

# Introducción a las Bases de Datos

## Fundamentos de Organización de Datos

### Práctica 1

Creación, consulta y mantenimiento de archivos secuenciales - Algorítmica Básica.

1. Realizar un algoritmo que cree un archivo de números enteros no ordenados y permita incorporar datos al archivo. Los números son ingresados desde teclado. La carga finaliza cuando se ingresa el número 30000, que no debe incorporarse al archivo. El nombre del archivo debe ser proporcionado por el usuario desde teclado.
2. Realizar un algoritmo, que utilizando el archivo de números enteros no ordenados creado en el ejercicio 1, informe por pantalla cantidad de números menores a 1500 y el promedio de los números ingresados. El nombre del archivo a procesar debe ser proporcionado por el usuario una única vez. Además, el algoritmo deberá listar el contenido del archivo en pantalla.
3. Realizar un programa que presente un menú con opciones para:
  - a. Crear un archivo de registros **no ordenados** de empleados y completarlo con datos ingresados desde teclado. De cada empleado se registra: número de empleado, apellido, nombre, edad y DNI. Algunos empleados se ingresan con DNI 00. La carga finaliza cuando se ingresa el String 'fin' como apellido.
  - b. Abrir el archivo anteriormente generado y
    - i. Listar en pantalla los datos de empleados que tengan un nombre o apellido determinado, el cual se proporciona desde el teclado.

- ii. Listar en pantalla los empleados de a uno por línea.
- iii. Listar en pantalla los empleados mayores de 70 años, próximos a jubilarse.

**NOTA:** El nombre del archivo a crear o utilizar debe ser proporcionado por el usuario.

4. Agregar al menú del programa del ejercicio 3, opciones para:

- a. Añadir uno o más empleados al final del archivo con sus datos ingresados por teclado. Tener en cuenta que no se debe agregar al archivo un empleado con un número de empleado ya registrado (control de unicidad).
- b. Modificar la edad de un empleado dado.
- c. Exportar el contenido del archivo a un archivo de texto llamado "todos\_empleados.txt".
- d. Exportar a un archivo de texto llamado: "faltaDNIEmpleado.txt", los empleados que no tengan cargado el DNI (DNI en 00).

**NOTA:** Las búsquedas deben realizarse por número de empleado.

5. Realizar un programa para una tienda de celulares, que presente un menú con opciones para:

- a. Crear un archivo de registros no ordenados de celulares y cargarlo con datos ingresados desde un archivo de texto denominado "celulares.txt". Los registros correspondientes a los celulares deben contener: código de celular, nombre, descripción, marca, precio, stock mínimo y stock disponible.
- b. Listar en pantalla los datos de aquellos celulares que tengan un stock menor al stock mínimo.
- c. Listar en pantalla los celulares del archivo cuya descripción contenga una cadena de caracteres proporcionada por el usuario.
- d. Exportar el archivo creado en el inciso a) a un archivo de texto denominado "celulares.txt" con todos los celulares del mismo. El archivo de texto generado

podría ser utilizado en un futuro como archivo de carga (ver inciso a), por lo que debería respetar el formato dado para este tipo de archivos en la NOTA 2.

**NOTA 1:** El nombre del archivo binario de celulares debe ser proporcionado por el usuario.

**NOTA 2:** El archivo de carga debe editarse de manera que cada celular se especifique en tres líneas consecutivas. En la primera se especifica: código de celular, el precio y marca, en la segunda el stock disponible, stock mínimo y la descripción y en la tercera nombre en ese orden. Cada celular se carga leyendo tres líneas del archivo "celulares.txt".

6. Agregar al menú del programa del ejercicio 5, opciones para:

- a. Añadir uno o más celulares al final del archivo con sus datos ingresados por teclado.
- b. Modificar el stock de un celular dado.
- c. Exportar el contenido del archivo binario a un archivo de texto denominado: "SinStock.txt", con aquellos celulares que tengan stock 0.

**NOTA:** Las búsquedas deben realizarse por nombre de celular.

7. Realizar un programa que permita:

- a) Crear un archivo binario a partir de la información almacenada en un archivo de texto. El nombre del archivo de texto es: "novelas.txt". La información en el archivo de texto consiste en: código de novela, nombre, género y precio de diferentes novelas argentinas. Los datos de cada novela se almacenan en dos líneas en el archivo de texto. La primera línea contendrá la siguiente información: código novela, precio y género, y la segunda línea almacenará el nombre de la novela.

- b) Abrir el archivo binario y permitir la actualización del mismo. Se debe poder agregar una novela y modificar una existente. Las búsquedas se realizan por código de novela.

**NOTA:** El nombre del archivo binario es proporcionado por el usuario desde el teclado.

***IMPORTANTE: Se recomienda implementar los ejercicios prácticos en Dev-Pascal. El ejecutable puede descargarse desde la plataforma Moodle.***

# Introducción a las Bases de Datos

## Fundamentos de Organización de Datos

### Práctica 2

#### Archivos Secuenciales ordenados - Algorítmica Clásica

1. Una empresa posee un archivo con información de los ingresos percibidos por diferentes empleados en concepto de comisión, de cada uno de ellos se conoce: código de empleado, nombre y monto de la comisión. La información del archivo se encuentra ordenada por código de empleado y cada empleado puede aparecer más de una vez en el archivo de comisiones.

Realice un procedimiento que reciba el archivo anteriormente descrito y lo compacte. En consecuencia, deberá generar un nuevo archivo en el cual, cada empleado aparezca una única vez con el valor total de sus comisiones.

**NOTA:** No se conoce a priori la cantidad de empleados. Además, el archivo debe ser recorrido una única vez.

2. El encargado de ventas de un negocio de productos de limpieza desea administrar el stock de los productos que vende. Para ello, genera un archivo maestro donde figuran todos los productos que comercializa. De cada producto se maneja la siguiente información: código de producto, nombre comercial, precio de venta, stock actual y stock mínimo. Diariamente se genera un archivo detalle donde se registran todas las ventas de productos realizadas. De cada venta se registran: código de producto y cantidad de unidades vendidas. Se pide realizar un programa con opciones para:
  - a. Actualizar el archivo maestro con el archivo detalle, sabiendo que:
    - Ambos archivos están ordenados por código de producto.
    - Cada registro del maestro puede ser actualizado por 0, 1 ó más registros del archivo detalle.
    - El archivo detalle sólo contiene registros que están en el archivo maestro.
  - b. Listar en un archivo de texto llamado "*stock\_minimo.txt*" aquellos productos cuyo stock actual esté por debajo del stock mínimo permitido.

3. A partir de información sobre la alfabetización en la Argentina, se necesita actualizar un archivo que contiene los siguientes datos: nombre de provincia, cantidad de personas alfabetizadas y total de encuestados. Se reciben dos archivos detalle provenientes de dos agencias de censo diferentes, dichos archivos contienen: nombre de la provincia, código de localidad, cantidad de alfabetizados y cantidad de encuestados. Se pide realizar los módulos necesarios para actualizar el archivo maestro a partir de los dos archivos detalle.

**NOTA:** Los archivos están ordenados por nombre de provincia y en los archivos detalle pueden venir 0, 1 ó más registros por cada provincia.

4. Se cuenta con un archivo de productos de una cadena de venta de alimentos congelados. De cada producto se almacena: código del producto, nombre, descripción, stock disponible, stock mínimo y precio del producto.  
Se recibe diariamente un archivo detalle de cada una de las 30 sucursales de la cadena. Se debe realizar el procedimiento que recibe los 30 detalles y actualiza el stock del archivo maestro. La información que se recibe en los detalles es: código de producto y cantidad vendida. Además, se deberá informar en un archivo de texto: nombre de producto, descripción, stock disponible y precio de aquellos productos que tengan stock disponible por debajo del stock mínimo. Pensar alternativas sobre realizar el informe en el mismo procedimiento de actualización, o realizarlo en un procedimiento separado (analizar ventajas/desventajas en cada caso).

**Nota:** todos los archivos se encuentran ordenados por código de productos. En cada detalle puede venir 0 o N registros de un determinado producto.

