

CURSO 2022-2023
CICLO SUPERIOR DE DESARROLLO DE APLICACIONES WEB
IES AGUADULCE

Bases de Datos

Francisco Javier Sueza Rodríguez

2 de noviembre de 2022

Índice general

1 Almacenamiento de la Información	4
1.1 Introducción	4
1.2 Los Ficheros de Información	5
1.2.1 ¿Que es un Fichero?	5
1.2.2 Tipos de Ficheros	5
1.2.3 Los Soportes de la Información	6
1.2.4 Métodos de Acceso a Ficheros	7
1.2.4.1 Ficheros Secuenciales	7
1.2.4.2 Ficheros de Acceso Directo	8
1.2.4.3 Ficheros Indexados	9
1.2.4.4 Otros Tipos de Organización	10
Glosario	11
Bibliografía	12

Índice de figuras

1.2.1	Clasificación de ficheros según su función	6
1.2.2	Tipos de ficheros por método de acceso	7
1.2.3	Estructura de un fichero secuencial	7
1.2.4	Modos de acceso a un registro en archivos de acceso directo	8
1.2.5	Ficheros indexados	9

Tema 1

Almacenamiento de la Información

En este primer tema, vamos a estudiar los conceptos básicos sobre el almacenamiento de la información, así como de las bases de datos y los SGBD (Sistemas de Gestión de Bases de Datos), pero en primer lugar, vamos a hacer una introducción más detallada sobre que consideramos información y el contenido de este módulo.

1.1 Introducción

Si pensamos cualquier en cualquier aspecto de nuestra vida cotidiana, o si analizamos la mayoría de ámbitos de actividad, nos encontramos que la utilización de bases de datos esta ampliamente extendida. Estás, y los datos contenidos en ellas, serán imprescindibles para llevar a cabo multitud de acciones.

Algunas de las situaciones en las que es necesario el uso de bases de datos son las siguientes:

- Cuando seleccionamos un canal de la TDT.
- Al utilizar la agenda del móvil para realizar una llamada telefónica.
- Cuando utilizamos un cajero automático.
- Cuando acudimos a la consulta del médico.
- Al inscribirnos en un curso, plataforma online, etc...
- Si utilizas el GPS.
- Cuando reservamos unas localidades en un evento deportivo.
- Cuando consultamos cualquier información en internet.
- Al solicitar un certificado de un organismo oficial.

Como vemos, el gran volumen de datos que manejamos y sus innumerables posibilidades hacen necesaria la existencia de técnicos perfectamente formados y capaces de trabajar con ellos.

Este módulo profesional se centra, precisamente, en las **Bases de Datos** y su uso en el desarrollo de aplicaciones. En esta primera unidad, comenzaremos conociendo los primeros sistema basados en ficheros para el almacenamiento y gestión de la información. Seguidamente, se desarrollarán los conceptos y definiciones básicas relacionados con las bases de datos, viendo también sus modelos y tipos. Más adelante conocer los sistemas gestores de bases de datos y finalmente, veremos las herramientas reales con las que llevar a caso dicha gestión.

1.2 Los Ficheros de Información

En esta sección vamos a hablar de los ficheros de información, en que consiste, que tipos nos podemos encontrar, métodos de acceso y parámetros de utilización.

1.2.1 ¿Que es un Fichero?

En la década de los setenta, los procesos básicos relacionados con una empresa se centraban en la contabilidad y facturación. Las necesidades de almacenamiento y gestión de la información podían satisfacerse con un número relativamente reducido de archivos de papel agrupados y ordenados, los típicos ficheros clásicos.

Con la primera informatización, se paso del papel al ordenador, pudiendo acceder a los datos de forma mucho más rápida. Los ordenadores adaptaron sus herramientas para que se asemejaran a las que los usuarios utilizaban manualmente, de forma que en informática también empezó a hablarse de ficheros, carpetas, formularios, etc...

