
Problem Set 6 Solutions

Solution to **Problem 1: What Kind of Gate is This?**

Solution to Problem 1, 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:

$$\begin{aligned}p(B_0) &= 0.75 \\p(B_1) &= 0.25\end{aligned}$$

The input information I_{in} is given by equation 5.10 in the notes, and is equal to

$$\begin{aligned}I_{in} &= p(A_{00}) \log_2 \left(\frac{1}{p(A_{00})} \right) + p(A_{01}) \log_2 \left(\frac{1}{p(A_{01})} \right) + \\&\quad p(A_{10}) \log_2 \left(\frac{1}{p(A_{10})} \right) + p(A_{11}) \log_2 \left(\frac{1}{p(A_{11})} \right) \\&= 4 \times (0.25) \log_2 \left(\frac{1}{0.25} \right) \\&= \log_2(4) \\&= 2 \text{ bits}\end{aligned}\tag{6-1}$$

The output information I_{out} is given by the same equation

$$\begin{aligned}I_{out} &= p(B_0) \log_2 \left(\frac{1}{p(B_0)} \right) + p(B_1) \log_2 \left(\frac{1}{p(B_1)} \right) \\&= (0.75) \log_2 \left(\frac{1}{0.75} \right) + (0.25) \log_2 \left(\frac{1}{0.25} \right) \\&= 0.75 \times 0.415 + 0.25 \times 2 \\&= 0.81 \text{ bits}\end{aligned}\tag{6-2}$$

The noise N is calculated via equation 7.14. Since c_{ij} is either 1 or 0, that means that each $\frac{c_{ij} \log_2(1/c_{ij})}{c_{ij}}$ 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 is calculated via equation 7.25, with $N = 0$:

$$\begin{aligned}L &= I_{in} - I_{out} \\&= 2 - 0.8112 \\&= 1.19\end{aligned}\tag{6-3}$$

The mutual information M is equal to I_{out} when there is no noise.

Solution to Problem 1, part b.

Figure 6-2 is a box diagram for the defective "Greater" gate.

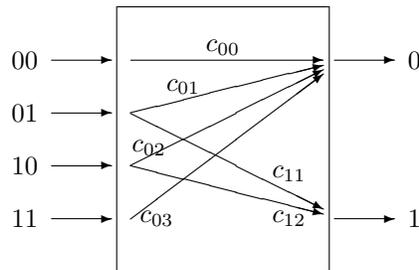


Figure 6-2: A defective "Greater" gate

and the following is the transition matrix:

$$\begin{bmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \end{bmatrix} = \begin{bmatrix} 1 & 0.1 & 0.8 & 1 \\ 0 & 0.9 & 0.2 & 0 \end{bmatrix} \tag{6-4}$$

Solution to Problem 1, part c.

Given that the result is 1 ...

- i. The probability that the 1 was produced by the input (0 1) is

$$\begin{aligned} \frac{c_{11}}{c_{11} + c_{12}} &= \frac{0.9}{0.9 + 0.2} \\ &= 0.82 \end{aligned} \tag{6-5}$$

- ii. The probability that it was produced by the input (1 0) is

$$\begin{aligned} \frac{c_{12}}{c_{11} + c_{12}} &= \frac{0.2}{0.9 + 0.2} \\ &= 0.18 \end{aligned} \tag{6-6}$$

- iii. The probability that it was produced by the input (1 1) is zero, since c_{13} is 0.

Solution to Problem 1, part d.

The input information I_{in} 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

$$\begin{aligned} B_0 &= c_{00}A_0 + c_{01}A_1 + c_{02}A_2 + c_{03}A_3 \\ &= 1 \times 0.25 + 0.1 \times 0.25 + 0.8 \times 0.25 + 1 \times 0.25 \\ &= 0.25 + 0.025 + 0.2 + 0.25 \\ &= 0.725 \end{aligned} \tag{6-7}$$

$$\begin{aligned} B_1 &= c_{10}A_0 + c_{11}A_1 + c_{12}A_2 + c_{13}A_3 \\ &= 0 \times 0.25 + 0.9 \times 0.25 + 0.2 \times 0.25 + 0 \times 0.25 \\ &= 0.225 + 0.05 \\ &= 0.275 \end{aligned} \tag{6-8}$$

Then the output information I_{out} is given by equation 7.12:

$$\begin{aligned}
 I_{out} &= \sum_j p(B_j) \log_2 \left(\frac{1}{p(B_j)} \right) \\
 &= 0.725 \log_2 \left(\frac{1}{0.725} \right) + 0.275 \log_2 \left(\frac{1}{0.275} \right) \\
 &= 0.725 \log_2 \left(\frac{1}{0.725} \right) + 0.275 \log_2 \left(\frac{1}{0.275} \right) \\
 &= 0.85
 \end{aligned} \tag{6-9}$$

Solution to Problem 1, part e.

