

Slides based on "Database Management Systems" 3rd ed, Ramakrishnan and Gehrke

Relational Algebra



- ightarrow Selects a subset of rows from relation
- $ightarrow rac{{
 m Projection}}{\pi} \,\, \pi\,\,$ Deletes unwanted columns from relation
- <u>Cross-product</u> X Allows us to combine several relations
- Join Combines several relations using conditions
- ▶ <u>Division</u> ÷ A bit more complex, will cover later on
- ▶ <u>Set-difference</u> <u>Union</u> U <u>Intersection</u> ∩
- <u>Renaming</u> P Helper operator, does not derive new result, just renames relations and fields

$\rho(R(F),E)$

F contains oldname →newname pairs

Operator Precedence

In decreasing order of priority:

- 1. <u>Selection</u> σ <u>Projection</u> π
- 2. Cross-product χ Join \bowtie
- 3. <u>Set-difference</u> <u>Intersection</u> ∩

4. <u>Union</u> U

÷

	Sailors					Boats	
<u>sid</u>	sname	rating	age		bid	name	color
22	dustin	7	45.0		101	interlake	red
31	lubber	8	55.5		101	clipper	green
58	rusty	10	35.0		105	cupper	green

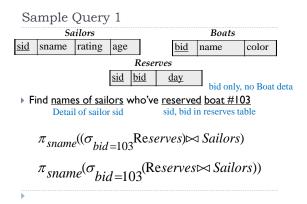
Reserves

<u>sid</u>	<u>bid</u>	<u>day</u>
22	101	10/10/96
58	103	11/12/96

Join Reserves 🖂 Sailors

R	Reserves				Sailors				
<u>sid</u>	<u>bid</u>	<u>day</u>		<u>sid</u>	sname	rating	age		
22	101	10/10/	96	22	dustin	7	45.0		
58	103	11/12/	96	31	lubber	8	55.5		
Reser	ves 🖂 Sa	uilors		58	rusty	10	35.0		
sid	bid	day	sname	rating	age				
22	101	10/10/9 6	dustin	7	45.0	_			
58	103	11/12/9 6	rusty	10	35.0				

Each sid in Reserves is filled out with Sailor attributes



Example Schema

	Sı	ailors				Boats		
sid	sname	rating	age		bid	name	color	
22	dustin	7	45.0		101	interlake	red	
31	lubber	8	55.5		101	clipper	green	
58	rusty	10	35.0		105	enpper	Siccu	
	Reserves							
		si	d bid	dav	/			

		<u></u>
22	101	10/10/96
58	103	11/12/96

Sa	ample	Query	y 2						
	Sa	ilors					Boats		
sid	sname	rating	age			bid	name	color	
			ŀ	Reser	rves				
			sid b	id	d	ay			
► F	 Find names of sailors who've reserved a red boat Detail of sailor sid sid, bid Detail of boat bid 								
$\pi_{sname}(\pi_{sid}((\pi_{bid}(\sigma_{color='red'}B)) \bowtie R) \bowtie S)$									
π	sname((o	color≓	red ^{, Bo}	oats)	\bowtie	Rese	rves⊠ Sai	lors)	

Sample Query 2

Find names of sailors who've reserved a red boat Detail of sailor sid sid, bid... Detail of boat bid

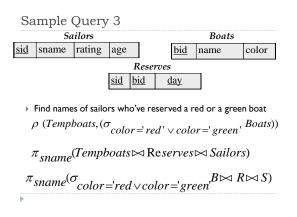
One way that's right:

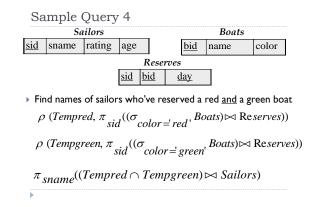
 $\pi_{sname}^{((\sigma_{color='red'}Boats)\bowtie Reserves\bowtie Sailors)}$

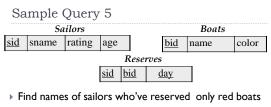
... but this next is Wrong!: Watch out for precedence! $\pi_{sname}\sigma_{color='red'}Boats \bowtie Reserves \bowtie Sailors$

 $(\pi_{sname}(\sigma_{color='red'}Boats)) \bowtie Reserves \bowtie Sailors$

empty!







 ρ (Tempred, π_{sid} (($\sigma_{color='red'}$ Boats) \bowtie Reserves))

 ρ (Tempothers, π_{sid} (($\sigma_{color} <>'red'^{Boats}$) \bowtie Reserves))

 $\pi_{sname}((Tempred-Tempothers) \bowtie Sailors)$

Time to try it yourself...

- > Try the exercises on the handed-out sheet
- You can confer with neighbors—this is not graded
- Turn in completed paper for the 3 points
- Lab sheet (Solution) (Solution posted later)
- Note: you need to attend class to get credit for this work—it is a form of class participation.

