Solutions to Exercise Set #3

1. Consider any binary tree with leaves at depth $l_1, \ldots, l_M$. Show that

$$\sum_{i=1}^{M} 2^{-l_i} \leq 1.$$

Solution: First note that for any binary tree with leaves at depth $l_1, \ldots, l_M$, there is a corresponding binary prefix code with codeword length $l_1, \ldots, l_M$. We briefly prove that for any binary prefix code, the codeword lengths $l_1, \ldots, l_M$ must satisfy the inequality

$$\sum_{i=1}^{M} 2^{-l_i} \leq 1. \tag{1}$$

Define the length of the longest codeword as $l_{\max} = \max_i l_i$. Now consider all nodes of the tree at depth $l_{\max}$. A codeword at depth $l_{\max}$ has $2^{l_{\max}-l_i}$ leaves. Since the total number of leaves at depth $l_{\max}$ is $2^{l_{\max}}$. We must have

$$\sum_{i=1}^{M} 2^{l_{\max}-l_i} \leq 2^{l_{max}}.$$

Therefore, (1) is satisfied.

2. Consider an i.i.d. source with alphabet $\{A, B, C, D\}$ and probabilities $\{0.1, 0.2, 0.3, 0.4\}$.

   (a) Find a Tunstall code that encodes the source phrases into binary codewords of length 4. Compute the average number of binary code symbols per source symbol and compare it to the entropy (with the appropriate base).

   (b) Now suppose the Tunstall code you found in part (a) is followed by a binary Huffman code. Compute the average number of binary code symbols per source symbol.

   (c) Find a Tunstall code that encodes the source phrases into quaternary code words of length 2. Compute the average number of quaternary code symbols per source symbol and compare it to the entropy (with the appropriate base).

   (d) Suppose the Tunstall code found in part (c) is followed by a quaternary Huffman code. Compute the average number of quaternary code symbols per source symbol and compare it to the entropy (with the appropriate base).

Solution:

(a) Refer to scanned pages for procedure. The codeword mappings can be given as the following.

$$\begin{array}{llll} A \rightarrow 0000 & BA \rightarrow 0001 & BB \rightarrow 0010 & BD \rightarrow 0011 \\ BC \rightarrow 0100 & DA \rightarrow 0101 & DB \rightarrow 0110 & DC \rightarrow 0111 \\ CA \rightarrow 1000 & CB \rightarrow 1001 & CD \rightarrow 1010 & CC \rightarrow 1011 \\ DDA \rightarrow 1100 & DDB \rightarrow 1101 & DDD \rightarrow 1110 & DDC \rightarrow 1111 \end{array}$$

(b) In order to encode the Tunstall code given in the previous problem into a binary Huffman code, we sort the Tunstall codewords with respect to their probabilities and perform Huffman coding on those probabilities. Continued on scanned pages.

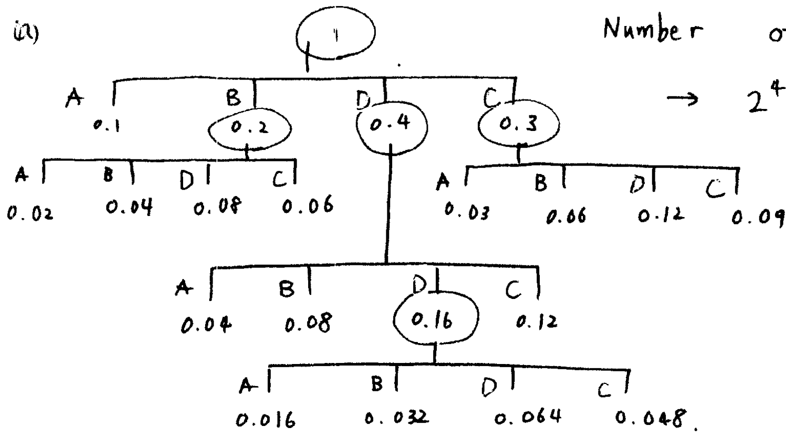
(c) The total number of codewords needed for a quaternary code of length 2 is $M = 42 = 16$. Hence we can use the results found in part (a) with different codeword mappings. The codeword mappings can be given as the following.

