Projektovanje baza podataka



Projektovanje šeme relacione BP

Anomalije ažuriranja, normalne forme

Sadržaj



- Anomalije ažuriranja
- Osnovni projektantski kriterijumi
- Normalne forme i normalizacija



Motivacija za potrebu projektovanja šeme BP

- šema univerzalne relacije (*U*, *OGR*) se, praktično, ne može implementirati, jer
 - skupovi *U* i *OGR* su preglomazni
 (i do nekoliko hiljada obeležja i ograničenja)
 - nemoguće je (*U*, *OGR*) sagledati u celini, a to nema ni logičkog smisla
 - univerzalna relacija bi bila, takođe, prevelika
 - javljaju se anomalije ažuriranja
 - javlja se nepotrebna redundansa podataka



Vrste anomalija ažuriranja

- anomalije upisa
- anomalije brisanja
- anomalije modifikacije (redundanse)





Anomalije upisa

- F
- moraju se, pri pokušaju upisa podataka o jednom entitetu, znati vrednosti obeležja svih povezanih entiteta
- potrebno je zadati sve vrednosti obeležja ključa
 - a neke od njih, međutim, nisu poznate u trenutku upisa

Anomalije brisanja

- brisanjem jedne torke gube se, na neželjen način, informacije o različitim realnim entitetima
 - koji su povezani sa entitetom kojeg reprezentuje brisana torka





- Anomalije modifikacije (redundanse)
 - modifikacija vrednosti obeležja istog realnog entiteta obavlja se na više mesta u relaciji
 - jer se nepotrebno ponavlja na više mesta u relaciji
 - često, za takvu operaciju, zahteva se prolaz kroz celu relaciju



Primer

- semantika obeležja koja postoje u univerzalnom skupu obeležja
 - BRI broj indeksa
 - IME ime studenta
 - PRZ prezime studenta
 - BPI broj položenih ispita
 - OZP oznaka predmeta
 - NAP naziv predmeta
 - NAS prezime nastavnika
 - OCE ocena na ispitu



Primer



Student

BRI	IME	PRZ	BPI	OZP	NAP	NAS	OCE
159	Ivo	Ban	13	P1	Mat	Han	09
159	Ivo	Ban	13	P2	Fiz	Kun	08
013	Ana	Tot	09	P1	Mat	Pap	06
119	Eva	Kon	15	P3	Hem	Kiš	07
159	Ivo	Ban	13	P3	Hem	Kiš	10
119	Eva	Kon	15	P1	Mat	Han	09
159	Ivo	Ban	13	P4	Mat	Car	10
037	Eva	Tot	01	P4	Mat	Car	10



- Primer nekih anomalija ažuriranja
 - Ne mogu se upisati podaci o novom studentu, dok student ne položi makar jedan ispit
 - Brisanjem poslednjeg položenog ispita nekog studenta, gube se i osnovni podaci (*IME*, *PRZ* i *BPI*) o studentu
 - Promena prezimena jednog studenta se mora sprovesti u više od jedne torke



Primer

$$F = \{BRI \rightarrow IME + PRZ + BPI, IME + PRZ \rightarrow BRI, OZP \rightarrow NAP, NAS \rightarrow OZP + NAP, BRI + OZP \rightarrow OCE + NAS\}$$

- šema relacije Student ima četiri ključa
 - $K_1 = BRI + NAS$, $K_2 = IME + PRZ + NAS$,
 - $K_3 = BRI + OZP$, $K_4 = IME + PRZ + OZP$
- Pojam ključa
 - vrlo bitan za sagledavanje anomalija ažuriranja

Sadržaj



- Anomalije ažuriranja
- Osnovni projektantski kriterijumi
- Normalne forme i normalizacija



Osnovni projektantski kriterijumi

- U teoriji RMP polazi se od pretpostavke
 - da jedna šema relacije (*U*, *OGR*) predstavlja inicijalni
 model realnog sistema

- Projektovanje šeme BP
 - može se vršiti dekomponovanjem (rastavljanjem)
 šeme relacije (*U*, *OGR*) na više drugih šema relacija relacione šeme BP (*S*, *I*)

Osnovni projektantski kriterijumi

- Šema BP (S, I) treba da zadovolji sledeće kriterijume u odnosu na (U, OGR)
 - (K1)

F

da predstavlja dekompoziciju ŠUR

$$(\forall N_i \in S)(R_i \neq \emptyset) \land \bigcup_{N_i \in S}(R_i) = U$$

- -(K2)
 - da se garantuje spojivost bez gubitaka informacija

$$r(\mathbf{U}, \mathbf{OGR}) = \triangleright \triangleleft_{N_i \in S}(r_i(R_i))$$

