

Aufgabe D1 – Multiple Choice (6 Punkte)

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



	Richtig	Falsch
Die Komplexität des Nested-Loop Joins ist stets höher als die des Merge-Joins.	<input type="checkbox"/>	<input type="checkbox"/>
Wenn die Daten vorab in sinnvoller Weise sortiert wurden, kann dies die Ausführung des Nested-Loop Joins beschleunigen.	<input type="checkbox"/>	<input type="checkbox"/>
Die Berechnung der Intervall-Grenzen ist beim Equi-Depth Histogramm aufwendiger als beim Equi-Width Histogramm.	<input type="checkbox"/>	<input type="checkbox"/>
Eine zustandsbehaftete Ausführung hat die Eigenschaft, dass der Zustand explizit erfasst und in einer Datenbank gespeichert wird.	<input type="checkbox"/>	<input type="checkbox"/>
Beim asynchronen Zugriff wird die Kontrolle an den Aufrufer zurückgegeben, sobald die letzte Kopie des Datenobjekts geschrieben wurde.	<input type="checkbox"/>	<input type="checkbox"/>
Eine sinnvolle Möglichkeit der Auflösung von Inkonsistenzen von mehreren Versionen des Einkaufswagen ist, ihre Schnittmenge zu berechnen.	<input type="checkbox"/>	<input type="checkbox"/>
Der Kommunikationsaufwand in strukturierten P2P-Systemen zur Ermittlung eines Anfrageergebnisses ist abhängig von der Anzahl der Datenobjekte in der Datenbank.	<input type="checkbox"/>	<input type="checkbox"/>
Vector Clocks sind Listen von (Knoten, Zähler)-Paaren, und es gibt eine solche Liste für jede Version eines Datenobjekts.	<input type="checkbox"/>	<input type="checkbox"/>
PIQL berücksichtigt Fremdschlüsselbeziehungen, um so den Aufwand zur Berechnung des Operator-Ergebnisses zu begrenzen.	<input type="checkbox"/>	<input type="checkbox"/>
PIQL berücksichtigt Kardinalitätsangaben im Datenbankschema, um so den Aufwand zur Berechnung des Operator-Ergebnisses zu begrenzen.	<input type="checkbox"/>	<input type="checkbox"/>
Im PIQL-System muß die Anwendung den Zustand zwischen Datenbank-Zugriffen verwalten, falls dies erforderlich sein sollte.	<input type="checkbox"/>	<input type="checkbox"/>
Der DataStop-Operator im PIQL-Kontext entspricht dem Top-Operator in anderen SQL-Dialekt.	<input type="checkbox"/>	<input type="checkbox"/>

Für jede richtige Antwort werden 0,5 Punkte gutgeschrieben, für jede falsche 0,5 Punkte abgezogen. Der Aufgabenteil wird insgesamt mit mind. 0 Punkten bewertet.

Aufgabe D2 – Normalformen (8 Punkte)

Gegeben sei das Relationenschema $R(A, B, C, D, E, F, G)$ und eine Menge F funktionaler Abhängigkeiten über R :

$$F = \{ \begin{array}{l} AE \rightarrow F, \\ B \rightarrow CE, \\ BC \rightarrow F, \\ C \rightarrow BE, \\ D \rightarrow AE, \\ E \rightarrow A, \\ EF \rightarrow G, \\ G \rightarrow A \end{array} \}$$

Bei allen Attributen handelt es sich um atomare Datentypen.

- a) Bestimmen Sie die Menge aller Schlüsselkandidaten von R .

(2,5 Punkte)

- b) Bestimmen Sie für alle Primattribute die Attributhülle. Was lässt sich aus diesen Attributhüllen bezüglich der höchst möglichen Normalform (1NF, 2NF, 3NF, BCNF) in der sich R befindet folgern? Begründen Sie kurz Ihre Antwort.

(1,5 Punkte)

- c) Finden Sie eine Zerlegung von R , die in dritter Normalform (3NF), abhängigkeitstreu, verbundtreu und minimal ist. Benutzen Sie für das Ergebnis Ihrer Zerlegung die Notation

$$S = \{(R_1, K_1), \dots, (R_n, K_n)\},$$

wobei R_i für Relationenschemata und K_i für die entsprechenden Schlüsselmengen stehen.

