

Compression LZW (Lempel Ziv Welch)

Le principe de la compression LZW est le suivant :

- Le dictionnaire est initialisé avec tous les caractères présents dans le texte.
- On a lu une chaîne 'w' présente dans le dictionnaire et on lit le caractère 'a' qui suit 'w' dans le texte.

On a deux possibilités :

- Si la chaîne 'wa' n'est pas présente dans le dictionnaire, alors on écrit le code correspondant à 'w' dans le fichier de sortie, on ajoute 'wa' dans le dictionnaire et 'a' devient la nouvelle valeur de 'w'.
- Si 'wa' est présente, alors elle devient la nouvelle valeur de 'w'.

Algorithme :

Initialiser dictionnaire D

W ← lire un caractère

Tant qu'il reste des caractères à lire faire

a ← lire caractère

si (w + a est dans D) alors w ← w + a

sinon

 écrire le code de w

 ajouter w + a dans D

 w ← a

fin si

fin tant que

écrire le code de w

Exemple 1: Compression

Entrée : XYZZX,XYZZX (dernier caractère : 4)

Solution :

Dictionnaire initial :

X : 1

Y : 2

Z : 3

, : 4

$W \leftarrow X$ et $a \leftarrow Y$, on ajoute XY : 5 au dictionnaire

Sortie = 1

$W = Y$ et $a \leftarrow Z$, on ajoute YZ : 6 au dictionnaire

Sortie = 12

$W = Z$ et $a = Z$, on ajoute ZZ : 7 au dictionnaire

Sortie= 123

$W = Z$ et $a = X$, on ajoute ZX : 8 au dictionnaire

Sortie = 1233

$W = X$ et $a = ,$, on ajoute X, : 9 au dictionnaire

Sortie : 12331

$W = ,$ et $a = X$ on ajoute ,X :10 au dictionnaire

Sortie : 123314

$W = X$ et $a = Y$ existe donc $w = XY$ et $a = Z$ on ajoute XYZ : 11 au dictionnaire

Sortie : 1233145

$W = Z$ et $a = Z$ existe donc $w = ZZ$ et $a = X$ on ajoute ZZX : 12 au dictionnaire

Sortie : 123314571

Il faut présenter le résultat sous forme de tableaux :

w	a
X	Y
Y	Z
Z	Z
Z	X
X	,
,	X
X	Y
XY	Z
Z	Z
ZZ	X

X	1
Y	2
Z	3
,	4
XY	5
YZ	6
ZZ	7
ZX	8
X,	9
,X	10
XYZ	11
ZZX	12

Sortie : 123314571

Décompression :

Le dictionnaire est construit au fur et à mesure de la décompression

Initialiser dictionnaire D

$a \leftarrow$ lire caractère

écrire D(a)

Tant qu'il reste des caractères à lire faire

$b \leftarrow$ lire caractère

 si b est dans D alors

$w \leftarrow D(b)$

 sinon

$w \leftarrow D(a) + \text{premier caractère de } D(a)$

 fin si

 écrire w

 ajouter D(a)+premier caractère de w à D

$a \leftarrow b$

fin tant que

Exemple 2: Décompression

Entrée : 99 97 103 116 97 258 262 97 (dernier caractère : 256)

Solution :

a	b	w
99	97	a
97	103	g
103	116	t
116	97	a
97	258	ag
258	262	aga
262	97	a
97		

Dictionnaire initial

NUL	0
...	
a	97
b	98
...	
	256
ca	257
ag	258
gt	259
ta	260
aa	261
aga	262
agaa	263

Sortie : cagtaagagaa