

Sistemas Operativos II

Práctica 2

Daniel Rivero
Dennis Quitaquís

Introducción

El objetivo de esta práctica es, basándonos en el código implementado en la práctica 1, añadir diversas funcionalidades para trabajar mejor con el árbol de palabras indexadas y utilizar la información del árbol para hacer un par de análisis mediante el programa gnuplot y tuberías.

Para ello desarrollaremos una interfaz sencilla que permite al usuario:

- Crear el árbol binario balanceado de palabras a partir de un fichero con la lista de archivos de los cuales extraer las palabras.
- Guardar el árbol en un fichero.
- Cargar el árbol desde un fichero.
- Analizar las palabras del árbol de diferentes maneras: para esto usaremos el programa gnuplot, que permite visualizar gráficas a partir de unos datos, y una pipe para enviarle la información al programa.

Memoria de la práctica

El programa que hemos implementado funciona, a grandes rasgos, de la siguiente manera:

Al iniciarse, imprime el menú de usuario por pantalla para elegir la opción deseada:

- Opción 1 - Creación del árbol: Esta opción crea el árbol usando el mismo proceso que se usaba en la práctica 1, es decir, rellenando una tabla hash por cada fichero con las palabras extraídas del fichero, y, al finalizar cada fichero, pasando la información de la tabla hash al árbol binario balanceado. Con este proceso, al finalizar la ejecución tenemos un árbol con todas las palabras de todos los ficheros.
- Opción 2 - Guardar árbol: Almacenamos la información de los nodos del árbol en un archivo mediante una función recursiva que recorre el árbol y utilizando la función fwrite para ir escribiendo en el fichero los datos de cada nodo.
- Opción 3 - Cargar árbol: Primero, si hay algún árbol cargado en el sistema lo eliminamos. A continuación, cargamos la información de cada nodo almacenado en el fichero indicado y vamos insertando los nodos en un árbol nuevo palabra por palabra.

- Opción 4 - Análisis del árbol: Carga un sub-menú en el que elegir que tipo de análisis queremos hacer.
 - Análisis de la probabilidad de que aparezca una palabra en cada uno de los ficheros: El usuario introduce la palabra deseada y la aplicación la busca en el árbol y, si está, calcula el porcentaje de aparición en cada archivo, lo ordena de mayor a menor y escribe los datos en un fichero. A continuación, mediante un pipe, envío al programa gnuplot el archivo con los datos y los parámetros necesarios para que muestre por pantalla un gráfico con esos datos.
 - Análisis de la probabilidad de que aparezca una palabras en los ficheros según su longitud: La aplicación analiza el árbol y cuenta cuantas palabras de cada longitud hay en los archivos, con lo cual calcula la probabilidad y lo escribe en un fichero. A continuación, mediante un pipe, envío al programa gnuplot el archivo con los datos y los parámetros necesarios para que muestre por pantalla un gráfico con esos datos.
- Opción 5 - Salir: Se libera la memoria del árbol y la aplicación finaliza su ejecución.

Para comprobar el correcto funcionamiento del programa hemos ejecutado la aplicación con todas sus opciones utilizando el árbol creado a partir de los 10 primeros ficheros de la base de datos grande (llista.cfg) y ejecuta todas las opciones correctamente.

En la carpeta “Proves” están los siguientes archivos:

- arbre_10fitxers.data: contiene el árbol almacenado de la ejecución con los 10 primeros archivos.
- log.txt: contiene el log de la ejecución, en él se puede ver toda la impresión por pantalla que hace el programa al ir haciendo las diferentes opciones.
- valgrind.txt: contiene el log de la ejecución hecha con valgrind. Se puede ver que, una vez ejecutadas todas las opciones y finalizada la aplicación, toda la memoria esta liberada y no hay ningún error.

Además de implementar las funcionalidades indicadas en esta práctica, también hemos solucionado los problemas de memoria de la práctica 1: los principales problemas por los cuales teníamos memoria sin liberar consistían en que no hacíamos correctamente el free de las listas de la tabla hash y de las keys de los nodos del árbol.