5. Suponga que trabaja en una oficina donde está montada una LAN (red local). La misma fue construida sobre una topología de red que conecta 5 máquinas entre sí y todas las máquinas se conectan con un servidor central. Semanalmente cada máquina genera un archivo de logs informando las sesiones abiertas por cada usuario en cada terminal y por cuánto tiempo estuvo abierta. Cada archivo detalle contiene los siguientes campos: cod\_usuario, fecha, tiempo\_sesion. Debe realizar un procedimiento que reciba los archivos detalle y genere un archivo maestro con los siguientes datos: cod\_usuario, fecha, tiempo\_total\_de\_sesiones\_abiertas.

Notas:

- Cada archivo detalle está ordenado por cod\_usuario y fecha.
- Un usuario puede iniciar más de una sesión el mismo día en la misma máquina, o inclusive, en diferentes máquinas.
- El archivo maestro debe crearse en la siguiente ubicación física: /var/log.

6. Se desea modelar la información necesaria para un sistema de recuentos de casos de covid para el ministerio de salud de la provincia de buenos aires.

Diariamente se reciben archivos provenientes de los distintos municipios, la información contenida en los mismos es la siguiente: código de localidad, código cepa, cantidad de casos activos, cantidad de casos nuevos, cantidad de casos recuperados, cantidad de casos fallecidos.

El ministerio cuenta con un archivo maestro con la siguiente información: código localidad, nombre localidad, código cepa, nombre cepa, cantidad de casos activos, cantidad de casos nuevos, cantidad de recuperados y cantidad de fallecidos.

Se debe realizar el procedimiento que permita actualizar el maestro con los detalles recibidos, se reciben 10 detalles. **Todos los archivos están ordenados por código de localidad y código de cepa.**

Para la actualización se debe proceder de la siguiente manera:

1. Al número de fallecidos se le suman el valor de fallecidos recibido del detalle.
2. Idem anterior para los recuperados.
3. Los casos activos se actualizan con el valor recibido en el detalle.
4. Idem anterior para los casos nuevos hallados.

Realice las declaraciones necesarias, el programa principal y los procedimientos que requiera para la actualización solicitada e informe cantidad de localidades con más de 50 casos activos (las localidades pueden o no haber sido actualizadas).

7. Se dispone de un archivo maestro con información de los alumnos de la Facultad de Informática. Cada registro del archivo maestro contiene: código de alumno, apellido, nombre, cantidad de cursadas aprobadas y cantidad de materias con final aprobado. El archivo maestro está ordenado por código de alumno.

Además, se tienen dos archivos detalle con información sobre el desempeño académico de los alumnos: un archivo de cursadas y un archivo de exámenes finales. El archivo de cursadas contiene información sobre las materias cursadas por los alumnos. Cada registro incluye: código de alumno, código de materia, año de cursada y resultado (solo interesa si la cursada fue aprobada o desaprobada). Por su parte, el archivo de exámenes finales contiene información sobre los exámenes finales rendidos. Cada registro incluye: código de alumno, código de materia, fecha del examen y nota obtenida. **Ambos archivos detalle están ordenados por código de alumno y código de materia**, y pueden contener **0, 1 o más registros por alumno** en el archivo maestro. Un alumno podría cursar una materia muchas veces, así como también podría rendir el final de una materia en múltiples ocasiones.

Se debe desarrollar un programa que **actualice el archivo maestro**, ajustando la cantidad de cursadas aprobadas y la cantidad de materias con final aprobado, utilizando la información de los archivos detalle. Las reglas de actualización son las siguientes:

- Si un alumno aprueba una cursada, se incrementa en uno la cantidad de cursadas aprobadas.
- Si un alumno aprueba un examen final (nota  $\geq 4$ ), se incrementa en uno la cantidad de materias con final aprobado.

Notas:

- **Los archivos deben procesarse en un único recorrido.**
- No es necesario comprobar que no haya inconsistencias en la información de los archivos detalles. Esto es, no puede suceder que un alumno apruebe más de una vez la cursada de una misma materia (a lo sumo la aprueba una vez), algo similar ocurre con los exámenes finales.

8. Se quiere optimizar la gestión del consumo de yerba mate en distintas provincias de Argentina. Para ello, se cuenta con un archivo maestro que contiene la siguiente información: código de provincia, nombre de la provincia, cantidad de habitantes y cantidad total de kilos de yerba consumidos históricamente.

Cada mes, se reciben 16 archivos de relevamiento con información sobre el consumo de yerba en los distintos puntos del país. Cada archivo contiene: código de provincia y cantidad de kilos de yerba consumidos en ese relevamiento. Un archivo de relevamiento puede contener información de una o varias provincias, y una misma provincia puede aparecer cero, una o más veces en distintos archivos de relevamiento.

Tanto el archivo maestro como los archivos de relevamiento están ordenados por código de provincia.

Se desea realizar un programa que actualice el archivo maestro en base a la nueva información de consumo de yerba. Además, se debe informar en pantalla aquellas provincias (código y nombre) donde la cantidad total de yerba consumida supere los 10.000 kilos históricamente, junto con el promedio consumido de yerba por habitante. Es importante tener en cuenta tanto las provincias actualizadas como las que no fueron actualizadas.

**Nota:** cada archivo debe recorrerse una única vez.

9. Se cuenta con un archivo que posee información de las ventas que realiza una empresa a los diferentes clientes. Se necesita obtener un reporte con las ventas organizadas por cliente. Para ello, se deberá informar por pantalla: los datos personales del cliente, el total mensual (mes por mes cuánto compró) y finalmente el monto total comprado en el año por el cliente. Además, al finalizar el reporte, se debe informar el monto total de ventas obtenido por la empresa.

El formato del archivo maestro está dado por: cliente (cod cliente, nombre y apellido), año, mes, día y monto de la venta. El orden del archivo está dado por: cod cliente, año y mes.



**Nota:** tenga en cuenta que puede haber meses en los que los clientes no realizaron compras. No es necesario que informe tales meses en el reporte.

10. Se necesita contabilizar los votos de las diferentes mesas electorales registradas por provincia y localidad. Para ello, se posee un archivo con la siguiente información: código de provincia, código de localidad, número de mesa y cantidad de votos en dicha mesa. Presentar en pantalla un listado como se muestra a continuación:

Código de Provincia

Código de Localidad	Total de Votos
---------------------	----------------

.....	.....
-------	-------

.....	.....
-------	-------

Total de Votos Provincia: \_\_\_\_

Código de Provincia

Código de Localidad	Total de Votos
---------------------	----------------

.....	.....
-------	-------

Total de Votos Provincia: \_\_\_\_

.....

Total General de Votos: \_\_\_\_

**NOTA:** La información está ordenada por código de provincia y código de localidad.

11. Se tiene información en un archivo de las horas extras realizadas por los empleados de una empresa en un mes. Para cada empleado se tiene la siguiente información: departamento, división, número de empleado, categoría y cantidad de horas extras realizadas por el empleado. Se sabe que el archivo se encuentra ordenado por departamento, luego por división y, por último, por número de empleado. Presentar en pantalla un listado con el siguiente formato:

Departamento

División

Número de Empleado	Total de Hs.	Importe a cobrar
--------------------	--------------	------------------

.....	.....	.....
-------	-------	-------

.....	.....	.....
-------	-------	-------

Total de horas división: \_\_\_\_

Monto total por división: \_\_\_\_

División

.....

Total horas departamento: \_\_\_\_

Monto total departamento: \_\_\_\_

Para obtener el valor de la hora se debe cargar un arreglo desde un archivo de texto al iniciar el programa con el valor de la hora extra para cada categoría. La categoría varía de 1 a 15. En el archivo de texto debe haber una línea para cada categoría con el número de categoría y el valor de la hora, pero el arreglo debe ser de valores de horas, con la posición del valor coincidente con el número de categoría.

12. La empresa de software 'X' posee un servidor web donde se encuentra alojado el sitio web de la organización. En dicho servidor, se almacenan en un archivo todos los accesos que se realizan al sitio. La información que se almacena en el archivo es la siguiente: año, mes, día, idUsuario y tiempo de acceso al sitio de la organización. El archivo se encuentra ordenado por los siguientes criterios: año, mes, día e idUsuario.

