# Theory I: Database Foundations

Jan-Georg Smaus (Georg Lausen)

26.7.2011

# Languages

## Paradigms

- Relational algebra
- Relational calculus
- SQL: not explicitly considered in this theory course!

# Relational Algebra

## Basic Operators

- delete attributes: Projection.
- select tuples: Selection.
- combine relations: Join.
- set operators: Union, Difference.

# Projection

Student

| MatrId | Name | Address | Semester |
|---|---|---|---|
| 1223 | Hans Eifrig | Seeweg 20 | 2 |
| 3434 | Lisa Lustig | Bergstraße 11 | 4 |
| 1234 | Maria Gut | Am Bächle 1 | 2 |

$\Downarrow$

Student'

| MatrId | Name |
|---|---|
| 1223 | Hans Eifrig |
| 3434 | Lisa Lustig |
| 1234 | Maria Gut |

## Projection on tuples

- Let $R(X)$ be a schema, where $X = \{A_1, \ldots, A_k\}$.
- Let $Y$ be a set of attributes, where $\emptyset \subset Y \subseteq X$.
- Let $\mu \in \mathrm{Tup}(X)$ be a tuple over $X$.
- $\mu[Y]$ is called projection of $\mu$ to $Y$:

$$\mu[Y] \in \mathrm{Tup}(Y),$$

$\mu[Y](A) = \mu(A), A \in Y.$

## Projection on relations

- Let $r \subseteq \text{Tup}(X)$ a relation and $Y \subseteq X$.
- $\pi[Y]r$ is called projection of $r$ to $Y$:

$$\pi[Y]r = \{\mu \in \text{Tup}(Y) \mid \exists \mu' \in r, \text{such that } \mu = \mu'[Y]\}.$$

### Example

$r =$

| $A$ | $B$ | $C$ |
|---|---|---|
| $a$ | $b$ | $c$ |
| $a$ | $a$ | $c$ |
| $c$ | $b$ | $d$ |

$\pi[A, C](r) =$

# Selection

Course

| CourseId | Institute | Title | Description |
|---|---|---|---|
| K010 | DBIS | Databases | Foundations of Databases |
| K011 | DBIS | Information Systems | Foundations of Information Systems |
| K100 | MST | Microsystems | Foundations of Microsystems |

$\Downarrow$

Course'

| CourseId | Institut | Title | Description |
|---|---|---|---|
| K100 | MST | Microsystems | Foundations of Microsystems |

## Selection condition

- Let $A, B \in X$, $a \in dom(A)$, and $\theta \in \{=, \neq, \leq, <, \geq, >\}$ a comparison operator.
- An (atomic) selection condition $\alpha$ (on $X$) is of the form $A\,\theta\,B$, resp. $A\,\theta\,a$, resp. $a\,\theta\,A$.
- A tuple $\mu \in \mathrm{Tup}(X)$ fulfills a selection condition $\alpha$, if $\mu(A)\,\theta\,\mu(B)$, resp. $\mu(A)\,\theta\,a$, resp. $a\,\theta\,\mu(A)$ hold.
- Atomic selection conditions can be generalized to formulas using $\wedge$, $\vee$, $\neg$, and (, ).

### Example

$$X = \{A, B, C\}.$$
$$\mu_1 = (A \to 2, B \to 2, C \to 1),\ \mu_2 = (A \to 2, B \to 3, C \to 2)$$
$$\alpha_1 = (A = B),\ \alpha_2 = ((B > 1) \wedge (C > 1))$$

## Selection

- Let $r \subseteq \mathrm{Tup}(X)$ be a relation and $\alpha$ a selection condition over $X$.
- $\sigma[\alpha]r$ is called selection of relation $r$ by $\alpha$:

$$\sigma[\alpha]r = \{\mu \in \mathrm{Tup}(X) \mid \mu \in r \wedge \mu \text{ fulfills } \alpha\}.$$

### Example

$$r = \begin{array}{ccc} A & B & C \\ \hline a & b & c \\ d & a & f \\ c & b & d \end{array}$$