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 defined in equation 7.22:

$$\begin{aligned}
 N &= \sum_i p(A_i) \sum_j c_{ji} \log_2 \left(\frac{1}{c_{ji}} \right) \\
 &= p(A_0) \sum_j c_{j0} \log_2 \left(\frac{1}{c_{j0}} \right) + p(A_1) \sum_j c_{j1} \log_2 \left(\frac{1}{c_{j1}} \right) + \\
 &\quad p(A_2) \sum_j c_{j2} \log_2 \left(\frac{1}{c_{j2}} \right) + p(A_3) \sum_j c_{j3} \log_2 \left(\frac{1}{c_{j3}} \right) \\
 &= p(A_0) \left(c_{00} \log_2 \left(\frac{1}{c_{00}} \right) + c_{10} \log_2 \left(\frac{1}{c_{10}} \right) \right) + p(A_1) \left(c_{01} \log_2 \left(\frac{1}{c_{01}} \right) + c_{11} \log_2 \left(\frac{1}{c_{11}} \right) \right) + \\
 &\quad p(A_2) \left(c_{02} \log_2 \left(\frac{1}{c_{02}} \right) + c_{12} \log_2 \left(\frac{1}{c_{12}} \right) \right) + p(A_3) \left(c_{03} \log_2 \left(\frac{1}{c_{03}} \right) + c_{13} \log_2 \left(\frac{1}{c_{13}} \right) \right) \\
 &= 0.25 \left(1 \log_2 \left(\frac{1}{1} \right) + 0 \log_2 \left(\frac{1}{0} \right) \right) + 0.25 \left(0.1 \log_2 \left(\frac{1}{0.1} \right) + 0.9 \log_2 \left(\frac{1}{0.9} \right) \right) + \\
 &\quad 0.25 \left(0.8 \log_2 \left(\frac{1}{0.8} \right) + 0.2 \log_2 \left(\frac{1}{0.2} \right) \right) + 0.25 \left(1 \log_2 \left(\frac{1}{1} \right) + 0 \log_2 \left(\frac{1}{0} \right) \right) \\
 &= 0 + 0.25 (0.1 \log_2 (10) + 0.9 \log_2 (1.11)) + 0.25 (0.8 \log_2 (1.25) + 0.2 \log_2 (5)) + 0 \\
 &= 0.30
 \end{aligned} \tag{6-10}$$

The loss L is defined as:

$$\begin{aligned}
 L &= N + I_{in} - I_{out} \\
 &= 0.30 + 2 - 0.81 \\
 &= 1.49
 \end{aligned} \tag{6-11}$$

The mutual information M is defined as:

$$\begin{aligned}
 M &= I_{in} - L \\
 &= I_{out} - N \\
 &= 2 - 1.49 \\
 &= 0.81 - 0.3 \\
 &= 0.51
 \end{aligned} \tag{6-12}$$

Solution to Problem 2: Cantabridgian Computer Competency

Solution to Problem 2, part a.

The probability $p(M)$ that any given student who took the exam is from MIT is

$$\begin{aligned} p(M) &= \frac{10,000}{10,000 + 20,000} \\ &= 0.33 \end{aligned} \tag{6-13}$$

Solution to Problem 2, part b.

The uncertainty U is the amount of information you would receive, on average, if you knew the result. This is the average information per event:

$$\begin{aligned} U &= p(M) \log_2 \left(\frac{1}{p(M)} \right) + p(H) \log_2 \left(\frac{1}{p(H)} \right) \\ &= 0.33 \times \log_2 \left(\frac{1}{0.33} \right) + 0.66 \times \log_2 \left(\frac{1}{0.66} \right) \\ &= 0.33 \times \log_2(3) + 0.66 \times \log_2(1.5) \\ &= 0.91 \text{ bits} \end{aligned} \tag{6-14}$$

Solution to Problem 2, part c.

First let us calculate the probabilities $p(H | C)$ and $p(M | C)$. We know that half of the Harvard students (10,000 students) and all of the MIT students (10,000 students) are competent.

$$\begin{aligned} p(H | C) &= \frac{10,000}{10,000 + 10,000} \\ &= 0.5 \\ &= p(M | C) \end{aligned} \tag{6-15}$$

The uncertainty in school if you are told that the student was deemed competent is the average information gained when we learn the school:

$$\begin{aligned} I_{avgC} &= p(M | C) \log_2 \left(\frac{1}{p(M | C)} \right) + p(H | C) \log_2 \left(\frac{1}{p(H | C)} \right) \\ &= 0.5 \times \log_2 \left(\frac{1}{0.5} \right) + 0.5 \times \log_2 \left(\frac{1}{0.5} \right) \\ &= 0.5 \times \log_2(2) + 0.5 \times \log_2(2) \\ &= 1 \text{ bit} \end{aligned} \tag{6-16}$$

Solution to Problem 2, part d.

