

Normalizacija sistema

Normalne forme nam omogućavaju da imamo “normalnu” strukturu relacija(tabela) u bazi podataka. To znači da će struktura nakon uvođenja ovih principa biti tako struktuirana da će izmene podataka biti daleko lakše, uticace pozitivno na performanse u ogromnoj većini slučajeva. (U slučaju da postoji određene tačke, pokazalo se da u određenim slučajevima vraćanje dela strukture na nernormalizovane forme ima bolje performanse).

Pozitivni rezultati uvođenja normalnih formi:

- + performanse - brža pretraga
- + lakša izmena podataka
- + manje memorije potrebno za čuvanje ovakve strukture
- + veća fleksibilnost u održavanju podataka od strane klijenata
- + sistem je modularan

Nenormalizovana tabela se prepoznaje po tome što sadrži više kolona od potrebnih da se zapamti što preciznije opisan entitet iz realnog sistema. Za logičko povezivanje podataka (predstavljanje veza između entiteta u sistemu) se koriste ključevi pomoću kojih se podaci povezuju u neku nenormalizovanu strukturu pogodnu korisniku, ali ne i dobro organizovanu strukturu baze podataka. Nenormalizovana tabela ima svojstvo da gubi broj kolona u procesu normalizacije, da kolone koje se “izbace” iz tabele koju normalizujemo prelaze u neke postojeće tabele ili stvaraju nove tabele.

I NF – PRVA NORMALNA FORMA

Omogućava da se za pojavljivanje jednog reda u tabeli u svim kolonama nalazi po tačno jedna logička vrednost:

Tabela nije u I NF:

Studenti_i_predmeti

student_id	student_ime	student_prezime	predmet_naziv
1	Mitar	Mirić	Programiranje 1 Programiranje 2

Tabela u I NF:

Studenti_i_predmeti

student_id	student_ime	student_prezime	broj_predavanja
1	Mitar	Mirić	Programiranje 1
2	Mitar	Mirić	Programiranje 2

II NF – DRUGA NORMALNA FORMA

Druga normalna forma podrazumeva da tabela ispunjava uslov za prvu normalnu formu.

Proširićemo tabelu kako bismo pamtili još jedan bitan podatak broj predavanja za određeni predmet.

Studenti_i_predmeti

student_id	student_ime	student_prezime	predmet_naziv	broj_predavanja
1	Mitar	Mirić	Programiranje 1	50
2	Mitar	Mirić	Programiranje 2	50

Atribute koji ne zavise od svih kandidata za ključ, već samo od dela kandidata za ključ ćemo izdvojiti u drugu tabelu.

Kandidati za ključ su student_id, student_ime, student_prezime i predmet_naziv. To znači da smo uzeli u razmatranje ove kolone da bi odabrali primarni ključ u tabeli – skup kolona čije će vrednosti za konkretni red jednoznačno odrediti sve ostale vrednosti u tabeli.

Ali broj predavanja logički zavisi

Na primer ako želimo da pratimo broj predavanja po predmetu to je bolje podatke o samom predmetu držati na drugom mestu jer ne zavise od svih kandidata za ključ već samo od određenih, u ovom primeru broj predavanja zavisi od samog predmeta koji je kandidat za ključ.

Studenti_i_predmeti

student_id	student_ime	student_prezime	predmet_naziv
1	Mitar	Mirić	Programiranje 1
2	Mitar	Mirić	Programiranje 2

Predmet

predmet_naziv	broj_predavanja
Programiranje 1	50
Programiranje 2	50

Ovo je takođe bolje uraditi iz razloga što na broj predavanja neće uticati student koji sluša dati predmet, već sam predmet.

III NF – TREĆA NORMALNA FORMA

Treća normalna forma podrazumeva da je tabela u drugoj normalnoj formi (samim tim je i u prvoj) i ako nema tranzitivnih zavisnosti.

Uvešćemo nove podatke od značaja za pamćenje kako bismo bolje pokazali primer druge normalne forme, mesto iz koga dolazi student i poštanski broj mesta. Druga normalna forma otklanja problem parcijalne zavisnosti.

Student

student_id	student_ime	student_prezime	predmet_naziv	poštanski_broj	mesto
1	Mitar	Mirić	Programiranje 1	11000	Beograd
2	Mitar	Mirić	Programiranje 2	11000	Beograd

Bitno nam je da pamtimo o studentu mesto iz koga dolazi kao i poštanski broj mesta, ali je sam naziv mesta vezan za poštanski broj. U ovom slučaju potrebno je izdvojiti mesto i poštanski broj u posebnu tabelu Mesto (kolona mesto može promeniti naziv u mesto_naziv na primer, jer će ime tabele biti Mesto / odabrano je namerno ime mesto_naziv). U tabeli Student će ostati samo poštanski broj kao ključni atribut preko koga je moguće doći do podataka o mestu.

