

Logische Optimierung

Wie kann logisch optimiert werden?


Betrachten wir nochmals das Ausgangsbeispiel von S. 4:

```
select Titel from Professoren, Vorlesungen
where Name='Popper' and PersNr=gelesenVon;
```

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

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

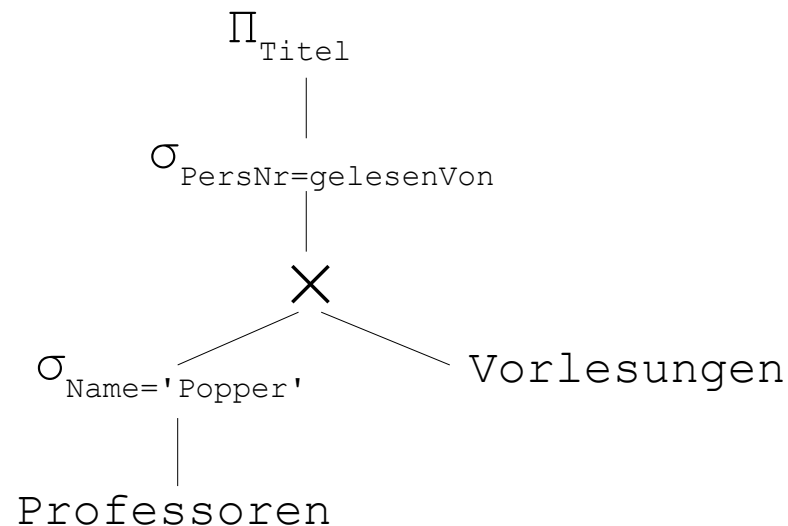
$\Pi_{\text{Titel}} (\sigma_{\text{Name='Popper' } \wedge \text{ PersNr=gelesenVon}} (\text{Professoren} \times \text{Vorlesungen}))$



Welche Schritte sind zur Berechnung des Ausdrucks notwendig?
Wie viele Tupel werden wann verarbeitet?
Könnte man diese Abfolge optimieren?

Wie kann logisch optimiert werden?

Optimierung durch Aufbrechen der Selektionen und deren Verschieben in den Ausdruck hinein:



$\Pi_{\text{Titel}} (\sigma_{\text{PersNr}=\text{gelesenVon}} (\sigma_{\text{Name}='Popper', (\text{Professoren}) \times \text{Vorlesungen}}))$

Äquivalenzen i.d. relationalen Algebra

1. Join, Vereinigung, Schnitt und Kreuzprodukt sind **kommutativ**:

$$R_1 \bowtie R_2 = R_2 \bowtie R_1$$

$$R_1 \cup R_2 = R_2 \cup R_1$$

$$R_1 \cap R_2 = R_2 \cap R_1$$

$$R_1 \times R_2 = R_2 \times R_1$$

2. Selektionen sind vertauschbar:

$$\sigma_p(\sigma_q(R)) = \sigma_q(\sigma_p(R))$$

Äquivalenzen i.d. relationalen Algebra

3. Join, Vereinigung, Schnitt und Kreuzprodukt sind assoziativ:

$$R_1 \bowtie (R_2 \bowtie R_3) = (R_1 \bowtie R_2) \bowtie R_3$$

$$R_1 \cup (R_2 \cup R_3) = (R_1 \cup R_2) \cup R_3$$

$$R_1 \cap (R_2 \cap R_3) = (R_1 \cap R_2) \cap R_3$$

$$R_1 \times (R_2 \times R_3) = (R_1 \times R_2) \times R_3$$

4. Konjunktionen in einer Selektionsbedingung können aufgebrochen bzw. nacheinander ausgeführte Selektionen durch Konjunktion zusammengefasst werden.

$$\sigma_{p_1 \wedge p_2 \wedge \dots \wedge p_n}(R) = \sigma_{p_1}(\sigma_{p_2}(\dots(\sigma_{p_n}(R))\dots))$$

Äquivalenzen i.d. relationalen Algebra

5. Geschachtelte Projektionen können eliminiert werden.

$$\Pi_{I_1}(\Pi_{I_2}(\dots(\Pi_{I_n}(R))\dots)) = \Pi_{I_1}(R)$$

nur sinnvoll wenn gilt:
 $I_1 \subseteq I_2 \subseteq \dots \subseteq I_n \subseteq R$

