

Name (1%): _____

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science
Department of Mechanical Engineering

6.050J/2.110J

Information and Entropy

Spring 2003

Issued: May 23, 2003, 1:30 PM

Final Exam

Due: May 23, 2003, 4:30 PM

Note: Please write your name at the top of each page in the space provided. The last page may be removed and used as a reference table for the calculation of logarithms.

Problem 1: Knowin' the Nomenclature (9%)

From the list of terms below, select the one which best fits each of the following statements. Make sure your answers to each part are different.

- a. _____ can have units of Joules/Kelvin.
- b. _____ is a fixed length code.
- c. The part of the heat-engine cycle without change in entropy is _____.
- d. _____ is a reversible image compression technique.
- e. The _____ is calculated by taking the number of ones in the XOR of two binary strings.
- f. The energy values of an electron wavefunction in a square well are _____, not _____.
- g. A thermodynamic quantity that is the same for two systems in contact is _____.
- h. The maximum efficiency of a heat engine is named after _____.
- i. _____ was the most interesting aspect of this course.

adiabatic	continuous	extensive	Jaynes	power
analog	CRC	Finlayson	JPEG	reversible
ASCII	DCT	GIF	Karnaugh	TCP
Avogadro	digital	Hamming distance	Maxwell	universal
bit	discontinuous	heat	macroscopic	work
Boltzmann's constant	discrete	Huffman code	microscope	
Carnot	energy	intensive	Morse code	
channel capacity	entropy	irreversible	MP-3	
compression	Euler	isothermal	parity	

Table F-1: List of Terms

Problem 2: Behaving Badly (10%)

The new NANNY gate has been designed to tell whether a child has been misbehaving. In its intended application, input A of the gate is set to 1 if and only if the nanny has reminded the child of an important thing to do, like “sit up straight.” Input B is set to 1 if and only if the child has actually done that thing, whether or not reminded. The gate calculates whether scolding is called for, i.e., it returns 1 if the nanny has made such a request and the child has not complied.

- a. Give the truth table and transition diagram for this gate.

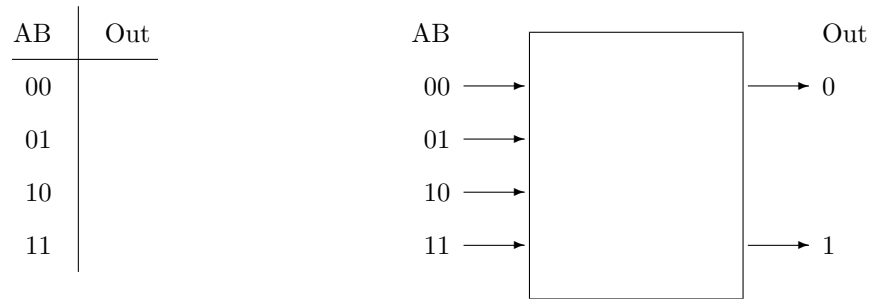


Figure F-1: NANNY Gate Truth Table and Transition Diagram

- b. Marketing research revealed that the nanny always reminds the child, and the child does the right thing 50% of the time. In this case, what are the probabilities of the four input events and two output events?.

AB=00: _____ AB=01: _____ AB=10: _____ AB=11: _____

Out=0: _____ Out=1: _____

- c. Nobody bought the product that used this gate, so the manufacturer is stuck with excess inventory. He wants to find other uses for the gate, and has asked you a number of questions about the gate.

Is it reversible? _____ Is it universal? _____ Is it deterministic? _____

- d. How many NANNY gates would you need to make a NOT gate? _____ How would you do it?

Problem 3: Parity, Code-rates, and Framing Errors (10%)

A particular application requires error detection. You have a long string of bits to transmit, and you decide to implement a parity scheme, where you insert an extra parity bit after each eight bits of data. If the number of bits is not a multiple of 8, the left over bits at the end get their own parity bit.

- a. What is the code rate of your scheme?

code rate: _____

- b. Sometimes, for some unfortunate reason, some of the bits are completely lost. This causes a loss of the bit, with no indication of its existence at all. For example, if the fourth bit in the string '00100' was erased completely, this would produce the string '0010', i.e., one bit shorter. With your error checking scheme is it always possible to detect this sort of error? Explain.

Problem 4: Architecture – Building Of Major Buildings (15%)

MIT has a history of erecting unusual buildings. Some of its buildings are innovative in a successful way, some are what the architect Frank Gehry calls “square boxes,” and the rest are just plain ugly. Call the probability that any building selected at random is in one of these three categories I , B , and U .

