

# Experts en Informatique

## Le codage de Huffman

# Qu'est ce que c'est?

David Huffman a donné son nom à une méthode de compression des données sans perte de qualité. Elle est fondée sur un principe de fréquence d'apparition des lettres dans un texte.

**Plus la lettre sera rare, plus son code sera long, et inversement,**

**Plus elle est fréquente plus le code correspondant sera court.**

Vous vous demandez sûrement à  
quoi sa sert...

Elle est utilisée en particulier pour la  
transmissions des messages, par télécopie,  
minitel et Internet.

# Comment ça marche?

- Sur nos ordinateurs, les nom des programmes correspondants se terminent en général par Zip.
- Dans chaque message, chaque symbole a une fréquence et leur position dans le texte n'a pas d'importance.
- Chaque symbole est codé en suite de 0 et de 1 de telle façon que les symboles que l'on retrouve souvent dans le texte soient codé par un code plus court que ceux qui sont plus rares.

# Le Codage: (Ex: Le mot BAC)

Imaginons que dans un texte, nous avons a coder le mot BAC.

Comme indication, nous savons que

A est codé 0111

B est codé 010

Et C est codé 10

Donc le mot BAC sera codé : 010011110

Car :010011110

Et ce, n'importe ou qu'il se placera dans le texte.

## Le Décodage :

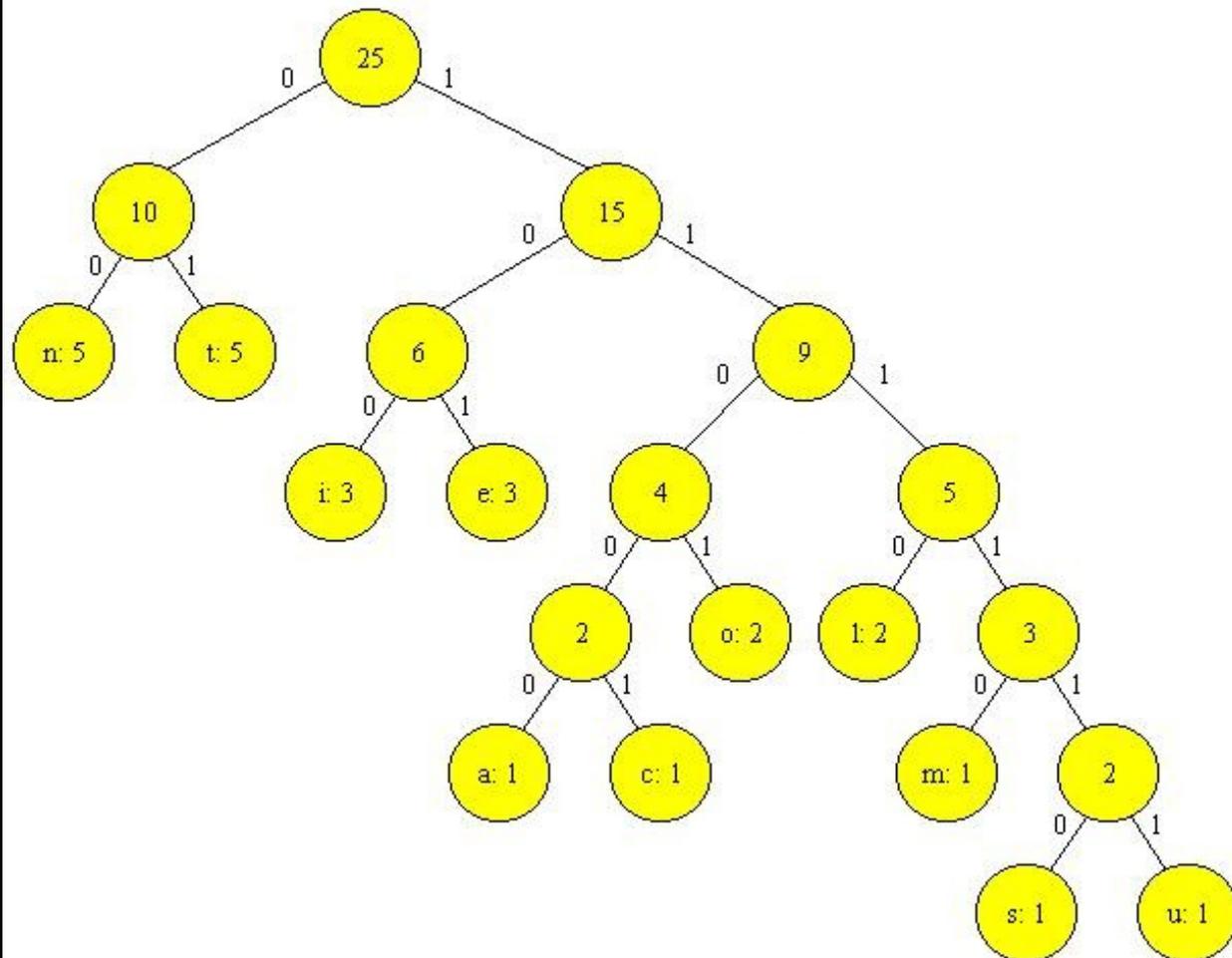
- Un mot est codé 11111010100.
- Comme indication, nous savons que :
- A: 1101
- B: 0
- C: 101
- D: 111
- E: 100

