

Увод у релационе базе података

Ненад Митић

Математички факултет
nenad.mitic@matf.bg.ac.rs

Пример релационе базе

База студената, предмета и испита - садржај
(дела) табеле испита

Indeks	Id_predmeta	Šk_godina	Oznaka_roka	Ocena	Dat_polaganja	Poeni
20140021	1001	2015	jan	9	20.01.2015	81
20140022	1001	2015	jan	8	20.01.2015	75
20130023	1001	2015	jan	8	20.01.2015	76
20130024	1001	2015	jan	10	20.01.2015	100
20140025	1001	2015	jan	6	20.01.2015	55
20140026	1001	2015	jan	5	20.01.2015	25
20130027	1001	2015	jan	8	20.01.2015	76
20140021	2001	2015	jan	10	25.01.2015	98
20140022	2001	2015	jan	9	25.01.2015	86
20140025	2001	2015	jan	5	25.01.2015	0
20140021	3001	2015	jan	7	27.01.2015	67
20130023	3001	2015	jan	5	27.01.2015	30
20140026	1001	2015	feb	7	10.02.2015	68
20140025	2001	2015	feb	6	10.02.2015	55
20140026	2001	2015	feb	7	10.02.2015	65
20140021	1021	2015	apr	7	03.04.2015	63
20130023	1021	2015	apr	10	03.04.2015	95
20140022	1021	2015	apr	5	03.04.2015	
20130024	1021	2015	apr	6		

Примери релационих оператора – рестрикција

Рестрикција (издвајање, селекција) издваја појединачне редове из табеле

DOSIJE WHERE Datum_upisa > '05.07.2014'

Indeks	Ime	Prezime	Datum_upisa	Mesto_rođenja
20140021	Milos	Peric	06.07.2014	Beograd
20140025	Marijana	Savkovic	06.07.2014	Kraljevo
20140026	Zorica	Miladinovic	06.07.2014	Vranje

Примери релационих оператора – пројекција

Пројекција издваја појединачне колоне из табеле
DOSIJE over Indeks, Ime, Prezime

Indeks	Ime	Prezime
20140021	Milos	Peric
20140022	Marijana	Savkovic
20130023	Sanja	Terzic
20130024	Nikola	Vukovic
20140025	Marijana	Savkovic
20140026	Zorica	Miladinovic
20130027	Milena	Stankovic

Примери релационих оператора - спајање

Спајање комбинује две табеле на основу заједничких вредности у заједничкој колони (у примеру је приказано природно спајање, део резултујуће табеле)

DOSIJE AND ISPIT OVER Indeks

Indeks	Ime	Prezime	Mesto_rodjenja	Datum_upisa	Id_predmeta	Školska_godina	Oznaka_roka	Oцена	Dat_polaganja	Bodovi
2014 0021	Milos	Peric	Beograd	06.07.2014	1001	2015	jan	9	20.01.2015	81
2014 0022	Marijana	Savkovic	Kraljevo	05.07.2014	1001	2015	jan	8	20.01.2015	75
20130023	Sanja	Terzic	Beograd	04.07.2013	1001	2015	jan	8	20.01.2015	76
20130024	Nikola	Vukovic		04.07.2013	1001	2015	jan	10	20.01.2015	100
2014 0025	Marijana	Savkovic	Kraljevo	06.07.2014	1001	2015	jan	6	20.01.2015	55
2014 0026	Zorica	Miladinovic	Vranje	06.07.2014	1001	2015	jan	5	20.01.2015	25
20130027	Milena	Stankovic			1001	2015	jan	8	20.01.2015	76
2014 0021	Milos	Peric	Beograd	06.07.2014	2001	2015	jan	10	25.01.2015	98
2014 0022	Marijana	Savkovic	Kraljevo	05.07.2014	2001	2015	jan	9	25.01.2015	86
20130023	Sanja	Terzic	Beograd	04.07.2013	2001	2015	jan	8	25.01.2015	72
20130024	Nikola	Vukovic		04.07.2013	2001	2015	jan	7	25.01.2015	63
2014 0025	Marijana	Savkovic	Kraljevo	06.07.2014	2001	2015	jan	5	25.01.2015	0
2014 0021	Milos	Peric	Beograd	06.07.2014	3001	2015	jan	7	27.01.2015	67
20130023	Sanja	Terzic	Beograd	04.07.2013	3001	2015	jan	5	27.01.2015	30
20130024	Nikola	Vukovic		04.07.2013	3001	2015	jan	6	28.01.2015	59
2014 0026	Zorica	Miladinovic	Vranje	06.07.2014	3001	2015	jan	6	28.01.2015	52
2014 0026	Zorica	Miladinovic	Vranje	06.07.2014	1001	2015	feb	7	10.02.2015	68
2014 0025	Marijana	Savkovic	Kraljevo	06.07.2014	2001	2015	feb	6	10.02.2015	55
2014 0026	Zorica	Miladinovic	Vranje	06.07.2014	2001	2015	feb	7	10.02.2015	65

