

RDBMS and SQL

Madhavan Mukund

<https://www.cmi.ac.in/~madhavan>

Lecture 3, 29 August 2024

Relational algebra — select and project

- List instructors from Physics department with salary above 90,000

$\sigma_{\text{dept_name} = \text{"Physics"} \wedge \text{salary} > 90000}(\text{Instructor})$

flexible syntax

ID	name	dept_name	salary
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
✓ 22222	Einstein	Physics	95000
32343	El Said	History	60000
✗ 35456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

Relational algebra — select and project

- List instructors from Physics department with salary above 90,000
- List names of instructors

$\pi_{\text{name}}(\text{Instructor})$


Project

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

Relational algebra — select and project

- List instructors from Physics department with salary above 90,000
- List names of instructors
- List names of instructors from Physics department with salary above 90,000

$\pi_{\text{name}}(\sigma_{\text{dept_name} = \text{Physics} \wedge \text{salary} > 90000}(\text{Instructor}))$



<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

Relational algebra — select and project

- List instructors from Physics department with salary above 90,000
- List names of instructors
- List names of instructors from Physics department with salary above 90,000

$\pi_{\text{name}}(\sigma_{\text{dept_name} = \text{Physics}}(\text{Instructor}))$
cannot swap
 \wedge
 $\text{salary} > 90000$

query: Table \rightarrow Table

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

Relational algebra — join

- List details of courses offered by instructors

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

<i>ID</i>	<i>course_id</i>	<i>sec_id</i>	<i>semester</i>	<i>year</i>
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Relational algebra — join

■ Instructor × Teaches

$$S \times T = \{ (s, t) \mid s \in S, t \in T \}$$

Instructor.ID	name	dept_name	salary	teaches.ID	course_id	sec_id	semester	year
✓ 10101	Srinivasan	Comp. Sci.	65000	✓ 10101	CS-101	1	Fall	2017
	10101	Srinivasan	Comp. Sci.	10101	CS-315	1	Spring	2018
	10101	Srinivasan	Comp. Sci.	10101	CS-347	1	Fall	2017
✗	10101	Srinivasan	Comp. Sci.	✗ 12121	FIN-201	1	Spring	2018
	10101	Srinivasan	Comp. Sci.	15151	MU-199	1	Spring	2018
	10101	Srinivasan	Comp. Sci.	22222	PHY-101	1	Fall	2017

	12121	Wu	Finance	10101	CS-101	1	Fall	2017
	12121	Wu	Finance	10101	CS-315	1	Spring	2018
	12121	Wu	Finance	10101	CS-347	1	Fall	2017
	12121	Wu	Finance	12121	FIN-201	1	Spring	2018
	12121	Wu	Finance	15151	MU-199	1	Spring	2018
	12121	Wu	Finance	22222	PHY-101	1	Fall	2017

Relational algebra — join

- $\sigma_{\text{Instructor.ID} = \text{teaches.ID}}(\text{Instructor} \times \text{Teaches})$

<i>Instructor.ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>	<i>teaches.ID</i>	<i>course_id</i>	<i>sec_id</i>	<i>semester</i>	<i>year</i>
10101	Srinivasan	Comp. Sci.	65000	10101	CS-101	1	Fall	2017
10101	Srinivasan	Comp. Sci.	65000	10101	CS-315	1	Spring	2018
10101	Srinivasan	Comp. Sci.	65000	10101	CS-347	1	Fall	2017
12121	Wu	Finance	90000	12121	FIN-201	1	Spring	2018
15151	Mozart	Music	40000	15151	MU-199	1	Spring	2018
22222	Einstein	Physics	95000	22222	PHY-101	1	Fall	2017
32343	El Said	History	60000	32343	HIS-351	1	Spring	2018
45565	Katz	Comp. Sci.	75000	45565	CS-101	1	Spring	2018
45565	Katz	Comp. Sci.	75000	45565	CS-319	1	Spring	2018
76766	Crick	Biology	72000	76766	BIO-101	1	Summer	2017
76766	Crick	Biology	72000	76766	BIO-301	1	Summer	2018
83821	Brandt	Comp. Sci.	92000	83821	CS-190	1	Spring	2017
83821	Brandt	Comp. Sci.	92000	83821	CS-190	2	Spring	2017
83821	Brandt	Comp. Sci.	92000	83821	CS-319	2	Spring	2018
98345	Kim	Elec. Eng.	80000	98345	EE-181	1	Spring	2017

Relational algebra — join

- Instructor \bowtie ~~Instructor.ID = teaches.ID~~ Teaches — cartesian project with select

