COMP-423B: Quiz 1 Solutions

Prepared by Matthew Wahab (T.A.)

General Comments

- For any questions/comments regarding the correction of the quiz: if I have made a mistake adding up your points, or you feel that your arguments were not understood, please contact me (Matt) at wahab@cim.mcgill.ca

1. **(4 points)**

   Marking:
   1 mark for each part of the question.

   Consider an alphabet with five symbols and probabilities:

   $$p(A_1) = .4, \quad p(A_2) = .35, \quad p(A_3) = .2, \quad p(A_4) = .05$$

   (a) Give an expression of the entropy defined by the $p(A_i)$'s.

   Solution:

   $$\begin{aligned} H &= \sum_{i=1}^{N=5} p(A_i) \log \frac{1}{p(A_i)} \\ &= 0.4 \log \frac{1}{0.4} + 0.35 \log \frac{1}{0.35} + 0.2 \log \frac{1}{0.2} + 0.05 \log \frac{1}{0.05} \end{aligned}$$

   (b) Construct a Huffman code such that, at each merge step, the child labelled 0 has probability less than or equal to the child labelled 1. Be sure to write the code words explicitly. A binary tree representation is not enough.

   Solution:

   $$C(A_1) = 0, \quad C(A_2) = 11, \quad C(A_3) = 101, \quad C(A_4) = 100$$

(c) Give an expression of the average code length of this Huffman code.

Solution:

$$\begin{aligned}\lambda_{avg} &= \sum_{i=1}^{N=5} \lambda_i p(A_i) \\ &= 1(0.4) + 2(0.35) + 3(0.2) + 3(0.05)\end{aligned}$$

(d) Decode the bit string 100101111010

Solution:

Traversing the tree, we decode:
$100 = A_4, 101 = A_3, 11 = A_2, 101 = A_3, 0 = A_1$.

Hence, $C^{-1}(100101111010) = A_4A_3A_2A_3A_1$

2. **(3 points)**

Marking:
1 mark for each part of the question. This question was fairly basic, so no partial marks were given, except for very small mistakes.

(a) Give an example of a prefix code on $\{A_1, A_2, \ldots, A_6\}$ with codeword lengths 1,3,3,3,4,4 respectively.

Solution:

$C(A_1) = 0, \quad C(A_2) = 100, \quad C(A_3) = 101, \quad C(A_4) = 110, \quad C(A_5) = 1110, \quad C(A_6) = 1111$

(this is just one example, the important thing is the codeword length)

(b) For what probabilities is this an optimal prefix code? For these probabilities, given an expression for the entropy.

Solution:

For an optimal prefix code we want $p(A_i) \approx 2^{-\lambda_i}$. This gives us the following probabilities:

$p(A_1) = 2^{-1},\ p(A_2) = 2^{-3},\ p(A_3) = 2^{-3},\ p(A_4) = 2^{-3},\ p(A_5) = 2^{-4},\ p(A_6) = 2^{-4}$

And an entropy of:

$$
\begin{aligned}
H &= \sum_{i=1}^{N=6} p(A_i) \log \frac{1}{p(A_i)} \\
&= 2^{-1} \log \frac{1}{2^{-1}} + 2^{-3} \log \frac{1}{2^{-3}} + 2^{-3} \log \frac{1}{2^{-3}} + 2^{-3} \log \frac{1}{2^{-3}} + 2^{-4} \log \frac{1}{2^{-4}} + 2^{-4} \log \frac{1}{2^{-4}}
\end{aligned}
$$

(c) Give an example of probabilities $p(A_i)$ for which your code is an optimal prefix code, but the average codeword length is greater than the entropy, *and* the probabilities *of each symbol* are different from (b).

Solution:

To do this we perturb the probabilities of the symbols by a small value (ϵ), maintaining the property $\sum_{i=1}^{N} p(i) = 1$ and that the ordering of the symbols (according to probability) doesn't change.

