Name (1%): Your 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**

Final Exam Solutions

Solution to Problem 1: Who's Who (14%)

Avogadro	Bayes	Boltzmann	Boole	Carnot	Gibbs	Hamming
Huffman	Jaynes	Joule	Kelvin	Kraft	Lempel	Maxwell
Morse	Reed	Schrödinger	Shannon	Solomon	Welsh	Ziv

Solution to Problem 1, part a.

Kelvin, a 19-th century physicist, has a unit of temperature named after him.

Solution to Problem 1, part b.

Reportedly Schrödinger secluded himself in his mountain cabin, with pearls in his ears to muffle the sound, and his girlfriend in his bed to inspire him, and came up with an equation to calculate the wave function of a quantum system.

Solution to Problem 1, part c.

The measure of how much two bit strings of equal length differ is named after Hamming.

Solution to Problem 1, part d.

While an MIT student in the 1940s, Kraft invented a famous inequality for his master's thesis.

Solution to Problem 1, part e.

Ziv, Lempel, Welsh was one of three engineers whose name is used for a lossless compression technique for the popular GIF image format.

Solution to Problem 1, part f.

A famous inequality is named after Gibbs, who received the first doctorate in engineering in America, and later was on the faculty at Yale University.

Solution to Problem 1, part g.

The Boltzmann constant is approximately 1.38×10^-23 Joules per Kelvin.

Solution to Problem 1, part h.

The military engineer Carnot showed that all reversible heat engines have the same efficiency.

Solution to Problem 1, part i.

The channel capacity theorem proved by ___Shannon___ states a possibility, not how to achieve it.

Solution to Problem 1, part j.

The algebra of binary numbers is named after the mathematician ___Boole___, who had been a childhood prodigy in Latin, publishing at the age of 12.

Solution to Problem 1, part k.

___Maxwell___ conceived a Demon to show the statistical nature of the Second Law of Thermodynamics.

Solution to Problem 1, part l.

___Jaynes___ promoted the Principle of Maximum Entropy as an unbiased way of assigning probabilities.

Solution to Problem 1, part m.

During an ocean journey ___Morse___ heard that electricity could be transmitted instantaneously, and in a fit of creativity invented a code to use it to transmit arbitrary information.

Solution to Problem 1, part n.

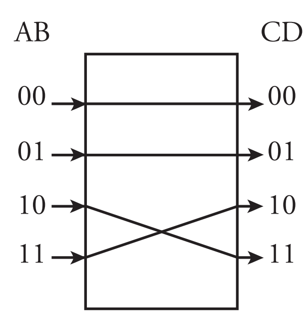
Over one weekend ___Huffman___ solved a problem posed in a problem set for an MIT graduate subject, thereby inventing the shortest possible variable length code.

Solution to Problem 2: Under Control (5%)

Your company has just purchased a large number of $C-NOT$ (controlled-not) logic gates which are really just XOR (exclusive or) gates with an extra output. They have two inputs A and B and two outputs C and D . One output, C, is merely a copy of A, and the other output is B if $A = 0$ or $NOT B$ if $A = 1$. In other words, the output D is B possibly run through a NOT gate depending on the input A. Your boss wants you to design all your circuits using only $C-NOT$ gates, so the company can save the cost of maintaining different components in its inventory. You wonder about the properties of this gate.

You start off by modeling the gate in terms of its transition probabilities, and noting some basic properties.

You know that a gate is reversible if its input can be inferred exactly from its output. Since you know nothing about how the gate might be used, you assume the four possible input combinations are equally probable.



Solution to Problem 2, part a.

In the diagram at the right, show the transition probabilities.

Solution to Problem 2, part b.

What is the input information in bits? ______2______

Solution to Problem 2, part c.

What is the output information in bits? ______2______

Solution to Problem 2, part d.

What is the loss in bits? ____0____

Solution to Problem 2, part e.

What is the noise in bits? ____0____

Solution to Problem 2, part f.

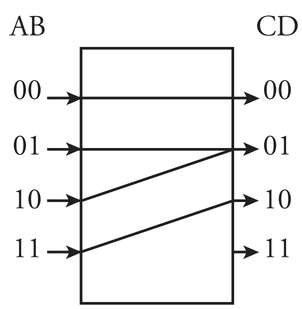
What is the mutual information in bits? ____2____

Solution to Problem 2, part g.

