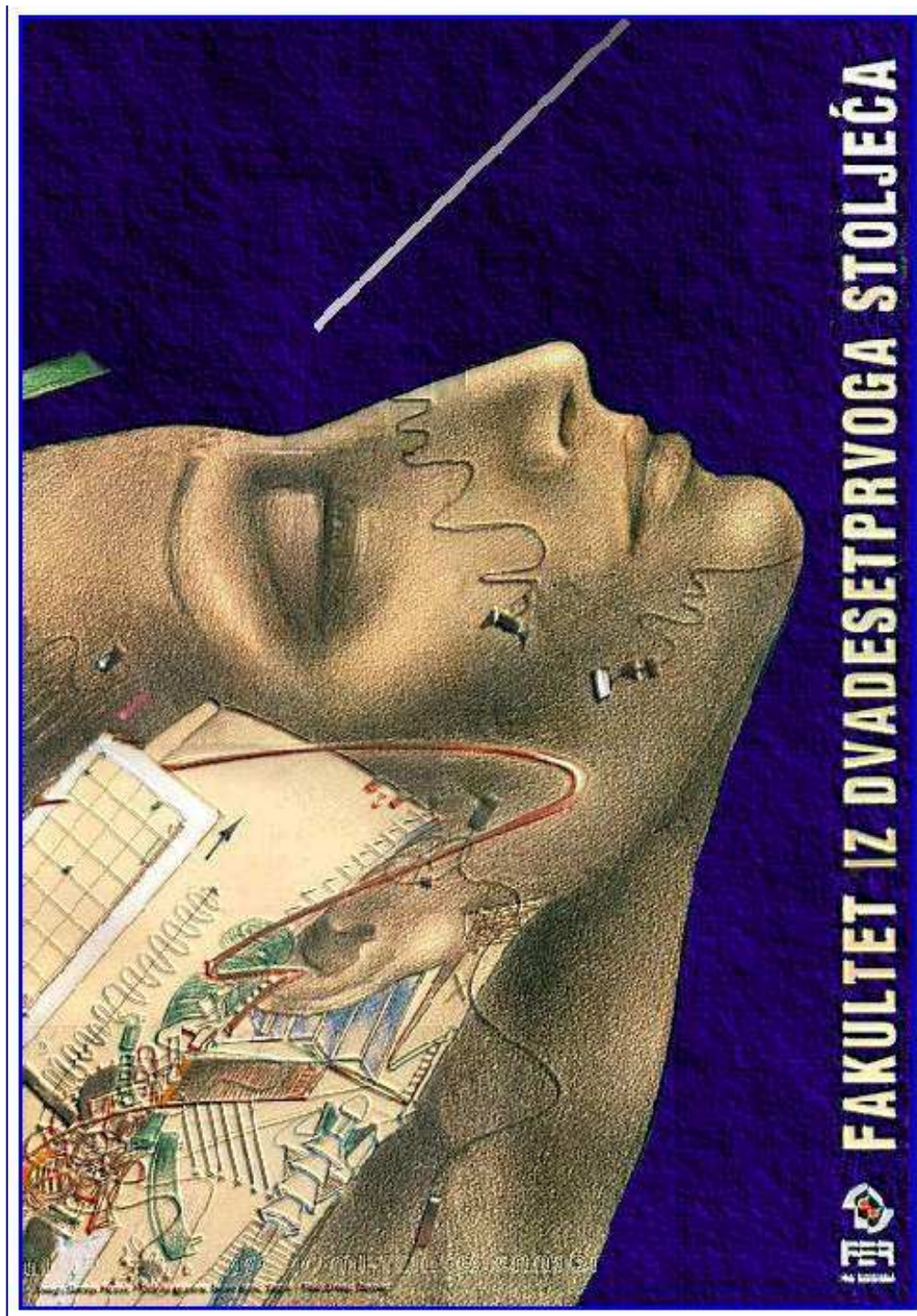


Napredni modeli i baze podataka

Predavanja

1. Uvod

Rujan, 2008.



Organizacija predmeta

Napredni modeli i baze podataka

Ciljevi predmeta

- Kolegij se nastavlja na kolegij Baze podataka s preddiplomskog studija
- Cilj mu je upoznati studente s najnovijim tehnologijama u području baza podataka i njihovim primjenama, kao što su
 - Objektne baze podataka,
 - Objektno-relacijske baze podataka,
 - XML baze podataka.
- Kolegij je usmjeren prema naprednim aplikacijama temeljenim na
 - Polustrukturiranim podacima,
 - Multimedijskim podacima,
 - Prostorno-vremenskim podacima,
 - Strukturama podataka koje se koriste u području bioinformatike, medicine i genetike.

Očekivani ishod učenja

- Studenti će biti sposobni oblikovati i implementirati baze podataka koje se temelje na objektnom, objektno-relacijskom, polustrukturiranom i prostorno-vremenskom modelu podataka.
- Stečeno znanje primijenit će se kroz primjere i različite manje aplikacije.

Potrebna predznanja

- Studenti trebaju biti upoznati sa:
 - Sustavima za upravljanje bazama podataka,
 - Relacijskim modelom,
 - Relacijskim bazama podataka.
- Trebali su savladati:
 - način oblikovanja relacijskih baza podataka,
 - oblikovanje modela entiteti-veze,
 - relacijska algebra, upitni jezik SQL,
 - osnove zaštite baza podataka.
 - modeliranje jednostavnijih baza podataka
 - postavljanje srednje složenih upita nad bazom podataka.

Predavači i asistenti

Prof.dr.sc. Mirta Baranović – nositeljica predmeta

Prof.dr.sc. Zdravko Galić (Geoprostorne baze podataka)

Doc.dr.sc. Strahil Ristov (Bioinformatika)

Mr.sc. Ljiljana Brkić

Mr.sc. Krešimir Križanović

Mr.sc. Igor Mekterović

Lidia Rován, dipl. ing.

Literatura

- H. Garcia-Molina, J. D. Ullman, J. Widom: **Database Systems: The Complete Book**, Prentice-Hall, 2005.
- A. Silberschatz, H.F. Korth, S. Sudarshan: **Database Systems Concepts**, 5th Edition, McGraw-Hill, 2005.
- M. Piattini and O. Diaz: **Advanced Database Technology and Design**, Artech House, 2000.
- Z. Galić: **Geoprostorne baze podataka**, Golden marketing, 2006.
- A. B. Chaudhri, A. Rashid, R. Zicari **XML Data Management**, Addison-Wesley, 2003
- R. Ramakrishnan, J. Gehrke: **Database Management Systems**, McGraw-Hill, 2003
- T. Sellis, A. U. Frank, S. Grumbach, R. H. Guting, M. Koubarakis **Spatio-Temporal Databases**, Springer, 2003

Upute i obavijesti

- na URL stranici predmeta

<http://www.fer.hr/predmet/nmbp>

- forum na stranici predmeta služi isključivo za međusobnu komunikaciju studenata

Organizacija nastave

PREDAVANJA

- *Powerpoint* prezentacije
 - datoteke u PDF formatu nalazit će se u repozitoriju datoteka na URL stranici predmeta
- Odabrani primjeri prikazivat će se pomoću prikladnih programskih alata

Organizacija nastave

SAMOSTALAN RAD

- Učenje -> Istraživanje (knjige, članci, Internet, slajdovi s predavanja samo kao okosnica)
- Zadaci za vježbu
- Kreiranje vlastitih primjera
- Projekti

Interakcija među sudionicima nastave

KONZULTACIJE

- unaprijed određeni termini konzultacija
- e-mail (*ime.prezime@fer.hr*)
- Termini će biti objavljeni na stranici predmeta.