(4 Punkte)

Aufgabe D3 – SQL (8 Punkte)

In einer SQL-Datenbank seien folgende Relationen definiert:

Player(player_id, name, team)

player_id	name	team
1	Lahm	GER
2	Klose	GER
3	Odonkor	GER
4	Villa	SPA
5	Van Nistelrooy	NED

Cup(cup_id, cup_name)

cup_id	cup_name
1	EM 2008
2	WM 2010
3	EM 2012

In einer weiteren Relation ist gespeichert, ob ein Spieler an einem Wettbewerb teilgenommen hat und wie viele Tore er dabei erzielt hat:

Participation(player_id, cup_id, goals)

player_id	cup_id	goals
1	1	1
1	2	0
1	3	1
2	1	2
2	2	4
2	3	1
4	1	4
4	2	5
5	1	2

Verwenden Sie in Ihrer Lösung ausschließlich SQL-Standard-konforme Ausdrücke, die in der Vorlesung vorgestellt wurden.

- a) Finden Sie eine SQL-Abfrage, die die Summe der geschossenen Tore pro Spieler über alle Wettbewerbe ausgibt. Speichern Sie das Ergebnis der Abfrage in einer View mit folgendem Schema:

GoalsPerPlayer(player_id, name, team, sum_goals)

(3 Punkte)

- b) Finden Sie eine SQL-Abfrage, die alle Spieler ausgibt, die an allen Wettbewerben teilgenommen haben. Die Abfrage soll dabei unabhängig vom Inhalt der Datenbank sein. Speichern Sie das Ergebnis der Abfrage in einer View mit folgendem Schema:

AlwaysParticipating(player_id, name)

(2,5 Punkte)

- c) Aufbauend auf der View aus Aufgabenteil a) wird eine zweite View wie folgt definiert:

```
CREATE VIEW GoalsPerTeam AS  
SELECT team, SUM(sum_goals) AS sum_goals FROM GoalsPerPlayer  
GROUP BY team;
```

Berechnen Sie die Ergebnisrelation folgender SQL-Abfrage für den gegebenen Inhalt der Datenbank:

```
SELECT t1.team, t1.sum_goals  
FROM GoalsPerTeam AS t1  
LEFT OUTER JOIN GoalsPerTeam AS t2  
ON t1.sum_goals < t2.sum_goals  
WHERE t2.team IS NULL;
```

Welches Informationsbedürfnis wird mit dieser Abfrage allgemein beantwortet?

(2,5 Punkte)

Aufgabe D4 – Transaktionen und Histories (8 Punkte)

Gegeben seien folgende drei Transaktionen:

$$T_1 = w_1[x] \ r_1[y] \ c_1$$

$$T_2 = w_2[y] \ w_2[x] \ w_2[z] \ c_2$$

$$T_3 = r_3[z] \ r_3[x] \ c_3$$

Aufbauend auf diesen Transaktionen seien folgende Histories definiert:

$$H_1 = w_1[x] \ w_2[y] \ w_2[x] \ w_2[z] \ c_2 \ r_1[y] \ c_1 \ r_3[z] \ r_3[x] \ c_3$$

$$H_2 = w_2[y] \ w_2[x] \ w_2[z] \ w_1[x] \ r_3[z] \ r_1[y] \ c_2 \ c_1 \ r_3[x] \ c_3$$

- a) Erzeugen Sie die Serialisierbarkeitsgraphen für beide Histories. Notieren Sie an den Kanten der Graphen alle Variablen, über die der Konflikt zustande kommt. Geben Sie jeweils für H_1 und H_2 an, ob die History serialisierbar ist.

(4 Punkte)

- b) Ermitteln Sie für beide Histories jeweils ob die Eigenschaften rücksetzbar (RC), kaskadenfrei rücksetzbar (ACA) und strikt (ST) erfüllt sind und bestimmen Sie die maximale Rücksetzbarkeitsklasse. Es muss dabei aus der Lösung ersichtlich werden, dass alle zu prüfenden Bedingungen berücksichtigt wurden.

(4 Punkte)