

# 1. Übungsblatt

## 3.0 VU Datenmodellierung

12. November 2014

### Allgemeines

In diesem Übungsteil werden Sie den Entwurf einer kleinen Datenbank üben, deren Überführung in das Relationenschema, sowie die relationale Algebra und den Relationenkalkül kennenlernen.

Lösen Sie die Beispiele **eigenständig**, denn bei der Prüfung und in der Praxis sind Sie auch auf sich alleine gestellt. Wir weisen Sie darauf hin, dass abgeschriebene Lösungen mit 0 Punkten beurteilt werden.

Geben Sie ein einziges PDF Dokument ab. Erstellen Sie Ihr Abgabedokument computerunterstützt. Wir akzeptieren keine gescannten handschriftlichen PDF-Dateien.

### Deadlines

<b>28.11.</b>	<b>06:55 Uhr</b>	Upload über den COURSEMANAGER
06.12.	08:00 Uhr	Feedback im COURSEMANAGER verfügbar
07.12.	23:59 Uhr	Reservierung eines Termins für das Tutorengespräch
07.12.	23:59 Uhr	Reservierung eines Termins für das Abgabegespräch

### Tutorengespräch

1. Sie müssen sich über den COURSEMANAGER zu einem Tutorengespräch anmelden. Bitte machen Sie das rechtzeitig, je später Sie sich anmelden, umso eingeschränkter ist das Terminangebot.
2. Wenn Sie kein Blatt abgegeben haben, sind Sie nicht zum Tutorengespräch zugelassen.
3. Sie kommen mit Ihrem Studierendenausweis zu der von Ihnen reservierten Zeit vorbei, und absolvieren das Tutorengespräch. Sie erhalten dabei wertvolles Feedback zu Ihrem Übungsblatt. Das Gespräch ist verpflichtend und dauert ca. 30 Minuten.

## Abgabegespräch

1. Sie müssen sich über den COURSEMANAGER zu einem Abgabegespräch anmelden. Bitte machen Sie das rechtzeitig, je später Sie sich anmelden, umso eingeschränkter ist das Terminangebot.
2. Wenn Sie kein Blatt abgegeben haben oder kein Tutorengespräch geführt haben, sind Sie nicht zum Abgabegespräch zugelassen.
3. Sie kommen mit Ihrem Studierendenausweis zu der von Ihnen reservierten Zeit vorbei, und absolvieren das Abgabegespräch. Stoffgebiet des Abgabegesprächs sind die mit dem Übungsblatt abgedeckten Themengebiete. Wir setzen voraus, dass Sie sich mit Ihrer korrigierten Abgabe auseinandergesetzt haben. Das Abgabegespräch hat Prüfungscharakter.
4. Sie absolvieren Ihr Abgabegespräch gemeinsam mit anderen KollegInnen. Das Gespräch dauert ca. 60 Minuten.
5. Sie können auf die Abgabe maximal 15 Punkte erreichen. Diese setzen sich wie folgt zusammen:  
**5 Punkte** auf das Übungsblatt  
**10 Punkte** auf das Abgabegespräch
6. Die Assistenten tragen die Punkte des Abgabegesprächs in den COURSEMANAGER ein und Sie sehen dort, wieviele Punkte Sie bekommen haben.
7. Falls Sie nicht zu Ihrem Gesprächstermin erscheinen, bekommen Sie automatisch 0 Punkte auf das Gespräch.

## EER

### **Aufgabe 1 (EER-Diagramm erstellen)** [0.5 Punkte]

Zeichnen Sie folgenden Sachverhalt in einem EER-Diagramm. Verwenden Sie dazu die (min, max)-Notation. Nullwerte sind nicht erlaubt. Manchmal kann es notwendig sein, zusätzliche künstliche Schlüssel einzuführen. Eine Unterstützung bei der Erstellung von EER-Diagrammen bietet das Tool dia (<http://live.gnome.org/Dia>). Achtung: im Diagramm Editor EER auswählen!

Die Geschäftsführung des Literaturverlags Leserate möchte eine Datenbank anlegen, um die wichtigsten Informationen über die Veröffentlichungen des Verlags zu speichern.

