

Arbres de codage

Algorithme de Huffman

Jean-Marc Vincent¹

¹Laboratoire LIG
Universités de Grenoble
Jean-Marc.Vincent@imag.fr

Claude Shannon (1916-2001)



Claude Elwood Shannon (30 avril 1916 à Gaylord, Michigan - 24 février 2001), ingénieur électrique, est l'un des pères, si ce n'est le père fondateur, de la théorie de l'information. Son nom est attaché à un célèbre "schéma de Shannon" très utilisé en sciences humaines, qu'il a constamment désavoué.

Il étudia le génie électrique et les mathématiques à l'Université du Michigan en 1932. Il utilisa notamment l'algèbre booléenne pour sa maîtrise soutenue en 1938 au MIT. Il y expliqua comment construire des machines à relais en utilisant l'algèbre de Boole pour décrire l'état des relais (1 : fermé, 0 : ouvert).

Shannon travailla 20 ans au MIT, de 1958 à 1978. Parallèlement à ses activités académiques, il travailla aussi aux laboratoires Bell de 1941 à 1972.

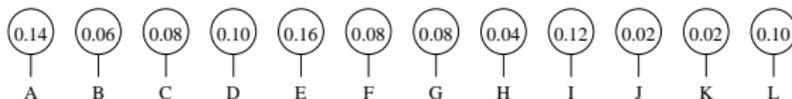
Claude Shannon était connu non seulement pour ses travaux dans les télécommunications, mais aussi pour l'étendue et l'originalité de ses hobbies, comme la jonglerie, la pratique du monocycle et l'invention de machines farfelues : une souris mécanique sachant trouver son chemin dans un labyrinthe, un robot jongleur, un joueur d'échecs (roi tour contre roi)...

Souffrant de la maladie d'Alzheimer dans les dernières années de sa vie, Claude Shannon est mort à 84 ans le 24 février 2001.

[Wikipedia](#)

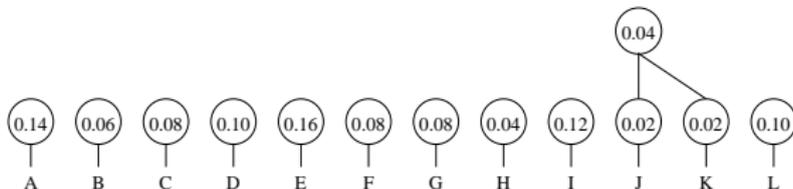
Algorithme de Huffman (1951)

A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10

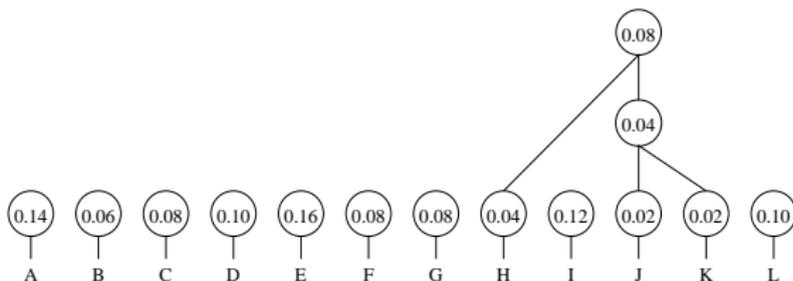


Algorithme de Huffman (1951)

A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10



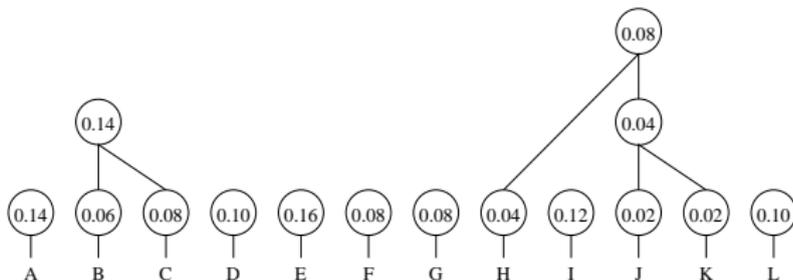
Algorithme de Huffman (1951)



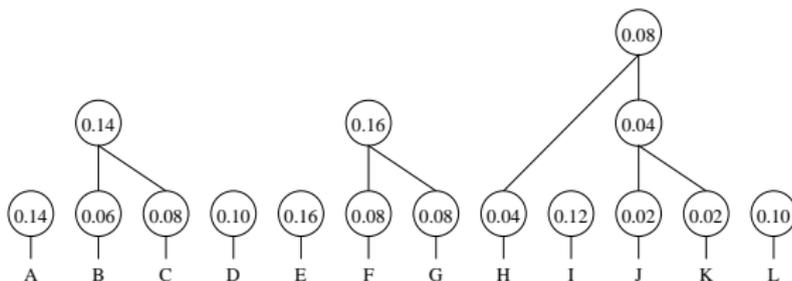
A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10

