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 **Due: March 25, 2006**

Problem Set 7 Solutions

Solution to Problem 1: Eating in the Dorms at HSU

Solution to Problem 1, part a.

The uncertainty without the knowledge of the gender in dormitory is simply:

$$U = P(V)\log_2 \frac{1}{P(V)} + P(H)\log_2 \frac{1}{P(H)} = 2/3\log_2(3/2) + 1/3\log_2(3) = 0.92\,\text{bits}$$

Solution to Problem 1, part b.

$U(\text{given } M) = 0$

Solution to Problem 1, part c.

$P(H|F) = 1/2$ and $P(V|F) = 1/2$ so $U(\text{given } F) = 1$ bit.

Solution to Problem 1, part d.

Learning the student is female increases your uncertainty about the dormitory. This seems strange, that learning one fact can increase your uncertainty of another. However, on average this is not so.

Solution to Problem 1, part e.

$$U(\text{given } M)P(M) + U(\text{given} F)P(F) = 1 \times 2/3 = 0.67\,\text{bits}.$$

This is less than the uncertainty before you know the gender, which is 0.92 bits. Thus uncertainty in dormitory is reduced on average by learning the gender.

Solution to Problem 1, part f.

Figure 7–2.

Solution to Problem 1, part g.

The combined transition matrix is the second transition matrix times the first transition matrix:

$$\begin{bmatrix} 0.75 & 0.5 \\ 0.25 & 0.5 \end{bmatrix}$$

Using the formula for noise:

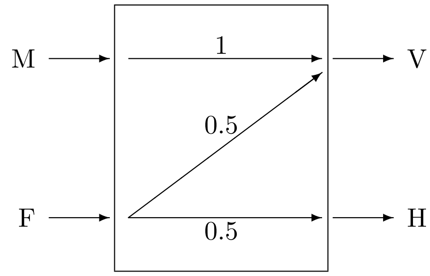

Figure 7–2: Your inference machine

$$
\begin{aligned}
N &= \frac{2}{3} \times 0 + \frac{2}{3}\left(0.5 \log_2 \frac{1}{0.5} + 0.5 \log_2 \frac{1}{0.5}\right) \\
&= 0.67 \text{ bits} \\
U_{in} &= U_{out} = 0.92 \text{ bits} \\
L &= N + U_{in} - U_{out} = 0.67 \text{ bits} \\
M &= U_{in} - L = 0.25 \text{ bits}
\end{aligned}
$$

This is apparently not very effective as a communications channel.

Solution to Problem 1, part h.

Just check whichever inequality you want. There are so many to choose from. In doing so, you may use the fact that the two process boxes have identical information-flow properties. Thus $L_1 = N_1 = L_2 = N_2 = 0.67$ bits and $M_1 = M_2 = 0.25$ bits. Therefore these four inequalities are all similar. For example, $M_1 - L_2 \leq M \leq M1 \leq U_{in,1}$ becomes $(0.25 - 0.67) \leq 0.05 \leq 0.25 \leq 0.92$.

Solution to Problem 2: NOR Gate

Solution to Problem 2, part a.

If each of the inputs is equally likely, then we just need to sum the probabilities of the inputs that converge at each output to find the output probabilities. Thus:

$$p(B_0) = 3 \times 1/2 \times 1/2 = 3/4$$
$$p(B_1) = 1/2 \times 1/2 = 1/4$$

The formula for the input information is given in the lecture notes as: $U = -\sum_i p(A_i) \log P(A_i)$. the input information is then equal to

$$\begin{aligned} I =& p(A_{00})\log_2\left(\frac{1}{p(A_{00})}\right) + p(A_{01})\log_2\left(\frac{1}{p(A_{01})}\right) + \\ &+ p(A_{10})\log_2\left(\frac{1}{p(A_{10})}\right) + p(A_{11})\log_2\left(\frac{1}{p(A_{11})}\right) \\ =& 4 \times (0.25)\log_2\left(\frac{1}{0.25}\right) \\ =& \log_2(4) \\ =& 2 \text{ bits} \end{aligned} \tag{7–1}$$

The output information J is given by the same equation

$$\begin{aligned} J =& p(B_0)\log_2\left(\frac{1}{p(B_0)}\right) + p(B_1)\log_2\left(\frac{1}{p(B_1)}\right) \\ =& (0.25)\log_2\left(\frac{1}{0.25)}\right) + (0.75)\log_2\left(\frac{1}{0.75}\right) \\ =& 0.811 \text{ bit} \end{aligned} \tag{7–2}$$

The noise N is calculated with $N = -\sum_i P(A_i)\sum_j c_{ji}\log c_{ji}$. Since c_{ij} is either 1 or 0, that means that each $c_{ij}\log_2\left(\frac{1}{c_{ij}}\right)$ term is zero by virtue of the logarithm (in the case of $c_{ij} = 1$) or by $c_{ij} = 0$. Thus the noise, N is zero.
The loss $L = N - J + I$, with $N = 0$:

$$\begin{aligned} L =& I - J \\ =& 2 - 0.81 \\ =& 1.19 \text{ bit} \end{aligned} \tag{7–3}$$

The mutual information M is equal to J when there is no noise.