Elementi ocjenjivanja (okvirno)

- aktivna prisutnost na predavanjima \Rightarrow 5 bodova
 - 3 projekta 3 x 6 bodova \Rightarrow 18 bodova
 - kontrolne zadaće 3 x 4 boda \Rightarrow 12 bodova
 - 1. međuispit \Rightarrow 15 bodova
 - 2. međuispit \Rightarrow 20 bodova
 - završni (ili ponovljeni završni) ispit \Rightarrow 30 bodova
 - MOGUĆI BROJ BODOVA UKUPNO \Rightarrow 100 bodova
-
- Za prolaznu ocjenu potrebno je ostvariti ukupno ≥ 50 bodova

Međuispiti i završni ispiti

- Na 1. međuispitu provjerava se znanje gradiva 1. nastavne cjeline
- Na 2. međuispitu provjerava se znanje gradiva 1. i 2. nastavne cjeline
- Na završnom ispitu i ponovljenom završnom ispitu provjerava se znanje čitavog gradiva

Način rada

- Diplomski studij = Istraživački studij
- Titula M.Sc. (Master of Science – Prema akreditaciji ASIIN)
- Završetkom studija studenti moraju biti sposobni za samostalan znanstveno-istraživački rad
- Rezultate istraživanja moći će primijeniti u praksi
- Način rada:
 - Više samostalnog rada
 - Vlastita inicijativa
 - Veća odgovornost za ostvarene rezultate

Ankete

- Studenti sudjeluju u vrednovanju kvalitete nastave i nastavnika putem ankete. Anketiranje se provodi za svaki predmet tri puta tijekom semestra.

Članak 20. stavak 5. Pravilnika o sveučilišnom preddiplomskom i diplomskom studiju na Sveučilištu u Zagrebu, Fakultetu elektrotehnike i računarstva

Okvirni raspored

- 11.9. Uvod
- 18.9. Objektne baze podataka
- 25.9. UML
- 2.10. Objektno-relacijske baze podataka
- 9.10. Objektno-relacijske baze podataka
- 30.10. XML baze podataka
- 6.11. XML baze podataka
- 13.11. Geoprostorne baze podataka
- 20.11. Geoprostorne i vremenske baze podataka
- 11.12. Multimedijske baze podataka
- 18.12. Web baze podataka, upravljanje podacima u Peer-to-Peer okruženju
- 8.1. Integracija podataka iz različitih izvora, semantički web
- 15.1. Modeli podataka u bioinformatiči



Uvod

"Način na koji prikupljate informacije,
upravljate njima i koristite ih, odredit će hoćete
li pobijediti ili izgubiti."

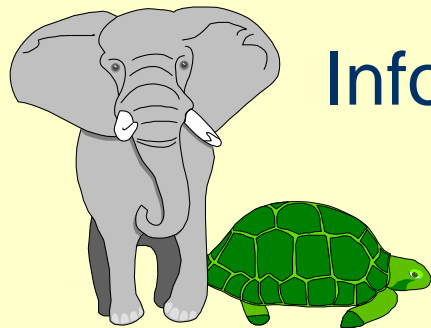
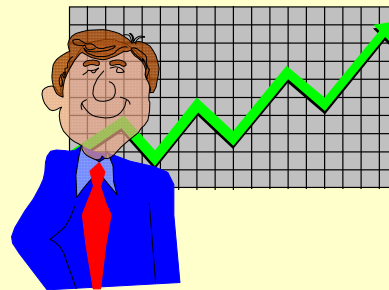
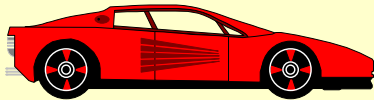
Bill Gates

Baza podataka (*Database*)

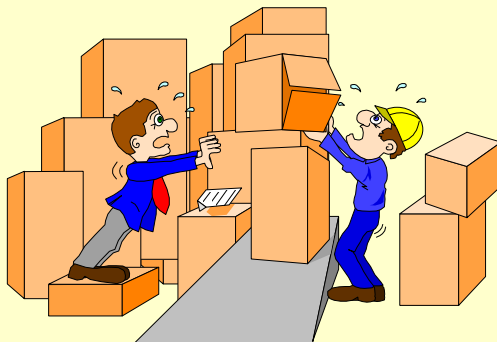
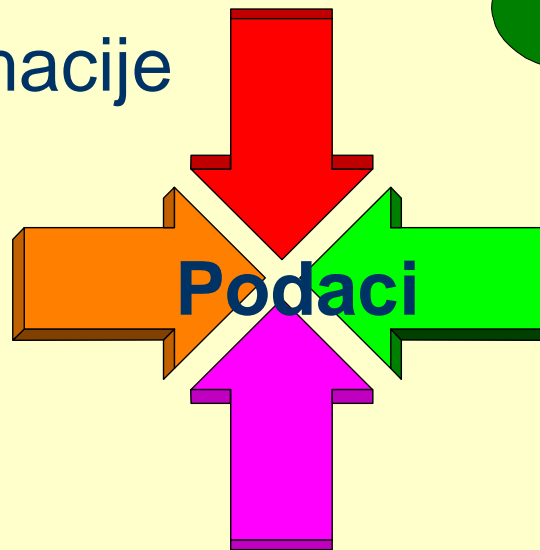
- BAZA PODATAKA je skup podataka koji su pohranjeni i organizirani tako da mogu zadovoljiti zahtjeve korisnika.
(*M. Vetter, 1981.*)
- BAZA PODATAKA je skup međusobno povezanih podataka, pohranjenih zajedno, uz isključenje bespotrebne zalihosti (redundancije), koji mogu zadovoljiti različite primjene. Podaci su pohranjeni na način neovisan o programima koji ih koriste. Prilikom dodavanja novih podataka, mijenjanja i pretraživanja postojećih podataka primjenjuje se zajednički i kontrolirani pristup. Podaci su strukturirani tako da služe kao osnova za razvoj budućih primjena.

(*J. Martin, 1979.*)

Stvarni svijet



Informacije



Entitet (*Entity*)

- bilo što, što ima suštinu ili bit i posjeduje značajke s pomoću kojih se može razlučiti od svoje okoline
- osobe: studenti, radnici, građani, ...
 - student Horvat Ivan (0036123456)
- ostala bića
 - Gupčeva Lipa, pas Lajka
- objekti: vozila, strojevi, uređaji, ulice, zgrade, mjesta, ...
 - Eiffelov toranj, cepelin Hindenburg
- apstraktni pojmovi: boje, predmeti nastavnog programa, ...
 - predmet Baze podataka u nastavnom programu FER-2
- događaji (nešto se desilo, dešava se ili se planira da će se desiti)
 - dana 24.11.2006. održana se proslava povodom 50. godišnjice FER-a
- povezanost među objektima, osobama, događajima, ...
 - student Horvat Ivan (0036123456) **stanuje u Zagrebu**
- nešto o čemu želimo prikupljati i pohranjivati podatke

Atribut (*Attribute*)

- Entitet posjeduje neka SVOJSTVA ili ATTRIBUTE koji ga karakteriziraju.
 - Za Informacijski sustav visokih učilišta (ISVU) važna svojstva studenta Horvat Ivana su:
 - Matični broj studenta (JMBAG)
 - Ime
 - Prezime
 - Datum rođenja, ...
- Izbor svojstava (atributa) koje ćemo pratiti ovisi o namjeni informacijskog sustava
 - Horvat Ivan u informacijskom sustavu MUP-a bit će karakteriziran i atributima:
 - Boja kose
 - Boja očiju
 - Otisak prsta, ...