- a. MIT used to have a committee devoted to campus renovations and new campus buildings, called the Central Committee for Campus Planning. If you are told that a building was selected by the CCCP for replacement, but do not know which one, and have no knowledge of how many buildings on campus are in each of the three categories, what is your uncertainty I_{cat} in bits of which category the building is in, and what the three probabilities I , B , and U ?

$I =$ _____ $B =$ _____ $U =$ _____ $I_{cat} =$ _____

- b. The Union of Social Architects (the USA) rates buildings according to aesthetic criteria. It assigns a numerical score of 0 to an ugly building, 1 to a square box, and 5 to an innovative building. In 1950 it reported that the “aesthetic index,” the average aesthetic measure of MIT buildings, was 1.2. With this additional knowledge, you are able to estimate what the probabilities I , B , and U were in 1950 more accurately. First, you observe that I , B , and U cannot all range between 0 and 1, but instead are bounded by minimum and maximum values.

I_{min} _____ I_{max} _____ B_{min} _____ B_{max} _____ U_{min} _____ U_{max} _____

- c. The USA and the CCCP had an ideological disagreement, since the CCCP did not agree with what the USA considered good or bad architecture: the CCCP always marked a “square box” building for replacement, and always replaced it with an ugly building. In response to pressure from the Cambridge City Council, students, and passers-by alike, (all of whom agreed with the USA) the CCCP and the USA entered into an agreement, called Mutual Architecture Determination, such that every time the CCCP chose one building to be replaced, the USA got to choose the next building that would be replaced. Since the USA always decided to replace an ugly building with an successfully innovative building, MIT’s aesthetic index rose from 1.2 in 1950 to 1.6 in 1990. Was the maximum uncertainty about the category of a building selected at random in 1990 larger or smaller than in 1950?

Larger/smaller than in 1950?: _____

Reasoning: _____

- d. This period, which came to be known as the Campus Overall Layout Determination With Architecture Replacement, lasted until the CCCP was disbanded in 1990. It was then replaced with several committees, each responsible for different areas of campus. Since then, the USA has not selected any buildings, and the new committees select buildings at random and replace them with square boxes. On average, over a long time, will the aesthetic index go up or down?

Aesthetic index up/down?: _____

Problem 5: Green Eggs and Hamming (15%)

The new MIT campus hotspot, Buzz, has a new breakfast menu with three items: Green Eggs, Pancakes or the Greencake Combo (Green Eggs and Pancakes together). Since this is MIT, customers are required to place their orders in binary strings.

- a. How long a binary string (in bits) is needed to encode one order?

bits: _____

- b. Give a suitable code using this number of bits.

Green Eggs: _____ Pancakes: _____ Greencake Combo: _____

- c. The chef has noticed that several orders have been received incorrectly, because of the high ambient noise level. She wants you to design a new code for the three orders that will allow all orders to be understood even if one of the bits is wrong. From your experience with information and entropy, you realize that what she wants is a code that will allow single-error correction. What is the Hamming distance required between any pair of codewords to achieve this?

Minimum Hamming Distance: _____

- d. Give a suitable code with this Hamming distance using five bits per order.

Green Eggs: _____ Pancakes: _____ Greencake Combo: _____

- e. The chef also has to transmit all the orders, in the order they are received, to the stock room. She has a noise-free channel for this purpose, but it is very expensive and she wants to encode the orders in the least number of bits. You recommend a Huffman code because they were invented at MIT. Do you expect the average Huffman code length to be less than, equal to, or more than your answer to part (a)?

average bit length is less/equal/more: _____

- f. Experience has shown that students order Green Eggs 30% of the time, Pancakes 50% of the time, and the Greencake combo 20% of the time. Give a Huffman code that takes advantage of these probabilities.

Green Eggs: _____ Pancakes: _____ Greencake Combo: _____

- g. What is the average code length of this code in bits?

average length: _____

Problem 6: Rational Spammers (15%)

You have just installed the state-of-the-art spam filter I8SPAM on your computer. This program scans each incoming e-mail message and renders a judgement on whether it is spam or not. Unfortunately, this software is not perfect. It has a 25% false positive rate (discarding a valid message as spam) and a 50% false negative rate (letting a spam message through). In this context, “positive” means the program decided the message was spam.

You decide to model this filter as a communications channel, in which the inputs are e-mail messages and the outputs are the I8SPAM filter judgements. Denote the probability of a message being spam as S and it being a valid message as $V = 1 - S$. Denote the probabilities of the filter giving an output of positive as P and negative as $N = 1 - P$. Your experience is that 20% of your incoming messages are spam (call this the “spam rate”).