Instructor.ID	name	dept_name	salary	teaches.ID	course_id	sec_id	semester	year
10101	Srinivasan	Comp. Sci.	65000	10101	CS-101	1	Fall	2017
10101	Srinivasan	Comp. Sci.	65000	10101	CS-315	1	Spring	2018
10101	Srinivasan	Comp. Sci.	65000	10101	CS-347	1	Fall	2017
12121	Wu	Finance	90000	12121	FIN-201	1	Spring	2018
15151	Mozart	Music	40000	15151	MU-199	1	Spring	2018
22222	Einstein	Physics	95000	22222	PHY-101	1	Fall	2017
32343	El Said	History	60000	32343	HIS-351	1	Spring	2018
45565	Katz	Comp. Sci.	75000	45565	CS-101	1	Spring	2018
45565	Katz	Comp. Sci.	75000	45565	CS-319	1	Spring	2018
76766	Crick	Biology	72000	76766	BIO-101	1	Summer	2017
76766	Crick	Biology	72000	76766	BIO-301	1	Summer	2018
83821	Brandt	Comp. Sci.	92000	83821	CS-190	1	Spring	2017
83821	Brandt	Comp. Sci.	92000	83821	CS-190	2	Spring	2017
83821	Brandt	Comp. Sci.	92000	83821	CS-319	2	Spring	2018
98345	Kim	Elec. Eng.	80000	98345	EE-181	1	Spring	2017

Natural
join

Relational algebra — set operations

- Union and intersection

Courses in 2017, 2018

$\pi_{course_id}(\sigma_{year=2017}(Teaches))$

$\pi_{course_id}(\sigma_{year=2018}(Teaches))$

Either 2017 or 2018

$\pi_c(\sigma_{\theta_1 \vee \theta_2}(r)) \equiv \pi_c(\sigma_{\theta_1}(r) \cup \sigma_{\theta_2}(r))$

Course in both years appears once

ID	course_id	sec_id	semester	year
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Teaches

Relational algebra — set operations

- Union and intersection

Both in 2017 & in 2018

$\sigma_{\theta_1 \wedge \theta_2}$?

$\pi_c(\sigma_{\theta_1}(r)) \cap \pi_c(\sigma_{\theta_2}(r))$

ID	course_id	sec_id	semester	year
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Relational algebra — set operations

- Union and intersection
- Complementation?

$\text{not}(r)$



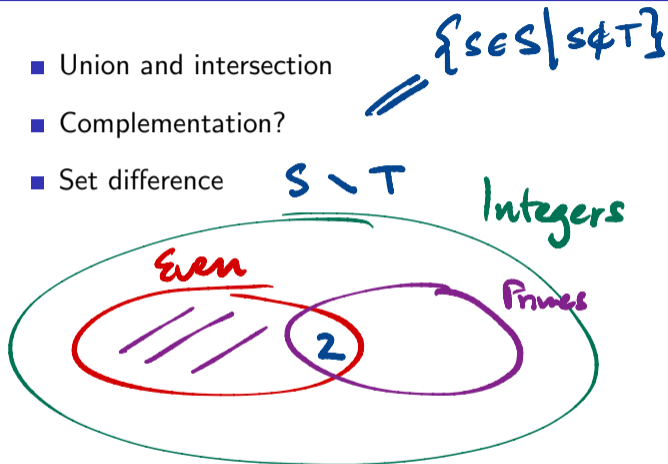
$\sigma_{\text{year}=2017}(\text{Teachers})$

All rows in table that don't
match condition

<i>ID</i>	<i>course_id</i>	<i>sec_id</i>	<i>semester</i>	<i>year</i>
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Relational algebra — set operations

- Union and intersection
- Complementation?
- Set difference

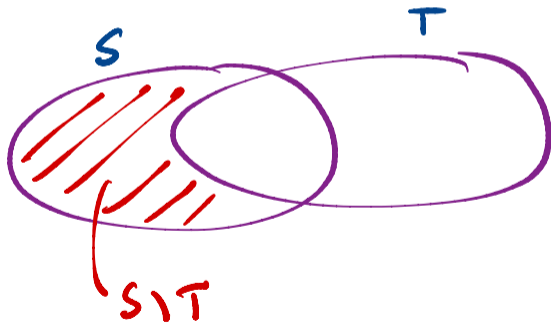


$$\text{Even} \setminus \text{Primes} = \{4, 6, 8, \dots\}$$
$$\text{Primes} \setminus \text{Even} = \{3, 5, 7, 11, \dots\}$$

ID	course_id	sec_id	semester	year
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Relational algebra — set operations

- Union and intersection
- Complementation?
- Set difference



<i>ID</i>	<i>course_id</i>	<i>sec_id</i>	<i>semester</i>	<i>year</i>
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Relational algebra — set operations

- Union and intersection
- Complementation?
- Set difference

Courses not taught in 2017

$\pi_{\text{course-id}}(\text{Teaches}) \setminus$

$\pi_{\text{course-id}}(\sigma_{\text{year}=2017}(\text{Teaches}))$

<i>ID</i>	<i>course_id</i>	<i>sec_id</i>	<i>semester</i>	<i>year</i>
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Relational algebra — set operations

- Union and intersection
- Complementation?
- Set difference

2017 but not 2018

$$\pi_c(\sigma_{\theta_1}(r)) \setminus \pi_c(\sigma_{\theta_2}(r))$$

<i>ID</i>	<i>course_id</i>	<i>sec_id</i>	<i>semester</i>	<i>year</i>
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Sets – ordering