6. Eine Selektion kann an einer Projektion vorbeigeschoben werden, falls die Projektion keine Attribute aus der Selektionsbedingung entfernt.

$$\Pi_I(\sigma_p(R)) = \sigma_p(\Pi_I(R)), \quad \text{falls } \text{attr}(p) \subseteq I$$

Äquivalenzen i.d. relationalen Algebra

7. Eine Selektion kann an einer Joinoperation/Kreuzprodukt vorbeigeschoben werden, falls sie nur Attribute **eines** der beiden Join-Argumente verwendet.

$$\sigma_p(R_1 \bowtie R_2) = \sigma_p(R_1) \bowtie R_2$$
$$\sigma_p(R_1 \times R_2) = \sigma_p(R_1) \times R_2$$

8. Vertauschen von Projektionen und Joins, wenn die Joinattribute bis zum Join erhalten bleiben.

$$\Pi_l(R_1 \bowtie_p R_2) = \Pi_l(\Pi_{l_1}(R_1) \bowtie_p \Pi_{l_2}(R_2))$$

Äquivalenzen i.d. relationalen Algebra

9. Selektionen können mit Mengenoperationen wie Vereinigung, Schnitt und Differenz vertauscht werden (distributiv).

$$\sigma_p(R \cup S) = \sigma_p(R) \cup \sigma_p(S)$$

$$\sigma_p(R \cap S) = \sigma_p(R) \cap \sigma_p(S)$$

$$\sigma_p(R - S) = \sigma_p(R) - \sigma_p(S)$$

10. Der Projektionsoperator kann mit der Vereinigung vertauscht werden (*nicht* Durchschnitt u. Differenz).

$$\Pi_l(R_1 \cup R_2) = \Pi_l(R_1) \cup \Pi_l(R_2)$$

Äquivalenzen i.d. relationalen Algebra

11. Eine Selektion und ein Kreuzprodukt können zu einem Join zusammengefasst werden, wenn die Selektionsbedingung eine Joinbedingung ist.

$$\sigma_{R1.A1=R2.A2}(R_1 \times R_2) = R_1 \bowtie_{R1.A1=R2.A2} R_2$$

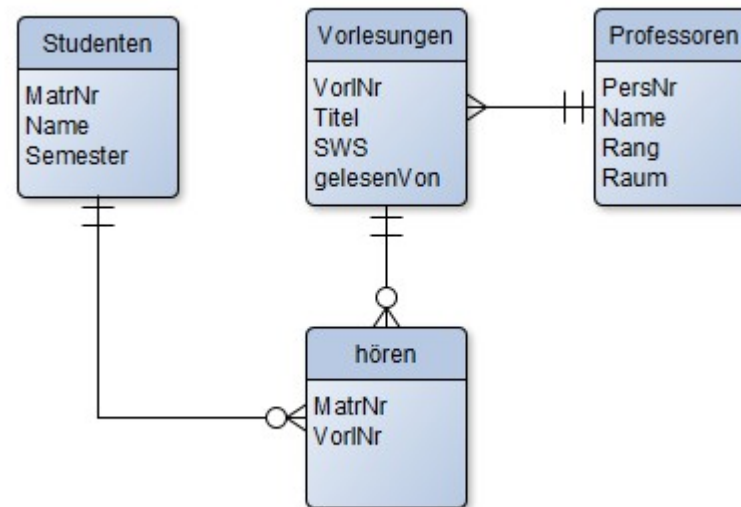
12. Die Join- und/oder Selektionsprädikate können mittels de Morgan's Regeln umgeformt werden. Damit kann auch eine Negation von außen nach innen geschoben werden.

$$\neg (p_1 \vee p_2) = \neg p_1 \wedge \neg p_2$$

$$\neg (p_1 \wedge p_2) = \neg p_1 \vee \neg p_2$$

Anwenden der Transformationsregeln

Beispiel zur Optimierung mit den Transformationsregeln:



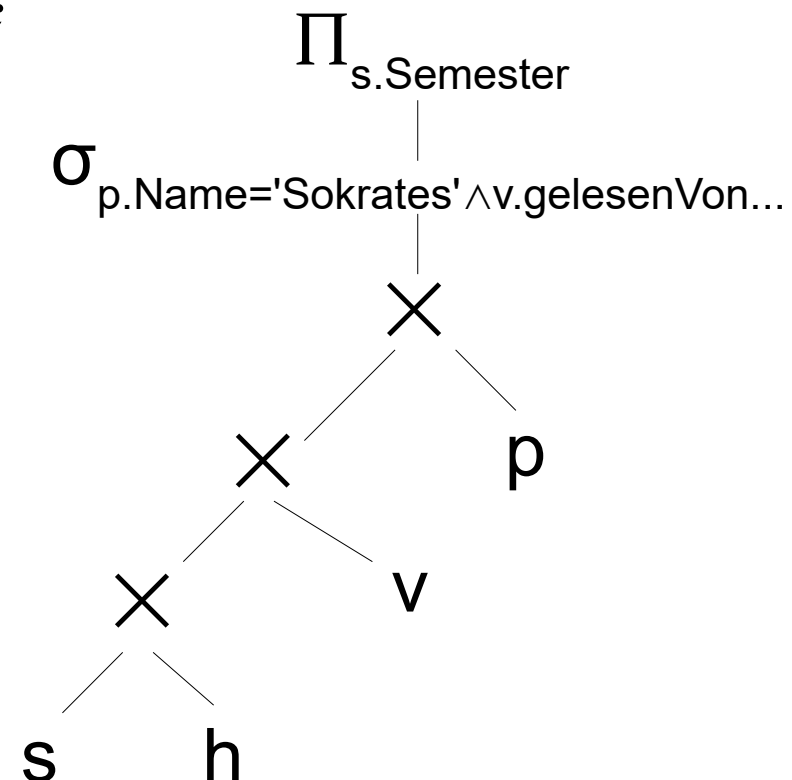
Erstellen Sie folgende Abfrage mit SQL, wobei vorerst keine Joins verwendet werden sollen sondern Kreuzprodukte:

In welchen Semestern sind die Studenten, die Vorlesungen von Sokrates hören?

Erstellen Sie eine Baumdarstellung der kanonischen Übersetzung!

Anwenden der Transformationsregeln

```
select distinct s.Semester  
from Studenten s, hören h,  
Vorlesungen v, Professoren p  
where p.Name = 'Sokrates' and  
v.gelesenVon = p.PersNr and  
v.VorlNr = h.VorlNr and  
h.MatrNr = s.MatrNr;
```

