

Tarea 5

(Fecha de entrega: hasta el Martes 2018-08-07 a las 23:30)

Huffman Code Compression

En esta tarea Ud. debe implementar un programa que genere la codificación de Huffman para un archivo de texto cualquiera. Su programa debe recibir dos parámetros para la ejecución:

- un nombre de archivo de entrada, que contendrá el texto sobre el cual se generará la estadística necesaria para generar la codificación;
- un nombre de un archivo de salida donde se muestre la codificación para los caracteres que contiene el archivo. El formato de este archivo es que en cada línea se indica la codificación de un carácter en orden de frecuencia descendente.

	Carácter	Cód ASCII (dec.)	codificación	frecuencia
Por ejemplo:	'e'	99	0101	12.5%
	' '	32	1011	12.2%
	'a'	61	1001	10.2%
	't'	116	11001	7.3%
	'\n'	10	00110	6.9%

para ir formando el árbol de codificación, en cada iteración se deben escoger los dos conjuntos árboles de menor frecuencia y fusionarse. Su solución debe ser tal que cada iteración tome tiempo del orden de $O(\log_2 n)$.

Como experimento se pide completar la siguiente tabla:

Nombre del archivo	Largo del archivo original	Largo del archivo comprimido
ElCidC1.txt		
Hamlet.txt		
Urfaust.txt		

Junto con el informe se deben entregar los archivos de salida con las codificaciones para cada uno de los archivos de entrada.

- los archivos de entrada están en <http://www.dcc.uchile.cl/nbaloian/cc3001/Tareas/Tarea5>

Condiciones de Entrega

- Esta tarea debe ser resuelta en Java.
- Es obligatoria la entrega de un informe junto con su tarea (Ver siguiente sección).
- Esta tarea es de carácter individual, cualquier caso de copia se evaluará con la nota minima.
- No olvide subir a U-cursos todos los archivos necesarios para que su tarea funcione correctamente.
- Debe subir los archivos de código fuente (*.java). Los archivos compilados (*.class) no serán evaluados.
- Cualquier duda respecto a la tarea puede ser consultada usando el foro del curso. No se aceptarán atrasos.

Informe

El informe debe describir el trabajo realizado, la solución implementada, los resultados obtenidos y las conclusiones o interpretaciones de estos. Principalmente debe ser breve, describiendo cada uno de los puntos que a continuación se indican:

- Portada: Indicando número de la tarea, fecha, autor, email, código del curso, etc.
- Introducción: Descripción breve del problema y su solución.
- Análisis del problema: Exponga en detalle el problema, los supuestos que pretende ocupar, casos de borde y brevemente la metodología usada para resolverlo.
- Solución del problema: Indique claramente los pasos que siguió para llegar a la solución del problema. Muestre mediante figuras y ejemplos qué es lo que realiza su código. Evite copiar todo el código fuente en el informe, sin embargo, puede mostrar las partes relevantes de éste.
- Modo de uso: Explicando brevemente cualquier dato necesario para la compilación y ejecución de su programa.
- Resultados: Entregue las salidas que se pide de sus programas., incluyendo gráficos y análisis de los resultados.