

KFK GC/DM Grid Computing - Datenbanksysteme	schriftliche Einzelpruefung	31.01.2005		1
--	--------------------------------	------------	--	---

### Aufgabe 1 [Relationale Abfragen: 25 Punkte]

Gegeben ist folgendes Relationenschema:

veranstaltung (titel, datum, groesse, art)

person (name, qualifikation, adresse)

mitwirkender (titel, name, honorar)

*mitwirkender.titel*  $\diamond$  *veranstaltung*, *mitwirkender.name*  $\diamond$  *person*

*veranstaltung.art* IN {'freiluft', 'indoor', 'kombiniert'}

Formulieren Sie die folgende Abfrage in Relationenalgebra:

a. (5 Punkte)

Ermitteln Sie alle Personen, die dieselbe Adresse wie eine zweite Person aufweisen. Gleichzeitig müssen beide Personen dieselbe Qualifikation haben. In der Ergebnisrelation sollen beide Namen der Personen sowie die zugehörige Adresse und Qualifikation aufscheinen.

Formulieren Sie die folgenden Abfragen in SQL99:

b. (8 Punkte)

Ermitteln Sie alle Mitwirkenden (name), wo zumindest eine Veranstaltungsart existiert, für die die entsprechende Person genau zweimal mitgewirkt hat.

c. (7 Punkte)

Ermitteln Sie die Titel aller Veranstaltungen, bei der genau 2 Personen mitgewirkt haben, wobei weiters darauf zu achten ist, dass das Gesamthonorar (=Summe der Honorare der beiden Mitwirkenden) größer als 125 sein muss.

KFK GC/DM Grid Computing - Datenbanksysteme	schriftliche Einzelpruefung	31.01.2005		2
--	--------------------------------	------------	--	---

d. (5 Punkte)

Gehen Sie davon aus, dass alle mitwirkenden Personen eine gesonderte Benachrichtigung für die bevorstehende Mitwirkung an einer Veranstaltung erhalten. Ermitteln Sie jene Veranstaltungen (titel), bei denen derartige Nachrichten an mehr als 10 Adressen versendet werden mussten.

**Aufgabe 2 [Datalog: 15 Punkte]**

Folgendes Relationenschema ist gegeben:

*stammbaum*(*kind\_name*, *mutter\_name*, *vater\_name*),  
*stammbaum.mutter\_name*  $\diamond$  *stammbaum*, *stammbaum.vater\_name*  $\diamond$  *stammbaum*.

Die Relation speichert für jedes Kind von wem dieses abstammt. Erstellen Sie ein Datalog-Programm für die folgenden Abfrage:

Ermitteln Sie alle Kinder bei denen beide Elternteile vom selben Ahnen abstammen.

**Aufgabe 3 [Datenstrukturen: 15 Punkte]**

Es handelt sich um variabel lange Records in einer fixen Längen-Repräsentation (Zeiger-Methode) mit folgender Filestruktur:

Hinweis: ↑3 bedeutet Zeiger auf Record 3

record	name	ort	kontostand	Pointer
0	Inwunz	NeuDehli	1	↑ 3
1	Franzi	Linz	250	
2	Seppi	LA	1000	↑ 4
3		Sidney	100000	↑ 6
4		Linz	500	↑ 10
5	Rudi	Barcelona	50	↑ 9
6		Laha	450	
7	Susi	Graz	2000	
8	Anna	Graz	100	
9		Graz	700	↑ 11
10		Wels	1000	
11		Wien	200	

Zeigen Sie die Veränderungen der Filestruktur durch folgende Aktionen, die der Reihe nach durchgeführt werden:

1. delete(Inwunz,Sidney,100000).
1. insert(Inwunz,Würzburg,99000).
2. insert(Rudi,Linz,200).
3. delete(Anna,Graz,100).

**Aufgabe 4 [Query Optimierung: 20 Punkte]**

Gegeben ist die folgende relationale Algebra Abfrage auf dem Relationenschema  $R1(A, B)$ ,  $R2(C, D)$ :

$$\pi(A, C)(\sigma_{(A>5 \wedge C<10)}(R1 \bowtie_{(A=C)} R2))$$

a. (2 Punkte)

Stellen Sie den Ausdruck grafisch dar.

b. (8 Punkte)

Führen Sie eine heuristische Optimierung mit Hilfe der in der Vorlesung angegebenen Äquivalenzumformungsregeln durch, mit dem Ziel den Abarbeitungsaufwand zu minimieren.

Stellen Sie den umgeformten Ausdruck grafisch dar und geben Sie die Reihenfolge der verwendeten Äquivalenzumformungsregeln an.

Begründen Sie Ihre Vorgangsweise.

c. (10 Punkte)

Nehmen Sie an, dass die Relation  $R1$  1000 und die Relation  $R2$  5000 Datensätze enthält, wobei die Blockgröße für beide Relationen 10 ist.

Für den Join wird das Nested-Loop Verfahren verwendet (Memorygröße 1 Block) und die Selektivität der Selektionen ist jeweils  $1/10$  (Annahme der Unabhängigkeit).

Geben Sie die Kosten (Anzahl von Blockzugriffen) für

- den ursprünglichen Ausdruck an und
- den umgeformten Ausdruck an.

Nehmen Sie an, dass die Abarbeitung der Ausdrücke Pipelining nützt.

KFK GC/DM Grid Computing - Datenbanksysteme	schriftliche Einzelpruefung	31.01.2005		5
--	--------------------------------	------------	--	---

**Aufgabe 5 [XML: 15 Punkte]**

Gegeben ist die folgende DTD:

```
<!DOCTYPE db [  
  <!ELEMENT Angestellte (Aname, Kinder*)>  
  <!ELEMENT Kinder (Name, Geburtstag)>  
  <!ELEMENT Geburtstag (Tag, Monat, Jahr)>  
  <!ELEMENT Aname( #PCDATA )>  
  <!ELEMENT Name( #PCDATA )>  
  <!ELEMENT Tag( #PCDATA )>  
  <!ELEMENT Monat( #PCDATA )>  
  <!ELEMENT Jahr( #PCDATA )>  
>
```

a. (5 Punkte)

Geben Sie ein gültiges XML Dokument an, das auf dieser DTD beruht und in dem alle Elemente der DTD vorkommen.

b. (5 Punkte) Schreiben Sie eine XPath Abfrage, die alle Angestellte liefert, die mehr als 2 Kinder haben.

c. (5 Punkte)

Schreiben Sie eine XQuery Abfrage, die alle Angestellte nach ihren Namen sortiert liefert, deren Kinder im "März" Geburtstag haben.

KFK GC/DM Grid Computing - Datenbanksysteme	schriftliche Einzelpruefung	31.01.2005		7
--	--------------------------------	------------	--	---

**Aufgabe 6 [Datagrid: 10 Punkte]**

- a. (5 Punkte) Was sind die Ziele und die Eigenschaften des Gridcomputing?
- b. (5 Punkte) Was versteht man unter Replikation im Datagrid und warum wird sie eingesetzt?