# Osnove baza podataka pitanja:

# 1. Istorijski razvoj automatizovane obrade podataka:

Nastankom ENIAC-a 1950ih nastaje termin automatizovana obrada podataka. Racunari su se u pocetku uglavnom koristili za resavanje nekih naucnih i racunskih problema nad relativnom malim skupom podataka, ali vremenom javlja se potreba za obradom velike kolicine podataka, kao sto je bila u finansijama npr. , i ubrzo krece koriscenje memorijskih jedinica i novih metoda pristupa za cuvanje ali i obradu podataka.

# 2. Klasicna obrada podataka(prednosti,mane,nastajanje modela podataka):

Klasicna obrada podataka podrazumeva cuvanje svih podataka u nekakvim fajlovima(txt npr).

Njena prednost jeste jednostavna realizacija i projektovanje za jednu aplikaciju, ali mane jesu to sto je vrlo komplikovano projketovati takav sistem za vise aplikacija koje bi koristile isti izvor pordatka(nepovezanost podataka),

to sto bi mi ako imamo vise aplikacija koje koriste iste podatke imali iste fajlove za svaku aplikaciju sto bi dovelo do postojanja vise datoteka sa istim podacima(redudanntnost podatka),

takodje mana je i cvrsta povezanost programa sa datotekom jer ako izmenimo fizicku ili logicku strukturu te datoteke moramo menjati logiku svakog programa koji je koristi.

Takodje vremenom kod slucaja da postoji vise programa moze se desiti da se dodje do neusaglasanosti izmedju datoteka gde za entitete sa istim obelezjima mozemo imati razlicite podatke.

Model podataka i strukutra datoteka nastaje iz zahteva tj iz nasih potreba nezavisno od samog programa na osnovu objekata i njhovih veza koje treba da podrzimo.

# 3. Koncepcija baza podataka(kada? Koji je cilj?):

Zbog samih problema koje je sadrzala klasicna obrada podataka krajem 50ih se odlucilo da bi sve podatke trebalo cuvati u nekakvoj bazi podatka. Sam koncept baze podataka ima za cilj da :

 Sve podatke jednog informacionog sistema treba integrisati u jednu fizicku strukutru podataka  Svi aplikacioni programi pri obradi baze podataka koriste standardne softverske alate nazvane sistemima za upravljanje bazom podataka(SUBP)

# 4. SUBP? (Sistem za upravljanje bazom podataka):

SUBP je program namenjen rokovodjenju bazom podataka odnosno skupom strukturiranih podatka kao i da izvrsava operacije nad njima. Obezbedjuju pretrazivanje,cuvanje,azuriranje, sigurnost podataka itd..Osnovni zahtev jeste ocuvanje integriteta podatka tj cuvanje tacnosti i koinzistetnosti istih.Neke od poznatijih su Oracle,MS SQL Server,Postgres,MySql

# 5. Poslovni zadaci Administratora baze podatka:

Poslove kao sto su regulisanje prava pristupa, pravljenje sigurnosnih kopija, koinzistentno sprovodjenje promena idt. Ne mozemo poveriti obicnom korisniku. Zbog toga se javlja pojam Administratora baze podataka. Administrator baze podataka je osoba odgovorna za specificiranje, projektovanje, implementiranje, efikasan rad i održavanje baze podataka. Administrator baze podataka radi na sledećim poslovima:

- Radi sa korisnicima pri uspostavljanju zahteva za bazu podataka
- Definise bazu podataka
- Radi sa programerima koji treba da pristupaju bazi podataka
- Odgovoran je za punjenje baze podacima
- Nadgleda rad baze podataka kako bi ustanovio da li je potrebna neka izmena ili drugacija organizacija podataka

# 6. Integritet baze podataka:

Integritet baze podataka je stanje baze podataka u kome su sve vrednosti podataka korektne tako da oslikavaju stanje realnog sveta u zadatoj preciznosti i tacnosti kao i da svi podaci postuju uzajamna pravila konzistentnosti tj. medjusobne odnose. Za odrzavanje integirteta BP zaduzen je administrator uz pomoc SUBP.

#### 7. Bezbednost BP:

Bezbednost baze podataka podrazumeva zastitu i specavanje od neovlascenog pristupa podacima kao i nedozvoljenog menjanja istih.Ona je potreba zbog zastite od nenamernih ali i namernih pristupa poverljivim podacima. Administratori moraju voditi racuna o pravima pristupa jer moze doci do konflikta pri pristupanju bazi imajuci na umu da vise razlicitih aplikacija moze da ima zajdnicku bazu podataka.

# 8. Principi rada sistema za upravljanje bazom podataka:

Softver za upravljanje bazom podataka je veoma slozen i sadrzi veliki broj programa i rutina. Poseduje komponente i mehanizme za

- upravljanje transakcijama
- upravljanje višekorisničkim režimom rada
- zaštitu BP od neovlašćenog pristupa
- zaštitu BP od uništenja / oštećenja
- implementaciju šeme BP u izabranom modelu podataka
- formiranje, korišćenje i ažuriranje BP
- upravljanje podacima

#### On se sadrzi od:

- Fizickog nivoa definise nacin na koji su podaci organizovani na spoljnim memorijama
- Konceptualnog nivoa (sema baze podataka)- definise logicku sturkturu baze podataka,sve podatke u sistemu,njihove odnose i veze.. Treba da omoguci upravljanje podacima kao zajednickim resursima u celom sistemu.
- Korisnickog nivoa definise logicku strukturu podataka pogodnu za programe.

# 9. Modeli podataka(Redudansa?):

Model podataka je skup medjusobno povezanih podataka koji opisuju objekte, njihove veze i osobine realnog sistema. U modelu podataka ne opisuje se potpuni skup znanja o sistemu, već se vrši odabir i opis relevantnih karakteristika sistema.

Redudansa predstavlja ponavljanje istih podataka u razlicitim datotekama.

# 10. Apstrakcije podataka:

Apstrakcija predstavlja kontrolisano dodavanje detalja u sistem tj. Skrivanje nepotrebnih detalja i ukljucivanje samo onih neophodnih. Pri izradi modela podataka koristimo sledece tipove apstrakcija:

- Klasifikacija objekta i uzrokovanje- Objekti koji imaju isti skup osobina (atributa i veza) i isto ponašanje mogu se predstaviti tipom ili klasom. Primer: STUDENT za svakog studenta na fakultetu.Uzrokovanje predstavlja prkazivanje jednog pojavljivanja(uzroka) datog tipa ili klase.
- Generalizacija i specijalizacija objekata apstrakcija u kojoj se skup sličnih tipova objekata predstalja opštijim generičkim tipom odnosno nadtipom. Slični su oni tipovi objekata koji imaju neke zajedničke osobine i ponašanje. Primer: STUDENT,PROFESOR,CISTACICA mogu se predstaviti kao GRADJANIN npr. Obrnut postupak generalizacije je specijalizacija.
- Agregacija i dekompozicija apstrakcija kojoj se skup objekata i
  njihovih međusobnih veza tretira kao novi, jedinstveni, agregirani
  tip koji je na visem nivou apstrakcije. Inverzan postupak agregaciji
  jeste dekompozicija.

