

FICHEROS Y BASES DE DATOS (E44)

3º INGENIERÍA EN INFORMÁTICA

Tema 9.

Reglas de Integridad

- 1.- Introducción.
- 2.- Claves Primarias.
- 3.- Regla de Integridad de Entidades.
- 4.- Claves Ajenas.
- 5.- Regla de Integridad Referencial.
- 6.- Reglas para las Claves Ajenas.

(Capítulo 12 del Date)

INTRODUCCIÓN

Reglas de Integridad

- Las bases de datos se descomponen en un conjunto de valores que reflejan la realidad.
- No todos los posibles conjuntos de valores tienen sentido.
- Por tanto resulta necesario ampliar la definición de las bases de datos para que incluya ciertas Reglas de Integridad.
- De este modo se intenta restringir los conjuntos de valores, para que se ajusten a valores posibles.
- La mayoría de las reglas de integridad son específicas de las bases de datos.
- Pero existen dos reglas generales asociadas al modelo relacional, que se refieren a las claves primarias y claves ajenas.

Comentarios Generales

- Todas las reglas de integridad se aplican sobre relaciones base, ya que son estas las que reflejan la realidad.
- Las dos reglas generales asociadas al modelo relacional, deben definirse como Metarreglas, ya que inducen otras reglas de integridad.
- Se puede definir una metarregla adicional, que aparece de modo implícito en el modelo relacional, asociada a los dominios.

INTRODUCCIÓN

Tablas de Ejemplo

```
CREATE DOMAIN S#           CHAR(5);
CREATE DOMAIN NOMBRE      CHAR(20);
CREATE DOMAIN SITUACION  SMALLINT;
CREATE DOMAIN CIUDAD     CHAR(15);
CREATE DOMAIN P#         CHAR(6);
CREATE DOMAIN COLOR      CHAR(6);
CREATE DOMAIN PESO       SMALLINT;
CREATE DOMAIN CANT       INTEGER;
```

```
CREATE TABLE S
( S#           DOMAIN (S#)           NOT NULL,
  SNOMBRE     DOMAIN (NOMBRE)       NOT NULL,
  SITUACION   DOMAIN (SITUACION)    NOT NULL,
  CIUDAD      DOMAIN (CIUDAD)       NOT NULL,
  PRIMARY KEY (S#) );
```

```
CREATE TABLE P
( P#           DOMAIN (P#)           NOT NULL,
  PNOMBRE     DOMAIN (NOMBRE)       NOT NULL,
  COLOR       DOMAIN (COLOR)        NOT NULL,
  PESO        DOMAIN (PESO)         NOT NULL,
  CIUDAD      DOMAIN (CIUDAD)       NOT NULL,
  PRIMARY KEY (P#) );
```

```
CREATE TABLE SP
( S#           DOMAIN (S#)           NOT NULL,
  P#           DOMAIN (P#)           NOT NULL,
  CANT         DOMAIN (CANT)         NOT NULL,
  PRIMARY KEY (S#, P#),
  FOREIGN KEY (S#) REFERENCES S,
  FOREIGN KEY (P#) REFERENCES P );
```

CLAVES PRIMARIAS

Definición

- Una Superclave es un conjunto de atributos que identifican de modo único las tuplas de una relación.
- Por definición el conjunto formado por todos los atributos de una relación se define como una superclave.
- Una Clave Candidata es el menor subconjunto de atributos de una superclave que sigue siendo un identificador único.
- En una relación pueden existir diferentes claves candidatas que se compongan de un número diferente de atributos.
- De todas las claves candidatas se elige una que será la Clave Primaria.
- El resto de claves candidatas se definen como Claves Alternativas.

Propiedades

- Por lo anterior, existen dos propiedades básicas asociadas a una clave candidata:
 - Unicidad. No existen dos tuplas que posean el mismo valor de la clave candidata.
 - Minimalidad. No se puede eliminar ningún atributo de la clave candidata sin destruir la unicidad de la clave candidata.
- Estas propiedades son heredadas por las claves primarias.

