

Traitement Automatique des Langues (corrections)

Exercice 1 : Coder/décoder

```
1 function encode(ch : chaine, dec : entier) : chaine
2 variables
3     code : chaine
4     i, new_ascii, ascii_base : entier
5     new_char : caractere
6 Debut
7     ascii_base <- Caractere_vers_Ascii('A')
8     pour i allant de 0 a Longueur(ch) - 1 faire
9         // décalage par rapport à zero :
10        new_ascii <- Caractere_vers_Ascii(ch[i]) - ascii_base + dec
11        // homogénéisation et décalage :
12        new_ascii <- new_ascii mod 26 + ascii_base
13        new_char <- Ascii_vers_Caractere(new_ascii)
14        code <- code + new_char
15    fin pour
16    retourner code
17 Fin
```

```
1 function decode(code : chaine, dec : entier) : entier
2 Debut
3     retourner encode(code, 26 - (dec mod 26))
4 Fin
```

```
1 Algorithme Trouver le codage Cesar
2 variables
3     code, reponse : chaine
4     dec : entier
5 Debut
6     code <- Saisie("Message codé :")
7     reponse <- 'N'
8     dec <- 1
9     tant que(reponse != 'O' et reponse != 'o' et dec < 26) faire
10        Ecrire(decode(code, dec))
11        reponse <- Saisie("C'est la bonne ? ('O' : oui, 'N' : non)")
12        dec <- dec + 1
13    fin tant que
14    si (reponse = 'o' ou reponse = 'O') alors
15        Ecrire("Trouvé : decalage " + EnChaine(dec - 1));
16    fin si
17 Fin
```

Exercice 2 : Fréquence des lettres

```
1 function monMajuscule(car : caractere) : caractere
2 Debut
3   si ("éèëëÊÊÊÊ".indexOf(car) != -1) alors car <- "E"
4   sinon si ("àâäåÀÂÃÄ".indexOf(car) != -1) alors car <- "A"
5   sinon si ("îïíîÎÏĨ".indexOf(car) != -1) alors car <- "I"
6   sinon si ("ôöóòÔÕÖ".indexOf(car) != -1) alors car <- "O"
7   sinon si ("ùûüúÛÜÛ".indexOf(car) != -1) alors car <- "U"
8   sinon si ("çÇ".indexOf(car) != -1) alors car <- "C"
9   sinon si ("ñÑ".indexOf(car) != -1) alors car <- "N"
10  fin si
11  retourner car;
```

```
1 function frequence(texte : chaine) : tableau de caracteres
2 variables
3   freq : tableau de 26 caracteres
4   i, ascci, pos : entier
5 Debut
6   pour i allant de 0 a Taille(freq) - 1 faire
7     freq[i] <- 0
8   fin pour
9   pour i allant de 0 a Longueur(text) - 1 faire
10    ascci <- Caractere_vers_Ascii(monMajuscule(text[i]))
11    pos <- ascci - 65
12    si (pos >= 0 et pos < 26)
13      freq[pos] <- freq[pos] + 1
14    fin pour
15    retourner freq;
16 Fin
```