Réponse :

11111010100

Nous obtenons finalement le mot DABE.

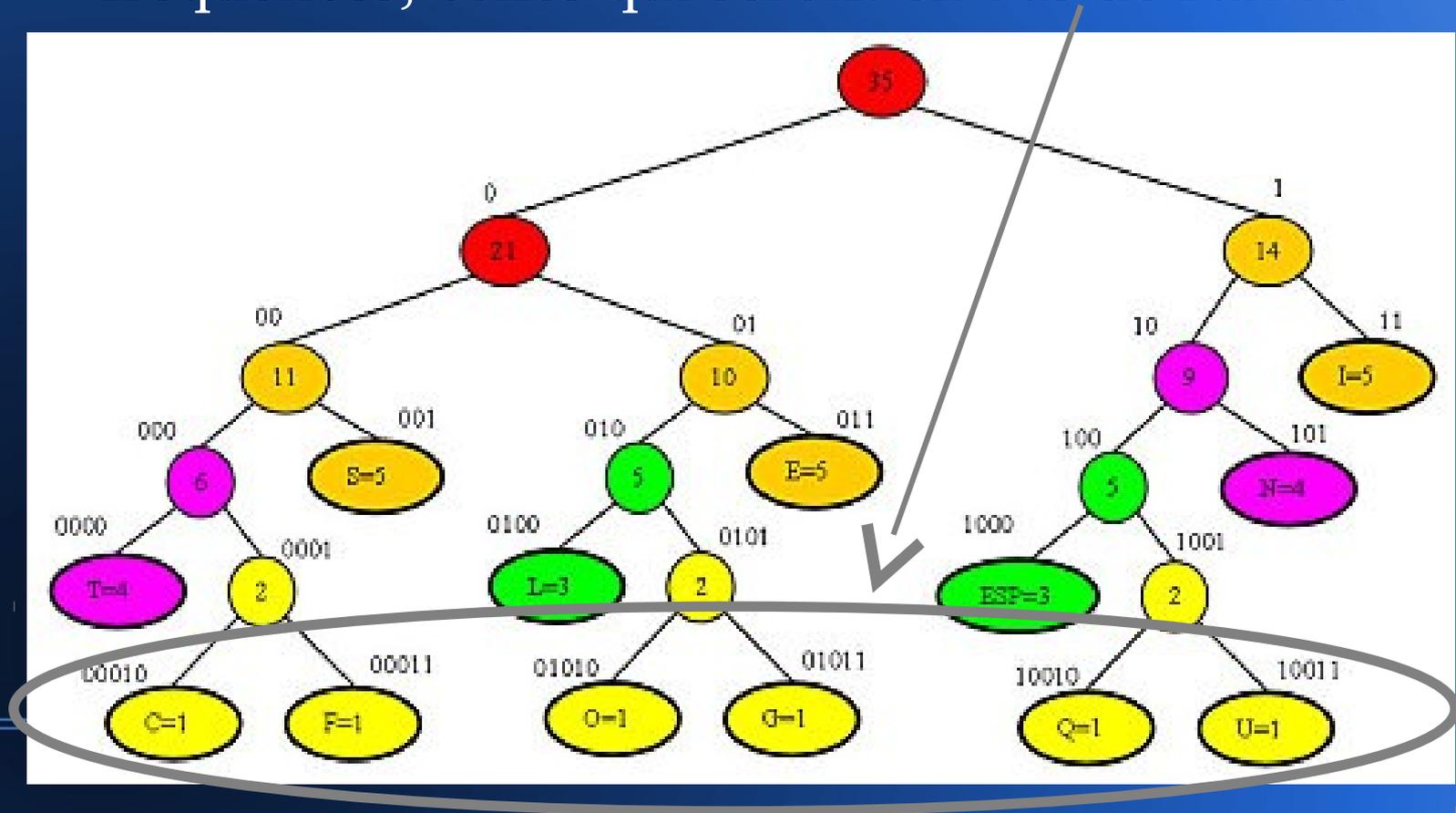
# L'arbre de Huffman...



