

## Gruppe A

Bitte tragen Sie **SOFORT** und **LESERLICH** Namen und Matrikelnr. ein und legen Sie Ihren Studentenausweis bereit.

PRÜFUNG AUS DATENBANKSYSTEME (181.038)			18. 3. 1999
Kennnr.	Matrikelnr.	Familienname	Vorname

Arbeitszeit: 120 Minuten. Die Aufgaben sind auf den Angabeblättern zu lösen; Zusatzblätter werden nicht gewertet.

### Aufgabe 1:

(25)

Zu entwerfen ist das Schema für die Datenbank einer Messeverwaltung.

Unterstreichen Sie je Relation einen Schlüssel. Verwenden Sie nur die vorgegebenen Attributnamen. (Diese sind nur bei ihrer jeweils ersten Erwähnung angeführt.) Führen Sie keine zusätzlichen Attribute ein und verwenden Sie möglichst wenige Relationen. Die Datenbank unterstützt keine Nullwerte. Für eine Relation, die nicht in 3NF ist, gibt es keine Punkte!

Die Messe wird von Ausstellern benutzt, um ihre Waren vorzustellen. Jeder Aussteller hat einen Namen (ANAME) und eine Adresse (AADR) sowie eine eindeutige Ausstellernummer (ANR).

Das Messegelände ist unterteilt in Hallen, welche eindeutige Namen haben (HNAME). Jede Halle hat eine Grundfläche (HF) und ein Thema (TH) und ist in eine Menge von Parzellen unterteilt. Außerdem hat sie einen Haupteingang (siehe unten), wo sich der Informationsstand befindet.

Eine Parzelle (zusammenhängende Grundfläche innerhalb der Halle) ist belegt mit einer Reihe von Ständen. Die Parzelle wird mit einem Buchstaben (PB) bezeichnet (eindeutig pro Halle) und hat ebenfalls eine Flächenangabe (PF) sowie eine Stockwerksangabe (ST).

Stände haben eine innerhalb der Halle eindeutige Nummer (SNR). Jedem Stand ist ein Hauptaussteller zugeordnet sowie möglicherweise eine Menge von Nebenausstellern (NANR), wobei für jeden Nebenaussteller sein Prozentanteil an den Gebühren für den Stand bekannt ist (NAPZ).

Über das Messegelände fährt Gratisbus, der auf einer bestimmten Route alle Hallen miteinander verbindet. Die Eingänge aller Hallen haben eine eindeutige Nummer (ENR)(pro Halle), und für den Bus ist bekannt, von welchem Eingang er zum nächsten fährt.

austeller ( <u>ANR</u> , ANAME, AADR	)
halle ( <u>HNAME</u> , HF, TH, ENR	)
parzelle ( <u>PB</u> , <u>HNAME</u> , PF, ST	)
stand ( <u>SNR</u> , <u>HNAME</u> , ANR	)
naussteller ( <u>SNR</u> , <u>HNAME</u> , <u>NANR</u> , NAPZ	)
route ( <u>HNAME</u> , ENR, <u>HNAME1</u> , <u>ENR1</u>	)
(	)
(	)
(	)

**Aufgabe 2:**

(33)

In einer Relationalen Datenbank der FIS werden die Resultate der vergangenen Schiweltcupsaision gespeichert:

<code>sfahr(<u>sname</u>, geb_dat, geschlecht, land)</code>	<code>rennen(<u>rname</u>, typ, länge, land, geschlecht)</code>
<code>gefahren(<u>sname</u>, <u>rname</u>, typ, dg, zeit, platz)</code>	<code>punkte(<u>platz</u>,punkte)</code>

Ein/e FahrerIn wird identifiziert mit dem Namen, beschrieben durch Geburtsdatum, Geschlecht und Herkunftsland. Ein Rennen identifiziert sich durch den Namen und Typ (z.B. Wengen, Slalom), und wird beschrieben durch die Länge der Strecke, dem Land und für welches Geschlecht das Rennen ausgetragen wurde. Die Tabelle gefahren speichert zu jedem Rennen und den darin angetretenen FahrerInnen die Zeit und die erreichte Plazierung. Für Rennen, die nur aus einem Durchgang bestehen, steht im Attribut `dg` der Wert `‘‘G’’`, ansonsten `‘‘1’’`, `‘‘2’’` und `‘‘S’’` für die Gesamtwertung. Im Falle eines Ausscheidens steht in `platz` NULL. Die Relation Punkte speichert die Anzahl der Punkte, die für eine bestimmte Plazierung vergeben werden.

