

Gruppe A

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

PRÜFUNG AUS DATENBANKSYSTEME (181.038)			13. 3. 2002
Kennnr.	Matrikelnr.	Familienname	Vorname

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

Aufgabe 1: (20)

Modellieren Sie folgende Datenbank, die eine Reihe von Gebäuden verwaltet. Unterstreichen Sie je Relation einen Schlüssel. Verwenden Sie nur die vorgegebenen Attributnamen. (Diese sind üblicherweise 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!

Jedes Gebäude hat einen Namen (NAME) und eine eindeutige Adresse bestehend aus Straßennamen/-nummer (STR) und Postleitzahl (PLZ) sowie eine Reihe von Geschossen. Diese werden durch eine Nummer (NR) bezeichnet, wobei 0 für das Erdgeschoß steht, positive Zahlen für Obergeschoße, und negative Zahlen für Kellergeschoße.

Für jedes dieser Geschosse speichern wir die max. Belastung (N_PER_KG) und sämtliche Räume in diesem Geschöß. Jeder Raum hat dabei eine auf das jeweilige Geschöß bezogen eindeutige Nummer (RNR) sowie eine Größe in m² (M2).

Weiters wird für jeden Raum die Anzahl der Fenster (FENSTER) gespeichert sowie eine Liste von Möbelstücken, wobei jedes Möbelstück eine eindeutige Inventarnummer (INV), eine Kurzbeschreibung (KURZ) und einen Schätzwert (WERT) hat. Für jeden Büroraum speichern wir zusätzlich noch die Anzahl an Telefon- bzw. Datendosen (TELDOS, DATDOS).

Lifte haben eine vom jeweiligen Hersteller vergebene und auf diesen Hersteller bezogen eindeutige Nummer (LNR) und wir notieren für jeden Lift das unterste (UNTEN) sowie das oberste (OBEN) Stockwerk, das er bedient. Von den Liftherstellern speichern wir ihren eindeutigen Herstellercode (HERST), den Namen (HNAME) und die Notfalltelefonnummer (TEL).

gebäude (<u>NAME</u> , <u>STR</u> , <u>PLZ</u>) [3]
geschoß (<u>NR</u> , <u>STR</u> , <u>PLZ</u> , N_PER_KG) [3]
raum (<u>NR</u> , <u>STR</u> , <u>PLZ</u> , <u>RNR</u> , M2, FENSTER) [3]
möbelstück (<u>INV</u> , KURZ, WERT, NR, STR, PLZ, RNR) [4]
büroraum (<u>NR</u> , <u>STR</u> , <u>PLZ</u> , <u>RNR</u> , TELDOS, DATDOS) [3]
lift (<u>HERST</u> , <u>LNR</u> , UNTEN, OBEN) [2]
hersteller (<u>HERST</u> , HNAME, TEL) [2]
()
()
()

In der relationalen Datenbank einer Kaufhauskette werden folgende Informationen gespeichert:

<code>ware(<u>wnr</u>, name, einkaufspreis, verkaufspreis)</code>	<code>einkauf(<u>wnr</u>, <u>fnr</u>, <u>datum</u>, anzahl)</code>
<code>filiale(<u>fnr</u>, adresse, m2, jahresfixkosten)</code>	<code>verkauf(<u>wnr</u>, <u>fnr</u>, <u>datum</u>, anzahl)</code>

Jede Ware wird durch ihre Warenummer (kurz wnr) eindeutig identifiziert und durch ihren Namen, den Einkaufspreis (zu dem die Kaufhauskette die Ware bei ihren Lieferanten einkauft) und den Verkaufspreis (zu dem sie ihren Kunden die Ware verkauft) beschrieben. Alle Filialen haben eine eindeutige Filialnummer (fnr), eine Adresse, die Größe der Verkaufsfläche in Quadratmetern (m2) und Fixkosten pro Jahr. Wenn an einem Tag kein Einkauf oder Verkauf stattfand, so gibt es für diesen Tag keinen entsprechenden Eintrag (d.h. anzahl > 0).

a) Schreiben Sie eine Abfrage in SQL, die die Nummern und Adressen von Filialen selektiert, bei denen im Jahr 2001 die Brutto-Einnahmen (Summe über Verkaufspreis \times Anzahl) pro Quadratmeter über dem Durchschnitt (aller Filialen) der Brutto-Ausgaben (Fixkosten + Summe über Einkaufspreis \times Anzahl) pro Quadratmeter lag.

