OBLIKOVANJE RELACIJSKOG MODELA – II dio

IV predavanje

Dr.sc. Emir Mešković

Armstrongovi aksiomi

- Projektant sheme baze podataka specificira funkcijske zavisnosti koje su mi semantički očigledne, ali obično vrijede i brojne druge funkcijske zavisnosti koje mogu biti izvedene iz početnih funkcijskih zavisnosti
- Korištenjem Armstrongovih aksioma izvode se nove funkcijske zavisnosti
- Neka je R relacijska shema, neka su X, Y, Z i V skupovi atributa i neka vrijedi:
- $X \subseteq R$, $Y \subseteq R$, $Z \subseteq R$, $V \subseteq R$

A-1 REFLEKSIVNOST

- Ako je $Y \subseteq X$, tada vrijedi $X \to Y$
  - Možemo reći da uvijek vrijedi $X \to X$ – aksiom o trivijalnoj funkcijskoj zavisnosti
- $X \to X \Rightarrow \pi_X(\sigma_{X=x}(r))$ uvijek sadrži najviše jednu n-torku
  - Uvijek postoji $X \to X$ u r

os(*matBr*, *prezime*, *ime*, *postBr*, *nazivGrad*)

matBr	*prezime*	*ime*	*postBr*	*nazivGrad*
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

$\sigma_{postBr=71000}(os)$ = *oso* (matBr, *prezime*, *ime*, *postBr*, *nazivGrad*)

matBr	*prezime*	*ime*	*postBr*	*nazivGrad*
23456	Đurić	Maja	71000	Sarajevo
45678	Pirić	Damir	71000	Sarajevo

$\pi_{postBr}(oso)$ = *osob* (*postBr*)

(*postBr*)
71000

A-1 REFLEKSIVNOST

- Ako je $Y \subseteq X$, tada r zadovoljava $X \rightarrow Y$
- Iz $t_i(X) = t_j(X)$ i $Y \subseteq X$ proizilazi $t_i(Y) = t_j(Y)$
- Prema tome $\pi_Y(\sigma_{X=x}(r))$ uvijek sadrži najviše jednu n-torku

$X = \{prezime, ime\}$
$Y = \{prezime\}$

os

matBr	*prezime*	*ime*	*postBr*	*nazivGrad*
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

$\sigma_{prezime = \text{'Pirić'} \wedge ime = \text{'Damir'}}(os) = os_1$

os_1

matBr	*prezime*	*ime*	*postBr*	*nazivGrad*
12345	Pirić	Damir	75000	Tuzla
45678	Pirić	Damir	71000	Sarajevo

$\pi_{prezime}(os_1) = os_2$

prezime
Pirić

A-2 UVEĆANJE

- Ako u relaciji r vrijedi $X \rightarrow Y$ i ako je $Z \subseteq R$ tada vrijedi i $XZ \rightarrow Y$ u r
  - Možemo uvećati lijevu stranu funkcijske zavisnosti
- U relaciji $os(OS)$ vrijedi $matBr \rightarrow ime$ i $prezime \subset OS$
  - U relaciji os vrijedi i $matBr, prezime \rightarrow ime$

os(*matBr*,	*prezime*,	*ime*,	*postBr*,	*nazivGrad*)
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

A-2 UVEĆANJE

- Ako r zadovoljava $X \to Y$, tada $\pi_Y(\sigma_{X=x}(r))$ sadrži najviše jednu n-torku za svaku vrijednost x od X
- Ako je $Z \subseteq R$ tada je $\sigma_{XZ=xz}(r) \subseteq \sigma_{X=x}(r)$, odnosno $\pi_Y(\sigma_{XZ=xz}(r)) \subseteq \pi_Y(\sigma_{X=x}(r))$
- Prema tome $\pi_Y(\sigma_{XZ=xz}(r))$ ima najviše jednu n-torku
  - r mora zadovoljavati $XZ \to Y$

A-2 UVEĆANJE

$X = \{postBr\}$
$Y = \{nazivGrad\}$
$Z = \{matBr\}$

os(matBr,	*prezime,*	*ime,*	*postBr,*	*nazivGrad)*
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

$\sigma_{postBr\ =\ 71000}(os) = oso_1$

oso₁(matBr,	*prezime,*	*ime,*	*postBr,*	*nazivGrad)*
23456	Đurić	Maja	71000	Sarajevo
45678	Pirić	Damir	71000	Sarajevo

