Baze de date relaționale
Dependențe multivaluate

Nicolae-Cosmin Vârlan

October 24, 2018

Exemplu

Presupunem că persoana cu CNP $= 1$ a fost admisă la două facultăți și are permis de conducere pentru categoriile A și B:

r :

CNP	Admis la facult.	Are permis categ.
1	Informatică	A
1	Matematică	B

Deși anumite rânduri nu sunt scrise în tabelă, putem să intuim că persoana cu CNP $= 1$ a dat la Facultatea de Informatică și are permis de conducerea categoria B. Deci, deși în r nu există t-uplul $\langle$ 1,Informatica,B $\rangle$, ar trebui să existe și el (pentru că poate fi dedus din cele existente).

Care alt t-uplu mai poate fi dedus ?

Exemplu

$r:$

CNP	Admis la facult.	Are permis categ.
1	Informatică	A
1	Matematică	B
1	Informatică	B
1	Matematică	A

t-uplele marcate cu roșu ar putea lipsi, ele fiind redundante deoarece pot fi obținute din primele două t-uple.

Prin intermediul dependențelor funcționale pot afla la care coloane pot renunța astfel încât să le pot reface ulterior.

Prin intermediul dependențelor multivaluate pot afla la care linii pot renunța astfel încât să le pot reface ulterior.

Dependențe multivaluate - definiție

Fie $X, Y \subseteq U$. O dependență multivaluată este notată cu $X \twoheadrightarrow Y$.

Definition

Relația *r peste U satisface dependența multivaluată $X \twoheadrightarrow Y$* dacă pentru oricare două tuple $t_1, t_2 \in r$ satisfăcând $t_1[X] = t_2[X]$, există tuplele t_3 și t_4 din r, astfel încât:

- $t_3[X] = t_1[X], t_3[Y] = t_1[Y], t_3[Z] = t_2[Z]$;
- $t_4[X] = t_2[X], t_4[Y] = t_2[Y], t_4[Z] = t_1[Z]$

unde $Z = U - XY$ (Z mai este denumită și $rest$).

Exemplul 2 (mai formal)

r :

A	B	C	D			
a_1	b_1	c_1	d_1	t_1		t_1"
a_1	b_2	c_2	d_2	t_2		
a_1	b_1	c_1	d_2	t_3		t_2"
a_1	b_2	c_2	d_1	t_4		
a_2	b_3	c_1	d_1		t'_1, t'_4	
a_2	b_3	c_1	d_2		t'_2, t'_3	

r satisface $A \twoheadrightarrow BC$

Intrebare: cum alegem t_3", t_4" ?

Deoarece atunci când $t_1[A] = t_2[A]$ avem că:
$t_3[A] = t_1[A], t_3[BC] = t_1[BC], t_3[D] = t_2[D]$ și
$t_4[A] = t_2[A], t_4[BC] = t_2[BC], t_4[D] = t_1[D]$

Definiție echivalentă

Relația *r peste U satisface dependența multivaluată $X \twoheadrightarrow Y$*, dacă pentru orice $t_1, t_2 \in r$ cu $t_1[X] = t_2[X]$ avem că $M_Y(t_1[XZ]) = M_Y(t_2[XZ])$

unde $M_Y(t[XZ]) = \{t'[Y] | t' \in r, t'[XZ] = t[XZ]\}$ = valorile lui Y din diferite tuple în care XZ sunt egale (cu XZ-ul din parametru).

r :

A	B	C	D		
a_1	b_1	c_1	d_1	$=$	t_1
a_1	b_2	c_2	d_2	$=$	t_2
a_1	b_1	c_1	d_2		
a_1	b_2	c_2	d_1		
a_2	b_3	c_1	d_1		
a_2	b_3	c_1	d_2		