$$\begin{array}{llll} A \rightarrow 00 & BA \rightarrow 01 & BB \rightarrow 02 & BD \rightarrow 03 \\ BC \rightarrow 10 & DA \rightarrow 11 & DB \rightarrow 12 & DC \rightarrow 13 \\ CA \rightarrow 20 & CB \rightarrow 21 & CD \rightarrow 22 & CC \rightarrow 23 \\ DDA \rightarrow 30 & DDB \rightarrow 31 & DDD \rightarrow 32 & DDC \rightarrow 33 \end{array}$$

Continued on scanned pages.

(d) In order to encode the Tunstall code given in the previous problem into a binary Huffman code, we sort the Tunstall codewords with respect to their probabilities and perform Huffman coding on those probabilities. Continued on scanned pages.

(a)



Number of codewords needed

$\rightarrow$ $2^4 = 16$

Sum of internal node $= 2.06$.

$$\frac{\text{AVG \# of code symbols}}{\text{AVG \# of source symbols}} = \frac{4}{2.06} = 1.942.$$

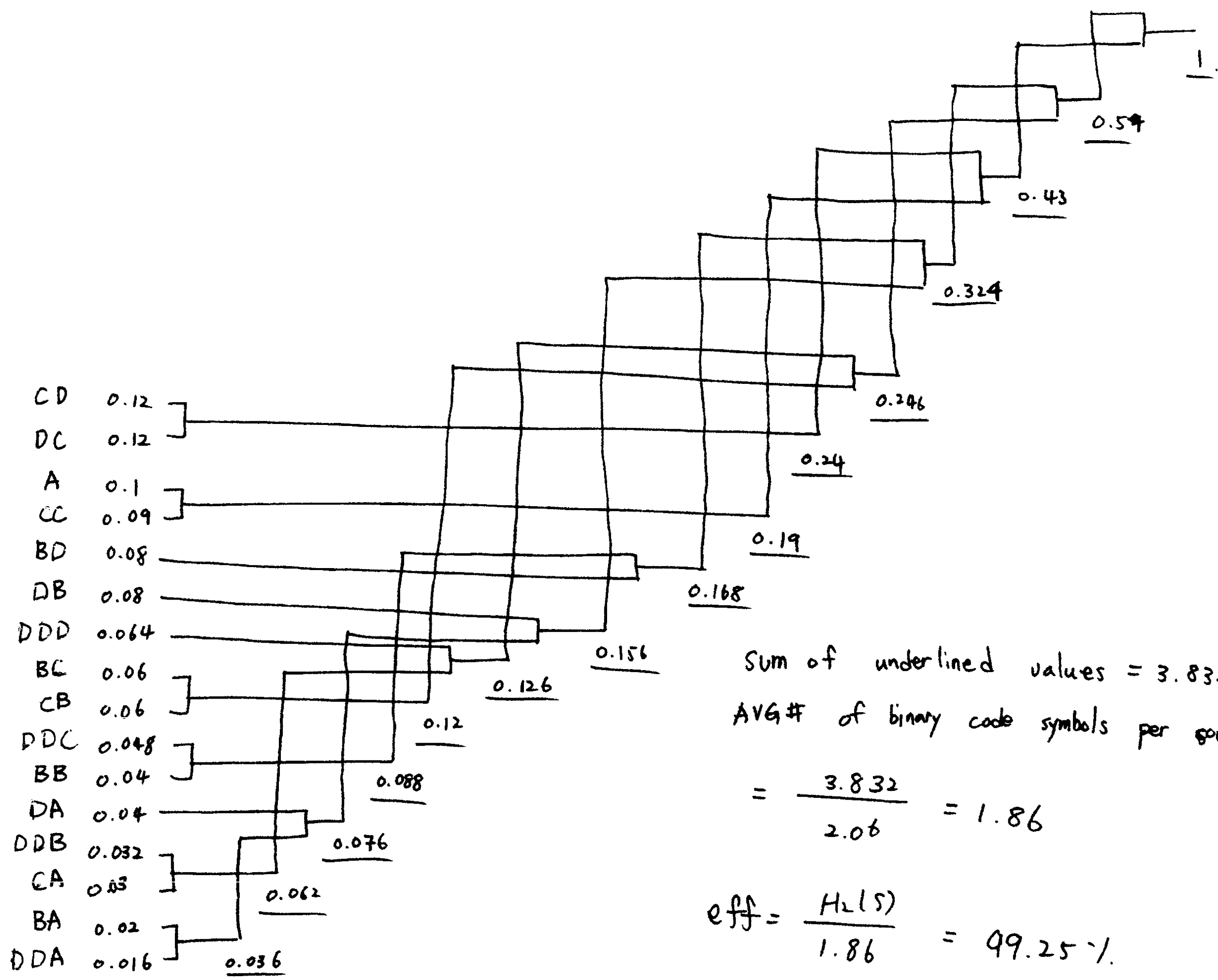
$$\text{eff} = \frac{H_2(S)}{1.942} = 95.06\%$$