Se debe realizar un procedimiento que genere un informe en pantalla, para ello se indicará el año calendario sobre el cual debe realizar el informe. El mismo debe respetar el formato mostrado a continuación:

Año : ---

Mes:-- 1

día:-- 1

idUsuario 1 Tiempo Total de acceso en el dia 1 mes 1

-----

idUsuario N Tiempo total de acceso en el dia 1 mes 1

Tiempo total acceso dia 1 mes 1

-----

día N

idUsuario 1 Tiempo Total de acceso en el dia N mes 1

-----

idUsuario N Tiempo total de acceso en el dia N mes 1

Tiempo total acceso dia N mes 1

Total tiempo de acceso mes 1

-----

Mes 12

día 1

idUsuario 1 Tiempo Total de acceso en el dia 1 mes 12

-----

idUsuario N Tiempo total de acceso en el dia 1 mes 12

Tiempo total acceso dia 1 mes 12

-----

día N  
 idUsuario 1    Tiempo Total de acceso en el día N mes 12  
 -----  
 idUsuario N    Tiempo total de acceso en el día N mes 12  
 Tiempo total acceso día N mes 12  
 Total tiempo de acceso mes 12  
 Total tiempo de acceso año

Se deberá tener en cuenta las siguientes aclaraciones:

- El año sobre el cual realizará el informe de accesos debe leerse desde el teclado.
- El año puede no existir en el archivo, en tal caso, debe informarse en pantalla “año no encontrado”.
- Debe definir las estructuras de datos necesarias.
- El recorrido del archivo debe realizarse una única vez procesando sólo la información necesaria.

13. Suponga que usted es administrador de un servidor de correo electrónico. En los logs del mismo (información guardada acerca de los movimientos que ocurren en el server) que se encuentra en la siguiente ruta: /var/log/logmail.dat se guarda la siguiente información: nro\_usuario, nombreUsuario, nombre, apellido, cantidadMailEnviados. Diariamente el servidor de correo genera un archivo con la siguiente información: nro\_usuario, cuentaDestino, cuerpoMensaje. Este archivo representa todos los correos enviados por los usuarios en un día determinado. Ambos archivos están ordenados por nro\_usuario y se sabe que un usuario puede enviar cero, uno o más mails por día.

- a. Realice el procedimiento necesario para actualizar la información del log en un día particular. Defina las estructuras de datos que utilice su procedimiento.
- b. Genere un archivo de texto que contenga el siguiente informe dado un archivo detalle de un día determinado:

nro\_usuarioX.....cantidadMensajesEnviados

.....

nro\_usuarioX+n.....cantidadMensajesEnviados

Nota: tener en cuenta que en el listado deberán aparecer **todos** los usuarios que existen en el sistema. Considere la implementación de esta opción de las siguientes maneras:

- i- Como un procedimiento separado del punto a).
- ii- En el mismo procedimiento de actualización del punto a). Qué cambios se requieren en el procedimiento del punto a) para realizar el informe en el mismo recorrido?

14. Una compañía aérea dispone de un archivo maestro donde guarda información sobre sus próximos vuelos. En dicho archivo se tiene almacenado el destino, fecha, hora de salida y la cantidad de asientos disponibles. La empresa recibe todos los días dos archivos detalles para actualizar el archivo maestro. En dichos archivos se tiene destino, fecha, hora de salida y cantidad de asientos comprados. Se sabe que los archivos están ordenados por destino más fecha y hora de salida, y que en los detalles pueden venir 0, 1 ó más registros por cada uno del maestro. Se pide realizar los módulos necesarios para:

- a. Actualizar el archivo maestro sabiendo que no se registró ninguna venta de pasaje sin asiento disponible.
- b. Generar una lista con aquellos vuelos (destino y fecha y hora de salida) que tengan menos de una cantidad específica de asientos disponibles. La misma debe ser ingresada por teclado.

**NOTA:** El archivo maestro y los archivos detalles sólo pueden recorrerse una vez.

15. Se desea modelar la información de una ONG dedicada a la asistencia de personas con carencias habitacionales. La ONG cuenta con un archivo maestro conteniendo información como se indica a continuación: Código pcia, nombre provincia, código de localidad, nombre de localidad, #viviendas sin luz, #viviendas sin gas, #viviendas de chapa, #viviendas sin agua, #viviendas sin sanitarios.

Mensualmente reciben detalles de las diferentes provincias indicando avances en las obras de ayuda en la edificación y equipamientos de viviendas en cada provincia. La información de los detalles es la siguiente: Código pcia, código localidad, #viviendas con luz, #viviendas construidas, #viviendas con agua, #viviendas con gas, #entrega sanitarios.

Se debe realizar el procedimiento que permita actualizar el maestro con los detalles recibidos, se reciben 10 detalles. **Todos los archivos están ordenados por código de provincia y código de localidad.**

Para la actualización del archivo maestro, se debe proceder de la siguiente manera:

- Al valor de viviendas sin luz se le resta el valor recibido en el detalle.
- Idem para viviendas sin agua, sin gas y sin sanitarios.
- A las viviendas de chapa se le resta el valor recibido de viviendas construidas

La misma combinación de provincia y localidad aparecen a lo sumo una única vez.

Realice las declaraciones necesarias, el programa principal y los procedimientos que requiera para la actualización solicitada e informe cantidad de localidades sin viviendas de chapa (las localidades pueden o no haber sido actualizadas).

16. La editorial X, autora de diversos semanarios, posee un archivo maestro con la información correspondiente a las diferentes emisiones de los mismos. De cada emisión se registra: fecha, código de semanario, nombre del semanario, descripción, precio, total de ejemplares y total de ejemplares vendidos.

Mensualmente se reciben 100 archivos detalles con las ventas de los semanarios en todo el país. La información que poseen los detalles es la siguiente: fecha, código de semanario y cantidad de ejemplares vendidos. Realice las declaraciones necesarias, la llamada al procedimiento y el procedimiento que recibe el archivo maestro y los 100 detalles y realice la actualización del archivo maestro en función de las ventas registradas. Además deberá informar fecha y semanario que tuvo más ventas y la misma información del semanario con menos ventas.

**Nota:** Todos los archivos están ordenados por fecha y código de semanario. No se realizan ventas de semanarios si no hay ejemplares para hacerlo

17. Una concesionaria de motos de la Ciudad de Chascomús, posee un archivo con información de las motos que posee a la venta. De cada moto se registra: código, nombre, descripción, modelo, marca y stock actual. Mensualmente se reciben 10 archivos detalles con información de las ventas de cada uno de los 10 empleados que trabajan. De cada archivo detalle se dispone de la siguiente información: código de moto, precio y fecha de la venta. Se debe realizar un proceso que actualice el stock del archivo maestro desde los archivos detalles. Además se debe informar cuál fue la moto más vendida.

**NOTA:** Todos los archivos están ordenados por código de la moto y el archivo maestro debe ser recorrido sólo una vez y en forma simultánea con los detalles.

18. Se cuenta con un archivo con información de los casos de COVID-19 registrados en los diferentes hospitales de la Provincia de Buenos Aires cada día. Dicho archivo contiene: código de localidad, nombre de localidad, código de municipio, nombre de municipio, código de hospital, nombre de hospital, fecha y cantidad de casos positivos detectados. El archivo está ordenado por localidad, luego por municipio y luego por hospital.

Escriba la definición de las estructuras de datos necesarias y un procedimiento que haga un listado con el siguiente formato:

Nombre: Localidad 1

Nombre: Municipio 1

Nombre Hospital 1.....Cantidad de casos Hospital 1

.....

Nombre Hospital N.....Cantidad de casos Hospital N

Cantidad de casos Municipio 1

.....

Nombre Municipio N

Nombre Hospital 1.....Cantidad de casos Hospital 1

.....  
 Nombre Hospital N.....Cantidad de casos Hospital N  
 Cantidad de casos Municipio N  
 Cantidad de casos Localidad 1  
 -----

Nombre Localidad N

Nombre Municipio 1

Nombre Hospital 1.....Cantidad de casos Hospital 1

.....

