भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology
Hyderabad

Source Coding

Teaching Learning Center
IIT Hyderabad

Theresh Babu Benguluri and G V V Sharma*

Contents

***Abstract*—This manual introduces Shannon's source coding theorem.**

1 Huffman Code

Problem 1.1. $X \in \mathcal{X} = \{x_1, x_2, \ldots, x_5\}$ and $Pr\{X = x_1\} = Pr\{X = x_2\} = 0.25; Pr\{X = x_4\} = Pr\{X = x_5\} = 0.25; Pr\{X = x_3\} = 0.2$. Construct the Huffman Code for this set.

Solution: The following steps are outlined in Fig. 1.1 resulting in Table I.

1) List the symbols in decreasing order of probability.
2) Fuse the two symbols with the lowest probabilities and add their probablities.
3) Assign 0,1 to the branches.
4) Reorder probabilities and start from step 2
5) Repeat the steps until you get 1.

Problem 1.2. Compute the *average code length* in Problem 1.1.

*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in.

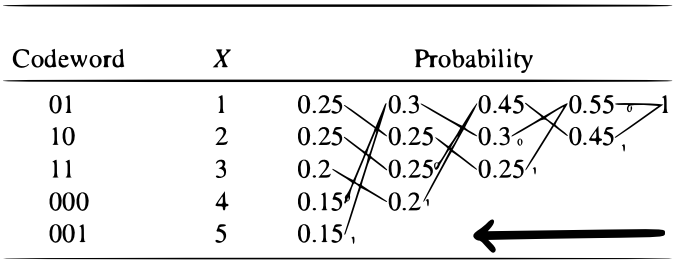


Fig. 1.1

HUFFMAN CODING			
symbol	Code	Length	Prob
x_1	01	2	0.25
x_2	10	2	0.25
x_3	11	2	0.2
x_4	000	3	0.15
x_5	001	3	0.15

TABLE I: Huffman Tree

Solution: Average Code length $\bar{L}$,

$$\begin{aligned}\bar{L} &= \sum_i p_i l_i \\ &= (3 \times 0.15) + (2 \times 0.2) + (3 \times 0.15) + (2 \times 0.25) \\ &\quad + (2 \times 0.25) \\ &= 0.45 + 0.4 + 0.45 + 0.5 + 0.5 \\ \bar{L} &= 2.3 \text{ bits per symbol}\end{aligned}$$

Definition 1.1. The *Entropy* of a code is defined as

$$H(X) = E\left[\log_2\left(\frac{1}{p(X)}\right)\right] = \sum_i p(x_i) \log_2\left[\frac{1}{p(x_i)}\right]. \tag{1.2.1}$$

Problem 1.3. Find the entropy of the above code.

Solution:

$$H(X) = -\sum_i p_i \log_2 p_i = -(2 \times 0.15 \log_2 0.15 + 2 \times 0.25 \log_2 0.25 + 0.2 \log_2 0.2) = 2.2855 \text{ bits per symbol} \quad (1.3.1)$$

Problem 1.4. Verify that

$$\bar{L} = H(X) + \epsilon \quad (1.4.1)$$

2 Average code length and Entropy

Definition 2.1. A single variable function f is said to be convex if

$$f[\lambda x + (1-\lambda) y] \leq \lambda f(x) + (1-\lambda) f(y), \quad (2.0.1)$$

for $0 < \lambda < 1$.

Problem 2.1. Execute the following python script. Is $\ln x$ convex or concave?

```
import numpy as np
import matplotlib.pyplot as plt


#Plotting log(x)
x = np.linspace(1,8,50)#points on
    the x axis
f=np.log(x)#Objective function
plt.plot(x,f,color=(1,0,1))
plt.grid()
plt.xlabel('$x$')
plt.ylabel('$\ln x$')

#Convexity/Concavity
a = 2
b = 7
lamda = 0.4
c = lamda *a + (1-lamda)*b
f_a = np.log(a)
f_b = np.log(b)

f_c = np.log(c)
f_c_hat = lamda *f_a + (1-lamda)*
    f_b

#Plot commands
plt.plot([a,a],[0,f_a],color
    =(1,0,0),marker='o',label="$f(a)
    $")
plt.plot([b,b],[0,f_b],color
    =(0,1,0),marker='o',label="$f(b)
    $")
plt.plot([c,c],[0,f_c],color
    =(0,0,1),marker='o',label="$f(\
    lambda a +(1-\lambda)b)$")
plt.plot([c,c],[0,f_c_hat],color
    =(1/2,2/3,3/4),marker='o',label=
    "$\lambda f(a) +(1-\lambda)f(b)
    $")
plt.plot([a,b],[f_a,f_b],color
    =(0,1,1))
plt.legend(loc=2)
#plt.savefig('../figs/1.1.eps')
plt.show()#Reveals the plot
```

Definition 2.2. If p_i and q_i are two probability mass functions

$$D(p \parallel q) = \sum_i p_i \log_2 \frac{p_i}{q_i} \quad (2.1.1)$$

is defined as the *relative entropy or Kullback Leibler distance Or KL Divergence* between the two probability mass functions p_i and q_i.