# 11. Komponente modela podataka:

Model podataka sadrzi tri komponente:

- **Strukturna komponenta** cini je skup primitivnih koncepata i skup pravila za izgradnju slozenih koncepata.
- Integritetnu kompomnentu cini je skup uslova integriteta tj ogranicenja. Uslovi integriteta oganicavaju pojave baze podataka na samo one koje zadovoljavaju neko od ogranicenja kao sto je dozovoljena vrednost podataka,dozvoljenih veza izmedju objekata i sl.
- Operacijsku komponentu cini je skup koncepata koji omogucavaju interpretaciju dimaickih karakteristika skupa podataka i promenu stanja podataka u bazi podataka u skladu sa promenom stanja u relanom sistemu. Primer moze biti upis godine jednog studenta koji zelimo da se izmeni kao sto se izmenilo i u realnom sistemu.

# 12. Generacije modela podataka:

Istorisjiski gledano postoje 3 generacije modela podataka:

- Modeli podataka prve generacije Svaki konvecnioni programski
  jezik koji ima zaseban model podataka. U nekim jezicima je
  moguce delimicno primeniti apstrakciju generalizacije kao i dodati
  ogranicenja nad vrednostima odrejenih tipova eksplicitno. Samim
  tim modeli podataka prve generacije i njihovi programski jezici
  nisu dovoljno pogodni za modelovanje realnog sistema, te im je
  prakticna primena jako ogranicena.
- Modeli podataka druge generacije-Za prezentaciju podataka sadrze koncepte kao sto su: stablo,set,relacija i dr. U osnovi koriste iste apstrakcije kao modeli prve generacija pa ne postoji poptuna podrska za koncepte agregacije i generalizacije. Ovde je moguce eksplicitno definisati ogranicenja na vrsti podatka. Uprkos tome sto jesu semanticki bogatiji od modela prve genracije i dalje ne mogu dati zadovoljavajuci prikaz realnog sistema. Ovoj generaciji pripadaju: funkcionali model podataka, herarhijski, mrezni, relacioni ...
- Modeli podataka trece generacije Sadrze koncepte generalizacije i agregacije. Podrzavaju sve vrste apstrakcija, poseduju mogucnost eksplicitnog definisanja ogranicenja na vrednosti podataka. Dobro modeliraju realni sistem. Korisniku su razumljivi. Ovoj generaciji pripadaju: model objektiveze (MOV), semanticki model podataka, prosireni relacioni model...

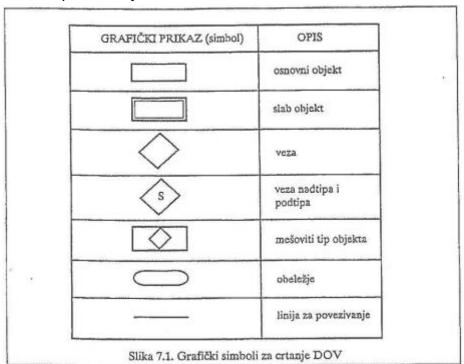
# 13. Model objekti veze(ER-model):

Najpopularniji i u praksi najčešće korišćeni. Koristi se kao grafički jezik za projektovanje konceptualne šeme baze podataka.

## • Konceprualna sema:

Konceptualna sema predstavlja takav model realnog sistema i njegove baze podataka koji ne zavisi od konkretnog SUBP. Može se lako prevesti u šemu baze podataka na kojoj je SUBP zasnovan. Konceptualna šema MOV predstavlja se uz pomoć dijagrama Objekti-Veze (DOV):

- -Tip objekta se predstavlja pravougaonikom sa upisanim nazivom.
- -Atributi elipsom sa upisanim nazivom povezani sa odgovarajućim tipom objekta ili tipom veze.
- -Atributi kojima se identifikuje objekat se podvlače.
- -Broj tipova objekata koji učestvuje u vezi definiše red veze
- -Veza se predstavlja rombom.



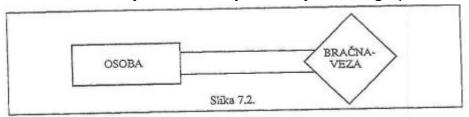
#### Strukturna komponenta:

Strukturnom komponentom definise se skup primitivnih koncepata i skup pravila za izgradnju slozenih koncepata. Primitivni koncepti u MOV su objekat, veza(izmedju objekata) i oblezje objekta.Koristeci apstrakciju klasifikacije pojedini objekti se mogu klasificirati u tipove objekata.Objekti se koji se indetifikuju nezavisno od ostalih objekata u sistemu se nazivaju osnovni tj jaki objekti i oni se predstavljaju pravougaonikom sa nazivom unutar njega.

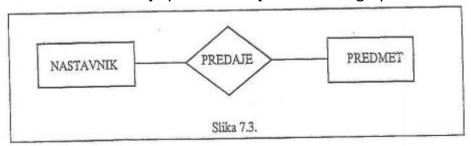
#### • Koncept veze:

Veza u modelu opsije nacin povezivanja objekata. Broj tipova koji ucestvuju u vezi definise red veze. Mozemo razlikovati rekurzivne, binarne i veze viseg reda.

Rekurzivna veza - je veza izmedju dva objekta istog tipa.



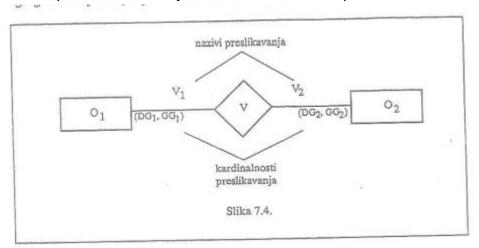
**Binarna veza** – cine je povezani objekti razlicitog tipa.



**Veze viseg reda -** mogu se predstaviti pomocu binarnih veza ili se takva veza tretira kao agregacija.

#### • Kardinalnost:

Informacije o prirodi odnosa izmedju objekata iz realnog sveta u modelu podataka iskazuje se kardinalitetom tipa veze.