Is this gate reversible (yes or no)? ____yes____

Solution to Problem 3: Out of Control (5%)

You have concluded that the $C - NOT$ gate from Problem 2 will not satisfy your needs, and look around for another. The incoming inspector tells you that some of the gates don't work right, and calls them $\mathbf{NOT - C - NOT}$ gates. They behave like $C-NOT$ gates except that they have a defective power supply which keeps the C output from being 1 when $D = 1$ even though the $C - NOT$ logic might call for them both to be 1 . This inspector thinks these gates are worthless but asks you want you think. You repeat your analysis for this gate, again assuming the four possible input combinations are equally probable.



Solution to Problem 3, part a.

In the diagram at the right, show the transition probabilities.

Solution to Problem 3, part b.

What is the input information in bits? ____2____

Solution to Problem 3, part c.

What is the output information in bits? ____1.5____

Solution to Problem 3, part d.

What is the loss in bits? ____0.5____

Solution to Problem 3, part e.

What is the noise in bits? ____0____

Solution to Problem 3, part f.

What is the mutual information in bits? ____1.5____

Solution to Problem 3, part g.

Is this gate reversible (yes or no)? ____No____

Solution to Problem 4: MIT Customer Complaint Department (15%)

You have recently been elected President of the UA, and it is your job to transmit student complaints to MIT President Susan Hockfield so they (hopefully) can be addressed rapidly. According to the UA's research, all student complaints fall into one of six categories, with percentages shown:

%	Complaint
50%	Not enough homework
30%	Campus dining options too diverse
10%	Tuition too low
5%	Administration too attentive
5%	Classes too easy

Unfortunately, Hockfield doesn't have much time, so she instructs you to only send very short messages to her. Because you've taken 6.050 you know about coding schemes, so you decide to encode the complaints above in a Huffman code.

Solution to Problem 4, part a.

Design a Huffman code for the complaints above.

Complaint	Code
Not enough homework	1
Campus dining options too diverse	01
Tuition too low	001
Administration too attentive	0001
Classes too easy	0000

Solution to Problem 4, part b.

What is the average number of bits to send one complaint?

Average # of bits/complaint: ____1.8____

Solution to Problem 5: Zero Power Cooler (30%)

Having landed a high-paying job in Boston, you decide to build a new house for yourself. You vow to use only energy-efficient techniques, based on what you know about information and entropy. Of course your house needs heating in the winter, cooling in the summer, and beverage refrigeration all year long. You decide to design a device that cools the beverage storage room without requiring any work and therefore no electricity bill.

Of course you are familiar with the Carnot efficiency results for reversible heat exchangers. You know that ordinarily heat engines, heat pumps, and refrigerators operate between two reservoirs at two different temperatures, and either produce or require work. But in your case you have reservoirs at three different temperatures (storage room, house, and outside). You wonder whether a reversible heat exchange system could take heat out of the storage room and either put heat into or take heat from the other two reservoirs, with no work.

Solution to Problem 5, part a.

First consider the winter. Model the outside environment as a large heat reservoir at $275K$ (about 2^oC or 35^oF). You want your house to be at $295K$ (about 22^oC or 72^oF) and the storage room at $280K$ (about 7^oC or 44^oF).

Is such a system possible, at least in principle? Or would it violate the Second Law of Thermodynamics?

Yes, it is possible

If it is possible, give the heat that would be exchanged with each of the three reservoirs if 1 Joule of heat is taken from the storage room at $280K$. If it is not possible, briefly explain why.

The key idea here is to extract work from one process and introduce it in the other. For example, we can extract work from the heat flow from the storage room to the outside. Then we have:

$$Q_{outside} = Q_{1,storage} - W \tag{F–3}$$

$$\frac{W}{Q_{1,storage}} = \eta = \frac{T_{storage} - T_o}{T_{storage}} = .0179 \tag{F–4}$$

We want to reintroduce this work in the house as heat, so

$$Q_{house} = Q_{2,storage} + W \tag{F–5}$$

$$\frac{W}{Q_{house}} = \eta = \frac{T_{house} - T_{storage}}{T_{house}} = 0.051 \tag{F–6}$$

