# Today

- Review Relational Algebra
- Class Participation Exercise (time permitting)

What is an "Algebra"?

- Mathematical system consisting of:
  - *Operators*: symbols denoting procedures that construct new values from given values (eg. +, -, *, /)
  - *Operands*: variables or values from which new values can be constructed. An operand is a value that is acted upon by an operator (eg. 3 + 7, 3 and 7 are operands)

What is Relational Algebra?

- An algebra whose operands are relations or variables that represent relations

- Input: 1 or 2 instances of relation. Output: an instance of output relation.

- Form basis for other query languages (eg. SQL) and implementation

1 or 2 relations → Operator → New relation
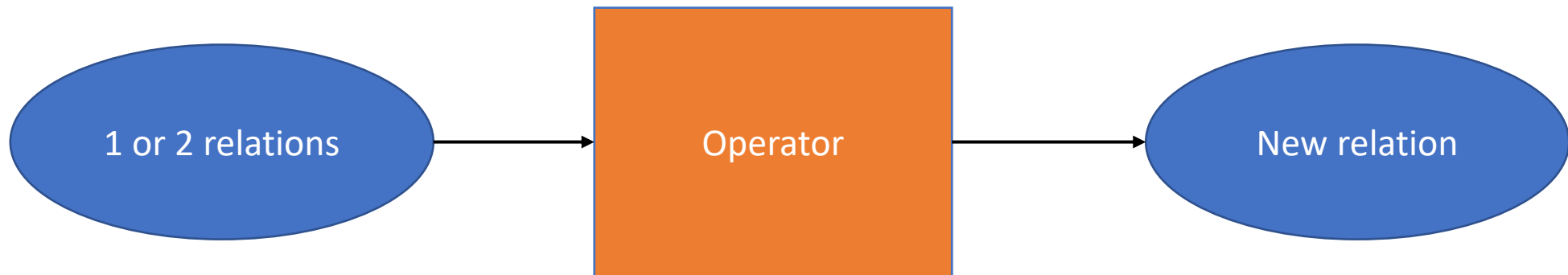
Operators

- Set Operations:
  - Union
  - Intersection
  - Difference
  - Cross Product (Cartesian Product)
- Selection
- Projection
- Renaming
- Joins
- *Division*

Schema

Sailors(<u>sid</u>: integer, sname: string, rating: integer, age: real)

Boats(<u>bid</u>: integer, bname: string, color: string)

Reserves (<u>sid</u>: integer, <u>bid</u>: integer, day: date)

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.0
58	Rusty	10	35.0

Instance S1 of Sailors

sid	sname	rating	age
28	yuppy	9	35.0
31	Lubber	8	55.0
44	guppy	5	35.0
58	Rusty	10	35.0

Instance S2 of Sailors

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

Instance R1 of Reserves

Union

- $R \cup S$ returns an instance of a relation containing all tuples that occur in *either R* or *S*

- *R* and *S* must be ***union-compatible***
  - Same number of fields
  - Correponding fields, taken in order, must have the same domains

- By convention, inherits names from *R*

- No duplicates

Union

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.0
58	Rusty	10	35.0

Instance S1 of Sailors

sid	sname	rating	age
28	yuppy	9	35.0
31	Lubber	8	55.0
44	guppy	5	35.0
58	Rusty	10	35.0

Instance S2 of Sailors

sid	sname	rating	age
22	Dustin	7	45.0
28	yuppy	9	35.0
31	Lubber	8	55.0
44	guppy	5	35.0
58	Rusty	10	35.0

S1 U S2

Intersection

- $R \cap S$ returns an instance of a relation containing all tuples that occur in *both R* or *S*

- *R* and *S* must be union-compatible

- By convention, inherits names from *R*

- No duplicates

Intersection

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.0
58	Rusty	10	35.0

Instance S1 of Sailors

sid	sname	rating	age
28	yuppy	9	35.0
31	Lubber	8	55.0
44	guppy	5	35.0
58	Rusty	10	35.0

Instance S2 of Sailors

sid	sname	rating	age
31	Lubber	8	55.0
58	Rusty	10	35.0

$S1 \cap S2$

Set-Difference

- R – S returns all tuples that occur in R but not S
- Must be union-compatible
- By convention, inherits names from *R*

Set-Difference

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.0
58	Rusty	10	35.0

Instance S1 of Sailors

sid	sname	rating	age
28	yuppy	9	35.0
31	Lubber	8	55.0
44	guppy	5	35.0
58	Rusty	10	35.0

Instance S2 of Sailors

sid	sname	rating	age
22	Dustin	7	45.0

$S1 - S2$

Cross Product

- R3 = R1 X R2
  - Pair each tuple t1 of R1 with each tuple t2 of R2
  - Concatenation t1, t2 is a tuple in R3
  - Schema of R3 is the attributes of R1 followed by the attributes of R2, in order
  - Beware of attributes having the same name in R1 and R2: use R1.A and R2.A
  - Sometimes called "Cartesian Product"
  - Total number of tuples is equal to |R1| x |R2| where |R| is the number of tuples in instance relation R

Cross Product

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.0
58	Rusty	10	35.0

Instance S1 of Sailors

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

Instance R1 of Reserves

S1.sid	sname	rating	age	R1.sid	bid	day
22	Dustin	7	45.0	22	101	10/10/96
22	Dustin	7	45.0	58	103	11/12/96
31	Lubber	8	55.0	22	101	10/10/96
31	Lubber	8	55.0	58	103	11/12/96
58	Rusty	10	35.0	22	101	10/10/96
58	Rusty	10	35.0	58	103	11/12/96

S1 X R1

Selection σ

- $R1 = \sigma_c(R2)$
  - C is a condition (as in "if" statements) that refers to attributes of R2
  - R1 returns all tuples of R2 that satisfy C
  - Analogous to WHERE clause in SQL

