

OBLIGATORIO 1 EDA 2022

PROTOTIPO DE MANEJADOR DE BASE DE DATOS

Se desea implementar un sistema manejador de base de datos que permita crear y mantener tablas, almacenar datos en ellas y realizar ciertas operaciones sobre los datos.

Una **base de datos** es un conjunto de tablas que poseen nombres únicos que permiten identificarlas. Una **tabla** es, desde el punto de vista lógico, un conjunto de **filas** y **columnas (campos)** donde se distribuye la información.

Ejemplo:

Tabla Personas

Nombre	Apellido	CI
María	Castaña	0.000.001
Juan	Pérez	2.256.698
Daniel	Ficha	3.333.444
Laura	Pérez	2.123.328

Cada columna tiene un nombre que debe ser único dentro de las columnas de la tabla y contiene datos de una misma naturaleza. Si tomamos como ejemplo la tabla de personas, cada columna representa un atributo de las personas. Luego las filas contienen los datos de cada persona. Una **tupla** es la colección de datos presentes en una fila de la tabla, considerando el orden en que se presentan los datos. Esto nos lleva a que dos tablas con todas sus columnas iguales pero dispuestas en distinto orden se consideren tablas distintas. En otras palabras, una tabla es un conjunto de tuplas y una tupla es un conjunto de datos, debiendo cumplirse que para una misma tabla todas las tuplas tienen la misma cantidad de datos de los mismos tipos y en el mismo orden. No se permite que en una tabla existan dos tuplas idénticas.

Existen tres formas de **calificar** a una columna de una tabla: NOT EMPTY, PRIMARY KEY y ANY. NOT EMPTY significa que la columna no admite campos vacíos. Es decir no admite campos con valor EMPTY. El calificador PRIMARY KEY puede aplicarse a 0 o 1 columnas de una tabla. PRIMARY KEY indica que los valores de la columna no pueden ser vacíos y son únicos. Entonces, se puede identificar unívocamente a un tupla por el valor del campo correspondiente a la columna PRIMARY KEY. ANY se considera un calificador neutro, es decir, no restringe los valores de la columna. El valor EMPTY se acepta para cualquier columna, independientemente de su tipo de datos, siempre que la columna no esté calificada como NOT EMPTY ni como PRIMARY KEY. Notar que en el ejemplo previo de la tabla de personas, Apellido no podría ser un campo PRIMARY KEY, pero sí el campo CI (de acuerdo a las tuplas del ejemplo).

Por motivos de simplicidad se considerarán sólo dos posibles tipos de datos almacenables en una tabla: strings e integers. Un string es una secuencia de caracteres donde se excluyen los siguientes caracteres: el mayor (>), el menor (<), el igual (=) y el dos puntos (:).

En resumen, el sistema deberá poder almacenar y administrar tablas que cumplan el siguiente esquema:

- Un nombre que identifica a la tabla de manera única
- Las columnas, de las que se conoce su:
 - Nombre (que la identifica dentro de la tabla)
 - Tipo de datos (string, integer)
 - Calificador que se aplica a esa columna

Consideraciones Generales

El sistema “**deberá**” implementar de forma eficiente las operaciones sobre la base de datos. En particular, teniendo en cuenta la posible existencia en las tablas de una clave primaria (columna PRIMARY KEY).

En caso de que alguna operación, de las que se describen más adelante, genere una tabla con tuplas repetidas, se deberán eliminar las tuplas repetidas, dejando tan sólo una de éstas. Note que de no hacerlo se estaría violando la definición de tabla.

Si la aplicación de alguna operación, de las que se describen más adelante, deja a las tablas de la base de datos en un estado inconsistente o la operación no está permitida por su especificación o por las reglas antes mencionadas se establecerá una situación de error en la cual:

- Permanecerá invariante el estado de la base de datos
- Se mostrará un mensaje de error adecuado y clarificador

A continuación, se describen las operaciones del sistema.

Operaciones sobre la Base de Datos

CREAR TABLA – createTable(nombreTabla)

Descripción: Crea una nueva tabla vacía (sin columnas ni tuplas) en la base de datos con nombre: nombreTabla, siempre que no exista ya una tabla con dicho nombre.

ELIMINAR TABLA – dropTable(nombreTabla)

Descripción: Elimina la tabla de nombre nombreTabla de la base de datos, si ésta existe, y las tuplas que la misma posee.

OPERACIONES PARA MODIFICAR UNA TABLA:

– addCol (nombreTabla, nombreCol, tipoCol, calificadorCol)

Descripción: Agrega a la tabla de nombre nombreTabla, si ésta existe, una nueva columna al final de nombre nombreCol, si ésta no existe, tipo tipoCol y calificador calificadorCol. Si la tabla tiene tuplas, el nuevo campo tendrá el valor EMPTY en cada tupla. Por lo tanto, en el caso en que la tabla tenga tuplas no es válido que se agregue un calificador distinto de ANY. Tampoco es válido que calificadorCol sea PRIMARY KEY si existe ya una columna con dicho calificador en la tabla nombreTabla.

– dropCol (nombreTabla, nombreCol)

Descripción: Elimina de la tabla de nombre nombreTabla, si ésta existe, la columna de nombre nombreCol, si ésta existe. Si la tabla tiene tuplas, entonces se eliminará de éstas el campo correspondiente a la columna eliminada. Si la tabla tenía una única columna de nombre nombreCol entonces quedará como tabla vacía.

