課堂作業3 solution

1. 給定p1=1/3,p2=1/4, p3,=1/5, p4=1/6，以及p5=1/20。求其Huffman編碼，並將求得之平均編碼長度與其熵值做比較，並計算編碼冗贅(coding redundancy)。

Ans: 熵值=2.14 bits，Huffman平均編碼長度=2.22 bits

1. Huffman樹可能不只一個。試以p1=1/3,p2=1/3, p3=1/6, p4=1/6為例，畫出兩種不同的Huffman樹，其中一棵Huffman樹需是編碼長度變異最小。驗證這兩種Huffman樹有相同的平均編碼長度。
2. 利用huff-enc.exe壓縮下列檔案，並說明比較結果。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 檔案 | 原始大小(Bytes) | 壓縮後  (Bytes) | 壓縮比 | 位元率  (Bit Rate) |
| world95.txt | 3,005,020 | 1935693 | 1.552 | 5.153 |
| water\_lilies.bmp | 1,440,054 | 1275792 | 1.129 | 7.087 |
| EXCEL.EXE | 34,482,256 | 26615945 | 1.296 | 6.175 |
| Bach-PartitaEmajor-44kHz-Stereo-16bit.wav | 9,580,594 | 8981514 | 1.067 | 7.500 |

1. 根據HUFF.C編解碼的方式，將以下編碼結果(採16進制)解碼，

61 66 04 03 00 00 02 01 00 e0 fa d3

其中第一個及第二個位元組 61及66表示符號源內的符號為{a,b,c,d,e,f}，61為a的ASCII code，66為f的ASCII code。後面緊接著每個符號出現的次數，出現次數分別為4,3,0,0,2,1，從d7 90開始為原始內容的Huffman編碼結果。

1. 看懂HUFF.C作法寫出a,b,c,d,e,f的Huffman碼。
2. 畫出huffman樹。
3. 將huffman編碼結果e0 fa d3解碼還原出原始內容。

Ans:

(a)

node='a' count= 4 child\_0= 0 child\_1= 0 Huffman code=11

node='b' count= 3 child\_0= 0 child\_1= 0 Huffman code=10

node='e' count= 2 child\_0= 0 child\_1= 0 Huffman code=00

node='f' count= 1 child\_0= 0 child\_1= 0 Huffman code=010

node=256 count= 1 child\_0= 0 child\_1= 0 Huffman code=011

node=257 count= 2 child\_0='f' child\_1=256

node=258 count= 4 child\_0='e' child\_1=257

node=259 count= 7 child\_0='b' child\_1='a'

node=260 count= 11 child\_0=258 child\_1=259

(b)

(c)

e0 fa d3

11 10 00 00 11 11 10 10 11 010 011

a b e e a a b b a f EOS

* 解碼結果：abeeaabbaf