Relational Algebra

Set-theoretic operations:

Two relations are **union-compatible** if they have the same number of attributes and the domains of the corresponding attributes in the two relations are the same.

Consider two relations $r(R)$ and $s(S)$ that are union-compatible (normally R = S).

Union: $r \cup s = \{t | t \in r \text{ or } t \in s\}$.

Difference: $r - s = \{t | t \in r \text{ and } t \notin s\}$

Intersection: $r \cap s = \{t | t \in r \text{ and } t \in s\}$

Cartesian Product: $r(R)$ and $s(S)$ on any schemes R and S.

$$r \times s = \{t_1.t_2 | t_1 \in r \text{ and } t_2 \in s\},$$

where, $t1.t2$ is the concatenation of tuples t_1 and t_2 to form a larger tuple.

Example: set operations

r

A	B
a	b
a	c
b	d

s

A	B
a	c
a	e

$r \cup s$

A	B
a	b
a	c
b	d
a	e

$r - s$

A	B
a	b
b	d

$r \cap s$

A	B
a	c

$r \times s$

r.A	r.B	s.A	s.B
a	b	a	c
a	b	a	e
a	c	a	c
a	c	a	e
b	d	a	c
b	d	a	e

Relation-theoretic operations

Consider $r(R)$ and $s(S)$, two relations, where $R = (A_1, ..., A_n)$ and $S = (B_1, ..., B_m)$

Rename: $r(C_1, ..., C_n) = \{t | t \in r\}$ with schema $(C_1, ..., C_n)$.

Select: $\sigma_F(r) = \{t | t \in r$ and t satisfies $F\}$.

where F is a selection criteria involving constants and attributes of r. (will discuss in examples how F is constructed)

Project: $\pi_{D_1,...,D_p}(r) = \{t[D_1, ..., D_p] | t \in r\}$

where D_i is one of $A_1, ..., A_n$.

theta-Join: $r \bowtie_F s = \{t | (\exists u \in r)(\exists v \in s)(t = u.v$ and F is satisfied by u and $v)\}$

where F is a conjunction of formulas relating attributes of r with attributes of s. (will discuss in examples how F is constructed)

Natural Join: $r \bowtie s = \{t | (\exists u \in r)(\exists v \in s)(t[R] = u$ and $t[S] = v)\}$

Division: Assume $B_1, ..., B_m \subset A_1, ..., A_n$.

$$r \div s = \{t | (\forall u \in s)(t.u \in r)\}$$

Examples: relation-theoretic operations

r

A	C	D
a	c	d
a	e	f
a	g	h
b	c	d
b	g	h
c	c	d
c	e	f

$\sigma_{A='b' \text{ or } C='c'}(r)$

A	C	D
a	c	d
b	c	d
b	g	h
c	c	d

$\pi_A(r)$

A
a
b
c

r

A	C	D
a	c	d
a	e	f
a	g	h
b	c	d
b	g	h
c	c	d
c	e	f

t

B	C	D
b	c	d
b	e	f

$r \bowtie_{r.A=t.B} t$

A	C	D	B	C	D
b	c	d	b	c	d
b	c	d	b	e	f
b	g	h	b	c	d
b	g	h	b	e	f

r

A	C	D
a	c	d
a	e	f
a	g	h
b	c	d
b	g	h
c	c	d
c	e	f

t

B	C	D
b	c	d
b	e	f

$r \bowtie t$

A	C	D	B
a	c	d	b
a	e	f	b
b	c	d	b
c	c	d	b
c	e	f	b

r

A	C	D
a	c	d
a	e	f
a	g	h
b	c	d
b	g	h
c	c	d
c	e	f

s

C	D
c	d
e	f

$r \div s$

A
a
c

Basic Relational Algebra Operations

- Basic set: union, difference, Cartesian product, rename, select, and project.
- none of them can be expressed in terms of the others.
- intersection, theta-join, natural join, and division can be expressed in terms of the basic operators as follows:

Intersection: $r \cap s = r - (r - s)$

theta Join: $r \bowtie_F s = \sigma_F(r \times s)$

Natural Join: $r \bowtie s = \pi_{R \cap S}(\sigma_F(r \times s))$
where F is a selection condition which indicates that the tuple values under the common attributes of r and s are equal.

Division: $r \div s = \pi_{R-S}(r) - \pi_{R-S}((\pi_{R-S}(r) \times s) - r)$

Even though relation schemes are defined as sequences, they are treated as sets in these equalities for simplicity.

An explanation for the equality for division is in order!

- First, all candidate tuples for the result are calculated by the expression

  $$\pi_{R-S}(r)$$

- Next, these candidate tuples are combined with all tuples of s in the following expression

  $$\pi_{R-S}(r) \times s$$

  to give a relation containing all combinations of candidate tuples with all tuples of s.

- Since we are looking for tuples under the scheme $R - S$ which combine with all tuples of s and are also present in r, if we subtract r from the previous expression, we will get all the combinations of tuples that are "missing" in r.

  $$(\pi_{R-S}(r) \times s) - r$$

- By projecting these tuples on $R - S$, we get all those tuples that should not go to the result in the following expression.

  $$\pi_{R-S}((\pi_{R-S}(r) \times s) - r)$$

- Finally, we subtract this set from the set of all candidate tuples and obtain the output relation of the division operator.

  $$r \div s = \pi_{R-S}(r) - \pi_{R-S}((\pi_{R-S}(r) \times s) - r)$$