

# Compression

Introduction:

Pour des raisons de stockage et de vitesse de transmission on est obligé d'utiliser des algorithmes de compression, notamment dans le domaine des images et des réseaux. Il existe 2 type de compression :

**réversible**                    l'information décompressée sera identique à l'information compressée  
**non réversible**                l'information décompressée est différente de l'information compressée,

## Type RLC (Run Length Coding)

On ne code pas individuellement chaque élément mais le couple valeur nombre de répétition. exemple AAAAAAAAAA devient 9A. Ce codage simple peut être couplé avec d'autres types de compressions.(Codage Horizontal, Vertical, Zigzag) Codage RLE utilisé dans les formats TIFF,PCX,BMP.

## Type VLC (Variable Length Coding)

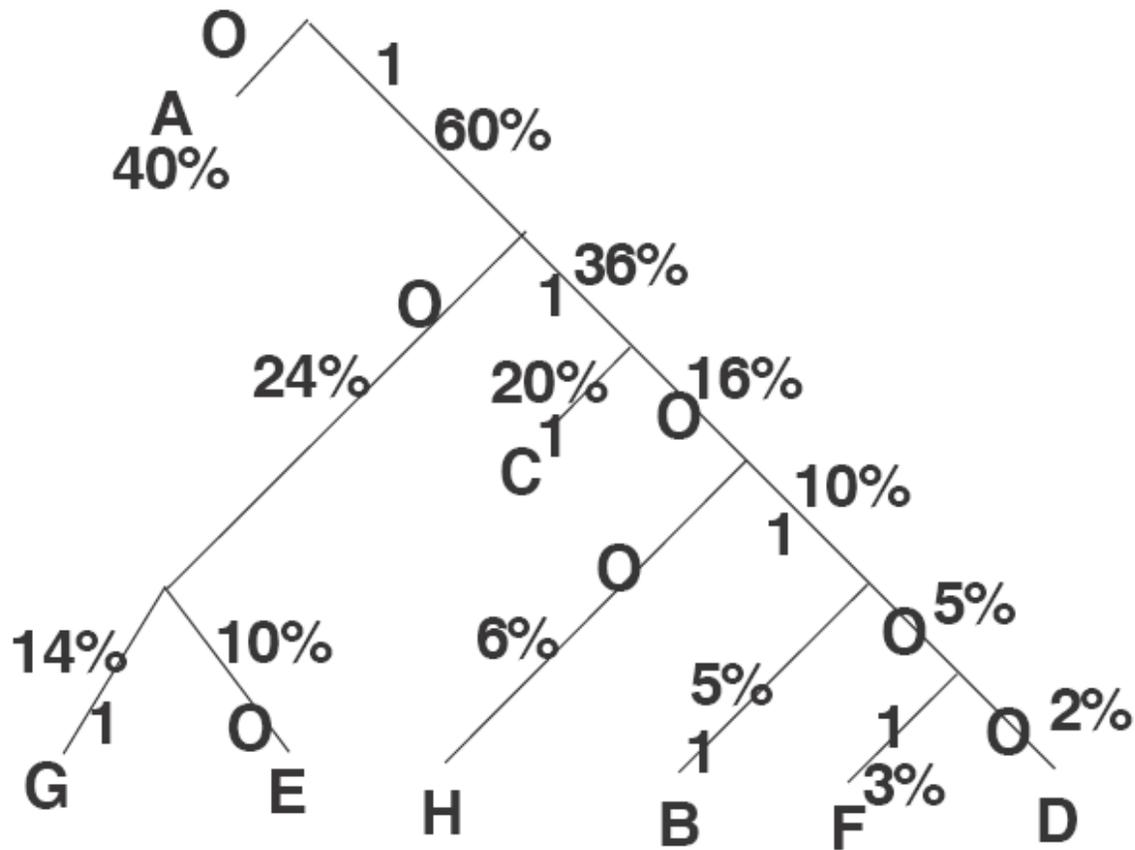
On recherche la probabilité d'apparition de chaque élément constituant le message puis on code les éléments en fonction de leur probabilité. On code l'élément qui a la plus grande probabilité avec le code le plus petit. Cette méthode nécessite une analyse préalable. On peut faire référence à un dictionnaire spécifique préétabli, en fonction du langage utilisé. (Non Adaptatif Codage de Huffman 1952)

## Codage LZW 1984(Lempel Ziv Welch)

Adaptatif ,il crée un dictionnaire au cours de l'analyse du fichier en codant des séquences d'éléments. Ce codage utilisé dans les formats GIFF et TIFF et dans les compresseurs PKZIP,ARJ,LHA

Arbre de Huffman :considérons un texte composé avec 8 éléments A,B,C,D,E,F,G,H avec les probabilités suivantes

$p(A)=40\%$ , $p(B)5\%$ , $p(C)20\%$ , $p(D)2\%$ , $p(E)10\%$ , $p(F)3\%$ , $p(G)14\%$ , $p(H)6\%$



Le codage de vient A=0 ,B=11011, C=111, D=110100, E=100, F=110101, G=101, H=1100  
 longueur moyenne du message 2.51 à comparer avec 3 bits normalement nécessaire.