La información que empezó a tratarse en los ordenadores debía ser almacenada para su posterior recuperación, consulta y procesamiento. El elemento que se creo para almacenar esta información fue el **fichero** o **archivo**.

Podemos definir un **fichero** como el **conjunto de información relacionada**, tratada como un todo y organizada de **forma estructurada**. Es una secuencia de dígitos binarios que organiza información relacionada con el mismo aspecto.

Los ficheros están formados por **registros lógicos** que contienen información relativa a un mismo elemento u objeto (por ejemplo, información de un usuario). A su vez, los registros están divididos en **campos** que tienen cada una de las informaciones elementales que forman un registro (por ejemplo, nombre de usuario, email,...).

Los datos están almacenados de forma que se pueda añadir, suprimir, actualizar y consultar, individualmente, en cualquier momento.

Como los ficheros suelen ser muy grandes, solo se puede llevar parte de ellos a la memoria principal para procesarlos. La cantidad de información que es transferida entre el soporte en el que se almacena el fichero y la memoria del ordenador, en solo una operación de lectura/escritura, se llama **registro físico** o **bloque**.

Normalmente en cada operación de lectura/escritura se transfieren varios registros de un fichero, es decir, un bloque suele contener varios registros lógicos. Al número de registros que entran en un bloque se le llama **factor de blocaje**, y a la operación de agrupar varios registros en un mismo bloque se conoce como **bloque de registros**.

1.2.2 Tipos de Ficheros

Según la función que vaya a desempeñar un fichero, estos pueden ser clasificados de varias maneras:

- (a) **Ficheros Permanentes:** contiene información relevante para una aplicación. Es decir, los datos necesarios para el funcionamiento de ésta. Tiene un período de permanencia en el sistema amplio. Se subdividen en:
 - **Ficheros Maestros:** contiene el estado actual de los datos que pueden modificarse desde la aplicación. Es la parte central de aplicación, su núcleo.

- **Ficheros Constantes:** son aquellos que incluyen datos fijos de la aplicación. No suelen ser modificados y se accede a ellos para la realización de consultas.
 - **Ficheros Históricos:** contiene datos que fueron considerados como actuales en un período o situación anterior. Se utilizan para la reconstrucción de situaciones o estados concretos.
- (b) **Ficheros Temporales:** se utilizan para almacenar datos que son útiles para una parte de la aplicación. Son generados a partir de datos de ficheros permanentes y tienen un período corto de existencia. Estos pueden ser:
- **Ficheros Intermedios:** almacenan resultados de una aplicación que serán usados por otra.
 - **Ficheros de Maniobras:** almacenan datos de una aplicación que no pueden ser mantenidos en memoria por falta de espacio.
 - **Ficheros de Resultados:** almacenan datos que van a ser transferidos a un dispositivo de salida.

En la siguiente figura podemos ver un esquema con esta clasificación.

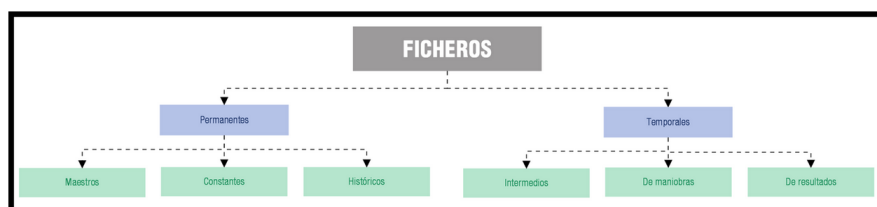


Figura 1.2.1: Clasificación de ficheros según su función

1.2.3 Los Soportes de la Información

Los ficheros se almacenan en soportes de información manejados por periféricos del ordenador, que permiten leer y grabar datos en el soporte. Los soportes más utilizados son las **cintas magnéticas** y los **discos** (magnéticos, ópticos o magneto-ópticos).