Here again, let us first let us calculate the probabilities $p(H | I)$ and $p(M | I)$. We know that half of the Harvard students (10,000 students) and none of the MIT students (0 students) are incompetent.

$$\begin{aligned}
 p(H | I) &= \frac{10,000}{0 + 10,000} \\
 &= 1
 \end{aligned}
 \tag{6-17}$$

$$p(M | I) = 0
 \tag{6-18}$$

The uncertainty in school if you are told that the student was deemed incompetent is the average information gained when we learn the school:

$$\begin{aligned}
 I_{avgI} &= p(M | I) \log_2 \left(\frac{1}{p(M | I)} \right) + p(H | I) \log_2 \left(\frac{1}{p(H | I)} \right) \\
 &= 0 + 1 \times \log_2 \left(\frac{1}{1} \right) \\
 &= 0 \text{ bits}
 \end{aligned}
 \tag{6-19}$$

Not a surprising answer. If we know the student is incompetent, we know they are from Harvard, not MIT.

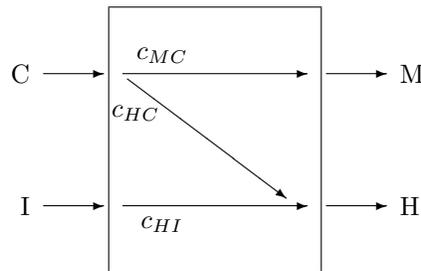
Solution to Problem 2, part e.

The average uncertainty on learning the competency of the student is:

$$\begin{aligned}
 I_{out} &= p(I)I_{avgI} + p(C)I_{avgC} \\
 &= 0.33 \times 0 + 1 \times 0.66 \\
 &= 0.66 \text{ bits}
 \end{aligned}
 \tag{6-20}$$

Solution to Problem 2, part f.

A process box diagram for the inference process would look as shown in Figure 6-3.



(b)

Figure 6-3: Cambridge computer competency exam inference machine

and the following is the transition matrix:

$$\begin{bmatrix} c_{HI} & c_{MI} \\ c_{HC} & c_{MC} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \end{bmatrix}
 \tag{6-21}$$

Solution to Problem 2, part g.

The transition matrix of the whole system is:

$$\begin{aligned}
\begin{bmatrix} c_{HH} & c_{MH} \\ c_{HM} & c_{MM} \end{bmatrix} &= \begin{bmatrix} c_{IH} & c_{IM} \\ c_{CH} & c_{CM} \end{bmatrix} * \begin{bmatrix} c_{HI} & c_{HC} \\ c_{MI} & c_{MC} \end{bmatrix} \\
&= \begin{bmatrix} 0.5 & 0 \\ 0.5 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0.5 \\ 0 & 0.5 \end{bmatrix} \\
&= \begin{bmatrix} 0.5 & 0.25 \\ 0.5 & 0.75 \end{bmatrix}
\end{aligned} \tag{6-22}$$

The noise is computed as follows:

$$\begin{aligned}
N &= \sum_i p(A_i) \sum_j c_{ji} \log_2 \left(\frac{1}{c_{ji}} \right) \\
&= p(H) \sum_j c_{jH} \log_2 \left(\frac{1}{c_{j0}} \right) + p(M) \sum_j c_{j1} \log_2 \left(\frac{1}{c_{jM}} \right) + \\
&= p(H) \left(c_{IH} \log_2 \left(\frac{1}{c_{IH}} \right) + c_{CH} \log_2 \left(\frac{1}{c_{CH}} \right) \right) + p(M) \left(c_{IM} \log_2 \left(\frac{1}{c_{IM}} \right) + c_{CM} \log_2 \left(\frac{1}{c_{CM}} \right) \right) + \\
&= 0.66 \left(0.5 \log_2 \left(\frac{1}{0.5} \right) + 0.5 \log_2 \left(\frac{1}{0.5} \right) \right) + 0.33 \left(0 \log_2 \left(\frac{1}{0} \right) + 1 \log_2 \left(\frac{1}{1} \right) \right) + \\
&= 0.66 (0.5 \log_2 (2) + 0.5 \log_2 (2)) + 0 \\
&= 0.66 \text{ bits}
\end{aligned} \tag{6-23}$$

The loss is as follows:

$$\begin{aligned}
L &= I_{in} - I_{out} + N \\
&= 0.91 - 0.91 + 0.66 \\
&= 0.66 \text{ bits}
\end{aligned} \tag{6-24}$$

The mutual information is as follows:

$$\begin{aligned}
M &= I_{out} - N \\
&= 0.91 - 0.66 \\
&= 0.25 \text{ bits}
\end{aligned} \tag{6-25}$$

MIT OpenCourseWare
<http://ocw.mit.edu>

6.050J / 2.110J Information and Entropy
Spring 2008

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.