پایگاه داده ها

مدل رابطه ای – جبر رابطه ای

مدل رابطه‌ای

- هر پایگاه داده‌های رابطه‌ای مبتنی بر مدل رابطه‌ای است
  - از مجموعه‌ای از جداول برای نمایش داده‌ها و روابط میان آن‌ها استفاده می‌کند
  - مفهوم جدول در پایگاه داده‌های رابطه‌ای مشابهت زیادی با مفهوم رابطه در ریاضیات دارد
    - رابطه ***account***

account_number	*branch_name*	*balance*
A-101	Downtown	500
A-102	Perryridge	400
A-201	Brighton	900
A-215	Mianus	700
A-217	Brighton	750
A-222	Redwood	700
A-305	Round Hill	350

مدل رابطه‌ای

- مجموعه‌های D_1، D_2، ...، D_n را در نظر بگیرید
  - هر زیرمجموعه از $D_1 \times D_2 \times \ldots \times D_n$ یک رابطه است
  - هر رابطه مجموعه‌ای از nتایی‌های $(a_1, a_2, \ldots, a_n)$ است به طوری که هر $a_i \in D_i$
  - مثال
    - فرض کنید *CustomerName* مجموعه نام‌های مشتریان، *CustomerStreet* مجموعه نام‌های خیابان‌ها و *CustomerCity* مجموعه نام‌های شهرها باشد

CustomerName = {Jones, Smith, Curry, Lindsay, …}

CustomerStreet = {Main, North, Park, …}

CustomerCity = {Harrison, Rye, Pittsfield, …}

مدل رابطه‌ای

- r رابطه‌ای روی $CustomerName \times CustomerStreet \times CustomerCity$ است

r = { (Jones, Main, Harrison),
(Smith, North, Rye),
(Curry, North, Rye),
(Lindsay, Park, Pittsfield) }

- هر جدول در واقع یک رابطه است
  - از اصطلاحات رابطه، چندگانه و خصیصه می‌توان به جای اصطلاحات جدول، سطر و ستون استفاده کرد
- دامنه یک خصیصه
  - مجموعه مقادیر مجاز برای آن خصیصه

مدل رابطه‌ای

- مقادیر خصیصه‌ها باید تجزیه ناپذیر باشند
  - مقدار یک خصیصه می‌تواند یک شماره حساب باشد، اما نمی‌تواند مجموعه‌ای از شماره‌های حساب باشد
- یک دامنه تجزیه ناپذیر است اگر همه اعضایش تجزیه ناپذیر باشند
- مقدار یک خصیصه ممکن است ناشناخته باشد یا این که موجود نباشد
  - تهی (Null)

شِمای رابطه

- هر شِمای رابطه شامل فهرستی از خصیصه‌ها و دامنه‌های متناظر با آن خصیصه‌ها است
  - مثال
    - شِمای رابطه ***account***

$$Account_schema = (account_number, branch_name, balance)$$

- ***account***(***Account_schema***) بیان می‌کند که ***account*** رابطه‌ای روی شِمای ***Account_schema*** است

شِمای رابطه

- مثال
  - رابطه ***branch***

branch_name	*branch_city*	*assets*
Brighton	Brooklyn	7100000
Downtown	Brooklyn	9000000
Mianus	Horseneck	400000
North Town	Rye	3700000
Perryridge	Horseneck	1700000
Pownal	Bennington	300000
Redwood	Palo Alto	2100000
Round Hill	Horseneck	8000000

  - شِمای رابطه ***branch***

***Branch_schema* = (*branch_name*, *branch_city*, *assets*)**

پایگاه داده‌های رابطه‌ای

- هر پایگاه داده‌های رابطه‌ای شامل تعدادی رابطه است
  - هر رابطه بخشی از اطلاعات را ذخیره می‌کند
  - مثال
    - پایگاه داده‌های سیستم بانکی
      - رابطه ***account***
        - اطلاعات حساب‌های بانکی
      - رابطه ***customer***
        - اطلاعات مشتریان
      - رابطه ***depositor***
        - وابستگی بین مشتریان و حساب‌های بانکی

