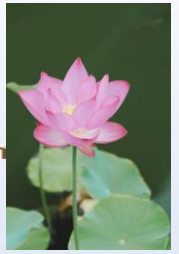
2.7 Native Relational Operations



1. The Projection

[1]Def 2.7.1 ***Projection***: Head(R) = A1, . . . , An

T := R[Aj, . . . , Ak]

- T is subset of R with Limit columns & remove duplicate rows.
- 参考书：R中选出若干属性组成新关系，从 ***“列”*** 角度操作.

$$\pi_A(R) = \{t[A] \mid t \epsilon R\}$$

[2]Example 2.7.1: (p48) List all customer names from the customers.

CN := CUSTOMERS [CNAME]

Table CN

CNAME
TIP TOP
Basics
Allied
ACME

Cus

cid	cname		dicsnt
c001	Tip Top	D	10.00
c002	Basics	D	12.00
c003	Allied	D	8.00
c004	ACME	D	8.00
c005	ACME	Kyoto	0.00

2.7 Native Relational Operations

2. The Selection

[1] Definition 2.7.2 ***Selection*** **Head(R) = A1, . . . , An**

R where *C*

C **is a logical condition that can be determined from the values of a single row of *C*.**

✓Ai *a* Aj : where *a* is a constant with " <, >, =, <=, >=, <> "

✓*C* also can be withed form : C and C', C or C', not C

✓字符数据比较：左边起第一个不相同字符的ASCII码比较。

- 参考书：从R中选择满足给定条件的元组。

  "行" 角度的运算, F为选择条件式. 也称限制(Restriction).

$$\sigma_F(R) = \{t \mid t \epsilon R \wedge F(t) = true\}$$

2.7 Native Relational Operations

[2]Example 2.7.2 (p50) Find all products stored in Dallas that cost more than $0.50

PRODUCTS where city = 'Dallas' and price >= 0.5

$\sigma_{city='Dallas'\ and\ price>=0.5}(Products)$

Products				
pid	pname	city	quantity	price
p01	comb	Dallas	111400	0.50
p02	brush	Newark	203000	0.50
p03	pen	Duluth	150600	1.00

[3]Example : Give cid and cname of all customers living in Dallas with discount greater than 8.

T1 := CUSTOMERS where city = 'Dallas' and discnt > 8

T2 := T1[cid, cname]

T :=(CUSTOMERS where city = 'Dallas' and discnt > 8) [cid, cname]

$\pi_{cid,cname}(\sigma_{city='Dallas'\ and\ discnt>8}(customers))$

Customers			
cid	cname	city	dicsnt
c001	Tip Top	Duluth	10.00
c002	Basics	Dallas	12.00
c003	Allied	Dallas	8.00
c004	ACME	Duluth	8.00
c005	ACME	Kyoto	0.00

2.7 Native Relational Operations

[4]Example 2.7.3 (p50) **Want to retrieve all pairs of agents, both with a percentage commission at least 6%, & both stationed in same city.**

- ✓ *Find out* agents with a percentage commission at least 6%

  M := AGNETS where percent >= 6

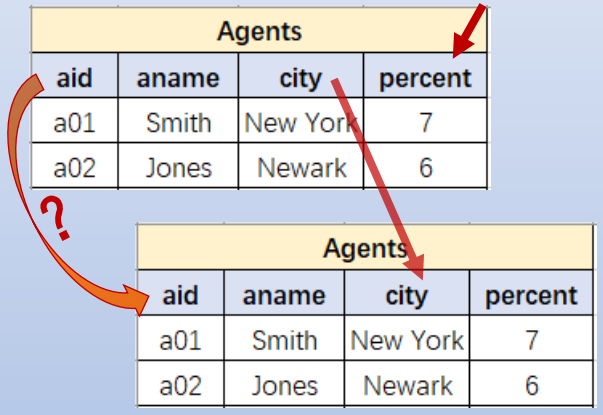
- ✓ *Find out* agents who stationed in the same city

  L := AGNETS where percent >= 6

  PAIRS := M x L where L.city = M.city

- ✓ *Except* product by themselves, So a simply form

  PAIRS := M x L where L.city = M.city and L.aid < M.aid



In Other form. $M := \sigma_{percent \geq 6}(agents)$

Solution-1 better

$L := M$

$P := \sigma_{M.city=L.city\ and\ M.cid<L.cid}(M \times L)$

2.7 Native Relational Operations

Solution-2
Bad ?

