

Hauptdiplomklausur

Datenbanksysteme I

Wintersemester 2002/2003

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	10	
Aufgabe 2	15	
Aufgabe 3	4	
Aufgabe 4	20	
Aufgabe 5	15	
Aufgabe 6	14	
Aufgabe 7	9	
Aufgabe 8	13	
	100	

1. (je 2 Punkte)

Folgende Aussagen sind fehlerhaft. Geben Sie jeweils an, wie die Aussage richtig lauten muß. (Eine einfache Verneinung der Aussage gibt keine Punkte!)

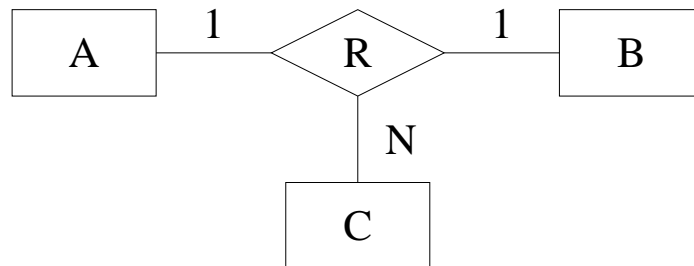
(a) Mit logischer Datenunabhängigkeit bezeichnet man die Abschirmung von Benutzern und Anwendungen von Änderungen an der physischen Speicherung der Datenbank.

(b) Das Ergebnis des Implementationsentwurfs ist ein E-R-Schema.

(c) Ein Relationenschema ist in 2NF, wenn alle Nichtschlüsselattribute (NSA) funktional von allen Kandidatenschlüsseln abhängen.

(d) Beim Zweiphasen-Sperrprotokoll (2PL) darf eine Transaktion nach Freigabe einer Sperre auf einem Datenobjekt nur noch Sperren auf anderen Datenobjekten anfordern.

- (e) Bei einer dreiwertigen N:1:1 Beziehung (siehe Abbildung unten) werden bei der Umsetzung in ein relationales Schema die Schlüssel der Entitäten der 1-Zweige (A und B) als Fremdschlüssel in Relation der Entität des N-Zweiges (C) übernommen.



2. (a) (5 Punkte)

Gegeben sei ein Relationenschema $\mathcal{R}(A, B, C, D, E, F)$ mit folgenden funktionalen/mehrwertigen Abhängigkeiten $\mathcal{F}_{\mathcal{R}} = \{AE \rightarrow BCDF, F \twoheadrightarrow A\}$. Geben Sie zunächst alle Kandidatenschlüssel des Schemas an. 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 $\mathcal{F}_{\mathcal{S}} = \{A \rightarrow B, B \rightarrow CD, C \rightarrow A, D \rightarrow E, E \rightarrow CF, F \rightarrow D\}$. Geben Sie zunächst alle Kandidatenschlüssel des Schemas an. 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 Abhängigkeiten $\mathcal{F}_{\mathcal{T}} = \{AB \rightarrow E, A \rightarrow D, C \rightarrow F\}$.

Geben Sie zunächst alle Kandidatenschlüssel des Schemas an. In welcher höchsten Normalform befindet sich \mathcal{T} ? Zerlegen Sie das Schema in 3NF (gehen Sie dabei davon aus, daß $\mathcal{F}_{\mathcal{T}}$ eine kanonische Überdeckung ist).

3. (a) (2 Punkte)

Welche Wirkung hat die Klausel

...

`A integer references R on update cascade`

...

im `create table` Befehl der Relation S beim Ändern von Tupeln in Relation R ?

(b) (2 Punkte)

Welche Wirkung hat die Klausel

...

`A integer references R on delete set null`

...

im `create table` Befehl der Relation S beim Löschen von Tupeln in Relation R ?

4. Gegeben sei folgender Auszug aus dem relationalen Schema zur Verwaltung einer Cart-Bahn.

- Kunde(KundenNr, Name, Adresse, TelNr)
- Fahrzeug(FzNr, Typ, letzteWartung)
- Bahn(Name, Ort)
- gemietet(AuftragsNr, KundenNr, BahnNr, Datum, von, bis)
- mitFz(AuftragsNr, FzNr)

Die Attribute *KundenNr* und *BahnNr* in der Relation *gemietet* sind Fremdschlüssel und beziehen sich auf das Attribut *KundenNr* in der Relation *Kunde* beziehungsweise auf das Attribut *BahnNr* in der Relation *Bahn*. Die Attribute *von* und *bis* geben an, von welcher Uhrzeit bis zu welcher Uhrzeit der Kunde die Bahn gemietet hat. Gehen Sie davon aus, daß Sie bei der Subtraktion von Datumsangaben (bzw. Uhrzeiten) voneinander die Ergebnisse in Tagen (bzw. Stunden) angegeben werden. In der Relation *mitFz* sind beide Attribute Fremdschlüssel, die sich auf die Relationen *gemietet* und *Fahrzeug* beziehen.

Formulieren Sie folgende Anfragen in SQL:

(a) (6 Punkte)

Prüfen Sie, ob ein Fahrzeug existiert, daß vermietet wurde, obwohl die letzte Wartung zum Mietzeitpunkt mehr als ein Jahr (365 Tage) zurückliegt.