کلیدها

- در یک رابطه هیچ دو چندگانه‌ای نباید مقادیر دقیقاً یکسانی برای تمام خصیصه‌ها داشته باشند
- کلید در مدل رابطه‌ای
  - متمایزکردن چندگانه‌های داخل یک رابطه
    - ابرکلید (Superkey)
    - کلید کاندید (Candidate Key)
    - کلید اصلی (Primary Key)
    - کلید خارجی (Foreign Key)

کلیدها

- ابرکلید
  - مجموعه‌ای از یک یا چند خصیصه که امکان شناسایی منحصر به فرد هر چندگانه را فراهم می‌کند
  - مثال
    - رابطه ***customer***

customer_name	*customer_street*	*customer_city*
Adams	Spring	Pittsfield
Brooks	Senator	Brooklyn
Curry	North	Rye
Glenn	Sand Hill	Woodside
Green	Walnut	Stamford
Hayes	Main	Harrison
Johnson	Alma	Palo Alto
Jones	Main	Harrison
Lindsay	Park	Pittsfield
Smith	North	Rye
Turner	Putnam	Stamford
Williams	Nassau	Princeton

    - ابرکلیدها
      - {***customer_name***} و {***customer_name, customer_street***}

کلیدها

- در این مثال فرض شده است که هیچ دو مشتری دارای نام یکسان نمی‌باشند
- در دنیای واقعی از خصیصه‌ای مانند ***customer_id*** به جای ***customer_name*** برای شناسایی منحصر به فرد مشتریان استفاده می‌شود

- **کلید کاندید**
  - **ابرکلیدی که هیچ زیرمجموعه محض آن ابرکلید نباشد**
  - **مثال**
    - **رابطه *customer***
      - **کلید کاندید**
        - ***{customer_name}***

کلیدها

- کلید اصلی
  - یک کلید کاندید که توسط طراح پایگاه داده‌ها برای شناسایی چندگانه‌های داخل یک رابطه استفاده می‌شود
  - کلید اصلی باید به گونه‌ای انتخاب شود که مقادیر خصیصه‌هایش هرگز تغییر نکند
    - آدرس پست الکترونیکی منحصر به فرد است، اما امکان تغییر آن وجود دارد
- کلید خارجی
  - دو شِمای رابطه‌ای R_1 و R_2 را در نظر بگیرید
    - هر زیرمجموعه از خصیصه‌های شِمای R_2 که در شِمای R_1 کلید اصلی باشد، کلید خارجی نامیده می‌شود