CLAVES PRIMARIAS

Comentarios

- En una relación aparece al menos una clave candidata que es la clave primaria.
- Si aparecen varias claves candidatas, la elección de la clave primaria queda fuera de la definición del modelo relacional.
- En principio, las claves alternativas se asocian a conceptos derivados de la clave primaria.
- El concepto de clave primaria tiene mayor importancia sobre las relaciones base.
- Como las claves son conjunto de atributos, lo más correcto sería representarlas con la terminología de conjuntos, es decir, con { }.
- La utilización de un índice o algún otro tipo de elemento para permitir un acceso rápido, queda fuera del modelo relacional.
- En la práctica se suele incluir alguna ruta de acceso especial asociada a la clave primaria.
- No es bueno definir una clave primaria como simplemente clave, ya que puede llevar a confusiones no deseadas.
- En primer lugar porque existen diferentes tipos de claves asociadas al modelo relacional.
- Además, el término clave se suele utilizar en la práctica para referirse a productos que manejan índices, aunque no tengan nada que ver con una clave primaria.

CLAVES PRIMARIAS

Importancia

- Las claves primarias son fundamentales para el funcionamiento del sistema relacional.
- Permiten localizar las tuplas de una relación, de modo similar a como las direcciones permiten localizar los datos en la memoria.
- Si se elimina el concepto de clave primaria el sistema resultante puede presentar diferentes anomalías.

Notas Finales

- Debido al manejo de las claves primarias, parece que el modelo relacional requiere un Direccionamiento Asociativo.
- Dicha afirmación es sólo válida a nivel lógico.
- A nivel físico, no es necesaria la utilización de un esquema de memoria asociativo.
- Los dispositivos de E/S y las estructuras de almacenamiento convencionales permiten una correcta implementación del nivel físico.

REGLA DE INTEGRIDAD DE ENTIDADES

Definición

- Ningún Componente de la Clave Primaria de una Relación Base puede aceptar Nulos.
- En el modelo relacional Nulo se refiere a ausencia de valor.
- En SQL dicha definición varía ligeramente.

Justificación

- Las tuplas de una relación base corresponden a entidades de la realidad.
- Las entidades se identifican de modo único en la realidad, y también deben de poderse identificar en el modelo relacional.
- Dicha identificación es realizada por las claves primarias.
- Si una clave primaria contiene un nulo quiere decir que no se puede aplicar la definición de clave primaria sobre la entidad asociada.
- Por lo tanto, esta entidad no se puede identificar, lo que entra en contradicción con la definición de entidad.
- Este razonamiento también es válido para claves primarias parcialmente nulas.
- Así pues, la regla también se puede formular:
 - En una base de datos relacional, no se puede almacenar información sobre algo que no se pueda identificar.

REGLA DE INTEGRIDAD DE ENTIDADES

Comentarios

- Las claves primarias compuestas debe de ser no nulas en su totalidad, es decir, ninguno de sus atributos puede ser nulo.
- La regla de integridad de entidades se aplica sobre relaciones base.
- Otros tipos de relaciones pueden tener claves primarias que si admitan nulos.
- La regla no se aplica a las claves alternativas, aunque éstas deben de seguir cumpliendo la propiedad de unicidad.
- En algunos casos se confunde la regla de integridad de entidades con la propiedad de unicidad, cuando ambos conceptos son totalmente diferentes.

CLAVES AJENAS

Definición

- Una Clave Ajena es un conjunto de atributos de una relación R2 cuyos valores o son completamente nulos o coinciden con los de la clave primaria de una relación R1.
- De un modo gráfico se representa mediante una flecha entre las dos relaciones que se etiqueta por el conjunto de atributos que la definen,

$$R_2 \xrightarrow{\text{Cto.de Atributos}} R_1 .$$

- La definición de clave ajena no se cumple al revés, es decir, no todos los valores de la clave primaria deben aparecer como clave ajena.

