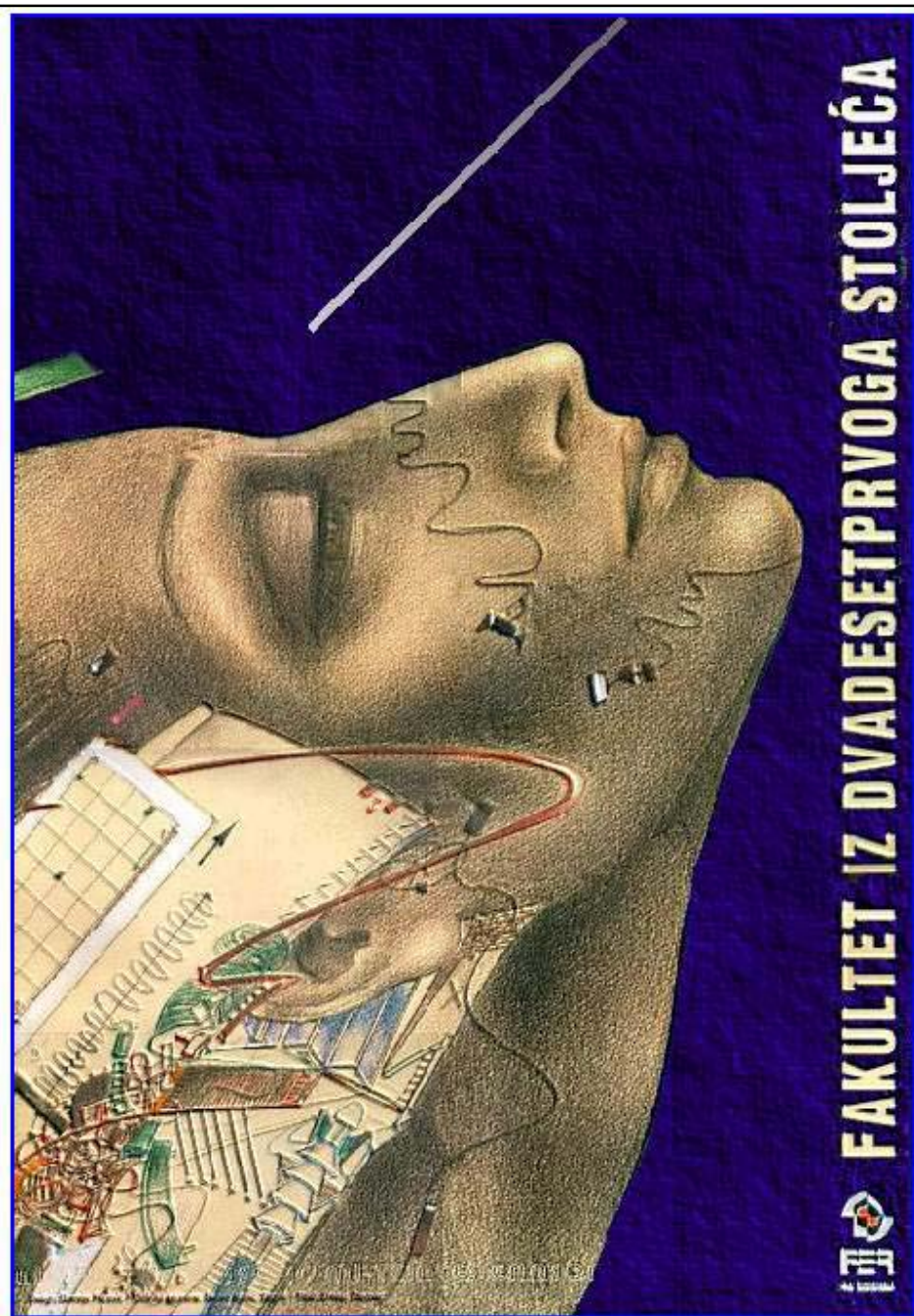


# Baze podataka

Predavanja  
travanj 2008.

