

Normalización

Clase Práctica 2

SPI y SPDF

Rosana Matuk

Departamento de Computación - Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Base de Datos
2do. Cuatrimestre 2014

Esquema General

- 1** Introducción
- 2** Sin pérdida de información
- 3** Preservación de Dependencias Funcionales

Esquema General

1 Introducción

2 Sin pérdida de información

3 Preservación de Dependencias Funcionales

Dos Características de una Buena Descomposición

- Sin Pérdida de Información
- Preservación de Dependencias Funcionales

Esquema General

1 Introducción

2 Sin pérdida de información

3 Preservación de Dependencias Funcionales



Esquema General

2 Sin pérdida de información

- Introducción
- Descomposición binaria
- Algoritmo del Tableau

Definición

Si R es un esquema de relación descompuesto en los esquemas R_1, R_2, \dots, R_k y D es un conjunto de dependencias, decimos que la **descomposición es sin pérdida de información (SPI)** con respecto a D , si para toda relación r para R que satisfaga D :

$$r = \pi_{R_1}(r) \bowtie \pi_{R_2}(r) \bowtie \dots \bowtie \pi_{R_k}(r)$$

Estrategias para comprobar SPI

Descomposiciones de dos esquemas

Descomposición binaria

Descomposiciones en más de dos esquemas

Algoritmo del Tableau

Esquema General

2 Sin pérdida de información

- Introducción
- Descomposición binaria
- Algoritmo del Tableau

Teorema de la Descomposición Binaria

La descomposición ρ de R , $\rho = (R_1, R_2)$ es SPI respecto a un conjunto de dependencias funcionales F sí y sólo sí:

F^+ contiene la DF: $R_1 \cap R_2 \rightarrow (R_1 - R_2)$

o

F^+ contiene la DF: $R_1 \cap R_2 \rightarrow (R_2 - R_1)$

Ejemplo

Sea $R = ABC$ y $F = \{A \rightarrow B\}$.

Pregunta: La descomposición de R en AB y AC es SPI?

Ejemplo

Sea $R = ABC$ y $F = \{A \rightarrow B\}$.

Pregunta: La descomposición de R en AB y AC es SPI?

Resp: Sí. $AB \cap AC = A$, $AB - AC = B$, y $A \rightarrow B$ está en F^+ .

Ejemplo

Sea $R = ABC$ y $F = \{A \rightarrow B\}$.

Pregunta: La descomposición de R en AB y AC es SPI?

Resp: Sí. $AB \cap AC = A$, $AB - AC = B$, y $A \rightarrow B$ está en F^+ .

Tarea para el hogar:

La descomposición de R en AB y BC es SPI?

Esquema General

2 Sin pérdida de información

- Introducción
- Descomposición binaria
- Algoritmo del Tableau



Definición de Tableau

Dado $R = (A_1, \dots, A_n)$, un tableau T para una descomposición $\rho = (R_1, \dots, R_k)$ de R se define de la siguiente forma:

- 1** T tiene n columnas, una para cada atributo de R
- 2** T tiene k filas, una para cada esquema de ρ
- 3** Dadas la fila i y la columna j (esquema R_i y atributo A_j), el contenido del tableau será:

$$\begin{array}{c}
 a_j \text{ si } A_j \in R_i \\
 \text{O} \\
 b_{ij} \text{ si } A_j \notin R_i
 \end{array}$$

Los a_j se denominan símbolos distinguidos, y los b_{ij} no distinguidos.

Algoritmo del Tableau

INPUT: Un esquema de relación R , un conjunto de dependencias funcionales F , y una descomposición ρ . OUTPUT: Una decisión de si ρ es SPI.

Construir el Tableau T

mientras haya cambios sobre T

para cada $df\ X \rightarrow Y \in F$

 buscar filas que coincidan en todos los símbolos de X

 Si se encontrasen dos filas, igualar los símbolos para los atributos de Y . Cuando se igualan 2

 símbolos, si alguno de ellos es a_j , asignarle al otro a_j .

