Ayudantía 1

Álgebra Relacional y Consultas de Primer Orden

Benjamín Saavedra — IIC2413 — Bases de Datos

Notación (recordatorio)

- Para una relación R, $\mathsf{att}(R)$ denota su conjunto de atributos.
- **Operadores base del álgebra relacional** (para relaciones R, S, atributos A, B y constante $a \in \mathbf{Data}$):
  - $R \cup S,\ R \cap S,\ R \setminus S$ (definidos si $\mathsf{att}(R) = \mathsf{att}(S)$)
  - $\sigma_{A=a}(R),\ \sigma_{A=B}(R)$, y su generalización $\sigma_\theta(R)$ con θ fórmula booleana de igualdades $(\wedge, \vee, \neg)$, además de $\sigma_{A\sim a}, \sigma_{A\sim B}$ con $\sim \in \{<, >, \leq, \geq\}$.
  - $\rho_{A\to B}(R)$ (definido si $A \in \mathsf{att}(R)$, $B \notin \mathsf{att}(R)$)
  - $\pi_{B_1,\ldots,B_k}(R)$
  - $R \bowtie S$ (join natural: combina tuplas *compatibles*, i.e. que coinciden en $\mathsf{att}(R) \cap \mathsf{att}(S)$; se cumple $\mathsf{att}(R \bowtie S) = \mathsf{att}(R) \cup \mathsf{att}(S)$)
  - $R \times S$ (caso particular de $\bowtie$ cuando $\mathsf{att}(R) \cap \mathsf{att}(S) = \varnothing$)
- **Consultas de primer orden**: $Q(x_1, \ldots, x_n) \leftarrow \varphi$, donde φ es una fórmula con $\mathsf{FV}(\varphi) \subseteq \{x_1, \ldots, x_n\}$ construida a partir de $x = a$, $x = y$, $R(t_1, \ldots, t_k)$ mediante $\wedge, \vee, \neg, \exists x, \forall x$.
- Recuerde que $\forall x.\, \varphi \equiv \neg\exists x.\, \neg\varphi$: esto es clave para traducir consultas del tipo *"para todos"* / *"ninguno"*.

Consejos generales

- Antes de escribir la fórmula/expresión, identifique *qué relaciones necesita combinar* y con qué atributo se conectan (llave que comparten).
- Para consultas del tipo *"que tengan dos o más ..."* necesitará comparar **dos testigos distintos** de una misma relación (dos variables existenciales con $\neg(\cdot = \cdot)$ en lógica; dos copias renombradas de la relación con ρ en álgebra).
- Para consultas del tipo *"que NO tengan ningún ..."* necesitará una negación de un existencial en lógica, o una diferencia $\setminus$ en álgebra.
- Para consultas del tipo *"para TODOS los ..."* piense en la equivalencia $\forall x.\varphi \equiv \neg\exists x.\neg\varphi$; en álgebra esto se traduce típicamente en dos diferencias anidadas sobre un "universo de candidatos" y una relación de evidencia (vea el Ejercicio A.1(h)).
- En álgebra, cuando necesite comparar dos tuplas de una misma relación, **renombre una copia con** ρ antes de aplicar $\bowtie$ o $\times$; de lo contrario los atributos compartidos forzarán una igualdad no deseada (o el operador quedará indefinido).

1 Parte A — Consultas en contexto

Recuerde el formato esperado: para cada consulta, entregue (i) una consulta de primer orden $Q(x_1, \ldots, x_n) \leftarrow \varphi$, y (ii) una expresión de álgebra relacional equivalente.

1.1 Contexto: Cine

Un cine registra información sobre sus películas, salas y funciones en el siguiente esquema:

$$\begin{aligned}
\textbf{Pelicula}(\mathsf{pid}, \mathsf{titulo}, \mathsf{duracion}) \quad & \mathsf{ar} = 3 \\
\textbf{Genero}(\mathsf{pid}, \mathsf{genero}) \quad & \mathsf{ar} = 2 \\
\textbf{Sala}(\mathsf{sid}, \mathsf{tipo}, \mathsf{capacidad}) \quad & \mathsf{ar} = 3 \\
\textbf{Funcion}(\mathsf{pid}, \mathsf{sid}, \mathsf{dia}, \mathsf{hora}) \quad & \mathsf{ar} = 4 \\
\textbf{Confiteria}(\mathsf{sid}, \mathsf{item}, \mathsf{precio}) \quad & \mathsf{ar} = 3
\end{aligned}$$

donde **Genero** es multivaluada (una película puede tener varios géneros), **Funcion**$(pid, sid, dia, hora)$ indica que la película *pid* se exhibe en la sala *sid* el día *dia* a la hora *hora*, y **Confiteria**$(sid, item, precio)$ indica que en la sala *sid* se vende el ítem de comida *item* a *precio* (puede variar entre salas).

