

Hauptdiplomklausur Datenbanksysteme I Sommersemester 2002

Name:
Vorname:
Matrikelnummer:
Studienfach:

Wichtige Hinweise:

1. Prüfen Sie Ihr Klausurexemplar auf Vollständigkeit (15 Seiten).
2. Es sind keine Hilfsmittel zugelassen.
3. Die Klausur dauert 100 Minuten.
4. Jede Aufgabe ist auf dem zugehörigen Aufgabenblatt (und ggf. auf separaten Lösungsblättern) zu bearbeiten.
5. Vermerken Sie Ihren Namen und Ihre Matrikelnummer auf jedem Aufgaben- (bzw. Lösungsblatt). Blätter ohne Namens- und Matrikelnummerangabe werden nicht bewertet.
6. Das Deckblatt sowie alle Aufgabenblätter (evtl. Lösungsblätter) sind abzugeben.

	maximale Anzahl Punkte	erreichte Anzahl Punkte
Aufgabe 1	5	
Aufgabe 2	15	
Aufgabe 3	5	
Aufgabe 4	8	
Aufgabe 5	12	
Aufgabe 6	15	
Aufgabe 7	12	
Aufgabe 8	12	
Aufgabe 9	8	
Aufgabe 10	8	
	100	

1. (je 1 Punkt)

Geben Sie für folgende Aussagen an, ob sie wahr oder falsch sind.

Bei einer falschen Antwort werden Punkte abgezogen, die Gesamtpunktzahl kann jedoch nicht unter 0 Punkte sinken.

	wahr	falsch
Jede strikte Historie ist auch rücksetzbar.	<input type="checkbox"/>	<input type="checkbox"/>
Jede serialisierbare Historie ist auch rücksetzbar.	<input type="checkbox"/>	<input type="checkbox"/>
Wenn die gleichen Daten mit einem B -Baum und mit einem B^+ -Baum indexiert werden, dann ist die Höhe des B^+ -Baums immer kleiner gleich der Höhe des B -Baums.	<input type="checkbox"/>	<input type="checkbox"/>
Jeder Superschlüssel ist auch ein Schlüssel	<input type="checkbox"/>	<input type="checkbox"/>
Eine Zerlegung von \mathcal{R} in \mathcal{R}_1 und \mathcal{R}_2 ist verlustlos $\Leftrightarrow (\mathcal{R}_1 \cap \mathcal{R}_2) \rightarrow \mathcal{R}_1$ oder $(\mathcal{R}_1 \cap \mathcal{R}_2) \rightarrow \mathcal{R}_2$	<input type="checkbox"/>	<input type="checkbox"/>

2. (a) (5 Punkte)

Gegeben sei ein Relationenschema $\mathcal{R}(A, B, C, D, E, F)$ mit folgenden funktionalen Abhängigkeiten $\mathcal{F}_{\mathcal{R}} = \{AC \rightarrow D, BE \rightarrow A, C \rightarrow F\}$. In welcher höchsten Normalform befindet sich \mathcal{R} ?

(b) (5 Punkte)

Gegeben sei ein Relationenschema $\mathcal{S}(A, B, C, D, E, F)$ mit folgenden funktionalen Abhängigkeiten $\mathcal{F}_{\mathcal{S}} = \{A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow AF, BC \rightarrow E\}$. In welcher höchsten Normalform befindet sich \mathcal{S} ?

(c) (5 Punkte)

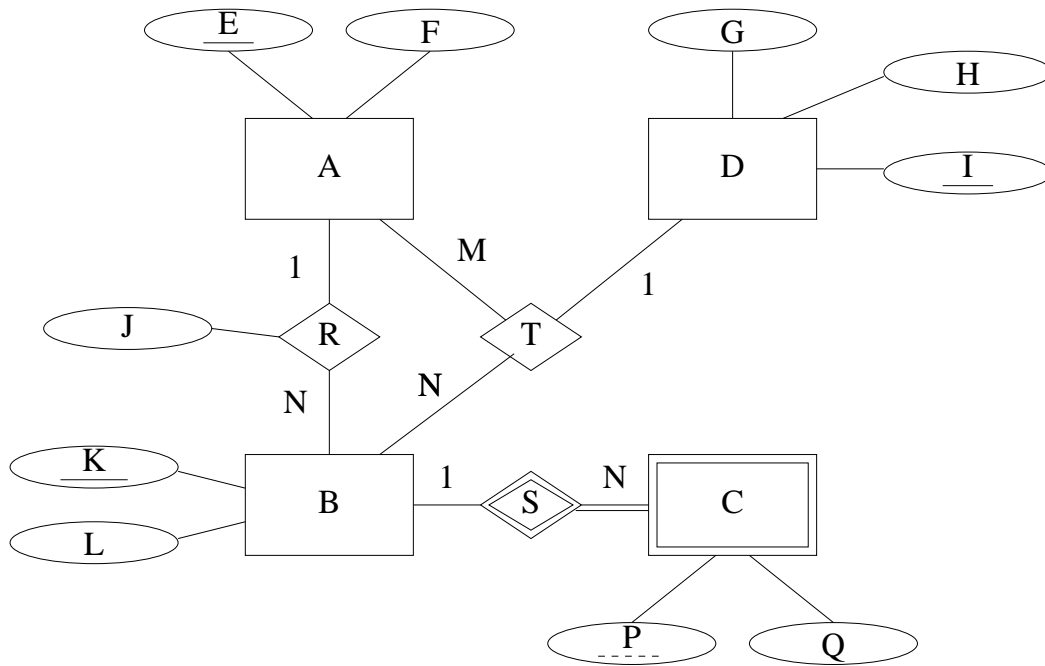
Gegeben sei ein Relationenschema $\mathcal{T}(A, B, C, D, E, F)$ mit folgenden funktionalen/mehrwertigen Abhängigkeiten $\mathcal{F}_{\mathcal{T}} = \{A \rightarrow BE, B \rightarrow CD, E \rightarrow F, ABC \twoheadrightarrow DEF\}$. In welcher höchsten Normalform befindet sich \mathcal{T} ?