(c)

$$\frac{\text{AVG \# of code symbols}}{\text{AVG \# of source symbols}} = \frac{2}{2.06} = 0.9709$$

$$\text{eff} = \frac{H_4(S)}{0.9709} = 95.07\%$$

(b)



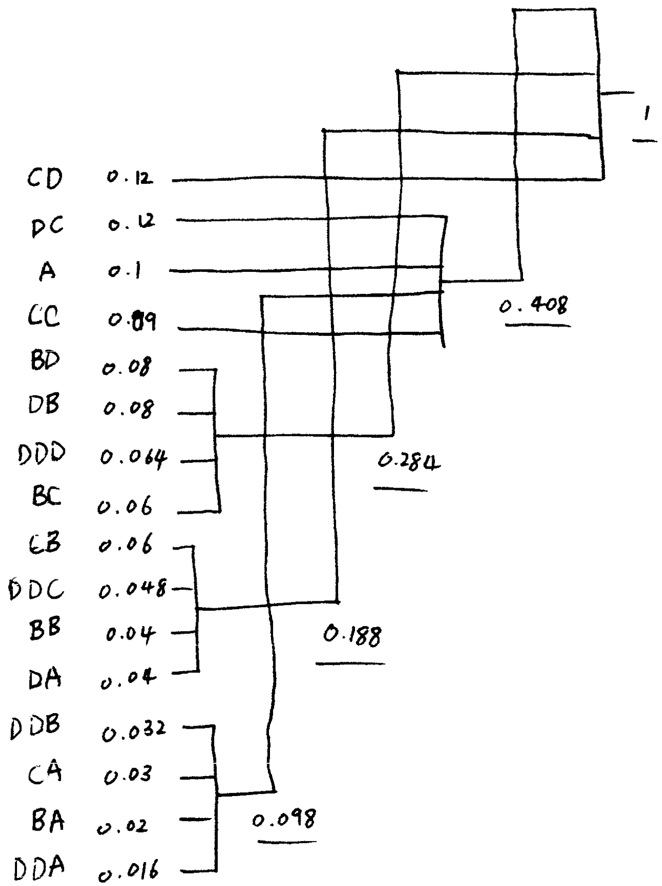
Sum of underlined values = 3.832

AVG # of binary code symbols per source symbol

$$= \frac{3.832}{2.06} = 1.86$$

$$\text{eff} = \frac{H_2(S)}{1.86} = 99.25\%$$

(d)



sum of underlined values = 1.978

AVG # of code symbols per source symbol

$$= \frac{1.978}{2.06} = 0.9602$$

$$\text{eff} = \frac{H_4(S)}{0.9602} = 96.13\%$$