These equations explain the heat flow, in addition we are told that the heat extracted from the storage room is 1J, this means that, according to our setting: $Q_{1,storage} + Q_{2,storage} = 1J$. *Rearranging the equations to express everything in terms of one variable*

$$W = .051Q_{house} \tag{F–7}$$

$$Q_{2,storage} = .949Q_{house} \tag{F–8}$$

$$Q_{1,storage} = 2.847Q_{house} \tag{F–9}$$

$$Q_{outside} = 2.796Q_{house} \tag{F–10}$$

$$Q_{1,storage} + Q_{2,storage} = 3.796Q_{house} = 1 \tag{F–11}$$

(F–12)

replacing the value of Q_{house}

$$Q_{house} = .2634J \quad \text{(F–13)}$$
$$W = .0134J \quad \text{(F–14)}$$
$$Q_{2,storage} = .25J \quad \text{(F–15)}$$
$$Q_{1,storage} = .75J \quad \text{(F–16)}$$
$$Q_{outside} = .7366J \quad \text{(F–17)}$$
$$\text{(F–18)}$$

Solution to Problem 5, part b.

Next, consider the summer, when the outside temperature is $300K$ (about 27^oC or 80^oF). The house and storage room temperatures are the same as for the winter, namely $295K$ and $280K$. You wonder whether a reversible heat exchange system can be designed to take energy out of the storage area without requiring any work.

Is such a system possible, at least in principle? Or would it violate the Second Law of Thermodynamics? Yes it is possible

If it is possible, give the heat that would be exchanged with each of the three reservoirs if 1 Joule of heat is taken from the refrigerator at $280K$. It is is not possible, briefly explain why.

Here the solutions is a bit tricky, we mostly depend on assuming that the room is maintained at the temperature of 295 K, by an external air conditioning. Then the problem is much easier. We can, for example, leave the heat pump storage-house from the previous part, unachanged and reverse the heat engine storage outside. The new Heat flow equations then look like:

$$Q_{outside} = Q_{1,storage} - W \quad \text{(F–19)}$$
$$\frac{W}{Q_{out}} = \eta = \frac{T_o - T_{storage}}{T_o} = .0667 \quad \text{(F–20)}$$
$$Q_{house} = Q_{2,storage} + W \quad \text{(F–21)}$$
$$\frac{W}{Q_{house}} = \eta = \frac{T_{house} - T_{storage}}{T_{house}} = 0.051 \quad \text{(F–22)}$$
$$Q_{2,storage} - Q_{1,storage} = 1J \quad \text{(F–23)}$$

expressing everything in terms of Q_{house} again

$$W = .051Q_{house} \quad \text{(F–24)}$$
$$Q_{2,storage} = 0.949Q_{house} \quad \text{(F–25)}$$
$$Q_{1,storage} = 0.763Q_{house} \quad \text{(F–26)}$$
$$Q_{outside} = 0.814Q_{house} \quad \text{(F–27)}$$
$$Q_{2,storage} - Q_{1,storage} = .186Q_{house} = 1 \quad \text{(F–28)}$$
$$\text{(F–29)}$$

replacing the value of Q_{house}

$$Q_{house} = 5.376J \quad \text{(F–30)}$$
$$W = .274J \quad \text{(F–31)}$$
$$Q_{2,storage} = 5.102J \quad \text{(F–32)}$$
$$Q_{1,storage} = 4.102J \quad \text{(F–33)}$$
$$Q_{outside} = 4.376J \quad \text{(F–34)}$$
$$\text{(F–35)}$$

Solution to Problem 6: Casino (20%)

On your vacation you find yourself in a gambling casino that caters to nerds. One of the games catches your attention because you wonder about the probabilities. In this game, you pay a nickel (5 cents) to play, and then a coin is chosen at random from a big jar and given to you. The jar contains pennies (worth 1 cent), nickels (5 cents), and quarters (25 cents). This game, unlike slot machines, has some payout each time you play. You ask the owner of the casino about the probabilities P, N, and Q of the three coins (penny, nickel, and quarter) being selected, but the only information he gives you is that the average payout is 4 cents per play. You notice that the jar is so large that P, N, and Q remain the same all day long.

Solution to Problem 6, part a.

