

Geodetske baze podataka

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

1

Sadržaj kolegija

- ♦ Cilj današnjeg predavanja: kratak pregled sadržaja kolegija "Geodetske baze podataka"
- ♦ Sadržaji svih kolegija na Geodetskom fakultetu dostupni su na adresi:
<http://www.geof.hr/gf/ects/ects.pdf>

2002/2003

Medak: Geodetske baze podataka

2

Cilj kolegija

- ♦ ovladati temeljima suvremenih baza podataka s posebnim naglaskom na prostorne (geo-) podatke
- ♦ usvojiti osnove samostalnog rukovanja komercijalnim i slobodnim programskim paketima za baze podataka (relacijski model, ER-model)
- ♦ steći osnovna znanja o konceptima baza podataka koje će postati industrijski standard u vrijeme početka profesionalne karijere (objektno-orijentirane baze podataka)

2002/2003

Medak: Geodetske baze podataka

3

Uvod

- ♦ Pojam baze podataka, definicija i kratak povijesni pregled
- ♦ Dizajn baza podataka:
 - konceptualni
 - implementacijski
 - fizički dizajn
- ♦ Fizička organizacija podataka

2002/2003

Medak: Geodetske baze podataka

4

Baze prostornih podataka

- ♦ Definicije i standardi prostornih podataka
- ♦ Osnovni elementi:
 - entiteti
 - objekti
 - tipovi objekata
 - atributi
 - vrijednosti atributa
 - slojevi

2002/2003

Medak: Geodetske baze podataka

5

Relacijske baze podataka

- ♦ Osnovni elementi: tablice i relacije, atributi i ključevi
- ♦ Osnove relacijske algebre: selekcija, projekcija, produkt, povezivanje (**join**), razlika, unija, presjek
- ♦ Upitni jezici (SQL)
- ♦ ER-dijagram
- ♦ Normalne forme relacijskih baza podataka
- ♦ Primjene u geodeziji

2002/2003

Medak: Geodetske baze podataka

6

Sustav za upravljanje bazom podataka (DBMS)

- ♦ Integritet podataka
- ♦ Transakcije
- ♦ ACID:
 - *atomicity*
 - *consistency*
 - *isolation*
 - *durability*
- ♦ Višekorisničke baze podataka

2002/2003

Medak: Geodetske baze podataka

7

Uvod u objektno-orijentirani dizajn

- ♦ Osnove OO programskih jezika i baza podataka:
 - nasljeđivanje ili *inheritance*
 - višeočlivenost ili *polymorphism*
 - prikrivanje (unutarnjeg ustroja) ili *encapsulation*
- ♦ Primjene u geodeziji (kompleksnost prostornih podataka i procesa)
- ♦ Uvod u UML (*Unified Modeling Language*): dijagrami klasa

2002/2003

Medak: Geodetske baze podataka

8

Budućnost baza podataka

- ♦ Objektno-relacijski model
- ♦ Objektno-orijentirane baze podataka
- ♦ Deduktivne baze podataka (Prolog)
- ♦ Prostorno-vremenske baze podataka
- ♦ Baze prostornih podataka na Internetu:
 - JDBC (Java DataBase Connectivity)
 - XML (eXtended Markup Language)
 - GML (Geography Markup Language)

2002/2003

Medak: Geodetske baze podataka

9

Program vježbi

- ♦ Auditorne vježbe
- ♦ Računalne vježbe (podjela u grupe)
- ♦ Relacijske baze podataka, SQL (MS Access, *freeware* Interbase)
- ♦ Deduktivne baze podataka, predikatna logika (Prolog)

2002/2003

Medak: Geodetske baze podataka

10

Cilj kolegija

- ♦ ovladati temeljima suvremenih baza podataka s posebnim naglaskom na prostorne (geo-) podatke
- ♦ usvojiti osnove samostalnog rukovanja komercijalnim i slobodnim programskim paketima za baze podataka (relacijski model, ER-model)
- ♦ steći osnovna znanja o konceptima baza podataka koje će postati industrijski standard u vrijeme početka profesionalne karijere (objektno-orijentirane baze podataka)

2002/2003

Medak: Geodetske baze podataka

11

Baze podataka – motivacija i definicija

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

12

Sadržaj

- ♦ Osnovni pojmovi i definicije:
 - baza podataka
 - informacijski sustav
 - informacija
 - podatak
- ♦ Modeli podataka
- ♦ Sustav za upravljanje bazom podataka
- ♦ Razine apstrakcije pri dizajnu baze
- ♦ Logički modeli podataka
- ♦ Relacijski model

2002/2003

Medak: Geodetske baze podataka

13

Tipični problemi u obradi informacija bez baze podataka

- ♦ Redundancija i inkonzistentnost
- ♦ Ograničena mogućnost pristupa podacima
- ♦ Problemi pri višekorisničkoj upotrebi podataka
- ♦ Gubitak podataka
- ♦ Povreda integriteta podataka
- ♦ Sigurnosni problemi
- ♦ Visoki troškovi razvoja aplikacija

2002/2003

Medak: Geodetske baze podataka

14

Baza podataka i sustavi

- ♦ **Baza podataka** je centralno mjesto **informacijskog sustava**. Poshranjeni podaci u bazi podataka opisuju trenutno stanje dijela realnog svijeta za koji je i razvijen informacijski sustav, naravno na način pogodan za računalnu obradu.
- ♦ **Informacijski sustav** je uvijek podsustav nekog organizacijskog sustava, a svrha mu je prikupljanje, obrada, pohranjivanje i distribucija informacija, koje su potrebne za praćenje rada i upravljanje tim organizacijskim sustavom ili nekim njegovim podsustavom.

2002/2003

Medak: Geodetske baze podataka

15

Informacija i podatak

- ♦ **INFORMACIJA** je znanje koje primatelju opisuje nove činjenice. To znanje se materijalizira u obliku **PODATAKA**, simbola koji služe za prikaz informacija u svrhu spremanja, prijenosa i obrade.
- ♦ Informacija je i obrađeni podatak koji za primatelja ima karakter novosti, otklanja neizvjesnost i služi kao podloga za odlučivanje.
- ♦ **PODATAK** je skup znakova u memoriji koji prikazuje jedan ili više elemenata informacije.

2002/2003

Medak: Geodetske baze podataka

16

Baza podataka - definicija

- ♦ **BAZA PODATAKA** je skup međusobno povezanih podataka pohranjenih bez nepotrebne zalihosti s ciljem da na optimalni način posluže u raznim primjenama. Podaci se spremaju neovisno o programima koji ih koriste, zajedničkim pristupom dodaju se novi podaci te mijenjaju i premještaju postojeći.
- ♦ Podaci se pohranjuju u bazu podataka na jedan organizirani način koristeći odgovarajući **MODEL PODATAKA**.

2002/2003

Medak: Geodetske baze podataka

17

Model podataka

- ♦ **formalni sustav** koji mora imati barem sljedeće tri komponente:
 1. **Skup objekata** koji su osnovni elementi baze podataka;
 2. **Skup operacija** koje možemo izvoditi nad objektima definiranim pod 1) i kojima se mogu pretraživati, dobivati i modificirati podaci o tim objektima;
 3. **Skup općih pravila integriteta** podataka koja implicitno ili eksplicitno definiraju skup konzistentnih stanja podataka ili promjena stanja, ili oboje i koja su općenita u smislu da su primjenjiva na bilo koju bazu podataka koja koristi taj model.

2002/2003

Medak: Geodetske baze podataka

18

Model podataka se koristi za:

- ♦ razvijanje sustava za upravljanje bazom podataka,
- ♦ razvijanje programskih jezika za rad s podacima u bazi podataka,
- ♦ razvijanje općih teorija oblikovanja baza podataka,
- ♦ istraživanje svojstava podataka, odnosno svojstava baza podataka.

2002/2003

Medak: Geodetske baze podataka

19

Sustav za upravljanje bazom podataka DBMS - Database Management System

- ♦ je **programski sustav** koji osigurava **osnovne funkcije odabranog modela podataka** u postupku **kreiranja i korištenja** baze podataka. Sastoji se od integrirane kolekcije programske podrške koja omogućava:
 - **opis i manipulaciju podacima** pomoću posebnog jezika (posebnih jezika),
 - **visoki nivo sučelja** prema podacima nezavisan od strukture podataka u računalu,
 - **efikasno korištenje i razumijevanje informacija** pohranjenih u bazi podataka, zahvaljujući skupu programskih alata (pomagala).

