

Corrigé de l'exercice 5 du *TP 2 système et shell*

Le projet Gutenberg met à disposition des utilisateurs des textes du domaine public en format numérique (txt, epub ...) sous licence libre (voir The Gutenberg License).

Le texte brut du “Tour du monde en 80 jours” écrit par Jules Verne est disponible à partir de l'[URL][URL] <http://www.gutenberg.org/ebooks/800.txt.utf-8>.

1. Ouvrir un terminal *shell* et choisir comme répertoire courant `~/TP-SHELL`.
2. Créer un un répertoire `Phileas` puis entrer dans ce répertoire.

```
junier@fredportable:~$ mkdir TP-SHELL
junier@fredportable:~$ cd TP-SHELL/
junier@fredportable:~/TP-SHELL$ mkdir Phileas
junier@fredportable:~/TP-SHELL$ cd Phileas
```

3. Consulter l'aide de la commande `wget` avec `wget --help` ou `man wget` puis télécharger le fichier contenant le texte du “Tour du monde en 80 jours” au format `txt`.

```
junier@fredportable:~/TP-SHELL/Phileas$ wget http://www.gutenberg.org/ebooks/800.txt.utf-8
--2020-08-18 10:54:47-- http://www.gutenberg.org/ebooks/800.txt.utf-8
Résolution de www.gutenberg.org (www.gutenberg.org)... 152.19.134.47, 2610:28:3090:3000:0:bad
Connexion à www.gutenberg.org (www.gutenberg.org)|152.19.134.47|:80... connecté.
requête HTTP transmise, en attente de la réponse... 302 Found
Emplacement : http://www.gutenberg.org/cache/epub/800/pg800.txt [suivant]
--2020-08-18 10:54:48-- http://www.gutenberg.org/cache/epub/800/pg800.txt
Réutilisation de la connexion existante à www.gutenberg.org:80.
requête HTTP transmise, en attente de la réponse... 200 OK
Taille : 461968 (451K) [text/plain]
Enregistre : «800.txt.utf-8»
```

```
800.txt.utf-8
```

```
100%[=====
```

```
2020-08-18 10:54:50 (253 KB/s) - «800.txt.utf-8» enregistré [461968/461968]
```

```
junier@fredportable:~/TP-SHELL/Phileas$ ls
800.txt.utf-8
```

4. Renommer le fichier en `tour-du-monde-80-jours.txt`.

```
junier@fredportable:~/TP-SHELL/Phileas$ mv 800.txt.utf-8 tour-du-monde-80-jours.txt
junier@fredportable:~/TP-SHELL/Phileas$ ls
tour-du-monde-80-jours.txt
```

5. Afficher le nombre de lignes, le nombre de mots, le nombre de caractères et le nombre d'octets de `tour-du-monde-80-jours.txt` avec des options bien choisies de la commande `wc`. Comment peut-on expliquer que le

nombre de caractères est inférieur au nombre d'octets ? Vérifier l'encodage du fichier avec la commande `file tour-du-monde-80-jours.txt`.

```
junier@fredportable:~/TP-SHELL/Phileas$ echo "Nombre de lignes " ; wc -l tour-du-monde-80-jours.txt
Nombre de lignes
9660 tour-du-monde-80-jours.txt
junier@fredportable:~/TP-SHELL/Phileas$ echo "Nombre de mots" ; wc -w tour-du-monde-80-jours.txt
Nombre de mots
70443 tour-du-monde-80-jours.txt
junier@fredportable:~/TP-SHELL/Phileas$ echo "Nombre de caractères" ; wc -c tour-du-monde-80-jours.txt
Nombre de caractères
461968 tour-du-monde-80-jours.txt
junier@fredportable:~/TP-SHELL/Phileas$ echo "Nombre d'octets" ; wc -m tour-du-monde-80-jours.txt
Nombre d'octets
450383 tour-du-monde-80-jours.txt
junier@fredportable:~/TP-SHELL/Phileas$ file tour-du-monde-80-jours.txt
tour-du-monde-80-jours.txt: UTF-8 Unicode (with BOM) text, with CRLF line terminators
```