Veza V izmedju pojava tipova objekata O1 i O2 definiše dva tipa preslikavanja:

V<sub>1</sub>:O<sub>1</sub> → P(O<sub>2</sub>) preslikavanje sa skupa pojavljivanja O<sub>1</sub> u skup pojavljivanja O<sub>2</sub> (O<sub>1</sub> je domen, a O<sub>2</sub> kodomen preslikavanja),

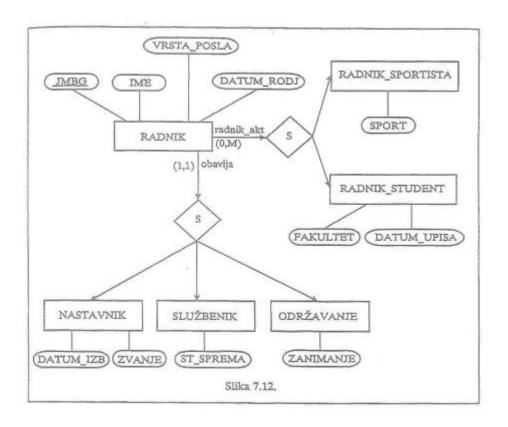
V<sub>2</sub>:O<sub>2</sub> → P(O<sub>1</sub>) (inverzno) preslikavanje sa skupa pojavljivanja O<sub>2</sub> u skup pojavljivanja O<sub>1</sub> (O<sub>2</sub> je domen, a O<sub>1</sub> kodomen preslikavanja),

gde su: V<sub>1</sub> i V<sub>2</sub> - nazivi preslikavanja,
P(O) - partitivni skup skupa O.

Za svaki od preslikavanja definise se par (DG,GG) gde DG(Donja granica) predstavlja najmanji a GG(Gornja granica) najveci broj elementa skupa u koji se preslikava jedan element skupa originala. Par (DG,GG) predstavlja kardialnost preslikavanja. Ako donja granica ima vrednost nula onda je to parcijalna(tj opciona)veza, dok ako je taj broj vce od nule onda je ta veza totalna(obavezna). Gornja granica moze ima bilo koju vrednost vecu od jedan koja nije poznata i to se oznacava sa M. Isto tako mora se ispostovati da DG <= GG.

# Nadtip – podtip:

Kao sto vec znamo generalizacija ja apstrakcija u kojoj se skup slcinih tipova objekata tretira kao jedan genericki tip objekata tj. Nadtip.Specijalizacija predstavlja inverzan postupak u kome se za neki tip objekta definisu njegovi podtipovi koji imaju neka samo njima specificna obelezja, veze ili operacije. Podtipovima se onda pripisuju sva obelezja koja ima i njhov nadtip a moze se dodati i ona njima specificna. Mozemo reci da je neki objekat podtip drugog objekta ako je svaka njegova pojava ujedno i pojava tog drugog objekta.



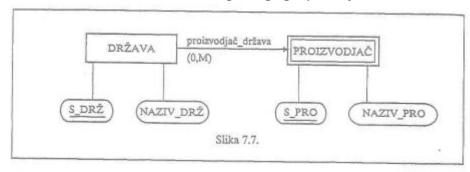
Takodje s obzirom na DG i GG mogu se definisati dve vrste specijalizacije:

onda kada svako pojavljivanje tipa spezijalziuje se u tacno jedan podtip(DG=GG=1)

i kada jedno pojavljivanje nadtipa moze uzrokovati vise pojavljivanja podtipa(DG=0, GG>1 tj. GG= M).

# Slabi objekat

Tip slabog objekta je takav tip objekta koji nema vlastitih obelezja za indetifikaciju, vec koristi obelezja nekog drugog tipa objekta koji moze biti jak(osnovni) ili slab tip. Kaze se da tip slabog objekta indefitikaciono zavisi od nekog drugog tipa objekta

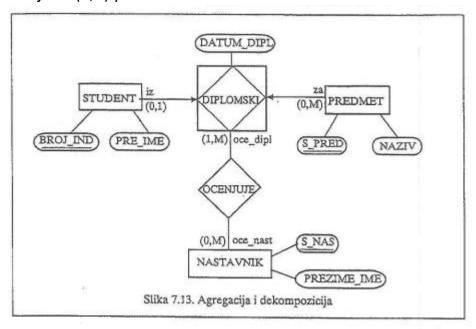


Oznacavaju se upisivanjem njegovog naziva u dvostruki pravougaonik, i samim tim da slab obejakt ne moze postojati bez njegovog nadredjenog objekta podrazumeva se da ce donja granica preslikavanja biti 1. To znaci da jedna pojva tipa slabog objekta u vezi je samo sa jednim tipom nadredjenog objekta. Ako se u bazi podataka brise pojava nekog objekta tada se brisu i sva pojavljivanja njemu indetifikacioni zavisnih slabih objekata.

## • Agregacija:

Na MOV vezu je moguce uspostaviti samo izmedju dva tipa objekta. Kada bi nam bilo potrebno da tip veze ima neka obelazja ili da se uspostavi veza između tipa veze i tipa objekta tada tip veze postaje mesoviti tip objekt-veza. Mesoviti tip obejkta moze imati i obelezja i povezati se sa drugim tipovima objekta. Prevodjenje tipa veze u tip obejkta izvodimo apstrakcijom agregacije u kojoj se veza izmedju dva ili vise tipova objekta tretira kao objekat na visem nivou apstrakcije. Agregatni tip objekta se razlikuje od ostalih objekata po tome sto nema svoj indetifikator vec ga indetifikuju objekti koje on agregira. Suprotan proces

agregaciji bio bi dekompozicija i njeno preslikavanje je uvek trivijalno(1,1) pa se na modelu i ne navodi.



#### 14. Proces izrade MOV:

Izrada MOV za dati realni sistem se obicno odvija u sledecim koriacima:

- -Izrada MOV po delovima
- -Integracija delova u jednu celinu koja se naziva globalnim modelom podataka
- -Uklkucivanje drugih semantickih detalja u model kao sto su uslovi integireta

Proces izrade MOV je iterativan tj. Svaka naknadna saznanja i zahtevi cesto mogu dovseti do izmene postojeceg resenja. Takodje prilikom izrade postoje razni obrasci za koje se preporucuje da ih koristimo.

#### Integiranje podmodela:

Kada smo vec zavrsili izradu po delovima tj. Podmodelima pristupa se integrisanju tih podemodela u jedan globalni model podataka. U tom procesu cesto moze doci do novih saznanja pa se vrlo cesto javlja potreba za izmenom u pojednim podmodelima.Pre same integracije treba proveriti da li se u pojedinacnim modelima javljaju identicni objekti ili neki sinonimi ili hononimi.Sve iste ili slicne tipove treba proedstaviti jednim globalnim modelom tipa kao sto i sve sinonime treba zameniti sa jednom recju a hononime promeniti tako da svaki tip ima novi jedinstven naziv. Tek nakon ove faze mozemo preci na pravu implementaciju i koriscenje svih koncepata apstrakcije.