کلیدها

- مثال
  - دو شِمای رابطه‌ای زیر را در نظر بگیرید

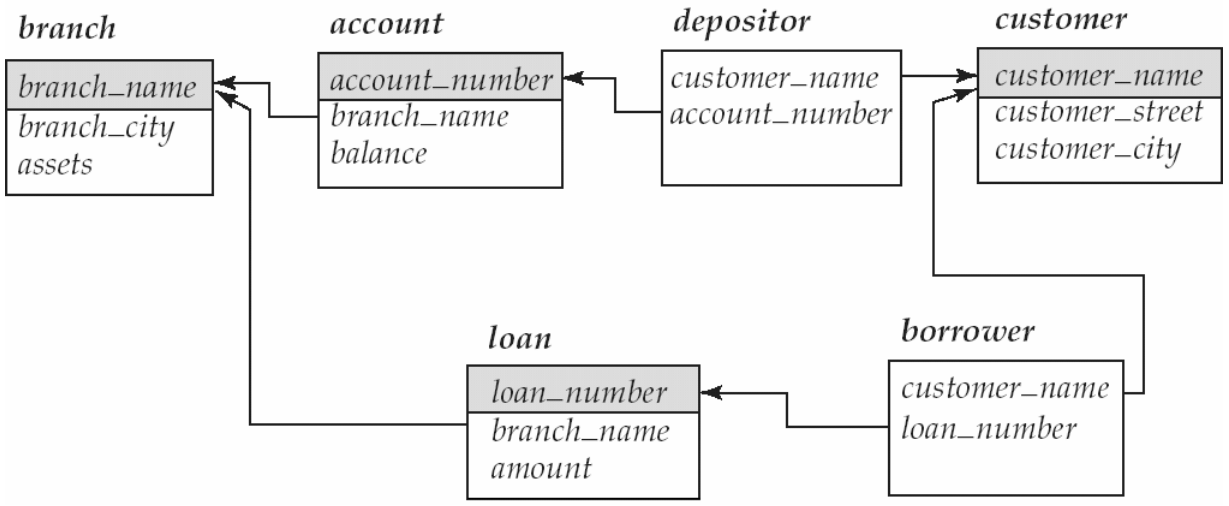
Account_schema = (*account_number*, *branch_name*, *balance*)
Branch_schema = (*branch_name*, *branch_city*, *assets*)

  - خصیصه *branch_name* در *Branch_schema* کلید اصلی است
    - هیچ دو شعبه‌ای دارای نام یکسان نمی‌باشند
  - خصیصه *branch_name* در *Account_schema* کلید خارجی است

- تنها مقادیر مشاهده شده در خصیصه‌های کلید اصلی رابطه مرجع ممکن است در خصیصه‌های کلید خارجی رابطه رجوع کننده مشاهده شوند

نمودار شِما

- هر شِمای پایگاه داده‌ها همراه با وابستگی‌های بین کلیدهای اصلی و خارَجی را می‌توان با استفاده از یک نمودار نمایش داد
  - مثال
    - نمودار شِما برای سیستم بانکی



زبان پرس و جو

- زبانی که با استفاده از آن کاربران اطلاعات مورد نظر خود را از پایگاه داده‌ها درخواست می‌کنند
  - جبر رابطه‌ای
  - حساب رابطه‌ای

جبر رابطه‌ای

- عملگرهای اصلی
  - انتخاب (Select)
    - σ
  - پرتو (Project)
    - Π
  - اجتماع (Union)
    - $\cup$
  - تفاضل (Difference)
    - $-$
  - ضرب دکارتی (Cartesian Product)
    - $\times$

عملگر انتخاب

- چندگانه‌هایی را از یک رابطه انتخاب می‌کند که شرط یا شرایط خاصی را برآورده می‌کنند
  - برای انتخاب چندگانه‌هایی از رابطه r که شرط p را برآورده می‌کنند، از نشانه‌گذاری $\sigma_p(r)$ استفاده می‌شود
- مثال
  - رابطه *loan* را در نظر بگیرید

loan_number	*branch_name*	*amount*
L-11	Round Hill	900
L-14	Downtown	1500
L-15	Perryridge	1500
L-16	Perryridge	1300
L-17	Downtown	1000
L-23	Redwood	2000
L-93	Mianus	500

عملگر انتخاب

- چندگانه‌هایی را از رابطه ***loan*** انتخاب کنید که نام شعبه آن‌ها **Perryridge** است

$$\sigma_{branch_name = \text{“Perryridge”}} (loan)$$

loan_number	*branch_name*	*amount*
L-15	Perryridge	1500
L-16	Perryridge	1300

- چندگانه‌هایی را از رابطه ***loan*** انتخاب کنید که نام شعبه آن‌ها **Perryridge** است و میزان وام آن‌ها از ۱۲۰۰ دلار بیشتر است

$$\sigma_{branch_name = \text{“Perryridge”} \wedge \ amount > 1200} (loan)$$

عملگر پرتو

- برای ایجاد رابطه‌ای استفاده می‌شود که تنها شامل بخشی از خصیصه‌های رابطه دیگر باشد
  - برای ایجاد رابطه‌ای که تنها شامل خصیصه‌های A_1، A_2، ...، A_k از رابطه r باشد از نشانه‌گذاری $\Pi_{A_1, A_2, ..., A_k}(r)$ استفاده می‌شود
- مثال
  - شماره و میزان همه وام‌ها را فهرست کنید

$$\Pi_{loan_number,\ amount}(loan)$$

loan_number	*amount*
L-11	900
L-14	1500
L-15	1500
L-16	1300
L-17	1000
L-23	2000
L-93	500

