Datenbanken I (T2INF2004)

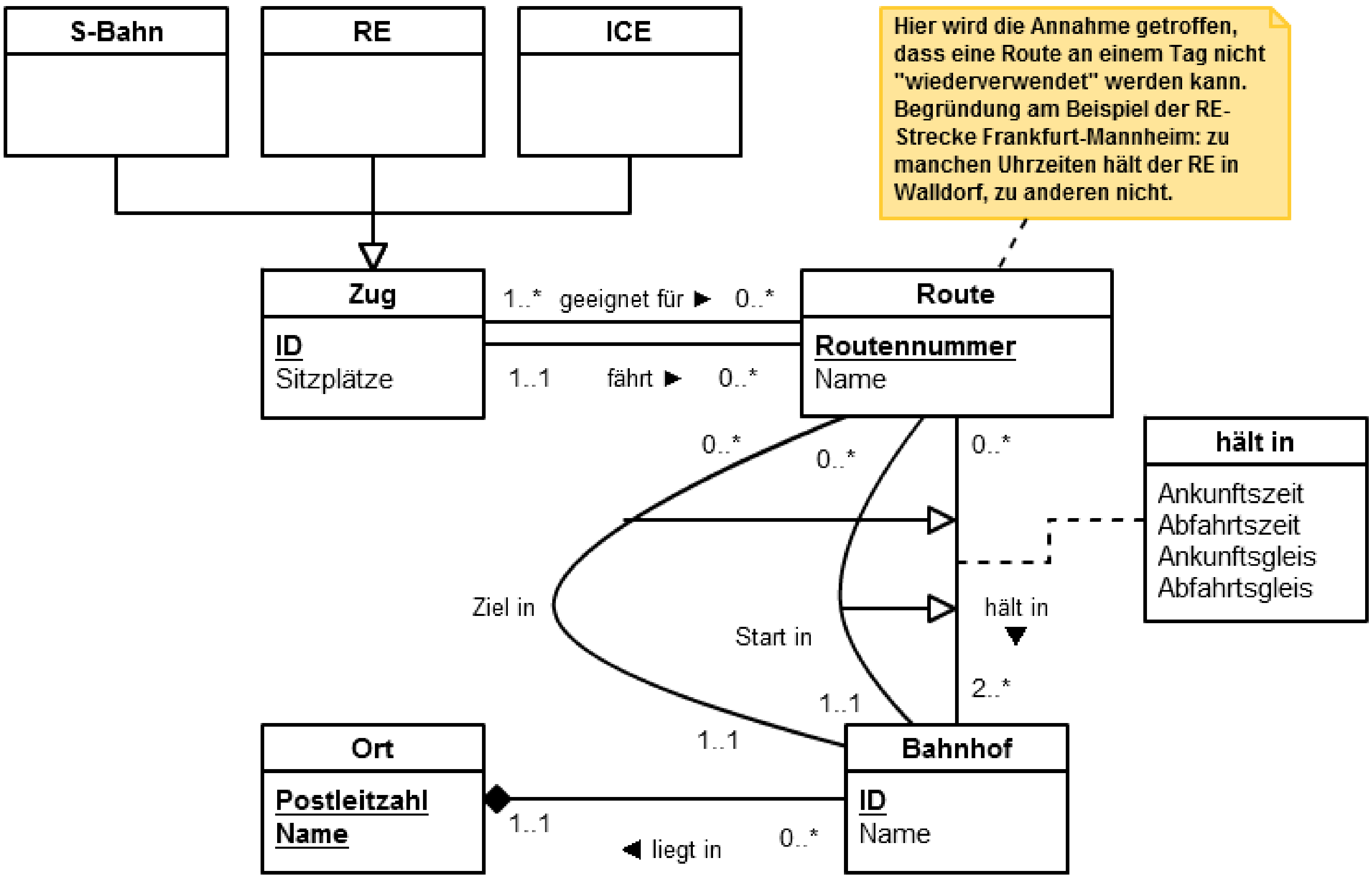
Foliensatz 4: Mengenlehre, Logik und Relationenalgebra

Uli Seelbach, DHBW Mannheim, 2023

Foliensatz freundlicherweise zur Verfügung gestellt von Mirko Schick

Hausaufgabe Foliensatz 3

DB-Entwurf mit UML, Lösungsvorschlag



Hausaufgabe Foliensatz 3

Schritt 1 ohne Zusammenfassen

S-Bahn: {[ID: integer, ...]}   RE: {[ID: integer, ...]}   ICE: {[ID: integer, ...]}

Zug: {[ID: integer, Sitzplätze: integer]}

Zug_geeignet_fuer_route: {[ZugID:integer, RouteID: integer]}

Route: {[Routennummer: integer, Name: String]}   gefahren_von: {[RouteID: integer, ZugID: Integer NOT NULL]}

Start_in: {[Routennummer: integer, Bahnhof: Integer]}

Ziel_in: {[Routennummer: integer, Bahnhof: Integer]}

Hält_in: {[Routennummer: integer, Bahnhof: Integer, Ankunftszeit: Date,

Ankunftsgleis: String, Abfahrtszeit: Date, Abfahrtsgleis: String]}

Bahnhof: {[ID: integer, Name: String, PLZ: String NOT NULL, Ort: String NOT NULL]}

Ort: {[PLZ: String, Ort: String]}

Hausaufgabe Foliensatz 3

Schritt 2 mit Zusammenfassen

NOT NULL bei Route.Start_in bzw. Ziel_in schlechte Idee. Warum eigentlich?

S-Bahn: {[ID: integer, ...]}

RE: {[ID: integer, ...]}

ICE: {[ID: integer, ...]}

Zug: {[ID: integer, Sitzplätze: integer]}

Zug_geeignet_fuer_route: {[ZugID:integer, RouteID: integer]}

Route: {[Routennummer: integer, Name: String, gefahren_von: integer NOT NULL, Start_in: integer, Ziel_in: integer]}

Hält_in: {[Routennummer: integer, Bahnhof: Integer, Ankunftszeit: Date, Ankunftsgleis: String, Abfahrtszeit: Date, Abfahrtsgleis: String]}

Bahnhof: {[ID: integer, Name: String, PLZ: String NOT NULL, Ort: String NOT NULL]}

Ort: {[PLZ: String, Ort: String]}

Wenn ein FK-Value (teilweise) NULL ist, dann muss der Wert in der Referenztabelle nicht vorkommen. In diesem Fall beziehen sich Route und Hält_in aufeinander... Wo also zuerst ein INSERT? Irgendwo wird immer ein Wert fehlen...

Datenbanken — in der Theorie
Relationenalgebra

- Präsentiert 1970 von Edgar Codd
- Basiert auf der mathematischen Modellierung von Tabellen in Form von Relationen
- Um gewünschte Informationen aus einem Pool von Relationen zu extrahieren, bedarf es definierter Vorgehen → Relationenalgebra
- Dazu sollten aber ein paar Grundlagen aus der Mengenlehre und Logik wiederholt werden

Mengenlehre

Definition einer Menge

"Unter einer Menge verstehen wir eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.“

Georg Cantor (1845-1918), Begründer der Mengenlehre

- Eine **Menge** besteht aus **Elementen**
- $e \in M$ bedeutet dann, dass e der Menge M angehört

Mengenlehre

Darstellung von Mengen

Generell auf zwei Wegen möglich	
Explizit durch Aufzählen der Elemente	
generisch	{x1, x2, …}
Beispiel	{2,3,5,7}
Implizit durch Angeben eines Ausdrucks A zur Bildung der Menge	
generisch	{x | A(X) }
Beispiel	{x | „x ist eine Primzahl kleiner 10“}

Mengenlehre
Besonderheiten

- Die Menge mit 0 Elementen wird „leere Menge“ genannt
- Die Reihenfolge der Elemente einer Menge spielt keine Rolle
  → {1,2,3} ≡ {2,3,1}
  → Wenn das eine Rolle spielen soll, benötigt man „Listen“
- Die Anzahl nicht unterscheidbarer Elemente ist irrelevant (eine Menge enthält jedes Element nur einmal)
  → {1,2,3,3,3,3} ≡ {2,3,1}
  → Bei Bedarf: Multimengen, Notation z. B. mit Doppelklammer: {{1,2,3,3,3}}

Mengenlehre
Kardinalität

- Kardinalität oder Mächtigkeit drückt die Anzahl der Elemente einer Menge aus
- Geschrieben: |M|
- Leere Menge: Kardinalität 0

?

Kardinalität folgender Mengen?
- {1,2,3,4,5}
- {1,2,33,4,55,55}
- {{1,2,3},{4,5}}

Mengenlehre
Operatoren

- A heißt Teilmenge von B, wenn jedes Element von A auch Element von B ist.

$$\boxed{A \subseteq B}$$

- Zwei Mengen A und B sind gleich, wenn A Teilmenge von B und B Teilmenge von A ist.

$$\boxed{A = B}$$

- A heißt echte Teilmenge von B, wenn A Teilmenge von B ist, aber wenn A nicht gleich B ist.

$$\boxed{A \subset B}$$

 Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Mengenlehre

Operatoren

Durchstreichen eines Mengenvergleichsoperators bzw. des Elementoperators bedeutet Negieren der jeweiligen Beziehung:

$A \nsubseteq B$	A ist <u>nicht</u> Teilmenge von B
$A \neq B$	A und B sind <u>ungleich</u>
$A \not\subset B$	A ist <u>keine</u> echte Teilmenge von B
$e \notin B$	e ist <u>kein</u> Element von B

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Mengenlehre

Operatoren

- Die Mengenlehre kennt drei grundlegende Operationen, mit denen man zwei Mengen verknüpfen kann:

Vereinigung	$A \cup B$	$=_{def}$	$\{ e \mid e \in A$ **oder** $e \in B \}$
Durchschnitt	$A \cap B$	$=_{def}$	$\{ e \mid e \in A$ **und** $e \in B \}$
Differenz	$A \setminus B$	$=_{def}$	$\{ e \mid e \in A$ **und** $e \notin B \}$

- Beispiele dazu:

$$\{1, 2\} \cup \{2, 3\} = \{1, 2, 3\}$$

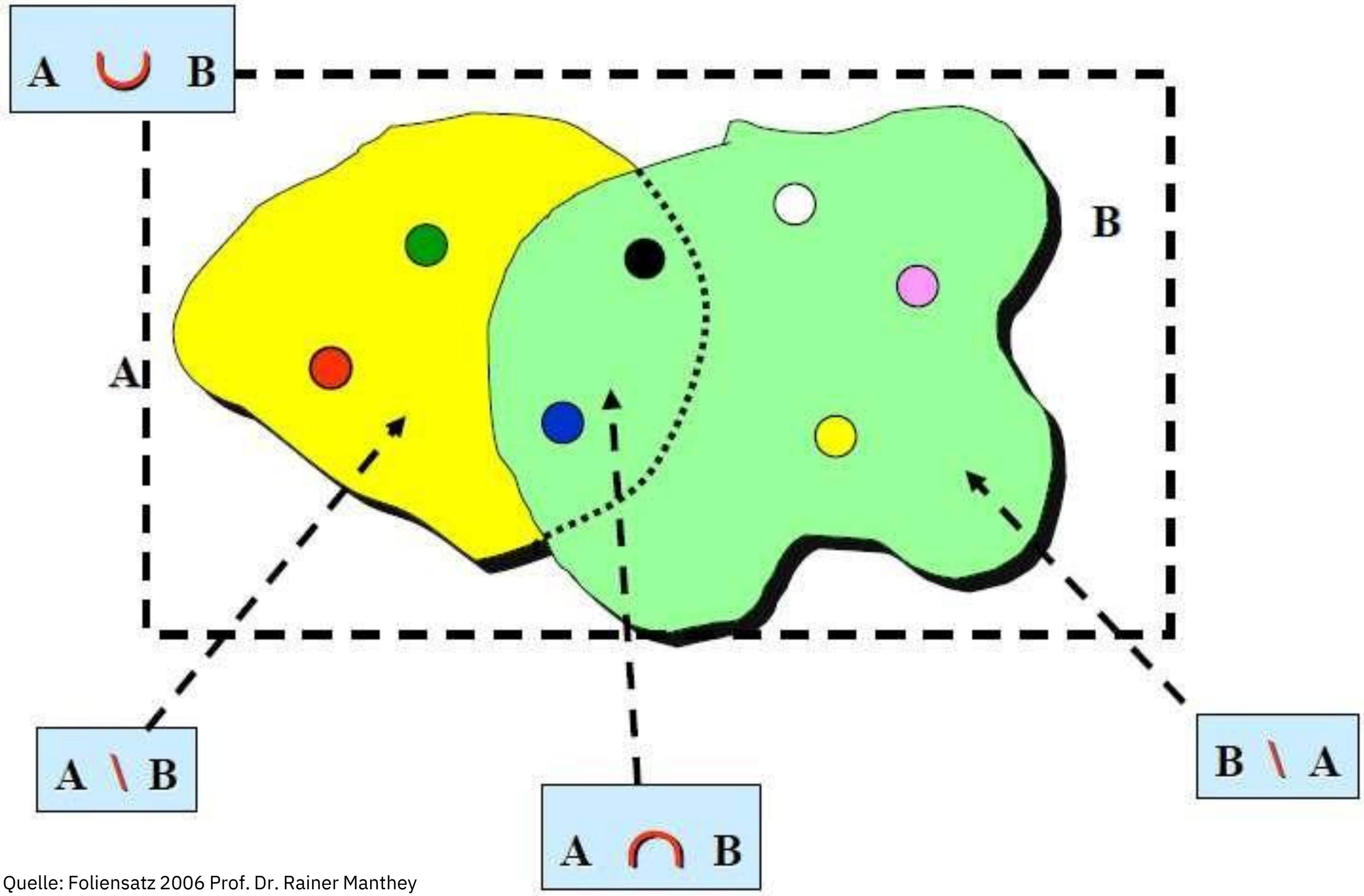
$$\{1, 2\} \cap \{2, 3\} = \{2\}$$

$$\{1, 2\} \setminus \{2, 3\} = \{1\}$$

 Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Mengenlehre

Beispiel für Anwendung von Mengenoperatoren



 Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Mengenlehre

Kartesisches Produkt / Tupelbegriff

- weitere zweistellige Grundoperation der Mengenlehre:

$$\boxed{A \times B = \{ (a, b) \mid a \in A \text{ und } b \in B \}}$$

(kartesisches) Produkt

- verallgemeinerte Produktbildung für n Mengen ($n \geq 2$) :