Skup entiteta (*Entity Set*)

- Slični entiteti se svrstavaju u skupove entiteta
- Slični su oni entiteti kojima se promatraju ista svojstva
- Svi entiteti koji su članovi istog skupa entiteta imaju iste attribute
- "atributi entiteta" \Leftrightarrow "atributi skupa entiteta"
 - atributi skupa entiteta PREDMET
 - sifPred
 - nazPred
 - ectdBod
 - nastProg
 - atributi skupa entiteta STUDENT
 - jmbag
 - ime
 - prezime
 - datRod

Domena i vrijednost atributa (*Domain, Attribute Value*)

- Za svaki entitet, atribut poprima vrijednosti iz određenog skupa vrijednosti koji predstavlja **domenu** tog atributa
 - domene atributa za skup entiteta PREDMET
 - sifPred: skup šifara predmeta - cijelih brojeva iz intervala [1,999999]
 - nazPred: skup naziva predmeta - nizova znakova duljine do 80 znakova
 - ectsBod: skup vrijednosti ECTS bodova - realnih brojeva iz intervala [0.5, 20.0] (s jednom znamenkom iza decimalne točke)
 - nastProg: skup oznaka nastavnih programa - nizova znakova duljine do 5 znakova
 - vrijednosti atributa za entitet **Baze podataka (31503)**
 - sifPred: 31503
 - nazPred: Baze podataka
 - ectsBod: 6.0
 - nastProg: FER-2

Skup entiteta - prikaz u obliku tablice

Skup entiteta PREDMET

atributi

entiteti

sifPred	nazPred	ectsBod	nastProg
31503	Baze podataka	6.0	FER-2
1228	Baze podataka	5.0	FER-1
19670	Matematika 1	7.0	FER-2
21006	Fizika 1	6.0	FER-2
90	Baze podataka	5.0	ETF-4

vrijednosti atributa

Identifikatori entiteta, ključevi

- Skupove atributa čije vrijednosti jednoznačno određuju entitet u promatranom skupu entiteta (dakle ne postoje dva entiteta s posve istim vrijednostima tih atributa) nazivamo IDENTIFIKATORIMA ili KLJUČEVIMA SKUPA ENTITETA.
- Primjer:** u skupu entiteta PREDMET prikazanom na slici:

sifPred	nazPred	ectsBod	nastProg
31503	Baze podataka	6.0	FER-2
1228	Baze podataka	5.0	FER-1
19670	Matematika 1	7.0	FER-2
21006	Fizika 1	6.0	FER-2
90	Baze podataka	5.0	ETF-4

- skup atributa { sifPred } **jest** ključ skupa entiteta
- skup atributa { nazPred } **nije** ključ skupa entiteta
- skup atributa { nazPred, nastProg } **jest** ključ skupa entiteta

Modeliranje stvarnog svijeta

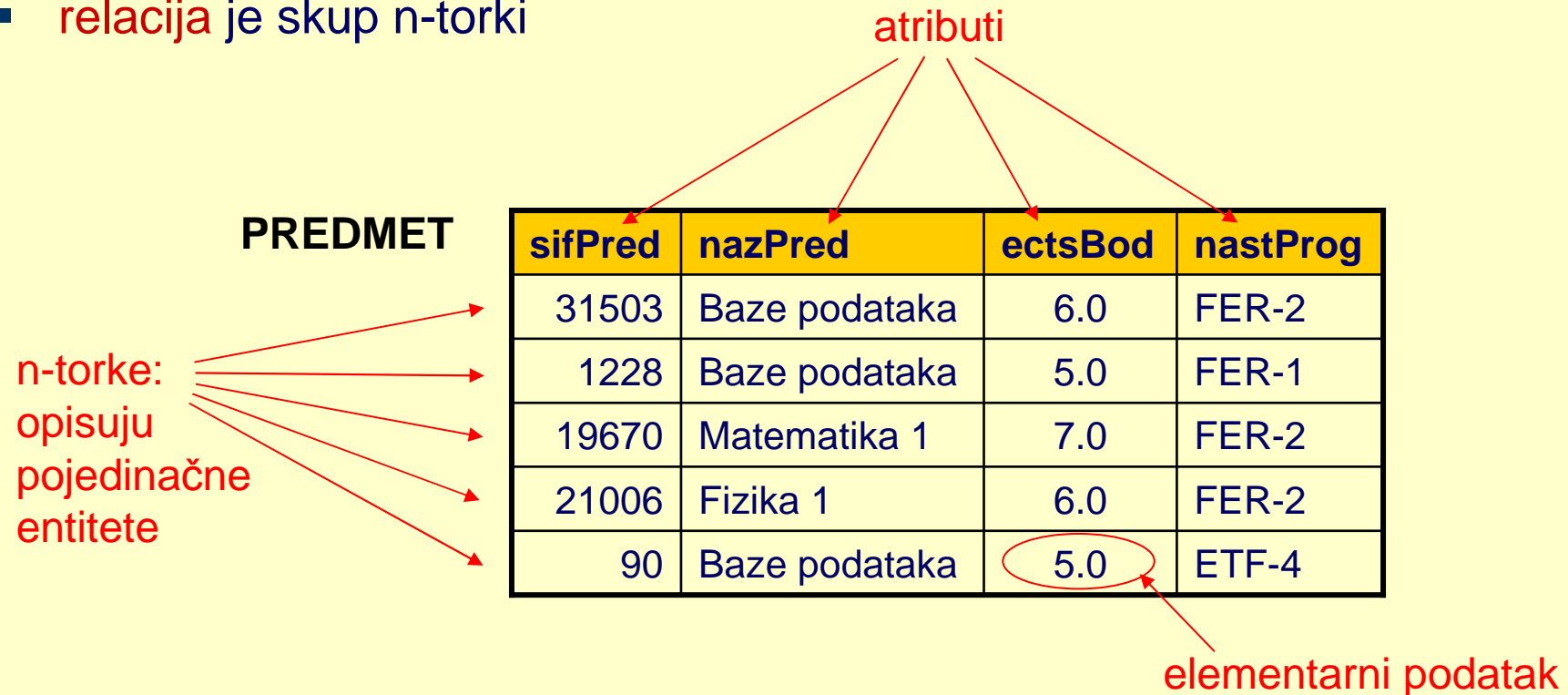
- Modeliranje stvarnog svijeta predstavlja preslikavanje stvarnog svijeta u oblik pogodan za računalnu obradu;
 - Baza podataka nekog informacijskog sustava predstavlja sliku stvarnog organizacijskog sustava;
 - Stvarni svijet, zbog njegove složenosti, ne možemo prikazati sa svim detaljima;
 - Stvarni svijet predstavlja se pojednostavnjenim, nadomjesnim modelom uz pomoć nekog formalnog sustava;
 - **Model podataka** je formalni sustav koji koristimo kod modeliranja baza podataka
- **Model podataka** je integrirani skup koncepata za opisivanje i upravljanje podacima, vezama među podacima i ograničenjima nad podacima

Model podataka (*Data Model*)

- **Model podataka je formalni sustav** koji se sastoji od:
 - **skupa objekata** - osnovnih elemenata (konceptata) baze podataka
 - **skupa operacija** koje se provode nad objektima
 - **skupa integritetskih ograničenja** (*integrity constraints*)
 - implicitno ili eksplicitno definiraju skup konzistentnih stanja podataka, promjena stanja, ili oboje
- **Povijesni razvoj modela podataka:**
 - Hijerarhijski model
 - Mrežni model
 - Relacijski model
 - ER model
 - Objektni model
 - Objektno-relacijski model
 - Polustrukturirani model

Relacijski model podataka – objekti

- elementi skupa objekata u relacijskom modelu podataka su **relacije**
- **relacija** je skup n-torki



- **shema relacije** obuhvaća naziv relacijske sheme (PREDMET) i skup atributa: (sifPred, nazPred, ectsBod, nastProg)

Relacijski model podataka – operacije

- operacija "selekcija" u relacijskom modelu podataka:

PREDMET predmet	sifPred	nazPred	ectsBod	nastProg
	31503	Baze podataka	6.0	FER-2
	1228	Baze podataka	5.0	FER-1
	19670	Matematika 1	7.0	FER-2
	21006	Fizika 1	6.0	FER-2
	90	Baze podataka	5.0	ETF-4

$\sigma_{\text{ectsBod}=6.0}$ (**predmet**)

sifPred	nazPred	ectsBod	nastProg
31503	Baze podataka	6.0	FER-2
21006	Fizika 1	6.0	FER-2

- ostale operacije u relacijskom modelu podataka
 - unija, razlika, presjek, projekcija, ...

Relacijski model podataka – integritetska ograničenja

- pravilo domenskog integriteta u relacijskom modelu podataka:

predmet

sifPred	nazPred	ectsBod	nastProg
31503	Baze podataka	6.0	FER-2
1228	Baze podataka	5.0	FER-1
19670	Matematika 1	7.0	FER-2
21006	Fizika 1	6.0	FER-2
90	Baze podataka	5.0	ETF-4

- domenski integritet:
 - vrijednost atributa sifPred mora biti iz intervala [1, 999999]
 - vrijednost atributa ectsBod mora biti iz intervala [0.5, 20.0]
 - ...
- ostala integritetska ograničenja u relacijskom modelu podataka:
 - entitetski, referencijski, ...