Instancia de ejemplo (opcional, para verificar sus respuestas):

Pelicula

pid	titulo	duracion
1	Matrix	136
2	Coco	105
3	Nosferatu	132
4	Intensamente	95

Genero

pid	genero
1	Accion
1	CienciaFiccion
2	Animacion
2	Comedia
3	Terror
4	Animacion

Sala

sid	tipo	capacidad
1	IMAX	200
2	2D	100
3	3D	150

Funcion

pid	sid	dia	hora
1	1	Viernes	20:00
1	2	Sabado	18:00
2	2	Viernes	16:00
3	1	Sabado	22:00
4	3	Viernes	15:00

Confiteria

sid	item	precio
1	Palomitas	3500
1	Bebida	2000
2	Palomitas	2800
2	Nachos	3000
3	Bebida	2200

Ejercicio A.1. Para cada consulta, entregue (i) una consulta de primer orden y (ii) una expresión de álgebra relacional equivalente.

(a) Obtenga el título y la duración de todas las películas del género "Terror".

(b) Obtenga los títulos de las películas que tienen alguna función en una sala de tipo "IMAX".

(c) Obtenga el identificador y el tipo de las salas donde se vende el ítem "Palomitas".

(d) Obtenga los títulos de las películas que pertenecen a dos o más géneros distintos.

(e) Obtenga los identificadores de las salas que no tienen ninguna función programada.

(f) Obtenga los títulos de las películas que nunca se han exhibido (no tienen ninguna función).

(g) Obtenga los títulos de las películas que se exhiben tanto en la sala 1 como en la sala 2.

(h) *(Desafío)* Obtenga los identificadores de las salas donde se exhiben **todas** las películas del género "Terror".

1.2 Contexto: Universidad

Una universidad registra la siguiente información:

$$\begin{aligned} \textbf{Estudiante}(\mathsf{rut}, \mathsf{nombre}, \mathsf{carrera}) \quad & \mathsf{ar} = 3 \\ \textbf{Curso}(\mathsf{codigo}, \mathsf{nombre}, \mathsf{creditos}) \quad & \mathsf{ar} = 3 \\ \textbf{Inscribe}(\mathsf{rut}, \mathsf{codigo}, \mathsf{semestre}) \quad & \mathsf{ar} = 3 \\ \textbf{Dicta}(\mathsf{codigo}, \mathsf{profesor}) \quad & \mathsf{ar} = 2 \\ \textbf{Requisito}(\mathsf{codigo}, \mathsf{prerrequisito}) \quad & \mathsf{ar} = 2 \end{aligned}$$

donde **Inscribe**($rut, codigo, semestre$) indica que el estudiante rut se inscribió en el curso $codigo$ en el semestre $semestre$; **Dicta**($codigo, profesor$) es multivaluada (un curso puede tener más de un profesor); y **Requisito**($codigo, prerrequisito$) indica que para inscribir $codigo$ se necesita haber aprobado $prerrequisito$ (ambos son códigos de curso).

Instancia de ejemplo (opcional):

Estudiante

rut	nombre	carrera
1	Ana	IngCivil
2	Ben	Computacion
3	Zoe	Computacion

Curso

codigo	nombre	creditos
IIC2413	BasesDeDatos	10
IIC2233	ProgAvanzada	10
IIC1103	IntroProg	10

Inscribe

rut	codigo	semestre
1	IIC1103	2025-1
2	IIC2233	2025-1
2	IIC2413	2025-2
3	IIC2413	2025-2
3	IIC2233	2024-2

Dicta

codigo	profesor
IIC2413	Riveros
IIC2413	Soto
IIC2233	Riveros
IIC1103	Perez

Requisito

codigo	prerrequisito
IIC2413	IIC2233
IIC2233	IIC1103

Ejercicio A.2. Para cada consulta, entregue (i) una consulta de primer orden y (ii) una expresión de álgebra relacional equivalente.

(a) Obtenga los nombres de los estudiantes de la carrera "Computacion".

(b) Obtenga los nombres de los cursos dictados por el profesor "Riveros".

(c) Obtenga los nombres de los estudiantes inscritos en el curso "IIC2413".