Algorithme de Huffman (1951)

A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10

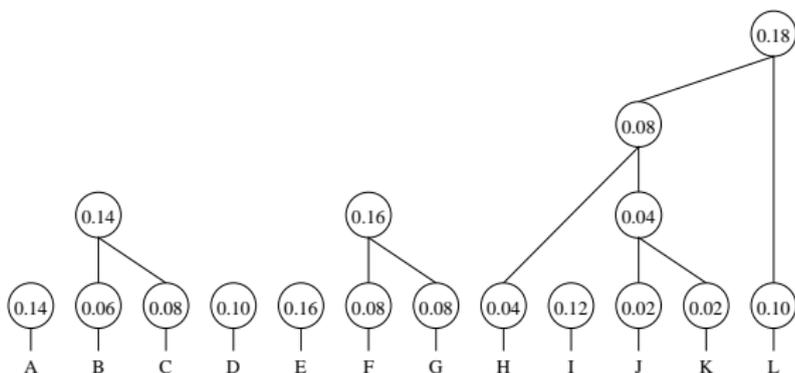


Algorithme de Huffman (1951)



A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10

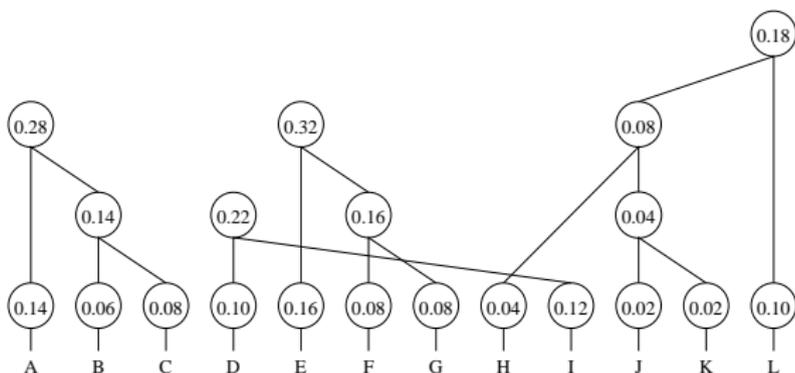
Algorithme de Huffman (1951)



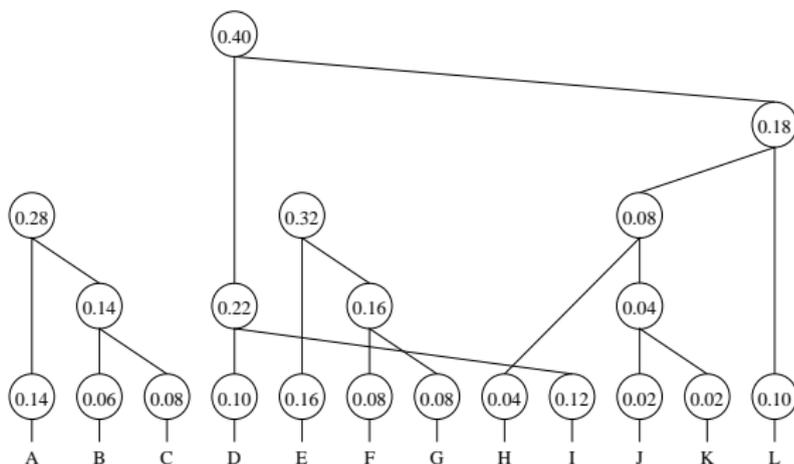
A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10

Algorithme de Huffman (1951)

A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10

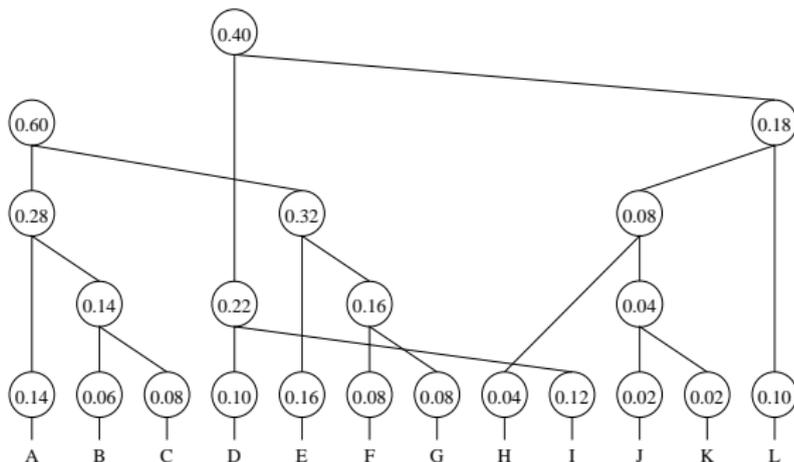


Algorithme de Huffman (1951)



