

Datenbanksysteme (f. InformatikerInnen)	schriftliche Einzelprüfung	03.11.2005	1
---	----------------------------	------------	---

Aufgabe 1 [Relationale Abfragen: 25 Punkte]

Gegeben ist folgendes Relationenschema:

buch (titel, verlag, datum, isbn, kategorie, preis)

person (name, email, adresse)

autor (titel, name, honorar)

autor.titel \diamond *buch*, *autor.name* \diamond *person*

buch.kategorie IN {'sachbuch', 'epik', 'lyrik'}

Formulieren Sie die folgende Abfrage in Relationenalgebra:

- a. (8 Punkte) Ermitteln Sie Titel und Verlag aller Bücher, bei denen zumindest 3 Autoren mitgewirkt haben.

Formulieren Sie die folgenden Abfragen in SQL99:

- b. (5 Punkte)

Ermitteln Sie die email-Adresse (email) aller Autoren, die sowohl bei Büchern in der Kategorie 'epik', als auch bei Büchern in der Kategorie 'lyrik' mitgewirkt haben, nicht aber bei Büchern der Kategorie 'sachbuch'.

- c. (7 Punkte)

Ermitteln Sie die ISBN-Nummern (isbn) derjenigen Buchpaare - d.h. jeweils 2 Bücher -, die denselben Verlag und verschiedene Kategorien aufweisen, wobei weiters darauf zu achten ist, dass zumindest eine Person existiert, die bei beiden Büchern als Autor mitgewirkt hat.

- d. (5 Punkte)

Ermitteln Sie Name und Adresse aller Autoren (name,adresse), deren Honorar kleiner ist als die Summe der Preise aller Bücher, an denen sie mitgewirkt haben.

Aufgabe 2 [Datalog: 15 Punkte]

Folgendes Relationenschema ist gegeben:

teil(name, sub_teil),

teil.sub_teil \diamond *teil*.

Die Relation speichert eine Stückliste (also gibt an, aus welchen Teilen ein größeres Teil zusammengesetzt wird).

Leider wurde beim Erfassen der Stückliste ein Fehler gemacht und Ihre Aufgabe ist es jetzt, mittels Datalog-Abfragen alle Teile zu finden, in deren Stückliste das Teil selbst (direkt oder indirekt) wieder als Subteil auftritt.

Aufgabe 3 [Integrität: 15 Punkte]

Welche Mechanismen kennen Sie, um die Datenintegrität einer Datenbank zu gewährleisten. Beschreiben Sie die Konzepte allgemein und geben Sie wo möglich kurze Beispiele in SQL an.

Aufgabe 4 [Query Optimierung: 20 Punkte]

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

$\pi(A, C)(\sigma_{(A>0 \wedge B<5)}(R1 \bowtie_{(A=B)} R2))$

- 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 und
- den umgeformten Ausdruck an.

Datenbanksysteme (f. InformatikerInnen)	schriftliche Einzelprüfung	03.11.2005		2
---	----------------------------	------------	--	---

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

Aufgabe 5 [XML: 15 Punkte]

Gegeben ist die folgende DTD:

```
<!DOCTYPE db [
  <!ELEMENT db (Studierende*)>
  <!ELEMENT Studierende (Name, Studium*)>
  <!ELEMENT Studium (Bezeichnung, BeginnDatum)>
  <!ELEMENT BeginnDatum (Tag, Monat, Jahr)>
  <!ELEMENT Name( #PCDATA )>
  <!ELEMENT Bezeichnung( #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 Studierende liefert, die mehr als 2 Studien betreiben.

c. (5 Punkte)

Schreiben Sie eine XQuery Abfrage, die alle Studierende nach ihren Namen sortiert liefert, deren Studien den "März" als Beginndatum haben.

Aufgabe 6 [Datagrid: 10 Punkte]

a. (5 Punkte) Was sind die Ziele und die Eigenschaften des Gridcomputing?

b. (5 Punkte) Was versteht man unter einem Broker im Grid und warum wird er eingesetzt?