Klausur zum Kurs 1665 Datenbanksysteme

Beachten Sie bitte die folgenden Hinweise:

- Die Bearbeitungszeit beträgt 3 Stunden.
- Es sind maximal 104 Punkte erreichbar.
- Als **Hilfsmittel** sind Schreibuntensilien und unbeschriebene Papierbögen zugelassen.
- Schreiben Sie mit Tinte oder Kugelschreiber.
- Füllen Sie das **Deckblatt** aus und heften Sie es zusammen mit der **Aufgabenstellung** vor ihre **Lösungsblätter**. Geben Sie alles zusammen nach Beendigung der Klausur ab.
- Schreiben Sie auf jedes Lösungsblatt oben rechts ihren Namen, ihre Matrikelnummer und ggf., zu welcher Aufgabe es gehört.
- Überprüfen Sie bitte die Vollständigkeit der Aufgabenstellungen: Die Klausur umfaßt das Deckblatt, eine Teilnahmebescheinigung, dieses Blatt und 10 Aufgaben.
- Die korrigierte Klausur erhalten Sie in 4 - 6 Wochen zusammen mit einer Musterlösung zurück.
- Sollten Sie eine Klausurteilnahmebescheinigung benötigen, füllen Sie bitte den entsprechenden Vordruck aus und heften Sie ihn hinter das Deckblatt. Die Bescheinigung wird von der FernUniversität abgestempelt und Ihnen zusammen mit der korrigierten Klausur zugeschickt.

Wir wünschen Ihnen viel Erfolg!

Aufgabe 1: Relationenalgebra - Relationenkalkül **(4 Punkte)**

Erläutern Sie den Unterschied zwischen der Relationenalgebra und dem Relationenkalkül (max. 5 Sätze) (je 2 Punkte)

Aufgabe 2: Entity-Relationship-Diagramm (10 Punkte)

Modellieren Sie ein ER-Diagramm zur Verwaltung der Tour de France.

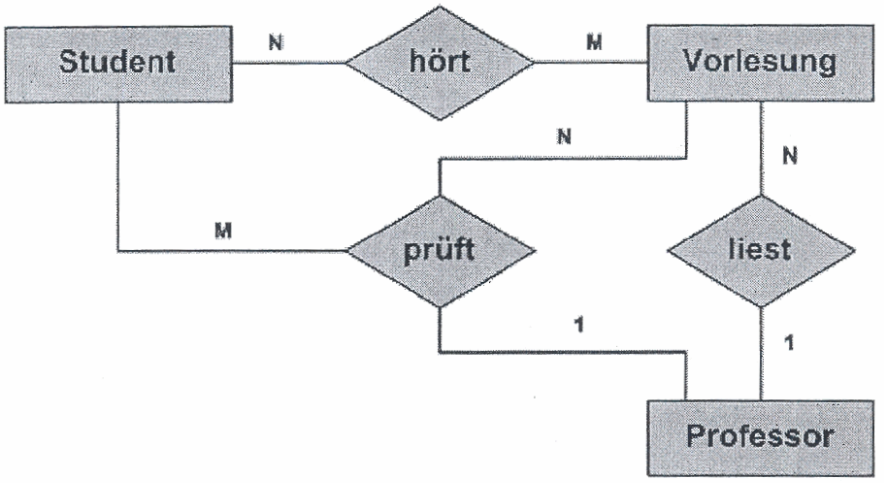
Es gibt Fahrer, die sowohl durch ihren Namen als auch eindeutig durch ihre Rückennummer gekennzeichnet sind, die während der Tour gleich bleibt. Jeder Fahrer hat eine Nationalität.

Fahrer gehören immer zu einem Team. Jedes Team besteht aus einer festen Anzahl von Fahrern. Teams sind durch den eindeutigen Teamnamen gekennzeichnet und haben auch eine Nationalität. In jedem Team erfüllt genau ein Fahrer die Rolle des Kapitäns. Daneben wird jedes Team durch genau einen sportlichen Leiter betreut.

Die Tour de France besteht aus mehreren Etappen. Eine Etappe hat eine Nummer, die als Schlüssel dient, ein Datum, einen Startort, einen Zielort und eine Länge. Zwischen Fahrern und Etappen gibt es eine Beziehung, die angibt, welche Fahrer an welchen Etappen teilgenommen haben. Zusätzlich wird der erreichte Platz notiert. Jeder Fahrer nimmt an wenigstens einer Etappe teil.

