Math 4280

Additional Problem 1/30/08.

Let S be a source with probabilities $1/3, 1/3, 1/4, 1/12$

a) What is the average word length of a binary optimal code for S?

We find the average length of a binary Huffman code, which is optimal.

Let $S = \{s_1, s_2, s_3, s_4\}$
$\quad 1/3, 1/3, 1/4, 1/12$

Let $s_5 = s_3 \vee s_4$; $p_5 = 1/4 + 1/12 = 1/3$

$S' = \{s_1, s_2, s_5\}$
$\quad 1/3, 1/3, 1/3$

Let $s_6 = s_2 \vee s_5$ $\quad p_6 = 1/3 + 1/3 = 2/3$

$S'' = \{s_6, s_1\}$
$\quad 2/3 \; 1/3$

Let $s_7 = s_6 \vee s_5$ $\quad p_7 = 2/3 + 1/3 = 1$

So $L(C) = p_5 + p_6 + p_7 = 1/3 + 2/3 + 1 = 2$

b) What are the possible word lengths l_1, l_2, l_3, l_4 for a binary optimal code for S?

The above reduction of S gives

$s_6 \longrightarrow 0$
$s_1 \longrightarrow 1$
$s_2 \longrightarrow 00$
$s_5 \longrightarrow 01$
$s_3 \longrightarrow 010$
$s_4 \longrightarrow 011$

So $C = \{1, 00, 010, 011\}$ is an optimal code for S with word lengths

$l_1 = 1, \; l_2 = 2, \; l_3 = l_4 = 3$

1/30/08

There is another way to reduce S.

$S = \{s_1, s_2, s_3, s_4\}$

$s_5 = s_3 \vee s_4 \quad p_5 = 1/4 + 1/12 = 1/3$ as before

$S' = \{s_1, s_2, s_5\}$

$1/3, 1/3, 1/3$

Now let $s_6 = s_1 \vee s_2 \quad p_6 = 1/3 + 1/3 = 2/3$, since s_1, s_2 & s_5 all have prob $1/3$ we can combine any two of them to make s_6.

$S'' = \{s_6, s_5\}$

$s_7 = s_6 \vee s_5 \quad p_7 = 1.$

Now
$s_6 \longrightarrow 0$
$s_5 \longrightarrow 1$
$s_1 \longrightarrow 00$
$s_2 \longrightarrow 01$
$s_3 \longrightarrow 10$
$s_4 \longrightarrow 11$

So $C = \{00, 01, 10, 11\}$ is an optimal code with ... word lengths $l_1 = l_2 = l_3 = l_4 = 2$.

We must also show that these sets of word lengths $\{1, 2, 3, 3\}$ and $\{2, 2, 2, 2\}$ are the only possibilities.

Suppose that $l_1 \le l_2 \le l_3 \le l_4$ are word lengths of an optimal code for S.

Then since optimal codes are instantaneous $\sum_{i=1}^{4} \frac{1}{2^{l_i}} \le 1$ and since the average length of an optimal code is 2,

$$\frac{1}{3} l_1 + \frac{1}{3} l_2 + \frac{1}{4} l_3 + \frac{1}{12} l_4 = 2.$$

$l_1 = \min(l_1, l_2, l_3, l_4)$ so $l_1 \le 2$, the average word length. Since l_1 is a positive integer $l_1 = 1$ or $l_1 = 2$.

1/30/08

Case 1. $l_1 = 1$,
If $l_2 = 1$ also then

$$\sum_{i=1}^{4} \frac{1}{2^{l_i}} = \frac{1}{2} + \frac{1}{2} + \frac{1}{2^{l_3}} + \frac{1}{2^{l_4}} = 1 + \frac{1}{2^{l_3}} + \frac{1}{2^{l_4}} > 1,$$

so the code can't be instantaneous.
Thus $l_1 = 1$ and $l_2 \geq 2$,

If $l_1 = 1$ and $l_2 = 2$ and $l_3 = 2$ then

$$\sum_{i=1}^{4} \frac{1}{2^{l_i}} = \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{2^{l_4}} = 1 + \frac{1}{2^{l_4}} > 1$$

so $l_3 \geq 3$ and as $l_4 \geq l_3 \geq 3$ $l_4 \geq 3$.

If $l_1 = 1$, $l_2 = 2$, $l_3 = 3$, $l_4 = 3$ is the sol'n we know about.

If $l_1 = 1$ and $l_2 > 2 \Rightarrow l_2 \geq 3 \Rightarrow l_3 \geq 3 \Rightarrow l_4 \geq 3$

$\Rightarrow \frac{1}{3} l_1 + \frac{1}{3} l_2 + \frac{1}{4} l_3 + \frac{1}{12} l_4 \geq \frac{1}{3} + 1 + \frac{3}{4} + \frac{1}{4} = 2 + \frac{1}{3} > 2$ so can't be optimal

Thus the only sol'n with $l_1 = 1$ is 1, 2, 3, 3.

Case 2 $l_1 = 2$.

Since $l_1 \leq l_2 \leq l_3 \leq l_4 \Rightarrow l_2 \geq 2$, $l_3 \geq 2$, $l_4 \geq 2$.

