

# Compression & Algo du Texte

– M1 –  
2015-2016

---

## Feuille 3 : LZ78 & bzip2

---

► **Exercice 1** ◀ Soit  $u = bbaabaabbbabbabba$ .

- Encodez  $u$  par l'algorithme de LZ78 avec un dictionnaire de taille illimitée.
- Encodez  $u$  par l'algorithme de LZW avec un dictionnaire de taille illimitée.
- On considère qu'on a un dictionnaire de taille max 5, qu'on vide si on n'a plus de place. Quel est le nouvel encodage par LZ78 ?

► **Exercice 2** ◀ Décodez le message suivant, codé avec LZ78 (dico illimité):

$[0, 't', 0, 'o', 0, 'b', 0, 'e', 2, 'r', 0, 'n', 2, 't', 1, 'o', 3, 'e']$

► **Exercice 3** ◀ Décodez le message suivant, codé avec LZW (dico illimité) sur l'alphabet  $A = \{a, b\}$ , où on initialise avec  $a \mapsto 1$  et  $b \mapsto 2$ :

$[1, 2, 1, 5, 4, 4, 5]$

► **Exercice 4** ◀

- Calculez la transformée de Burrows-Wheeler  $v$  de  $u = ananas$ .
- Appliquez Move-to-Front à  $v$ .

► **Exercice 5** ◀ On suppose que l'alphabet de départ est  $A = \{i, m, p, s\}$ , et qu'en appliquant Burrows-Wheeler puis Move-to-Front on arrive à

$[1, 3, 4, 0, 4, 4, 4, 3, 0, 1, 0]$

Reconstituez  $u$ .

► **Exercice 6** ◀ On considère le mot  $u = (abcdefgh)^n$ .

- Quelle est la transformée de Burrows-Wheeler  $v$  de  $u$  ?
- Calculez Move-To-Front( $v$ ).
- Encodez le résultat avec l'algorithme de Huffman. Quel est la taille du message compressé ?

► **Exercice 7** ◀ Dans cet exercice on voit comment calculer la fonction  $\phi$  efficacement. Soit  $v$  la transformée de Burrows-Wheeler de  $u$ . Soit  $A$  l'alphabet de  $v$ .

1. La première étape consiste, pour chaque lettre  $a \in A$ , à stocker dans  $C[a]$  le nombre de positions dans  $v$  où la lettre est strictement plus petite que  $a$ . Expliquez comment le faire en temps linéaire.
2. Si en position  $j$  dans  $v$  on a la  $i$ -ème occurrence de la lettre  $a$ , justifiez que  $\phi(j) = C[a] + i - 1$
3. En déduire un algorithme linéaire pour calculer  $\phi$ .