Al principio se usaban tambores de cinta magnética, similares en tamaño a un disco de vinilo, funcionaban de manera similar a los antiguos casetes, pero al tener un tamaño mucho más grande permitían almacenar mucha más información, permitiendo el acceso a esta de forma secuencial.

Posteriormente los medios de almacenamiento fueron evolucionando a la par que el hardware, en concreto con la aparición de los disquetes y el disco duro. Estos dispositivos ya permitían el acceso aleatorio a los datos.

Por lo tanto, podemos distinguir dos tipos de dispositivos de almacenamiento de datos:

- **Soportes de Acceso Directo a Datos:** son los más empleados y el acceso a datos se hace de forma directa, pudiendo colocarlos en la posición que más nos interese.
- **Soporte de Acceso Secuencial:** se suele usar en copias de seguridad y si deseamos leer un dato que está a mitad de la cinta, tendremos que leer todo lo que hay hasta llegar a esa posición.

Si quieres aprender más sobre las características de cintas y discos, puedes consultar los enlaces siguientes:

- [Cintas magnéticas de almacenamiento de datos](#)
- [Discos magnéticos y Discos ópticos](#)

1.2.4 Métodos de Acceso a Ficheros

A medida que la tecnología ha ido evolucionando, el acceso a la información ha ido variando mucho. Los objetivos fundamentales de estas variaciones son los siguientes:

- Proporcionar acceso rápido a los registros.
- Conseguir economizar el almacenamiento.
- Facilitar la actualización de los registros.
- Permitir que la estructura refleje la organización real de la información.

Los ficheros se pueden clasificar según como se organiza en la memoria principal, o dicho de otra forma, los métodos de acceso al fichero, que podemos ver en la siguiente figura.

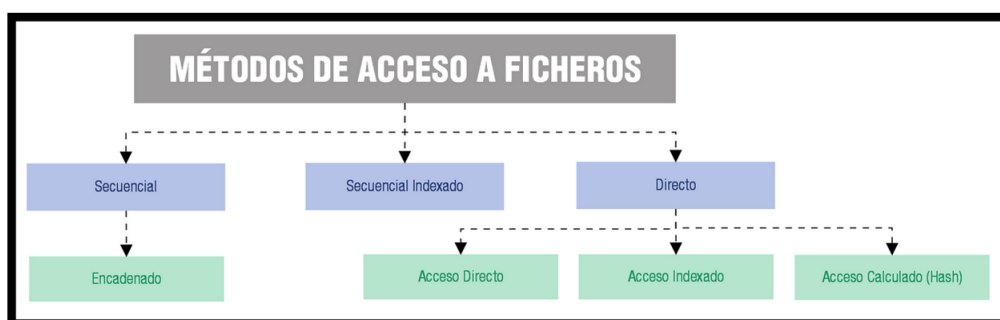


Figura 1.2.2: Tipos de ficheros por método de acceso

Las organizaciones secuencial, de acceso aleatorio o directo e indexado, son las más comunes. En esta sección vamos a explicar las características de cada uno de los métodos de acceso a ficheros.

1.2.4.1 Ficheros Secuenciales

Un **fichero secuencial** se caracteriza porque sus registros están almacenados de forma continua, de forma que la única manera de acceder a él, es leyendo un registro detrás de otro hasta el final. En los ficheros secuenciales hay una marca que indica el final del fichero, suele denominarse **EOF** (End Of File). Así, para detectar el final de fichero solo es necesario encontrar esta marca.

Este tipo de fichero puede usar dispositivos de almacenamiento de acceso secuencial, como cintas magnéticas, aunque también se utilizan en los CD de audio y DVD de vídeo, en los que la música y las imágenes se almacenan en un espiral continua.

Los registros almacenados se identifican por medio de la información ubicada en uno de sus campos, que se denomina **clave** o **llave**. Si se ordena un archivo secuencial por su clave es más rápido realizar operaciones sobre él.



