(1,1) = 81/256

$$Secondorderentropy = -\frac{49}{256} * log_2 \frac{49}{256} - \frac{63}{256} * log_2 \frac{63}{256} - \frac{63}{256} * log_2 \frac{63}{256} - \frac{81}{256} * log_2 \frac{81}{256}$$

3) 2 bits needed per symbol. Occurrence of 00: 4, 01: 2, 10: 4, 11: 6. With 16 total symbols meaning
Pr(00) = 1/4, Pr(01) = 1/8, Pr(10) = 1/4, Pr(11) = 3/8.
Which leads us to

$$L = \frac{1}{4} * 2 + \frac{1}{8} * 2 + \frac{1}{4} * 2 + \frac{3}{8} * 2 = 2$$

6 Example 2

Given input sequence XZZXYYXXYZYY, and assume that the initial dictionary is: 0 X, 1 Y, 2 Z

Output: 0 2 2 0 1 1 0 6 2 7

Code:
3 XZ
4 ZZ
5 ZX
6 XY
7 YY
8 YX
9 ZY