

# 1. Aufgabenblatt (SS 2023)

6.0 VU Datenbanksysteme

## Informationen zum Aufgabenblatt

### Inhalt

Dieser Übungsteil behandelt

- die Modellierung mittels EER Diagrammen
- die Übersetzung von EER Diagrammen in ein relationales Schema
- die relationalen Abfragesprachen Relationale Algebra und Relationenkalkül (Semantik der Operatoren sowie das Verfassen von Anfragen).

Das Aufgabenblatt enthält 8 Aufgaben, auf welche insgesamt 15 Punkte erreicht werden können.

### Deadline

bis 20.04., 14:00Uhr: Upload der Abgabe über TUWEL

### Abgabe

Bitte laden Sie ein ein einziges .pdf Dokument (maximal 10MB) in TUWEL hoch.

Mehr Informationen, Vorgaben und Hinweise zu den abzugebenden Dateien finden Sie in TUWEL direkt bei der Abgabe. Bitte beachten Sie vor allem die Hinweise zu handschriftlichen Abgaben (Einschränkungen und Hinweise zur Lesbarkeit der Abgabe!).

Lösen Sie die Aufgaben eigenständig. Wir weisen Sie darauf hin, dass sämtliche abgeschriebene Lösungen mit 0 Punkten beurteilt werden (sowohl das “Original” als auch die “Kopie”).

### Bei Fragen ...

Bei Fragen (inhaltlicher oder organisatorischer Art) verwenden Sie gerne das TUWEL Forum, kommen Sie zur Sprechstunde und beobachten Sie die Ankündigungen in TUWEL für mögliche weitere Angebote.

Wir laden Sie ein im Forum aktiv mit Ihren Kolleg:innen zu diskutieren. Unserer Erfahrung nach hilft dies allen Beteiligten das Verständnis des Stoffs zu vertiefen. Unterlassen Sie es bitte, (Teile von) Lösungen zu “aktiven” Lösungsblättern zu posten.s

## Aufgaben: EER-Diagramme

### Aufgabe 1 (Semantik von ER Diagrammen)

[1 Punkt]

- (a) Betrachten Sie das folgende ER-Diagramm:



Demonstrieren Sie die Bedeutung der mittels (min, max)-Notation angegebenen Multiplizitäten, indem Sie zuerst eine Liste von Musiker:innen und Liedern angeben, und anschließend wer von diesen Musiker:innen welche Lieder singt.<sup>1</sup> Achten Sie dabei darauf, alle durch das Modell erlaubten Fälle abzudecken.

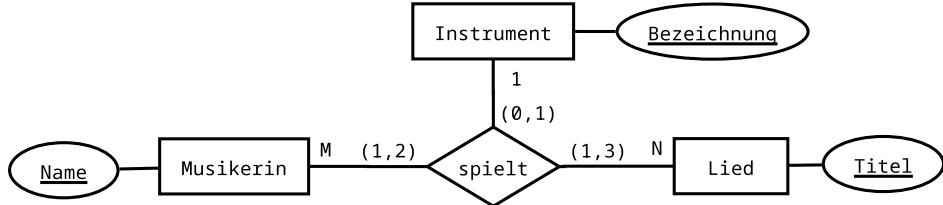
Die Lösung kann in drei einfachen Listen angegeben werden, z.B.:

**Musiker:innen:** Jaskier, The Doctor

**Lieder:** "The Rains of Castamere", "The Road Goes Ever On"

**singt:** A singt B; C singt E

- (b) Betrachten Sie das folgende ER-Diagramm:



In dem Diagramm wird sowohl die Notation mittels Funktionalitäten, als auch die (min,max)-Notation verwendet. (Dies geschieht hier zu Übungszwecken und ist in der Praxis nicht üblich.) Das Diagramm enthält dadurch mehr Informationen als wenn nur eine der beiden Notationen verwendet worden wäre. Demonstrieren Sie diesen Umstand, indem Sie

- Eine konkrete Situation angeben, welche nach den angegebenen Funktionalitäten zulässig wäre, jedoch durch die (min, max)-Notation verboten ist.
- Eine konkrete Situation angeben, welche nach der (min, max)-Notation zulässig ist, jedoch durch die angegebenen Funktionalitäten verboten ist.