3. (5 Punkte)

Zeigen Sie: jedes Relationenschema in 3NF ist auch in 2NF.

4. (8 Punkte)

Setzen Sie folgendes E-R-Diagramm in ein relationales Schema um:



5. Auf der Relation *Personal* sind folgende Trigger definiert:

```
create trigger mehrgehalt
after update on Personal
for each row
when (old.Abteilung = 'Vorstand')
begin
    set new.Gehalt = old.Gehalt;
end
```

```
create trigger abteilungda
before insert on Personal
for each row
begin
    if new.Abteilung not in
        (select Name
         from Abteilung)
    then
        -- Abbruch mit Fehlermeldung
        print(Fehler!)
        exit(-1)
    end if;
end
```

```
create trigger abteilungweg
after delete on Personal
for each row
begin
    if (select count(*)
        from Personal
        where Abteilung = old.Abteilung) = 0
    then
        (delete from Abteilung
         where Name = old.Abteilung)
    end if;
end
```

Die Relationen *Personal* und *Abteilung* beinhalten folgende Tupel:

Personal				Abteilung	
<u>PersNr</u>	Name	Gehalt	Abteilung	<u>Name</u>	Ort
1	Liese Müller	50.000	Personal	Finanzen	Frankfurt
2	Fritz Schmidt	40.000	Finanzen	Personal	Mainz
3	Karl-Heinz Meier	45.000	Finanzen	Vorstand	Frankfurt
4	Jane Doe	150.000	Vorstand		
5	Johann Schröder	55.000	Personal		

(a) (je 3 Punkte)

Folgende Operationen werden auf der Relation *Personal* ausgeführt. Geben Sie jeweils an, zu welchem Ergebnis die jeweilige Operation führt. Gehen Sie davon aus, daß alle Operationen auf den ursprünglichen Inhalt von *Personal* und *Abteilung* (wie oben angegeben) angewendet werden.

i. delete from Personal where PersNr = 4;

ii. insert into Personal values(6, 'Karl König', 60.000, 'Technik');

iii. update Personal set Gehalt = 56.000 where PersNr = 1;

(b) (3 Punkte)

Mit welchen anderen Mitteln kann der Trigger *abteilungda* realisiert werden? Welche Zeilen müssen dafür in den Create-Table-Befehlen der Relationen *Personal* bzw. *Abteilung* hinzugefügt werden?

6. Gegeben sei wieder das relationale Schema aus Aufgabe 5:

- Personal(PersNr, Name, Gehalt, Abteilung)
- Abteilung(Name, Ort)

Das Attribut *Abteilung* in der Relation *Personal* ist ein Fremdschlüssel und bezieht sich auf das Attribut *Name* in der Relation *Abteilung*.

Formulieren Sie folgende Anfragen in SQL:

(a) (6 Punkte)

Geben Sie die Personalnummern und Namen aller Personen an, die in einer Abteilung in Frankfurt oder in Mannheim arbeiten und mehr als 80.000 verdienen.

(b) (9 Punkte)

Geben Sie die Personalnummer und den Namen derjenigen Person an, deren Gehalt den Medianwert bildet, d.h. deren Gehalt genau in der Mitte liegt. Um das ganze zu vereinfachen, nehmen Sie an, daß es eine ungerade Anzahl von Personen gibt und daß alle Gehälter verschieden sind (d.h. gesucht ist die Person für die die Anzahl der Leute die mehr verdienen = Anzahl der Leute die weniger verdienen).