عملگر پرتو

- مثال
  - رابطه *customer* را در نظر بگیرید

customer_name	*customer_street*	*customer_city*
Adams	Spring	Pittsfield
Brooks	Senator	Brooklyn
Curry	North	Rye
Glenn	Sand Hill	Woodside
Green	Walnut	Stamford
Hayes	Main	Harrison
Johnson	Alma	Palo Alto
Jones	Main	Harrison
Lindsay	Park	Pittsfield
Smith	North	Rye
Turner	Putnam	Stamford
Williams	Nassau	Princeton

- نام‌های مشتریانی را فهرست کنید که در **Pittsfield** زندگی می‌کنند

$$\Pi_{customer_name}\left(\sigma_{customer_city = \text{“Pittsfield”}}\left(customer\right)\right)$$

عملگر اجتماع

- برای این که اجتماع دو رابطه امکان‌پذیر باشد
  - هر دو رابطه باید تعداد خصیصه‌های یکسانی داشته باشند
  - دامنه خصیصه‌های متناظر در هر دو رابطه باید یکسان باشد
- اجتماع دو رابطه r و s با $r \cup s$ نمایش داده می‌شود
- مثال
  - نام‌های مشتریانی را فهرست کنید که حساب بانکی و یا وام بانکی دارند

$$\Pi_{customer_name}(depositor) \cup \Pi_{customer_name}(borrower)$$

عملگر تفاضل

- فرض کنید دو رابطه r و s داده شده باشد
  - عبارت $r - s$ رابطه‌ای ایجاد می‌کند که تنها شامل چندگانه‌های رابطه r است
  - برای این که تفاضل دو رابطه r و s امکان‌پذیر باشد، هر دو رابطه باید با هم سازگار باشند
- مثال
  - نام‌های مشتریانی را فهرست کنید که حساب بانکی دارند، اما وام بانکی دریافت نکرده‌اند

$$\Pi_{customer_name}(depositor) - \Pi_{customer_name}(borrower)$$

عملگر ضرب دکارتی

- برای ترکیب اطلاعات هر دو رابطه دلخواه استفاده می‌شود
- ضرب دکارتی دو رابطه r و s با $r \times s$ نمایش داده می‌شود
- مثال
  - دو رابطه ***loan*** و ***borrower*** را در نظر بگیرید

loan_number	*branch_name*	*amount*
L-11	Round Hill	900
L-14	Downtown	1500
L-15	Perryridge	1500
L-16	Perryridge	1300
L-17	Downtown	1000
L-23	Redwood	2000
L-93	Mianus	500

customer_name	*loan_number*
Adams	L-16
Curry	L-93
Hayes	L-15
Jackson	L-14
Jones	L-17
Smith	L-11
Smith	L-23
Williams	L-17

