Relational Algebra (Part IV)

R & G, Chapter 4

Join + Set Operations Quiz

- **Quiz of last lecture: Find the name of sailors with age over 20 who have not reserved a red boat**
  - Find sids of sailors with age over 20 as set S1
  - Find sids of sailors who have reserved a red boat as set S2
  - Take the set difference of S1 and S2
  - Join with sailors, and return name of the sailors from join result

– Answer:

$$\rho(S1,\pi_{sid}(\sigma_{age>20}(Sailors)))$$

$$\rho(S2,\pi_{sid}((\sigma_{color='red'}(Boats))\bowtie Reserves)$$

$$\pi_{sname}(Sailors\bowtie (S1-S2))$$

Find the name of sailors with age over 20 who have not reserved a red boat

- **Solution 1**

$$\rho(T1, \pi_{sid}(\sigma_{age>20}(Sailors) \bowtie Boats \bowtie Reserves))$$

$$\rho(T2, \pi_{sid}(\sigma_{color='red'}(Boats) \bowtie Reserves \bowtie Sailors))$$

$$\pi_{sname}(Sailors \bowtie (T2 - T1))$$

- Is this solution correct?
- This solution is wrong because T1 does not include those sailors (e.g., Lubber) who have not reserved any boat

Sid	Bid	day
22	101	10/10/96
58	103	11/12/96

Reserves

Bid	Bname	Color
101	Interlake	Blue
102	Interlate	Red
103	Clipper	Green
104	Marine	red

Boats

Sid	Sname	Rating	Age
22	Dustin	7	45.0
31	Lubber	8	55.5
58	Rusty	10	35.5

Sailors

Find the name of sailors with age over 20 who have not reserved a red boat

- **Solution 2**

$$\pi_{sname}(\sigma_{age>20}(Sailors) \bowtie \sigma_{color \neq 'red'}(Boats) \bowtie Reserves)$$

  - Is this solution correct?
  - This solution is wrong because it may return the sailors (e.g., Dustin) who have reserved a red boat and a non-red boat.

Sid	Bid	day
22	101	10/10/96
58	103	11/12/96
22	102	12/10/96

Reserves

Bid	Bname	Color
101	Interlake	Blue
102	Interlate	Red
103	Clipper	Green
104	Marine	red

Boats

Sid	Sname	Rating	Age
22	Dustin	7	45.0
31	Lubber	8	55.5
58	Rusty	10	35.5

Sailors

Division

- Useful for expressing "for all" queries like:
  Find sids of sailors who have reserved all boats.
- Division A/B: Attributes of B MUST be a subset of attrs of A.

Sno	pno
S1	P1
S1	P2
S1	P3
S1	P4
S2	P1
S2	P2
S3	P2
S4	P2
S4	P4

A

Pno
P2
P4

B

Sno
S1
S4

A/B

Division (I)

- **Find the names of sailors who've reserved all boats**
  - The way of thinking:
    - find the set of of sailors (with sid, bid) who have reserved some boats.
    - find the set B of all boats (with bid).
    - take A / B to get (sid of) sailors who have reserved all boats

Why do we need the projections on Reserves & Boats?

$$\rho(Tempsids,(\pi_{sid,bid}(\text{Re}serves))/(\pi_{bid}(Boats)))$$

$$\pi_{sname}(Tempsids \bowtie Sailors)$$

Division (II)

- **Find the names of sailors who've reserved all 'Interlake' boats**
  - The way of thinking:
    - find the set A of sailors (with sid, bid)who have reserved some boats.
    - find the set B of all 'Interlake' boats (with bid).
    - take A / B to get (sid of) sailors who have reserved all 'Interlake' boats

$$\rho(T1,(\pi_{sid,bid}(\mathrm{Reserves}))/\pi_{bid}(\sigma_{bname='Interlake'}(Boats)))$$

$$\pi_{sname}(T1 \bowtie Sailors)$$

Division (III)