7. Folgende Einträge befinden sich in der Log-Datei eines Datenbanksystems als es zum Absturz kommt:

19	T_1	-	BOT	-	0
20	T_2	-	BOT	-	0
21	T_1	P_B	$B:=B+10$	$B:=B-10$	19
22	T_2	P_A	$A:=A-20$	$A:=A+20$	20
23	T_2	-	COMMIT	-	22

Die Seiten auf denen die Datenelemente A und B gespeichert sind, sehen folgendermaßen aus:

P_A	P_B
$\overline{\text{LSN}=15}$ $A = 28$	$\overline{\text{LSN}=21}$ $B = 12$

Der Wiederanlauf wird angestoßen. Gehen Sie davon aus, daß es während des Wiederanlaufs keine weiteren Abstürze des Systems gibt.

- (a) (2 Punkte)

Was ist das Ergebnis der Analyse-Phase?

- (b) (5 Punkte)

Führen Sie die Redo-Phase durch (d.h. geben Sie die konkreten Schritte an).

(c) (5 Punkte)

Führen Sie die Undo-Phase durch (geben Sie auch hier die konkreten Schritte an).

8. Welche Eigenschaften haben die folgenden Historien jeweils (seriell, serialisierbar, rücksetzbar, vermeidet kaskadierendes Rücksetzen (ACA), strikt, keine der vorherigen)?

(a) (4 Punkte)

$r_3[x] \ w_1[x] \ r_3[y] \ w_2[x] \ c_1 \ c_2 \ r_3[x]$

(b) (4 Punkte)

$r_1[x] \ r_2[y] \ w_1[x] \ c_1 \ r_2[x] \ w_2[z] \ c_2$

(c) (4 Punkte)

$r_1[x] \ w_1[z] \ r_2[x] \ r_2[z] \ w_1[y] \ w_2[y] \ c_2 \ c_1$

9. (a) (3 Punkte)

Was sind die drei grundlegenden Korrektheitsanforderungen an eine Fragmentierung in einem verteilten Datenbanksystem?

(b) (5 Punkte)

Da in einem verteilten Datenbanksystem Tupel von einer Station auf eine andere „umziehen“ können, wurden die TIDs folgendermaßen modifiziert:

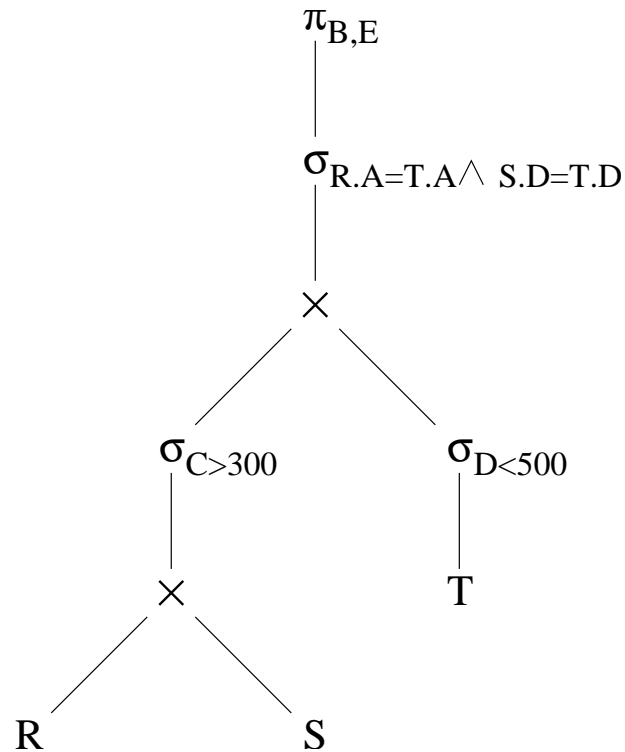
TID

ServerID	SeitenNr	SlotNr
----------	----------	--------

Zusätzlich zur Seitennummer und Slotnummer gibt es jetzt noch eine ServerID. Was ist das Hauptproblem an diesem Verfahren in bezug auf die Performance?

10. (8 Punkte)

Der Optimierer eines Datenbanksystems ist noch nicht ganz ausgereift und produziert für eine Anfrage folgenden Operatorbaum:



Die Relationen besitzen folgendes Schema:

- $R(A, B, C)$
- $S(D, E, F)$
- $T(A, D)$

Weiterhin kennt das Datenbanksystem folgende Statistiken über die Relationen:

Größe von R : 100 Tupel

Größe von S : 10.000 Tupel

Größe von T : 100.000 Tupel

Selektivität von $\sigma_{C>300} = \frac{1}{10}$

Selektivität von $\sigma_{D<500} = \frac{1}{10}$

Selektivität von $R \bowtie T = \frac{1}{1000}$

Selektivität von $S \bowtie T = \frac{1}{1000}$

Verbessern Sie den oben abgebildeten Operatorbaum.

Name: _____ Matr.Nr.: _____