Selection σ

sid	sname	rating	age
28	yuppy	9	35.0
31	Lubber	8	55.0
44	guppy	5	35.0
58	Rusty	10	35.0

Instance S2 of Sailors

sid	sname	rating	age
28	yuppy	9	35.0
58	Rusty	10	35.0

$$\sigma_{rating > 8}(S2)$$

Projection π

- $R1 = \pi_L(R2)$
  - *L* is a list of attributes from the schema of R2
  - R1 is constructed by looking at each tuple of R2, extracting the attributes on list L, in the order specified, and creating from those components a tuple from R1
  - Eliminate duplicates, if any
  - Analogous to SELECT clause

Projection π

sid	sname	rating	age
28	yuppy	9	35.0
31	Lubber	8	55.0
44	guppy	5	35.0
58	Rusty	10	35.0

Instance S2 of Sailors

age
35.0
55.0

$$\pi_{age}(S2)$$

sname	rating
yuppy	9
Rusty	10

$$\pi_{sname,rating}(\sigma_{rating > 8}(S2))$$

Joins

- Conditional Joins: R3 = R1 $\bowtie_c$ R2
  - Take the product of R1 X R2
  - Then apply σ_c to the result
  - ie. $R3 = \sigma_c(R1\ x\ R2)$
  - Result of cross product is usually much larger than result of join, so most implementations do not materialize the underlying cross product

Joins

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.0
58	Rusty	10	35.0

Instance S1 of Sailors

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

Instance R1 of Reserves

S1.sid	sname	rating	age	R1.sid	bid	day
22	Dustin	7	45.0	22	101	10/10/96
22	Dustin	7	45.0	58	103	11/12/96
31	Lubber	8	55.0	22	101	10/10/96
31	Lubber	8	55.0	58	103	11/12/96
58	Rusty	10	35.0	22	101	10/10/96
58	Rusty	10	35.0	58	103	11/12/96

$S1 \times R1$

Joins

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.0
58	Rusty	10	35.0

Instance S1 of Sailors

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

Instance R1 of Reserves

S1.sid	sname	rating	age	R1.sid	bid	day
22	Dustin	7	45.0	58	103	11/12/96
31	Lubber	8	55.0	58	103	11/12/96

$S1 \bowtie_{S1.sid < R1.sid} R1$

Equijoin

- For join conditions that contain only equalities between attributes of the two relations, there will be a duplicate column for each equality

- In this case an additional projection is done to drop this extra column

- Resulting schema contains fields of R1 followed by fields of R2 that do not appear in the join conditions

Equijoin

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.0
58	Rusty	10	35.0

Instance S1 of Sailors

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

Instance R1 of Reserves

S1.sid	sname	rating	age	bid	day
22	Dustin	7	45.0	101	10/10/96
58	Rusty	10	35.0	103	11/12/96

$S1 \bowtie_{S1.sid = R1.sid} R1$

Natural Join

- For a special case of an Equijoin where the join condition contains equalities of all attributes in common between two relations, we can simply remove join condition

- Called a **Natural Join**

- Denoted R3 = R1 ⋈ R2

- Previous $S1 \bowtie_{S1.sid = R1.sid}$ R1 is actually a Natural Join since the only common field is sid

- If two relations have no attributes in common, the natural join is simply the cross product

Renaming ρ

- ρ operator allows renaming of relation and attributes
- Useful when fields from different tables have the same name or before certain set operations
- $R' = \rho_{R'(A1 \rightarrow A1', \ldots An \rightarrow An')}(R)$
- R' is a relation with attributes A1',…An' and the same tuples as R

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

S3 (Sailors)

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

R2 (Reserves)

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

B1 (Boats)

Q1: Find the names of sailors who have reserved boat 103

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

S3 (Sailors)

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

R2 (Reserves)

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

B1 (Boats)

Q1: Find the names of sailors who have reserved boat 103

A1: $\pi_{sname}(\sigma_{bid=103}(Reserves) \bowtie Sailors)$

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

S3 (Sailors)

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

R2 (Reserves)

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

B1 (Boats)

Q2: Find the names of sailors who have reserved a red boat.

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

S3 (Sailors)

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

R2 (Reserves)

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

B1 (Boats)

Q2: Find the names of sailors who have reserved a red boat.

A2: $\pi_{sname}(\sigma_{color=red}(Boats) \bowtie Reserves \bowtie Sailors)$

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

S3 (Sailors)

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

R2 (Reserves)

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

B1 (Boats)

Q3: Find the names of sailors who have reserved a red boat.

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

S3 (Sailors)

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

R2 (Reserves)

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

B1 (Boats)

Q3: Find the names of sailors who have reserved a red or a green boat.

A3: $\rho\left(tempboats, (\sigma_{color=red}(Boats)) \cup (\sigma_{color=green}(Boats))\right)$

$\pi_{sname}(tempboats) \bowtie Reserves \bowtie Sailors)$

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

S3 (Sailors)

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

R2 (Reserves)

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

B1 (Boats)

Q4: Find the colors of boats reserved by Lubber.

Q5: Find the names of sailors who have reserved a red and a green boat.

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

S3 (Sailors)

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

R2 (Reserves)

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

B1 (Boats)

A4: $\pi_{color}(\sigma_{sname=Lubber}(Sailors) \bowtie Reserves \bowtie Boats)$

A5: $\rho(tempred, \pi_{sid}((\sigma_{color=red}(Boats)) \bowtie Reserves))$

$\rho(tempgreen, \pi_{sid}((\sigma_{color=green}(Boats)) \bowtie Reserves))$

$\pi_{sname}(tempred \cap tempgreen) \quad \bowtie Sailors)$