

---

# Klausur der Lehrveranstaltung “Datenbanksysteme”

04. August 2017

---

Hier bitte Namensschild aufkleben

## Hinweis

Kleben Sie zunächst das Namensschild auf das Titelblatt und versehen Sie alle Antwortblätter mit Ihrem Namen und Ihrer Matrikelnummer. Schreiben Sie die Lösungen nur auf das jeweilige zur Aufgabe passende Antwortblatt im Anhang der Klausur. Sie können die Antwortblätter auf beiden Seiten beschreiben. Sollte Ihnen wider Erwarten der Platz auf einem Blatt nicht ausreichen, so können Sie von der Klausuraufsicht weitere Blätter erhalten, die Sie wiederum mit Ihrem Namen und Ihrer Matrikelnummer kennzeichnen. Geben Sie alle Blätter dieser Klausur, auch die Aufgabenblätter, wieder ab; nachgereichte Blätter können nicht korrigiert werden.

Aufgabe:	1	2	3	4	$\Sigma$
Punkte:					
Note:					

## Aufgabe 1: ER-Modellierung (15 Punkte)

Gegeben ist das in Abbildung 1 dargestellte ER-Modell. Die Schlüsselattribute sind durch Unterstreichen gekennzeichnet. Die Kardinalitätsangaben sind *Teilnehmerkardinalitäten*.

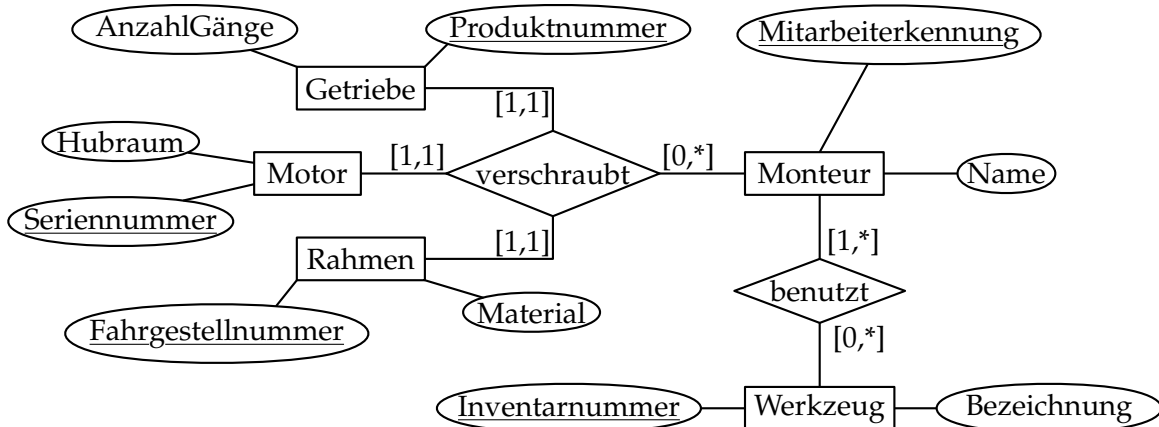


Abbildung 1: Gegebenes ER-Modell

- a) Bitte markieren Sie auf dem Antwortblatt, auf welche der Entitäten in Abbildung 1 die folgenden Aussagen zutreffen. Jede korrekt ausgefüllte Zeile (Aussage) wird mit 0,5 Punkten bewertet.

(2 Punkte)

	Werkzeug	Monteur	Motor	Getriebe	Rahmen
Hat einen Primärschlüssel	■	■	■	■	■
Steht mit dem Entitätstyp <i>Monteur</i> in Beziehung	■	■	■	■	■
Steht in keiner reflexiven Beziehung	■	■	■	■	■
Kann Entitäten enthalten, die mit keiner anderen Entität in Beziehung stehen	■	■	■	■	■

- b) Überführen Sie das ER-Modell aus Abbildung 1 kapazitätserhaltend in ein relationales Modell. Optimieren Sie die Relationen dabei soweit wie möglich mit den in der Vorlesung vorgestellten Methoden. Bitte geben Sie nur die Attribute an, nicht aber deren Datentyp. Kennzeichnen Sie Primärschlüssel durch Unterstreichen und Fremdschlüssel durch Pfeile zu dem entsprechenden Schlüssel. Verwenden Sie die auf dem Antwortblatt vorgegebene Struktur. Dort müssen nicht alle Felder ausgefüllt werden.