a) Die Länderwertung gibt zu jedem Land und Geschlecht die Summe aller in einer Saison erreichten Punkte an. (Achtung bei Rennen mit 2 Durchgängen.) Schreiben Sie eine Abfrage in SQL, die zu jedem Land das stärkere Geschlecht in der Länderwertung ausgibt gemeinsam mit den Punkten. [10]

```
select s.geschlecht, s.land, sum(punkte)
from sfahr s, gefahren g, punkte p,
where s.sname = g.sname and
      g.platz = p.platz and
      g.dg = ‘‘G’’ or g.dg = ‘‘S’’
group by s.geschlecht, s.land
having sum(punkte) > (select sum(punkte)
                      from sfahr s1, gefahren g1, punkte p1
                      where s1.sname = g1.sname and
                            g1.platz = p.platz and
                            g1.dg = ‘‘G’’ or g1.dg = ‘‘S’’ and
                            s1.land = s.land and
                            g1.geschlecht != s.geschlecht
                      group by s1.geschlecht, s1.land);
```

b) Schreiben Sie eine Abfrage in SQL, die den Namen jener FahrerInnen angibt, die es geschafft haben, sich im 2. Lauf in der Gesamtwertung um mehr als 5 Plätze zu verbessern und wie oft es ihnen gelungen ist. [6]

```
select s.sname, count (*)
from gefahren g1, gefahren g2
where g1.sname = g2.sname and
      g1.rname = g2.rname and
      g1.dg = ‘‘1’’ and g2.dg = ‘‘S’’ and
      g1.platz > g2.platz + 5
group by g1.sname;
```

c) Schreiben Sie eine Abfrage in Relationaler Algebra, die Name, Geburtsdatum und Nationalität jener FahrerInnen ausgibt, die in dieser Saison an allen Rennen (ihres Geschlechts) teilgenommen haben. [9]

$$\pi_{sname, gebdat, land} sfahr \bowtie$$

$$\left[ \left( \pi_{rname, sname} gefahren \div \pi_{rname} \sigma_{geschlecht = "w"} rennen \right) \cup \right. \\ \left. \left( \pi_{rname, sname} gefahren \div \pi_{rname} \sigma_{geschlecht = "m"} rennen \right) \right]$$

d) Schreiben Sie ein stratifiziertes Datalog-Programm, das im Prädikat a(NAME) den Namen jener SchifahrerInnen angibt, die nur Auslandssiege gefeiert haben. [8]

```

sieger(NAME, LAND) :- sfahr(NAME, _, G, _), rennen(R, T, _, LAND, G), gefahren(NAME, R, T, 'G', _, '1').
sieger(NAME, LAND) :- sfahr(NAME, _, G, _), rennen(R, T, _, LAND, G), gefahren(NAME, R, T, 'S', _, '1').

inland_sieger(NAME) :- sieger(NAME, LAND), sfahr(NAME, _, _, LAND).

a(NAME) :- sieger(NAME, _), non inland_sieger(NAME).

```

**Aufgabe 3:** (7)

Gegeben sind die Relationenschemata  $p(ABE)$ ,  $q(CDE)$  und  $r(ABCE)$ .

Optimieren Sie den nachfolgenden algebraischen Ausdruck:

$$\pi_{DE} \sigma_{(A=1 \vee C=4) \wedge D < 1} \left[ \pi_{AB} (r - (p \bowtie r)) \times \pi_{CDE} (p \bowtie q) \right]$$

$$\pi_{DE} \sigma_{(A=1 \vee C=4)} \left[ \pi_{AB} (r - (p \bowtie r)) \times (\pi_{EP} \bowtie \sigma_{D < 1} q) \right]$$

**Aufgabe 4:** (15)

Bestimmen Sie für jedes der Relationenschemata  $R_i(R_i, F_i)$ ,  $i \in \{1, 2, 3\}$ , alle Schlüssel. Geben Sie eine verlustfreie und abhängigkeitsreue Zerlegung in 3NF mit möglichst wenig Relationen an. Unterstreichen Sie in jeder Relation der Zerlegung einen Schlüssel.

