Introduction to Relational Model and Relational Algebra

Relational Model Uni DB

- intstructor(ID, name, dept_name → department, salary)
- course(id, title, dept_name → department, credits)
- prereq(course_id → course, prereq_id → course)
- department(name, building, budget)
- section(course_id, id, semester, year, (building, room_number) → classroom, time_slot_id)
- teaches(instructor_ID → instructor, (course_id, sec_id, semester, year) → section)
- student(ID, name, dept_name → department, total_credit)
- takes(student_ID → student, (course_id, section_id, semester, year) → section, grade)
- advisor(student_id → student, instructor_id → instructor)
- classroom(building, room_number, capacity)
- time_slot(id, day, start_time, end_time)

Relational Algebra

Select Operation

- Information of all instructors from the physics department:

$$\sigma_{\texttt{dept\_name="Physics"}}(\texttt{instructor})$$

```
select *
from instructor
where dept_name = 'Physics'
```

Table 1: 2 records

id	name	dept_name	salary
22222	Einstein	Physics	95000
33456	Gold	Physics	87000

- Information of all instructors with salaries greater than 90,000 $:

$$\sigma_{\texttt{salary > 90000}}(\texttt{instructor})$$

```
select *
from instructor
where salary > 90000
```

Table 2: 2 records

id	name	dept_name	salary
22222	Einstein	Physics	95000
83821	Brandt	Comp. Sci.	92000

- Information about all instructors from the physics department with salaries greater than 90000:

$$\sigma_{\texttt{dept\_name = 'Physics'} \wedge \texttt{salary > 90000}}(\texttt{instructor})$$

```
select *
from instructor
where salary > 90000 and dept_name = 'Physics'
```

Table 3: 1 records

id	name	dept_name	salary
22222	Einstein	Physics	95000

- comparison of two different attributes of the **same** relation is possible, e.g. all departments whose name is the same as their building name:

$$\sigma_{\texttt{dept\_name = building}}(\texttt{department})$$

```
select *
from department
where name = building
```

Table 4: 0 records

name	building	budget

Project Operation

- list ID, name and salary information of all instructors:

$$\Pi_{\texttt{ID, name, salary}}(\texttt{instructor})$$

```
select i.id, i.name, i.salary
from instructor i
```

Table 5: Displaying records 1 - 10

id	name	salary
10101	Srinivasan	65000
12121	Wu	90000
15151	Mozart	40000
22222	Einstein	95000
32343	El Said	60000
33456	Gold	87000
45565	Katz	75000
58583	Califieri	62000
76543	Singh	80000
76766	Crick	72000

- expressions of attributes are allowed, e.g. montly salaries:

$$\Pi_{\texttt{ID, name, salary/12}}(\texttt{instructor})$$

```
select id, name, salary / 12 as month_salary
from instructor
```

Table 6: Displaying records 1 - 10

id	name	month_salary
10101	Srinivasan	5416.667

id	name	month_salary
12121	Wu	7500.000
15151	Mozart	3333.333
22222	Einstein	7916.667
32343	El Said	5000.000
33456	Gold	7250.000
45565	Katz	6250.000
58583	Califieri	5166.667
76543	Singh	6666.667
76766	Crick	6000.000

Composition of Relational Operations

- find the names of all instructors in the Physics department

$$\Pi_{\texttt{name}}(\sigma_{\texttt{dept\_name = 'Physics'}}(\texttt{instructor}))$$

```
select name
from instructor
where dept_name = 'Physics'
```

Table 7: 2 records

name
Einstein
Gold

Cartesian (Cross) Product

let $r[R]$ and $s[S]$. If $R \cap S = \emptyset$, then $r \times s$ is simply:

$$(r \times s)[R \cup S] := \{t[R \cup S] \mid t[R] \in r \wedge t[S] \in s\}$$

If $R \cap S \neq \emptyset$, equally named attributes must be distinguished. Let

$$R \tilde{+} S := R \oplus S \bigcup_{x \in R \cap S} \{R.x, S.x\}$$

Then,

$$(r \times s)[R \tilde{+} S] := \{t[R \tilde{+} S] \mid t[(R \quad S) \cup \bigcup_{x \in R \cap S} \{R.x\}] \in t[R] \wedge t[(S \quad R) \cup \bigcup_{x \in R \cap S} \{S.x\}] \in t[S]\}$$

Problem when $r \times r$. We must use rename.