(5 Punkte)

c) Zeichnen Sie für die folgenden Sachverhalte jeweils ein kleines Beispiel. Jede korrekte Zeichnung wird mit einem Punkt bewertet.

- (i) Zeichnen Sie ein EER-Modell mit einer Entität *A*, welches ein optionales Attribut *x* besitzt.
- (ii) Zeichnen Sie ein EER-Modell mit einer Entität *A* und einer mehrfachen Spezialisierung von *A* zu *A1* und *A2*.
- (iii) In Abbildung 2 ist ein ER-Modell mit Standardkardinalitäten gegeben. Überführen Sie das ER-Modell in ein Modell mit entsprechenden Teilnehmerkardinalitäten in Intervallnotation.



Abbildung 2: ER-Modell zu Aufgabe c.iii

(3 Punkte)

- d) Überführen Sie das in Abbildung 3 gegebene ER-Modell in ein EER-Modell. Erweitern Sie das EER-Modell zusätzlich um die Entität *Motorrad* mit dem Attribut *Lenkerhöhe*. Modellieren Sie auch die nachfolgenden Eigenschaften und Beziehungen. Jedes Motorrad ist ein *Kraftfahrzeug*. Ein *Kraftfahrzeug* kann entweder ein *Motorrad* oder ein *Auto* sein, jedoch nicht beides. Es gibt auch Kraftfahrzeuge, die weder Auto noch Motorrad sind.

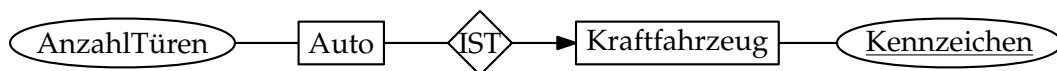


Abbildung 3: ER-Modell zu Aufgabe d

(5 Punkte)

## Aufgabe 2 – SQL (12 Punkte)

Der Stromanbieter „Energy4You“ gestaltet seine Datenhaltung mithilfe einer SQL-Datenbank. In der Datenbank werden die angebotenen Stromtarife (Relation „Tarif“), die Kunden des Anbieters (Relation „Kunde“) sowie die von den Kunden selbst abgelesenen Zählerstände (Relation „Zählerstand“) gespeichert. Zur Vereinfachung nehmen wir an, dass die Vertragslaufzeit ein Jahr beträgt, ein Vertragsjahr immer am 1. Januar eines Jahres beginnt und zu diesem Stichtag (1. Januar) das jährliche Ablesen der Zählerstände erfolgt.

Ein Ausschnitt aus dem Inhalt der Relationen folgt:

Tarif		Kunde					
ID	Preis	ID	VName	NName	Zähler	Tarif	Abschlag
Öko	0,28	0	Derek	Shepherd	97182	Online	32,00
ÖkoPlus	0,265	1	Meredith	Grey	29174	Online	33,00
Online	0,25	2	Alex	Karev	63916	Öko	52,00

Zählerstand		
Zähler	Jahr	Stand
63916	2016	13200
63916	2017	17600
20174	2016	22500
20174	2017	24800
97182	2016	24800

Erläuterungen:

- Nehmen Sie sinnvolle Datentypen und Schlüsseldefinitionen an.
- Die Daten in den Tabellen sind exemplarisch, geben Sie für die unten gestellten Aufgaben aber immer generelle Lösungen an! Alle Aufgaben lassen sich mit SQL-Konstrukten lösen, die in der Vorlesung vorgestellt wurden.
- Das Attribut „Preis“ in der Relation „Tarif“ beschreibt den Strompreis in Euro pro kWh. Der Stand der Zähler (Attribut „Stand“ in Relation „Zählerstand“) hat die Einheit kWh.
- Die in der Relation „Kunde“ gespeicherten Abschlagszahlungen sind monatlich zu entrichten.

