

1^{er} Parcial Diseño y Gestión de Bases de Datos (06/02/2024)-Recuperación

Apellidos, Nombre:.....

1. (0,4 puntos) Indica cuáles de las siguientes afirmaciones son verdaderas y cuáles falsas. Cada respuesta mal contestada resta 0,05 puntos.
 - a. La arquitectura de niveles de un SGBD se compone de 3 niveles: esquema lógico, esquema físico y esquema interno. (F)
 - b. El SGBD sólo puede ejecutar operaciones de actualización de datos requeridas por los usuarios si los bloques afectados por la actualización se encuentran en memoria principal. (V)
 - c. Los privilegios discrecionales genéricos son los que permiten a cualquier usuario (publico) acceder a objetos concretos de la base de datos. (F)
 - d. La concesión de privilegios fijos se basa en niveles de seguridad como Top Secret, Secret, Confidential y Unclassified. (V)
2. (0,3 puntos) Indica cuáles son los factores que intervienen en el tiempo de respuesta de la ejecución de consultas sobre una base de datos.

- La gestión de los bloques en disco.
- La gestión de los buffers en memoria principal.
- La gestión de los ficheros que implementan las estructuras de datos de la BD:
 - organización de ficheros.
 - técnicas de indexación utilizadas.

3. (0,4 puntos) Sean los siguientes supuestos:
 - Sea **T** una tabla con un atributo **A** que no es clave (que no tiene restricción de unicidad).
 - **T** tiene 10.000 filas.
 - En el atributo **A** de **T** aparecen 50 valores distintos con una distribución uniforme.
 - **T** se implementa como un fichero, **F**, ordenado sobre **A** con un factor de bloque = 10.
 - Se define un índice multinivel en **F** sobre el atributo **A** con una altura =2.

¿Cuál es el coste, en número de accesos a bloque, para una operación de selección con condición de igualdad sobre **A** usando el índice?

En el índice hay que leer un bloque por nivel (2).

Si tenemos 50 valores distintos de **A** almacenados de manera consecutiva en los bloques del fichero (fichero ordenado por **A**) con una distribución uniforme, hay 200 registros con cada valor (10.000/50) dado que en cada bloque caben 10 registros, habrá que leer 20 bloques para recuperar esos 200 registros.

Coste total: $2 + \lceil (10000/50)/10 \rceil = 22$

4. (0,4 puntos) Sea **F** un fichero, **A** y **B** dos campos del fichero e **I** un índice definido sobre **A**. Si el fichero tiene 500 filas con un factor de bloque de 10, indica, para cada uno de los supuestos que aparecen la siguiente tabla qué tipo de índice será **I** (primario, secundario o de agrupación), si será disperso o denso y cuántas entradas tendrá el índice en cada caso.

Supuestos	Tipo	Denso/Disperso	Nº Entradas
a) F está ordenado por A y A es un campo clave en F	Primario	Disperso	50: Una entrada por cada bloque de datos
b) F está ordenado por A y A no es un campo clave en F. A tiene 100 valores diferentes.	Agrupación	Disperso	100: Una entrada por cada valor distinto de A.
c) F está ordenado por B. A es un campo clave en F. B es un campo clave en F.	Secundario	Denso	500: una entrada por cada valor de B
d) F está ordenado por B. A no es un campo clave en F. B es un campo clave en F. A tiene 100 valores diferentes.	Secundario	Disperso	100: Una entrada por cada valor distinto de A.
e) F está ordenado por B. A es un campo clave en F y B no es un campo clave en F	Secundario	Denso	500: una entrada por cada valor de A
f) F está ordenado por B. A no es un campo clave en F. B no es un campo clave en F. A tiene 100 valores diferentes.	Secundario	Disperso	100: Una entrada por cada valor distinto de A.
g) F está desordenado. A es un campo clave en F.	Secundario	Denso	500: una entrada por cada valor de A
h) F está desordenado. A no es un campo clave en F	Secundario	Disperso	100: Una entrada por cada valor distinto de A.

5. (0,2 puntos) Explica brevemente en qué consiste el método de concatenación $R \bowtie_{R.A=S.A} S$ por direccionamiento calculado

- Crear una copia, implementada con direccionamiento calculado, de la tabla más pequeña (p.e. S), con una función de dispersión f sobre A.
- Recorrer R (la tabla mayor) en un bucle, y para cada registro r de R aplicar f sobre A para acceder a un cubo (o bloque) de la copia de S donde buscar los registros coincidentes ($r.A=s.A$).

6. (0,2 puntos) Explica cómo se realiza la ejecución de un árbol de consulta.

