Name: ____

Massachusetts Institute of Technology

Department of Electrical Engineering and Computer Science
Department of Mechanical Engineering

6.050J/2.110J **Information and Entropy** **Spring 2005**

Issued: April 12, 2005, 12:00 PM **Quiz** **Due: April 12, 2005, 1:00 PM**

Note: Please make sure you follow this instructions

- **Write your name** at the top of **each** page in the space provided ⟶ **worth 1%** of the grade.
- You may bring and use any notes you want on **one 8 1/2 x 11 sheet of paper**. You may also use a hand calculator, but the quiz is designed so that you do not need one *(see the table of logarithms on the last page which you may tear off and use for reference).*
- Put your answers on this sheet and turn it in along with any calculations you do on other pieces of paper. Be sure to put your name on all pages you turn in.

Problem 1: (25%)

In computing cumulative ratings (MIT's version of GPA) grades of A count as 5, grades of B as 4, C as 3, and so on. One semester Professor Plum gave only grades of A, B, and C in his course "Process of Elimination." Later he remembered that the average grade he assigned, using the numerical scale above, was 4.30 but he did not have a clue about the actual number of A, B, or C grades. Without making any further assumptions, you want to estimate the fraction of A, B, and C grades (i.e., the probabilities that a student selected at random would have received a particular grade). For simplicity you can refer to these unknown probabilities as A, B, and C, rather than $p(A)$, $p(B)$, and $p(C)$.

a. What values of A, B, and C are possible?

$\underset{A_{\text{Min}}}{\_\_\_\_\_\_\_\_} \leq A \leq \underset{A_{\text{Max}}}{\_\_\_\_\_\_\_\_}$ $\underset{B_{\text{Min}}}{\_\_\_\_\_\_\_\_} \leq B \leq \underset{B_{\text{Max}}}{\_\_\_\_\_\_\_\_}$ $\underset{C_{\text{Min}}}{\_\_\_\_\_\_\_\_} \leq C \leq \underset{C_{\text{Max}}}{\_\_\_\_\_\_\_\_}$

b. Naturally you decide to use the Principle of Maximum Entropy to estimate A, B, and C. Express the entropy as a function of a single probability (any one of the three will do). You do not need to actually solve for the probabilities.

Entropy = ____

Problem 2: (20%)

According to the Registrar, Mrs. White, the grades that semester were actually 50% A, 30% B, and 20% C. You decide to produce an encoding to store the thousands of grades assigned by Professor Plum efficiently. You want to use a Huffman code for this purpose, because you can represent the grades using fewer bits on average than two per grade, which would be required for a fixed-length code.

a. You calculate the entropy per assigned grade of this source of data. It is one of the following. Circle which one it is.

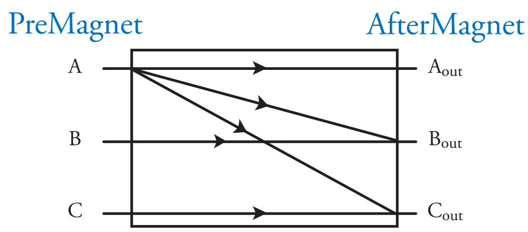
0.92 bits 1.12 bits 1.32 bits 1.49 bits 1.74 bits 2.32 bits

b. Give your Huffman Code and the resulting average codeword length

A _________ B _________ C _________ Average Length _________

Problem 3: (20%)

Professor Plum's teaching assistant, Miss Scarlet, accidentally damaged a floppy disk containing the grades (she did it in the lab with a magnet). Many random errors were introduced – 40% the grades of A were shifted to B, and 20% of the grades of A became C. Grades of B and C were not affected.



The police investigator, Colonel Mustard, treated this incident as a communication channel with errors. Using the probability distribution from Problem 2 ($p(A) = 0.5$, $p(B) = 0.3$, and $p(C) = 0.2$), he calculated the probability distribution of the grades on the damaged disk A_{out}, B_{out}, and C_{out} and the resulting entropy at the output J. He reasoned that if the output information is the same as the input information there was nothing lost and no need to make any corrections, so the case could be closed. What did he find? Do you agree with his reasoning?

$p(A_{out})$ _________ $p(B_{out})$ _________ $p(C_{out})$ _________ J _________

Do you agree? _________ Why?:____

____

Problem 4: (34%)

a. One of the students, Mr. Green, was told that this defective floppy disk had his grade as B. He asked you to estimate the probability that his originally assigned grade had been A, B, or C. You should make use of the original probability distribution, the error probabilities caused by the accident, and the fact that his grade is now shown as a B.

$p(A)$ _________ $p(B)$ _________ $p(C)$ _________

b. You decide to review Colonel Mustard's conclusions. Calculate the residual uncertainty U about the original grade for each of the three possible output grades,$U_{before}(A_{out})$, $U_{before}(B_{out})$, and$U_{before}(C_{out})$, and from those the loss L and mutual information M.

$U_{before}(A_{out})$_________ $U_{before}(B_{out})$_________ $U_{before}(C_{out})$_________

L_________ M _________

Logarithm and Entropy Table

This page is provided so that you may rip it off the quiz to use as a separate reference table. In Table 1, the entropy $S = p\log(1/p) + (1-p)\log_2(1/(1-p))$.

p	1/8	1/5	1/4	3/10	1/3	3/8	2/5	1/2	3/5	5/8	2/3	7/10	3/4	4/5	7/8
$\log_2(\frac{1}{p})$	3.00	2.32	2.00	1.74	1.58	1.42	1.32	1.00	0.74	0.68	0.58	0.51	0.42	0.32	0.18
S	0.54	0.72	0.81	0.88	0.92	0.95	0.97	1.00	0.97	0.95	0.92	0.88	0.81	0.72	0.54

Table 1: Table of logarithms in base 2 and entropy in bits