- a) Bestimmen Sie für jeden Tarif, wie viele Kunden diesen Tarif nutzen.  
(2,5 Punkte)
- b) Bestimmen Sie den für das Jahr 2016 gemeldeten Zählerstand (Ausgabe: Zählerstand.Zähler und Zählerstand.Stand) aller Kunden, die einen Tarif nutzen, der „Öko“ in seinem Namen (Tarif.ID) hat.  
(2,5 Punkte)
- c) Bestimmen Sie die Namen (NName) aller Kunden, die einen überdurchschnittlichen Abschlagsbetrag haben. Dabei sollen nur Kunden in die Berechnung des Durchschnittswertes eingehen, die ihren Zählerstand für das Jahr 2017 bereits mitgeteilt haben.  
(2 Punkte)
- d) Erstellen Sie eine Sicht mit dem Namen „Verbrauch16“ und dem Schema (Zähler, Verbrauch), die den im Jahr 2016 erfassten Stromverbrauch je Zähler beschreibt. Existiert die Sicht in der Datenbank bereits, soll die bereits bestehende Sicht gelöscht werden. Es sollen nur Zähler in der Sicht aufgeführt werden, für die bereits ein Stand für das Jahr 2017 von den Kunden gemeldet wurde.  
(3 Punkte)
- e) Berechnen Sie die ID's der drei Kunden, die die höchste Rückzahlung für 2016 erhalten. Geben Sie die ID sowie die Höhe der Rückzahlung aus.  
(2 Punkte)

### Aufgabe 3 – Relationentheorie (15 Punkte)

- a) Gegeben sei die Relation  $R_a(A, B, C, D, E, G, H)$  mit der folgenden Menge an funktionalen Abhängigkeiten:

$$\begin{aligned} F = \{ & CDG \rightarrow EB \\ & AB \rightarrow C \\ & A \rightarrow D \\ & ABC \rightarrow E \\ & B \rightarrow E \\ & CDE \rightarrow A \} \end{aligned}$$

- (i) Welche der folgenden Attributmengen sind Schlüssel(-kandidaten) für  $R_a$  mit den funktionalen Abhängigkeiten  $F$ ?

Attributmenge	Schlüssel?
$\{BGH\}$	ja / nein
$\{ABGH\}$	ja / nein
$\{CDGH\}$	ja / nein
$\{ADEGH\}$	ja / nein

(2 Punkte)

- (ii) Transformieren Sie  $R_a$  mit dem Synthesalgorithmus in 3NF. Geben Sie die Zwischenergebnisse der Schritte des Algorithmus an.

(4 Punkte)

- (iii) Ist die Zerlegung  $R_{a1}(A, B, C, E), R_{a2}(C, D, E, G, H)$  verbundtreu? Begründung!

(3 Punkte)

**(9 Punkte)**

- b)** Gegeben sei die Relation  $R_b(A, B, C, D, E)$  mit der folgenden Menge an funktionalen Abhängigkeiten:

$$F = \{ \begin{array}{l} A \rightarrow E \\ B C \rightarrow A \\ D E \rightarrow B \end{array} \}$$

- (i)** Berechnen Sie die Hüllen der folgenden Attributmengen bzgl.  $F$ .

Attributmenge	Hülle
$\{A\}$	
$\{B, C\}$	
$\{D, E\}$	
$\{A, C, D\}$	
$\{B, C, D\}$	
$\{C, D, E\}$	

(2 Punkte)

- (ii)** Ist  $R_b$  in 3NF? Begründung!

(1 Punkt)

- (iii)** Welche der folgenden Abhängigkeiten verletzen die BCNF. Begründen Sie Ihre Antwort.

Abhängigkeit	Verletzt BCNF?	Begründung
$A \rightarrow E$	Ja / Nein	
$B C \rightarrow A$	Ja / Nein	
$D E \rightarrow B$	Ja / Nein	

(3 Punkte)

**(6 Punkte)**

## Aufgabe 4 – Transaktionen und Join-Algorithmen (18 Punkte)

Gegeben sind die folgenden partiellen Historien:

$$H_1 = w_1[x] \ w_2[x] \ w_3[x] \ w_4[x] \ c_1$$

$$H_2 = w_3[x] \ w_1[x] \ r_1[y] \ w_2[x] \ c_1 \ r_3[x] \ w_3[y] \ w_2[y] \ c_2$$

$$H_3 = w_3[x] \ w_1[x] \ r_1[y] \ w_2[x] \ c_1 \ r_3[x] \ w_3[y] \ w_2[y] \ c_3$$

$$H_4 = w_3[x] \ w_1[x] \ r_1[y] \ w_2[x] \ c_1 \ r_3[x] \ w_3[y] \ c_3 \ w_2[y] \ a_2$$

$$H_5 = r_2[x] \ w_3[z] \ r_2[z] \ w_1[z] \ r_1[y] \ w_2[y] \ c_1 \ r_3[x] \ w_3[y] \ c_3$$

$$H_6 = w_4[z] \ w_3[x] \ r_2[z] \ w_1[x] \ w_2[x] \ r_1[y] \ r_3[x] \ r_2[y] \ w_3[y] \ r_4[z] \ c_1 \ a_2 \ c_3 \ c_4$$

$$H_7 = r_4[z] \ r_2[y] \ w_1[x] \ w_4[z] \ r_2[z] \ w_2[z] \ w_3[x] \ r_1[y] \ c_1 \ r_3[x] \ w_3[y] \ c_3$$

- a) Benennen Sie die Begriffe, die hinter der Abkürzung ACID stehen.  
(2 Punkte)
- b) Geben Sie die Serialisierbarkeitsgraphen  $SG(H_1)$ ,  $SG(H_2)$ ,  $SG(H_3)$  und  $SG(H_4)$  an. Beschriften Sie die Kanten der Serialisierbarkeitsgraphen mit den Objekten der konfligierenden Operationen. Begründen Sie für jede Historie, ob sie serialisierbar ist.  
(4 Punkte)
- c) Geben Sie für die Committed Projections  $C(H_5)$ ,  $C(H_6)$  und  $C(H_7)$  jeweils an, ob diese zu  $C(H_3)$ , d. h. zur Committed Projection der Historie  $H_3$ , konfliktäquivalent sind. Begründen Sie Ihre Antwort.  
(3 Punkte)
- d) (i) Geben Sie die Definition der Rücksetzbarkeitsklasse STRICT an.  
(ii) Impliziert Konfliktäquivalenz die Zugehörigkeit zu den gleichen Rücksetzbarkeitsklassen? Begründen Sie.  
(iii) Geben Sie an, in welchen Rücksetzbarkeitsklassen  $H_1$ ,  $H_3$  und  $H_6$  liegen.  
*Hinweis: Verwenden Sie für Ihre Antwort die Tabelle auf dem Lösungsblatt.*  
(4 Punkte)



- e) Gegeben seien die Relationen  $R_1(\underline{A}, B, C, D)$  und  $R_2(A, \underline{B}, C, D)$ .  $A$  ist Schlüssel der Relation  $R_1$ ,  $B$  ist Schlüssel der Relation  $R_2$ . Die Relationen sind nach dem Attribut  $B$  sortiert. Das Attribut  $C$  hat für jedes Tupel in der jeweiligen Relation den gleichen Wert. Die Relation  $R_1(\underline{A}, B, C, D)$  enthält  $n$  Tupel und hat eine Gesamtgröße von 512 MB, und  $R_2(A, \underline{B}, C, D)$  enthält  $m$  Tupel und hat eine Gesamtgröße von 60 MB.

$R_1(\underline{A}, B, C, D)$	<u>A</u>	B	C	D	}	m
	23	5	4	25		
	7	5	4	5		
	...	...	...	...		
	188	10233	4	50		
$R_2(A, \underline{B}, C, D)$	A	<u>B</u>	C	D	}	n
	19	2	88	5		
	19	3	88	2		
	...	...	...	...		
	138	1118	88	1		

- (i) Gehen Sie davon aus, dass der in der Vorlesung vorgestellte Merge-Join-Algorithmus genutzt wird. Wie ist die Komplexität des Merge-Joins der beiden Relationen im O-Kalkül (Landau-Symbole), bezogen auf die Anzahl der Tupel  $n$  und  $m$ ? Geben Sie jeweils die Komplexität für den Fall an, dass das Join-Attribut  $A$ ,  $B$  bzw.  $C$  ist. (3 Punkte)
- (ii) Gehen Sie davon aus, dass für die Join-Operation Hauptspeicher in der Größe von 64 MB zur Verfügung steht. Die Blockgröße sei 4 MB. Welche Kosten bezüglich der Blockzugriffe entstehen für einen Equi-Join von  $R_1$  und  $R_2$  über das Attribut  $D$  bei Durchführung eines Block-Nested-Loop Joins? (2 Punkte)

**(5 Punkte)**