# Relationale Algebra

Thomas Heimrich



Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum

# Formale Sprachen

## relationale Algebra

Die relationale Algebra ist *prozedural* orientiert. Sie beinhaltet implizit einen Abarbeitungsplan für die Anfrage. Die rel. Algebra ist wichtig für die Optimierung von Anfragen.

## Relationenkalkül

Das Relationenkalkül ist eine deklerative Sprache. Man sagt, *welche* Daten man erhalten will. Man sagt **nicht** *wie* die Anfrage ausgewertet werden soll.

Beide Sprachen sind abgeschlossen, d.h. die Ergebnisse der Anfragen sind wieder Relationen.

# Beispielrelationen

Professoren

| **PersNr** | Name | Raum |
|---|---|---|
| 1013 | Klitschko | C301 |
| 8516 | Ali | C302 |
| 1005 | Maske | C303 |
| 5313 | Schulz | C304 |

Studenten

| **MatrNr** | Name | Semester |
|---|---|---|
| 2401 | Alf | 713 |
| 2402 | Frodo | 4 |
| 2403 | Heidi | 2 |
| 2404 | Harry | 1 |
| 2405 | Ronja | 4 |

Vorlesungen

| **VorlNr** | Titel | SWS | gehaltenVon |
|---|---|---|---|
| 107 | Ethik | 4 | 1013 |
| 286 | Logik | 6 | 8516 |
| 456 | Sport | 4 | 1005 |
| 117 | Kunst | 6 | 1013 |
| 118 | Joga | 4 | 1005 |

voraussetzen

| **Vorgänger** | **Nachfolger** |
|---|---|
| 107 | 117 |
| 107 | 286 |
| 456 | 117 |
| 286 | 118 |

hören

| **MatrNr** | **VorlNr** |
|---|---|
| 2401 | 107 |
| 2402 | 107 |
| 2401 | 456 |


Rel. Algebra

Grundlagen

Beispielrelationen

rel. Algebra
- Selektion
- Projektion
- Vereinigung
- Differenz
- Kreuzprod.
- Umbenennung

Definition der rel. Algebra

Verbundoperatoren
- Natürlicher Verbund
- Allgemeiner Verbund

Operatorbaum

# Selektion

Bei der Selektion werden diejenigen Tupel einer Relation ausgewählt, die das sogenannte *Selektionsprädikat* erfüllen. Die Selektion wird mit $\sigma$ bezeichnet und hat das Selektionsprädikat als Subskript.

## Beispiel für Selektion

Anfrage:

$$\sigma_{\text{Semester}>13}(\text{Studenten})$$

Ergebnis:

Studenten

| **<u>MatrNr</u>** | Name | Semester |
|---|---|---|
| 2401 | Alf | 713 |



Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum

# Projektion

Während die Selektion einzelne Zeilen auswählt, werden bei der Projektion Spalten extrahiert. Die Projektion wird mit dem Operatorsymbol $\Pi$ bezeichnet und enthält die Menge der Attributnamen im Subskript.

## Beispiel für Projektion

Anfrage:

$$\Pi_{\text{Semester}}(\text{Studenten})$$

Ergebnis:

Studenten

| Semester |
|---|
| 713 |
| 4 |
| 2 |
| 1 |

Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum

# Vereinigung

Zwei Relationen mit gleichem Schema – d.h. mit gleichen Attributnamen und Attributtypen – kann man durch Vereinigung zu einer Relation zusammenfassen.

## Beispiel für Vereinigung

Anfrage:

$$\Pi_{\text{Name}}(\text{Studenten}) \bigcup \Pi_{\text{Name}}(\text{Professoren})$$

Ergebnis:

| Name |
|---|
| Alf |
| Frodo |
| Heidi |
| Harry |
| Ronja |
| Klitschko |
| Ali |
| Maske |
| Schulz |



Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum

# Mengendifferenz

Für die zwei Releationen $R$ und $S$ mit gleichem Schema ist die Mengendifferenz:

$$R - S$$

definiert als die Menge der Tupel, die in $R$ aber nicht in $S$ vorkommen.

## Beispiel für Mengendifferenz

Anfrage:

$$\Pi_{\text{MatrNr}}(\text{Studenten}) \; - \; \Pi_{\text{MatrNr}}(\text{hören})$$

Ergebnis:

| MatrNr |
| --- |
| 2403 |
| 2404 |
| 2405 |

# Kartesisches Produkt (Kreuzprodukt)

Das Kreuzprodukt zweier Releationen $R$ und $S$ wird als

$$R \times S$$

gebildet und enthält alle $|R| * |S|$ möglichen Paare von Tupeln aus $R$ und $S$. Das Schema der Ergebnisrelation also **sch**($R \times S$), ist die Vereinigung der Attribute aus **sch**($R$) und **sch**($S$).

$$\textbf{sch}(R \times S) = \textbf{sch}(R) \cup \textbf{sch}(S) = \mathcal{R} \cup \mathcal{S}$$

Das Schema der Relationen $R$ und $S$ wird auch mit $\mathcal{R}$ bzw. $\mathcal{S}$ bezeichnet.



Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum

# Beispiel für kartesisches Produkt

Anfrage:

$$\text{Professoren} \times \text{hören}$$

Ergebnis:

| PersNr | Name | Raum | MatrNr | VorlNr |
|---|---|---|---|---|
| 1013 | Klitschko | C301 | 2401 | 107 |
| 1013 | Klitschko | C301 | 2402 | 107 |
| 1013 | Klitschko | C301 | 2401 | 456 |
| 8516 | Ali | C302 | 2401 | 107 |
| 8516 | Ali | C302 | 2402 | 107 |
| 8516 | Ali | C302 | 2401 | 457 |
| 1005 | Maske | C303 | 2401 | 107 |
| 1005 | Maske | C303 | 2402 | 107 |
| 1005 | Maske | C303 | 2401 | 456 |
| 5313 | Schulz | C304 | 2401 | 107 |
| 5313 | Schulz | C304 | 2402 | 107 |
| 5313 | Schulz | C304 | 2401 | 456 |



Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum

# Umbenennung von Relationen und Attributen

Manchmal ist es notwendig, dieselbe Relation mehrfach in einer Anfrage zu verwenden. Dazu wird – zumindest logisch– eine vollständige zusätzliche Kopie der Relation erzeugt. In diesem Fall muss zumindest eine der Relationen umbenannt werden.

In Anfragen können auch Attributnamen mehrfach vorkommen. Auch hier muss dann eine Umbenennung erfolgen

Für Umbenennungen wird der Operator $\rho$ verwendet. Im Subscript steht der neue Name der Relation bzw. des Attributs.

## Beispiele für Umbenennung

$$\rho_{V1}(\text{voraussetzen})$$

$$\Pi_{\text{V1.Vorgänger}}(\sigma_{\text{V2.Nachfolger=118} \wedge \text{V1.Nachfolger=V2.Vorgänger}}$$
$$(\rho_{\text{V1}}(\text{voraussetzen}) \times \rho_{\text{V2}}(\text{voraussetzen})))$$

Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum

# Umbenennung von Attributnamen

Im Beispiel wird das Attribut *Vorgänger* der Relation *voraussetzen* umbenannt.

$$\rho_{\text{Voraussetzung} \leftarrow \text{Vorgänger}}(\textit{voraussetzen})$$



Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum

# Definition der relationalen Algebra

Die bisher eingeführten Operatoren sind ausreichend, um die relationale Algebra formal definieren zu können. Die Basisausdrücke der rel. Algebra sind entweder Relationen der Datenbank oder konstante Relationen.

## Aufbau von Ausdrücken der rel. Algebra

Ein allgemeiner Relationenalgebra-Ausdruck wird aus „kleineren" Algebraausdrücken kombiniert. Seien $E_1$ und $E_2$ relationale Algebraausdrücke, dann sind die folgenden Ausdrücke auch gültige Algebraausdrücke:

- $E_1 \cup E_2$, wobei $E_1$ und $E_2$ das gleiche Schema haben müssen.
- $E_1 - E_2$, die Schemas müssen auch gleich sein
- $E_1 \times E_2$
- $\sigma_P(E_1)$, mit einem Prädikat $P$ über den Attributen von $E_1$
- $\Pi_S(E_1)$, mit einer Attributliste $S$, deren Attribute in dem Schema von $E_1$ vorkommen
- $\rho_V(E_1)$ und $\rho_{A \leftarrow B}(E_1)$, wobei $B$ ein Attributname der Relation $E_1$ ist, und $A$ nicht als Attributname in $E_1$ vorkommt.


Rel. Algebra

Grundlagen

Beispielrelationen

rel. Algebra
- Selektion
- Projektion
- Vereinigung
- Differenz
- Kreuzprod.
- Umbenennung

Definition der rel. Algebra

Verbundoperatoren
- Natürlicher Verbund
- Allgemeiner Verbund

Operatorbaum

# Natürlicher Verbund

Der natürliche Verbund zweier Relationen $R$ und $S$ wird mit $R \bowtie S$ bezeichnet.

