

Méthodes Empiriques et Langages de Script

TP 3

Gabriele A. Musillo
musillo4@etu.unige.ch

November 21, 2005

1 Exercice 1: les structures de contrôle

- considérez le pseudo-code de la boucle `repeat-until` suivante et simulez-la au moyen d'une boucle PERL `do-while`:

```
REPEAT
  a_statement;
UNTIL ($p == 0)
```

- formulez la conditionnelle équivalente au bloc suivant:

```
$b = 1;
while ($b && $e) {
  a_statement;
  $b = 0;
}
while ($b) {
  another_statement;
  $b = 0;
}
```

- codez un programme qui permet d'afficher des patterns carrés $N \times N$ telles que:

```
+ + + + +
- + + + +
- - + + +
- - - + +
- - - - +
(un pattern 5 × 5)
```

- codez un programme qui permet d'afficher n'importe quelle pattern telle que:

```

- + + + + + -
- - + + + - -
- - - + - - -
- - - - - - -
- - - + - - -
- - + + + - -
- + + + + + -

```

(une pattern de hauteur 3)

Indication: remarquez et exploitez les symétries de cette pattern.

2 Exercice 2: les n -grammes

Un n -gramme est une séquence de n mots. Vous supposerez que 2 mots sont séparés par un ou plusieurs blancs.

Codez un programme qui lit un fichier de texte et un entier n (≥ 2) et qui écrit, sur la sortie standard, les n -grammes ainsi que leur fréquence.

3 Exercice 3: l'entropie de Shannon

L'entropie $h(\sigma)$ d'un symbole σ est définie de la manière suivante:

$$h(\sigma) = -\log_2 P(\sigma)$$

La probabilité $P(\sigma)$ d'un symbole σ est égale à la fréquence relative du symbole σ . La fréquence relative d'un symbole est le rapport de sa fréquence et de la somme des fréquences de tous les symboles.

L'entropie d'un ensemble de symboles \mathcal{A} est définie ainsi:

$$H(\mathcal{A}) = -\sum_{x \in \mathcal{A}} P(x) \log_2 P(x)$$

Codez un programme qui calcule l'entropie h de chaque symbole apparaissant dans un fichier de texte ainsi que l'entropie H de l'ensemble des symboles du fichier.

PERL met à disposition la fonction `log($n)` qui retourne le logarithme de base e du nombre $\$n$. Pour convertir `log($n)` en base 2, il suffit d'appliquer la formule `log_2($n) = log($n)/log(2)`.