What values of P, N, and Q (nonnegative, no larger than 1) are possible?

$$\underset{P_{\text{Min}}}{\underline{\quad 0.25 \quad}} \leq P \leq \underset{P_{\text{Max}}}{\underline{\quad 0.875 \quad}} \qquad \underset{N_{\text{Min}}}{\underline{\quad 0 \quad}} \leq N \leq \underset{N_{\text{Max}}}{\underline{\quad 0.75 \quad}} \qquad \underset{Q_{\text{Min}}}{\underline{\quad 0 \quad}} \leq Q \leq \underset{Q_{\text{Max}}}{\underline{\quad 0.125 \quad}}$$

Solution to Problem 6, part b.

You decide to estimate P, N, and Q using the Principle of Maximum Entropy (if the owner made this choice it would be exciting to nerd customers, because they would gain the most information when learning which coin is chosen). (b) Start this estimate by eliminating P and N and writing the entropy (your uncertainty about the coin to be selected) as a function of Q:

$$\text{Entropy} = \underline{\quad -\left[Q\log(Q) + (0.75 - 6Q)\log(0.75 - 6Q) + (0.25 + 5Q)\log(0.25 + 5Q)\right] \quad}$$

Without a calculator you can't find the maximum of this expression, so instead you guess that it is about 0.5 bits.

Then you realize that the owner need not stock the jar using maximum entropy. He could set the probabilities so as to maximize the number of quarters paid out, while still keeping an average payout of 4 cents (this choice would be exciting to non-nerd customers hoping to strike it rich).

Solution to Problem 6, part c.

With this strategy, what are P, N, Q, and the entropy (to two decimal places)?
P = ___.875___, N = ___0___, Q = ___.125___, Entropy = ___.544 bits___.

Solution to Problem 6, part d.

Was your earlier quick guess for the entropy (when you thought the Principle of Maximum Entropy was used by the owner) a good one? Explain your answer.

it was not a good guess, because we have just found a guess with higher entropy, and so, that guess could not have been the maximum entropy one

Solution to Problem 6, part e.

(Extra credit) Another strategy for the owner would be to stock the jar with no quarters at all, so Q = 0 (this choice would appeal only to nerd customers). For this strategy, what are P, N, Q, and the entropy (to two decimal places)?

P = .25 , N = .75 , Q = 0 , Entropy = .815 bits .

Solution to Problem 7: The Telephone Game (30%)

The "Telephone Game" illustrates how correct information gets converted into false rumors. In the game one person (Alice) sends a message to another (Bob) through a chain of humans, each of whom potentially corrupts the message. Thus Alice tells person #1, who then tells person #2, and so on until the last person in the chain tells Bob. Then Bob and Alice announce their versions of the message, normally accompanied by amazement at how different they are.

Consider the case where a single bit is being passed and each person in the chain has a 20% probability of passing on a bit different from the one she received. Thus we model each person in the chain as a symmetric binary channel as shown in Figure F–3.

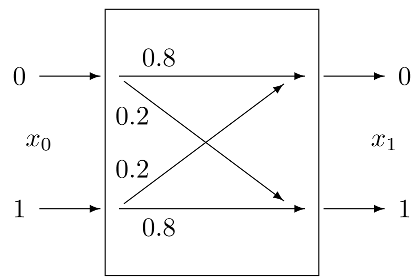


Figure F–3: Simple model of a person

The game is being demonstrated for you at a party one day. Alice and Bob take their positions at opposite ends of the chain, and Alice whispers the value of x_0 to person #1. You know that person #1, like the other members of the chain, has a 20% probability of changing the bit she hears. Parts *a.* and *b.* concern the model for this person, and parts *c.* and *d.* concern the behavior of the chain.

Solution to Problem 7, part a.