2002/2003

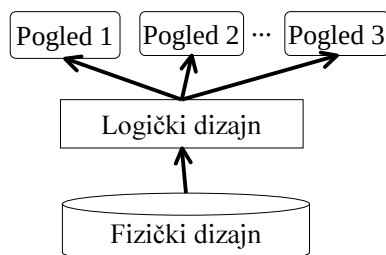
Medak: Geodetske baze podataka

20

Razine apstrakcije

Neovisnost podataka:

- fizička neovisnost
- logička neovisnost

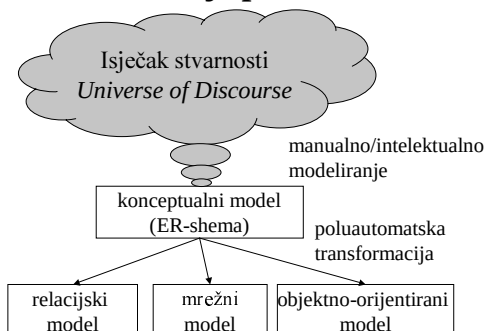


2002/2003

Medak: Geodetske baze podataka

21

Modeliranje podataka

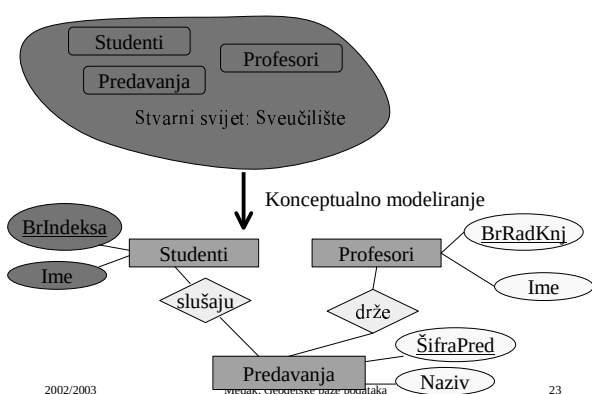


2002/2003

Medak: Geodetske baze podataka

22

Mali primjer



2002/2003

Medak: Geodetske baze podataka

23

Logički modeli podataka

- ♦ Mrežni model
- ♦ Hijerarhijski model
- ♦ Relacijski model
- ♦ Objektno-orijentirani model
- ♦ Deduktivni model

2002/2003

Medak: Geodetske baze podataka

24

Relacijski model

Student		slusa		Predavanje	
BrInd	Ime	BrInd	ŠifraPr	ŠifraPr	Naziv
5555	Ivana	5555	2007	2003	Geoinformatika II
5556	Marko	5556	2007	2007	Geodetske baze podataka
...	...	...	...	...	...

Select Ime

From Student, slusa, Predavanje

Where Student.BrInd = slusa.BrInd **and**
slusa.SifraPr = Predavanje.SifraPr **and**
Predavanje.Naziv = 'Geoinformatika II';

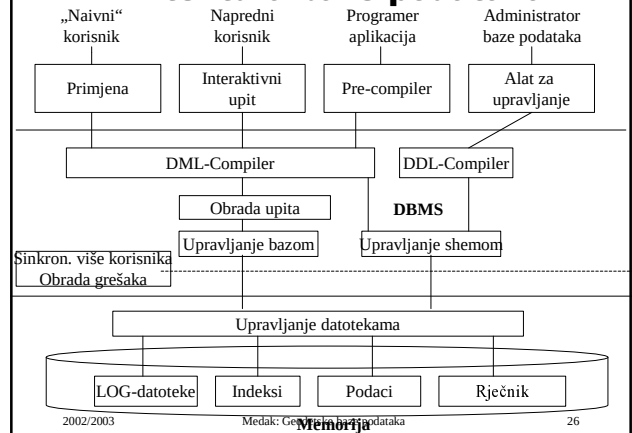
update Predavanje

set Naziv = 'Geodetske baze podataka'

where SifraPr = 2007;

25

Arhitektura baze podataka



26

Baze podataka – modeli i elementi

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

27

Sadržaj

- ♦ Flat files vs. baze podataka – primjer geopodataka
- ♦ Arhitektura tri razine:
 - interni model
 - konceptualni model
 - eksterni model
- ♦ Osnovni elementi baze podataka: entiteti, atributi i odnosi (veze)

2002/2003

Medak: Geodetske baze podataka

28

Ravne datoteke (flat files)

- ♦ **pisma:** slova, riječi, odjeljci
- ♦ **crteži:** točke, linije, simboli, tekst, šrafure
- ♦ **hipertekst:** slike, audio, video, veze (linkovi)
- ♦ obrađuju se sekvencijalno – baza podataka nije nužna
- ♦ bez ili s vrlo malo strukture

2002/2003

Medak: Geodetske baze podataka

29

Primjer prostornih podataka

- ♦ rasterska slika $\Rightarrow$ flat file (ravna datoteka)
- ♦ Atlas Europe $\Rightarrow$ strukturirani objektni prikaz
 - || Švicarska | Bern | 7 mil | njem. fran. tal. | $x_1/y_1 \dots x_k/y_k$
 - || Njemačka | Berlin | 82 mil | njem. | $x_{k+1}/y_{k+1} \dots x_m/y_m$
 - || Austrija | Beč | 8 mil | njem. | $x_{m+1}/y_{m+1} \dots x_n/y_n \dots$
- ♦ preduvjeti:
 - dogovor oko znaka za razdvajanje (||) i dr.
 - problemi oko koordinata granica (otoci, "rupe"...)

2002/2003

Medak: Geodetske baze podataka

30

Ravna datoteka

- ♦ Nedostaci ravne datoteke:
 - nemoguć naknadni detaljniji opis granice (progušćavanjem koordinata)
 - dodavanje novog atributa izuzetno složeno
- ♦ Ravna datoteka je pogodna za statičke primjene (sadržaj i funkcionalnost se nikad ne mijenja)
- ♦ Npr. europske države se ne pojavljuju samo u atlasu Europe nego i u gospodarstvu Europe, uspostava veze u ravnoj datoteci nemoguća...

2002/2003

Medak: Geodetske baze podataka

31

Baza podataka

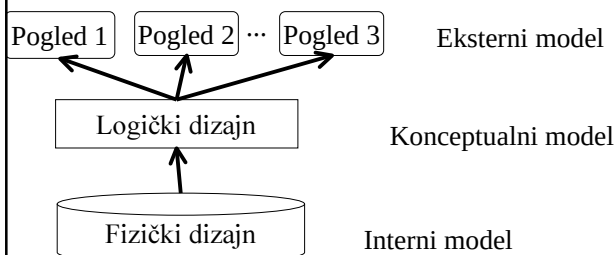
- ♦ Što manje pojedina aplikacija “zna” o internoj navigaciji baze – tim bolje!
- ♦ Općenito, aplikacije bi se trebale baviti pitanjem:
 - ♦ ŠTO je potrebno?dok DBMS odgovara na pitanje
 - ♦ KAKO realizirati zahtjev?

2002/2003

Medak: Geodetske baze podataka

32

Modeli baze podataka (arhitektura tri razine)



2002/2003

Medak: Geodetske baze podataka

33

Konceptualni model

- ♦ **Konceptualni model baze podataka definira logičku strukturu integralne baze podataka.**
- ♦ **Neovisan je** o sustavu za upravljanje bazom podataka, odnosno računalu na kojem će se implementirati.

2002/2003

Medak: Geodetske baze podataka

34

Eksterni model

- ♦ **Eksterni model baze podataka definira strukture podataka za pojedine korisnike i spada u logičke modele baze podataka.**
- ♦ **Struktura podataka za pojedinog korisnika predstavlja derivat konceptualnog modela.**

2002/2003

Medak: Geodetske baze podataka

35

Interni model

- ♦ **Interni model baze podataka spada u fizičke modele** kojim se definira način i mjesto pohranjivanja određenog sadržaja u memoriju.

2002/2003

Medak: Geodetske baze podataka

36

Fizička i logička neovisnost

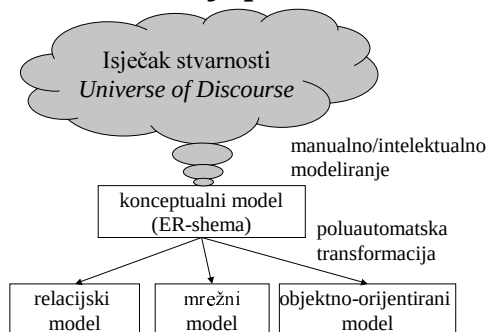
- ♦ osigurana arhitekturom na tri razine
- ♦ Pod **fizičkom nezavisnošću podataka** podrazumijevamo mogućnost **promjene fizičkog opisa baze podataka (fizičke sheme) bez promjene konceptualnog opisa (konceptualne sheme)**.
- ♦ Pod **logičkom nezavisnošću podataka** podrazumijevamo da **mijenjanje konceptualnog opisa (konceptualne sheme) ne zahtijeva promjenu onih eksternih opisa (podshema) na koje se te promjene ne odnose, odnosno odgovarajućih aplikacijskih programa koji koriste navedene podsheme**.