Nombre Hospital N.....Cantidad de casos Hospital N

Cantidad de casos Municipio 1  
 -----

Nombre Municipio N

Nombre Hospital 1.....Cantidad de casos Hospital 1

.....

Nombre Hospital N.....Cantidad de casos Hospital N

Cantidad de casos Municipio N

Cantidad de casos Localidad N

Cantidad de casos Totales en la Provincia

Además del informe en pantalla anterior, es necesario exportar a un archivo de texto la siguiente información: nombre de localidad, nombre de municipio y cantidad de casos del municipio, para aquellos municipios cuya cantidad de casos supere los 1500. El formato del archivo de texto deberá ser el adecuado para recuperar la información con la menor cantidad de lecturas posibles.

**NOTA: El archivo debe recorrerse solo una vez.**

19. A partir de un siniestro ocurrido se perdieron las actas de nacimiento y fallecimientos de toda la provincia de buenos aires de los últimos diez años. En pos de recuperar dicha información, se deberá procesar 2 archivos por cada una de las 50 delegaciones distribuidas en la provincia, un archivo de nacimientos y otro de fallecimientos y crear el archivo maestro reuniendo dicha información.

Los archivos detalles con nacimientos, contendrán la siguiente información: nro partida nacimiento, nombre, apellido, dirección detallada (calle, nro, piso, depto, ciudad), matrícula del médico, nombre y apellido de la madre, DNI madre, nombre y apellido del padre, DNI del padre.

En cambio, los 50 archivos de fallecimientos tendrán: nro partida nacimiento, DNI, nombre y apellido del fallecido, matrícula del médico que firma el deceso, fecha y hora del deceso y lugar.

Realizar un programa que cree el archivo maestro a partir de toda la información de los archivos detalles. Se debe almacenar en el maestro: nro partida nacimiento, nombre, apellido, dirección detallada (calle, nro, piso, depto, ciudad), matrícula del médico, nombre y apellido de la madre, DNI madre, nombre y apellido del padre, DNI del padre y si falleció,

además matrícula del médico que firma el deceso, fecha y hora del deceso y lugar. Se deberá, además, listar en un archivo de texto la información recolectada de cada persona.

**Nota:** Todos los archivos están ordenados por nro partida de nacimiento que es única. Tenga en cuenta que no necesariamente va a fallecer en el distrito donde nació la persona y además puede no haber fallecido.

**IMPORTANTE:** Se recomienda implementar los ejercicios prácticos en Dev-Pascal. El ejecutable puede descargarse desde la plataforma moodle.

# Introducción a las Bases de Datos

## Fundamentos de Organización de Datos

### Práctica 3

#### Parte 1: Bajas en archivos

1. Modificar el ejercicio 4 de la práctica 1 (programa de gestión de empleados), agregándole una opción para realizar bajas copiando el último registro del archivo en la posición del registro a borrar y luego truncando el archivo en la posición del último registro de forma tal de evitar duplicados.
2. Definir un programa que genere un archivo con registros de longitud fija conteniendo información de asistentes a un congreso a partir de la información obtenida por teclado. Se deberá almacenar la siguiente información: nro de asistente, apellido y nombre, email, teléfono y D.N.I. Implementar un procedimiento que, a partir del archivo de datos generado, elimine de forma lógica todos los asistentes con nro de asistente inferior a 1000.

Para ello se podrá utilizar algún carácter especial situándolo delante de algún campo String a su elección. Ejemplo: '@Saldaño'.

3. Realizar un programa que genere un archivo de novelas filmadas durante el presente año. De cada novela se registra: código, género, nombre, duración, director y precio. El programa debe presentar un menú con las siguientes opciones:



- a. Crear el archivo y cargarlo a partir de datos ingresados por teclado. Se utiliza la técnica de lista invertida para recuperar espacio libre en el archivo. Para ello, durante la creación del archivo, en el primer registro del mismo se debe almacenar la cabecera de la lista. Es decir un registro ficticio, inicializando con el valor cero (0) el campo correspondiente al código de novela, el cual indica que no hay espacio libre dentro del archivo.
- b. Abrir el archivo existente y permitir su mantenimiento teniendo en cuenta el inciso a), se utiliza lista invertida para recuperación de espacio. En particular, para el campo de "enlace" de la lista (utilice el código de novela como enlace), se debe especificar los números de registro referenciados con signo negativo, . Una vez abierto el archivo, brindar operaciones para:
  - i. Dar de alta una novela leyendo la información desde teclado. Para esta operación, en caso de ser posible, deberá recuperarse el espacio libre. Es decir, si en el campo correspondiente al código de novela del registro cabecera hay un valor negativo, por ejemplo -5, se debe leer el registro en la posición 5, copiarlo en la posición 0 (actualizar la lista de espacio libre) y grabar el nuevo registro en la posición 5. Con el valor 0 (cero) en el registro cabecera se indica que no hay espacio libre.
  - ii. Modificar los datos de una novela leyendo la información desde teclado. El código de novela no puede ser modificado.
  - iii. Eliminar una novela cuyo código es ingresado por teclado. Por ejemplo, si se da de baja un registro en la posición 8, en el campo código de novela del registro cabecera deberá figurar -8, y en el registro en la posición 8 debe copiarse el antiguo registro cabecera.
- c. Listar en un archivo de texto todas las novelas, incluyendo las borradas, que representan la lista de espacio libre. El archivo debe llamarse "novelas.txt".

**NOTA:** Tanto en la creación como en la apertura el nombre del archivo debe ser proporcionado por el usuario.

4. Dada la siguiente estructura:

```
type  
  
    reg_flor = record  
  
        nombre: String[45];  
        codigo: integer;  
    end;  
  
    tArchFlores = file of reg_flor;
```

Las bajas se realizan apilando registros borrados y las altas reutilizando registros borrados. El registro 0 se usa como cabecera de la pila de registros borrados: el número 0 en el campo código implica que no hay registros borrados y -N indica que el próximo registro a reutilizar es el N, siendo éste un número relativo de registro válido.

- a. Implemente el siguiente módulo:

*{Abre el archivo y agrega una flor, recibida como parámetro manteniendo la política descrita anteriormente}*

```
procedure agregarFlor (var a: tArchFlores ; nombre: string;  
codigo:integer);
```

- b. Liste el contenido del archivo omitiendo las flores eliminadas. Modifique lo que considere necesario para obtener el listado.

5. Dada la estructura planteada en el ejercicio anterior, implemente el siguiente módulo:

*{Abre el archivo y elimina la flor recibida como parámetro manteniendo la política descrita anteriormente}*

```
procedure eliminarFlor (var a: tArchFlores; flor:reg_flor);
```

6. Una cadena de tiendas de indumentaria posee un archivo maestro no ordenado con la información correspondiente a las prendas que se encuentran a la venta. De cada prenda se registra: cod\_prenda, descripción, colores, tipo\_prenda, stock y precio\_unitario. Ante un eventual cambio de temporada, se deben actualizar las prendas a la venta. Para ello reciben un archivo conteniendo: cod\_prenda de las

prendas que quedarán obsoletas. Deberá implementar un procedimiento que reciba ambos archivos y realice la baja lógica de las prendas, para ello deberá modificar el stock de la prenda correspondiente a valor negativo.

Adicionalmente, deberá implementar otro procedimiento que se encargue de efectivizar las bajas lógicas que se realizaron sobre el archivo maestro con la información de las prendas a la venta. Para ello se deberá utilizar una estructura auxiliar (esto es, un archivo nuevo), en el cual se copien únicamente aquellas prendas que no están marcadas como borradas. Al finalizar este proceso de compactación del archivo, se deberá renombrar el archivo nuevo con el nombre del archivo maestro original.