## 7. Oblikovanje sheme relacijske baze podataka (2. dio)



---

# Normalizacija

# Postupci normalizacije

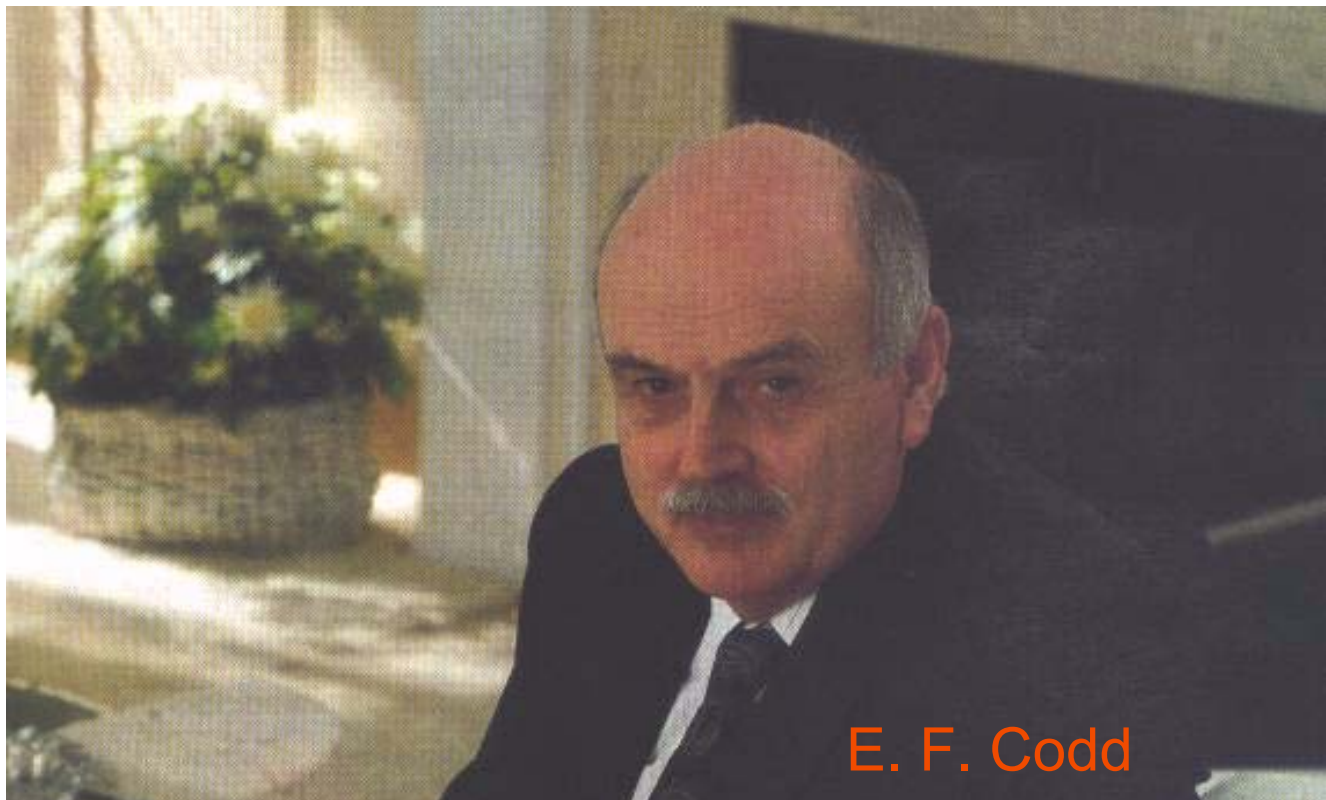
---

- Uočena znanja o međusobnim **funkcijskim zavisnostima** atributa relacije koriste se u postupcima normalizacije.
- **Cilj:**
  - **ukloniti redundanciju**
    - anomalije unosa, izmjene i brisanja
    - neracionalno korištenje prostora za pohranu
  - **spriječiti pojavu lažnih n-torki**
- Postupci normalizacije omogućavaju da se postupno, **točno definiranom metodom**, odredi dobra zamjena za loše koncipiranu relacijsku shemu

**E. F. Codd:**

**“Normalized data base structure: A brief tutorial”**

*Proc. ACM SIGFIDET Workshop on Data Description, Access and Control, 1971*



E. F. Codd

"I called it normalization because then-President Nixon was talking a lot about normalizing relations with China. I figured that if he could normalize relations, so could I."  
(*'A "FIRESIDE" CHAT'*, DBMS, Dec. 1993)

# Normalne forme

---

- Prva normalna forma - 1 NF
  - Druga normalna forma - 2 NF
  - Treća normalna forma - 3 NF
  - Boyce-Coddova normalna forma - BCNF
- Temelje se na FUNKCIJSKIM ZAVISNOSTIMA
- 

- Četvrta normalna forma - 4NF
- Temelji se na VIŠEZNAČNIM ZAVISNOSTIMA
- Projekcijsko-spojna normalna forma - PJNF
- Temelji se na SPOJNIM ZAVISNOSTIMA

# Postupci normalizacije

---

- Dekompozicija
  - početne relacije (relacijske sheme) se dekomponiraju na temelju uočenih funkcijskih zavisnosti
  
- Sinteza
  - zadan je skup atributa i nad njima skup funkcijskih zavisnosti iz kojih se sintetiziraju relacijske sheme koje zadovoljavaju 3NF

# Dekompozicija relacijske sheme (relacije)

- Dekompozicijom (razlaganjem) relacijska shema  $R$  zamjenjuje se shemama  $R_1, R_2, \dots, R_n$ ,  $R_i \subseteq R$ , pri čemu vrijedi  $R = R_1 \bowtie R_2 \dots R_n$
- Dekompozicijom se relacija  $r(R)$  zamjenjuje relacijama  $r_1(R_1), r_2(R_2), \dots, r_n(R_n)$ , pri čemu je  $r_i(R_i) = \pi_{R_i}(r)$ , za  $i = 1, \dots, n$
- Relacija  $r(R)$  se dekomponira na relacije  $r_1(R_1), r_2(R_2), \dots, r_n(R_n)$  **bez gubitaka informacija** (*lossless decomposition*) ako vrijedi:

$$r_1(R_1) \bowtie r_2(R_2) \bowtie \dots \bowtie r_n(R_n) = r(R)$$

odnosno

$$\pi_{R_1}(r) \bowtie \pi_{R_2}(r) \bowtie \dots \bowtie \pi_{R_n}(r) = r(R)$$



# Dekompozicija relacije - primjer

- Zadana je relacija:

r(R)

A	B	C	D
a1	b1	c1	d1
a2	b2	c1	d1

- Relaciju r(R) dekomponirati na relacije

- $r_1(R_1)$ ,  $R_1 = \{ A, C \}$
- $r_2(R_2)$ ,  $R_2 = \{ B, C \}$
- $r_3(R_3)$ ,  $R_3 = \{ C, D \}$

$r_1(R_1)$

A	C
a1	c1
a2	c1

$r_2(R_2)$

B	C
b1	c1
b2	c1

$r_3(R_3)$

C	D
c1	d1

- Je li dekompozicija obavljena bez gubitaka informacija?

$r_1(R_1) \bowtie r_2(R_2) \bowtie r_3(R_3)$

A	B	C	D
a1	b1	c1	d1
a1	b2	c1	d1
a2	b1	c1	d1
a2	b2	c1	d1

⇒ NE



# Razlaganje relacije bez gubitaka na dvije projekcije

- Relacija se bez gubitaka razlaže na svoje dvije projekcije ako:
  - projekcije imaju zajedničke atribute
  - zajednički atributi su ključ u barem jednoj od projekcija

PRIMJER:

osoba	matBr	prez	ime	postBr	nazMj
	11234	Novak	Josip	21000	Split
	12345	Horvat	Ivan	10000	Zagreb
	23456	Kolar	Ana	31000	Osijek
	34567	Novak	Josip	10000	Zagreb

$$\text{osoba}_1 = \pi_{\text{matBr, prez, ime, postBr}} (\text{osoba})$$

$$\text{mjesto} = \pi_{\text{postBr, nazMj}} (\text{osoba})$$

- Hoće li se relacija osoba dekomponirati bez gubitaka informacija na relacije osoba<sub>1</sub> i mjesto? Odnosno, vrijedi li:

$$\text{osoba} \equiv \text{osoba}_1 \bowtie \text{mjesto}$$

# Primjer razlaganja relacije na dvije projekcije

osoba<sub>1</sub>

matBr	prez	ime	postBr
11234	Novak	Josip	21000
12345	Horvat	Ivan	10000
23456	Kolar	Ana	31000
34567	Novak	Josip	10000

