ECE 6605 INFO. THEORY 9/3/03

HOMEWORK # 1 SOLUTIONS DR. MCLAUGHLIN 1/4

1. a) $H(X)$, $H(Y)$? First find the pmf's for the r.v.'s X & Y

$P(x)$

$x=a$	$x=b$
3/8	5/8

$$H(X) = -\frac{3}{8}\log_2\left(\frac{3}{8}\right) - \frac{5}{8}\log_2\left(\frac{5}{8}\right) = 0.9544$$

$P(y)$

$y=1$	$y=2$	$y=3$
1/2	3/8	1/8

$$H(Y) = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{3}{8}\log_2\left(\frac{3}{8}\right) - \frac{1}{8}\log_2\left(\frac{1}{8}\right) = 1.4056$$

For convenience we will skip to part c) first

c)

	$Y=1$	$Y=2$	$Y=3$
$X=a$	1/4	1/8	0
$X=b$	1/4	1/4	1/8

from the table we find that

$$H(X,Y) = -3\left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right) - 2\left(\frac{1}{8}\right)\log_2\left(\frac{1}{4}\right) = 2.25$$

b) we know from the chain rule

$$H(X,Y) = H(X) + H(Y|X) \Rightarrow H(Y|X) = H(X,Y) - H(X)$$

$$H(Y|X) = 2.25 - 0.9544 = 1.2956$$

Similarly

$$H(X,Y) = H(Y) + H(X|Y) \Rightarrow H(X|Y) = H(X,Y) - H(Y)$$

$$H(X|Y) = 2.25 - 1.4056 = 0.8444$$

We can also do this the long way

$$\Rightarrow p(x=a|y=1) = \frac{p(x=a, y=1)}{p(y=1)} = \frac{1/4}{1/2} = 1/2 \quad \text{similarly}$$

$$p(x=a|y=2) = \frac{1/8}{3/8} = 1/3$$

$$p(x=a|y=3) = \emptyset$$

$$p(x=b|y=1) = \frac{1/4}{1/2} = 1/2$$

$$p(x=b|y=2) = \frac{1/4}{3/8} = 2/3$$

$$p(x=b|y=3) = \frac{1/8}{1/8} = 1$$

$$\Rightarrow H(X|Y) = -\left(\frac{1}{4}\right)\log_2\left(\frac{1}{2}\right) - \left(\frac{1}{8}\right)\log_2\left(\frac{1}{3}\right) - \left(\frac{1}{4}\right)\log_2\left(\frac{1}{2}\right) - \left(\frac{1}{4}\right)\log_2\left(\frac{2}{3}\right) - \left(\frac{1}{8}\right)\cancel{\log_2(1)}^{\emptyset}$$

$$= 0.8444$$

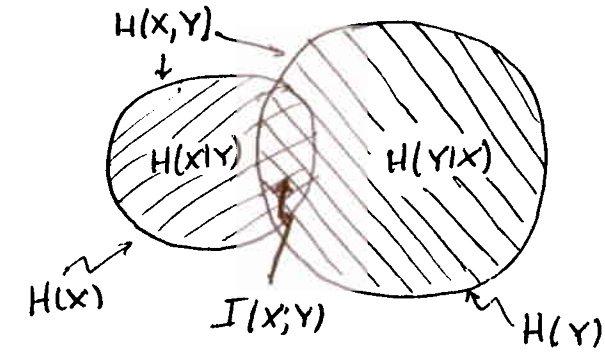
we can do the same for $H(Y|X)$

d) $\Rightarrow H(Y) - H(Y|X) = 1.4056 - 1.2956 = 0.11$

e) $I(X;Y) = H(Y) - H(Y|X) = 0.11$ from d). Also notice

$= H(X) - H(X|Y) = 0.9544 - 0.8444 = 0.11$

f)



Notice $I(X;Y)$ is the intersection of $H(X)$ & $H(Y)$ in our case the smallest area (0.11) and $H(X,Y) = H(X) + H(Y|X)$ is the whole shaded area (2.25) in this example.

PROBLEM #2

a) Let's call the variables T and B (top and bottom) then $I(T;B) = H(T) - H(T/B)$ now we know the value of $H(T) = -2(\frac{1}{2})\log_2(\frac{1}{2}) = \log_2 2 = 1$ since the coin is fair. Now $H(T/B)$ is the uncertainty in T given B. However since the coin has only two sides once B is given then we immediately know T so $H(T/B) = 0$. Therefore

$I(T;B) = H(T) - H(T/B) = 1 - 0 = 1$ bit.

Another way to look at it is to obtain a table for $P(T=t, B=h)$. We have two values t=tails and h=heads.