2002/2003

Medak: Geodetske baze podataka

37

Modeliranje podataka



2002/2003

Medak: Geodetske baze podataka

38

Elementi modela baze podataka

- ♦ **Entitet** je individualni primjer elementa stvarnog svijeta ili njegova isječka.
- ♦ **Entitet** je bilo što o čemu želimo **prikupljati i pohranjivati** informacije.
- ♦ **Tip entiteta** je **skup entiteta s istim svojstvima**.
- ♦ **Atribut** je svojstvo koje posjeduje neki entitet.
- ♦ Primjer: najjednostavniji geometrijski element je točka opisana s barem dvije (nerazdvojne) koordinate (y i x)

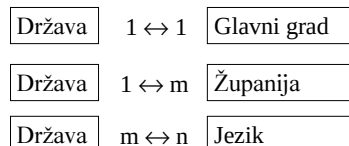
2002/2003

Medak: Geodetske baze podataka

39

Odnos između entiteta

- ♦ ovisan o odabranom modelu podataka
- ♦ odnos ponekad može biti i entitet!
- ♦ binarni odnosi: 1:1, 1: m, m:n

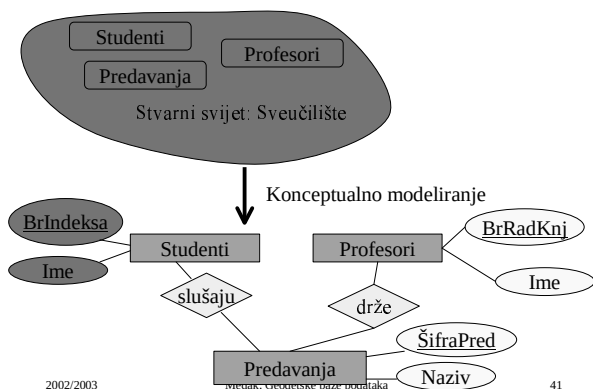


2002/2003

Medak: Geodetske baze podataka

40

Primjer



2002/2003

Medak: Geodetske baze podataka

41

Baze podataka – fizička organizacija i modeli

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

42

Sadržaj

- ♦ Logička i fizička organizacija podataka
- ♦ Pretraživanje podataka
- ♦ Indeksi
- ♦ Primarni i sekundarni ključevi
- ♦ Modeli baza podataka
- ♦ Hijerarhijski model
- ♦ Mrežni model
- ♦ Relacijski model

2002/2003

Medak: Geodetske baze podataka

43

Logička organizacija podataka

- ♦ **Zapis (record)**: niz podataka povezan s jednim entitetom (npr. podaci o jednom studentu)
- ♦ **Polje (field)**: mjesto rezervirano za jedan element zapisa (npr. mjesto za prezime studenta)
 - polje varijabilne duljine (npr. komentar, popis ocjena)
 - polje fiksne duljine (npr. godina upisa na studij)
 - ključ jednoznačno definira zapis (npr. broj indeksa, JMBG)
- ♦ **Datoteka (file)**: niz zapisa obično istog tipa (npr. popis studenata Geodetskog fakulteta)

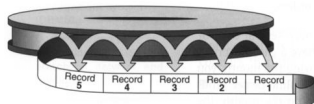
2002/2003

Medak: Geodetske baze podataka

44

Fizička organizacija podataka

- ♦ Datoteke su fizički spremljene na disk tako da blokovi na disku čuvaju zapise.
- ♦ Ako je kapacitet bloka manji od veličine zapisa, zapis će zauzeti dva ili više blokova.
- ♦ Ako je kapacitet bloka veći od veličine zapisa, blok će sadržavati više zapisa (najčešći slučaj).



2002/2003

Medak: Geodetske baze podataka

45

Metode pristupa podacima

- ♦ Otvaranje (open) priprema datoteku za pristup
- ♦ Traženje (find) lociranje bloka na disku koji sadrži zapis koji zadovoljava postavljene uvjete
- ♦ Čitanje (read) kopira zapis u radnu memoriju
- ♦ Brisanje (delete) briše zapis i sprema rezultat na disk
- ♦ Izmjena (modify) mijenja polje u zapisu i sprema rezultat na disk
- ♦ Dodavanje (insert) čita blok u radnu memoriju, dodaje novi zapis i sprema rezultat na disk
- ♦ Zatvaranje (close) prekida pristup datoteci

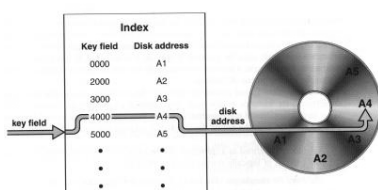
2002/2003

Medak: Geodetske baze podataka

46

Pretraživanje podataka

- ♦ jednostavno i brzo pretraživanje (pronalaženje) - jedna od temeljnih funkcija baze podataka



2002/2003

Medak: Geodetske baze podataka

47

Metode pretraživanja

- ♦ neuređene datoteke i linearno pretraživanje
 - jednostavno dodavanje novih podataka
 - složen direktni pristup (npr. naći ocjene studenta prema imenu: potrebno pretraživanje svih zapisa dok se ne pronađe upravo onaj zapis u kojem se polje slaže s uvjetom),
 - za n podataka prosječno je potrebno učitati $n/2$ zapisa
 - linearno pretraživanje se još naziva **brute-force** (sirova snaga)
- ♦ uređene datoteke i binarno pretraživanje
 - broj zapisa se u svakom koraku prepolovi
 - za n podataka potrebno je maksimalno $\log_2(n)$ koraka, npr. za 1000 studenata poredanih po prezimenu dovoljno je 10 pristupa.

2002/2003

Medak: Geodetske baze podataka

48

Indeksiranje podataka (1)

- ♦ Dobra strana: ubrzava pristup zapisima
- ♦ Loša strana: potreban dodatan prostor za spremanje indeksa
- ♦ Usporedba: knjigu bez indeksa pojmova bi trebalo svaki put pretražiti od prve do zadnje stranice
- ♦ Indeks je uređena datoteka čiji zapisi sadrže dva polja:
 - polje indeksa (sadrži sortirane vrijednosti indeksnog polja)
 - polje pokazivača (sadrži adrese blokova na disku koje imaju vrijednost indeksa)
- ♦ Pretraživanje indeksne datoteke je binarno.

2002/2003

Medak: Geodetske baze podataka

49

Indeksiranje podataka (2)

- ♦ Pronalazak traženog indeksa vodi nas do zapisa koji sadrži pokazivač na zapis s ostalim poljima o traženom entitetu.

DATOTEKA			
BrojInd	Prezime	Ime	Vježbe
5756	Bogner	Hrvoje	+
5876	Gojmerac	Darija	+
...			
5881	Adžić	Ivana	+
5887	Čustović	Aida	+
5907	Gradišer	Lovro	+
5910	Babić	Luka	+
...			
5959	Čanak	Luka	+
5972	Bevanda	Slavica	+

2002/2003

Medak: Geodetske baze podataka

50

Ključevi

- ♦ **Ključ** je minimalni skup atributa potreban za jednoznačnu određivanje zapisa.
- ♦ **Primarni ključ** je s entitetom kojeg opisuje uvijek povezan asocijacijom 1:1 ili m:1.
- ♦ Primjer: ime države može biti napisano na više različitih jezika (Germany, Deutschland, Njemačka), a ipak se odnosi na jedan te isti zapis u bazi podataka.
- ♦ **Sekundarni ključ** je bilo koji atribut koji nije kandidat za primarni ključ.
- ♦ **Strani ključ (foreign key)** je atribut koji služi kao primarni ključ drugog entiteta (npr. kratica države kao atribut grada)

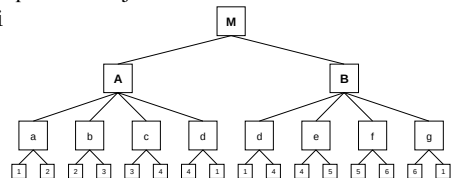
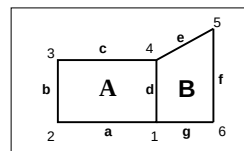
2002/2003

Medak: Geodetske baze podataka

51

Hijerarhijski DBMS

- ♦ “vlasnici” i “članovi”: jedan član može imati samo jednog vlasnika
- ♦ putovi pretraživanja su fiksni



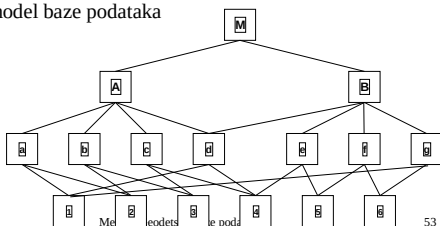
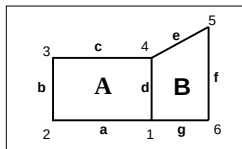
2002/2003

Medak: Geodetske baze podataka

52

Mrežni DBMS

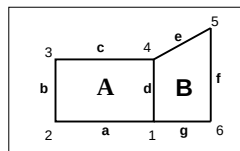
- ♦ opći slučaj hijerarhijskog modela
- ♦ odnosi definirani eksplicitno
- ♦ aplikacija mora poznavati interni model baze podataka



2002/2003

53

Geometrija i relacijski DBMS



Karta A B

Poligoni	A	a	b	c	d
	B	d	e	f	g

Linije

	a	1	2
b	2	3	
c	3	4	
d	4	1	
e	4	5	
f	5	6	
g	6	1	

Točke

	1	x ₁	y ₁
2	x ₂	y ₂	
3	x ₃	y ₃	
4	x ₄	y ₄	
5	x ₅	y ₅	
6	x ₆	y ₆	

Zgrada_ID	Parcela	Vlasnik	Godina	Tip
955	785			
589	294			
610	455			

Parcela_ID	Vlasnik	Povrsina	Adresa
434			
453			
455	Mirko M.	480	Aleja A22

2002/2003

Medak: Geodetske baze podataka

54

Relacijske baze podataka

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

55

Sadržaj

- ♦ Relacijske baze podataka
- ♦ Relacije
- ♦ Posebnosti RDBMS
- ♦ Normalne forme
- ♦ Temelji relacijske algebre

2002/2003

Medak: Geodetske baze podataka

56

Relacijske baza podataka

- ♦ Svi su podaci spremljeni u tabličnom obliku.
- ♦ Svakom tipu podataka odgovara jedna tablica.

