
POJAM NORMALIZACIJE, F-NE ZAVISNOSTI I NF



NORMALIZACIJA

- Normalizacija je postupak **logičkog projektovanja baze podataka** kojim se **odstranjuju anomalije u njenom održavanju**.
- Redudansa (ponavljanje, dupliranje) podataka je jedan od glavnih uzroka pojave anomalija.
- Određena struktura baze podataka se smatra dobrom ako joj je logička redudansa minimalna.

NENORMALIZOVANA RELACIJA I ANOMALIJE ODRŽAVANJA (2)

Nenormalizovana relacija je relacija koja poseduje neke vrednosti atributa koje nisu "atomske", odnosno relacija koja poseduje "grupe sa ponavljanjem".

STUDENT

BI	IME	SEM	ŠSMER	IMERUK	POLOZIO (ŠPRED-NAZPRED-OCENA)
21	ZORAN	5	01	BATA	{121, MATEMAT, 7} {323, BAZEPOD, 8} {056, SOCIOL, 8}
77	ANA	7	01	BATA	{056, MARKSIZ, 10} {121, MATEMAT, 5}
36	PERA	4	02	MIKA	{323, BAZEPOD, 8} {456, ELEKTRON, 9} {442, FIZIKA, 6} {056, SOCIOL, 8}

Jedna n-torka ove nenormalizovane relacije bi se mogla zapisati ovako:

(21, ZORAN, 5, 01, BATA, { (121, MATEMAT, 7), (323, BAZEPOD, 8), (056, SOCIOL, 8) })

NENORMALIZOVANA RELACIJA I ANOMALIJE ODRŽAVANJA (3)

- **Anomalije u dodavanju:**

Ako je u novom nastavnom planu definisan novi predmet, ne mogu se ubaciti podaci o tom predmetu dok ga neki student ne položi.

Ili, ako se otvori neki novi smer, ne mogu se ubaciti podaci o tom smeru dok ga neki student ne upiše.

- **Anomalije u brisanju:**

Ako je jedan predmet (FIZIKA) položio samo jedan student (PERA) i ako se on ispiše sa fakulteta, odnosno izbaci odgovarajuca n-torka, gube se i sve informacije o tom predmetu.

Ako je taj student bio i jedini student na nekom smeru, gube se i sve informacije o tom smeru.

- **Anomalije u ažuriranju:**

Ako se promeni naziv nekog predmeta ili rukovodioc nekog smera, to se mora uciniti na onoliko mesta koliko je studenata položilo taj predmet, odnosno koliko je studenata upisano na dati smer.

- **Izveštavanje za datu strukturu:**

- Data struktura relacije je veoma pogodna za izveštaj na zahtev za izdavanje uverenja o položenim ispitima.
- "Prikaži listu predmeta, imena svih studenata koji su ga položili i prosečnu ocenu na predmetu", za datu strukturu relacije zahtevao bi znatno složeniji program.

ČEMU SLUŽI NORMALIZACIJA

- Postupkom normalizacije logicka struktura baze podataka se dovodi u takav oblik (ili, drugim recima, relacije se dovode u normalne forme) u kome se **izbegavaju anomalije u održavanju i problemi u izveštavanju.**

PRVA NORMALNA FORMA (1NF)

- Relacija R je u Prvoj normalnoj formi (1NF) ako su sve vrednosti njenih atributa atomske.
- Termin normalizovana relacija se koristi za relacije koje su u 1NF.
- Svi upitni jezici zasnovani na relacionoj algebri zahtevaju normalizovane relacije.

PRVA NORMALNA FORMA (1NF)

- **Relacija R je u Prvoj normalnoj formi (1NF) ako su sve vrednosti njenih atributa atomske.**
- Termin **normalizovana** relacija se koristi za relacije koje su u 1NF.
- Svi upitni jezici zasnovani na relacionoj algebri zahtevaju (horizontalno) normalizovane relacije.

STUDENT

BI	IME	SEM	ŠSMER	IMERUK	ŠPRED	NAZPRED	OCENA
21	ZORAN	5	01	BATA	121	MATEMAT	7
21	ZORAN	5	01	BATA	323	BAZEPOD	8
21	ZORAN	5	01	BATA	056	SOCIOL	8
77	ANA	7	01	BATA	056	MARKSIZ	10
77	ANA	7	01	BATA	121	MATEMAT	5
36	PERA	4	02	MIKA	323	BAZEPOD	8
36	PERA	4	02	MIKA	456	ELEKTRON	9
36	PERA	4	02	MIKA	442	FIZIKA	6
36	PERA	4	02	MIKA	056	SOCIOL	8

FUNKCIONALNE (FUNKCIJSKE) ZAVISNOSTI

Data je relacija R sa atributima X i Y , moguće složenim. Atribut Y je funkcionalno zavisan od atributa X (ili X funkcionalno određuje Y),

$$R.X \rightarrow R.Y,$$

ako i samo ako svakoj vrednosti X odgovara jedna i samo jedna vrednost Y .

Primer.

STUDENTI (BI, IME, SEM, ŠSMER, IMERUK) (I)

BI \rightarrow IME

BI \rightarrow SEM

BI \rightarrow ŠSMER

BI \rightarrow IMERUK

Atribut Y relacije R je funkcionalno zavisan od atributa X relacije R ako i samo ako kad god dve n -torke relacije R imaju istu x -vrednost one moraju imati istu y -vrednost.

ZADOVOLJENOST F-NE ZAVISNOSTI

Ako u relaciji R važi funkcionalna zavisnost $X \rightarrow Y$, onda se kaže da relacija R *zadovoljava* tu funkcionalnu zavisnost.

