

## Aufgabe 1 – ER-Modellierung (14 Punkte)

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

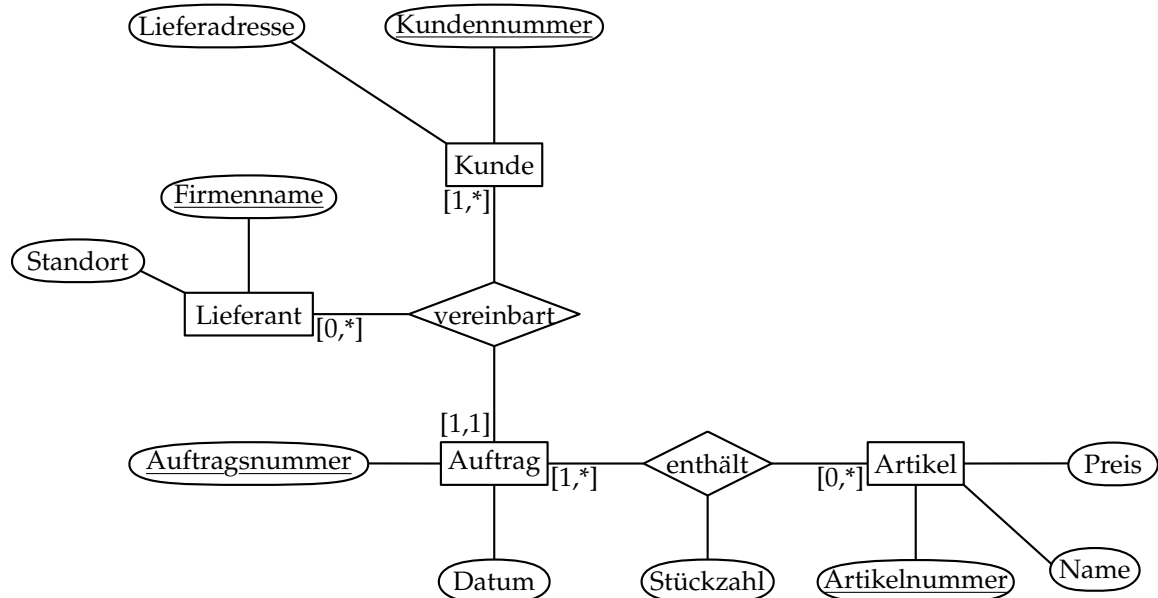


Abbildung 1: ER-Modell 'Lieferungen'

- a) Geben Sie auf dem Antwortblatt für die folgenden Aussagen an, ob diese *wahr* oder *falsch* sind. Für jede richtige Antwort gibt es 1/2 Punkt, für falsch beantwortete Antworten wird 1/2 Punkt abgezogen. Für nicht beantwortete Aufgaben werden weder Punkte vergeben noch abgezogen. Diese Aufgabe wird nicht mit weniger als 0 Punkten bewertet. (4 Punkte)

	WAHR	FALSCH
Ein Auftrag muss mindestens einen Artikel enthalten.	■	■
Jeder Artikel kann von mindestens einem Lieferanten geliefert werden.	■	■
Ein Kunde kann einen Auftrag nur mit <b>einem</b> Lieferanten vereinbaren.	■	■
Die Beziehung <i>vereinbaren</i> kann semantikerhaltend durch drei zweistellige Beziehungen dargestellt werden.	■	■
<i>Preis</i> ist ein Schlüsselkandidat für <i>Artikel</i> .	■	■
Es gibt im Allgemeinen Beziehungstypen, die nur mit <b>einem</b> Entitätstyp in Beziehung stehen.	■	■
Ein Entity-Relationship-Modell besteht ausschließlich aus Entitäten und Beziehungen.	■	■
Eine Partitionierung kann durch eine IST-Beziehung modelliert werden.	■	■

- b)** Geben Sie auf dem Antwortblatt für die Beziehung *vereinbart* im ER-Modell aus Abbildung 1 die entsprechenden Standardkardinalitäten an.  
(2 Punkte)
- c)** 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 Regeln. Geben Sie nur die Attribute an, nicht aber deren Datentyp. Geben Sie auch die Schlüssel und Fremdschlüssel an.  
(5 Punkte)
- d)** Erweitern Sie das Modell in Abbildung 1 mithilfe der Methoden der EER-Modellierung auf dem Antwortblatt wie folgt: *Kunden* können *Privatkunden* und *Großkunden* sein. Insbesondere kann ein Kunde auch sowohl Privatkunde als auch Großkunde sein. Privatkunden haben einen *Vornamen* und *Nachnamen*, während Großkunden durch einen gewährten *Rabatt* charakterisiert werden.  
(3 Punkte)