Примери релационих оператора – спајање

Код спајања се у резултујућој табели

- заједничке вредности јављају само једном у реду
- јављају све комбинације упарених вредности (нпр. 20140021 вредност за Indeks)
- не јављају се редови са вредностима које се не налазе у обе табеле које учествују у спајању

Карактеристике релационе базе

- Релациони системи захтевају само **да се подаци прикажу** кориснику у облику табеле (начин смештања и чувања на медијумима није специфициран)
- **Информациони принцип**: Целокупан информациони контекст базе се приказује на тачно један начин као експлицитне вредности у позицијама врста и колона табеле

Карактеристике релационе базе – наставак

Последица Информационог принципа: нема показивача који међусобно повезују табеле

- показивачи могу да постоје на физичком нивоу
- разлика у односу на нерелационе моделе где су показивачи део структуре која се приказује корисницима и које корисник користи у обради

Ефекат релационих оператора

- Резултат примене сваког оператора је табела - особина затворења релационих система
- Резултат примене оператора је истог типа као и његов аргумент \implies могу да се пишу угнеждени релациони изрази

Ефекат релационих оператора

- Све операције се примењују на цео скуп истовремено а не само на појединачни ред
- Резултат операције није никада појединачни ред већ је увек комплетна табела која садржи скуп редова

Формални поглед на релациони модел

Релациони модел се састоји од:

- Отвореног скупа **скаларних типова** (који укључује и тип логичких вредности *boolean*)
- **Генератора релационих типова** и њихове одговарајуће интерпретације
- Могућности дефинисања **релационих променљивих** за генерисане релационе типове
- Операције **релационе доделе** којом се додељују релационе вредности дефинисаним релационим променљивим
- Отвореног скупа општих **релационих оператора** ("релациона алгебра") за извођење релационих вредности из других релационих вредности
 - 1 → Релациона алгебра и релациони рачун
 - 2 → Додатни оператори релационе алгебре

Терминологија

Codd је при формулисању принципа релационих база увео нову терминологију која се разликовала од неодређене терминологије тог времена

Термин	Значење
Релација	математички термин за табелу
Торка	ред у табели
Атрибут	колона у табели
Кардиналност	број торки
Степен	број атрибута
Примарни кључ	атрибут или комбинација атрибута који јединствено идентификују табелу
Домен	скуп важећих вредности (типова)

Терминологија

Indeks	Ime	Prezime	Mesto_rođenja	Datum_upisa
20140021	Milos	Peric	Beograd	06.07.2014
20140022	Marijana	Savkovic	Kraljevo	05.07.2014
20130023	Sanja	Terzic	Beograd	04.07.2013
20130024	Nikola	Vukovic		04.07.2013
20140025	Marijana	Savkovic	Kraljevo	06.07.2014
20140026	Zorica	Miladinovic	Vranje	06.07.2014
20130027	Milena	Stankovic		

Релација	је комплетна табела (име релације DOSIJE)
Торка	је сваки ред у табели
Атрибути	табеле су <i>Indeks, Ime, Prezime, Mesto_rođenja, Datum_upisa</i>
Кардиналност	табеле је 7 (табела поседује 7 торки)
Степен	табеле је 5 (табела има 5 атрибута)
Примарни кључ	табеле је <i>Indeks</i>
Домен	у табели су скуп могућих бројева индекса (INDEKS), скуп могућих имена (IME), ...

