

Gruppe A

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

PRÜFUNG AUS DATENBANKSYSTEME			16. 10. 1998
Kenn-	Matrikelnr.	Familienname	Vorname

Arbeitszeit: 120 Minuten. Die Aufgaben sind auf den Angabebblätter zu lösen.

Aufgabe 1:

(25)

Zu entwerfen ist das Schema für die Datenbank der US-Luftwaffe zur Verwaltung von UFO-Sichtungen.

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. Für eine Relation, die nicht in 3NF ist, gibt es keine Punkte!

Die Datenbank enthält Berichte über Ereignisse. Ein Bericht wird identifiziert durch eine eindeutige Aktennummer (B#) und beschrieben durch Datum (DATUM), Uhrzeit (UZ), eine Geheimhaltungsstufe (GHS) und eine Beschreibung (BTEXT).

Die Datenbank enthält Informationen über verschiedene Personen. Eine Person wird identifiziert durch die Sozialversicherungsnummer (SVNR) und beschrieben durch ihren Namen (NAME), Geburtsdatum (GD), und Adresse (ADR).

Personen können Mitarbeiter sein - in diesem Fall ist bekannt, ab welcher Geheimhaltungsstufe sie Zugriff auf Dokumente haben. Für jeden Mitarbeiter ist bekannt, an welchen Berichten er mitgearbeitet hat.

Für jeden Bericht und für jeden Zeugen (Person) des Ereignisses gibt es genau eine Zeugenaussage. Die Aussage wird beschrieben durch den Inhalt (ATEXT), das Datum, an dem die Aussage gemacht wurde (ADATUM), und den interviewenden Mitarbeiter (ISVNR).

Weiters enthält die Datenbank folgende Informationen über Radarstationen: ihre eindeutige geographische Position (LAENGE, BREITE), und die Typnummer des Radargeräts (TNR).

Radarstationen liefern Radarmeldungen, die ein Radarbild (BILD) und eine Zeitangabe (MZEIT) enthalten und pro Station mit einer Meldungsnummer (MNR) durchnummeriert sind. Für jede Meldung ist gespeichert, welchen Berichten sie möglicherweise zugeordnet werden kann, sowie für jede Zuordnung, welcher Mitarbeiter die Zuordnung vorgenommen hat.

bericht	(<u>B#</u>	,	DATUM	,	UZ	,	GHS	,	BTEXT)
person	(<u>SVNR</u>	,	NAME	,	GD	,	ADR)		
mitarbeiter	(<u>SVNR</u>	,	GHS)						
bearbeitet	(<u>SVNR</u>	,	<u>B#</u>)						
aussage	(<u>B#</u>	,	<u>SVNR</u>	,	ATEXT	,	ADATUM	,	ISVNR)
radar	(<u>LAENGE</u>	,	<u>BREITE</u>	,	TNR)				
meldung	(<u>LAENGE</u>	,	<u>BREITE</u>	,	<u>MNR</u>		BILD	,	MZEIT)
zuordnung	(<u>LAENGE</u>	,	<u>BREITE</u>	,	<u>MNR</u>	,	<u>B#</u>	,	SVNR)
	()

Aufgabe 2:

(32)

In einer Relationalen Datenbank der Wiener Verkehrsbetriebe werden Informationen über Fahrplan und Fuhrpark gespeichert:

<code>fahrzeug(<u>fznr</u>, baujahr, typ)</code>	<code>linie(<u>lnr</u>, b_von, b_bis, s_von, s_bis)</code>
<code>einsatz(<u>fznr</u>,<u>lnr</u>,<u>datum</u>)</code>	<code>strecke(<u>station1</u>, <u>station2</u>, <u>lnr</u>, fzeit)</code>

Ein Fahrzeug wird identifiziert durch eine Fahrzeugnummer (**fznr**) und beschrieben durch das Baujahr (**baujahr**) und seinen Typ (**typ**) i.e. Tram, Bus, UBahnzug. Eine Linie (z.B. 13A) identifiziert sich durch die Liniennummer und speichert Information über die Betriebszeit (**b_von**, **b_bis**) als auch über Anfangs- und Endstation (**s_von**, **s_bis**). Die einzelnen Teilstrecken einer Linie und die für diese Teilstrecke benötigte Fahrtzeit sind in der Relation **strecke** gespeichert. Außerdem steht in der Relation **einsatz**, an welchem Tag welches Fahrzeug auf welcher Linie eingesetzt ist.

