DB: Datenbanken

Relationale Algebra

Prof. Dr. Ludger Martin

Gliederung

- Einführung
- * Relationenorientierte Operationen
- Mengenoperationen
- Beispiele
- * Relationale Algebra und relationale Sprachen

Einleitung

- * 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.
- Die relationale Algebra ist abgeschlossen, d.h. die Ergebnisse der Anfragen sind wieder Relationen.

* Die Selektion realisiert eine Auswahl einer Tupelmenge aus einer Relation (*Zeilenauswahl*). Der *Selektionsoperator* $\sigma_P(r)$ (Sigma) selektiert die Menge der Tupel aus der Relation r, für die das Selektionsprädikat P wahr ist.

* Ein Selektionsprädikat kann Verknüpfungen mit Hilfe der logischen Operatoren AND, OR und

NOT enthalten.

* $\sigma_{einzelpreis>13}(b_beinhaltet_m)$

a_nr	be_nr	anzahl	einzelpreis
0001-E	BE-0020	1	4.99
0007-B	BE-0020	1	11.9
0008-B	BE-0010	2	8 9
0010-T	BE-0049	1	14.9
0012-T	BE-0023	1	14.9
0017-V	BE-0010	1	14.9
0018-V	BE-0056	3	17.9

- * Die Projektion realisiert eine Auswahl einer Attributliste aus einer Relation (*Spaltenauswahl*). Der *Projektionsoperator* $\pi_A(r)$ (Pi) führt die Projektion auf die Attribute der Liste A aus der Relation r durch.
- * $\pi_{anr,einzelpreis}(b_beinhaltet_m)$

a nr	be nr	anzahl	einzelpreis
0001-E	8E-0020	1	4,99
0007-B	E-0020	1	11.9
0008-B	BE-0010	2	8.9
0010-T	8E-0049	1	14.9
0012-T	BE-0023	1	14.9
0017-V	BE-0010	1	14.9
0018-V	BE-0056	3	17.9

Operationen können auch kombiniert werden

$$\pi_{anr,einzelpreis}(\sigma_{einzelpreis>13}(b_beinhaltet_m))$$

alternativ

$$\sigma_{\textit{einzelpreis}>13}(\pi_{\textit{anr},\textit{einzelpreis}}(b_\textit{beinhaltet}_m))$$

a_nr	be_nr	anzahl	einzelpreis
0001-E	BE-0020	1	4.99
0007-B	BE-0020	1	11.9
0008-B	BE-0010	2	8.9
0010-T	BE-0049	1	14.9
0012-T	BE-0023	1	14.9
0017-V	BE-0010	1	14.9
0018-V	BE-0056	3	17.9

- * Der innere Verbund oder auch inner join realisiert das Zusammenfügen von Relationen. Der Verbundoperator $r_1 \bowtie_P r_2$ verkettet Tupel aus Relation r_1 und Relation r_2 miteinander, wenn sie das Verbundprädikat P erfüllen.
- * $b_beinhaltet_m \bowtie_{anr=anr} medienartikel$

a_nr	be_nr	anzahl	einzelpreis	titel	jahr
0001-E	BE-0020	1	4.99	LaTeX in 21 Tagen	2004
0007-B	BE-0020	1	11.9	Simplify Your Life	2004
0008-B	BE-0010	2	8.9	Ich bin dann mal weg	2006
0010-T	BE-0049	1	14.9	The Best of Michael Jackson	1990
0012-T	BE-0023	1	14.9	Piece By Piece	2005
0017-V	BE-0010	1	14.9	Pirates of the Caribbean	2004
0018-V	BE-0056	3	17.9	The Da Vinci Code	2006

- * Als natürlicher Verbund (natural join) wird ein Verbund bezeichnet, der zwei Relationen r_1 und r_2 verknüpft, die gleichnamige Attribute enthalten. Die explizite Angabe des Verbundprädikats entfällt: $r_1 \bowtie r_2$
- * Der äußere Verbund verkettet Tupel aus Relation r_1 und Relation r_2 miteinander, wenn sie das Verbundprädikat P erfüllen. Eine Relation wird vollständig übernommen.
 - *Linker äußere Verbund (left outer join): $r_1 \bowtie_P r_2$
 - *Rechter äußere Verbund (right outer join): $r_1 \bowtie_P r_2$

- * Die Umbenennung ändert den Namen eines Attributes einer Relation und damit dessen Relationenschema. Der Operator $\rho_{Aneu \leftarrow Aalt}(r)$ (Rho) benennt in der Relation r das Attribut Aalt in Aneu um.
- * $\rho_{name \leftarrow titel}(medienartikel)$
- * Der Operator $\rho_s(r)$ benennt die Relation r in s um.
- * $\rho_{ma}(medienartikel)$

a_nr	name	jahr
0001-E	LaTeX in 21 Tagen	2004
0003-E	Harry Potter III	1997
0005-B	Der Schwarm	2004
0006-B	LaTeX kurz und gut	2004
0007-B	Simplify Your Life	2004
0008-B	Ich bin dann mal weg	2006
0009-T	Back to Bedlam	2005
0010-T	The Best of Michael Jackson	1990
0011-T	Music	2000
0012-T	Piece By Piece	2005
חכוחה	Diretas of the Caribbana	2004

