Ayudantía 5

Benjamín Saavedra — IIC2413 — Bases de Datos

Notación y algoritmos (recordatorio)

Dependencias funcionales y llaves.

- Para una relación R, $\mathsf{att}(R)$ denota su conjunto de atributos.
- R **satisface** $\bar{A} \to \bar{B}$ (notado $R \models \bar{A} \to \bar{B}$) si para toda $\bar{\tau}_1, \bar{\tau}_2 \in R$: si $\pi_{\bar{A}}(\bar{\tau}_1) = \pi_{\bar{A}}(\bar{\tau}_2)$ entonces $\pi_{\bar{B}}(\bar{\tau}_1) = \pi_{\bar{B}}(\bar{\tau}_2)$.
- $\bar{A} \to \bar{B}$ es **trivial** si $\{\bar{B}\} \subseteq \{\bar{A}\}$ (siempre se cumple).
- $\bar{A}$ es **superllave** de R si $R \models \bar{A} \to \bar{A}^c$, donde $\{\bar{A}^c\} = \mathsf{att}(R) \setminus \{\bar{A}\}$: es decir, $\bar{A}$ determina *todos* los demás atributos.
- $\bar{A}$ es **llave** de R si es una superllave **minimal**: ningún subconjunto propio de $\{\bar{A}\}$ es superllave.

Clausura. La clausura $\bar{A}^+$ de $\bar{A}$ según $\mathcal{F}$ es $\{B \mid \mathcal{F} \models \bar{A} \to B\}$.

Algoritmo 1 (clausura). `input`: $\mathcal{F}$, $\bar{A}$.

1. reglas $\leftarrow \mathcal{F}$; clausura $\leftarrow \{\bar{A}\}$
2. **repeat**: si $(\bar{B} \to \bar{C}) \in$ reglas y $\{\bar{B}\} \subseteq$ clausura, entonces clausura $\leftarrow$ clausura $\cup\{\bar{C}\}$ y reglas $\leftarrow$ reglas $\setminus\{\bar{B} \to \bar{C}\}$
3. **until** no hay cambios; **return** clausura

Descomposiciones. Una descomposición de R en $\bar{A}$ y $\bar{B}$ (con $\mathsf{att}(R) = \{\bar{A}\} \cup \{\bar{B}\}$) son $S = \pi_{\bar{A}}(R)$, $T = \pi_{\bar{B}}(R)$. Es **sin pérdida** si $R = S \bowtie T$ (siempre se cumple $R \subseteq S \bowtie T$; falta la otra contención). Las dependencias **proyectadas** $\mathcal{F}_S$ son las DFs de $\mathcal{F}$ cuyos atributos (izquierdos y derechos) están todos en $\{\bar{A}\}$ (análogo para $\mathcal{F}_T$); la descomposición **preserva las dependencias** si $\mathcal{F}_S \cup \mathcal{F}_T \models \mathcal{F}$.

BCNF y 3NF.

- R está en **BCNF** si para toda DF no trivial $\bar{A} \to \bar{B}$ de R: $\bar{A}$ es superllave.
- Un atributo es **primario** si pertenece a alguna llave. R está en **3NF** si para toda DF no trivial $\bar{A} \to \bar{B}$: $\bar{A}$ es superllave, **o** todo atributo de $\bar{B}$ fuera de $\bar{A}$ es primario. BCNF $\Rightarrow$ 3NF (no vale el recíproco).
- **Algoritmo toBCNF**: mientras R no esté en BCNF, tome $\bar{A} \to \bar{B} \in \mathcal{F}$ con $\bar{A}$ no superllave; sea $\bar{A}^- := (\mathsf{att}(R) \setminus \bar{A}^+) \cup \{\bar{A}\}$; descomponga en $R^+ = \pi_{\bar{A}^+}(R)$ y $R^- = \pi_{\bar{A}^-}(R)$; recurse en cada una (con sus DFs proyectadas).
- Un conjunto $\mathcal{F}$ es **minimal** (cobertura mínima) si: (1) toda DF tiene un solo atributo a la derecha; (2) ninguna DF es redundante; (3) ningún atributo del lado izquierdo de ninguna DF es redundante.
- **Algoritmo de descomposición en 3NF**: dado $\mathcal{F}$ minimal, por cada $\bar{A} \to B \in \mathcal{F}$ cree la relación $R_{\bar{A}B} = \pi_{\bar{A},B}(R)$ con sus DFs proyectadas; agregue además $R_{\bar{K}}$ para una llave $\bar{K}$ de R.

Consejos generales

- Para decidir si $\bar{A}$ es (super)llave, calcule $\bar{A}^+$: es superllave si y sólo si $\bar{A}^+ = \mathsf{att}(R)$.
- Un atributo que **nunca** aparece al lado derecho de ninguna DF debe pertenecer a **toda** llave. Úselo para acotar rápidamente los candidatos a llave.
- Al aplicar toBCNF, recuerde que $\bar{A}^+$ se calcula con **todo** $\mathcal{F}$, no sólo con la DF elegida.

