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 Solutions

Due: April 7, 2006

Solution to Problem 1: Star War ComScan

Solution to Problem 1, part a.

$$P_e = P(C, H_0) + P(I, H_1).$$

Solution to Problem 1, part b.

$$P_e(I, B_0) = P(H_1|B_0),$$
$$P_e(C, B_0) = P(H_0|B_0).$$

Solution to Problem 1, part c.

He should choose I, if $P_e(I, B_0) < P_e(C, B_0)$ and C, otherwise. We will denote this dicision making process by:

$$P(H_1|B_0) \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} P(H_0|B_0).$$

Solution to Problem 1, part d.

The above expression gives the explicit decision making strategy for B_0. Similarly, for B_1, we have:

$$P_e(I, B_1) = P(H_1|B_1),$$
$$P_e(C, B_1) = P(H_0|B_1),$$
$$P(H_1|B_1) \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} P(H_0|B_1).$$

As we see here, both expressions look the same and we can merge them as follows:

$$P(H_1|B) \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} P(H_0|B) \quad B = B_0, B_1.$$

It has the following, nice interpretation: given any output of the system, you have to find the hypothesis that is more likely under that condition. For instance, if $P(H_1|B) > P(H_0|B_1)$, it is more likely that hypothesis H_1 is true, so he will call for help.

Solution to Problem 1, part e.

Using Bayes' rule, for $B = B_0, B_1$ we have

$$P(H_1|B) \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} P(H_0|B)$$

$$\frac{P(B|H_1)P(H_1)}{P(B)} \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \frac{P(B|H_0)P(H_0)}{P(B)}$$

$$P(B|H_1)P(H_1) \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} P(B|H_0)P(H_0)$$

$$\frac{P(B|H_1)}{P(B|H_0)} \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \frac{P(H_0)}{P(H_1)}.$$

Solution to Problem 1, part f.

- B_0: Han should call for help.

$$\frac{P(B_0|H_1)}{P(B_0|H_0)} \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \frac{P(H_0)}{P(H_1)}$$

$$\frac{0.3}{0.95} \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \frac{0.2}{0.8}$$

$$0.32 \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} .2$$

- B_1: again, Han should call for help.

$$\frac{P(B_1|H_1)}{P(B_1|H_0)} \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \frac{P(H_0)}{P(H_1)}$$

$$\frac{0.7}{0.05} \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \frac{0.2}{0.8}$$

$$\frac{0.7}{0.05} \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \frac{0.2}{0.8}$$

$$14 \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} .2$$

Thus, Han should immediately call Wedge since independent of the output of comscan his P_e is minimized by choosing C.

Solution to Problem 1, part g.

As before, we have

$$\frac{P(B|H_1)}{P(B|H_0)} \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \frac{1-p}{p} = \eta,$$

where η changes between infinity and zero, when p goes from 0 to 1. Therefore,

- For B_0:

$$0.32 \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \eta$$

- For B_1:

$$14 \underset{\text{decide } C}{\overset{\text{decide } I}{\lessgtr}} \eta$$

Accordingly, we have the following cases:

- $\eta < 0.32$ ($p > 0.758$): Always decide C.
- $0.32 < \eta < 14$ ($0.067 < p < 0.758$): Decide C if B_1 and I if B_0.
- $\eta > 14$ ($p < 0.067$): Always decide I.

Solution to Problem 1, part h.

Depending on your final goal, there are several ways, for example you can try to increase $P(C|H_1)$, i.e. probability of calling when you are really in danger, while keeping $P(C|H_0)$, probability of calling when you are not in danger, below a certain chosen threshold. If you work it out, you will see that you will not need any information about the prior $P(H)$ in order to use a scheme like this. Note that increasing $P(C|H_1)$ and decreasing $P(C|H_0)$ are competing goals and you cannot acheive both simultaneously.

Solution to Problem 2: Lots of Pages to Read..