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

Instructor

Same relation

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
22222	Einstein	Physics	95000
12121	Wu	Finance	90000
32343	El Said	History	60000
45565	Katz	Comp. Sci.	75000
98345	Kim	Elec. Eng.	80000
76766	Crick	Biology	72000
10101	Srinivasan	Comp. Sci.	65000
58583	Califieri	History	62000
83821	Brandt	Comp. Sci.	92000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
76543	Singh	Finance	80000

Instructor, unsorted

Sets – duplicates

List all depts

TL dept_name (Instructor)

Sorted by dept name

Bio
CS
CS
CS
EE
Finan
Finc
:

ID	name	dept_name	salary
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics ✗	87000
45565	Katz	Comp. Sci. ✗	75000
58583	Califieri	History ✗	62000
76543	Singh	Finance ✗	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci. ✗	92000
98345	Kim	Elec. Eng.	80000

Instructor

A challenge

- List instructors who teach more than one course

$\sigma, \pi, \bowtie, \cup, \cap, \setminus$

Self-Join

Teachers \bowtie Teachers



ID	course_id	sec_id	semester	year
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Teachers

Relational algebra — assignment, renaming

Left \leftarrow Teaches
Right \leftarrow Teachers } Assignment

Left \bowtie Right
Left-ID = Right-ID
 \wedge
Left-Course-id \neq Right-Course-id

ID	course_id	sec_id	semester	year
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Teaches

Renaming - ρ rho

ρ_{left} (Teaches)

\bowtie left-ID = Right-ID
 \wedge

left-Course-ID \neq Right-Course-ID

ρ_{right} (Teache)

ID	course_id	sec_id	semester	year
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Teaches

Renaming can also rename columns

ρ left (ID-L, course-id-L, ...) (Teaches)

ID	course_id	sec_id	semester	year
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Teaches

Relational algebra — Examples

Instructor (ID, name, dept_name, salary)

- Find all faculty members from Physics who earn more than **at least one** faculty member from Comp.Sci.

I_1, I_2 two
copies of instructor
Columns are renamed

$\pi_{\substack{\text{Id-1,} \\ \text{Name-1}}} (I_1 \bowtie I_2)$

$\theta: \text{dept_name-1} = \text{Physics}$
 \wedge
 $\text{dept_name-2} = \text{CS}$
 \wedge
 $\text{salary-1} > \text{salary-2}$

Relational algebra — Examples

Instructor (ID, name, dept_name, salary)

- Find all faculty members from Physics who earn more than **at least one** faculty member from Comp.Sci.

$I_1 = \sigma_{dept_name = Physics} (Instr.)$

$I_2 = \sigma_{dept_name = CS} (Instr.)$

$\pi_{id-1} \left(I_1 \bowtie_{salary-1 > salary-2} I_2 \right)$
Name-1

Relational algebra — Examples

Instructor (ID, name, dept_name, salary)

- Find all faculty members from Physics who earn more than **every** faculty member from Comp.Sci.

Effectively
computing max

PhyFac \ [List Phy fac with salary
≤ at least one CS fac]

`Instructor (ID, name, dept_name, salary)`

- Find the faculty member(s) with the minimum salary.

Exercise

Relational algebra — Examples

family(ID,name,gender)

relation(ID1,ID2,relationship)

parent, sibling

- Compute the relation

sister(ID1,ID2) — ID1 is a sister of ID2

$$\pi_{ID_1, ID_2} \left(\text{family} \bowtie_{\substack{ID=ID_1 \\ \wedge \\ \text{gender} = F \\ \wedge \\ \text{relationship} = \text{sibling}}} \text{relation} \right)$$

Relational algebra — Examples

family(ID,name,gender)

relation(ID1,ID2,relationship)

- Compute `grandparent(ID1,ID2)` —
ID1 is grandparent of ID2

$R_1 \leftarrow \text{relation}$
 $R_2 \leftarrow \text{relation}$

$ID_1 \rightarrow ID_3 \rightarrow ID_2$

$R_1 \bowtie R_2$

$R_1.ID_2 = R_2.ID_1$
 \wedge
 $R_1.relationship = \text{Parent}$
 \wedge
 $R_2.relationship = \text{Parent}$

Relational algebra — Examples

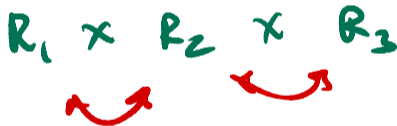
`family(ID,name,gender)`

`relation(ID1,ID2,relationship)`

- Compute

`greatgrandparent(ID1,ID2) — ID1`

is greatgrandparent of `ID2`



Relational algebra — Examples

family(ID,name,gender)

relation(ID1,ID2,relationship)

- Can you compute ancestor(ID1,ID2) in general?

How many generations?

Transitive Closure

Not possible

Not expressible

First Order Logic = Predicate Logic

$R_1 \times R_2$

$R_1 \times R_2 \times R_3$

⋮

$R_1 \times R_2 \times R_2 \dots \times R_n$