– alterCol (nombreTabla, nombreCol, tipoColNuevo, calificadorColNuevo, nombreColNuevo)

Descripción: Modifica de la tabla de nombre nombreTabla, si ésta existe, la columna de nombre nombreCol, si ésta existe, quedando esta columna con el nuevo tipo de datos tipoColNuevo, calificador calificadorColNuevo y nombre nombreColNuevo, si éste último no es el nombre de otra columna de la tabla. Si la tabla tiene tuplas, los valores de la columna modificada deberán satisfacer las nuevas características (tipo de dato y calificador). El tipo de datos sólo puede cambiar de integer a

string y en este caso se deberá realizar la conversión de tipo de la columna especificada en todas las tuplas de la tabla.

Operaciones para la Edición de Datos

INSERTAR TUPLA – insertInto(nombreTabla, columnasTupla, valoresTupla)

Descripción: Inserta en la tabla de nombre nombreTabla, si ésta existe, una tupla con los valores dados en valoresTupla para las columnas indicadas en columnasTupla, si los nombres de las columnas existen, los valores son del tipo adecuado y satisfacen los calificadores correspondientes a cada columna. Si no se indican todas las columnas se inserta EMPTY en las otras. Por lo tanto, la operación se permite sólo si las columnas que no se indican tienen el calificador ANY. Los nombres de las columnas en columnasTupla y los valores de valoresTupla se separan con el uso del caracter dos puntos (:) y deben corresponderse uno a uno. Esto es, el nombre de columna *i* en columnasTupla con el valor en la posición *i* de valoresTupla. Si la tupla a insertar pertenece a la tabla, la operación no tendrá efecto.

ELIMINAR TUPLA – delete(nombreTabla, condicionEliminar)

Descripción: Elimina de la tabla de nombre nombreTabla, si ésta existe, todas las tuplas que cumplen la condición condiciónEliminar. En caso de que la condición sea “”, se eliminan todas las tuplas de la tabla. Si ninguna tupla cumple la condición, la operación no tendrá efecto.

El formato de las condiciones es: *columna operador valor* (sin espacios en blanco intermedios). Los operadores a utilizar son: = “igual”, <> “Distinto”, > “Mayor” y < “Menor”.

Por ejemplo, Sexo=Masculino, Edad<18, Código<>20, Apellido>Perez.

Para comparar strings con el operador > o < se utilizará el orden lexicográfico habitual.

El valor EMPTY puede usarse en una condición. La condición *columna*=EMPTY resulta verdadera para una tupla si, y sólo si, el valor de la columna en dicha tupla es EMPTY; *columna*<>EMPTY resulta verdadera para una tupla si, y sólo si, el valor de la columna en dicha tupla es distinto de EMPTY. En cualquier otro caso, una condición que involucre el valor EMPTY resulta ser falsa. Asimismo, toda condición que no involucre al valor EMPTY resultará falsa al ser instanciada por una tupla que tenga el valor EMPTY en la columna de la condición.

MODIFICAR TUPLA

– update(nombreTabla, condicionModificar, columnaModificar, valorModificar)

Descripción: Modifica en la tabla de nombre nombreTabla, si ésta existe, el valor de las tuplas en la columna de nombre columnaModificar, si ésta existe, que cumplen la condición condiciónModificar. En la columna especificada de las tuplas que cumplen la condición se asigna el valor valorModificar, siempre que este valor sea del tipo adecuado y satisfaga el calificador de la columna especificada. En caso de que la condición sea “”, la operación tiene aplicación sobre todas las tuplas de la tabla.

Operación entre Tablas

SELECCIÓN – selectWhere (nombreTabla1, condicion, nombreTabla2)

Descripción: Dado un nombre de tabla nombreTabla1 y una condición, genera una nueva tabla en la base de datos de nombre nombreTabla2, si nombreTabla1 existe y nombreTabla2 no existe, con las tuplas de la tabla nombreTabla1 que cumplan la condición. En caso de que la condición sea “”, la operación tiene aplicación sobre todas las tuplas de la tabla. La condición respeta el formato descripto anteriormente para condiciones (ver operación delete).

Operaciones para la Impresión de Información

LISTAR TABLAS – printTables()

Descripción: Imprime los nombres de las tablas de la base de datos del sistema, ordenados alfabéticamente de menor a mayor.

LISTAR ESQUEMA – printMetadata(nombreTabla)

Descripción: Imprime el esquema de la tabla de nombre nombreTabla, si ésta existe. Es decir, imprime el nombre de la Tabla, los nombres de sus columnas en el orden correspondiente, indicando para cada columna su tipo de datos y calificador si lo tuviera.

LISTAR TABLA – printDataTable (nombreTabla, ordenadaPor)

Descripción: Imprime las tuplas de la tabla de nombre nombreTabla, si ésta existe, ordenados de acuerdo a las columnas especificadas en el parámetro ordenadaPor. Los nombres de las columnas se expresan en el formato *columna₁:columna₂: ... :columna_n*. Las tuplas se muestran ordenadas por *columna₁* (de menor a mayor) y si dos tuplas coinciden en el valor de la columna *columna₁*, entonces éstas se ordenan (de menor a mayor) por el campo *columna₂*, y así sucesivamente según los campos especificados, de izquierda a derecha, en ordenadaPor. Si el parámetro ordenadaPor es “”, imprime las tuplas en cualquier orden, salvo que la tabla tenga un campo PRIMARY KEY, en cuyo caso las tuplas se imprimen ordenadas (de menor a mayor) según dicho campo.

ENTREGA JUEVES 6 DE OCTUBRE A LAS 23:55 HS