Arhitektura baze podataka

Arhitektura baze podataka

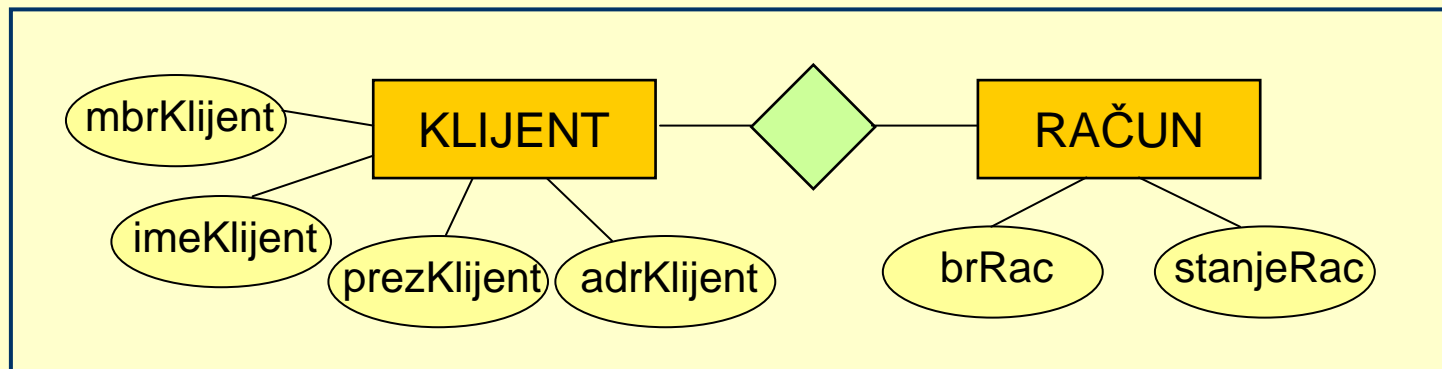


Arhitektura baze podataka

- Shema (struktura) baze podataka se opisuje na tri razine apstrakcije:
 - Na konceptualnoj razini opisuje se
 - **KONCEPTUALNA SHEMA**
 - Na unutarnjoj razini opisuje se
 - **INTERNA SHEMA**
 - Na vanjskoj razini opisuju se
 - **EKSTERNE SHEME**
- Jedna baza podataka ima jednu konceptualnu, jednu internu i (najčešće) više eksternih shema
- Shema baze podataka se *relativno* rijetko mijenja
- Sadržaj ili instanca baze podataka (skup svih podataka baze podataka u određenom trenutku) se ČESTO mijenja

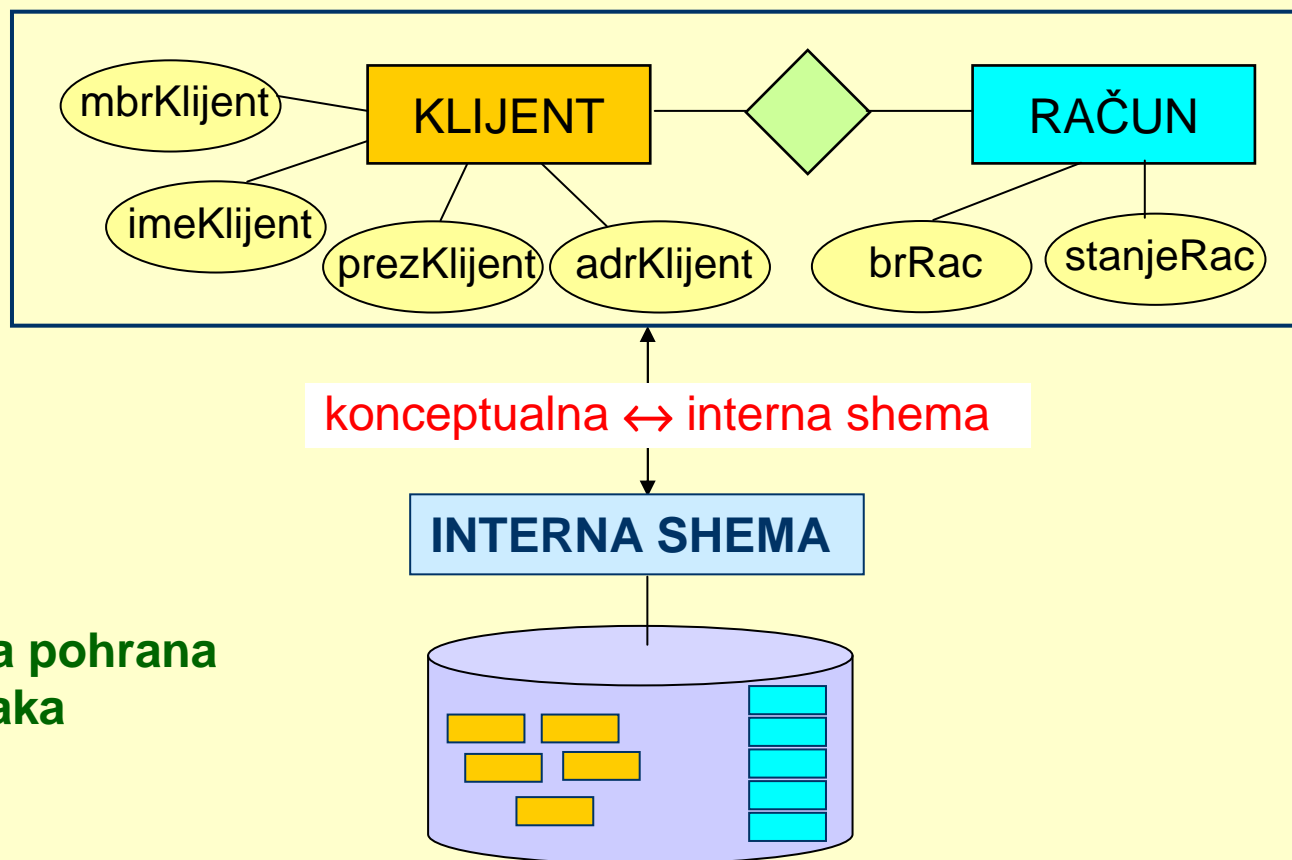
Konceptualna shema

- često se koristi i naziv LOGIČKA SHEMA
- sadrži opis svih entiteta i veza, atributa, domena i integritetska ograničenja
- konceptualna shema se može opisati korištenjem modela podataka, npr. relacijskog ili ER modela