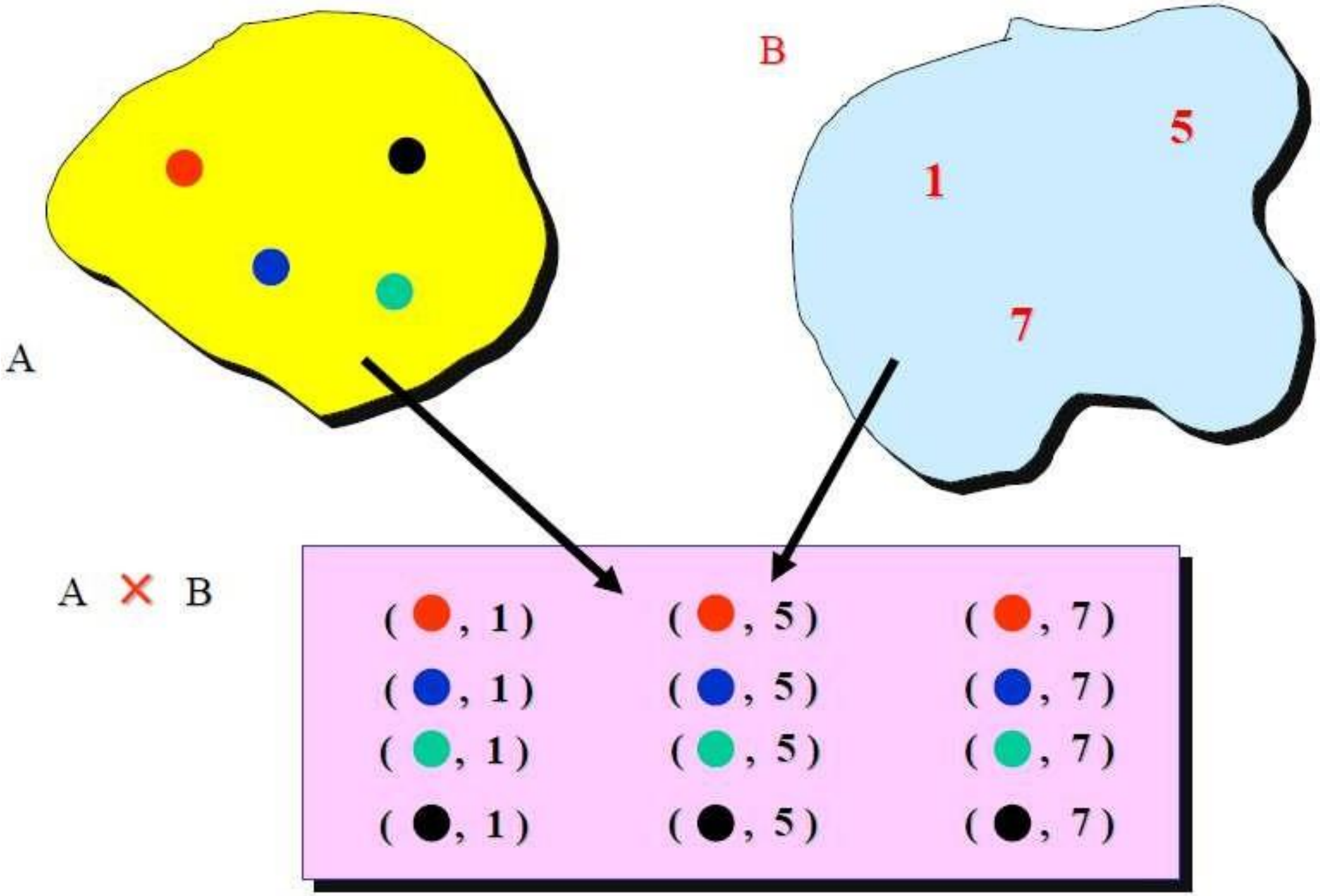
$$\boxed{A_1 \times \ldots \times A_n = \{ (a_1, \ldots, a_n) \mid a_i \in A_i \}}$$

- Elemente des Produkts von n Mengen heißen (n)-Tupel.

- spezielle Bezeichnungen für Tupel:
  - $n = 2$: Paare
  - $n = 3$: Tripel
  - $n = 4$: Quadrupel

 Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Mengenlehre
Produktbildung zweier Mengen



Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Mengenlehre
Tupel und Relationen

- Mengen von Tupeln werden als Relationen bezeichnet
- Jede Teilmenge des kartesischen Produktes aus $A \times B \times \ldots \times Z$ ist eine Relation über A, B, ..., Z:

$$R \subseteq A \times B \times \ldots \times Z$$

- Relationen werden zur Veranschaulichung meist tabellenförmig dargestellt

Logik
Geschichte

- Grundlegend sind die beiden Gebiete
  - Aussagenlogik (formalisiert von Boole, Mitte 19. Jhd.)
    - Lehre von Aussagen und deren Verknüpfungen
    - Aussage: wahr oder falsch.
  - Prädikatenlogik (formalisiert von Frege, Ende 19. Jhd.)

Logik
Aussagenlogik

- Mehrere Aussagen könne zusammengesetzt werden und bilden eine neue Aussage → dies kann anhand von Wahrheitstabellen dargestellt werden
- Einstellige und mehrstellige Junktoren
- Zweistellige Junktoren sind beispielsweise die Konjugation, Äquivalenz etc.
- Einstelliger Junktor zum Beispiel die Negation

Logik
Junktoren

- Zweistellige Junktoren der Aussagenlogik:

Konjunktion	$\wedge$	"und"	(auch: &)
Disjunktion	$\vee$	"oder"	(auch: |)
Implikation	$\Rightarrow$	"wenn . . , dann . . "	
Äquivalenz	$\Leftrightarrow$	"genau dann . . , wenn . . "	

- Einstelliger Junktor:

Negation	$\neg$	"nicht"

 Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Logik

zusammengesetzte Aussagen

- Wahrheitswerte zusammengesetzter Aussagen lassen sich systematisch aus den Wahrheitswerten ihrer Teilaussagen herleiten, z.B.:

$$\boxed{\overbrace{\overbrace{(\underbrace{(5<6)}_{\text{wahr}} \wedge \underbrace{(6<7)}_{\text{wahr}})}^{\text{wahr}} \Rightarrow \overbrace{(5<7)}^{\text{wahr}}}^{\text{wahr}}}$$

- Mit einer ganz analogen Konstruktion lässt sich aber auch dieser Aussage der Wahrheitswert 'wahr' zuordnen:

!

$$\boxed{((5<6) \wedge (a \neq b)) \Rightarrow (5<7)}$$

- Fazit: Teilaussagen (wahrer) zusammengesetzter Aussagen müssen nicht unbedingt "etwas miteinander zu tun" haben.

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Logik

Wahrheitstafeln

- Junktoren sind also syntaktische "Werkzeuge", mit denen sich die Bedeutung ("Semantik") von zusammengesetzten Aussagen aus der Bedeutung der Teilaussagen herleiten lässt.

- Wie dies zu geschehen hat, wird i.a. durch sogenannte Wahrheitstafeln festgelegt:

z.B.: Wahrheitstafel für Konjunktion

A	B	$A \wedge B$
W	W	W
W	F	F
F	W	F
F	F	F

W: wahr
F: falsch

- wie folgt zu lesen: Wenn A wahr ist und B falsch, dann hat die Aussage $A \wedge B$ den Wahrheitswert falsch.

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra
Einführung

- Was ist eine „Algebra“ oder algebraische Struktur?
  - System von Operatoren (korrekterweise „Verknüpfungen“ genannt), die auf einer bestimmten (im Regelfall nichtleeren) Trägermenge operieren: Inputparameter stammen aus dieser Trägermenge, Resultat ebenfalls!
  - Daraus folgt: Operatoren können auf Resultate anderer Operatoren angewendet werden (Schachtelung)
  - Beispiele
    - x+2x=3x (elementare Algebra)
    - true OR false = true (Boolesche Algebra)
    - Mengenalgebra
    - Kennen Sie jQuery? Es gibt Parallelen!
  - Siehe auch: http://de.wikipedia.org/wiki/Algebraische_Struktur

- Relationale Algebra weist viele Parallelen mit der Mengenalgebra auf, aber
  es handelt sich um ganz besondere Mengen: Tupelmengen
  - Diese benötigen ein anderes Verhalten einiger Operatoren

Relationale Algebra
Vergleich mit Mengenalgebra

- Da es sich bei Relationen um Mengen handelt, sind Operatoren der Mengenalgebra ebenfalls anwendbar:

Operator	Beispiel
Vereinigung	**R ∪ S**
Differenz	**R - S**
Durchschnitt	**R ∩ S = R - (R - S)**
Produkt	**R × S**

- Probleme in der Relationalen Algebra:

<table>
<tr><th>Operator</th><th>Problem</th></tr>
<tr><td>Vereinigung</td><td rowspan="3">Eventuell sind Relationen aufgrund der Tupel nicht vereinbar / vergleichbar</td></tr>
<tr><td>Differenz</td></tr>
<tr><td>Durchschnitt</td></tr>
<tr><td>Produkt</td><td>Attribute sind ggf. gleich benannt und werden somit mehrdeutig</td></tr>
</table>

Relationale Algebra
Vereinigungsverträglichkeit

Problem: Alle Verknüpfungen von Relationen mit Mengenoperatoren sind wieder Mengen, aber nicht alle sind auch wieder Relationen!

R:

PersNr	Name

S:

MatrNr	Name	Fach

$R \cup S$ ist keine Relation!

$\Rightarrow$ nur Relationen „gleichen Typs“ können vereinigt / geschnitten / subtrahiert werden

a) gleiche Stelligkeit

b) Namensgleichheit der entsprechenden Spalten

- ggf. Umbenennung von Spalten erforderlich
- Umbenennung mittels Hilfsoperator ρ (griech. rho) , z.B.: $\rho_{A \leftarrow B}(R)$

c) Typgleichheit der Spalten (gleiche Wertebereiche d_i)

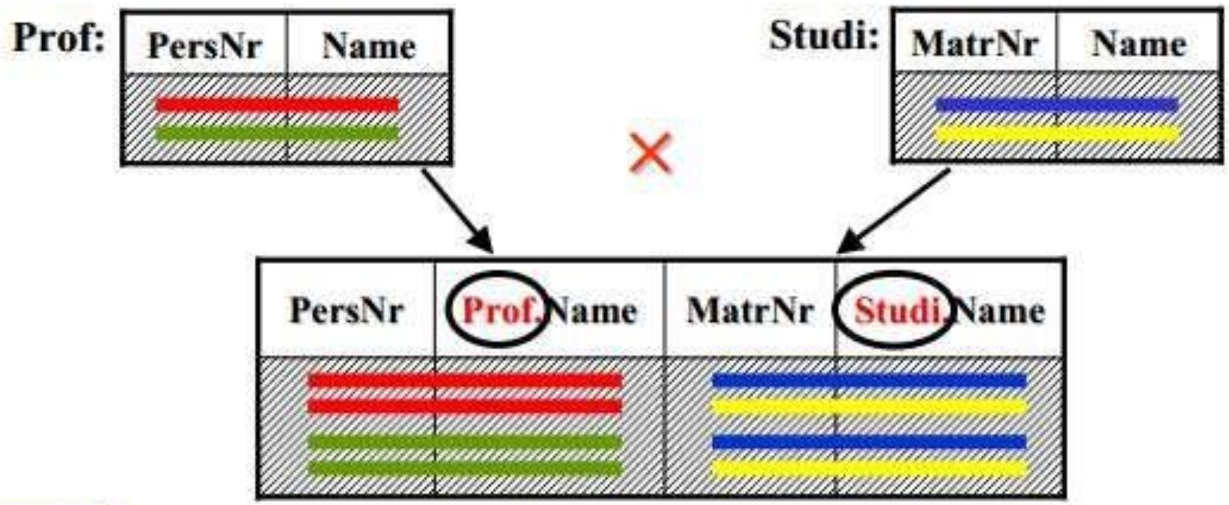
Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra
Attributmehrdeutigkeite

bei Produktbildung: R × S

Identische Attribute in beiden Operanden machen Umbenenung im Relationenschema erforderlich!