Relacija R je dinamičkog sadržaja, što znači da njen sadržaj u raznim momentima može biti različit. *Stanje* r relacije R je njen sadržaj u nekom momentu. Kazaćemo da stanje r relacije R *zadovoljava* FZ $X \rightarrow Y$ ako ta funkcionalna zavisnost važi u relaciji R u momentu koji definiše njeno stanje r . Sada se za relaciju R može reći da *zadovoljava* FZ $X \rightarrow Y$ ako svako njeno stanje r *zadovoljava* tu FZ.

PITANJE

Da li dato stanje relacije zadovoljava sledeću f-nu zavisnost:

a) $\{C\} \rightarrow \{A\}$

b) $\{A, D\} \rightarrow \{C\}$

c) $\{B\} \rightarrow \{E\}$

A	B	C	D	E
1	1	1	1	1
1	2	2	2	1
2	1	3	3	1
2	1	4	3	1
3	2	5	1	1

PITANJE

Da li dato stanje relacije zadovoljava sledeću f-nu zavisnost:

a) $\{C\} \rightarrow \{A\}$

Važi. Postojanje $3 \rightarrow 2$, $4 \rightarrow 2$ nije u suprotnosti sa zahtevom definicije

b) $\{A, D\} \rightarrow \{C\}$

c) $\{B\} \rightarrow \{E\}$

A	B	C	D	E
1	1	1	1	1
1	2	2	2	1
2	1	3	3	1
2	1	4	3	1
3	2	5	1	1

PITANJE

Da li dato stanje relacije zadovoljava sledeću f-nu zavisnost:

a) $\{C\} \rightarrow \{A\}$

Važi. Postojanje $3 \rightarrow 2$, $4 \rightarrow 2$ nije u suprotnosti sa zahtevom definicije

b) $\{A, D\} \rightarrow \{C\}$

Ne važi jer postoji $(2,3) \rightarrow 3$, $(2,3) \rightarrow 4$

c) $\{B\} \rightarrow \{E\}$

A	B	C	D	E
1	1	1	1	1
1	2	2	2	1
2	1	3	3	1
2	1	4	3	1
3	2	5	1	1

PITANJE

Da li dato stanje relacije zadovoljava sledeću f-nu zavisnost:

a) $\{C\} \rightarrow \{A\}$

Važi. Postojanje $3 \rightarrow 2$, $4 \rightarrow 2$ nije u suprotnosti sa zahtevom definicije

b) $\{A, D\} \rightarrow \{C\}$

Ne važi jer postoji $(2,3) \rightarrow 3$, $(2,3) \rightarrow 4$

c) $\{B\} \rightarrow \{E\}$

Važi.

A	B	C	D	E
1	1	1	1	1
1	2	2	2	1
2	1	3	3	1
2	1	4	3	1
3	2	5	1	1

DEFINICIJA KLJUČA I NADKLJUČA

- Atribut X moguće složeni, je **nadključ** neke relacije R ako i samo **ako funkcionalno određuje sve ostale attribute relacije R .**
- Atribut X , moguće složeni, je **ključ relacije R** ako je **nadključ** relacije R , a **nijedan njegov pravi podskup nema tu osobinu.**

FUNKCIONALNE (FUNKCIJSKE) ZAVISNOSTI

- $X \rightarrow Y$ je **trivijalna** akko za važi da je $Y \subseteq X$
- $X \rightarrow Y$ je **potpuna** akko ne postoji ni jedan pravi podskup Z od X , za koji važi $Z \rightarrow Y$
- $X \rightarrow Y$ je **parcijalna** akko za $X \rightarrow Y$ postoji pravi podskup Z od X , za koji važi $Z \rightarrow Y$
- $X \rightarrow Y$ je **tranzitivna** akko za $X \rightarrow Y$ postoji Z različito od X i Y za koje važi $X \rightarrow Z$ i $Z \rightarrow Y$

X, Y, Z - podskupovi skupa atributa relacije R

POTPUNE FUNKCIONALNE ZAVISNOSTI

Atribut Y relacije R je **potpuno funkcionalno zavisan** od kolekcije atributa X relacije R ako je funkcionalno zavisan od kolekcije atributa X , a nije funkcionalno zavisan ni od jednog pravog podskupa kolekcije atributa X .

Primere potpune i nepotpune funkcionalne zavisnosti možemo pokazati na relaciji PRIJAVA:

(a) $\{BI, \check{S}PRED\} \rightarrow \{OCENA\}$

$\{BI\} \not\rightarrow \{OCENA\}$

$\{\check{S}PRED\} \not\rightarrow \{OCENA\}$

atribut OCENA je potpuno funkcionalno zavisnan od $\{BI, \check{S}PRED\}$.