Grad_ID	Grad_ime	Grad_geom
SBG	Salzburg	...
MUC	Muenchen	...
GRZ	Graz	...
ZRH	Zuerich	...
...	...	...

Drzava_ID	Drzav_ime	Drzav_geom
GER	Njemačka	...
AUT	Austrija	...
SUI	Švicarska	...
CRO	Hrvatska	...
...	...	...

2002/2003

Medak: Geodetske baze podataka

57

Relacije

- ♦ Temelj relacijskih baza podataka: **povezivanje entiteta različitog tipa**.
- ♦ Odnos grada i države moguće je i implicitno definirati: gradovi kao točke, države kao poligoni, provjeriti da li je točka u poligonu.
- ♦ Druga mogućnost je definicija predmetnog odnosa (relacije) uz pomoć nove tablice

Grad_ID	Drzava_ID
MUC	GER
GRZ	AUT
SBG	AUT
ZRH	SUI
...	...

2002/2003

Medak: Geodetske baze podataka

58

Posebnosti relacijskih DB

- ♦ niti **redosljed zapisa** niti **redosljed polja** u tablici **nije** važan,
- ♦ **položaj (redosljed) identičnih vrijednosti nije** važan (npr. AUT u prethodnoj tablici),
- ♦ relacija se uspostavlja preko **primarnih ključeva** entitetnih tablica

2002/2003

Medak: Geodetske baze podataka

59

Posljedice (1)

- ♦ **interna organizacija** neke tablice potpuno **neovisna** od drugih tablica (npr. sortiranje)
- ♦ korisnikov upit se bavi samo pitanjem **što?**, dok baza podataka (RDBMS) rješava problem **kako?**
- ♦ primjena **neovisna** o unaprijed zadanim putevima pristupa podacima (hijerarhijski i mrežni model)
- ♦ tablice su bliske strukturama koje susrećemo u svakodnevnom životu (računi, kalendar, ...)

2002/2003

Medak: Geodetske baze podataka

60

Posljedice (2)

- ♦ tablice je moguće jednostavno kombinirati, mijenjati i ispitivati: **relacijska algebra** je matematički precizno definirana
- ♦ upitni jezik sadrži **malen broj elemenata i lagano ga je naučiti** (SQL)
- ♦ tablice se lako **proširuju** novim podacima
- ♦ tablice dozvoljavaju **prikaze specifične za pojedinog korisnika** (eksterni model)
- ♦ postoje jednostavna pravila za spremanje podataka bez redundancije (normalne forme)

2002/2003

Medak: Geodetske baze podataka

61

Normalne forme

- ♦ Realizacija relacijskih baza podataka ostavlja veliki prostor za izbor entiteta, atributa i relacija.
- ♦ Ciljevi: što jednostavnije održavanje, konzistentnost, stabilna struktura podataka.
- ♦ Nepotrebno spremljeni podaci – nestabilna struktura podataka – anomalije baza podataka
- ♦ Opasnost: pri izmjeni podataka dolazi do neželjenih sporednih efekata
- ♦ Relacijski model: jednostavna provjera pridržavanja normalnih formi

2002/2003

Medak: Geodetske baze podataka

62

1. normalna forma

- ♦ **Uvjeti:**
 - veza između podataka na **logičkoj**, a ne na fizičkoj razini (adrese na disku),
 - za svaki tip entiteta postoji **jedan primarni ključ**,
 - svako **polje** unutar jednog entiteta ima **jednoznačno ime** koje se ne ponavlja.
- ♦ Prva dva uvjeta su u RDBMS-u neminovna, dok treći uvjet ne dozvoljava sljedeći primjer:
- ♦ Drzava:
Drzava_ime, Glavni_grad, Zupanija1, Zupanija2, ...

2002/2003

Medak: Geodetske baze podataka

63

2. normalna forma

- ♦ **Uvjeti:**
 - moraju biti ispunjeni **uvjeti 1. normalne forme**,
 - svako **polje** unutar jednog entiteta **funkcionalno potpuno ovisi o ključu**, a da pri tome samo nije ključ.
- ♦ Primjer: podizanje cijene nekretnina (parcela) samo u turističkim regijama, (parcele jednoznačno definirane unutar katastarskih općina)
- Cijena:** Opcina_ID, Parcela_ID, Iznos , Regija
- Opcina_u_regiji:** Opcina_ID, Regija

2002/2003

Medak: Geodetske baze podataka

64

3. normalna forma

- ♦ **Uvjeti:**
 - moraju biti ispunjeni **uvjeti 2. normalne forme**,
 - ne postoji tranzitivna zavisnost **nijednog polja o nekom ključu**.
- ♦ Primjer: Vezu županije – države – glavni gradovi “razbili” smo u dvije tablice:
Drzava, Glavni_grad Zupanija, Drzava
Odnos smo mogli prikazati jednom tablicom:
Zupanija_u_drzavi: Zupanija, Drzava, Glavni_grad

2002/2003

Medak: Geodetske baze podataka

65

Normalne forme

- ♦ Postoje i viši stupnjevi normalnih formi.
- ♦ Normalne forme su skup pravila korisnih za modeliranje općih slučajeva baza podataka.
- ♦ Kod kompleksnih tipova podataka ponekad su neophodni kompromisi – odstupanje od idealnog rješenja.
- ♦ Takvi kompleksni tipovi podataka posebno se često javljaju u geoinformatici.
- ♦ Rezultat: objektno-relacijske baze podataka.

2002/2003

Medak: Geodetske baze podataka

66

Temelji relacijske algebre

- ♦ Relacije opisuju i podatke i njihove odnose.
- ♦ Rezultat operacije nad relacijama je ponovo relacija (slično kao i kod brojeva).
- ♦ Relacijska algebra je zatvoreni i konzistentni skup pravila koja su primjenjiva na relacije.

2002/2003

Medak: Geodetske baze podataka

67

Elementarne operacije

- ♦ $\cup$ unija relacija
- ♦ $-$ razlika relacija
- ♦ $\cap$ presjek relacija
- ♦ π projekcija
- ♦ σ selekcija
- ♦ $\times$ Kartezijev produkt relacija
 - ♦ $(A \text{ join})$

2002/2003

Medak: Geodetske baze podataka

68

Relacijska algebra

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

69

Temelji relacijske algebre

- ♦ Relacije opisuju i podatke i njihove odnose.
- ♦ Rezultat operacije nad relacijama je ponovo relacija (slično kao i kod brojeva).
- ♦ Relacijska algebra je zatvoreni i konzistentni skup pravila koja su primjenjiva na relacije.

2002/2003

Medak: Geodetske baze podataka

70

Elementarne operacije

- ♦ $\cup$ unija relacija
- ♦ $-$ razlika relacija
- ♦ $\cap$ presjek relacija
- ♦ π projekcija
- ♦ σ selekcija
- ♦ $\times$ Kartezijev produkt relacija
(∞ join)

2002/2003

Medak: Geodetske baze podataka

71

Unija relacija ($C = A \cup B$)

- ♦ koncept sukladan uniji skupova
- ♦ rezultirajuća relacija sadrži sve zapise iz obje tablice pri čemu se zapisi koji se ponavljaju spremaju samo jedanput
- ♦ UVJET: ulazne relacije (tablice) moraju biti kompatibilne, tj. moraju imati identična polja
- ♦ Primjer: pri udruživanju geodetskih točaka koordinate istih točaka u različitim tablicama mogu biti različite

2002/2003

Medak: Geodetske baze podataka

72

Razlika relacija ($C = A - B$)

- ♦ rezultat su svi zapisi sadržani u relaciji A, a koji nisu u relaciji B
- ♦ UVJET: kompatibilnost relacija
- ♦ Primjer: baza geodetskih točaka, tablica A sadrži sve točke, a tablica B samo izjednačene točke; razlikom A-B dobijemo točke koje tek treba izjednačiti
- ♦ operacija u standardnom programskom jeziku vrlo složena, u RDBMS vrlo jednostavna

2002/2003

Medak: Geodetske baze podataka

73

Presjek relacija ($C = A \cap B$)

- ♦ rezultat sadrži samo zapise koji postoje i u tablici A i u tablici B
- ♦ Primjer:
 - A: tablica zemalja u kojima se govori engleski
 - B: tablica europskih zemalja
 - $C = A \cap B$: tablica europskih zemalja u kojima se govori engleski

2002/2003

Medak: Geodetske baze podataka

74

Projekcija ($B = \pi A$)

- ♦ izbor stupaca neke tablice
- ♦ Primjer: tablica s podacima o geodetskim točkama (n, y, x, H, naziv, tip_tocke, ...), a želimo samo broj točke i koordinate (n, y, x)
- ♦ izbor podataka za eksterni model gdje isti podaci ne moraju ili ne smiju biti vidljivi svim korisnicima baze podataka

2002/2003

Medak: Geodetske baze podataka

75

Selekcija ($\sigma B = A$)