## Golobalni model podataka treba da bude:

Globalni model podataka mora biti:

- -kompletan(sadrzi sva znanja iz svih podmodela)
- -neredundatan(ne sadrzi visestruke predstave istih znacenja)
- -koinzistentan(ne sadrzi modjusobno protivrecna znacenja tj da na razlicitim podemodelima imamo medjusobno protivrecne tvrdnje o nasem sistemu)

# 15. Relacioni model podataka:

Koncept realcionog modela podataka je prvi put objavljenja u clanku Edgara Codd-a iz 1970 godine. Nastao je sa ciljem da se rese problemi koji su imali tada poznati modeli podataka. Zasnva se na relacionoj algebri. Osnovni pojam realcionog modela je realcija. Relacija se moze posmatrati iz dva aspekta <u>znacenje</u> i <u>sadrzaj</u>. Znacenje realcije se naziva intezijom i formalno se iskazuje semom relazije. Sadrzaj relacije naziva se ekstenzija i iskazuje se tabelom podataka ili pojavom seme relacije.

## Relaciona sema(Sema relacije):

Sema realcije tj. Zaglavlje tabele nam daje apstraktnu ideju o tome sta moze biti u nasoj tabeli.Sema realcije definise samu strukturu baze.Ona se definise nazivom,skupom obelezja i ogranicenjima.

#### Relacija:

Prema matematickom shvatanju relacija nad neka dva skupa je podskup njihovog dekartovog proizvoda. Relacija zapravo predstavlja citavu strukturu(tabelu) u kojoj se podaci predstavljaju i organizovana je u redove i kolone. Svaki red naziva se n-torka a svaka kolona n-torke atribut. Ona nije uredjena tako da redosled n-torki nije bitan.

# Koje uslove treba da ispuni skup obelezja da bi bio koandidat za kljuc seme relacije:

S obzirom da je svaka n-torka relacije jedinstvena to znaci da ce postojati jedinstveni indetifkator istih. Taj indetifikator se naziva kandidatom kljuca. Da bi jedan skup obelezja bio kandidat mora ispuniti sledece uslova:

Jedinstvenost - Ni u jednoj ekstenziji relacije ne postoje n-torke koje imaju jednake vrednosti svih njegovih obelezja.

Minimalnost – broj obelezja u kljucu bi trebalo svesti na miminum.

Jedan od kandidata za kljuce se bira za primarni, a oni koji ispune uslov jedinstvenosti ali ne i uslov minimalnosti se nazivaju super kljuc(npr. jmbg i ime).

### Spoljni kljuc:

Spoljni kljuc su ona obelezja ili skup obelezja koji nisu primarni kljucevi date tabele vec se nalaze u nekoj drugoj semi relacije.(Neke medjutable npr)

# 16. Integritetna komponenta relacionog modela podataka: Problem nedostajucih informacija:

Model baze podataka polazi od predpostavke da su sve relevante informacije poznate i da se mogu uneti u bazu podataka. U praksi ova pretpostavka cseto nije ispunjena. Taj problem se naziva problem nedostajucih informacija. Ovaj problem se pojavljuje kada npr. u trenutku unosa podataka u bazu nisu poznati neki podaci (adresa boravka, broj telefona i sl.) Nekada ce biti moguce te informacije dobiti naknadno ali nekada to nije moguce kao npr naci ime supruga osobi koja nije u braku.

#### **Null vrednosti:**

Na mesto koje treba upisati stvarnu vrednost obelezja koje nam nije poznato mi upisujemo null vrednost. Null nam oznacava da ta vrednost nije poznata. To moze komplikovati operacijsku komponentu pa se gleda da se izbegnu kada god je to moguce.

#### Integritet baze podataka(kako se obezbedjuje,podela):

Pod integritetom baze podataka podrazumeva se istinitost i ispravnost podataka u njoj.On se definise uslovima ili pravilima integrita nad sadrzajem baze podataka organicavajuci je na neka dozvoljena stanja.Uslove integriteta mozemo podeliti u dve grupe:

Opsti uslovi integriteta – vaze u svim relacionim bazama podataka.

Korisnicka pravila integriteta – specificna za pojedine aplikativne sisteme,tj u oblasti primene.

#### Od cega zavise korisnicki uslovi integriteta:

Oni zavise od korisnika, specifikacije. mogucnosti...

#### Opsti uslovi integriteta koji vaze u svim realcionim bazma:

Pravila integriteta mogu biti staticka i dinamicka. Statickim pravilima definisemo uslove koji moraju vaziti pre i posle izvrsavanja bilo koje operacije nad bazom, dok dinamickim pravilima definisane su procedure u relacionom modelu koje ce nam garantovati ostvarenje uslova

integriteta. Relacioni model podataka postuje sledeca staticka praviala integriteta:

- Integiretet objekta
- Referencijalni integritet.

# Sta je i kako glasi inetgritet objekta:

Integritet objekta vezan je za sam pojam primarnog kljuca seme relacije . S obzirom na osobine primarnog kljuca sledi definicija da vrednost primarnog kljuca kao celine kao i svaka njegova komponenta ne sme biti jednaka null vrednosti. To se ne sme desiti jer nam primarni kljuc omogucava lako i efikasno adresiranje n-torke pa ako se tu nadje null vrednost ne bi smo mogli to ostvariti.

# Sta je i kako glasi referencijalni uslov integriteta:

Referencijalni integritet predstavlja poseban slučaj opštijeg uslova integriteta, koji se naziva zavisnost sadržavanja. Zavisnost sadrzavanja definise ograncenje u smislu da se u nekoj relaciji ne moze upisati pozvivajuca n-torka ako u drugoj relaciji ne postoji bar jedna ciljana n-torka. Pri brisanju ciljane n-torke moraju se brisati i sve pozivajuce n-torke. Na primer ako imamo tabelu nastavnik koja bi pozivala katedru na kojoj predaje preko sifre katedre ako bi obrisali katedru onda nam sifra kat iz nastanika ne pokazuje ninasta i narusava referencijalni integritet. Taj problem nam resava SUBP koji ce nam prijaviti gresku u tom slucaju.

# 17. Operacijska komponenta realcionog modela:

#### Na cemu se zasniva:

Operacijska komponenta omogucava manipulisanje podacima u bazi podataka sto se pre svega odnosi na psotavljanje upita i azuriranje baze podataka. Smisao operacija nad semom baze podataka jeste u izmeni postojece seme relacije kao i u dodavanju novih realcija. Svi postojeci jezici relacione baze podataka kao sto je SQL su razvijeni na osnovu:

- -relacione algebre
- -relacionog racuna

#### Definicija operacija relacione algebre:

Operacije relacione algebre mogu se podeliti u dve grupe.

Prvu grupu čine operacije matematičke teorije skupova:

- -UNIJA,
- -RAZLIKA.