7. Se cuenta con un archivo que almacena información sobre especies de aves en vía de extinción, para ello se almacena: código, nombre de la especie, familia de ave, descripción y zona geográfica. El archivo no está ordenado por ningún criterio. Realice un programa que permita borrar especies de aves extintas. Este programa debe disponer de dos procedimientos:
  - a. Un procedimiento que dada una especie de ave (su código) marque la misma como borrada (en caso de querer borrar múltiples especies de aves, se podría invocar este procedimiento repetidamente).
  - b. Un procedimiento que compacte el archivo, quitando definitivamente las especies de aves marcadas como borradas. Para quitar los registros se deberá copiar el último registro del archivo en la posición del registro a borrar y luego eliminar del archivo el último registro de forma tal de evitar registros duplicados.
    - i. Implemente una variante de este procedimiento de compactación del archivo (baja física) donde el archivo se trunque una sola vez.
8. Se cuenta con un archivo con información de las diferentes distribuciones de linux existentes. De cada distribución se conoce: nombre, año de lanzamiento, número de

versión del kernel, cantidad de desarrolladores y descripción. El nombre de las distribuciones no puede repetirse. Este archivo debe ser mantenido realizando bajas lógicas y utilizando la técnica de reutilización de espacio libre llamada **lista invertida**. Escriba la definición de las estructuras de datos necesarias y los siguientes procedimientos:

- a. **BuscarDistribucion:** módulo que recibe por parámetro el archivo, un nombre de distribución y devuelve la posición dentro del archivo donde se encuentra el registro correspondiente a la distribución dada (si existe) o devuelve -1 en caso de que no exista..
- b. **AltaDistribucion:** módulo que recibe como parámetro el archivo y el registro que contiene los datos de una nueva distribución, y se encarga de agregar la distribución al archivo reutilizando espacio disponible en caso de que exista. (El control de unicidad lo debe realizar utilizando el módulo anterior). En caso de que la distribución que se quiere agregar ya exista se debe informar "ya existe la distribución".
- c. **BajaDistribucion:** módulo que recibe como parámetro el archivo y el nombre de una distribución, y se encarga de dar de baja lógicamente la distribución dada. Para marcar una distribución como borrada se debe utilizar el campo cantidad de desarrolladores para mantener actualizada la lista invertida. Para verificar que la distribución a borrar exista debe utilizar el módulo BuscarDistribucion. En caso de no existir se debe informar "Distribución no existente".

## Parte 2: Actualización maestro/detalle, reportes y merge con archivos no ordenados

**Para los ejercicios de esta parte de la práctica, teniendo en cuenta que los archivos no están ordenados por ningún criterio, puede resultar necesario recorrer los archivos más de una vez. La idea es resolver los ejercicios sin ordenar los archivos dados, y comparar la eficiencia (en cuanto al número de lecturas/escrituras) de la solución brindada en esta práctica respecto a la solución para el mismo problema considerando los archivos ordenados.**

1. El encargado de ventas de un negocio de productos de limpieza desea administrar el stock de los productos que vende. Para ello, genera un archivo maestro donde figuran todos los productos que comercializa. De cada producto se maneja la siguiente información: código de producto, nombre comercial, precio de venta, stock actual y stock mínimo. Diariamente se genera un archivo detalle donde se registran todas las ventas de productos realizadas. De cada venta se registran: código de producto y cantidad de unidades vendidas. Resuelve los siguientes puntos:
  - a. Se pide realizar un procedimiento que actualice el archivo maestro con el archivo detalle, teniendo en cuenta que:
    - i. Los archivos no están ordenados por ningún criterio.
    - ii. Cada registro del maestro puede ser actualizado por 0, 1 ó más registros del archivo detalle.
  - b. ¿Qué cambios realizaría en el procedimiento del punto anterior si se sabe que cada registro del archivo maestro puede ser actualizado por 0 o 1 registro del archivo detalle?
2. Se necesita contabilizar los votos de las diferentes mesas electorales registradas por localidad en la provincia de Buenos Aires. Para ello, se posee un archivo con la siguiente información: código de localidad, número de mesa y cantidad de votos en dicha mesa. Presentar en pantalla un listado como se muestra a continuación:

Código de Localidad	Total de Votos
.....	.....
.....	.....
Total General de Votos:	.....

#### NOTAS:

- La información en el archivo no está ordenada por ningún criterio.
- Trate de resolver el problema sin modificar el contenido del archivo dado.
- Puede utilizar una estructura auxiliar, como por ejemplo otro archivo, para llevar el control de las localidades que han sido procesadas.

3. Suponga que trabaja en una oficina donde está montada una LAN (red local). La misma fue construida sobre una topología de red que conecta 5 máquinas entre sí y todas las máquinas se conectan con un servidor central. Semanalmente cada máquina genera un archivo de logs informando las sesiones abiertas por cada usuario en cada terminal y por cuánto tiempo estuvo abierta. Cada archivo detalle contiene los siguientes campos: `cod_usuario`, `fecha`, `tiempo_sesion`. Debe realizar un procedimiento que reciba los archivos detalle y genere un archivo maestro con los siguientes datos: `cod_usuario`, `fecha`, `tiempo_total_de_sesiones_abiertas`.

Notas:

- Los archivos detalle no están ordenados por ningún criterio.
- Un usuario puede iniciar más de una sesión el mismo día en la misma máquina, o inclusive, en diferentes máquinas.

# Introducción a las Bases de Datos

## Fundamentos de Organización de Datos

### Práctica 4

#### **Políticas para la resolución de underflow:**

**Política izquierda:** se intenta redistribuir con el hermano adyacente izquierdo, si no es posible, se fusiona con hermano adyacente izquierdo.

**Política derecha:** se intenta redistribuir con el hermano adyacente derecho, si no es posible, se fusiona con hermano adyacente derecho.

**Política izquierda o derecha:** se intenta redistribuir con el hermano adyacente izquierdo, si no es posible, se intenta con el hermano adyacente derecho, si tampoco es posible, se fusiona con hermano adyacente izquierdo.

**Política derecha o izquierda:** se intenta redistribuir con el hermano adyacente derecho, si no es posible, se intenta con el hermano adyacente izquierdo, si tampoco es posible, se fusiona con hermano adyacente derecho.

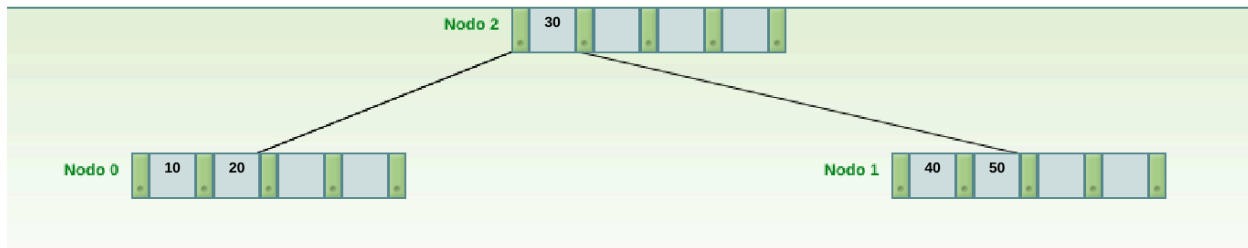
**Casos especiales:** en cualquier política si se tratase de un nodo hoja de un extremo del árbol debe intentarse redistribuir con el hermano adyacente que el mismo posea.

#### **Aclaración:**

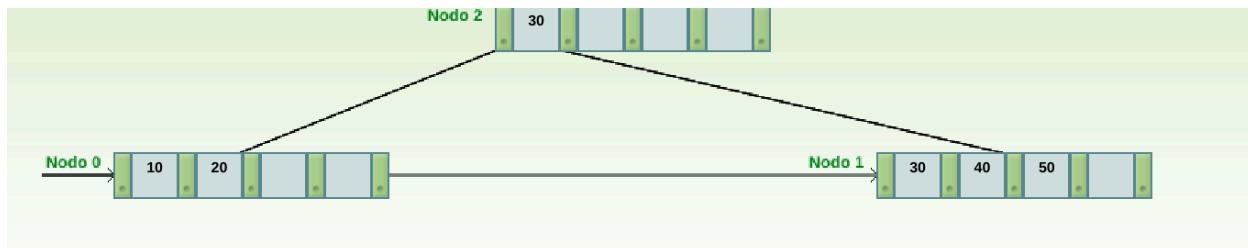
- En caso de underflow lo primero que se intenta **SIEMPRE** es redistribuir y el hermano adyacente se encuentra en condiciones de ceder elementos si al hacerlo no se produce underflow en el.
- En caso de overflow **SIEMPRE** se genera un nuevo nodo. Las claves se distribuyen equitativamente entre el nodo desbordado y el nuevo.