$\pi_{nazivGrad}(oso_1) = oso_2$

(nazivGrad)
Sarajevo

$\sigma_{postBr\ =\ 71000\ \wedge\ matBr\ =\ 23456}(os) = oso_3$

oso₃(matBr,	*prezime,*	*ime,*	*postBr,*	*nazivGrad)*
23456	Đurić	Maja	71000	Sarajevo

$\pi_{nazivGrad}(oso_3) = oso_4$

(nazivGrad)
Sarajevo

A-3 TRANZITIVNOST

- Ako u relaciji r vrijedi $X \rightarrow Y$ i $Y \rightarrow Z$ tada vrijedi i $X \rightarrow Z$ u r
- $X \rightarrow Z$ je tranzitivna zavisnost
- U relaciji *os*(*OS*) vrijedi *matBr* $\rightarrow$ *postBr* i *postBr* $\rightarrow$ *nazivGrad*
  - U relaciji *os* vrijedi i *matBr* $\rightarrow$ *nazivGrad*

os(*matBr*,	*prezime*,	*ime*,	*postBr*,	*nazivGrad*)
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

A-3 TRANZITIVNOST

- Neka r zadovoljava $X \to Y$ i $Y \to Z$
- Neka su t_1 i t_2 n-torke iz r
  - $t_1(X) = t_2(X) \to t_1(Y) = t_2(Y)$
  - $t_1(Y) = t_2(Y) \to t_1(Z) = t_2(Z)$
- Zbog toga
  - $t_1(X) = t_2(X) \to t_1(Z) = t_2(Z)$
  - r zadovoljava $X \to Z$

A-3 TRANZITIVNOST

$X = \{matBr\}$
$Y = \{postBr\}$
$Z = \{nazivGrad\}$

	os(*matBr*,	*prezime*,	*ime*,	*postBr*,	*nazivGrad*)
	12345	Pirić	Damir	75000	Tuzla
t_1	23456	Đurić	Maja	71000	Sarajevo
	34567	Pejić	Dino	72000	Zenica
	45678	Pirić	Damir	71000	Sarajevo
t_2	23456	Đurić	Goran	71000	Sarajevo

$$t_1(matBr) = t_2(matBr) \rightarrow t_1(postBr) = t_2(postBr)$$
$$t_1(postBr) = t_2(postBr) \rightarrow t_1(nazivGrad) = t_2(nazivGrad)$$

- Zbog toga
  - $t_1(matBr) = t_2(matBr) \rightarrow t_1(nazivGrad) = t_2(nazivGrad)$
  - *os* zadovoljava $matBr \rightarrow nazivGrad$

Pravila koja proizilaze iz Armstrongovih aksioma

- Neka je *R* relacijska shema, neka su *X*, *Y*, *Z* i *V* skupovi atributa i neka vrijedi:
- $X \subseteq R$, $Y \subseteq R$, $Z \subseteq R$, $V \subseteq R$
- *Pravilo unije (pravilo o aditivnosti)*
- *Pravilo dekompozicije (pravilo o projektivnosti)*
- *Pravilo o pseudotranzitivnosti*

P-1 Pravilo unije

- Ako u relaciji r vrijedi $X \rightarrow Y$ i $X \rightarrow Z$, tada vrijedi i $X \rightarrow YZ$ u r ($X \rightarrow Y \cup Z$)

os(*matBr*,	*prezime*,	*ime*,	*postBr*,	*nazivGrad*)
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

- U relaciji *os* (*OS*) vrijedi *matBr* $\rightarrow$ *prezime* i *matBr* $\rightarrow$ *ime*
  - U relaciji *os* vrijedi i *matBr* $\rightarrow$ *ime*, *prezime*

Dokaz pravila unije

- Ako r zadovoljava $X \rightarrow Y$ i $X \rightarrow Z$, tada $\pi_Y(\sigma_{X=x}(r))$ i $\pi_Z(\sigma_{X=x}(r))$ imaju najviše jednu n-torku za svaku vrijednost x od X
- Ako $\pi_{YZ}(\sigma_{X=x}(r))$ ima više od jedne n-torke, tada bi barem jedna od projekcija $\pi_Y(\sigma_{X=x}(r))$ i $\pi_Z(\sigma_{X=x}(r))$ morala imati više od jedne n-torke
- Budući da $\pi_Y(\sigma_{X=x}(r))$ i $\pi_Z(\sigma_{X=x}(r))$ imaju najviše jednu n-torku
  - r zadovoljava $X \rightarrow YZ$