Osnovni projektantski kriterijumi

- Šema BP (S, I) treba da zadovolji sledeće kriterijume u odnosu na (U, OGR)
 - -(K3)
 - skup svih ograničenja da bude ekvivalentan polaznom skupu
 ograničenja OGR

$$\bigcup_{N_i \in S} (O_i) \cup I \equiv OGR$$

- -(K4)
 - da se otklone sve anomalije ažuriranja

- U praksi je, često, nemoguće, ili nepotrebno strogo ispoštovati kriterijume K1-K4
 - kriterijumi se mogu, po potrebi, "oslabiti"



Osnovni projektantski kriterijumi

- Pojmovi relacionog MP, važni za projektovanje relacione šeme BP
 - funkcionalna zavisnost $X \rightarrow Y$, $XY \subseteq U$ $(\forall u, v \in r)(u[X] = v[X] \Rightarrow u[Y] = v[Y])$
 - projekcija skupa fz \boldsymbol{F} na skup obeležja $X \subseteq \boldsymbol{U}$ $\boldsymbol{F}|_{X} = \{V \rightarrow W \mid \boldsymbol{F} \models V \rightarrow W \land VW \subseteq X\}$
 - projekcija relacije r na skup obeležja $X \subseteq R$ $\pi_X(r(R)) = \{t[X] \mid t \in r(R)\}$
 - prirodni spoj relacija $r_1(R_1, \mathbf{F}_1)$ i $r_2(R_2, \mathbf{F}_2)$ $r_1 \triangleright \triangleleft r_2 = \{t \mid t[R_1] \in r_1 \land t[R_2] \in r_2\}$

Sadržaj



- Anomalije ažuriranja
- Osnovni projektantski kriterijumi
- Normalne forme i normalizacija

- Moguće je izbegavanje, ili u idealnom slučaju, potpuno uklanjanje anomalija ažuriranja
- Šema BP treba da zadovolji kriterijum odgovarajuće normalne forme
- Postoji sedam normalnih formi
 - 1NF, 2NF, 3NF, BCNF
 - 4NF, 5NF (PJNF), DKNF
- Za praksu su najbitnije prve četiri
 - temelje se na pojmovima fz i ključa

- Prva normalna forma (1NF)
 - šema relacije N(R, O) je u 1NF ako
 - R sadrži samo elementarna obeležja
 - za svaku pojavu r(N) važi da su sve vrednosti svih obeležja iz R atomarne
 - ne predstavljaju niz, ili skup drugih vrednosti iz domena obeležja
- Šema BP (S, I) je u 1NF ako su sve šeme relacija skupa S u 1NF

Normalne forme i normalizacija

Primer

r	N	1	Р
	n_1	i ₁	p_1
	n_1	i ₁	p_2
	n_2	i ₁	p_1
	n_3	i ₂	p_3

r	N	1/	P
	n_{\uparrow}	<i>i</i> ₁	(p_1, p_2)
	n_1	i_1	(p_3, p_4)
	p_2	17	(p_1, ω)
	n_3	i_2	(p_3, ω)



Druga normalna forma (2NF)

- šema relacije N(R, F) sa skupom ključeva K je u 2NF ako je
 - u 1NF i
 - ako je svako neprimarno obeležje u potpunoj funkcionalnoj zavisnosti od svakog ključa

$$(\forall A \in R \setminus Kpr)(\forall X \in K)(\forall Y \subset X)(F \neq Y \rightarrow A)$$

 $-\mathit{Kpr} = \bigcup_{X \in \mathit{K}}(\mathit{K})$ - skup primarnih obeležja šeme relacije N

 Šema BP (S, I) je u 2NF ako su sve šeme relacija skupa S u 2NF

Normalne forme i normalizacija

- Primer
 - ¬(2NF)

Student({BRI, PRZ, IME, BPI, OZP, NAR},

 $\{BRI \rightarrow PRZ + MNE + BPI, OZP \rightarrow NAP\}$

- K = {BRI+OZP}
- BRI+OZP→NAP
 - NAP neprimarno obeležje
 - nepotpuna fz
 - sledi iz OZP→NAP
- BRI+OZP→PRZ+IME+BPI
 - PRZ, IME, BPI neprimarna obeležja
 - nepotpuna fz
 - sledi iz BRI→PRZ+IME+BPI

Normalne forme i normalizacija

- Treća normalna forma (3NF)
 - šema relacije N(R, F) sa skupom ključeva K je u 3NF ako je
 - u 1NF i
 - ako je svako neprimarno obeležje samo u netranzitivnoj funkcionalnoj zavisnosti od svakog ključa

$$(\forall A \in R \setminus Kpr)(\forall X \in K)(\forall Y \subseteq R \setminus \{A\})(F |= Y \rightarrow A \Rightarrow F |= Y \rightarrow X)$$

 $-\mathit{Kpr} = \bigcup_{K \in \mathcal{R}} (K)$ - skup primarnih obeležja šeme relacije N

 Šema BP (S, I) je u 3NF ako su sve šeme relacija skupa S u 3NF

Normalne forme i normalizacija

- Primer
 - 2NF i ¬(3NF)

Student({BRI, PRZ, IME, SOD, NAQ},

 $\{BRI \rightarrow PRZ + ME + SOD, SOD \rightarrow NAO\}$

- $K = \{BRI\}$
- BRI→NAO
 - tranzitivna FZ
 - BRI je ključ
 - NAO je neprimarno obeležje
 - sledi iz BRI→SOD i SOD→NAO i $\neg(SOD$ →BRI)

Normalne forme i normalizacija

- Boyce-Codd normalna forma (BCNF)
 - šema relacije N(R, F) sa skupom ključeva K je u BCNF ako je
 - u 1NF i
 - svaka netrivijalna funkcionalna zavisnost bilo kog atributa mora sadržati ključ s leve strane

$$(\forall A \in R)(\forall Y \subseteq R \setminus A)(F |= Y \rightarrow A \Rightarrow (\exists X \in K)(X \subseteq Y))$$

 Šema BP (S, I) je u BCNF ako su sve šeme relacija skupa S u BCNF

Normalne forme i normalizacija

- Primer
 - -3NFi (BCNF)

Poveravanje($\{OZP, NAP, OZN\}, \{OZP \rightarrow NAP, NAP \rightarrow OZP\}$)

- *K* = {*OZP*+*OZN*, *NAP*+*OZN*}
- OZP+OZN→NAP
 - OZP+OZN je ključ
 - nepotpuna fz
 - sledi iz OZP→NAP

- Odnos uslova normalnih formi
 - 1NF je potreban uslov za sve više normalne forme
 - ugrađen u definicije uslova svih ostalih normalnih formi
 - $-3NF \Rightarrow 2NF = \boxed{}$
 - dokaz
 - $\neg 2NF \Rightarrow \neg 3NF$
 - $-\neg 2NF \Rightarrow (\exists A \in R \setminus Kpr)(\exists X \in K)(\exists Y \subset X)(F \models Y \rightarrow A)$
 - $-\Rightarrow$ Za takve X, Y i A važi: $X\rightarrow Y$, $Y\rightarrow A$ i $\neg(Y\rightarrow X)$
 - $-\Rightarrow X \rightarrow A$ je tranzitivna fz od ključa $\Rightarrow \neg 3NF$
 - komentari, za sve netrivijalne fz iz F⁺
 - 2NF i 3NF zabranjuju postojanje nepotpunih fz neprimarnih obeležja od ključa
 - 2NF dozvoljava postojanje fz između neprimarnih obeležja
 - 3NF zabranjuje postojanje fz između neprimarnih obeležja
 - » sve fz neprimarnih obeležja sadrže ključ s leve strane

- Odnos uslova normalnih formi
 - Alternativna (ekvivalentna) formulacija uslova 3NF
 - 1NF i
 - svaka netrivijalna funkcionalna zavisnost bilo kog neprimarnog atributa mora sadržati ključ s leve strane

$$(\forall A \in R \setminus Kpr)(\forall Y \subseteq R \setminus \{A\})(F |= Y \rightarrow A \Rightarrow (\exists X \in K)(X \subseteq Y))$$

- dokaz, obratom po kontrapozicji (⇒)
 - $-A \in R \setminus Kpr, Y \subseteq R \setminus A$, važi $F \models Y \rightarrow A \land (\forall X \in K)(X \not\subseteq Y)$
 - $\Rightarrow (\forall X \in K)(F \mid= X \rightarrow Y \land F \mid= Y \rightarrow A \land F \not\models Y \rightarrow X)$
 - » inače bi Y sadržao ključ, što je suprotno $(\forall X \in K)(X \not\subseteq Y)$
 - $-\Rightarrow F \models X \rightarrow A$ je tranzitivna fz, što je kontradiktorno def. 3NF
- dokaz, obratom po kontrapozicji (⇐)
 - $-A \in R \setminus Kpr, Y \subseteq R \setminus A$, važi $F \models Y \rightarrow A \land F \not\models Y \rightarrow X$
 - $-\Rightarrow A \in R \setminus Kpr, Y \subseteq R \setminus A$, važi $F \models Y \rightarrow A \land (\forall X \in K)(X \not\subseteq Y)$
 - \Rightarrow postoji fz $Y \rightarrow A$, takva da Y ne sadrži ni jedan ključ kontrad.

- Odnos uslova normalnih formi
 - BCNF \Rightarrow 3NF
 - $(\forall A \in R)(\forall Y \subseteq R \setminus A)(F \mid Y \to A \Rightarrow (\exists X \in K)(X \subseteq Y))$
 - $\Rightarrow (\forall A \in R \setminus Kpr)(\forall Y \subseteq R \setminus \{A\})(F \models Y \rightarrow A \Rightarrow (\exists X \in K)(X \subseteq Y))$ » pošto je R \ Kpr ⊆ R
 - → 3NF, po prethodnom tvrđenju
 - komentari, za sve netrivijalne fz iz F⁺
 - 2NF zabranjuje postojanje nepotpunih fz neprimarnih obeležja od ključa
 - 2NF dozvoljava postojanje fz između neprimarnih obeležja
 - 3NF zabranjuje postojanje fz između neprimarnih obeležja
 » sve fz neprimarnih obeležja sadrže ključ s leve strane
 - 3NF dozvoljava postojanje fz primarnih obeležja, koje ne sadrže ključ s leve strane
 - BCNF zabranjuje postojanje fz koje ne sadrže ključ s leve strane



Odnos uslova normalnih formi

Zabrane / Normalne forme	1NF	2NF	3NF	BCNF
nizovi ili skupovi vrednosti obeležja, umesto jedne vrednosti iz domena	NE	NE	NE	NE
nepotpune fz neprimarnih obeležja od ključa iz <i>F</i> ⁺	DA	NE	NE	NE
netrivijalne fz neprimarnih obeležja koje ne sadrže ključ s leve strane – tranzitivne fz neprimarnih obeležja iz <i>F</i> ⁺	DA	DA	NE	NE
netrivijalne fz iz F+ koje ne sadrže ključ s leve strane	DA	DA	DA	NE



- Normalizacija
 - postupak projektovanja takvog skupa šema relacija
 - kod kojeg su u celosti, ili delimično, ostvareni osnovni projektantski kriterijumi (K1-K4) i
 - zadovoljena je odgovarajuća normalna forma

- Dve metode normalizacije
 - Metoda dekompozicije
 - Metoda sinteze



Metoda dekompozicije

- postupak sistematskog rastavljanja šeme relacije na po dve šeme relacije
- postupak započinje od univerzalne šeme relacije
- postupak završava kada se obezbedi da se anomalije ažuriranja u potpunosti, ili u traženoj meri izbegnu
 - postizanjem željene normalne forme (do 5NF)



Metoda dekompozicije

- rastavljanje na osnovu fz (do BCNF) i drugih zavisnosti do viših normalnih formi
 - višeznačnih zavisnosti, do 4NF
 - zavisnosti spoja, do 5NF
- garantuje se očuvanje polaznog skupa obeležja
- garantuje se obezbeđenje spoja bez gubitaka
- ne garantuje se očuvanje polaznog skupa fz
 - može se garantovati očuvanje polaznog skupa fz, do zadovoljenja 3NF
 - može doći do narušavanja polaznog skupa fz, pri prelasku iz 3NF u BCNF



Metoda sinteze

- postupak sintetizovanja ("sklapanja") skupa šema relacija, na osnovu definisanog skupa fz i skupa obeležja
- može se garantovati očuvanje polaznog skupa obeležja
- može se garantovati obezbeđenje spoja bez gubitaka, ukoliko je to potrebno
- garantuje se očuvanje polaznog skupa fz
- garantuje se očuvanje uslova 3NF
 - anomalije ažuriranja izbegavaju se do nivoa, definisanog uslovom 3NF

Završne napomene



- Kada je broj obeležja mali (do 20) primena normalizacije se čini nepotrebnom
 - jer bi iste rezultate dao i intuitivni ekpertski pristup
- Kada je broj obeležja veliki
 - manuelna primena normalizacije je teška i podložna greškama
 - ekspertski pristup daje teško predvidive rezultate
- Rešenje
 - primena CASE / MDSD alata
 - projektovanje formalnih specifikacija konceptualne šeme BP
 - niz transformacija formalnih specifikacija za dolazak do
 - implementacione šeme BP
 - interne šeme BP / opisa šeme BP u jeziku ciljnog SUBP



Literatura

- Pavle Mogin, Ivan Luković, Miro Govedarica:
 Principi projektovanja baza podataka
 - Glava 1 i Poglavlje 2.1

Sadržaj



- Anomalije ažuriranja
- Osnovni projektantski kriterijumi
- Normalne forme i normalizacija



Pitanja i komentari



jektovanje baza podataka



Projektovanje šeme relacione BP

Anomalije ažuriranja, normalne forme