Kunde(KundenNr, Name, Adresse, TelNr)

Fahrzeug(FzNr, Typ, letzteWartung)

Bahn(Name, Ort)

gemietet(AuftragsNr, KundenNr, BahnNr, Datum, von, bis)

mitFz(AuftragsNr, FzNr)

(b) (6 Punkte)

Geben Sie die Gesamtzahl der Stunden an, die der Kunde mit der KundenNr 123 im Zeitraum vom '01.10.2002' bis zum '31.10.2002' Bahnen gemietet hatte.

Kunde(KundenNr, Name, Adresse, TelNr)

Fahrzeug(FzNr, Typ, letzteWartung)

Bahn(Name, Ort)

gemietet(AuftragsNr, KundenNr, BahnNr, Datum, von, bis)

mitFz(AuftragsNr, FzNr)

(c) (8 Punkte)

Geben Sie alle Kunden an, die im Durchschnitt mehr als 20 Fahrzeuge pro Auftrag anmieten.

6. In einem B⁺-Baum werden die Einträge eines Wörterbuchs gespeichert. Abbildung 1 zeigt den Baum nach Einfügen der Wörter:

Daten, Datenblock, Datenverlust, DBA, DBMS, DBS
 5 10 12 3 4 3

Die Zahlen unter den Wörtern gibt jeweils die Länge des Wortes in Anzahl Zeichen an. Gehen Sie davon aus, daß auf jeder Seite des Baumes bis zu 20 Zeichen und beliebig viele Zeiger Platz finden.

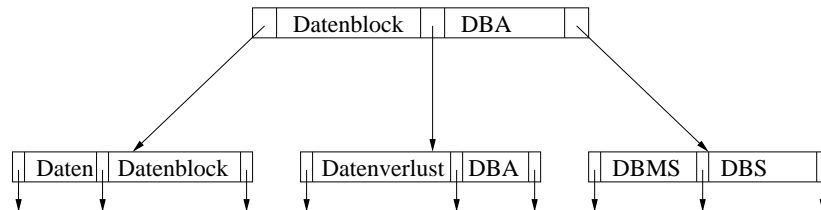


Abbildung 1: B⁺-Baum nach Einfügen

- (a) (7 Punkte)

Skizzieren Sie den Baum nach Einfügen des Wortes Datenbank.

(b) (7 Punkte)

Wie sähe der Baum nach Einfügen des Wortes *Datenbank* aus, wenn ein Präfix-B⁺-Baum vorläge?

7. Ein verteiltes Datenbanksystem benutzt das Quorum-Concensus Verfahren, um bei der Synchronisation auf replizierten Daten bei einer Änderungsoperation nicht alle Kopien eines Datenelements sperren zu müssen.

(a) (3 Punkte)

Angenommen das Datenelement A wird auf 7 Servern (S_1 bis S_7) als Kopie gehalten. Die Server S_1 und S_2 haben eine Ausfallwahrscheinlichkeit von jeweils 0.1%, Server S_3 von 0.2%, Server S_4 von 0.3%, Server S_5 von 0.4%, und die Server S_6 und S_7 von jeweils 0.6%. Wie würden Sie die Gewichte (Stimmen) 1,1,3,4,6,12,12 den Servern zuordnen? Begründen Sie Ihre Antwort kurz.

(b) (2 Punkte)

Geben Sie ein möglichst kleines gültiges Schreibquorum $Q_w(A)$ und ein möglichst kleines gültiges Lesequorum $Q_r(A)$ an.

(c) (2 Punkte)

Was ist bei einer Gewichte Verteilung von 1,1,3,4,6,12,12 die kleinste Anzahl von Rechnern mit denen das System noch lauffähig ist?

(d) (2 Punkte)

Was würde sich in (c) ändern, wenn alle Server das Gewicht 1 hätten? (D.h. was ist bei einer Gewichte Verteilung von 1,1,1,1,1,1,1, die kleinste Anzahl von Rechnern mit denen das System noch lauffähig ist?)

8. Der Optimierer eines Datenbanksystems soll folgende SQL-Anfragen in eine relationale Algebra übersetzen. Da Duplikateliminierung teuer ist, existiert in dieser Algebra ein Projektionsoperator π^D , der Duplikate nicht ausfiltert (d.h. diese Algebra arbeitet auf Multimengen).

Bei der Übersetzung welcher der folgenden SQL-Anfragen kann der schnellere Operator π^D eingesetzt werden, bei welchen ist der ursprüngliche Projektionsoperator π nötig? Begründen Sie Ihre Antwort jeweils kurz.

Die Relationen besitzen folgendes Schema:

- $R(\underline{A}, B, C)$
- $S(\underline{C}, D, E)$

(a) (2 Punkte)

```
select distinct R.B
from R;
```

(b) (2 Punkte)

```
select distinct S.C
from S;
```

(c) (3 Punkte)

```
select distinct R.A, S.D
from   R,S
where  R.C = S.C;
```

(d) (3 Punkte)

```
select distinct R.C, S.E
from   R,S
where  R.C = S.C;
```

(e) (3 Punkte)

```
select distinct R.B, S.C
from   R,S
where  R.C = S.C;
```