$M_Y(t1[AD]) = M_Y(t2[AD]) = \{(b_1, c_1), (b_2, c_2)\}$

Observații

- Dacă r satisface dependența funcțională $X \to Y$, atunci pentru orice $t \in r$, avem $M_Y(t[XZ]) = \{t[Y]\}$.
- Dacă r satisface dependența funcțională $X \to Y$, atunci r satisface și dependența multivaluată $X \twoheadrightarrow Y$.
- Dacă r satisface dependența multivaluată $X \twoheadrightarrow Y$, atunci putem defini o funcție $\psi : r[X] \to \mathcal{P}(r[Y])$, prin $\psi(t[X]) = M_Y(t[XZ]), \forall t \in r$ (returnează valorile diferite din proiecția pe Y). Când r satisface $X \to Y$, atunci $\psi : r[X] \to r[Y]$ (deoarece valorile pe Y nu sunt diferite în cadrul dependenței funcționale).

Proprietăți ale dependențelor multivaluate

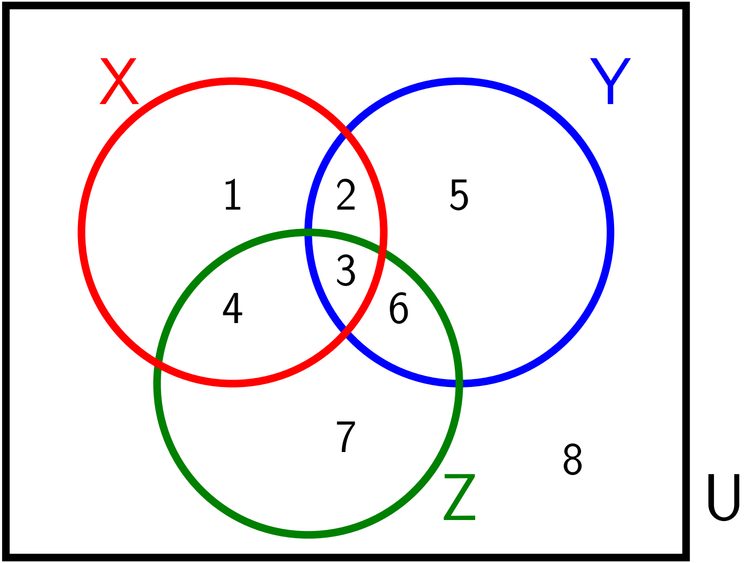
MVD0 (Complementariere) Fie $X, Y, Z \subseteq U$, asfel încât $XYZ = U$ și $Y \cap Z \subseteq X$. Dacă r satisface $X \twoheadrightarrow Y$, atunci r satisface $X \twoheadrightarrow Z$.

MVD1 (Reflexivitate) Dacă $Y \subseteq X$, atunci orice relație r satisface $X \twoheadrightarrow Y$.

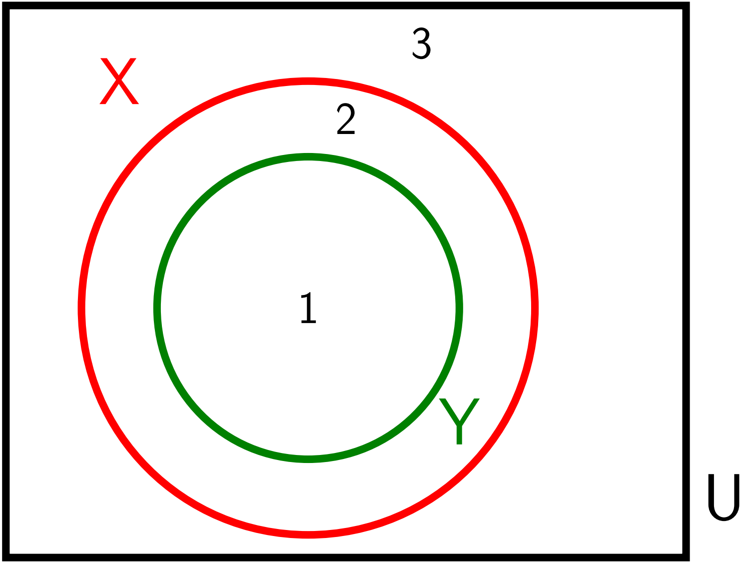
MVD2 (Extensie) Fie $Z \subseteq W$ și r satisface $X \twoheadrightarrow Y$. Atunci r satisface $XW \twoheadrightarrow YZ$

