Database Management Systems (DBMS)

Lec 6: Relational model of data (Cont.)

Ramesh K. Jallu

IIIT Raichur

Date: 9/2/21

Recap

- Two other variations of the join operation
  - Equijoin and Natural join

- The division operation (÷)

Today's plan

- One exercise problem for division operation
- Relational Calculus

Another example for the division operation

Q: Find employee(s) SSN(s) who work on *all* projects handled by department number 5

PROJECT

Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

$\rho_{\text{TEMP}}(\text{Pno})(\pi_{\text{Pnumber}}(\sigma_{\text{Dnum}=5}(\text{PROJECT})))$

Pno
1
2
3

$\text{EMP\_DEP5} \leftarrow \Pi_{\text{Essn,Pno}}(\text{WORKS\_ON})$

EMP_DEP5

Essn	Pno
123456789	1
123456789	2
666884444	3
453453453	1
453453453	2
333445555	2
333445555	3
333445555	10
333445555	20
999887777	30
999887777	10
987987987	10
987987987	30
987654321	30
987654321	20
888665555	20

WORKS_ON

Essn	Pno	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

X = {Essn,Pno}, Y={Pno}, Z = {Essn}

$$T_1 \leftarrow \pi_Z(\text{EMP\_DEP5})$$
$$T_2 \leftarrow \pi_Z((\text{TEMP} \times T_1) - \text{EMP\_DEP5})$$
$$T \leftarrow T_1 - T_2$$

$\text{EMP\_DEP5} \div \text{TEMP}$

Result: Empty relation

Relational query languages

- **Query**: A question or an inquiry – it causes some data to be retrieved from the database
- **Query language**: a language in which a user requests information from the database
- **Categories**
  a. Imperative (or procedural) query language: Relational Algebra
  b. Declarative (or non-procedural) query language: Relational Calculus

Relational Calculus (RC)

- Like RA, RC is a language for defining new relations in a DB
- Unlike RA, RC is a ***non-procedural***
- Uses the concepts of formal logic to express relational queries
- The **expressive power** of both the languages is ***identical***
- Two well-known formulations of RC are
  - Tuple relational calculus (TRC)
  - Domain relational calculus (DRC)

Tuple relational Calculus

- A query is expressed as a set, say ***R***, {***t*** | ***P*(*t*)**}
  - ***R*** contains all the tuples ***t*** such that predicate ***P*** is true for ***t***

- We use similar notations ***t[A]*** and ***t* ∈ *R*** to denote the value of ***t*** in on attribute ***A*** and to denote tuple ***t*** is in ***R***, respectively

Example queries

- Q_1: Find all employees whose salary is above 30,000/-

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

$$\{t \mid t \in \text{EMPLOYEE} \wedge t[\text{Salary}] > 30000\}$$

$$\sigma_{\text{Salary} > 30000}(\text{EMPLOYEE})$$

Example queries

$$\{t \mid t \in \text{EMPLOYEE} \wedge t[\text{Salary}] > 30000\}$$

- The previous query selects all the attributes from EMPLOYEE
- To retrieve a specific attributes we use “there exists” from mathematical logic
- Notation: $\exists\, t \in \boldsymbol{R}(Q(t))$

Example queries

- Q_2: Find the first and last names all employees whose salary is above 30,000/-

$$\{t \mid \exists\, s \in \text{EMPLOYEE}(\ s[\text{Salary}] > 30000 \wedge t[\text{Fname}] = s[\text{Fname}] \wedge t[\text{Lname}] = s[\text{Lname}])\}$$

  - We read the preceding expression as “The set of all tuples ***t*** such that there exists a tuple ***s*** in relation **EMPLOYEE** for which the value of ***s*** for the **Salary** attribute is greater than 30,000/-, the values of ***t*** and ***s*** for the **Fname** attribute are equal, and the values of ***t*** and ***s*** for the **Lname** attribute are equal.”

Example queries

- Q_3: Retrieve the birth date and address of the employee (or employees) whose name is John B. Smith.

$\{t \mid \exists s \in \text{EMPLOYEE}(t[\text{Bdate}] = s[\text{Bdate}] \wedge t[\text{Address}] = s[\text{Address}] \wedge s[\text{Fname}] = \text{'John'} \wedge s[\text{Mint}] = \text{'B.'} \wedge s[\text{Lname}] = \text{'Smith'})\}$

- Q_4: List the first name and address of all employees who work for the ‘Research’ department

Example queries

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

$$\{t \mid \exists s \in \text{EMPLOYEE}(t[\text{Fname}] = s[\text{Fname}] \wedge t[\text{Address}] = s[\text{Address}] \wedge \exists u \in \text{DEPARTMENT}(u[\text{Dname}] = \text{'Research'} \wedge u[\text{Dnumber}] = s[\text{Dno}]))\}$$

Example queries

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

- Q_5: Retrieve the Social Security numbers of all employees who either work in department 5 or directly supervise an employee who works in department 5

$$\{t \mid \exists\, s \in \text{EMPLOYEE}(t[\text{Ssn}] = s[\text{Ssn}] \wedge s[\text{Dno}] = 5) \vee$$
$$\exists\, u \in \text{EMPLOYEE}(t[\text{Super\_ssn}] = u[\text{Super\_ssn}] \wedge u[\text{Dno}] = 5)\}$$

Example queries

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

WORKS_ON

Essn	Pno	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

- Q_6: Retrieve the names of all employees who do not work on any project
  - Equivalent relational algebra expression?

$$\{t \mid \exists\, s \in \text{EMPLOYEE}(t[\text{Ssn}] = s[\text{Ssn}]) \wedge \neg\, \exists\, u \in \text{WORKS\_ON}(t[\text{Essn}] = u[\text{Essn}])\}$$

Example queries

- Q_7: Find all employees SSNs who work on all projects handled by department number 5
  - We use implication (⇒)
  - P ⇒ Q means “if P is true, then Q must be true.”
  - P ⇒ Q is logically equivalent to ¬P ∨ Q
  - Notation: ∀ t ∈ R(Q(t)), I.e., Q is true for all tuples *t* in relation R

{*t* | ∃ *r* ∈ EMPLYOYEE(*r*[Ssn] = *t*[Ssn]) ∧ (∀ *u* ∈ PROJECT(*u*[Dnum] = 5 ⇒ ∃ *s* ∈ WORKS_ON(*t*[Essn] = *s*[EssN] ∧ *s*[Pno] = *u*[Pnumber]))}

Example queries

$$\{t \mid \exists\, r \in \text{EMPLYOYEE}(r[\text{Ssn}] = t[\text{Ssn}]) \wedge (\forall\, u \in \text{PROJECT}(u[\text{Dnum}] = 5 \Rightarrow \exists\, s \in \text{WORKS\_ON}(t[\text{Essn}] = s[\text{Essn}] \wedge s[\text{Pno}] = u[\text{Pnumber}]))\}$$

- We interpret this expression as “The set of all employees (i.e., (Ssn) tuples *t*) such that, for all tuples *u* in the PROJECT relation, if the value of *u* on attribute Dnum is 5, then there exists a tuple in the WORKS_ON relation that includes the employee Essn and the project number.”

Thank you!