M × L

Agents			
aid	aname	city	percent
a01	Smith	New York	7
a02	Jones	Newark	6

Join

Agents			
aid	aname	city	percent
a01	Smith	New York	7
a02	Jones	Newark	6

same(=, >,<,.....)

3. The Join

[1]Def 2.7.4 ***Join***：R×S中“R的A属性组值”与“S的B属性组值”比较，挑出满足“*tr[A] θ ts[B]*”的元组，构成新关系.

$$R \underset{A\theta B}{\infty} S = \{t_r \frown t_s \mid t_r \in R \wedge t_s \in S \wedge t_r[A]\ \theta\ t_s[B]\}$$

特1：***Equijoin***，选A、B属性组中值相等的元组构成新关系.

$$R \underset{A=B}{\infty} S = \{t_r \frown t_s \mid t_r \in R \wedge t_s \in S \wedge t_r[A] = t_s[B]\}$$

特2：***Natural Join***，等值连接的基础上把重复列去掉.

$$R \infty S = \{t_r \frown t_s \mid t_r \in R \wedge t_s \in S \wedge t_r[A] = t_s[B]\}$$

2.7 Native Relational Operations

R	
A	B1
a1	b1
a1	b2
a2	b1

S	
B1	C
b1	c1
b1	c2
b1	c3
b2	c4

R join S on R.B1>S.B1			
A	B1	B1	C
a1	b2	b1	c1
a1	b2	b1	c2
a1	b2	b1	c3

R join S on R.B1=S.B1			
A	B1	B1	C
a1	b1	b1	c1
a1	b1	b1	c2
a1	b1	b1	c3
a2	b1	b1	c1
a2	b1	b1	c2
a2	b1	b1	c3
a2	b2	b2	c4

R natural join S		
A	B1	C
a1	b1	c1
a1	b1	c2
a1	b1	c3
a2	b1	c1
a2	b1	c2
a2	b1	c3
a2	b2	c4

3. The Join

[1]Def 2.7.4 ***Join***: R×S中“R的A属性组值”与“S的B属性组值”比较，挑出满足“*tr[A] θ ts[B]*”的元组，构成新关系.

$$R \underset{A\theta B}{\infty} S = \{t_r \frown t_s | t_r \in R \wedge t_s \in S \wedge t_r[A]\ \theta\ t_s[B]\}$$

特1：***Equijoin***，选A、B属性组中值相等的元组构成新关系.

$$R \underset{A=B}{\infty} S = \{t_r \frown t_s | t_r \in R \wedge t_s \in S \wedge t_r[A] = t_s[B]\}$$

特2：***Natural Join***，等值连接的基础上把重复列去掉.

$$R \infty S = \{t_r \frown t_s | t_r \in R \wedge t_s \in S \wedge t_r[A] = t_s[B]\}$$

2.7 Native Relational Operations

【综上有】： Cartesian Product ⊃ Join ⊃ Equijoin ⊃ Natural Join 真子集

[2]另一说法： Cartesian Product、Equijoin、Natural Join区别

1) *自然连接* 一定是*等值连接*，但*等值连接*不一定是*自然连接*。

2) *等值连接* 分量值相等，不一定是公共属性；而 *自然连接* 要求相等的分量须是公共属性(相容).