- ♦ izbor zapisa koji udovoljavaju postavljenim uvjetima
- ♦ Primjer: parcela s površinom i vrijednošću

Broj_parcele	Povrsina	Vrijednost
271/2	1215 m ²	€ 40.000.-
271/4	1149 m ²	€ 28.000.-
271/3	1080 m ²	€ 37.000.-

2002/2003

Medak: Geodetske baze podataka

76

Selekcija ($\sigma B = A$)

- ♦ selekcijom se mogu izabrati samo one parcele koje imaju površinu veću od 1100 m² i cijenu manju od 30.000 €
- ♦ rezultat:

Broj_parcele	Povrsina	Vrijednost
271/4	1149 m ²	€ 28.000.-

2002/2003

Medak: Geodetske baze podataka

77

Selekcija ($\sigma B = A$)

- ♦ ponekad nas zanimaju samo pojedini stupci iz zapisa koji zadovoljavaju uvjete pa se izraz selekcija primjenjuje i za kombinaciju selekcije i projekcije
- ♦ uvjeti mogu biti numerički (=, >, <, ...), ali i alfanumerički
- ♦ upotrebljivost relacijskog upitnog jezika uvelike ovisi o mogućnostima selekcije

2002/2003

Medak: Geodetske baze podataka

78

Kartezijski produkt ($C = A \times B$)

- ♦ Rezultat je tablica u kojoj je svaki zapis iz prve tablice kombiniran sa svakim zapisom iz druge tablice
- ♦ Primjer: Aerodromi i gradovi

Aerodrom_ID	Grad
PLS	Zagreb
FRA	Frankfurt
VIE	Beč

2002/2003

Medak: Geodetske baze podataka

79

Kartezijski produkt ($C = A \times B$)

Aerodrom_ID_OD	Grad_OD	Aerodrom_ID_DO	Grad_DO
PLS	Zagreb	PLS	Zagreb
PLS	Zagreb	FRA	Frankfurt
PLS	Zagreb	VIE	Beč
FRA	Frankfurt	PLS	Zagreb
FRA	Frankfurt	FRA	Frankfurt
FRA	Frankfurt	VIE	Beč
VIE	Beč	PLS	Zagreb
VIE	Beč	FRA	Frankfurt
VIE	Beč	VIE	Beč

2002/2003

Medak: Geodetske baze podataka

80

Join ($C = A \bowtie B$)

- ♦ poseban oblik Kartezijskog produkta
- ♦ u svakoj se tablici (relaciji) bira stupac (polje) preko čijih se vrijednosti uspostavlja veza
- ♦ tako se iz svih matematički mogućih kombinacija zapisa biraju samo oni zapisi kod kojih su vrijednosti u pridruženim poljima podudarni

2002/2003

Medak: Geodetske baze podataka

81

Join ($C = A \bowtie B$)

Aerodrom_ID	Grad
PLS	Zagreb
FRA	Frankfurt
VIE	Beč

Grad	Geometrija
Frankfurt	xxx/yyy
Beč	xxx/yyy
Rijeka	xxx/yyy
Zagreb	xxx/yyy

- ♦ veza (join) se ostvaruje preko polja **Grad**, redoslijed zapisa je sporedan (!)
- ♦ na taj se način podaci mogu geokodirati, tj. povezati njihov prostorni opis s drugim atributima (npr. prostorne jedinice sa statističkim bazama podataka, prijavom boravka, vlasništvom itd.)

2002/2003

Medak: Geodetske baze podataka

82

Zaključak

- ♦ Opisane operacije dolaze uvijek u kombinaciji
 1. iz više tablica **selektiramo** samo one zapise koji udovoljavaju uvjetima
 2. napravimo pridruživanje (**join**)
 3. eliminiramo **projekcijom** nepotrebna polja
- ♦ Na taj se način iz vrlo malog broja osnovnih operacija relacijske algebre može izvesti veliki broj kombinacija za obradu i analizu podataka

2002/2003

Medak: Geodetske baze podataka

83

Sljedeće predavanje

- ♦ Opisanu relacijsku algebru je moguće jednostavno prevesti u jezik za pristup i manipulaciju podacima pohranjenih u relacijske baze podataka.
- ♦ Primjer za takav jezik je **Structured Query Language (SQL)**

2002/2003

Medak: Geodetske baze podataka

84

Relacijski upitni jezici

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

85

Sadržaj

- ♦ Povijest relacijskih upitnih jezika
- ♦ Tipovi podataka
- ♦ Naredbe za definiranje i promjenu sheme
- ♦ Elementarne operacije
- ♦ Višetablični upiti
- ♦ Agregiranje i grupiranje

2002/2003

Medak: Geodetske baze podataka

86

Povijesni pregled

- ♦ 1970-ih: IBM – DBMS System R – upitni jezik SEQUEL (Structured English Query Language)
- ♦ 1980-ih: SEQUEL preimenovan u SQL, komercijalni sustav SQL/DS, pojava novih relacijskih DBMS-a (Oracle, Informix), potreba za standardizacijom upitnog jezika
- ♦ 1986: prvi SQL-standard ANSI-komisije
- ♦ 1989: prva revizija standarda
- ♦ 1992: proširenje na **SQL-92 (SQL 2)**
- ♦ 2002: daljnje proširenje na SQL 3 još u pripremnj fazi

2002/2003

Medak: Geodetske baze podataka

87

SQL-92

- ♦ upitni jezik zasnovan na relacijskoj algebri
- ♦ teorijske postavke pojednostavljene zbog jednostavnije uporabe i veće brzine
- ♦ deklarativni jezik (što? a ne kako?)
- ♦ jezik nije samo “upitni”, nego služi i za definiciju i manipulaciju podacima
- ♦ SQL ne obrađuje relacije u matematičkom smislu, nego tablice
- ♦ zapis i polje postaju redak i stupac
- ♦ osim manipulacije tablicama, SQL omogućuje definiciju integriteta podataka, prava pristupa i kontrolu transakcija

2002/2003

Medak: Geodetske baze podataka

88

Tipovi podataka

- ♦ Tri temeljna tipa podataka:
 - brojevi (numeric)
 - znakovni nizovi (character)
 - datum (date)
- ♦ Brojevi mogu biti cjelobrojni (integer) i decimalni (float)
- ♦ Znakovni nizovi mogu biti fiksne duljine (character) ili varijabilne duljine (varchar)
- ♦ Dodatno blob (binary large object) ili raw za binarne tipove podataka (npr. u CAD-primjenama)

2002/2003

Medak: Geodetske baze podataka

89

Definiranje sheme baze podataka

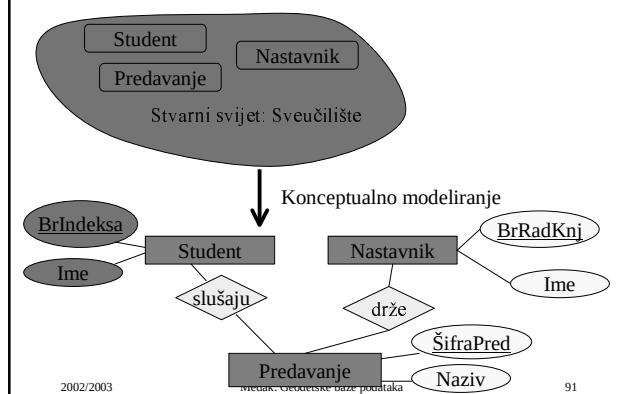
- ♦ nova tablica:
create table Nastavnik
(BrRadKnj **integer not null**
Ime **varchar(10) not null**
Titula **character(4));**
- ♦ popis atributa i tipa podatka za svaki atribut
- ♦ “not null” primjer definicije uvjeta integriteta podataka (nije moguće zadati nastavnika bez imena)

2002/2003

Medak: Geodetske baze podataka

90

Primjer



2002/2003

Medak: Geodetske baze podataka

91

Izmjena sheme baze podataka

- ♦ pretpostavimo da u tablicu želimo dodati broj kancelarije za svakog nastavnika:
alter table Nastavnik
add column Soba integer;
- ♦ pretpostavimo da je 10 znakova bilo pogrešna procjena za duljinu imena nastavnika:
alter table Nastavnik
modify column Ime varchar(30);
- ♦ posljednja naredba ne isključuje ranije zadani uvjet integriteta (not null)

2002/2003

Medak: Geodetske baze podataka

92

Elementarne operacije

- ♦ dodavanje redaka (punjenje podacima):
insert into Nastavnik
values (21366, 'Medak', 'doc', 77);
- ♦ izbor podataka na temelju kriterija:
select BrRadKnj, Ime
from Nastavnik
where Titula = 'doc';
- ♦ **select** određuje atribute koji će se pojaviti u rezultatu
- ♦ **from** određuje iz koje se tablice traže podaci
- ♦ **where** određuje uvjet kojeg zapis mora udovoljiti

2002/2003

Medak: Geodetske baze podataka

93

Elementarne operacije

- ♦ snaga SQL-a je u bliskosti naredbi s razgovornim jezikom; najčešće je dovoljan prijevod naredbe s engleskog ("izaberi broj i ime nastavnika čija je titula docent")
- ♦ rezultat prethodnog upita bila je lista bez utvrđenog redoslijeda
- ♦ SQL omogućuje sortiranje prema bilo kojem atributu:
select BrRadKnj, Ime, Titula
from Nastavnik
order by Titula desc, Name asc;
- ♦ **ascending** = a-z, 1-9, **descending** = z-a, 9-1
- ♦ eliminacija duplikata:
select distinct Titula
from Nastavnik;

