

1. Neka je relacija Porudžbine data tabelom:

P(sifkup,ime,adresa,količina,sifartikla,naziv,kvalitet,cena).

Da li je data relacija normalizovana? Da li postoje anomalije? Odgovor obrazložiti.

2. Neka je relacija Programer data tabelom:

P(šifprogramera,ime,prezime,broj kancelarije,telefon kancelarije).

Da li je data relacija normalizovana? Da li postoje anomalije? Odgovor obrazložiti.

ZAKLJUČAK:

1. Projektovanju baze podataka može se prići E-R modelom, ali i na drugi način. Na primer, prostim skupljanjem svih raspoloživih podataka.

2. Logički model ovako dobijene baze podataka mora se organizovati tako da ne sadrži redundanciju da bi se sistem mogao efikasno koristiti (pretraživati, ažurirati, brisati itd.)

3. Tabele sa sirovo nabacanim podacima retko kad zadovoljavaju sve uslove kako bi baza funkcionisala nesmetano => NORMALIZACIJA

Tehnički postupak normalizacije svodi se na relacione operacije projekcije i spajanja, i to na:

- na dekompoziciju tabela generisanjem vertikalnih podskupova (operacijom projekcije) i
- na generisanje novih relacija iz dvaju ili više dobijenih vertikalnih podskupova (operacijom spajanja).
- Ovakav pristup izrade baze podataka naziva se još i vertikalna normalizacija. U sklopu vertikalne normalizacije definisano je do sada 5+1 normalnih formi, od kojih su najvažnije druga i treća, na kojima se često normalizacija i zaustavlja.

Pored vertikalne postoji i horizontalna normalizacija koja nije teorijski dovršena, i vezana je za distribuirane baze.

Prednosti normalizacije:

Smanjenje fizičkog prostora za smeštanje podataka

Bolja organizacija podataka

Promene podataka se rade na nivou samo jednog sloga

Mane normalizacije

Fizički prostor diska je danas jeftin – malo bitan (izuzev kod velikih baza tipa Data warehouses)

Minimizacija dovodi do visoke granularnosti. SQL JOIN QUERY se sporo izvršavaju

Nastaju šeme relacija koje karakteriše visoka kompleksnost sa stanovišta dizajnera i programera

Problem gubitka informacija

- Međusobne zavisnosti atributa nisu umetnute veze već veze koje postoje u realnom svetu i koje se preslikavaju u informacioni sistem. Tako da ako dekompozicijom ugrozimo te veze informacioni sistem prestaje da bude verodostojan.

- Relacija se dekomponuje bez opasnosti gubitaka funkcionalnih zavisnosti samo ako se dekompozicija vrši prema funkcionalnoj zavisnosti koja ne ide od kandidata ključa.

U teoriji, trebalo bi krenuti od jedinstvene relacione šeme (nekad se naziva i univerzalna šema, ili U) koja sadrži sve attribute baze podataka.

Potom se rekurzivno primenjuju pravila kako bi se dobio skup sve više i više

normalizovanih šema pod-relacija.

Kada su sve šeme u trećoj normalnoj formi, onda je cela baza normalizovana.

Otklanjanje anomalija

Jedan od načina rešavanja anomalija je dekompozicija relacija, pri čemu se jedna relacija zamenjuje sa više drugih relacija tako da se uklone anomalije. Mogu se prepoznati nekoliko pravila za normalizaciju:

Polja treba da budu prostih tipova.

Svaki slog – red u tabeli bi trebalo da ima jedinstveni identifikator – primarni ključ

Primarni ključ treba da bude jednostavan, stabilan i kratak

Svako polje bi trebalo da obezbedi neku dodatnu informaciju o slogu u kome primarni ključ služi za identifikaciju

Informacije u tabelama ne bi trebalo da se pojavljuju na više mesta, već samo ne jednom.

Ako su X i Y atributi, onda se notacija  $X \rightarrow Y$  čita kao „X funkcionalno određuje Y“ ili „Y je funkcionalno zavisno od X“.

Ovo znači da ako dobijemo tabelu sa podacima i vrednost X, da možemo da odredimo Y.

Nadključ uvek funkcionalno određuje sve attribute u relaciji (kao i sebe). Ovo je „dobra“ FZ. „Loša“ FD se dešava kada postoji atribut ili skup atributa koji predstavljaju nadključ za neke attribute, ali ne i za celu relaciju. Ovakvi atributi se nazivaju podključevi (subkey) relacije. Većina pravila normalizacije znači da se eliminišu redundantni podaci u bazi podataka. Podključevi uvek znače redundantne podatke.

Ako u tabelama ne postoje podključevi, onda se ona nalazi u trećoj normalnoj formi.