Schreiben Sie eine Abfrage in SQL, die die Liniennummer, Start- und Endstation jener Straßenbahnen angibt, deren Linien nach 23 Uhr noch in Betrieb sind. [7]

```
select distinct l.lnr, l.s_von, l.s_bis
from linie l, einsatz e, fahrzeug f
where l.b_bis > 23 and
      f.typ = Strassenbahn and
      l.lnr = e.lnr and
      e.fznr = f.fznr;
```

Schreiben Sie eine Abfrage in SQL, die die Linie ausgibt, die die meisten Stationen hat, und die Anzahl der auf dieser Linie heute eingesetzten Fahrzeuge. [10]

```
select l.lnr, count(*) as anzahl
from linie l, einsatz e
where l.lnr = e.lnr and
      e.datum = today and
      l.lnr = (select s.lnr
               from station s
               group by s.lnr
               having count(*) >= all (select count(*)
                                       from station
                                       group by lnr))
group by l.lnr;
```

c) Schreiben Sie eine Abfrage in Relationaler Algebra, die alle Linien (deren Nummer, Start und Endstation) angibt, die auf ihrem Weg nicht am Karlsplatz vorbeikommen. [7]

$$\pi_{lnr,s_von,s_bis} linie - \pi_{lnr,s_von,s_bis} (linie \bowtie \sigma_{(station1 \vee station2)=Karlsplatz} strecke)$$

d) Schreiben Sie ein stratifiziertes Datalog-Programm, das im Prädikat **ges_fzeit**(LNR,FZEIT) zu jeder Linie die Fahrzeit ausgibt, die für die gesamte Strecke benötigt wird. Gehen Sie davon aus, daß der Aufenthalt in den Stationen nicht berücksichtigt wird. Das Prädikat **sum**(X,Y,Z) ($X = Y + Z$) steht zur Verfügung. [8]

```
fahrt(LNR,S1,S2,FZEIT):- linie(LNR,B1,B2,S1,S), strecke(S1,S2,LNR,FZEIT).
fahrt(LNR,S1,S2,FZEIT):- fahrt(LNR,S1,S,FZ1), strecke(S,S2,LNR,FZ2)
                           sum(FZEIT,FZ1,FZ2).

ges_fzeit(LNR,FZEIT)    :- fahrt(LNR,S1,S2,FZEIT), linie(LNR,B1,B2,S1,S2),
```

Aufgabe 3:

Gegeben sind die Relationenschemata $p(ABCD)$, $q(AC)$ und $r(BCD)$.

Optimieren Sie den nachfolgenden algebraischen Ausdruck:

$$\pi_{CD} \sigma_{C=1 \vee D=1} [(p \bowtie q) \cup (p - (p \bowtie r))]$$

$$\pi_{CD} (\pi_{ACD} \sigma_{C=1 \vee D=1} p \bowtie q) \cup \pi_{CD} [\sigma_{C=1 \vee D=1} p - (\sigma_{C=1 \vee D=1} p \bowtie \sigma_{C=1 \vee D=1} r)]$$

Aufgabe 4:

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

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

$$F1 = \{ D \rightarrow ACF, B \rightarrow E, C \rightarrow B \} \quad F2 = \{ BCD \rightarrow E, CD \rightarrow A \} \quad F3 = \{ B \rightarrow E, A \rightarrow F, EF \rightarrow CD \}$$

RS	Schlüssel	Zerlegung in 3NF (Einen Schlüssel in jeder Relation unterstreichen)					
R1	D	R1-1 <u>BE</u>	R1-2 <u>CB</u>	R1-3 <u>D</u> ACF	R1-4		
R2	BCDF	R2-1 <u>BCDE</u>	R2-2 <u>CDA</u>	R2-3 <u>BCDF</u>	R2-4		
R3	AB	R3-1 <u>EFCD</u>	R3-2 <u>AF</u>	R3-3 <u>BE</u>	R3-4 <u>AB</u>		

Aufgabe 5: (6)
Gegeben ist das Relationenschema ABCDEF und die Menge **F** von funktionalen Abhängigkeiten. Bestimmen Sie eine minimale Überdeckung.

$$\mathbf{F} = \{ A \rightarrow C, B \rightarrow A, C \rightarrow DE, B \rightarrow D, DE \rightarrow F, E \rightarrow F \}.$$

$$A \rightarrow C, B \rightarrow A, C \rightarrow DE, E \rightarrow F$$

Aufgabe 6: (7)
Die nachfolgende Parallelausführung 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_1	T_2	T_3	a	b	c
10	20	30	RT=WT=0	RT=WT=0	RT=WT=0
write b			WT= 20	WT= 10	
	write a	read b		RT= 30	
write a	read b		WT= 20	RT= 30	
	read c				RT= 20
		write a	WT= 30		
read c	write c				WT= 20
read b					RT= *
		write c		RT=	WT= 30

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]

$$T_2 < T_3 < T_1$$

Aufgabe 7: (8)
Gegeben ist ein Datenbankschema `person(name, stadt, strasse, nummer)` und `haus(stadt, strasse, nummer, wert)`.

a) Welcher Ausdruck der relationalen Algebra entspricht dem folgenden Datalog-Prädikat: [4]
 $p(N,W) \text{ :- } person(N,S1,S2,Nr), \text{ haus}(S1,S2,Nr,W).$

$$\pi_{name, wert}(person \bowtie haus)$$

b) Zeichnen Sie $DEP(person, haus, p)$. [2]

c) Vervollständigen Sie: [2]

Das Datalogprogramm ist stratifiziert, weil keine Negation vorkommt.

Gesamtpunkte: 100