(b) $\{BI, \check{S}PRED\} \rightarrow \{NAZPRED\}$

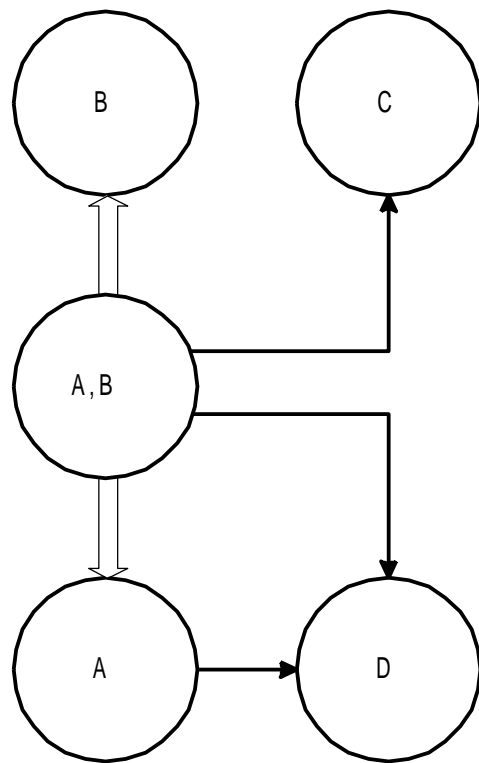
$\{BI\} \not\rightarrow \{NAZPRED\}$

$\{\check{S}PRED\} \dashrightarrow \{NAZPRED\}$

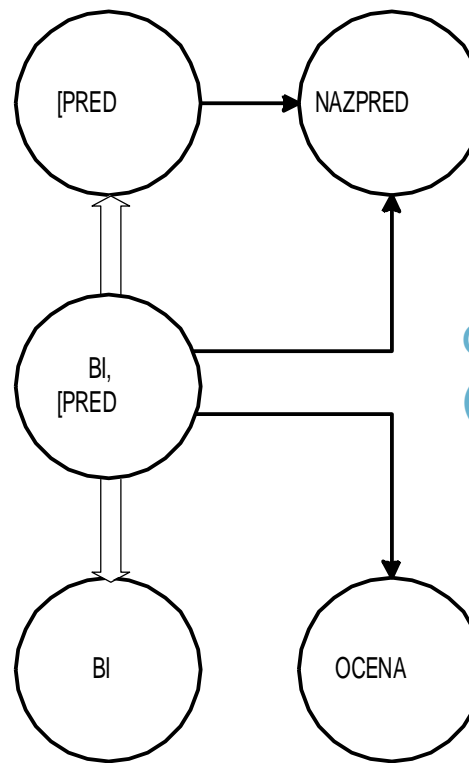
atribut NAZPRED je nepotpuno funkcionalno zavisan od složenog atributa BI, ŠPRED, jer je funkcionalno zavisan i od njega i od jednog njegovog dela, od ŠPRED.

POTPUNE FUNKCIONALNE ZAVISNOSTI

C je potpuno funkcionalno zavisno od složenog atributa (A,B), a D nije.



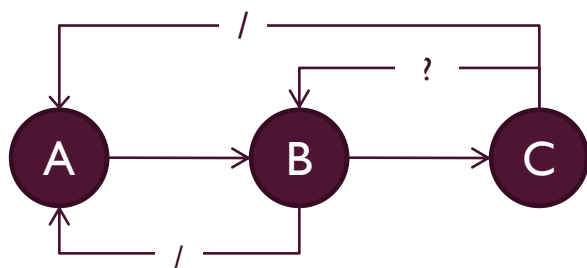
sa \implies su označene trivijalne funkcije



OCENA je potpuno zavisna od (BI, ŠPRED), a NAZPRED nije.

TRANZITIVNE FUNKCIONALNE ZAVISNOSTI

Atribut C je **tranzitivno funkcionalno zavisan od atributa A** ako je funkcionalno zavisan od A i ako je funkcionalno zavisan od nekog atributa B koji je i sam funkcionalno zavisan od A.

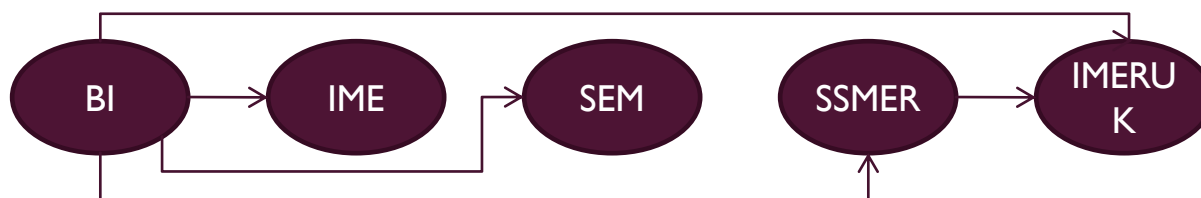


Tranzitivna funkcionalna zavisnost je redundantna, pa je nije neophodno pamtiti u bazi podataka.

$f_1: BI \rightarrow \check{S}SMER$

$f_2: \check{S}SMER \rightarrow IMERUK$

$f_3 = f_1 \circ f_2 : BI \rightarrow IMERUK$



NORMALNE FORME

- Relacija R je u **Drugoj normalnoj formi (2NF)** ako i samo ako je u **1NF** i svi njeni neključni atributi **potpuno funkcionalno** zavise od primarnog ključa.
- Relacija R je u **Trecoj normalnoj formi (3NF)** ako i samo ako je u **2NF** i ako svi njeni neključni atributi **netranzitivno funkcionalno** zavise od ključa.
- **Boyce-Codd-ova normalna forma (BCNF)**
- 4NF
- 5NF



VERTIKALNA NORMALIZACIJA



VERTIKALNA NORMALIZACIJA

- Cilj: formiranje relacionih šema koje ispunjavaju uslove zahtevane normalne forme.