## Aufgabe 2 – SQL (15 Punkte)

Seit kurzem ist das KIT Partner der ‚Space Exploration Missions‘ im Verbund mit verschiedenen Weltraumorganisationen. Hierbei sollen so viele Daten wie möglich der Öffentlichkeit zur Verfügung gestellt werden. Insbesondere qualitativ hochwertige Photos finden in der Öffentlichkeit viel Beachtung. Sie wurden nun damit beauftragt, für eine grafische Oberfläche die folgenden Abfragen zu entwerfen bzw. Aufgaben zu lösen. Die relationale Datenbank ‚Sky Survey‘ enthält verschiedene stellare Objekte und ihre Eigenschaften. Bisher besteht sie aus drei Relationen: Stars, Planets und Photos (von Planeten).

Nehmen Sie die Relationen aus Abbildung 2 mit sinnvollen Datentypen und Schlüsseln an. Die Daten in den Relationsinstanzen sind exemplarisch. Geben Sie für die unten gestellten Aufgaben immer generelle Lösungen an!

**Erläuterungen:** Alle Aufgaben lassen sich mit SQL-Konstrukten lösen, die in der Vorlesung vorgestellt wurden. Die Teilaufgaben sind voneinander unabhängig. Die Spalten, die mit ID benannt sind, beschreiben jeweils einen Primärschlüssel. Für Fremdschlüssel wird die Konvention Name der Relation gefolgt von ‚ID‘ verwendet.

- a) Geben Sie die Namen aller Planeten unseres Sonnensystems aus (Stars.Name = ‚Sun‘). (2 Punkte)
- b) Selektieren Sie alle Photos (Attribut ‚Photo‘) vom Mars mit maximalem Zoom(-Level). (3 Punkte)
- c) Planeten, auf denen potentiell Leben existiert, haben die folgenden Eigenschaften:
- Der Mutterstern ist vom Typ ‚G2V‘ oder ‚M5.5Ve‘.
  - Die Masse des Planeten (‚Mass‘) liegt zwischen 0.5 und 2 (Erdbmassen).
  - Der Durchmesser des Planeten (‚Diameter‘) beträgt 0.7 bis 1.2 (Erddurchmesser).

Erstellen Sie eine Sicht (planet\_name, star\_name, mass, diameter, distance) mit allen Planetenkandidaten, die diese Eigenschaften erfüllen.

(3 Punkte)

- d) Geben Sie für jeden Planeten dessen ID, Name und die (minimale) Distanz zum nächsten Planeten desselben Sonnensystems an. Sollte ein Planet der einzige in einem System sein, soll er nicht Teil des Ergebnisses sein.  
*Hinweise:* Gehen Sie vereinfachend von den folgenden Annahmen aus. ‚Distance‘ in der Relation Planet ist die Distanz des Planetenmittelpunkts von dem zugehörigen Stern. Die Umlaufbahnen sind jeweils kreisförmig in einer Ebene. Abstände zwischen Planeten beziehen sich auf die Abstände ihrer Umlaufbahnen. (3 Punkte)
- e) Die Relation ‚Stars‘ enthält Redundanzen im Attribut ‚Type‘, die zu Update-Anomalien führen könnten. Normalisieren Sie die Relation auf geeignete Art und Weise, sodass dieses Problem nicht mehr auftritt. Verwenden Sie für diese Umstrukturierung geeignete SQL-Statements.

**Hinweis:** Bei dieser Teilaufgabe dürfen Sie Konstrukte wie ROWNUM oder Auto\_Increment benutzen.

(4 Punkte)

Stars		
ID	Name	Type
1	'SUN'	'G2V'
2	'Proxima Centauri'	'M5.5Ve'
3	'Alpha Centauri A'	'G2V'
4	'Alpha Centauri B'	'K1V'
5	'Luhman 16A'	'L8.1'
6	'Luhman 16B'	'T1.2'
7	'Sirius A'	'A1V'
8	'Wolf 359'	'M6.0V'
9	'Kepler-186'	'M5.5Ve'