2002/2003

Medak: Geodetske baze podataka

94

Višetablični upiti

- ♦ Povezivanje tablica – tko predaje Geodetske baze podataka:
select Ime, Titula
from Nastavnik, Predavanje
where BrRadKnj = drzi and Naziv="GBP";
- ♦ prijevod: "Izaberi ime i titulu iz kombinacije nastavnika i predavanja, pri čemu je vrijednost stupca **Drži** jednaka vrijednosti stupca **BrRadKnj** i naziv predavanja je GBP."
- ♦ Obrada upita se odvija u tri faze:
 1. **Kartezijev produkt** tablica Nastavnik i Predavanje
 2. za svaki redak produkta se ispituje **uvjet** postavljen pod **where**
 3. **projekcijom** se biraju pod **select** zadani stupci (atributi)

2002/2003

Medak: Geodetske baze podataka

95

Višetablični upiti - join

- ♦ Povezivanje tablica – koji studenti slušaju koja predavanja:
select Ime, Naziv
from Student, Sluša, Predavanje
where Student.BrInd = Slusa.BrInd and
Sluša.SifraPred = Predavanje.SifraPred ;
- ♦ elegantniji zapis uz pomoć varijabli:
select s.Ime, p.Naziv
from Student s, Sluša h, Predavanje p
where s.BrInd = h.BrInd and
h.SifraPred = p.SifraPred ;

2002/2003

Medak: Geodetske baze podataka

96

Agregiranje

- ♦ Agregativne funkcije izvode operaciju na skupu zapisa i svode skup vrijednosti na jednu vrijednost:
- ♦ Primjeri:
 - avg (prosjeck skupa brojeva)
 - max, min (najveći, odnosno najmanji element skupa)
 - sum (zbroy svih vrijednosti)
 - count (broj vrijednosti)
- ♦ Srednji broj semestara svih studenata je tako:
- ♦ **select avg (Semestar)**
from Student ;

2002/2003

Medak: Geodetske baze podataka

97

Grupiranje ...

- ♦ Uz pomoć naredbe **group by** moguće je npr. izračunati opterećenje nastavnika:
- ♦ **select Drzi, sum (Satnica)**
from Predavanje
group by Drzi ;
- ♦ Objašnjenje: u rezultatu su svi zapisi u tablici koje imaju jednaku vrijednost za atribut Drzi grupirani u jedan, te je za svaki izračunat zbroj satnica.
- ♦ Uz pomoć naredbe **having** moguće je postaviti dodatan uvjet kojeg mora ispuniti grupa definirana s **group by**

2002/2003

Medak: Geodetske baze podataka

98

...Grupiranje

- ♦ Primjer: grupirati predavanja prema nastavnicima i za svaku grupu izračunati prosjek satnice, te uzeti u obzir samo one nastavnike koji imaju više od tri sata tjedno :
- ♦ **select Drzi, sum (Satnica)**
from Predavanje
group by Drzi
having avg (Satnica) > 3;

2002/2003

Medak: Geodetske baze podataka

99

Zaključak

- ♦ SQL predstavlja standardizirani oblik komunikacije korisnika s relacijskim sustavom za upravljanje bazom podataka (RDBMS)
- ♦ Naredbe napisane u SQL-u su daleko moćnije od grafičkih korisničkih sučelja za upisivanje upita prema primjeru (Query By Example)
- ♦ SQL-92 je podržan od velike većine komercijalnih i nekomercijalnih proizvoda
- ♦ SQL 3 će predstavljati bitno proširenje standarda novim funkcijama koje proširuju relacijski model u smjeru objektno-relacijskog modela

2002/2003

Medak: Geodetske baze podataka

100

Sljedeće predavanje

- ♦ Relacijski model se zasniva na teorijski konzistentnom skupu pravila relacijske algebre i idealan je za podatke tablične naravi (trgovina, bankarstvo, proračuni, personalni podaci)
- ♦ Podaci složenije strukture (CAD, prostorni podaci, dinamički procesi) teško se mogu modelirati u relacijskom modelu.
- ♦ Zbog toga u posljednje vrijeme dobivaju na značenju objektno-orijentirane metode, programski jezici i baze podataka.
- ♦ Za razumijevanje OO-tehnologije važno je razumjeti osnovna obilježja: nasljeđivanje (generalizacija i specijalizacija), višeooblčnost i prikrivanje (encapsulation)

2002/2003

Medak: Geodetske baze podataka

101

Transakcije, konzistentnost i sigurnost baza podataka

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

102

Transakcije

- ♦ **Transakcija** je skup operacija koji prevodi bazu podataka iz jednog konzistentnog stanja u drugo konzistentno stanje.
- ♦ Primjer: plaćanje s jednog računa uključuje barem dvije operacije: umanjiti iznos na jednom računu i uvećati iznos na drugom. Ako je prva operacija izvršena, a druga ne, novac bi mogao nestati.
- ♦ DBMS mora zadovoljiti četiri tzv. ACID uvjeta da bi omogućio pravilan rad s transakcijama

2002/2003

Medak: Geodetske baze podataka

103

ACID

- ♦ **atomicity** – atomičnost transakcije: transakciju nije moguće dijeliti na elementarnije transakcije. Ili su izvršene sve operacije unutar transakcije ili nijedna.
- ♦ **consistency** – konzistentnost: baza podataka mora uvijek zadržati konzistentno stanje
- ♦ **isolation** – izolacija: dok transakcija nije izvršena, njezine operacije ne utječu na druge, paralelne operacije
- ♦ **durability** – trajnost: nakon što je transakcija izvršena, promjene ostaju trajno pohranjene, čak i u slučaju “pada” sustava.

2002/2003

Medak: Geodetske baze podataka

104

Konzistentnost baza podataka

- ♦ Za pojedine attribute u tablicama moguće je u SQL-u definirati točno određene uvjete (constraints) koje vrijednosti atributa moraju zadovoljiti
- ♦ Primjer prvi: NOT NULL onemogućuje prazno polje
- ♦ Primjer drugi:

```
CREATE TABLE Posao (  
    ...  
    min_placa NUMERIC NOT NULL,  
    max_placa NUMERIC NOT NULL,  
    ...  
    CHECK (min_placa < max_placa);
```

2002/2003

Medak: Geodetske baze podataka

105

Sigurnost podataka ...

- ♦ **Fizička sigurnost:** više kopija na različitim mjestima (u različitim zgradama), naredba SHADOW kreira istovjetnu kopiju
- ♦ **Logička sigurnost:** osigurati autorizirani pristup bazi podataka (korisničko ime + lozinka)
- ♦ U SQL-u moguće definirati autorizaciju pristupa svakog korisnika za pravo izvođenja pretraživanja odabranih tablica:
GRANT SELECT ON Student TO medak;

2002/2003

Medak: Geodetske baze podataka

106

...Sigurnost podataka

- ♦ Za naredbe koje mijenjaju vrijednosti u tablicama (INSERT, UPDATE, REFERENCE) moguće je izabrati pojedine stupce za koje se dozvoljava pristup:
GRANT UPDATE (Datum_ispita, Ocjena)
ON Student
TO medak;
- ♦ Sve je dozvole pristupa moguće povući naredbom **REVOKE**

2002/2003

Medak: Geodetske baze podataka

107

Uvod u objektno-orijentirane metode

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

108

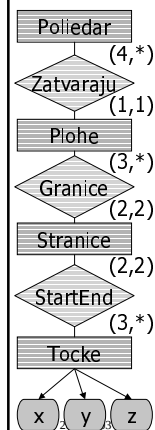
Sadržaj

- ♦ Nedostaci relacijskog modela
- ♦ Prednosti objektno-orientiranog modela
- ♦ Obilježja objekata
- ♦ Nasljeđivanje
- ♦ Komercijalni produkti
- ♦ Objektno-relacijske baze podataka

2002/2003

Medak: Geodetske baze podataka

109



Poliedar				
PolyID	Masa	Materijal	...	
cubo#5	25.765	Zeljezo	...	
tetra#7	37.985	Staklo	...	
...				
Ploha			...	
PlohaID	PolyID	Površina		
f1	cubo#5	...		
f2	cubo#2	...		
...	...	...		
f6	cubo#5	...		
...				
Stranice				
StranicaID	F1	F2	P1	P2
k1	f1	f4	p1	p4
k2	f1	f2	p2	p3
...				
Tocke				
TockaID	X	Y	Z	
p1	0.0	0.0	0.0	
p2	1.0	0.0	0.0	
geodetske baze podataka				