$R(A, B, C, D)$

$A, B \rightarrow A, B, C, D$

$A \rightarrow C$



$R_1(A, B, D)$

$A, B \rightarrow A, B, D$

$R_2(A, C)$

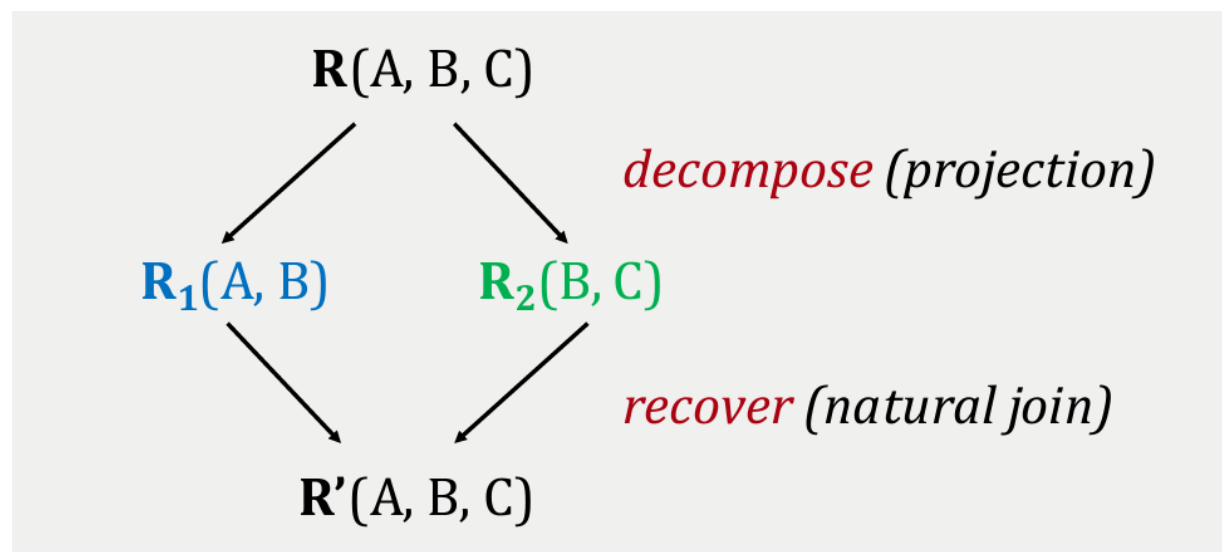
$A \rightarrow C$

POSTUPCI VERTIKALNE NORMALIZACIJE

- Postupci vertikalne normalizacije:
 1. **normalizacija dekompozicijom,**
 2. **normalizacija sintezom.**
- **Normalizacija dekompozicijom** započinje od proizvoljne nenormalizovane relacije šeme i izvodi se u koracima.
 - Svakim korakom normalizacije relaciona šema se prevodi u višu normalnu formu, tako da se polazni skup atributa deli u dva skupa i od svakog formira posebna relaciona šema.
 - Svaki korak normalizacije mora biti reverzibilan (dekompozicija bez gubljenja informacija).
- **Normalizacija sintezom** polazi od skupa atributa i od skupa zavisnosti zadatih na tom skupu atributa. Postupak se ne izvodi u koracima već se direktno formiraju relacione šeme koje ispunjavaju uslove zahtevane normalne forme.

DEKOMPOZICIJA RELACIJA BEZ GUBLJENJA INFORMACIJA

- Relacija R se dekomponuje u svoje projekcije bez gubljenja informacija ako prirodno spajanje tako dobijenih projekcija dovodi do polazne relacije.



Dekompozicijom se ne gube informacije ako za svaku instancu relacije R važi da je $R=R'$.

PRIMER DEKMPOZICIJE KOJA NE OBEZEĐUJE ČUVANJE INFORMACIJA

PRIJAVA	BI	ŠPRED	NAZPRED	OCENA
	21	121	MATEMAT	7
	21	323	BAZEPOD	8
	21	056	SOCIOL	8
	77	056	SOCIOL	10
	77	121	MATEMAT	5
	36	323	BAZEPOD	8
	36	456	ELEKTRON	9
	36	442	FIZIKA	6
	36	056	SOCIOL	8

$$R3 = R1 [OCENA * OCENA] R2$$

R3	BI	ŠPRED	NAZPRED	OCENA
	21	121	MATEMAT	7
	21	323	BAZEPOD	8
	21	056	SOCIOL	6
	77	056	SOCIOL	10
	77	121	MATEMAT	5
	36	323	BAZEPOD	9
	36	456	ELEKTRON	9
	36	442	FIZIKA	6
	36	056	SOCIOL	8

RI	BI	ŠPRED	OCENA	R2	NAZPRED	OCENA
	21	121	7		MATEMAT	7
	21	323	8		BAZEPOD	8
	21	056	8		SOCIOL	8
	77	056	10		SOCIOL	10
	77	121	5		MATEMAT	5
	36	323	8		ELEKTRON	9
	36	456	9		FIZIKA	6
	36	442	6			
	36	056	8			

	21	323	SOCIOL	8
	36	056	BAZEPOD	8
	21	056	FIZIKA	6
	36	442	SOCIOL	6
	36	323	ELEKTRON	9
	36	465	BAZEPOD	9

DEKOMPOZICIJA RELACIJA BEZ GUBLJENJA INFORMACIJA

- Heath-ova teorema

Relacija $R(A,B,C)$, gde su A , B i C podskupovi atributa, u kojoj važi $R.A \rightarrow R.B$ može se uvek dekomponovati u svoje projekcije $R_1(A,B)$ i $R_2(A,C)$ bez gubljenja informacija.

- Za navedeni primer relacije PRIJAVA dekompozicija bez gubljenja informacija bi bila dekompozicija u projekcije $R_1(BI, \check{S}PRED, OCENA)$ i $R_2(\check{S}PRED, NAZPRED)$

NORMALIZACIJA DEKOMPOZICIJOM

- Kada je dekompozicija dobra?
 - **Svi atributi** iz početne relacije se moraju pojaviti **i u novim relacijama**.
 - Dekompozicija je dobra samo u slučaju da se **ne gube f-ne zavisnosti**.
 - Sve FZ polaznog skupa moraju da budu očuvane (direktno ili mogućim izvodjenjem iz skupa relacija dobijenih dekompozicijom)
 - Ako u novodobijenim projekcijama nastalim razbijanjem osnovne relacije postoji **zajednički atribut**, on **mora da bude ključ u bar jednoj** od novodobijenih relacija.