An Example of Self-Joins

Sailors

<u>sid</u>	sname	rating	age
22	dustin	7	45.0
31	lubber	8	55.5
58	rusty	10	35.0

►	Find	sailors	with	maximum	age

- No max operator in RA... (SQL has this)
- Need a trick: self join with "left" age smaller than "right" age
- This will list rows for all ages for left side but the max age row(s)
- ▶ Then use all-sailors this list

An Example of Self-Join: cross-product with rows eliminated by condition

sid1	sname1	rating1	age1	sid2	sname2	rating2	age2
22	ductin	7	45.0	52	ductin	7	45.0
22	dustin	7	45.0	31	lubber	8	55.5
22	dustin	7	45.0	58	rusty	10	35.0
31	lubber	8	55 5	22	dustin	7	45.0
31	1 1 1	0	33.3	21	1 1 1	,	55.5
51	lubber	0	55.5	<u> 91</u>	luober	8	55.5
31	lubber	8	55.5	58	rusty	10	35.0
58	rusty	10	35.0	22	dustin	7	45.0
58	rusty	10	35.0	31	lubber	8	55.5
58	rusty	10	35.0	58	rustv	10	35.0
	n condit	ion:"lof		<u> </u>	er than "	-	

An Example of Self-Join: Max ages

 ρ (S1,Sailors)

 ρ (S2,Sailors)

$$\rho(TempJoin(1 \rightarrow f1,2 \rightarrow f2,3 \rightarrow f3,4 \rightarrow f4),$$

$$S1 \bowtie_{S1.age < S2.age} S2)$$

 ρ (LeftHalf, $\pi_{f1,f2,f3,f4}$ TempJoin)

 Finally, subtract the resulting left hand side from the initial relation, and you get sailors with maximum ages

Final result is

Sailors-LeftHalf

More on Natural Joins

Natural Joins match all same-named columns

Consider two tables T1 and T2:

T1(<u>id1</u>, attr1, city) T2(<u>id2</u>, id1, attr2, city)

- \blacktriangleright Probably want to join on id1, a key for T1 showing up in both tables
- But TI and T2 have id1 and city in common, so a natural join TI ⋈ T2 matches both
- If we don't want non-key columns matched like this
 - We can use a theta join with an explicit condition: T1 ⋈_{r1.idl=r2.idl} T2
 - Or project out city before one of the joins

Consider the Example Schema, modified to have a name attribute for two entities

	/ Sa	ilors		 _	Boats	
<u>sid</u>	name	rating	age	bid	name	color
22	dustin	7	45.0		interlake	red
31	lubber	8	55.5	101	clipper	green
58	rusty	10	35.0	105	enpper	green

Reserves								
sid	<u>bid</u>	<u>day</u>						
22	101	10/10/96						
58	103	11/12/96						

Sample Query 2 on modified schema

sid bid

Find names of sailors who've reserved a red boat

Reserves

Old solution for unmodified schema: returns empty table here
 π_{sname}((σ_{color='red}, Boats) ⋈ Reserves ⋈ Sailors)
 Here we can project out boat names before join to Sailors
 π_{name}(((π_{bid} σ_{color='red}, B)⋈ R) ⋈ S)

<u>day</u>

Boats

color

bid name

Sailors

rating age

sid name

ĥ

Sample Query 2 on modified schema

Sailors					Boats				
sid	name	rating	age		<u>bid</u>	name	color		
	Reserves								
			sid bio	<u>1</u> d	lay				
		L			-	_			
► Fi		es of sail of sailor				ed a red bo			
	Old solution:								
π_s	$\pi_{sname}((\sigma_{color='red'}^{Boats}) \bowtie \text{Reserves} \bowtie \text{ Sailors})$								
Returns an empty relation!									
► lť	It's looking for matches on name as well as bid, sid								

A (1 10°°	sid	sname	rating	age
Another self join:	28	yuppy	9	35.0
Close competitors	31	lubber	8	55.5
	44	guppy	5	35.0
	58	rusty	10	35.0

 Find pairs of sailors (sids) with ratings that differ by no more than one.

$\rho(S1, Sailors) \rho(S2, Sailors)$	sid1	sid2
ρ (TempJoin(1 \rightarrow sid1,5 \rightarrow sid2),	28	28
$S1 \mapsto S1 \xrightarrow{S1 \times ating \leq S2 \times ating + 1^S1 \times ating \geq S2 \times ating - 1} S2$	28	31
$\pi_{sid1,sid2}^{TempJoin}$		58
sid1, sid2	31	28
We don't want a lot of these results		

	sid	sname	rating	age	
Another self join	28	yuppy	9	35.0	
	31	lubber	8	55.5	
	44	guppy	5	35.0	
	58	rusty	10	35.0	

Find pairs of different sailors (sids) with ratings that differ by no more than one, listing each unordered pair once.

 ρ (S2,Sailors) ρ (S1, Sailors)

 $\rho(TempJoin(1 \rightarrow sid1, 5 \rightarrow sid2),$ $S1 \bowtie_{S1 rating <= S2 rating -1 \land S1 . sid < S2 . sid} S2)$

$\pi_{sid1,sid2}^{TempJoin}$	sid1	sid2	1
Stat, sta 2That's better!	28	31	
	28	58	

Lil	ke Que	ery 0							
Sailors					Boats				
<u>sid</u>	sname	rating	age		<u>bid</u> name color				
	Reserves								
			sid bi	id c	lay_				
 Find <u>colors of boats</u> with <u>names starting with C</u> Detail of Boat Another Boat detail (assume lowercase names) 									
$\pi_{color}^{((\sigma_{name \ge c'^{name < d'}}Boats)}$									
	π_{co}	lor ⁰ n	ame>	='c'^n	ame	<'d' ^{Boat}	S		
•									

Lil	ke Que	ery 3								
	Sa	ilors				Boats				
sid	sname	rating	age			bid name color				
			ŀ	Reser	rves					
			sid b	id	<u>d</u>	la <u>y</u>				
ו פ ק	boat. List sailor ren (<i>Tempb</i>	names a ted both oats, (σ	nd the color color=	boat boat <i>! rec</i>	t col ts) l' ∨	lor (tv <i>coloi</i>	a red or a wo rows if r='green' ⋈ Sailor	the Boats))		

•