Mit "konkrete Situation" ist gemeint dass Sie, ähnlich wie in Unteraufgabe (a), eine entsprechende Anzahl an Entitäten jedes Typs (Musikerin, Instrument, Lied) definieren, sowie die nötigen Beziehungen zwischen diesen Entitäten. Wir empfehlen die Beziehungen als Tripel anzugeben, also z.B. als **spielt:** (M1, I1, L1), (M2, I2, L2)

Achten Sie in beiden Unteraufgaben darauf, dass in den Beziehungen nur Entitäten vorkommen, welche Sie zuvor auch definiert haben.

<sup>1</sup> Alle Personen, Lieder und Beziehungen können künstlich/erfunden/... sein. Es muss nicht die Realität abgebildet werden, sondern es soll die Bedeutung des Diagramms veranschaulicht werden! Auch S1, S2, ... und L1, L2, ... sind zulässige Musikerinnen und Lieder! Gilt genauso für Aufgabe (b).

**Aufgabe 2 (EER-Diagramm erstellen)**

[3 Punkte]

*Pirat Weißhaar muss immer öfter feststellen, dass ihm Details zu seiner Piraten-Organisation entfallen. Auch wenn nicht klar ist, ob das der wachsenden Größe seiner Organisation oder seines Alters geschuldet ist, so ist klar, dass es eine technische Lösung für dieses Problem braucht. Auf der Suche nach solch einer Lösung sollen in einem ersten Schritt die abzubilden Informationen in einem EER-Diagramm gesammelt werden. Für die Erstellung wählt Weißhaar Sie und macht Ihnen ein Angebot, das Sie nicht ablehnen können.*

Zeichnen Sie aufgrund der vorliegenden Informationen (siehe nächste Seite) ein EER-Diagramm. Verwenden Sie dabei die in der Vorlesung vorgestellte Schreibweise, sowie die (min,max)-Notation. Es sind keine NULL-Werte erlaubt (jede Entität eines Typs muss alle Attribute des Entitätstyps besitzen), und Redundanzen sollen vermieden werden. Manchmal kann es notwendig sein, zusätzliche künstliche Schlüssel einzuführen.

Sollten die Werte der (min, max)-Notation in einigen Fällen nicht eindeutig aus der Angabe hervorgehen, dann treffen Sie sinnvolle Annahmen (und dokumentieren diese). (In praktischen Szenarien ist es allerdings ratsam, sich um eine vollständige Spezifierung dieser Fälle zu bemühen, anstatt einfach eine Annahmen zu treffen.)

Eine Unterstützung bei der Erstellung von EER-Diagrammen bietet das Tool dia (<http://wiki.gnome.org/Apps/Dia>, binaries unter <http://dia-installer.de>; Achtung: im Diagramm Editor ER auswählen!). Eine Alternative ist draw.io (<https://www.diagrams.net/>) – neben der Online-Anwendung unter diagrams.net direkt im Browser können Sie sich auch eine Stand-Alone Variante herunterladen. Sie können das EER-Diagramm aber natürlich mit jeder beliebigen Software erstellen – Achten Sie jedoch darauf, das in der Vorlesung vorgestellte Format zu verwenden.

**Beschreibung des zu modellierenden Sachverhalts:**

Weißhaar verfügt über mehrere Schiffe. Jedes Schiff hat einen eindeutigen Namen (NAME), eine Segelfläche (SEGEL) und einen Tiefgang (TIEFE). Handelt es sich um ein Geisterschiff, dann wird zu diesem Schiff die Anzahl an Geistern (GEISTER) sowie der Zeitpunkt der Geisterstunde (GEISTERSTUNDE) gespeichert.

Auf jedem Schiff gibt es mindestens einen Papagei. Diese haben einen Namen (NAME). Auf einzelnen Schiffen ist dieser einmalig, unterschiedliche Schiffe sprechen sich untereinander aber nicht ab – Papageien auf verschiedenen Schiffen können also den gleichen Namen haben. Zu jedem Papagei ist außerdem seine Farbe (FARBE) sowie die Anzahl an Wörtern bekannt, die er sprechen kann (WOERTER). Papageien gibt es nur auf Schiffen, und sie gehen mit ihrem Schiff unter.