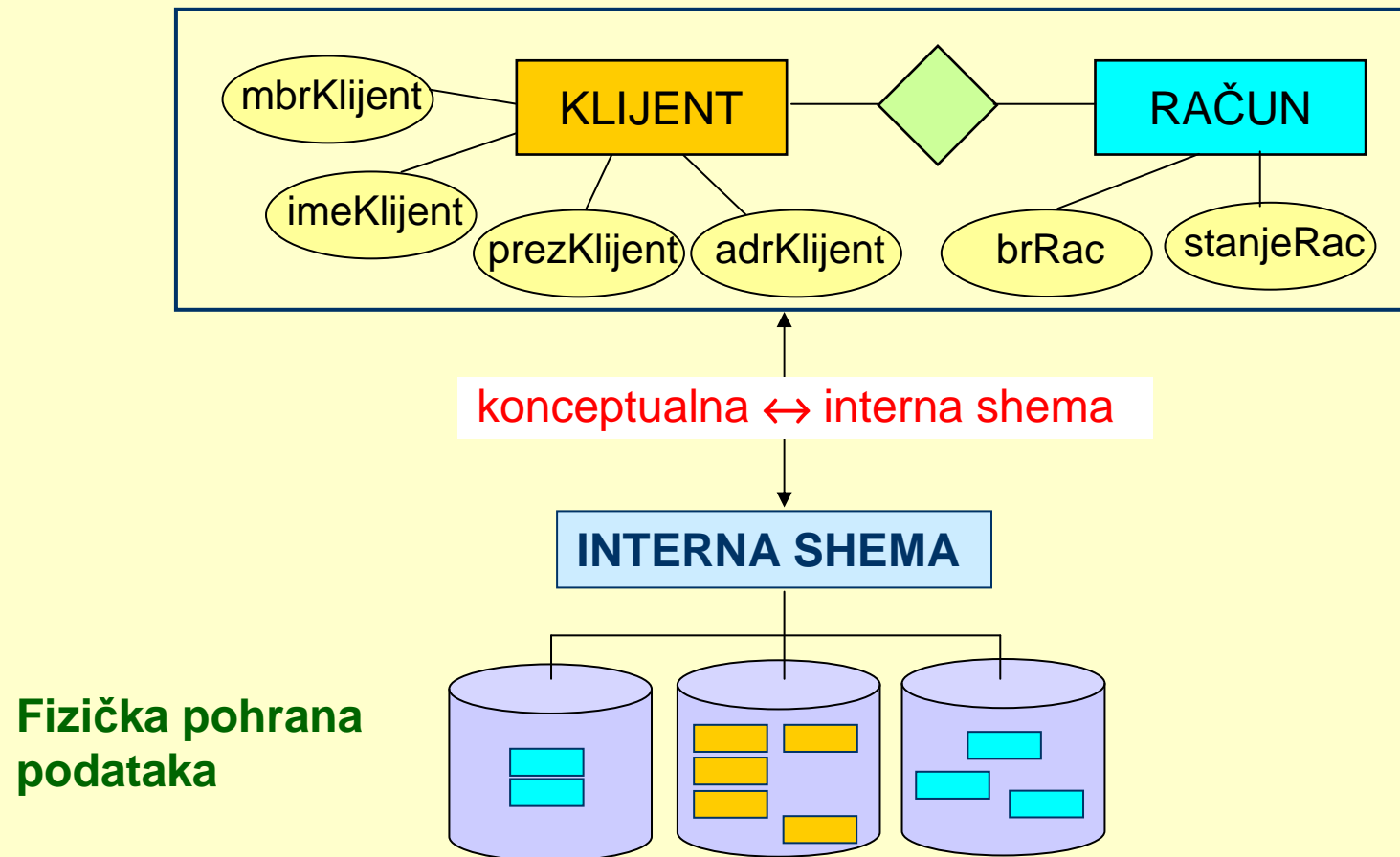
Interna shema

- opisuje detalje fizičke strukture pohrane i metode pristupa podacima: kako su podaci pohranjeni i koje se metode koriste za pristup podacima