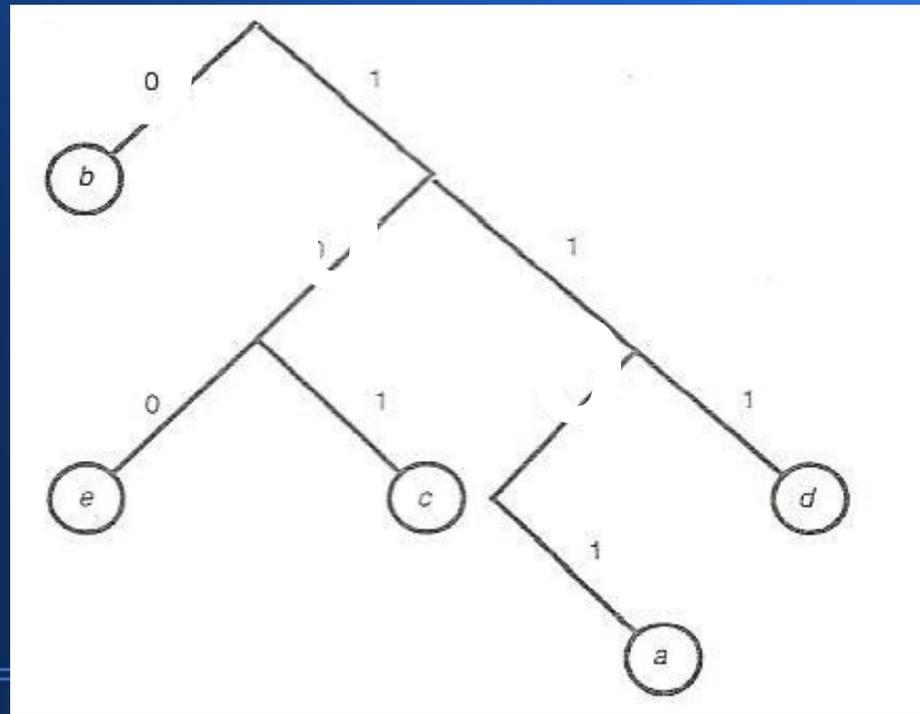
Une méthode simple pour décoder le message est de noter la « règle de codage » sous la forme d'un arbre.

# Pour le construire...

Nous commençons par les plus petites fréquences, celles qui seront en bas de l'arbre



Chaque bit correspond à une descente sur une branche suivant la règle suivante : 0 implique une descente à gauche, 1 une descente à droite. Chaque bit nous fait donc passer d'un nœud (le premier étant la racine) à un autre.



Après lecture d'un certain nombre de bits, nous parvenons à une feuille étiquetée par une lettre. Nous l'écrivons et recommençons à la racine de l'arbre avec le bit suivant.

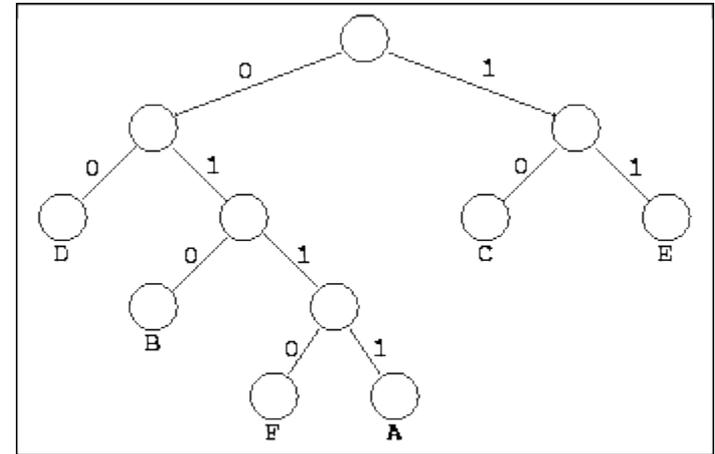


FIG. 2.2 - *Arbre de décodage.*