Figura 1.2.3: Estructura de un fichero secuencial

Algunas características de este tipo de ficheros son las siguientes:

- La **lectura** siempre se realiza **hacia adelante**.
- Son ficheros **monousuario**, no permiten el acceso simultáneo de varios usuarios.
- Tiene una **estructura rígida de campos**. Todos los registros deben aparecer en orden, es decir, la posición de los campos en el registros siempre debe ser la misma.
- El **modo de apertura** del fichero, condiciona la lectura o escritura.
- **Aprovechan al máximo** el soporte de **almacenamiento**, no dejando huecos vacíos.
- Se pueden **grabar** en **cualquier tipo** de **soporte**, tanto secuenciales como direccionales.
- Todos los **lenguajes de programación** contiene instrucciones para **trabajar** con este tipo de ficheros.
- No se pueden **insertar registros** en los que están **ya grabados**.

1.2.4.2 Ficheros de Acceso Directo

En este tipo de archivos se puede acceder a un registro indicando la posición relativa del registros dentro del archivo, o a través de una **clave** que forma parte del registro como un **campo** más. Estos archivos deben almacenarse en dispositivos de memoria masiva con acceso directo como los discos magnéticos.

Cada uno de los registros se guarda en una posición física, que dependerá del espacio disponible en memoria masiva, por lo que la distribución es aleatoria dentro del soporte de almacenamiento. Para acceder a la posición física del registros se utiliza una posición o índice, de forma que no es necesario recorrer todo el fichero para encontrar un determinado registros.

Esta **dirección física** se obtendrá tras la aplicación de una **transformación** específica a la **clave**. Según como sea esta transformación, existen tres modos de acceso diferente, como podemos ver en la siguiente figura.

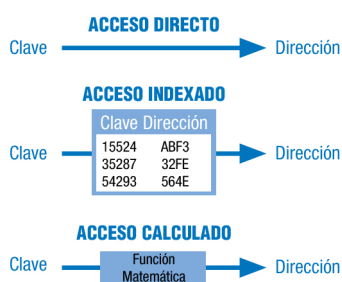


Figura 1.2.4: Modos de acceso a un registro en archivos de acceso directo

El método más rápido es el **acceso directo**, donde la clave coincide con la dirección física del registro, teniendo ésta que ser una posición válida dentro del rango de direcciones físicas.

La **medida de posicionamiento** básico del puntero en el fichero es el **byte**, dependiendo del tipo de codificación de caracteres que empleemos (**Unicode**, **ANSI**), se usarán 1 o 2 bytes por carácter respectivamente. Teniendo esto en cuenta, el puntero avanza de 1 en 1 o de 2 en 2 bytes para poder leer o escribir cada carácter.

Otras **características** de este tipo de ficheros son las siguientes:

- **Posicionamiento inmediato.**
- **Registros de longitud fija.**
- **Apertura** del fichero en **modo mixto**, para lectura y escritura.
- Permiten **múltiples usuarios** al mismo tiempo.
- Los **registros se borran** colocando un cero en la posición que ocupan.
- Permiten la utilización de **algoritmos de compactación** de huecos.
- Los archivos se **crean** con un **tamaño definido**, es decir, con un máximo de registros definidos durante su creación.
- Esta organización solo es posible en **soportes direccionales**.
- Se **usan** cuando el **acceso a datos** de un registro se hace siempre empleando la **misma clave** y la **velocidad de acceso** al registro es lo que más importa.
- Permiten la **actualización de registros** en el mismo fichero, sin necesidad de copiarlo.
- Permiten realizar **procesos de actualización en tiempo real**.

1.2.4.3 Ficheros Indexados

Se basan en el uso de **índices**, que permiten el acceso a un registros sin tener leer el fichero entero. Estos índices son similares a los de los libros, si nos interesa leer un capítulo podemos recurrir al índice donde se nos dice en que página comienza y acaba dicho capítulo.

