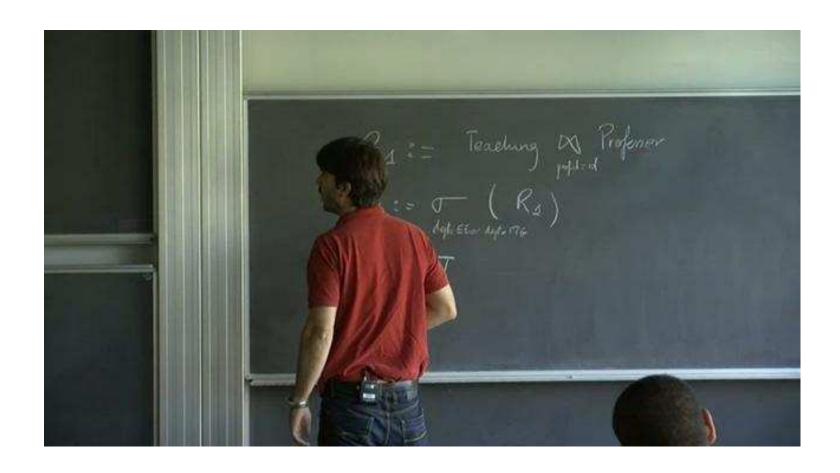
# Álgebra relacional

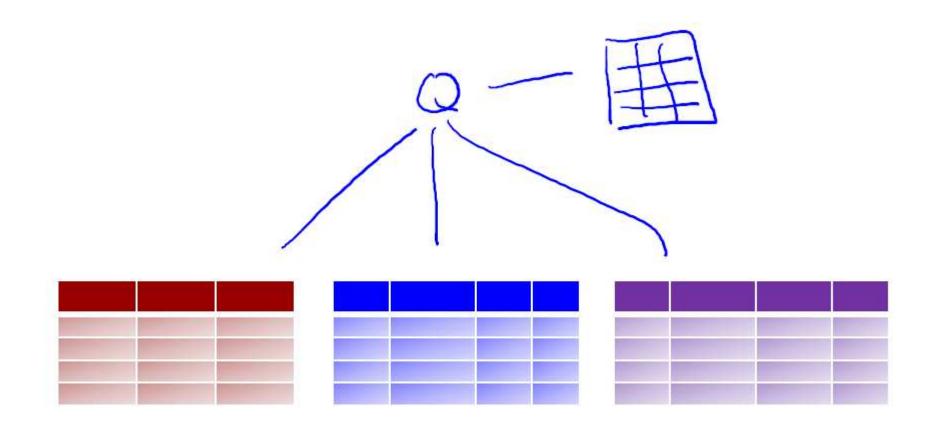
Ing. Sergio Molina Base de datos @UVG



Lenguaje formal



Query = consulta



Resultado de un query = relación

### Base de datos (universidades)

#### Definamos

- Universidad (uNombre, ubicación, total)
- Estudiante (eID, eNombre, promedio)
- Asignación (eID, uNombre, materia, nota)

Ţ	Jniversida	ıd		Estudiant	te	Asignación			
uNom	ubicac	total	Estudiante eID eNom prom			eID	uNom	mat	not

### Base de datos (universidades)

- Query trivial
  - El nombre de la relación: Estudiante
  - Me retorna la tabla de la relación Estudiante
  - Usar operadores para filtrar, partir y combinar

Ţ	Jniversida	ıd		Estudiant	te	Asignación			
uNom	ubicac	total	Estudiante eID eNom prom			eID	uNom	mat	not

### **Operador SELECT**

- Selecciona filas específicas: σ<sub>(condición)</sub> Relación
  - Estudiantes con promedio > 90  $\sigma_{prom > 90}$  Estudiante
  - Asignaciones de 'BD' en 'UVG'
     σ<sub>(uNom = 'UVG' ∧ mat = 'BD')</sub> Asignación

	Universidad			Estudiant	Asignación				
uNom	ubicac	total	eID	eNom	prom	eID	uNom	mat	not

### Operador PROJECT

- Selecciona columnas específicas:  $\pi_{A_1,..,A_n}$  Relación
  - ID y notas de todas las asignaciones.

