

Baze podataka
Katedra za računarstvo
Elektronski fakultet u Nišu

Analiza šeme relacione BP

Prof.dr Leonid Stoimenov

Pregled

- ▶ Uvod
- ▶ Anomalije kod ažuriranja
- ▶ Dekompozicija



Šta je analiza šeme relacione baze podataka?

- ▶ Proces određivanja kvaliteta projektovanih šema relacija
- ▶ U teoriji baza podataka definisani su formalizmi za analizu kvaliteta šeme relacije i **prevođenje** šeme relacije u **ekvivalentnu bolju šemu**
- ▶ Postoje dva takva formalizma:
 - ▶ Logičke zavisnosti i
 - ▶ Normalne forme



Kvalitet projektovane šeme relacije

- ▶ Kvalitet se može posmatrati sa
 - ▶ logičkog (sa strane korisnika)
 - ▶ fizičkog nivoa (sa strane sistema)
- ▶ Na **logičkom nivou** od šeme se zahteva da bude jasna i laka za razumevanje
- ▶ Na **fizičkom nivou** od šeme se zahteva da obezbedi ažurni skup podataka uz minimalan utrošak računarskih resursa



Kvalitet projektovane šeme relacije

- ▶ Za merenje kvaliteta šeme relacije koriste se sledeći faktori:
 - ▶ semantika atributa relacije
 - ▶ redukovanje redundantnih vrednosti u torkama
 - ▶ redukovanje NULL vrednosti u torkama
 - ▶ isključivanje lažnih torki
-



Logičke zavisnosti

- ▶ Zavisnosti izražavaju činjenicu da jedan objekat zavisi od drugih objekata
- ▶ U relacionom modelu podataka **objekti** su **atributi** i **torke**
- ▶ U šemi relacije $\mathbf{R}(\mathbf{R};\mathbb{F})$ nad skupom atributa $\mathbf{R}=(A_1, A_2, \dots, A_n)$ mogu postojati različita ograničenja zavisnosti kao posledica osobina i pravila funkcionisanja realnog sistema koji se ovom relacijom modelira
 - ▶ najčešće se susreću **funkcionalne** i **višeznačne** zavisnosti



Anomalije pri obradi relacija (1)

- ▶ Nepoželjno ponašanje relacija pri izvođenju operacija ažuriranja naziva se **anomalijom ažuriranja**
- ▶ Na osnovu operacije koja indukuje anomalije ažuriranja razlikujemo:
 - ▶ anomalije unosa
 - ▶ anomalije brisanja
 - ▶ anomalije modifikacije



Anomalije pri obradi relacija (2)

- ▶ Anomalije ažuriranja ćemo razmotriti na primeru relacije radsek nad šemom relacije:

RADSEK(RMBR, RIME, RLD, SBR, SNAZIV, SRUK)

- ▶ Ova relacija sadrži attribute dva entiteta: **RADNIK** i **SEKTOR**
 - ▶ Entitet **RADNIK** je opisan atributima {RMBR, RIME, RLD}, a entitet **SEKTOR** atributima {SBR, SNAZIV, SRUK}
 - ▶ U šemi relacije **RADSEK** postoje sledeće funkcionalne zavisnosti nad atributima:
 $f1: RMBR \rightarrow \{RIME, RLD, SBR\}$
 $f2: SBR \rightarrow \{SNAZIV, SRUK\}$
- ▶ Pri izvođenju operacija nad ovom relacijom javljaju se sve tri vrste anomalija



Primer loše projektovane relacije

RADSEK(RMBR, RIME, RLD, SBR, SNAZIV, *SRUK*)

RADSEK

RMBR	RIME	RLD	SBR	SNAZIV	<i>SRUK</i>
111	Laza	100000	10	Projektni biro	333
222	Mara	60000	10	Projektni biro	333
333	Nada	150000	10	Projektni biro	333
444	Djura	50000	20	Kontrola	null

