

# Desarrollo de un primer programa en Python: índice por palabras para ficheros de texto plano

Factor de ponderación [0-10]: 10

## 0.1. Objetivo

El objetivo principal de esta práctica es aprender Python e introducirse en el uso de algunas de las funciones y métodos que tendrán que manejarse durante las prácticas de Compiladores.

Un buen manual de referencia (incluyendo una introducción informal) de Python se encuentra en [8].

[9] es un manual similar, pero más académico.

En [10] se encuentra disponible un libro sobre Python.

La referencia fundamental es el site oficial del lenguaje, [7].

## 0.2. Descripción

La práctica consiste en escribir en Python un programa que lea de un fichero de texto plano las palabras que contiene. Estas palabras serán introducidas en memoria junto al número de línea en la que aparece cada palabra. Posteriormente se volcará esta estructura (en un fichero y/o por pantalla) ordenada por palabras, de tal forma que tengamos un índice, como el que se muestra a modo de ejemplo en el Listado 1.

According	10, 11, 48
Acknowledgement	19, 25
Acknowledgment	17, 13, 27

Listado 1: Ejemplo de índice generado por el programa

## 0.3. Requisitos de codificación

Cada fichero de código fuente de los que se utilicen en la práctica ha de llevar comentarios de cabecera. También es imprescindible que cada función lleve comentarios de cabecera con una breve indicación de la operación que implementa.

## 0.4. Detalles de implementación

El indexador debe leer un fichero de entrada en formato de texto plano. Debe descartar los espacios y signos de puntuación (puntos, comas, etc). El resultado se volcará a un fichero también de texto plano. Tanto el fichero de entrada como el de salida serán suministrados en línea de comando al ejecutar el programa.

Un buen ejemplo de entrada/salida en Python se encuentra en [3].

Es imprescindible utilizar sentencias `try` y `except` para el manejo de ficheros (al menos).

Debe también aprender a utilizar los potentes bucles de que dispone Python de modo que pueda iterar sobre las líneas y las palabras del fichero (y sobre las listas). En [1] puede verse un ejemplo de cómo hacerlo.

Al recoger las palabras del fichero debe procesárselas para eliminar todos los caracteres de puntuación no deseados. Para ello pueden usarse las funciones de cadenas de Python. Véase un ejemplo en [4].

Pueden emplearse las estructuras de datos que consideren más conveniente, pero se recomienda usar listas. Véase [5] a modo de ejemplo.

Pero sobre todo, diccionarios [2].

Una referencia sobre ordenación en Python puede estudiarse en [6].

Para aprobar la práctica basta con que la salida ordenada se escriba en un fichero de texto plano.

La práctica puede completarse de forma opcional (para una mejor nota) si se vuelca a un fichero en formato HTML como la tabla del ejemplo (si es Ud. muy purista puede hacerlo con capas y hojas de estilo), añadiendo un enlace a un buscador web para que localice esa palabra.

Pueden experimentar si desean aprender más con `os.system` que permite invocar a un navegador. Se podría usar una sentencia similar a:

```
os.system('konqueror ' + Path + '/' + Salida + '.html').
```

# Bibliografía

- [1] Bucles en Python, [http://www.mclibre.org/consultar/python/lecciones/le05\\_for.html](http://www.mclibre.org/consultar/python/lecciones/le05_for.html)
- [2] Diccionarios en Python, [http://es.diveintopython.org/odbchelper\\_dict.html](http://es.diveintopython.org/odbchelper_dict.html)
- [3] Entrada/Salida en Python, <http://mundogeek.net/archivos/2008/04/02/python-entrada-salida-ficheros/>
- [4] Funciones de cadenas en Python, <http://luauf.com/2008/08/24/strings-functions-en-python/>
- [5] Listas en Python, [http://es.diveintopython.org/odbchelper\\_list.html](http://es.diveintopython.org/odbchelper_list.html)
- [6] Orden en Python, <http://wiki.python.org/moin/HowTo/Sorting/>
- [7] Python Official Site, <http://www.python.org/doc/>
- [8] Referencia 1, <http://glud.udistrital.edu.co/fileadmin/Archivo/Documentacion/Traducciones/python-tut/html/tut/tut.html>
- [9] Referencia 2, <http://www.mclibre.org/consultar/python/>
- [10] Referencia 3, <http://www.scribd.com/doc/2898872/Python-para-todos>