Engineering Tripos Part IIA THIRD YEAR

3F1: Signals and Systems

INFORMATION THEORY

Examples Paper

1. Let the joint probability mass function of two binary random variables X and Y be given in the following table:

x, y	$P_{XY}(x, y)$
0,0	0.2
0,1	0.3
1,0	0.1
1,1	0.4

Compute the following quantities. Answers may be expressed numerically or in terms of the binary entropy function

$$h(x) = -x \log x - (1-x)\log(1-x).$$

(a) $H(X)$

(b) $H(Y)$

(c) $H(X|Y)$

(d) $H(Y|X)$

(e) $H(XY)$

(f) $I(X;Y)$

Hint: Use the chain rule in two different ways in the last two questions and use a calculator to verify that you got the same result both times.

2. Let an N-ary random variable X be distributed as follows:

$$\begin{cases} P_X(1) & = 1-p \\ P_X(k) & = \frac{p}{N-1} \text{ for } k = 2, 3, \ldots, N. \end{cases}$$

Express the entropy of X in terms of the binary entropy function $h(.)$.

3. A discrete memoryless source has an alphabet of eight letters, $x_i, i = 1, 2, \cdots, 8$ with probabilities 0.25, 0.20, 0.15, 0.12, 0.10, 0.08, 0.05 and 0.05.

(a) Use the Huffman encoding to determine a binary code for the source output.

(b) Determine the average codeword length L.

(c) Determine the entropy of the source and hence its efficiency.

4. Show that for statistically independent events

$$H(X_1, X_2, \cdots, X_n) = \sum_{i=1}^{n} H(X_i)$$

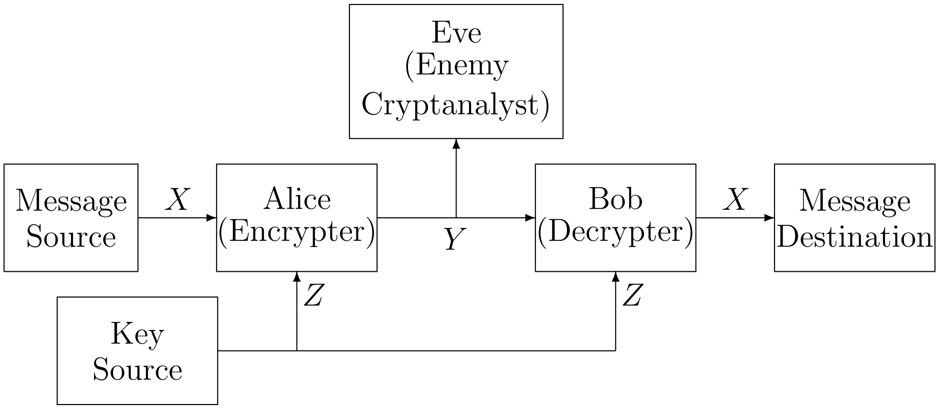
5. A five-level non-uniform quantizer for a zero-mean signal results in the 5 levels $-b, -a, 0, a, b$ with corresponding probabilities of occurrence $p_{-b} = p_b = 0.05$, $p_{-a} = p_a = 0.1$ and $p_0 = 0.7$.

(a) Design a Huffman code that encodes one signal sample at a time and determine the average bit rate per sample.

(b) Design a Huffman code that encodes two output samples at a time and determine the average bit rate per sample.

(c) What are the efficiencies of these two codes?

6. While we cover in 3F1 and 4F5 the application of Shannon's theory to data compression and transmission, Shannon also applied the concepts of entropy and mutual information to the study of secrecy systems. The figure below shows a cryptographic scenario where Alice wants to transmit a secret plaintext message X to Bob and they share a secret key Z, while the enemy Eve has access to the public message Y.



(a) Write out two conditions using conditional entropies involving X,Y and Z to enforce the deterministic encryptability and decryptability of the messages.

Hint: the entropy of a function given its argument is zero, e.g., for any random variable X, $H(f(X)|X) = 0$.

(b) Shannon made the notion of an "unbreakable cryptosystem" precise by saying that a cryptosystem provides perfect secrecy if the enemy's observation is statistically independent of the plaintext, i.e., $I(X;Y) = 0$. Show that this implies Shannon's much cited bound on key size

$$H(Z) \geq H(X),$$

i.e., perfect secrecy can only be attained if the entropy of the key (and hence its compressed length) is at least as large as the entropy of the secret plaintext.

(c) Vernam's cipher assumes a binary secret plaintext message X with any probability distribution $P_X(0) = p = 1 - P_X(1)$ and a binary secret key Z that's uniform $P_Z(0) = P_Z(1) = 1/2$ and independent of X. The encrypter simply adds the secret key to the plaintext modulo 2, and the decrypter by adding the same key to the ciphertext can recover the plaintext. Show that Vernam's cipher achieves perfect secrecy, i.e., $I(X;Y) = 0$.

7. What is the entropy of the following continuous probability density functions?

(a) $P(x) = \begin{cases} 0 & x < -2 \\ 0.25 & -2 < x < 2 \\ 0 & x > 2 \end{cases}$

(b) $P(x) = \frac{\lambda}{2} e^{-\lambda|x|}$

(c) $P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$

8. Continuous variables X and Y are independent and normally distributed with standard deviation $\sigma = 1$.

$$P(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad\qquad P(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

A variable Z is defined by $z = x + y$. What is the mutual information of X and Z?

9. A symmetric binary communications channel operates with signalling levels of ±2 volts at the detector in the receiver, and the rms noise level at the detector is 0.5 volts. The binary symbol rate is 100 kbit/s.

(a) If the output of this channel is quantised by a two-level quantiser with threshold 0, determine the probability of error on the resulting channel and hence, based on mutual information, calculate the theoretical capacity of this channel for error-free communication.

(b) If the binary signalling were replaced by symbols drawn from a continuous process with a Gaussian (normal) pdf with zero mean and the same mean power at the detector, determine the theoretical capacity of this new channel, assuming the symbol rate remains at 100 ksym/s and the noise level is unchanged.

(c) *(Computer Exercise)* In MATLAB/Octave, plot the two capacities above in function of the Signal to Noise ratio on a scale from -5dB to 15dB. A third channel of interest that is closely related to the two channels studied is the channel with binary signal levels (as in (a)) but continuous output (as in (b)). The capacity of this channel can only be computed numerically. You may use the following approximation:

```
s2 = 4*10^(snr/10);
eta = linspace(-20,20,1e5);
x=exp(-(eta-s2/2).^2./(2*s2))/sqrt(2*pi*s2).*log(1+exp(-eta))/log(2);
C = 1-trapz(x)*(eta(2)-eta(1));
```

Compare the capacity of the three channels and discuss the practical implications of your findings.

Numerical Answers

3. b) 2.83 bits; c) 2.798 bits, 98.9%

5. a) 1.6 bit / sample; b) 1.465 bit / sample; c) 91.05%, 99.44%

7. a) $\log_2(4) = 2$; b) $\log_2(2e/\lambda)$; c) $\log_2(\sigma\sqrt{2\pi e})$

8. 0.5 bit

9. a) $p_e = 3.17 \,.\, 10^{-5}$, 99.9481 kbit/s.; b) 204.37 kbit/s.

Jossy Sayir, January 21, 2014.