2.5 - Huffman màxim

Solució:

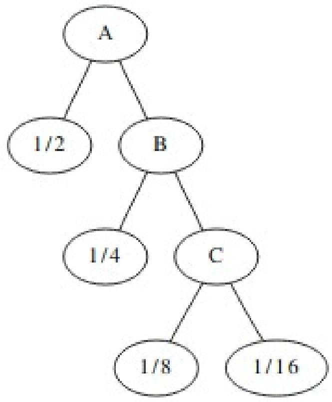
Considerem el següent conjunt de freqüències:[1]

$$f_{n-i} = \frac{1}{2^i}, \quad \text{per a } i \in \{1, \ldots, n-1\}$$

per a un alfabet de mida n. La paraula codificada més llarga tindrà longitud $n-1$. La suma acumulada de totes les freqüències fins a la $(j-1)$-èsima freqüència més petita (incloent aquesta) serà menor que f_j, és a dir,

$$\sum_{i=1}^{j-1} f_i < f_j. \tag{1}$$

Aleshores, per construcció, l'arbre serà una "columna vertebral" amb una única fulla a cada subarbre intermig (és a dir, un arbre de mínim factor de balanceig, o un arbre binari complet de màxima alçada).



De fet, la condició de l'Equació 1 es pot relaxar, perquè el que volem a cada iteració de l'algorisme de Huffman és que el subarbre més alt creat sigui un fill de l'arbre que crearem a aquella iteració (però no importa si és el fill esquerre o el dret). Per exemple: a la primera iteració crearem l'arbre de freqüència $f_1 + f_2$, que té alçada 1; a la següent iteració voldrem que aquest subarbre sigui un dels fills del que crearem i, per tant, haurà de tenir una de les dues freqüències més petites de la cua de prioritat. Per tal que això passi, requerim que $f_1 + f_2 < f_4$ (perquè $f_3 < f_4, \ldots, f_n$). De manera que el nou arbre tindrà ara freqüència $f_1 + f_2 + f_3$ (i serà l'arbre més alt, i un dels dos de freqüència més petita, de la cua de prioritat).

Per tant, per tal de construir l'arbre més alt possible, en realitat només necessite que les freqüències $f_1, \ldots, f_n$ tinguin la següent propietat:

$$\forall k : 2 \le k \le n-2 : \left(\sum_{i=1}^{k} f_i \right) < f_{k+2}.$$

Una assignació de freqüències que compleix això és la proposada anteriorment, però n'hi ha infinites més (p.ex., per $n = 6$: $\{1, 2, 3, 5, 8, 13\}$).

[1] No hi ha una única solució. N'hi ha d'altres assignacions de freqüències que permeten obtenir el mateix tipus d'arbre.