1 Parte A: Ejercicio integrador de diseño

1.1 Contexto: Arriendo de bicicletas públicas

Un sistema de bicicletas públicas registra sus arriendos en una única relación:

$$\textbf{PrestamoBici}(\mathsf{bicicleta}, \mathsf{estacion}, \mathsf{comuna}, \mathsf{usuario}, \mathsf{fecha}) \qquad \mathsf{ar} = 5$$

Se sabe que: cada bicicleta tiene una **estación base** fija; cada estación pertenece a una **comuna** fija; y, para simplificar, cada usuario arrienda **a lo más una** bicicleta por día (es decir, usuario y fecha juntos determinan qué bicicleta se usó).

Instancia de ejemplo (opcional):

bicicleta	estacion	comuna	usuario	fecha
B12	E5	Providencia	U1	2026-09-01
B12	E5	Providencia	U2	2026-09-02
B07	E5	Providencia	U1	2026-09-03
B30	E9	Ñuñoa	U3	2026-09-01

Ejercicio A.1.

(a) Escriba las DFs descritas en el enunciado usando la notación formal.

(b) Encuentre la(s) llave(s) de **PrestamoBici** (indique qué atributos deben estar en toda llave y por qué). *Hint:* el Ejercicio B.2 demuestra formalmente un atajo para esta parte: puede usarlo aquí sin volver a demostrarlo.

(c) ¿Está **PrestamoBici** en BCNF? ¿En 3NF? Indique todas las violaciones encontradas.

(d) Aplique el algoritmo toBCNF completo, paso a paso.

(e) Encuentre una cobertura mínima de las DFs y aplique el Algoritmo de descomposición en 3NF. Compare el resultado con (d): ¿son las mismas relaciones?

(f) En la relación original **PrestamoBici** (sin descomponer), dé un ejemplo concreto de anomalía de **actualización**: ¿qué pasa si la estación E5 cambia de comuna? ¿Se resuelve este problema con su descomposición?

2 Parte B: Propiedades

Ejercicio B.1 (¿cuándo es $\varnothing$ una llave?).

(a) Demuestre que $\varnothing$ es siempre **minimal**, trivialmente: $\varnothing$ no tiene ningún subconjunto propio, así que la condición de minimalidad ("ningún subconjunto propio es superllave") se cumple vacuamente. Concluya que "$\varnothing$ es llave" y "$\varnothing$ es superllave" son equivalentes.

(b) Suponga que $\mathcal{F}$ no contiene ninguna regla con lado izquierdo vacío. Demuestre, a partir del **mecanismo** del Algoritmo 1, que en ese caso $\varnothing^+ = \varnothing$. Concluya que $\varnothing$ es superllave de R si y sólo si $\text{att}(R) = \varnothing$.

Hint: mire el cuerpo del `repeat` con clausura $= \varnothing$: ¿qué regla $\bar{B} \to \bar{C}$ podría disparar, si toda regla de $\mathcal{F}$ tiene $\{\bar{B}\} \neq \varnothing$?

(c) Muestre, en cambio, que si $\text{att}(R) = \{A\}$ (un solo atributo), entonces $\{A\}$ (no $\varnothing$) sí es siempre llave de R, para cualquier $\mathcal{F}$.

Ejercicio B.2. Sea R una relación con conjunto de DFs $\mathcal{F}$, y sea $A \in \texttt{att}(R)$ un atributo que **nunca** aparece al lado derecho de ninguna DF de $\mathcal{F}$ (es decir, no existe $\bar{B} \to \bar{C} \in \mathcal{F}$ con $A \in \{\bar{C}\}$).

(a) Demuestre, a partir del **mecanismo** del Algoritmo 1 (no de la definición semántica de $\models$), que para todo conjunto de atributos $\bar{X}$:

$$A \notin \{\bar{X}\} \quad \Longrightarrow \quad A \notin \bar{X}^+.$$

Hint: mire el cuerpo del `repeat`: ¿de qué única manera puede un atributo entrar a `clausura` en una iteración? ¿Puede A entrar alguna vez, dado lo que se supone sobre $\mathcal{F}$?

(b) Concluya, tomando el contrarrecíproco de (a), que si $\bar{X}$ es una **superllave** de R, entonces $A \in \{\bar{X}\}$. (Es decir: A pertenece a toda superllave, y por lo tanto a toda llave, de R.)

Comentario. Esta es exactamente la herramienta que permite resolver el Ejercicio A.1(b) sin probar los 32 subconjuntos posibles de atributos: basta con mirar una vez los lados derechos de $\mathcal{F}$, sin calcular ninguna clausura.

Ejercicio B.3* (monotonía de la clausura).