Međutim, ova normalna forma dozvoljava podključeve u nekim izuzetnim slučajevim, a normalna forma koja striktno ne dozvoljava podključeve je BCNF (Boyce-Coddova normalna forma).

3. Dati su šema relacije šema relacije pozajmica

P (SIFN, SIFC, DATUM, DANA, SIFK, NAZIVN, SIFO, NAZIVO)

i skup funkcijskih zavisnosti:

$SIFN, SIFC, DATUM \rightarrow DANA$	$SIFN, SIFC, DATUM \rightarrow SIFK$	
$SIFN, SIFC, DATUM \rightarrow NAZIVN$	$SIFN, SIFC, DATUM \rightarrow SIFO$	
$SIFN, SIFC, DATUM \rightarrow NAZIVO$	$SIFK \rightarrow SIFN$	
$SIFN \rightarrow SIFO$	$SIFN, SIFC \rightarrow SIFN$	$SIFO \rightarrow NAZIVO$

Identifikuj tipove funkcionalnih zavisnosti.

REŠENJE:

superključne

Neka je R šema relacije, a X, Y i Z su podskupovi od R.

Funkcijska zavisnost  $X \rightarrow Y$  je **superključna** ako važi  $X \rightarrow R$

$SIFN, SIFC, DATUM \rightarrow DANA$	$SIFN, SIFC, DATUM \rightarrow SIFK$
$SIFN, SIFC, DATUM \rightarrow NAZIVN$	$SIFN, SIFC, DATUM \rightarrow SIFO$
$SIFN, SIFC, DATUM \rightarrow NAZIVO$	

trivijalna

Funkcijska zavisnost  $X \rightarrow Y$  je **trivijalna** ako važi  $Y \subseteq X$

$SIFN, SIFC \rightarrow SIFN$

totalne

Funkcijska zavisnost je **totalna** ako ne postoji ni jedan pravi podskup Z od X za koji važi  $Z \rightarrow Y$  odnosno:  $X \rightarrow Y \wedge \neg \exists Z (Z \subset X \wedge Z \rightarrow Y)$

SIFN, SIFC, DATUM  $\rightarrow$  DANA      SIFN, SIFC, DATUM  $\rightarrow$  SIFK  
SIFK  $\rightarrow$  SIFN      SIFN  $\rightarrow$  SIFO      SIFO  $\rightarrow$  NAZIVO

parcijalne: Funkcijska zavisnost je **parcijalna** ako postoji neki pravi podskup Z od X za koji važi  $Z \rightarrow Y$  odnosno:  $X \rightarrow Y \wedge \exists Z (Z \subset X \wedge Z \rightarrow Y)$

SIFN, SIFC, DATUM  $\rightarrow$  SIFO

tranzitivne

Funkcijska zavisnost  $X \rightarrow Y$  je **tranzitivna** ako postoji neko Z, različito od X i Y za koje važi:  $X \rightarrow Z$  i  $Z \rightarrow Y$

SIFN, SIFC, DATUM  $\rightarrow$  NAZIVO

- 1NF      Atributi prostih, skalarnih tipova, atomični
- 2NF      Nema parcijalne zavisnosti ne-ključnog atributa od kandidat ključa
- 3NF      Nema tranzitivne zavisnosti ne-ključnog atributa od kandidat ključa
- BC NF    Nema dodatnih tranzitivnih zavisnosti

### Postupak normalizacije

Neka polazna relaciona šema nije u određenoj normalnoj formi, ako u skupu funkcijskih zavisnosti F postoji bar jedna koja narušava definiciju normalne forme

U svakom koraku normalizacije:

uočava se jedna takva zavisnost ( $X \rightarrow Y$ )

vrši se dekompozicija u cilju uklanjanja takve zavisnosti

Ako je u polaznoj važilo  $X \rightarrow Y$ , dekompozicijom nastaju dve relacije, u prvoj se gube atributi Y, a druga nastaje nad atributima X i Y

Ne dozvoljava se gubitak podataka

4. Dati su relacija AUTOR(SIFA, SIFN, IME, KOJI) i skup funkcijskih zavisnosti  $F = \{SIFA, SIFN \rightarrow IME, KOJI \rightarrow SIFA \rightarrow IME\}$ . Dovedi relaciju do 2NF.

REŠENJE: Setimo se da relacija je u 2NF ako svi atributi, koji nisu primarni ključevi, zavise totalno od (od celog) primarnog ključa

Problem. Postoji parcijalna zavisnost:  $SIFA \rightarrow IME$  koja isključuje SIFN.

Dakle, u postupku normalizacije treba  $SIFA \rightarrow IME$  izdvojiti u posebnu šemu relacije, a iz polazne šeme relacije ukloniti desnu stranu te zavisnosti, tj. IME.