$OSOBA_1 = \{ \text{matBr}, \text{prez}, \text{ime}, \text{postBr} \}$

$K_{OSOBA_1} = \{ \text{matBr} \}$

mjesto

postBr	nazMj
21000	Split
10000	Zagreb
31000	Osijek

$MJESTO = \{ \text{postBr}, \text{nazMj} \}$

$K_{MJESTO} = \{ \text{postBr} \}$

$OSOBA_1 \cap MJESTO = \{ \text{postBr} \}$

$\Rightarrow \text{osoba} \equiv \text{osoba}_1 \triangleright \triangleleft \text{mjesto}$

osoba

matBr	prez	ime	postBr	nazMj
11234	Novak	Josip	21000	Split
12345	Horvat	Ivan	10000	Zagreb
23456	Kolar	Ana	31000	Osijek
34567	Novak	Josip	10000	Zagreb

# Prva normalna forma (1NF)

---

- Definicija:

Relacijska shema je u **1NF** ako:

- domene atributa sadrže samo jednostavne (nedjeljive) vrijednosti
  - vrijednost svakog atributa je samo jedna vrijednost iz domene tog atributa
  - neključni atributi relacije funkcijski ovise o ključu relacije
- Shema baze podataka  $\mathbf{R} = \{ R_1, R_2, \dots, R_n \}$  je u 1NF ako je svaka relacijska shema  $R_1, R_2, \dots, R_n$  u 1NF

# Prva normalna forma - primjer

- Poduzeće evidentira podatke o radnicima
  - $RADNIK = \{ matBr, prezime, ime, datRod, sifOdjel \}$

radnik (RADNIK)	matBr	prezime	ime	datRod	sifOdjel
	1111	Novak	Ivan	28.12.1970	50
	1121	Kolar	Iva	16.10.1965	30
	1133	Horvat	Krešo	19.03.1978	50

$K_{RADNIK} = \{ matBr \}$

- domene svih atributa sadrže jednostavne (nedjeljive) vrijednosti
- vrijednost svakog atributa je samo jedna vrijednost iz domene tog atributa
- neključni atributi relacije funkcijski ovise o ključu relacije

⇒ Relacijska shema RADNIK je u 1NF

# Prva normalna forma - primjer

- Poduzeće evidentira podatke o radnicima i njihovoj djeci - korisnicima zdravstvenog osiguranja.
  - $RADNIK_1 = \{ matBr, prezime, ime, imenaDjece \}$

$radnik_1(RADNIK_1)$	matBr	prezime	ime	imenaDjece
$K_{RADNIK_1} = \{ matBr \}$	1111	Novak	Ivan	Jasna, Vedran
	1121	Kolar	Iva	Ivan
	1133	Horvat	Krešo	Petar, Ana, Ivan

- domena atributa *imenaDjece* ne sadrži jednostavne (nedjeljive vrijednosti)

⇒ Relacijska shema  $RADNIK_1$  nije u 1NF

# Prva normalna forma - primjer

- Poduzeće evidentira podatke o radnicima i njihovoj djeci - korisnicima zdravstvenog osiguranja.
  - $RADNIK_2 = \{ matBr, prezime, ime, imeDj, datRodDj \}$

$K_{RADNIK_2} = \{ matBr \}$	radnik <sub>2</sub> (RADNIK <sub>2</sub> )				
	matBr	prezime	ime	imeDj	datRodDj
	1111	Novak	Ivan	Jasna Vedran	21.01.1995 13.12.1997
	1121	Kolar	Iva	Ivan	23.03.2000
	1133	Horvat	Krešo	Petar Ana Ivan	22.02.1998 19.09.2000 05.11.2002

- domene sadrže jednostavne vrijednosti, ali vrijednost atributa *imeDj* nije uvijek samo jedna vrijednost iz domene tog atributa (isto vrijedi i za atribut *datRodDj*)

⇒ Relacijska shema RADNIK<sub>2</sub> nije u 1NF

# Normalizacija na 1NF - izdvajanjem atributa u posebnu relaciju

- u posebnu relaciju izdvaja se skup atributa koji se ponavlja s jednakom kratnošću, zajedno s ključem originalne relacije

$RADNIK_2 = \{ matBr, prezime, ime, imeDj, datRodDj \}$       $K_{RADNIK_2} = \{ matBr \}$

radnik<sub>3</sub> (RADNIK<sub>3</sub>)

matBr	prezime	ime
1111	Novak	Ivan
1121	Kolar	Iva
1133	Horvat	Krešo

$RADNIK_3 = \{ matBr, prezime, ime \}$

$K_{RADNIK_3} = \{ matBr \}$

dijete (DIJETE)

matBr	imeDj	datRodDj
1111	Jasna	21.01.1995
1111	Vedran	13.12.1997
1121	Ivan.	23.03.2000
1133	Petar	22.02.1998
1133	Ana	19.09.2000
1133	Ivan	05.11.2002

$DIJETE = \{ matBr, imeDj, datRodDj \}$

$K_{DIJETE} = \{ matBr, imeDj \}$

- operacija je izvedena bez gubitaka informacija - relacijske sheme imaju zajedničke attribute (matBr), zajednički atributi su ključ u RADNIK<sub>3</sub>



# Normalizacija na 1NF - promjenom ključa

$RADNIK_2 = \{ matBr, prezime, ime, imeDj, datRodDj \}$        $K_{RADNIK_2} = \{ matBr \}$

