COS221
L12 - Relational Calculus
(Chapter 6 in Edition 6 and Chapter 8 in Edition 7)

Linda Marshall

24 March 2023

Overview

- Relational Calculus is another formal query language for the relational model.
- Two variations exist
  - tuple relational calculus, and
  - domain relational calculus
- In both variations, a declarative expression is written to specify a retrieval request.
- No description of how, or in what order to evaluate the query is given, just what is to be retrieved. Relational calculus is therefore a non-procedural language. That is, unlike with relational algebra where steps were given, no steps are given here.

Overview

- Queries written in relational algebra can be written in relational calculus and vice versa. This means the expressive power of the two languages is the same, leading to the definition of relational completeness.
- A relational query language L is relationally complete if any query expressed in L can be expressed in relational calculus.
- Relational calculus has a basis in mathematical logic and SQL has some of its foundations in tuple relational calculus

Tuple Relational Calculus

- Tuple relational calculus specifies a number of *tuple variables*. A tuple variable, ranges over a database relation - the *range relation*. That is, it gets its value from any individual tuple in the relation.
- A query has the general form:
  $\{t | COND(t)\}$
  where t is a tuple variable and $COND(t)$ is a boolean expression involving t. The result of the query is all the tuples that satisfy the condition, that is the selected combinations satisfy the query. A set of attributes can be retrieved and are referred to as the *requested attributes*.

Tuple Relational Calculus

- For example:
  $\{t|EMPLOYEE(t)$ AND $t.Salary > 50000\}$
  where the range relation of the query is $EMPLOYEE(t)$ and the condition to be satisfied is $t.Salary > 50000$
- This query retrieves **all** the attributes of each selected EMPLOYEE tuple t. The following query selects specific attributes:
  $\{t.Fname, t.Lname|EMPLOYEE(t)$ AND $t.Salary > 50000\}$

Tuple Relational Calculus

- The general expression for tuple relational calculus is therefore given A condition or formula is made up of atoms which can be one of the following:
  $\{t_1.A_j, t_2.A_k, ..., t_n.A_m | COND(t_1, t_2, ..., t_n, t_{n+1}, t_{n+2}, ..., t_{n+m})\}$
  where A_i is an attribute on which t_i , the tuple variable, ranges and $COND$ is a condition/formula.
- A condition or formula is made up of atoms which can be one of the following:
  - $R(t)$
  - $(t_i.A$ op $t_j.B)$ where op is a comparison operator, t_i and t_j are tuple variables and A and B are attributes on which the tuple variable respectively range.
  - $(t_i.A$ op $c)$ or $(c$ op $t_j.B)$ where c is a constant.

Tuple Relational Calculus

- Each atom evaluates to a truth value.
- A formula is made up of one or more atoms connected with logical operator AND, OR and. NOT.
- The following rules are defined recursively:
  - Rule 1 - Every atom is a formula
  - Rule 2 - For two formulae, F_1 and F_2, (F_1 AND F_2), (F_1 OR F_2), (NOT F_1) and (NOT F_2) are also formulas and adhere to the standard boolean truth values.

Tuple Relational Calculus

- Two additional special quantifier symbols, the universal quantifier ($\forall$) and the existential quantifier ($\exists$) are defined.
- A tuple variable, t, is bound if it appears in an ($\exists t$) or ($\forall t$), otherwise it is *free*. The following rules determine whether t is free or bound:
  - tuple variables that are atoms are free
  - by adding logical connectors (AND, OR and NOT) the status (free or bound) of the tuple variable participating in the connection dictates the status of the tuple variable in the connection.
  - All free occurrences of t in F are bound in F' for $F' = (\exists)(F)$ or $F' = (\forall t)(F)$

Tuple Relational Calculus

- For example:
  $F_1 : d.Dname = \text{'Research'}$
  $F_2 : (\exists t)(d.Number = t.Dno)$
  $F_3 : (\forall d)(d.Mgr\_ssn =' 333445555'$
  from F_1 and F_2, d is free, in F_2, t is bound to $(\exists)$, and in F_3, d is bound to $(\forall)$
- The following rules can now be added:
  - Rule 3 - If F is a formula, then so is $F' = (\exists t)(F)$. F' is true if F evaluates to true for at least one tuple assigned to free occurrences of t in F.
  - Rule 4 - If F is a formula, then so is $F' = (\forall t)(F)$. F' is true if F evaluates to true for every tuple assigned to free occurrences of t in F.

