

Álgebra Relacional

Dra. Amparo López Gaona

Posgrado en Ciencia e Ingeniería de la Computación
Fac. Ciencias, UNAM

Conjunto de operaciones usadas para manipular relaciones.

Estas operaciones toman relaciones como operandos y regresan relaciones que a su vez pueden ser manipuladas. \Rightarrow *MR es cerrado*.

Los operadores:

- Unión, diferencia, intersección. Con el significado usual en conjuntos, aplicado a relaciones.
- Selección. Selecciona ciertas tuplas de una relación.
- Proyección. Selecciona ciertas columnas de una relación.
- Productos y joins. Composición de relaciones.
- Renombrado de relaciones y atributos.

Selección

- Notación: $\sigma_{predicado}(R)$
- Selecciona tuplas que satisfacen un predicado dado.
Predicado: Operadores: ($>$, $<$, $>=$, $<=$, $=$, $<>$, \wedge , \vee , \neg)
Operandos: atributos o constantes.

- Relación R:

| A | B | C | D |
|---|---|----|----|
| a | a | 1 | 7 |
| a | b | 5 | 7 |
| b | b | 12 | 3 |
| b | b | 23 | 10 |

- $\sigma_{A=B \wedge D > 5}(R)$

| A | B | C | D |
|---|---|----|----|
| a | a | 1 | 7 |
| b | b | 23 | 10 |

Proyección

- Notación: $\pi_{\text{lista de atributos}}(R)$
- Es una tabla obtenida de R al eliminar los atributos no-especificados. En la tabla resultante aparecen los atributos en el mismo orden que en la lista.
- Los renglones duplicados se eliminan.

- Relación R:

| A | B | C | D |
|---|---|----|----|
| a | a | 1 | 7 |
| a | b | 5 | 7 |
| b | b | 12 | 3 |
| b | b | 23 | 10 |

- $\pi_{A,D}(R)$

| A | D |
|---|----|
| a | 7 |
| b | 3 |
| b | 10 |

- Notación: $R \cup S$
- Es la tabla que contiene las tuplas de la primera relación además de las tuplas de la segunda relación.
- Al adaptar los operadores de conjuntos a relaciones se debe asegurar que exista *compatibilidad* entre ellas.
 - Tienen el mismo grado.
 - Los atributos tienen el mismo nombre.
 - El dominio del atributo- i de R es el mismo que el dominio del atributo- i en $S, \forall i$

• Relación R:

| A | B |
|---|---|
| a | 1 |
| a | 2 |
| b | 1 |

S:

| A | B |
|---|---|
| a | 2 |
| b | 3 |

$R \cup S$

| A | B |
|---|---|
| a | 1 |
| a | 2 |
| b | 1 |
| b | 3 |

Diferencia

- Notación: $R - S$
- Crea una tabla con las tuplas que están en la relación R pero no en S.
- Operación válida entre relaciones compatibles.

• Relación R:

| A | B |
|---|---|
| a | 1 |
| a | 2 |
| b | 1 |

S:

| A | B |
|---|---|
| a | 2 |
| b | 3 |

• $R - S$

| A | B |
|---|---|
| a | 1 |
| b | 1 |

Producto Cartesiano

- Notación: $R \times S$
- Permite combinar información de cualquier par de relaciones.
 $R \times S = \{tq | t \in r \text{ and } q \in s\}$.
- Si R y S tienen atributos en común es necesario renombrarlos. Para evitar ambigüedades se precede el nombre del atributo con el nombre de la relación.

• Relación R:

| A | B |
|---|---|
| 1 | 2 |
| 3 | 4 |

S:

| B | C | D |
|---|----|----|
| 2 | 5 | 6 |
| 4 | 7 | 8 |
| 9 | 10 | 11 |

• $R \times S$

| A | R.B | S.B | C | D |
|---|-----|-----|----|----|
| 1 | 2 | 2 | 5 | 6 |
| 1 | 2 | 4 | 7 | 8 |
| 1 | 2 | 9 | 10 | 11 |
| 3 | 4 | 2 | 5 | 6 |
| 3 | 4 | 4 | 7 | 8 |
| 3 | 4 | 9 | 10 | 11 |

Join natural

- Notación: $R \bowtie S$

- $R (X_1, X_2, \dots, X_m, Y_1, Y_2, \dots, Y_n)$

$$S (Y_1, Y_2, \dots, Y_n, Z_1, Z_2, \dots, Z_p)$$

Relación con atributos X, Y, Z y poblado por el conjunto de tuplas que tienen igual valor de Y en R y en S.