	$B=t$	$B=h$
$T=t$	0	½
$T=h$	½	0

$H(T,B) = -2(\frac{1}{2})\log_2(\frac{1}{2}) = \log_2 2 = 1$

but $H(T) = \log_2 2 = 1 = H(B)$

$\Rightarrow H(T,B) = H(B) + H(T|B) \Rightarrow H(T|B) = H(T,B) - H(B)$

$\Rightarrow H(T|B) = 1 - 1 = 0$

$\Rightarrow I(T;B) = H(T) - H(T|B) = 1 - 0 = \underline{\underline{1}}$

PROBLEM #2

b) Call the sides T & F then we want to find $I(T;F) = H(T) - H(T|F)$

First assume the die is fair, that is $P(T=i) = 1/6$ for $i = 1, 2, \ldots, 6$.

Also the die is a "standard" die so that opposite faces add up to 7, meaning if we know the top face we also know the bottom. This assumption does not change the way we solve the problem only the final value of the solution.

To calculate $H(T)$ notice we have 6 equiprobable outcomes $\Rightarrow H(T) = -6\left(\frac{1}{6}\right)\log_2\left(\frac{1}{6}\right) = \log_2 6$.

Now for $H(T|F)$ notice that if we know the front face F then T can take one of four values with equal probability (since the die is "standard", if it weren't) it could take one of five values), then we get

$H(T|F) = -4\left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right) = \log_2 4$

$\Rightarrow I(T;F) = \log_2 6 - \log_2 4 = \log_2 \frac{3}{2} = \log_2 3 - 1$

PROBLEM #3

a) No. The chain rule says $I(X;Y_1,Y_2) = I(X;Y_1) + I(X;Y_2|Y_1)$ we can have $I(X;Y_2|Y_1) > 0$. The problem reminds us of the probability fact that pair-wise independence of three random variables is not sufficient to guarantee that all three are mutually independent.

b) Let Y_1 and Y_2 be independent fair coin flips and let $X = Y_1$ XOR Y_2. X is pairwise independent of Y_1 and $Y_2 \Rightarrow I(X,Y_1) = 0$ and $I(X,Y_2) = 0$.

However X is not independent of (Y_1,Y_2) since X is uniquely determined by the pair, therefore

$I(X;Y_1,Y_2) > 0$

PROBLEM #4

using the chain rule we write

$$H(X, g(X)) = H(X) + H(g(X)|X) = H(g(X)) + H(X|g(X))$$

however $H(g(X)|X) = 0$ since once X is given there is no uncertainty about $g(X)$ therefore

$$H(X) = H(g(X)) + H(X|g(X))$$ and since

$H(X|g(X)) \geq 0$ we get

$$H(X) \geq H(g(X))$$

equality holds iff $H(X|g(X)) = 0$ which requires X to be a function of $g(X)$ or what is equivalent requires $g(\cdot)$ to be one-to-one.

PROBLEM #5

First analyze the single noisy detection

$Z = X + N$ where $var[N] = \sigma^2$ & $E[N] = 0$

$E[Z] = E[X] + E[N] = E[X]$

$$\Rightarrow Var[Z] = E[Z^2] - (E[Z])^2$$

$$= E[(X+N)(X+N)] - (E[X])^2$$

$$= E[X^2] + 2\cancel{E[XN]}^{0} + E[N^2] - (E[X])^2$$

$$= \underbrace{E[X^2] - (E[X])^2} + \underbrace{E[N^2] - (\underset{0}{E[N]})^2}$$

$$= var[X] + \sigma^2$$

now for $Y = (Y_1 + Y_2)/2 = X + N_1/2 + N_2/2$

$$\Rightarrow E[Z] = E[X] + \cancel{E[N_1]}^{0}/2 + \cancel{E[N_2]}^{0}/2 = E[X]$$

$$\Rightarrow Var[Y] = E[Y^2] - (E[Y])^2$$

$$= E[(X + \tfrac{N_1}{2} + \tfrac{N_2}{2})(X + \tfrac{N_1}{2} + \tfrac{N_2}{2})] - (E[X])^2$$

$$= E[X^2] + \cancel{E[XN_1]}^{0} + \cancel{E[XN_2]}^{0} + \cancel{E[\tfrac{N_1N_2}{2}]}^{0} + E[\tfrac{N_1^2}{4}] + E[\tfrac{N_2^2}{4}] - (E[X])^2$$

$$= \underbrace{E[X^2] - (E[X])^2} + \underbrace{\tfrac{1}{4}E[N_1^2] - (\underset{0}{E[N_1]})^2} + \tfrac{1}{4}(\underbrace{E[N_2^2] - (\underset{0}{E[N_2]})^2}_{\sigma^2})$$

$$= var[X] + \frac{\sigma^2}{4} + \frac{\sigma^2}{4} = var[X] + \frac{\sigma^2}{2}$$

so the noise variance is halved from the previous case