$RADNIK_4 = \{ matBr, prezime, ime, imeDj, datRodDj \}$

$K_{RADNIK_4} = \{ matBr, imeDj \}$

radnik<sub>4</sub> (RADNIK<sub>4</sub>)

matBr	prezime	ime	imeDj	datRodDj
1111	Novak	Ivan	Jasna	21.01.1995
1111	Novak	Ivan	Vedran	13.12.1997
1121	Kolar	Iva	Ivan	23.03.2000
1133	Horvat	Krešo	Petar	22.02.1998
1133	Horvat	Krešo	Ana	19.09.2000
1133	Horvat	Krešo	Ivan	05.11.2002

# Druga normalna forma (2NF)

---

- Definicija:

Relacijska shema  $R$  je u **2NF** ako je u 1NF i ako je

- **svaki atribut iz zavisnog dijela potpuno funkcijski ovisan o svakom ključu relacije**

- Shema baze podataka  $\mathbf{R} = \{ R_1, R_2, \dots, R_n \}$  je u 2NF ako je svaka relacijska shema  $R_1, R_2, \dots, R_n$  u 2NF

# Potpuna funkcijska zavisnost

- Skup atributa Y **potpuno je funkcijski ovisan** o skupu atributa X relacijske sheme R ako:
  - Y funkcijski ovisi o X
  - ne postoji pravi podskup od X koji funkcijski određuje Y

PRIMJER:

Zadan je skup FZ  $F = \{ ABC \rightarrow DE, E \rightarrow F \}$ .

Je li  $\{ D, E \}$  potpuno funkcijski ovisan o  $\{ A, B, C \}$  ?

Da, jer ne postoji skup  $Z \subset \{ A, B, C \}$  takav da  $Z \rightarrow \{ D, E \}$

# Nepotpuna funkcijska zavisnost

Zadana je relacijska shema  $R$  i skupovi atributa  $X$  i  $Y$  iz  $R$ , tj.  $X \subseteq R$ ,  $Y \subseteq R$ . Neka u  $R$  vrijedi FZ  $X \rightarrow Y$ .

FZ  $X \rightarrow Y$  je **nepotpuna** ako postoji skup atributa  $Z$  koji je **podskup od  $X$** , za koji vrijedi  $Z \rightarrow Y$

odnosno

FZ  $X \rightarrow Y$  je **nepotpuna** ako  $(\exists Z) (Z \subset X) : Z \rightarrow Y$

PRIMJER:

Zadan je skup FZ  $F = \{ ABC \rightarrow D, BC \rightarrow E, E \rightarrow D \}$ .

Je li  $\{ D \}$  potpuno funkcijski ovisan o  $\{ A, B, C \}$  ?

Ne, jer postoji skup  $\{ B, C \} \subset \{ A, B, C \}$  takav da  $\{ B, C \} \rightarrow \{ D \}$

# Druga normalna forma - primjer

$RADNIK_4 = \{ matBr, prezime, ime, imeDj, datRodDj \}$       $K_{RADNIK_4} = \{ matBr, imeDj \}$

radnik <sub>4</sub> (RADNIK <sub>4</sub> )	matBr	prezime	ime	imeDj	datRodDj
	1111	Novak	Ivan	Jasna	21.01.1995
	1111	Novak	Ivan	Vedran	13.12.1997
	1121	Kolar	Iva	Ivan.	23.03.2000
	1133	Horvat	Krešo	Petar	22.02.1998
	1133	Horvat	Krešo	Ana	19.09.2000
	1133	Horvat	Krešo	Ivan	05.11.2002

- relacijska shema  $RADNIK_4$  zadovoljava 1NF
- **postoji FZ:**    **matBr** → **prezime ime**
- **matBr imeDj** → **prezime ime**    je nepotpuna FZ!

⇒ Relacijska shema  $RADNIK_4$  nije u 2NF

# Normalizacija na 2NF

- Normalizacijom na 2NF nastaju:
  - relacijska shema koja sadrži skup atributa koji su bili nepotpuno funkcijski ovisni o ključu i dio ključa o kojem su potpuno funkcijski ovisni
  - relacijska shema koja sadrži ključ originalne relacije i skup atributa koji su potpuno funkcijski ovisni o ključu

$RADNIK_5 = \{ matBr, prezime, ime \}$

$K_{RADNIK_5} = \{ matBr \}$

radnik<sub>5</sub> (RADNIK<sub>5</sub>)

matBr	prezime	ime
1111	Novak	Ivan
1121	Kolar	Iva
1133	Horvat	Krešo

$DIJETE = \{ matBr, imeDj, datRodDj \}$

$K_{DIJETE} = \{ matBr, imeDj \}$

dijete (DIJETE)

matBr	imeDj	datRodDj
1111	Jasna	21.01.1995
1111	Vedran	13.12.1997
1121	Ivan.	23.03.2000
1133	Petar	22.02.1998
1133	Ana	19.09.2000
1133	Ivan	05.11.2002

# Normalizacija na 2NF

Neka su X, Y, Z, V atributi ili skupovi atributa. Zadana je relacijska shema  $R = XYZV$  i na njoj skup funkcijskih zavisnosti  $F = \{ XY \rightarrow ZV, X \rightarrow Z \}$ . Ključ relacije  $K_R = XY$ . R je u 1NF. **Zadovoljava li R 2NF?**

- funkcijska zavisnost  $XY \rightarrow Z$  je nepotpuna  
**R ne zadovoljava 2NF**

Normalizacijom na 2NF shema R se zamjenjuje shemama:

$$R_1 = XZ$$

$$R_2 = XYV$$

$$K_{R_1} = X$$

$$K_{R_2} = XY$$

Relacija  $r(R)$  se normalizacijom na 2NF zamjenjuje projekcijama:

$$r_1 = \pi_{XZ}(r) \quad r_2 = \pi_{XYV}(r)$$

- operacija je izvedena bez gubitaka informacija - relacijske sheme imaju zajedničke attribute (X), zajednički atributi su ključ u  $R_1$ .



