

Übung zur Vorlesung *Grundlagen: Datenbanken* im WS19/20

Christoph Anneser, Moritz Sichert, Lukas Vogel (gdb@in.tum.de)

<https://db.in.tum.de/teaching/ws1920/grundlagen/>

Blatt Nr. 02

Tool zum Üben der relationalen Algebra <https://ira.db.in.tum.de/>.

Hausaufgabe 1

Formulieren Sie folgende Anfragen auf dem bekannten Universitätsschema in der relationalen Algebra, sowie im Tupel- und Domänenkalkül:

- Finden Sie die *Vorlesungen*, die keine Hörer haben.
- Finden Sie die *Studenten*, die alle *Vorlesungen* hören.

Lösung:

- Finden Sie die *Vorlesungen*, die keine Hörer haben.

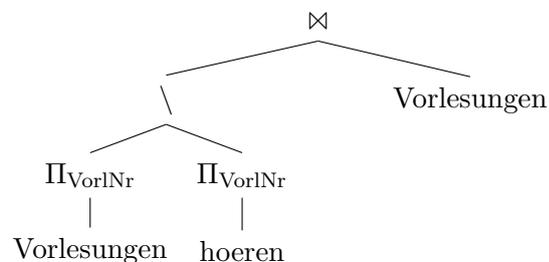
Formulierung in relationaler Algebra

$$(\Pi_{\text{VorlNr}}(\text{Vorlesungen}) \setminus \Pi_{\text{VorlNr}}(\text{hören})) \bowtie \text{Vorlesungen}$$

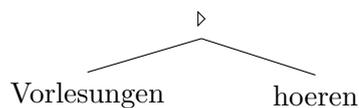
oder

$$\text{Vorlesungen} \triangleright \text{hören}$$

In Operatorbaumdarstellung:



oder



Formulierung im Tupelkalkül

$$\{v \mid v \in \text{Vorlesungen} \wedge \nexists h \in \text{hören}(v.\text{VorlNr} = h.\text{VorlNr})\}$$

oder

$$\{v \mid v \in \text{Vorlesungen} \wedge \forall h \in \text{hören}(v.\text{VorlNr} \neq h.\text{VorlNr})\}$$

Formulierung im Domänenkalkül

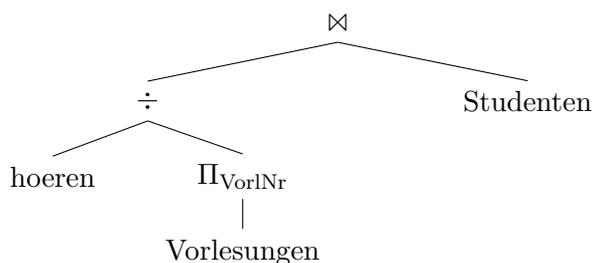
$$\{[v,t,s,g] \mid [v,t,s,g] \in \text{Vorlesungen} \wedge \exists m([m,v] \in \text{hören})\}$$

b) Finden Sie die *Studenten*, die alle *Vorlesungen* hören.

Formulierung in relationaler Algebra

$$(\text{hoeren} \div \Pi_{\text{VorlNr}}(\text{Vorlesungen})) \bowtie \text{Studenten}$$

In Operatorbaumdarstellung:



Formulierung im Tupelkalkül

$$\{s \mid s \in \text{Studenten} \wedge \forall v \in \text{Vorlesungen}(\exists h \in \text{hoeren}(v.\text{VorlNr} = h.\text{VorlNr} \wedge s.\text{MatrNr} = h.\text{MatrNr}))\}$$

Formulierung im Domänenkalkül

$$\{[m,n,s] \mid [m,n,s] \in \text{Studenten} \wedge \forall v,t,sws,g([v,t,sws,g] \in \text{Vorlesungen} \Rightarrow [m,v] \in \text{hören})\}$$

Hausaufgabe 2

Beantworten Sie mittels relationaler Algebra:

- Geben Sie einen Ausdruck an, der die Relation $\neg\text{hoeren}$ erzeugt. Diese enthält für jeden Studenten und jede Vorlesung, die der Student **nicht** hört einen Eintrag mit Matrikelnummer und Vorlesungsnummer.
- Finden Sie alle Studenten, die keine Vorlesung hören. Geben Sie zwei verschiedene Lösungen an.