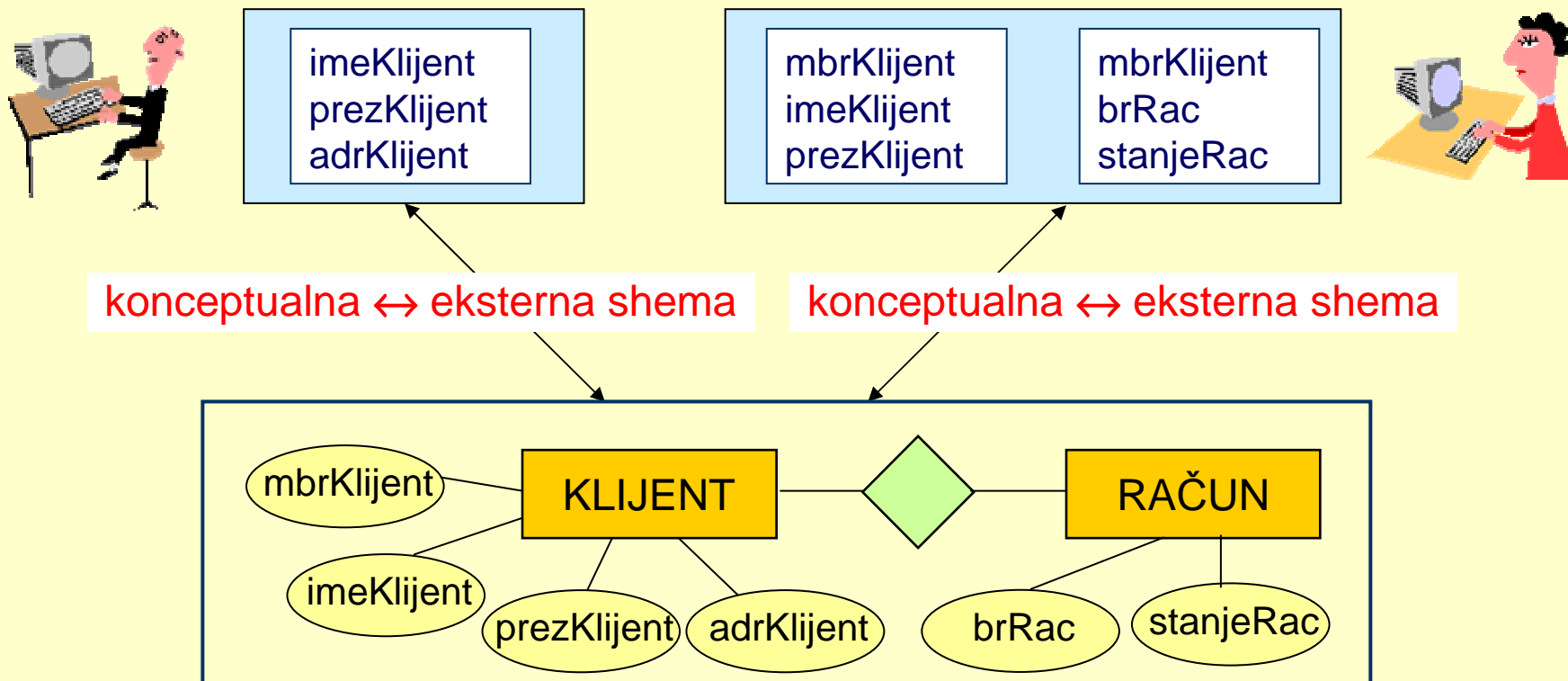
Fizička nezavisnost podataka

- izmjena interne sheme ne utječe na konceptualnu shemu



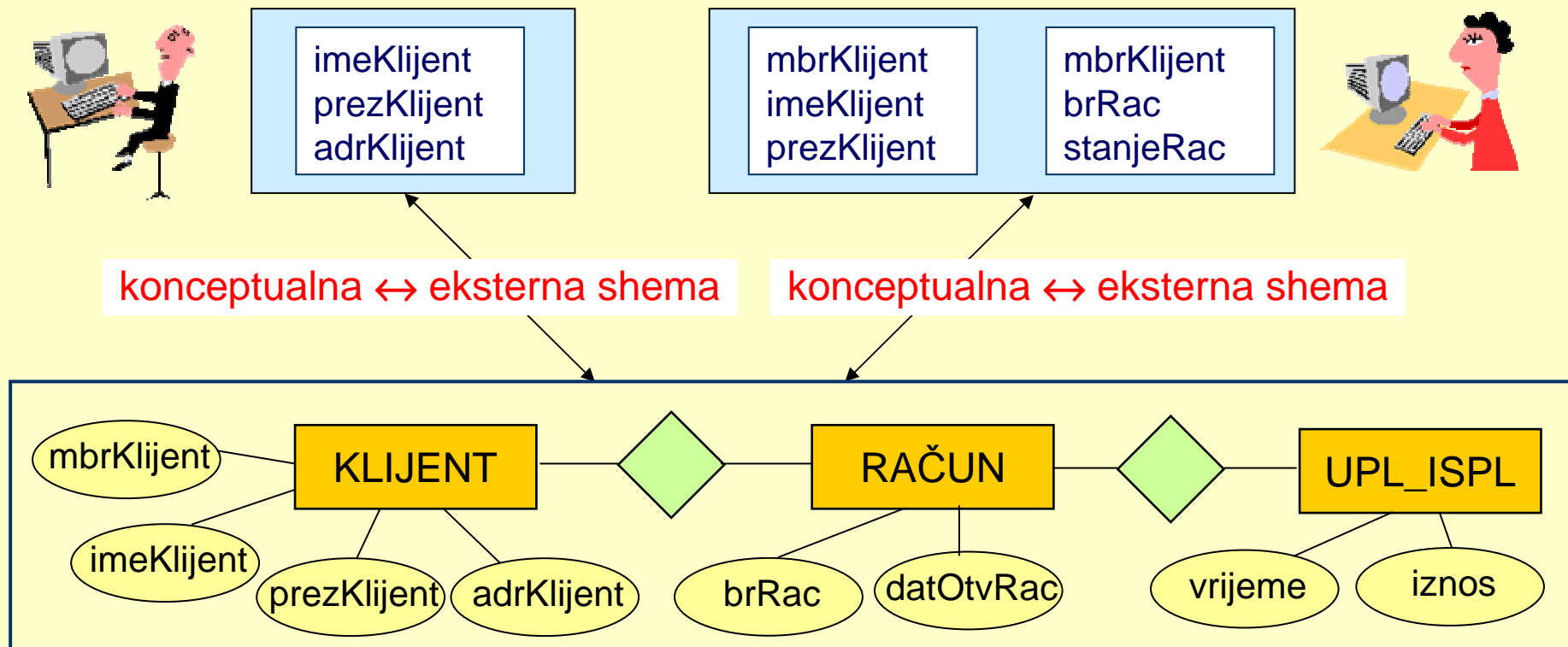
Eksterna shema

- eksterna shema opisuje "pogled" na dio baze podataka koji je namijenjen specifičnoj grupi korisnika
- osnova za opis eksternih shema je konceptualna shema



Logička nezavisnost podataka

- izmjena konceptualne sheme ne mora izazvati izmjenu eksternih shema → izmjena konceptualne sheme ne utječe na korisnike i aplikacijske programe koji ih koriste



Sustav za upravljanje bazom podataka - SUBP

Database Management System - DBMS

Sustav za upravljanje bazom podataka

Sustav za upravljanje bazom podataka - SUBP (*Database Management System - DBMS*) je **programski sustav** koji omogućava upravljanje podacima u bazi podataka.

SUBP se temelji na odabranom modelu podataka.

Prema modelu podataka na kojem se temelje, **SUBP**-ove dijelimo na:

- hijerarhijske,
- mrežne,
- relacijske,
- objektno-relacijske,
- objektno-orijentirane.

Sustav za upravljanje bazom podataka

- sakriva od korisnika detalje fizičke pohrane podataka
- osigurava logičku i fizičku nezavisnost podataka
- omogućuje definiciju i rukovanje podacima
 - DDL - Data Definition Language
 - DML - Data Manipulation Language
- obavlja funkciju zaštite podataka
 - integritet podataka
 - pristup podacima - autorizacija, sigurnost
 - kontrola paralelnog pristupa
 - obnova u slučaju razrušenja
- obavlja optimiranje upita

Jezici baze podataka

- DDL (*Data Definition Language*)
 - omogućava imenovanje i opis entiteta, atributa, veza i pripadnih ograničenja integriteta i pravila sigurnosti
 - koristi se za definiranje nove sheme baze podataka ili modificiranje postojeće
 - obavljanje DDL operacija rezultira izmjenom sadržaja rječnika podataka (metapodataka)
- DML (*Data Manipulation Language*)
 - omogućava korištenje skupa operacija za rukovanje podacima u bazi podataka
 - upitni jezik (*query language*)
 - ne sasvim korektno, pojam se koristi ne samo za operacije dohvata podataka, već i za operacije izmjene, brisanja i unosa podataka
 - proceduralni jezici, neproceduralni jezici

Napredni modeli i baze podataka

Zašto su modeli važni?



<http://en.wikipedia.org>



<http://www.people.com>



<http://www.voodoo.cz>

Modeli

- How are statisticians like artists? They both fall in love with their models.
- Meteorology is still somewhat an art. Like all artists, meteorologists tend to ***fall in love with their models.***
- Database designers ...

Zašto su modeli (baza podataka) važni?

- Predstavljaju temelj na kojem se gradi programska potpora ili informacijski sustav
- Temelji moraju biti stabilni – ako nisu – sve što je nad njima izgrađeno uruši će se
- ? Što znači stabilnost modela?
- ? Znači li to da se modeli ne mogu razvijati?
- Jezgra modela mora biti stabilna – nepromjenjiva
- Oko jezgre se dodaju novi elementi koje zahtijevaju nove primjene



Modeli podataka \leftrightarrow modeli baza podataka

- Model podataka je formalni sustav – definira koncepte - objekte, operacije i ograničenja integriteta
- Model podataka može biti:
 - Relacijski,
 - Entiteti – veze,
 - Objektno orijentirani,
 - Objektno-relacijski,
 - Polustrukturirani,
 - Model za pohranu i upravljanje nizovima i serijama, ...
- Model baze podataka je pojednostavnjena slika nekog segmenta stvarnog svijeta
- Model baze podataka temelji se na nekom modelu podataka (formalnom sustavu)
- Npr. Model baze podataka sustava visokih učilišta temeljen je na objektno-relacijskom modelu podataka

Koji su modeli podataka dobri?

- ? Koji su modeli podataka (formalni sustavi) dobri?
- ? Postoje li dobri i loši modeli podataka?
- ? Postoje li najbolji modeli podataka?
 - Ne postoje apsolutno dobri ili loši modeli podataka
 - Neki modeli podataka primjenjiviji su za neke primjene

- Potrebno je voditi računa o primjenjivosti modela u nekom području i o njegovim ograničenjima

DBMS Matrix

ad hoc upiti	relacijska baza	objektno relacijska baza
nema ad hoc upita	datoteka	objektna baza
	jednostavni podaci	složeni podaci

M. Stonebreaker, D. Moore:

Object-relational DBMSs – the next great wave, Morgan Kaufmann, 1996

Modeli podataka

- Osnovni objekti
- Operacije
- Integritetska ograničenja

Osnovni objekti

- Potrebno je analizirati moć izražavanja/ekspresivnost, mogućnost opisivanja objekata i pojava iz stvarnog svijeta s pomoću objekata promatranog modela podataka
- Ukoliko model podataka ne sadrži objekte kojima bismo na prirodan način opisali objekte i/ili pojave iz stvarnog svijeta, morat ćemo ih “simulirati” uz pomoć postojećih objekata modela podataka
- Problemi: kompleksnost, smanjena ekspresivnost, smanjena razumljivost modela baze podataka
- Primjeri:
 - U relacijskom modelu ne postoje objekti kojima bismo opisali veze među entitetima – opisujemo ih na isti način kao i entitete. Posljedice?
 - U relacijskom modelu ne postoji mehanizam nasljeđivanja

...Osnovni objekti

- Vrste objekata
 - Složenost objekata
 - Mogućnost definiranja vlastitih tipova podataka
 - Nasljeđivanje
 - Apstraktni tipovi podataka
 - Virtualni tipovi podataka
-
- Podržanost u alatima za oblikovanje modela baza podataka i izgradnju programske potpore - CASE (Computer Aided Software Engineering) alati, razvojna okruženja (framework), platforme za projektiranje i izgradnju informacijskih sustava/programske potpore

Operacije

- Razina jezika za upravljanje podacima (3. generacija, 4. generacija, 5. generacija, ..)
 - Proceduralnost
 - Ekspresivnost
 - Učahurivanje
 - Ponovna iskoristivost
-
- Podržanost u alatima za oblikovanje modela baza podataka i izgradnju programske potpore

Integritetska ograničenja

- Integritet – točnost, ispravnost
- Mogućnost definiranja pravila za očuvanje integriteta na razini baze podataka (ne u primijenjenoj programskoj podršci/aplikacijama)
- Uobičajena ograničenja:
 - Entitetski integritet, pravila ključa, pravila stranog ključa, ograničenje domene, ograničenja odnosa među podacima
 - Poslovna pravila

Meta podaci

- Podaci o podacima
- U bazama podataka – podaci kojima opisujemo podatke sadržane u bazi podataka
- Zapisani su u rječniku podataka ili proširenom rječniku podataka
- Svrha – pohranjivanje znanja o podacima
 - Korištenje pri izradi programske potpore – nazivi ekranskih formi, polja u ekranskim formama, poruke korisnicima – npr. poruke o pogreškama
 - Dokumentacija
 - Olakšavanje održavanja baze podataka i programske potpore, migracije podataka, integracije podataka

Predmet u akademskoj godini



Visoko učilište

36

Fakultet elektrotehnike i računarstva

Sveučilište u Zagrebu

Predmet

34539

Napredni modeli i baze podataka

Akademska godina

2008

/

2009

Deaktivacija oslobođenja na studomatu

Ne

2 / 2

Opis predmeta

Literatura

Izvođači

Oznaka jezika

en

Naziv jezika

Engleski

Opis predmeta

The goal of the course is to familiarize students with key concepts and issues related to emerging database technologies, object-oriented databases, object relational and XML databases. Course focuses on advanced applications based on semistructured, multimedia and spatio-temporal data, as well as data structures used in bioinformatics, life sciences and other.