# Treća normalna forma (3NF)

---

- Definicija:

Relacijska shema je u **3NF** ako je u 1NF i ako:

- niti jedan atribut iz zavisnog dijela nije tranzitivno funkcijski ovisan o bilo kojem ključu relacije
- Shema baze podataka  $\mathbf{R} = \{ R_1, R_2, \dots, R_n \}$  je u 3NF ako je svaka relacijska shema  $R_1, R_2, \dots, R_n$  u 3NF

# Tranzitivna funkcijska zavisnost

---

Zadano je:

- relacijska shema  $R$ ,
- skupovi atributa  $X \subseteq R$ ,  $Y \subseteq R$ ,  $Z \subseteq R$

Skup atributa  $Z$  je tranzitivno ovisan o  $X$  ako vrijedi:

- $X \rightarrow Y$ ,  $Y \not\rightarrow X$  i  $Y \rightarrow Z$

# Tranzitivna funkcijska zavisnost - primjer

---

Zadana je relacijska shema  $R = \{ A, B, C, D, E, F, G \}$  i skup FZ  $F = \{ AB \rightarrow CD, D \rightarrow EF, CD \rightarrow ABG \}$

EF je tranzitivno funkcijski ovisan o AB, jer

- $AB \rightarrow D, D \not\rightarrow AB, D \rightarrow EF$

G nije tranzitivno funkcijski ovisan o AB, jer iako

- $AB \rightarrow CD, CD \rightarrow G$
- nije zadovoljen uvjet  $CD \not\rightarrow AB$

# Treća normalna forma - primjer

OSOBA={ matBr, prez, ime, postBr, nazMjesto }     $K_{OSOBA}=\{ \text{matBr} \}$

osoba (OSOBA)	matBr	prez	ime	postBr	nazMjesto
	1111	Novak	Ivan	10000	Zagreb
	1121	Kolar	Iva	31000	Osijek
	1133	Horvat	Krešo	10000	Zagreb

- Relacijska shema OSOBA zadovoljava 1NF.
  - vrijedi FZ:     $\text{matBr} \rightarrow \text{postBr}$
  - vrijedi FZ:     $\text{postBr} \rightarrow \text{nazMjesto}$
  - ne vrijedi FZ:  $\text{postBr} \rightarrow \text{matBr}$
- $\text{matBr} \rightarrow \text{nazMjesto}$  je tranzitivna zavisnost !
- Relacijska shema OSOBA ne zadovoljava 3NF.

# Normalizacija na 3NF

OSOBA={ matBr, prez, ime, postBr, nazMjesto }     $K_{OSOBA}=\{ \text{matBr} \}$

Normalizacijom na 3NF nastaju:

- relacijska shema koja sadrži skup atributa relacijske sheme OSOBA koji su tranzitivno ovisni o ključu (*nazMjesto*) te srednji skup atributa uočene tranzitivne zavisnosti (*postBr*)
- relacijska shema koja sadrži ključ relacijske sheme OSOBA (*matBr*) i neključne attribute relacijske sheme OSOBA koji nisu tranzitivno ovisni o ključu

MJESTO={ postBr, nazMjesto }

$K_{MJESTO}=\{ \text{postBr} \}$

mjesto (MJESTO)

postBr	nazMjesto
10000	Zagreb
31000	Osijek

OSOBA<sub>1</sub>={ matBr, prezime, ime, postBr }

$K_{OSOBA_1}=\{ \text{matBr} \}$

osoba<sub>1</sub> (OSOBA<sub>1</sub>)

matBr	prezime	ime	postBr
1111	Novak	Ivan	10000
1121	Kolar	Iva	31000
1133	Horvat	Krešo	10000

# Normalizacija na 3NF

Neka su X, Y, Z, V atributi ili skupovi atributa. Zadana je relacijska shema  $R = XYZV$  i na njoj skup funkcijskih zavisnosti  $F = \{ X \rightarrow YZV, Z \rightarrow V \}$ . Ključ relacije  $K_R = X$ . R je u 1NF. **Zadovoljava li R 3NF?**

- funkcijska zavisnost  $X \rightarrow V$  je tranzitivna  
R ne zadovoljava 3NF

Normalizacijom na 3NF shema R se zamjenjuje shemama:

$$R_1 = XYZ$$
$$K_{R_1} = X$$

$$R_2 = ZV$$
$$K_{R_2} = Z$$

Relacija  $r(R)$  se normalizacijom na 3NF zamjenjuje projekcijama:

$$r_1 = \pi_{XYZ}(r) \quad r_2 = \pi_{ZV}(r)$$

- operacija je izvedena bez gubitaka informacija - relacijske sheme imaju zajedničke attribute (Z), zajednički atributi su ključ u  $R_2$ .

# Treća normalna forma - komentar

Normalizacija na 2NF nije nužni preduvjet za provođenje normalizacije na 3NF jer se nepotpune FZ mogu promatrati kao tranzitivne FZ.

**Primjer:** zadana je shema  $R = XYZV$  i na njoj skup funkcijskih zavisnosti  $F = \{ XY \rightarrow ZV, X \rightarrow Z \}$ . Ključ relacije  $K_R = XY$ .  $R$  je u 1NF, ali nije u 2NF jer postoji nepotpuna FZ  $XY \rightarrow Z$ . Međutim, postoji i tranzitivna funkcijska zavisnost  $XY \rightarrow Z$  ( $XY \rightarrow X \wedge X \rightarrow Z$ ).

Normalizacijom na 3NF shema  $R$  se zamjenjuje shemama:

$$R_1 = XZ$$

$$K_{R_1} = X$$

$$R_2 = XYV$$

$$K_{R_2} = XY$$

$R_1$  i  $R_2$  su u 2NF i 3NF

**Preporuka: normalizaciju ipak obavljati postupno**

**$1NF \Rightarrow 2NF \Rightarrow 3NF$**



# Normalizacija na 3NF - primjer

$OSOBA_2 = \{ \text{matBr}, \text{prezime}, \text{ime}, \text{JMBG} \}$

osoba2 (OSOBA2)	matBr	prez	ime	JMBG
	1111	Novak	Ivan	1403970330103
	1121	Kolar	Iva	2812968310267
	1133	Horvat	Krešo	0301979320319

- postoji FZ:  $\text{matBr} \rightarrow \text{prez ime JMBG}$
- postoje FZ:  $\text{JMBG} \rightarrow \text{prez ime}$  i  $\text{JMBG} \rightarrow \text{matBr}$
- $\text{matBr}$  i  $\text{JMBG}$  su mogući ključevi
- Relacijska shema  $OSOBA_2$  zadovoljava 3NF.