Por lo tanto, deberá existir una **zona de registros** en los que se encuentren los datos del archivo y una **zona de índices**, que contiene la tabla con las claves de los registros y las posiciones donde se encuentran. La tabla de índices esta ordenada por campos clave.

La tabla de índices será cargada en la memoria principal para realizar en ella la búsqueda de la fila correspondiente a la clave del registros a encontrar, proporcionando así la dirección donde se encuentra el registro. Una vez localizada la dirección, solo es necesario acceder el dispositivo de almacenamiento y colocarlos en la dirección indicada.

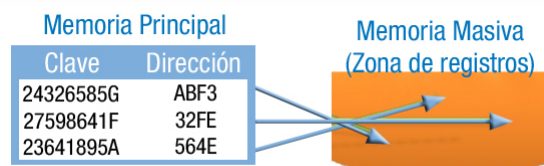


Figura 1.2.5: Ficheros indexados

Las características mas relevantes de los ficheros indexados son las siguientes:

- El diseño de registros tiene que tener un campo o campos, que permita identificar cada registro de forma única, es decir, no puede haber 2 registros que tengan la misma información en él. A este campo se le llama **campo clave** y es el que va a servir de índice. Un mismo fichero puede tener varios campos clave, pero al menos uno de ellos **no puede tener valores duplicados** y se llama **clave principal**. A las restantes se les llama **claves alternativas**.

- Permite usar el modo de **acceso secuencial** y el modo de **acceso directo** para leer la información que guardan sus registros.
- Para acceder a estos ficheros usando el modo de **acceso directo**, se hace conociendo el **contenido del campo clave** que queremos localizar. Con esta información el sistema operativo puede consultar el índice y conocer la posición dentro del fichero.
- Para acceder a este tipo de ficheros usando el **modo secuencial**, los registros son **leídos** por el contenido del **campo clave**, independientemente del orden en el que fueron grabados, ya que el acceso se hace a través del índice, que para hacer más fácil la búsqueda de registros, permanece siempre ordenado por campos clave.
- Solamente puede **almacenarse** en un **medio direccionable**, ya que sino no podría usar el modo de acceso directo.

1.2.4.4 Otros Tipos de Organización

Además de los tipos de organización de ficheros que ya hemos visto, existen otros como los **ficheros secuenciales indexados** o los **ficheros de acceso calculado**, los cuales pasamos a describir a continuación.

(a) Ficheros Secuenciales Indexados:

También llamados parcialmente indexados, al igual que los ficheros indexados existe una **zona de índices** y otra **zona de registros de datos**, pero esta última se encuentra **dividida** en **segmentos** ordenados.

En la zona de índices, cada fila hace referencia a cada uno de los segmento. La clave corresponde al último registros del segmento y el índice al registro inicial. Una vez que se accede al primer registro del segmento, dentro de él se localiza (de forma secuencial) el registro buscado.

Las principales características de este tipo de ficheros son:

- Permite el **acceso secuencial**. Esto es muy interesante cuando la tasa de accesos es alta. En el acceso secuencial además los registros se leen ordenados por el campo clave.
- Permite el **acceso directo a registros**. Realmente **emula** este tipo de acceso, empleando para ello las tablas de índices. Primero busca la clave en el área de índices y luego va a leer al área de datos en la dirección que indica la tabla.
- Se pueden **actualizar** los **registros** en el **mismo fichero**, sin necesidad de crear uno nuevo de copia en el proceso de actualización.

Glosario

ANSI El Instituto Nacional Estadounidense de Estándares, más conocido como ANSI (por sus siglas en inglés: American National Standards Institute), es una organización sin fines de lucro que supervisa el desarrollo de estándares para productos, servicios, procesos y sistemas en los Estados Unidos.. 8

Unicode El Estándar Unicode es un estándar de codificación de caracteres diseñado para facilitar el tratamiento informático, transmisión y visualización de textos de múltiples lenguajes y disciplinas técnicas además de textos clásicos de lenguas muertas.. 8

Bibliografía