2 / 2

Meta podaci

- Koji meta podaci se pohranjuju?
- Izbor meta podataka ovisi o primijenjenom modelu podataka
- Relacijski model: entiteti, atributi, domene, primarni ključevi, alternativni ključevi, strani ključevi
- ER model: entiteti, veze, spojnost i preslikavanje, atributi entiteta, atributi veza, ključevi entiteta, ključevi veza, domene atributa, hijerarhijski odnos entiteta (generalizacija/specijalizacija)

Oblikovanje ER modela

- **definiranje entiteta**
 - ime, opis, komentar
- **definiranje veza**
 - ime, opis, komentar, entiteti koje povezuje, preslikavanje
- **definiranje atributa entiteta**
 - za svaki atribut: ime, opis, komentar, domena
 - definirati ključeve, provjeriti da li zadovoljava 3NF
- **definiranje atributa veza**
 - za svaki atribut: ime, opis, komentar, domena
 - definirati ključeve, provjeriti da li zadovoljava 3NF

Zadatak

- Oblikovati model **meta baze podataka** koji će omogućiti pohranu i pretraživanje podataka o nekoj (bilo kojoj ?) bazi podataka

Relacijski model podataka

Formalne definicije

Relacijska shema, relacija


Relacijska shema R (intenzija) je imenovani skup atributa:
 $R = \{ A_1, A_2, \dots, A_n \}$ ili $R = A_1, A_2, \dots, A_n$

Relacija r (ekstenzija) definirana nad relacijskom shemom R je konačan skup n-torki, označava se s

$r(R)$ ili $r(A_1, A_2, \dots, A_n)$ predstavlja trenutnu vrijednost.

Primjer:

STUDENT = mbrStudent, prezimeStudent, imeStudent, slikaStudent

student(STUDENT) = { <123456, Bošnjak, Neven, 


student(mbrStudent, prezimeStudent, imeStudent, slikaStudent)

123456	Bošnjak	Neven
234567	Katić	Maja
345678	Šuša	Matko




- Oznaka $t(A)$ predstavlja vrijednost koju atribut A poprima u n-torki t . $t(A)$ se naziva **A-vrijednost n-torke t**.

Primjer:

$t = \langle 123456, \text{Bošnjak, Neven, }  \rangle$
 $t(\text{prezimeStudent}) = \text{Bošnjak}$

- Neka je $X \subseteq R$. n-torka t reducirana na skup atributa X se naziva **X-vrijednost n-torke t i označava s $t(X)$**

Primjer:

$t = \langle 234567, \text{Katić, Maja, }  \rangle$
 $\text{prezimeStudent, imeStudent} \subset \text{STUDENT}$
 $t(\text{prezimeStudent, imeStudent}) = \langle \text{Katić, Maja} \rangle$

Karakteristike relacije

- Stupanj - broj stupaca (atributa)
- Kardinalnost - broj n-torki (tuple)

Operacije s relacijama

- Relacijska algebra
- Predikatni račun

Relacijska algebra

unija	\cup
presjek	\cap
razlika	\setminus
Kartezijev produkt	$*$
dijeljenje	$/$
projekcija	π
selekcija	σ
spajanje	\bowtie

- Primjer: $a = \sigma_{X=x \ \& \ Y=y} (d \cap (b \cup c))$
- proceduralnost - navođenje redoslijeda operacija

Predikatni račun

$$r = \{ t \mid f(t) \}$$

t je varijabla koja predstavlja:

- n-torke - n-torski račun
- domene - domenski račun

Primjer:

$$a = \{ t \mid (d(t) \wedge ((b(t) \vee c(t))) \wedge t(X) = x \wedge t(Y) = y \}$$

- neproceduralnost - navođenje predikata koje n-torke moraju zadovoljavati

Operacije relacijske algebre

- **Osnovne operacije**
- presjek, unija i razlika
- primijenjuju se nad relacijama koje su **unijski kompatibilne**
 - relacije istog stupnja
 - korespondentni atributi definirani nad istim domenama

Primjer relacija

R1=(Prezime, Ime, PostBr)

	Novak	Ivan	10000
r1	Seljan	Nada	10040
	Marić	Ana	10000

R2=(Prezime, Ime, Pbr)

	Kolar	Petar	10020
r2	Novak	Ivan	10000

Stupanj: d1 = 3

d2 = 3

Kardinalnost: m1 = 3

m2 = 2

Relacije su unijski kompatibilne.

Preimenovanje atributa

Neka su A i B atributi i neka je r relacija definirana na shemi R , $A \in R$, $B \notin R \setminus A$.

Neka je $\text{DOM}(A) = \text{DOM}(B)$. Neka je $R' = (R \setminus A) \cup B$.

Operacija preimenovanja atributa A u B u relaciji $r(R)$ označava se s

$\delta_{A \leftarrow B}(r)$, a definira pomoću izraza:

$$\delta_{A \leftarrow B}(r) = r'(R') = \{ t' \mid t \in r, t'(R \setminus A) = t(R \setminus A) \wedge t'(B) = t(A) \}$$

Neka su A_1, A_2, \dots, A_k različiti atributi u R i neka su B_1, B_2, \dots, B_k različiti atributi koji nisu članovi skupa $R \setminus (A_1, A_2, \dots, A_k)$.

Neka je $\text{DOM}(A_i) = \text{DOM}(B_i)$, za $1 \leq i \leq k$.

Simultano preimenovanje atributa A_1, A_2, \dots, A_k u attribute B_1, B_2, \dots, B_k u relaciji r označava se s

$$\delta_{A_1, A_2, \dots, A_k \leftarrow B_1, B_2, \dots, B_k}(r)$$

Unijska kompatibilnost relacija

Dvije relacije r i s definirane na shemama R i S ,
odnosno njihove sheme R i S su unijski kompatibilne
ukoliko postoji 1:1 preslikavanje

$$f : R \rightarrow S, \quad f(A_i) = B_j, \quad f^{-1}(B_j) = A_i,$$

pri čemu je $A_i \in R, B_j \in S, \text{DOM}(A_i) = \text{DOM}(B_j)$.

Unija relacija

Neka su r i s unijski kompatibilne relacije definirane na shemama R i S . Operacija unije relacija r i s označava se $s \cup r \cup s$, a definira izrazom

$$r \cup s = q(R) = \{ t \mid t \in r \vee t \in s \} \text{ ako je } R = S$$

$$r \cup s = q(R) = \{ t \mid t \in r \vee t \in \delta_{Y \leftarrow X} s(S) \} \text{ ako je } R \neq S$$

gdje je $X \subseteq R \setminus S$, $Y \subseteq S \setminus R$,

$$Y = f(X), \quad X = f^{-1}(Y),$$

i $f : X \rightarrow Y$ je restrikcija od $f : R \rightarrow S$ i $f^{-1} : S \rightarrow R$

Presjek relacija

Neka su r i s unijski kompatibilne relacije definirane na shemama R i S .