$K1_{OSOBA2} = \{ \text{matBr} \}$

$K2_{OSOBA2} = \{ \text{JMBG} \}$

# Normalizacija na 3NF - dodatna razmatranja

Neka su  $X, Y, Z, V$  atributi ili skupovi atributa. Zadana je relacijska shema  $R = XYZV$  i na njoj skup funkcijskih zavisnosti  $F = \{ X \rightarrow YZV, Z \rightarrow V, Z \rightarrow X \}$ .  $R$  je u 1NF. Neka je ključ  $K_R = X$ .

- vrijedi  $X \rightarrow Z$  i  $Z \rightarrow V$ , ali  $X \rightarrow V$  nije tranzitivna FZ jer vrijedi i  $Z \rightarrow X$
- Zbog  $X \rightarrow Z$  i  $Z \rightarrow X$  funkcijsku zavisnost  $X \rightarrow V$  nije potrebno ukloniti jer u tom slučaju nema redundancije.
- $Z$  je također mogući ključ u  $R$

$$K1_R = X$$

$$K2_R = Z$$

$X$  i  $Z$  su mogući ključevi.

- Relacijska shema  $R$  zadovoljava 3NF.

# 1. primjer normalizacije

- Zadana je relacijska shema:

ISPIT = { matBr, prez, ime, sifPred, nazPred, datlsp, ocj, sifNas, prezNas }

i trenutna vrijednost relacije ispit(ISPIT):

ispit (ISPIT)

matBr	prez	ime	sifPred	nazPred	datlsp	ocj	sifNas	prezNas
1111	Novak	Ivan	1001	Mat-1	29.01.06	1	1111	Pašić
1111	Novak	Ivan	1001	Mat-1	05.02.06	3	1111	Pašić
1111	Novak	Ivan	1003	Fiz-1	28.06.06	2	3333	Horvat
1111	Novak	Ivan	1002	Mat-2	27.06.06	4	2222	Brnetić
1234	Kolar	Petar	1001	Mat-1	29.01.06	3	2222	Brnetić

- funkcijske zavisnosti odrediti na temelju značenja podataka
- odrediti primarni ključ relacije (tako da bude zadovoljen uvjet 1NF prema kojem neključni atributi funkcijski ovise o ključu)
- postupno normalizirati relacijsku shemu ISPIT na 2NF i 3NF

# 1. primjer normalizacije - 1NF

ispit (ISPIT)

matBr	prez	ime	sifPred	nazPred	datlsp	ocj	sifNas	prezNas
1111	Novak	Ivan	1001	Mat-1	29.01.06	1	1111	Pašić
1111	Novak	Ivan	1001	Mat-1	05.02.06	3	1111	Pašić
1111	Novak	Ivan	1003	Fiz-1	28.06.06	2	3333	Horvat
1111	Novak	Ivan	1002	Mat-2	27.06.06	4	2222	Brnetić
1234	Kolar	Petar	1001	Mat-1	29.01.06	3	2222	Brnetić

- Određivanje ključa: ako se (pogrešno) pretpostavi da je  $K = \{ \text{matBr} \}$   
Bi li tada postojali neključni atributi koje ključ funkcijski ne određuje?
- $\text{matBr} \rightarrow \text{prez ime}$   
**međutim:**
- $\text{matBr} \nrightarrow \text{sifPred}$        $\text{matBr} \nrightarrow \text{nazPred}$   
 $\text{matBr} \nrightarrow \text{datlsp}$        $\text{matBr} \nrightarrow \text{ocj}$   
 $\text{matBr} \nrightarrow \text{sifNas}$        $\text{matBr} \nrightarrow \text{prezNas}$

# 1. primjer normalizacije - 1NF

ispit (ISPIT)

matBr	prez	ime	sifPred	nazPred	datlsp	ocj	sifNas	prezNas
1111	Novak	Ivan	1001	Mat-1	29.01.06	1	1111	Pašić
1111	Novak	Ivan	1001	Mat-1	05.02.06	3	1111	Pašić
1111	Novak	Ivan	1003	Fiz-1	28.06.06	2	3333	Horvat
1111	Novak	Ivan	1002	Mat-2	27.06.06	4	2222	Brnetić
1234	Kolar	Petar	1001	Mat-1	29.01.06	3	2222	Brnetić

- Ako se pretpostavi  $K = \{ \text{matBr}, \text{sifPred}, \text{datlsp} \}$

Bi li tada postojali **neključni atributi koje ključ funkcijski ne određuje?**

- $\text{matBr sifPred datlsp} \rightarrow \text{prez ime nazPred ocj sifNas prezNas}$

postoji li skup  $X \subset \{ \text{matBr}, \text{sifPred}, \text{datlsp} \}$  za kojeg vrijedi  $X \rightarrow R$  ?

$\Rightarrow$  **NE**  $\Rightarrow \{ \text{matBr}, \text{sifPred}, \text{datlsp} \}$  je mogući ključ

$$K_{\text{ISPIT}} = \{ \text{matBr}, \text{sifPred}, \text{datlsp} \}$$

- **zadovoljen je uvjet 1NF prema kojem neključni atributi funkcijski ovise o ključu**

# 1. primjer normalizacije - 2NF

ispit (ISPIT)

matBr	prez	ime	sifPred	nazPred	datlsp	ocj	sifNas	prezNas
1111	Novak	Ivan	1001	Mat-1	29.01.06	1	1111	Pašić
1111	Novak	Ivan	1001	Mat-1	05.02.06	3	1111	Pašić
1111	Novak	Ivan	1003	Fiz-1	28.06.06	2	3333	Horvat
1111	Novak	Ivan	1002	Mat-2	27.06.06	4	2222	Brnetić
1234	Kolar	Petar	1001	Mat-1	29.01.06	3	2222	Brnetić

- Postoje li neključni atributi koji ne ovise o čitavom ključu nego samo o dijelu ključa?
  - $\text{matBr} \rightarrow \text{prez ime}$