Terminología

- Se dice que una clave ajena representa una Referencia a la tupla donde se encuentra el valor correspondiente de la clave primaria.
- La Restricción Referencial acota el dominio de la clave ajena a los valores actuales de la clave primaria.
- El problema de la Integridad Referencial trata de asegurar que la restricción referencial se cumpla.
- La relación donde se encuentra la clave ajena se denomina Relación Referencial.
- La Relación Referida u Objetivo indica la relación donde se encuentra la clave primaria.

CLAVES AJENAS

Comentarios

- La clave primaria y la clave ajena asociada deben de tomar valores en el mismo dominio.
- La clave ajena puede formar parte de la clave primaria de la relación a la que pertenece.
- Una relación referida puede ser también una relación referencial respecto de otro conjunto de atributos.
- De este modo se puede definir una Ruta Referencial entre diferentes relaciones de la base de datos,

$R_n \quad R_{n-1} \quad \dots \quad R_2 \quad R_1.$

- Si alguna de las relaciones aparece más de una vez en la ruta referencial se dice que aparecen Ciclos Referenciales,

$R_n \quad \dots \quad R_2 \quad R_1 \quad R_n.$

- La Relación Autorreferencial es un caso particular del anterior, donde coinciden la relación referida y la relación referencial
- Las claves ajenas pueden admitir nulos, a diferencia de las claves primarias.
- La integridad referencial permite conectar de modo coherente las diferentes entidades de la base de datos, definiendo interrelaciones entre las tuplas.
- Otro tipo de interrelaciones se definen entre los dominios de dos atributos, que al cambiar el esquema podría generar una clave ajena.

REGLA DE INTEGRIDAD REFERENCIAL

Definición

- La Base de Datos no debe contener Valores de Clave Ajena sin concordancia.
- Es decir, para cualquier valor no nulo de la clave ajena existe un valor asociado en la clave primaria de la relación objetivo.
- La regla se aplica a la relación existente entre clave ajena y clave primaria, y no entre clave ajena y clave candidata.
- La definición de clave ajena y de integridad referencial están estrechamente relacionadas de modo que al hablar de uno se referencia a la otra.

REGLAS PARA LAS CLAVES AJENAS

Estados de la Base de Datos

- La regla de integridad referencial se define sobre un estado de la base de datos.
- De modo que en un estado determinado se puede cumplir o no la regla.
- La regla no dice como se puede evitar la aparición de estados incorrectos.
- Una posibilidad sería que el sistema rechazara toda operación que violara la regla.
- Otra posibilidad, más interesante, es que el sistema aceptara estas operaciones, pero que actuara de modo conveniente para que se garantizara la transición a un estado legal.
- Es tarea del diseñador de la base de datos indicar que operaciones se deben rechazar y cuales requieren operaciones adicionales, u operaciones de compensación.
- Existen tres cuestiones que se debe plantear en este asunto:
 - ¿La clave ajena puede aceptar nulos?
 - ¿Qué ocurre cuando se realiza un borrado de la referencia de una clave ajena?
 - ¿Qué ocurre si se modifica la clave primaria del objetivo de la referencia de una clave ajena?
- De nuevo, todos estos aspectos quedan fuera del modelo relacional.

REGLAS PARA LAS CLAVES AJENAS

Claves Ajenas y los Nulos

- La aceptación de nulos en una clave ajena se debe ajustar a los requerimientos.
- Es decir, si se permite que la ruta referencial asociada a una tupla se rompa.

Borrado de Claves Primarias

- La decisión depende de los requerimientos, existiendo tres posibilidades:
 - Restringida (Restricted). No se puede borrar una tupla si está referenciada.
 - Se Propaga (Cascades). Se borra la tupla junto con las tuplas que la referencian.
 - Anula (Nullifies). Se borra la tupla y se pone a nulo la clave ajena de las tuplas que la referencian, siempre y cuando los acepte.

