

Juan Ignacio Falcone

Sintaxis y Semántica de los Lenguajes 171.555-0 - <u>jfalcone@frba.utn.edu.ar</u> Github: <u>FalcFalcFalc</u> - <u>Repositorio del TP1</u>

TP 1 - Expresiones regulares extendidas en Bash

26 de junio de 2021

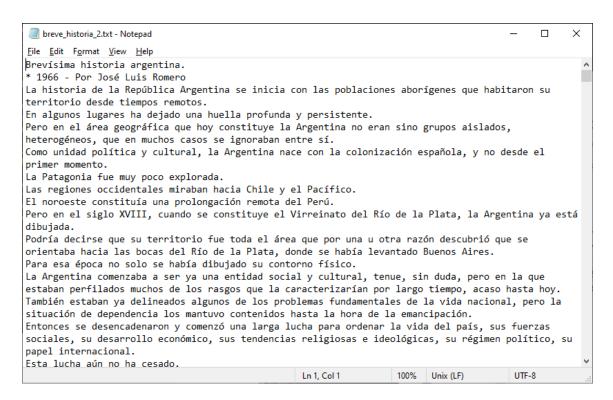
Objetivos

- 1) Crear un repositorio de git
- 2) Utilizar los comandos grep y sed en Bash, de Ubuntu, sobre una sección del libro de José Luis Romero, Breve Historia Argentina, para:
 - a) Reemplazar cada punto por un punto y salto de línea y generar un nuevo archivo
 - b) Borrar todas las líneas en blanco
 - c) Generar un nuevo archivo con las modificaciones aplicadas en A y B
 - d) Listar todas las oraciones que contengan la palabra 'guerra' sin distinguir entre mayúsculas o minúsculas
 - e) Mostrar líneas que comiencen con 'A' y terminen, tanto con 's' o con 's.'
 - f) Enumerar las oraciones en las que se nombre la palabra 'peronismo'
 - g) Enumerar las oraciones que contengan tanto 'Sarmiento' como 'Rosas'
 - h) Mostrar oraciones que contengan fechas del siglo XIX
 - i) Borrar la primera palabra de cada letra y guardar el resultado en un archivo
 - j) Enumerar todos los archivos de la extensión '.txt' en el directorio del ejecutable
- 3) Investigar y explicar los siguientes conceptos
 - a) Variable
 - b) Sentencias Condicionales

- c) Sentencias Cíclicas
- d) Subprocesos

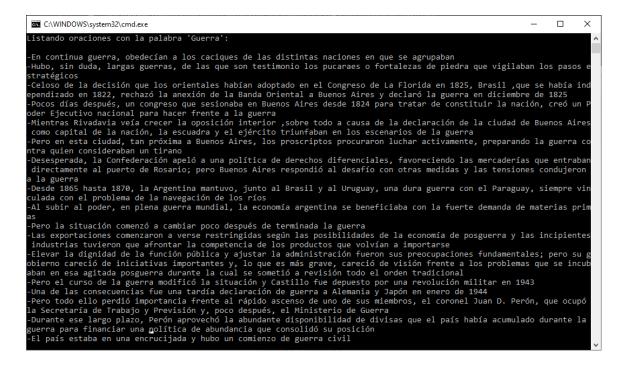
Punto 2

- A. En el ejecutable de bash, con respecto al punto A, se está ejecutando el comando sed para convertir todos los '.' en '.\n', donde se guarda el resultado en un archivo llamado breve_historia_1.5.
- B. Nuevamente utilizamos el comando sed con lo almacenado en breve_historia_1.5 para eliminar todas las líneas vacías, reconocidas como '^\$'.
- C. El resultado de B se guarda en un archivo nuevo, breve_historia_2. Nótese como no hay líneas vacías, ni puntos seguidos.



D. En el punto D, buscamos las oraciones que contengan la palabra guerra sin distinguir entre mayúsculas o minúsculas. Para reconocer estas oraciones, utilicé la siguiente expresión regular extendida:

Me gustaría explicarla ya que se utilizan variaciones de esta expresión dos veces más en el trabajo práctico: la expresión busca una letra mayúscula acompañada de cualquier cantidad de cualquier tipo de carácter exceptuando puntos, a menos que el punto sea precedido por alguna mayúscula (para poder incluir segundos nombres abreviados), seguidos de la sucesión de letras que conforme la palabra guerra, en cualquier combinación de mayúsculas o minúsculas, y luego cualquier cantidad de símbolos hasta encontrar un punto no precedido por una mayúscula. Nos dá como resultado la siguiente pantalla:



E. En el punto E se nos pide encontrar las líneas que comiencen con la letra A mayúscula y terminen con 's' o con 's.' No se encuentran coincidencias, y por eso nos devuelve vacío.

```
C:\WINDOWS\system32\cmd.exe — — X

Listando lineas que empiecen con 'A' y terminen con 's' o 's.':
```

F. Para el punto F, contamos la cantidad de ocurrencias de la palabra 'peronismo'. Son cinco las veces que aparece en el archivo, y son las siguientes:

```
C:\WINDOWS\system32\cmd.exe — — X

Cuantas veces aparece la palabra 'peronismo': 5

(...) que podía competir con el * PERONISMO *, se había dividido en (...)
(...) triunfado gracias al apoyo del * PERONISMO * creó a su alrededor (...)
(...) El * PERONISMO * parecía tener mayoría y (...)
(...) pues, o la proscripción del * PERONISMO * o su (...)
(...) del «Frente» que agrupaba al * PERONISMO *, a la Unión Cívica (...)
```

G. En el punto G, teníamos que escribir una expresión regular que encontrase oraciones que contengan tanto la palabra Sarmiento como la palabra Rosas, siendo esta la expresión a la que llegué:

```
"[A-Z]([^\.]|[A-Z]\.)*((Sarmiento([^\.]|[A-Z]\.)*?Rosas)|(Rosas([^\.]|[A-Z]\.)*?Sarmiento?))([^\.]|[A-Z]\.)*\."
```

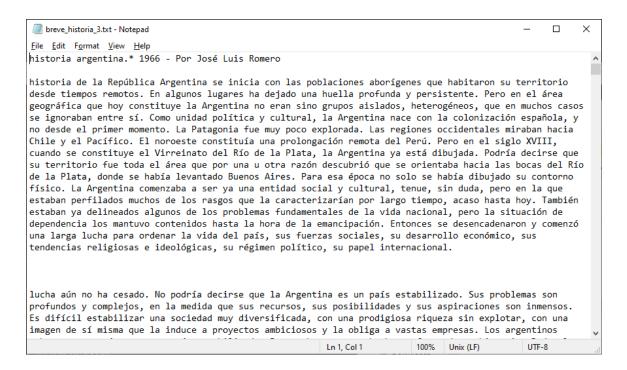
La expresión resulta parecida a la utilizada en el punto D debido a que nos piden también oraciones acá, solo que en vez de buscar en el medio la palabra 'Guerra', busca. La coincidencia encontrada fue una sola y fue esta:

 H. Para el punto H debíamos hallar oraciones que contengan fechas del siglo XIX. Para eso utilicé la expresión regular utilizada

```
"[A-Z]+([^\.]|[A-Z]\.)*+18[0-9]{2}+([^\.]|[A-Z]\.)*\."
```

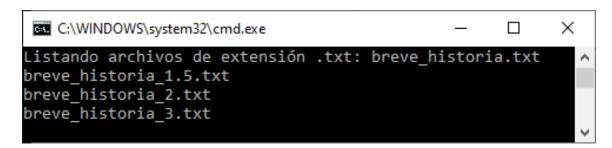
Una vez más, la expresión es parecida a la del punto G. En este caso, busca algún número de 4 cifras cuyas primeras dos sean 1 y 8 respectivamente.

I. En el punto I se pedía eliminar la primera palabra de cada oración y guardar el resultado en un nuevo archivo. Este fue el resultado:



J. El punto J exigía un listado de archivos de tipo '.txt'. Para realizarlo, se nos aconsejó usar un pipeline con el comando ls. El comando resultó siendo así:

Is | grep "[^\.]*\.txt"



Punto 3

Variables

Una variable es un **espacio de memoria** al que podemos referenciar y llamar para **almacenar**, **modificar y leer datos** durante la ejecución del programa.

Para establecer una variable, uno debe escribir el nombre de la variable, el símbolo igual, '=' y el valor que se le desea asignar. Para acceder a la variable, uno debe colocar el símbolo de dinero, '\$', a la izquierda del nombre de la variable.