**En el caso de órdenes impares se debe promocionar la clave o la copia (en árbol B+) que se encuentra en la posición del medio.**

Ejemplo árbol B, orden 5

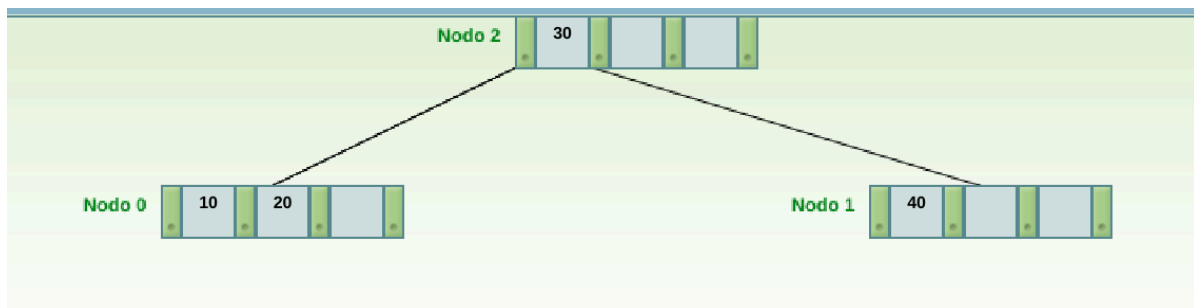


Ejemplo árbol B+, orden 5



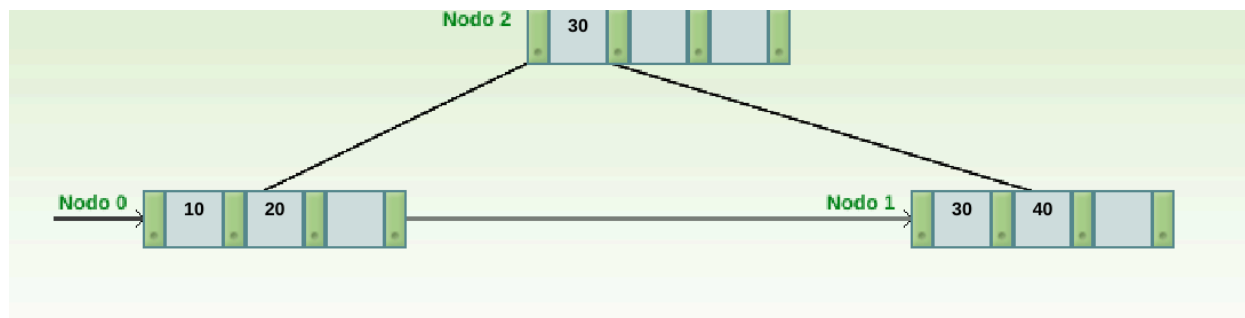
**En el caso de órdenes pares se elige la menor de las claves mayores o su copia (en árbol B+) para promocionar.**

Ejemplo árbol B, orden 4:



Ejemplo árbol B+, orden 4:





## Parte 1: Archivos de datos, índices y árboles B

1. Considere que desea almacenar en un archivo la información correspondiente a los alumnos de la Facultad de Informática de la UNLP. De los mismos deberá guardarse nombre y apellido, DNI, legajo y año de ingreso. Suponga que dicho archivo se organiza como un árbol B de orden M.
  - a. Defina en Pascal las estructuras de datos necesarias para organizar el archivo de alumnos como un árbol B de orden M.
  - b. Suponga que la estructura de datos que representa una persona (registro de persona) ocupa 64 bytes, que cada nodo del árbol B tiene un tamaño de 512 bytes y que los números enteros ocupan 4 bytes, ¿cuántos registros de persona entrarían en un nodo del árbol B? ¿Cuál sería el orden del árbol B en este caso (el valor de M)? Para resolver este inciso, puede utilizar la fórmula  $N = (M-1) * A + M * B + C$ , donde N es el tamaño del nodo (en bytes), A es el tamaño de un registro (en bytes), B es el tamaño de cada enlace a un hijo y C es el tamaño que ocupa el campo referido a la cantidad de claves. El objetivo es reemplazar estas variables con los valores dados y obtener el valor de M (M debe ser un número entero, ignorar la parte decimal).
  - c. ¿Qué impacto tiene sobre el valor de M organizar el archivo con toda la información de los alumnos como un árbol B?
  - d. ¿Qué dato seleccionaría como clave de identificación para organizar los elementos (alumnos) en el árbol B? ¿Hay más de una opción?
  - e. Describa el proceso de búsqueda de un alumno por el criterio de ordenamiento especificado en el punto previo. ¿Cuántas lecturas de nodos se necesitan para encontrar un alumno por su clave de identificación en el peor y en el mejor de los casos? ¿Cuáles serían estos casos?
  - f. ¿Qué ocurre si desea buscar un alumno por un criterio diferente? ¿Cuántas lecturas serían necesarias en el peor de los casos?

2. Una mejora respecto a la solución propuesta en el ejercicio 1 sería mantener por un lado el archivo que contiene la información de los alumnos de la Facultad de Informática (archivo de datos no ordenado) y por otro lado mantener un índice al archivo de datos que se estructura como un árbol B que ofrece acceso indizado por DNI de los alumnos.
- Defina en Pascal las estructuras de datos correspondientes para el archivo de alumnos y su índice.
  - Suponga que cada nodo del árbol B cuenta con un tamaño de 512 bytes. ¿Cuál sería el orden del árbol B (valor de  $M$ ) que se emplea como índice? Asuma que los números enteros ocupan 4 bytes. Para este inciso puede emplear una fórmula similar al punto 1b, pero considere además que en cada nodo se deben almacenar los  $M-1$  enlaces a los registros correspondientes en el archivo de datos.
  - ¿Qué implica que el orden del árbol B sea mayor que en el caso del ejercicio 1?
  - Describa con sus palabras el proceso para buscar el alumno con el DNI 12345678 usando el índice definido en este punto.
  - ¿Qué ocurre si desea buscar un alumno por su número de legajo? ¿Tiene sentido usar el índice que organiza el acceso al archivo de alumnos por DNI? ¿Cómo haría para brindar acceso indizado al archivo de alumnos por número de legajo?
  - Suponga que desea buscar los alumnas que tienen DNI en el rango entre 40000000 y 45000000. ¿Qué problemas tiene este tipo de búsquedas con apoyo de un árbol B que solo provee acceso indizado por DNI al archivo de alumnos?
3. Los árboles B+ representan una mejora sobre los árboles B dado que conservan la propiedad de acceso indexado a los registros del archivo de datos por alguna clave, pero permiten además un recorrido secuencial rápido. Al igual que en el ejercicio 2, considere que por un lado se tiene el archivo que contiene la información de los alumnos de la Facultad de Informática (archivo de datos no ordenado) y por otro lado se tiene un índice al archivo de datos, pero en este caso el índice se estructura como un árbol B+ que ofrece acceso indizado por DNI al archivo de alumnos. Resuelva los siguientes incisos:
- ¿Cómo se organizan los elementos (claves) de un árbol B+? ¿Qué elementos se encuentran en los nodos internos y que elementos se encuentran en los nodos hojas?
  - ¿Qué característica distintiva presentan los nodos hojas de un árbol B+? ¿Por qué?
  - Defina en Pascal las estructuras de datos correspondientes para el archivo de alumnos y su índice (árbol B+). Por simplicidad, suponga que todos los nodos del árbol B+ (nodos internos y nodos hojas) tienen el mismo tamaño

- d. Describa, con sus palabras, el proceso de búsqueda de un alumno con un DNI específico haciendo uso de la estructura auxiliar (índice) que se organiza como un árbol B+. ¿Qué diferencia encuentra respecto a la búsqueda en un índice estructurado como un árbol B?
- e. Explique con sus palabras el proceso de búsqueda de los alumnos que tienen DNI en el rango entre 40000000 y 45000000, apoyando la búsqueda en un índice organizado como un árbol B+. ¿Qué ventajas encuentra respecto a este tipo de búsquedas en un árbol B?