Operacija presjeka relacija r i s označava se $s \cap r$, a definira izrazom:

$$r \cap s = q(R) = \{ t \mid t \in r \wedge t \in s \} \text{ ako je } R = S$$

$$r \cap s = q(R) = \{ t \mid t \in r(R) \wedge t \in \delta_{Y \leftarrow X} s(S) \} \text{ ako je } R \neq S$$

gdje je $X \subseteq R \setminus S$, $Y \subseteq S \setminus R$,

$$Y = f(X), \quad X = f^{-1}(Y),$$

i $f : X \rightarrow Y$ je restrikcija od $f : R \rightarrow S$ i $f^{-1} : S \rightarrow R$

Razlika relacija

Neka su r i s unijski kompatibilne relacije definirane na shemama R i S . Operacija razlike relacija r i s označava se $s \setminus r$, a definira izrazom:

$$r \setminus s = q(R) = \{ t \mid t \in r \wedge t \notin s \} \text{ ako je } R = S$$

$$r \setminus s = q(R) = \{ t \mid t \in r \wedge t \notin \delta_{Y \leftarrow X} s(S) \} \text{ ako je } R \neq S$$

gdje je $X \subseteq R \setminus S$, $Y \subseteq S \setminus R$,

$$Y = f(X), \quad X = f^{-1}(Y),$$

i $f : X \rightarrow Y$ je restrikcija od $f : R \rightarrow S$ i $f^{-1} : S \rightarrow R$

Spajanje n-torki

Neka su $a = (a_1, a_2, \dots, a_k)$ i $b = (b_1, b_2, \dots, b_m)$ n-torke.
Operacija spajanja n-torki označava se s $a \wedge b$, a
definira pomoću izraza

$$a \wedge b = (a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_m).$$

Kartezijev produkt relacija

Neka su r i s relacije. Operacija Kartezijev produkt relacija r i s označava se $r \times s$, a definira izrazom

$$r \times s = \{ (t_r \wedge t_s) \mid t_r \in r \wedge t_s \in s \}$$

Kartezijev produkt relacija ne mora nužno biti relacija.

Specifične operacije

- Unarne
 - Projekcija - izdvajanje stupaca (atributa)
 - Selekcija - izdvajanje redaka (n-torki)
- Binarne
 - Prirodno pridruživanje
 - Theta pridruživanje
 - Pridruživanje s izjednačavanjem

Projekcija relacije

$$s = \pi_{A_1, \dots, A_k} (r)$$

- Relacijska shema od s:

$$S = \{ A_1, \dots, A_k \}$$

- stupanj: $d_s = k$
- kardinalnost: $\text{card}(s) \leq \text{card}(r)$
 - obavlja se eliminacija duplikata

Projekcija relacije

Neka je r relacija definirana na shemi R i neka je X skup atributa, $X \subseteq R$.

Operacija projekcije relacije $r(R)$ na skup atributa X se označava s $\pi_X(r)$, a definira izrazom

$$\pi_X(r) = \rho(X) = \{ t(X) \mid t \in r \}$$

Operacija selekcije

$$s = \sigma_F (r)$$

s je relacija koja sadrži skup n-torki iz r koje zadovoljavaju formulu selekcije (predikat) F

Relacija s je: $s \subseteq r$

Relacijska shema od s: $S = R$

Formula selekcije F

- Neka je r relacija definirana na shemi R i neka su A i B atributi iz R .
- Neka je θ relacijski operator iz skupa $\{=, \neq, >, \geq, <, \leq\}$.
Neka je c konstanta iz skupa $\text{DOM}(A)$. Formula F je definirana rekurzivnim izrazom:
- $A \theta B$, $A \theta c$, $c \theta A$ su formule. Ove formule se nazivaju jednostavnim formulama (atomi).
- Ako su G i H formule, tada su $G \wedge H$, $G \vee H$, $\neg G$, $\neg H$ također formule
- Ništa drugo nije formula.

... Formula selekcije F

Neka je $R = A_1, A_2, \dots, A_k$ i neka je r relacija definirana na shemi R .

Formula F je primjenjiva na r ako su:

- konstante koje se pojavljuju u F iz skupa $\text{DOM}(A_1) \cup \text{DOM}(A_2) \cup \dots \text{DOM}(A_k)$ i
- ako su atributi koji se pojavljuju u F iz skupa R .

Selekcija relacije

Neka je r relacija definirana na shemi R i neka je F formula primjenjiva na r . Operacija selekcije nad relacijom r uz formulu selekcije F označava se s $\sigma_F(r)$, a definira izrazom

$$\sigma_F(r) = q(R) = \{ t \mid t \in r \wedge t \text{ zadovoljava } F \}$$

Pridruživanje relacija

- Theta pridruživanje
- Pridruživanje s izjednačavanjem
- Prirodno pridruživanje

Theta pridruživanje (θ – pridruživanje)

$S = r_1 \bowtie_{\theta} r_2$ θ je uvjet na temelju kojeg se obavlja pridruživanje

Primjer: ZRAKOPLOV=(Tip, Dolet) LINIJA=(Let, Udaljenost)

B747 6000

DC9 3000

B727 4500

CA-825 7200

A-224 3300

CA-878 4700

CA-224 2000

mogućnost = linija \bowtie zrakoplov
Udaljenost \leq Dolet

MOGUĆNOST=(Let, Udaljenost, Tip, Dolet)

A-224 3300 B747 6000

A-224 3300 B727 4500

CA-878 4700 B747 6000

CA-224 2000 B747 6000

CA-224 2000 DC9 3000

CA-224 2000 B727 4500

Theta pridruživanje relacija

Neka je θ relacijski operator iz skupa $\{=, \neq, >, \geq, <, \leq\}$. Neka su r i s relacije definirane na shemama R i S . Neka su A i B atributi, $A \in R$ i $B \in S$.

Operacija theta-pridruživanja relacija r i s na osnovi formule $A \theta B$ označava se s

$r \bowtie_{A \theta B} s$, a definirana je izrazom

$$r \bowtie_{A \theta B} s = \{ t_r \wedge t_s \mid t_r \in r \wedge t_s \in s \wedge t_r(A) \theta t_s(B) \}$$

Rezultat theta-pridruživanja relacija ne mora nužno biti relacija.

Umjesto jednostavne formule $A \theta B$, kao formula pridruživanja može se koristiti složena formula dobivena konjunkcijom jednostavnih formula oblika $A_i \theta B_j$, pri čemu je $A_i \in R$ i $B_j \in S$.

Pridruživanje s izjednačavanjem (*eng. equi-join*)

Pridruživanje relacija s izjednačavanjem je poseban oblik theta-pridruživanja u kojem se kao θ operator koristi isključivo operator jednakosti (=).

Prirodno pridruživanje relacija

Obavlja se na temelju jednakih vrijednosti istoimenih atributa

$R_1 =$	<u>Prezime</u>	<u>Ime</u>	<u>PostBr</u>
	Novak	Ivan	10000
r_1	Seljan	Nada	10040
	Marić	Ana	10000

$R_{61} =$	<u>PostBr</u>	<u>NazivGrad</u>
	10000	Zagreb
r_{61}	10020	Novi Zagreb

$$r_{12} = r_1 \triangleright \triangleleft r_{61}$$

$R_{12} =$	<u>Prezime</u>	<u>Ime</u>	<u>PostBr</u>	<u>NazivGrad</u>
	Novak	Ivan	10000	Zagreb
$r_{12}(R_{12})$	Marić	Ana	10000	Zagreb

Prirodno pridruživanje

Neka su r i s relacije definirane na shemama R i S .

Operacija prirodnog pridruživanja relacija r i s označava se $r \bowtie s$, a definira izrazom:

$$r \bowtie s = \pi_{R \cup S}(\sigma_{t(R)=t_r \wedge t(S)=t_s}(r \cup s)) = \{ t \mid t_r \in r \wedge t_s \in s, t(R) = t_r \wedge t(S) = t_s \}$$

Usporediti s theta pridruživanjem

$$r \bowtie_{A \theta B} s = \{ t_r \wedge t_s \mid t_r \in r \wedge t_s \in s \wedge t_r(A) \theta t_s(B) \}$$