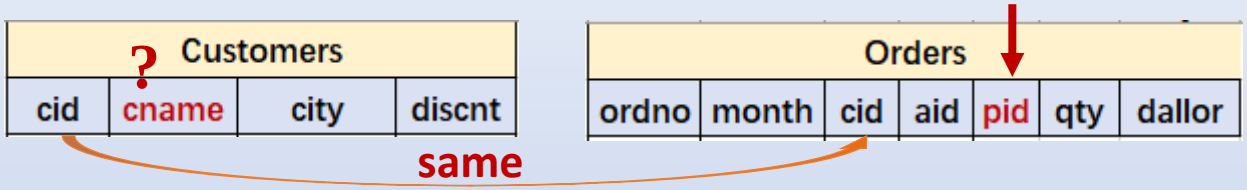
3) *自然连接* 两个关系中做比较的分量须是相同的属性组，且连接结果中要去掉重复的属性列。

[3]Example 2.7.6 (p54) Join 与 其它关系运算的关系：

R、S两个表没有公共列时	$R \infty S = R \times S$
R、S两个表的列完全一致时	$R \infty S = R \cap S$
R、S两个表有公共列A时	$R \infty S = (R \times S)$ where R.A=S.A

2.7 Native Relational Operations

[4]Example 2.7.7 (p55) Wish to find the names of customers who have ordered product p01.



T1 := Customers ∞ Orders

T2 := T1 where pid = 'p01'

CP01 := T2[cname]s

All rows

Customers.cid = Order.cid

选 rows ↓

T1 := Orders where pid = 'P01'

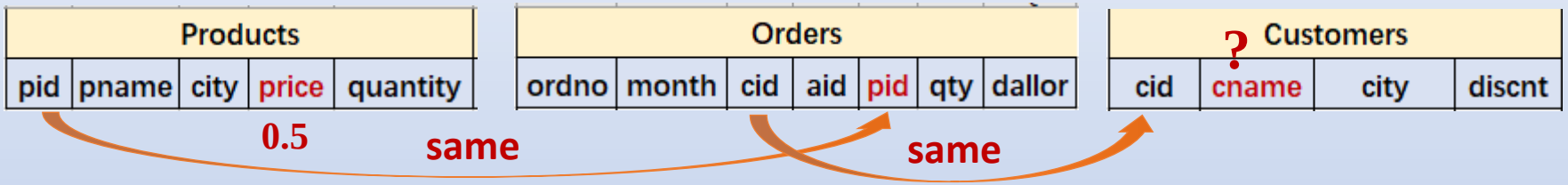
T2 := Customers ∞ T1

Better Solution

CP01 := (Customers ∞ (Orders where pid = 'p01')) [cname]

2.7 Native Relational Operations

[5]Example 2.7.8 (p56) Wish to pose a query to get cname who order at least one product costing $0.50.



Solution P1 := (PRODUCTS where price = 0.50) [pid]

O1 := (ORDERS ∞ P1) [cid]

C1 := (O1 ∞ CUSTMOTERS) [cname]

Order.pid=P1.pid

O1.cid=Customer.cid

Summary C1 := ((PRODUCTS where price=0.50) [pid])

∞ ORDERS ∞ CUSTOMERS) [cname]

2.7 Native Relational Operations

提问：如果查询所有学生选课的成绩(包括学生姓名，课程名称，成绩)

$$\Pi_{Sname,\ Cname,\ Score}(SC \bowtie Course \bowtie Student)$$

选课SC。但姓名？课名？

Student

S#	Sname	Ssex	Sage	D#	Sclass
98030101	张三	男	20	03	980301
98030102	张四	女	21	03	980301
98030103	张五	男	19	03	980301
98040201	王三	男	18	04	980402
98040202	王四	男	21	04	980402
98050104	林六	女	19	05	980501

SC

S#	C#	Score
98030101	001	92.0
98030101	002	85.0
98030101	003	88.0
98040202	002	90.5
98040202	003	80.0
98040202	001	55.0
98050104	003	56.0
98030102	001	54.0
98030102	002	85.0
98030102	003	48.0

Course

C#	Cname	Chours	Credit	T#
001	数据库	40	6	001
003	数据结构	40	6	003
002	高等数学	80	12	004

课堂练习：用英文教材中的关系运算符表达方式，查询选修数据库课的学生成绩表

学号、课程号可替换成sno、cno来表达

2.7 Native Relational Operations

[6] Relationship between "Product" and "Join" (结合律、交换律)

- $(R \times S) \times T = R \times (S \times T)$; $(R \infty S) \infty T = R \infty (S \infty T)$;
- $R \times S = S \times R$; $R \infty S = S \infty R$;

