

CC42A – BASES DE DATOS
Profesores: Claudio Gutiérrez, Gonzalo Navarro
Auxiliar: Mauricio Monsalve

Auxiliar 4

- Dependencias Multivariadas
- Normalización
- Álgebra Relacional (Nociones)

Dependencias Multivaluadas:

Definición: $X \twoheadrightarrow Y$ (X multidetermina Y) significa que si $t_1[X]=t_2[X]$, entonces se cumple que existe t_3 tal que $t_3[X]=t_1[X]$, $t_3[Y]=t_1[Y]$ y $t_3[Z]=t_2[Z]$, donde Z es R-X-Y.

Problemas:

1. ¿Una dependencia funcional es multivaluada? Mostrarlo.
2. Sea la dep. mul. $X \twoheadrightarrow Y$ y sea Z en R-X-Y. Mostrar que $X \twoheadrightarrow Z$.
3. Sea $R=(A,B,C)$ con $A \twoheadrightarrow B$. Se sabe que (a,b,\$), (a,c,\$) y (a,d,%) están en R. Inmediatamente, ¿qué otras tuplas están en R?
4. La cuarta forma normal (4FN) exige que toda dependencia multivaluada no trivial $X \twoheadrightarrow Y$ debe ser tal que X es superllave. ¿A qué FN se parece? ¿Qué algoritmo de descomposición inmediato puede tener? ¿Qué quiere decir, realmente, 4FN?

Normalización:

Algoritmos:

- **FNBC:** Si en R, $X \twoheadrightarrow Y$ evita FNBC, partir R en R-Y y en $X \cup Y$. Aplicar recursivamente hasta lograr FNBC en cada esquema. Esto no asegura conservar dependencias, pero sí la reunión (no aditiva).
- **3FN:** Para cada $X \twoheadrightarrow Y$ en un recubrimiento minimal, hacer un esquema XY. Si ninguno posee una llave, entonces hacer una relación con esa llave. Asegura reunión y dependencias.

Problemas:

1. Normalizar $R=(A,B,C)$ con $F=\{AB \twoheadrightarrow C, C \twoheadrightarrow B\}$ en FNBC. ¿Qué problema ocurre?
2. Normalizar $R=(A,B,C,D,E)$ con $F=\{A \twoheadrightarrow C, B \twoheadrightarrow D, C \twoheadrightarrow B\}$.

Álgebra Relacional:

Problemas:

1. ¿Qué operadores relacionales son conmutativos?
2. ¿Qué operador relacional sirve para enlazar relaciones luego de una normalización? ¿Qué condición debe ocurrir sobre los algoritmos de descomposición?
3. Sea $R=(\text{Persona}, \text{Idioma})$ y $S=(\text{Idioma})$. ¿Quiénes hablan todos los idiomas? (¿Qué se debe cumplir para prescindir de S?)
4. (Control 1 – Otoño 2004) Sean $R=(A,B,C)$ y $T=(B,D)$, R con n tuplas y T con m tuplas, cumpliéndose que $n>m>0$. Indicar el tamaño mínimo y máximo de las siguientes expresiones relacionales. Indicar, además, si cambian los resultados al saber que B es llave de R y de T, y cuáles son los cambios.
 - $R * T$
 - $\sigma_C(R)$
 - $R - (T - R)$
 - $R \div T$

Modificación: Caso en que existe una relación $S=(B,C)$ con \tilde{n} tuplas, $n>\tilde{n}>0$.

Solución:

Dependencias Multivaluadas:

1. Lo es. Basta ver que si dos tuplas tienen el mismo valor según X ($X \rightarrow Y$) entonces Y será el mismo. Una tercera tupla sería, en realidad, una de las anteriores.
2. Basta ver la simetría $Y \cdot Z$ de la definición.
3. Está (a,d,\$) y están (a,b,%) y (a,c,%).
4. Es todo como FNBC (incluso el algoritmo). Sin embargo como las únicas dependencias multivaluadas son de la forma SuperLlave $\rightarrow \rightarrow$ Algo, y se sabe que SuperLlave \rightarrow Algo, 4FN exige que toda dependencia multivaluada sea en realidad funcional y que no viole Boyce-Codd.

Normalización:

1. Se forma (A,B) y (B,C) en FNBC **pero no se preserva la dependencia** $AB \rightarrow C$. En casos así se prefiere 3FN ya que no pone en peligro la integridad del sistema.
2. Propuesto, sólo es aplicar el algoritmo de 3FN y de FNBC.

Álgebra Relacional:

1. Unión, intersección.
2. El JOIN (reunión). La condición sobre los algoritmos es que preserven la reunión, obviamente.
3. Si S son todos los idiomas, entonces la respuesta es $R \div S$. Para prescindir de S se necesitaría que existiese una persona, en R, que hablase todos los idiomas. En ese caso la respuesta es $R \div \pi_{Idioma}(R)$.
4. Veamos cada caso:
 - $R * T$: Join Natural. Cardinalidad: Entre 0 y n·m (caso en que B=b, en todas las tuplas, o sea, sólo posee un valor). Si B es llave, se reduce a 0 y m (pues $m = \text{Máx}\{n,m\}$).
 - $\sigma_c(R)$: Selección. Cardinalidad: Entre 0 y n. B no trasciende.
 - $R - (T - R)$: Sustracción. Como R y T tienen esquemas diferentes, entonces no es posible realizar esta operación a nivel relacional.
 - $R \div T$: División. Es que D no está en R, no se puede siquiera realizar la operación.
 - $R \div S$: División. Cardinalidad: Entre 0 y $\left\lfloor \frac{n}{\tilde{n}} \right\rfloor$. Si B es llave, entre 0 y 1, pues cada elemento de B aparece sólo 1 vez por tupla según sea R o S.