Planets					
StarID	ID	Name	Mass	Diameter	Distance
1	1	'Mercury'	0.06	0.382	0.39
1	2	'Venus'	0.82	0.949	0.72
1	3	'Earth'	1	1	1
1	4	'Mars'	0.11	0.532	1.52
1	5	'Jupiter'	317.8	11.209	5.2
1	6	'Saturn'	95.2	9.449	9.54
9	7	'Kepler-186f'	1.4	1.11	0.4

Photos		
PlanetID	Photo	Zoom
1	'Pic1.jpg'	0
1	'Pic2.jpg'	2
3	'Pic3.jpg'	0
3	'Pic4.jpg'	1
3	'Pic5.jpg'	2
2	'Pic6.jpg'	0
2	'Pic7.jpg'	1
6	'Pic8.jpg'	1
6	'Pic9.jpg'	2

Abbildung 2: Relationsinstanzen der relationalen Datenbasis 'Sky Survey'

### Aufgabe 3 – Relationentheorie, Join-Operatoren (15 Punkte)

- a) Gegeben sei eine Relation  $R(A, B, C, D, E, F)$  mit der folgenden Menge an funktionalen Abhängigkeiten:

$$F = \{ A \rightarrow BC \\ CD \rightarrow EF \\ AB \rightarrow E \}$$

- a1) Sind die folgenden Aussagen richtig oder falsch? Für jede richtige Antwort gibt es 1/2 Punkt und für jede falsche wird 1/2 Punkt abgezogen. Diese Teilaufgabe kann nicht mit weniger als 0 Punkten bewertet werden. Füllen Sie die Tabelle auf dem Antwortblatt aus.

Funktionale Abhängigkeiten	RICHTIG	FALSCH
$AD \rightarrow F$	■	■
$BD \rightarrow EF$	■	■
$A \rightarrow C$	■	■
$A \rightarrow E$	■	■
$AD \rightarrow EF$	■	■
$B \rightarrow A$	■	■

- a2) Bestimmen Sie die Menge aller Schlüssel der Relation  $R$ .
- a3) In welcher höchsten Normalform befindet sich  $R$ ? Gehen Sie dabei davon aus, dass alle Attribute atomare Domänen haben. Begründung!
- a4) Ist die Zerlegung von  $R$  in  $R_1(\underline{A}, B, C, E)$  und  $R_2(\underline{C}, \underline{D}, F)$  verbundtreu? Begründung! *Hinweis:* der Schlüssel ist jeweils unterstrichen.

(5 Punkte)

- b) Gegeben sei ein Kalkül- $K$  (Ableitungsregeln) zur Berechnung der vollständigen Hülle der funktionalen Abhängigkeiten:

$$KR_1 : \quad \{ \} \models A \rightarrow A \\ KR_2 : \quad \{ A \rightarrow B, B \rightarrow C \} \models A \rightarrow C \\ KR_3 : \quad \{ A \rightarrow B, A \rightarrow C \} \models A \rightarrow BC \\ KR_4 : \quad \{ A \rightarrow BC \} \models A \rightarrow B$$

Zur Erinnerung sei der in der Vorlesung vorgestellte RAP-Kalkül gegeben. Der Kalkül ist gültig, vollständig und minimal:

$$R : \quad \{ \} \models X \rightarrow X \\ A : \quad \{ X \rightarrow YZ, Z \rightarrow VW \} \models X \rightarrow YZV \\ P : \quad \{ X \rightarrow YZ \} \models X \rightarrow Y$$

- b1)** Leiten Sie die  $A$ -Regel des RAP-Kalküls aus dem  $K$ -Kalkül ab.
- b2)** Leiten Sie die  $KR_2$ - und  $KR_3$ -Regel des  $K$ -Kalküls aus dem RAP-Kalkül ab.
- b3)** Sind die folgenden Aussagen richtig oder falsch? Für jede richtige Antwort gibt es 1/2 Punkt und für jede falsche wird 1/2 Punkt abgezogen. Diese Teilaufgabe kann nicht mit weniger als 0 Punkten bewertet werden.

	RICHTIG	FALSCH
Der Kalkül $K$ ist gültig	■	■
Der Kalkül $K$ ist vollständig	■	■
Der Kalkül $K$ ist minimal	■	■

(5 Punkte)

- c)** Gegeben seien die Relationen  $R_1(X, Y, Z)$  und  $R_2(X, Y, Z)$ .  $X$  ist jeweils der Schlüssel der Relation. Die Relationen sind nach dem Attribut  $Y$  sortiert. Das Attribut  $Z$  hat für jeden Tupel den gleichen Wert. Die Relation  $R_1(X, Y, Z)$  enthält  $n$  Einträge, und  $R_2(X, Y, Z)$  enthält  $m$  Einträge.