Rename Operation

A whole relation can be renamed:

$$\beta_s(r)$$

Attributes of a relation can be renamed:

$$\beta_{b_1 \leftarrow a_1, b_2 \leftarrow a_2}(r)$$

Above the attributes a_1 and a_2 of r are renamed to b_1 and b_2.

Using rename we can perform cross product of a relation with itself:

$$\beta_s(r) \times r$$

sql version:

```
select *
from r, r as s
```

Example illustrating rename:

- Find the ID and name of all instructors who earn more than the instructor whose ID is 12121:

$$\Pi_{\texttt{i.ID, i.name}}\Big(\sigma_{\texttt{i.salary > wu.salary}}(\beta_{\texttt{i}}(\texttt{instructor} \times \beta_{\texttt{wu}}(\sigma_{\texttt{id = 12121}}(\texttt{instructor})))))\Big)$$

Join Operation

Natural Join

for $r[R]$ and $s[S]$ natural join is defined as:

$$r \bowtie s := \{t[R \cup S] \mid t[R] \in r \wedge t[S] \in s\}$$

e.g. $\texttt{instructor} \bowtie \texttt{teaches}$ gives all information about instructors and courses they teach:

θ-Join

General θ-join for a predicate θ is defined as:

$$r \bowtie_\theta s := \sigma_\theta(r \times s)$$

join can be expressed in terms θ-join with appropriate rename and projection operations.

Set Operations

for relations r and s with compatible schemes R and S (compatible means same arities and corresponding domains) simply

- $\texttt{r} \cup \texttt{s}$
- $\texttt{r} \cap \texttt{s}$
- $\texttt{r} \ \texttt{s}$

examples:

- courses offered in 2017 fall semester **or** in 2018 spring semester:

$$\Pi_{\texttt{course\_id}}(\sigma_{\texttt{semester = 'Fall'} \wedge \texttt{year = 2017}}(\texttt{section})) \cup$$
$$\Pi_{\texttt{course\_id}}(\sigma_{\texttt{semester = 'Spring'} \wedge \texttt{year = 2018}}(\texttt{section}))$$

- courses offered in 2017 fall semester **and** in 2018 spring semester:

$$\Pi_{\texttt{course\_id}}(\sigma_{\texttt{semester = 'Fall'} \wedge \texttt{year = 2017}}(\texttt{section})) \cap$$
$$\Pi_{\texttt{course\_id}}(\sigma_{\texttt{semester = 'Spring'} \wedge \texttt{year = 2018}}(\texttt{section}))$$

- courses offered in 2017 fall semester **but not** in 2018 spring semester:

$$\Pi_{\texttt{course\_id}}(\sigma_{\texttt{semester = 'Fall'} \wedge \texttt{year = 2017}}(\texttt{section}))$$
$$\Pi_{\texttt{course\_id}}(\sigma_{\texttt{semester = 'Spring'} \wedge \texttt{year = 2018}}(\texttt{section}))$$

Asssignment

For convenience we can name intermediate results of relational algebraic operations, by assigning them variable names:

$$\texttt{r} := \Pi_{\texttt{course\_id}}(\sigma_{\texttt{semester = 'Fall'} \wedge \texttt{year = 2017}}(\texttt{section}))$$
$$\texttt{s} := \Pi_{\texttt{course\_id}}(\sigma_{\texttt{semester = 'Spring'} \wedge \texttt{year = 2018}}(\texttt{section}))$$
$$\texttt{r} \cup \texttt{s}$$