Lösung:

a)

$$(\Pi_{\text{MatrNr}} \text{Studenten} \times \Pi_{\text{VorlNr}} \text{Vorlesungen}) - \text{hoeren}$$

b)

$$\text{Studenten} \triangleright \text{hoeren}$$

oder

$$\text{Studenten} - (\text{Studenten} \bowtie \text{hoeren})$$

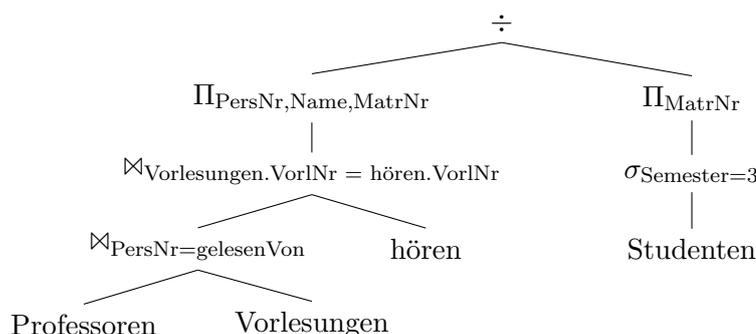
Hausaufgabe 3

Formulieren Sie folgende Anfragen auf dem bekannten Universitätsschema in der Relationalalgebra:

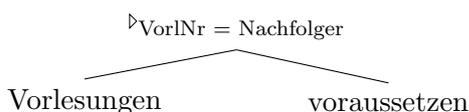
- Finden Sie alle bei den Drittsemestern beliebte Professoren. Ein Professor ist bei einem gegebenen Semester beliebt, wenn alle Studenten aus diesem Semester mindestens eine seiner Vorlesungen hören (aber nicht notwendigerweise alle dieselbe).
- Finden Sie alle Grundlagenvorlesungen. Eine Grundlagenvorlesung ist eine Vorlesung, die keine Voraussetzungen hat.
- Carnap will eine Seminararbeit einreichen. Er will in seiner Danksagung alle Professoren und ihre Assistenten erwähnen, deren Vorlesungen er hört. Geben Sie eine Anfrage an, die alle diese Namen ermittelt.

Lösung:

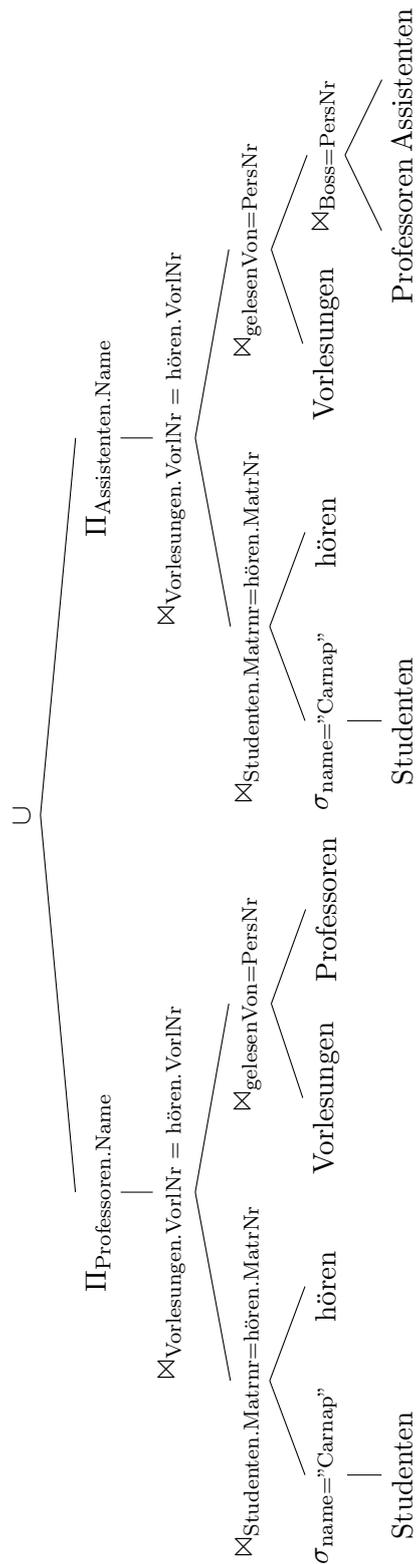
- Um Anfragen mit Allquantifizierung auszudrücken, kann man in der relationalen Algebra den Divisionsoperator verwenden. Wichtig ist hier, dass das Schema des Divisors eine Teilmenge des Schemas des Dividenden ist. Daher projizieren wir vor der Division den Dividenden auf die Matrikelnummer.



- Um Anfragen der Art "Finde alle Elemente für die es kein Pendant (Joinpartner gibt)" auszudrücken, verwendet man den Anti-Semi-Join.



- Wir suchen zunächst alle Professoren, deren Vorlesung Carnap besucht und dann (in einer beinahe identischen Anfrage) alle ihre Assistenten. Um die Namensliste zu bekommen, bilden wir dann die Vereinigungsmenge von beiden. Wichtig ist hier, jeweils auf den Namen zu projizieren, da für die Vereinigung beide Schemata gleich



sein müssen.