(a) Demuestre, también a partir del **mecanismo** del Algoritmo 1 (no de la definición semántica), que para cualesquiera conjuntos de atributos $\bar{X}, \bar{Y}$ (según el mismo $\mathcal{F}$):

$$\{\bar{X}\} \subseteq \{\bar{Y}\} \quad \Longrightarrow \quad \bar{X}^+ \subseteq \bar{Y}^+.$$

Hint: corra el Algoritmo 1 sobre $\bar{X}$ y sobre $\bar{Y}$ "en paralelo". Cada vez que una regla $\bar{B} \to \bar{C}$ dispara para $\bar{X}$ (es decir, $\{\bar{B}\} \subseteq$ clausura de $\bar{X}$ en ese momento), argumente que esa misma regla también puede dispararse para $\bar{Y}$ en algún momento de su ejecución. Formalice esto por inducción en el número de atributos que se han agregado a la clausura de $\bar{X}$.

(b) *(Aplicación: ¿para qué sirve esto?)* Suponga que ya sabe que $\bar{X} = \{A, B, C\}$ es una **superllave** de una relación R, y quiere verificar si además es una **llave**. La definición exige revisar los $2^3 - 2 = 6$ subconjuntos propios no vacíos de $\bar{X}$ ($\{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}$) y confirmar que **ninguno** es superllave.

Usando el resultado de (a), explique por qué en realidad **basta con revisar sólo 3 de ellos** (los subconjuntos "maximales" que se obtienen quitando un *único* atributo a $\bar{X}$, es decir $\{B, C\}$, $\{A, C\}$ y $\{A, B\}$) para concluir lo mismo sobre los 6. En particular: si ya verificó que $\{A, B\}$ **no** es superllave, ¿por qué eso le permite descartar también a $\{A\}$ y a $\{B\}$ **sin** calcular $\{A\}^+$ ni $\{B\}^+$?

(c) *(Otra aplicación: por qué basta con revisar $\mathcal{F}$ para verificar BCNF).* Las definiciones de BCNF y 3NF hablan de **toda** DF que R satisface, es decir, toda DF que se **deriva** de $\mathcal{F}$ (formalmente, $\mathcal{F} \models \bar{X} \to Y$), no sólo las que están escritas explícitamente en $\mathcal{F}$. Sin embargo, el algoritmo de verificación de BCNF visto en clases sólo recorre las reglas de $\mathcal{F}$. Demuestre que esto es válido: si **ninguna** regla $\bar{B} \to \bar{C} \in \mathcal{F}$ tiene $\bar{B}$ no-superllave, entonces **ninguna** DF derivada $\bar{X} \to Y$ (con $\mathcal{F} \models \bar{X} \to Y$ y $Y \notin \{\bar{X}\}$) puede tener $\bar{X}$ no-superllave tampoco.

Hint: suponga, por contradicción, que sí existe una DF derivada $\bar{X} \to Y$ con $\bar{X}$ no superllave. Como $Y \in \bar{X}^+$, en algún momento de la ejecución del Algoritmo 1 sobre $\bar{X}$ se disparó **alguna** regla $\bar{B} \to \bar{C} \in \mathcal{F}$ (una regla real, de las que sí están escritas) que agregó Y a la clausura; en ese instante, $\{\bar{B}\} \subseteq$ clausura parcial $\subseteq \bar{X}^+$. Use el resultado de (a) para concluir que $\bar{B}^+ \subseteq (\bar{X}^+)^+$. Luego use que $(\bar{X}^+)^+ = \bar{X}^+$ (la clausura de un conjunto ya cerrado no puede seguir creciendo: todo lo que se podía derivar ya se derivó) para concluir que $\bar{B}^+ \subseteq \bar{X}^+ \neq \mathsf{att}(R)$, es decir, que $\bar{B}$ **tampoco** es superllave, contradiciendo la hipótesis, ya que $\bar{B} \to \bar{C}$ sí está en $\mathcal{F}$.

(d) *(Corolario: una forma de "aumentación").* Usando el resultado de (a), demuestre que si $R \models \bar{X} \to Y$ (con Y un único atributo) y $\{\bar{X}\} \subseteq \{\bar{Z}\}$, entonces $R \models \bar{Z} \to Y$. Es decir: si un lado izquierdo ya alcanza para determinar Y, cualquier lado izquierdo **más grande** también alcanza.

Hint: traduzca la hipótesis y la conclusión a clausuras usando la Proposición de la Clase 08 ($R \models \bar{A} \to B$ sii $B \in \bar{A}^+$), y aplique (a) directamente: no hace falta volver a correr el Algoritmo 1.

Comentario. Esta es la razón por la que, en el Ejercicio A.1(b), alcanzó con probar $\{\mathsf{usuario}\}$ y $\{\mathsf{fecha}\}$ por separado: como $|\{\mathsf{usuario}, \mathsf{fecha}\}| = 2$, los "subconjuntos maximales" (quitar un atributo) **son** directamente los únicos subconjuntos propios no vacíos que existen: no hubo ningún atajo que aplicar de más porque la lista ya era mínima desde el principio. Con una llave de 3 o más atributos, en cambio, la diferencia entre "revisar todos los subconjuntos propios" y "revisar sólo los maximales" sí se nota.