 Si ellos son b_{ij} y b_{lj} , asignarle a ambos b_{ij} o b_{lj} .

Si hay una fila con todos símbolos distinguidos, retornar Sí

end (mientras)

Retornar No

Verificación SPI - Ejercicio 1

Sea $R = ABCDE$, $F = \{A \rightarrow B, D \rightarrow C\}$.

Decidir si la descomposición $\rho = \{ABC, CDE, ADE\}$ es SPI.

Verificación SPI - Ejercicio 1 - Tableau Inicial

$R = ABCDE$, $F = \{A \rightarrow B, D \rightarrow C\}$, $\rho = \{ABC, CDE, ADE\}$

	A	B	C	D	E
ABC	a_1	a_2	a_3	b_{14}	b_{15}
CDE	b_{21}	b_{22}	a_3	a_4	a_5
ADE	a_1	b_{32}	b_{33}	a_4	a_5

Verificación SPI - Ejercicio 1 - Tableau Intermedio

$$R = ABCDE, F = \{A \rightarrow B, D \rightarrow C\}, \rho = \{ABC, CDE, ADE\}$$

$$A \rightarrow B$$

	A	B	C	D	E
ABC	a_1	a_2	a_3	b_{14}	b_{15}
CDE	b_{21}	b_{22}	a_3	a_4	a_5
ADE	a_1	b_{32}	b_{33}	a_4	a_5

Verificación SPI - Ejercicio 1 - Tableau Intermedio

$$R = ABCDE, F = \{A \rightarrow B, D \rightarrow C\}, \rho = \{ABC, CDE, ADE\}$$

$$A \rightarrow B$$

	A	B	C	D	E
ABC	a_1	a_2	a_3	b_{14}	b_{15}
CDE	b_{21}	b_{22}	a_3	a_4	a_5
ADE	a_1	a_2	b_{33}	a_4	a_5

Verificación SPI - Ejercicio 1 - Tableau Intermedio

$$R = ABCDE, F = \{A \rightarrow B, D \rightarrow C\}, \rho = \{ABC, CDE, ADE\}$$

$$D \rightarrow C$$

	A	B	C	D	E
ABC	a_1	a_2	a_3	b_{14}	b_{15}
CDE	b_{21}	b_{22}	a_3	a_4	a_5
ADE	a_1	a_2	b_{33}	a_4	a_5

Verificación SPI - Ejercicio 1 - Tableau Final

$$R = ABCDE, F = \{A \rightarrow B, D \rightarrow C\}, \rho = \{ABC, CDE, ADE\}$$

$$D \rightarrow C$$

	A	B	C	D	E
ABC	a_1	a_2	a_3	b_{14}	b_{15}
CDE	b_{21}	b_{22}	a_3	a_4	a_5
ADE	a_1	a_2	a_3	a_4	a_5

Verificación SPI - Ejercicio 1 - Tableau Final

$R = ABCDE, F = \{A \rightarrow B, D \rightarrow C\}, \rho = \{ABC, CDE, ADE\}$

	A	B	C	D	E
ABC	a_1	a_2	a_3	b_{14}	b_{15}
CDE	b_{21}	b_{22}	a_3	a_4	a_5
ADE	a_1	a_2	a_3	a_4	a_5

Como hay una fila con todos símbolos distinguidos, ρ es SPI.

Verificación SPI - Ejercicio 2

Sea $R = ABCDEFGHI$,

$F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$.

Decidir si la descomposición $\rho = \{ABD, DEF, FGC, CHI\}$ es SPI.

