

13. März 2020

# Relationale Algebra

**Motivation:** Wir sind mit der relationalen Anfragesprache in der Lage das Schema einer Relation zu beschreiben. Wir können also neben der konzeptionellen Beschreibung durch das ER-Modell nun auch die Strukturbeschreibung der Datenbank realisieren. Was aber bisher auf Anwenderebene fehlt ist eine Beschreibung von Anfragen auf einer Datenbank. Dazu kann man SQL benutzen. Allerdings möchte man sich von der konkreten Datenbank-Anfragesprache lösen und eine abstrakte Möglichkeit in Form einer allgemeinen Anfragesprache schaffen. Dazu nutzt man die *relationale Algebra*, die den Prozess der Anfragen und deren Auswertung mithilfe von geeigneten Operatoren beschreibt. Es gibt noch andere Möglichkeiten, die Ergebnisse von Anfragen zu beschreiben. Das Relationenkalkül nutzt deklarative Ausdrücke, um anzugeben, welche Ergebnismenge man anhand festgelegter Prädikate erhält. Dies sind *deklarative Anfragesprachen* (z. B. relationales Tupelkalkül)

Insgesamt gibt es 6 Basisoperatoren:

- Selektion  $\sigma$
- Projektion  $\pi$
- Kreuzprodukt  $\times$
- Vereinigung  $\cup$
- Differenz  $-$
- Umbenennung  $\rho$

Diese Operatoren definieren eine Algebra. Die relationale Algebra ist abgeschlossen, d.h. sie liefert Relationen als Ergebnis der Auswertung.

Neben den Basisoperatoren gibt es noch eine Reihe spezieller Operatoren (z. B. für die Joins, siehe 1.7), die aber alle durch die Basisoperatoren dargestellt werden können.

## 1 Definitionen der Operatoren

Seien  $R_1, R_2, R_3$  Relationen mit folgenden Ausprägungen:

$R_1$		
A	B	C
$a_1$	$b_1$	$c_1$
$a_2$	$b_2$	$c_2$
$a_3$	$b_3$	$c_3$

z. B.

Vorlesungen			
VorlNr	Titel	SWS	gelesenVon
5001	Grundzüge	4	2137
5041	Ethik	4	2125
5043	Erkenntnistheorie	3	2126
5049	Mäeutik	2	2125
4052	Logik	4	2125
5052	Wissenschaftstheorie	3	2126
5216	Bioethik	2	2126
5259	Der Wiener Kreis	2	2133
5022	Glaube und Wissen	2	2134
4630	Die 3 Kritiken	4	2137

$R_2$		
E	F	G
$e_1$	$f_1$	$g_1$
$e_2$	$f_2$	$g_2$
$e_3$	$f_3$	$g_3$

z. B.

Professoren			
PersNr	Name	Rang	Raum
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	052
2134	Augustinus	C3	309
2136	Curie	C4	036
2137	Kant	C4	007

$R_3$		
H	I	J
$h_1$	$i_1$	$j_1$
$h_2$	$i_2$	$j_2$
$h_3$	$i_3$	$j_3$

z. B.

Assistenten			
PersNr	Name	Fachgebiet	Boss
3002	Platon	Ideenlehre	2125
3003	Aristoteles	Syllogistik	2125
3004	Wittgenstein	Sprachtheorie	2126
3005	Rhetikus	Planetenbewegung	2127
3006	Newton	Keplersche Gesetze	2127
3007	Spinoza	Gott und Natur	2134

## 1.1 Selektion

Bei einer Selektion werden diejenigen Tupel einer Relation ausgewählt, die einem bestimmten Prädikat genügen. Die Anfrage  $\sigma_P(R)$  liefert alle  $t \in R$ , die P erfüllen, also  $\sigma_{t.A > Wert}(R_1)$ .

Beispiel:

$\sigma_{SWS > 3}(Vorlesungen)$			
VorlNr	Titel	SWS	gelesenVon
5001	Grundzüge	4	2137
5041	Ethik	4	2125
4052	Logik	4	2125
4630	Die 3 Kritiken	4	2137

$\sigma_{SWS > 3}(Vorlesungen) \rightsquigarrow$

## 1.2 Projektion

Bei einer Projektion werden alle Tupel einer Relation ausgewählt und die Ergebnisrelation auf die angegebenen Attributwerte reduziert.  $\pi_{A_1, A_2, \dots, A_n}(R)$ ,  $n \in \mathbb{N}_0$  liefert alle  $t \in R$  mit  $t = (t.A_1, t.A_2, \dots, t.A_n)$ . Mit  $R_1$  würde z. B.  $\pi_A(R_1)$  alle Werte des Attributes A liefern.

Beispiel:

$\pi_{VorlNr, Titel}(Vorlesungen)$	
VorlNr	Titel
5001	Grundzuege
5041	Ethik
5043	Erkenntnistheorie
5049	Mäeutik
4052	Logik
5052	Wissenschaftstheorie
5216	Bioethik
5259	Der Wiener Kreis
5022	Glaube und Wissen
4630	Die 3 Kritiken

$\pi_{VorlNr, Titel}(Vorlesungen) \rightsquigarrow$

### 1.3 Vereinigung

Da Relationen Mengen von Tupel darstellen, können diese auch vereinigt werden. Dies setzt aber gleiche Schemata voraus. Sollten die Schemata verschieden sein aber gleiche Attributwerte (d.h. Wertebereiche) beinhalten, so kann man diese Attribute durch vor-gesetzte Projektion separieren und dann vereinigen.