Exercice 4 : Compter les mots

```
1 function est_alphanum(car : caractere) : booleen
2 variables
3   ascci : entier
4 Debut
5   ascci <- Caractere_vers_Ascii(monMajuscule(car))
6   retourner (ascci >= 65 et ascci <= 90) ou (ascci >= 48 et ascci <= 57)
7 Fin
```

```

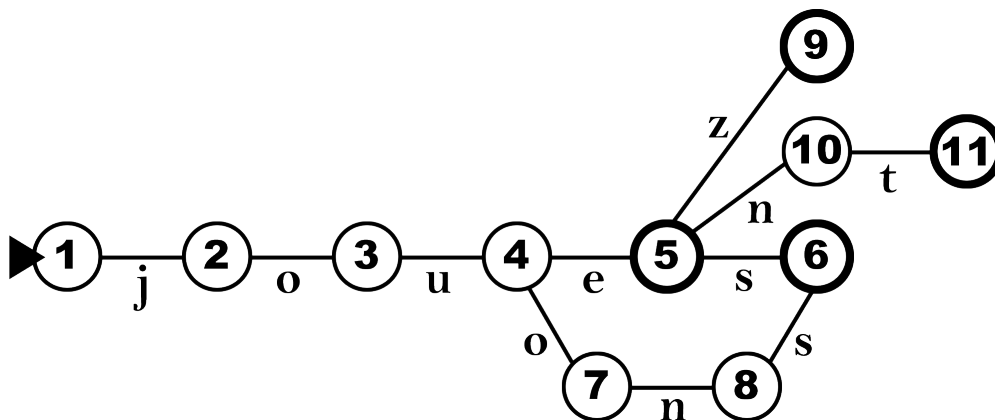
1 function mots(texte : chaine) : entier
2 variables
3   c, i : entier
4   mot : chaine
5 Debut
6   c <- 0
7   mot <- ""
8   pour i allant de 0 a Taille(freq) - 1 faire
9     freq[i] <- 0
10  fin pour
11  pour i allant de 0 a Longueur(texte) - 1 faire
12    si(est_alphanum(texte[i])) alors
13      mot <- mot + texte[i]
14    sinon
15      si(mot != "")
16        // faire quelque chose avec le mot trouvé
17        c <- c + 1
18        mot <- ""
19      finsi
20    fin si
21  fin pour
22  retourner c
23 Fin

```

Exercice 7 : Automate

- “chameau” $\rightarrow \{1, 2, 3, 4, 8, 9, 10, 7\}$
- Le mot “chamelle” ne fait pas partie du langage de l’automate, parce-que après la transition 8-9 l’automate bloque.
- Le mot “chato” ne fait pas partie du langage de l’automate, parce-que l’état 6, n’est pas final.
- {chameau, chacal, chat, chatte, chaton, chevre, chevreuil, chevrot, chiot, chien, chienne}
- Momoire :
 - Liste : $7 + 6 + 4 + 6 + 6 + 6 + 9 + 7 + 5 + 5 + 7 = 68$
 - Automate : $22 + 28 = 50$
- Temps :
 - “Chacal” :
 - * Liste : 13 comparaisons
 - chevre *vs.* chacal = 3 comparaisons
 - chat *vs.* chacal = 4 comparaisons
 - chacal *vs.* chacal = 6 comparaisons
 - * Automate : 10 comparaisons

- “Chiot” :
 - * Liste : 11 comparaisons
 - chevre *vs.* chiot = 3 comparaisons
 - chevrot *vs.* chiot = 3 comparaisons
 - chiot *vs.* chiot = 5 comparaisons
 - * Automate : 8 comparaisons
- “Chenille” :
 - * Liste : 12 comparaisons
 - chevre *vs.* chenille = 4 comparaisons
 - chevrot *vs.* chenille = 4 comparaisons
 - chevreuil *vs.* chenille = 4 comparaisons
 - * Automate : 6 comparaisons
- Six formes conjuguées du verbe “jouer” :



Exercice 8 : Chercher une chaîne correspondant à une expression rationnelle

- $a.^+tion$: {action, alimentation, ... }
- $ch.\{2-4\}$: {chat, chatte, ... }
- $ils?|elles?$: {il, ils, elle, elles}
- $(a|b)^*$: { ϕ , a, b, ab, ba, aab, bba, abb, ... }
- $a^*|b^*$: { ϕ , a, aa, a...a, b, bb, b...b }
- a^*b^* : { ϕ , b, ab, aab, a...ab, a, abb, ab...b, a...ab...b }
- ab^*a : {aa, aba, ab...ba }

Exercice 9 : Écrire des expressions rationnelles

- mots qui commencent par 'M', finissent par 'E' et dont la longueur est comprise entre 2 et 5 : $M.(0-3)E$
- les mots qui contiennent 2 't' consécutifs : $.*tt.*$
- les nombres entiers écrits en binaire : $(0|1)^+$
- les numéros de téléphone français : $/(+33/)(0-9)\{9\}$
- une date au format (JJ/MM/AAAA) : $(0-9)\{2\}//(0-9)\{2\}//(0-9)\{4\}$
- les adresses IP : $(0-9)\{3\}(/.(0-9)\{3\})\{3\}$
- les nombres décimaux (les entiers en font partie) : $(0-9)^+(/.(0-9)^+)?$