Verificación SPI - Ejercicio 2 - Tableau Inicial

$R = ABCDEFGHI$, $F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$,
 $\rho = \{ABD, DEF, FGC, CHI\}$

	A	B	C	D	E	F	G	H	I
ABD	a_1	a_2	b_{13}	a_4	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}
DEF	b_{21}	b_{22}	b_{23}	a_4	a_5	a_6	b_{27}	b_{28}	b_{29}
FGC	b_{31}	b_{32}	a_3	b_{34}	b_{35}	a_6	a_7	b_{38}	b_{39}
CHI	b_{41}	b_{42}	a_3	b_{44}	b_{45}	b_{46}	b_{47}	a_8	a_9

Verificación SPI - Ejercicio 2 - Tableau Intermedio

$R = ABCDEFGHI$, $F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$,
 $\rho = \{ABD, DEF, FGC, CHI\}$

$D \rightarrow H$

	A	B	C	D	E	F	G	H	I
ABD	a_1	a_2	b_{13}	a_4	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}
DEF	b_{21}	b_{22}	b_{23}	a_4	a_5	a_6	b_{27}	b_{28}	b_{29}
FGC	b_{31}	b_{32}	a_3	b_{34}	b_{35}	a_6	a_7	b_{38}	b_{39}
CHI	b_{41}	b_{42}	a_3	b_{44}	b_{45}	b_{46}	b_{47}	a_8	a_9

Verificación SPI - Ejercicio 2 - Tableau Intermedio

$R = ABCDEFGHI$, $F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$,
 $\rho = \{ABD, DEF, FGC, CHI\}$

$D \rightarrow H$

	A	B	C	D	E	F	G	H	I
ABD	a_1	a_2	b_{13}	a_4	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}
DEF	b_{21}	b_{22}	b_{23}	a_4	a_5	a_6	b_{27}	b_{18}	b_{29}
FGC	b_{31}	b_{32}	a_3	b_{34}	b_{35}	a_6	a_7	b_{38}	b_{39}
CHI	b_{41}	b_{42}	a_3	b_{44}	b_{45}	b_{46}	b_{47}	a_8	a_9

Verificación SPI - Ejercicio 2 - Tableau Intermedio

$R = ABCDEFGHI$, $F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$,
 $\rho = \{ABD, DEF, FGC, CHI\}$

$H \rightarrow AD$

	A	B	C	D	E	F	G	H	I
ABD	a_1	a_2	b_{13}	a_4	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}
DEF	b_{21}	b_{22}	b_{23}	a_4	a_5	a_6	b_{27}	b_{18}	b_{29}
FGC	b_{31}	b_{32}	a_3	b_{34}	b_{35}	a_6	a_7	b_{38}	b_{39}
CHI	b_{41}	b_{42}	a_3	b_{44}	b_{45}	b_{46}	b_{47}	a_8	a_9

Verificación SPI - Ejercicio 2 - Tableau Intermedio

$R = ABCDEFGHI$, $F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$,
 $\rho = \{ABD, DEF, FGC, CHI\}$

$H \rightarrow AD$

	A	B	C	D	E	F	G	H	I
ABD	a_1	a_2	b_{13}	a_4	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}
DEF	a_1	b_{22}	b_{23}	a_4	a_5	a_6	b_{27}	b_{18}	b_{29}
FGC	b_{31}	b_{32}	a_3	b_{34}	b_{35}	a_6	a_7	b_{38}	b_{39}
CHI	b_{41}	b_{42}	a_3	b_{44}	b_{45}	b_{46}	b_{47}	a_8	a_9

Verificación SPI - Ejercicio 2 - Tableau Intermedio

$R = ABCDEFGHI$, $F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$,
 $\rho = \{ABD, DEF, FGC, CHI\}$

$A \rightarrow B$

	A	B	C	D	E	F	G	H	I
ABD	a_1	a_2	b_{13}	a_4	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}
DEF	a_1	b_{22}	b_{23}	a_4	a_5	a_6	b_{27}	b_{18}	b_{29}
FGC	b_{31}	b_{32}	a_3	b_{34}	b_{35}	a_6	a_7	b_{38}	b_{39}
CHI	b_{41}	b_{42}	a_3	b_{44}	b_{45}	b_{46}	b_{47}	a_8	a_9

Verificación SPI - Ejercicio 2 - Tableau Intermedio

$R = ABCDEFGHI$, $F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$,
 $\rho = \{ABD, DEF, FGC, CHI\}$