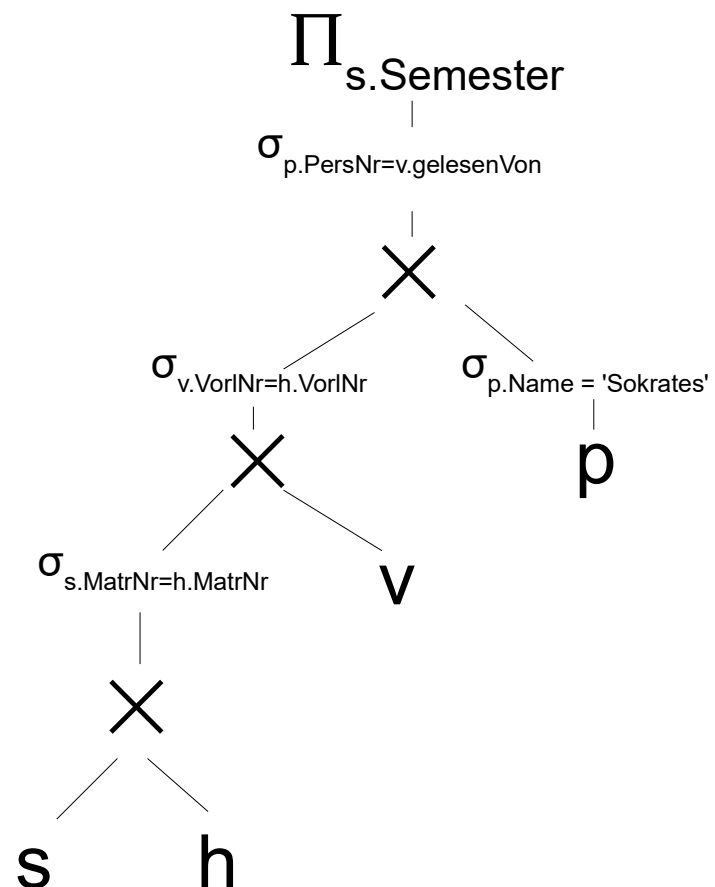
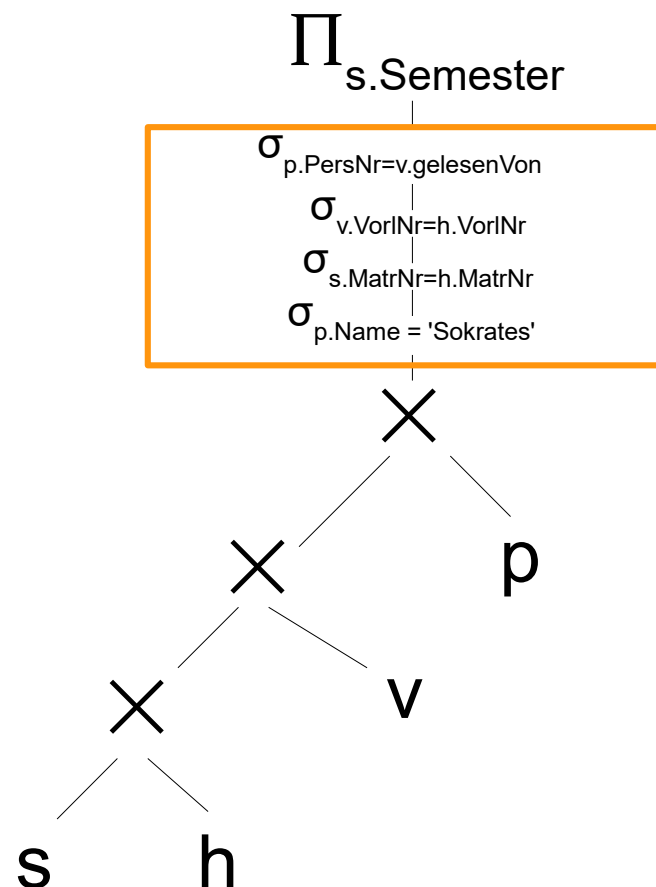


Aufbrechen der Selektion

Lt. Regel 4 können die Konjunktionsglieder der Selektion aufgebrochen werden.

Diese können einzeln innerhalb des Ausdrucks verschoben werden (Regeln 2, 6, 7 und 9).

Selektion so früh wie möglich, um gleich einen großen Teil der Tupel von der weiteren Verarbeitung auszuschließen!