## Definition von $R \bowtie S$

Wenn $R$ insgesamt $m + k$ Attribute $A_1, \ldots, A_m, B_1, \ldots, B_k$ und $S$ $n + k$ Attribute $B_1, \ldots, B_k, C_1, \ldots, C_n$ hat, dann hat $R \bowtie S$ die Stelligkeit $m + n + k$. Es wird angenommen, dass die Attribute $A_i$ und $C_j$ für $(1 \leq i \leq m, 1 \leq j \leq n)$ jeweils paarweise unterschiedlich sind – d.h. $R$ und $S$ haben nur die Attribute $B_1, \ldots, B_k$ als gleichbenannte Attribute. In diesem Fall ist das Ergebnis von $R \bowtie S$ wie folgt definiert:

$$R \bowtie S = \Pi_{A_1,\ldots,A_m,R.B_1,\ldots,R.B_k,C_1,\ldots,C_n}(\sigma_{R.B_1=S.B_1 \wedge \ldots \wedge R.B_k=S.B_k}(R \times S))$$

Logisch wird das Kreuzprodukt gebildet, aus dem dann nur diejenigen Tupel selektiert werden, deren Attributwerte für gleichbenannte Attribute der beiden Argumentrelationen gleich sind. Die gleichbenannten Attribute werden nur einmal in das Ergebnis übernommen.

# Allgemeiner Verbund

Der allgemeine Joinoperator, auch *Theta-Join* genannt, erlaubt die Spezifikation eines beliebigen Joinprädikats $\theta$.

## Definition des Theta-Join ($R \bowtie_\theta S$)

Gegeben sind die Relationen $R$ und $S$ mit den Attributen $A_1, \ldots, A_n$ und $B_1, \ldots, B_m$. Der Theta-Join ist definiert als:

$$R \bowtie_\theta S$$

Hierbei ist $\theta$ ein beliebiges Prädikat über den Attributen $A_1, \ldots, A_n, B_1, \ldots, B_m$.

Der Theta-Join ist also eine vereinfachte Formulierung des kartesischen Produkts gefolgt von einer Selektion:

$$R \bowtie_\theta S = \sigma_\theta(R \times S)$$


Rel. Algebra

Grundlagen

Beispielrelationen

rel. Algebra
- Selektion
- Projektion
- Vereinigung
- Differenz
- Kreuzprod.
- Umbenennung

Definition der rel. Algebra

Verbundoperatoren
- Natürlicher Verbund
- Allgemeiner Verbund

Operatorbaum

# Theta-Join – Beispiele

## Allgemeines Beispiel

Gegeben sind wieder die Relationen $R$ und $S$ mit den Attributen $A_1, \ldots, A_n$ und $B_1, \ldots, B_m$. Ein Theta-Join könnte wie folgt aussehen:

$$R \bowtie_{A_1 > B_1 \wedge A_2 = B_2 \wedge A_3 < B_5} S$$

## Was ermittelt diese Anfrage?

$$\Pi_{Studenten.Name, Professoren.Name}$$
$$(\text{Studenten} \bowtie_{\text{MatrNr=PersNr}} \text{Professoren})$$

# Weitere Join-Varianten

- left outer, right outer und full outer join
- Semi-Join. Ein Semi-Join arbeitet ähnlich einem natürlichen Verbund. Ein Verbund wird dabei aber nicht ausgeführt. Es werden nur die Tupel der linken oder der rechten Tabelle zurückgegeben, die einen Verbundpartner haben. D.h. man bekommt z.B. alle Tupel der linken Tabelle, für die ein Verbundpartner existiert.

$$L \ltimes R = \Pi_{\mathcal{L}}(L \bowtie R)$$

$\mathcal{L}$ bezeichnet die Menge der Attribute von $L$

# Operatorbaum

Bisher wurden Relationenalgebra-Ausdrücke immer „in-line“ dargestellt. Bei komplizierteren Anfragen ist eine sogenannte Operatorbaum-Darstellung üblich.

## Aufbau eines Operatorbaums

- In den Blättern stehen Basisrelationen der Datenbank oder konstante Relationen.
- Operatoren verarbeiten die Tupel der Relationen und erzeugen Zwischenergebnisse.
- Die Zwischenergebnisse werden dann zum nächsten Operator weitergeleitet.
- An der Wurzel des Baums entsteht das Anfrageergebnis.

Der Operatorbaum dient der Optimierung der Anfragen. Weiterhin wird sichtbar, welche Teile der Anfrage parallel abgearbeitet werden können.



Grundlagen
Beispielrelationen
rel. Algebra
Selektion
Projektion
Vereinigung
Differenz
Kreuzprod.
Umbenennung
Definition der rel. Algebra
Verbundoperatoren
Natürlicher Verbund
Allgemeiner Verbund
Operatorbaum