(meist systematisch durch "Vorschalten" des Relationsnamens: R.a)



Achtung!

Operation entspricht nicht exakt dem Produkt der Mengenalgebra !!
R x S für n- und m-stellige Relationen sollte eigentlich eine Menge von (2)-Tupeln sein – hier direkte Konkatenation der Operanden zu (n+m)-Tupel.

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra
Produktbildung

- Eigentlich ist das Produkt zweier Relationen <u>immer</u> eine zweistellige Relation, deren Elemente aber Tupel von Tupeln sind !

$$A \times B = \{ (a, b) \mid a \in A \text{ und } b \in B \}$$

- Wenn z.B. das Tupel (a,b) Element der zweistelligen Relation A und das Tupel (1,2,3) Element der dreistelligen Relation B ist, dann enthält – in der Mengenalgebra – das Produkt von A und B das Paar ((a,b) , (1,2,3)).

- In der Relationenalgebra ist die Produktbildung aber <u>anders</u> definiert als in der Mengenalgebra, denn hier werden die Tupelpaare der Operandenrelationen zusätzlich noch konkateniert, bevor sie ins Produkt aufgenommen werden:

$$\begin{array}{l} (a,b) \searrow \\ \qquad\qquad (a,b,1,2,3) \\ (1,2,3) \nearrow \end{array}$$

- In der RA ist also das Produkt einer n-stelligen und einer m-stelligen Relation eine (n+m)-stellige und nicht etwa eine 2-stellige Relation !

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra, Operatoren

Operatoren

σ
π
ρ
−
∪ / ∩
÷
x
$\bowtie_{\theta}$
$\bowtie$
a / a
⟕ / ⟖
⟗
▷

Auswahl
(Selektion)
Projektion
Umbenennung
Differenz
Vereinigung /
Durchschnitt Division
Kreuzprodukt (Kartesisches Produkt, Cross-Join)
Theta-Join
Natural Join
Left Semi-Join / Right Semi-Join
Left Outer-Join / Right Outer-Join Full Outer-Join
Left Anti-Join (Complement of Left Semi-Join)

Professoren

PERSNR	NAME	RANG	RAUM
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

Studenten

MATRNR	NAME	SEMESTER
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

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	2126

Vorlesungen

VORLNR	TITEL	SWS	GELESEN_VON
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

hören

MATRNR	VORLNR
25403	5022
26120	5001
27550	4052
27550	5001
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5001
29555	5022

voraussetzen

VORGÄNGER	NACHFOLGER
5001	5041
5001	5043
5001	5049
5041	5052
5041	5216
5043	5052
5052	5259

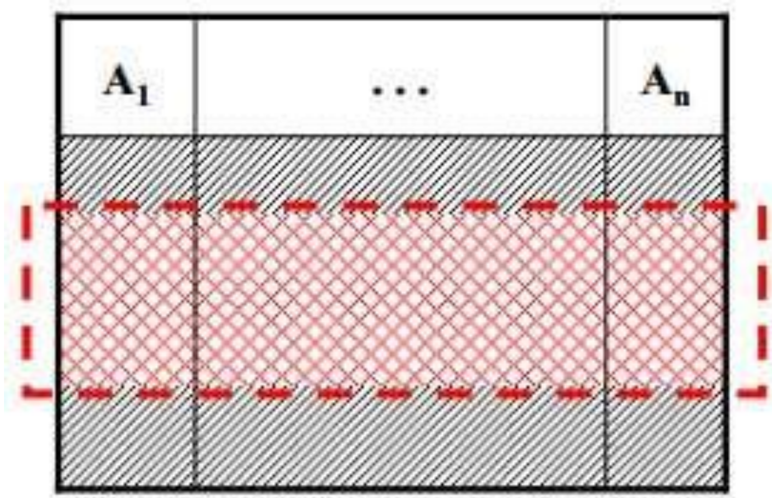
prüfen

MATRNR	VORLNR	PERSNR	NOTE
28106	5001	2126	1
25403	5041	2125	2
27550	4630	2137	2

Relationale Algebra, Operatoren

σ Auswahl (Selektion)

- Selektion (nicht mit „SELECT“ in SQL zu verwechseln!)
  - Darum werden wir diesen Operator in Zukunft „Auswahl“ nennen
  - Kann als Filter auf eine Relation betrachtet werden: Es werden die Tupel aus der Relation eliminiert, welche nicht der Bedingung genügen
- Allgemeine Anwendung:
  - $\sigma_{cond}(R)$
  - Cond kommt hier aus der Aussagenlogik: genutzt werden können Attribute, Konstanten, zum Datentyp passende Vergleichsoperatoren, Operatoren der booleschen Algebra und ggf. selbst definierte Funktionen enthalten
- Beispiel
  - $\sigma_{Alter>21}(\text{Studenten})$
  - $\sigma_{\neg Geschlecht=‚w‘ \land Alter>21}(\text{Studenten})$



Relationale Algebra, Operatoren

σ Auswahl (Selektion) :: BEISPIEL

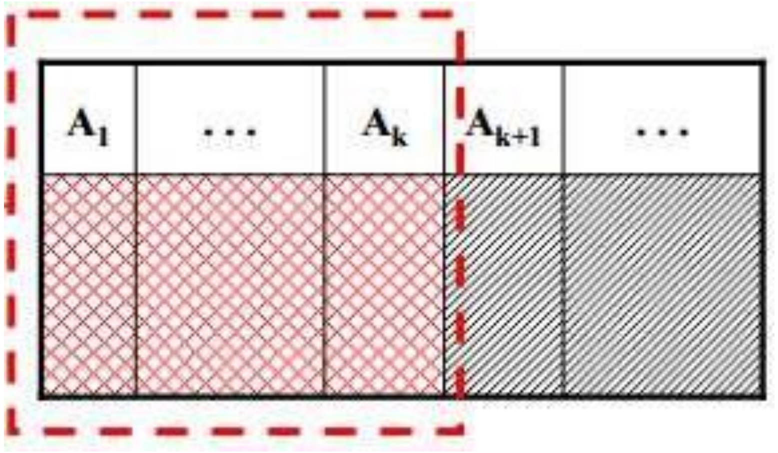
Studenten		
MatrNr	**Name**	**Semester**
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

$\sigma_{\text{Semester} > 10}$ (Studenten)		
MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12

Relationale Algebra, Operatoren

π Projektion

- Operator zum Ausblenden von Attributen durch explizite Angabe der gewünschten Spalten
- Die Kardinalität der Ergebnis-Relation ist immer kleiner oder gleich groß wie die Relation, auf die die Projektion angewendet wird
- Allgemeine Anwendung:
  - $\pi_{A_1,\ldots,A_k}(R)$
  - Eliminiert alle übrigen Attribute von R (alle außer $A_1,\ldots,A_k$)
  - Eliminiert ggf. Duplikattupel
- Beispiel
  - π_{Alter} (Studenten)
  - $\pi_{\text{Geschlecht, Alter}}$ (Studenten)



Relationale Algebra, Operatoren

π Projektion :: BEISPIEL

Professoren			
PersNr	Name	Rang	Raum
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

Π_{Rang}(Professoren)
Rang
C4
C3

Relationale Algebra, Operatoren

ρ Umbenennung

Operator zum Umbenennen von Attributen oder Relationen

Allgemeine Anwendung:

- $\rho_S(R)$ ➔ Umbenennung der Relation
- „S“ ist dann sozusagen ein Alias / eine temporäre Relation
- $\rho_{S(B_1, B_2, \dots B_n)}(R)$ ➔ Umbenennung der Relation und aller Attribute
- $\rho_{(B_1, B_2, \dots B_n)}(R)$ ➔ Umbenennung aller Attribute
- $\rho_{(A \leftarrow B_x, B \leftarrow B_y)}(R)$ ➔ Umbenennung zweier Attribute

Beispiel:

- $\rho_{Ergebnis}(\sigma_{Alter>21}(Studenten))$

→ Wir erinnern uns: Das Ergebnis jeder Operation ergibt wieder eine Relation (die auch 0 Tupel beinhalten kann)

Relationale Algebra, Operatoren

ρ Umbenennung :: BEISPIEL

Studenten		
MatrNr	**Name**	**Semester**
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

$\Pi_{\text{Name}}(\sigma_{\text{Semester} > 10}(\text{Studenten}))$
Name
Xenokrates
Jonas

$$\rho_{\text{Ergebnis(Langzeitstudent} \leftarrow \text{Name)}}(\Pi_{\text{Name}}(\sigma_{\text{Semester} > 10}(\text{Studenten})))$$

Ergebnis
Langzeitstudent
Xenokrates
Jonas

Relationale Algebra, Operatoren

Mengenoperationen ∪ ∩ –

- Vereinigung R ∪ S ergibt eine Relation mit allen Tupeln aus R und S, diese sind in der Ergebnisrelation paarweise verschiedenen
  - kommutativ und assoziativ

- Schnittmenge R ∩ S ergibt eine Relation mit allen Tupeln, die sowohl in R als auch in S vorhanden sind
  - kommutativ und assoziativ

- Differenz R – S ergibt eine Relation mit allen Tupeln, die in R, jedoch nicht in S vorhanden sind

Relationale Algebra, Operatoren

Mengenoperationen ∪ ∩ –   :: BEISPIEL

Student	
Vorname	**Nachname**
Roy	Redmond
Farrah	Barnard
Phebe	Blackburn
Jay-Jay	Whitley
Rui	Rodriquez
Ivo	Coles
Darcie	Mercer
Mi	Chan

Lehrer	
Vorname	**Nachname**
Alena	Riley
Anna	Brock
Kendal	Needham
Farrah	Barnard
Phebe	Blackburn

Student ⋃ Lehrer = Lehrer ⋃ Student	
Roy	Redmond
Farrah	Barnard
Phebe	Blackburn
Jay-Jay	Whitley
Rui	Rodriquez
Ivo	Coles
Darcie	Mercer
Mi	Chan
Alena	Riley
Anna	Brock
Kendal	Needham

Student – Lehrer Student \ Lehrer	
Roy	Redmond
Jay-Jay	Whitley
Rui	Rodriquez
Ivo	Coles
Darcie	Mercer
Mi	Chan

Lehrer – Student Lehrer \ Student	
Alena	Riley
Anna	Brock
Kendal	Needham

Student ⋂ Lehrer = Lehrer ⋂ Student	
Farrah	Barnard
Phebe	Blackburn

Relationale Algebra, Operatoren

÷ Division

- Dividiert man eine Relation R mit zwei Attributen A und B durch eine Relation S mit dem Attribut A, so erhält man eine Relation T mit dem Attribut B
  - In T.B sind alle Werte enthalten, die mit **<u>jedem</u>** Wert aus in S.A ein Tupel in der Relation R bilden

R	
M	V
m_1	v_1
m_1	v_2
m_1	v_3
m_2	v_2
m_2	v_3

÷

S
V
v_1
v_2

=

R ÷ S
M
m_1

- Allgemeine Anwendung:
  - T ← R ÷ S

- Beispiel

Prüfung	
Student	**Fach**
Horst	Mathe I
Horst	Datenbanken
Marie	Mathe I
Marie	Datenbanken
Marie	Informatik Grundlagen
Peter	Englisch
Mae	Mathe I
Mae	Informatik Grundlagen

