

- 作業內容

- 請將上一份作業的文本用 Huffman code 的方式以字為單位編碼產出 codebook

EX:

Input : aabbbbcc

Output(codebook) : (word) (code)

b	1
a	10
c	00

- 程式需求

- 要求用 argc 與 argv 的寫法
在 cmd 裡輸入 410xxxxxx.exe input.txt output.txt 就可以執行的寫法。輸入是文本(input.txt)，輸出亦為文本(output.txt)
- 讀取模式要用“rb”，讀取模式要用“wb”。
- 記得要排序，從編碼長度由短排到長。
- 記得加註解，切勿抄襲。

- 繳交格式

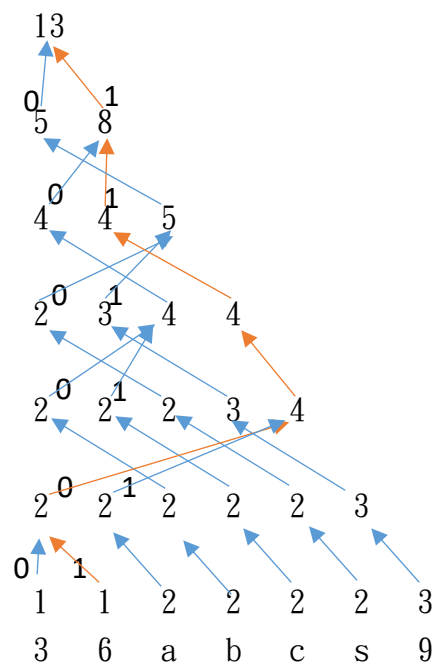
- 請交.c 檔，檔名是你的學號 Ex:410xxxxxx.c
- 不須壓縮，請直接上傳數位學苑
- 請務必照著格式繳交，不然可能會扣分。

- 提醒

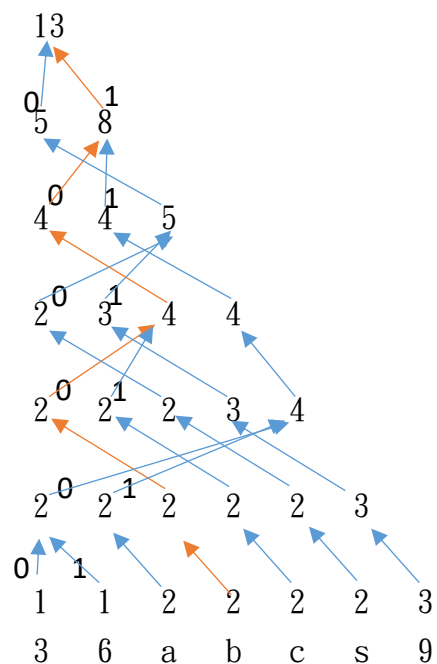
- 有任何問題，都歡迎討論

4. 由上往下將每個跟得 0/1 依序排列，找出每個 symbol 的編碼。

EX: symbol “6” 的編碼 : 1101



EX: symbol “b” 的編碼 : 100



■ 一種簡單的編碼演算法(你可以自己想，也可以用這個就好)

假設：

Count	8	9	12	13	26
Symbol	a	b	c	d	e
Position	0	1	2	3	4

此方法是以“位置”來做判別，將每一次最小兩項相加後所到的位置記錄下來。

68				0
26	42			1
17	25	26		1
12	13	17	26	2
8	9	12	13	26
a	b	c	d	e

“2110”就記錄了這棵樹長高過程中，我將其命名為編碼序。

如何通過編碼序來產生編碼？

a 的編碼

位置	最小兩項相加後的位置	過程	編碼
起始位置 → 0	2110	在位置 0，相加後到位置 2	0
2	2110	在位置 2，相加後到位置 0	X
0	2110	在位置 0，相加後到位置 1	0
1	2110	在位置 1，相加後到位置 0	1
0			

所以 a 的編碼就是 “100”

b 的編碼

位置	最小兩項相加後的位置	過程	編碼
起始位置 → 1	2110	在位置 1，相加後到位置 2	1
2	2110	在位置 2，相加後到位置 0	X
0	2110	在位置 0，相加後到位置 1	0
1	2110	在位置 1，相加後到位置 0	1
0			

所以 b 的編碼就是 “101”

c 的編碼

	位置	最小兩項相加後的位置	過程	編碼
起始位置	→2	2110	在位置 2，相加後到位置 0	X
	0	2110	在位置 0，相加後到位置 1	0
	1	2110	在位置 1，相加後到位置 1	1
	1	2110	在位置 1，相加後到位置 0	1
	0			
所以 c 的編碼就是 “110”				

d 的編碼

	位置	最小兩項相加後的位置	過程	編碼
起始位置	→3	2110	在位置 3，相加後到位置 1	X
	1	2110	在位置 1，相加後到位置 1	1
	1	2110	在位置 1，相加後到位置 1	1
	1	2110	在位置 1，相加後到位置 0	1
	0			
所以 d 的編碼就是 “111”				

e 的編碼

	位置	最小兩項相加後的位置	過程	編碼
起始位置	→4	2110	在位置 4，相加後到位置 3	X
	3	2110	在位置 3，相加後到位置 2	X
	2	2110	在位置 2，相加後到位置 0	X
	0	2110	在位置 0，相加後到位置 0	0
	0			
所以 e 的編碼就是 “0”				