$p(A_1) = 2^{-1} + \epsilon$
$p(A_2) = 2^{-3} - \epsilon$
$p(A_3) = 2^{-3} - \epsilon$
$p(A_4) = 2^{-3} + \epsilon$
$p(A_5) = 2^{-4} - \epsilon$
$p(A_6) = 2^{-4} + \epsilon$

where ϵ is some small fraction, say $\epsilon = \frac{1}{10000}$, for example.

3. **(5 points)**

Marking:
1 mark for each part of the question. This question was very straightforward (you know it or you don't), hence partial marks were given for only very minor mistakes.

Draw the binary tree representations of a:

(a) unary code

(b) Golomb code ($b = 4$),

(c) Elias code

for the numbers from $\{1, 2, \ldots, 10\}$.

Solution:

i	$Unary$	$Golomb, b = 4$	$Elias$
1	1	100	1
2	01	101	010
3	001	110	011
4	0001	111	00100
5	00001	0100	00101
6	000001	0101	00110
7	0000001	0110	00111
8	00000001	0111	0001000
9	000000001	00100	0001001
10	0000000001	00101	0001010
11	00000000001	00110	0001011
12	000000000001	00111	0001100
13	0000000000001	000100	0001101
14	00000000000001	000101	0001110
15	000000000000001	000110	0001111

4. **(2 points)**

Marking:
1 mark for each part of the question..

Use Jensen's inequality to derive an upper bound on the following:

$$(a) \qquad \sum_{i=N_1}^{N_2} \log \log i$$

$$(b) \qquad \sum_{i=1}^{N} \sqrt{i}$$

Solution:

(a) $\sum_{i=N_1}^{N_2} \log\log i$

$$\begin{aligned}
\sum_{i=N_1}^{N_2} \log\log i &= \sum_{i=N_1}^{N_2} \frac{(N_2 - N_1 + 1)}{(N_2 - N_1 + 1)} \log\log i \\
&= (N_2 - N_1 + 1) \sum_{i=N_1}^{N_2} \frac{1}{(N_2 - N_1 + 1)} \log\log i \\
&\leq (N_2 - N_1 + 1) \log(\sum_{i=N_1}^{N_2} \frac{1}{(N_2 - N_1 + 1)} \log i) \\
&\leq (N_2 - N_1 + 1) \log(\log(\sum_{i=N_1}^{N_2} \frac{1}{(N_2 - N_1 + 1)} i)) \\
&= (N_2 - N_1 + 1) \log(\log(\frac{1}{(N_2 - N_1 + 1)} \sum_{i=N_1}^{N_2} i)) \\
&= (N_2 - N_1 + 1) \log(\log(\frac{1}{(N_2 - N_1 + 1)} \frac{(N_1 + N_2)(N_2 - N_1 + 1)}{2})) \\
&= (N_2 - N_1 + 1) \log(\log(\frac{(N_1 + N_2)}{2}))
\end{aligned}$$

(b) $\sum_{i=1}^{N} \sqrt{i}$

$$\begin{aligned}
\sum_{i=1}^{N} \sqrt{i} &= \sum_{i=1}^{N} \frac{N}{N} \sqrt{i} \\
&= N \sum_{i=1}^{N} \frac{1}{N} \sqrt{i} \\
&\leq N \sqrt{\sum_{i=1}^{N} \frac{1}{N} i} \\
&= N \sqrt{\frac{1}{N} \sum_{i=1}^{N} i} \\
&= N \sqrt{\frac{1}{N} \frac{N(N+1)}{2}} \\
&= N \sqrt{\frac{(N+1)}{2}}
\end{aligned}$$

5. **(3 points)**

   Marking:
   1 mark for part (a) and 2 marks for part (b).

   Consider Bernoulli trials with probability p_0 for the occurence of a 0 and probability $1 - p_0$ for the occurence of a 1.

   (a) What is the entropy for a single Bernoulli trial?

