

Prof. Dr. Guido Moerkotte

B6, 29, Raum C0.10

68131 Mannheim

Telefon: (0621) 181-2582

Email: moer@db.informatik.uni-mannheim.de

Datenbanksysteme 1
Frühjahr-/Sommersemester 2010Klausur
18.06.2010

Hinweise

1. Überprüfen Sie bitte Ihr Klausurexemplar auf Vollständigkeit (17 Seiten).
2. Bearbeiten Sie die Aufgaben *ausschließlich* auf dem Aufgabenblatt der jeweiligen Aufgabe.
3. Schreiben Sie auf jedes Blatt, das bewertet werden soll, oben Ihren Namen und Ihre Matrikelnummer.
4. Verwenden Sie nur dokumentenechte Stifte (z. B. keinen Bleistift) und keine roten Stifte.
5. Es sind keine Hilfsmittel zugelassen.
6. Elektronische Geräte am Arbeitsplatz werden als Betrugsversuch betrachtet.

Korrekturzeile

Bitte *nicht* ausfüllen!

Aufgabe	1	2	3	4	5	6	7	Summe
Maximale Punktzahl	10	18	12	22	10	10	8	90
Erreichte Punktzahl								

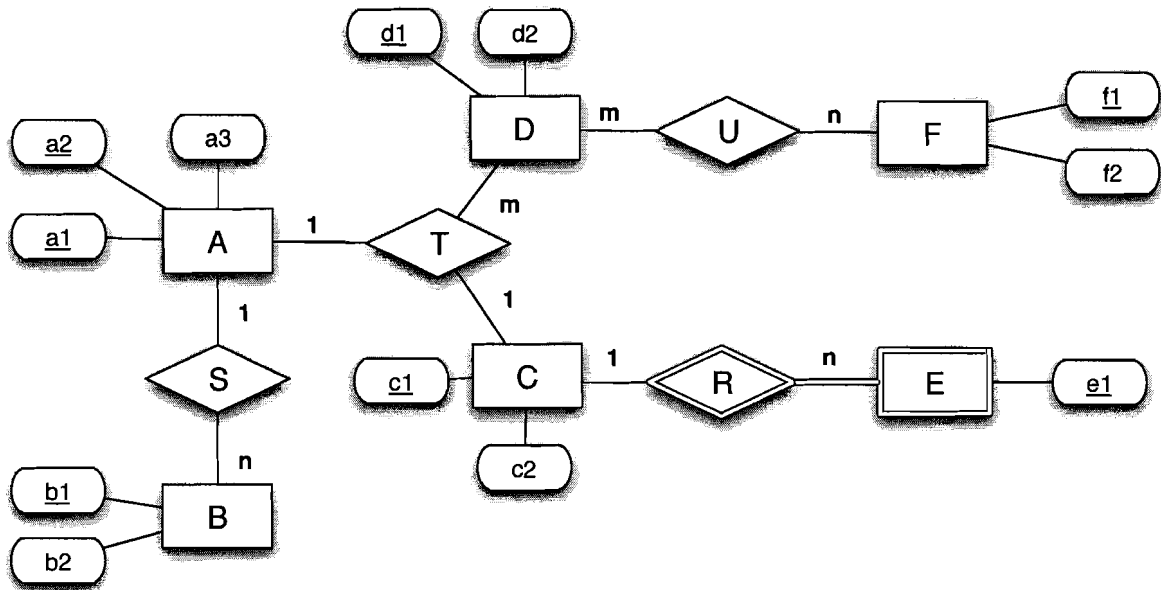
Name:

Matrikelnummer:

Aufgabe 1

10 Punkte

Setzen Sie folgendes ER-Diagramm in ein *verfeinertes*, relationales Schema um.



Name:

Matrikelnummer:

Aufgabe 2

18 Punkte

Ein Zeitschriftenverlag verwaltet seine Geschäftsprozesse mittels eines relationalen Datenbanksystems. Im Folgenden ist ein Ausschnitt des Datenbankschemas gegeben.