$$\pi_{R \cup S}(\sigma_{R.Y_1=S.Y_1 \wedge R.Y_2=S.Y_2 \wedge \dots \wedge R.Y_n=S.Y_n}(R \times S))$$

- | A | B |
|---|---|
| 1 | 2 |
| 3 | 4 |

 \bowtie

| B | C | D |
|---|----|----|
| 2 | 5 | 6 |
| 4 | 7 | 8 |
| 9 | 10 | 11 |

 =

| A | B | C | D |
|---|---|---|---|
| 1 | 2 | 5 | 6 |
| 3 | 4 | 7 | 8 |

- Si las relaciones R y S no tienen nombres de atributos en común entonces $A \bowtie B \equiv A \times B$

Theta Join

- Notación: $R \bowtie_{\theta} S$
- Equivalente al Join sólo que se permite usar cualquier condición de comparación. (θ).
- El resultado se construye:
 - Toma el $R \times S$
 - Selecciona sólo las tuplas que satisfacen θ

• $R \bowtie_{A < D} S =$

| A | B | C |
|---|---|---|
| 1 | 2 | 3 |
| 6 | 7 | 8 |
| 9 | 7 | 8 |

| B | C | D |
|---|---|----|
| 2 | 3 | 4 |
| 2 | 3 | 5 |
| 7 | 8 | 10 |

| A | R.B | R.C | S.B | S.C | D |
|---|-----|-----|-----|-----|----|
| 1 | 2 | 3 | 2 | 3 | 4 |
| 1 | 2 | 3 | 2 | 3 | 5 |
| 1 | 2 | 3 | 7 | 8 | 10 |
| 6 | 7 | 8 | 7 | 8 | 10 |
| 9 | 7 | 8 | 7 | 8 | 10 |

Intersección

- Notación: $R \cap S = R - (R - S)$
- Relación con las tuplas que están en R y en S también.
- Operación válida entre relaciones compatibles.

• Relación R:

| A | B |
|---|---|
| a | 1 |
| a | 2 |
| b | 1 |

S:

| A | B |
|---|---|
| a | 2 |
| b | 3 |

• $R \cap S$

| A | B |
|---|---|
| a | 2 |

Renombrado

- Notación: $\rho_x(E)$
 $\rho_x(A_1, A_2, \dots, A_n)(E)$
- Asigna nombre a la relación y/o a los atributos. No se obtiene una nueva relación.

- Relación R:

| A | B | C | D |
|---|---|---|---|
| 1 | 2 | 5 | 6 |
| 3 | 4 | 7 | 8 |

- $\rho_{w,x,y,z}(R)$

| w | x | y | z |
|---|---|---|---|
| 1 | 2 | 5 | 6 |
| 3 | 4 | 7 | 8 |

Combinación de operadores

Álgebra =

- 1 Argumentos básicos +
- 2 Formas de construcción de expresiones.

Para el álgebra relacional:

- 1 Los argumentos son las variables que representan las relaciones + tuplas constantes.
- 2 Las expresiones se construyen aplicando los operadores y paréntesis si se requieren.

Consulta es una expresión del álgebra relacional.

Esquema de la Base de Datos

Cuenta (nombreSucursal, numCta, saldo)

Sucursal (nombreSucursal, ciudad, activos)

Cliente (nombreCliente, calle, ciudad)

CtaCliente (nombreCliente, numCta)

Prestamo (nombreSucursal, numPrestamo, importe)

Prestatario (nombreCliente, numPrestamo)

Ejemplos de consultas

- 1 Encontrar la información de todos los préstamos realizados en la sucursal llamada "Fuentes Brotantes".

$\sigma_{\text{nombreSucursal}=\text{"FuentesBrotantes"}}(\text{Préstamo})$

- 2 Determinar el nombre de los clientes que viven en Guanajuato.

$\pi_{\text{nombreCliente}}(\sigma_{\text{ciudad}=\text{"Guanajuato"}}(\text{Cliente}))$

- 3 Nombre de los clientes del banco que tienen una cuenta, un préstamo o ambas cosas.

a) $\pi_{\text{nombreCliente}}(\text{Prestatario})$

b) $\pi_{\text{nombreCliente}}(\text{CtaCliente})$

La respuesta es: $\pi_{\text{nombreCliente}}(\text{ctaCliente}) \cup \pi_{\text{nombreCliente}}(\text{Prestatario})$

- 4 Relación de clientes que tienen abierta una cuenta pero no tienen ninguna de préstamo.

