Relational Algebra

(Chapter 3.2)

Peter Revesz

CSCE 413/813
Computer Science and Engineering
University of Nebraska – Lincoln

Relational Database Query Languages

- Relational Algebra
- SQL
- NoSQL

Each has some advantages and disadvantages.

Relational Algebra

Disadvantage:

- Less abstract than SQL

Advantages:

- It consists of a set of basic operators that are simple to implement. Each operator takes as input table(s) and give as output a single table.
- It provides a basis of implementing SQL. In a relational database system a query processor translates each SQL query into some sequence of is built on top of relational algebra.

Relational Algebra Operators

- The basic operators are:
  - Projection
  - Selection
  - Rename
  - Intersection
  - Union
  - Difference
  - Product
  - Natural Join

Hospital Database

Patient

Name	ID	CM	KG
Anderson	100	200	130
Brown	111	150	50
Davis	222	190	90
Edwards	333	160	90
Ford	345	165	100
Hardy	454	175	70
Johnson	567	170	50
Smith	755	180	120

Doctor

Name	ID	Age	Specialty
Cheney	987	50	pediatry
Hardy	454	53	osteopathology
McBride	377	36	radiology
Miller	300	60	neurology
Moss	244	30	neurology
Nelson	400	76	cardiology
Oltman	181	56	urology
Paine	266	45	cardiology
Pepper	555	42	cardiology
Snow	500	65	radiology

Visit

PID	DID	Month	Day	Year
100	181	5	20	2008
100	555	6	30	2009
111	987	8	20	2009
111	987	5	28	2010
222	266	9	12	2007
222	400	5	20	2008
222	555	5	20	2008
333	987	6	23	2009
345	300	5	16	2009
454	244	6	10	2010
567	377	2	20	2010
567	454	5	28	2010
755	987	6	23	2009

Projection (π)

Finds in a table those columns that are specified.

Example: π_{Name} Patient

Name
Anderson
Brown
Davis
Edwards
Ford
Hardy
Johnson
Smith

Selection (σ)

-
- Finds in a table those rows that satisfy a condition.

  Example: $\sigma_{\text{Name = 'Anderson'}}$ Patient

Name	ID	CM	KG
Anderson	100	200	130

Combining Operators Using Parentheses

Example: Find the centimeter height of patient Anderson.

$\pi_{CM}\ (\sigma_{Name = 'Anderson'}\ Patient)$

CM
200

Example: Find the centimeter height and kilogram weight of patient Smith.

CM	KG
180	120

Giving a Name to a Returned Relation

- Example: Find the planets smaller than the Earth.
  Call the output relation 'Small_Planet'.

  $Small\_Planet \leftarrow \sigma_{Mass < 1} Planet$

Small_Planet

Name	Mass	Period
Mercury	0.06	0.24
Venus	0.82	0.62
Mars	0.11	1.88

- Example: Find the planets slower than Earth.
  Call the output relation `Slow_Planet'.

  $Slow\_Planet \leftarrow \sigma_{Period > 1} Planet$

Slow_Planet

Name	Mass	Period
Mars	0.11	1.88
Jupiter	317.80	11.86
Saturn	95.2	29.46
Uranus	14.6	84.01
Neptune	17.2	164.8

Intersection (∩), Union (U) and Difference (\)

- Small and Slow

  Small_Planet ∩ Slow_Planet

Name	Mass	Period
Mars	0.11	1.88

- Small or Slow

  Small_Planet U Slow_Planet

Name	Mass	Period
Mercury	0.06	0.24
Venus	0.82	0.62
Mars	0.11	1.88
Jupiter	317.80	11.86
Saturn	95.2	29.46
Uranus	14.6	84.01
Neptune	17.2	164.8

- Planet minus Slow

  Planet \ Slow_Planet

Name	Mass	Period
Earth	1.00	1.00
Jupiter	317.80	11.86
Saturn	95.2	29.46
Uranus	14.6	84.01
Neptune	17.2	164.8

Product (X) and Natural Join (⋈)

Let us introduce two new relations:

Astronaut

FirstName	Age
Claire	40
Gregory	38
Sophia	42

Orbit

Name	Radius
Mercury	0.39
Venus	0.72
Earth	1.00
Mars	1.52
Jupiter	5.20
Saturn	9.54
Uranus	19.22
Neptune	30.06

Product (X) and Natural Join (⋈)

- Product of Small_Planet and Astronaut

  Small_Planet X Astronaut

Name	Mass	Period	FirstName	Age
Mercury	0.06	0.24	Claire	40
Mercury	0.06	0.24	Gregory	38
Mercury	0.06	0.24	Sophia	42
Venus	0.82	0.62	Claire	40
Venus	0.82	0.62	Gregory	38
Venus	0.82	0.62	Sophia	42
Mars	0.11	1.88	Claire	40
Mars	0.11	1.88	Gregory	38
Mars	0.11	1.88	Sophia	42

- Natural Join of Planet and Orbit

  Planet ⋈ Orbit

Name	Mass	Period	Radius
Mercury	0.06	0.24	0.39
Venus	0.82	0.62	0.72
Earth	1.00	1.00	1.00
Mars	0.11	1.88	1.52
Jupiter	317.80	11.86	5.20
Saturn	95.2	29.46	9.54
Uranus	14.6	84.01	19.22
Neptune	17.2	164.8	30.06

Relational Algebra Queries - Summary

Definition 3.2.1 A relational algebra query is defined recursively as follows:

1. Each relation name R_i is a relational algebra expression. (This will just return the relation R_i as a result.)

2. If e_1 is a relational algebra expression, then $\Pi_L(e_1)$, where L is a subset of the attributes of the resulting relation, $\sigma_F(e_1)$, where F is a selection condition, and $\rho_C(e_1)$, where C is a renaming specification, are also relational algebra expressions.

3. If e_1 and e_2 are relational algebra expressions that result in relations that have the same set and order of attributes, then $(e_1 \cap e_2)$, $(e_1 \cup e_2)$, and $(e_1 \setminus e_2)$ are also relational algebra expressions.

4. If e_1 and e_2 are relational algebra expressions that result in relations with no common attributes, then $(e_1 \times e_2)$ is also a relational algebra expression.

5. If e_1 and e_2 are relational algebra expressions, then $(e_1 \bowtie e_2)$ is a relational algebra expression.