



Übung zur Vorlesung *Einsatz und Realisierung von Datenbanken im SoSe22*

Alice Rey, Maximilian {Bandle, Schüle}, Michael Jungmair (i3erdb@in.tum.de)

<http://db.in.tum.de/teaching/ss22/impldb/>

Blatt Nr. 05

Hinweise Die Datalogaufgaben können auf <https://datalog.db.in.tum.de/> getestet werden. Auf der Seite kann unter *examples* ein entsprechender Datensatz geladen werden. Die neuen IDB Regeln sollten am Ende der EDB definiert und dann im Query-Eingabefeld abgefragt werden.

Zusätzlich zu der in der Vorlesung vorgestellten Syntax hier noch eine Kurzübersicht der Vergleichsoperatoren: $X < Y, Y > X$ (kleiner, größer), $X =< Y, X >= Y$ (kleiner oder gleich, größer oder gleich), $X = Y, X \neq Y$ (gleich, ungleich), $not(pred(X, Y))$ (existiert nicht $pred(X, Y)$).

Hausaufgabe 1

Definieren Sie das Prädikat $sg(X, Y)$ das für “same generation” steht. Zwei Personen gehören zur selben Generation, wenn Sie mindestens je ein Elternteil haben, das derselben Generation angehört.

Verwenden Sie beispielsweise die folgende Ausprägung einer ElternKind Relation. Das erste Element ist hier das Kind, das zweite ein Elternteil.

```
parent(c, a).  
parent(d, a).  
parent(d, b).  
parent(e, b).  
parent(f, c).  
parent(g, c).  
parent(h, d).  
parent(i, d).  
parent(i, e).  
parent(f, e).  
parent(j, f).  
parent(j, h).  
parent(k, g).  
parent(k, i).
```

- Definieren Sie das Prädikat in Datalog.
- Demonstrieren Sie die naive Ausführung des Prädikats.
- Erläutern Sie das Vorgehen bei der seminaiven Auswertung.

Gruppenaufgabe 2

Ist folgendes Datalog-Programm stratifiziert?

$$\begin{aligned}p(X, Y) & :- q_1(Y, Z), \neg q_2(Z, X), q_3(X, P). \\q_2(Z, X) & :- q_4(Z, Y), q_3(Y, X). \\q_4(Z, Y) & :- p(Z, X), q_3(X, Y).\end{aligned}$$

Ist das Programm sicher – unter der Annahme, dass p, q_1, q_2, q_3, q_4 IDB- oder EDB-Prädikate sind?

Hausaufgabe 3

Gegeben sei folgende Faktenbasis, die einen direkten azyklischen Graphen (DAG) darstellt.

```
kante(1,2).
kante(2,3).
kante(3,4).
kante(2,5).
kante(5,3).
```

1. Geben Sie in Datalog ein Prädikat `pfad(V,N,L)` an, das alle möglichen Pfade von V nach N mit Länge L ausgibt.
2. Geben Sie nun das Prädikat `kuerzestePfade(V,N,L)` an, das pro Beginn V und Ziel N nur den kürzesten Pfad ausgibt.
3. Bestimmen Sie nun den längsten kürzesten Pfad `laengsterkuerzesterPfad(L)`.
4. Erstellen Sie in SQL eine rekursive CTE `pfad(V,N,L)`, die die Länge aller Pfade im DAG ausgibt.
5. Basierend auf `pfad(V,N,L)`, geben Sie die Länge des längsten kürzesten Pfades aus.

Hausaufgabe 4

Gehen Sie von folgender kombinierter Fragmentierung der in Abbildung 1 dargestellten Relation *Professoren* aus:

Professoren						
PersNr	Name	Rang	Raum	Fakultät	Gehalt	Steuerklasse
2125	Sokrates	C4	226	Philosophie	85000	1
2126	Russel	C4	232	Philosophie	80000	3
2127	Kopernikus	C3	310	Physik	65000	5
2133	Popper	C3	52	Philosophie	68000	1
2134	Augustinus	C3	309	Theologie	55000	5
2136	Curie	C4	36	Physik	95000	3
2137	Kant	C4	7	Philosophie	98000	1

Abbildung 1: Beispielausprägung der um drei Attribute erweiterten Relation *Professoren*

1. Zuerst erfolgt eine vertikale Fragmentierung in

$$\begin{aligned}\text{ProfVerw} & := \Pi_{\text{PersNr,Name,Gehalt,Steuerklasse}}(\text{Professoren}) \\ \text{ProfS} & := \Pi_{\text{PersNr,Name,Rang,Raum,Fakultät}}(\text{Professoren})\end{aligned}$$

2. Das Fragment Profs wird weiter horizontal fragmentiert in

$$\begin{aligned}\text{TheolProfs} &:= \sigma_{\text{Fakultät} = \text{'Theologie'}}(\text{Profs}) \\ \text{PhysikProfs} &:= \sigma_{\text{Fakultät} = \text{'Physik'}}(\text{Profs}) \\ \text{PhiloProfs} &:= \sigma_{\text{Fakultät} = \text{'Philosophie'}}(\text{Profs})\end{aligned}$$

Übersetzen Sie aufbauend auf dieser Fragmentierung die folgende SQL-Anfrage in die kanonische Form.

```
select Name, Gehalt Rang
from Professoren
where Gehalt > 80000;
```

Optimieren Sie diesen kanonischen Auswertungsplan durch Anwendung algebraischer Transformationsregeln (Äquivalenzen).

Gruppenaufgabe (wird nicht in der Übung besprochen)

Schreiben Sie zu dem U-Bahn-Netz-Beispiel auf der Datalog Seite (unter Examples) folgende Anfragen in Datalog:

1. Erstellen Sie den Stationsplan für den U-Bahnhof Fröttmanning, der alle Stationen, die ohne Umstieg erreichbar sind, auflistet.
2. Erstellen Sie für Garching-Forschungszentrum einen Plan, der alle erreichbaren Stationen, die minimale Anzahl an Umstiegen und Stops auflistet. Beschreiben Sie Ihren Ansatz ausführlich.