$\pi_{\text{nombreCliente}}(\text{CtaCliente}) - \pi_{\text{nombreCliente}}(\text{Prestatario})$

- 5 Todos los clientes que tienen un préstamo y una cuenta abierta.

$\pi_{\text{nombreCliente}}(\text{Prestatario}) \cap \pi_{\text{nombreCliente}}(\text{CtaCliente})$

$\pi_{\text{nombreCliente}}(\text{Prestatario} \bowtie \text{CtaCliente})$

... Ejemplos de consultas

- 1 Nombre de los clientes que tienen un préstamo en la sucursal de Fuentes Brotantes.

$$\pi_{\text{nombreCliente}}(\sigma_{\text{Prestatario.numPréstamo}=\text{Préstamo.numPréstamo}}(\sigma_{\text{nombreSucursal}=\text{"FuentesBrotantes"}}(\text{Prestatario} \times \text{Préstamo})))$$

- 2 Nombre de los clientes que tienen un préstamo y el importe del mismo.

$$\pi_{\text{nombreCliente},\text{importe}}(\text{Prestatario} \bowtie \text{Préstamo})$$

- 3 Nombre de los clientes con préstamo mayor a cinco mil pesos.

$$\pi_{\text{nombreCliente}}(\sigma_{\text{importe}>5000}(\text{Préstamo} \bowtie \text{Prestatario}))$$

- 4 Nombre de los clientes que tienen una cuenta con saldo menor a \$3,000 y que no tienen préstamo.

$$\pi_{\text{nombreCliente}}(\text{CtaCliente} \bowtie (\sigma_{\text{importe}<3000}(\text{Cuenta}))) \\ - \pi_{\text{nombreCliente}}(\text{Prestatario})$$

Operaciones de Actualización (Borrado)

$R \leftarrow R - e$ con $e =$

- Una tupla constante especificada
- Resultado de una consulta en álgebra relacional.

Restricciones: Tupla con el mismo grado que la R y con valores en los dominios adecuados.

1. Borrar los préstamos con importe entre 0 y 8000 pesos.

$$R_1 \leftarrow \sigma_{importe \geq 0 \wedge importe \leq 8000}(Prestamo)$$
$$Prestamo \leftarrow Prestamo - R_1$$
$$Prestatario \leftarrow Prestatario -$$
$$\pi_{nombreCliente, numPrestamo}(Prestatario \bowtie R_1)$$

2. Borrar las cuentas de Guanajuato.

$$r_1 \leftarrow \sigma_{ciudad="Guanajuato"}(Cuenta \bowtie Sucursal)$$
$$r_2 \leftarrow \pi_{nombreSucursal, numCta, saldo}(r_1)$$
$$Cuenta \leftarrow Cuenta - r_2$$

Operaciones de Actualización (Inserción)

$R \leftarrow R \cup e$ con $e =$

- Una tupla constante especificada
- Resultado de una consulta en álgebra relacional.

Restricciones: Tupla con el mismo grado que la R y con valores en los dominios adecuados.

- 1 Insertar a Pérez con \$240,000 en la cuenta 973 de Cuernavaca
 $Cuenta \leftarrow Cuenta \cup \{("Cuernavaca", 973, 240000)\}$
 $CtaCliente \leftarrow CtaCliente \cup \{("Perez", 973)\}$
 $Cliente \leftarrow Cliente \cup \{("Perez", "Bugambilias", "Cuernavaca")\}$
- 2 Ofrecer un préstamo de \$40,000 a todos los clientes de Tijuana que tienen una cuenta en el banco. El número de la cta es el del préstamo.

$r_1 \leftarrow (\sigma_{ciudad="Tijuana"}(Cliente \bowtie CtaCliente) \bowtie Cuenta)$

$r_2 \leftarrow \pi_{nombreSucursal, numCta}(r_1)$

$Prestamo \leftarrow Prestamo \cup (r_2 \times \{40000\})$

$Prestatario \leftarrow Prestatario \cup \pi_{nombreCliente, numPrestamo}(r_1)$

Operaciones de Actualización (Actualización)

Borrado + Inserción.

- 1 Disminuir al saldo de la cuenta C973 de Pérez el 10 % .

$$r_1 \leftarrow \sigma_{numCta=C973}(Cuenta)$$

$$Cuenta \leftarrow Cuenta - r_1$$

$$r_2 \leftarrow \pi_{saldo*1,10}(r_1)$$

$$Cuenta \leftarrow Cuenta \cup \{("Cuernavaca", C973) \times r_2\}$$

- 2 Aumentar todos los saldos en un 5 % .

$$Cuenta \leftarrow \pi_{nombreSucursal,numCta,saldo*1,05}(Cuenta)$$

- 3 Disminuir 6 % a las cuentas con saldo mayor que 20000 y a las demás 5 %.

$$r_1 \leftarrow \pi_{nombreSucursal,numCta,saldo*1,06}(\sigma_{saldo>20000}(Cuenta))$$

$$r_2 \leftarrow \pi_{nombreSucursal,numCta,saldo*1,05}(\sigma_{saldo\leq 20000}(Cuenta))$$

$$Cuenta \leftarrow r_1 \cup r_2$$

Funciones de Agregación

Permiten combinar tuplas de una relación para producir un valor “agregado” .

Las más comunes son: `sum`, `avg`, `count`, `min`, `max`.

Ejemplo: $sum_{atributo}(R)$ crea una relación con un único atributo que es la suma de los valores de atributo en la relación R.

Álgebra relacional como lenguaje para definir restricciones

El tercer aspecto importante de un MD es la habilidad de restringir los datos almacenados en la BD.

Formas en álgebra relacional para expresar restricciones:

- Sea R una expresión en álgebra relacional(AR), entonces $R = \emptyset$ es una restricción que dice “no hay tuplas en el resultado de R ”.
- Si R y S son expresiones en AR, entonces $R \subseteq S$.

Ambas expresiones son equivalentes:

- $R \subseteq S$ es equivalente a $R - S = \emptyset$.
- $R = \emptyset$ puede escribirse como $R \subseteq S$

Restricciones de integridad referencial

Aseguran que un valor que aparece en un contexto también aparece en otro contexto relacionado.

En general, si un valor v para un atributo A de alguna tupla en una relación R , entonces se espera que v también aparezca en un atributo B de alguna tupla de otra relación S .

En álgebra se expresa como $\pi_A(R) \subseteq \pi_B(S)$ o de manera equivalente $\pi_A(R) - \pi_B(S) = 0$.

... Restricciones de integridad referencial

Ejemplos:

- Departamento(nombre, numDepto, CURPJefe, fechaIngresoJefe)
Proyecto(nombre, claveProy, ubicacionP, claveDepto)
Es válido suponer que la clave de departamento en el proyecto sea la de uno existente.
Esta restricción se expresa en álgebra como:

$$\pi_{claveDepto}(Proyecto) \subseteq \pi_{numDepto}(Departamento)$$

- TrabajaEn(tituloPeli, anioPeli, trabajador)
Pelicula(titulo, anio, duracion, genero, nombreEstudio, CURPPro)
Se desea asegurar que cada película mencionada en la primera relación también aparezca en la segunda.

$$\pi_{tituloPeli,anioPeli}(TrabajaEn) \subseteq \pi_{titulo,anio}(Pelicula)$$

Restricciones de llave

Departamento(nombre, numDepto, CURPJefe, fechaIngresoJefe)

Significa que no hay dos departamentos con el mismo número.

$D1 \leftarrow \rho_{D1}(\text{Departamento})$

$D2 \leftarrow \rho_{D2}(\text{Departamento})$

$\sigma_{D1.\text{numDepto}=D2.\text{numDepto} \text{ AND } D1.\text{nombre} \neq D2.\text{nombre}}(D1 \times D2) = \emptyset$

Restricciones de dominio

Se desea restringir el valor de sexo para los empleados:

$$\sigma_{\text{sexo} \neq 'F' \text{ AND } \text{sexo} \neq 'M'}(\text{Empleado}) = \emptyset$$

Otras restricciones

Se desea restringir el sueldo de un ejecutivo a ser mínimo \$100,000.00

Ejecutivo(nombre, direccion, CURP, ganancias)

Estudio(nombre, direccion, CURPPres)

$$\sigma_{ganancias < 100000}(Estudio \bowtie_{CURP=CURPPres} Ejecutivo) = \emptyset$$

O de manera equivalente

$$\pi_{CURPPres}(Estudio) \subseteq \pi_{CURP}(\sigma_{ganancias \geq 100000}(Ejecutivo))$$