$A \rightarrow B$

	A	B	C	D	E	F	G	H	I
ABD	a_1	a_2	b_{13}	a_4	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}
DEF	a_1	a_2	b_{23}	a_4	a_5	a_6	b_{27}	b_{18}	b_{29}
FGC	b_{31}	b_{32}	a_3	b_{34}	b_{35}	a_6	a_7	b_{38}	b_{39}
CHI	b_{41}	b_{42}	a_3	b_{44}	b_{45}	b_{46}	b_{47}	a_8	a_9

Verificación SPI - Ejercicio 2 - Tableau Final

$R = ABCDEFGHI$, $F = \{A \rightarrow B, CD \rightarrow F, H \rightarrow AD, I \rightarrow C, D \rightarrow H\}$,
 $\rho = \{ABD, DEF, FGC, CHI\}$

	A	B	C	D	E	F	G	H	I
ABD	a_1	a_2	b_{13}	a_4	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}
DEF	a_1	a_2	b_{23}	a_4	a_5	a_6	b_{27}	b_{18}	b_{29}
FGC	b_{31}	b_{32}	a_3	b_{34}	b_{35}	a_6	a_7	b_{38}	b_{39}
CHI	b_{41}	b_{42}	a_3	b_{44}	b_{45}	b_{46}	b_{47}	a_8	a_9

Como no hay ninguna fila con todos símbolos distinguidos, y aunque sigamos iterando nuevamente por todas las dependencias, ninguna alterará el tableau, ρ NO es SPI.

Esquema General

1 Introducción

2 Sin pérdida de información

3 Preservación de Dependencias Funcionales

Esquema General

3 Preservación de Dependencias Funcionales

- Introducción

- Ejercitación

Preservación de Dependencias Funcionales

Dados un esquema de relación R , una descomposición $\rho = (R_1, \dots, R_k)$, y un conjunto F de dependencias funcionales.

$\pi_Z(F)$: proyección de F sobre un conjunto de atributos Z

Conjunto de dependencias $X \rightarrow Y$ en F^+ tal que $XY \subseteq Z$

Testeo (orden exponencial)

La descomposición ρ preserva F si $F^+ = (\bigcup_{i=1}^k \pi_{R_i}(F))^+$

Testeo Polinomial de Preservación de Dependencias Funcionales

Dados un esquema de relación R , una descomposición $\rho = (R_1, \dots, R_k)$, y un conjunto F de dependencias funcionales.

Para toda dependencia funcional $X \rightarrow Y \in F$:

Verificar que se preserva $X \rightarrow Y$:

$Z = X$

while Z cambia

for $i = 1$ to k do

/* clausura con respecto a F */

$Z = Z \cup ((Z \cap R_i)^+ \cap R_i)$

Si $Y \notin Z$ retornar No

Retornar Sí

Esquema General

3 Preservación de Dependencias Funcionales

■ Introducción

■ Ejercitación

Preservación de Dependencias Funcionales: Ejercicio

Sean $R = ABCDE$ y $F = \{AB \rightarrow C, A \rightarrow D, D \rightarrow E, E \rightarrow C\}$.

Decidir si la descomposición $\rho = \{AD, DE, ECB\}$ es sin pérdida de dependencias funcionales (SPDF).

Preservación de Dependencias Funcionales: Resolución Ejercicio

$$R = ABCDE$$

$$F = \{AB \rightarrow C, A \rightarrow D, D \rightarrow E, E \rightarrow C\}$$

$$\rho = \{AD, DE, ECB\}$$

Estrategia de Resolución:

Preservación de Dependencias Funcionales: Resolución Ejercicio

$R = ABCDE$

$F = \{AB \rightarrow C, A \rightarrow D, D \rightarrow E, E \rightarrow C\}$

$\rho = \{AD, DE, ECB\}$

Estrategia de Resolución:

Las dependencias $A \rightarrow D, D \rightarrow E, E \rightarrow C$ se preservan trivialmente (por qué?), y no es necesario aplicarles el algoritmo.

Le aplicaremos el algoritmo a la dependencia $AB \rightarrow C$ para ver si se preserva.