At first, you do not know what Alice has told person #1. Naturally, you express your state of knowledge in terms of the two probabilities $p(x_0 = 0)$ and $p(x_0 = 1)$. To avoid any unintended bias you use the Principle of Maximum Entropy to conclude that each of these probabilities is equal to 0.5. Then you calculate your uncertainty I_0 about the value of x_0, the output probabilities $p(x_1 = 0)$ and $p(x_1 = 1)$, and your uncertainty I_1 about the value of x_1. Then you calculate the channel noise N, channel loss L, and mutual information M, all in bits:

$p(x_0 = 0) =$ 0.5 $p(x_0 = 1) =$ 0.5 $I_0 =$ 1 bits

$p(x_1 = 0) =$ .5 $p(x_1 = 1) =$ .5 $I_1 =$ 1 bits

$N =$ .72 $L =$ .72 $M =$ .32

Next, consider the behavior of a cascade of independent identical channels representing the individuals passing the message, as illustrated in Figure F–4. Let x_k represent the output of the k-th channel.

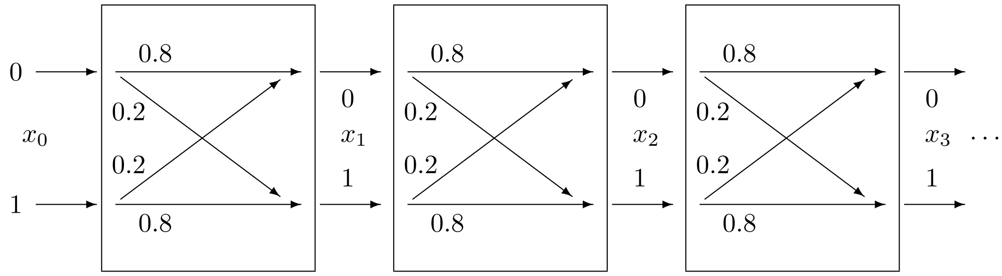


Figure F–4: Simple model of the telephone game

Solution to Problem 7, part b.

Now suppose that you know that Alice's bit is 0, so that $x_0 = 0$. Having just calculated probabilities for x_1, you wonder how much you know about the other values, x_2, x_3, x_4, $\ldots x_k \ldots$. Is it true that $p(x_k = 0) > 0.5$ for every channel output x_k? In other words, is every value passed along more likely to be 0 than 1? Write a paragraph defending your conclusion. If possible, make this an outline of a proof. You may find some version of the principle of induction helpful. (Pictures and equations are allowed in the paragraph, if needed.)

Yes, it is more likely to be zero than 1.

The way to see it is to acknowledge that at each step, noise will blur the output so that both outputs are possible, and probabilities of $x_k = 0$ decreases while that of $x_k = 1$ increases. In every new step the probability of $x_k = 0$ will decrease by 20% and increase by a 20% of the value of the complementary probability. That is, if p_k is the probability of $x_k = 0$, $p_{k+1} = p_k \times .8 + (1 - p_k) \times .2 = .6 \times p_k + .2$. As long as $p_k > 0.5$, $.6 \times p_k > .3$, and the resulting $p_{k+1} > .5$ as well. Sinc ewe know that $p_0 = 1$, we can automatically assert (by induction) that $p_k > .5$ for all values of k.

Solution to Problem 7, part c.

You conclude (correctly) that $p(x_k = 0)$ decreases as the message moves along the chain, i.e.,

$$p(x_0 = 0) > p(x_1 = 0) > p(x_2 = 0) > \ldots > p(x_k = 0) > p(x_{k+1} = 0) > \ldots \tag{F–36}$$

Let I_k be your uncertainty about x_k in bits. Does the uncertainty about x_k increase as you move through successive channels? In other words, is the following sequence of inequalities true?

$$I_0 < I_1 < I_2 < \ldots < I_k < I_{k+1} < \ldots \tag{F–37}$$

Write a paragraph defending your answer.

Yes it is true

In this case we can use part of the proof for the previous part. We would like to show that $p_k > p_{k+1}$, that this is true comes from replacing the value of p_{k+1} that we got in the previous problem statement

$$\begin{aligned} p_k &\overset{?}{>} .6 \times p_k + .2 \\ .4 \times p_k &\overset{?}{>} .2 \\ p_k &\overset{?}{>} .5 \qquad \text{(F–38)} \\ & \qquad\qquad \text{(F–39)} \end{aligned}$$

and this we know it is true from the solution to the previous part, so it follows that the probability p_k *decreases monotonically towards .5. In turn, we know that for a binary channel the maximum entropy is reached at p=.5, and that as we approach it the entropy grows monotonically, therefore it follows that* $I_k < I_{k+1}$*, which is what we wanted to prove.*