Релације и релационе променљиве

- 1 У дефиницији базе студената, предмета и испита, DOSIJE и ISPIT су релационе променљиве, односно променљиве чије су вредности релационе вредности
- 2 Нпр. нека DOSIJE има текућу вредност из које треба да се избрише ред за студента са бројем индекса 20100027:

```
DELETE DOSIJE WHERE Indeks =  
Indeks(20100027);
```

Релације и релационе променљиве

Стара вредност релационе променљиве DOSIJE је замењена новом релационом вредношћу

Indeks	Ime	Prezime	Mesto_rođenja	Datum_upisa
20140021	Milos	Peric	Beograd	06.07.2014
20140022	Marijana	Savkovic	Kraljevo	05.07.2014
20130023	Sanja	Terzic	Beograd	04.07.2013
20130024	Nikola	Vukovic		04.07.2013
20140025	Marijana	Savkovic	Kraljevo	06.07.2014
20140026	Zorica	Miladinovic	Vranje	06.07.2014

Оператор брисања је скраћеница за релациони оператор доделе

```
DOSIJE := DOSIJE WHERE NOT (Indeks = Indeks(20130027));
```


Основни релвар-и и погледи

- Оригинални релвар-и = основни релвар-и
- Њихове вредности = основне релације
- Изведене релације = релације које нису основне али могу да се добију из основних путем релационих израза

Основни релвар-и и погледи (наставак)

- Основни релвар-и морају да имају име
- Релационе базе имају механизам за формирање основних релвар-а
- У SQL-у су основни релвар-и табеле. Њихово формирање се врши наредбом

```
CREATE TABLE <име_основног_релвара> ...
```

Основни релвар-и и погледи (наставак)

- Релациони системи подржавају и именоване релвар-е који су изведене релације - погледе
- Вредност погледа је (тренутни) резултат извршавања одређеног релационог израза који се наводи при формирању погледа

```
CREATE VIEW Studenti_iz_Kraljeva AS  
  (DOSIJE WHERE Mesto_rodjenja='Kraljevo') {Indeks, Ime, Prezime}
```

Студенти_из_Краљева

Indeks	Ime	Prezime
20140022	Marijana	Savkovic
20140025	Marijana	Savkovic

Основни релвар-и и погледи (наставак)

Над погледима могу да се врше операције као над основним табелама

```
(Studenti_iz_Kraljeva WHERE Indeks > 20140023){Ime, Prezime}
```

Ime	Prezime
Marijana	Savkovic

При извршавању систем модификује упит у

```
((DOSIJE WHERE Mesto_rodjenja='Kraljevo') {Indeks, Ime, Prezime}) WHERE Indeks > 20100023){Ime, Prezime}
```

Односно упрошћава у

```
(DOSIJE WHERE Mesto_rodjenja='Kraljevo' AND Indeks > 20100023){Ime, Prezime}
```


Основни релвар-и и погледи (наставак)

- Основни релвар-и представљају податке који су физички смештени у бази података ('заиста постоје')
- Погледи представљају различит начин гледања на 'реалне податке' (не постоје, виртуални релвар-и)

Оптимизација, каталог и трансакције

- 1 Оптимизацију (аутоматску) врши компонента СУБП названа оптимизатор
- 2 Каталог ('речник') се и сам састоји од релвар-а \implies каталогу се приступа на исти начин као и другим релационим променљивим
- 3 Трансакција је 'логичка јединица посла' која обично укључује више операција над базом