4. Precedence of Relational Operations (p53)

Precedence	*Operators*	*Symbols*
Highest	PROJECT	R[A]
	SELECT	R where C
	TIMES	×
	JOIN, DIVIDEBY	∞, ÷
	INTERSECTION	∩
Lowest	UNION, DIFFERENCE	∪, −

2.7 Native Relational Operation

R		S	
X	Y	Y	Z
x1	y5	y2	z1
x2	y2	y2	z3
x1	y2	y5	z1

5. The Division Operation

[1] Definition 2.7.5 ***Division*** Consider two tables R and S

$$\mathbf{Head}(R) = X1 \dots Xn\ Y1 \dots Ym = X, Y$$

$$\mathbf{Head}(S) = Y1 \dots Ym\ Z1 \dots Zp = Y, Z$$

X→Y的象集

Y的投影集

P
X
?

则关系R(X,Y)和S(Y,Z)做除运算记为

$$P(X) = R(X,Y) \div S(Y,Z) = \{t_r[X] | t_r \in R \wedge W_X(R) \supseteq \pi_Y(S)\}$$

- 象集$W_X(R)$：关系R(X, Y)中当tr[X] =x1时，x1在R中的象集为R[X]上值为x1的诸元组, 在Y属性值的子集.

$$W_X(R) = \{t_r[Y] \mid t_r \in R, t_r[X] = x1\}$$

- S上的投影：

$$\pi_Y(S) = S[Y]$$

Wx1(R)	{y2,y5}	
Wx2(R)	{y2}	{y2,y5}

2.7 Native Relational Operations

P	
A	B
?	?

S
c1

[2] Example 2.7.9 (p58) 两种解法:

- Sol-1：书中的解法
- Sol-2：R(A,B,C)、S(C)，则R(A,B,C)÷S(C) = P(A,B)

a) 求S与R有*公共属性的投影*，即：$\pi_C(S) = \{c1\}$

b) 求R上*非公共属性的象集* $W_{(A,B)}(R)$, 因(A, B)在R中有三种取值, 故有3个象集:

$W_{(a1,b1)}(R) = \{c_1, c_5\}$

$W_{(a2,b1)}(R) = \{c_1, c_2\}$

$W_{(a1,b2)}(R) = \{c_1, c_2, c_3, c_4\}$

R		(a,b)→c
A	B	C
a1	b1	c1
a2	b1	c1
a1	b2	c1
a1	b2	c2
a2	b1	c2
a1	b2	c3
a1	b2	c4
a1	b1	c5

c) 检查 $W_{(A,B)}(R) \supseteq \pi_C(S)$ 成立者，得:

$$P(A,B) = R(A,B,C) \div S(C) = \{(a_1, b_1), (a_2, b_1), (a_1, b_2)\}$$

2.7 Native Relational Operations

[3] Example 2.7.10 (p59) We want to extract (select) the customers who have placed orders for all products that ordered by C006.

a) Products that ordered by customer C006

PC6 := (ORDERS where cid = 'C006') [pid]

b) All Order Projection: CP := ORDERS [cid, pid]

c) who have placed orders for all products that ordered by C006.

T := CP ÷ PC6 →→ T[cid]

Orders						
ordno	month	cid	aid	pid	qty	dollars
1011	jan	c001	a01	p01	1000	450
1019	feb	c001	a02	p02	400	180
1017	feb	c002	a06	p03	600	540
1013	jan	c006	a03	p03	1000	880
1026	may	c006	a05	p01	800	704

