Information Theory—Solution of exam paper, 5 December 2002

Important! Answers are not complete without sufficient reasoning.

Problem 1 *Give the chain rule for the joint entropy of the random variables $X_1, X_2, \ldots, X_n$. What does it state if the random variables are independent?*

Solution:
Chain rule: $H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i \mid X_1, \ldots, X_{i-1})$
If the random variables are independent, then $H(X_i \mid X_1, \ldots, X_{i-1}) = H(X_i)$, thus $H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i)$.

Problem 2 *Give the definition of the entropy rate of an information source. State the source coding theorem, lower and upper bounds for the expected codeword length per symbol.*

Solution:
The entropy rate of an information source $\mathbf{X} = X_1, X_2, \ldots$ is defined by $H(\mathbf{X}) = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n)$ whenever the limit exists.

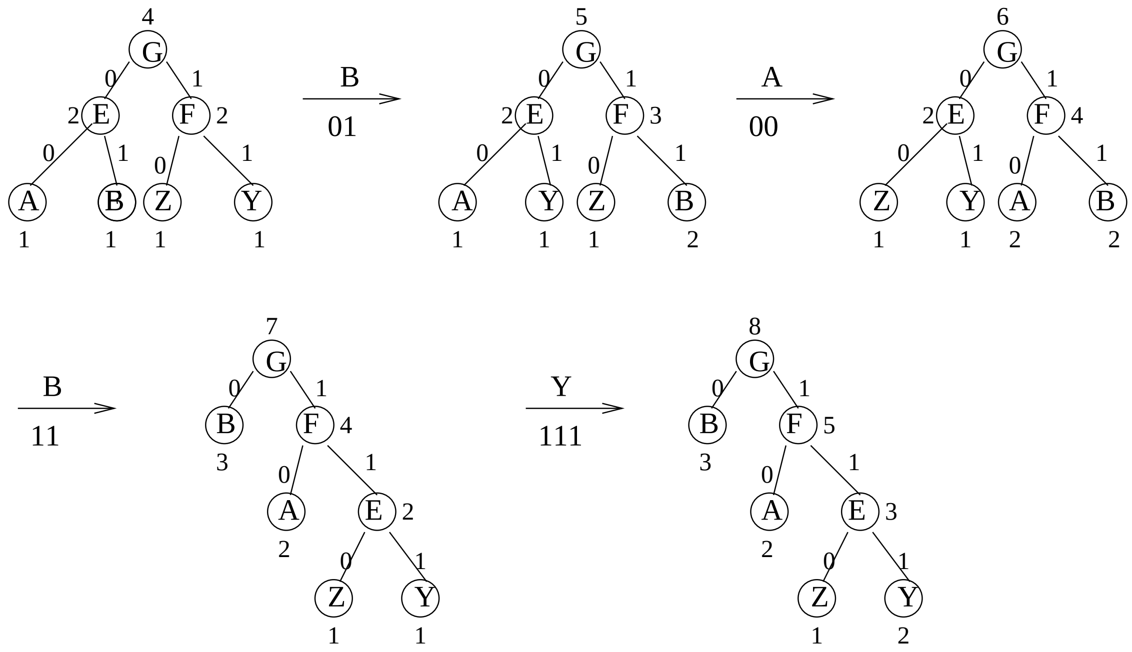
Source coding theorem: If the stationary source $\mathbf{X}$ is encoded with a uniquely decodable code $C : \mathcal{X}^k \mapsto \mathcal{D}^*$ then for the expected codeword length per symbol

$$L = \frac{1}{k}\mathbf{E}\{l(X_1, X_2, \ldots, X_k)\} \geq \frac{H(\mathbf{X})}{\log D}.$$

On the other hand, if k is large enough then there exists a code with L arbitrarily close to the lower bound $\frac{H(\mathbf{X})}{\log D}$.

Problem 3 *Assume that the source alphabet is $\mathcal{X} = \{A, B, Z, Y\}$. Find a possible adaptive binary Huffman code for the word BABY. Describe also the Huffman tree with the weights after each step.*

Solution:
Above the arrows are the input symbols, under the arrows are the corresponding codewords that are sent.



Whenever a symbol is read, it moves to the end of its block (the node of the symbol and the last node in the block is exchanged, also in the tree), then its weight is increased by one (it moves to the next block), and then we consider its parent etc. Initial blocks: $(A, B, Z, Y)(E, F)(G)$

B: its codeword according to the (actual) initial tree is 01, the last node in its block is Y, thus they are exchanged in the tree. The weight of B becomes 2. Its new parent is F, it is the last in its block, thus it can stay, and its weight is increased by one. Its parent is G, its weight is increased as well. The new blocks: $(A, Y, Z)(B, E)(F)(G)$

A: its codeword according to the actual tree is 00, the last node in its block is Z, thus they are exchanged in the tree. The weight of A becomes 2. Its new parent is F, it is the last in its block, thus it can stay, and its weight is increased by one. Its parent is G, its weight is increased as well. The new blocks: $(Z, Y)(A, B, E)(F)(G)$

B: its codeword according to the actual tree is 11, the last node in its block is E, thus they are exchanged in the tree. The weight of B becomes 3. Its new parent is G, its weight is increased as well. The new blocks: $(Z, Y)(A, E)(B)(F)(G)$

Y: its codeword according to the actual tree is 111, it is the last in its block, thus it can stay. The weight of Y becomes 2. Its parent is E, it is the last in its block, thus it can stay, and its weight is increased by one. Its parent is F, it is the last in its block, thus it can stay as well, and its weight is increased by one. Its parent is G, its weight is increased as well. The new blocks: $(Z)(Y, A)(E, B)(F)(G)$.