We know that $l_1 = l_2 = l_3 = l_4 = 2$ is our other sol'n.

If any of l_2, l_3 or l_4 is strictly greater than 2, then so is l_4, in this case

$$\frac{1}{3} l_1 + \frac{1}{3} l_2 + \frac{1}{4} l_3 + \frac{1}{12} l_4 > \frac{2}{3} + \frac{2}{3} + \frac{2}{4} + \frac{2}{12} = 2$$

So no other solutions with $l_1 = 2$.

1/30/08

(c) Describe all the binary optimal codes for S.

<u>Case 1</u> $l_1 = l_2 = l_3 = l_4 = 2$.

In this case there is only one possible set of code words $\{00, 01, 10, 11\}$, and any one-to-one map from $\{s_1, s_2, s_3, s_4\}$ to $\{00, 01, 10, 11\}$ gives a code. There are 4! such maps.

<u>Case 2</u> $l_1 = 1$, $l_2 = 2$, $l_3 = l_4 = 3$.

First we count the possible sets of code words: There are $\binom{2}{1} = 2$ ways to pick the word of length 1, leaving 2 words of length 2 & 4 words of length 3. There are $\binom{2}{1} = 2$ ways to pick the word of length 2, leaving 2 words of length 3. There is $\binom{2}{2} = 1$ one way to pick the words of length 3. So there are $2 \cdot 2 \cdot 1 = 4$ sets of code words.

(Explicitly $\{0, 10, 110, 111\}$, $\{0, 11, 100, 101\}$, $\{1, 00, 010, 011\}$, $\{1, 01, 000, 001\}$)

a map from $S = \{s_1, s_2, s_3, s_4\}$ to one of these sets gives an optimal code if and only if $p_i > p_j$ implies $l_i \leq l_j$.

Since $p_1 = p_2 = 1/3$, s_1 can map to the word of length 1 or the word of length 2, 2 choices, then s_2 must map to the other one, 1 choice. Since $l_3 = l_4 = 3$ s_3 can map to either word of length 3, 2 choices, and s_4 must then map to the other one, 1 choice.
So for each set of code-words, there are $2 \cdot 1 \cdot 2 \cdot 1 = 4$, optimal codes.

This gives $4 \cdot 4 = 16$ optimal codes.

B

1/30/08

(d) Which of the binary optimal codes for S are Huffman codes.

<u>Case 1</u> $l_1 = l_2 = l_3 = l_4$.

Binary Huffman codes for these word lengths arise from the following reduction of the source.

$S = \{s_1, s_2, s_3, s_4\}$
$\quad\;\; 1/3 \quad 1/3 \quad 1/4 \quad 1/12$

$s_5 = s_3 \vee s_4$, $p_5 = 1/4 + 1/12 = 1/3$

$S' = \{s_5, s_1, s_2\}$
$\quad\;\; 1/3, \; 1/3 \;\; 1/3$

$s_6 = s_1 \vee s_2 \quad p_6 = 1/3 + 1/3 = 2/3$

$S'' = \{s_6, s_5\}$
$\quad\;\; 2/3, \; 1/3$

$s_7 = s_5 \vee s_6 \quad p_7 = 1$.

Since $s_5 = s_3 \vee s_4$ and $s_6 = s_1 \vee s_2$, any binary Huffman code will have s_1 and s_2 starting with the same bit and s_3 and s_4 starting with the same bit.

4 choices for s_1
1 choice for s_2
2 choices for s_3
1 choice for s_4

So $4 \cdot 1 \cdot 2 \cdot 1 = 8$ binary Huffman codes.

(d)

Case 2 $\quad l_1 = 1,\ l_2 = 2,\ l_3 = l_4 = 3.$

the binary Huffman code with these word lengths come from the following source reduction

$$S = \{s_1, s_2, s_3, s_4\}$$
$$\tfrac{1}{3},\ \tfrac{1}{3},\ \tfrac{1}{4},\ \tfrac{1}{12}$$
$$s_5 = s_3 \vee s_4$$

$S' = \{s_1, s_2, s_5\}$ $\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}$ $s_6 = s_2 \vee s_5$ $S'' = \{s_6, s_1\}$ $\tfrac{2}{3}\ \ \tfrac{1}{3}$ $S''' = \{s_7\}$	or $\quad S' = \{s_2, s_1, s_5\}$ $\tfrac{1}{3}\ \tfrac{1}{3}\ \tfrac{1}{3}$ $s_6 = s_1 \vee s_5$ $S'' = \{s_6, s_2\}$ $\tfrac{2}{3}, \tfrac{1}{3}$ $S''' = \{s_7\}$

In constructing a binary Huffman code from this reduction, there are 2 choices for C''
2 choices for C'
& 2 choices for C.

So $2 \cdot 2 \cdot 2 = 8$ codes

The same 8 choices in making C using this reduction. So there are $8 + 8 = 16$, binary Huffman codes, i.e., all the optimal codes with word lengths $l_1 = 1$, $l_2 = 2$, $l_3 = l_4 = 4$.