	Universidad			Estudiant	Asignación				
uNom	ubicac	total	eID	eNom	prom	eID	uNom	mat	not

### Ambos operadores juntos

Seleccionar filas y columnas específicas:

$$\sigma_{(condición)}$$
 (Expresión)  $\pi_{A_1,..,A_n}$  (Expresión)

ID y nombre de estudiantes con promedio > 90

 $\pi_{eID, eNom} (\sigma_{prom > 90} Estudiante)$ 

	Universidad			Estudiant	Asignación				
uNom	ubicac	total	eID	eNom	prom	eID	uNom	mat	not

### Duplicados

- Algebra relacional (sets): se eliminan duplicados
- SQL (multi-sets, bags): no elimina duplicados
- Lista de materias con notas:  $\pi_{mat, not}$  Asignación

Ţ	<b>Jniversida</b>	ad		Estudiant	te	Asignación			
uNom	ubicac	total	Estudiante eID eNom prom			eID	uNom	mat	not

### Producto cartesiano (cross)

- Combina dos relaciones
  - El esquema resultante es la unión del esquema de las dos relaciones.
  - El contenido son todas las combinaciones de tuplas de esas relaciones.

Estudiante × Asignación

Ţ	Universidad			Estudiant	<u>Asignación</u>				
uNom	ubicac	total	eID	eID eNom prom			uNom	mat	not

### Producto cartesiano (cross)

 Nombres de estudiantes con promedio > 90 asignados a 'BD' y que sacaron > 95

 $\pi_{eNom}(\sigma_{(Estudiante.eID = Asignación.eID \land prom > 90 \land mat = 'BD' \land not > 95)}(Estudiante \times Asignación))$ 

	Universidad			Estudiant		Asignación			
uNom	ubicac	total	eID eNom prom			eID	uNom	mat	not

#### Natural Join

- Producto cartesiano con características
  - Refuerza igualdad en atributos con mismo nombre
  - Elimina una copia de atributos duplicados

	Universidad uNom ubicac total			Estudiant	Asignación eID uNom mat not				
uNom	ubicac	total	eID eNom prom		eID	uNom	mat	not	

#### Natural Join

• Nombres de estudiantes con promedio > 90 asignados a 'BD' y que sacaron > 95 en universidades con > 1,000 estudiantes.

 $\pi_{\mathsf{eNom}}(\sigma_{(\mathsf{prom}\ >\ 90\ \land\ \mathsf{mat}\ =\ \mathsf{'BD'}\ \land\ \mathsf{not}\ >\ 95\ \land\ \mathsf{total}\ >\ 1,000)}\ (\textit{Estudiante}\ \rhd\ \lhd\ (\textit{Asignación}\ \rhd\ \lhd\ \textit{Universidad}))$ 

	Universidad uNom ubicac total			Estudiant	Asignación				
uNom	ubicac	total	eID eNom prom		eID	uNom	mat	not	

#### Natural Join

Operación binaria (uso de paréntesis)

 $\textit{Expresion} \ 1 \ \trianglerighteq \ \triangleleft \ \textit{Expresion} \ 2 \ \equiv \pi_{\texttt{Esquema}} \ (\texttt{E}_1) \ \ U \ \ \texttt{Esquema} \ (\texttt{E}_2) \\ (\sigma_{(\texttt{E}_1,\texttt{A}_1 \ = \ \texttt{E}_2,\texttt{A}_1 \ \land \ \texttt{E}_1,\texttt{A}_2 \ = \ \texttt{E}_2,\texttt{A}_2 \ \land \ \ldots)} \\ (\texttt{Exp1} \ \times \ \texttt{Exp2}))$ 

- Útil si se tiene set up correcto de esquemas
- No da poder expresivo: conveniente en notación

Ur	Universidad		Estudiante			Asignación eID uNom mat not			
uNom	Nom ubicac total			eNom	prom	eID	uNom	mat	not

#### Theta Join

- Operación básica implementada en DBMS
  - Término JOIN usualmente se refiere a theta join

$$Exp_1 \triangleright \triangleleft_{\theta} Exp_2 \equiv \sigma_{\theta} (Exp_1 \times Exp_2)$$

Ţ	<b>Jniversida</b>	ad		Estudiant	te	Asignación			
uNom	ubicac	total	Estudiante eID eNom prom			eID	uNom	mat	not

### Operador Unión

- Lista de nombres de universidades y estudiantes
  - Los anteriores asocian horizontalmente
  - Aquí buscamos una asociación vertical
- Nota: Para unión se debe tener el mismo esquema (mismo nombre de atributo)
  - Aquí no es así pero luego lo corregiremos

 $\pi_{uNom}$  Universidad U  $\pi_{eNom}$ Estudiante

Universidad				Estudiant	Asignación				
uNom	ubicac	total	eID eNom prom			eID	uNom	mat	not

### Operador Diferencia

• IDs de estudiantes que no tienen asignada ninguna materia

 $\pi_{eID}$ Estudiante —  $\pi_{eID}$ Asignación

<u>Universidad</u>				Estudiant	Asignación				
uNom	ubicac	total	Estudiante eID eNom prom			eID	uNom	mat	not