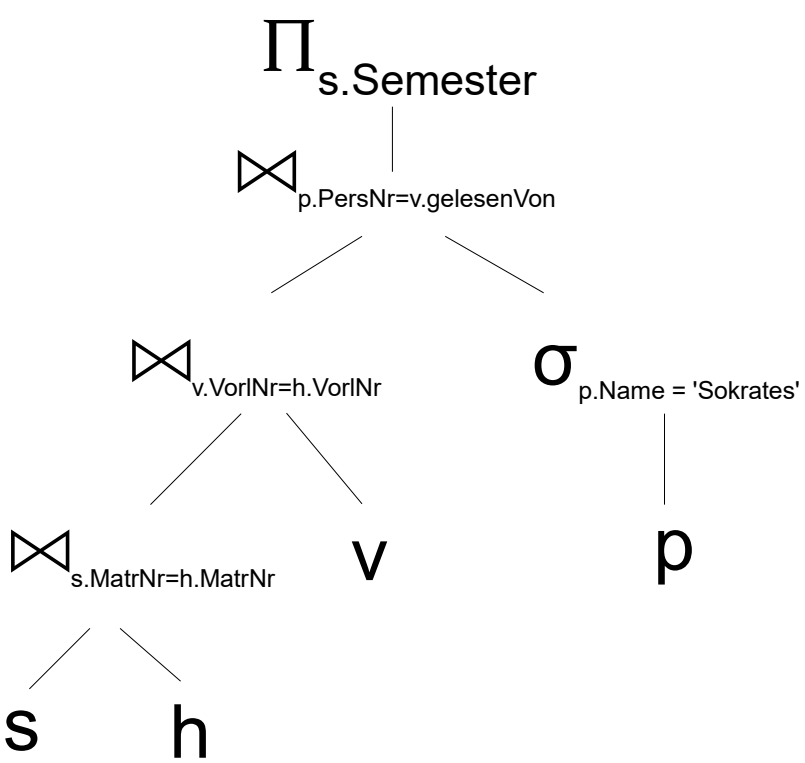


Zusammenfassen von Selektionen und Kreuzprodukten zu Joins

Joins sind Kreuzprodukten vorzuziehen, da letztere das Zwischenergebnis stark aufblähen würden.

→ Umwandeln in Joins (Regel 11)

Joins statt Kreuzprodukten!



hören		Studenten			Professoren	
MatrNr	VorlNr	MatrNr	Name	Semester	PersNr	Name
26120	5001	24002	Xenokrates	18	2125	Sokrates
27550	5001	25403	Jonas	12	2126	Russel
27550	4052	26120	Fichte	10	2127	Kopernikus
28106	5041	26830	Aristoxenos	8	2133	Popper
28106	5052	27550	Schopenhauer	6	2134	Augustinus
28106	5216	28106	Carnap	3	2136	Curie
28106	5259	29120	Theophrastos	2	2137	Kant
29120	5001	29555	Feuerbach	2		
29120	5041					
29120	5049					
29555	5022					
25403	5022					
29555	5001					

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

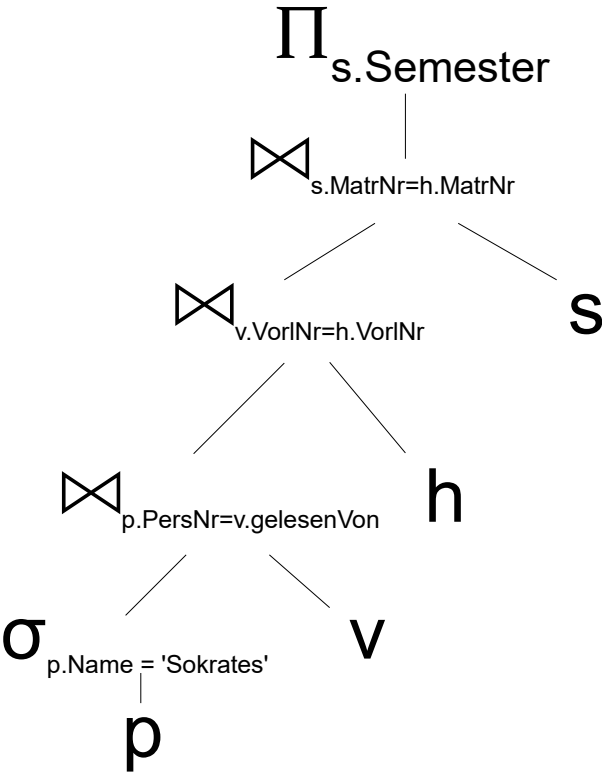
?

Wie viele Tupel beinhaltet das jeweilige Zwischenergebnis an den Joins?

Bestimmen der Join-Reihenfolge


Mit Hilfe der Kommutativität und der Assoziativregel (3) kann die Joinreihenfolge angepasst werden.

Reihenfolge finden mit
geringen Zwischengrößen!



hören		Studenten			Professoren	
MatrNr	VorlNr	MatrNr	Name	Semester	PersNr	Name
26120	5001	24002	Xenokrates	18	2125	Sokrates
27550	5001	25403	Jonas	12	2126	Russel
27550	4052	26120	Fichte	10	2127	Kopernikus
28106	5041	26830	Aristoxenos	8	2133	Popper
28106	5052	27550	Schopenhauer	6	2134	Augustinus
28106	5216	28106	Carnap	3	2136	Curie
28106	5259	29120	Theophrastos	2	2137	Kant
29120	5001	29555	Feuerbach	2		
29120	5041					
29120	5049					
29555	5022					
25403	5022					
29555	5001					