[10]

```
select f.fnr, f.adresse
from ware w, filiale f, verkauf v
where w.wnr = v.wnr and f.fnr = v.fnr and
      v.datum between '1.1.2001' and '31.12.2001'
group by f.fnr, f.adresse, f.m2
having sum(w.verkaufspreis * v.anzahl)/f.m2 >
      ( select avg(ausgabenproquadrat)
        from ( select (sum(w1.einkaufspreis * e.anzahl)
                      + f1.jahresfixkosten)/f1.m2 as ausgabenproquadrat
                from ware w1, filiale f1, einkauf e
                where w1.wnr = e.wnr and f1.fnr = e.fnr and
                      e.datum between '1.1.2001' and '31.12.2001'
                group by f1.fnr, f1.m2, f1.jahresfixkosten ) );
```

b) Schreiben Sie eine Abfrage in SQL, die die Nummern und Namen von Waren sowie bestimmte Verkaufstage für diese Waren ausgibt, so dass am jeweiligen Tag in *jeder* Filiale zumindest 100 Stück davon verkauft wurden.

[10]

```
select distinct w.wnr, w.name, v.datum
from ware w, verkauf v
where w.wnr = v.wnr and not exists ( select f.fnr
                                     from filiale f
                                     where f.fnr not in ( select v2.fnr
                                                           from verkauf v2
                                                           where v.datum = v2.datum and
                                                                 v.wnr = v2.wnr and
                                                                 v2.anzahl >= 100 ) );
```

Alternative:

```
select w.wnr, w.name, v.datum
from ware w, verkauf v
where w.wnr = v.wnr and v.anzahl >= 100
group by w.wnr, w.name, v.datum
having count(v.fnr) = (select count(*) from filiale);
```

c) Schreiben Sie dieselbe Abfrage wie in b) in Relationaler Algebra. [8]

Abkürzungen:

$X = \text{ware} \bowtie \text{verkauf}$

$Y = \pi_{f_{nr}}(\text{filiale})$

$Z = \sigma_{anzahl \geq 100}(X)$

Resultat: $\pi_{f_{nr}, wnr, name, datum}(Z) \div Y =$

$\pi_{f_{nr}, wnr, name, datum}(\sigma_{anzahl \geq 100}(\text{ware} \bowtie \text{verkauf})) \div \pi_{f_{nr}}(\text{filiale})$

Aufgabe 3: (12)

Bestimmen Sie für jedes der Relationenschemata $R_i(R_i, F_i)$, $i \in \{1, 2\}$, 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: $R_i = MNOPQR$, $i \in \{1, 2\}$

$F_1 = \{ M \rightarrow QR, N \rightarrow OM, R \rightarrow NP \}$

$F_2 = \{ R \rightarrow NO, M \rightarrow QR, ORQ \rightarrow N \}$

RS	Schlüssel	Zerlegung in 3NF (Einen Schlüssel in jeder Relation unterstreichen)			
R1	M, N, R ..	R1-1 <u>MNOPQR</u> ...	R1-2	R1-3	R1-4
R2	MP	R2-1 <u>RNO</u>	R2-2 <u>MQR</u>	R2-3 <u>MP</u>	R2-4

Aufgabe 4: (6)

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

$F = \{ T \rightarrow U, TS \rightarrow Q, Q \rightarrow TS, PQ \rightarrow R, Q \rightarrow S, U \rightarrow SU \}$

$T \rightarrow U, T \rightarrow Q, Q \rightarrow T, PQ \rightarrow R, U \rightarrow S$

Aufgabe 5:

(8)

Die nachfolgende Parallelausführung von Transaktionen wird mithilfe von Zeitstempeln synchronisiert.

a) Geben Sie die Werte der Zeitstempel an (soweit möglich); markieren Sie einen Transaktionsabbruch mit '*'. [5]

T_1	T_2	T_3	a	b	c
99	20	5	RT=WT=0	RT=WT=0	RT=WT=0
		read a	RT= 5		
write b				WT= 99	
	write c				WT= 20
		write c			WT= 20
	write b			WT= 99	
read b				RT= 99	
	write b			WT= *	
		write a	WT= 5		
	read a		RT=		

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

$$T_3 < T_1 < T_2$$

Aufgabe 6:

(6)

Welche der beiden Operationen Projektion und Selektion ist im Sinne der Optimierung von Ausdrücken in Relationaler Algebra, wie wir sie kennengelernt haben, "teurer" und sollte daher wenn möglich *nach* der anderen dieser beiden Operationen durchgeführt werden?

Begründen Sie Ihre Antwort!

Die Projektion ist teurer als die Selektion, da wir hier Duplikate eliminieren müssen bzw. führt letztere tendenziell zu kleineren Zwischenergebnissen.

Gesamtpunkte: 80