Preservación de Dependencias Funcionales: Resolución Ejercicio

$R = ABCDE$

$F = \{AB \rightarrow C, A \rightarrow D, D \rightarrow E, E \rightarrow C\}$

$\rho = \{AD, DE, ECB\}$

Queremos verificar que se preserva $AB \rightarrow C$

Verificar que se preserva $X \rightarrow Y$:

$Z = X$

while Z cambia

for i = 1 to k do

$Z = Z \cup ((Z \cap R_i)^+ \cap R_i)$

Si $Y \not\subseteq Z$ retornar No

$Z = AB$

Preservación de Dependencias Funcionales: Resolución Ejercicio

$R = ABCDE$

$F = \{AB \rightarrow C, A \rightarrow D, D \rightarrow E, E \rightarrow C\}$

$\rho = \{AD, DE, ECB\}$

Queremos verificar que se preserva $AB \rightarrow C$

Verificar que se preserva $X \rightarrow Y$:

$Z = X$

while Z cambia

for i = 1 to k do

$Z = Z \cup ((Z \cap R_i)^+ \cap R_i)$

Si $Y \not\subseteq Z$ retornar No

$$\begin{aligned}
 Z &= Z \cup ((Z \cap R_1)^+ \cap R_1) = \{A, B\} \cup ((\{A, B\} \cap \{A, D\})^+ \cap \{A, D\}) \\
 &= \{A, B\} \cup ((A)^+ \cap \{A, D\}) \\
 &= \{A, B\} \cup (\{A, D, E, C\} \cap \{A, D\}) \\
 &= \{A, B, D\}
 \end{aligned}$$

C no está incluido en Z; seguimos...

Preservación de Dependencias Funcionales: Resolución Ejercicio

$R = ABCDE$

$F = \{AB \rightarrow C, A \rightarrow D, D \rightarrow E, E \rightarrow C\}$

$\rho = \{AD, DE, ECB\}$

Queremos verificar que se preserva $AB \rightarrow C$

Verificar que se preserva $X \rightarrow Y$:

$Z = X$

while Z cambia

for i = 1 to k do

$Z = Z \cup ((Z \cap R_i)^+ \cap R_i)$

Si $Y \not\subseteq Z$ retornar No

$$\begin{aligned}
 Z &= Z \cup ((Z \cap R_2)^+ \cap R_2) = \{A, B, D\} \cup ((\{A, B, D\} \cap \{D, E\})^+ \cap \{D, E\}) \\
 &= \{A, B, D\} \cup ((D)^+ \cap \{D, E\}) \\
 &= \{A, B, D\} \cup (\{D, E, C\} \cap \{D, E\}) \\
 &= \{A, B, D, E\}
 \end{aligned}$$

C no está incluido en Z; seguimos...

Preservación de Dependencias Funcionales: Resolución Ejercicio

$R = ABCDE$

$F = \{AB \rightarrow C, A \rightarrow D, D \rightarrow E, E \rightarrow C\}$

$\rho = \{AD, DE, ECB\}$

Queremos verificar que se preserva $AB \rightarrow C$

Verificar que se preserva $X \rightarrow Y$:

$Z = X$

while Z cambia

for $i = 1$ to k do

$Z = Z \cup ((Z \cap R_i)^+ \cap R_i)$

Si $Y \not\subseteq Z$ retornar No

$Z = Z \cup ((Z \cap R_3)^+ \cap R_3) =$

$\{A, B, D, E\} \cup ((\{A, B, D, E\} \cap \{E, C, B\})^+ \cap \{E, C, B\})$

$= \{A, B, D, E\} \cup ((EB)^+ \cap \{E, C, B\})$

$= \{A, B, D, E\} \cup (\{E, B, C\} \cap \{E, C, B\})$

$= \{A, B, D, E, C\}$

Ahora sí C está incluido en Z : la dependencia $AB \rightarrow C$ se preserva

Bibliografía

Referencia

Jeffrey D. Ullman "Principles of Database and Knowledge-base systems", Volumen I, Computer Science Press, 1988 (*capítulo 7*).