- Empezar con los nodos hoja.
- Ejecutar un nodo interno (operador) cuando sus operandos (nodos-hijo) están disponibles (tablas).
- Reemplazar el nodo interno por el resultado de la operación.
- La ejecución termina al llegar al nodo raíz y realizar el último reemplazamiento

7. Sea el siguiente esquema relacional:

R(A:entero, B:entero, C:boolean, D:entero)

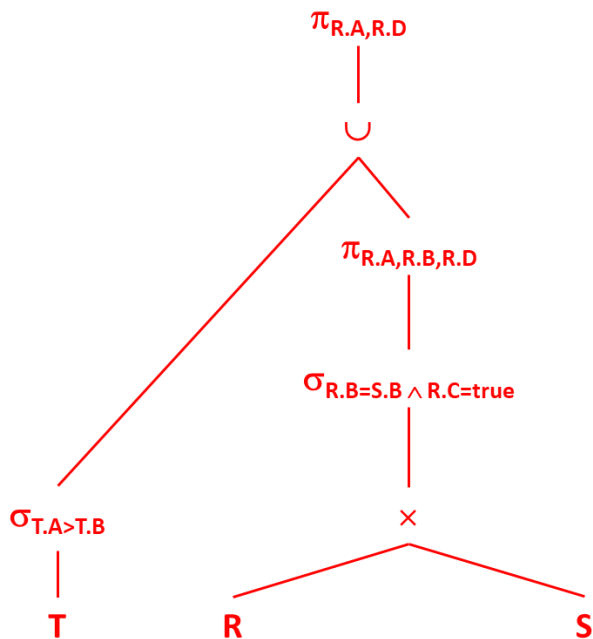
S(E:entero, F:char(10))

T(A:entero, B:entero, D: entero)

Y sea la siguiente consulta en algebra relacional:

$$\pi_{A,D}(\sigma_{A>B} T \cup \pi_{A,B,D} \sigma_{B=E \wedge C=true}(R \times S))$$

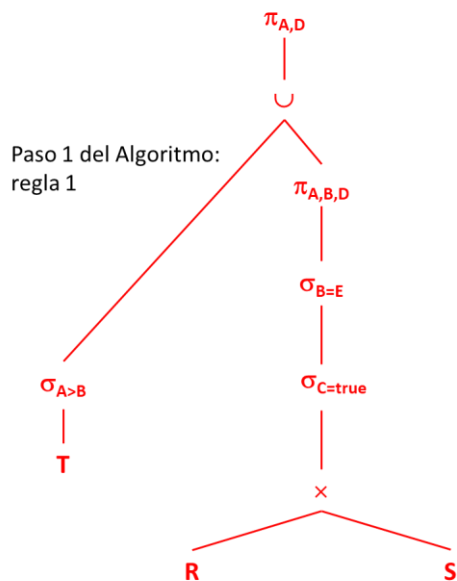
a) (0,2 puntos) Construye el árbol de consulta asociado.



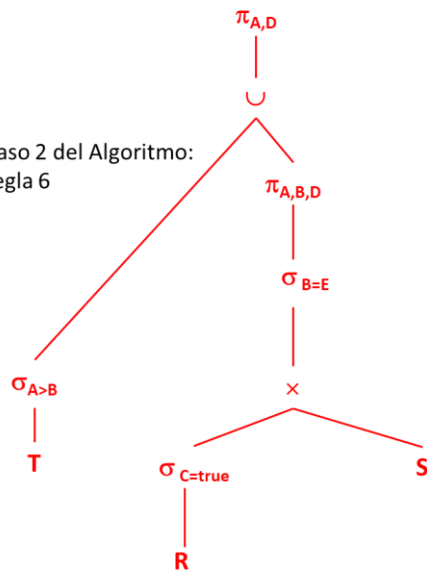
- b) (0,1 puntos) Esta pregunta respóndela después de resolver el apartado c). Escribe cuál sería la expresión de álgebra relacional equivalente al último árbol generado en la pregunta siguiente.

$$\pi_{A,D} \sigma_{A>B} T \cup \pi_{A,D} (\sigma_{C=true} R \bowtie_{B=E} S)$$

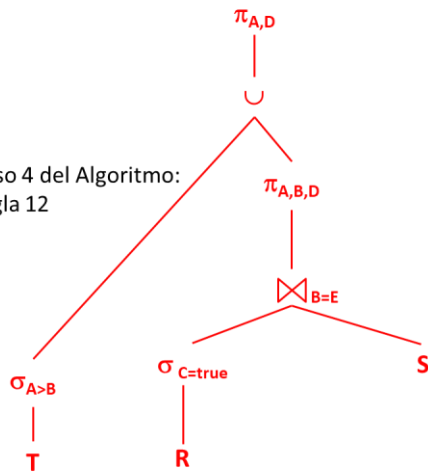
- c) (0,4 puntos) Optimiza el árbol de consulta obtenido, indicando cada una de las reglas utilizadas y dibujando el árbol resultante de la aplicación de cada una. Usa el algoritmo que se te proporciona.



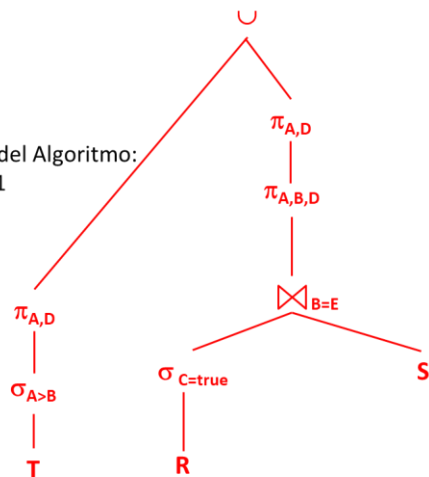
Paso 2 del Algoritmo:
regla 6



Paso 4 del Algoritmo:
regla 12



Paso 5 del Algoritmo:
regla 11



Paso 5 del
Algoritmo: regla 3