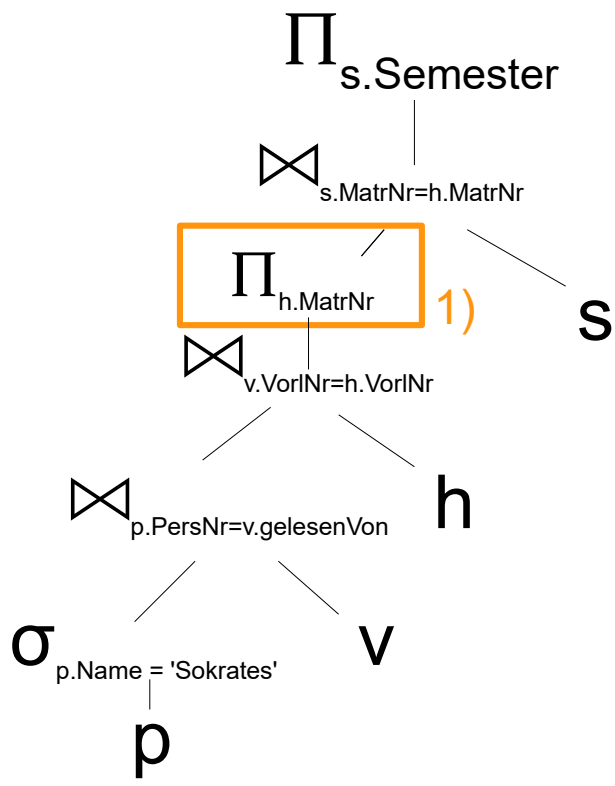
Vorlesungen			
VorlNr	Titel	SWS	gelesen von
5001	Grundzüge	4	2137
5041	Ethik	4	2125
5043	Erkenntnistheorie	3	2126
5049	Mäeutik	2	2125
4052	Logik	4	2125
5052	Wissenschaftstheorie	3	2126
5216	Bioethik	2	2126
5259	Der Wiener Kreis	2	2133
5022	Glaube und Wissen	2	2134
4630	Die 3 Kritiken	4	2137

 Wie viele Tupel beinhaltet jetzt das jeweilige Zwischenergebnis an den Joins?

Verschieben/Einfügen von Projektionen

Durch Anwendung der Regeln 5, 6, 8 und 10.

- Elimination von Duplikaten, die durch Projektion entstanden
- Reduktion der Tupelgröße (Aufwändig → Sinn prüfen!)



hören		Studenten			Professoren	
MatrNr	VorlNr	MatrNr	Name	Semester	PersNr	Name
26120	5001	24002	Xenokrates	18	2125	Sokrates
27550	5001	25403	Jonas	12	2126	Russel
27550	4052	26120	Fichte	10	2127	Kopernikus
28106	5041	26830	Aristoxenos	8	2133	Popper
28106	5052	27550	Schopenhauer	6	2134	Augustinus
28106	5216	28106	Carnap	3	2136	Curie
28106	5259	29120	Theophrastos	2	2137	Kant
29120	5001	29555	Feuerbach	2		
29120	5041					
29120	5049					
29555	5022					
25403	5022					
29555	5001					

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

1)

Zwischenergebnis							
p.PersNr	p.Name	v.VorlNr	v.Titel	v.SWS	v.gelesenVon	h.MatrNr	h.VorlNr
2125	Sokrates	5041	Ethik	4	2125	28106	5041
2125	Sokrates	5041	Ethik	4	2125	29120	5041
2125	Sokrates	5049	Mäeutik	2	2125	29120	5049
2125	Sokrates	4052	Logik	4	2125	27550	4052

Optimierungsheuristik f. Kanonische NF

- 1) Aufbrechen von Selektionen
- 2) Verschieben der Selektionen soweit wie möglich nach unten im Operatorbaum
- 3) Zusammenfassen von Selektionen und Kreuzprodukten zu Joins
- 4) Reihenfolge der Joins festlegen, so dass möglichst kleine Zwischenergebnisse erreicht werden
- 5) Evtl. Einfügen von Projektionen
- 6) Verschieben der Projektionen soweit wie möglich nach unten im Operatorbaum