

Control 2 CC42A

Prof. C. Gutiérrez / Auxs. M. San Martín, P. Rossel

18 Octubre 2006

Considere la relación `Emp(empID, sexo, salario, edad, deptID, jefeID)`, donde `empID` es llave primaria, y `jefeID` referencia la tabla `Emp`, y representa el jefe del empleado correspondiente.

Pregunta 1 Exprese en Algebra Relacional la consulta:

1. Todas los pares de empleados (`e1`, `e2`) de hombres - mujeres tales que tienen la misma edad, trabajan en el mismo departamento, y la mujer gana menos.
2. Los departamentos que tienen hombres y mujeres si en la empresa hay hombres y mujeres (y si en la empresa sólo hay hombres o sólo mujeres, todos).

Pregunta 2 Exprese en SQL la consulta o indique porqué no puede expresarse:

1. Los departamentos que tienen más empleados mujeres que hombres.
2. Indique una consulta que entregue todos los subordinados del empleado con ID 7675. (Nota: Los subordinados de un empleado *A* también son subordinados del jefe de *A*.)

Pregunta 3 Considere el siguiente esquema de una tabla de una base de datos que guarda la información de usuarios de un servidor web muy popular, donde fecha y hora son las del momento en que el usuario se logea cada vez.

`Usuarios(login, passwd, fechaLg, horaLg, id-archivo-visitado)`

1234	gr54	2006-3-12	12:00	345
3778	aabc	2005-5-21	19:21	222
3778	aabc	2005-5-21	19:21	57
1234	gr54	2006-1-1	12:15	345
1234	pq22	2006-5-14	02:05	21

..etc..

Al administrador le interesan dos tipos de consultas:

- (a) “Cuáles fueron los archivos visitados despues del dia Y”
- (b) “Los días que se ha logeado el usuario X” (cada fecha y hora distinta indica un login distinto).

1. ¿Por cuál atributo conviene ordenar el archivo de datos para optimizar la suma de los tiempos que toman ambas consultas (sin índices)? Justifique.
2. Indique qué índices y de qué tipo (hash/B+, agrupado/no) crear para mejorar la performance general del administrador. Justifique.

Pregunta 4 Sobre la misma tabla de la pregunta anterior, considere la consulta “El usuario (login) que se ha logeado más veces” (recuerde que cada fecha y hora distinta indica un login distinto).

Indique cuál alternativa es más eficiente para la consulta (justifique estimando el tiempo):

1. El archivo de datos está ordenado por (login, fechaLg, horaLg) y no hay índices.
2. Hay índice B+ agrupado sobre login.
3. Hay índice de hash sobre (login, fechaLg, horaLg)