- Zeitschrift(zID, titel, rID)
- Autor(aID, name, buero)
- Redakteur(rID, name, buero)
- Kunde(kID, name, adresse)
- Beitrag(aID, zID, titel, anzWoerter)
- erhaelt(kID, zID, preis)

Dabei besitzt jede *Zeitschrift* des Verlages nur einen Redakteur (*rID*). Ein *Autor* kann für mehrere Zeitungen *Beiträge* schreiben. Der *Kunde erhaelt* (durch ein Abonnement) eine oder mehrere *Zeitschriften* des Verlages.

Formulieren Sie die folgenden Anfragen in SQL.

Aufgabe 2 a)

2 Punkte

Für eine neue Werbekampagne wird das gesamte Zeitschriftenangebot des Verlages benötigt. Bestimmen Sie die Titel aller Zeitschriften des Verlages. Die Ergebnisse sollen alphabetisch sortiert ausgegeben werden.

Name:

Matrikelnummer:

Aufgabe 2 b)

3 Punkte

- Zeitschrift(zID, titel, rID)
- Autor(aID, name, buero)
- Redakteur(rID, name, buero)
- Kunde(kID, name, adresse)
- Beitrag(aID, zID, titel, anzWoerter)
- erhaelt(kID, zID, preis)

Der neue Vorstand des Verlages möchte alle Autoren kennenlernen und wünscht sich vor den einzelnen Gesprächen eine Zusammenstellung der Arbeiten der einzelnen Autoren.

Ermitteln Sie die Namen aller Autoren und dazugehörig die Titel aller Zeitschriften, in denen der jeweilige Autor Beiträge veröffentlicht hat. Bestimmen Sie nicht nur den Namen der Zeitschrift, sondern auch die Gesamtzahl aller von dem Autor geschriebenen Wörter pro Zeitschrift. Sortieren Sie das Ergebnis aufsteigend nach den Namen des Autors und absteigend nach der Anzahl der Wörter pro Zeitschrift.

Name:

Matrikelnummer:

Aufgabe 2 c)

7 Punkte

- Zeitschrift(zID, titel, rID)
- Autor(aID, name, buero)
- Redakteur(rID, name, buero)
- Kunde(kID, name, adresse)
- Beitrag(aID, zID, titel, anzWoerter)
- erhaelt(kID, zID, preis)

Das Management des Verlages interessiert sich für die erfolgreichsten Zeitschriften ihres Verlages. Für Erfolg hat das Management zwei Indikatoren:

- der höchste Gesamterlös (Attribut *preis* der Relation *erhaelt* über alle Kunden) und
- die höchste Auflage (Anzahl der Abonnements = Anzahl der Kunden, welche die Zeitschrift *erhalten*).

Bestimmen Sie mit einer Anfrage die Namen der erfolgreichsten Zeitschriften. Hinweis: Das Ergebnis der Anfrage kann aus zwei Datensätzen bestehen, falls die Zeitschrift mit der höchsten Auflage nicht die meisten Einnahmen erzielt.

Ergänzen Sie den Lückentext:

```
with zeitschriften_im_abo as (  
  
select titel ,  
  
as gesamt_erloes,  
  
as anzahl_abo  
  
from zeitschrift z, erhaelt e  
where  
  
)  
select titel  
from zeitschriften_im_abo  
where  
anzahl_abo =  
  
gesamt_erloes =  
  
;
```

Name:

Matrikelnummer:

Aufgabe 2 d)

6 Punkte

- Zeitschrift(zID, titel, rID)
- Autor(aID, name, buero)
- Redakteur(rID, name, buero)
- Kunde(kID, name, adresse)
- Beitrag(aID, zID, titel, anzWoerter)
- erhaelt(kID, zID, preis)

Der Verlag möchte seine wichtigsten Kunden ermitteln, um ihnen einen Treuebonus zukommen zu lassen.

Bestimmen Sie Namen und Adresse aller Kunden, die alle Zeitschriften des Verlages erhalten.