Transforming the Universal and Existential Quantifiers

- The following equivalences exist
  - $(\forall x)(P(x)) \equiv NOT(\exists x)(NOT(P(x))$
  - $(\exists x)(P(x)) \equiv NOT(\forall x)(NOT(P(x))$
- The following implications are true
  - $(\forall x)(P(x)) \implies (\exists x)(P(x))$
  - $NOT(\exists x)(P(x)) \implies NOT(\forall x)(P(x))$

Safe and Unsafe Expressions

- Expressions in relational calculus must be safe. That is, it must yield a finite number of tuples in its result.
- The following example is unsafe, it yields all tuples in the universe that are not in EMPLOYEE.
  $\{t|NOT(EMPLOYEE(t))\}$

Domain Relational Calculus

- Domain calculus was proposed after the development of QBE (Query-by-Example).
- Domain relational calculus differs from tuple relational calculus in that instead of variables ranging over tuples, variables range over single values from domains of attributes.
- To form a relation of degree n from a query, n domain variables are required - one per attribute.
- An expression in domain relational calculus takes the form:
  $\{x_1, x_2, ..., x_n | COND(x_1, x_2, ..., x_n, x_{n+1}, x_{n+2}, ..., x_{n+m}\}$
- The properties of a condition or formula are similar to those of Tuple Relational Calculus.
- Note, when referring to relations, the commas can be dropped. That is:
  $\{x_1, x_2, ..., x_n | R(x_1, x_2, ..., x_n)$ AND ... $\} \equiv$
  $\{x_1, x_2, ..., x_n | R(x_1 x_2 ... x_n)$ AND ... $\}$

Domain Relational Calculus

- An example: *List the birth date and address of employee whose name is 'John B Smith'*
  $\{u, v \mid (\exists q)(\exists r)(\exists s)(\exists t)(\exists w)(\exists x)(\exists y)(\exists z)(EMPLOYEE(qrstuvwxyz)$
  $q = \text{'John'}$ AND $r = \text{'B'}$ AND $s = \text{'Smith'})\}$
  - 10 variables are required for the EMPLOYEE relation. Of these 10, only u and v are free.
  - A shorthand notation for this query is:
    $\{u, v \mid (EMPLOYEE(\text{'John'},\text{'B'},\text{'Smith'}, t, u, v, w, x, y, z)\}$

Examples

Using the COMPANY database to write the following queries in RA, RC (tuple and domain)

- Query 1: Retrieve the name and address of all employees who work for the 'Research' department.
- Query 2: For every project located in 'Stafford', list the project number, the controlling department number, and the departments manager's last name, address and birth date.
- Query 3: Find the names of employees who work on all the projects controlled by department number 5.
- Query 4: Make a list of project numbers for projects that involve an employee who's last name is 'Smith', either as a worker or as a manager of the department that controls the project.
- Query 5: List the names of all employees with two or more dependents.
- Query 6: Retrieve the names of employees who have no dependents.
- Query 7: List the names of managers who have at least one

Examples

Query 1: Retrieve the name and address of all employees who work for the 'Research' department.

Figure 3.6
One possible database state for the COMPANY relational database schema.

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

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

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

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

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

SQL: SELECT

RA: R1 ← σ Dname = 'Research' (DEPARTMENT)

R2 ← R1 ⋈ Dnumber = Dno EMPLOYEE

R ← π Fname, Minit, Lname, Address (R2)

(R1)

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	...	...

T_RC:

{t.Fname, t.Lname, t.Minit, t.Address | EMPLOYEE(t) AND (∃d)(DEPARTMENT(d) AND t.Dno = d.Dnumber AND d.Dname = 'Research')}

→ SELECT ①
FROM EMPLOYEE AS t, DEPARTMENT AS d
WHERE ③

$$\{q,r,s,u \mid (\exists_3)(\exists_2)(\exists_m)\ \text{EMPLOYEE}(qrstuvwxyz) \text{ AND } \text{DEPARTMENT}(lmno) \text{ AND } l = \text{'Research'} \text{ AND } m = 3\}$$

Examples

Query 2: For every project located in 'Stafford', list the project number, the controlling department number, and the departments manager's last name, address and birth date.

Figure 3.6
One possible database state for the COMPANY relational database schema.

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

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

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

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

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

```
SELECT Pnumber, Dnum, Lname,
       Adress, Bdate
FROM   EMPLOYEE(e), DEPARTMENT(d),
       PROJECT(p)
WHERE  Plocation = 'Stafford'
AND    Dnum = Dnumber
AND    mgr_ssn = Ssn;
```

(*) command: *, **, ***

```
D_RC  (j)(k)(m)(n)(t)
{i,k,s,u,v | PROJECT(hijk) AND
  DEPARTMENT(lmno) AND
  EMPLOYEE(qrstuvwxy3)
  AND j = 'Stafford'
  AND k = m
  AND n = t }
```

$SP \leftarrow \sigma_{Plocation = 'Stafford'} (PROJECT)$

$CD \leftarrow (SP) \bowtie_{Dnum = Dnumber} (DEPARTMENT)$

$PD \leftarrow CD \bowtie_{mgrSsn = Ssn} EMPLOYEE$

$R \leftarrow \pi_{\textcircled{2}} PD$

Examples

Query 7: List the names of managers who have at least one dependent.

Figure 3.6
One possible database state for the COMPANY relational database schema.

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

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

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

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

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse