

FICHEROS Y BASES DE DATOS (E44)

3º INGENIERÍA EN INFORMÁTICA

Tema 11.

Cálculo Relacional

- 1.- Introducción.
- 2.- Cálculo Relacional Orientado a Tuplas.
- 3.- Cálculo Relacional vs Álgebra Relacional: Algoritmo de Reducción de Codd.
- 4.- Ejemplo de Cálculo Relacional.
(Capítulo 14 del Date)

INTRODUCCIÓN

El Álgebra Relacional y el Cálculo Relacional

- El álgebra y el cálculo relacional son formas equivalentes de establecer la base formal de la parte manipulativa del modelo relacional.
- El álgebra construye una relación mediante la utilización de unos operadores.
- En el cálculo se definen las características de la relación, y el SGBD define las operaciones a aplicar para construir la relación.
- Para la consulta, "Número y ciudad de los proveedores que suministran la parte P2", el álgebra debería realizar,
 - Reunión natural de S y SP por S#.
 - Restringir el resultado para P# = P2.
 - Proyectar el resultado sobre S# y CIUDAD.

mientras que en el cálculo se formularía como,

- Obtener S# y CIUDAD para los proveedores tales que exista un envío en SP con el mismo valor en S# y con el valor P2 en P#.
- Se puede decir que el cálculo es descriptivo y que el álgebra es prescriptivo.
- También se puede decir que el cálculo plantea el problema y que el álgebra lo resuelve.
- Y de un modo muy informal, que el álgebra se fundamenta en procedimientos, y el cálculo no es de procedimientos

INTRODUCCIÓN

Panorama General del Cálculo

- El cálculo relacional se fundamenta en una rama de la lógica matemática denominada Cálculo de Predicados.
- Como ocurre entre el álgebra relacional y el matemático, el cálculo de predicados se ha modificado ligeramente para su utilización en el modelo relacional.
- Una característica fundamental del cálculo relacional es la denominada Variable de Tuplas o Variable de Recorrido.
- Una variable de tupla se asocia a una relación, y sus únicos valores permitidos son las tuplas de dicha relación.
- De tal modo que si T es una variable de tupla de relación R, en un instante determinado el valor de T corresponde con una tupla de R.
- En SQL, no hace falta definir explícitamente una variable de tuplas, aunque para entender el modo en el cuál se evalúa una sentencia es necesario utilizar variables de tupla.

```
SELECT S.S#  
FROM S  
WHERE S.CIUDAD = 'Londres' ;
```

- Si la consulta es compleja, sí resulta necesario definir variables de tuplas.
- Se le suele denominar como Cálculo Relacional Orientado a las Tuplas, para diferenciarlo del Orientado a los Dominios.

CÁLCULO DE TUPLAS

Gramática BNF

definición_de_recorrido

$::=$ RANGE OF variable IS
lista_con_comas_de_elems_de_recorrido

lista_con_comas_de_elems_de_recorrido

$::=$ elemento_de_recorrido |
elemento_de_recorrido ,
lista_con_comas_de_elems_de_recorrido

elemento_de_recorrido

$::=$ relación | (expresión)

expresión

$::=$ lista_con_comas_de_elems_de_objetivo
[WHERE fórmula_bien_formada]

lista_con_comas_de_elems_de_objetivo

$::=$ elemento_objetivo |
elemento_objetivo ,
lista_con_comas_de_elems_de_objetivo

elemento_objetivo

$::=$ [atributo =] variable . atributo | variable

fórmula_bien_formada

$::=$ comparación |
NOT fórmula_bien_formada |
comparación AND fórmula_bien_formada |
comparación OR fórmula_bien_formada |
IF comparación THEN
fórmula_bien_formada |
EXISTS variable [fórmula_bien_formada] |
FORALL variable [fórmula_bien_formada]

CÁLCULO DE TUPLAS

Comentarios sobre la Gramática

- Los paréntesis se definen como elementos del lenguaje, sin tener ningún significado adicional.
- En cambio, los corchetes indican opcionalidad y el carácter "|" marca la separación de las posibles interpretaciones de un símbolo.
- Las categorías "relación", "variable" y "atributo" son los Identificadores, es decir, son una categoría terminal de la gramática, y representan el nombre de una relación, de una variable de tuplas y de un atributo.
- La categoría "comparación" representa una fórmula bien formada, o bien una operación de comparación escalar sencilla, cuyos operandos son atributos o literales,

comparación

$::= \text{escalar comparador escalar}$

comparador

$::= < \mid <= \mid = \mid <> \mid >= \mid >$

escalar

$::= \text{atributo} \mid \text{literal}$

- En una comparación, un atributo se identifica como "variable . atributo".
- La categoría fórmula_bien_formada (fbf) se analiza completamente más adelante.
- Se definen modificaciones de la gramática para reducir el exceso de parentizado.

CÁLCULO DE TUPLAS

Variables de Tuplas

- Una Variable de Tupla se define mediante una proposición de la forma siguiente,

RANGE OF T IS X1, X2, . . . , Xn

donde

- T es la variable de tuplas definida.
- X1, X2, . . . , Xn son nombres de relación, o bien una expresión del cálculo de tuplas entre paréntesis.
- Si se define a Ri como la relación obtenida al evaluar la expresión Xi, entonces,
 - Las relaciones Ri deben ser compatibles respecto de la unión, es decir, todas tienen la misma cabecera.
 - La variable T toma valores en la unión de todas las relaciones Ri.
- Normalmente, una variable de tupla se asocia a una única relación, aunque la definición sea más general.
- En estos casos, el nombre de las variables de tupla se construye añadiendo una X, Y o Z al nombre de la relación asociada, tal y como sigue,

RANGE OF	SX	IS	S
RANGE OF	SY	IS	S
RANGE OF	SZ	IS	S
RANGE OF	PX	IS	P
RANGE OF	SPX	IS	SP

CÁLCULO DE TUPLAS

Concepto de Variables Libres y Acotadas

- Cada Ocurrencia, o Aparición, de una Variable de Tupla en una fórmula bien formada puede ser Libre o Acotada.
- Será Acotada si la variable se asocia a un cuantificador y la fórmula donde se encuentra la ocurrencia forma parte de la fórmula bien formada del cuantificador.

$$\text{EXISTS } x (x > 3)$$

- Será Libre si no es Acotada.

$$(x > 3)$$

- En una expresión, una misma variable de tupla puede aparecer como libre y como acotada en diferentes fórmulas bien formadas.

$$\text{EXISTS } x (x > 3) \text{ AND } (x > 0)$$

- En estos casos, se dice que la ocurrencia libre y la ocurrencia acotada se refieren a dos variables de tupla diferentes.
- Las ocurrencias acotadas se asocian a Variables Simuladas, y su único objetivo es conectar la expresión entre paréntesis con el cuantificador asociado.
- Por esta razón, se podría sustituir la ocurrencia de una variable acotada por la ocurrencia de otra variable, sin que con ello se cambie el sentido de la expresión,

$$\text{EXISTS } y (y > 3) \text{ AND } (x > 0)$$

CÁLCULO DE TUPLAS

Sintaxis de Variables Libres y Acotadas

- Una variable de tupla aparece en una fórmula bien formada como la variable simulada de un cuantificador, o como la referencia a uno de sus atributos, T.A, donde
 - T es una variable de tupla.
 - A es un atributo de las relaciones asociadas.
- En una ocurrencia sencilla, (T.A < U.A), todas las ocurrencias de variables de tupla son libres.
- En las fórmulas bien formadas (f) y NOT (f), las ocurrencias de variables de tupla son libres o acotadas según lo sean en f.
- En las fórmulas bien formadas f AND g y f OR g, las ocurrencias de variables de tupla son libres o acotadas según lo sean en f y g.
- Las ocurrencias de T que sean libres en f serán acotadas en EXISTS T (f) y en FORALL T (f). El resto de ocurrencias de variables de tupla son libres o acotadas según lo sean en f.
- El cuantificador EXISTS se lee como "existe", mientras que el cuantificador FORALL se lee como "para todos".
- El FORALL se puede expresar mediante EXISTS,

$$\text{FORALL } x (g) \quad \text{NOT EXISTS } x (\text{NOT } g)$$

- También existe la expresión IF - THEN, que no es primitiva, y cuya funcionalidad es la siguiente,

$$\text{IF } f \text{ THEN } g \quad (\text{NOT } f) \text{ OR } g$$

CÁLCULO DE TUPLAS

Listas Objetivo

- Una Lista Objetivo es una lista de Elementos Objetivo separador mediante comas, en la que cada elemento es una variable de tupla, o bien una expresión como la siguiente,

$$[X =] T.A$$

donde

- T es una variable de tupla, y A es uno de sus atributos asociados.
- X es un nombre de atributo para el atributo correspondiente en el resultado de evaluar la lista objetivo.
- Si se excluye X, el atributo correspondiente hereda el nombre del atributo de la variable de tupla utilizada.
- La referencia a X se utiliza para evitar la aparición de dos atributos en el resultado con el mismo nombre.
- Una variable de tupla puede aparecer como elemento objetivo, en cuyo caso representa todos los atributos de la relación asociada.
- Algunos ejemplos de listas objetivo aparecen a continuación,

SX.S#

SNUM = SX.S#

SX

SX.S#, SX.CIUDAD, SPX.P#

CÁLCULO DE TUPLAS

Expresiones

- Una expresión del cálculo de tuplas es una expresión de la forma,

lista_objetivo [WHERE f]

donde f es una fórmula bien formada.

- Una expresión que cumpla que
 - Referencie a las variables de tupla T, U, . . . , V.
 - Sean X1, X2, . . . , Xn el conjunto de atributos del resultado.

se evalúa como sigue,

- Se calcula el producto cartesiano de las relaciones de las variables de tupla, es decir, $T \times U \times \dots \times V$.
 - Si existe la condición WHERE, se eliminan las tuplas que no la cumplan.
 - Se proyecta el resultado sobre los atributos del resultado X1, X2, . . . , Xn.
- Algunos de los ejemplos de expresiones se muestran seguidamente,

SX.S#

SX.S# WHERE SX.CIUDAD = 'Londres'

SNUM = SX.S# WHERE SX.CIUDAD = 'Londres'

SX.S#, SX.CIUDAD

WHERE EXISTS SPX

(SPX.S# = SX.S# AND SPX.P# = 'P2')

CÁLCULO DE TUPLAS Y ÁLGEBRA

Equivalencia del Cálculo y del Álgebra

- El álgebra relacional y el cálculo de tuplas son dos formalismos equivalentes.
- Codd demostró esta equivalencia con su Algoritmo de Reducción, que genera una expresión del álgebra semánticamente equivalente a la expresión del cálculo de la que se parte.
- Con el objetivo de comprender el algoritmo, se analiza las fases de evaluación de una expresión del cálculo.
 - 1.- Para cada variable de tupla que aparece en la expresión del cálculo se obtiene la relación asociada, restringiendo las tuplas que aparecen si fuera posible.
 - 2.- Calcular el producto cartesiano de las relaciones resultantes.
 - 3.- Restringir el producto de acuerdo con las "condiciones de reunión" del WHERE.
 - 4.- Aplicar los cuantificadores de derecha a izquierda, del modo siguiente,
 - Los EXISTS producen una proyección que elimina los atributos de la relación asociada a la variable de tupla del cuantificador.
 - Los FORALL producen una división respecto de la relación asociada a la variable de tupla del cuantificador
 - 5.- Proyectar para obtener los atributos de la lista_objetivo de la expresión del cálculo.

CÁLCULO DE TUPLAS Y ÁLGEBRA

Algoritmo de Codd

- Las fases de la evaluación de una expresión del cálculo puede realizarse mediante la utilización del álgebra, tal y como se muestra a continuación.
- 1.- Obtener una relación intermedia para cada variable de tupla, como sigue,
 - Utilizar el operador UNION para calcular la unión de todas las relaciones sobre las que opera la variable de tupla.
 - Aplicación de una Restricción para eliminar las tuplas no deseadas.
- 2.- Cálculo del Producto Cartesiano de las relaciones intermedias.
- 3.- Aplicación de una Restricción sobre las condiciones de reunión.
- 4.- Inclusión de Proyecciones y Divisiones, con el objetivo de sustituir los cuantificadores.
- 5.- Cálculo de la Proyección que genera la relación resultado.
- De modo similar, se podría definir un algoritmo de conversión de una expresión del cálculo de tuplas a una expresión del álgebra.

EJEMPLO DEL ALGORITMO DE CODD

Relaciones Utilizadas

S

S#	SNOMBRE	SITUACION	CIUDAD
S1	Salazar	20	Londres
S2	Jaimes	10	París
S3	Bernal	30	París
S4	Corona	20	Londres
S5	Aldana	30	Atenas

SPJ

S#	P#	J#	CANT
S1	P1	J1	200
S1	P1	J4	700
S2	P3	J1	400
S2	P3	J2	200
S2	P3	J3	200
S2	P3	J4	200
S2	P3	J5	500
S2	P3	J6	600
S2	P3	J7	400
S2	P5	J2	800
S3	P3	J1	100
S3	P4	J2	200
S4	P6	J3	500
S4	P6	J7	300
S5	P2	J2	200
S5	P2	J4	100
S5	P5	J5	500
S5	P5	J7	100
S5	P6	J2	200
S5	P1	J4	100
S5	P3	J4	200
S5	P4	J4	800
S5	P5	J4	400
S5	P6	J4	500

P

P#	PNOMBRE	COLOR	PESO	CIUDAD
P1	Tuerca	Rojo	12	Londres
P2	Perno	Verde	17	París
P3	Birlo	Azul	17	Roma
P4	Birlo	Rojo	14	Londres
P5	Leva	Azul	12	París
P6	Engrane	Rojo	19	Londres