110

Medak: Geodetske baze podataka

110

Nedostaci relacijskog modela

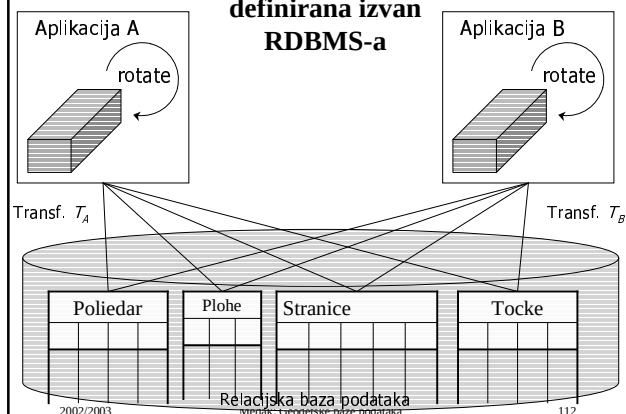
- ♦ Segmentiranje: jedan objekt se obično segmentira na više relacija, čijim se pridruživanjem (join) ponovno uspostavlja cjelina
- ♦ Umjetni ključevi: nužno uvođenje dodatnih atributa kao ključeva (xxxID)
- ♦ Nemoguće je definirati ponašanje (behavior), npr. promjena mjerila, rotacija i translacija, volumen, masa nisu definirani u logičkom modelu, nego ovise o konkretnoj aplikaciji.
- ♦ Eksterno programersko sučelje: nužno povezivanje relacijskog jezika (SQL) s nekim drugim programskim jezikom (C, Java), što vodi dupliciranju rješenja istih operacija (npr. rotacija poliedra)

2002/2003

Medak: Geodetske baze podataka

111

Operacija rotate definirana izvan RDBMS-a

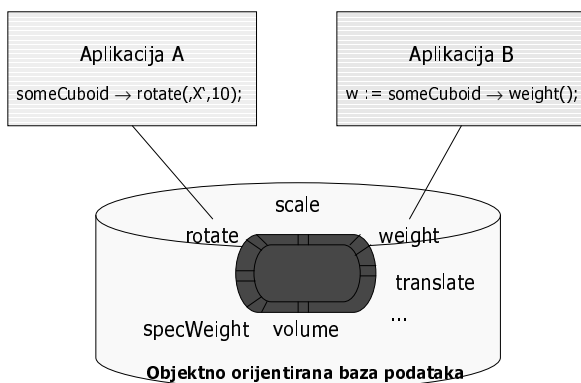


2002/2003

Medak: Geodetske baze podataka

112

Prednosti objektno-orientiranog modela



2002/2003

Medak: Geodetske baze podataka

113

Prednosti OO-modela

- ♦ Prikrivanje informacija (encapsulation)
- ♦ Ponovna upotrebljivost (reusability)
- ♦ Operacije realizirane direktno u jeziku objektnog modela
- ♦ Nasljeđivanje (inheritance)
- ♦ Višeobličnost (polymorphism)

2002/2003

Medak: Geodetske baze podataka

114

Obilježja objekata

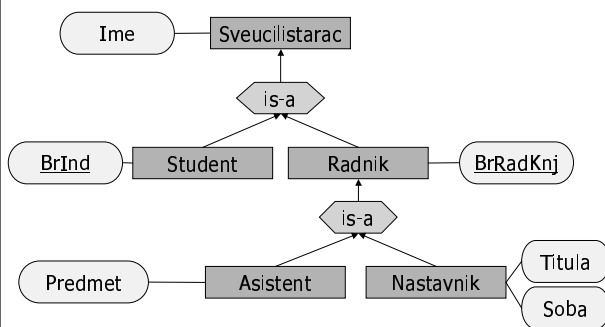
- ♦ Relacijski model: entiteti u obliku zapisa, zapisi se sastoje iz konačnog broja polja čije su vrijednosti elementarne (ne mogu se sastojati od složenih struktura) i nepromjenjive (npr. broj 2)
- ♦ U OO-modelu objekt ima tri dijela:
 - **Identitet**: svaki objekt ima jedinstveni identifikator, koji se za vrijeme životnog ciklusa objekta ne mijenja
 - **Tip (Klasa)**: određuje strukturu i ponašanje objekta. Konkretni objekti nastaju kao instance tipa objekta.
 - **Vrijednost (Stanje)**: u nekom trenutku životnog ciklusa objekt ima određeno stanje dato vrijednostima atributa i postojećih veza s drugim objektima

Večernji list, 27. XI 2002.

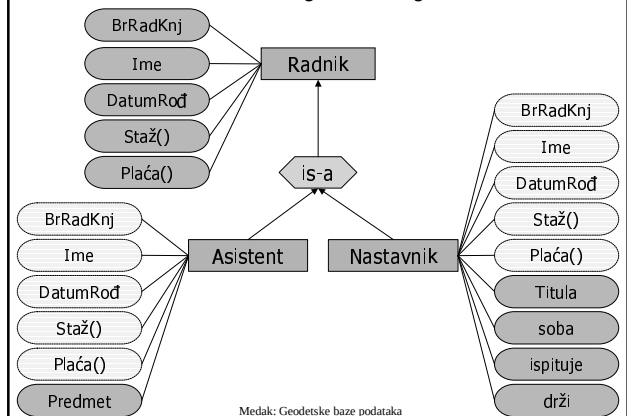
UKRALI IDENTITET 30.000 LJUDI

- ♦ NEW YORK – Državno tužiteljstvo u New Yorku optužilo je jučer trojicu ljudi za “krađu identiteta” oko 30 tisuća osoba kojima je nanescena šteta od gotovo 3 milijuna dolara. Trojka je tijekom tri godine krala podatke iz **baza podataka** kreditnog odjela tvrtke Ford, kreditne agencije Experian i još nekih tvrtki. Podatke o osobama preprodavali su za 60 dolara. “Kupci” novog identiteta mogli su, lažno se predstavljajući, nesmetano tražiti duplikate kreditnih ili bankovnih kartica, podizati kredite. Državni tužitelj na Manhattanu kaže da su na takav način pribavili oko 15 tisuća kreditnih kartica, te da je to **najveća krađa identiteta u povijesti SAD-a**.

Nasljeđivanje



Nasljeđivanje atributa



OODBMS na tržištu

- | | |
|------------------|-----------|
| ♦ GemStone | ♦ Ontos |
| ♦ Illustra | ♦ OpenOBD |
| ♦ Itasca | ♦ POET |
| ♦ MATISSE | ♦ UniSQL |
| ♦ O ₂ | ♦ Statice |
| ♦ Objectivity/DB | ♦ Versant |
| ♦ ObjectStore | |
| ♦ Ontos | |

Objektno-relacijske baze podataka

Funkcionalno proširenje relacijskog modela:

- ♦ **Skupovni atributi**: dozvoljen atribut koji sadrži više istovrsnih podataka (ocjene studenta)
- ♦ **Definiranje tipova** od strane korisnika
- ♦ **Reference** na pojedine zapise (umjesto stranih ključeva), reference za predavanja za svakog studenta
- ♦ **Identitet objekata**
- ♦ **Nasljeđivanje**
- ♦ **Operacije** (u SQL se podacima ne mogu pridružiti operacije)

Budućnost baza podataka

- ♦ Objektno-relacijski koncepti prisutni kod većine najnovijih verzija komercijalnih (Oracle 9i, Informix 9.2), i nekomercijalnih baza podataka (Postgresql 7.2)
- ♦ SQL3 će inkorporirati objektno-relacijske koncepte u novi standard
- ♦ “Čisti” objektno-orientirani modeli imaju perspektivu u specijalnim primjenama

2002/2003

Medak: Geodetske baze podataka

121

Objektno-relacijske baze podataka

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

122

Sadržaj

- ♦ Potreba za proširenjem relacijskog modela
- ♦ Geodetske primjene: prostorni podaci
- ♦ Međunarodni standardi
- ♦ Primjeri (tipovi podataka i operacije):
 - Oracle Spatial
 - PostgreSQL
- ♦ Zaključak

2002/2003

Medak: Geodetske baze podataka

123

Objektno-relacijske baze podataka

Funkcionalno proširenje relacijskog modela:

- ♦ **Skupovni atributi:** dozvoljen atribut koji sadrži više istovrsnih podataka (ocjene studenta)
- ♦ **Definiranje tipova** od strane korisnika
- ♦ **Reference** na pojedine zapise (umjesto stranih ključeva), reference za predavanja za svakog studenta
- ♦ **Identitet objekata**
- ♦ **Nasljeđivanje**
- ♦ **Operacije** (u SQL se podacima ne mogu pridružiti operacije)

2002/2003

Medak: Geodetske baze podataka

124

Geodetske primjene: proširenja za prostorne podatke

- ♦ Operacije nad prostornim podacima u SQL-92 vrlo složene (površina poligona)
- ♦ Tipovi prostornih podataka – problem standarda (Open Geodata Interoperability Specification)
- ♦ Operacije nad prostornim podacima – problem kompleksnosti (ažuriranje baze) i efikasnosti (brzina obrade upita)
- ♦ Primjeri:
 - Oracle Spatial (komercijalna DB)
 - PostgreSQL (open source DB)

2002/2003

Medak: Geodetske baze podataka

125

Standardi za prostorne podatke

- ♦ Open GIS: *Open Geodata Interoperability Specification* (Otvorena specifikacija za interoperabilnost geopodataka),
 - UML (*Unified Modeling Language*), Java
- ♦ OGC: Open GIS Consortium
- ♦ ISO: Međunarodna organizacija za standardizaciju