Modificación de Claves Primarias

- La decisión depende de los requerimientos, existiendo tres posibilidades:
 - Restringida (Restricted). Una tupla no se puede modificar si está referenciada.
 - Se Propaga (Cascades). Se modifica la tupla junto con las tuplas que la referencian.
 - Anula (Nullifies). Se modifica la tupla y se pone a nulo la clave ajena de las tuplas que la referencian, siempre y cuando los acepte.

REGLAS PARA LAS CLAVES AJENAS

Inclusión en el Modelo

- En la definición de una clave ajena se incluyen los atributos que lo forman y la relación objetivo.
- También será necesario especificar, para cada una de ellas, como contestar a las preguntas anteriores.
- Una posible sintaxis sería la siguiente:

```
FOREIGN KEY (clave_ajena) REFERENCES objetivo  
NULLS [ NOT ] ALLOWED  
DELETE OF objetivo efecto  
UPDATE OF clv_prim_objetivo efecto
```

donde se cumpliría que:

- La variable *clave_ajena* es el conjunto de atributos asociado a la clave que se está definiendo.
- La variable *objetivo* corresponde a la relación referida de la clave ajena.
- La variable *efecto* toma uno de los valores siguientes: RESTRICTED, CASCADES, NULLIFIES.
- La variable *clv_prim_objetivo* corresponde a la clave primaria de la relación referida.

REGLAS PARA LAS CLAVES AJENAS

Ejemplo de Definición de Claves Ajenas

```
CREATE TABLE S
( S#          DOMAIN (S#)          NOT NULL,
  SNOMBRE    DOMAIN (NOMBRE)       NOT NULL,
  SITUACION  DOMAIN (SITUACION)    NOT NULL,
  CIUDAD     DOMAIN (CIUDAD)       NOT NULL,
  PRIMARY KEY (S#) );
```

```
CREATE TABLE P
( P#          DOMAIN (P#)          NOT NULL,
  PNOMBRE    DOMAIN (NOMBRE)       NOT NULL,
  COLOR      DOMAIN (COLOR)        NOT NULL,
  PESO       DOMAIN (PESO)         NOT NULL,
  CIUDAD     DOMAIN (CIUDAD)       NOT NULL,
  PRIMARY KEY (P#) );
```

```
CREATE TABLE SP
( S#          DOMAIN (S#)          NOT NULL,
  P#          DOMAIN (P#)          NOT NULL,
  PRIMARY KEY (S#, P#),
  FOREIGN KEY (S#) REFERENCES S
    NULLS NOT ALLOWED
    DELETE OF S  CASCADES
    UPDATE OF S# CASCADES,
  FOREIGN KEY (P#) REFERENCES P
    NULLS NOT ALLOWED
    DELETE OF P  CASCADES
    UPDATE OF P# CASCADES);
```

REGLAS PARA LAS CLAVES AJENAS

Comentarios

- A parte de las opciones comentadas, existen otras que permiten asegurar la integridad referencial. Algunas de ellas son;
 - Preguntar al usuario.
 - Almacenar la información en otro fichero.
 - Modificación de la clave ajena.
- Para incluir otras opciones en el modelo la variable *efecto* podría ser la llamada a un procedimiento definido por el usuario.
- Las operaciones de borrado y modificación son atómicas pero pueden producir efectos laterales.
- La propagación del borrado de una tupla puede producir la eliminación de diferentes tuplas en las relaciones asociadas a su ruta referencial.
- Una situación similar se podría presentar en el proceso de modificación, cuando las claves ajenas formaran parte de la clave primaria.
- Cuando aparecen ciclos referenciales, para permitir la inserción de la primera fila, se debe de cumplir una de las condiciones siguientes,
 - Al menos una clave ajena del ciclo debe permitir nulos.
 - No se pueden verificar las restricciones hasta que no se haya completado la transacción.