Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
Department of Mechanical Engineering

6.050J/2.110J **Information and Entropy** **Spring 2006**

Issued: March 21, 2006 **Problem Set 7** **Due: March 25, 2006**

Problem 1: Eating in the Dorms at HSU

All the students attending Humongous State University (HSU) live in one of two dormitories, Hill (H) and Vale (V) – 1/3 are in Hill and 2/3 are in Vale. Hill is an all–female dormitory, while Vale has exactly half male (M) and half female (F) students. Note that H/V form a partition and M/F form another partition.

The HSU housing office has surveyed the students about satisfaction with the dormitory eating arrangements, and you have been asked to study the results. All students responded. Unfortunately, the survey writer forgot to ask for dormitory, so there is no direct way to tell if a sample is from a Hill student or a Vale student! However, there was at least a box for gender.

Having aced 6.050/2.110 as a freshman back at MIT, you decide to model the survey as a nondeterministic process, where the input is the dormitory (H/V) and the output is gender (M/F). Then guessing the dormitory is similar to what is done in communication channels, where the input is inferred from knowing the output.

a. What is the uncertainty (measured in bits to two decimal places) of dormitory if you do not yet know the gender of the student whose returned questionnaire you have selected at random?

b. What is the uncertainty in dormitory if you are told that the student is male?

c. What is the uncertainty if you discover that the student is female?

d. What is so unusual about learning that a student is female compared to not knowing the student's gender at all?

e. What is the uncertainty, on average, of dormitory once you learn the gender of the student?

The housing office needs to correlate survey results with dormitory, but doesn't want to conduct another survey. To help them, you decide to design a probabilistic inference machine to guess H/V given M/F. Your design is quite simple: if the probability of a student living in Hill given that the student is male $P(H|M)$ were 3/4 (clearly it's not), then your machine would guess H with probability 3/4 every time you give it M. Your inference machine can be modeled as a process.

f. Design the inference machine. In other words, give $P(H|M)$ (the probability of H given M), $P(H|F)$, etc. Present your answer in the form of a process probability diagram, like the ones discussed in lecture and in Chapter 7 of the notes.

Suddenly you realize that you can think of the combination of the survey and your inference machine as a binary communications channel: a student taking the survey "sends" his or her dormitory information through this channel by specifying only gender. You wonder how good a communications channel this might be.

g. Evaluate the performance of this channel by computing the noise, loss, and mutual information of data carried across the combined system. As usual, your answer should be in bits rounded to two decimal places.

h. Verify one of the four inequalities for M in terms of M_1 and M_2 at the end of Chapter 7 of the notes (Equation 7.38, 7.39, 7.40, or 7.41, your choice).

Problem 2: NOR Gate

The process model covered this week can be used for both deterministic systems, whose output is determined by the input, and by nondeterministic systems. Let's use it to describe the action of a NOR gate. A properly working NOR gate has the model in Figure 7–1.

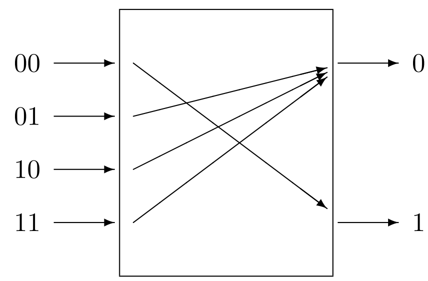

Figure 7–1: NOR gate model

Assume each of the four possible inputs is equally likely, for example if the input had been obtained by two independent coin tosses.

a. Calculate the two output probabilities $p(B_0)$ and $p(B_1)$, the input information I, the output information J, the noise N, loss L, and mutual information M.

You have purchased some NOR gates from Mitch's Microelectronics, which has low prices but no refunds. The order arrives with a note saying that, due to a "slight manufacturing error," your NOR gates do not behave exactly in the normal way. Specifically:

- the outputs for inputs (11) and (00) are correct
- the output for input (10) will be wrong 30% of the time
- the output for input (01) will be wrong 10% of the time

Since you cannot return them and you don't have time to buy new gates before your project is due, you decide to make the best of things, and start by analyzing these defective NOR gates. Continue to assume that all four inputs are equally probable.

b. Draw a probability diagram which models the defective NOR gate as a process. Include the transition probabilities in your diagram.

c. What are the input information I and the output information J (in the correct units)?

d. What are the noise N, the loss L, and the mutual information M? Is this process noisy, lossy, both, or neither?

Turning in Your Solutions

If you used MATLAB for this problem set, you may have some M-files and a diary. Name the M-files with names like `ps7p1.m`, `ps7p2.m`, and name the diary `ps7diary`. You may turn in this problem set by e-mailing your written solutions, M-files, and diary to 6.050-submit@mit.edu. Do this either by attaching them to the e-mail as text files, or by pasting their content directly into the body of the e-mail (if you do the latter, please indicate clearly where each file begins and ends). If you have figures or diagrams you may include them as graphics files (GIF, JPG or PDF preferred) attached to your email. Alternatively, you may turn in your solutions on paper in room 38-344. The deadline for submission is the same no matter which option you choose.

Your solutions are due 5:00 PM on Friday, March 25, 2006. Later that day, solutions will be posted on the course website.