

## Universidad de Guadalajara

# CENTRO UNIVERSITARIO DE CIENCIAS EXACTAS E INGENIERÍAS

Actividad de la asignatura.: Seminario de Solución de Problemas

de Estructuras de Datos II

Maestro: Janette Araceli Castellanos Barajas

Alumno: Miguel Vladimir Fierros Mora

**Grado:** Tercer semestre

Horario: Lunes y Miércoles de 11:00 AM- 12:55 PM

**Ciclo:** 2022 B

Fecha: 02/11/2022

# COMPRESION Y DESCOMPRESIÓN CON ARCHIVOS BINARIOS Y DE TEXTO

#### **INDICACIONES:**

El estudiante deberá entregar el código portable que permita comprimir y descomprimir la información en un archivo mediante serialización con TDA, para la clase ya trabajada anteriormente de "CLIENTES", con las operaciones Agregar, Imprimir, Buscar, Modificar, Eliminar, Indexar y Ordenar alfabéticamente.

## "LA COMPRESIÓN"

## ¿Qué es la compresión de archivos?

La compresión es el método por el cual por medio de ciertos algoritmos en un programa se consigue reducir el espacio de memoria que tiene que ocupar un archivo. El objetivo principal de comprimir un archivo es volverlo más pequeño para que ocupe menos espacio en las áreas donde se vaya a ocupar. Al hacerlo no afectas el contenido o la estructura del archivo, simplemente reduces el espacio que ocupa. El proceso inverso de la compresión es la descompresión.

Estos procesos se suelen realizar cuando:

- Se necesita ahorrar espacio: en este caso, primero comprime los archivos más grandes y que no utilizas muy a menudo.
- Se necesita compartir varios archivos para compartirlos.

## Ejemplos de los métodos y tipos de compresión que se usan:

Los archivos ZIP: Este método se trata de un formato de compresión que no tiene pérdida y es bastante simple al igual que comprime los archivos de forma separada. ZIP ofrece un proceso de compresión bastante rápido y con un ratio de compresión aceptable, si bien este no destaca por ser de los más altos y en ese sentido es claramente inferior a otros formatos de compresión como RAR o 7z. Al igual que otros, el formato ZIP es compatible con un sistema de cifrado simétrico de clave publica, lo que lo convierte en un sistema débil ante ataques de fuerza bruta o diccionario.

#### Los archivos GZIP:

El formato de compresión Gzip es un sistema de compresión adoptado por el proyecto GNU y que es muy ampliamente utilizado en plataformas Unix y Linux. Gzip utiliza el algoritmo de compresión Deflate, una combinación de Lempel-Zib (LZ77) y la codificación del algoritmo de Huffman. Se trata de un algoritmo de compresión que no posee pérdida y muy extendido no solo en zip y gzip, sino también en el formato de imágenes PNG. Por defecto, en GNU/Linux comprime los archivos con la extensión .gz. El ratio de compresión de Gzip es notablemente más alto que el del formato zip, siendo igualmente rápido durante el proceso de compresión. Es útil para comprimir archivos de forma rápida, con una ratio de compresión aceptable pero no de los más altos.

#### Los archivos RAR:

El formato de compresión RAR es un sistema de compresión propietario que se ha vuelto muy popular al ser el formato que provee el software de compresión WinRAR, que se volvió ampliamente utilizado en entornos Windows. Este formato reúne las ventajas de ZIP, pero ofreciendo un proceso de compresión bastante más optimizado además de ser compatible con opciones avanzadas como el uso de la compresión sólida, un método de compresión especial mediante el cual todos los archivos son tratados como si fuera un flujo continuo de datos. Entre otras características están también la compatibilidad con el cifrado AES de 128 bits, etc.