Auch alle Pirat:innen haben einen Namen (NAME), darüber hinaus ist ihr Geburtsdatum (GEBDAT), ihr Herkunftsland (LAND) sowie die Anzahl an Schlägereien, in denen sie verwickelt waren, (STREIT) bekannt. Bei der Aufnahme neuer Pirat:innen wird jeweils penibel darauf geachtet, niemals zwei Personen aus dem selben Land aufzunehmen die am selben Tag Geburtstag haben.

Jeder Papagei wird von mindestens einem und maximal drei Pirat:innen gepflegt, wobei kein:e Pirat:in mehr als zwei Papageien pflegen muss.

Es muss vermerkt werden, welcher Pirat/welche Piratin auf welchem Schiff welche Position einnimmt. Positionen haben einen eindeutigen Titel (TITEL), ein ebenso eindeutiges Abzeichen (ABZEICHEN) sowie eine Aufgabenbeschreibung (AUGABE). Auf jedem Schiff muss zumindest eine Position von einer Piratin/einem Piraten eingenommen werden.

Jeder Position steht monatlich eine bestimmte Menge (MENGE) von mindestens einer Rumsorte zu. Rumsorten werden durch die Kombination aus dem Jahr ihrer Herstellung (JAHR) und dem Hersteller (QUELLE) unterschieden. Außerdem führt jedes Schiff verschiedene (jedoch mindestens drei) Sorten Rum mit sich, wobei zu jeder Sorte die mitgeführte Menge (MENGE) bekannt ist.

Weißhaars Organisation operiert von verschiedenen Inseln aus. Jede Insel hat sowohl eine eindeutige Bezeichnung (BEZ) als auch eindeutige Koordinaten (KORD). Des weiteren ist jeweils die Größe (M2) bekannt.

Die verschiedenen Inseln lassen sich in zwei Gruppen einteilen: Schatzinseln und Inseln welche als Basis genutzt werden. Für Basen ist die Anzahl der auf der Insel befindlichen Bars (BARS) und Duschen (DUSCHEN) interessant, sowie welche Schiffe dort stationiert sind (jedes Schiff ist in maximal einer Basis stationiert).

Für Schatzinseln ist die Anzahl der dort befindlichen Schätze bekannt (SCHAETZE). Auf jeder Schatzinsel gibt es außerdem mindestens einen Hinweis auf eine der anderen Schatzinseln, wobei es zu jeder Schatzinsel auf mindestens einer, aber nie mehr als vier, anderen Schatzinseln Hinweise gibt.

**Aufgabe 3 (Überführung ins Relationenschema)**

[2 Punkte]

Überführen Sie das EER-Diagramm aus Abbildung 1 in ein Relationenschema. Das Schema muss unter der Annahme modelliert werden, dass in keinem Attribut Nullwerte erlaubt sind (Sie können dabei annehmen, dass alle für einen Entitätstyp angegebenen Attribute für alle Entitäten dieses Typs existieren; d.h. die Definiertheit sämtlicher Attribute ist 100%). Verwenden Sie unter Einhaltung dieser Einschränkung möglichst wenig Relationen. Unterstreichen Sie sämtliche Primärschlüssel, schreiben Sie die Fremdschlüsselelemente kursiv und stellen Sie sicher, dass ein Fremdschlüssel eindeutig der passenden Relation zugeordnet werden kann.

