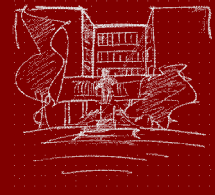


[P273] 4.2  
Пројектовање база података



Саша Малков  
Универзитет у Београду  
Математички факултет  
2023/2024

[P273]  
Пројектовање база података  
Саша Малков



Тема 7.1

Релациони модел података  
(структурни део)

Релациони модел података је већ изучаван током  
студија,  
Релациони модел података



у оквиру предмета *Релационе базе података (И)*  
или *Увод у релационе базе података (МР)*

- `assert( students.foreach(s: s.isRelationalModelExpert()) )`
- Ради комплетности ћемо ипак проћи кроз основне елементе модела
  - на начин који је донекле прилагођен контексту
- На испиту се очекује темељно познавање релационог модела

Значај релационог модела података



- Релациони модел података је из више разлога најзначајнији модел података који је до сада развијен:
  - Потпуна формализација
    - основ за формално изучавање и унапређивање
    - непротивречан
    - недвосмислено прецизан у примени
    - погодан за аутоматско обрађивање и оптимизовање
    - и многе друге битне последице
  - Јединствена репрезентација података и метаподатака
    - све врсте информација се представљају на исти начин – релацијама
    - све врсте метаподатака се представљају на исти начин – релацијама
  - У високој мери стандардизован
    - релативно висока преносивост
  - Веома заступљен у пракси
    - велики број различитих система и разноврсних примена
    - изузетно добро проверен и документован

## Кратко подсећање на историју



- Edgar F. Codd, **A relational model of data for large shared data banks**, *Comm.ACM*, 13(6), 1970.
- *“Релациони модел података обезбеђује представљање података само њиховом природном структуром, тј. без увођења било каквих додатних претпоставки о структури ради машинске репрезентације. У складу са тиме, представља основу за језике вишег нивоа која ће омогућити максималну независност између програма на једној и машинске репрезентације на другој страни.”*

## Кратко подсећање на историју (2)



- *IBM System R*
  - први прототип, започет 1974.
- *Oracle*
  - прва комерцијална имплементација, 1979.
- Најраспрострањенији РСУБП (*DB-Engines, Nov. 2024*)
  - *Oracle*
  - *MySQL*
  - *MS SQL Server*
  - *PostgreSQL*
  - *IBM DB2*
  - *Snowflake*
  - *SQLite*
  - *MS Access*
  - *MariaDB*

## Делови релационог модела



- Релациони модел чине:
  - Структурни део релационог модела
    - начин моделирања података
  - Манипулативни део релационог модела
    - начин руковања моделираном подацима
  - Интегритетни део релационог модела
    - начин обезбеђивања ваљаности података

## Структурни део релационог модела



- Основна идеја је да се *све* моделира релацијама
  - *све* значи и *ентитети* и *односи*
    - појмови ентитета и односа не могу до краја да се раздвоје, зато што неке ствари из једног угла могу да изгледају као односи а из другог као ентитети, али је уобичајено да се посматрају одвојено
    - *ентитети* су сви различити постојећи елементи (*“објекти”*) система које моделирамо базом података
    - *односи* су значајни међусобни односи двају или више ентитета (или других односа)
  - најпре ћемо да размотримо моделирање ентитета
    - то је често и једина ствар која се теоријски разматра
    - зато што сваки однос може да се представи као ентитет

## Математичке основе



- Релациони модел је у *истиносној* формално математички заснован
- Овде ћемо навести оквире те формализације
- За више информација погледати
  - књигу Гордана Павловић-Лажетић, *Увод у релационе базе података, 2. изд. Математички факултет, 1999.*
  - као и друге изворе, наведене на крају презентације

## Модел релације



- Свакој математичкој релацији одговара тачно један скуп објеката који задовољавају релацију
- Тај скуп се често употребљава као математички модел релације
- Ако је  $\rho$   $n$ -арна релација и њен домен означен са  $Dom(\rho)$
- Онда је један модел релације  $\rho$  скуп:
 
$$model(\rho) = \{ (a_1, a_2, \dots, a_n) \mid (a_1, a_2, \dots, a_n) \in Dom(\rho) \wedge \rho(a_1, a_2, \dots, a_n) \}$$
- Приметимо да важи:
 
$$model(\rho) \subset Dom(\rho)$$

## Представљање релација скуповима



- За све  $n$ -торке из  $Dom(\rho)$  важи:
 
$$\rho(a_1, a_2, \dots, a_n) \Leftrightarrow (a_1, a_2, \dots, a_n) \in model(\rho)$$
- Због наведене еквиваленције често се уместо ознаке  $model(\rho)$  употребљава и само ознака релације  $\rho$  у истом контексту:
 
$$\rho = \{ (a_1, a_2, \dots, a_n) \mid (a_1, a_2, \dots, a_n) \in Dom(\rho) \wedge \rho(a_1, a_2, \dots, a_n) \}$$
- Т.ј. релације често представљамо управо помоћу њиховог скуповног модела, без могућности да то изазове неспоразуме