## **ACTIVIDAD Y RESULTADOS**

## Explicación de la labor a realizar:

Este programa tenía el propósito de hacer posible el manejo de datos con el uso de archivos de tipo binarios y de texto (ARCHIVOS.bin, ARCHIVOS.txt) para practicar y lograr la compactación de archivos con el uso del algoritmo de huffman, por lo que se buscaba conseguir la capacidad de hacer uso de estos con el fin de poder almacenar datos de una manera funcional y práctica; de tal forma que, aún si se cierra el programa una o varias veces por parte del uso del usuario, los datos podrían ser salvados o reconstruidos en algún otro dispositivo, sistema o base de datos, así de esta forma se podría acceder en cualquier momento y lugar a la información que se haya registrado, de la misma forma que se pueda lograr crear un archivo binario con uso de serialización desde 0 y de índices simples ordenados. Este programa igualmente permitiría la edición de los archivos en el momento, o sea, durante la ejecución se podrían seguir manipulando los archivos.

Este programa formará parte de la base que conformará el proyecto final, a cerca del manejo de datos referentes a la venta de pinturas en cierta sucursal. En este caso el programa busca emplear el uso de la serialización y compresión con la clase clientes, que ya se había utilizado antes, lo cual puede ayudarnos a organizar la información que se registre de algunos campos para tener una mayor facilidad de comprensión con los mismos. Continuando con esto, se busca que el programa pudiera realizar las funciones de "CAPTURA", "IMPRESIÓN DE DATOS", "BUSCAR", "ELIMINAR", "MODIFICAR", "INDEXAR", "ORDENAR" y "DESCOMPRIMIR", esta última refiriendo a abrir el archivo compactado durante el proceso. Esto será con los datos de la clase "FACTURAS", estos cambios se podrán reflejar en la consola, así como en los archivos de binarios y de texto.

#### Menú del programa:

🔳 D:\DESCARGAS 2\CËDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDII.exe

```
REGISTRO DE CLIENTES

- RELLENA EL SIGUIENTE FORMULARIO PARA REGISTRAR LOS DATOS DE LOS REGISTROS DENTRO DEL ARCHIVO DE CLIENTES.

* EL PROGRAMA TE MOSTRARÁ UN MENÚ PARA QUE REALICES DISTINTAS ACCIONES CON LOS DATOS DE LOS CLIENTES QUE YA HAYAS REGISTRADO.

1.- CAPTURAR
2.- MOSTRAR
3.- BUSCAR
4.- MODIFICAR
5.- EL ININAR
6.- GENERAR INDICES CON LAS CLAVES
7.- ORDENAR INDICE
8.- LEER ARCHIVO COMPACTADO
9.- SALIR

SELECCIONA LA OPCION (1-9):
```

## La función capturar:

Esta función permitirá que el usuario pueda crear un archivo binario desde cero al momento de iniciar el programa, al mismo tiempo que permitiría editar uno que ya esté existente con el mismo nombre en el dispositivo, pues el software añade los datos que se registren consecutivamente al final de lo que se haya registrado de un uso anterior del mismo programa, esto pasa cuando se vuelve a hacer uso el programa, esto también se verá reflejado tanto en consola como en el archivo de binario. Esta misma función de manera interna permitirá que existan 3 archivos, uno denominado "CLIENTES2.bin", otro como "CLIENTES2.txt" y otro con un índice llamado "INDICE\_CLIENTES.bin", así como uno denominado como "COMPACTADO\_HUFFMAN.bin", estos poseen un nombre similar al de los anteriores códigos que desarrollamos ya que en algún punto todos estos coexistirán y formarán un solo conjunto dando forma al proyecto final. Estos mismos darán la capacidad al usuario de ver los datos reflejados tanto en consola como en el archivo de registro final, esto de forma limpia y ordenada.

```
DADESCARGAS 2/CÉDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN_FIERROS_MORA_MIGUEL_VLADIMIR_SSPEDILexe

CAPTURAR DATOS DE LOS CLIENTES

PERSONA 5

ESCRIBA EL FOLIO DEL CLIENTE:
DVFFDVDFV

ESCRIBE EL NOMBRE DE LA PERSONA:
LUIS MANUEL AVILA

¿CUÁL ES EL RED DEL CLIENTE?
DCSCDOSCOSCOSC

¿CUÁL ES LA TARIFA A PAGAR?
12,380

¿CUÁL ES LA TARIFA A PAGAR?
12,380

¿CUÁL ES EL TELEFONO DEL CLIENTE?
3324214322

HORA: 3:34:5 FECHA:16/11/2022

¿CUÁL ES EL DOMICILIO DEL ESTABLECIMENTO?
JALISCO, TLAQUEPAQUE, CP 45620, COL PLAN DE ORIENTE, CALLE FLORES #75
```