Es gelte:  $R_i = ABCDEF$  ,  $i \in \{1, 2, 3\}$

$F_1 = \{ A \rightarrow BD, BD \rightarrow E, D \rightarrow AC \}$ 
 $F_2 = \{ A \rightarrow CD, B \rightarrow CD, CD \rightarrow EF, EF \rightarrow C \}$

$F_3 = \{ C \rightarrow D, E \rightarrow F, F \rightarrow A, E \rightarrow A, F \rightarrow B \}$

RS	Schlüssel	Zerlegung in 3NF (Einen Schlüssel in jeder Relation unterstreichen)			
R1	FA, FD ...	R1-1 <u>DEC</u> .....	R1-2 <u>AB</u> .....	R1-3 <u>FAD</u> .....	R1-4 .....
R2	AB .....	R2-1 <u>CDEF</u> .....	R2-2 <u>ACD</u> .....	R2-3 <u>BCD</u> .....	R2-4 <u>AB</u> .....
R3	CE .....	R3-1 <u>FAB</u> .....	R3-2 <u>EF</u> .....	R3-3 <u>CD</u> .....	R3-4 <u>CE</u> .....

**Aufgabe 5:**

(6)

Gegeben ist das Relationenschema ABCDEF und die Menge **F** von funktionalen Abhängigkeiten. Bestimmen Sie eine minimale Überdeckung.

$$\mathbf{F} = \{ ABC \rightarrow D, AB \rightarrow E, E \rightarrow D, D \rightarrow F, CD \rightarrow F, E \rightarrow D \}.$$

AB  $\rightarrow$  E, D  $\rightarrow$  F, E  $\rightarrow$  D

**Aufgabe 6:**

(7)

Die nachfolgende parallele Ausführung von Transaktionen werde mit Hilfe von Zeitstempeln synchronisiert.

a) Geben Sie die Werte der Zeitstempel an (soweit möglich) und markieren Sie einen Transaktionsabbruch mit '\*'. [4]

T <sub>1</sub>	T <sub>2</sub>	T <sub>3</sub>	a	b	c
80	20	40	RT=WT=0	RT=WT=0	RT=WT=0
	write b			WT= 20	
read b		write a	WT= 40		
		read b		RT= 80	
	write a		WT= 40		
		read c			RT= 40
write a			WT= 80		
		write c			WT= 40
	read c				RT= *
	read b			RT=	
write c					WT= 80

b) Welche Reihenfolge wird zwischen den Transaktionen durch die Zeitstempel festgelegt ( $T_i < T_j \equiv "T_i \text{ vor } T_j"$ )? Annahme: Eine abgebrochene Transaktion wird sofort nach dem Abbruch wieder gestartet. [3]

T<sub>3</sub> < T<sub>1</sub> < T<sub>2</sub>

**Aufgabe 7:**

(7)

a) Kreuzen Sie die richtigen Antworten an (3 richtig angekreuzte Antworten: 2 Punkte) [4]

- ☐ Aufgabe der Recovery ist es, das Cache mit dem Hauptspeicher konsistent zu machen
- ☒ Recovery mit Schattenspeicher benötigt mehr Platz in der Datenbank selbst als mit Logprotokoll
- ☐ Recovery mit Checkpoints bedeutet, daß kein UNDO über den Checkpoint hinaus nötig ist
- ☒ UNDO und REDO sollen idempotent sein

b) Ist der Outer Join eine Erweiterung des Natural Join oder des Equijoin? [1]

Der Outer Join ist eine Erweiterung des Natural Join

c) Warum ist Effektserialisierbarkeit nur hinreichend, aber nicht notwendig für Serialisierbarkeit? [2]

Sie ist notwendig für Serialisierbarkeit, wenn vor jedem Schreibzugriff auch ein Lesezugriff durchgeführt wird (Constraint Write Assumption).

Gesamtpunkte: 100