عملگر ضرب دکارتی

- رابطه ***borrower × loan***

customer_name	borrower. loan_number	loan. loan_number	branch_name	amount
Adams	L-16	L-11	Round Hill	900
Adams	L-16	L-14	Downtown	1500
Adams	L-16	L-15	Perryridge	1500
Adams	L-16	L-16	Perryridge	1300
Adams	L-16	L-17	Downtown	1000
Adams	L-16	L-23	Redwood	2000
Adams	L-16	L-93	Mianus	500
Curry	L-93	L-11	Round Hill	900
Curry	L-93	L-14	Downtown	1500
Curry	L-93	L-15	Perryridge	1500
Curry	L-93	L-16	Perryridge	1300
Curry	L-93	L-17	Downtown	1000
Curry	L-93	L-23	Redwood	2000
Curry	L-93	L-93	Mianus	500
Hayes	L-15	L-11		900
Hayes	L-15	L-14		1500
Hayes	L-15	L-15		1500
Hayes	L-15	L-16		1300
Hayes	L-15	L-17		1000
Hayes	L-15	L-23		2000
Hayes	L-15	L-93		500
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
Smith	L-23	L-11	Round Hill	900
Smith	L-23	L-14	Downtown	1500
Smith	L-23	L-15	Perryridge	1500
Smith	L-23	L-16	Perryridge	1300
Smith	L-23	L-17	Downtown	1000
Smith	L-23	L-23	Redwood	2000
Smith	L-23	L-93	Mianus	500
Williams	L-17	L-11	Round Hill	900
Williams	L-17	L-14	Downtown	1500
Williams	L-17	L-15	Perryridge	1500
Williams	L-17	L-16	Perryridge	1300
Williams	L-17	L-17	Downtown	1000
Williams	L-17	L-23	Redwood	2000
Williams	L-17	L-93	Mianus	500

ترکیب عملگرهای رابطه‌ای

- مثال
  - نام‌های مشتریانی را فهرست کنید که از شعبه Perryridge وام بانکی دریافت کرده‌اند
    - پرس و جوی ۱

$$\Pi_{customer_name}(\sigma_{branch_name = \text{“Perryridge”}}(\sigma_{borrower.loan_number = loan.loan_number}(borrower \times loan)))$$

    - پرس و جوی ۲

$$\Pi_{customer_name}(\sigma_{loan.loan_number = borrower.loan_number}((\sigma_{branch_name = \text{“Perryridge”}}(loan)) \times borrower))$$

ترکیب عملگرهای رابطه‌ای

- تمرین
  - نام‌های مشتریانی را فهرست کنید که از شعبه Perryridge وام بانکی دریافت کرده‌اند، اما در آن شعبه حساب بانکی ندارند
  - نام‌های مشتریانی را فهرست کنید که از شعبه Perryridge وام بانکی دریافت کرده‌اند، اما در هیچ شعبه‌ای حساب بانکی ندارند

جبر رابطه‌ای

- عملگرهای اضافی
  - اشتراک (Intersect)
    - ∩
  - پیوند طبیعی (Natural Join)
    - ⋈
  - تقسیم (Division)
    - ÷

عملگر اشتراک

- برای این که اشتراک دو رابطه امکان‌پذیر باشد
  - هر دو رابطه باید تعداد خصیصه‌های یکسانی داشته باشند
  - دامنه خصیصه‌های متناظر در هر دو رابطه باید یکسان باشد
- اشتراک دو رابطه r و s با $r \cap s$ نمایش داده می‌شود

$$r \cap s = r - (r - s)$$

- مثال
  - نام‌های مشتریانی را فهرست کنید که هم حساب بانکی و هم وام بانکی دارند

$$\Pi_{customer_name}(depositor) \cap \Pi_{customer_name}(borrower)$$

عملگر پیوند طبیعی

- پرس و جوهایی که در آن‌ها یک عمل انتخاب برروی نتیجه یک عمل ضرب دکارتی اعمال می‌شود را می‌توان ساده‌تر بیان کرد
  - عملگر پیوند طبیعی ($\bowtie$)
- مثال
  - نام‌های مشتریانی که وام بانکی دریافت کرده‌اند را به همراه شماره وام و میزان وام آن‌ها فهرست کنید

$$\Pi_{customer_name,\ loan.loan_number,\ amount} (\sigma_{borrower.loan_number = loan.loan_number} (borrower \times loan))$$