6. Les commandes `du` et `zip` permettent respectivement d'afficher la taille d'un fichier et de compresser un fichier. Consulter leurs pages de manuel avec `man du` | `less` et `man zip` | `less`. La commande `less` est un *pager* qui permet d'afficher une page à la fois dans le terminal.
 - Afficher la taille du fichier en kilo-octets avec la commande `du -h tour-du-monde-80-jours.txt`.
 - Compresser la fichier avec la commande `zip`. Quel est le taux de compression ?
 - Avec la commande `head`, afficher les dix premières lignes des fichiers `tour-du-monde-80-jours.txt` et `tour-du-monde-80-jours.zip`. Que peut-on remarquer ?

```
junier@fredportable:~/TP-SHELL/Phileas$ zip --help
Copyright (c) 1990-2008 Info-ZIP - Type 'zip -L' for software license.
Zip 3.0 (July 5th 2008). Usage:
zip [-options] [-b path] [-t mmdyyy] [-n suffixes] [zipfile list] [-xi list]
The default action is to add or replace zipfile entries from list, which
can include the special name - to compress standard input.
If zipfile and list are omitted, zip compresses stdin to stdout.
-f freshen: only changed files      -u update: only changed or new files
-d delete entries in zipfile         -m move into zipfile (delete OS files)
-r recurse into directories          -j junk (don't record) directory names
-0 store only                        -l convert LF to CR LF (-ll CR LF to LF)
-1 compress faster                   -9 compress better
-q quiet operation                   -v verbose operation/print version info
-c add one-line comments             -z add zipfile comment
-@ read names from stdin              -o make zipfile as old as latest entry
-x exclude the following names       -i include only the following names
-F fix zipfile (-FF try harder)     -D do not add directory entries
```

```
-A  adjust self-extracting exe  -J  junk zipfile prefix (unzipsfx)
-T  test zipfile integrity      -X  eXclude eXtra file attributes
-y  store symbolic links as the link instead of the referenced file
-e  encrypt                     -n  don't compress these suffixes
-h2 show more help
```

```
junier@fredportable:~/TP-SHELL/Phileas$ man zip
junier@fredportable:~/TP-SHELL/Phileas$ zip tour-du-monde-80-jours.zip tour-du-monde-80-jours.txt
adding: tour-du-monde-80-jours.txt (deflated 63%)
```

```
junier@fredportable:~/TP-SHELL/Phileas$ ls -l
total 620
-rw-rw-r-- 1 junier junier 461968 août  1 10:43 tour-du-monde-80-jours.txt
-rw-rw-r-- 1 junier junier 171746 août 18 11:02 tour-du-monde-80-jours.zip
```

```
junier@fredportable:~/TP-SHELL/Phileas$ head -n5 tour-du-monde-80-jours.txt
The Project Gutenberg EBook of Le Tour du Monde en 80 Jours, by Jules Verne
```

This eBook is for the use of anyone anywhere at no cost and with almost no restrictions whatsoever. You may copy it, give it away or re-use it under the terms of the Project Gutenberg License included

```
junier@fredportable:~/TP-SHELL/Phileas$ head -n5 tour-du-monde-80-jours.zip #caractères illi
junier@fredportable:~/TP-SHELL/Phileas$ file tour-du-monde-80-jours.zip
tour-du-monde-80-jours.zip: Zip archive data, at least v2.0 to extract
```

7. Consulter la page de manuel de la commande `tac` avec `man | less tac`.
En une seule commande, créer un fichier `tour-du-monde-80-jours-inverse.txt` où toutes les lignes du fichier initial sont recopiées à l'envers.

```
junier@fredportable:~/TP-SHELL/Phileas$ tac tour-du-monde-80-jours.txt > tour-du-monde-80-jours-inverse.txt
```

8. Dans `tour-du-monde-80-jours.txt`, avec la commande `grep` et des options bien choisies :
 - Compter le nombre d'occurrences du mot `phileas`. On doit trouver 330.
 - Afficher le numéro de ligne du fragment de texte `*** START OF`. Vérifier avec un éditeur de textes.
 - Afficher le numéro de ligne du fragment de texte `*** END OF`. Vérifier avec un éditeur de textes.

```
junier@fredportable:~/TP-SHELL/Phileas$ grep -c -i "phileas" tour-du-monde-80-jours.txt
330
```

```
junier@fredportable:~/TP-SHELL/Phileas$ grep -n "*** START" tour-du-monde-80-jours.txt
19:*** START OF THIS PROJECT GUTENBERG EBOOK LE TOUR DU MONDE EN 80 JOURS ***
9332:*** START: FULL LICENSE ***
```

```
junier@fredportable:~/TP-SHELL/Phileas$ grep -n "*** START OF" tour-du-monde-80-jours.txt
19:*** START OF THIS PROJECT GUTENBERG EBOOK LE TOUR DU MONDE EN 80 JOURS ***
```

```
junier@fredportable:~/TP-SHELL/Phileas$ grep -n "*** END OF" tour-du-monde-80-jours.txt
9302:*** END OF THIS PROJECT GUTENBERG EBOOK LE TOUR DU MONDE EN 80 JOURS ***
```

- En une seule ligne de commandes, créer un fichier texte `tour-du-monde-80-jours-brut.txt` qui contient toutes les lignes comprises entre celles commençant par `*** START OF` et `*** END OF`, les deux bornes exclues.

```
junier@fredportable:~/TP-SHELL/Phileas$ tail -n+20 tour-du-monde-80-jours.txt | head -n-359
```