Vrši se dekompozicija:

AUTOR1 (SIFA, SIFN, KOJI)	{SIFA, SIFN $\rightarrow$ KOJI}
AUTOR2 (SIFA, IME)	{SIFA $\rightarrow$ IME}

Dekompozicijom polazne relacije uklonjena je neželjena parcijalna funkcijska zavisnost

5. Dati su relacija o naslovima koja sadrži sve podatke o naslovima, oblastima i autorima i skup funkcijskih zavisnosti F. Dovedi relaciju do 3NF.

NASLOV(SIFN,SIFA,KOJI,NAZIVN,IME,SIFO,NAZIVO),  
F={ SIFN,SIFA → KOJI,NAZIVN,IME,SIFO,NAZIVO  
SIFN → NAZIVN,SIFO                      SIFA → IME  
SIFO → NAZIVO}

REŠENJE:

Problem u ovom primeru je tranzitivna zavisnost NAZIVO od SIFN, preko SIFO.

( SIFN → NAZIVN,SIFO    SIFO → NAZIVO).

Setimo se da relacija je u 3NF ako nad njom ne postoji ni jedna funkcijska zavisnost, po kojoj neki ne-ključni atribut tranzitivno zavisi od bilo kog kandidat-ključa

Uslovi za 3NF (dozvoljene funkcijske zavisnosti):

zavisnost  $X \rightarrow Y$  je trivijalna tj.  $Y \subseteq X$

zavisnost je superključna, tj.  $X \rightarrow R$

Y je deo kandidat ključa, tj.  $\exists Z(Y \subseteq Z \wedge Z \rightarrow R)$

Sledi normalizacija primenom kriterijuma za 3NF.

NASLOV(SIFN,SIFA,KOJI,NAZIVN,IME,SIFO,NAZIVO)

Uklanjam zavisnost SIFO → NAZIVO i dobija se

**NASLOV1**(SIFN,SIFA,KOJI,NAZIVN,IME,SIFO)    {SIFO → NAZIVO}

NASLOV(SIFN,SIFA,KOJI,NAZIVN,IME,SIFO)

Uklanjam zavisnost SIFN → NAZIVN,SIFO i dobija se

**NASLOV2**(SIFN,SIFA,KOJI,NAZIVN,IME,SIFO)    {SIFN → NAZIVN,SIFO}

NASLOV(SIFN,SIFA,KOJI,IME)

Uklanjam zavisnost SIFA → IME i dobija se

**NASLOV3**(SIFN,SIFA,KOJI,IME)    {SIFA → IME}

**NASLOV4**(SIFN,SIFA,KOJI)    {SIFN,SIFA → KOJI}

6. Data je šema relacije o pozajmicama, koja sadrži podatke o naslovima, članovima i knjigama:

POZAJMICA(SIFN,SIFC,DANA,NAZIVN,SIFK)

kao i odgovarajući skup funkcijskih zavisnosti:

F={ SIFN,SIFC,DANA → NAZIVN,SIFK  
SIFK → SIFN    SIFN → NAZIVN}

Dovedi relaciju do BCNF.

REŠENJE: Setimo sa da BCNF ima strožiju definiciju u odnosu na 3NF, jer se eliminišu dodatne tranzitivne zavisnosti.

Preciznije, šema relacije R je u BCNF ako svaka funkcijska zavisnost  $X \rightarrow Y$  koja važi nad njom zadovoljava jedan od uslova:

zavisnost je trivijalna tj.  $Y \subseteq X$

zavisnost je superključna, tj.  $X \rightarrow R$

Normalizacijom se dobija sledeća dekompozicija:

POZAJMICA(SIFN,SIFC,DANA,NAZIVN,SIFK)

Uklanjam zavisnost SIFN → NAZIVN

**POZAJMICA1**(SIFN,NAZIVN)

POZAJMICA(SIFN,SIFC,DANA,SIFK)

Uklanjamo zavisnost  $SIFK \rightarrow SIFN$

**POZAJMICA2**(SIFK,SIFN)

**POZAJMICA3**(SIFK,SIFC,DATUM)

7. Dati su šema relacije R (A, B, C, D, E) i skup funkcijskih zavisnosti  $F = \{ A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A \}$ . Odrediti skup kandidata ključeva date šeme.

REŠENJE:  $A^+ = A|BC|D|E$

$B^+ = B|D$

$C^+ = C$

$D^+ = D$

$E^+ = E|A|BC|D$

$(BC)^+ = BC|D|E|A$

$(BD)^+ = BD$

$(CD)^+ = CD|E|A|B$

$KK = \{A, E, BC, CD\}$