Loše projektovana relacija

RADSEK

RMBR	RIME	RLD	SBR	SNAZIV	SRUK
111	Laza	100000	10	Projektni biro	333
222	Mara	60000	10	Projektni biro	333
333	Nada	150000	10	Projektni biro	333
444	Djura	50000	20	Kontrola	null

RADNIK

RMBR	RIME	RLD	SBR
111	Laza	10000	10
222	Mara	60000	10
333	Nada	150000	10
444	Djura	50000	20

Dobro projektovane relacije

SEKTOR

SBR	SNAZIV	SRUK
10	Projektni biro	333
20	Kontrola	null



Anomalije unosa (1)

Relacija **RADSEK**(**RMBR**, RIME, RLD, SBR, SNAZIV, SRUK)
poseduje anomalije unosa koje se ogledaju u sledećem :

- ▶ U relaciju **RADSEK** se **ne mogu** uneti podaci o sektoru ukoliko u taj sektor nije raspoređen nijedan radnik.
- ▶ Ukoliko bi se takav unos dozvolio tada bi atributi RMBR, RIME i RLD imali nedefinisanu (NULL) vrednost, što nije dozvoljeno budući da je RMBR primarni ključ relacije



Anomalije unosa (2)

- ▶ Relacija **RADSEK(RMBR, RIME, RLD, SBR, SNAZIV, SRUK)** **poseduje anomalije unosa** koje se ogledaju u sledećem :
 - ▶ Pri unosu podataka o novom radniku potrebno je, uz konkretne vrednosti za sve attribute entiteta RADNIK, uneti konkretne vrednosti za sve attribute entiteta SEKTOR.
 - ▶ Pri svakom unosu moramo voditi računa da podaci o sektoru budu konzistentni sa već unetim podacima o tom sektoru.
 - ▶ Ponavljanje podataka o sektoru u torkama svih radnika tog sektora govori o **prisustvu redundance u relaciji**.
-



Anomalije unosa (3)

RADSEK

RMBR	RIME	RLD	SBR	SNAZIV	SRUK
111	Laza	100000	10	Projektni biro	333
222	Mara	60000	10	Projektni biro	333
333	Nada	150000	10	Projektni biro	333
444	Djura	50000	20	Kontrola	Null
?	?	?	30	Razvoj	Null
555	Pera	20000	30	Razvoj	Null

Anomalije unosa (4)

- ▶ Opisane anomalije su posledica **tranzitivne funkcionalne zavisnosti** među atributima relacije:
 $RMBR \rightarrow \{RIME, RLD, SBR\},$
 $SBR \rightarrow \{SNAZIV, SRUK\},$
SBR nije podskup ključa relacije
 - ▶ Anomalije se mogu otkloniti **dekompozicijom** šeme relacije **RADSEK** tako da se **svaka funkcionalna zavisnost** predstavi posebnom relacijom
 - ▶ Rezultat ove dekompozicije su sledeće šeme relacija:
RADNIK(RMBR, RIME, RLD, SBR)
SEKTOR(SBR, SNAZIV, SRUK)
 - ▶ Može se uočiti da svaka nova šema relacije sadrži samo attribute jednog entiteta
-



Anomalije brisanja

- ▶ Relacija **RADSEK**(**RMBR**, RIME, RLD, SBR, SNAZIV, **SRUK**) poseduje anomalije brisanja koje se ogledaju u sledećem:
 - ▶ Ako iz relacije **RADSEK** **obrišemo sve radnike** koji rade u nekom sektoru, tada ćemo **obrisati i podatke o tom sektoru**. Na taj način se iz baze podataka gubi informacija o sektorima koji trenutno nemaju raspoređene radnike.

