Relational Algebra

Tom Kelliher, CS 318

Feb. 11, 2002

1 Administrivia

Announcements

Assignment

Read 6.2–3.

From Last Time

E/R design example.

Outline

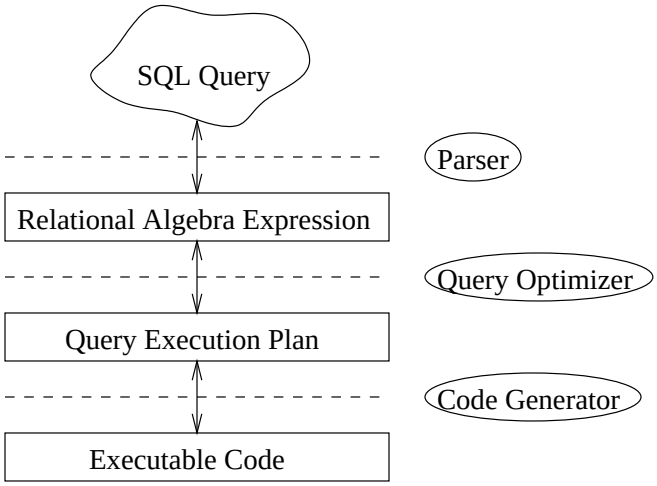
1. Relational algebra's role.
2. Primitive RA operators.
3. Derived RA operators.

Coming Up

Queries and updates in SQL.

2 Relational Algebra's Role

Query processing:



3 Primitive Relational Algebra Operators

Generally, each operator takes:

1. One or two relations.
2. A conditional expression.

and produces a relation.

3.1 Select

1. Used to filter out tuples from a relation.
2. Example:
$$\sigma_{\text{Dept}='\text{CS}'}(\text{Professor})$$
3. General form:
$$\sigma_{selection-condition}(relation\text{–}name)$$

4. *Selection-condition* can be extended with the usual boolean operators.

5. Only real restriction: Attributes named in the selection-condition must belong to the relation.

3.2 Project

1. Used to filter out attributes from a relation.

2. Example:

$$\pi_{\text{CrsCode,Semester}}(\text{Teaching})$$

3. Combination example:

$$\pi_{\text{Name}}(\sigma_{\text{DeptId}='\text{CS}'}(\text{Professor}))$$

3.3 Union

1. Relations must be union-compatible:

   (a) Same number of attributes.

   (b) Attribute names are the same in both relations.

   (c) Attribute domains match.

2. Example:

$$\pi_{\text{CrsCode,Semester}}(\sigma_{\text{Grade}='\text{C}'}(\text{Transcript})) \cup \pi_{\text{CrsCode,Semester}}(\sigma_{\text{CrsCode}='\text{MAT123}'}(\text{Teaching}))$$

3.4 Set Difference

1. Union-compatability required.

2. A − B: Everything in A and not in B.

   Let A = $(1, 3, 5, 7, 9)$ and B = $(6, 7, 8, 9, 10)$. A − B = $(1, 3, 5)$ and B − A = $(6, 8, 10)$.

3.5 Cartesian Product

1. Used to combine two relations. **Expensive** to compute.

2. A $\times$ B is the set of all tuples which can be formed by combining a tuple of A with a tuple of B.

   Let A be:

Name	**Price**
Hat	15
Coat	75

   and B be:

Name	**Size**
Green	Small
Red	Petite
Magenta	Medium

   A $\times$ B is:

Name	**Price**	**Name**	**Size**
Hat	15	Green	Small
Hat	15	Red	Petite
Hat	15	Magenta	Medium
Coat	75	Green	Small
Coat	75	Red	Petite
Coat	75	Magenta	Medium

   Whoops. To rename the attributes, we could write:

$$\text{A} \times \text{B}\ [\text{Item},\ \text{Price},\ \text{Color},\ \text{Size}]$$

3.6 Examples

Assume we have the schemas Student and Transcript:

Id	Name	Status	Address

StuId	CrsCode	Semester	Grade

Write relational algebra expressions to determine the following:

1. The names of all seniors.

2. The name and address of the student whose Id is 123456789.

3. The names of all juniors and those students who Id is between 111111111 and 222222222, inclusive.

4. The names of all students who have taken CS318.

5. The names of all sophomores who received A's during the F2001 semester.

6. The names and Ids of any students who have never taken a course.

4 Derived Relational Algebra Operators

4.1 Join

1. Theta join:

$$\mathrm{A} \bowtie_{\mathrm{join-condition}} \mathrm{B} = \sigma_{\mathrm{join-condition}}(\mathrm{A} \times \mathrm{B})$$

   Example:

$$\sigma_{\mathrm{1id}}(\mathrm{Employee}\ [\mathrm{1id},\ \mathrm{1sal},\ \mathrm{1sup}] \bowtie_{\mathrm{1sal>2sal\ AND\ 1sup=2id}} \mathrm{Employee}\ [\mathrm{2id},\ \mathrm{2sal},\ \mathrm{2sup}])$$

   (Ids of all those who earn more than their supervisors.)

2. Equijoin: only comparison operator is =.

   Important because they "reconstruct" the database: Get names of all students who have taken CS318.

3. Natural join: equijoin on intersection of attributes where the redundant attributes in each tuple are projected out.

4.2 Division

1. Let $< s >$ and $< t >$ be sets of attributes and let relation A = $(< s >< t >)$ and relation B = $(< t >)$. Then:

$$\text{A/B} = \{< s > \mid < s > \times B \subseteq A\}$$

2. Example. Let A be:

Student	**Course**
Jessie	CS220
Jessie	CS240
John	CS240

Let B be:

Course
CS220
CS240

A/B = (Jessie).

3. Derivation of division A/B:

$$\begin{aligned} \text{T}_1 &= \pi_{<a>}(A) \times B \\ T_2 &= \pi_{<a>}(T_1 - A) \\ A/B &= \pi_{<a>}(A) - T_2 \end{aligned}$$