cid→购物象集

c001	p01
	p02
c002	p03
c006	p03
	p01

÷

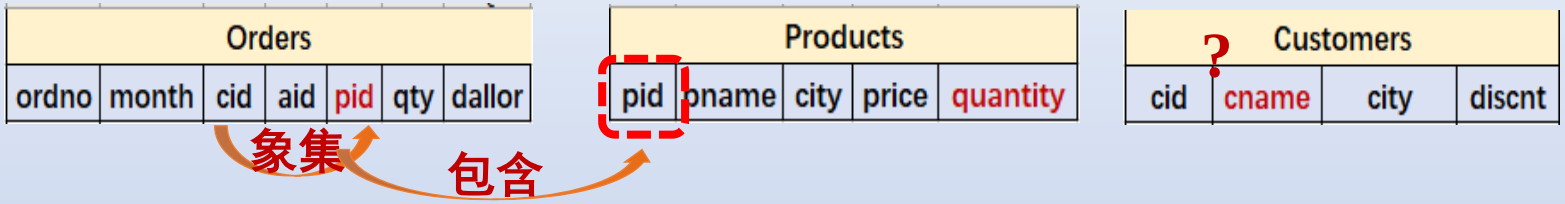
c001
c006

c006→购物集

找买了
C006所购全部商品
的顾客

2.7 Native Relational Operations

[4] Example 2.7.11 (p60) **Find the customers Who have placed orders for all the products.**



a) All products of pid P := PRODUCTS [pid]

b) All orders CP := ORDERS [cid, pid]

c) summary ((CP ÷ P) ∞ CUSTOMER[cid]) [CNAME]

2.7 Native Relational Operations

[4] Example

a) All Course of c# P := Course [C#]

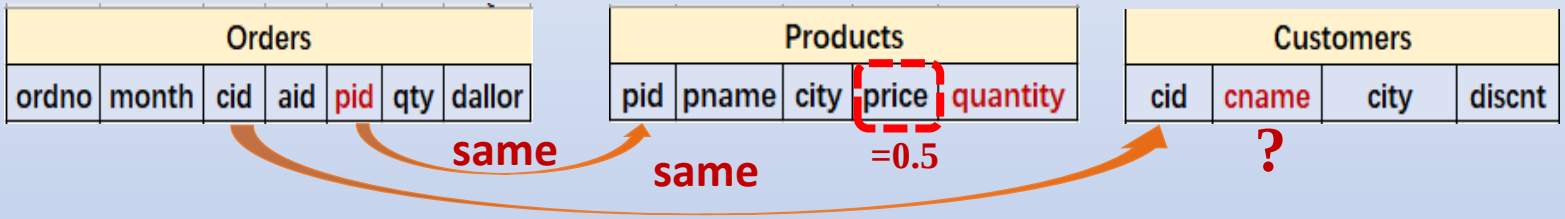
b) All C# of each Student CP := SC [S#, C#]

c) summary *T*:= (CP ÷ P)

2.8 Illustrative Examples

Set alias for CAP are C:=Customers, A:=Agents, P:=Products, O:=Orders

1. **Example 2.9.1** We want to get the names of customers who order at least one product priced at $0.50 .



a) What Products have priced at $0.50?

T1:= (Products where price=0.50)[pid]

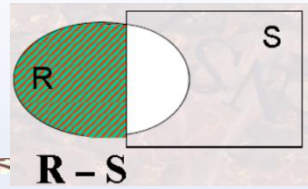
b) Who ordered these product? T2 := Orders[cid, pid] ∞ T1

c) What name of them? T := (T2 ∞ Customers[cid]) [cname]

Summary C:=Customers, A:=Agents, P:=Products, O:=Orders

T:= (P where price=0.50) ∞ O[cid, pid] ∞ C[cid]

2.8 Illustrative Examples



2. Example 2.9.2 Get cids of all customers who do not place any order through agent a03.

? Orders						
ordno	month	cid	aid	pid	qty	dallor

except

Orders						
ordno	month	cid	aid	pid	qty	dallor
1025	may	c002	a03	p02	800	450

a) a03 placed whose order ? T := (O where aid = 'a03') [cid]

b) cid that don't place any order through a03?

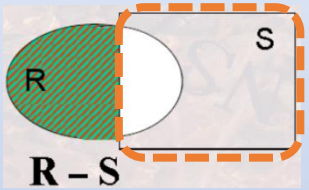
O [cid] - T 即 O[cid] - (O where aid = 'a03') [cid]

或 ((O where aid <> 'a03') [cid] ∞ Customers) [cid, cname]

2.8 Illustrative Examples

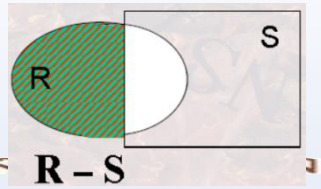
3.Example 2.9.3 Get cids who place orders only through agent a03.