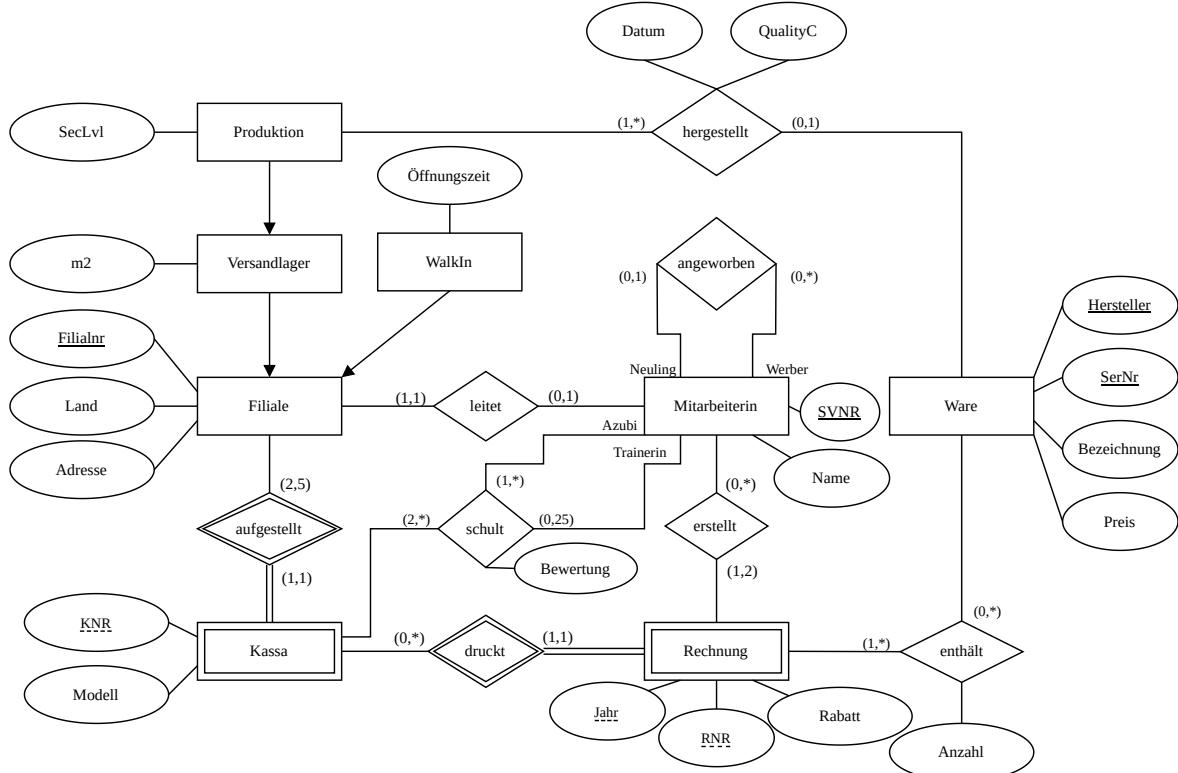


Abbildung 1: EER-Diagramm zu Aufgabe 3

Beispiel, wie eine Antwort aussehen könnte/welche Informationen verlangt sind:

Tabelle1 (T1PK, T1Attribut1, T1Attribut2)

Tabelle2 (T2PK1, T2PK2, T2Attribut1 : T1PK)

Tabelle3 (T3PK, T3A1, T3A2, T3Attribut3)

FK: (T3A1, T3A2) : (Tabelle2.T2PK1, Tabelle2.T2PK2)

## Aufgaben: Relationale Algebra - Relationenkalkül

Um Ihnen die Erstellung Ihrer Abgabe zu den folgenden Aufgaben zu erleichtern, haben wir unter <http://dbai.tuwien.ac.at/education/dm/resources/symbols.html> eine Liste mit den wichtigsten Symbolen der relationalen Algebra zusammengestellt. Sie können diese per copy/paste in Ihr Word/LibreOffice/OpenOffice/... Dokument einfügen. Zusätzlich sind die entsprechenden LATEX Befehle vermerkt.

### Aufgabe 4 (Auswerten)

[1 Punkt]

Gegeben sind die folgenden vier Relationen.

Tournaments		
tname	format	start
Tata Steel Chess	classical	2023-01-13
Airthings Masters	rapid	2023-02-03
Norway Chess	classical	2022-05-30
Superbet Chess Classic	classical	2022-05-03
Sinquefield Cup	classical	2022-09-01
Aimchess Rapid	rapid	2022-10-14

Player	
name	country
Magnus Carlsen	Norway
Ding Liren	China
Anish Giri	Netherlands
Jan-K. Duda	Poland

Winner	
tournament	name
Tata Steel Chess	Anish Giri
Airthings Masters	Magnus Carlsen
Norway Chess	Magnus Carlsen
Aimchess Rapid	Jan-K. Duda

Games		
winner	loser	date
Magnus Carlsen	Anish Giri	2022-01-30
Ding Liren	Jan-K. Duda	2022-01-27
Anish Giri	Ding Liren	2023-01-24
Jan-K. Duda	Anish Giri	2022-11-16
Levon Aronian	Magnus Carlsen	2022-09-22

Geben Sie das Ergebnis der folgenden Anfragen über diesen Relationen an.