Aufgabe 4: Queries **(14 Punkte)**

Betrachtet sei die Datenbank für einen Universitätsbetrieb. Es ergibt sich folgendes ER-Diagramm (ohne Attribute):



Hieraus resultiert das folgende Schema:

Student (MatrNr, Name, Semester)
Vorlesung (VorlNr, Titel, SWS)
Professor (ProfNr, Name, Rang, Raum)
hört (MatrNr, VorlNr)
liest (ProfNr, VorlNr)
prüft (MatrNr, VorlNr, ProfNr, Note)

Formulieren Sie auf die Datenbank folgende Anfragen in SQL.

a) Listen Sie die Matrikelnummern und Namen aller Studenten im ersten Semester. (1 Punkt)

b) Zählen Sie alle Studenten mit Semesterzahlen zwischen 7 und 10. (1 Punkt)

c) Listen Sie alle Vorlesungen, die im Titel den String Datenbank enthalten, klein oder groß geschrieben. (2 Punkte)

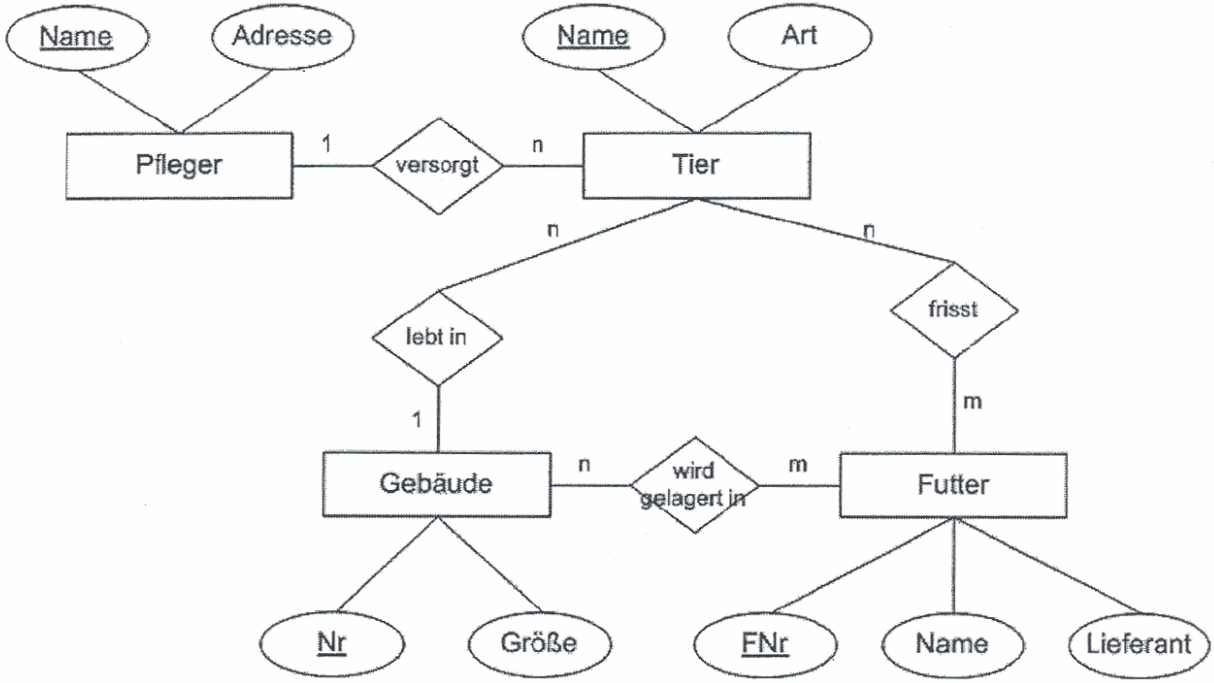
d) Listen Sie die Namen der Studenten mit ihren Vorlesungstiteln aufsteigend sortiert nach Namen und Vorlesungsnummer. (3 Punkte)

e) Listen Sie die Namen aller C4-Professoren und die Summe der SWS ihrer Vorlesungen. (3 Punkte)

f) Listen Sie solche Studenten, die alle 4-stündigen Vorlesungen hören. (4 Punkte)

Aufgabe 3: Datenmodellierung **(12 Punkte)**

Gegeben sei das folgende ER-Diagramm:



a) Erzeugen Sie aus dem gegebenen ER-Diagramm ein relationales DB-Schema. Begründen Sie Ihr Vorgehen.

b) Geben Sie die SQL Befehle für die Erzeugung der Datenbanktabellen des Schemas an.