Primjer za pravilo unije

$X = \{matBr\}$
$Y = \{ime\}$
$Z = \{prezime\}$

os(matBr,	*prezime,*	*ime,*	*postBr,*	*nazivGrad)*
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

$\pi_{ime}(\sigma_{matBr\ =\ 12345}(os)) = oso_1$

(ime)
Damir

$\pi_{prezime}(\sigma_{matBr\ =\ 12345}(os)) = oso_2$

(prezime)
Pirić

$\pi_{ime,\ prezime}(\sigma_{matBr\ =\ 12345}(os)) = oso_3$

(ime,	*prezime)*
Damir	Pirić

P-2 Pravilo dekompozicije

- Ako u relaciji r vrijedi $X \rightarrow Y$ i ako je $V \subseteq Y$, tada vrijedi i $X \rightarrow V$ u r
- *Ili*
- Ako u relaciji r vrijedi $X \rightarrow YZ$, tada vrijedi i $X \rightarrow Y$

os(<u>*matBr*,</u>	<u>*prezime*,</u>	<u>*ime*,</u>	<u>*postBr*,</u>	<u>*nazivGrad*</u>)
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

- U relaciji *os* (*OS*) vrijedi *matBr* $\rightarrow$ *ime*, *prezime*
  - U relaciji *os* vrijedi i *matBr* $\rightarrow$ *prezime* i *matBr* $\rightarrow$ *ime*

Dokaz pravila dekompozicije

- Ako r zadovoljava $X \rightarrow YZ$, tada $\pi_{YZ}(\sigma_{X=x}(r))$ ima najviše jednu n-torku za svaku vrijednost x od X
- $\pi_Y(\pi_{YZ}(\sigma_{X=x}(r))) = \pi_Y(\sigma_{X=x}(r))$
- Budući da $\pi_Y(\sigma_{X=x}(r))$ može imati najviše jednu n-torku
  - r zadovoljava $X \rightarrow Y$

P-3 Pravilo o pseudotranzitivnosti

- Ako u relaciji *r* vrijedi $X \rightarrow Y$ i $VY \rightarrow Z$, tada vrijedi i $XV \rightarrow Z$ u *r*

os(<u>*matBr*,</u>	<u>*prezime*,</u>	<u>*ime*,</u>	<u>*postBr*,</u>	<u>*nazivGrad*,</u>	<u>*adresa*</u>)
12345	Pirić	Damir	75000	Tuzla	Franjevačka 5
23456	Đurić	Maja	71000	Sarajevo	Kulina bana 8
34567	Pejić	Dino	72000	Zenica	Aske Borića bb
45678	Pirić	Damir	71000	Sarajevo	Ferhadija 14

- U relaciji *os* (*OS*) vrijedi $matBr \rightarrow nazivGrad$ i $nazivGrad, adresa \rightarrow postBr$
  - U relaciji *os* vrijedi i $matBr, adresa \rightarrow postBr$

Dokaz pravila o pseudotranzitivnosti

- Neka r zadovoljava $X \rightarrow Y$ i $VY \rightarrow Z$
- Neka su t_1 i t_2 n-torke iz r
  - $t_1(X) = t_2(X) \rightarrow t_1(Y) = t_2(Y)$
  - $t_1(VY) = t_2(VY) \rightarrow t_1(Z) = t_2(Z)$
- Iz $t_1(XV) = t_2(XV)$ može se izvesti
  - $t_1(X) = t_2(X)$
  - $t_1(V) = t_2(V)$ $\quad \Rightarrow \quad t_1(Y) = t_2(Y)$
- Odnosno:
- $t_1(XV) = t_2(XV) \Rightarrow t_1(VY) = t_2(VY) \Rightarrow t_1(Z) = t_2(Z)$
  - r zadovoljava $XV \rightarrow Z$

Primjer korištenja aksioma i pravila

- Uz pretpostavku da na relacijskoj shemi R = { A, B, C, D, E, F } vrijedi skup funkcijskih zavisnosti F = {A $\rightarrow$ CD, AB $\rightarrow$ E, D $\rightarrow$ F }, dokazati da vrijedi funkcijska zavisnost AB $\rightarrow$ EF