- a. Fill in the process-flow diagram below

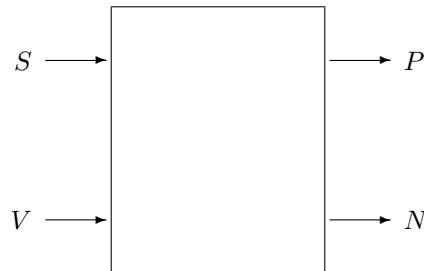


Figure F-2: Spam Filter Transition Diagram

- b. What are P and N ?

$$P = \underline{\hspace{2cm}} \quad N = \underline{\hspace{2cm}}$$

- c. Continuing with this model of the spam filter as a channel, determine the input information I_{in} , the output information I_{out} , the noise N , the loss L , and the mutual information M , all expressed in bits.

$$I_{in} = \underline{\hspace{1cm}} \quad I_{out} = \underline{\hspace{1cm}} \quad N = \underline{\hspace{1cm}} \quad L = \underline{\hspace{1cm}} \quad M = \underline{\hspace{1cm}}$$

- d. Your friend is a Course XV major and wonders whether your filter is worthwhile in terms of saving you time. She estimates that it costs you one minute to look at each message and discard it if you judge it to be spam. She also guesses that for each valid message that is discarded you will spend an extra five minutes getting the lost information another way. She points out that the spam filter is actually costing you time, rather than saving you time. You respond by saying that the spam is increasing, and it soon will be more than 20% of your incoming messages. When the spam rate is higher than some amount, the filter will save you time. What is that amount, above which the filter saves you time, assuming the false positive and negative rates remain what they are today?

$$\text{spam rate: } \underline{\hspace{2cm}}$$

Problem 7: I Am... Getting... So Hot, I Want to Open the Refrigerator Door (10%)

You have all experienced it. You open the door of your kitchen refrigerator and feel the room where you are standing cool down a little. One hot summer day Ben Bittiddle, thinking of this experience, decided to cool his bedroom down. He bought a kitchen refrigerator and put it in the corner of his bedroom, and turned it on but left the door open.

- a. Will the room cool down as a result?

yes/no: _____

- b. Why or why not?

As a second experiment, he moved the refrigerator next to his bed, oriented it so that the front was facing the bed. Again he turned it on but left the door open so he could look inside it from his pillow.

- c. Will the area of his bed cool down as a result?

yes/no: _____

- d. Why or why not?

Problem 8: A Dirty Sort of Business (15%)

You have just been hired by I.U. (Information Unlimited, Ltd.), founded by your illustrious TA, Mark, to work on their next generation heat engines, called Mark's Unlimited Device (MUD). Because their technology is based on quantum computation, all their products are reversible. Their Automotive Products Division makes a reversible heat engine that charges the car battery using heat from the muffler (600 Kelvin) discharging into the interior of the car at ambient temperature (300 Kelvin). The device charges the battery and at the same time heats the car for winter comfort. As a heat engine its efficiency is exactly the Carnot efficiency.

- a. Calculate the Carnot efficiency for MUD.

η for MUD: _____

The marketing department, aware that the automotive market is very price-sensitive, has asked for a cheaper version with, if necessary, reduced efficiency. Mark, knowing that a heat engine cannot have greater efficiency than the Carnot efficiency but could have less, wants the staff to design a reversible heat engine for this purpose, with 25% efficiency, and wants to call it Carnot's Reduced Unlimited Device (CRUD). It has not actually been made yet, but he is already looking for additional applications for CRUD. Ben Bitdiddle, a fellow engineer, has suggested combining it with MUD to make an air conditioner so the car can be used in the summer. He plans to run CRUD as a reverse heat engine on the energy generated by MUD. That is, MUD will be used normally, to extract heat from the muffler and put it in the interior of the car, and CRUD will run in reverse, used to extract heat from the interior of the car and put it back on the muffler. His power analysis was done assuming that 1 watt of power was generated by MUD, and all that power was then used by CRUD.

- b. You are asked to work on the design of this product, which we will call Product X (some people wanted to call it CRUDdy MUD, but the upper management disapproved). Complete the power analysis table below, considering that CRUD is run backward (i.e., extracting heat from the interior of the car), making, for example, heat into the car interior for MUD positive, and heat 'into' the interior for CRUD negative. Give your answers in Watts.

	MUD	CRUD	Total
Heat into car interior			
Heat extracted from Muffler			
Work input	-1	1	0

Table F-2: Product X heat extraction data, in Watts

- c. Something is bothering you about Product X. Explain. If something is wrong, be as precise as you can in locating the error.

Logarithm and Entropy Table

This page is provided so that you may rip it off the exam to use as a separate reference table. In Table F-3, the entropy $S = p \log_2(1/p) + (1 - p) \log_2(1/(1 - p))$.

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
$\log_2(1/p)$	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 F-3: Table of logarithms in base 2 and entropy in bits