

Paradigmas de la Programación

Laboratorio 6: Scripting

Laura Alonso Alemany

Cristian Cardellino

Ezequiel Orbe

En este laboratorio desarrollaremos un proyecto utilizando un lenguaje de scripting, para lo cual implementaremos el corrector ortográfico (*spellchecker*) que hicimos en los Laboratorios 2, 3 y 5 en uno de los lenguajes de scripting más conocidos: **Bash**.

Características de la presentación

- El trabajo es en grupo: máximo: **dos integrantes**, mínimo: **dos integrantes**.
- Fecha de Entrega: Hasta las 23:59:59 del 12/06/2015.
- Formato de entrega:
 1. Empaquetar el directorio `spellchecker` en un archivo llamado `<dni>-lab-6.tar.gz`, donde `<dni>` es el DNI de alguno de los integrantes del grupo.
 2. Enviar el archivo `<dni>-lab-6.tar.gz` por mail a la dirección:

`paradigmas@famaf.unc.edu.ar`

El título del mail debe decir: `lab-6`. En el cuerpo del mail se deben indicar los integrantes (nombre completo y DNI).

3. Los trabajos enviados fuera de término serán evaluados de acuerdo a la siguiente Política de Entrega Tardía:
 - Hasta 1 día después del deadline: -20% de la nota.
 - Hasta 2 días después del deadline: -40% de la nota.
 - Hasta 3 días después del deadline: -60% de la nota.
 - Hasta 4 días después del deadline: -80% de la nota.
 - Más de 4 días después del deadline: el trabajo se considera NO entregado y se debe recuperar.
4. Los trabajos enviados que no cumplan con las condiciones de presentación no serán considerados.

Software Requerido

Para realizar este laboratorio necesitaremos el siguiente software:

- Bash (viene instalado por defecto en Linux)

Código Inicial

Para realizar este laboratorio, utilizaremos el código disponible en:

`http://cs.famaf.unc.edu.ar/materias/paradigmas/node/204`

Al extraer el contenido del archivo, tenemos la siguiente estructura de directorios:

```
spellchecker
└─ spellchecker.sh
```

1 Especificaciones del Corrector Ortográfico

El corrector ortográfico (spellchecker) que desarrollaremos deberá utilizar un diccionario de palabras conocidas, el cual se cargará en memoria desde un archivo cuando el programa se inicie, y un diccionario de palabras ignoradas, el cual inicialmente estará vacío.

Dado un documento de entrada, el spellchecker copiará a un documento de salida todas las palabras y los signos de puntuación, consultando interactivamente al usuario sobre cada palabra desconocida.

Dada una palabra desconocida, el usuario podrá aceptarla, ignorarla o reemplazarla.

Si el usuario acepta la palabra, la misma se agregará al diccionario, si el usuario ignora la palabra, esta ocurrencia y las subsiguientes se ignorarán. Finalmente, si el usuario decide reemplazar la palabra, el sistema permitirá que el usuario ingrese una nueva palabra que reemplazará a la palabra desconocida en el documento de salida.

Una vez que todo el documento de entrada haya sido procesado, el programa guardará el diccionario en el mismo archivo desde el cual lo cargó al inicio y dejará disponible un nuevo archivo que contenga el documento de salida.

1.1 Funcionalidades deseadas

1. Leer un diccionario de palabras conocidas desde un archivo.
2. Tener en memoria un diccionario de palabras ignoradas.
3. Leer un documento de entrada, palabra por palabra.
4. No proponer acciones para palabras conocidas.
5. Proponer tres acciones posibles para palabras desconocidas: **aceptar**, **ignorar**, **reemplazar**.
6. En caso de aceptar una palabra desconocida: incorporarla al diccionario de palabras conocidas y que las subsiguientes ocurrencias de la palabra sean tratadas como conocidas.
7. En caso de ignorar una palabra desconocida: incorporarla al diccionario de palabras ignoradas e ignorar la palabra y sus subsiguientes ocurrencias.
8. En caso de reemplazar una palabra desconocida: en el documento de salida se escribirá la palabra por la que se reemplaza, en lugar de la original.
9. Actualizar el diccionario de palabras conocidas incluyendo las palabras aceptadas.
10. Escribir un documento de salida con las palabras originales o sus modificaciones, según las acciones realizadas por el usuario.

2 Consignas

- Implementar en **Bash** el spellchecker.
- La implementación no puede utilizar los spellcheckers desarrollados en los laboratorios anteriores.
- Se pueden utilizar todas las herramientas disponibles en un terminal de Linux (grep, read, touch, echo, expresiones regulares, etc). Algunas herramientas muy útiles para procesamiento de documentos en **Bash** línea por línea son **sed**, **awk** y **perl** (con las opciones **-n** o **-p** además de **-e**).
- Se penalizará la duplicación de código. Si ya existe un programa que me permite realizar una tarea, lo utilizo, no codifico de nuevo.
- Es fundamental leer la documentación de Bash.