(1)Only through a03 ? *(2)R: through a03 & other* *(3)S: not through a03*



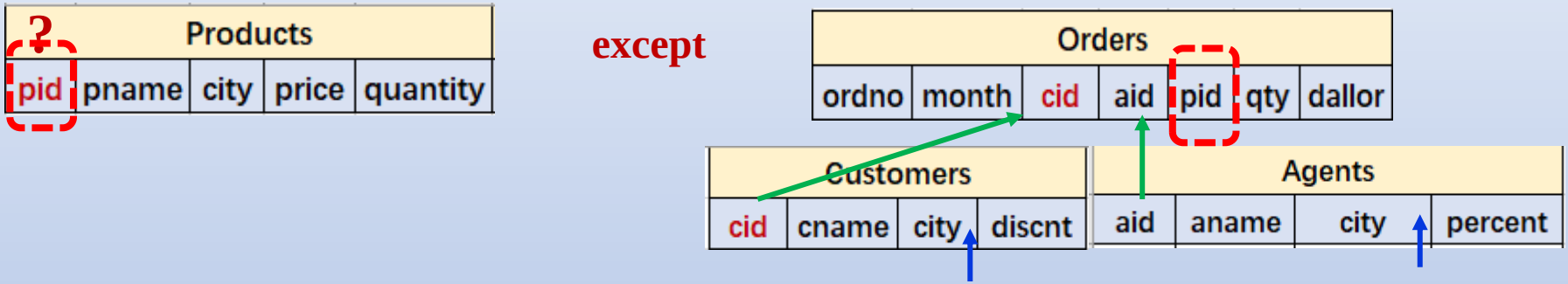
O [cid] - 'not T' 即 O[cid] - (O where aid <> 'a03') [cid]

2.8 Illustrative Examples



4.Example 2.9.4 Want to find products that have never been ordered by a customer based in "New York" through an agent based in "Boston".

(找pids，它们从未被New York的顾客通过Boston代理商订购过)



a) which products have been ordered by a customer based in "New York" through an agent based in "Boston"?

$$temp := ((C\ where\ city =' New\ York')[cid] \infty O \infty A\ where\ city =' Boston')[pid]$$

b) which products have never been ordered by a customer based in "New York" through an agent based in "Boston"?

$$P[pid] - temp$$

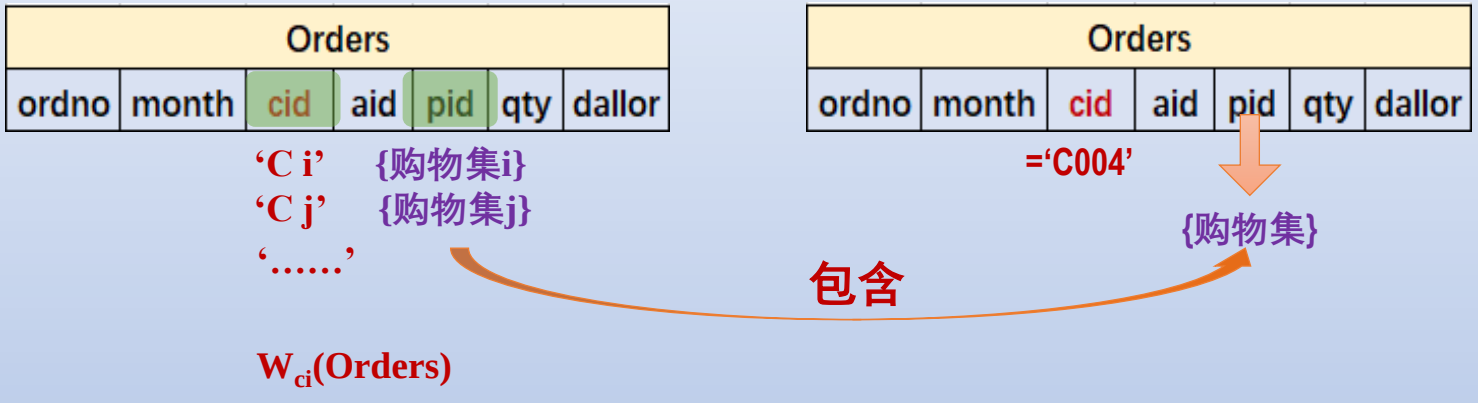
2.8 Illustrative Examples

c004订购商品集

包含

cid所购商品象集

5.Example 2.9.7 Get cids of customers who take orders on at least that set of products ordered by c004. (找aids，他们所作订单至少包含了c004顾客所订的产品)



a) which products are ordered by c004 ?