Детаљније о овим карактеристикама у другом делу курса

Домени

- Домен = тип података
- уграђен (системски предефинисан)
 - INTEGER
 - CHAR
 - ...
 - кориснички дефинисан
 - INDEKS
 - IME
 - GOD_RODJENJA
 -

Домени (наставак)

- Провера типова. Строга типизираност
 - `Dosije.Indeks + Ispit.ocena` → погрешно!
 - `Ispit.Godina_roka * Ispit.Ocena` → исправно!
- Дефинисање новог типа
 - TYPE <име типа> <могуће репрезентације>
 - TYPE INDEKS REPREZENT (INTEGER);
 - TYPE IME REPREZENT (CHAR);

Домени - SQL подршка

- Naredbe
 - CREATE DOMAIN
 - ALTER DOMAIN
 - DROP DOMAIN (RESTRICT, CASCADE)

- У DB2 ове наредбе не постоје. Постоје
 - 1 CREATE TYPE
 - 2 ALTER TYPE
 - 3 DROP TYPE

- Истовремено се дефинишу и одговарајуће функције и оператори поређења

- Нема строге типизираниости али постоји основни облик провере типова

Домени - SQL подршка (наставак)

SQL подржава следеће релационе домене

- 1 Бројеви (*numbers*)
- 2 Ниске карактера (*character strings*)
- 3 Ниске битова (*bit strings*)
- 4 Датуми (*dates*)
- 5 Времена (*times*)
- 6 Комбинације датума и времена (*timestamps*)
- 7 Интервали година/месец (*year/month intervals*)
- 8 Интервали дан/време (*day/time intervals*)
- 9 ...

Релације

Нека је дат скуп од n типова или домена T_i ($i=1,2,\dots,n$), при чему не морају сви типови да буду међусобно различити. R је релација над тим типовима ако се састоји од два дела, заглавља и тела, где важи:

- 1 **Заглавље** је скуп од n атрибута облика $A_i : T_i$ где су A_i (који сви морају бити различити) имена атрибута релације R , а T_i одговарајућа имена типова ($i = 1, 2, \dots, n$)
- 2 **Тело** је скуп од m торки t где је t скуп компоненти облика $A_i : v_i$ у којима је v_i вредност типа T_i

m се назива кардиналност, а n степен (арност) релације R

Релација = Табела?

Нису једнаке, јер

- Табела може да садржи дуплиране редове а релације не
- Редови у табели су уређени у редоследу од врха ка дну, док за релацију то не важи
- Колоне у табели су уређене у редоследу са лева у десно, док за релацију то не важи

Релације и празан скуп

- Релација која има празан скуп торки
 - непразно заглавље
 - тело је празно
- Релација која има празну торку
 - празно заглавље
 - тело са једном торком без компоненти

Типови релације

Дефиниција типа релације има следећи облик
`RELATION {<листа атрибута раздвојених зарезима>}`

атрибут је уређен пар облика

`<име атрибута> <име типа>`

Пример:

`RELATION {Индекс INDEKS, Име IME, Презиме PREZIME,
God_rodjenja GOD_RODJENJA, Mesto_rodjenja
MESTO_RODJENJA}`

Атрибути и типови података

- Сваки атрибут има неки тип
- Сваки тип може да буде уграђен или кориснички дефинисан
- Сваки тип може да буде атомски (скаларан) или учаурен (није скаларан)
- Атрибути релације могу да буду произвољног типа

Недостајуће вредности

У свакодневной пракси се често јавља проблем недостатка података:

- '...датум рођења непознат...'
- '...из изборне јединице XX недостају подаци...'
- ...