MVD3 (Tranzitivitate) Dacă r satisface $X \twoheadrightarrow Y$ și $Y \twoheadrightarrow Z$, atunci r satisface $X \twoheadrightarrow Z - Y$

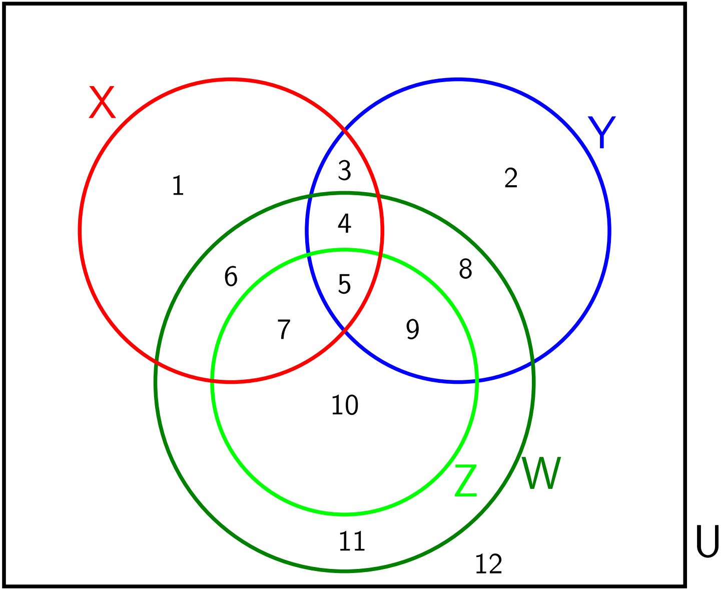
MVD0



MVD1



MVD2



Proprietăți ale dependențelor multivaluate

MVD4 (Pseudotranzitivitate) Dacă r satisface $X \twoheadrightarrow Y$ și $YW \twoheadrightarrow Z$, atunci r satisface și $XW \twoheadrightarrow Z - YW$.

MVD5 (Uniune) Dacă r satisface $X \twoheadrightarrow Y$ și $X \twoheadrightarrow Z$ atunci r satisface $X \twoheadrightarrow YZ$.

MVD6 (Descompunere) Dacă r satisface $X \twoheadrightarrow Y$ și $X \twoheadrightarrow Z$, atunci r satisface $X \twoheadrightarrow Y \cap Z$, $X \twoheadrightarrow Y - Z$, $X \twoheadrightarrow Z - Y$

Proprietăți mixte ale dependențelor multivaluate

FD-MVD1. Dacă r satisface $X \to Y$, atunci r satisface și $X \twoheadrightarrow Y$.

FD-MVD2. Dacă r satisface $X \twoheadrightarrow Z$ și $Y \to Z'$, cu $Z' \subseteq Z$ și $Y \cap Z = \emptyset$, atunci r satisface $X \to Z'$.

FD-MVD3. Dacă r satisface $X \twoheadrightarrow Y$ și $XY \twoheadrightarrow Z$, atunci r satisface $X \to Z - Y$.