## Пример



- Нека имамо
  - црвену, плаву, белу и зелену кутију
  - и у њима лопте, коцке и плочице
- Бинарна релација  $kutijaSadrzi(K,P)$  је задовољена ако кутија  $K$  садржи  $P$
- Њен домен је:
 
$$Dom(kutijaSadrzi) = \{ crvena, plava, bela, zelena \} \times \{ lopte, kocke, ploceice \}$$
- Нека наредни модел релације описује шта се налази у којој кутији:
 
$$kutijaSadrzi = \{ (plava, lopte), (plava, kocke), (zelena, ploceice), (crvena, kocke) \}$$
- Исказ  $kutijaSadrzi(plava, kocke)$  је тачан
- Исказ  $kutijaSadrzi(zelena, lopte)$  није тачан
- Искази  $kutijaSadrzi(zuta, kocke)$  и  $kutijaSadrzi(zelena, knjige)$  нису дефинисани, зато што аргументи нису у домену релације

## Пример (2)

- Релације (односно њихови модели) са коначним бројем елемената могу да се представе и помоћу табела (таблица):

kutijaSadrži	
BOJA	SADRŽAJ
plava	lopte
plava	kočke
zelena	pločice
crvena	kočke



## Бесконачне релације и скупови

- Релације са бесконачним бројем елемената можемо да представљамо формалним записивањем скупа који представља модел релације
- На пример, релацију *neparan* можемо да представимо, па и да дефинишемо, скупом:
 
$$neparan = \{ n \mid n \in \mathbb{N} \wedge (\exists k \in \mathbb{N})(n = 2k - 1) \}$$
- У области база података бесконачне релације нису посебно занимљиве
  - колико год да тежимо да наша складишта података буду велика, она су свакако коначна
  - ипак, оне могу да имају примену у "базама знања" и системима за аутоматско закључивање



## Ентитети и атрибути

- Нека је  $E$  неки скуп *енџитиџита*
  - Енџитиџита* називамо све различите постојеће елементе ("објекте") система који посматрамо, које моделирамо базом података
- Кажемо да се скуп ентитета карактерише коначним скупом атрибута  $A(E) = \{A_1, \dots, A_n\}$ , у ознаци  $E(A_1, \dots, A_n)$ , ако:
  - Сваки атрибут  $A_i$  представља функцију која слика ентитете у одговарајући домен атрибута  $D_i$ :
 
$$A_i: E \rightarrow D_i$$
  - Сваки атрибут  $A_i$  има јединствен назив  $t_i$
- За сваки ентитет  $e \in E$ , вредност функције  $A_i(e) \in D_i$  представља *вредност атрибута*  $A_i$



## Ентитети и атрибути - пример

- Нека је  $E$  скуп радника који се карактерише атрибутима:
  - ime* (ниска од 20 знакова)
  - prezime* (ниска од 25 знакова)
  - osnova\_plate* (број облика 7.2).
- Скуп свих атрибута одређује функцију:
 
$$\alpha(x) = (ime(x), prezime(x), osnova_plate(x))$$

IME	PREZIME	OSNOVA_PLATE
Dragana	Pantić	45000
Marko	Marković	40000
Dragana	Pantić	45000
Jelena	Popović	43000
Dragiša	Đukić	47000

\* Приметимо да за различите ентитете можемо да добијемо једнаке слике





## Ентитети и атрибути (2)

- Скуп свих атрибута одређује функцију:  

$$\alpha(e) = (A_1(e), \dots, A_n(e))$$
- Слика скупа ентитета функцијом  $\alpha$  је скуп:  

$$R = \alpha(E)$$
- Подсетимо се, функција  $\alpha$  је инјективна ако за све различите оригинале даје различите слике, тј. ако важи:  

$$\alpha(e) = \alpha(u) \text{ ако } e = u$$
- Ако је описана функција  $\alpha$  инјективна, онда кажемо да скуп атрибута (и одговарајућа функција  $\alpha$ ) *добро карактеришу* скуп ентитета  $E$