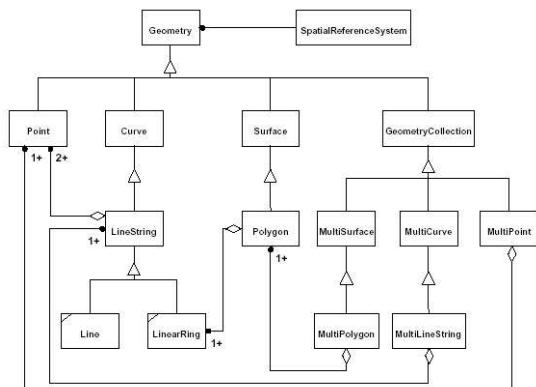
- ♦ ISO/TC 211
Geographic information/
Geomatics

2002/2003

Medak: Geodetske baze podataka

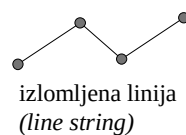
126

OGC - hijerarhija klasa u UML-u

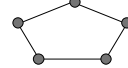


Oracle Spatial: model prostornih podataka

♦ tipovi prostornih podataka:



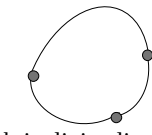
izlomljena linija
(line string)



poligon



krivulja
(arc string)



krivuljni poligon

2002/2003

Medak: Geodetske baze podataka

128

Oracle Spatial: operacije nad prostornim podacima

- ♦ operacije SDO_RELATE, SDO_FILTER i SDO_WITHIN_DISTANCE provjeravaju topološke odnose između dva prostorna objekta:
- contains (sadrži)
 - coveredby (prekriven)
 - covers (prekriva)
 - disjoint (razdvojen)
 - equal (jednak)
 - inside (unutar)
 - overlap (prekriva)
 - touch (dodiruje)

2002/2003

Medak: Geodetske baze podataka

129

Oracle Spatial: Primjer upita

♦ Naći županije koje sijeku okvir prozora na ekranu

```

select Zupanija.Ime
from Zupanija, Prozori
where MDSYS.SDO_RELATE (Zupanija.geometry,
Prozori.geometry, 'mask = OVERLAP') =
'TRUE';
    
```

2002/2003

Medak: Geodetske baze podataka

130

PostgreSQL: geometrijski tipovi podataka

- | | | |
|-------------|--|-----------------------|
| ♦ point | (x,y) | točka |
| ♦ line | ((x ₁ , y ₁), (x ₂ , y ₂)) | pravac |
| ♦ lseg | ((x ₁ , y ₁), (x ₂ , y ₂)) | dužina |
| ♦ box | ((x ₁ , y ₁), (x ₂ , y ₂)) | pravokutnik |
| ♦ path | ((x ₁ , y ₁), ...) | zatvorena poliliniija |
| ♦ open path | [(x ₁ , y ₁), ...] | otvorena poliliniija |
| ♦ polygon | ((x ₁ , y ₁), ...) | poligon |
| ♦ circle | < (x, y), r > | krug |

2002/2003

Medak: Geodetske baze podataka

131

PostgreSQL: operacije nad prostornim podacima

- | | | |
|----|---------------------------|---|
| + | translacija pravokutnika | $box \times point \rightarrow box$ |
| * | promjena mjerila | $box \times integer \rightarrow box$ |
| # | presjek dužina | $lseg \times lseg \rightarrow point$ |
| # | broj vrhova poligona | $polygon \rightarrow integer$ |
| && | preklapanje poligona | $polygon \times polygon \rightarrow bool$ |
| << | da li je s lijeve strane? | $geom \times geom \rightarrow bool$ |
| >> | da li je s desne strane | $geom \times geom \rightarrow bool$ |
| <^ | da li je ispod? | $geom \times geom \rightarrow bool$ |
| >^ | da li je iznad? | $geom \times geom \rightarrow bool$ |
| ?# | presjek | $path \times geom \rightarrow bool$ |

2002/2003

Medak: Geodetske baze podataka

132

PostgreSQL: Primjer upita

- ♦ Naći ceste koje prolaze općinom X

select Cesta.naziv

from Opcine, Ceste

where Opcine.opcinaID = 'X'

And Opcine.geometry ?# Ceste.geometry;

2002/2003

Medak: Geodetske baze podataka

133

Zaključak

- ♦ Objektno-relacijski koncepti prisutni kod većine najnovijih verzija komercijalnih (Oracle 9i, Informix 9.2), i nekomercijalnih baza podataka (Postgresql 7.3)
- ♦ Za geodetske primjene standardizacija prostornih podataka ključni preduvjet efikasne primjene baza podataka u svakodnevnom radu.
- ♦ Napori ISO/TC211 i OGC-a vrijedni pažnje geodetskih stručnjaka i prije samog donošenja standarda.

2002/2003

Medak: Geodetske baze podataka

134

Deduktivne baze podataka

doc. dr. sc. Damir Medak
dmedak@geof.hr
www.geof.hr/~dmedak

2002/2003

Medak: Geodetske baze podataka

135

Sadržaj

- ♦ Baze znanja i logika
- ♦ Činjenice i pravila
- ♦ Deduktivne baze podataka – logičke kombinacije relacija
- ♦ SWI-Prolog
- ♦ Primjer: obiteljsko stablo u Prologu
- ♦ Zaključak

2002/2003

Medak: Geodetske baze podataka

136

Baze znanja i logika

- ♦ Znanje nije jednostavno definirati.
- ♦ Znanje predstavlja sljedeću kariku u nizu:
podaci – informacije – znanje
- ♦ Stara definicija: “istinito vjerovanje koje je moguće opravdati”, tj. znanje je moguće potkrijepiti argumentom (npr. Zagreb je glavni grad Republike Hrvatske.)
- ♦ Razvoj formalne logike omogućio računalni pristup problematici izvođenja zaključaka iz poznatih činjenica.

2002/2003

Medak: Geodetske baze podataka

137

Činjenice i pravila

- ♦ **Činjenica** (eng. *fact*) u formalnoj logici je atomarna izjava koja se ne može dijeliti, tj. prikazati kao logička kombinacija jednostavnijih izjava
- ♦ Činjenice su specifične:
 - “Sokrat je smrtnik” je činjenica
 - “Svi ljudi su smrtnici” nije činjenica
- ♦ **Pravilo** (eng. *rule*) je uvjetna izjava koja uključuje kvantificiranje: $\forall$ “za svaki”, $\exists$ “postoji”
- ♦ Logički simboli: $\Rightarrow$ “implikacija”,
 $\wedge$ “i”, $\vee$ “ili”, $\neg$ “ne”

2002/2003

Medak: Geodetske baze podataka

138

“Tradicionalni” pristup

- ♦ Činjenice su tip znanja kojim se bave tradicionalne baze podataka (npr. relacija GLAVNI_GRAD bi imala polja GRAD i DRZAVA)
- ♦ Tradicionalne baze podataka se bave jednostavnim činjenicama, ali je vrlo teško definirati **pravila** tipa:
“Svi letovi za Dubrovnik idu preko Zagreba”
- ♦ Kod geodetskih baza prostornih podatak velik broj informacija je implicitno pohranjen

2002/2003

Medak: Geodetske baze podataka

139

Deduktivne baze podataka

- ♦ Dedukcija: izvođenje zaključaka iz premisa u logičkim sustavima
- ♦ Primjer:
 - ♦ Svi ljudi su smrtnici,
 - ♦ Sokrat je čovjek
 - ⇒ Sokrat je smrtnik.
- ♦ Zračni promet:
A: $\forall x \forall y$ (jedno_presjedanje) $\Leftrightarrow \exists z . (let(x,z) \wedge let(z,y))$

2002/2003

Medak: Geodetske baze podataka

140

Usporedba RDB i DDB pri računanju upita

- | | |
|---|---|
| <ul style="list-style-type: none">♦ Relacijski model<ul style="list-style-type: none">– SQL manipulira složenom strukturom podataka, čijim pretraživanjem dobivamo odgovor– jezik za manipulaciju je odvojen od podataka | <ul style="list-style-type: none">♦ Deduktivni model<ul style="list-style-type: none">– izračunava vrijednost logičkih izraza (istina ili neistina) ili traži vrijednosti za koje je logički izraz istinit– činjenice i pravila su dio jezika, tj. nema razlike između podataka i jezika za njihovu manipulaciju |
|---|---|

2002/2003

Medak: Geodetske baze podataka

141

Problemi

- ♦ osiguravanje trajnosti zapisa u deduktivnu bazu
- ♦ konstrukcija logičkog jezika koji će imati mogućnosti prelaziti okvire radne memorije

2002/2003

Medak: Geodetske baze podataka

142

Zaključak

- ♦ Deduktivne baze podataka obećavajuća tehnologija za baze podataka utemeljene na znanju
- ♦ Kompaktan format zapisa pravila koja zamjenjuju tisuće redaka
- ♦ Jedinstvenost podataka i jezika za manipulaciju
- ♦ Problemi s trajnošću podataka

2002/2003

Medak: Geodetske baze podataka

143

Prolog

- ♦ logički sustav, odnedavno Open Source
- ♦ www.swi-prolog.org
- ♦ iznimno popularan alat u sveučilišnoj zajednici
- ♦ široki spektar primjena
- ♦ Primjer: obiteljsko stablo

2002/2003

Medak: Geodetske baze podataka

144