

Datenbanksysteme (f. InformatikerInnen)	schriftliche Einzelprüfung	16.03.2006	1
---	----------------------------	------------	---

Aufgabe 1 [Relationale Abfragen: 25 Punkte]

Gegeben ist folgendes Relationenschema:

ort (ortsname, einwohnerzahl, flaeche)

person (sozialversicherungsnummer, vorname, zuname, geburtsdatum)

bereist (ortsname, sozialversicherungsnummer, datum, bezahlt)

bereist.ortsname \diamond *ort*, *bereist.sozialversicherungsnummer* \diamond *person*

bereist.bezahlt IN {'privat', 'firma', 'sponsor'}

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

- (8 Punkte) Ermitteln Sie Sozialversicherungsnummer und Zuname aller Personen, die den Ort "Wien" höchstens einmal bereist haben.
(Hinweis: höchstens einmal ist äquivalent zu nicht mindestens zweimal)

Formulieren Sie die folgenden Abfragen in SQL99:

- (5 Punkte) Ermitteln Sie die Sozialversicherungsnummer aller Personen, die mindestens einmal und immer nur auf Kosten eines Sponsors gereist sind.
- (7 Punkte) Ermitteln Sie die Ortsnamen aller Orte, die die gleiche Einwohnerzahl aber verschiedene Flächen haben, wobei zumindest eine Person existiert, die beide Orte bereist hat.
- (5 Punkte) Ermitteln Sie Ortsname und Einwohnerzahl aller Orte, die jeweils weniger als 100 mal besucht wurden.

Aufgabe 2 [Datalog: 15 Punkte]

Folgendes Relationenschema ist gegeben:

person(name, alter),

infiziert(name1, name2),

infiziert.name1 \diamond *person*,

infiziert.name2 \diamond *person*.

Die Relation "infiziert" speichert die Ausbreitung einer virösen Infektion. Ein Eintrag ist so zu interpretieren, dass die Person mit "name1" die Person mit "name2" angesteckt hat.

Finden Sie eine Datalog-Abfrage, die die Namen aller Personen älter als 30 Jahre findet, die sich direkt oder indirekt bei "Herbert Mustermann" angesteckt haben.

Aufgabe 3 [Objektrelationale Datenbanken: 15 Punkte]

Welche Erweiterungen bieten objektrelationale Datenbanksysteme im Vergleich zu den relationalen Systemen?

Erklären Sie insbesondere die Begriffe:

Tabellenvererbung

Referenztypen

Path Expression

Nesting und Unnesting

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):

$\sigma_{(A>0 \wedge E<5)}(R1 * (R2 * R3))$

- (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 1000, die Relation R2 5000 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.

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

Aufgabe 5 [XML: 15 Punkte]

Gegeben ist die folgende DTD:

```
<!DOCTYPE Lager [  
  <!ELEMENT Lager (Regal*)>  
  <!ELEMENT Regal (Nummer, Hoehe, Fach*)>  
  <!ELEMENT Fach (Inhalt, Einzelgewicht, Anzahl)>  
  <!ELEMENT Nummer( #PCDATA )>  
  <!ELEMENT Hoehe( #PCDATA )>  
  <!ELEMENT Inhalt( #PCDATA )>  
  <!ELEMENT Einzelgewicht( #PCDATA )>  
  <!ELEMENT Anzahl( #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 Regal-Elemente liefert, die höchstens 10 Fächer haben.

c. (5 Punkte) Schreiben Sie eine XQuery Abfrage, die die Nummern aller Regale sortiert nach Höhe liefert, die zumindest ein Fach haben, in dem das Einzelgewicht des Inhalts 500kg übersteigt.

Aufgabe 6 [Serialisierbarkeit: 10 Punkte]

a. (5 Punkte) Beschreiben Sie das Two-Phase Locking Protokoll und die Varianten Strict Two-Phase Locking und Rigorous Two-Phase Locking

b. (5 Punkte) Verhindert Two-Phase Locking Deadlocks? Wenn ja, beschreiben Sie, warum das so ist. Andernfalls geben Sie ein Beispiel für einen Schedule mit Two-Phase Locking, bei dem ein Deadlock auftritt.