Student

student_id	student_ime	student_prezime	predmet_naziv	poštanski_broj
1	Mitar	Mirić	Programiranje 1	11000
2	Mitar	Mirić	Programiranje 2	11000

Mesto

poštanski_broj	mesto_naziv
11000	Beograd

Možemo zaključiti da će studenti slušati puno predmeta na fakultetu i da će se nazivi predmeta mnogo puta ponavljati u jednoj tabeli. Ovo će za posledicu imati “teže” održavanje sistema – ako je došlo do greške u nazivu predmeta izmena naziva predmeta bi trebala da se desi u onoliko redova u koliko se pojavljuje naziv samog predmeta, kod svih studenata. Ako bismo podatak naziv predmeta izdvojili u posebnu tabelu i naziv predstavili preko kolone id, dobili bismo daleko lakše održavanje u pogledu izmene i pretrage podataka. Nakon izmene struktura će izgledati ovako:

Student

student_id	student_ime	student_prezime	predmet_id	poštanski_broj
1	Mitar	Mirić	1	11000
2	Mitar	Mirić	2	11000

Predmet

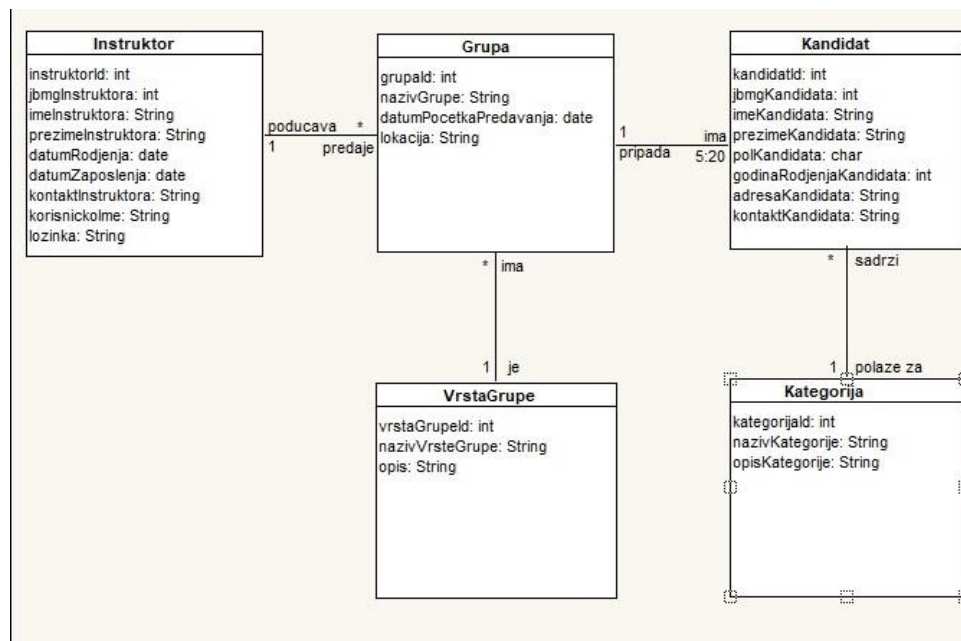
predmet_id	predmet_naziv
1	Programiranje 1
2	Programiranje 2

Ako pravimo strukturu baze podataka i posmatramo tabele kao entitete iz realnog sveta, u tabelama je potrebno praviti tabele na takav način da one sadrže najmanji mogući skup informacija i da se pomoću veza u sistemu (spoljnih ključeva) može doći do ostalih podataka.

Rezultat normalizacije će imati veći broj tabela u sistemu, sistem će postati modularan. Moći će da pamti podatke na struktuiraniji način koji će biti lakši za imenu, čuvanje, biće fleksibilniji.

Slede dva primera modela koja rešavaju problem sistema za evidenciju rada auto škole. Razlika između naredna 2 modela je ta što je drugi model otvoreniji da pamti više podataka kada govorimo o informacionom sistemu za vođenje autoškole. Oba sistema nisu pokrila sve procese koji postoje u realnom sistemu. Analizirajte oba sistema da uočite razlike i prednostni i mane oba modela.

Ovaj model je u 3NF, ali kako se zahtevi menjaju, tako se i sistemi šire, prosto rečeno potrebno je da pamtimo više podataka. Ako bismo nove podatke dodavali u postojeće tabele, napravili bismo šumu nelogički povezanih celina, a ovo bi nas puno koštalo prilikom održavanja. Ne bi se videla šuma od drveća.



Struktura koja predstavlja minimalne logičke celine uvezane na način da se moći do svih bitnih informacija za korisnike sistema, a sistem je dovoljno fleksibilan za korisničke zahteve i izmene.