## Представљање ентитета релацијама

Ако је  $\alpha$  инјективна функција и сваки од атрибута је атомичног типа, онда кажемо да је слика  $R = \alpha(E)$

*релација  $R$  са атрибутима  $A(R) = \{A_1, \dots, A_n\}$ ,  
 доменом релације  $Dom(R) = D_1 \times \dots \times D_n$   
 и називима атрибута  $Kol(R) = (t_1, \dots, t_n)$*

и скуп ентитета  $E$  са атрибутима  $A(E) = \{A_1, \dots, A_n\}$  моделирамо релацијом  $R$

- функција  $\alpha(E)$  је *добра карактеризација* скупа ентитета  $E$
- формално, дефинишемо  $R$  тако да је  $model(R) = \alpha(E)$
- атрибут је *атомичан* (т.ј. атомичног типа) ако његов домен не представља производ других домена, тј. ако је *скаларног* типа



## Записивање

- Ако је  $\alpha(e) = (a_1, \dots, a_n)$ , онда  $n$ -торку  $(a_1, \dots, a_n)$ , која формално представља “модел ентитета  $e$ ”, често називамо и “ентитетом  $e$ ”
- $A_i(e)$  записујемо и као  $e.t_i$
- Најчешће не правимо разлику између релације  $R$  и одговарајућих релација добијених пермутовањем атрибута
  - т.ј. атрибуте реферишемо користећи имена а не редне бројеве
  - због тога имена атрибута у релацији морају да буду јединствена



## Представљање релације табелом

- Релацију  $R = \alpha(E)$  често представљамо, па и називамо, *табелом*:
  - Колоне табеле одговарају атрибутима  $A_1, \dots, A_n$
  - Називи колона одговарају називима атрибута  $t_1, \dots, t_n$
  - Врсте табеле одговарају *шоркама* релације, тј. ентитетима



## “Релација” или “табела”

- У контексту теоријског разматрања, као и концептуалног и логичког моделирања уобичајено је да се користи термин “*релација*”
- У контексту физичког моделирања обично се користи термин “*табела*”



## Представљање ентитета – пример

- Релација не сме да има једнаке елементе

RADNIK		
IME	PREZIME	OSNOVA_PLATE
Dragana	Pantić	42000
Marko	Marković	40000
Dragana	Pantić	45000
Jelena	Popović	43000
Dragiša	Đukić	47000

- Ако је  $\alpha(x)=(ime_x, prezime_x, osnova\_plate_x)$ , онда ћемо  $n$ -торку  $(ime_x, prezime_x, osnova\_plate_x)$  у записима обично да изједначавамо са  $x$
- $ime(x)$  ћемо да записујемо као  $x.ime$  (и слично за остале атрибуте)
- Нећемо разликовати релацију *RADNIK* од релација добијених пермутовањем атрибута
  - као ни табеле које се разликују само по редоследу редова



## Релациона база података

- *Релациона база података* је скуп релација
- *Опис релације* чине домен релације и називи атрибута
- *Релациона схема* је скуп описа релација које чине базу података



## Пример базе података – типови

- Уведимо описно следеће скупове:
  - $N(a)$  = “скуп целих бројева записаних са највише  $a$  декадних цифара”;
  - $N(a,b)$  = “скуп бројева који у декадном запису са леве стране децималне запете имају највише  $a$ , а са десне стране највише  $b$  цифара”
  - $C(a)$  = “скуп ниски дужине до  $a$  знакова”
  - $D$  = “скуп датума (у неком подразумеваном опсегу)”
- Наведени скупови се најчешће употребљавају као домени (тј. типови) атрибута релација
- Сви атрибути у наредном примеру релационе базе података имају за домен неки од наведених скупова

## Пример базе података – схема

- $Kol(ORG\_JEDINICA) = ('org\_jed\_id', 'nad\_org\_jed\_id', 'naziv')$
- $Dom(ORG\_JEDINICA) = N(3) \times N(3) \times C(60)$
- $Kol(RADNIK) = ('radnik\_id', 'org\_jed\_id', 'ime', 'prezime', 'pol', 'dat\_zaposlenja', 'osnova\_plate')$
- $Dom(RADNIK) = N(4) \times N(3) \times C(30) \times C(50) \times C(1) \times D \times N(10,2)$
- $Kol(KAT\_STAZA) = ('od\_godine', 'do\_godine', 'staz\_bonus', 'naziv')$
- $Dom(KAT\_STAZA) = N(2) \times N(2) \times N(10,2) \times C(60)$
- $Kol(RESENJE) = ('rad\_id', 'proj\_id', 'dat\_pocetka', 'dat\_prestanka', 'opis')$
- $Dom(RESENJE) = N(4) \times N(3) \times D \times D \times C(60)$
- $Kol(PROJEKAT) = ('projekat\_id', 'proj\_bonus', 'naziv')$
- $Dom(PROJEKAT) = N(3) \times N(10,2) \times C(60)$