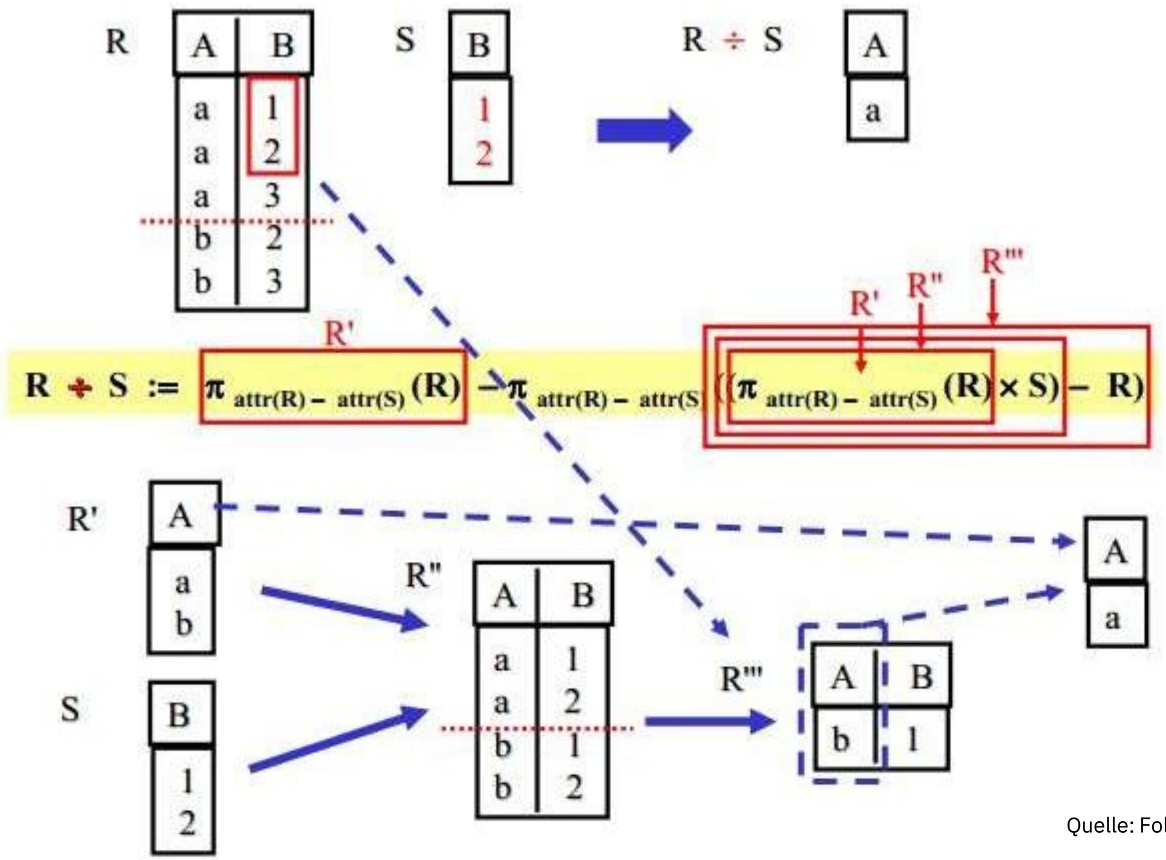
Pflichtprüfungen
Fach
Mathe I
Informatik Grundlagen

Prüfung ÷ Pflichtprüfungen
Student
Marie
Mae

Relationale Algebra, Operatoren
÷ Division

- Kann aus anderen Operatoren abgeleitet werden:

$$R \div S := \pi_{R'}(R) - \pi_{R'}((\pi_{R'}(R) \times S) - R)$$



Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra, Operatoren

Beispiel: Lösen Sie folgende Fragen mit relationaler Algebra!

Welchen Rang haben die Professoren in den 200er-Räumen?

!

$$\pi_{Rang}(\sigma_{\text{Raum} \geq 200 \wedge \text{Raum} < 300}(\text{Professoren}))$$

Relationale Algebra, Operatoren

Übung: Lösen Sie folgende Fragen mit relationaler Algebra!

In welchen Räumen sind Sokrates, Russel oder Kopernikus? (*)

Wie sind die Titel der Vorlesungen mit 2 SWS?

Welche Vorlesungen (Nr) hat 28106 besucht, die 29120 nicht besucht hat?

Welche Matrikelnummern haben bereits die Vorlesungen 5041 und 5052 gehört?

Bearbeitungszeit: 15 Minuten

(*) In natürlicher Sprache „und“, in der Sprache der Logik „oder“

Relationale Algebra, Operatoren

Lösungen: 1 und 2

? In welchen Räumen sind Sokrates, Russel oder Kopernikus?

$$\pi_{Raum}(\sigma_{Name='Sokrates' \vee Name='Kopernikus' \vee Name='Russel'}(Professoren))$$

π_{Raum}

|

$\sigma_{Name='Sokrates' \vee Name='Kopernikus' \vee Name='Russel'}$

|

Professoren

? Wie sind die Titel der Vorlesungen mit 2 SWS?

$$\pi_{Titel}(\sigma_{SWS=2}(Vorlesungen))$$

π_{Titel}

|

$\sigma_{SWS=2}$

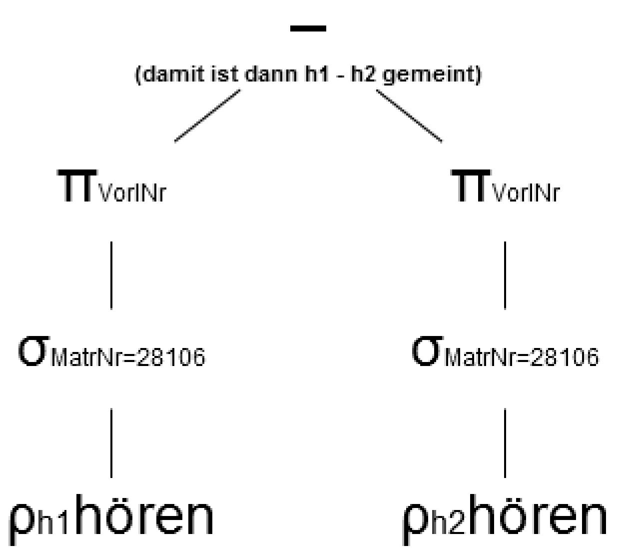
|

Vorlesungen

Relationale Algebra, Operatoren

Lösungen: 3

Welche Vorlesungen (Nr) hat 28106 besucht, die 29120 nicht besucht hat?



$$\pi_{VorlNr}(\sigma_{MatrNr=281062}(\rho_{h1}\text{hören})) - \pi_{VorlNr}(\sigma_{MatrNr=281062}(\rho_{h1}\text{hören}))$$

Relationale Algebra, Operatoren

Lösungen: 4

Welche Matrikelnummern haben bereits die Vorlesungen 5041 und 5052 gehört?

$\div$

(Relation mit mehr Attributen durch Relation mit weniger Attributen)

hören

π_{VorlNr}

$\sigma_{VorlNr=5041 \vee VorlNr=5052}$

Vorlesungen

$$\text{hören} \div (\pi_{VorlNr}\sigma_{VorlNr=5041 \vee VorlNr=5052}(\text{Vorlesungen}))$$

Professoren

PERSNR	NAME	RANG	RAUM
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

Studenten

MATRNR	NAME	SEMESTER
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

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	2126

Vorlesungen

VORLNR	TITEL	SWS	GELESEN_VON
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

hören

MATRNR	VORLNR
25403	5022
26120	5001
27550	4052
27550	5001
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5001
29555	5022

voraussetzen

VORGÄNGER	NACHFOLGER
5001	5041
5001	5043
5001	5049
5041	5052
5041	5216
5043	5052
5052	5259

prüfen

MATRNR	VORLNR	PERSNR	NOTE
28106	5001	2126	1
25403	5041	2125	2
27550	4630	2137	2

Relationale Algebra, Operatoren II

Joins

- x Kreuzprodukt (Kartesisches Produkt, Cross-Join) Theta-
- $\bowtie_{\theta}$, ⟕$_{\theta}$, … Join
- ⋈ Natural Join („normaler“)
- a▷a Left Semi-Join / Right Semi-Join Left
- ⟕ / ⟖ Outer-Join / Right Outer-Join Full Outer-
- ⟗ Join
- D Left Anti-Join (Complement of Left Semi-Join)

- Eine Relation R mit r-Tupeln & eine Rel. S mit s-Tupeln werden mit einer Join- Operation zu einer Rel. mit t-Tupeln kombiniert, wobei t >= min(r,s), t <= r+s

- Im Alltag mit „SQL“ haben vor allem die rot gekennzeichneten Joins Relevanz
  - Persönliche Meinung des Dozenten…

- Alle Joins, die keine Outer-Joins sind, sind Inner-Joins

- Alle nicht-Theta-Joins (bis auf X), vergleichen gleichnamige Attribute und werden als Natural Joins bezeichnet

- Achtung bei Reihenfolge der Join-Operationen!

Relationale Algebra, Operatoren II

Joins

×	Kreuzprodukt (Kartesisches Produkt, Cross-Join)
$\bowtie_{\theta}$, ⟕$_{\theta}$, …	Theta-Join
$\bowtie$	Natural Join
a▷a	Left semi-join
⟕ / ⟖	Left / right outer join
⟗	Full outer join
▷	Left anti-join

- Eine Relation R mit r-Tupeln & eine Rel. S mit s-Tupeln werden mit einer Join-Operation zu einer Rel. mit t-Tupeln kombiniert, wobei t >= min(r,s), t <= r+s
- Im Alltag mit „SQL“ haben vor allem die rot gekennzeichneten Joins Relevanz
- Alle Joins, die keine Outer-Joins sind, sind Inner-Joins
- Alle nicht-Theta-Joins (bis auf X), vergleichen gleichnamige Attribute und werden als Natural Joins bezeichnet
- Achtung bei Reihenfolge der Join-Operationen!

Relationale Algebra, Operatoren II

× Kreuzprodukt

- Ein Kreuzprodukt aus R und S enthält alle |R|*|S| möglichen Tupel-Kombinationen beider Relationen. Das Schema der Ergebnisrelation attr(RxS) entspricht attr(R) ∪ attr(S), wobei allen Attributnamen der Relationsname bei Bedarf vorangestellt wird (wichtig bei identischen Namen)
- Allgemeine Anwendung:
  - $R \times S$
- Beispiel
  - Studenten × schreibt_prüfung_in × Studienfach

Relationale Algebra, Operatoren II

× Kreuzprodukt :: BEISPIEL

Professoren			
PersNr	**Name**	**Rang**	**Raum**
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

hören	
MatrNr	**VorlNr**
26120	5001
27550	5001
27550	4052
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5022
25403	5022
29555	5001

Professoren × hören

Professoren				hören	
PersNr	**Name**	**Rang**	**Raum**	**MatrNr**	**VorlNr**
2125	Sokrates	C4	226	26120	5001
...	...	...	...	...	...
2125	Sokrates	C4	226	29555	5022
...	...	...	...	...	...
2137	Kant	C4	7	29555	5001

Relationale Algebra, Operatoren II

$\bowtie_\theta$ Theta-Join

- Das Kreuzprodukt hat den Nachteil, dass sehr viele „unnötige" Tupel entstehen, obwohl eigentlich nur zueinander passende interessieren
  - Darum erfolgt in einem nächsten Schritt eine Selektion
- Ein Theta-Join aus R und S ist die verkürzte Schreibweise aus einem Kreuzprodukt mit anschließender Auswahl, die den „Joinpartner" spezifiziert
  - $R \bowtie_\theta S = \sigma_\theta (R \times S)$
- Allgemeine Anwendung:
  - $R \bowtie_\theta S$

$R \bowtie_\Theta S$ mit $\Theta = (A \leq C \wedge S.B > 0)$

R

A	B
1	a
3	b
2	a

S

B	C
5	2
0	1
1	1

A	R.B	S.B	C
1	a	5	2
1	a	1	1
2	a	5	2

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra, Operatoren II

⋈ Natural Join

- Der Natural Join ist wie auch der Theta-Join eine ableitbare Operation der RA, denn derselbe Effekt ließe sich auch durch Kombination von Projektion, Selektion und Kreuzprodukt erreichen
- Im Gegensatz zu einem Theta-Join werden hier jedoch automatisch gleichnamige Attribute „gematched" (R.a=S.a) und doppelte Attributnamen in der Ergebnisrelation ausgeblendet
- Allgemeine Anwendung:
  – R ⋈ S

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

attr(R) – attr(S)			attr(R) ∩ attr(S)			attr(S) – attr(R)		
A_1	...	A_m	B_1	...	B_k	C_1	...	C_n
				gemeinsame Attribute				

Hierbei bezeichnet attr(R) die Attributmenge der Relation R.

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra, Operatoren II

⋈ Natural Join :: BEISPIEL

Studenten		
MatrNr	**Name**	**Semester**
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
...	...	...
29555	Feuerbach	2

hören	
MatrNr	**VorlNr**
26120	5001
25403	5022
28106	4052
...	...
29555	5001

Vorlesungen			
VorlNr	**Titel**	**SWS**	**gelesen Von**
5001	Grundzüge	4	2137
5041	Ethik	4	2125
5022	Glaube und Wissen	2	2134
...	...	...	...
4630	Die 3 Kritiken	4	2137

(Studenten ⋈ hören) ⋈ Vorlesungen

(Studenten ⋈ hören) ⋈ Vorlesungen						
MatrNr	**Name**	**Semester**	**VorlNr**	**Titel**	**SWS**	**gelesenVon**
26120	Fichte	10	5001	Grundzüge	4	2137
25403	Jonas	12	5022	Glaube und Wissen	2	2134
28106	Carnap	3	4052	Wissenschftstheorie	3	2126
...	...	...	...	...	...	...