A := (O where cid = 'c004') [pid]

b)who take orders at least that set of products ordered by c004?

T:=O[cid,pid]÷A 即 O[aid,pid]÷(O where cid='c004')[pid]

2.8 Illustrative Examples

Orders表结构：
1个订货记录仅购1种商品

6.Example 2.9.8(1) Get cids of customers who order p01 and p07.

(O where pid = 'p01')[cid] ∩ (O where pid = 'p07')[cid]

Wrong : (O where pid = 'p01' *and* pid = 'p07')[cid]

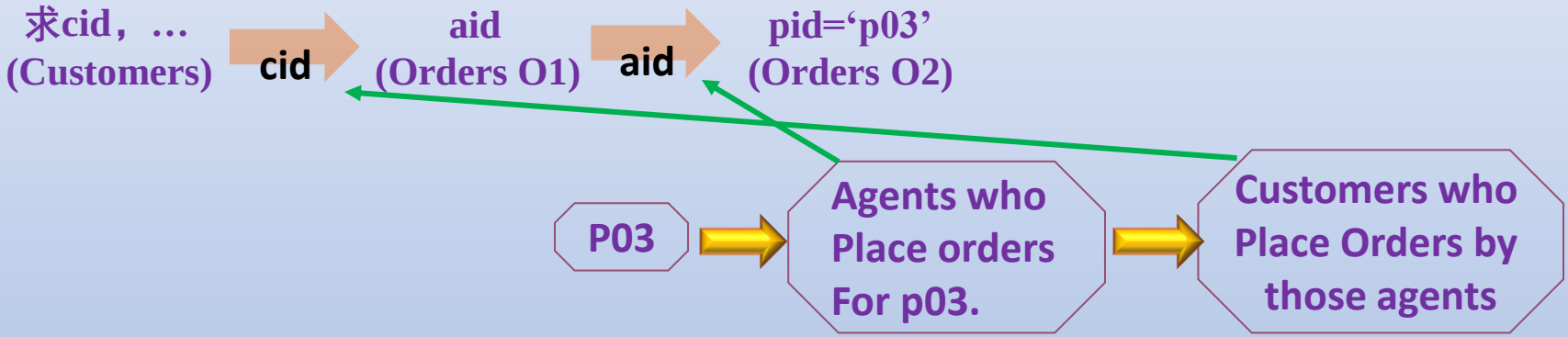
7. Example 2.9.8(2) Get cids of customers who order p01 or p07.

T:= (O where pid = 'p01' or pid = 'p07') [cid]

2.8 Illustrative Examples

8. Example 7 Get cids of customers who place an order through at least one agent who placed an order for product p03.

(找cids，他们的代理商销售过p03产品)



$$O1 := Orders$$
$$O2 := Orders$$

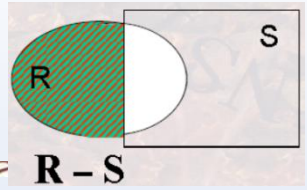
a) The agent who placed an order for product p03.

$$PC3 := (O2\ where\ pid =' p03')[aid]$$

b) Customers who place an order through these agents.

$$X := (PC3 \infty O1)[cid, cname]$$ (aid 等值连接)

2.8 Illustrative Examples



9.Example 2.9.12 Retrieve pids of all products that are not ordered by any customers living in a city beginning with the letter "D".

a) the sense of "products not ordered"

$$P[pid] - (O \infty (C \text{ where } \ldots))[pid]$$

b) a condition for living in city beginning with the letter "D".