## Пример базе података – схема (2)

RADNIK	
RADNIK_ID	N(4)
ORG_JED_ID	N(3)
IME	C(30)
PREZIME	C(50)
POL	C(1)
DAT_ZAPOSLENJA	D
OSNOVA_PLATE	N(10,2)

ORG_JEDINICA	
ORG_JED_ID	N(3)
NAD_ORG_JED_ID	N(3)
NAZIV	C(60)

PROJEKAT	
PROJEKAT_ID	N(3)
PROJ_BONUS	N(10,2)
NAZIV	C(60)

RESENJE	
RAD_ID	N(4)
PROJ_ID	N(3)
DAT_POCETKA	D
DAT_PRESTANKA	D
OPIS	C(60)

KAT_STAZA	
OD_GODINE	N(3)
DO_GODINE	N(3)
STAZ_BONUS	N(10,2)
NAZIV	C(60)

## Пример базе података – садржај (1)

RADNIK						
RADNIK_ID	ORG_JED_ID	IME	PREZIME	POL	DAT_ZAPOSLENJA	OSNOVA_PLATE
100	10	Petar	Perić	M	26.03.1991.	42000
101	11	Marko	Marković	M	30.03.1990.	43000
102	-	Lazar	Lazić	M	21.04.1981	28000
103	11	Milica	Milić	Ž	30.06.1979.	45000
104	30	Dragica	Dragić	Ž	10.07.1988.	44000
105	30	Gorica	Bojić	Ž	20.07.2003.	39000
106	30	Vladana	Mladić	Ž	05.01.2007.	29000
107	40	Đurđa	Perić	Ž	29.03.1978.	35000
108	40	Đura	Lazić	M	18.03.1983.	42000

- о недефинисаним вредностима ћемо другом приликом, за сада претпоставити да проширују сваки од домена

## Пример базе података – садржај (2)

ORG_JEDINICA		
ORG_JED_ID	NAD_ORG_JED_ID	NAZIV
40	-	Softver
30	40	Projektovanje
11	10	Dokumentacija
10	40	Planiranje
50	40	Analiza
60	40	Dizajn i modeliranje
70	40	Kodiranje



## Пример базе података – садржај (3)

KAT_STAZA			
OD_GODINE	DO_GODINE	STAŽ_BONUS	NAZIV
1	5	2000	Početna
6	20	5000	Srednja
21	50	8000	Završna

PROJEKAT		
PROJEKAT_ID	PROJ_BONUS	NAZIV
100	-	Crveni
150	-	Ljubičasti
200	-	Plavi
250	1500	Zeleni
300	2000	Žuti
250	1500	Narandžasti
900	-	Crni



## Пример базе података – садржај (4)

REŠENJE				
RAD_ID	PROJ_ID	DAT_POČETKA	DAT_PRESTANKA	OPIS
100	200	01.01.1991.	-	privremeno učešće
101	200	01.01.1991.	-	-
108	200	01.01.1984.	-	-
104	100	01.01.1989.	31.12.1998.	-
105	100	01.01.2004.	-	-
106	100	01.01.2008.	-	-
107	100	01.01.1981.	-	-
100	150	01.01.1999.	31.12.1998.	Koordinator
101	150	01.01.1972.	31.12.1998.	-
102	150	01.01.1972.	31.12.1998.	-
108	150	01.01.1972.	31.12.1998.	-
104	150	01.01.1986.	31.12.1998.	-
105	150	01.01.1998.	31.12.1998.	-
107	150	01.01.1999.	-	Projektant
108	250	01.01.1999.	-	-
108	300	01.01.1999.	-	-
108	350	01.01.1999.	-	-
105	250	01.01.1998.	-	-
107	350	01.01.1998.	-	-



## Кључ и Наткључ

- Појам *кључа* има веома важну улогу у релационом моделу
  - Сада ћемо га увести мало поједностављено, формално нешто касније
- **Напоменљиво**  $K$  релације  $R$  је подскуп њених атрибута  $K \subseteq A(R)$  за који важи:
  - скуп атрибута  $K$  једнозначно одређује торке релације
    - т.ј. ако знамо вредности тих атрибута онда је једнозначно одређена торка
- **Кључ** (или *кључ-кандидат*) релације  $R$  је наткључ  $K$  за који важи:
  - не постоји прави подскуп скупа  $K$  који је наткључ релације  $R$
- Приметимо:
  - Свака релација има најмање кључ, а може да их има више
  - Сваки наткључ садржи бар један кључ