NORMALIZACIJA DEKOMPOZICIJOM

- Ukloniti sve redundantne f-ne zavisnosti, tj. one koje se mogu izvesti iz drugih.
 - Odrediti nereduantno pokrivanje skupa FZ.
- Određivanje kandidata za ključ
 - Sporedni (neključni) atributi su svi oni koji **ne pripadaju bilo kom kandidatu za ključ.**
- Dekompozicija prema f-im zavisnostima od interesa.

Za ovo nam je potrebna teorija zavisnosti!



TEORIJA ZAVISNOSTI



ARMSTRONGOVE AKSIOME IZVOĐENJA

- Aksiome izvođenja su pravila koja utvrđuju važenje nekih funkcionalnih zavisnosti u relaciji R, na osnovu važenja nekih drugih funkcionalnih zavisnosti u R.
- Mogu se izvesti na osnovu osobina f-nih zavisnosti, pa u svojoj suštini nisu prave aksiome.
- Armstrongove aksiome (u modifikovanom obliku) su:
 - Aksioma **refleksivnosti**:
Funkcionalna zavisnost $X \rightarrow X$ uvek važi.
 - Aksioma **proširenja**:
Ako u relaciji R važi $X \rightarrow Y$ i ako je Z podskup skupa $A_{tr}(R)$, onda važi i $X, Z \rightarrow Y$.
 - Aksioma **aditivnosti**:
Ako u relaciji R važi $X \rightarrow Y$ i $X \rightarrow Z$, onda važi i $X \rightarrow Y, Z$.
 - Aksioma **projektivnosti**:
Ako u relaciji R važi $X \rightarrow Y, Z$, onda važi i $X \rightarrow Y$ i $X \rightarrow Z$.
 - Aksioma **tranzitivnosti**:
Ako u relaciji R važi $X \rightarrow Y$ i $Y, Z \rightarrow W$, onda važi i $X \rightarrow Z$.
 - Aksioma **pseudotranzitivnosti**:
Ako u relaciji R važi $X \rightarrow Y$ i $Y, Z \rightarrow W$, onda važi i $X, Z \rightarrow W$.

IZVOĐENJE F-NIH ZAVISOSTI

- Neka je zadat skup F funkcionalnih zavisnosti jedne relacije:
 - Kako konstruisati skup svih funkcionalnih zavisnosti koje se mogu izvesti iz F po Armstrongovim aksiomama?
 - Da li je zadata FZ logička posledica skupa F ? (Da li je izvodiva iz F)

ZATVORENJE SKUPA F-IH ZAVISNOSTI

- Neka je F neki skup FZ relacije R . **Zatvorenje od F** , u oznaci F^+ ; jeste najmanji skup FZ koji sadži F i takav da se primenom Armstrongovih aksioma na F^+ ne dobija nijedna FZ koja već nije u F^+ .

- Primer

$R(A,B,C)$

$F = \{A \rightarrow B, A \rightarrow C\}$

$F^+ = \{A \rightarrow A, A \rightarrow B, A \rightarrow C,$
 $AB \rightarrow A, AB \rightarrow B, AB \rightarrow C,$
 $AB \rightarrow AB, AB \rightarrow BC, AB \rightarrow AC,$
 $AB \rightarrow ABC,$
 $AC \rightarrow \dots,$
 $ABC \rightarrow ABC\}$

Da bismo odredili KK...
Najčeće nepotrebno.

HEATH-OVA TEOREMA PREPRIČANA

- Neka je data relacija $R(A)$ sa skup f-nih zavisnosti F i njena dekompozicija na $R_1(A_1)$ i $R_2(A_2)$
- Data dekompozicija je dekompozicija bez gubljenja informacija akko se u zatvorenju F^+ nalazi bar jedna od narednih f-nih zavisnosti:
 - $A_1 \cap A_2 \rightarrow A_1$
 - $A_1 \cap A_2 \rightarrow A_2$

ZATVORENJE X NAD F

- Neka je R skup atributa, X neki njegov podskup, a F skup f-ih zavisnosti nad R . **Zatvorenje skupa atributa X nad skupom f-ih zavisnosti F** , u oznaci X^+ ili $(X,F)^+$, čini skup atributa Y koji odgovara desnoj strani zavisnosti $X \rightarrow Y$ u zatvaraču skupa funkcionalnih zavisnosti sa maksimalnim Y .

$$(X, F)^+ := \{A_i \mid X \rightarrow A_i \text{ se može izvesti iz } F\}$$

Neformalno: Maksimalan skup atributa f-no zavisnih od X .

- Kako doći do zatvorenja skupa atributa X nad skupom f-nih zavisnosti F ?

```
unused := F
closure := X
do {
  for( $Y \rightarrow Z \in \text{unused}$ ) {
    if( $Y \subseteq \text{closure}$ ) {
      unused := unused \ { $Y \rightarrow Z$ }
      closure := closure U Z
    }
  }
} while (unused and closure did not change)
return closure
```

ZATVORENJE X NAD F

$$F = \left\{ \begin{array}{l} \{A\} \rightarrow \{B, C\}, \\ \{B\} \rightarrow \{E\}, \end{array} \quad \begin{array}{l} \{E\} \rightarrow \{C, G\}, \\ \{C, D\} \rightarrow \{E, G\} \end{array} \right\}$$

Odrediti $Y = (\{A,B\},F)^+ / (A,B)^+$

Neka je $X = \{A,B\}, Y = (\{A,B\},F)^+$

Algoritam:

$$Y = \{A,B\}$$

dodajemo C, zbog $\{A\} \rightarrow \{B, C\}$

$$Y = \{A,B,C\}$$

dodajemo E, zbog $\{B\} \rightarrow \{E\}$

$$Y = \{A,B,C,E\}$$

dodajemo G, zbog $\{E\} \rightarrow \{C,G\}$

$$Y = \{A,B,C,E,G\}$$

$$(\{A,B\},F)^+ = \{A,B,C,E,G\}$$

ČEMU SLUŽI $(X, F)^+$

- Za dat skup f-ih zavisnosti F možemo odrediti da li $X \rightarrow Y$ pripada F^+
Trebalo bi proveriti da li $Y \subseteq (X, F)^+$
- Da li je dati skup atributa X nadključ relacije R
Trebalo bi proveriti da li $(X, F)^+ =$ skup svih atributa relacije R

POKRIVANJE SKUPA FZ

Definicija 8.3 Dva skupa FZ F i F' relacije R su *ekvivalentna* (u oznaci $F \equiv F'$) ako je $F^+ = F'^+$. Ako je $F \equiv F'$, onda je F *pokrivanje* za F' (F' je *pokrivanje* za F).

Definicija 8.4 Skup FZ F je *neredundantan* ako ne sadrži pravi podskup $F' \subset F$ takav da je $F' \equiv F$; u suprotnom je skup F *redundantan*. Skup F je *neredundantno pokrivanje* skupa FZ G ako je F pokrivanje skupa G i F je neredundantan skup FZ.

KANONIČKO POKRIVANJE SKUPA FZ

Definicija 8.6 *Kanoničko pokrivanje* skupa FZ G je neredundantno pokrivanje F koje zadovoljava sledeća dva uslova:

- a) svaka FZ u skupu F je oblika $X \rightarrow A$, gde je A jedan atribut;
- b) leva strana svake FZ iz F oslobođena je *nebitnih* atributa, tj. atributa čije udaljenje iz leve strane te FZ ne menja zatvorenje skupa FZ G .

KANONIČKO POKRIVANJE SKUPA FZ

$A \rightarrow BC$

$B \rightarrow C$

$A \rightarrow B$

$AB \rightarrow C$

$AC \rightarrow D$



$A \rightarrow B$

$A \rightarrow C$

$B \rightarrow C$

~~$A \rightarrow B$~~

$AB \rightarrow C$

$AC \rightarrow D$



$A \rightarrow B$

$A \rightarrow C$

$B \rightarrow C$

~~$A \rightarrow B$~~

$AB \rightarrow C$

~~$AC \rightarrow D$~~

$A \rightarrow D$



$A \rightarrow B$

$A \rightarrow C$

$B \rightarrow C$

~~$A \rightarrow B$~~

~~$AB \rightarrow C$~~

~~$AC \rightarrow D$~~

$A \rightarrow D$



$A \rightarrow B$

~~$A \rightarrow C$~~

$B \rightarrow C$

~~$A \rightarrow B$~~

~~$AB \rightarrow C$~~

~~$AC \rightarrow D$~~

$A \rightarrow D$



NORMALIZACIJA DEKOMPOZICIJOM



NORMALIZACIJA DEKOMPOZICIJOM

- Ukloniti sve redundantne f-ne zavisnosti, tj. one koje se mogu izvesti iz drugih.
 - Odrediti neredundantno pokrivanje skupa FZ.
- Određivanje kandidata za ključ
 - Sporedni (neključni) atributi su svi oni koji **ne pripadaju bilo kom kandidatu za ključ**.
- Dekompozicija prema f-im zavisnostima od interesa.
- Kada je dekompozicija dobra?
 - Svi atributi iz početne relacije se moraju pojaviti i u novim relacijama.
 - Dekompozicija je dobra samo u slučaju da se ne gube f-ne zavisnosti.
 - Sve FZ polaznog skupa moraju da budu očuvane (direktno ili mogućim izvodjenjem iz skupa relacija dobijenih dekompozicijom)
 - Ako u novodobijenim projekcijama nastalim razbijanjem osnovne relacije postoji zajednički atribut, on mora da bude ključ u bar jednoj od novodobijenih relacija.

ODREĐIVANJE KK

- Ukloniti trivijalne funkcijske zavisnosti iz F
- Rasporediti attribute u četiri grupe:
 - i. Elementi koji se ne nalaze ni u jednoj funkcijskoj zavisnosti u F
 - ii. Elementi koji se nalaze isključivo sa leve strane funkcijskih zavisnosti u F
 - iii. Elementi koji se nalaze isključivo sa desne strane funkcijskih zavisnosti u F
 - iv. Elementi koji se nalaze sa obe strane funkcijskih zavisnosti u F
- Izvršiti uniju grupa i i ii i odrediti zatvarač skupa atributa
- A) Ako se u koraku III dobio KK to je jedini KK.
B) Ako se u koraku III nije dobio KK, onda se kombinuje unija grupa i i ii sa atributima grupe iv. Kombinacije se vrše iterativno po broju atributa.

ODREĐIVANJE KK

$R(A, B, C, D, E, F) \quad F = \{AB \rightarrow C, C \rightarrow B, CD \rightarrow AD, BD \rightarrow AE\}$.

- Ukloniti trivijalne funkcionalne zavisnosti iz F

$F = \{AB \rightarrow C, C \rightarrow B, CD \rightarrow A, BD \rightarrow AE\}$

- Rasporediti attribute u četiri grupe:

i. Elementi koji se ne nalaze ni u jednoj f-oj zavisnosti u F

F

ii. Elementi koji se nalaze isključivo sa leve strane f-ne zavisnosti u F

D

iii. Elementi koji se nalaze isključivo sa desne strane f-ne zavisnosti u F

E

iv. Elementi koji se nalaze sa obe strane f-ne zavisnosti u F

A, B, C

- Izvršiti uniju grupa i i ii i odrediti zatvarač skupa atributa

$(F, D)^+ \rightarrow F, D$

- Kombinovanje unija grupa i i ii sa atributima grupe iv.