(d) Obtenga los nombres de los estudiantes que están inscritos tanto en "IIC2413" como en "IIC2233" (no necesariamente en el mismo semestre).

(e) Obtenga los nombres de los cursos que no tienen ningún prerrequisito.

(f) Obtenga los nombres de los estudiantes que no están inscritos en ningún curso.

(g) Obtenga los nombres de los cursos dictados por más de un profesor.

(h) *(Desafío)* Obtenga los nombres de los estudiantes que están inscritos en **todos** los cursos que dicta el profesor "Riveros".

2 Parte B — Nuevos operadores del álgebra relacional

El álgebra relacional que vimos en clases se construye a partir de un conjunto *base* de operadores: $\cup, \cap, \setminus, \sigma, \rho, \pi, \bowtie$. En la práctica es común definir operadores adicionales que faciliten escribir consultas. Para que un operador nuevo sea parte del álgebra relacional (y no la extienda a un lenguaje más expresivo), debemos demostrar que puede *simularse* usando sólo los operadores base: es decir, dar una expresión equivalente y demostrar la igualdad de las relaciones resultantes (típicamente por doble contención, o por "si y sólo si").

Ejercicio B.1 (semijoin). Para relaciones R, S se define el *semijoin* de R con S, denotado $R \ltimes S$, como la relación con $\mathsf{att}(R \ltimes S) = \mathsf{att}(R)$ tal que

$$R \ltimes S \;:=\; \{\bar{\tau} \in R \mid \exists \bar{\sigma} \in S.\ \bar{\tau} \sim \bar{\sigma}\}$$

(recuerde que $\bar{\tau} \sim \bar{\sigma}$ significa que $\bar{\tau}$ y $\bar{\sigma}$ son *compatibles*: coinciden en los atributos que comparten). Es decir, $R \ltimes S$ son las tuplas de R que tienen al menos un "compañero" compatible en S. Proponga una expresión equivalente en términos de $\pi, \bowtie$, y demuéstrela.

Ejercicio B.2 (anti-join). Se define el *anti-join* de R con S, denotado $R \rhd S$, como la relación con $\mathsf{att}(R \rhd S) = \mathsf{att}(R)$ tal que

$$R \rhd S \;:=\; \{\bar{\tau} \in R \mid \nexists \bar{\sigma} \in S.\ \bar{\tau} \sim \bar{\sigma}\}$$

es decir, las tuplas de R que **no** tienen ningún compañero compatible en S.

(a) Demuestre que $R \rhd S = R \setminus (R \ltimes S)$.

(b) Use $\rhd$ para reescribir de forma más compacta su respuesta a la consulta A.1(e) ("salas sin función").

3 Parte C — Propiedades del álgebra relacional

Para cada propiedad: si es **verdadera**, demuéstrela (por extensión de la definición de cada operador, mostrando que ambos lados tienen los mismos atributos y las mismas tuplas); si es **falsa**, entregue relaciones concretas $R, S, \ldots$ (con sus atributos y algunas tuplas) donde ambos lados de la igualdad den resultados distintos.

C.1. (Conmutatividad de la selección) Para toda relación R y fórmulas θ_1, θ_2 (cuyos atributos pertenecen a $\mathsf{att}(R)$):

$$\sigma_{\theta_1}\big(\sigma_{\theta_2}(R)\big) \;=\; \sigma_{\theta_2}\big(\sigma_{\theta_1}(R)\big) \;=\; \sigma_{\theta_1 \wedge \theta_2}(R).$$

C.2. (Proyección distribuye sobre la unión) Sean R, S con $\mathsf{att}(R) = \mathsf{att}(S)$:

$$\pi_{B_1,\ldots,B_k}(R \cup S) \;=\; \pi_{B_1,\ldots,B_k}(R) \cup \pi_{B_1,\ldots,B_k}(S).$$

C.3. (Proyección y diferencia) Sean R, S con $\mathsf{att}(R) = \mathsf{att}(S)$. Demuestre o refute:

$$\pi_{B_1,\ldots,B_k}(R \setminus S) \;=\; \pi_{B_1,\ldots,B_k}(R) \setminus \pi_{B_1,\ldots,B_k}(S).$$

C.4. (Conmutatividad y asociatividad del join) Para relaciones R, S, T cualesquiera:

$$R \bowtie S = S \bowtie R \qquad \text{y} \qquad (R \bowtie S) \bowtie T = R \bowtie (S \bowtie T).$$