(a)

$$\left( \sigma_{format=rapid}(\text{Tournaments}) \times \rho_{tname \leftarrow tournament}(\text{Winner}) \right) \bowtie \text{Player}$$

(b)

$$\pi_{winner, date} \left( \sigma_{name=loser} (\text{Winner} \times \text{Games}) \right)$$

(c)

$$\begin{aligned} & \{ [w.\text{country}, l.\text{country}] \mid w \in \text{Player} \wedge l \in \text{Player} \wedge \\ & \quad \exists g \in \text{Games} (g.\text{winner} = w.\text{name} \wedge g.\text{loser} = l.\text{name} \wedge \\ & \quad \quad \quad g.\text{date} < 2023-01-01) \} \end{aligned}$$

**Aufgabe 5 (Äquivalenzen)**

[2 Punkte]

Gegeben sind die folgenden Relationenschemata  $R(\underline{ABC})$ ,  $S(\underline{BCD})$ ,  $T(\underline{ADF})$  (es gelten keine FKs) sowie die untenstehenden Paare  $p_i, q_i$  an Ausdrücken der Relationalen Algebra. Geben Sie für jedes Paar folgendes an:

- Überprüfen Sie, ob die jeweiligen Ausdrücke äquivalent sind (also ob Sie über allen möglichen Ausprägungen der Schemata immer das gleiche Ergebnis liefern). Sie können dabei davon ausgehen, dass die Ausprägungen keine NULL-Werte enthalten.
  - Begründen Sie Ihre Antwort mit einer kurzen **Erklärung**.
  - Falls die beiden Ausdrücke *nicht* äquivalent aber gültige Abfragen sind, geben Sie zusätzlich noch ein **Gegenbeispiel** an. (Ein Gegenbeispiel besteht aus konkreten Ausprägungen der beteiligten Relationenschemata sowie den Ergebnissen beider Ausdrücke über diesen Ausprägungen.) Das Gegenbeispiel kann entfallen wenn einer der beiden Ausdrücke kein gültiger Ausdruck der Relationalen Algebra ist. In diesem Fall reicht die Erklärung aus.
- (a)  $p_1: \sigma_{A=1}(R) \bowtie (\pi_{B,C}(S) \bowtie \rho_{B \leftarrow A}(T))$  und  
 $q_1: (\sigma_{A=1}(R) \bowtie \rho_{B \leftarrow A}(\pi_{B,C}(S) \bowtie T))$
- (b)  $p_2: \sigma_{C>0}(\pi_{B,C}(R)) \cap \sigma_{C<10}(\pi_{B,C}(R))$  und  
 $q_2: \pi_{B,C}(\sigma_{C>0 \vee C<10}(R))$

**Aufgabe 6 (Größenabschätzung)**

[2 Punkte]

Gegeben sind die Relationenschemata  $R(\underline{ABCE})$ ,  $S(\underline{ABC})$ , und  $T(\underline{ACDE})$  (es gelten keine FKs) sowie eine Ausprägung für jedes dieser Schemata. Von den Ausprägungen ist nur die Anzahl der enthaltenen Tupel bekannt, nicht aber die Tupel selber. Im folgenden bezeichnen  $|R|$ ,  $|S|$  und  $|T|$  jeweils die Anzahl der Tupel in den Ausprägungen von  $R$ ,  $S$  und  $T$ .

- Betrachten Sie die untenstehenden Ausdrücke der Relationalen Algebra. Geben Sie die minimale bzw. maximale mögliche Größe (= Anzahl der Tupel) der Ergebnisse an, die bei der Auswertung dieser Ausdrücke über den gegebenen Ausprägungen entstehen können. Verwenden Sie die gegebenen Werte für  $|R|$ ,  $|S|$  und  $|T|$ .  
(D.h., überlegen Sie sich, welche möglichen Ausprägungen der gegebenen Größe das größte bzw. kleinste Ergebnis erzeugen würden.)
- Geben Sie sowohl für das minimale als auch für das maximal mögliche Ergebnis jeweils **konkrete Ausprägungen** mit der gegebenen Anzahl von Tupeln an, auf welchen die Ausdrücke tatsächlich Ergebnisse der von Ihnen berechneten Größe erzeugen. Geben Sie jeweils außerdem das Ergebnis mit an.
- Begründen Sie kurz, weshalb keine andere Ausprägung ein kleineres bzw. größeres Ergebnis erzeugt.