Hausaufgabe 4

Gegeben seien die beiden Relationen $R : \{[a_1, \dots, a_n]\}$ und $S : \{[b_1, \dots, b_m]\}$. Geben Sie die folgenden Ausdrücke im Tupel- und Domänenkalkül an:

- a) $Q_1 := R \bowtie_{a_1=b_1} S$
- b) $Q_2 := R \bowtie_{a_1=b_1} S$
- c) $Q_3 := R \bowtie_{a_1=b_1} S$
- d) $Q_4 := R \bowtie_{a_1=b_1} S$

Lösung:

Bitte beachten Sie, dass in dieser Aufgabe ausschließlich allgemeine *Theta*joins ($\bowtie_{\Theta}, \bowtie_{\Theta}, \dots$) verwendet werden. Gemäß Definition werden somit alle Attribute der beiden Eingaberelationen in die Ausgabere Relation projiziert, einschließlich der Attribute, welche in der Joinbedingung enthalten sind. Lediglich bei *natürlichen* Joins, wo implizit eine Gleichheitsbedingung für alle gleichnamigen Attribute erfüllt sein muss, werden gleichnamige Attribute nicht doppelt in die Ausgabere Relation projiziert. Siehe hierzu auch Folie Kapitel 3, "Andere Join-Arten".

- a) $Q_1 := R \bowtie_{a_1=b_1} S$

Formulierung im Tupelkalkül

$$Q_1 := \{[r.a_1, \dots, r.a_n, s.b_1, \dots, s.b_m] \mid r \in R \wedge s \in S \wedge r.a_1 = s.b_1\}$$

Da der Joinoperator Tupel aus verschiedenen Relationen verbindet, müssen für die Ergebnismenge neue Tupel mithilfe des Tupelkonstruktors konstruiert werden: $[\text{attribut}_1 : \text{wert}_1, \dots, \text{attribut}_n : \text{wert}_n]$. Die oben verwendete Tupelkonstruktion $[r.a_1, \dots]$ ist eine verkürzte Schreibweise für $[a_1 : r.a_1, \dots]$ und kann verwendet werden, wenn der Attributname im neuen Tupel unverändert bleibt.

Im Falle des Thetajoins kann auch die Tupelkonkatenation $t_1 \circ t_2$ verwendet werden:

$$Q_1 := \{r \circ s \mid r \in R \wedge s \in S \wedge r.a_1 = s.b_1\}$$

Formulierung im Domänenkalkül

$$Q_1 := \{[a_1, \dots, a_n, b_1, \dots, b_m] \mid [a_1, \dots, a_n] \in R \wedge [b_1, \dots, b_m] \in S \wedge a_1 = b_1\}$$

oder

$$Q_1 := \{[a_1, \dots, a_n, b_1 : a_1, b_2, \dots, b_m] \mid [a_1, \dots, a_n] \in R \wedge [a_1, b_2, \dots, b_m] \in S\}$$

- b) $Q_2 := R \bowtie_{a_1=b_1} S$

Formulierung im Tupelkalkül

$$Q_2 := Q_1 \cup \{[r.a_1, \dots, r.a_n, b_1 : \text{null}, \dots, b_m : \text{null}] \mid r \in R \wedge \nexists s \in S(r.a_1 = s.b_1)\}$$

Formulierung im Domänenkalkül

$$Q_2 := Q_1 \cup \{[a_1, \dots, a_n, b_1 : \text{null}, \dots, b_m : \text{null}] \mid [a_1, \dots, a_n] \in R \wedge \exists c_2, \dots, c_m ([a_1, c_2, \dots, c_m] \in S)\}$$

c) $Q_3 := R \bowtie_{a_1=b_1} S$

Formulierung im Tupelkalkül

$$Q_3 := \{s \mid s \in S \wedge \exists r \in R (r.a_1 = s.b_1)\}$$

Formulierung im Domänenkalkül

$$Q_3 := \{[b_1, \dots, b_m] \mid [b_1, \dots, b_m] \in S \wedge \exists a_2, \dots, a_n ([b_1, a_2, \dots, a_n] \in R)\}$$

d) $Q_4 := R \ltimes_{a_1=b_1} S$

Formulierung im Tupelkalkül

$$Q_4 := \{s \mid s \in S \wedge \nexists r \in R (r.a_1 = s.b_1)\}$$

Formulierung im Domänenkalkül

$$Q_4 := \{[b_1, \dots, b_m] \mid [b_1, \dots, b_m] \in S \wedge \nexists a_2, \dots, a_n ([b_1, a_2, \dots, a_n] \in R)\}$$