

Datenbanksysteme (f. InformatikerInnen)	schriftliche Einzelprüfung	11.05.2006	1
-----------------------------------------	----------------------------	------------	---

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

Gegeben ist folgendes Relationenschema:

fahrzeug (seriennummer, marke, type, baujahr)

person (zuname, vorname, adresse, telefon)

probefahrt (zuname, seriennummer, datum, km)

*probefahrt.seriennummer*  $\diamond$  *fahrzeug*, *probefahrt.zuname*  $\diamond$  *person*

Formulieren Sie die folgende Abfrage in Relationenalgebra (ohne Erweiterungen für Aggregatfunktionen zu verwenden):

- (8 Punkte) Ermitteln Sie Zuname und Telefonnummer (telefon) aller Personen, die das Fahrzeug mit der Seriennummer "4711" höchstens einmal probegefahren haben.  
(Hinweis: höchstens einmal ist äquivalent zu nicht mindestens zweimal)

Formulieren Sie die folgenden Abfragen in SQL99:

- (5 Punkte) Ermitteln Sie die Zunamen aller Personen, die mindestens einmal probegefahren sind, aber bei keiner Fahrt mehr als 50 km zurückgelegt haben.
- (7 Punkte) Ermitteln Sie Paare von Seriennummern aller Fahrzeuge, die das gleiche Baujahr aber verschiedene Marken haben, wobei zumindest eine Person existiert, die beide Fahrzeuge probegefahren hat.
- (5 Punkte) Ermitteln Sie die Seriennummer aller Fahrzeuge, die jeweils mehr als 1000 km probegefahren wurden.

### Aufgabe 2 [Datalog: 15 Punkte]

Folgendes Relationenschema ist gegeben:

*person*(name, alter),

*sendetan*(name1, name2),

*sendetan.name1*  $\diamond$  *person*,

*sendetan.name2*  $\diamond$  *person*.

Die Relation "sendetan" speichert die Ausbreitung von Kettenbriefen. Ein Eintrag ist so zu interpretieren, dass die Person mit "name1" den Kettenbrief an die Person mit "name2" weitergeschickt hat.

Finden Sie eine Datalog-Abfrage, die die Namen aller Personen unter 20 Jahren findet, die (direkt oder indirekt) den von "Maria Musterfrau" initiierten Kettenbrief erhalten haben.

### Aufgabe 3 [Speicher und Indexstrukturen: 15 Punkte]

Was versteht man unter variabel langen Records. Welche Methoden zu deren Repräsentierung haben Sie in der Vorlesung kennengelernt. Geben Sie jeweils geeignete Beispiele und Vor- und Nachteile der beschriebenen Methoden an.

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

Gegeben ist die folgende relationale Algebra Abfrage auf dem Relationenschema  $R1(A, D)$ ,  $R2(B, D)$ ,  $R3(B, E)$ :

$\sigma_{(A < 3 \wedge B < 5)}(R1 * (R2 * R3))$

- (2 Punkte)

Stellen Sie den Ausdruck grafisch dar.

- (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.

- (10 Punkte)

Nehmen Sie an, dass die Relation  $R1$  2000, die Relation  $R2$  1000 und die Relation  $R3$  3000 Datensätze enthält, wobei die Blockgröße für alle Relationen 10 ist.

Für den Join wird das Block Nested-Loop Verfahren verwendet (Memorygröße 12 Blöcke), die Ergebnisse eines Joins enthalten nicht mehr Tupel als die kleinere Relation, und die Selektivität der Selektionen ist jeweils 1/10 (Annahme der Unabhängigkeit).

Geben Sie die Kosten (Anzahl von Blockzugriffen) für

- den ursprünglichen Ausdruck und
- den umgeformten Ausdruck an.

Datenbanksysteme (f. InformatikerInnen)	schriftliche Einzelprüfung	11.05.2006		2
-----------------------------------------	----------------------------	------------	--	---

Nehmen Sie an, dass Zwischen- und Endergebnisse materialisiert werden müssen.

### Aufgabe 5 [XML: 15 Punkte]

Gegeben ist die folgende DTD:

```
<!DOCTYPE Kiste [
  <!ELEMENT Kiste (Karton*)>
  <!ELEMENT Karton (Hoehe, Breite, Laenge, Gewicht, Artikel*)>
  <!ELEMENT Artikel (Nummer, Bezeichnung)>
  <!ELEMENT Hoehe( #PCDATA )>
  <!ELEMENT Breite( #PCDATA )>
  <!ELEMENT Laenge( #PCDATA )>
  <!ELEMENT Gewicht( #PCDATA )>
  <!ELEMENT Nummer( #PCDATA )>
  <!ELEMENT Bezeichnung( #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 Kisten-Elemente liefert, die höchstens 10 Kartons beinhalten.

c. (5 Punkte) Schreiben Sie eine XQuery Abfrage, die das Gesamtgewicht aller Kartons liefert, die höchstens drei Artikel beinhalten.

### Aufgabe 6 [Transaktionskonzept: 10 Punkte]

Beschreiben Sie das Konzept der Serialisierbarkeit. Welche Formen kennen Sie? Geben Sie jeweils ein Beispiel an.