

Laboratorio # 3

Compresión de mensajes y archivos

Objetivo

Comprender la codificación Huffman y aritmética (con enteros).

Descripción

Implementar una herramienta que permita comprimir y descomprimir mensajes con los dos métodos y explicación del proceso.

Entradas y salidas

- El programa solicitará al usuario la opción de comprimir o descomprimir un mensaje, el método a utilizar y si será modo automático o no.
- Compresión automática:
 - o Si selecciona Huffman deberá solicitar el mensaje y dar como salida el mensaje codificado, la tabla de frecuencias, el código de cada caracter y la tasa de compresión.
 - o Si selecciona Aritmética deberá solicitar el mensaje y dar como salida el mensaje codificado, la tabla de frecuencias, los valores calculados de k , T y n y la tasa de compresión.
 - o Para ambos casos debe mostrar el resultado en consola y en un archivo de salida llamado `codificacion#.log` (el # indica un número que solicitará al usuario) indicando el método, modo y salidas solicitadas.
- Compresión no automática:
 - o Si selecciona Huffman deberá solicitar el mensaje y la tabla de frecuencias, si algún caracter no está en la tabla mostrará un mensaje de error, la salida será igual a la indicada en el modo automático.
 - o Si selecciona Aritmética deberá solicitar el mensaje, la tabla de frecuencias y k , si algún caracter no está en la tabla o el valor de k no cumple con la condición de ser al menos $4T$, debe mostrar un mensaje de error, la salida será igual a la indicada en el modo automático.
 - o Al igual que en el modo automático, debe mostrar los resultados en consola y en un archivo llamado `codificacion#.log`.
- Descompresión:
 - o Para ambos casos solicitará el nombre de un archivo, este será alguno de los que generó en la codificación (`codificacion#.log`).
 - o De este obtendrá los datos necesarios para decodificar el mensaje.
 - o Debe mostrar el mensaje decodificado en consola.
- Explicación del proceso:
 - o Para cada operación deberá generar un segundo archivo de nombre `codificacionProceso#.log` o `decodificacionProceso#.log` con la explicación paso a

paso del proceso de codificación o decodificación, debe incluir una descripción general y las fórmulas a utilizar para los cálculos y/o criterios de ordenamiento, etc.

- Codificación Huffman:
 - Mostrar tabla de frecuencias a utilizar
 - Mostrar cada paso de cómo se seleccionan los caracteres con menor frecuencia y se va generando el código binario de cada uno
 - Ir indicando en cada paso los caracteres que se seleccionaron y el reordenamiento que se haya realizado
- Decodificación Huffman:
 - Mostrar tabla de frecuencias
 - Mostrar árbol binario de codificación
 - Por cada carácter indicar cómo se recorrió el árbol para llegar a decodificar.
- Codificación Aritmética:
 - Mostrar tabla de frecuencias y valores de k , R , T , n .
 - Puede utilizar una tabla, cada fila sería una iteración o una regla de expansión, recuerde indicar R , $R/2$, $R/4$ y $3R/4$ así como todos los cálculos de l , u , s y m .
- Decodificación Aritmética:
 - Mostrar tabla de frecuencias y valores de k , R , T , n .
 - Puede utilizar una tabla similar a la codificación.

Criterios de Evaluación

1. Codificación Huffman: 20 puntos
2. Decodificación Huffman: 20 puntos
3. Codificación Aritmética: 20 puntos
4. Decodificación Aritmética: 20 puntos
5. Ingreso de datos por el usuario y verificación de errores: 10 puntos
6. Documentación: 10 puntos

Los criterios 1 al 4 incluyen el mensaje correctamente codificado/decodificado, salida a consola, salida a archivo (para la codificación) y salida a archivo de la explicación del proceso.

Documentación

1. Código fuente bien documentado
2. Archivos de prueba
3. Manual del usuario: este debe incluir carátula, descripción de la aplicación y conclusiones

Plazo

La entrega debe realizarse en un plazo máximo de dos semanas a partir de la fecha de asignación.