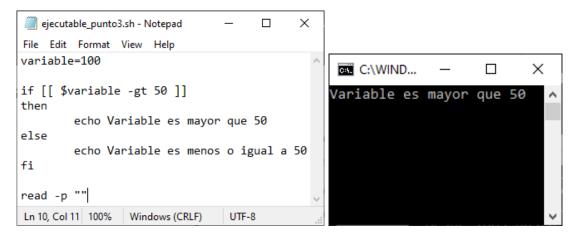
```
utnso@Juani-Notebook:/mnt/c/Users/juani$ valor=3
utnso@Juani-Notebook:/mnt/c/Users/juani$ echo $valor

utnso@Juani-Notebook:/mnt/c/Users/juani$ valor=valor+valor
utnso@Juani-Notebook:/mnt/c/Users/juani$ echo $valor
valor+valor
utnso@Juani-Notebook:/mnt/c/Users/juani$ valor=3
utnso@Juani-Notebook:/mnt/c/Users/juani$ valor=$valor+$valor
utnso@Juani-Notebook:/mnt/c/Users/juani$ echo $valor
```

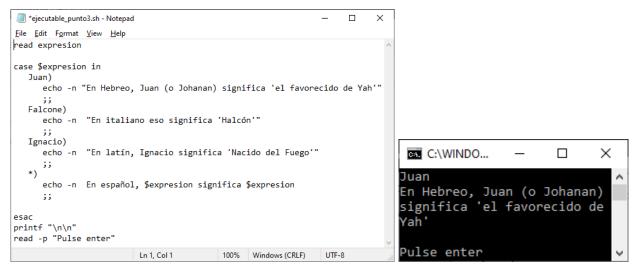
Sentencias Condicionales

Las sentencias condicionales se pueden dividir generalmente en dos: las sentencias IF y las Case.

La sentencia **if** permite al programa tomar decisiones binarias. Lo mínimo que exige para funcionar es una **condición** y un **set de instrucciones** a seguir en caso de que la condición sea verdadera. Adicionalmente, se le puede agregar un set de instrucciones para el caso en el que la condición es falsa.



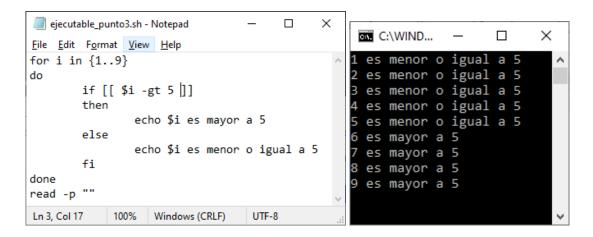
Cuando se usa **case** (también llamadas **switch**), el programa puede tomar muchos caminos distintos. Exige una **expresión** a juzgar, y **sets de instrucciones para cada caso**. Se puede incluir también un caso default, especificado por un asterisco (*) para ejecutar para todos los casos que no se hayan especificado.



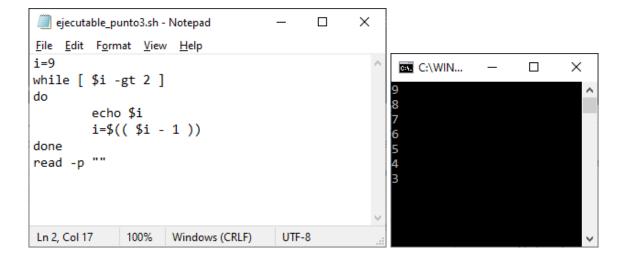
Sentencia Cíclica

Para establecer una sentencia cíclica, se pueden usar tanto la sentencia **For** como la **While**, que funcionan de distintas maneras:

For hace un **recorrido**, toma elementos de una lista y **pasa uno por uno** hasta terminar.



While por otro lado, **repite el set de instrucciones** que se le dé, siempre y cuando la **condición sea verdadera**. Una vez que la **condición es falsa**, no entra nuevamente a la repetición y **termina**. Idealmente, en el set de instrucciones debería haber una parte en la que los parámetros de la condición cambian, porque de lo contrario, la sentencia podría generar una repetición infinita.



Subprogramas

Los subprogramas son secuencias de instrucciones que pueden ser llamados en cualquier momento de la ejecución. Al ser invocadas, realizan sus instrucciones, y al finalizar, continúan con la línea siguiente en donde la invocaron.

```
*ejecutable_punto3.sh - Notepad
                                                                             X
File Edit Format View Help
preguntar(){
  printf "\nIngrese un numero para acumular, o ingrese -1 para salir\n\n"
acumulado=0
preguntar
read i
while [[ $i -ne -1 ]]
    acumulado=$(( $acumulado + $i))
    preguntar
    read i
done
printf "La sumatoria de los IDs es $acumulado\n"
read -p "Enter para Salir"
                         Ln 15, Col 1
                                           100% Windows (CRLF)
                                                                 UTF-8
```

```
Ingrese un numero para acumular, o ingrese -1 para salir

Ingrese un numero para acumular, o ingrese -1 para salir

Ingrese un numero para acumular, o ingrese -1 para salir

Ingrese un numero para acumular, o ingrese -1 para salir

Ingrese un numero para acumular, o ingrese -1 para salir

Ingrese un numero para acumular, o ingrese -1 para salir

Ingrese un numero para acumular, o ingrese -1 para salir
```