- -PRESEK,
- -DEKARTOV PROIZVOD,

i u relacionom modelu podataka se mogu primeniti pošto je relacija definisana kao skup n-torki.

Drugu grupu operacija čine operacije koje su posebno razvijene za relacioni model podataka:

- -SELEKCIJA,
- -PROJEKCIJA
- -SPOJ.

Rezultat svake operacije nad nekom relacijom ili skupom relacija je nova relacija nad kojom se dalje mogu primenjivati operacije relacione algebre.

### Binanrne operacije(UNIJA, PRESEK I RAZLIKA):

Mogu se izvoditi samo na relacijama koje su medjusobno uporedive tj kompatibiline.Relacije su kompatibilne ako imaju isti broj oblezja i imaju domen iste vrste. Medju tim operacijama spdaju:

## Operacija unije:

Unija izmedju dve relacije je skup n-torki koji su sadrzani u jednoj od relacija ili i obe relacije  $r \cup p = \{t \mid t \in r \lor t \in p\}.$ 

Ona ima osobine komutativnosti i asocijativnosti(AuB = BuA, (AuB)uC = Au(BuC))

#### Operacija preseka:

Presek dve kompatibilne relacije je realcija koja sadrzi sve n-torke koje su istovremeno ementi i prve i druge relacije.  $r \cap p = \{t \mid t \in r \land t \in p\}$ .

```
r(ABC) p(ABC) rop=r(ABC)
abc abc abc
defabf efg
efg
abg
```

#### Operacija razlike:

Razlika izmedju dve kompatibilne ralcije je skup svih n-torki sadrzanih u prvoj koji nisu sadrzani u onoj drugoj.  $r-p = \{t \mid t \in r \land t \notin p\}.$ 

# Operacija dekartov proizvod:

Dekartov proizvod dve realcije predstavlja skup n-torki koje su nastale kao rezlutat spajanja svake pojedine n-torke sadrzane u relaciji jedan sa svakom pojedinom n-torkom iz relacije dva.  $rxp = \{(t_n, t_n) | t_n \in P\}$ .

Relaciona sema dobijena dekartovim prozivodom za obelezja ima uniju skupa obelezja obe relacione seme. Ako je presek skupova obelezja prazan skup tada dekartov prozivod jeste relacija a u suprotno to ne bi bilo tacno jer bi se pojavljivala ista obelezja sto nije dozvoljeno. Za razliku od dekartovog proizvoda nad skupovima dekartov proizvod nad relacijama je komutativan jer se redosled n-torki moze menjati.

a) Neka vredi  $R \cap P = \emptyset$ . Dekartov proizvod  $r \times p$  je relacija.

b) Neka vredi R∩P ≠ Ø. Dekartov proizvod r x p nije relacija.

# Operacija selekcija:

Selekcija je unarna operacija kojom se iz relacije izvdvaja odredjeni podskup n-torki koji ce zadovoljavati odredjeni uslov koji smo mi definisali.  $\sigma_{\mathbf{r}}(\mathbf{r}) = \{t \mid t \in \mathbf{r} \land t \text{ zadovoljava } \mathbf{r}\}.$ 

Kriterijum za selekciju moze sadrzati: konstante,nazive obelezja relacije nad kojom se selekcija izvodi, operatore poredjenja(= ,  $\neq$  , < ,  $\leq$  , > ,  $\geq$  ),logicke operatore i slcino.

# Operacija projekcija:

Projekcija predstavlja unarnu operaciju u kojoj se iz date relacije izdvajaju pojedine kolone:  $\pi_{x}(r) = \{t[X] \mid t \in r\}.$ 

$$r (A B C) (A B)$$
  $\pi_{AB}(r) = r'(A B)$   
 $a b c$   $a b$   $a b$   
 $d e f$   $d e$   
 $a b f$   $a b$ 

## Operacija spoj:

Operacija spoj je slozena binarna operacija za koju se moze reci da se izvrsava u tri koraka:

- 1. Formiranje dekartovog proizvoda relacija.
- Iz dekartovog proizvoda se izdvajaju n-torke koje zadovoljavaju postavljene uslove
- 3. Iz tabele dobijene u drugom koraku izdvajaju se odredjene kolone(samo u slucaju prirodnog spoja)

r (A B) p (C D)

a 2 4 d

e 5 1 f

f 2

1. Korak: 
$$(rxp) = r^2$$
 (A B C D)

a 2 4 d

a 2 1 f

e 5 4 d

e 5 1 f

f 2 4 d

f 2 1 f

2. Korak:  $\sigma_{8 < c}(r^i) = r^n$  (A B C D)

a 2 4 d

f 2 1 f

Thetha spoj( $\theta$  - spoj) predstavlja najopstiji oblik operacije spoja Teta spoj r[Ai  $\theta$  Bi] je skup n-torki koji zadovoljava sledece uslove:

- -podskup je dekartovog proizvoda tih relacije
- -svaki element tog podskupa zadovoljava uslov izdvajanja

$$r[A_i \Theta B_i]p = \{(t_r, t_p) | t_r \in r \land t_p \in p \land t_r[A_i] \Theta t_p[B_i\}.$$

Theta spoj ce biti relacija ukoliko je i njegov dekartov spoj relacija. U tom slucaju spoj se moze definistai na osnovu dekartovog proizovda i selekcije  $r[A_i \ominus B_j] p = \sigma_{A_i \ominus B_i} (r \times p)$ .

Postoje dva slucaja teta spoja: kada uslov izdvajanja ne postoji(tada operacija prelazi u dekartov prozivod), kada je operator znak jednakosti pa se to naziva spoj sa izjednacavanjem(equi-join).

$$r(A B)$$
  $p(C D)$   $r[B = C]p = q(A B C D)$   
a 6 1 d a 6 6 f  
b 1 6 f b 1 1 d  
e 3 e 1

# Operacija prirodnog spoja:

Prirodni spoj je u tesnoj vezi sa teta spojem. Kao rezultat dobije se tabela sa najmenje dve kolone indeticnih vrednosti obelezja, a cesto su i nazivi identicni. Da bi smo tu tabelu transofrmisali u realciju moramo odstraniti kopije obelezja zajdeno sa odgovarajucim zaglavljima tabele. Ova dodatna operacija se izvodi u okviru operacije prirodni spoj. Prirodnim spojem dve relacije spajaju se međusobno n-torke tih relacija na osnovu vrednosti obeležja koja se nalaze u obe šeme relacija. Operaciju prirodnog spoja mozemo izraziti pomocu jednostavnih operacija: Dekartovog prozivoda, selekcije i projekcije.