Aufgabe 5: Normalformen **(10 Punkte)**

Betrachten Sie die folgende Relation „ZOOTIERE“:

TIER_NR	NAME	ABT_NAME	PFLEGER
14	Bongo	Gehege1	Krause
17	Clara	Haus3	Müller
25	Admiral	Gehege1	Krause
29	Streik	Gehege1	Krause
40	Herbert	Gehege2	Meyer
67	Anton	Haus3	Müller
81	RedBull	Haus3	Müller
89	Lenzen	Gehege2	Meyer

Es gelten die folgenden funktionalen Abhängigkeiten:

TIER_NR → ABT_NAME
TIER_NR → NAME
ABT_NAME → PFLEGER

a) In welcher Normalform befindet sich die Relation? Begründen Sie Ihre Antwort. (5 Punkte)

b) Erzeugen Sie aus der Relation „TIER“ eine Fd-treue Zerlegung in 3. Normalform. (5 Punkte)

Aufgabe 6: Funktionale Abhängigkeiten **(10 Punkte)**

Betrachten Sie die folgende Relation:

Tupel-Nr	A	B	C
#1	10	b1	c1
#2	10	b2	c2
#3	11	b4	c1
#4	12	b3	c4
#5	13	b1	c1
#6	14	b3	c4

Welche der folgenden Abhängigkeiten *kann* angesichts der obigen Relationsausprägung gelten? Falls eine Abhängigkeit nicht gilt, erklären Sie warum dies der Fall ist, indem Sie die Tupel spezifizieren, die die Verletzung verursachen.

(i) A → B
(ii) B → C
(iii) C → B
(iv) B → A
(v) C → A
(vi) A → C

Aufgabe 7: Anfrageoptimierung **(14 Punkte)**

Im Skript werden einige Heuristiken eingeführt, mit denen die algebraische Optimierung von Anfragen durchgeführt werden kann.

Betrachtet sei die Datenbankeiner Leihwagenfirma mit den folgenden Relationen:

Auto (Typ, Hersteller, StammfilialeNr, Kennzeichen)
Filiale (StammfilialeNr, FilialeStraße, FilialeStadt)
Kunde (Name, Straße, Wohnort, AusweisNr)
Vertrag (AusweisNr, Kennzeichen, Datum)

An diese Datenbank wird folgende Anfrage gestellt:

Auto [Typ = ´Golf´] [Kennzeichen = Kennzeichen] Vertrag
[AusweisNr = AusweisNr] Kunde [Datum < 12.03.2004] [Wohnort = ´Hagen´]
[Straße = ´Universitätsstr. 1´] [Name = ´Bratbäcker´ AND Hersteller = ´VW´]
[Datum, Typ, Hersteller, Name, Wohnort]

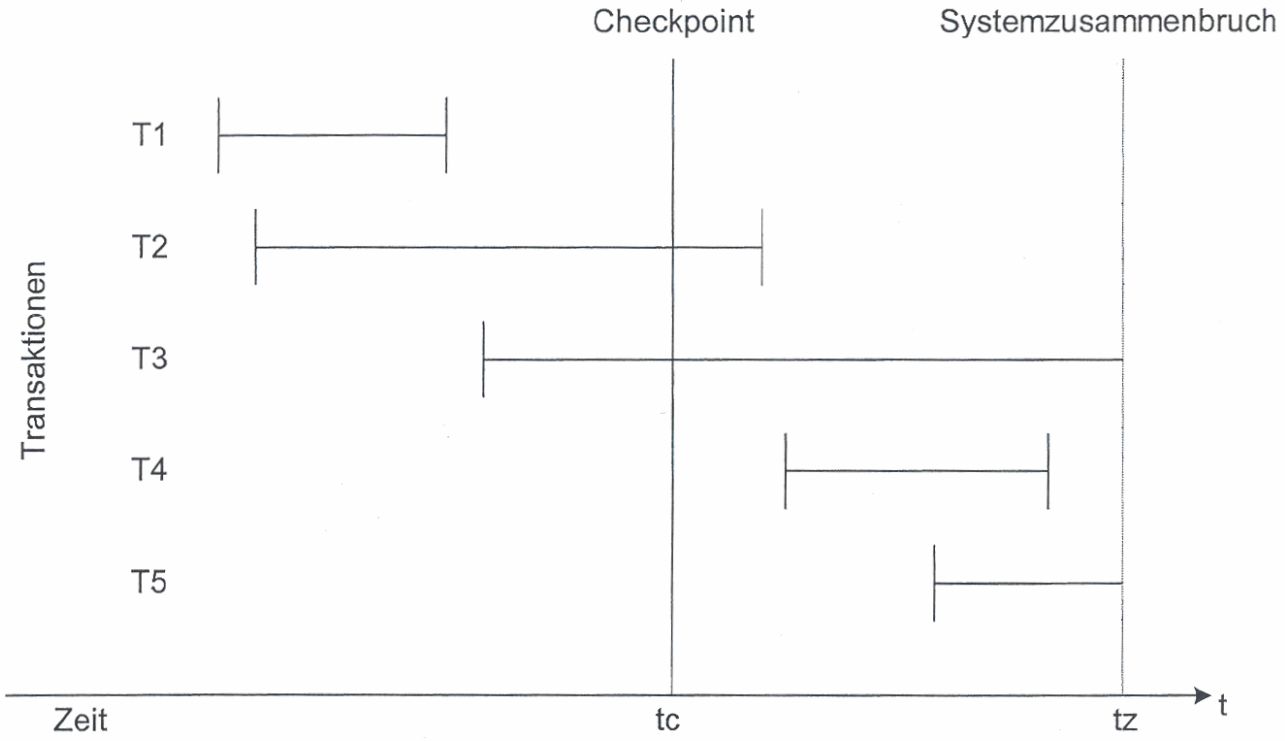
a) Zeichnen Sie für den obigen Ausdruck den entsprechenden Operatorbaum. (2 Punkte)