$(F, D, A)^+ \rightarrow F, D, A$ (nije KK)

$(F, D, B)^+ \rightarrow F, D, B, A, E, C$ (KK)

$(F, D, C)^+ \rightarrow F, D, C, B, A, E$ (KK)

DRUGA NORMALNA FORMA (2NF)

Relacija R je u Drugoj normalnoj formi (2NF) ako i samo ako je u 1NF i svi njeni neključni atributi potpuno i funkcionalno zavise od primarnog ključa.

- Svodenje na 2NF vrši se dekompozicijom na taj način što u jednoj projekciji ostavlja primarni ključ i svi atributi koji su potpuno funkcionalno zavisni od njega, a u drugim projekcijama se realizuju one funkcionalne zavisnosti koje su prouzrokovale nepotpune funkcionalne zavisnosti.

SVOĐENJE NA 2NF POSTUPKOM DEKOMPOZICIJE

$R(\underline{A}, B, C, D, E)$

Neka već znamo KK.

$A \rightarrow C$

$B \rightarrow D$

Kako teče dekompozicija:

1. $A, B \rightarrow C$ je parcijalna zavisnost i zato
 $R(A, B, D)$ KK(A,B) i $B \rightarrow D$
 $R_1(A, C)$ $A \rightarrow C$
2. $A, B \rightarrow D$ je parcijalna zavisnost i zato
 $R(A, B)$
 $R_2(B, D)$
3. $R_3 = R$

Konačno $R_1(A, C)$ $R_2(B, D)$ $R_3(A, B)$

Algoritam 2NF normalizacije

Ulaz: relacija R sa atributima $\{A_1, \dots, A_n\}$ i skupom FZ F ;

Izlaz: skup R_1, \dots, R_i ($i \geq 1$) projekcija relacije R , koje su sve u 2NF i za koje važi: ako je $i = 1$, onda je $R_1 \equiv R$; ako je $i > 1$, onda je $R = R_1 * R_2 * \dots * R_i$;

```
BEGIN
  i := 1;
  WHILE relacija R nije u 2NF DO
    BEGIN
      uočiti parcijalnu FZ  $X \rightarrow A$ , ( $X \subseteq Atr(R)$ ,  $A \in Atr(R)$ )
      sporednog atributa  $A$  od ključa  $X$ ;
      neka je  $X = X'X''$ , gde je  $X' \rightarrow A$  potpuna FZ;
      neka je  $Z = Atr(R) \setminus (X \cup \{A\})$ ;
      zameniti relaciju  $R$  njenim projekcijama  $R[XZ]$ ,  $R[X'A]$ ;
       $R_i := R[X'A]$ ;
      i := i + 1;
      R := R[XZ]
    END;
  R_i := R
END.
```

TREĆA NORMALNA FORMA (3NF)

Relacija R je u Trećoj normalnoj formi (3NF) ako i samo ako je u 2NF i ako svi njeni neključni atributi **netranzitivno funkcionalno** zavise od primarnog ključa.

Svodenje na 2NF vrši se dekompozicijom na taj način što u jednoj projekciji ostavlja primarni ključ i svi atributi netranzitivno zavisni atributi, a u drugim projekcijama se realizuju funkcionalne zavisnosti koje su dovele do tranzitivnih zavisnosti.

SVOĐENJE NA 3NF POSTUPKOM DEKOMPOZICIJE

$R(\underline{A}, B, C, D)$

$B \rightarrow C$

$C \rightarrow D$

Neka već znamo KK. Kako teče dekompozicija:

1. $A \rightarrow C, D$ je tranzitivna zavisnost, zbog

$B \rightarrow C, D$ i zato

$R(A, B)$

$R_1(B, C, D) \quad B \rightarrow C, \quad C \rightarrow D$

2. $A, B \rightarrow D$ je parcijalna zavisnost i zato

$R(A, B)$

$R_2(B, D)$

3. $R_3 = R$

Konačno $R_1(A, C) \quad R_2(B, D) \quad R_3(A, B)$

Algoritam 3NF normalizacije

Ulaz: relacija R sa atributima $\{A_1, \dots, A_n\}$ i skupom FZ F ;

Izlaz: skup R_1, \dots, R_m ($m \geq 1$) projekcija relacije R , koje su sve u 3NF i za koje važi: ako je $m = 1$, onda je $R_1 \equiv R$; ako je $m > 1$, onda je $R = R_1 * R_2 * \dots * R_m$;

```
BEGIN
```

```
   $i := 1$ ;
```

```
  WHILE relacija  $R$  nije u 3NF, DO
```

```
    BEGIN
```

```
      uočiti tranzitivnu zavisnost  $X \rightarrow Y, Y \rightarrow A, Y \not\rightarrow X$  sporednog atributa
```

```
       $A$  od ključa  $X$ , za koju je  $Y \rightarrow A$  potpuna funkcionalna zavisnost;
```

```
      neka je  $Z = Atr(R) \setminus (X \cup \{A\})$ ;
```

```
      zameniti relaciju  $R$  projekcijama  $R[YA], R[XZ]$ ;
```

```
       $R_i := R[YA]$ ;
```

```
       $i := i + 1$ ;
```

```
       $R := R[XZ]$ 
```

```
    END;
```

```
     $R_i := R$ 
```

```
  END.
```



BCNF, IVNF IVNF



BOYCE-CODD-OVA NORMALNA FORMA (BCNF)

- Kodova originalna definicija 3NF nije uzimala u obzir slučajeve kada
 - relacija ima više od jednog kandidata za ključ
 - složeni kandidati za ključeve se preklapaju
- Ovi slučajevi su obuhvaćeni Bojs-Kodovom normalnom formom

BOYCE-CODD-OVA NORMALNA FORMA (BCNF)

- Relacija sa tzv "preklapajucim" kandidatima za ključ (dva ili više složenih kandidata za ključ koji imaju barem jedan zajednicki atribut).