Solution to Problem 2, part a.

i. We will need to rearrange Equations 8–1 and 8–2. First we get the following from 8–2

$$P(D) = \frac{16}{9} - \frac{7}{3}P(K) - \frac{5}{3}P(T) \tag{8–3}$$

and substituting $P(S)$ into 8–1 we have

$$\begin{aligned} 1 &= P(D) + P(K) + P(T) \\ &= \frac{16}{9} - \frac{7}{3}P(K) - \frac{5}{3}P(T) + P(K) + P(T) \\ &= \frac{16}{9} - \frac{4}{3}P(K) - \frac{2}{3}P(T) && (8\text{–}4) \\ P(T) &= \frac{7}{6} - 2P(K) && (8\text{–}5) \end{aligned}$$

Once $P(D)$ and $P(T)$ are determined, $P(K)$ is determined as well:

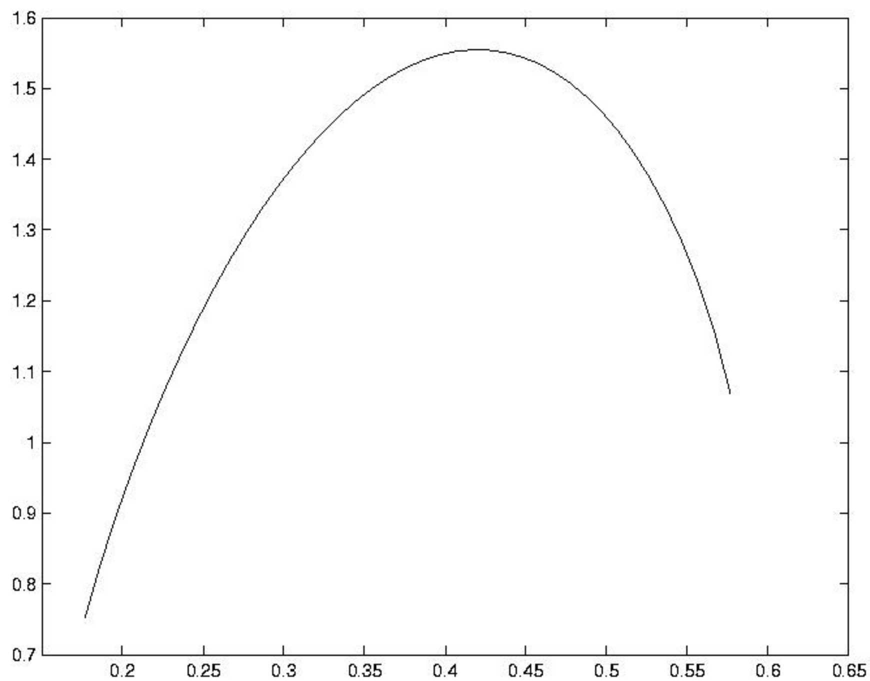
$$\begin{aligned} P(K) &= \frac{1}{6} + P(D) \\ &= \frac{7}{12} - \frac{1}{2}P(T) && (8\text{–}6) \end{aligned}$$

And so, since $P(D)$, $P(T)$, and $P(K)$ must be between 0 and 1, we see that $P(K)$ must be between $\frac{1}{6}$ and $\frac{7}{12} = 0.7619$.

ii. The equation for the entropy is as follows:

$$\begin{aligned} H &= P(K)\log_2\left(\frac{1}{P(K)}\right) + P(D)\log_2\left(\frac{1}{P(D)}\right) + P(T)\log_2\left(\frac{1}{P(T)}\right) \\ &= P(K)\log_2\left(\frac{1}{P(K)}\right) + (\frac{7}{12} - 2P(K))\log_2\left(\frac{1}{\frac{7}{12} - 2P(K)}\right) + \\ &\quad (P(K) - \frac{1}{6})\log_2\left(\frac{1}{P(K) - \frac{1}{6}}\right) \end{aligned} \quad (8\text{–}7)$$

in order to find the maximum we may either take the derivative and equal it to zero or plot in the range of valid values of $P(K)$. With this formula, you obtain the following plot:



iii. The maximum entropy of $H = 1.555$ bits is at $P((K) = 0.42$, which gives values of $P(T) = 0.33$ and $P(D) = 0.25$.

Solution to Problem 2, part b.

The entropy should be less than (a-iii) because, after all, that value was calculated with a procedure known as the Principle of Maximum Entropy. The maximum value of $P(K)$ consistent with the constraints is $7/12$.

Solution to Problem 2, part c.

If $P(E)$ is $7/12$, then $P(S) = 0$ and $P(U) = 5/12$. The entropy at this point is $H = .98$ bits.

Solution to Problem 2, part d.

Since the students have to read one book the first week, the minimum number of pages is 900. The maximum number of pages is 2100. The entropy is zero in both cases. Any distribution that puts full weight into one of the variables will achieve zero entropy.