

Huffman Code

Huffman Code 是一種很有效的編碼壓縮法，它根據一篇明文中的各個字母出現的機率將字母轉換成二進制的編碼，編碼方式如下：

假設一個明文為「mississippis」

m 出現的頻率 $1/12$

p 出現的頻率為 $2/12$

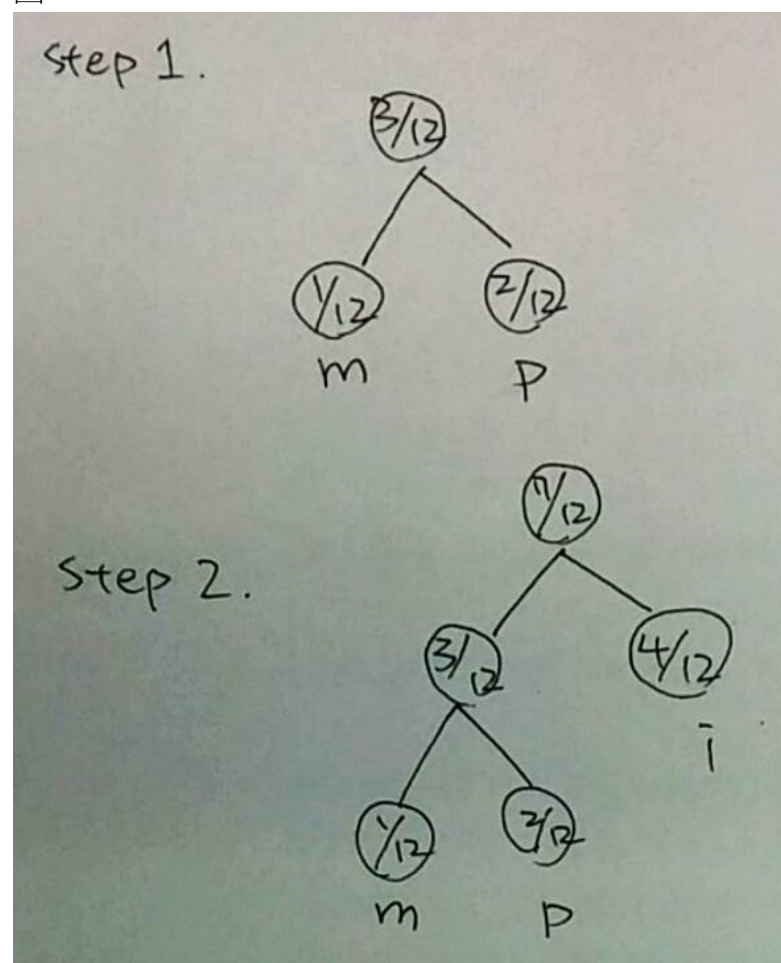
i 出現的頻率 $4/12$

s 出現的頻率 $5/12$

我們將頻率由小到大排列為 $\{1/12, 2/12, 4/12, 5/12\}$

我們每次都選出其中最小的來建成 tree，所以我們會先選 $1/12$ ，再來是 $2/12$ ，並將其頻率相加變成 $3/12$ 並放回數列中，接下來我們選出 $3/12$ 與 $4/12$ 相加變成 $7/12$ ，並放回數列中，接下來我們選出 $5/12$ 與 $7/12$ 相加得到 1，完成 Huffman Tree 的建構，注意我們每次都會將頻率較小的那一個放在左子樹，頻率較大的放在右子樹。

最後我們根據建好的樹，左樹給 0 右樹給 1 規則，為每個字母編碼，步驟如附圖：



m:100

p:101