Problem 2.2. Using (2.1.1) and (2.0.1), show that

$$D(p \parallel q) \geq 0 \quad (2.2.1)$$

Problem 2.3. Let

$$q_i = \frac{2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}} \quad (2.3.1)$$

Show that q_i forms probability distribution.

Problem 2.4. Using K-L divergence, show that

$$\bar{L} \geq H(X) + \epsilon, \quad \epsilon > 0 \quad (2.4.1)$$

and find the value of ϵ. This is the converse of Shannon's *source coding theorem.*

3 Shannon Source Coding Theorem

Problem 3.1. Using Lagrange Multipliers, solve

$$\min_{l_i} \bar{L} = \sum_{i=0}^{M-1} p_i l_i \quad (3.1.1)$$

$$\text{such that } \sum_{i=0}^{M-1} 2^{-l_i} \leq 1. \quad (3.1.2)$$

Solution: Let

$$F = \sum_{i=0}^{M-1} p_i l_i + \lambda \left(\sum_{i=0}^{M-1} 2^{-l_i} - 1 \right) \tag{3.1.3}$$

Then,

$$\frac{dF}{dl_i} = 0, \quad i = 0, 1, \ldots M-1 \tag{3.1.4}$$

$$\implies p_i + \lambda 2^{-l_i}(-1)\ln 2 = 0 \tag{3.1.5}$$

$$\implies l_i = \log_2 \left[\frac{\lambda \ln 2}{p_i} \right] \tag{3.1.6}$$

To find λ,

$$\sum_{i=0}^{M-1} 2^{-\log_2 \left[\frac{\lambda \ln 2}{p_i} \right]} = 1 \tag{3.1.7}$$

$$\implies \lambda = \frac{1}{\ln 2} \tag{3.1.8}$$

Substituting in (3.1.6),

$$\boxed{l_i = \log_2 \frac{1}{p_i}} \tag{3.1.9}$$

Problem 3.2. Show that

$$\boxed{H(X) \leq \bar{L} \leq H(X) + 1} \tag{3.2.1}$$

Solution: Since,

$$\log_2 \frac{1}{p_i} \leq \lceil \log_2 \frac{1}{p_i} \rceil \leq \log_2 \frac{1}{p_i} + 1, \tag{3.2.2}$$

$$\sum_{i=0}^{M-1} p_i \log_2 \frac{1}{p_i} \leq \sum_{i=0}^{M-1} p_i \lceil \log_2 \frac{1}{p_i} \rceil \leq \sum_{i=0}^{M-1} p_i \log_2 \frac{1}{p_i} + 1 \tag{3.2.3}$$

resulting in (3.2.1).

3.1 Achievability

Problem 3.3. Consider the sequence, $\bar{X} = X_0, \ldots, X_{n-1}$ If the symbols are i.i.d, show that

$$H(\bar{X}) = H(X_0) + \cdots + H(X_{n-1}) \tag{3.3.1}$$

Problem 3.4. Let the *rate* of a code

$$R \triangleq \lim_{n \to \infty} \frac{\bar{L_n}}{n}. \tag{3.4.1}$$

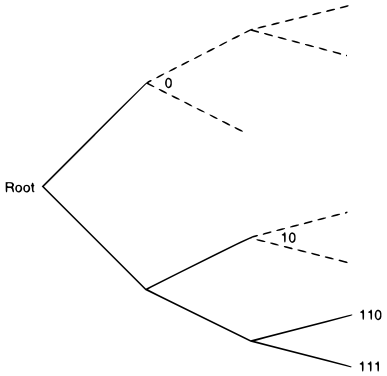
where $\bar{L_n}$ is the average length of the code. Show that

$$\boxed{R \approx H(X)}. \tag{3.4.2}$$

4 Kraft's Inequality

Problem 4.1. For a prefix code, show that

$$\sum_{i=0}^{M-1} 2^{-l_i} \leq 1 \tag{4.1.1}$$



- Consider a Binary tree in which each node has 2 children.Let the branches of the tree represent the symbols of the codeword.
- Let l_{max} be the length of the longest codeword of the set of codewords.
- A codeword at depth l_i has $2^{l_{max}-1}$ leaves underneath itself at depth l_{max}.The sets of leaves under codewords are disjoint.The total number of leaves under codewords are less than or equal to $2^{l_{max}}$.Thus we have

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

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

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

- Therefore,*(Krafts Ineqality:)* for any prefix free code the codeword lengths $l_0, \ldots, l_{M-1}$ must satisfy the ineqality

$$\boxed{\sum_{i=0}^{M-1} 2^{-l_i} \leq 1} \tag{4.1.2}$$

- Conversely, If l_i satisfy condition above,there exist a prefix free code with codeword lengths l_i.