COMP2610 - Information Theory: Assignment 2: Solutions

Australian National University

2018

I Arithmetic and Stream Coding [40 points]

1. a [2 points]

Symbol	Cumulative Probability	Binary Interval
a	$\frac{1}{4}$	$[0.0, 0.01)$
b	$\frac{1}{2}$	$[0.01, 0.10)$
c	$\frac{3}{4}$	$[0.10, 0.11)$
d	1	$[0.11, 1)$

b [4 points]

From (a), midpoints of intervals are $0.001, 0.011, 0.101, 0.111$ giving the code

$$
\begin{aligned}
a &\to 001 \\
b &\to 011 \\
c &\to 101 \\
d &\to 111
\end{aligned}
$$

abab is then encoded as 001011001011.

c [3 points]

A Huffman code for this ensemble would encode each letter with 2 bits, encoding abab with a binary string of length 8 rather than of length 12.

2. a [2 points]

$P(x = b) = \frac{1}{4}$

b [2 points]

$P(x = a|baaa) = \frac{3+1}{4+4} = \frac{1}{2}$

c [2 points]

$P(x = b|ac) = \frac{1+1}{4+2} = \frac{1}{6}$

3. [10 points]

Following the algorithm presented in the lectures

Step 1

Counts $= (1, 1, 1, 1)$, $x = a$ giving

$I = [0, \frac{1}{4})$, $b' = 00$ and $b = 00$

Step 2

Counts $= (2, 1, 1, 1)$, $x = c$ giving

$I = [\frac{3}{20}, \frac{4}{20}) = [001\overline{0011}, 001\overline{1001})$, $b' = 1$ and $b = 001$

Step 3

Counts $= (2, 1, 2, 1)$, $x = b$ giving

$I = [\frac{20}{120}, \frac{21}{120}) = [001\overline{01}, 001\overline{0110})$, $b' = 0$ and $b = 0010$

Final Step

$I = [\frac{41}{240}, \frac{21}{120}) = [0010\overline{1011}, 001\overline{0110})$, $b' = 1011$ giving a final code of 00101011. As a sanity check, under this model we have $\lceil \log(P(acb)) \rceil = 7$ meaning a code of length 8 is expected.

4. a [3 points]

   There are 32 symbols hence a Huffman code will use 5 bits for each symbol, giving a code of length 5000bits.

   b [6 points]

   Arithmetic codes are very near optimal hence we have

$$L(aaa\ldots a) \approx -\log(P(aaa\ldots a))$$

   Under the first model we have

$$-\log(P(aaa\ldots a)) = \sum_{i=0}^{998} \log(\frac{32+i}{1+i}) \approx 196$$

   And under the second model we have

$$-\log(P(aaa\ldots a)) = \sum_{i=0}^{998} \log(\frac{1032+i}{1000+i}) \approx 31$$

   Here I have been lazy and used Mathematica to do the sums, however one could also have used an integral approximation

$$\sum_{i=0}^{998} \log(\frac{32+i}{1+i}) \approx \int_0^{998} \log(\frac{32+x}{1+x})dx$$

   c [6 points]

   LZ78 will encode the string $aaa\ldots a$ as $(0, a), (1, a), (2, a)\ldots(n, a)$ and so on. We require that

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} = 1000$$

   Giving $n \approx 45$. Note at each step we are (approximately) required to encode the index, which is a number between O and 45, as well as the letter a. Assuming we use a uniform code for symbols it will take 5 bits to encode each a and roughly $\lceil \log(45) \rceil \approx 6$ bits to encode the index. Hence we use roughly $(5 + 6) * 45 \approx 496$ bits. This is an upper bound for the actual number of bits used.

II Noisy Channel Coding [40 points]

1. a [5 points]

   We have a markov chain

$$X \to Y \xrightarrow{f} Y'$$

   Hence by the data processing theorem, for any distribution on the input X we have $I(X;Y) \geq I(X;Y')$, meaning $C(Q) \geq C'(Q)$ and the saleperson is wrong.

   b [5 points]

   We have a markov chain

$$X' \xrightarrow{f} X \to Y$$

   And again by the data processing theorem, for any distribution on the input X' we have $I(X';Y) \leq I(X;Y)$. To see this, remember that you can use bayes rule on the above chain to get a markov chain

$$Y \to X \to X'$$

   That gives the same joint distribution over X', X, Y. Maximizing over the input distributions yields $C(Q) \geq C'(Q)$, meaing the salesperson is wrong again.