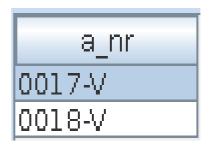
- * Der Vereinigungsoperator $r_1 \cup r_2$ übernimmt aus Relation r_1 und r_2 alle Tupel in die Ergebnisrelation und entfernt doppelte Tupel.
- * $\pi_{anr}(video) \cup \pi_{anr}(dvd)$

a_nr
0017-V
0018-V
0019-V
0020-V
0013-D
0014-D
0015-D
0016-D

- * Der Differenzoperator r_1-r_2 entnimmt aus Relation r_1 die Tupel, die nicht in Relation r_2 enthalten sind.
- * $\pi_{anr}(medienartikel) \pi_{anr}(video)$

a_nr
0001-E
0003-E
0005-B
0006-B
0007-B
0008-B
0009-T
0010-T
0011-T
0012-T
0013-D
0014-D
0015-D
0016-D

- * Der Durchschnittsoperator $r_1 \cap r_2$ übernimmt die Tupel, die sowohl in Relation r_1 als auch in Relation r_2 enthalten sind.
- * $\pi_{anr}(video) \cap \pi_{anr}(b_beinhaltet_m)$



- * Das kartesische Produkt (Kreuzprodukt) $r_1 \times r_2$ multipliziert die Tupel der Relationen r_1 mit den Tupeln der Relationen r_2 .
- ⋆ video × b beinhaltet m

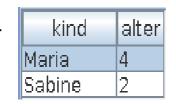
a_nr	spra	a_nr	be_nr	anzahl	einzelpreis
0017-V	en US	0001-E	BE-0020	1	4.99
0018-V	de DE	0001-E	BE-0020	1	4.99
0019-V	de DE	0001-E	BE-0020	1	4.99
0020-V	de DE	0001-E	BE-0020	1	4.99
0017-V	en US	0007-B	BE-0020	1	11.9
0018-V	de DE	0007-B	BE-0020	1	11.9
0019-V	de DE	0007-B	BE-0020	1	11.9
0020-V	de DE	0007-B	BE-0020	1	11.9
0017-V	en US	0008-B	BE-0010	2	8.9
00187A	de DE	0008-B	BE-0010	2	8.9

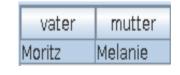
- * Die Division $r_1 \div r_2$ erzeugt eine Relation r(X), die alle Tupel t[X] aus $r_1(Z)$ beinhaltet, welche in r_1 in Kombination mit jedem Tupel aus $r_2(Y)$ erscheinen, wobei gilt: $Z = X \cup Y$ (X, Y, Z Attributmengen)
- Die Division ist definiert durch

$$r_1 \div r_2 = \pi_Z(r_1) - \pi_Z((\pi_Z(r_1) \times r_2) - r_1)$$

★ Beispiel Division: eltern÷kind

vater	mutter	kind	alter
Franz	Helga	Harald	5
Franz	Helga	Maria	4
Franz	Ursula	Sabine	2
Moritz	Melanie	Gertrud	7
Moritz	Melanie	Maria	4
Moritz	Melanie	Sabine	2
Peter	Christina	Robert	9





* Dividiert man *eltern* durch *kind*, so erhält man als Ergebnis eine Relation, die nur noch diejenigen Ehepaare enthält, die sowohl eine Tochter Maria mit Alter 4, als auch eine Tochter Sabine mit Alter 2 haben.

Beispiele

* Welcher Regisseur (Vor- und Nachname) arbeitet für welches Studio?

$$\pi_{name, vorname, studio}(person \bowtie regisseur)$$

★ Welche Artikel (Titel) sind teurer als 7€?

$$\pi_{titel}(medienartikel \bowtie \sigma_{preis>7}(angebot))$$

Relationale Algebra und relationale Sprachen

* In SQL werden die Operatoren durch entsprechende Sprachelemente dargestellt.

```
Projektion: \pi_{x,y,z}(R)
                           SELECT x, y, z FROM R
Selektion: \sigma_{y < 5}(R)
                           SELECT * FROM R WHERE x < 5
Join: R \bowtie_{R,nr=S,nr} S
                           SELECT R.*, S.* FROM R
                             INNER JOIN S ON R.nr = S.nr
Umbenennung: \rho_{nummer \leftarrow nr}(S)
                           SELECT nr AS nummer FROM R
Vereinigung: R \cup S
                           SELECT * FROM R UNION
                              SELECT * FROM S
Differenz: R-S
                           SELECT * FROM R EXCEPT
                              SELECT * FROM S
Durchschnitt: R \cap S
                           SELECT * FROM R INTERSECT
                             SELECT * FROM S
Kartesisches P.: R \times S
                           SELECT R.*, S.* FROM R, S
```

EXCEPT und INTERSECT nicht von MySQL unterstützt

Relationale Algebra und relationale Sprachen

- * Eine DB-Sprache ist dann relational vollständig, wenn sie mindestens die relationenorientierten Operatoren *Projektion* und *Selektion* und die mengenorientierten Operatoren *Vereinigung*, *Differenz* und *kartesisches Produkt* umsetzt.
 - → SQL wird daher als relational vollständig bezeichnet

Literatur

- * Thomas Kudraß: Taschenbuch Datenbanken, Hanser, 2007
- * Wikipedia: Relationale Algebra, https://de.wikipedia.org/wiki/Relationale_Algebra