## Después de su uso:

🔳 D:\DESCARGAS 2\CËDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDI

## La función mostrar:

Esta función fue diseñada con el fin principal de mostrar en tiempo real los datos que estén guardados dentro del archivo llamado "CLIENTES2.bin", el cual es de tipo binario, de tal forma que si hay registros de datos en el archivo los imprimirá de una forma ordenada en consola conforme a la llegada de estos al registro del programa, esto para que el usuario los visualice de una forma práctica y concisa. De la misma forma, esta función ayudará a que el usuario se dé cuenta de si existen o no datos en el documento y también lo hará si es que no existe dicho archivo.

## Si existe un registro:

## Si no existe el registro:



## La función buscar:

Esta función le permite al programa poder mostrar el registro de una de los registros de la clase "CLIENTES" al momento en que el usuario lo solicita en consola durante la ejecución del programa, esto lo realizará siempre y cuando el registro exista en la misma carpeta del programa, de lo contrario le notificará al usuario de que dicho elemento no existe en el archivo o que el archivo ni siquiera ha sido creado.

#### Resultados de su funcionamiento:

#### Si existe

#### Si no existe el dato:

D:\DESCARGAS 2\CEDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDII.exe

BUSQUEDA DE REGISTROS

DIGITA EL FOLIO DEL CLIENTE:
fcdfcs

DATO NO ENCONTRADO

#### Si no existe el dato:

D:\DESCARGAS 2\CËDIGOS YA TERMINADOS\ARCHIV(

NO EXISTE UN REGISTRO DE DATOS

Presione una tecla para continuar . . .

## La función editar:

Esta función se implementó con el propósito de darle al usuario la capacidad de editar los datos de alguno de los registros que estén dentro del archivo de tipo binario. Esta acción se verá reflejada cuando se muestren datos en consola y en el archivo binario, así como de sus derivados, esto siempre y cuando se encuentren los datos que se soliciten por el usuario, de lo contrario el mismo programa notificará al usuario de que no existe el registro o que el archivo ni siquiera existe.

#### Resultados de su funcionamiento:

## **Durante el proceso**

## Si existe el registro:

## Si no existe el registro:

D:\DESCARGAS 2\CEDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDII.exe

EDICIÓN DE REGISTROS

DAME EL FOLIO DEL CLIENTE: cscscdscsdcsdc

DATO NO ENCONTRADO

Presione una tecla para continuar . . .

## Si no existe el registro completo:

D:\DESCARGAS 2\CËDIGOS YA TERMINADOS\ARCHIV(

NO EXISTE UN REGISTRO DE DATOS

Presione una tecla para continuar . . .

## Después de su uso:

■ D:\DESCARGAS 2\CËDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDII.exe

## La función eliminar:

Esta función fue diseñada para que el usuario pueda tener la capacidad de eliminar algún registro que esté existente en el archivo, esto lo borrara definitivamente de la visualización de consola, así como del archivo de texto que haya generado el software. Esto se hará siempre y cuando exista el registro, de lo contrario este notificará al usuario de que este no existe.

#### Resultados de su funcionamiento:

## Durante el proceso de uso

🔳 D:\DESCARGAS 2\CËDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDII.exe

#### Si no existe el dato:

■ D:\DESCARGAS 2\CEDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDII.exe

```
DAME EL FOLIO DEL CLIENTE: dsdsdcsddsc

DATO NO ENCONTRADO

Presione una tecla para continuar . . .
```

## Si no existe el registro:

🔳 D:\DESCARGAS 2\CËDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDII.exe

```
DAME EL FOLIO DEL CLIENTE: dsdsdcsddsc

DATO NO ENCONTRADO

Presione una tecla para continuar . . .
```