- **Find the names of sailors who have reserved boats of all colors**
  - The way of thinking:
    - find the set A of sailors (with sid, color)who have reserved some boats.
    - find the set B of all colors of boats.
    - take A / B to get (sid of) sailors who have reserved boats of all colors
- **Solution**

$$\pi_{sname}((\pi_{sid,color}(\text{Re}serves \bowtie Boats)/\pi_{color}(Boats)) \bowtie Sailors)$$

Aggregate Queries by Relational Algebra

- **Relational algebra does not provide any aggregate function in general.**

- **Relational algebra handles aggregate queries in a complicated way**

Aggregate Queries

- **Find the names of sailors who have reserved at least two different boats**
  - The way of thinking:
    - *Find sailors S1 who reserved at least one boat*
    - *From S1, find sailors who reserved at least two boats*
    - *Select sailors with two different boats reserved*

Sailors who have reserved at least ONE boat →

$$\rho(R,(\pi_{sid,sname,bid}(Reserves \bowtie Sailors))$$

Sailors who have reserved at least TWO boats →

$$\rho(RPairs,(1\text{->}sid1,2\text{->}sname1,3\text{->}bid1,4\text{->}sid2,5\text{->}sname2,6\text{->}bid2),R \bowtie_{Sid1=Sid2} R)$$

Renaming to eliminate duplicate attributes

$$\pi_{sname1}(\sigma_{bid1\neq bid2} RPairs)$$

Same sailor but two different boats

Running Example (1/3)

Reserves

Sid	Bid	day
22	101	10/10/96
58	103	11/12/96
22	102	12/10/96

Boats

Bid	Bname	Color
101	Interlake	Blue
102	Interlate	Red
103	Clipper	Green
104	Marine	red

Sailors

Sid	Sname	Rating	Age
22	Dustin	7	45.0
31	Lubber	8	55.5
58	Rusty	10	35.5

$$\rho(R,(\pi_{sid,sname,bid}(Reserves \bowtie Sailors)$$

Sailors who have reserved at least ONE boat

Sid	Sname	bid
22	Dustin	101
22	Dustin	102
58	Rusty	103

R

Running Example (2/3)

Sid	Sname	bid
22	Dustin	101
22	Dustin	102
58	Rusty	103

R

$$\rho\,(RPairs,(1\text{->}sid1,2\text{->}sname1,3\text{->}bid1,4\text{->}sid2,\ 5\text{->}sname2,6\text{->}bid2),R \times R)$$

$Sid1=Sid2$

Sid1	Sname1	bid1	Sid2	Sname2	bid2
22	Dustin	101	22	Dustin	101
22	Dustin	101	22	Dustin	102
22	Dustin	102	22	Dustin	101
22	Dustin	102	22	Dustin	102
58	Rusty	103	58	Rusty	103

RPairs

Running Example (3/3)

Sid1	Sname1	bid1	Sid2	Sname2	bid2
22	Dustin	101	22	Dustin	101
22	Dustin	101	22	Dustin	102
22	Dustin	102	22	Dustin	101
22	Dustin	102	22	Dustin	102
58	Rusty	103	58	Rusty	103

RPairs

$$\pi_{sname1}\ (\sigma_{bid1 \neq bid2}\ RPairs)$$

- **Returns "Dustin"**

Further Thinking

- **Find the names of sailors who have reserved at least two different boats**

Sailors who have reserved at least ONE boat →

$$\rho(R,(\pi_{sid,sname,bid}(Reserves \bowtie Sailors))$$

Sailors who have reserved at least TWO boats →

$$\rho(RPairs,(1->sid1,2->sname1,3->bid1,4->sid2,5->sname2,6->bid2),R \times R)$$

Renaming to eliminate duplicate attributes

$$\pi_{sname1}(\sigma_{bid1 \neq bid2} RPairs)$$

Same sailor but two different boats

- Questions:

(1) Can we remove 'sname' from R?

(2) Can we change equal-join in Rpairs to be cross-product?

(3) How can we write the query: Find the names of sailors who have reserved at least three different boats