# Baze de date
# Dependențe funcționale

Nicolae-Cosmin Vârlan

# Egalitatea a două tuple

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

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 $X \to Y$.

O relație $r$ peste $U$ satisface *dependența funcțională* $X \to Y$ 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 $\emptyset \to Y$ 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 $X \to \emptyset$

Dacă $r$ satisface $X \to Y$, atunci există o funcție $\varphi : r[X] \to r[Y]$ definită prin $\varphi(t) = t'[Y]$, unde $t' \in r$ și $t'[X] = t \in r[X]$.
Dacă $r$ satisface $X \to Y$ spunem că $X$ determină funcțional pe $Y$ în $r$.

# Exemplu

Fie relația $r$ peste mulțimea de atribute
$U = \{nume, l(nume), data\_nastere, zodie, varsta\}$

$r$ :

| $nume$ | $l(nume)$ | $data\_nastere$ | $zodie$ | $varsta$ |
|---|---|---|---|---|
| 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
$U = \{nume, l(nume), data\_nastere, zodie, varsta\}$

$r$ :

| $nume$ | $l(nume)$ | $data\_nastere$ | $zodie$ | $varsta$ |
|---|---|---|---|---|
| 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 |
| ... | ... | ... | ... | ... |

- $nume \to l(nume)$
- $data\_nastere \to varsta$
- $data\_nastere \to zodie$
- $nume \to zodie$ - *discuție*

# Proprietăți ale dependențelor funcționale

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

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

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

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

# Proprietăți ale dependențelor funcționale

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

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

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

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

# Dependențe funcționale - consecință și acoperire

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

Fie $\Sigma^* = \{X \to Y | \Sigma \models X \to Y\}$. Fie $\Sigma_1 =$ mulțime de dependențe funcționale. $\Sigma_1$ constituie o *acoperire* pentru $\Sigma^*$ dacă $\Sigma_1^* = \Sigma^*$.

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

# Proprietăți ale dependențelor funcționale

## Propoziție

*Pentru orice mulțime $\Sigma$ de dependențe funcționale există o acoperire $\Sigma_1$ pentru $\Sigma^*$, astfel încat toate dependențele din $\Sigma_1$ sunt de forma $X \to A$, $A$ fiind un atribut din $U$.*

## Propoziție

$\Sigma \models X \to Y$ *dacă și numai dacă* $\Sigma \models X \to B_j$ *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 $\Sigma$ o mulțime de dependențe funcționale.

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

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

# Reguli de deducere (la nivel sintactic)

Conform proprietăților FD1-FD6 putem defini regulile:

FD1f: $\frac{Y \subseteq X}{X \to Y}$

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

FD3f: $\frac{X \to Y,\ Y \to Z}{X \to Z}$

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

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

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

## Propoziție

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

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

## Propoziție

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

Idei de demonstratie:

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

# Axiomele lui Armstrong

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

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

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

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

A3: $\frac{A_1, ... A_m \to B_1, ... B_r, \quad B_1, ... B_r \to C_1, ... C_p}{A_1 ... A_m \to C_1, ... C_p}$

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.*

Notație:

$$\Sigma^+_{\mathcal{R}} = \{X \to Y | \Sigma \vdash_{\mathcal{R}} X \to Y\}$$

## Propoziție

*Fie $\mathcal{R}'_1$ si $\mathcal{R}'_2$ doua multimi de reguli astfel incat $\mathcal{R}'_1$ se exprima prin $\mathcal{R}'_2$ si invers. Atunci $\Sigma^+_{\mathcal{R}'_1} = \Sigma^+_{\mathcal{R}'_2}$ pentru orice multime $\Sigma$ de dependente functionale.*

Consecinta: $\Sigma^+_{\mathcal{R}_1} = \Sigma^+_{\mathcal{R}_A}$

Fie $X \subseteq U$ si $\mathcal{R}$ o multime de reguli de inferenta. Notam cu

$$X^+_{\mathcal{R}} = \{A | \Sigma \vdash_{\mathcal{R}} X \rightarrow A\}$$

## Lema

$\Sigma \vdash_{\mathcal{R}} X \rightarrow Y$ *daca si numai daca* $Y \subseteq X^+_{\mathcal{R}_1}$.

### Lema

*Fie* $\Sigma$ *o multime de dependente functionale si* $\sigma : X \rightarrow Y$ *o dependenta functionala astfel incat* $\Sigma \nvdash_{\mathcal{R}_1} X \rightarrow Y$. *Atunci exista o relatie* $r_\sigma$ *ce satisface toate dependentele functionale din* $\Sigma$ *si* $r_\sigma$ *nu satisface* $X \rightarrow Y$.

### Theorem

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

- $r_0$ *satisface* $\tau$, $\forall \tau \in \Sigma^+_{\mathcal{R}_1}$ *si*
- $r_0$ *nu satisface* $\gamma$, $\forall \gamma \notin \Sigma^+_{\mathcal{R}_1}$

# Bibliografie


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