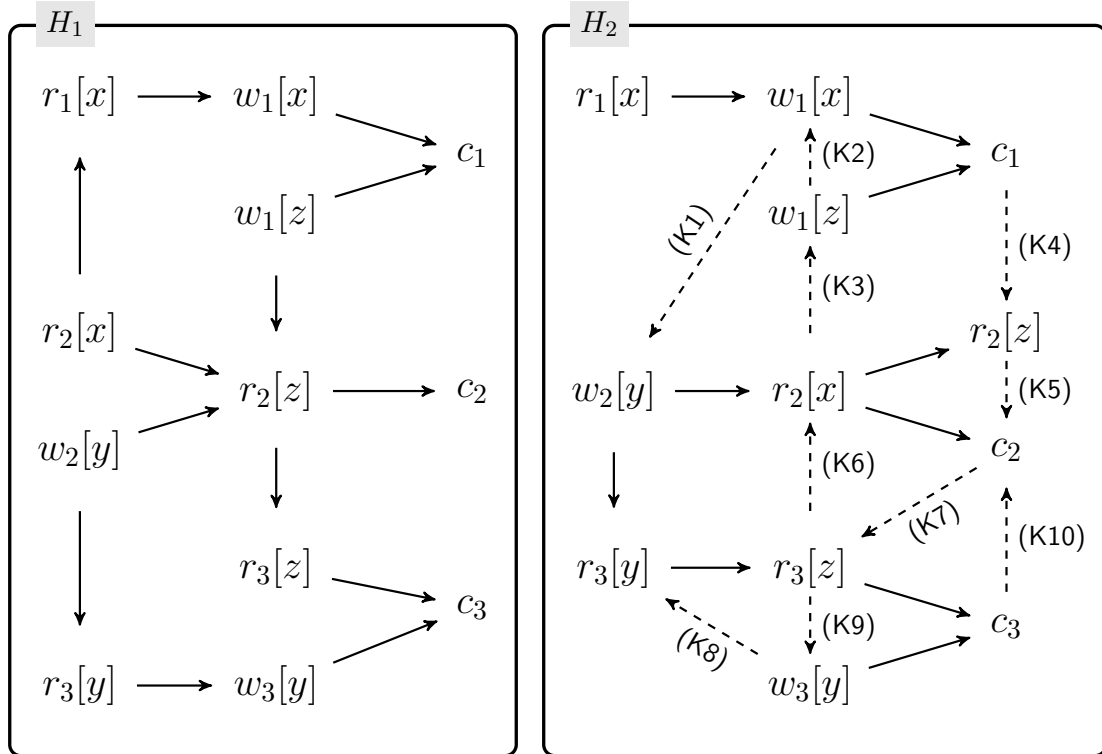
	<table border="1" style="border-collapse: collapse; text-align: center;"> <thead> <tr><th><math>X</math></th><th><math>Y</math></th><th><math>Z</math></th></tr> </thead> <tbody> <tr><td>23</td><td>5</td><td>4</td></tr> <tr><td>7</td><td>5</td><td>4</td></tr> <tr><td>...</td><td>...</td><td>...</td></tr> <tr><td>188</td><td>233</td><td>4</td></tr> </tbody> </table>	$X$	$Y$	$Z$	23	5	4	7	5	4	...	...	...	188	233	4	$\left. \vphantom{\begin{matrix} R_1(X, Y, Z) \\ R_2(X, Y, Z) \end{matrix}} \right\} m$	<table border="1" style="border-collapse: collapse; text-align: center;"> <thead> <tr><th><math>X</math></th><th><math>Y</math></th><th><math>Z</math></th></tr> </thead> <tbody> <tr><td>19</td><td>2</td><td>4</td></tr> <tr><td>25</td><td>3</td><td>4</td></tr> <tr><td>...</td><td>...</td><td>...</td></tr> <tr><td>138</td><td>118</td><td>4</td></tr> </tbody> </table>	$X$	$Y$	$Z$	19	2	4	25	3	4	...	...	...	138	118	4	$\left. \vphantom{\begin{matrix} R_1(X, Y, Z) \\ R_2(X, Y, Z) \end{matrix}} \right\} n$
$X$	$Y$	$Z$																																
23	5	4																																
7	5	4																																
...	...	...																																
188	233	4																																
$X$	$Y$	$Z$																																
19	2	4																																
25	3	4																																
...	...	...																																
138	118	4																																