Потреба да се

- индикатор о недостатку вредности чува у бази
- на одговарајући начин врши обрада таквих података

Недостајуће вредности (наставка)

- Најчешти прихваћен приступ прихваћен и у пракси је коришћење 'недостајуће вредности' (NULL), односно тровалентне логике
- *Codd* је предложио коришћење 4-валентне логике јер постоје две врсте недостајућих вредности:
 - вредност је непозната
 - вредност није применљива, вредност не постоји,
 -

3VL логика

Три вредности

- Тачно
- Нетачно
- Непознато (код *DATE*-а *UNK* од *unknown*)

Оператори: I, ILI, NE, MOŽDA (AND, OR, NOT, MAYBE)

3VL логика (наставак)

AND	t	u	f	OR	t	u	f	NOT	
t	t	u	f	t	t	t	t	t	f
u	u	u	f	u	t	u	u	u	u
f	f	f	f	f	t	u	f	f	t

Оператор МОЖДА

MAYBE	
t	f
u	t
f	f

Оператор је потребан нпр.
због оваквих упита:

*Приказати све запослене
који су можда били, али за
које није сасвим поуздано
да су били, програмери
рођени пре 25. јануара 1997
године са платом мањом од
150.000 на дан 30.09.2022.*

Развој релационих продуката

	System R	INGRES
Аутор	IBM San Jose Res. Lab 1974 - 1979	UC Berkeley касне 1970 - ране 1980
Рачунар	IBM System 370	DEC PDP
Оперативни систем	VM/CMS	UNIX
Упитни језик	SQL	QUEL
Матични (хост) језик	COBOL, PL/1	COBOL, PASCAL, C, FORTRAN, BASIC
Комерцијални производ	DB2, SQL/DS	Комерцијални INGRES
Дистрибуирана обрад	R*	Дистрибуирани INGRES
ОО проширења	Starburst	POSTGRES

Компоненте SQL-а

- 1 DDL подскуп SQL-а
- 2 DML подскуп SQL-а
- 3 DCL подскуп SQL-а

DDL подскуп SQL-а

Најважније наредбе

- 1 CREATE
- 2 ALTER
- 3 DROP
- 4 DECLARE

DML подскуп SQL-а

Најважније наредбе

- 1 SELECT
- 2 INSERT
- 3 UPDATE
- 4 DELETE
- 5 MERGE

DCL подскуп SQL-а

Најважније наредбе

- 1 GRANT
- 2 REVOKE
- 3 ...

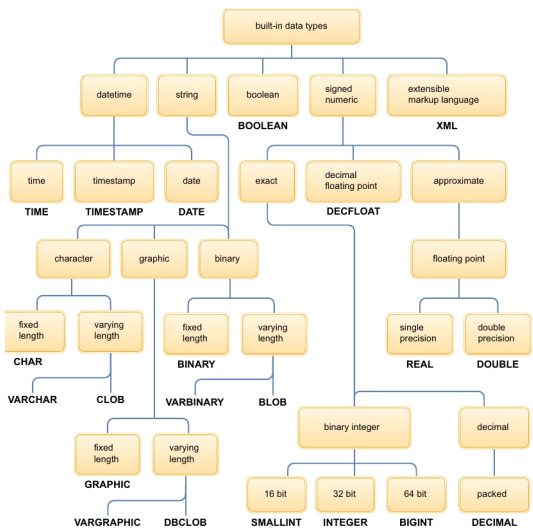
SQL подршка за недостајуће вредности

- SQL при примени WHERE клаузуле на табелу T елиминира све редове за које израз наведен у WHERE има вредност нетачно или недефинисано
- Тест за null: IS [NOT] NULL
- Не важи еквиваленција
 $r \text{ IS NOT NULL}$ и $\text{NOT} (r \text{ IS NULL})$
- Остали начини обраде NULL биће дати кроз примере

Додатак: Неки типови података у SQL/DB2

- ниске карактера
- нумерички подаци
- датумско-временски подаци
- константе
- графичке ниске
- кориснички дефинисани типови података
- *rowid* вредности
- бинарне ниске
- LOB
- XML
- ...

Предефинисани типови података у DB2



Додатак: Специјални регистри у DB2

- current application encoding scheme
- current date
- current degree
- current locale lc_ctype
- current optimization hint
- current packageset
- current path
- current precision
- current rules
- current server
- current sqlid
- current time
- current timestamp
- current timezone
- user