$\text{student} = \pi_{\text{matBr, prez, ime}}(\text{ispit})$

$\text{ispit}_1 = \pi_{\text{matBr, sifPred, nazPred, datlsp, ocj, sifNas, prezNas}}(\text{ispit})$

$K_{\text{STUDENT}} = \{ \text{matBr} \}$

student (STUDENT)

matBr	prez	ime
1111	Novak	Ivan
1234	Kolar	Petar

2NF, 3NF: O.K.

ispit<sub>1</sub> (ISPIT<sub>1</sub>)

$K_{\text{ISPIT}_1} = \{ \text{matBr, sifPred, datlsp} \}$

matBr	sifPred	nazPred	datlsp	ocj	sifNas	prezNas
1111	1001	Mat-1	29.01.06	1	1111	Pašić
1111	1001	Mat-1	05.02.06	3	1111	Pašić
1111	1003	Fiz-1	28.06.06	2	3333	Horvat
1111	1002	Mat-2	27.06.06	4	2222	Brnetić
1234	1001	Mat-1	29.01.06	3	2222	Brnetić

# 1. primjer normalizacije - 2NF (nastavak)

ispit<sub>1</sub> (ISPIT<sub>1</sub>)

- Postoje li neključni atributi koji ne ovise o čitavom ključu nego samo o dijelu ključa?

matBr	sifPred	nazPred	datlsp	ocj	sifNas	prezNas
1111	1001	Mat-1	29.01.06	1	1111	Pašić
1111	1001	Mat-1	05.02.06	3	1111	Pašić
1111	1003	Fiz-1	28.06.06	2	3333	Horvat
1111	1002	Mat-2	27.06.06	4	2222	Brnetić
1234	1001	Mat-1	29.01.06	3	2222	Brnetić

- sifPred → nazPred

predmet =  $\pi_{\text{sifPred, nazPred}}(\text{ispit}_1)$

ispit<sub>2</sub> =  $\pi_{\text{matBr, sifPred, datlsp, ocj, sifNas, prezNas}}(\text{ispit}_1)$

$K_{\text{PREDMET}} = \{ \text{sifPred} \}$

predmet (PREDMET)	
sifPred	nazPred
1001	Mat-1
1003	Fiz-1
1002	Mat-2

2NF, 3NF: O.K.

ispit<sub>2</sub> (ISPIT<sub>2</sub>)

$K_{\text{ISPIT}_2} = \{ \text{matBr, sifPred, datlsp} \}$

matBr	sifPred	datlsp	ocj	sifNas	prezNas
1111	1001	29.01.06	1	1111	Pašić
1111	1001	05.02.06	3	1111	Pašić
1111	1003	28.06.06	2	3333	Horvat
1111	1002	27.06.06	4	2222	Brnetić
1234	1001	29.01.06	3	2222	Brnetić

2NF: O.K.



# 1. primjer normalizacije - 3NF

ispit<sub>2</sub> (ISPIT<sub>2</sub>)

matBr	sifPred	datlsp	ocj	sifNas	prezNas
1111	1001	29.01.06	1	1111	Pašić
1111	1001	05.02.06	3	1111	Pašić
1111	1003	28.06.06	2	3333	Horvat
1111	1002	27.06.06	4	2222	Brnetić
1234	1001	29.01.06	3	2222	Brnetić

- Postoje li neključni atributi koji tranzitivno ovise o ključu?

- matBr sifPred datlsp → sifNas      sifNas → prezNas

nastavnik =  $\pi_{\text{sifNas, prezNas}}(\text{ispit}_2)$

ispit<sub>3</sub> =  $\pi_{\text{matBr, sifPred, datlsp, ocj, sifNas}}(\text{ispit}_2)$

$K_{\text{NASTAVNIK}} = \{ \text{sifNas} \}$

nastavnik (NASTAVNIK)

sifNas	prezNas
1111	Pašić
3333	Horvat
2222	Brnetić

3NF: O.K.

ispit<sub>3</sub> (ISPIT<sub>3</sub>)       $K_{\text{ISPIT}_3} = \{ \text{matBr, sifPred, datlsp} \}$

matBr	sifPred	datlsp	ocj	sifNas
1111	1001	29.01.06	1	1111
1111	1001	05.02.06	3	1111
1111	1003	28.06.06	2	3333
1111	1002	27.06.06	4	2222
1234	1001	29.01.06	3	2222

3NF: O.K.

# 1. primjer normalizacije - 3NF

student (STUDENT)

matBr	prez	ime
1111	Novak	Ivan
1234	Kolar	Petar

$K_{\text{STUDENT}} = \{ \text{matBr} \}$

predmet (PREDMET)

sifPred	nazPred
1001	Mat-1
1003	Fiz-1
1002	Mat-2

$K_{\text{PREDMET}} = \{ \text{sifPred} \}$

nastavnik (NASTAVNIK)

sifNas	prezNas
1111	Pašić
3333	Horvat
2222	Brnetić

$K_{\text{NASTAVNIK}} = \{ \text{sifNas} \}$

ispit<sub>3</sub> (ISPIT<sub>3</sub>)

matBr	sifPred	datlsp	ocj	sifNas
1111	1001	29.01.06	1	1111
1111	1001	05.02.06	3	1111
1111	1003	28.06.06	2	3333
1111	1002	27.06.06	4	2222
1234	1001	29.01.06	3	2222

$K_{\text{ISPIT}_3} = \{ \text{matBr}, \text{sifPred}, \text{datlsp} \}$

- Relacijska shema baze podataka STUSLU:

$\text{STUSLU} = \{ \text{STUDENT}, \text{PREDMET}, \text{NASTAVNIK}, \text{ISPIT}_3 \}$

- Shema baze podataka STUSLU zadovoljava 3NF

## 2. primjer normalizacije

---

Zadana je relacijska shema  $R = ABCDEFGH$  i na njoj skup funkcijskih zavisnosti

$F = \{ ABC \rightarrow DEFGH, A \rightarrow D, BC \rightarrow FGH, FG \rightarrow H \}$ .

Domene atributa sadrže samo jednostavne vrijednosti, vrijednost svakog atributa je samo jedna vrijednost iz domene tog atributa.