Relationale Algebra, Operatoren II

⋉ Left Semi-Join / ⋊ Right Semi-Join

- Semi-Join (lat. "semi": halb):
  - Teilrelationenbildung eines der beiden Join-Operanden
  - zwei Varianten: linker und rechter Semi-Join
  - Nur diejenigen Tupel des ausgewählten Joinoperanden werden ausgewählt, die "einen Joinpartner" besitzen.
  - symbolische Notation: $R \ltimes S$ (linker Semi-Join, rechter analog)

- Beispiel: (natürlicher) linker Semi-Join $R \ltimes S$

R

A	B
1	2
3	1
2	5

S

B	C
5	2
2	1
2	1

$R \ltimes S$

A	B
1	2
2	5

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra, Operatoren II

▷ Left Anti-Join (Complement of Left Semi-Join)

- Im Gegenteil zum Left Semi-Join werden hier nur die Tupel angezeigt, die keinen Join-Partner besitzen (wurde von Codd nicht eingeführt und ist deshalb in vielen Bücher nicht einmal erwähnt...).
  Wichtig für Fragen wie „welche Studenten haben noch keine Vorlesung besucht“
- Right Anti-Join analog
- Allgemeine Anwendung:
  - R $\overline{\ltimes}$ S
  - R S (wenn als Komplement des Semi-Join verstanden)

Beispiel:

R

A	B
1	2
3	1
2	5

S

B	C
5	2
2	1
2	1

R $\overline{\ltimes}$ S

A	B
3	1

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra, Operatoren II

⟕ Left Outer Join
⟖ Right Outer Join
⟗ Full Outer Join

- Ein Outer Join erweitert den schon bekannten Theta-Join so, dass zusätzlich zu allen Tupeln, die die Join-Bedingung erfüllen, auch jene enthalten sind, die sie nicht erfüllen. Die Attribute der jeweils anderen Input-Relation werden dann mit NULL-Werten aufgefüllt
- Das Wort „Left“, „Right“ oder „Full“ gibt an, aus welcher Relation auf jeden Fall alle Werte enthalten sein müssen
  - Der Full Outer-Join ist die Vereinigungsmenge aus Left- und Right-Join
- Allgemeine Anwendung:
  - R ⟕ S
- Wozu? Es gehen keine Tupel ohne entsprechenden Join-Partner „verloren“

Relationale Algebra, Operatoren II

⟕ / ⟖ Left / Right Outer Join :: BEISPIEL

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

⟕

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat				
A	B	C	D	E
a_1	b_1	c_1	d_1	e_1
a_2	b_2	c_2	-	-

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

⟖

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat				
A	B	C	D	E
a_1	b_1	c_1	d_1	e_1
-	-	c_3	d_2	e_2

Relationale Algebra, Operatoren II

⟗ Full Outer-Join :: BEISPIEL

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
		c3

⟗

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat				
A	B	C	D	E
a_1	b_1	c_1	d_1	e_1
a_2	b_2	c_2	-	-
-	-	c_3	d_2	e_2

?

Technisch „korrekter" wäre folgende Darstellung rechts mit vorangestellten Relationennamen, wie wir sie vom Theta-Join kennen. Warum haben wir bei allen NATURAL OUTER JOINS einen Informationsverlust, bei THETA OUTER JOINS jedoch nicht?

Resultat					
L.A	L.B	L.C	R.C	R.D	R.E
a_1	b_1	c_1	c_1	d_1	e_1
a_2	b_2	c_2	-	-	-
-	-	c3	c_3	d_2	e_2

Relationale Algebra, Operatoren II

Outer-Joins auch mit Theta verwendbar :: BEISPIEL

$$R \underset{\Theta}{⟗} S \text{ mit } \Theta = (R.A \leq S.C \wedge S.B > 0)$$

R

A	B
1	a
3	b
2	a

S

B	C
5	2
0	1
1	1

A	R.B	S.B	C
1	a	5	2
1	a	1	1
2	a	5	2
3	b	Null	Null
Null	Null	0	1

- Vorweg: So wird das auch in SQL gemacht… NATURAL OUTER JOINS werden Sie in der Praxis nie sehen

Relationale Algebra, Operatoren II
Joins

- Falls das Theta bei einem Join auf Gleichheit (welcher Art auch immer) beruht (R.a=S.a $\wedge$ R.b=S.b), spricht man von einem Equi-Join („equals“)
- Falls kein Theta angegeben wird, spricht man von einem Natural Join
- Alle Joins, die keine Outer-Joins sind, sind Inner-Joins (bei einem Outer-Join werden NULL-Werte hinzugefügt)
- Doppelte Attributnamen werden in der Regel ausgeblendet (außer beim Kreuzprodukt)
- Theta-Joins und Kreuzprodukt sind kommutativ und assoziativ
  - Outer-Joins, Semi-Joins und Anti-Joins sind nicht kommutativ!
- Das Kreuzprodukt wird in der Praxis ebenfalls als Join (Cross-Join) bezeichnet. So kann man das Kreuzprodukt auch als speziellen Theta-Join ansehen.

?

Wie kann man ein Kreuzprodukt als „Theta-Join“ schreiben?

Relationale Algebra
logische Optimierung

- Für die „praktische" Auswertung ist nicht jeder äquivalente komplexe algebraische Ausdruck „gleich teuer"
- In der RA gibt es einige Äquivalenzen, die man ausnutzen kann
  - Unnötige Projektionen löschen, Projektionen zu früh wie möglich anwenden, um „Datenmüll" zu vermeiden
  - Redundante Selektionen löschen / zusammenfassen
  - Für Join-Operationen möglichst kleine Tabellen oder möglichst selektive conditions verwenden (viele Operatoren assoziativ: Reihenfolge dort egal)
- Beispiel: In welchen Semestern sind Sokrates' Studenten?
  - Wir verwenden hier der Einfachheit halber nur die ersten Buchstaben der Relationen:
    - S ← Studenten
    - H ← hören
    - V ← Vorlesungen
    - P ← Professoren

Relational Algebra

logische Optimierung, Beispiel

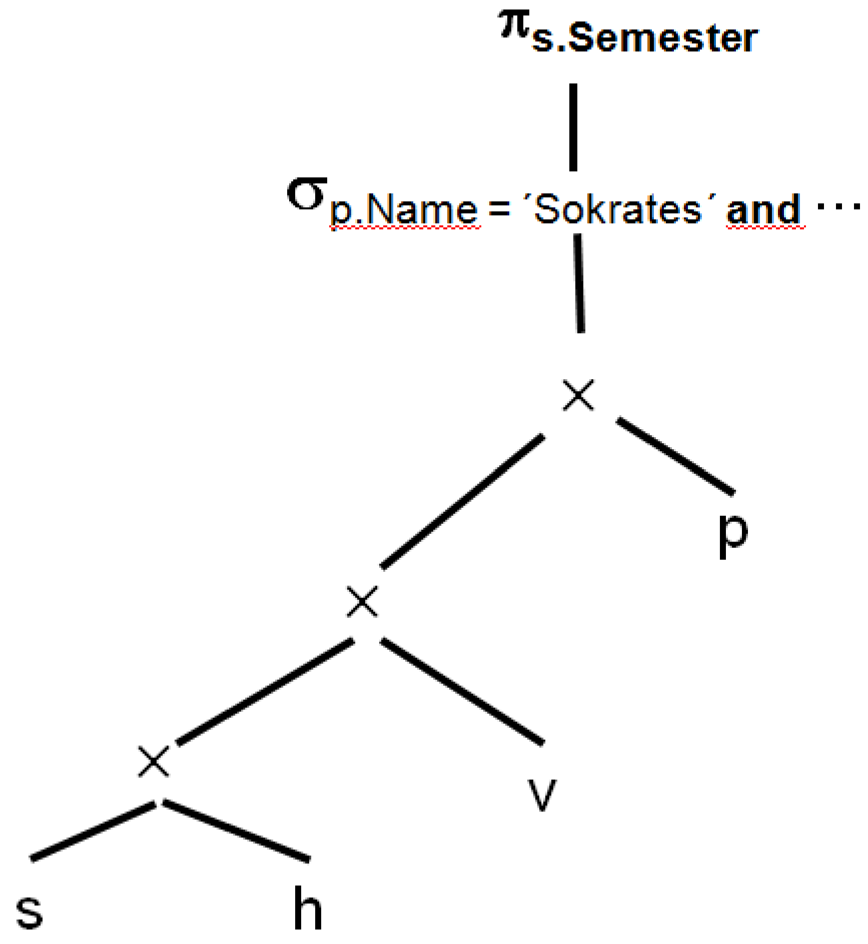
- In welchen Semestern sind Sokrates‘ Studenten?

$$\Pi_{\text{Semester}} \Big(\sigma_{\text{p.Name} = \text{´Sokrates´} \wedge \text{v.gelesenVon} = \text{p.PersNr} \wedge \text{v.VorlNr} = \text{h.VorlNr} \wedge \text{h.MatrNr} = \text{s.MatrNr}} (p \times v \times h \times s) \Big)$$

- Alternativ kann man das auch als Baum darstellen, bei komplexen Abfragen erhöht das die Übersichtlichkeit!!!

