

Baze de date relaționale

Dependențe multivaluate

Nicolae-Cosmin Vârlan

Exemplu

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

	CNP	Admis la facult.	Are permis categ.
$r :$	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 $\langle 1, \text{Informatica}, B \rangle$, ar trebui să existe și el (pentru că poate fi dedus din cele existente).

Care alt t -uplu mai poate fi dedus ?

Exemplu

	CNP	Admis la facult.	Are permis categ.
$r :$	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)

	A	B	C	D			
$r :$	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 $A \rightarrow BC$

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

	A	B	C	D	
	a_1	b_1	c_1	d_1	$= t_1$
	a_1	b_2	c_2	d_2	$= t_2$
$r :$	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	
	<hr/>				
	$M_Y(t_1[AD]) = M_Y(t_2[AD]) = \{(b_1, c_1), (b_2, c_2)\}$				

Observații

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

Dependențe multivaluate (exercițiu)

Arătați că $AC \twoheadrightarrow BD$:

	A	B	C	D	E
	8	1	2	0	4
	8	9	2	2	9
$r :$	9	3	2	4	9
	8	1	2	0	9
	8	9	2	2	4
	9	3	2	4	4

Când $r[AC] = \{(8, 2)\}$ avem $r[BD] = \{(1, 0), (9, 2)\}$ și $r[E] = \{(4), (9)\}$. Găsim toate produsele carteziane dintre cele 3 ?

Când $r[AC] = \{(9, 2)\}$ avem $r[BD] = \{(3, 4)\}$ și $r[E] = \{(4), (9)\}$. Găsim toate produsele carteziane ?

DA (este MVD)

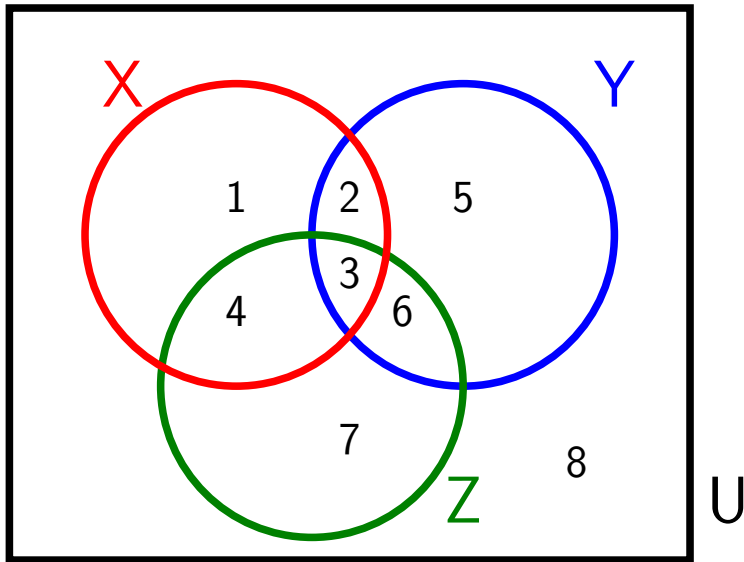
Proprietăți ale dependențelor multivaluate

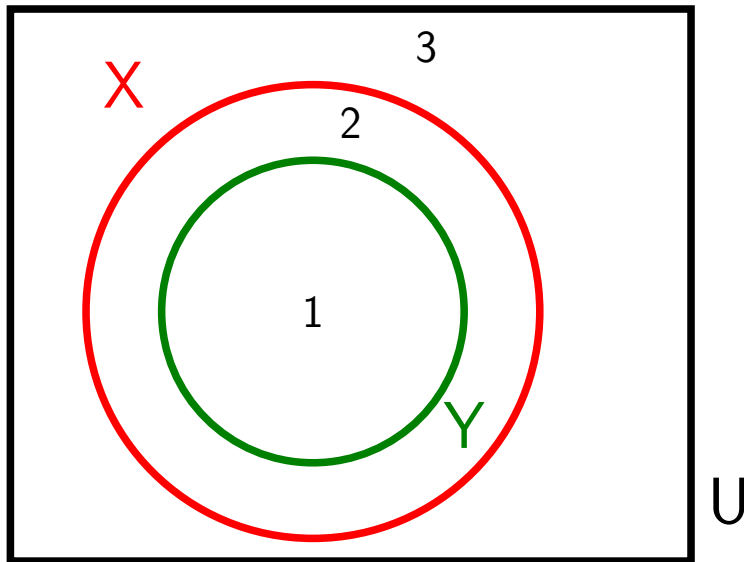
MVD0 (**Complementarizare**) Fie $X, Y, Z \subseteq U$, astfel încât $XYZ = U$ și $Y \cap Z \subseteq X$. Dacă r satisface $X \twoheadrightarrow Y$, atunci r satisface $X \twoheadrightarrow Z$.

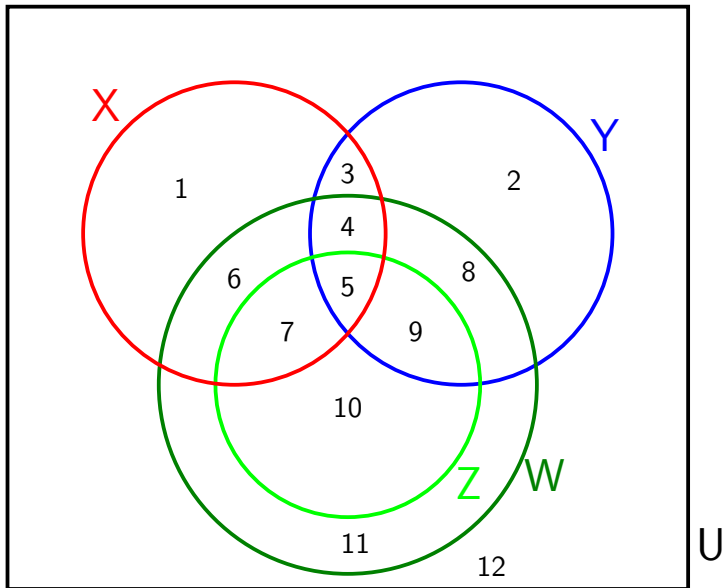
MVD1 (**Reflexivitate**) Dacă $Y \subseteq X$, atunci orice relație r satisface $X \twoheadrightarrow Y$.

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 $X \twoheadrightarrow Z - Y$







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 $X \rightarrow Y$, atunci r satisface și $X \twoheadrightarrow Y$.

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

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

Reguli de inferență

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

$$\text{MVD1f: } \frac{Y \subseteq X}{X \rightarrow Y}$$

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

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

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

Reguli de inferență

$$\text{MVD5f: } \frac{X \rightarrow Y, X \rightarrow Z}{X \rightarrow YZ}$$

$$\text{MVD6f: } \frac{X \rightarrow Y, X \rightarrow Z}{X \rightarrow Y \cap Z, X \rightarrow Y - Z, X \rightarrow Z - Y}$$

$$\text{FD-MVD1f: } \frac{X \rightarrow Y}{X \rightarrow Y}$$

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

$$\text{FD-MVD3f: } \frac{X \rightarrow Y, XY \rightarrow Z}{X \rightarrow Z - Y}$$

Propoziție

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

Propoziție

Fie $\mathcal{R}_{FM} = \{FD1f - FD3f^1, MVD0f - MVD3f, FD - MVD1f - FD - MVD3f\}$. 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 $MVD0f - MVD3f$

Bibliografie

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