Odrediti primarni ključ relacije (tako da bude zadovoljen uvjet 1NF prema kojem neključni atributi funkcijski ovise o ključu), te shemu postupno normalizirati na 2NF i 3NF.

## 2. primjer normalizacije

---

$R = ABCDEFGH$

$F = \{ ABC \rightarrow DEFGH, A \rightarrow D, BC \rightarrow FGH, FG \rightarrow H \}$

- **Odrediti primarni ključ relacije**

Vrijedi li  $ABC \rightarrow DEFGH$  ?

**DA**

postoji li skup  $X \subset ABC$  za kojeg vrijedi  $X \rightarrow R$  ?

**NE**

$\Rightarrow ABC$  je mogući ključ i može se odabrati kao primarni ključ sheme  $R$ .

$R = ABCDEFGH$

$K_R = ABC$

$R$  je u 1NF

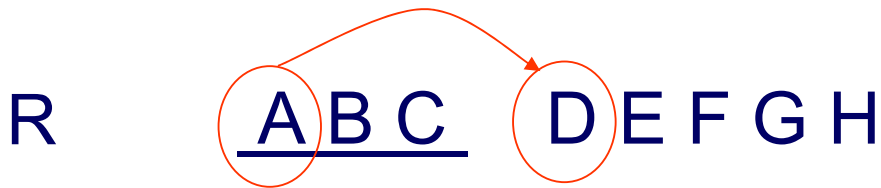
## 2. primjer normalizacije - 2NF

$R = ABCDEFGH$        $K_R = ABC$

$F = \{ ABC \rightarrow DEFGH, A \rightarrow D, BC \rightarrow FGH, FG \rightarrow H \}$

- **Normalizacija na 2NF**

Svi atributi iz zavisnog dijela moraju biti **potpuno** funkcijski ovisni o ključu.



- $ABC \rightarrow D$  je nepotpuna FZ, jer vrijedi  $A \rightarrow D$

R nije u 2NF

Normalizacijom na 2NF se R zamjenjuje shemama:

$R_1 = AD$

$K_{R_1} = A$

$R_1$  je u 2NF

$R_2 = ABCEFGH$

$K_{R_2} = ABC$

$R_2$  nije u 2NF

## 2. primjer normalizacije - 2NF (nastavak)

$$R_1 = AD$$

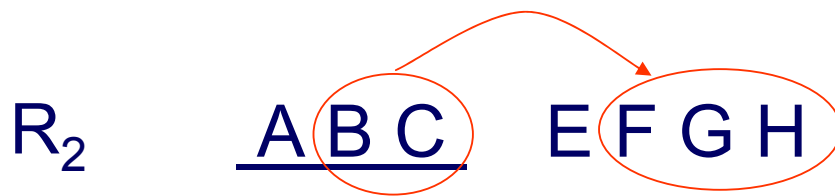
$$K_{R_1} = A$$

$$R_2 = ABCEFGH$$

$$K_{R_2} = ABC$$

$$F = \{ ABC \rightarrow DEFGH, A \rightarrow D, BC \rightarrow FGH, FG \rightarrow H \}$$

Svi atributi iz zavisnog dijela moraju biti **potpuno** funkcijski ovisni o ključu.



- $ABC \rightarrow FGH$  je nepotpuna FZ, jer vrijedi  $BC \rightarrow FGH$       $R_2$  nije u 2NF

Normalizacijom na 2NF se  $R_2$  zamjenjuje shemama:

$$R_{21} = BCFGH$$

$$K_{R_{21}} = BC$$

$$R_{21} \text{ je u 2NF}$$

$$R_{22} = ABCE$$

$$K_{R_{22}} = ABC$$

$$R_{22} \text{ je u 2NF}$$

## 2. primjer normalizacije - 3NF

$R_1 = AD$

$K_{R_1} = A$

$R_1$  je u 3NF

$R_{21} = BCFGH$

$K_{R_{21}} = BC$

$R_{21}$  nije u 3NF

$R_{22} = ABCE$

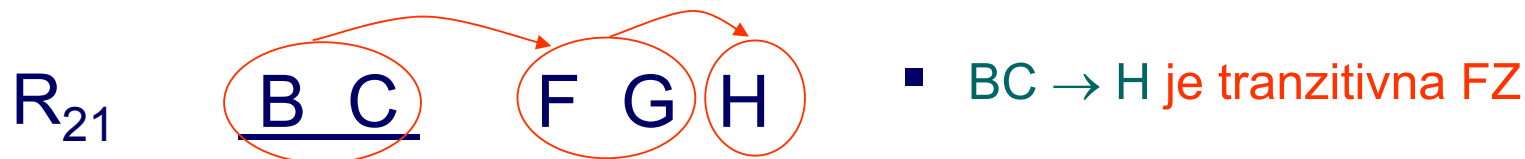
$K_{R_{22}} = ABC$

$R_{22}$  je u 3NF

$F = \{ ABC \rightarrow DEFGH, A \rightarrow D, BC \rightarrow FGH, FG \rightarrow H \}$

- **Normalizacija na 3NF**

Niti jedan atribut iz zavisnog dijela ne smije biti tranzitivno ovisan o ključu.



Normalizacijom na 3NF se  $R_{21}$  zamjenjuje shemama:

$R_{211} = BCFG$

$K_{R_{211}} = BC$

$R_{211}$  je u 3NF

$R_{212} = FGH$

$K_{R_{212}} = FG$

$R_{212}$  je u 3NF

**Shema baze podataka u 3NF sastoji se od relacijskih shema:**

$R_1, R_{22}, R_{211}$  i  $R_{212}$

# Boyce-Coddova normalna forma - BCNF

---

- Definicija:

Relacijska shema R je u **BCNF** ako je u 1NF i ako niti jedan atribut nije tranzitivno funkcijski ovisan o bilo kojem ključu relacije

- Iako je BCNF stroža od 3NF, rijetki su slučajevi da je relacijska shema u 3NF, a da istovremeno nije i u BCNF.
- Normalizaciju na BCNF nije nužno provoditi.
- **Smatra se da shema baze podataka ima dobra svojstva ako zadovoljava 3NF.**