   Solution:

$$H = p_0 \log(\frac{1}{p_0}) + (1 - p_0) \log(\frac{1}{(1 - p_0)})$$

   (b) What is the entropy for the run lengths?

$$\begin{aligned}
H &= -\sum_{i=1}^{\infty} p_i \log(p_i) \\
&= -\sum_{i=1}^{\infty} p_0^{i-1}(1 - p_0) \log(p_0^{i-1}(1 - p_0)) \\
&= -\sum_{i=1}^{\infty} p_0^{i-1}(1 - p_0) \log(p_0^{i-1}) - \sum_{i=1}^{\infty} p_0^{i-1}(1 - p_0) \log(1 - p_0) \\
&= -(1 - p_0) \log(p_0) \sum_{i=1}^{\infty} (i - 1) p_0^{i-1} - (1 - p_0) \log(1 - p_0) \sum_{i=1}^{\infty} p_0^{i-1} \\
&= -(1 - p_0) \log(p_0)(\frac{p_0}{(1 - p_0)^2}) - (1 - p_0) \log(1 - p_0) \frac{1}{(1 - p_0)} \\
&= \frac{-p_0 \log(p_0)}{1 - p_0} - \log(1 - p_0) \\
&= \frac{p_0 \log(\frac{1}{p_0})}{1 - p_0} + \log(\frac{1}{1 - p_0})
\end{aligned}$$

6. **(3 points)**

   Marking:
   1 point for the correct λ_i, 2 points for an example of the code. Part marks were given.

   Consider a probability function

$$p(i) = i2^{-(i+1)}$$

on the positive integers $p \in (1, 2, 3, ...)$.

Define a code $C(i)$ such that the average codeword length is as close to the entropy as possible. Be sure to specify $\lambda(i)$ for this code.

Solution: We want $p(i) \approx 2^{-\lambda_i}$. Solving for λ_i we get:

$$\lambda_i = (i+1) - \log(i)$$

The main idea is that we have a unary code but there are little "hickups" that occur at every power of 2 that slow down the code. I've drawn up an example of what the code would look like for $i \in (1, ..., 16)$ below.

i	λ_i
1	2.000000
2	2.000000
3	2.415037
4	3.000000
5	3.678072
6	4.415037
7	5.192645
8	6.000000
9	6.830075
10	7.678072
11	8.540568
12	9.415037
13	10.299560
14	11.192645
15	12.093109
16	13.000000
17	13.912537
18	14.830075
19	15.752072
20	16.678072

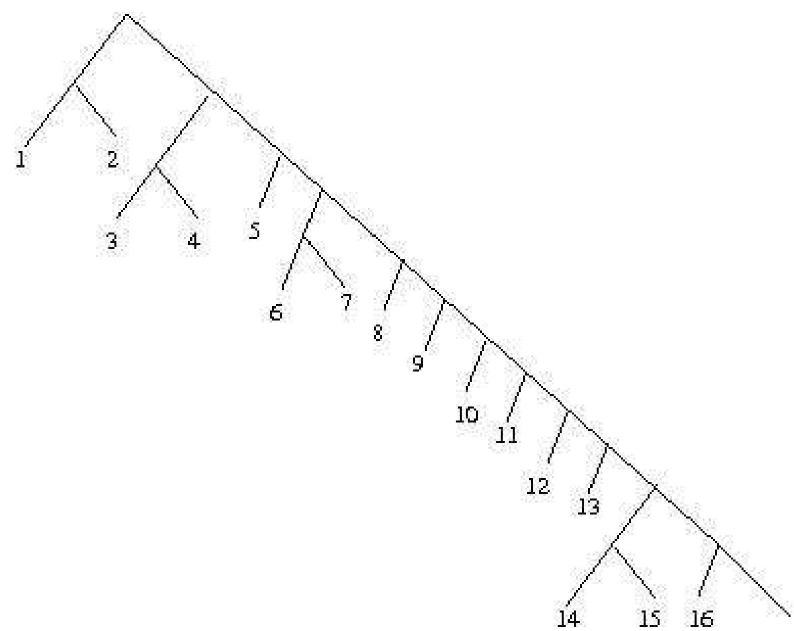

Figure 1: Example of code for $i \in (1, ..., 16)$