PRIJAVA(BI, ŠPRED, NAZPRED, OCENA)

BI, ŠPRED ---> NAZPRED

BI, ŠPRED ---> OCENA

ŠPRED ---> NAZPRED

NAZPRED ---> ŠPRED

KK(BI, ŠPRED) i KK(BI, NAZPRED)

- **Determinanta** relacije R je bilo koji atribut, prost ili složen, od koga neki drugi atribut u relaciji potpuno funkcionalno zavisi.
- Relacija R je u Boyce-Codd-ovoj normalnoj formi (BCNF) ako i samo ako su **sve determinante u relaciji i kandidati za ključ**.

BOYCE-CODD-OVA NORMALNA FORMA (BCNF)

PRIJAVA(BI, ŠPRED, NAZPRED, OCENA)

BI, ŠPRED \rightarrow NAZPRED, OCENA (D) (KK)

BI, NAZPRED \rightarrow ŠPRED, OCENA (D) (KK)

ŠPRED \rightarrow NAZPRED (D)

NAZPRED \rightarrow ŠPRED (D)

KK(BI, ŠPRED) i KK(BI, NAZPRED)

- Dekompozicijom, pri kojoj se iz relacije "izvlace" projekcije sa onim determinantama koje nisu kandidati za ključ, relacija se svodi na BCNF.

PRIJAVA1(BI, ŠPRED, OCENA)

PREDMET(ŠPRED, NAZPRED)

- Svaka relacija koja je u BCNF je sigurno i u 2NF i 3NF. Obrnuto ne važi.

DEKOMPOZICIJA NA ZAVISNE I NEZAVISNE PROJEKCIJE

- Projekcije R1 i R2 relacije R su nezavisne ako i samo ako važi sledeće:
 1. Svaka funkcionalna zavisnost u R se može logički dedukovati iz funkcionalnih zavisnosti u R1 i R2 i
 2. Zajednicki atribut relacija R1 i R2 je kandidat za ključ barem u jednoj od njih.

VIŠENAČNE ZAVISNOSTI I ČETVRTA NORMALNA FORMA

PREDMET NASTAVNIK KNJIGA

INF-SIST	BRANKO	MARTIN
	KRCA	DATE
	BAJA	
SIST-ANAL	VLADAN	DEMARCO
		SARSON

Jedan predmet predaje više nastavnika. Za jedan predmet se koristi više knjiga. Ne postoji nikakva veza između nastavnika i knjiga.

■ Normalizovana

PREDMET NASTAVNIK KNJIGA

INF-SIST	BRANKO	MARTIN
INF-SIST	BRANKO	DATE
INF-SIST	KRCA	MARTIN
INF-SIST	KRCA	DATE
INF-SIST	BAJA	MARTIN
INF-SIST	BAJA	DATE
SIST-ANAL	VLADAN	DEMARCO
SIST-ANAL	VLADAN	SARSON

VIŠENAČNE ZAVISNOSTI I ČETVRTA NORMALNA FORMA

RASP(PREDMET, NASTAVNIK) i UDŽBENIK(PREDMET, KNJIGA)

RASP	PREDMET	NASTAVNIK	UDŽBENIK	PREDMET	KNJIGA
	INF-SIST	BRANKO		INF-SIST	MARTIN
	INF-SIST	KRCA		INF-SIST	DATE
	INF-SIST	BAJA		SIST-ANAL	DEMARCO
	SIST-ANAL	VLADAN		SIST-ANAL	SARSON

- Dekompozicija je bez gubljenja informacija.
- Veze koje postoje između atributa ove relacije nazivaju se višeznacnim vezama.

VIŠENAČNE ZAVISNOSTI I ČETVRTA NORMALNA FORMA

- U relaciji $R(A, B, C)$ postoji višeznacna zavisnost $A \twoheadrightarrow B$ ako za datu vrednost A , postoji skup od nula, jedne ili više vrednosti B , a taj skup vrednosti ni na koji način ne zavisi od vrednosti atributa C . Atributi A , B i C mogu biti složeni.

- U relaciji $R(A, B, C)$ postoji višeznacna zavisnost $A \twoheadrightarrow B$ ako i samo ako kad god u njoj postoje n -torke $\langle a, b, c \rangle$ i $\langle a, b', c' \rangle$,

postoje takode i n -torke

$\langle a, b, c' \rangle$ i $\langle a, b', c \rangle$.

Atributi A , B i C mogu biti složeni.

- U prethodnom primeru važile su

PREDMET \twoheadrightarrow NASTAVNIK i PREDMET \twoheadrightarrow KNJIGA.

VIŠENAČNE ZAVISNOSTI I ČETVRTA NORMALNA FORMA

- Relacija R je u Četvrtoj normalnoj formi (4NF) ako i samo ako kad god postoji višeznačna funkcionalna zavisnost, na primer $A \twoheadrightarrow B$, tada svi atributi relacije moraju takode biti funkcionalno zavisni od A.
- Relacija R je u 4NF ako u njoj nisu date dve (ili više) nezavisne višeznačne činjenice.

ZAVISNOSTI SPAJANJA I PETA NORMALNA FORMA (5NF)

- U relaciji $R(X, Y, \dots, Z)$ postoji zavisnost spajanja ako i samo ako relacija R rezultuje iz prirodnog spajanja njenih projekcija po X, Y, \dots, Z , gde su X, Y, \dots, Z podskupovi atributa relacije R .
- Relacija R je u Petoj normalnoj formi ako i samo ako se svaka zavisnost spajanja može pripisati kandidatu za ključ.