Beispiel:

$$\text{Abfrage} := \pi_{\text{Name}}(\text{Assistenten}) \cup \pi_{\text{Name}}(\text{Professoren}) \rightsquigarrow$$

Abfrage
Name
Platon
Aristoteles
Wittgenstein
Rhetikus
Newton
Spinoza
Sokrates
Russel
Kopernikus
Popper
Augustinus
Curie
Kant

Man beachte, dass bei der Vereinigung eine Duplikateeliminierung stattfinden muss.

### 1.4 Mengendifferenz

Nach dem gleichen Prinzip kann man auch eine Differenz  $R_1 - R_2$  definieren.

Beispiel: Finde ein eigenes Beispiel!

### 1.5 Kartesisches Produkt

Das Kreuzprodukt zweier Relationen  $R_1 \times R_2$  wird entsprechend der (aus der Mathematik bekannten) Art und Weise gebildet, indem jedes Element  $t \in R_1$  mit jedem Element  $t \in R_2$  kombiniert wird. Die Ergebnisrelation ist dabei diejenige Relation, die alle Attribute aus  $R_1$  und  $R_2$  beinhaltet. Seien  $\mathcal{R}$  das Schema einer Relation  $R$ . Dann ist  $\mathcal{R}_1 \times \mathcal{R}_2$  die Vereinigung der Attributwerte von  $R_1$  und  $R_2$ .

Beispiel:

$Professoren \times Vorlesungen \rightsquigarrow$

<i>Professoren × Vorlesungen</i>							
PersNr	Name	Rang	Raum	VorlNr	Titel	SWS	gelesenVon
2125	Sokrates	C4	226	5001	Grundzuege	4	2137
2126	Russel	C4	232	5041	Ethik	4	2125
...	...	...	...	...	...	...	...

## 1.6 Umbenennung

Manchmal ist es notwendig Relationen umzubenennen. Zum einen, weil die Relation in komplexen Abfragen mehrfach verwendet wird.

$\rho_{neuerName}(R)$  benennt in der Relation  $R$  in die Relation  $neuerName$  um.

Beispiel: klar

Dieser Operator kann aber auch für die Umbenennung von Attributnamen in Relationen verwendet werden. So wird mit  $\rho_{neuerName \leftarrow alterName}(R)$  das Attribute  $alterName$  in  $neuerName$  umbenannt.

## 1.7 weitere Operatoren

Division  $\div$ , Mengendurchschnitt  $\cap$ , (linker) Semi-Join  $\bowtie$ , (rechter) Semi-Join  $\bowtie$ , outer (äußerer) Join (linker äußerer Join  $\bowtie$  rechter äußerer Join  $\bowtie$ )

Aufgabe: Studiere im Kemper/Eickler die entsprechenden Seiten zu den Joins.

## 2 Übungen

1. Aufgabe: Formuliere die Anfragen aus der Klausur in relationale Algebra

Ergebnisse:  $\{[\underline{EID}: \text{integer}, \text{Platz}: \text{integer}, \text{Ergebniswert}: \text{float}]\}$

Athleten:  $\{[\underline{\text{TeilnID}}: \text{integer}, \text{Name}: \text{string}, \text{Nation}: \text{string}]\}$

Helfer:  $\{[\underline{\text{HelferID}}: \text{integer}, \text{Name}: \text{string}]\}$

Austragungsorte:  $\{[\underline{\text{Adresse}}: \text{string}, \underline{\text{Stadt}}: \text{string}]\}$

Veranstaltungen:  $\{[\underline{\text{VID}}: \text{integer}, \text{Termin}: \text{date}, \text{Disziplin}: \text{string}]\}$



findet\_statt: {[VID: integer, Adresse: string, Stadt: string]}  
nimmt\_teil: {[VID: integer, TeilnID: integer, EID: integer]}  
bewertet: {[HelferID: integer, VID: integer]}  
betreut: {[HelferID: integer, VID: integer]}

- a) Gesucht sind die Namen aller deutschen Teilnehmer.
- b) Gesucht ist der Austragungsort des Marathonlaufes.

2. Aufgabe: Welches Ergebnis liefern die folgenden Anfragen in der Relationenalgebra?  
Formuliere auch in SQL!:

- a)  $\sigma_{Semester > 10}(Studenten)$
- b)  $\pi_{Rang}(Professoren)$
- c) Professoren x hören
- d) c formuliert als Join
- e) Finde MatrNr der Studenten, die alle vierstündigen Vorlesungen hören
- f) Welcher Professor liest „Mäeutik“ ?
- g) Welche Studenten hören welche Vorlesungen?

3. Aufgabe: Wer ist der Chef von Heinz, formuliert in relationaler Algebra. ;-)