4. Dado el siguiente algoritmo de búsqueda en un árbol B:

```

procedure buscar(NRR, clave, NRR_encontrado, pos_encontrada, resultado)
var clave_encontrada: boolean;
begin
  if (nodo = null)
    resultado := false; {clave no encontrada}
  else
    posicionarYLeerNodo(A, nodo, NRR);
    claveEncontrada(A, nodo, clave, pos, clave_encontrada);
    if (clave_encontrada) then begin
      NRR_encontrado := NRR; { NRR actual }
      pos_encontrada := pos; { posicion dentro del array }
      resultado := true;
    end
    else
      buscar(nodo.hijos[pos], clave, NRR_encontrado, pos_encontrada, resultado)
    end;
end;

```

Asuma que el archivo se encuentra abierto y que, para la primera llamada, el parámetro NRR contiene la posición de la raíz del árbol. Responda detalladamente:

- a. PosicionarYLeerNodo(): Indique qué hace y la forma en que deben ser enviados los parámetros (valor o referencia). Implemente este módulo en Pascal.
- b. claveEncontrada(): Indique qué hace y la forma en que deben ser enviados los parámetros (valor o referencia). ¿Cómo lo implementaría?
- c. ¿Existe algún error en este código? En caso afirmativo, modifique lo que considere necesario.
- d. ¿Qué cambios son necesarios en el procedimiento de búsqueda implementado sobre un árbol B para que funcione en un árbol B+?

5. Defina los siguientes conceptos:

- Overflow
- Underflow
- Redistribución
- Fusión o concatenación

En los dos últimos casos, ¿cuándo se aplica cada uno?

6. Suponga que se tiene un archivo que contiene información de los empleados de una empresa. De cada empleado se mantiene la siguiente información: DNI, legajo, nombre completo y salario. Considere que se mantiene además un índice, organizado como árbol B de orden 4, que provee acceso indizado a los empleados por su DNI. Grafique como queda el archivo de empleados (archivo de datos) y el archivo índice (árbol B) tras la inserción de los siguientes registros:

DNI	Legajo	Nombre y apellido	Salario
35.000.000	100	Juan Perez	\$250000
40.000.000	101	Lucio Redivo	\$400000
32.000.000	102	Nicolás Lapro	\$720000
28.000.000	103	Luis Scola	\$2000000
26.000.000	104	Andres Nocioni	\$1500000
37.000.000	105	Facundo Campazzo	\$1200000
25.000.000	106	Emanuel Ginobili	\$5000000
23.000.000	107	Pepe Sanchez	\$1000000
21.000.000	108	Alejandro Montecchia	\$800000
36.000.000	109	Marcos Delia	\$300000
45.000.000	110	Leandro Bolmaro	\$600000

Notas:

- Grafique los estados de ambos archivos (datos e índice) cuando ocurren cambios relevantes en el índice como la creación de un nuevo nodo.
- Además de graficar los archivos con sus respectivas estructuras internas, resulta útil que grafique la vista del archivo índice como un árbol B.

## Parte 2: Operaciones en árboles B y B+

**Para los siguientes ejercicios debe:**

- **Indicar los nodos leídos y escritos en cada operación.**
- **Todas las operaciones deben estar claramente justificadas, enunciando las mismas indefectiblemente tal cual se presenta en la materia.**

- Los números de nodo deben asignarse en forma coherente con el crecimiento del archivo. La reutilización de nodos libres se debe efectuar con política LIFO (último en entrar, primero en salir).
- Para los siguientes ejercicios sólo interesa graficar los estados del árbol B que representa el índice (no es necesario dibujar la estructura interna de los archivos como si se solicitó en el ejercicio 6).

7. Dado el siguiente árbol B de orden 5, mostrar cómo quedaría el mismo luego de realizar las siguientes operaciones: +320, -390, -400, -533. Justificar detalladamente cada operación indicando lecturas y escrituras en orden de ocurrencia. Para la resolución de underflow debe utilizar **política a izquierda**. Graficar cada operación por separado.

2: 0 (220) 1 (390) 4 (455) 5 (541) 3

0: (10)(150) 1: (225)(241)(331)(360) 4: (400)(407) 5: (508)(533) 3: (690)(823)

8. Dado el siguiente árbol B de orden 4, mostrar cómo quedaría el mismo luego de realizar las siguientes operaciones: +5, +9, +80, +51, -50, -92.

Política de resolución de underflows: derecha.

2: 0 (66) 1

0: (22)(32)(50)

1: (77) (79) (92)

9. Dado el siguiente árbol B de orden 6, mostrar cómo quedaría el mismo luego de realizar las siguientes operaciones: +15, +71, +3, +48, -56, -71.

Política de resolución de underflows: derecha o izquierda.

0: (34) (56) (78) (100) (176)

10. Dado el siguiente árbol B de orden 5, mostrar cómo quedaría el mismo luego de realizar las siguientes operaciones: +450, -485, -511, -614.

Política de resolución de underflows: derecha.

2: 0 (315) 1 (485) 4 (547) 5 (639) 3

0: (148)(223) 1: (333)(390)(442)(454) 4: (508)(511) 5: (614)(633) 3: (789)(915)

11. Dado un árbol B de orden 5 y con política izquierda, para cada operación dada:

- Dibuje el árbol resultante
- Explique las decisiones tomadas
- Escriba las lecturas y escrituras

Operaciones: -76, -400, +900, +12

nodo 2: 3 i 0(76)4(300)1(600)3

nodo 0: 4 h(21)(45)(46)(70)

nodo 4: 2 h(100)(200)

nodo 1: 1 h(400)

nodo 3: 2 h(700)(800)

12. Dada las siguientes operaciones, mostrar la construcción paso a paso de un árbol B de orden 4: +50, +70, +40, +15, +90, +120, +115, +45, +30, +100, +112, +77, -45, -40, -50, -90, -100.

Política de resolución de underflows: izquierda o derecha.

13. Dadas las siguientes operaciones, mostrar la construcción paso a paso de un árbol B de orden 5:

Política de resolución de underflows: izquierda.

+80, +50, +70, +120, +23, +52, +59, +65, +30, +40, +45, +31, +34, +38, +60, +63, +64, -23, -30, -31, -40, -45, -38.

14. Dado el siguiente árbol B de orden 6, mostrar cómo quedaría el mismo luego de realizar las siguientes operaciones: +300, +577, -586, -570, -380, -460.

Política de resolución de underflows: izquierda o derecha.

2: 0 (216) 1 (460) 4 (570) 5 (689) 3 (777) 6

0: (100)(159)(171)

5: (586)(599)(615)(623)(680)

1: (222)(256)(358)(380)(423)

3: (703)(725)

4: (505)(522)

6: (789)(915)(1000)

15. Dada las siguientes operaciones, mostrar cómo se construye el árbol B de orden 4: +65, +89, +23, +45, +20, +96, +10, +55, -23, +110, +50, -10, +25, -50, -45, +120, +130, +70, +75, +73, +100, -120, -110.

Política de resolución de underflows: derecha.

16. Dado el siguiente árbol B + de orden 4 y con política de resolución de underflows a derecha, realice las siguientes operaciones indicando lecturas y escrituras en el orden de ocurrencia. Además, debe describir detalladamente lo que sucede en cada operación. +80, -400.

4: 0 (340) 1 (400) 2 (500) 3

0: (11)(50)(77) 1: (340)(350)(360) 2: (402)(410)(420) 3: (520)(530) -1

17. Dado el siguiente árbol B+ de orden 4, mostrar como quedaría el mismo luego de realizar las siguientes operaciones: +120, +110, +52, +70, +15, -45, -52, +22, +19, -66, -22, -19, -23, -89.

Política de resolución de underflows: derecha.

2: 0 (66) 1

0: (23)(45) 1

1: (66) (67) (89)

18. Dada las siguientes operaciones, mostrar la construcción paso a paso de un árbol B+ de orden 4:

+67, +56, +96, +10, +28, +95, +16, +46, +23, +36, +120, +130, +60, +57, -96, -67, -95, -60, -120, -57, -56.

Política de resolución de underflows: derecha o izquierda.