**Beispiel:** Für die Relation  $V(\underline{XY})$  mit  $|V| = 3$  und den Ausdruck  $\pi_Y(V)$  erhalten wir:

[Minimum: 1 Tupel | Maximum: 3 Tupel]

Konkrete Beispiele sind

V	
X	Y
1	1
2	1
3	1

Ergebnis	
Y	Y
	1

(Minimum) und

V	
X	Y
1	1
2	2
3	3

Ergebnis	
Y	Y
	1
	2
	3

(Maximum)

Ein kleineres Ergebnis als 1 ist nicht möglich, da die Projektion auf einer nicht leeren Relation keine leere Relation zum Ergebnis haben kann. Ebenso ist kein größeres Ergebnis als 3 möglich, da das Ergebnis der Projektion niemals mehr Tupel enthält als die Relation auf die die Projektion angewandt wird.

**Aufgaben:**

- $q_1: \pi_{A,C}(R) \bowtie \pi_{A,B,C,D}(S \bowtie \rho_{F \leftarrow E} T)$  (mit  $|R| = 2$ ,  $|S| = 3$  und  $|T| = 4$ )
- $q_2: (S - \rho_{B \leftarrow A, A \leftarrow B}(S)) \cup \pi_{A,B,C}(R \ltimes T)$  (mit  $|R| = 4$ ,  $|S| = 4$  und  $|T| = 4$ )

**Aufgabe 7 (Abfragesprachen)**

[1 Punkt]

Gegeben sind die Relationenschemata  $R(\underline{ABC}E)$ ,  $S(\underline{CD})$ , und  $T(\underline{BE}F)$ .

In den folgenden Aufgaben ist eine Abfrage in einer der Sprachen aus der Vorlesung gegeben. Die Abfrage soll in die angegebene(n) andere(n) Sprache(n) übersetzt werden.

- (a) Übersetzen Sie die Abfrage

$$\sigma_{A \neq 1}(R - (R \ltimes S))$$

in den Tupel- und den Domänenkalkül.

- (b) Übersetzen Sie die Abfrage

$$\{[b, c, f] \mid \exists e, d, d'([c, d] \in S \wedge [c, d'] \in S \wedge d < d' \wedge [b, e, f] \in T)\}$$

in die Relationale Algebra.

- (c) Übersetzen Sie die Abfrage

$$\{[r.A, r.C] \mid r \in R \wedge \forall t \in T (\exists r' \in R (r.A = r'.A \wedge r.C = r'.C \wedge r.B = t.B \wedge r.E = r.E))\}$$

in den Domänenkalkül.

**Aufgabe 8 (Formalisieren von Anfragen)**

[3 Punkte]

Eine Online-Händler verwendet eine Datenbank um seinen Bestand, Kund:innen und Verkäufe zu verwalten. Die Datenbank hat folgendes Schema (Primärschlüssel sind unterstrichen, Fremdschlüssel sind kursiv geschrieben):

Product	<u>PId</u> , name, description, stock, price
Customer	<u>CId</u> , userName
Review	(PId : Product.PId, CId : Customer.CId, rating, text)
Order	<u>OId</u> , CId : Customer.CId, price)
SoldProduct	(OId : Order.OId, PId : Product.PId, amount)
Wishlist	(PId : Product.PId, CId : Customer.CId)

Formulieren Sie die unten beschriebenen Abfragen jeweils sowohl in der **relationaler Algebra**, dem **Tupelkalkül** und dem **Domänenkalkül**.

- (a) Geben Sie die Kund:innen aus, die mindestens ein Produkt bestellt und mit “1 Stern” bewertet haben.
- (b) Geben Sie die Namen von jenen Produkten aus, die sich auf mindestens zwei verschiedenen Wishlists befinden.
- (c) Geben Sie die Bestellungen aus, bei denen die Kundin/der Kunde keine der in der Bestellung inkludierten Produkte bewertet hat.