- A $\rightarrow$ CD (*pravilo dekompozicije*) $\Rightarrow$ A $\rightarrow$ C $\wedge$ A $\rightarrow$ D
- A $\rightarrow$ D $\wedge$ D $\rightarrow$ F (*aksiom o tranzitivnosti*) $\Rightarrow$ A $\rightarrow$ F
- A $\rightarrow$ F (*aksiom o uvećanju*) $\Rightarrow$ AB $\rightarrow$ F
- AB $\rightarrow$ E $\wedge$ AB $\rightarrow$ F (*pravilo unije*) $\Rightarrow$ AB $\rightarrow$ EF

Pravilo o akumulaciji

- Dodatno pravilo koje omogućuje "algoritamski" pristup rješavanju sličnih zadataka

- Ako u relaciji r vrijedi $X \rightarrow VZ$ i $Z \rightarrow W$, tada vrijedi i $X \rightarrow VZW$ u r

- $X \rightarrow VZ$ (*pravilo dekompozicije*) $\Rightarrow X \rightarrow V \wedge X \rightarrow Z$
- $X \rightarrow Z \wedge Z \rightarrow W$ (*aksiom o tranzitivnosti*) $\Rightarrow X \rightarrow W$
- $X \rightarrow VZ \wedge X \rightarrow W$ (*pravilo unije*) $\Rightarrow X \rightarrow VZW$

Primjer korištenja pravila o akumulaciji

- Uz pretpostavku da na relacijskoj shemi R = { A, B, C, D, E } vrijedi skup funkcijskih zavisnosti F = {A → BD, B → C, D → E }, dokazati da vrijedi funkcijska zavisnost AE → AC
  - Označimo lijevu stranu FZ s X (X=AE), a desnu stranu FZ s Y (Y=AC)
  - Dokaz se izvodi primjenom aksioma o refleksivnosti, pravila o akumulaciji i pravila o dekompoziciji
  - 1. korak: X → X
    - AE → AE
  - 2. korak: pomoću pravila akumulacije "uvećavati desnu stranu FZ" sve dok desna strana ne sadrži Y
    - AE → AE ∧ A → BD ⇒ AE → AEBD
    - AE → AEBD ∧ B → C ⇒ AE → AEBDC
  - 3. korak: kad (i ako) desna strana sadrži Y primijeniti pravilo dekompozicije
    - AE → AEBDC ⇒ AE → AC

Primjer korištenja pravila o akumulaciji (za vježbu)

- Uz pretpostavku da na relacijskoj shemi R = { L, M, N, P, Q, R } vrijedi skup funkcijskih zavisnosti F = {Q → R, M → PQ, PQL → N }, dokazati da vrijedi funkcijska zavisnost MLR → QN

- Uz pretpostavku da na relacijskoj shemi R = { L, M, N, P, Q, R } vrijedi skup funkcijskih zavisnosti F = {Q → R, M → PQ, PQL → N }, dokazati da vrijedi funkcijska zavisnost MQ → LN

Priroda funkcijskih zavisnosti - podsjetnik

- Postojanje funkcijske zavisnosti ne može se dokazati na osnovu postojećih podataka u relaciji
- Analizom podataka moguće je pretpostaviti da funkcijska zavisnost postoji
- Dokaz za postojanje funkcijske zavisnosti treba tražiti u značenju pojedinih atributa
- Funkcijske zavisnosti neće se promijeniti promjenom sadržaja relacije
- Dodavanjem ili ukidanjem funkcijskih zavisnosti **mijenja se model baze podataka**
- Prilikom konstrukcije modela javljaju se redundatne funkcijske zavisnosti koje je potrebno ukloniti

Potpuna funkcijska zavisnost

- Zadana je relacijska shema R i skupovi atributa X i Y iz R (X, $Y \subseteq R$)
- Neka u R vrijedi funkcijska zavisnost $X \rightarrow Y$
- Atribut ili skup atributa Y **potpuno je funkcijski ovisan** o atributu ili skupu atributa X relacijske sheme R ako:
  - Y funkcijski ovisi o čitavom X
  - Ne postoji pravi podskup od X koji funkcijski određuje Y

