



Ciclo 1

Semana 7

Archivos de Texto

Lectura 2 – Organización de Archivos



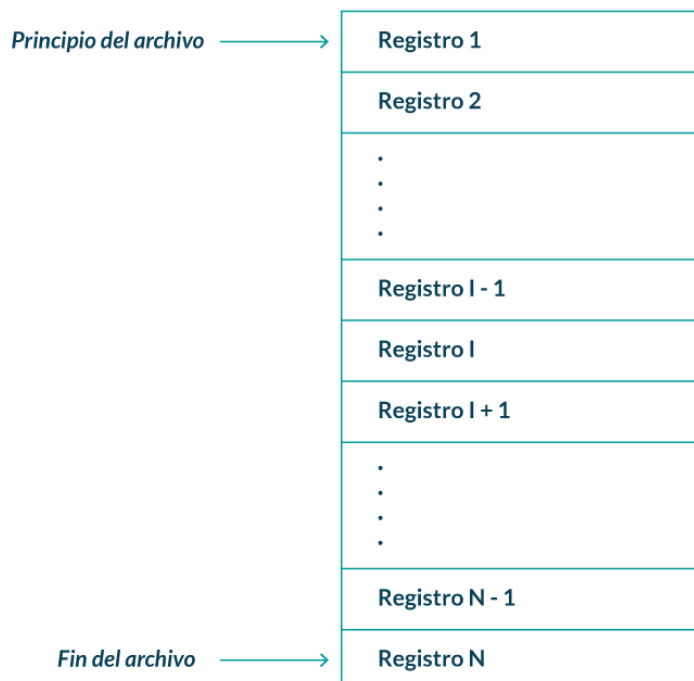
| Organización de Archivos

Según las características del soporte empleado y el modo que se han organizado los registros, se consideran dos tipos de acceso a los registros de un archivo:

- **Acceso secuencial:** según el orden de almacenamiento de sus registros, uno tras otro.
- **Acceso directo:** implica el acceso a un registro determinado, sin que implique la consulta de los registros precedentes. Es posible solo con soportes direccionables.

La organización de un archivo define la forma en que los registros se disponen sobre el soporte de almacenamiento, o también la forma en que se estructuran los datos en un archivo. En general se tienen tres organizaciones fundamentales:

Organización secuencial: es una sucesión de registros almacenados consecutivamente sobre el soporte externo, de tal modo que para acceder a un registro n es obligatorio pasar por los $n-1$ registros que le preceden. Los registros se graban consecutivamente cuando el archivo se crea y se debe acceder consecutivamente cuando se leen dichos registros.





El orden físico en que fueron grabados los registros es el orden de lectura de los mismos. Todos los tipos de dispositivos de memoria auxiliar soportan esta organización.

Organización directa: ocurre cuando el orden físico no corresponde con el orden lógico. Los datos se ubican en el archivo y se accede a ellos directamente mediante su posición, es decir, el lugar relativo que ocupan. Tiene la ventaja que los registros se pueden leer y escribir en cualquier orden y posición, y su acceso es muy rápido. Tiene el inconveniente de que necesita programar la relación existente entre el contenido de un registro y la posición que ocupa. Por su funcionamiento pueden ocurrir huecos libres entre registros. La correspondencia entre clave y dirección debe poder ser programada y la determinación de la relación entre el registro y su posición física se obtiene con una fórmula. Para que un archivo sea de organización directa:

- Debe ser almacenado en un soporte direccionable.
- Los registros deben tener un campo clave que identifique cada registro de forma única.
- Debe haber una correspondencia entre los posibles valores de la clave y las direcciones disponibles sobre el soporte.

Organización secuencial indexada: Un diccionario es un archivo secuencial, cuyos registros son las entradas y cuyas claves son las palabras definidas por las entradas. Para buscar una palabra (una clave) no se busca secuencialmente de la 'a' a la 'z', sino que se abre el diccionario por la letra inicial de la palabra. Por ejemplo, si se desea buscar 'índice', se abre por la letra I y en su primera página se busca la cabecera de página hasta encontrar la página más próxima a la palabra, buscando a continuación palabra a palabra hasta encontrar 'índice'. El diccionario es un ejemplo típico de archivo secuencial indexado con dos niveles de índices, el nivel superior para las letras iniciales y el nivel menor para las cabeceras de página. En un computador las letras y las cabeceras de páginas se guardarán en un archivo de índices independiente de las entradas del diccionario, por consiguiente, un archivo secuencial indexado consta de un archivo índice y un archivo de datos. Un archivo está organizado en esta forma si:

- El tipo de sus registros contiene un campo clave identificador.
- Los registros están situados en un soporte direccionable por el orden de los valores indicados por la clave.
- Un índice para cada posición direccionable, la dirección de la posición y el valor de la clave; en esencia, el índice contiene la clave del último registro y la dirección de acceso al primer registro del bloque.



Un archivo en organización secuencial indexada consta de las siguientes partes:

- Área de datos o primaria: contiene los registros en forma secuencial y está organizada en secuencia de claves sin dejar huecos intercalados.
- Área de índices: es una tabla que contiene los niveles de índice, la existencia de varios índices enlazados se denomina nivel de indexación.
- Área de desbordamiento o excedentes: utilizada, si fuese necesario, para las actualizaciones.

El área de índices es equivalente, en su función, al índice de un libro. En ella se refleja el valor de la clave identificativa más alta de cada grupo de registros del archivo y la dirección de almacenamiento del grupo. Estos archivos secuenciales indexados presentan las siguientes ventajas:

- Acceso rápido.
- El sistema de gestión de archivos se encarga de relacionar la posición de cada registro con su contenido mediante la tabla de índices.

Y presentan las siguientes desventajas:

- Desaprovechamiento de espacio por dejar huecos intermedios cada vez que se actualiza el archivo.
- Se necesita espacio adicional para el área de índices.

Los soportes que se utilizan para esta organización son los que permiten el acceso directo, discos magnéticos.