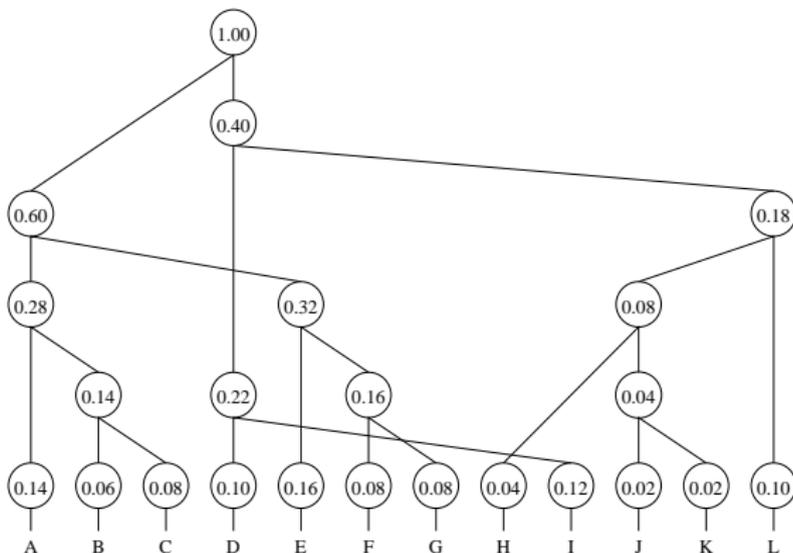
A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10

Algorithme de Huffman (1951)



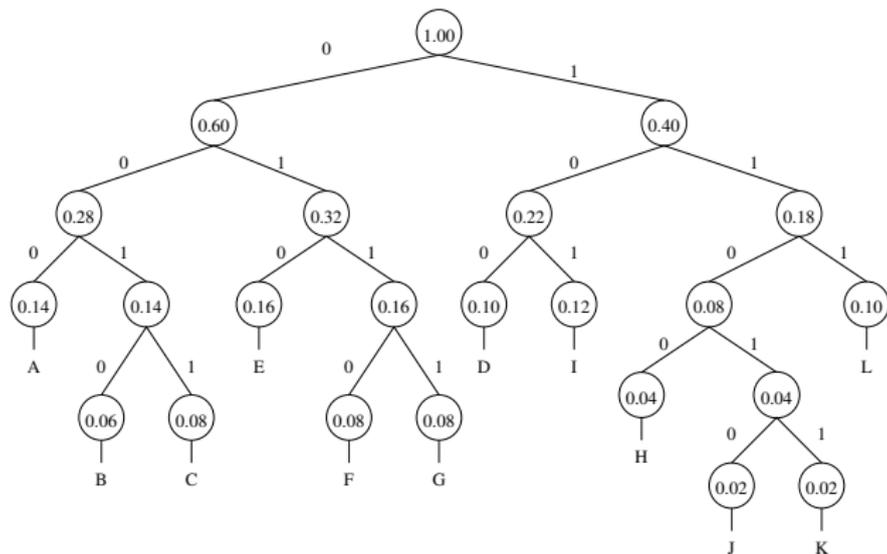
A	0.14
B	0.06
C	0.08
D	0.10
E	0.16
F	0.08
G	0.08
H	0.04
I	0.12
J	0.02
K	0.02
L	0.10

Algorithme de Huffman (1951)



- A 0.14
- B 0.06
- C 0.08
- D 0.10
- E 0.16
- F 0.08
- G 0.08
- H 0.04
- I 0.12
- J 0.02
- K 0.02
- L 0.10

Algorithme de Huffman (1951)



A	0.14	000
B	0.06	0010
C	0.08	0011
D	0.10	100
E	0.16	010
F	0.08	0110
G	0.08	0111
H	0.04	1100
I	0.12	101
J	0.02	11010
K	0.02	11011
L	0.10	111

Codage optimal : L-moy = 3.42, Entropie = 3.38

Profondeur = $-\log_2(\text{probabilité})$

Généralisation Lempel-Ziv,...

Algorithme de Huffman (1951) : Implantation

```
{Alphabet à coder  $C_1, \dots, C_k$  : Caractères}
{Proportion des caractères  $p_1, \dots, p_k$ }
file_priorité Q= $\emptyset$  {File à priorité de nœuds pondérés}
pour i=1 ; i  $\leq$  k ; i++ faire
    z=nouveau_nœud()
    z.etiquette= $C_i$  ; z.poids= $p_i$ 
    z.gauche=NIL ; z.droit=NIL
fin pour
pour i=1 ; i  $\leq$  k-1 ; i++ faire
    z=nouveau_nœud()
    z.gauche= Extrait_min(Q) ; z.droit=Extrait_min(Q)
    z.poids=z.gauche.poids + z.droit.poids
    Insérer(Q,z)
fin pour
retourner Extrait_min(Q)
```