Relational Calculus

CSE 132A Winter 2016. 22 Jan 2016

Existential and Universal Quantifiers

$\exists b \in \text{Books}[\ldots]$
There exists a book b where this condition holds.
Analagous to "FROM Books b WHERE ..." in SQL.

$\exists b \in \text{Books}[\forall c \in \text{Customers}[\exists p \in \text{Purchases}[\ldots]]]$
There exists a book, where for each customer, there exists a purchase where this condition holds.

Tuples

Tuples are like rows in a database

$\exists b \in \text{Books}[b(\text{title}) \ldots]$

$\exists b \in \text{Books}[\forall c \in \text{Customers}[\exists p \in \text{Purchases}[$
$c(\text{customerid}) \ldots p(\text{customerid}) \ldots p(\text{bookid}) \ldots b(\text{bookid})]]]$

Similar to b.title and p.price

SELECT Analogy

$$\{r : \text{bookid} \mid \exists b \in \text{Books}[b(\text{title}) = \text{'Twilight'}$$
$$\wedge r(\text{bookid}) = b(\text{bookid})]\}$$

$$\{r : \text{title} \mid \exists b \in \text{Books}[\forall c \in \text{Customers}[\exists p \in \text{Purchases}[$$
$$c(\text{customerid}) = p(\text{customerid}) \wedge p(\text{bookid}) = b(\text{bookid})$$
$$]] \wedge b(\text{title}) = r(\text{title})]\}$$

Formal Definition

The general form of the relational calculus is:

$$\{t|P(t)\}$$

t is a free variable. P is the formula. A tuple variable is a *free variable* unless it is quantified by $\exists, \forall$. If quantified, it is *bound variable*. Free variable are

- not in scope of any quantifier
- free variables are the parameters of the formula

$$t \in Purchases \wedge \exists c \in Customers[t(custid) = c(custid)]$$

Atoms

- $t \in R$, where t is a tuple variable, R is a relation
- $t(x)\Theta u(y)$, where t, u are tuple variables, x, y are the attributes. Θ is comparison operator $(<, \leq, =, \neq, \geq, >)$.
- $t(x)\Theta c$, where c is a constant.

Formulae

- An atom P is a formula
- $\neg P, P_1 \wedge P_2, P_1 \vee P_2$
- $P_1 \rightarrow P_2$
- $P(t)$ contains a free tuple t, and.

$$t \in R[P(t)]$$
$$\forall t \in R[P(t)]$$

Equivalence

- $P_1 \wedge P_2 \equiv \neg(\neg P_1 \vee \neg P_2)$
- $\forall t \in R[P(t)] \equiv \neg t \in R[\neg P(t)]$
- $P_1 \rightarrow P_2 \equiv \neg P_1 \vee P_2$

Limitations

Aggregation

Note that the tuple relational calculus does not have any equivalent of the aggregate operation, but it can be extended to support aggregation. Extending the tuple relational calculus to handle arithmetic expressions is straightforward.

Example

Schema

Consider the following database schema for a BOOKSTORE database:

- Books (bookid, title, author, year)
- Customers (customerid, name, email)
- Purchases (customerid, bookid, year)
- Reviews (customerid, bookid, rating)
- Pricing (bookid, format, price)

Example 1

SQL to Relational Calculus

Find books (show their titles) written by 'EDMUND MORGAN' since year 1990.

SELECT title
FROM Books
WHERE author = 'EDMUND MORGAN'
AND year >= 1990;

Example 1

SQL to Relational Calculus

SELECT title
FROM Books
WHERE author = 'EDMUND MORGAN'
AND year $>=$ 1990;

$\{r : \text{title} \mid \ldots\}$

Example 1

SQL to Relational Calculus

SELECT title
FROM Books
WHERE author = 'EDMUND MORGAN'
AND year >= 1990;

$\{r : \text{title} \mid \exists b \in \text{Books}[...]\}$

Example 1

SQL to Relational Calculus

SELECT title
FROM Books
WHERE author = ’EDMUND MORGAN’
AND year $>=$ 1990;

$\{r : \text{title} \mid \exists b \in$
$\text{Books}[b(\text{author}) = \text{'EDMUND MORGAN'} \wedge b(\text{year}) \geq 1990]\}$
DONE?

Example 1

SQL to Relational Calculus

SELECT title
FROM Books
WHERE author = 'EDMUND MORGAN'
AND year $>=$ 1990;

$$\{r : \text{title} \mid \exists b \in \text{Books}[b(\text{author}) = \text{'EDMUND MORGAN'} \wedge b(\text{year}) \geq 1990 \wedge r(\text{title}) = b(\text{title})]\}$$

Example 2

Relational Calculus to SQL

What are the titles of the newest books?

Example 2

Relational Calculus to SQL

What are the titles of the newest books?

$\{r : \text{title} \mid \ldots\}$

Example 2
Relational Calculus to SQL

What are the titles of the newest books?

$$\{r : \text{title} \mid \exists b \in \text{Books}[...]\}$$

Example 2
Relational Calculus to SQL

What are the titles of the newest books?

$$\{r : \text{title} \mid \exists b \in \text{Books}[\forall o \in \text{Books}[\ldots]]\}$$

Example 2
Relational Calculus to SQL

What are the titles of the newest books?