- c1)** Wie ist die Komplexität des Merge-Joins der beiden Relationen im O-Kalkül (Landau-Symbole), bezogen auf die Anzahl der Tupel  $n$  sowie  $m$ , bei Einsatz des aus der Vorlesung bekannten Verfahrens? Geben Sie jeweils die Komplexität für den Fall an, dass das Join-Attribut  $X$ ,  $Y$  bzw.  $Z$  ist.
- c2)** Gehen Sie nun davon aus, dass die Relationen nicht sortiert vorliegen, und für die Relation  $R_1$  ein Hash-Index über das Attribut  $Y$  eingerichtet ist. Wie würden Sie die Join-Operation zwischen  $R_1$  und  $R_2$  über das Attribut  $Y$  realisieren, um diesen Hash-Index zu nutzen? Was wäre der Aufwand im O-Kalkül?

(5 Punkte)

## Aufgabe 4 – Nebenläufigkeit, Transaktionen, Replikation (16 Punkte)

- a) Gegeben sei die folgende History  $H_1$  sowie eine noch unvollständige History  $H_2$ . Die gestrichelten Pfeile, beschriftet mit (K1) bis (K10), kennzeichnen mögliche Ausführungsreihenfolgen der Operationen zur Ergänzung von  $H_2$ .



- a1) Konstruieren Sie für  $H_1$  den Serialisierbarkeitsgraphen  $SG(H_1)$ . Begründen Sie, ob  $H_1$  serialisierbar ist.
- a2) Gesucht ist eine Teilmenge der möglichen Ausführungsreihenfolgen (K1) bis (K10) der Operationen für  $H_2$ . Wählen Sie die Teilmenge so, dass die resultierende History vollständig **und** konfliktäquivalent zu  $H_1$  ist.
- a3) Analysieren Sie die maximale Rücksetzbarkeitsklasse von  $H_1$ , und geben Sie die dabei zusätzlich notwendigen Annahmen über die Ausführungsreihenfolge der Operationen an.

(10 Punkte)

**b)** Gehen Sie von einer Datenbasis mit fünf Kopien (Replikaten) R1, R2, R3, R4 und R5 aus; die Replikation auf fünf Knoten im Netzwerk wird mit dem *Quorum Consensus*-Protokoll kontrolliert. Weiterhin ist  $p = 3$  (Lese-Quorum) und  $q = 3$  (Schreibquorum) gegeben. Die Datenbasis enthält zwei Variablen  $x$  und  $y$ , die über diese 5 Kopien repliziert sind. Initial enthalten alle Kopien  $x=0$  und  $y=0$ . Nehmen Sie an, dass die folgenden sieben Schritte genau in der vorgegebenen Reihenfolge stattfinden:

1. Die Knoten 4 und 5 und damit die entsprechenden Kopien sind nicht verfügbar.
2. Es findet ein Update statt,  $x := 1$ .
3. Die Kopien R4 und R5 sind wieder verfügbar, aber jetzt sind R2 und R3 nicht verfügbar.
4. Es findet ein Update statt,  $y := 1$ .
5. Die Kopien R2 und R3 sind wieder verfügbar, aber nun ist R1 nicht verfügbar.
6. Es findet eine Leseoperation  $read(x)$  statt; die Kopien R2, R3 und R4 werden für das Lese-Quorum ausgewählt.
7. Es findet eine Leseoperation  $read(y)$  statt; die Kopien R3, R4 und R5 werden für das Lese-Quorum ausgewählt.

**b1)** Geben Sie den Inhalt der fünf Kopien **nach der ersten und nach der zweiten** Update-Operation an. Erklären Sie kurz Ihre Antwort.

**b2)** Geben sie das jeweilige Resultat von beiden Lese-Operationen an.

**b3)** Die Kopie R1 ist die einzige, die in beiden Update-Operationen beteiligt war, d. h. sie ist die einzige, die direkte Kenntnis über beide Updates hat. Wodurch wird garantiert, dass in beiden Fällen die Leseoperationen das richtige Ergebnis liefern, obwohl die Kopie R1 während dieser Leseoperationen nicht verfügbar ist?

**b4)** Welche anderen möglichen Wertekombinationen  $(p,q)$ , **außer**  $(p=3, q=3)$ , könnten generell für das Quorum-Consensus-Protokoll bei fünf Replikaten gewählt werden?

(6 Punkte)