

UNIDAD TEMÁTICA 2 – ARBOLES GENÉRICOS, TRIES y ARBOLES B

Prácticos Domiciliarios Individuales - 4

ESCENARIO (el indicado en el TA2)

Dado el programa resultante del **Trabajo de Aplicación 2**, se requiere verificar el correcto funcionamiento de las siguientes operaciones:

- Insertar una nueva palabra en este diccionario (se hará para algunas palabras de los textos contenidos)
- Buscar las ocurrencias de una cierta palabra.
- Imprimir el índice construido, indicando, para cada palabra del índice, la página en que se encuentra.

EJERCICIO 1

Desarrolla las siguientes funcionalidades y programas, incluyendo casos de prueba para verificar el correcto funcionamiento:

1. Agregar, para cada palabra en el índice (trie), una estructura de datos (seleccionar una de las ya tratadas en el curso – listas, árboles, etc. -) que almacene las páginas del texto en que esta palabra se encuentra, y modificar el código existente del trie – si resulta necesario - para adecuarse a la nueva estructura del trie (nodo).
2. Crear un método llamado **“indizarLibro”**, que, tomando como entrada un libro (archivo de texto), lo recorra y apunte en el trie las páginas (del libro) en que se encuentra cada palabra del índice. A tales efectos, cada página se ha de considerar como el equivalente de **50** líneas.
3. Crear un método **“imprimirIndice”**, que listará por consola, en orden lexicográfico, cada palabra contenida en el índice, y a continuación las páginas en que esta palabra se encuentra (separadas por espacios).
4. Archivos de prueba: “palabrasindiceprueba.txt” y “libro prueba.txt”. Armar el trie con las palabras contenidas en el archivo “palabrasindiceprueba” y luego, recorriendo el archivo “libro prueba”, indizar las ocurrencias.

Una vez probadas las funcionalidades, ejecutar el programa para indizar un libro completo utilizando:

1. archivo de entrada **“PalabrasIndice.txt”**, con el cual se ha de construir el trie
2. archivo de entrada **“libro.txt”** a usar para indizar el libro usando el trie ya cargado.
3. Se deberá ejecutar el programa incluyendo **“imprimirIndice”**

NOTA IMPORTANTE: SE DEBEN RETIRAR **TODOS LOS CARACTERES ESPECIALES** – QUE NO SEAN LETRAS- DE LAS PALABRAS DEL LIBRO, A EFECTOS DE ASEGURAR LA CORRECTA EJECUCIÓN. TAMBIÉN SE DEBE CONVERTIR TODO EL TEXTO A MINÚSCULAS.

EJERCICIO 2

Modificar el método “**buscar**” para, dada una palabra como argumento, devolver:

- Si existe o no en el índice
- Cantidad de comparaciones realizadas
- **Palabra y Páginas en que se encuentra la palabra (separados por espacios)**

NOTA IMPORTANTE: SE DEBEN RETIRAR TODOS LOS CARACTERES ESPECIALES – QUE NO SEAN LETRAS- DE LAS PALABRAS DEL LIBRO, A EFECTOS DE ASEGURAR LA CORRECTA EJECUCIÓN. TAMBIÉN SE DEBE CONVERTIR TODO EL TEXTO A MINÚSCULAS.

RESPONDER LAS PREGUNTAS QUE SE MOSTRARÁN EN PANTALLA, TOMANDO COMO ENTRADA LOS ARCHIVOS INDICADOS POR LA CATEDRA

EJERCICIO 3

El propósito de este ejercicio es revisar lo aprendido y desarrollado, analizando las estructuras y algoritmos en forma comparativa con otras posibles soluciones para el problema planteado.

Los temas a analizar son:

1. Dado el escenario planteado, ¿puede resolverse mediante las estructuras de datos vistas anteriormente (en las Unidades Temáticas anteriores – Listas, Colas, Arboles)? (se debe discutir la factibilidad de uso de cada uno de los TDA ya vistos en el curso, y describir en forma abstracta cómo se habría de implementar). Para cada uno de estos TDA, ¿CUÁLES son sus ventajas y desventajas comparativas con respecto al TRIE implementado?
2. ¿Cuál sería la recomendación del Equipo si el (los) archivos a indizar fueran pequeños (por ejemplo, unas decenas de líneas)

GUARDAR ESTAS NOTAS!!!!

RESPONDER LAS PREGUNTAS QUE SE MOSTRARÁN EN PANTALLA, TOMANDO COMO ENTRADA LOS ARCHIVOS INDICADOS POR LA CATEDRA