Relationale Algebra

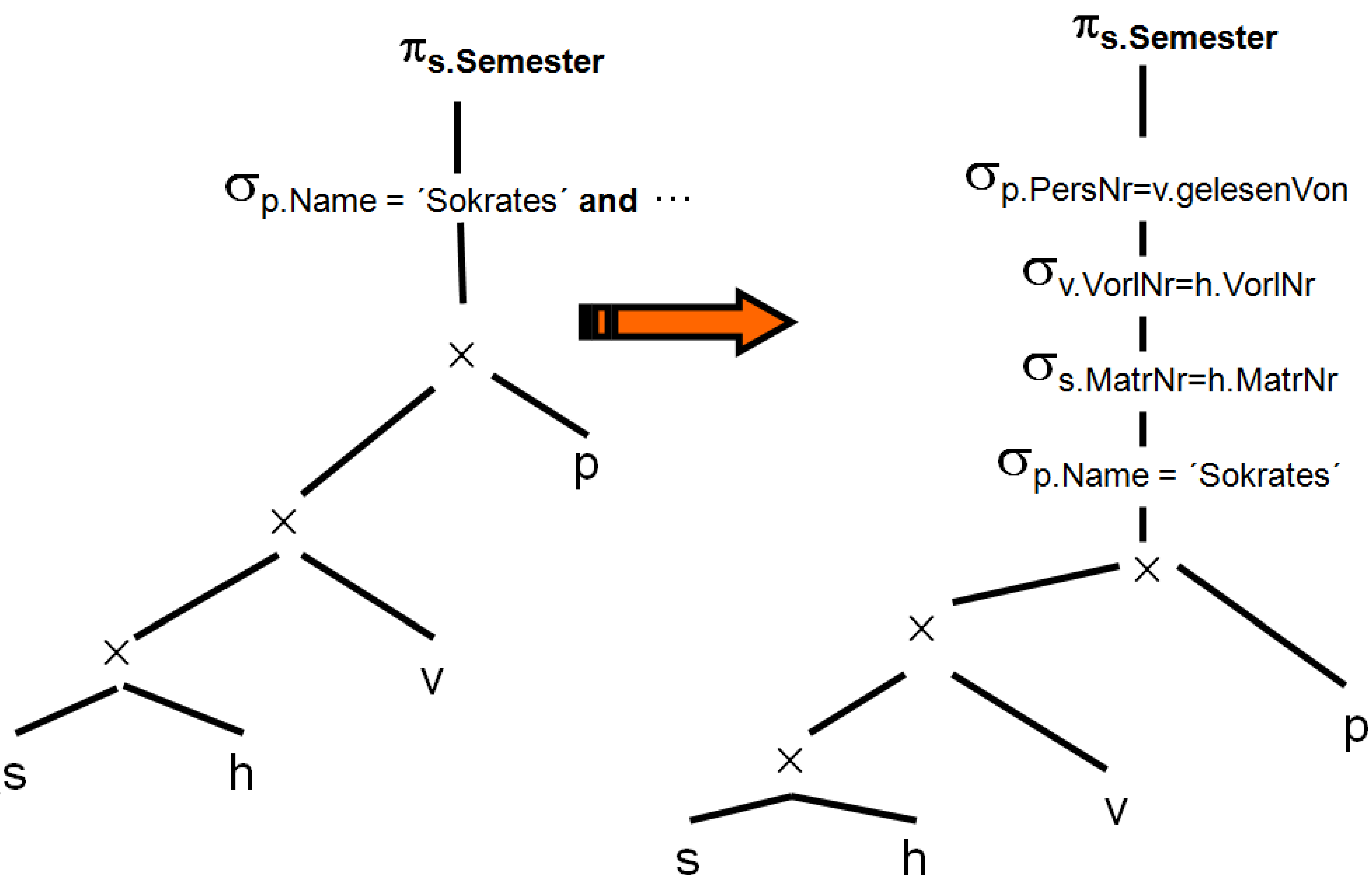
logische Optimierung, Beispiel



Relationale Algebra

logische Optimierung, Beispiel

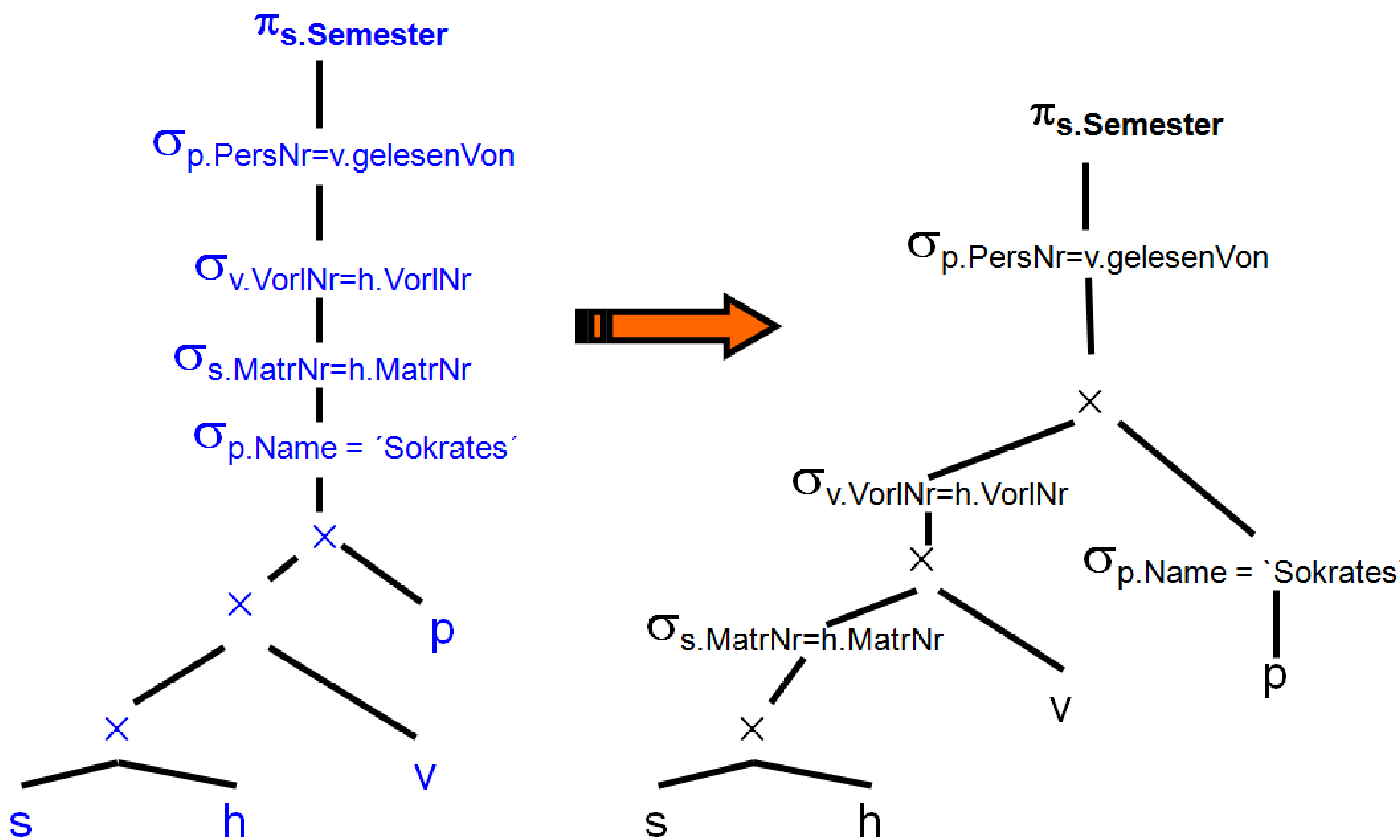
Aufspalten der Selektionsprädikate



Relationale Algebra

logische Optimierung, Beispiel

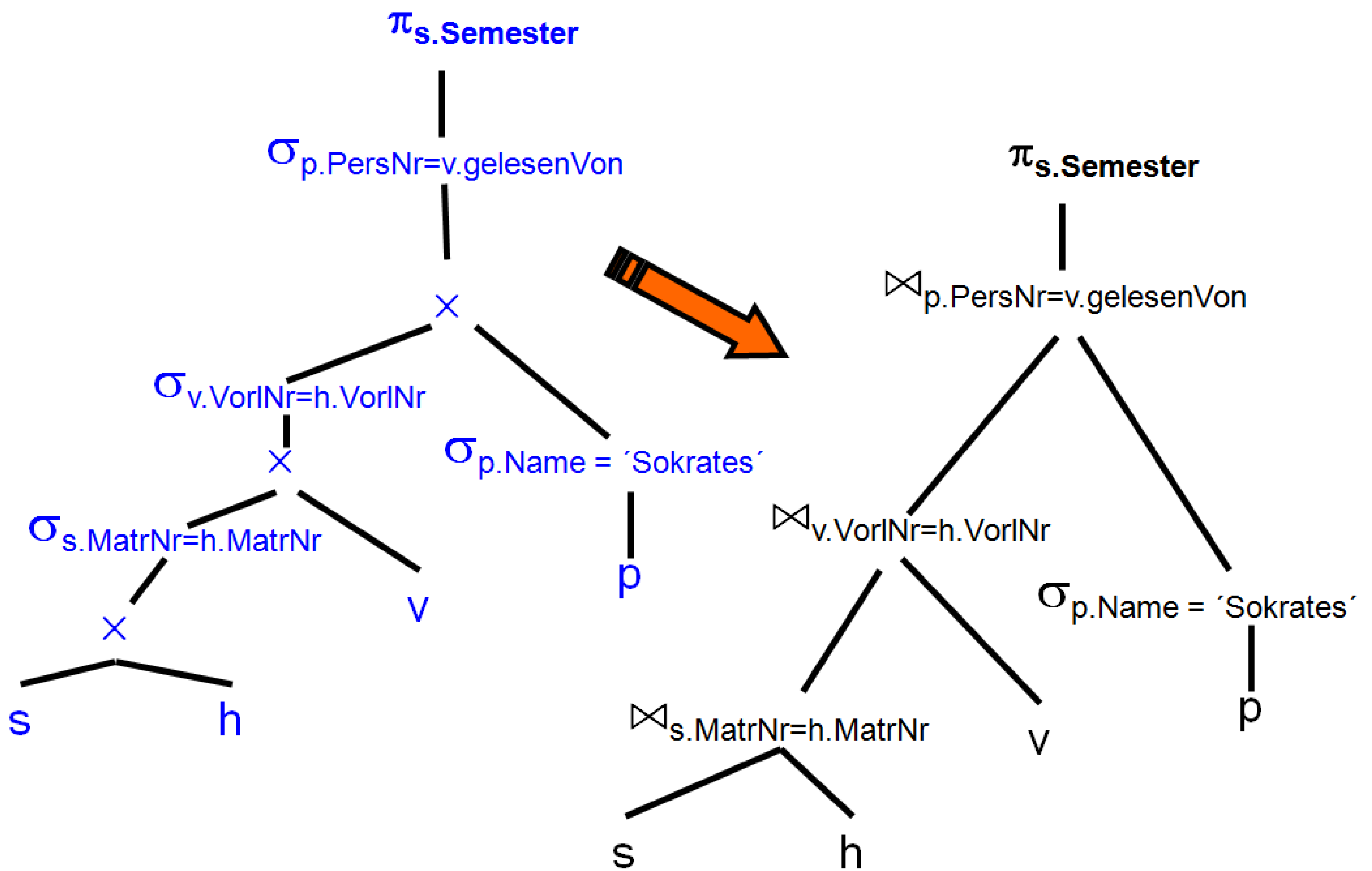
Verschieben der Selektionsprädikate „Pushing Selections“



Relationale Algebra

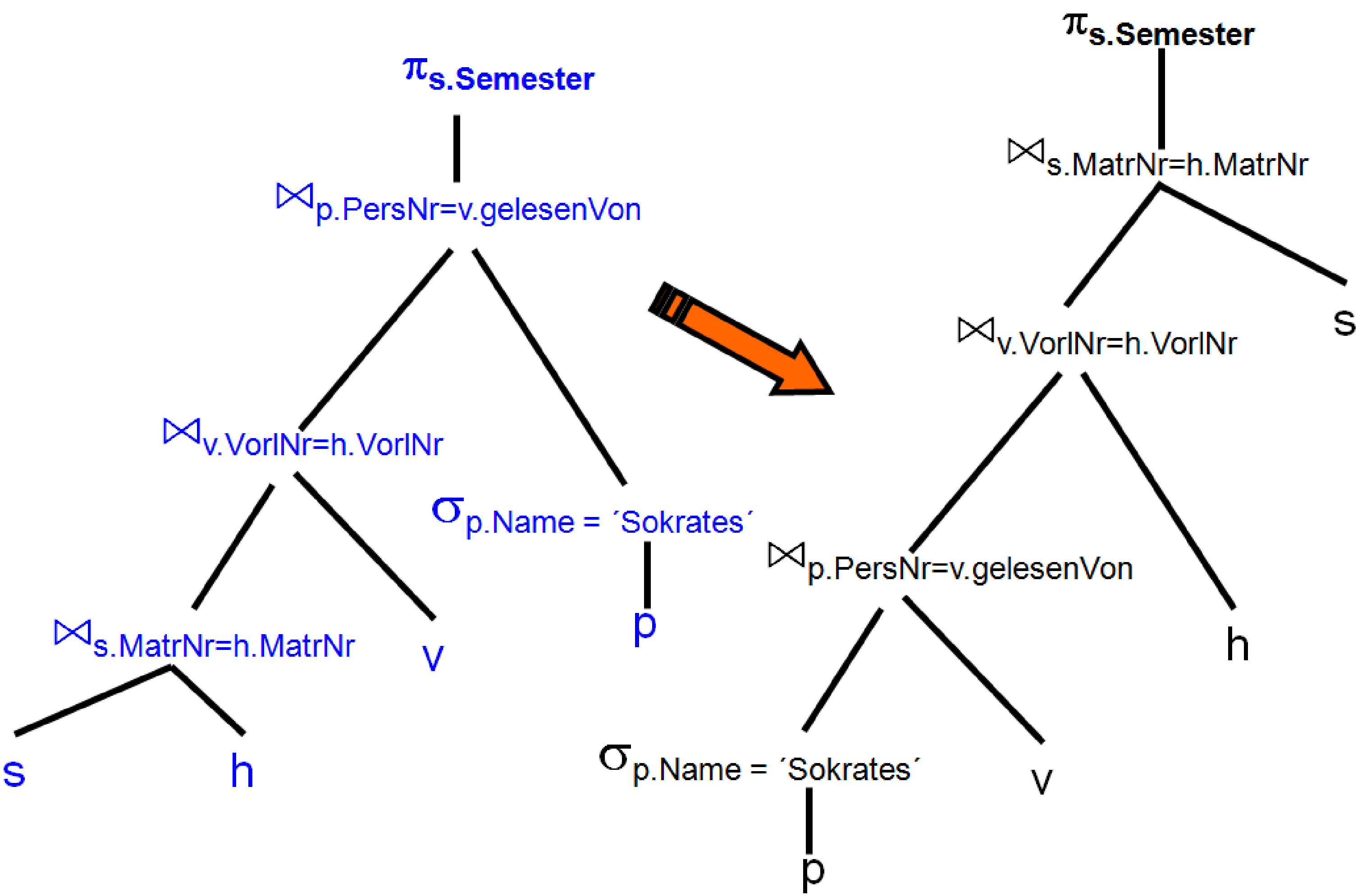
logische Optimierung, Beispiel

Zusammenfassung von Selektionen und Kreuzprodukten zu Joins



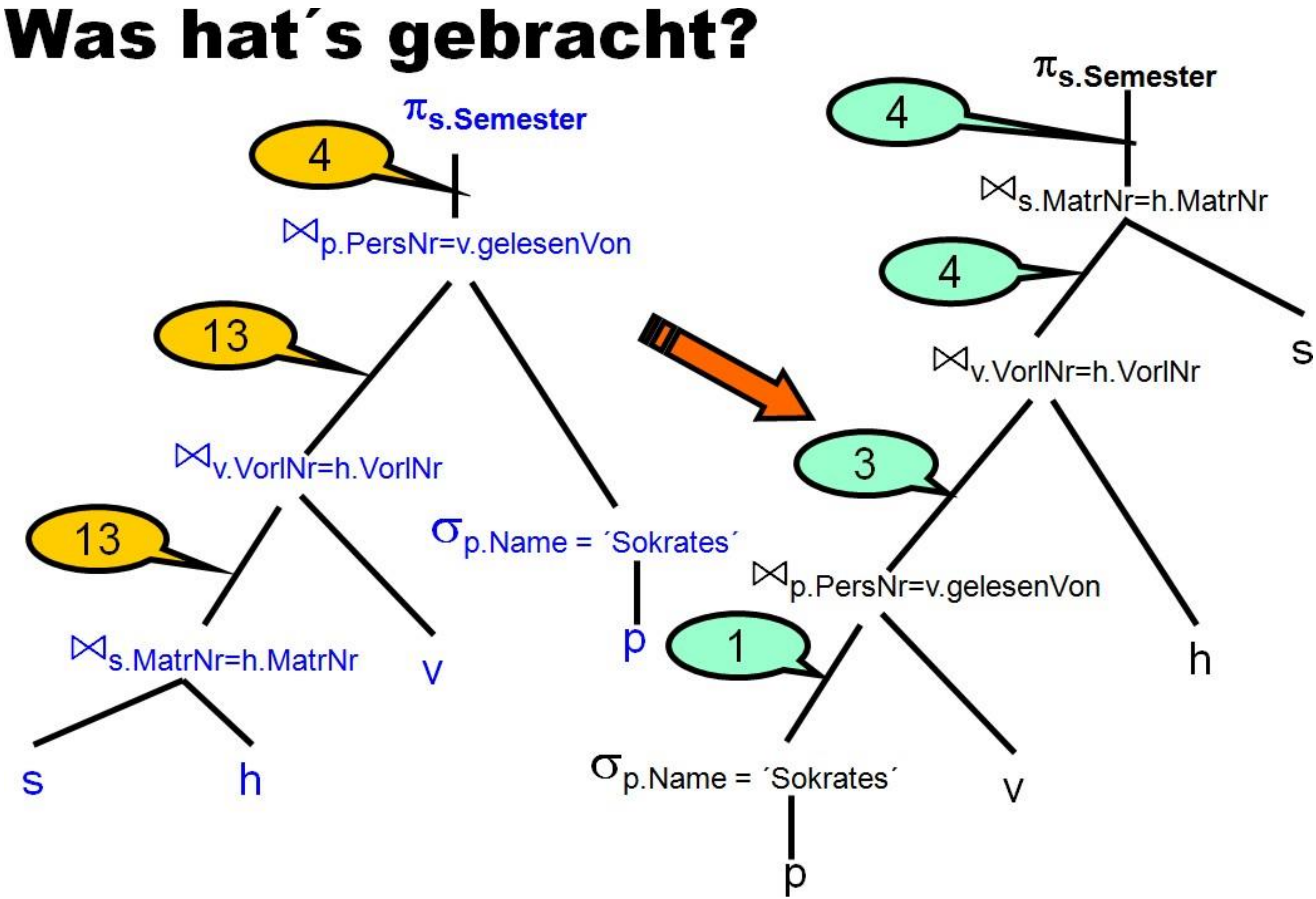
Relationale Algebra
logische Optimierung, Beispiel

