

# Bases de Datos

## Práctico 4: Procesamiento de Consultas

**Ejercicio 1:** Sea la tabla sucursal con esquema (nombre-sucursal, ciudad-sucursal, activos). Supóngase que hay un índice de árbol B+ disponible en ciudad-sucursal y que no hay más índices. ¿Cuál sería el mejor modo de manejar las siguientes selecciones con negaciones?

- a)  $\sigma_{\text{not}(\text{ciudad@sucursal} < \text{"Arganzuela"})}(\text{sucursal})$
- b)  $\sigma_{\text{not}(\text{ciudad@sucursal} = \text{"Arganzuela"})}(\text{sucursal})$
- c)  $\sigma_{\text{not}(\text{ciudad@sucursal} < \text{"Arganzuela"} \parallel \text{activo} < 5000)}(\text{sucursal})$

**Ejercicio 2:** Genere los árboles sintácticos de expresiones, es decir, el plan de ejecución, para las siguientes consultas:

- a)  $\sigma_{\text{precio} < 1000}(\Pi_{\text{color}; \text{precio}}(\text{pinturas}))$
- b)  $\Pi_{\text{color}, \text{precio}}(\sigma_{\text{precio} < 1000}(\text{pinturas}))$

Luego anote en los árboles generados el costo (en cantidad de bloques leídos y escritos) de cada operación, teniendo en cuenta los siguientes parámetros: La tabla pinturas tiene 10.000 registros, y cada bloque contiene 10 registros. Pero se pueden almacenar 100 registros conteniendo sólo color y precio. Se estima que 3/4 de las pinturas tienen un precio mayor o igual a 1000. Para ello, deberá indicar qué algoritmo utiliza. Considere las siguientes situaciones:

- a) Se tiene un índice primario en precio.
- b) Se tiene un índice secundario en precio.
- c) No se tiene un índice en precio.

Recuerde que la tabla temporal que escribe un operador no tiene índices.

**Ejercicio 3:** Dadas las tablas r1(A, B, C) y r2(C, D, E) con las siguientes propiedades: r1 tiene 20.000 tuplas, r2 tiene 45.000 tuplas, 25 tuplas de r1 caben en un bloque y 30 tuplas de r2 que caben en un bloque. Estímese el número de accesos a bloques requeridos utilizando las siguientes estrategias para la reunión  $r1 \bowtie r2$ :

- a) Reunión en bucle anidado.
- b) Reunión en bucle anidado por bloques.
- c) Reunión por mezcla.

¿Y cuántos bloques se escriben? ¿Cambia este número la estrategia utilizada?

**Ejercicio 4:** Realice las mismas tareas que las indicadas en el ejercicio 2 para las siguientes consultas:

- a)  $\Pi_{\text{legajo}, \text{profe.nombre}}(\sigma_{\text{materia.nombre} = \text{"Intro a los Algoritmos"}}(\text{profe legajo} \bowtie \text{a\_carga materia}))$
- b)  $\Pi_{\text{legajo}, \text{profe.nombre}}(\text{profe legajo} \bowtie \text{a\_carga } \sigma_{\text{materia.nombre} = \text{"Intro a los Algoritmos"}}(\text{materia}))$

La tabla materia tiene 100 registros, ocupando 20 bloques, mientras que profe tiene 2.000 registros, ocupando 500 bloques. Considere los diferentes escenarios según la existencia de índice (ninguno, primario, o secundario) en las columnas relevantes de las tablas.

**Ejercicio 5:** Considere las consultas del ejercicio 2 y 4 y responda: ¿En qué situaciones tiene sentido primero ordenar, y en qué parte del árbol sintáctico pondría el ordenamiento? Dibuje el nuevo árbol con sus correspondientes valores de bloques leídos y escritos.

**Ejercicio 6:** Calcule el costo total (de entrada y salida) de las siguientes consultas, utilizando variables para aquellos datos desconocidos. La tabla de las consultas es acceso\_dispositivo(usuario, fecha, dispositivo), y tiene un índice primario en usuario. Justifique el costo en cada parte describiendo brevemente qué haría el algoritmo en uso.

- a)  $v(\Pi_{\text{dispositivo}}(\sigma_{\text{usuario} = \text{"raúl"}}(\text{acceso\_dispositivo})))$
- b)  $\text{usuario} \bowtie \text{count}(\text{dispositivo})(\text{acceso\_dispositivo})$
- c)  $\sigma_{\text{usuario} = \text{"Raúl"}}(\text{acceso\_dispositivo}) \Join \sigma_{\text{usuario} = \text{"Luisa"}}(\text{acceso\_dispositivo})$

**Ejercicio 7:** Escribir el pseudocódigo para un iterador que implementa reunión selectiva indexada por loop anidado, donde la relación externa es entubada. Su pseudocódigo debe definir abrir(), próximo() y cerrar(). Mostrar qué información de estado el iterador debe mantener entre llamadas.

**Ejercicio 8:** escribir un iterador en pseudocódigo para el operador selección  $\sigma_{A \geq v}(r)$  usando índice secundario en A. Definir las 3 operaciones del iterador y cómo se mantiene la información de estado del iterador.