

Baze de date Dependențe funcționale

Nicolae-Cosmin Vârlan

October 21, 2019

Egalitatea a două tuple

Considerăm $U = \{A_1, A_2 \dots A_n\}$ o mulțime de attribute și două tuple t_1 și t_2 construite peste această mulțime de attribute.

Spunem că *tuplele t_1 și t_2 sunt egale*, dacă și numai dacă

$$\pi_{A_i}[t_1] = \pi_{A_i}[t_2], \forall i \in \{1..n\}$$

Cu alte cuvinte, tuplele t_1 și t_2 sunt egale dacă ele sunt egale pe fiecare dintre componentele lor. Considerând că

$t_1 = (v_{11}, v_{12}, \dots v_{1n})$ și $t_2 = (v_{21}, v_{22}, \dots v_{2n})$, atunci $t_1 = t_2$ dacă și numai dacă $v_{11} = v_{21}, v_{12} = v_{22}, \dots, v_{1n} = v_{2n}$.

În restul cursului, vom înlocui notația $\pi_X[t]$ cu $t[X]$.

Dependențe funcționale

Fie $X, Y \subseteq U$. Vom nota o dependență funcțională cu $X \rightarrow Y$.

O relație r peste U satisface **dependența funcțională** $X \rightarrow Y$ dacă:

$$(\forall t_1, t_2)(t_1, t_2 \in r) \ t_1[X] = t_2[X] \Rightarrow t_1[Y] = t_2[Y]$$

$X = \emptyset$ avem $\emptyset \rightarrow Y$ dacă $(\forall t_1, t_2)(t_1, t_2 \in r)[t_1[Y] = t_2[Y]]$

$Y = \emptyset$ atunci orice $\forall r$ peste U avem că $X \rightarrow \emptyset$

Dacă r satisface $X \rightarrow Y$, atunci există o funcție $\varphi : r[X] \rightarrow r[Y]$ definită prin $\varphi(t) = t'[Y]$, unde $t' \in r$ și $t'[X] = t \in r[X]$.

Dacă r satisface $X \rightarrow Y$ spunem că X determină funcțional pe Y în r .

Exemplu

Fie relația r peste mulțimea de atribute

$$U = \{nume, l(nume), data_nastere, zodie, varsta\}$$

	<i>nume</i>	<i>l(nume)</i>	<i>data_nastere</i>	<i>zodie</i>	<i>varsta</i>
$r :$	Ion	3	20.02.1990	Pesti	28
	Vasile	6	24.02.1992	Pesti	26
	Maria	5	1.08.2014	Leu	4
	Cosmin	6	7.07.1978	Rac	40
	Maria	5	4.08.2010	Leu	8

Puteți depista dependențele funcționale ?

Exemplu

Fie relația r peste mulțimea de atribute

$$U = \{nume, l(nume), data_nastere, zodie, varsta\}$$

	<i>nume</i>	<i>l(nume)</i>	<i>data_nastere</i>	<i>zodie</i>	<i>varsta</i>
	Ion	3	20.02.1990	Pesti	28
	Vasile	6	24.02.1992	Pesti	26
$r :$	Maria	5	1.08.2014	Leu	4
	Cosmin	6	7.07.1978	Rac	40
	Maria	5	4.08.2010	Leu	8

- ▶ $nume \rightarrow l(nume)$
- ▶ $data_nastere \rightarrow varsta$
- ▶ $data_nastere \rightarrow zodie$
- ▶ $nume \rightarrow zodie$ - *discuție*

Proprietăți ale dependențelor funcționale

FD1. (**Reflexivitate**) Dacă $Y \subseteq X$, atunci r satisface $X \rightarrow Y$, $\forall r \in U$.

FD2. (**Extensie**) Dacă r satisface $X \rightarrow Y$ și $Z \subseteq W$, atunci r satisface $XW \rightarrow YZ$.

FD3. (**Tranzitivitate**) Dacă r satisface $X \rightarrow Y$ și $Y \rightarrow Z$, atunci r satisface $X \rightarrow Z$.

FD4. (**Pseudotranzitivitate**) Dacă r satisface $X \rightarrow Y$ și $YW \rightarrow Z$, atunci r satisface $XW \rightarrow Z$.