2. a [4 points]

   Q transmits a with no errors and b with error probability of 0.5. The opposite for Q'.

   b(i) [2 points]

   $P(y = a) = \frac{1}{2}(1 + p)$ and $P(y = b) = \frac{1}{2}(1 - p)$

   b(ii) [2 points]

   $H(Y) = H_2(\frac{1}{2}(1 + p))$

   b (ii) [2 points]

   $H(Y|x = a) = 0$ and $H(Y|x = b) = 1$, meaning $H(Y|X) = 1 - p$ meaning

$$I(X;Y) = H(Y) - H(Y|X) = H_2(\frac{1}{2}(1 + p)) - (1 - p)$$

   c [6 points]

   From previous question we have $I(X;Y)$ as a function of p. As $I(X,Y)$ is a concave function of p To maximize $I(X : Y)$ we solve

$$\partial_p I(X;Y) = 0$$

   yielding $p = 0.6$. This gives $C(Q) \approx 0.32$.

   d [4 points]

   Symmetry. If we interchange a and b then channel Q becomes Q'. Q' is the same as swapping a with b, using Q and then swapping back. The same holds for Q. As such both channels have the same capacity.

e (i) [3 points]

$$Q^* = \frac{1}{2}Q + \frac{1}{2}Q' = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}$$

e (ii) [3 points]

$Q*$ is a symmetric channel meaning $C(Q) = 1 - H_2(3/4) \approx 0.189$

f [4 points]

Use Q when $x = a$ and Q' otherwise.

III Combining Noisy Channels [20 points]

1. [4 points]

$Q = \begin{pmatrix} 1 & \frac{1}{4} \\ 0 & \frac{3}{4} \end{pmatrix}$ and $Q' = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$

We have $X \oplus X' = \{a, b, c, d\}$ and $Y \oplus Y' = \{1, 2, 3, 4\}$. This gives

$$Q \oplus Q' = \begin{pmatrix} 1 & \frac{1}{4} & 0 & 0 \\ 0 & \frac{3}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

$X \otimes X' = \{(a, c), (a, d), (b, c), (b, d)\}$ and $Y \otimes Y' = \{(1, 3), (1, 4), (2, 3), (2, 4)\}$. Note that

$$P(Y = y, Y' = y'|X = x, X' = x') = P(Y = y|X = x)P(Y' = y'|X' = x')$$

This gives

$$Q \otimes Q' = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{1}{12} & \frac{2}{12} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{12} & \frac{1}{12} \\ 0 & 0 & \frac{3}{12} & \frac{6}{12} \\ 0 & 0 & \frac{6}{12} & \frac{3}{12} \end{pmatrix}$$

2. [8 points]

For a channel of the form $Q \otimes Q'$ one has

$$P(X, X, Y, Y') = P(X, X')P(Y|X)P(Y'|X')$$

meaning $H(X, X, Y, Y') = H(X, X') + H(Y|X) + H(Y'|X')$. We have

$$\begin{aligned} I(X, X'; Y, Y') &= H(X, X') + H(Y, Y') - (X, X, Y, Y') \\ &= H(Y, Y') - H(Y|X) - H(Y'|X') \\ &\le H(Y) - H(Y|X) + H(Y') - H(Y'|X') \\ &= I(X; Y) + I(X'; Y') \end{aligned}$$

With equality when $P(X,X') = P(X)P(X')$, as in this case Y and y' are independent. Maximizing over $P(X)$ and $P(X')$ we obtain

$$C(Q \otimes Q') = C(Q) + C(Q')$$

3. [8 points]

Any distribution P on $X \oplus X'$ can be written as a combination of two distributions P_X on X and P'_X on X'

$$P(X \oplus X') = \alpha P(X) + (1-\alpha)P(X')$$

where $\alpha = P(z \in X)$, $z \in X \oplus X'$. It can be verified that

$$H(X \oplus X') = \alpha H(X) + (1-\alpha)H(X') + H_2(\alpha)$$

Meaning for channels of the form $Q \oplus Q'$.

$$\begin{aligned} H(X \oplus X') &= \alpha H(X) + (1-\alpha)H(X') + H_2(\alpha) \\ H(Y \oplus Y') &= \alpha H(Y) + (1-\alpha)H(Y') + H_2(\alpha) \\ H(X \oplus X', Y \oplus Y') &= \alpha H(X,Y) + (1-\alpha)H(X',Y') + H_2(\alpha) \end{aligned}$$

This gives

$$I(X \oplus X'; Y \oplus Y') = \alpha I(X;Y) + (1-\alpha)I(X';Y') + H_2(\alpha)$$

Maximizing over $P(X \oplus X')$ is the same as maximizing over $P(X)$, $P(X')$ and α. Maximizing over the first two gives

$$\phi(\alpha) = \alpha C(Q) + (1-\alpha)C(Q') + H_2(\alpha)$$

Solving $\partial_\alpha \phi(\alpha) = 0$ gives $\alpha = \frac{2^{C(Q)}}{2^{C(Q)}+2^{C(Q')}}$ which upon substitution gives

$$C(Q \otimes Q') = \log(2^{C(Q)} + 2^{C(Q')})$$