Optimierung der Joinreihenfolge
Kommutativität und Assoziativität ausnutzen



Relationale Algebra

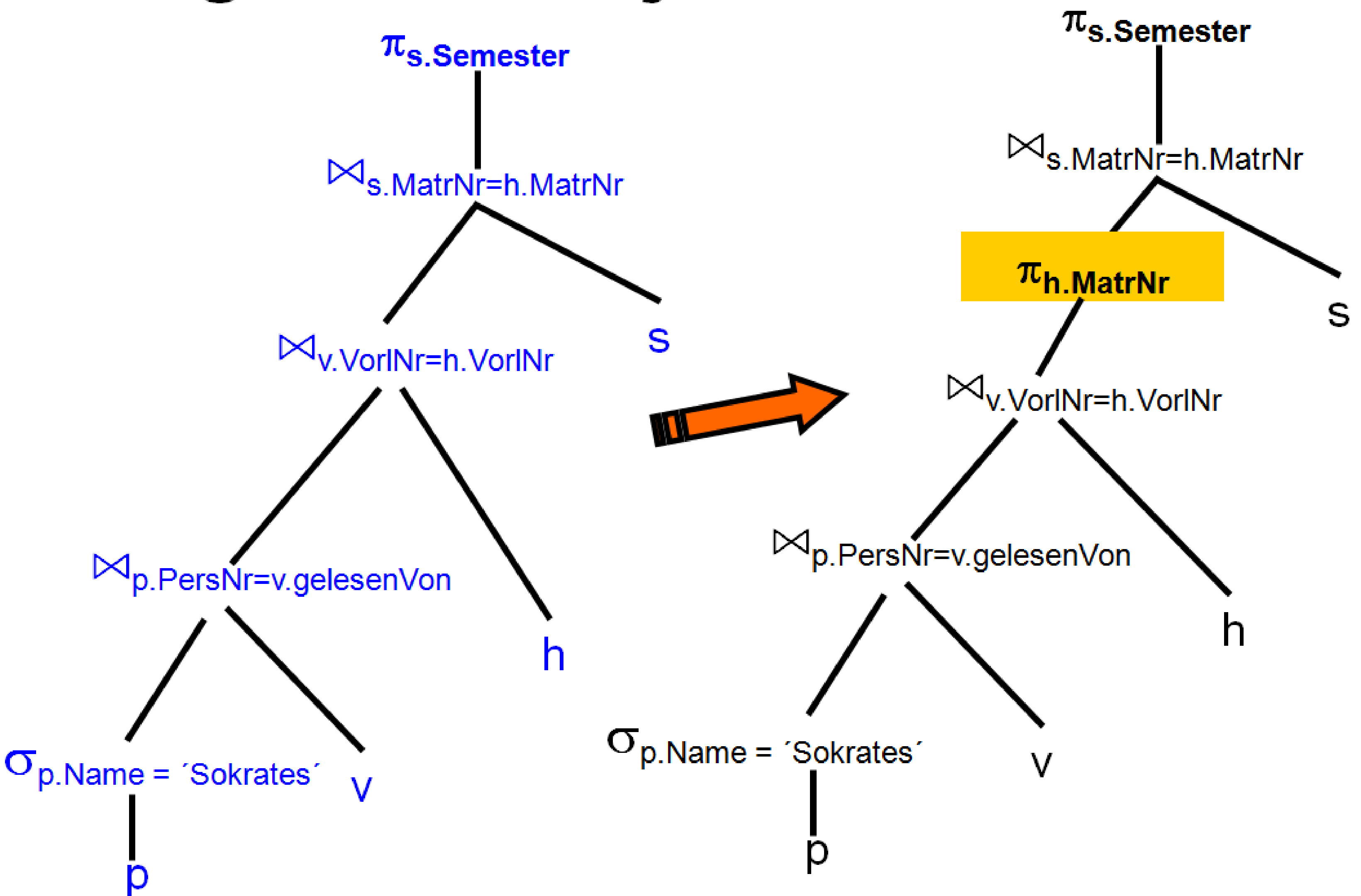
logische Optimierung, Beispiel



Relationale Algebra

logische Optimierung, Beispiel

Einfügen von Projektionen



Relationale Algebra
weitere Optimierungen

- Physische Optimierung
  - Indexstrukturen, Suchbäume
  - Daten vorsortiert auf der Festplatte ablegen
- Kostenmodelle
  - Heuristiken / Statistiken / Verteilungen

Relationale Algebra
weitere Optimierungen

- Optimierung der Anwendungen
  - keine sinnlosen Anfragen
  - nicht unbedingt selbst den Algorithmus definieren, sondern der DB die Arbeit „überlassen“ (muss nicht immer das beste Performance-Ergebnis bringen). Folgende häufig in der Webentwicklung anzutreffende „Technik“ zwingt selbst high-end DB-Server in die Knie:

```
$result = mysql_query("some-query");
while($cur = mysql_fetch_object($result)) {
    $secondResult = mysql_query("some-query that uses $cur->id");
    while($result2 = mysql_fetch_object($secondResult)) {
        echo $result2->name . "<br>";
    }
} // AUTSCH! Was ist hier schief gelaufen?!
```

Relationale Algebra, Operatoren II

Übung: Lösen Sie folgende Fragen mit relationaler Algebra!

Tipp: Kürzen Sie jede Relation mit den Anfangsbuchstaben ab.

? **Zeige die Assistenten (Namen) von Sokrates, deren Fachgebiet mit „S“ beginnt!**

? **Zeige alle Studentennamen, die weniger Semester haben als Schopenhauer!**

? **Zeige, welche Noten Professoren mit dem Rang C4 an die Studenten „Fichte“ bzw „Jonas“ vergeben haben**

? **Welche Studenten (Namen) haben bereits Ethik und Wissenschaftstheorie gehört?**

Bearbeitungszeit: 20 Minuten

Professoren

PERSNR	NAME	RANG	RAUM
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

Studenten

MATRNR	NAME	SEMESTER
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

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	2126

Vorlesungen

VORLNR	TITEL	SWS	GELESEN_VON
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

hören

MATRNR	VORLNR
25403	5022
26120	5001
27550	4052
27550	5001
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5001
29555	5022

voraussetzen

VORGÄNGER	NACHFOLGER
5001	5041
5001	5043
5001	5049
5041	5052
5041	5216
5043	5052
5052	5259

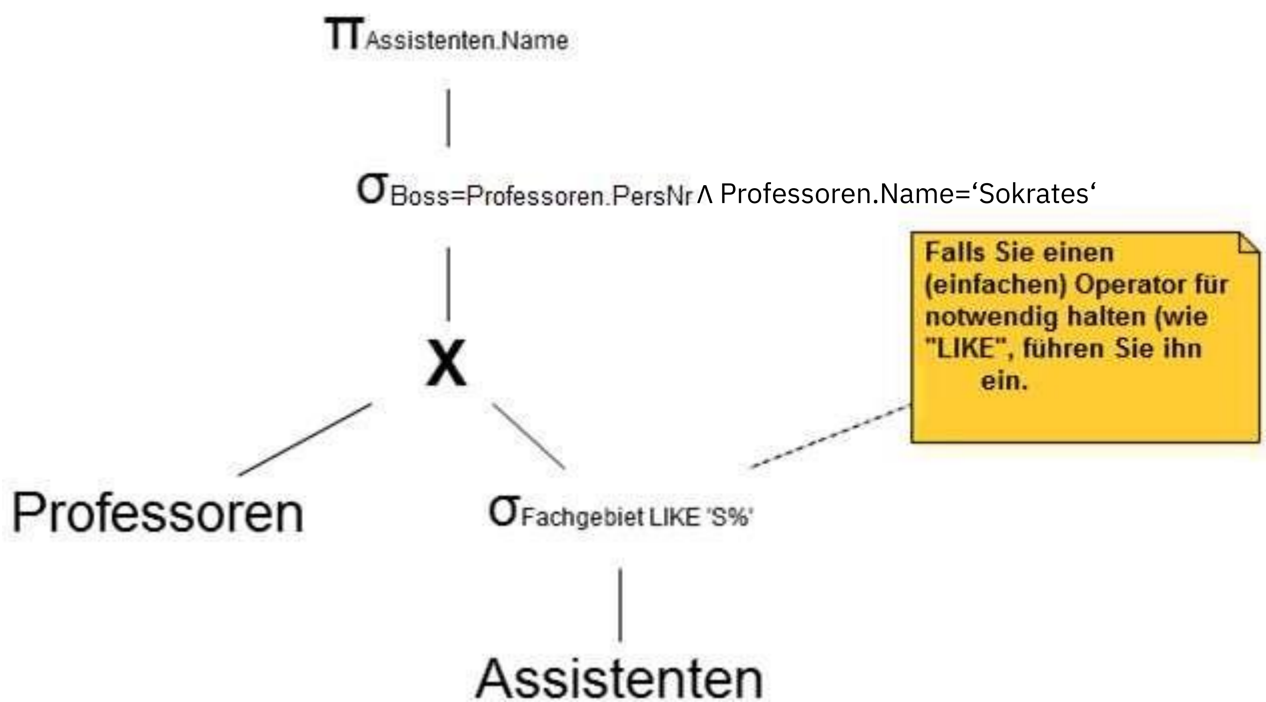
prüfen

MATRNR	VORLNR	PERSNR	NOTE
28106	5001	2126	1
25403	5041	2125	2
27550	4630	2137	2

Relationale Algebra, Operatoren II

Lösungen: 1 (nur noch als Operatorbäume)

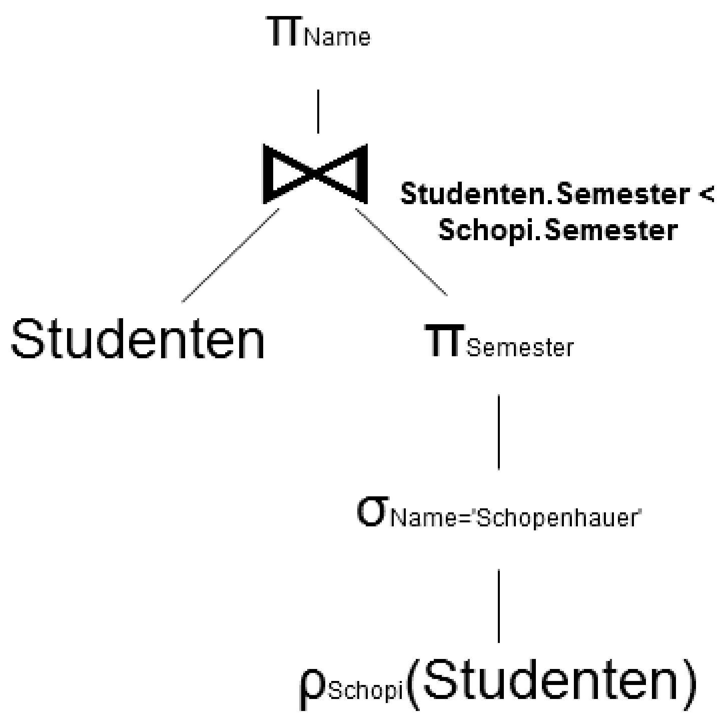
Zeige die Assistenten (Namen) von Sokrates, deren Fachgebiet mit „S“ beginnt!



