GATE-EC2023

Jay Vikrant EE22BTECH11025

Q65EC.2023: The frequency of occurrence of 8 symbols (a-h) is shown in the table below. A symbol is chosen and it is determined by asking a series of "yes/no" questions which are assumed to be truthfully answered. The average number of questions when asked in the most efficient sequence, to determine the chosen symbol, is

Symbols	Frequency of occurance
a	$\frac{1}{2}$
b	$\frac{1}{4}$
c	$\frac{1}{8}$
d	$\frac{1}{16}$
e	$\frac{1}{32}$
f	$\frac{1}{64}$
g	$\frac{1}{128}$
h	$\frac{1}{128}$

Solution:

Parameter	Value	Description
X	$1 \le X \le 8$	number of symbols
l	2	base of algorithm
$H(X)$	$\sum_i p_X(i) \log_l\left(\frac{1}{p_X(i)}\right)$	average number of question

$$H(X) = \sum_i p_X(i) \log_b\left(\frac{1}{p_X(i)}\right) \tag{1}$$

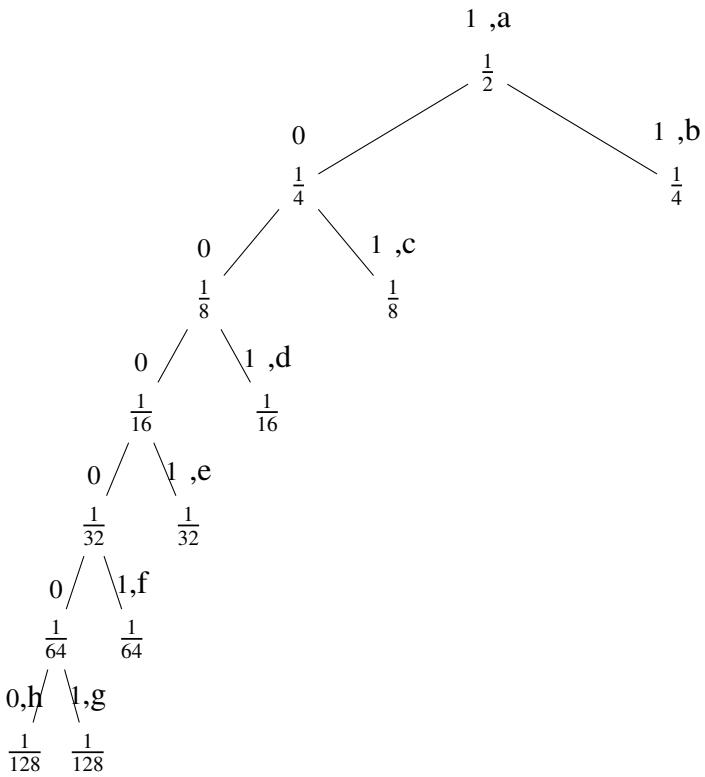
$$= \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + ... + \frac{1}{128}\log_2(128) \tag{2}$$

$$= 0.5 + 0.5 + 0.375 + ... + 0.0078125 \tag{3}$$

$$= 1.984375 \tag{4}$$

Now, finding the average using Huffman code, We start from a frequency of 1 and distribute it uniformly. The following conventions is used,

symbol	alloted bit
occured	1
not occured	0



Using the above binary table following code is generated;

Symbols	Frequency	Code	Size
a	$\frac{1}{2}$	1	0.5
b	$\frac{1}{4}$	01	0.25
c	$\frac{1}{8}$	001	0.125
d	$\frac{1}{16}$	0001	0.0625
e	$\frac{1}{32}$	00001	0.03125
f	$\frac{1}{64}$	000001	0.015625
g	$\frac{1}{128}$	0000001	0.0078125
h	$\frac{1}{128}$	0000000	0.0078125

TABLE 0
HUFFMAN TABLE

The average number of question = Weighted path length = 1.9844