Problem 4 *Let X be uniformly distributed on $[0,1]$. We quantize X with a two level quantizer Q_a that has quantization levels a and $2a$ ($0 \le a \le 1/2$). Find a for which the quantizer is optimal, i.e. the distortion is minimal. Give then the entropy of the quantizer. For what a is this entropy maximal?*

Solution:

According to the nearest neighbor quantization rule, in case of a two level quantizer with levels a and $2a$, X is quantized to a if it is smaller than $\frac{a+2a}{2} = \frac{3}{2}a$, and it is quantized to $2a$ if it is greater than $\frac{a+2a}{2} = \frac{3}{2}a$. Thus the two quantization regions are $[0, \frac{3}{2}a]$ and $(\frac{3}{2}a, 1]$. Using the definition, the distortion of the quantizer is

$$D(Q_a) = \mathbf{E}\{(X - Q(X))^2\} = \int_0^1 (x - Q_a(x))^2 f(x)dx = \int_0^{\frac{3}{2}a} (x-a)^2 dx + \int_{\frac{3}{2}a}^1 (x-2a)^2 dx = -\frac{9}{4}a^3 + 4a^2 - 2a + \frac{1}{3},$$

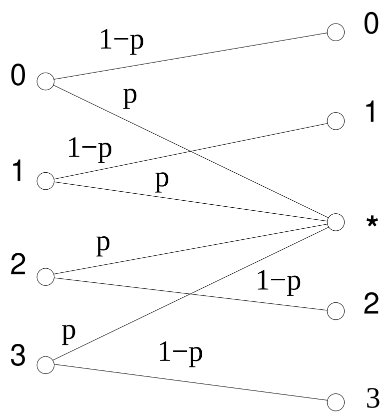
since X is uniform on $[0,1]$, thus the density function $f(x) = 1$ for all x. To minimize $D(Q_a)$, take its derivative:

$$D'(Q_a) = -\frac{27}{4}a^2 + 8a - 2.$$

It is 0 for $a_1 = \frac{16+2\sqrt{10}}{27}$ and $a_2 = \frac{16-2\sqrt{10}}{27}$. a_1 is greater than $\frac{1}{2}$, but a_2 is a good solution. It is about 0.358349.

Entropy: $Q(X)$ can take two values a and $2a$, so it is binary. $\mathbf{P}\{Q(X) = a\} = \mathbf{P}\{X \in [0, \frac{3}{2}a]\} = \int_0^{\frac{3}{2}a} f(x)dx = \int_0^{\frac{3}{2}a} 1dx = \frac{3}{2}a$. Thus $H(Q_a) = h(\frac{3}{2}a)$. It is maximal when $\frac{3}{2}a = \frac{1}{2}$, i.e. when $a = \frac{1}{3}$. And then $H(Q_a) = 1$.

Problem 5 *Find the capacity of the channel given by the following figure.*



Solution:
Let denote the input distribution by (p_0, p_1, p_2, p_3). Capacity:

$$C = \max_{p_0,p_1,p_2,p_3} I(U,V) = \max_{p_0,p_1,p_2,p_3} (H(V) - H(V \mid U)).$$

$H(V \mid U) = h(p)$, since for any input symbol, the output can take two values with probability p and $1-p$. To find $H(V)$, we need the distribution of V. $\mathbf{P}\{V=0\} = p_0(1-p)$, $\mathbf{P}\{V=1\} = p_1(1-p)$, $\mathbf{P}\{V=2\} = p_2(1-p)$, $\mathbf{P}\{V=3\} = p_3(1-p)$, $\mathbf{P}\{V=*\} = p$.

$$\begin{aligned}
H(V) &= \\
&= -\sum_i \mathbf{P}\{V=i\}\log\mathbf{P}\{V=i\} \\
&= -\left(p_0(1-p)\log p_0(1-p) + p_1(1-p)\log p_1(1-p) + p_2(1-p)\log p_2(1-p) + p_3(1-p)\log p_3(1-p) + p\log p\right) \\
&= -\left((1-p)(p_0+p_1+p_2+p_3)\log(1-p) + p\log p + p_0\log p_0 + p_1\log p_1 + p_2\log p_2 + p_3\log p_3\right) \\
&= h(p) + (1-p)H(p_0,p_1,p_2,p_3) = h(p) + (1-p)H(U).
\end{aligned}$$

Thus

$$C = \max_{p_0,p_1,p_2,p_3} (h(p) + (1-p)H(U) - h(p)) = \max_{p_0,p_1,p_2,p_3} (1-p)H(U).$$

$H(U)$ is maximal if the input is uniformly distributed, that is if $p_0 = p_1 = p_2 = p_3 = \frac{1}{4}$ and then $H(U) = \log 4 = 2$. Therefore $C = 2(1-p)$.