Relationale Algebra, Operatoren II

Lösungen: 2 (nur noch als Operatorbäume)

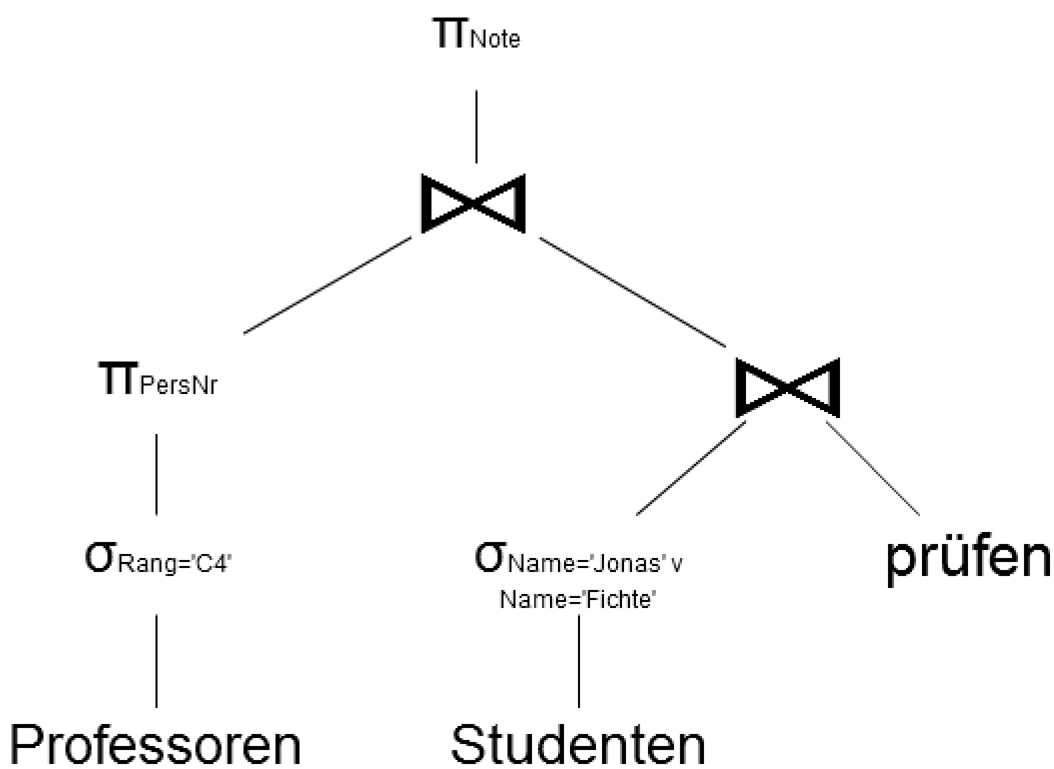
Zeige alle Studentennamen, die weniger Semester haben als Schopenhauer!



Relationale Algebra, Operatoren II

Lösungen: 3 (nur noch als Operatorbäume)

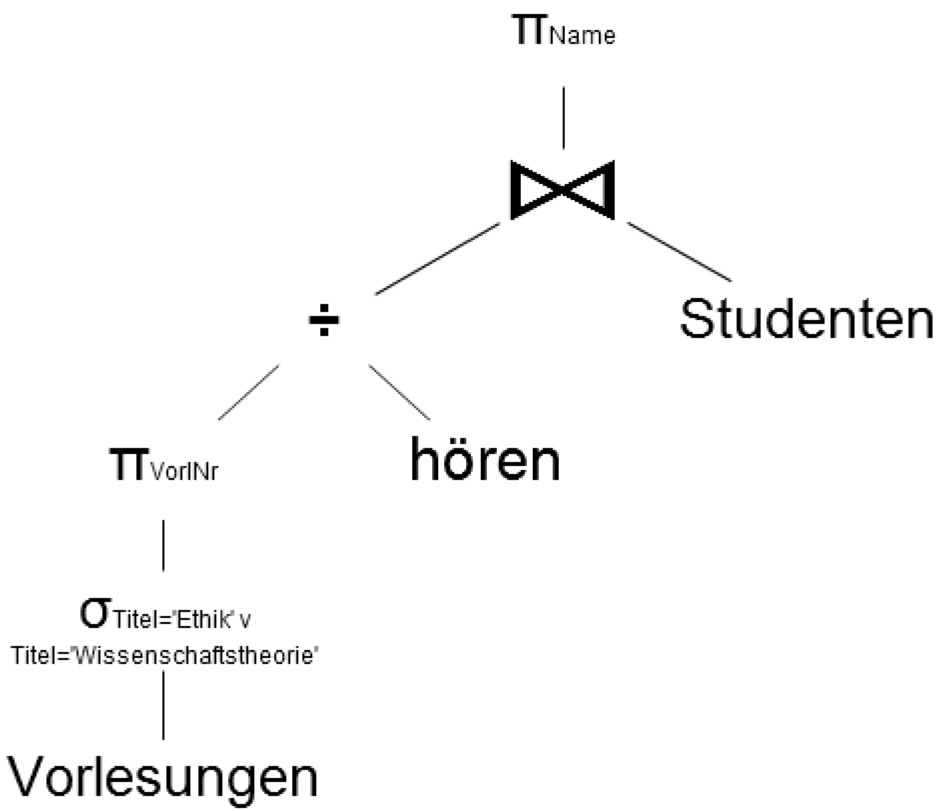
Zeige, welche Noten Professoren mit dem Rang C4 an die Studenten „Fichte“ bzw „Jonas“ vergeben haben



Relationale Algebra, Operatoren II

Lösungen: 4 (nur noch als Operatorbäume)

Welche Studenten (Namen) haben bereits Ethik und Wissenschaftstheorie gehört?



Das waren die „Grundlagen“...

wie geht es ggf. im Masterstudium weiter?

- Alternative zur RA: Formulierung von Anfragen durch logische Terme und Formeln

- wesentlicher Unterschied: Logiksprachen sind "deskriptiver" als algebraische Sprachen; Abarbeitungsreihenfolge ist aus Formeln meist nicht ersichtlich.

- In Codd's grundlegender Arbeit wurden zwei logische Kalküle für relationale Datenbanken vorgeschlagen, die heute die Grundlage der meisten "konkreten" Anfragesprachen bilden:

Tupelkalkül

(engl.: "tuple relational calculus", TRC)

Bereichskalkül

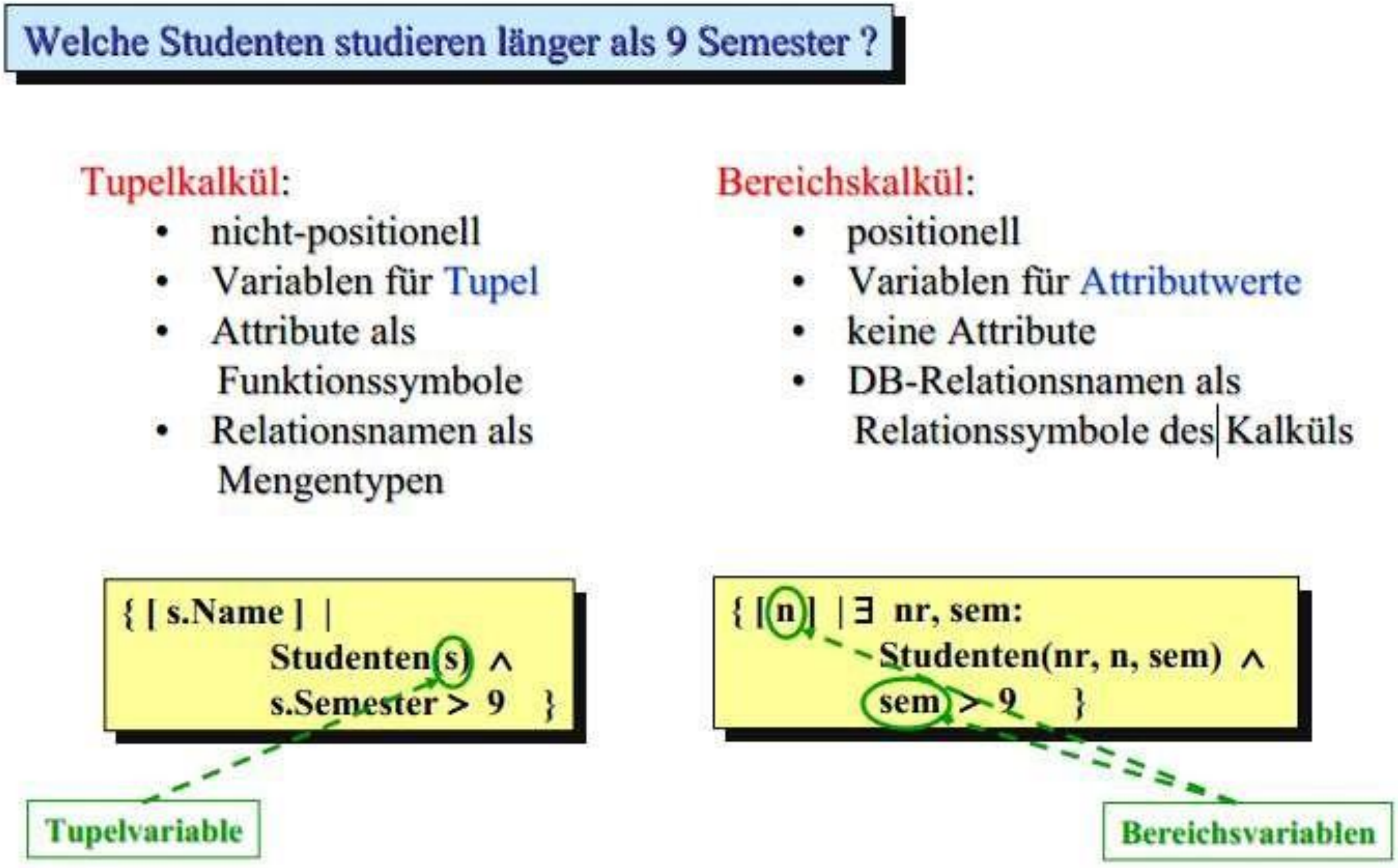
(engl.: "domain relational calculus", DRC)

bei Kemper/Eickler daher auch: Domänenkalkül

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Das waren die „Grundlagen“...

wie geht es ggf. im Masterstudium weiter?



 Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Relationale Algebra
Zusammenfassung

- Es gibt 6 Grundoperatoren, aus denen sich alle weiteren ableiten lassen:
  - Projektion, Auswahl (Selektion), Produkt, Vereinigung, Differenz, Umbenennung
  - Nun muss man sich bei der Definition des Produktes fragen, ob ein Outer- Join nicht auch dazugehört... siehe Hausaufgabe!
- Darüber hinaus viele abgeleitete Operatoren
- Eine Anfragesprache ist relational vollständig, wenn jeder Ausdruck der relationalen Algebra formuliert werden kann
  - Meist sind die Sprachen jedoch viel mächtiger
- Bereichs,- oder Tupelkalkül ein wenig mächtiger als RA, jedoch können Ausdrücke unsicher sein (Antwortmenge unendlich groß)
  - $\{ x \mid \neg R(x) \}$ → Alle möglichen Tupel, die nicht in Relation R vorkommen...
- Diverse theoretische Erweiterungen (Aggregate [Summe, ...], Rekursion, ...)

Relationale Algebra

Schwächen relational vollständiger Sprachen

- Es gibt trotzdem eine ganze Reihe sinnvoller Anfragen, die nicht in der RA formulierbar sind (und damit auch nicht in den sicheren Kalkülen).
- Um diese Fälle abzudecken, muss man die angeblich relational vollständigen Sprachen noch erweitern.
- Erweiterungen um arithmetische und Aggregatfunktionen (inklusive Gruppierungs- und Sortierungsoperatoren) sind in der Praxis unerlässlich.
- Boolesche Anfragen (Ja/Nein-Anfragen) lassen sich in der RA gar nicht formulieren: Dazu müsste man die RA z.B. um Vergleiche mit der leeren Menge erweitern.
- weiterer schwerwiegender Mangel: Rekursive Anfragen (wie z.B. nach der transitiven Hülle einer Relation) lassen sich ebenfalls nicht ausdrücken:

  Dazu wurde für die RA ein "Hüllenoperator" * vorgeschlagen.
  (Analoge Erweiterungen sind für die Kalküle möglich.)

"Relational vollständige" Sprachen sind nicht so vollständig, wie es scheint . . . !

Quelle: Foliensatz 2006 Prof. Dr. Rainer Manthey

Hausaufgaben
bis zur nächsten Vorlesung

@

(a) Zeigen Sie, dass sich der Left-Outer-Join aus anderen nicht-Outer-Join-Operatoren ableiten lässt.

Beachten Sie dabei, dass Sie die Konstante {[$NULL_1$, ..., $NULL_n$]} für eine Relation mit n Attributen verwenden können, die nur NULL-Werte enthält.

@

(b)Ermitteln Sie alle Vorlesungstitel, die (direkte oder indirekte) Voraussetzung von „Der Wiener Kreis" sind. Beachten Sie: es können Daten hinzukommen!

(c)Finden Sie die Professoren P, die eine Vorlesung halten, bei der alle direkten Vorgängervorlesungen (Voraussetzungen) von der selben Person Q gehalten werden (P und Q können gleich sein)

@

(d) Für Studenten müssen mögliche Betreuer der Bachelorarbeiten gefunden werden. Dies sind Assistenten der Professoren, bei denen ein Student schon mal Unterricht hatte. Erstellen Sie eine Übersicht aller möglichen Betreuer je Student. Falls es keinen gibt, soll NULL angezeigt werden.