## Моделирање односа

- На самом почетку смо нагласили да се односи моделирају на исти начин као и ентитети – релацијама
  - штавише, однос често може да се посматра као ентитет, па је то савим природно
- Ипак, описаћемо формално представљање бинарних односа, а за сложеније важе аналогни принципи
- Нека су ентитети  $e \in E$  и  $f \in F$  моделирани као
 
$$e = (x_1, \dots, x_n), f = (y_1, \dots, y_m),$$
 и нека могу да буду у односу  $w$ , који нема додатне атрибуте (тј. само знамо да ли су два ентитета у односу или нису)





## Моделирање односа (2)

- Ако су два ентитета  $e \in E$  и  $f \in F$  у неком односу  $w$ , онда то описујемо новом релацијом  $\rho(e, f)$ , чији је домен  $Dom(\rho) = E \times F$
- Ако су ентитети  $e \in E$  и  $f \in F$  моделирани као  $e = (x_1, \dots, x_n)$ ,  $f = (y_1, \dots, y_m)$ , онда њихов се њихов однос моделира као  $\rho(e, f) = (x_1, \dots, x_n, y_1, \dots, y_m)$



## Моделирање односа (3)

- Приметимо, ако су  $K_E$  и  $K_F$  неки кључеви релација  $E$  и  $F$ , онда скуп атрибута  $K_R = K_E \cup K_F$  представља кључ релације  $\rho$
- атрибути из  $K_R$  једнозначно одређују све атрибуте из  $E$  и  $F$ , па тиме и из  $\rho$
- Посебно, онда се однос може описати "суженом" релацијом  $\rho_1$  која не садржи све атрибуте из  $E$  и  $F$  већ само атрибуте скупа  $K_R$



## Моделирање односа са атрибутима

- У општем случају односи могу да имају додатне атрибуте
- На пример, у једној библиотеци однос *позајмица* између *књиџе* и *члана библиотеке* осим тога што га чине једна књига и један члан, може да има и додатне атрибуте, као на пример:
  - датум позајмице
  - датум враћања
- Тада кажемо да је однос карактерисан повезаним ентитетима и додатним атрибутима, које називамо *описним атрибутима односа*



## Интегрисани односи

- У неким случајевима однос између два ентитета (или односа) може да буде "интегрисан" у оквиру релације која моделира један од тих ентитета (или односа)
  - На пример, релација *Студент* може да садржи атрибуте релације *СтудијскиПрограм* (све или само кључ), чиме се описује који студијски програм је студент уписао
  - Али то је истовремено и однос између студента и студијског програма
- Однос између релација  $E$  и  $F$  може да буде *интегрисан* у оквиру релације  $E$  само ако сваком ентитету из те релације одговара тачно по један ентитет релације  $F$ 
  - (или највише један, ако су допуштене недефинисане вредности)



## Пример интегрисаног односа

- У претходном примеру базе података већ имамо и неке интегрисане односе:
  - Сваком раднику из *RADNIK* одговара тачно један ентитет из *ORG\_JEDINICA*
    - радник ради у тачно једној организационој јединици
  - Зато релација *RADNIK* садржи атрибут кључа релације *ORG\_JEDINICA*



## Структурни део модела - резиме

- Упознали смо основне концепте моделирања података у релационом моделу
  - релација, модел релације, домен релације
  - атрибут, скуп атрибута, кључ
  - формалан начин моделирања ентитета и односа

## Литература за тему



- Гордана Павловић-Лажетић, **Увод у релационе базе података, 2. изд.** *Математички факултет, 1999.*
  - доступно онлајн: <http://poincare.matf.bg.ac.rs/~gordana/urbp-2016.htm>
- Ramakrishnan, Gehrke, **Database Management Systems, 2.ed, 2000.**
- Codd, **A relational model of data for large shared data banks**, *Comm.ACM*, 13(6), 1970.
- Codd, **Extending the database relational model to capture more meaning**, *ACM ToDS*, 4(4), 1979.
- Codd **The Relational Model for Database Management – Version 2**, *AddisonWesley Publ. Inc.*, 1990.
- Darwen, Date, **The Third Manifesto**, 1995.
- IBM, **Database Administration Concepts and Configuration Reference**, 2012.