Primer: Neka su 
$$r(R)$$
 i  $p(P)$  relacije i neka vredi  $R = \{A, B, C\}$ ,  $P = \{B, C, D\}$  i  $T = \{A, B, C, D\}$ .

$$r(A B C) \quad p(B C D) \quad r \nabla p = q(A B C D)$$
a 2 4 2 7 d a 2 4 k
b 1 5 1 5 h b 1 5 h
c 2 7 2 4 k c 2 7 d

Mogu postojati dva slucaja prirodnog spoja:

-Kada je presek obelezja prazan skup i to znaci da one nemaju zajednickih obelezja pa je prirodni spoj tada jedank dekartovom proizvodu

Kada su skupovi obelezja jednaki(obe relacije su zadate na istoj realcionoj semi) tada je rezultat prirodnog spoja jednak preseku relacija.

# **SQL(osobine):**

Sql nije samo upitni jezik vec predstavlja kompletan jezik podataka koji sadrzi jezike za definisanje podataka,azuriranje,kontrolu koinzistentnosti,konkuretan rad i jezik za odrzavanje recnika podataka. Ugradjen je u većinu komercijalno raspoloživih SUBP, od personalnih do server racunara.Osnovne karaktreisktike sql su:

- Jednostavnost i jednobraznost pri koriscenju(relacije se kreiraju jednom nardbom i odmah su spremene za koriscenje)
- Mogućnost interaktivnog i klasičnog programiranja (koristeći SQL dobijaju se odgovori na trenutno postavljene zahteve ili se SQL blokovi ugradjuju u neki viši programski jezik)
- Neproceduralnost(On je proceduralan u velikoj meri jer se njime definise STA se zeli a ne KAKO doci do rezultata)

Sql naredbe se mogu grupisati u vise grupa a mi ih mozemo podeliti na :

- Naredbe za definisanje podataka (Data Definition Statements DDL) :
   CREATE ,ALTER,DROP...
- -Naredbe za manipulisanje podacima (Data Manipulation Statements DML): INSERT,SELECT,UPDATE,DELETE..
- -Naredbe za kontrolne (upravljacke) funkcije (Data Control Statements):

GRANT(korisnicka prava),REVOKE(oduzma prava),COMMIT(prenos dejstva transakcije u bazu),ROLLBACK(ponistavanje transakcije)

# 18. Cilj relacione beze podataka:

Cilj relacione baze podataka:

- -Nezavisnosnost strukture podataka od programa koje ih koriste i obrnuto
- -Minimalna redudansa
- -Jednostavnije koriscenje i upravaljane podacima
- -Koriscenje baze podataka od strane veceg broja korisnika

# 19. Anomalije odrzavanja podataka

Da bi omogucili jednostavno koriscenje i menjanje podataka moramo spreciti anomalije odrzavanja podataka. U tu grupu spadaju:

- Anomalija dodavanja: Javlja se kada se inforamcije o svojstvima
  jednog objekta u bazi cuvaju kao deo opisa nekog drugog
  objekta.Npr. ako imamo podatke o katedri ili predmetu u samoj tabeli
  nastavnik to bi nam onemogucilo da te objekte dodamo osim ako ne
  postoji bar jedan nastavnik koji ce tu radili ili predavati taj predmet.
- Anomalija brisanja: to je inverzna anomalija dodavanja. Prime bi bio da ako u opisu nastavnika nalazimo svojstva predemtta na kojem predaje. Svakim brisanjem nastavnika brise se i ta kopija predmeta koja ce nam mozda nekada biti potrebna.
- Anomalija menjanja(Azuriranja): javalja se u slucaju kada promene podataka o jednom objektu treba izvrsiti na vise kopija objekata. Kao i sa proslim primerima ako postoji podatak o predmetu na kojem radi nastavnik u samom njegovom opisu onda pri izmeni tog predmeta

treba to izvrsiti na onoliko mesta na koliko profesora predaje tu.Ako se to ne izvrsi imamo vise tvrdnji koje nisu istinite sto predstavlja nekoinzistentnu bazu podataka.

# 20. Normalizacija:

Osnovni cilj relacionih baza podataka je da odgovarajuća baza podataka: Ne sadrzi redundansu, Da se moze jednostavno koristiti i menjati Formalne kriterijume prema kojima se utvrdjuje da li model podataka ispunjava prethodne zahteve daju normalne forme.Normalizacija je proces provere uslova normalnih formi i po potrebi svodjenje seme relacije na oblik koji zavoljava iste. Ocekuje se reverzabilnost tj da ne sme doci do gubitka informacija sadrzanih u polaznoj relaciji i moguca je rekonstrukcija iste.

# Metode normalizacije:

Postoje dve tehnike normalizacije: vertikalna normalizacija i horizontalana.

Vertikalna normalizacija je proces u kome se nenormlaizvoana sema relacije transformise u skup manjih normalizovanih relacija. iz relacione šeme se izdvajaju obeletja koja stoje u nedozvoljenim odnosima sa ostalim obelezjima u semi. Zasniva se na operacokama projekcije i skupa. Projekcijom relaciju vertikalno razbijamo na dve ili vise manjih relacija dok nam prirodni spoj sluzi da bi dokazali reverbilnost tj da bi rekonstuisali polaznu relaciju. Ima dve varijante: Normalizacija sintezom (ne izvodi se u koracima), Normalizacija dekompozicijom (izovdi se u koracima)

Horiznotalna normalizacija je slucaj kada se relacija rastavlja na podskupove n-torki, odnosno fragmente relacije koji ispunjavaju odredjene uslove. Zasniva se na operacija selekcije i unije ali sama tehnika je jos uvek u razvoju.

#### Funkcionala zavisnost obelezja:

Funkcionalna zavisnost obelezja je ogranicenje koje definise vezu izmedju obelezja gde jedno obelezje moze tacno da odredi drugo obelezje. (Sva obelezja moraju biti zavisna od tacno jednog obelezja tj. On funkcionalno odredjuje ostala obelezja).

## Tranzitivna zavisnost oebelzja:

To je ogranicenje gde se odredjuje funkcionalana zavisnost nekljucnih elemenata sa drugim ne kljucim elementima. Primer:

Razlog za postojanje opisanih anomalija u relaciji NFM je već navedena FZ NASTAVNIK → MESTO koju nazivamo i *tranzitivna zavisnost* jer proizilazi iz zavisnosti S\_NAS → FAKULTET, FAKULTET → MESTO.

### Parcijalana zavisnost obelezja:

Specijalan slucaj tranzitivne i vazi samo ako neklucna obelezja zavise od delova kljuca tj kada je primarni kljuc slozen kljuc.

# Normalizacija dekompozicijom:

Zapocinje od proizvoljne nenormalizovane relacione seme i izvodi se u koracima. Svakim korakom normalizacije sema se prevodi u visu normalnu formu tako sto polazni skup obelezja delimo u dva skupa obelezja i od svakog se formira posebna relaciona sema. Svaki od koraka mora biti reverzibilan.