Nepotpuna funkcijska zavisnost

- Zadana je relacijska shema R i skupovi atributa X i Y iz R ($X, Y \subseteq R$)
- Neka u R vrijedi funkcijska zavisnost $X \rightarrow Y$
- Funkcijska zavisnost $X \rightarrow Y$ je **nepotpuna** ako postoji atribut ili skup atributa Z koji je podskup od X, za koji vrijedi $Z \rightarrow Y$
- Ili
- FZ $X \rightarrow Y$ je **nepotpuna** ako
- $(\exists Z)(Z \subset X) : Z \rightarrow Y$

Tranzitivna funkcijska zavisnost

- Zadano je:
  - Relacijska shema R
  - Skupovi atributa $X, Y, Z \in R$
  - Skup funkcijskih zavisnosti F

- Atribut Z je tranzitivno ovisan o X ako vrijedi:
  - $X \rightarrow Y$
  - $Y \nrightarrow X$ i
  - $Y \rightarrow Z$ (proizilazi iz F)
  - $Z \not\subset XY$

Ključ relacije

- Relacijom se opisuje skup entiteta
- Entitet je "nešto" što se prema svojim karakteristikama može razlučiti od okoline
- Ključ relacije sadrži one atribute koji omogućuju da se pojedini entiteti mogu razlučiti od okoline
- Ključ relacije je atribut ili skup atributa koji nedvosmisleno određuje n-torke relacije
- Ključ relacijske sheme R je skup atributa K, K $\subseteq$ R, koji ima sljedeća svojstva:
  - $K \rightarrow (R \setminus K)$ (također vrijedi i $K \rightarrow R$)
    - Ključ relacije ima svojstvo da funkcijski određuje atribute u preostalom dijelu relacijske sheme
  - Ne postoji $K' \subset K$ za kojeg vrijedi $K' \rightarrow R$
    - Ključ je **minimalan** skup atributa koji funkcijski određuje atribute u preostalom dijelu relacijske sheme

Ključ relacije - primjer

radnik(<u>*matBr*</u>, *prezime*, *ime*, *jmbg*)

matBr	prezime	ime	jmbg
12345	Pirić	Damir	1710977180025
23456	Đurić	Maja	1812982185011
34567	Pejić	Dino	0209979180016
45678	Pirić	Damir	2701981180001

- Ključ: K_{radnik} = {*matBr*}
- matBr $\rightarrow$ prezime
- matBr $\rightarrow$ ime
- matBr $\rightarrow$ jmbg

- Za: K = {*matBr, prezime*} također vrijedi
- K $\rightarrow$ { ime, jmbg }
- Ali K nije ključ jer postoji K' = {matBr}, K' $\subset$ K, za kojeg vrijedi
- K' $\rightarrow$ { prezime, ime, jmbg }

Ključevi relacije

- Mogući ključ (*candidate key*)
- Primarni ključ (*primary key*) odabire se jedan od mogućih ključeva
- Alternativni ključ (*alternate key*) ostali mogući ključevi

radnik

matBr	*prezime*	*ime*	*jmbg*
12345	Pirić	Damir	1710977180025
23456	Đurić	Maja	1812982185011
34567	Pejić	Dino	0209979180016
45678	Pirić	Damir	2701981180001

- Mogući ključevi: {*matBr*}, {*jmbg*}
- Primarni ključ: {*matBr*}
- Alternativni ključ: {*jmbg*}

Struktura relacije

- Relacijska shema se sastoji od
  - atributa koji su dio ključa (ključni atributi, ključni dio relacije)
  - atributa iz zavisnog dijela relacije (neključni atributi, neključni dio relacije)
- Ključ relacije ima svojstvo da **funkcijski određuje preostali dio relacije**

OSOBA(*matBr*, *prezime*, *ime*, *postBr*, *nazivGrad*)

matBr	*prezime*	*ime*	*postBr*	*nazivGrad*
12345	Pirić	Damir	75000	Tuzla
23456	Đurić	Maja	71000	Sarajevo
34567	Pejić	Dino	72000	Zenica
45678	Pirić	Damir	71000	Sarajevo

- Ključ: $K_{OSOBA} = \{matBr\}$
  - $matBr \rightarrow prezime$ $matBr \rightarrow ime$
  - $matBr \rightarrow postBr$ $matBr \rightarrow nazivGrad$