Proprietăți ale dependențelor funcționale

FD5. (**Uniune**) Dacă r satisface $X \rightarrow Y$ și $X \rightarrow Z$, atunci r satisface $X \rightarrow YZ$.

FD6. (**Descompunere**) Dacă r satisface $X \rightarrow YZ$, atunci r satisface $X \rightarrow Y$ și $X \rightarrow Z$.

FD7. (**Proiectabilitate**) Dacă r peste U satisface $X \rightarrow Y$ și $X \subset Z \subseteq U$, atunci $r[Z]$ satisface $X \rightarrow Y \cap Z$

FD8. (**Proiectabilitate inversă**) Dacă $X \rightarrow Y$ este satisfăcută de o proiecție a lui r , atunci $X \rightarrow Y$ este satisfăcută de r .

Dependențe funcționale - consecință și acoperire

Dacă Σ este o mulțime de dependențe funcționale peste U atunci spunem că $X \rightarrow Y$ *este consecință din Σ* dacă orice relație ce satisface toate dependențele din Σ satisface și $X \rightarrow Y$.

Notăție: $\Sigma \models X \rightarrow Y$

Fie $\Sigma^* = \{X \rightarrow Y \mid \Sigma \models X \rightarrow Y\}$. Fie $\Sigma_1 =$ mulțime de dependențe funcționale. Σ_1 constituie o *acoperire* pentru Σ^* dacă $\Sigma_1^* = \Sigma^*$.

Exercițiu: Fie $U = \{A, B, C, D, E, F\}$ și $\Sigma = \{A \rightarrow BD, B \rightarrow C, DE \rightarrow F\}$ găsiți cât mai multe elemente din $\Sigma^* - \Sigma$.

Proprietăți ale dependențelor funcționale

Propoziție

Pentru orice mulțime Σ de dependențe funcționale există o acoperire Σ_1 pentru Σ^ , astfel încat toate dependențele din Σ_1 sunt de forma $X \rightarrow A$, A fiind un atribut din U .*

Propoziție

$\Sigma \models X \rightarrow Y$ dacă și numai dacă $\Sigma \models X \rightarrow B_j$ pentru $j = \overline{1, h}$, unde $Y = B_1 \dots B_h$.

Reguli de deducere (la nivel sintactic)

Fie \mathcal{R} o mulțime de reguli de deducere pentru dependențe funcționale și Σ o mulțime de dependențe funcționale.

Spunem că $X \rightarrow Y$ este o *demonstrație* în Σ utilizând regulile \mathcal{R} și vom nota $\Sigma \vdash_{\mathcal{R}} X \rightarrow Y$, dacă există șirul $\sigma_1, \sigma_2, \dots, \sigma_n$, astfel încât:

- ▶ $\sigma_n = X \rightarrow Y$ și
- ▶ pentru $\forall i = \overline{1, n}$, $\sigma_i \in \Sigma$ sau există în \mathcal{R} o regulă de forma $\frac{\sigma_{j_1}, \sigma_{j_2}, \dots, \sigma_{j_k}}{\sigma_i}$, unde $j_1, j_2, \dots, j_k < i$.

Reguli de deducere (la nivel sintactic)

Conform proprietăților FD1-FD6 putem defini regulile:

$$\text{FD1f: } \frac{Y \subseteq X}{X \rightarrow Y}$$

$$\text{FD4f: } \frac{X \rightarrow Y, YW \rightarrow Z}{XW \rightarrow Z}$$

$$\text{FD2f: } \frac{X \rightarrow Y, Z \subseteq W}{XW \rightarrow YZ}$$

$$\text{FD5f: } \frac{X \rightarrow Y, X \rightarrow Z}{X \rightarrow YZ}$$

$$\text{FD3f: } \frac{X \rightarrow Y, Y \rightarrow Z}{X \rightarrow Z}$$

$$\text{FD6f: } \frac{X \rightarrow YZ, X \rightarrow YZ}{X \rightarrow Y}, \frac{X \rightarrow YZ, X \rightarrow YZ}{X \rightarrow Z}$$

Propoziție

Regulile FD4f, FD5f, FD6f se exprimă cu ajutorul regulilor FD1f, FD2f, FD3f.