#### Uslovi dekompozicije bez gubitka informacija:

Generalno gledano kod dekopozicije se po pravilu ne gube podaci vec ih po spajanju ima vise nego sto je bilo u polaznoj relaciji. Uslovi koje bi morali da ispunimo da dekompozicija bude bez gubitaka bi bili:

- Novonastala sema relacije mora biti podskup pocetne seme relacija i da unija tih relacija predstavlja skup obelezja pocetne seme.
- Da bi dekompozicija bila reverzivna presek projekcija dve manje seme relacija koje su nastale od pocetne ne sme biti prazan.
- Skup obelezja koja cine presek manjih relacija moraju biti kandidati za kluc u bar jednoj od relacija

# Vertikalna normalizacija dekompozicijom:

U kontekstu vertikalne normalziacije definisemo 6 normalnih formi:

- Prva normalna forma (1NF)
- Druga normalna forma (2NF)
- Treca normalna forma (3NF)
- Codd-ova normlana fomra(BCNF)
- Cetvrta normalna fomra
- Peta normalna forma

Kod postupka normalizacije semu prvo transformisemo u u 1NF pa zatim u 2NF,3NF i tako redom. Sto je redno broj forme veci to su i njeni uslovi stroziji i broj funkcionalnih zavisnosti se smanjuje.

#### **Uslovi ispunjenosti 1NF:**

Sema relacije je u prvoj normalnoj fazi ako i samo ako je domen svakog od njenih obelezja skup atomiranih vrednosti i sva nekljucna obelezja nisu funkcionalno zavisna od primarnog kljuca. S obzirom da u kontekstu modela relaciju definisemo kao neprazan skup dekartovog proizvoda atomiranih vrednosti sledi onda da je savaka sema u 1NF.

#### Uslovi ispunjenosti 2NF:

Relaciona sema se nalazi u 2NF ako je svako nekljucno obelezje potpuno funkcionalno zavisno od kluca relacije. To znaci da ne sme postojati parcijalna zavisnost i svi elementi moraju zavisiti od celog kljuca umesto da podatak mozemo definisati i samo delom kljuca.

#### **Uslovi ispunjenosti 3NF:**

Relaciona sema R nalazi se u 3NF ako je u 2NF i ako nijedno nekljucno obelezje u R nije tranzitivno zavisno od primarnog kljuca od R vec cu svi funkcionalno zavisni od njega.

# Postupak svodjenja seme relacije koja nije u 2NF u skup seme relacije koje su u 2NF:

Ako imamo tabelu kao u primeru dole koja ima identifikator i svaka kolona ima atomirane vrednosti tj. Svaka celija ima istu vrednost i svaka kolona ima isti tip vrednosti. Onda pokusavamo da je dekomponujemo na vise podtabela koje ce zadovoljiti uslov 2NF a to je da sva nekljucna obelezja moraju zavisiti od primarnog kljuca te tabele

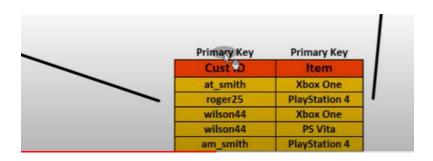
Cust ID	Cust Name	Item	Shipping Address	Newsletter	Supplier	Supplier Phone	Price
at_smith	Alan Smith	Xbox One	35 Palm St, Miami	Xbox News	Microsoft	(800) BUY-XBOX	250
roger25	Roger Banks	PlayStation 4	47 Campus Rd, Boston	PlayStation News	Sony	(800) BUY-SONY	300
wilson44	Evan Wilson	Xbox One	28 Rock Av, Denver	Xbox News	Microsoft	(800) BUY-XBOX	250
wilson44	Evan Wilson	PS Vita	28 Rock Av, Denver	PlayStation News	Sony	(800) BUY-SONY	200
am smith	Alan Smith	PlayStation 4	47 Campus Rd, Boston	PlayStation News	Sony	(800) BUY-SONY	300

## Nakon toga dobijamo dve tabele koje izgledaju ovako:

**Primary Key** 

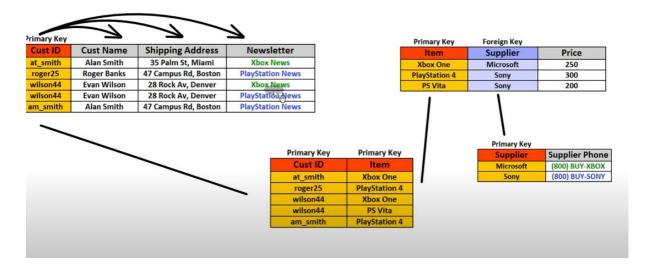
Cust ID	Cust Name	Shipping Address	Newsletter
at_smith	Alan Smith	35 Palm St, Miami	Xbox News
roger25	Roger Banks	47 Campus Rd, Boston	PlayStation News
wilson44	Evan Wilson	28 Rock Av, Denver	Xbox News
wilson44	Evan Wilson	28 Rock Av, Denver	PlayStation News
am_smith	smith Alan Smith 47 Campus Rd, Boston		PlayStation News

Item	Supplier	Supplier Phone	Price
Xbox One	Microsoft	(800) BUY-XBOX	250
PlayStation 4	Sony	(800) BUY-SONY	300
PS Vita	Sony	(800) BUY-SONY	200



Da bi proverili da li je doslo do gubitka mozemo uraditi uniju obelezja obe tabele i proveriti da li se poklapa sa pocetnom tabelom. Takodje mozemo da uradimo presek ili razliku kljucnih obelezja dve tabele i kao rezultat moramo dobiti kljucna obelezja one trece.

Zatim da bi ova tabela bila u 3NF moramo ukloniti tranizitivnu zavisnost tj moramo ukloniti sve nekljucana obelezja koja nisu funkcionalno zavisna od primarnog kljuca. Tada mozemo dobiti tabele kao sa slike:



# 21. Prevodjenje MOV u relacioni model podataka:

### Zasto se vrsi prevodjenje:

S obizrom da za sada ne postoji komerciajlno raspoloziv SUBP zasnovan na MOV to znaci da je praksi potrebno prevesti model radjen po MOV u neki drugi model podataka (najcesce relacioni).Kod prevodjenja se moze zapaziti da svakoj apstraktnoj operaciji bogatog modela (3 nivoa) odgovara procerdura semanticki manje bogatog modela (2 nivao relacioni).Ovaj postupak se moze automatizovati uz pomoc softverskih alataa za prevodjenje MOV-a u relacioni model.

# Opsta pravila:

Opšta pravila iz kojih se izvode sva pojedinačna su:

Obezbediti jedinstvenost ključa,

Izbeći višeznačna obeležja,

Izbegavati neprimenljiva svojstva,

Obezbediti minimalan skup šema relacija.