RADSEK

RMBR	RIME	RLD	SBR	SNAZIV	SRUK
111	Laza	100000	10	Projektni biro	333
222	Mara	60000	10	Projektni biro	333
333	Nada	150000	10	Projektni biro	333
444	Djura	50000	20	Kontrola	null

Anomalija modifikacije

- ▶ Relacija **RADSEK**(**RMBR**, RIME, RLD, SBR, SNAZIV, SRUK) poseduje anomalije modifikacije koje se ogledaju u sledećem:
 - ▶ Pri **promeni vrednosti** nekog od atributa entiteta **SEKTOR** (na primer, SRUK) potrebno je da se **izmeni vrednost** atributa SRUK u **torkama** svih radnika koji rade u tom sektoru.
 - ▶ Dok proces izmene traje postoji nekozistentnost u podacima. U našem primeru svi radnici koji rade u sektoru čiji atribut SRUK menjamo neće imati istu vrednost za SRUK. Ova pojava se naziva anomalijom modifikacije, a posledica je ranije pomenute tranzitivne zavisnosti među atributima relacije.
-



Rezime o anomalijama ažuriranja

- ▶ Razmatrane anomalije ažuriranja su posledica izvesnih **neželjenih struktura funkcionalnih zavisnosti** u šemi relacije
 - ▶ Eliminisanjem svih nepoželjnih struktura funkcionalnih zavisnosti eliminišu se i anomalije ažuriranja
 - ▶ Za šeme relacija koje su formirane pod ovim restrikcijama kaže se da su u **normalnoj formi**, a postupak kojim se šema relacije prevodi u neku normalnu formu naziva se **normalizacija**
 - ▶ Pri normalizaciji se koristi tehnika **dekompozicije**
 - ▶ Dekompozicija šeme relacije nije proizvoljna. Po pravilu se zasniva na određenoj strukturi funkcionalnih zavisnosti među atributima relacije
 - ▶ Dekompozicijom se svaka funkcionalna zavisnost među atributima šeme relacije izoluje u posebnu šemu relacije
-



Dekompozicija relacija

- ▶ **Dekompozicija** je proces zamene šeme relacije $R(R;F)$ skupom šema relacija $D=\{R_1(R_1;F_1), \dots, R_m(R_m;F_m)\}$ pri čemu moraju biti ispunjeni sledeći uslovi:
 1. **Očuvanje atributa:** Nijedan atribut nesme biti izgubljen, što znači da se svaki atribut iz R mora naći bar u jednoj šemi R_i dekompozicije D
 2. **Spoj-bez-gubitaka (neaditivni spoj)** - osigurava da se pri prirodnom spoju relacija dobijenih dekomponovanjem ne jave **lažne torke**. Svojstvo **spoj-bez-gubitaka** odnosi se na **gubitak informacije**, a ne na gubitak torki
 3. **Očuvanje zavisnosti.** Dekompozicija obezbeđuje očuvanje zavisnosti ako su sve funkcionalne zavisnosti šeme $(R;F)$ zadržane u D
- ▶ **Dekompozicija je dobra** ako je skup šema relacija D ekvivalentan šemi R i ako su isključene anomalije

Svojstvo spoj-bez-gubitaka za binarnu dekompoziciju

- ▶ Dekompozicija $D=\{\mathbf{R1}, \mathbf{R2}\}$ šeme relacije $\mathbf{R}(R;F)$ ima svojstvo **spoj-bez-gubitaka** u odnosu na skup funkcionalnih zavisnosti F u \mathbf{R} ako i samo ako važi:
 - (Uslov 1) $R1 \cap R2 \rightarrow R1 \in F^+$ ili
 - (Uslov 2) $R1 \cap R2 \rightarrow R2 \in F^+$, gde je:
- ▶ $R1 \cap R2 = Z$ skup zajedničkih atributa obe komponente
- ▶ (Uslov 1) Z je super ključ komponente $R1$
- ▶ (Uslov 2) Z je super ključ komponente $R2$
- ▶ Ako je $R1 \cap R2$ super ključ šeme relacije $\mathbf{R1}$ ili $\mathbf{R2}$, tada binarna dekompozicija ima svojstvo spoj-bez-gubitaka
 - ▶ Za proveru da li je $R1 \cap R2$ super ključ treba koristiti algoritam za nalaženje zatvarača skupa atributa

Analiza šeme

Pitanja ???