$$T1 := C \text{ where } C.city \geq' D' \text{ and } C.city <' E'$$

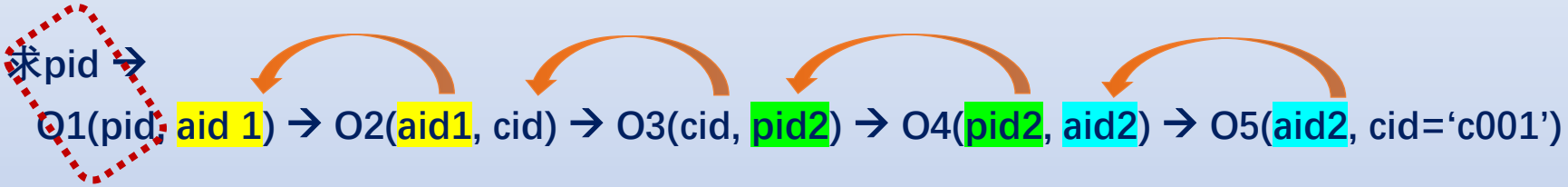
c) customers living in city beginning with the letter "D" to buy the product.

$$T2 := (T1 \infty O)[pid]$$

Summary: (cid 等值连接) $T := P[pid] - (T1 \infty O)[pid]$

2.8 Illustrative Examples

10. Example 2.9.11 List pids of products that are ordered through agents who placed an order for (different) customers who order at least one product from an agent who has placed an order for c001.



a) Agent who have placed orders for c001? (为c001服务的代理商)

$$A := (O \ where \ cid =' c001')[aid]$$

b) Customer who order product through these agent ? (被(a)服务过的顾客集)

$$X := orders \quad Y := X \quad Z := X \quad B := (X \infty A)[cid]$$ (aid等值连接)

c) Agent who have placed orders for b)'s cid ? (为(b)服务过的全体销售商)

$$C := (Y \infty B)[aid]$$ (cid等值连接, 投影出aid)

d) Products ordered through these agent are? (被(c)销售商代理的全部产品)

$$D := (C \infty P)[pid]$$

2.8 Illustrative Examples

1. Outer Join （P68）

NAME	*SYMBOL*	*EXAMPLE*
OUTER JOIN	or R OUTERJ S	$R \infty_{O} S$
LEFT OUTER JOIN	or R LOUTERJ S	$R \infty_{RO} S$
RIGHT OUTER JOIN	or R ROUTERJ S	$R \infty_{LO} S$
THETA JOIN	or R JN(∞) S	$R \infty_{A>B} S$

[1] Example: outer join. Given two tables R and S

R

A	B
a1	b1
a2	b2
a3	b3

S

B	C
b1	c1
b2	c2
b4	c3
b5	c4

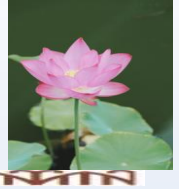
$R \infty_{LO} S$

A	B	C
a1	b1	c1
a2	b2	c2
a3	b3	null

$R \infty_{O} S$

A	B	C
a1	b1	c1
a2	b2	c2
a3	b3	null
null	b4	c3
null	b5	c4

2章 Review



实体-联系图 → 关系模型 → 关系结构 / 关系操作 / 操作约束

(对象-业务的抽象)

(形式化数据模型)

(模式定义)

关系操作 → 4种原生RA操作 / 4种扩展RA操作

4种原生RA操作 / 4种扩展RA操作 → 5种基本RA操作 / RA操作的优先级 / RA操作的等价性

5种基本RA操作 / RA操作的优先级 / RA操作的等价性 → DB模式上查询的RA表达 → SQL编程

操作约束 → 候选键，主键，主属性，非主属性(如,电话号,学号) / 外键 (如，学籍-选课记录，员工-家眷) / 其它约束 (如，课程成绩)

候选键，主键，主属性，非主属性(如,电话号,学号) / 外键 (如，学籍-选课记录，员工-家眷) / 其它约束 (如，课程成绩) → 实体完整性 / 参照完整性 / 用户自定义完整性 → 关系规则 → Rule1: 1NF / Rule2: 内容访问 / Rule3: row唯一 / Rule4: 实体完整/主键非空