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: April 4, 2006 **Problem Set 8** **Due: April 7, 2006**

Problem 1: Star War ComScan

Han Solo is running a dangerous smuggling mission to Kessel in his spaceship *Millennium Falcon*. However, unknown to Han, the bounty hunter *Bobba Fett* is following him and is determined to get Han once and for all. All of a sudden, a *Bothan spy* tips Han off about Bobba Fett. Han's past experience shows that he cannot perfectly trust these spies and luckily, the Millenium Falcon has a *Bounty-Hunter-Detector ComScan*, which alerts Han of any hostile ships that come in range.

Han knows that usually there is a probability 0.05 that the comscan alarms just because it has mistaken some flying objects around Falcon with the enemies (false positive output.) On the other hand, the sensitivity of comscan to the presence of the Bounty-Hunter spacecrafts is unfortunately not perfect, that is, there is a probability 0.3 that it would fail to alarm in the case the enemy is approaching to Falcon (false negative output.) At this point, Han wonders whether he should ask *Wedge Antilles*, the war hero, for immediate help, calls him only if the comscan alarms or simply ignore the threat.

Believe it or not, Han spent a year at an MIT space chapter as an exchange student during his undergraduate time at the Imperial Academy and there, he took 6.050J/2.110J. Thus, he has enough mathematical tools to deal with this dilemma in a scientific way. Firstly, he denotes the event that Bobba is coming by H_1 and that he is not coming by H_0, reflecting the fact that he considers those two as his hypotheses about the real condition. Also, he denotes the output of comscan by B and the two events of alarming or not by B_1 and B_0, respectively. Correspondingly, comscan would be characterized by the conditional probabilities of the form $P(B|H)$, i.e., probabilities $P(B_1|H_0) = 0.05$, $P(B_0|H_0) = 0.95$, $P(B_0|H_1) = 0.3$ and $P(B_1|H_1) = 0.7$. You can see how Han uses the model he learned in units 7 and 8 of 6.050J/2.110J to come up with this transition matrix.

Now, Han has to take an appropraite strategy by selecting a certain criterion that should be met while making his decision whether to call Wedge for help or not. Han thinks he can decide on that in such a way to minimize the probaility of error, namely, the probability of calling for help while Bobba is not coming or not calling while he is actually coming. Let us call the two events "calling for help" or "ignoring the warning", C and I, respectively. Accordingly, Han has to choose C or I, based on his knowledge of the occurance of B_0 or B_1.

a. Express probability of error P_e in terms of the joint probabilities of Han's decision (C or I) and the hypothesis H (H_0 or H_1.)

The entire decision makeing problem now reduces to the appropriate assignment of C and I to each of two possible outputs B of comscan. Let us now consider the case when some B (either B_0 or B_1,) has occured. In this situation, Han has to decide which of C or I to choose in a way that he would minimize P_e. Clearly, if he accomplishes that for both cases B_0 and B_1, he is done.

b. If Han assignes I to the event B_0 what is the probability of error in terms of conditional probabilities of H given B_0? What if he assignes C to that event?

c. Based on the result of the previous part, state the strategy Han should take to achieve his goal in decision making.

d. Write down the explicit decision making strategy for both cases B_0 and B_1 based on the conditional probabilities of H given B. Try to intuitively explain the result you have found (**Hint:** Try to include the term likely in your explanation!)

Suppose that based on his past exprience Han thinks the Bothan spy is 80% trustful, that is, in 20% times he lies. Han thinks that fact can be used to find values for prior probabilities on his hypotheses, $P(H_0)$ and $P(H_1)$.

e. Using Bayes' rule, express the decision rule in terms of the probabilities B conditioned on H, and the prior probabilities $P(H_0)$ and $P(H_1)$.

f. According to the values Han knows for all those probabilities, precisely specify how he should make his decision and explain the result.

g. If we denote $P(H_1)$ by p, explain how Han's decision would change according to the value of p.

h. **Extra Credit:** Suppose that the spy did not inform Han about Bobba and the only knowledge Han has is the output of comscan, that is to say, he cannot assign any prior probabilities for $P(H)$. Suggest another criterion for decision making that Han can use to make his decision here. Note that the probability of error criterion that we used in this problem can be only defined in terms of randomness of H as you can clearly see in your answer to part (a).

Problem 2: Lots of Pages to Read...

You are taking 21R.12X "Way Too Much Reading." The catalog description says that students read an average of 1600 pages per week for the class. The first week the professor has selected three different reading assignments:

- Fyodor Dostoyevsky, *The Brothers Karamazov*, 900 pages,
- Leo Tolstoy, *War and Peace*, 1500 pages,
- Donald E. Knuth, *The Art of Computer Programming*, Volumes 1-3, 2100 pages total.

Each student is randomly assigned one of these books (the three volumes of Knuth together count as one book). The professor has hired a former 6.050J/2.110J student to write a program to make random assignments consistent with the statement in the catalog description, that on average a student will read 1600 pages per week. There is some confusion, however, about whether this average is to be the goal each week (averaging the assignments over all students who survived the previous week) or the goal for each student (averaging over all weeks in the semester).

Initially, the professor says the average is to be met each week. You have not yet received your assignment for the first week, and of course wonder what it will be. Since you don't know with certainty, you express what knowledge you do have in terms of probabilities – the probabilities p(D) (Dostoyevsky), p(T) (Tolstoy), and p(K) (Knuth). Since you will be assigned exactly one book, you know that the events D, T, and K form a partition, so

$$p(D) + p(T) + p(K) = 1 \qquad (8\text{–}1)$$

Furthermore, you know that the average number of pages assigned is 1600:

$$900p(D) + 1500p(T) + 2100p(K) = 1600 \qquad (8\text{–}2)$$

Noticing that there are many possible probability distributions consistent with these two constraints, you decide to find the distribution that uses only the information you have. In other words, you use the principle of maximum entropy. (This approach is consistent with your experience that usually professors reveal the minimum amount of information about their courses.)

Note: To do this problem, you may want to use MATLAB. If you do and then hand in your assignment on paper, please write the MATLAB statement you used. Remember to show all work. Note that `log2(x)` is not implemented correctly on all MATLAB versions, so use `log(x)/log(2)` instead. The commands `solve` and `diff` can be used to find where a function's derivative is zero; to see the help files for these operations, type `help sym/solve` or `help sym/diff`.

a. Recall that all probabilities must lie in the interval between 0 and 1.

   i. What range of $p(K)$, the probability that you are assigned the Knuth book, is consistent with these constraints?

   ii. Plot the entropy of the probability distribution over this range as a function of $p(K)$. (If you turn in the problem set on paper, be sure to include this graph!)

   iii. Find the maximum entropy, and give the value of all three probabilities at this maximum.

The professor guesses that most of the students in his class would prefer to read Knuth, so he asks for another probability distribution that is consistent with the constraints but that maximizes p(K).

b. Before calculating the entropy in this case, say whether it could be less than, equal to, or greater than the entropy calculated in part (a-iii) above.

c. Now find the other two probabilities, and the entropy.

The student programmer points out that the constraint on the average could be fulfilled by averaging over the entire semester, so there is no need to adhere to it the very first week. He suggests a wider range of options in assigning the probabilities.

d. What are the minimum and maximum number of pages of reading (averaged over all students in the class) the professor can assign? Describe in words how these situations are achieved. What are the three probabilities in these casee? What are the corresponding values of entropy?

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 `ps8p1.m`, `ps8p2.m`, and name the diary `ps8diary`. 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, April 7, 2006. Later that day, solutions will be posted on the course website.