Name:

Matrikelnummer:

Aufgabe 3

12 Punkte

Gegeben sei folgende SQL-Anfrage:

```
select distinct z. titel , b. titel  
from Zeitschrift z, Beitrag b, Autor a  
where z.zId = b.zid and b.aid = a.aid and a.name = '$Michael_Moss'$
```

Aufgabe 3 a)

2 Punkte

Formulieren Sie die SQL-Anfrage in Worten.

Aufgabe 3 b)

4 Punkte

Überführen Sie die SQL-Anfrage in einen *kanonischen* Operatorbaum.

Name:

Matrikelnummer:

Aufgabe 3 c)

6 Punkte

Transformieren Sie das Ergebnis der vorigen Teilaufgabe mittels Heuristiken der logischen Optimierung. Geben Sie den optimierten Operatorbaum an. Berücksichtigen Sie dabei folgende Informationen:

$$\begin{aligned} |Autor| &= 10^3 \\ |Beitrag| &= 10^6 \\ |Zeitung| &= 20 \\ \text{selectivity}(a.name = 'MichaelMoss') &= \frac{1}{1000} \\ \text{selectivity}(b.aID = a.aID) &= \frac{1}{1000} \\ \text{selectivity}(z.zID = b.zID) &= \frac{1}{20} \end{aligned}$$

Name:

Matrikelnummer:

Aufgabe 4

22 Punkte

Gegeben sei ein Relationenschema $\mathcal{R}(A, B, C, D, E, F)$. Lösen Sie jede Teilaufgabe unter Verwendung der jeweiligen funktionalen Abhängigkeiten $\mathcal{F}_{\mathcal{R}}$.

1. Bestimmen Sie alle Kandidatenschlüssel von $\mathcal{R}(A, B, C, D, E, F)$.
2. Bestimmen Sie die höchste Normalform, in der sich \mathcal{R} befindet.
3. Wenn das Schema nicht mindestens in dritter Normalform ist, zerlegen Sie es verlustfrei und abhängigkeitsbewahrend in die dritte Normalform.

Aufgabe 4 a)

$$\mathcal{R}(A, B, C, D, E, F)$$
$$\mathcal{F}_{\mathcal{R}} = \{A \rightarrow BCE, ABF \rightarrow D, C \rightarrow BF, B \rightarrow E\}.$$

Name:

Matrikelnummer:

Aufgabe 4 b)

$$\mathcal{R}(A, B, C, D, E, F)$$
$$\mathcal{F}_{\mathcal{R}} = \{A \rightarrow CE, B \rightarrow AD, ADC \rightarrow BF, CE \rightarrow F, E \rightarrow FC\}.$$

Name:

Matrikelnummer:

Aufgabe 4 c)

$$\mathcal{R}(A, B, C, D, E, F)$$
$$\mathcal{F}_{\mathcal{R}} = \{A \rightarrow BC, B \rightarrow C, C \rightarrow ADE, AD \rightarrow EF\}.$$

Name:

Matrikelnummer:

Aufgabe 4 d)

$$\mathcal{R}(A, B, C, D, E, F)$$
$$\mathcal{F}_{\mathcal{R}} = \{AF \rightarrow BE, CD \rightarrow E, CDF \rightarrow B, CF \rightarrow AE, D \rightarrow F, E \rightarrow C\}.$$

Name:

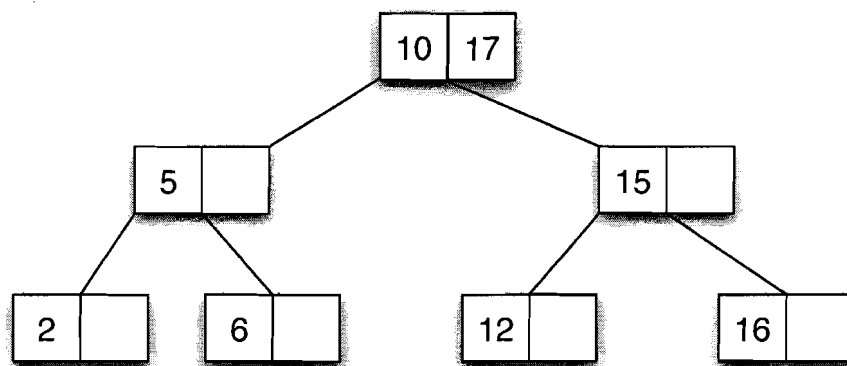
Matrikelnummer:

Aufgabe 5

10 Punkte

Gegeben sei je ein Baum. Geben Sie an, ob es sich dabei um einen gültigen B-Baum handelt. Falls es sich um einen B-Baum handelt, bestimmen Sie den Grad k . Bringen Sie die eingetragenen Schlüssel in eine Reihenfolge, so dass aus einem leeren B-Baum der jeweilige B-Baum entsteht.

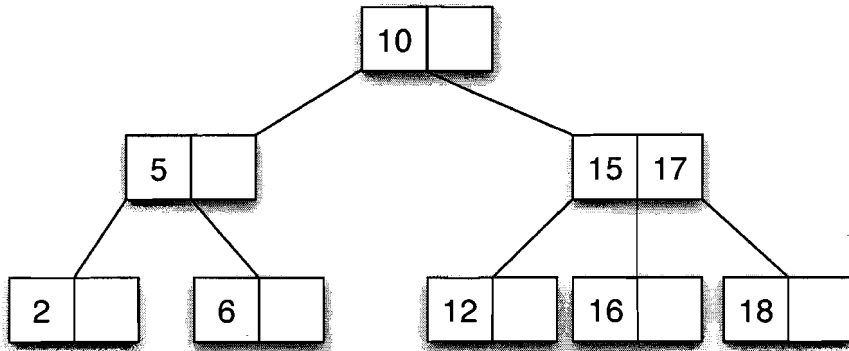
Aufgabe 5 a)



Name:

Matrikelnummer:

Aufgabe 5 b)



Name:

Matrikelnummer:

Aufgabe 6

10 Punkte

Geben Sie für jede Historie an, ob sie serialisierbar ist, ob sie kaskadierendes Rücksetzen vermeidet, rücksetzbar oder strikt ist.

Aufgabe 6 a)

2 Punkte

$r1[x] <_H w3[x] <_H w3[y] <_H r2[y] <_H w2[x] <_H r2[x] <_H c3 <_H c1 <_H c2$

Aufgabe 6 b)

2 Punkte

$r2[x] <_H w3[y] <_H r2[y] <_H w2[y] <_H a2 <_H r3[x] <_H w1[x] <_H c1 <_H w3[x] <_H c3$

Aufgabe 6 c)

2 Punkte

$w1[x] <_H r2[x] <_H w1[y] <_H r3[x] <_H w2[z] <_H w2[y] <_H c1 <_H c2 <_H r3[z] <_H c3$

Aufgabe 6 d)

2 Punkte

$w1[x] <_H w1[y] <_H w2[z] <_H c1 <_H r2[x] <_H r3[x] <_H w2[y] <_H c2 <_H r3[z] <_H c3$

Name:

Matrikelnummer:

Aufgabe 6 e)

2 Punkte

$$r1[y] <_H w1[x] <_H c1 <_H r3[y] <_H w2[y] <_H w3[y] <_H r3[x] <_H r2[x] <_H c3 <_H c2$$

Name:

Matrikelnummer:

Aufgabe 7 8 Punkte

Beantworten Sie folgende Fragen! Eine Begründung ihrer Antwort ist *nicht* notwendig.

Aufgabe 7 a) 1 Punkte

Wofür steht die Abkürzung OLTP?

Aufgabe 7 b) 1 Punkte

Wofür steht die Abkürzung OLAP?

Aufgabe 7 c) 1 Punkte

Nennen Sie einen nicht obsoleten OLTP-Benchmark der Transaction Processing Council.

Aufgabe 7 d) 1 Punkte

Nennen Sie einen nicht obsoleten OLAP-Benchmark der Transaction Processing Council.

Aufgabe 7 e) 1 Punkte

Welche Relationen werden in einem star schema normalisiert und welche denormalisiert?

Aufgabe 7 f) 3 Punkte

Erläutern Sie den Unterschied zwischen einem row store und einem column store.