$\sigma[B = b](r) =$

## Union and difference

- Let $X$ be a set of attributes and $r \subseteq \text{Tup}(X), s \subseteq \text{Tup}(X)$ two relations.
- $$\begin{aligned} r \cup s &= \{\mu \in \text{Tup}(X) \mid \mu \in r \vee \mu \in s\}. \\ r - s &= \{\mu \in \text{Tup}(X) \mid \mu \in r, \text{where } \mu \notin s\}. \end{aligned}$$

### Example

$$r = \begin{array}{ccc} A & B & C \\ \hline a & b & c \\ d & a & f \\ c & b & d \end{array} \qquad s = \begin{array}{ccc} A & B & C \\ \hline b & g & a \\ d & a & f \end{array} \qquad r \cup s =$$

$$r = \begin{array}{ccc} A & B & C \\ \hline a & b & c \\ d & a & f \\ c & b & d \end{array} \qquad s = \begin{array}{ccc} A & B & C \\ \hline b & g & a \\ d & a & f \end{array} \qquad r - s =$$

# Join

Student

| MatrId | Name |
|---|---|
| 1223 | Hans Eifrig |
| 3434 | Lisa Lustig |
| 1234 | Maria Gut |

Registration

| MatrId | CourseId |
|---|---|
| 1223 | K010 |
| 1234 | K010 |

Course

| CourseId | Title |
|---|---|
| K010 | Databases |
| K011 | Information System |
| K100 | Microsystems |

$\Downarrow$

Student'

| MatrId | Name | CourseId | Title |
|---|---|---|---|
| 1223 | Hans Eifrig | K010 | Databases |
| 1234 | Maria Gut | K010 | Databases |

## Join

- For sets of attributes $X, Y$, we may also write $XY$ instead of $X \cup Y$.
- Let $r \subseteq \mathrm{Tup}(X), s \subseteq \mathrm{Tup}(Y)$.
- The (natural) join $\bowtie$ of $r$ and $s$ is defined:

$$r \bowtie s = \{\mu \in \mathrm{Tup}(XY) \mid \mu[X] \in r \wedge \mu[Y] \in s\}.$$

### Example

$r =$

| $A$ | $B$ | $C$ |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 6 |

$s =$

| $C$ | $D$ |
|---|---|
| 3 | 1 |
| 6 | 2 |
| 4 | 5 |

$r \bowtie s =$

### Observation about join

If $X_1 \cap X_2 = \emptyset$, then $r_1 \bowtie r_2 = r_1 \times r_2$.

## Generalization of join

Let $X_i$, $1 \leq i \leq n$ be sets of attributes.

$$\bowtie_{i=1}^{n} r_i = \{\mu \in \mathrm{Tup}(\cup_{i=1}^{n} X_i) \mid \mu[X_i] \in r_i, 1 \leq i \leq n\}.$$

## Basic Operators

- Selection, projection, union, difference, and join are the basic operators of relational algebra.
- The valid expressions of the relational algebra can be defined inductively.
- We could define other useful operators.

## The relational algebra as query language

- In the algebra expressions we have seen, the operations are applied to relation instances (small letters $r$, $s$, ...), not relation names (capital letters $R$, $S$, ...).
- One can also build expressions based on the relation names. These expressions are then called queries and must be evaluated wrt. a database instance $\mathcal{I}$. We write $\mathcal{I}(Q)$ for the result of this evaluation, the answer. That is, to obtain $\mathcal{I}(Q)$, one has to replace every relation name $R$ occurring in $Q$ by the relation instance $\mathcal{I}(R)$.
- $\mathcal{I}(Q)$ is again a relation. Recall that a query is formally given as a mapping (transformation) from a database instance to a relation instance.
- Not all computable transformations can be expressed in the relational algebra. Example: transitive closure.

## Equivalence

Two algebra expressions $Q, Q'$ are called equivalent, $Q \equiv Q'$, if for any instance $\mathcal{I}$ of a database:

$$\mathcal{I}(Q) = \mathcal{I}(Q').$$

### Examples

Let $\mathrm{attr}(\alpha)$ be the attributes in $\alpha$ and let $R, S, T \ldots$ be relation names whose formats are $X, Y, Z$.

- $Z \subseteq Y \subseteq X \Longrightarrow \pi[Z](\pi[Y]R) \equiv \pi[Z]R.$
- $X = Y \Longrightarrow R \cap S \equiv R \bowtie S.$