### Operador Diferencia

- Ids y nombres de estudiantes que no tienen asignada ninguna materia
  - No podemos meterlo en el listado ya que no coinciden los atributos en ambos lados

 $\pi_{\text{eID. eNom}}((\pi_{\text{eID}}\textit{Estudiante} - \pi_{\text{eID}} Asignación) \rhd \lhd \text{Estudiante})$ 

Universidad				Estudiant	Asignación				
uNom	ubicac	total	eID	eNom	prom	eID	uNom	mat	not

### Operador Intersección

- Nombres que son tanto de universidad como de estudiantes
  - Ambas expresiones deben tener el mismo esquema

 $\pi_{uNom}$  Universidad  $\Pi$   $\pi_{eNom}$  Estudiante

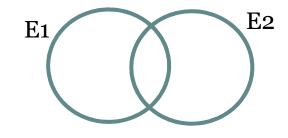
Universidad			Estudiante			Asignación			
uNom	ubicac	total	eID	eNom	prom	eID	uNom	mat	not

## Operador Intersección

No agrega poder expresivo (derivada)

$$E_1 \cap E_2 \equiv E_1 - (E_1 - E_2)$$

$$E_1 \cap E_2 \equiv E_1 \triangleright \triangleleft E_2$$
(mismo esquema)



Universidad			Estudiante			Asignación			
uNom	ubicac	total	eID eNom prom			eID	uNom	mat	not

### Operador Renombrar PR(A1,..., An)(E)

- Unificar esquemas para operadores (antes)
  - Lista de nombres de universidades y estudiantes  $\rho_{U(nombre)}(\pi_{uNom}Universidad)$  U  $\rho_{U(nombre)}(\pi_{eNom}Estudiante)$
- Quitar ambigüedad en "self-joins" (R join R)
  - Pares de universidades en la misma ubicación

 $\sigma_{ubicac = ubicac}(Universidad \times Universidad) \longrightarrow \sigma_{s1 = s2}(\rho_{C_1(n1,u1,t1)}(Universidad) \times \rho_{C_2(n2,u2,t2)}(Universidad))$   $\sigma_{n1 < n2}(\rho_{C_1(n1,u,t1)}(Universidad) \rhd \lhd \rho_{C_2(n2,u,t2)}(Universidad))$ 

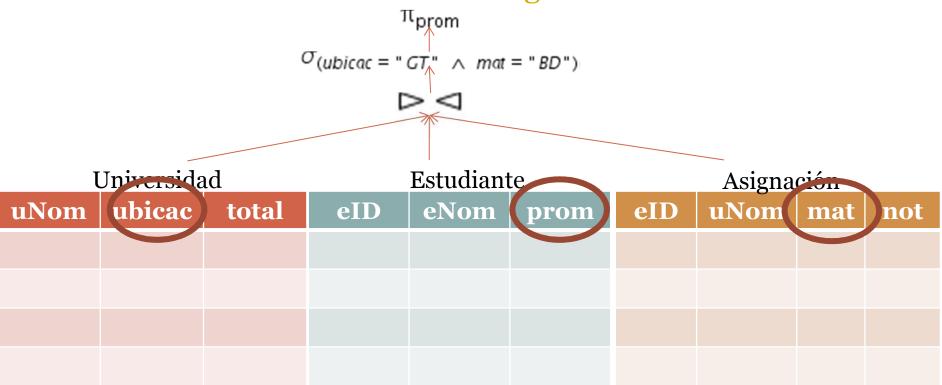
### Notación alternativa

- Asignaciones
  - Parejas de universidades en misma ubicación
    - $C_1$ : =  $\rho_{n1,u,t1}$ Universidad
    - $C_2$ : =  $\rho_{n2,u,t2}$  Universidad
    - $CP: = C_1 \triangleright \triangleleft C_2$
    - Ans: =  $\sigma_{(n_1 < n_2)} CP$

Universidad			Estudiante			Asignación			
uNom	ubicac	total	eID eNom prom			eID	uNom	mat	not

#### Notación alternativa

- Árbol de expresiones: visualizar la estructura
  - Promedios de estudiantes asignados a "BD" en "GT"



### En conclusión...

- Base
  - Relación, SELECT, PROJECT, producto cartesiano, unión, diferencia y renombrar.
- Abreviaciones (de las bases)
  - Natural join, theta join e intersección

Universidad				Estudiant	Asignación				
uNom	ubicac	total	Estudiante eID eNom prom			eID	uNom	mat	not