$$\Pi_{customer_name,\ loan_number,\ amount} (borrower \bowtie loan)$$

عملگر پیوند طبیعی

- مثال
  - دو رابطه زیر را در نظر بگیرید

A	*B*	*C*	*D*
α	1	α	*a*
β	2	γ	*a*
γ	4	β	*b*
α	1	γ	*a*
δ	2	β	*b*

$\bowtie$

B	*D*	*E*
1	*a*	α
3	*a*	β
1	*a*	γ
2	*b*	δ
3	*b*	ξ

=

A	*B*	*C*	*D*	*E*
α	1	α	*a*	α
α	1	α	*a*	γ
α	1	γ	*a*	α
α	1	γ	*a*	γ
δ	2	β	*b*	δ

عملگر پیوند طبیعی

- مثال
  - نام‌های مشتریانی را فهرست کنید که در شعبه‌های Downtown و Uptown دارای حساب بانکی هستند

$$\Pi_{customer_name}(\sigma_{branch_name = \text{“Downtown”}}(depositor \bowtie account)) \cap$$
$$\Pi_{customer_name}(\sigma_{branch_name = \text{“Uptown”}}(depositor \bowtie account))$$

  - نام‌های مشتریانی را فهرست کنید که هم حساب بانکی و هم وام بانکی دارند

$$\Pi_{customer_name}(borrower \bowtie depositor)$$

عملگر پیوند طبیعی

- تمرین
- برای مشتریانی که حساب بانکی دارند و در شهر Harrison زندگی می‌کنند، نام‌های شعبه‌های بانکی‌شان را فهرست کنید

عملگر تقسیم

- برای پرس و جوهایی مناسب است که در آن‌ها عبارت "برای همه" وجود دارد
- فرض کنید r و s دو رابطه روی شِماهای R و S باشند

$$R = (A_1, \ldots, A_m, B_1, \ldots, B_n)$$
$$S = (B_1, \ldots, B_n)$$

- رابطه $r \div s$ رابطه‌ای روی شِمای $R - S = (A_1, \ldots, A_m)$ است

$$r \div s = \{ t \mid t \in \prod_{R-S}(r) \wedge \forall u \in s, tu \in r \}$$

که در آن tu چندگانه‌ای است که از الحاق چندگانه‌های t و u به دست آمده است

عملگر تقسیم

- مثال
  - دو رابطه زیر را در نظر بگیرید

A	*B*	*C*	*D*	*E*
α	a	α	a	1
α	a	γ	a	1
α	a	γ	b	1
β	a	γ	a	1
β	a	γ	b	3
γ	a	γ	a	1
γ	a	γ	b	1
γ	a	β	b	1

÷

D	*E*
a	1
b	1

=

A	*B*	*C*
α	a	γ
γ	a	γ

عملگر تقسیم

- مثال
  - نام‌های مشتریانی را فهرست کنید که در همه شعبه‌های مستقر در شهر Brooklyn حساب بانکی دارند
    - نام‌های همه شعبه‌های مستقر در Brooklyn

$$\Pi_{branch_name}\ (\sigma_{branch_city = \text{“Brooklyn”}}\ (branch))$$

    - نام‌های مشتریان و نام‌های شعبه‌هایی که مشتریان در آن‌ها حساب بانکی دارند

$$\Pi_{customer_name,\ branch_name}\ (depositor \bowtie account)$$

    - پرس و جوی نهایی

$$\Pi_{customer_name,\ branch_name}\ (depositor \bowtie account) \div \Pi_{branch_name}\ (\sigma_{branch_city = \text{“Brooklyn”}}\ (branch))$$

عملگر تقسیم

- نتیجه پرس و جوی نهایی

customer_name	*branch_name*
Hayes	Perryridge
Johnson	Downtown
Johnson	Brighton
Jones	Brighton
Lindsay	Redwood
Smith	Mianus
Turner	Round Hill

÷

branch_name
Brighton
Downtown

=

customer_name
Johnson