

Ejercicio 1 (4 puntos): Supongamos que tenemos la siguiente base de datos:

estudiante (id estudiante, nombre, carrera)

curso (id curso, nombre curso, docente)

inscripción (id estudiante, id curso, fecha inscripción, nota)

Asumir la siguiente información sobre las tablas:

Tabla	Cantidad de tuplas	Otras informaciones
estudiante	10000	
curso	500	
inscripción	40000	V(id estudiante) = 9000 V(id curso) = 450

Sea la consulta: $\Pi_{\text{nombre}} (\sigma_{\text{nota} > 8} (\text{estudiante} \bowtie \text{inscripción}))$

Asumir que se tiene un índice primario en *id estudiante* de *inscripción* y *estudiante* está ordenado por *nombre*. Asumir que en un nodo del índice entran 314 claves de búsqueda. Asumir que un estudiante puede hacer como máximo 20 cursos. Asumir que en un bloque de estudiante entran 56 tuplas.

Se pide:

1. Calcular la altura del árbol B+ del índice de *inscripción*. (usar máximo 2 líneas).
2. Calcular la cantidad de bloques de la tabla *estudiante* asumiendo que una tupla de estudiante consume 73 B y que un bloque ocupa 4096 bytes. (usar máximo una línea)
3. Considerar el operador físico para la reunión natural que emplea el índice sobre *inscripción*. Calcular su costo en transferencia de bloques en el peor caso (usar máximo dos líneas).

Ejercicio 2 (3 puntos): Considerar la reunión natural de tres tablas: S, T y U. Por simplicidad asumimos que cada una tiene 1000 tuplas. Asumimos:

$S(b, c)$	$T(c, d)$	$U(d, a)$
$V(S, b) = 100$		$V(U, a) = 50$
$V(S, c) = 500$	$V(T, c) = 20$	
	$V(T, d) = 50$	$V(U, d) = 1000$

Nombre _____

Se trabaja con el algoritmo de programación dinámica que encuentra el mejor árbol de reunión profunda a la izquierda. Usar lo siguiente:

	$\{S\}$	$\{T\}$	$\{U\}$
Size	1000	1000	1000
Cost	0	0	0
Best plan	S	T	U

	$\{S, T\}$	$\{S, U\}$	$\{T, U\}$
Size	2000	1,000,000	1000
Cost	0	0	0
Best plan	$S \bowtie T$	$S \bowtie U$	$T \bowtie U$

1. Indicar cómo calcular el costo de la reunión del mejor plan de $X - \{r\}$ y r , donde $X = \{S, T, U, W\}$, W es una tabla adicional que no mostramos (usar 2 líneas) - 40%
2. Indicar cómo hacer para elegir el mejor plan para $\{S, T, U\}$. Solo árboles de reunión profunda a la izquierda son contemplados (usar dos líneas) - 30%
3. Considerando $\{S, T, U\}$ ¿Cuál es el mejor plan? (solo dar el resultado en 1 línea). 30 %

Ejercicio 3 (3 puntos): Suponga que está diseñando un sistema de recuperación de información basado en un índice invertido. Se trabaja con un lenguaje de consultas de proximidad. El mismo incluye el operador SAME, cuya semántica exige que los términos aparezcan en el mismo párrafo. No es posible implementar correctamente el operador SAME si el índice invertido no accede a información sobre los párrafos. Se pide:

1. Definir una estructura eficiente para la lista de ocurrencias de un término
¿Cómo son las tuplas de esa lista? (Usar máximo dos líneas).
2. Definir lógicamente el significado del operador SAME. (usar máximo 4 líneas)