$\{r : \text{title} \mid \exists b \in \text{Books}[\forall o \in \text{Books}[b(\text{year}) \geq o(\text{year})]]\}$
DONE?

Example 2

Relational Calculus to SQL

What are the titles of the newest books?

$$\{r : \text{title} \mid \exists b \in \text{Books}[\forall o \in \text{Books}[b(\text{year}) \geq o(\text{year})]$$
$$\wedge b(\text{title}) = r(\text{title})]\}$$

Example 2

Relational Calculus to SQL

What are the titles of the newest books?

$$\{r : \text{title} \mid \exists b \in \text{Books}[\neg\exists o \in \text{Books}[b(\text{year}) < o(\text{year})] \\ \wedge b(\text{title}) = r(\text{title})]\}$$

Example 2

Relational Calculus to SQL

What are the titles of the newest books?

$$\{r : \text{title} \mid \exists b \in \text{Books}[\neg\exists o \in \text{Books}[b(\text{year}) < o(\text{year})]$$
$$\wedge b(\text{title}) = r(\text{title})]\}$$

```
SELECT b.title
FROM Books b
WHERE ...
```

Example 2

Relational Calculus to SQL

What are the titles of the newest books?

$$\{r : \text{title} \mid \exists b \in \text{Books}[\neg \exists o \in \text{Books}[b(\text{year}) < o(\text{year})]$$
$$\wedge b(\text{title}) = r(\text{title})]\}$$

```
SELECT b.title
FROM Books b
WHERE NOT EXISTS (SELECT
  FROM Books o
  WHERE b.year < o.year
);
```

Exercise 1

Find books (show their titles, authors and prices) that are on 'CIVIL WAR' (i.e., the title field contains 'CIVIL WAR'), available in 'AUDIO' format.

$$\{t : title, author, price | \exists b \in Books$$
$$[(b(title) = t(title) \wedge b(author) = t(title) \wedge b(price) = t(price)) \wedge$$
$$b(title) =' \%CIVILWAR\%' \wedge$$
$$\exists p \in Pricing[p(bookid) = b(bookid) \wedge p(format) =' AUDIO']$$
$$]\}$$

Exercise 2

Find the set of books purchased by 'JOHN CHAMBERS'

$$\{t|t \in Books \wedge \exists p \in Purchases, \exists c \in Customers$$
$$[t(bookid) = p(bookid) \wedge c(customerid) = p(customerid) \wedge$$
$$c(name) = \text{'JOHN CHAMERS'}$$
$$]\}$$

We don't need to get the subset of columns of Books t.

Exercise 3

Find customers (show their names and email addresses) who purchased more than one book in year 2003.
* Do it without aggregation.

$$
\begin{aligned}
&\{t : name, email \mid \exists c \in Customers \\
&\quad [(t(name) = c(name) \wedge t(email) = c(email)) \wedge \\
&\quad \exists p_1, p_2 \in Purchase[p_1 \neq p_2 \wedge p_1(customerid) = c(customerid) \wedge \\
&\quad\quad p_2(customerid = c(customerid)) \wedge \\
&\quad\quad p_1(year) = 2003 \wedge p_2(year) = 2003] \\
&\ ]\}
\end{aligned}
$$

Exercise onward

The schema about battleships and the battles they fought in:

- Ships(name, yearLaunched, country, numGuns, gunSize, displacement)
- Battles(ship, battleName, result)

Exercise 4

Which battleships launched before 1930 had 16-inch guns? List their names, their country, and the number of guns they carried.

$$\{t : names, country, numGuns | \exists s \in Ships[$$
$$t(name) = s(name) \wedge t(country) = s(country)$$
$$\wedge\, t(numGuns) = s(numGuns) \wedge$$
$$s(yearLaunched) < 1930 \wedge s(gunSize) = 16$$
$$]\}$$

Exercise 5

Which battleship(s) had the guns with the largest gun size?

$$
\begin{aligned}
&\{t \mid t \in Ships \wedge \\
&\quad \forall s \in Ships[t(gunSize) \geq s(gunSize) \\
&]\}
\end{aligned}
$$

Exercise 6

Which battleships had the guns with the second largest gun size. More precisely, find the ships whose gun size was exceeded by only one gun size, no matter how many other ships had that larger gun size.

$$\{t | t \in Ships \wedge$$
$$\exists m \in Ships \Big[\forall s \in Ships [m(gunSize) \geq s(gunSize)] \wedge$$
$$t(gunSize) < m(gunSize) \wedge \forall s' \in Ships [s'(gunSize) < m(gunSize)$$
$$\rightarrow t(gunSize) \geq s'(gunSize)]$$
$$\Big] \}$$

Exercise 7

List all the pairs of countries that fought each other in battles. List them with the country that comes first in alphabetical order first[1].

$$
\begin{aligned}
\{t : & c_1, c_2 \mid t(c_1) < t(c_2) \wedge \\
& \exists s, s' \in Ships, \exists b, b' \in Battles \Big[\\
& s(country) = t(c_1) \wedge b(ship) = s(name) \wedge \\
& s'(country) = t(c_2) \wedge b'(ship) = s'(name) \wedge \\
& b(battleName) = b'(battleName) \\
\Big] \}
\end{aligned}
$$

[1] may contains duplicated entries

Reference

1. "*Database Systems Concepts*" by Silberschatz, Korth and Sudarshan, 6th edition, McGraw-Hill.