# Pojedinacna pravila za objekte:

Pravila za glase:

- Svaki objekat iz DOV postaje sema relacije.Ime tipa objekta postaje ime seme relacije.Obelezja tipa su obelezja sema relacija.Za onsnovne objekte indetifikator objekata postaje primarni kljuc seme relacije.
- Identifikator nadredjenog objekta postaje jedno od obelezja šeme relacije koja odgovara slabom objektu, i postaje deo kljuca šeme relacije slabog objekta, zajedno sa obeležjima koja jednoznačno identifikuju pojavljivanje "slabog" objekta u okviru pojavljivanja njemu nadredjenog objekta.
- Nadtip i podtip takodje postaju sema relacije.Ime seme relacije i obelezja jednaka su imenu objekta i obelezjima samog objekta.

#### Pojedinacna pravila za veze:

Pravila za pojedinacne binarne veze glase:

<u>Veza 1:1</u>(Obavezna sa obe strane) :

Sama veza i oba objekta koji u njoj ucestvuju postaju jedna šema relacije.

Obeležja seme relacije su svi atributi jednog i drugog objakta i atributi veze (odnosno agregacije).

Kandidati za kljuc u ovoj šemi relacije su identifikatori i jednog i drugog objekta koji su vezi.

# Veza 1:1(Sa jedne strane obavezna):

Dva objekta u vezi i sama veza daju dve seme relacija s tim sto se identifikator jednog od objekata koji su u vezi uvrsti u obelezja druge seme relacije .

Veza se predstavlja spoljnim kljucem u onoj šemi relacije u kojoj vrednost tog obeležja mora biti data.

# Veza 1:1(Opciona sa obe strane):

Kreiraju se tri šeme relacija.

Obeležja u semi relacije koja odgovara vezi su i identifikatori objekata koji su u vezi i oba su kandidati za kljuc.

Za primarni kljuc se bira jedan od indentifikatora objekata koji učestvuju u vezi

# - **Veza 1:M**(Sa jedne strane obavezna):

Ne postaje posebna sema relacije vec identifikator objekta za koji je GG=M postaje obeležje šeme relacije koja odgovara objektu sa strane za koju je GG=1.

Veza izmedju jakog i slabog objekta kao i veza izmedju podtipa i nadtipa ne postaje posebna sema relacije.

#### Veza 1:M(Sa jedne strane opciona):

Postaje posebna sema relacije

Obeležja te seme relacije su identifikatori objekata koji su u vezi, a kljuc seme relacije je identifikator objekta za koji je GG=1

Agregacija mesovit tip objekt veze se prevodi kao odgovarajuci tip veze, a obelezje agregacije ide u semu relacije kojom se resava veza.

# 22. Osnove obrade transakcije:

# Sta je transakcija:

Transakcija je vremenski uredjeni niz nedeljivih radnji nad bazom podataka koje u celini ne remete uslove integriteta. Predstavlju neku logicku jedinicu rada nad bazom podataka. Ona transformise bazu iz jednog koinzistentnog stanja u drugo koinzistentno stanje. Bitno je da transakcija bude tako oformljena da usled njenog prekida ne dodje do promene same baze podataka.

## Sql naredbe za rukavanje transakcijama:

Commit(siganalizira o uspesnom zavrsetku transakcije, gde se onda dejstva transakcije mogu sacuvati i baza ostaje u koinzistentnom stanju) Rollback(signalizira da je transakcija neuspesno izvrsena i u ovom slucaju treba ponisititi njeno dejstvo)

# Paralelno izvrsavanje transakcija:

Sistem za upravljanje bazom podataka treba da omoguci paralelan rad vecem broju korisnika nad istim podacima sto znaci paralelno izvrsavanje vise transakcija odjednom.Da bi to bilo moguce mora se obezbebditi neko adekvatno upravljanje izvrsavanjem transakcija kako one ne bi nepozeljno mesale svoja dejstva jedna na drugu.Zbog mesanja transakcija moze nastati pogresan rezulatat jer pri mesanju operacija iz dve korektne transakcije moze nastati nekorektan rezulata.

Ti ptoblemi mogu biti:

- Problem izgubljenog azuriranja
- Problem zavisnosti od privremenih azuriranja
- Narusavanja serijabilnosti

#### Problem izgubljenog azuriranja:

Ako bi uzeli u obzir dve transakcije A i B. Transakcija A cita zapis u jednom vremenskom trenutku dok B u drugom. A ce onda nakon toga u odredjenom trenuntku azurirati podatke na osnovu onih procitanih na pocetku dok ce be to isto uratiti ali u malo kasnijem vremenskom intervalu cime se delovanje transakcije A ponistava jer B u trenutku svog izvrsavanja nije znala za nju i njeno dejstvo.

#### Problem pri zavisnosti od privremenog azuriranja:

To je slucaj kada bi na primer kada bi transakciji A dozvolili citanje a jos gore i azuriranje zapisa koji su privremeno azurirani od transakcije B pri cemu se taj zapis jos nije trajno upisao u bazu. Ako u nekom momentu transakcija B ponisiti svoje privremena azuriranja dolazimo u situcaiju gde je transakcija A izvrsila operacija nad bazom sa podacima koji nikada nisu ni postojali u njoj.

# Osobine transakicja su:

Koinsitencija Izolacija Atomicnost Trajnost

## Protokoli zakljucavanja:

Uobicajeni nacin resavanja problema pralelnog izvrsavanja jesu protkoloi zakljucavanja. Osnovna ideja zakljucavanja jeste u tome da transakcija koja zahteva garancije da neki objeka za koji je zainteresovana ne moze biti azuriran u nekom vremenu. Takav objekat se zakljucava i sprecava se njegovo azuriranje od drugih transakcija. Na taj nacin transakcija raspolaze sa stabilnim podacima o objektu koliko to ona zahteva. Zakljucavanja moze biti direkntno (sami navodimo) i medjusobno.

# Restauracija konzistentnog stanja baze podataka:

Administrator transakcija u sebi ima implementiran protokol nedeljivog kompletiranja transakcije kojim se sva ažuriranja BP upisuju u log transakcije. Mehanizam upisa u log transakcije može biti:

- Upis novih vrednosti BP u log transakcije (after image)
- Upis starih vrednosti BP u log transakcije (before image) Svaka tačka sinhronizacije se unosi log transakcije. Logovi pojedinačnih transakcija se kombinuju u jedinstven sistemski log koji nazivamo dnevnik ažuriranja, u kojem su pokazivačima povezani zapisi koji pretstavljaju log jedne transakcije.

Dnevnik ažuriranja omogućuje oporavak BP u kombinaciji sa sigurnosnom kopijom BP.