Solution to Problem 2, part b.

Figure 7–3 is a box diagram for the gate with defects.

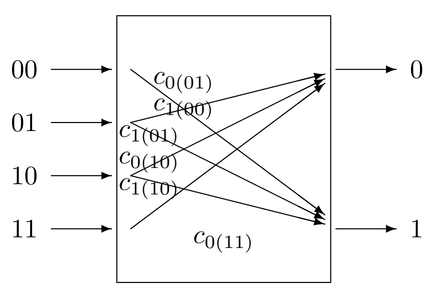

Figure 7–3: The gate with defects

and the following is the transition matrix:

$$\begin{bmatrix} c_{0(00)} & c_{0(01)} & c_{0(10)} & c_{0(11)} \\ c_{1(00)} & c_{1(01)} & c_{1(10)} & c_{1(11)} \end{bmatrix} = \begin{bmatrix} 0 & 0.9 & .7 & 1 \\ 1 & 0.1 & .3 & 0 \end{bmatrix} \tag{7–4}$$

Solution to Problem 2, part c.

The input information I is the same as before, i.e., 2 bits. To calculate the output information first we calculate B_0 and B_1 by noting that each input is equally likely

$$\begin{aligned}
P(B_0) =& c_{0(00)}P(A_{00}P + c_{0(01)}P(A_{01}) + c_{0(10)}P(A_{10}) + c_{0(11)}P(A_{11}) \\
=& \frac{1}{4} \times (0 + .9 + .7 + 1) \\
=& \frac{2.6}{4} \\
=& 0.65 && \text{(7–5)}
\end{aligned}$$

$$P(B_1) = 0.35 \qquad \text{(7–6)}$$

Then the output information J is computed as before:

$$\begin{aligned}
J =& \sum_j p(B_j) \log_2 \left(\frac{1}{p(B_j)} \right) \\
=& -0.35 \log_2 0.35 - 0.65 \log_2 0.65 \\
=& .93 && \text{(7–7)}
\end{aligned}$$

Solution to Problem 2, part d.

The process is both noisy and lossy, since there both fan-outs from the inputs and fan-ins to the outputs. The noise N is:

$$\begin{aligned}
N =& \sum_i p(A_i) \sum_j c_{ji} \log_2 \left(\frac{1}{c_{ji}} \right) \\
=& p(A_{00}) \sum_j c_{j(00)} \log_2 \left(\frac{1}{c_{j(00)}} \right) + p(A_{01}) \sum_j c_{j(01)} \log_2 \left(\frac{1}{c_{j(01)}} \right) + \\
& + p(A_{10}) \sum_j c_{j(10)} \log_2 \left(\frac{1}{c_{j(10)}} \right) + p(A_{11}) \sum_j c_{j(11)} \log_2 \left(\frac{1}{c_{j(11)}} \right) \\
=& p(A_{00}) \left(c_{0(00)} \log_2 \left(\frac{1}{c_{0(00)}} \right) + c_{1(00)} \log_2 \left(\frac{1}{c_{1(00)}} \right) \right) + p(A_{01}) \left(c_{0(01)} \log_2 \left(\frac{1}{c_{0(01)}} \right) + c_{1(01)} \log_2 \left(\frac{1}{c_{1(01)}} \right) \right) + \\
& + p(A_{10}) \left(c_{1(10)} \log_2 \left(\frac{1}{c_{1(10)}} \right) \right) + p(A_{11}) \left(c_{0(11)} \log_2 \left(\frac{1}{c_{0(11)}} \right) + c_{1(11)} \log_2 \left(\frac{1}{c_{1(11)}} \right) \right) \\
=& \frac{1}{4} \times 1 + \frac{1}{4} \times .881 + \frac{1}{4} \times .469 \\
=& 0.338 \text{ bits} && \text{(7–8)}
\end{aligned}$$

The loss L is defined as:

$$\begin{aligned}
L =& I - (J - N) \\
=& 2 - (0.338) \\
=& 1.403 \text{ bits} && \text{(7–9)}
\end{aligned}$$

The mutual information M is defined as:

$$
\begin{aligned}
M &= I - L \\
&= J - N \\
&= 2 - 1.403 \\
&= 0.597 \text{ bits}
\end{aligned} \tag{7–10}
$$