Notăm cu $\mathcal{R}_1 = \{\text{FD1f, FD2f, FD3f}\}$,
și cu $\mathcal{R}_2 = \mathcal{R}_1 \cup \{\text{FD4f, FD5f, FD6f}\}$

Propoziție

Regulile FD4f, FD5f, FD6f se exprimă cu ajutorul regulilor FD1f, FD2f, FD3f.

Idei de demonstrație:

- ▶ FD4f: Se aplica FD2f pentru $X \rightarrow Y$ și $W \subseteq W$ iar din rezultat și din $YW \rightarrow Z$ prin FD3f se obține rezultatul;
- ▶ FD5f: Se aplica FD2f pentru $X \rightarrow Y$ și $X \subseteq X$ și la fel pentru $X \rightarrow Z$ și $Y \subseteq Y$ apoi FD3f (tranzitivitatea) între rezultate;
- ▶ FD6f: din FD1f avem ca $YZ \rightarrow Y$ și $YZ \rightarrow Z$ și din FD3f rezulta $X \rightarrow Y$ și $X \rightarrow Z$

Axiomele lui Armstrong

Armstrong a definit (în *Dependency structures of database relationships* Proc. IFIP 74, Amsterdam, 580-583) următoarele reguli de inferență (numite *Axiomele lui Armstrong*):

$$A1: \frac{}{A_1 \dots A_n \rightarrow A_i}, i = \overline{1, n}$$

$$A2: \frac{A_1, \dots, A_m \rightarrow B_1, \dots, B_r}{A_1 \dots A_m \rightarrow B_j}, j = \overline{1, r}$$

$$\frac{A_1, \dots, A_m \rightarrow B_j, j = \overline{1, r}}{A_1 \dots A_m \rightarrow B_1, \dots, B_r}$$

$$A3: \frac{A_1, \dots, A_m \rightarrow B_1, \dots, B_r, \quad B_1, \dots, B_r \rightarrow C_1, \dots, C_p}{A_1 \dots A_m \rightarrow C_1, \dots, C_p}$$

unde A_i, B_j, C_k sunt atribute. Notăm $\mathcal{R}_A = \{A1, A2, A3\}$.

Obs: regula A3 este de fapt FD3f (tranzitivitatea).

Propoziție

Regulile din \mathcal{R}_1 se exprimă prin cele din \mathcal{R}_A și invers.

Notăție:

$$\Sigma_{\mathcal{R}}^+ = \{X \rightarrow Y \mid \Sigma \vdash_{\mathcal{R}} X \rightarrow Y\}$$

Propoziție

Fie \mathcal{R}'_1 și \mathcal{R}'_2 două mulțimi de reguli astfel încât \mathcal{R}'_1 se exprima prin \mathcal{R}'_2 și invers. Atunci $\Sigma_{\mathcal{R}'_1}^+ = \Sigma_{\mathcal{R}'_2}^+$ pentru orice mulțime Σ de dependente funcționale.

Consecința: $\Sigma_{\mathcal{R}_1}^+ = \Sigma_{\mathcal{R}_A}^+$

Fie $X \subseteq U$ și \mathcal{R} o mulțime de reguli de inferență. Notăm cu

$$X_{\mathcal{R}}^+ = \{A \mid \Sigma \vdash_{\mathcal{R}} X \rightarrow A\}$$

Lema

$\Sigma \vdash_{\mathcal{R}} X \rightarrow Y$ dacă și numai dacă $Y \subseteq X_{\mathcal{R}}^+$.

Lema

Fie Σ o multime de dependente functionale si $\sigma : X \rightarrow Y$ o dependenta functionala astfel incat $\Sigma \not\models_{\mathcal{R}_1} X \rightarrow Y$. Atunci exista o relatie r_σ ce satisface toate dependentele functionale din Σ si r_σ nu satisface $X \rightarrow Y$.

Theorem

Fie Σ o multime de dependente functionale. Atunci exista o relatie r_0 ce satisface exact elementele lui $\Sigma_{\mathcal{R}_1}^+$, adica:

- ▶ r_0 satisface τ , $\forall \tau \in \Sigma_{\mathcal{R}_1}^+$ si
- ▶ r_0 nu satisface γ , $\forall \gamma \notin \Sigma_{\mathcal{R}_1}^+$

Bibliografie

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