J

J#	JNOMBRE	CIUDAD
J1	Clasificador	París
J2	Perforadora	Roma
J3	Lectora	Atenas
J4	Consola	Atenas
J5	Compaginador	Londres
J6	Terminal	Oslo
J7	Cinta	Londres

EJEMPLO DEL ALGORITMO DE CODD

Enunciado

"Obtener Nombres y Ciudades de los Proveedores que suministran por lo menos 50 Unidades de cada parte a por lo menos un proyecto de Atenas"

Expresión del Cálculo de Tuplas

```
SX.SNOMBRE, SX.CIUDAD
WHERE EXISTS JX FORALL PX EXISTS SPJX
( JX.CIUDAD = 'Atenas' AND
  JX.J#     = SPJX.J#     AND
  PX.P#     = SPJX.P#     AND
  SX.S#     = SPJX.S#     AND
  SPJX.CANT >= 50 )
```

Resultado del Algoritmo de Codd

```
T1 := ( S RENAME S# AS SNUM, CIUDAD AS SCIUDAD )
T2 := ( P RENAME P# AS PNUM, CIUDAD AS PCIUDAD )
T3 := ( J RENAME J# AS JNUM, CIUDAD AS JCIUDAD )
T4 := ( SPJ WHERE CANT >= 50 )
T5 := ( T3 WHERE JCIUDAD = 'Atenas' )

T6 := ( T1 TIMES T2 TIMES T4 TIMES T5 )

T7 := ( T6 WHERE J# = JNUM AND P# = PNUM )
T8 := ( T7 WHERE S# = SNUM )

T9 := T8 [S , P , J]
T10 := ( T9 DIVIDEBY T2 )
T11 := T10[SNUM , SNOMBRE, SITUACION , SCIUDAD]
T12 := T11[SNOMBRE , SCIUDAD]
```

CÁLCULO DE TUPLAS Y ÁLGEBRA

Comentarios Finales

- El algoritmo de reducción de Codd, justifica la elección de los ocho operadores básicos definidos en el álgebra.
- Estos operadores son la base de un lenguaje objetivo para la realización de una expresión del cálculo.
- También se puede decir que un lenguaje es relacionalmente completo si sus expresiones permiten obtener cualquier relación definida en las expresiones del cálculo relacional.
- Pero el patrón de referencia suele ser el álgebra, ya que resulta más sencillo demostrar que un lenguaje posee operadores con la misma capacidad que las cinco primitivas del álgebra.
- El término relacionalmente completo se suele extender de modo que una sola expresión del cálculo puede expresarse mediante una sola expresión del lenguaje.
- De este modo, el álgebra y el cálculo son la base para el diseño de lenguajes que ofrezcan esta capacidad de expresión sin necesidad de utilizar bucles.
- También sería interesante que los lenguajes fueran completos computacionalmente, y para ello se incluyeron los operadores SUMMARIZE y EXTEND del álgebra.
- El cálculo también podría extenderse de modo similar.

EJEMPLO DE CÁLCULO RELACIONAL

Base de Datos del Video Club

Título de Películas tales que no existe un préstamo de una cinta de esa película.

```
PELICULAX.TITULO WHERE  
  NOT EXISTS PRESTAMOX ( EXISTS CINTAX  
    ( PRESTAMOX.CODCINTA = CINTAX.CODCINTA AND  
      PELICULAX.CODPELI = CINTAX.CODPELI ) )
```

Título de Películas tales que para todos los préstamos no existe una cinta de esa película que sea la cinta prestada.

```
PELICULAX.TITULO WHERE  
  FORALL PRESTAMOX ( NOT EXISTS CINTAX  
    ( PRESTAMOX.CODCINTA = CINTAX.CODCINTA AND  
      PELICULAX.CODPELI = CINTAX.CODPELI ) )
```

Título de Películas tales que no existe una cinta de esa película que haya sido prestada.

```
PELICULAX.TITULO WHERE  
  NOT EXISTS CINTAX  
    ( PELICULAX.CODPELI = CINTAX.CODPELI AND  
      EXISTS PRESTAMOX  
        ( PRESTAMOX.CODCINTA = CINTAX.CODCINTA ) )
```

Título de Películas tales que para todas las cintas de esa película no existe un préstamo que la incluya.

```
PELICULAX.TITULO WHERE  
  FORALL CINTAX  
    IF PELICULAX.CODPELI = CINTAX.CODPELI THEN  
      ( NOT EXISTS PRESTAMOX  
        ( PRESTAMOX.CODCINTA = CINTAX.CODCINTA ) )
```