#### Después de su uso:

🔟 D:\DESCARGAS 2\CËDIGOS YA TERMINADOS\ARCHIVOS DE EDII\HUFFMAN\HUFFMAN\_FIERROS\_MORA\_MIGUEL\_VLADIMIR\_SSPEDII.exe

```
REGISTRO DE CLIENTES
    FOLIO-----: VDVDFVD
    PRECIO A PAGAR------ DFVDS

DOMICILIO DE LA SUCURSAL----- SDCSFVCSV
     HORA Y FECHA DEL REGISTRO-----: HORA: 1:34:21 FECHA: 16/11/2022
                                   ---: SFVSVCFS
    NOMBRE DEL CLIENTE------: FERNANDO MONTES DE OCA
RFC DEL CLIENTE------ DCDSCDSCS
    TELEFOND DEL CLIENTE------: 3324637411

PRECIO A PAGAR-----: 13232

DOMICILIO DE LA SUCURSAL----: JALISCO, ALLENDE, CP 3432, COL LOMAS VERDES, CALLE LEOPARDO #2432

HORA Y FECHA DEL REGISTRO----: HORA: 3:52:23 FECHA: 16/11/2022
                                      DVDFVDFV
    NOMBRE DEL CLIENTE----:
                                      CSVV
DSCDS
    FOI TO-----
                  -----: DVFFDVDFV
    NOMBRE DEL CLIENTE----
                                      LUIS MANUEL AVILA
    Presione una tecla para continuar . . .
```

## La función indexar:

Esta función se creó derivado de la práctica de índices simples, esto lo implemente para que el usuario pudiese tener la capacidad de indexar cada uno de los registros existentes, en este caso, los folios de los registros, ya que me resulta practico que el usuario pueda tener acceso a esta forma de visualizar datos. Este proceso permitirá mostrar un índice de estos, pues a los registros se les coloca un número de acuerdo con el número de registros que estén existentes y al orden en que llegan. Esto generará un archivo de tipo binario llamado "INDICE\_CLIENTES", por lo que este es el archivo que existirá y tendrá dentro los registros de las claves en el orden en que fueron ingresados en los archivos.

#### Resultados de su funcionamiento:



#### La función ordenar:

Está función se hizo con el propósito de hacer que el índice del archivo binario con nombre "INDICE\_CLIENTES" ya generado anteriormente se pueda ordenar de manera alfabética en consola y en un nuevo archivo de tipo binario con el nombre "INDICE BINARIO3". En este caso mantendrá su número de registro y estarán en orden para una mayor facilidad de búsqueda.

#### Resultados de su funcionamiento:



## **CONCLUSIÓN**

El desarrollo y funcionamiento de este programa me ha permitido tener la oportunidad de estudiar y comprender el uso y el proceso del manejo junto a la creación de los archivos de tipo binario que son dinámicos en la programación, al mismo tiempo que me permitió comprender el uso y manejo de la serialización junto a la compactación, esto cuando se hace uso del lenguaje de "C++", así como cuando se hace uso de los índices simples con serialización incluida. Este mismo programa me ayudo a darme cuenta de la gran importancia de todo este proceso al momento de desarrollar un programa que incluye bases de datos y de hacer uso administrado de distintos tipos de datos así como de registros cuando se trata de buscar datos, ya que me pude dar cuenta de que estos permiten que un sinfín de programas tengan un funcionamiento estable y por la misma razón los programas, aún si son cerrados, puedan seguir y guardar un proceso con distintos tipos de funciones ya prediseñadas. Pese a que son un tanto complicados de entender, creo que logré captar la idea del funcionamiento de estos y sobre todo aprendí a recrearlos por mí mismo. Haciendo énfasis en el proceso de compresión y descompresión, creo yo que es de lo más complejo que he llegado a manejar, pues me llevo bastante tiempo de estudio, pero sin duda reconozco que fue un reto importante, ya que me ha dado el tiempo para aprender conceptos que considero que me servirán a futuro. Siguiendo con lo ya mencionado, también fui capaz de entender el uso de los accesos a los archivos tanto de texto como binarios, en este caso, considero que pude lograr el propósito de organizar de forma presentable y estable la información con el uso de los índices simples y de los delimitadores, por lo tanto, puedo concluir con la realización de este programa que se enfocaba en la compresión y descomprensión de archivos.