Reguli de inferență

MVD0f: $$\frac{XYZ{=}U,\ \ Y{\cap}Z{\subseteq}X,\ \ X{\twoheadrightarrow}Y}{X{\twoheadrightarrow}Z}$$

MVD1f: $$\frac{Y{\subseteq}X}{X{\twoheadrightarrow}Y}$$

MVD2f: $$\frac{Z{\subseteq}W,\ \ X{\twoheadrightarrow}Y}{XW{\twoheadrightarrow}YZ}$$

MVD3f: $$\frac{X{\twoheadrightarrow}Y,\ \ Y{\twoheadrightarrow}Z}{X{\twoheadrightarrow}Z{-}Y}$$

MVD4f: $$\frac{X{\twoheadrightarrow}Y{,}YW{\twoheadrightarrow}Z}{XW{\twoheadrightarrow}Z{-}YW}$$

Reguli de inferență

MVD5f: $\dfrac{X \twoheadrightarrow Y,\ X \twoheadrightarrow Z}{X \twoheadrightarrow YZ}$

MVD6f: $\dfrac{X \twoheadrightarrow Y,\ X \twoheadrightarrow Z}{X \twoheadrightarrow Y \cap Z,\ X \twoheadrightarrow Y - Z,\ X \twoheadrightarrow Z - Y}$

FD-MVD1f: $\dfrac{X \rightarrow Y}{X \twoheadrightarrow Y}$

FD-MVD2f: $\dfrac{X \twoheadrightarrow Z,\ Y \rightarrow Z',\ Z' \subseteq Z,\ Y \cap Z = \emptyset}{X \rightarrow Z'}$

FD-MVD3f: $\dfrac{X \twoheadrightarrow Y,\ XY \twoheadrightarrow Z}{X \twoheadrightarrow Z - Y}$

Propoziție

Fie $\mathcal{R}$ o multime de reguli valide si γ o regula $\frac{\alpha_1,\alpha_2,\ldots\alpha_k}{\beta}$, astfel incat $\{\alpha_1,\ldots\alpha_k\} \vdash_{\mathcal{R}} \beta$, atunci si regula γ este valida.

Propoziție

Fie $\mathcal{R}_{FM} = \{FD1f \ - \ FD3f$[1]$, \quad MVD0f \ - \ MVD3f,$ $FD - MVD1f \ - \ FD - MVD3f\}$. Avem:

- *$FD - MVD3f$ se exprima cu celelalte regulid din $\mathcal{R}_{FM}$ si FD*
- *$MVD2f$ se exprima prin celelalte reguli din $\mathcal{R}_{FM}$.*

Propoziție

Regulile $MVD4f - MVD6f$ se exprima cu ajutorul regulilor $MVD0f - MVD3f$

[1] cele de la dependente functionale

Theorem

Fie Σ o multime de dependente functionale sau multivaluate si X o submultime de atribute. Atunci exista o partitie a lui $U - X$ notata prin $Y_1 \dots Y_k$, astfel incat pentru $Z \subseteq U - X$ avem $\Sigma \vdash_{\mathcal{R}_{FM}} X \twoheadrightarrow Z$ daca si numai daca Z este reuniunea unui numar de multimi din partitia $\{Y_1, \dots Y_k\}$

Definition

Pentru Σ o multime de dependente functionale sau multivaluate si X o submultime de atribute, numim *baza de dependenta pentru X cu privire la Σ* partitia $B(\Sigma, X) = \{\{A_1\} \dots \{A_h\}, Y_1 \dots Y_k\}$, unde $X = A_1, \dots A_h$, iar $Y_1, \dots Y_k$ este partitia construita in teorema precedenta.

Observatii

- Avem $\Sigma \vdash_{\mathcal{R}_{FM}} X \twoheadrightarrow Z$ daca si numai daca Z este o reuniune de elemente din partitia $B(\Sigma, X)$.
- Fie $X^*_\Sigma = \{A | \Sigma \vdash_{\mathcal{R}_{FM}} X \rightarrow A\}$. Atunci pentru orice $A \in X^*_\Sigma$ avem $\{A\} \in B(\Sigma, X)$.

Bibliografie

- Baze de date relaționale. Dependențe - *Victor Felea*; Univ. Al. I. Cuza, 1996