19. Dada las siguientes operaciones, mostrar la construcción paso a paso de un árbol B+ de orden 6:

+52, +23, +10, +99, +63, +74, +19, +85, +14, +73, +5, +7, +41, +100, +130, +44, -63, -73, +15, +16, -74, -52.

Política de resolución de underflows: izquierda.

20. Dado un árbol B+ de orden 4 y con política izquierda o derecha, para cada operación dada:

- Dibuje el árbol resultante.
- Explique brevemente las decisiones tomadas.
- Escriba las lecturas y escrituras,

Operaciones: +4, +44, -94, -104

nodo 7: 1 i 2(69)6  
 nodo 2: 2 i 0(30)1(51)3  
 nodo 6: 1 i 4(94)5  
 nodo 0: 3 h(5)(10)(20)->1  
 nodo 1: 2 h(30)(40)->3  
 nodo 3: 2 h(51)(60)->4  
 nodo 4: 2 h(69)(80)->5  
 nodo 5: 1 h(104)->-1

21. Dado el árbol B+ que se detalla más abajo, con orden 6, es decir, capacidad de 5 claves como máximo. Muestre los estados sucesivos al realizar la siguiente secuencia de operaciones: +159, -5 y -190, además indicar nodos leídos y escritos en el orden de ocurrencia. Política de resolución underflow derecha.

Nodo 2: 5, i, 0(10)1(60)3(115)4(145)5(179)6

Nodo 0: 2, h, (1)(5) -> 1

Nodo 1: 2, h, (34)(44) -> 3

Nodo 3: 2, h, (60)(113) -> 4

Nodo 4: 4, h, (120)(125)(131)(139) -> 5

Nodo 5: 5, h, (145)(153)(158)(160)(177) -> 6

Nodo 6: 2, h, (179)(190) -> -1

22. Dado un árbol B de orden 5 y con política izquierda o derecha, para cada operación dada:

- Dibuje el árbol resultante
- Explique detalladamente las decisiones tomadas
- Escriba las lecturas y escrituras

Operaciones: +165, +260, +800, -110,

Árbol:           Nodo 8: 1 i 2 (150) 7

Nodo 2: 1 i 0 (120) 3

Nodo 7: 2 i 4 (210)6(300)1

Nodo 0: 2 h (30)(110)

Nodo 3: 1 h (130)



Nodo 4: 4 h (160)(170)(180)(200)

Nodo 6: 4 h (220)(230)(240)(250)

Nodo 1: 4 h (400)(500)(600)(700)

23. Dado un árbol B+ de orden 5 y con política izquierda o derecha, para cada operación dada:

- Dibuje el árbol resultante
- Explique detalladamente las decisiones tomadas
- Escriba las lecturas y escrituras

Operaciones: +250, -300, -40

Árbol: nodo 8: 1 i 2(70)7

nodo2:1 i 0(50)4

nodo7: 4 i 5(90)6(120)3(210)9(300)1

nodo 0: 1 h(40)->4

nodo 4: 1 h(50)->5

nodo 5: 2 h(70)(80)->6

nodo 6: 2 h(90)(100)->3

nodo 3: 2 h(120)(200)->9

nodo 9: 4 h(210)(220)(230)(240)->1

nodo 1: 2 h(400)(500)-> -1

# Introducción a las Bases de Datos

## Fundamentos de Organización de Datos

### Práctica 5

#### Hashing (Dispersión)

##### Parte 1: Preguntas conceptuales

1. Defina el concepto de hashing (o dispersión). ¿Cómo se relaciona este concepto con archivos?
2. Explique el concepto de *función de dispersión*. Enumere al menos tres funciones de dispersión y explique brevemente cómo funciona cada una.
3. Explique los conceptos de *sinónimo*, *colisión* y *desborde (overflow)*. ¿Qué condición es necesaria en el archivo directo para que pueda ocurrir una colisión y no un desborde?
4. ¿Qué alternativas existen para reducir el número de colisiones (y por ende de desbordes) en un archivo organizado mediante la técnica de hashing?
5. Explique brevemente qué es la *densidad de empaquetamiento*. ¿Cuáles son las consecuencias de tener una menor densidad de empaquetamiento en un archivo directo?
6. Explique brevemente cómo funcionan las siguientes técnicas de resolución de desbordes que se pueden utilizar en hashing estático.
  - Saturación progresiva
  - Saturación progresiva encadenada
  - Saturación progresiva encadenada con área de desborde separada
  - Dispersión doble

## Parte 2: Dispersión extensible

**Nota:** (+) indica una operación de inserción, (-) indica una operación de baja.

7. Para las siguientes claves, realice el proceso de dispersión mediante el método de hashing extensible, sabiendo que cada nodo tiene capacidad para dos registros. El número natural indica el orden de llegada de las operaciones. **Se debe mostrar el estado del archivo para cada operación. Justifique brevemente ante colisión y desborde los pasos que realiza.**

1	+ Darin	00111111	2	+ Alterio	11110100
3	+ Sbaraglia	10100101	4	+ De la Serna	01010111
5	+ Altavista	01101011	6	+ Grandinetti	10101010
7	- Altavista	01101011	8	- Sbaraglia	10100101

8. Realice el proceso de dispersión mediante el método de hashing extensible, sabiendo que cada registro tiene capacidad para dos claves. El número natural indica el orden de llegada de las operaciones. **Se debe mostrar el estado del archivo para cada operación. Justifique brevemente ante colisión y desborde los pasos que realiza.**

1	+ Buenos Aires	....1001	2	+ San Juan	....0100
3	+ Entre Ríos	....1110	4	+ Corrientes	....0010
5	+ San Luis	....0101	6	+ Tucumán	....0111
7	+ Río Negro	....0011	8	+ Jujuy	....1111
9	+ Salta	....1010	10	- Río Negro	...0011

9. Para las siguientes claves, realice el proceso de dispersión mediante el método de hashing extensible, sabiendo que cada nodo tiene capacidad para dos registros. El número natural indica el orden de llegada de las operaciones. **Se debe mostrar el estado del archivo para cada operación. Justifique brevemente ante colisión y desborde los pasos que realiza.**

1	+ Tristana	11110010	2	+ Jarvan IV	00111010
3	+ Teemo	01010100	4	+ Annie	10100101
5	+ Ryze	10101110	6	+ Morgana	01101011
7	+ Garen	11001011	8	- Teemo	01010100

10. Para las siguientes claves, realice el proceso de dispersión mediante el método de hashing extensible, sabiendo que cada nodo tiene capacidad para dos registros. El número natural indica el orden de llegada de las operaciones. **Se debe mostrar el estado del archivo para cada operación. Justifique brevemente ante colisión y desborde los pasos que realiza.**

1	+ Guillermo.B	01100011	2	+ Gomez	00000001
3	+ Gustavo.B	01010110	4	+ Sosa	11110100
5	+ Enria	00110101	6	+ Guli	00101000
7	- Gustavo.B	01010110	8	- Sosa	11110100

11. Para las siguientes claves, realice el proceso de dispersión mediante el método de hashing extensible, sabiendo que cada nodo tiene capacidad para dos registros. El número natural indica el orden de llegada de las operaciones. **Se debe mostrar el estado del archivo para cada operación. Justifique brevemente ante colisión y desborde los pasos que realiza.**

1	+ Mansilla	01100010	2	+ Cetré	10001000
3	+ Ascacibar	01010111	4	+ Carrillo	11110101
5	+ Manyoma	00110100	6	+ Méndez	00101001
7	+ Alario	11000101	8	- Mansilla	01100010

12. Realice el proceso de dispersión mediante el método de hashing extensible, sabiendo que cada nodo tiene capacidad para dos claves. El número natural indica el orden de llegada de las operaciones. **Deberá explicar los pasos que realiza en cada operación y dibujar los estados sucesivos correspondientes (inclusive el estado inicial).**

1	+ Aconcagua	10100111	2	+ Kilimanjaro	10101010
3	+ Mont Blanc	00111110	4	+ Cervino	01101111
5	+ Etna	00110101	6	+ Chañi	11110000
7	+ Cho Oyu	01011101	8	+ Vinicunca	01011011
9	- Chañi	11110000	10	- Cervino	01101111