b) Optimieren Sie den angegebenen Ausdruck mit Hilfe der Ihnen bekannten Regeln (begründen Sie!) und geben Sie den optimierten Ausdruck an (6 Punkte). Zeichnen Sie den zugehörigen Operatorbaum (2 Punkte).

c) Beschreiben Sie möglichst präzise zwei jetzt noch durchführbare Optimierungen auf der physischen Ebene. (4 Punkte)

Aufgabe 8: Recovery **(10 Punkte)**

Betrachten Sie die folgende Situation:



Zum Zeitpunkt tz ereignet sich ein Systemzusammenbruch. Der letzte Checkpoint wurde zum Zeitpunkt tc geschrieben.

a) Wie sieht das Recovery für die obige Situation aus, wenn man davon ausgeht, dass das Schreiben eines Checkpoints das Herausschreiben aller veränderter Daten erzwingt? Was geschieht mit den einzelnen Transaktionen T1 bis T5? (5 Punkte)

Aufgabe 9: Serialisierbarkeit **(6 Punkte)**

Gegeben sind die im folgenden beschriebenen Logs.

a) L_1: $r_2(b)$ $r_4(d)$ $r_1(c)$ $r_4(a)$ $w_1(c)$ $r_3(b)$ $w_4(a)$ $w_4(d)$ $r_2(c)$ $r_2(d)$ $w_2(b)$ $r_3(a)$ $w_3(a)$

b) L_2: $r_1(d)$ $w_2(a)$ $r_3(b)$ $w_1(b)$ $r_2(d)$ $w_1(c)$ $r_3(a)$ $r_3(d)$ $r_2(c)$ $w_2(c)$

Sind die Transaktionen serialisierbar? Zeichnen Sie jeweils den zugehörigen Abhängigkeitsgraphen und geben Sie gegebenenfalls jeweils eine serielle Schedule an (Reihenfolge der Transaktionen). (jeweils 3 Punkte)

Aufgabe 10: Zugriffsschutz (14 Punkte)

Sie sind ein Datenbankadministrator und haben in einem relationelean DBMS die Tabellen **Angestellter** mit den Feldern **Name**, **Abteilung** und **Gehalt** erzeugt.
Zum Zwecke der Authorisierung haben Sie noch die Sichten **V_Angestelltenname**, mit **Name** als einzigem Attibut, sowie die Sicht **V_Abteilungsinfo** mit den Feldern **Abteilung** und **Duchrschnittsgehalt** zu definieren. Das Attribut **Duchschnittsgehalt** entspricht dem Durchschnittsgehalt pro Abteilung.

a) Geben Sie die SQL-Anweisungen an, mit denen Sie die beiden Sichten **V_Angestelltenname** und **V_Abteilungsinfo** definieren. (6 Punkte)

b) Sie möchten Ihren Stellvertreter mit dem Benutzernamen **Meier** dazu befähigen
   - Angestellte zu entlassen,
   - zu überprüfen, wer Angestellter ist,
   - die Durchschnittsgehälter der einzelnen Abteilungen einzusehen.

   Welche SQL-Befehle sind abzusetzen, damit Ihr Stellvertreter die oben genannten Rechte erhält? (6 Punkte)

c) Ihr Stellvertreter Meier soll weiterhin die Möglichkeit haben, andere Benutzer zu befähigen, die Sicht V_Angestelltenname zu lesen. Wie sieht der entsprechende SQL-Befehl aus? (2 Punkte)