Information Theory and Coding 2004,
Paper 8 Question 10 (JGD)

(*a*) (*i*) Entropy of the source, $H(X)$, is 1 bit. [2 marks]

(*ii*) Output probabilities are $p(y = 0) = (0.5)(1 - \epsilon) + (0.5)\epsilon = 0.5$ and $p(y = 1) = (0.5)(1-\epsilon)+(0.5)\epsilon = 0.5$. So the entropy of this distribution is $H(Y) = 1$ bit, just as for the entropy $H(X)$ of the input distribution. [2 marks]

(*iii*) Joint probability distribution $p(X, Y)$ is

$$\begin{pmatrix} 0.5(1-\epsilon) & 0.5\epsilon \\ 0.5\epsilon & 0.5(1-\epsilon) \end{pmatrix}$$

and the entropy of this joint distribution is $H(X,Y) = -\sum_{x,y} p(x,y)\log_2 p(x,y)$

$= -(1-\epsilon)\log(0.5(1-\epsilon))-\epsilon\log(0.5\epsilon) = (1-\epsilon)-(1-\epsilon)\log(1-\epsilon)+\epsilon-\epsilon\log(\epsilon)$

$= 1 - \epsilon\log(\epsilon) - (1-\epsilon)\log(1-\epsilon)$ [2 marks]

(*iv*) The mutual information is $I(X;Y) = H(X) + H(Y) - H(X,Y)$, which we can evaluate from the quantities above as: $1 + \epsilon\log(\epsilon) + (1-\epsilon)\log(1-\epsilon)$. [2 marks]

(*v*) In the two cases of $\epsilon = 0$ and $\epsilon = 1$ (perfect transmission, and perfectly erroneous transmission), the mutual information reaches its maximum of 1 bit and this is also then the channel capacity. [2 marks]

(*vi*) If $\epsilon = 0.5$, the channel capacity is minimal and equal to 0. [2 marks]

(*b*) The N binary code word lengths $n_1 \le n_2 \le n_3 \le \cdots \le n_N$ must satisfy the Kraft-McMillan Inequality in order to form a uniquely decodable prefix code:

$$\sum_{i=1}^{N} \frac{1}{2^{n_i}} \le 1$$

[3 marks]

(*c*) (*i*) The Karhunen-Loève transform. [1 mark]

(*ii*) The Karhunen-Loève transform decorrelates random vectors. Let the values of the random vector $\mathbf{v}$ represent the individual images in one file. All vector elements being linear combinations of five values means that for each file there exists an orthonormal matrix M such that each image vector $\mathbf{v}$ can be represented as $\mathbf{v} = M\mathbf{t}$, where $\mathbf{t}$ is a new random vector whose covariance matrix is diagonal and in which all but the first five elements are zero. The Karhunen-Loève transform provides this matrix

M by calculating the spectral decomposition of the covariance matrix of $\mathbf{v}$. The significant part of the transform result $M^\top \mathbf{v} = \mathbf{t}$ are only five numbers, which can be stored compactly for each image, together with the five relevant rows of M per file. [4 marks]

[This question relates to the section on correlation coding, as discussed in the course section on coding audiovisual signals.]