In der Datenbank soll gespeichert werden, welche Personen an den Veröffentlichungen mitarbeiten. Personen werden durch eine eindeutige ID (PID) identifiziert. Außerdem wird zu jeder Person der Name (NAME), das Geburtsdatum (GEBDAT) und eine Kurzbiographie für die Website des Verlags (KURZBIO) gespeichert.

Bei den Personen kann es sich um Autor/innen, Übersetzer/innen oder Lektor/innen handeln. Eine Person kann auch Autor/in und Übersetzer/in zugleich sein.

Bei Lektor/innen wird in der Datenbank erfasst, welche Sprachen sie beherrschen. Jede Sprache wird dabei durch einen eindeutigen Namen (NAME) identifiziert. Ein Lektor oder eine Lektorin kann beliebig viele Sprachen beherrschen. Bei Übersetzer/innen muss gespeichert werden, in welche Sprachen sie übersetzen können und aus welchen Originalsprachen sie übersetzen können. Es kann sein, dass jemand z.B. englische Texte ins Deutsche übersetzen kann, aber keine deutschen Texte ins Englische.

Für jedes Werk, das im Verlag Leseratte erscheint, wird gespeichert, welche Autor/innen es geschrieben haben. Dabei ist zu beachten, dass auch mehrere Autor/innen gemeinsam ein Werk schreiben können. Ein Werk wird durch den Titel (TITEL) und das Jahr, in dem es geschrieben wurde (JAHR), identifiziert. Es wird auch eine kurze Beschreibung für die Verlagswebsite (BESCHREIBUNG) gespeichert.

Ein Werk kann eine Übersetzung eines anderen Werks sein. In der Datenbank soll gespeichert werden, bei welchen Werken es sich um Übersetzungen handelt und wer sie übersetzt hat. Der Einfachheit halber nehmen wir an, dass es nicht möglich ist, dass mehrere Personen gemeinsam eine Übersetzung eines Werks erstellen. Von einem Werk kann es aber mehrere Übersetzungen geben, die auch von unterschiedlichen Übersetzer/innen sein können.

Zu den Werken wird außerdem noch gespeichert, zu welchen Genres (Science-Fiction-Roman, Gedichtband, ...) sie gehören. Jedes Werk wird einem oder mehreren Genres zugeordnet. Jedes Genre wird durch eine ID (GID) eindeutig identifiziert und hat außerdem noch einen Namen (NAME) und eine kurze Beschreibung (BESCHREIBUNG). Ein Genre kann (muss aber nicht) ein Subgenre von einem oder mehreren anderen Genres sein.

Von jedem Werk, das im Verlag Leseratte erscheint, gibt es eine oder mehrere Auflagen. Eine Auflage wird durch das Werk und durch eine zusätzliche Auflagennummer (NUMMER) eindeutig identifiziert. (Hinweis: Die Auflagennummer allein ist im Allgemeinen nicht eindeutig.) Außerdem soll das Jahr gespeichert werden, in dem die Auflage erscheint (JAHR), und die Lektor/innen, die an der Auflage mitgearbeitet haben. An jeder Auflage arbeitet mindestens ein Lektor/eine Lektorin mit.

Zu einer Auflage eines Werks kann es mehrere Publikationen geben. Publikationen können die Typen Taschenbuch, Hardcover und E-Book haben, wobei es zu einem Werk auch mehrere Publikationen vom selben Typ geben kann. Publikationen werden durch die Auflage und eine zusätzliche ID (PID) identifiziert. Außerdem wird noch die Seitenanzahl (SEITEN) und bei E-Books das Dateiformat (FORMAT) gespeichert. Bei manchen, aber nicht allen Publikationen gibt es einen Klappentext (TEXT), der auch in der Datenbank

gespeichert werden soll.

Schließlich sollen auch noch die Verträge, die der Verlag mit Personen abschließt, in die Datenbank aufgenommen werden. Zu jedem Vertrag enthält die Datenbank eine eindeutige ID (VID), das Abschlussdatum (DATUM), einen kurzen Beschreibungstext (BESCHREIBUNG) und – optional – eine Datei mit dem Vertragstext (DATEI). Außerdem soll gespeichert werden, mit welcher Person der Vertrag abgeschlossen wird. Dabei kann man davon ausgehen, dass Verträge immer mit genau einer Person abgeschlossen werden.

**Lösung:** Siehe Abbildung 1.

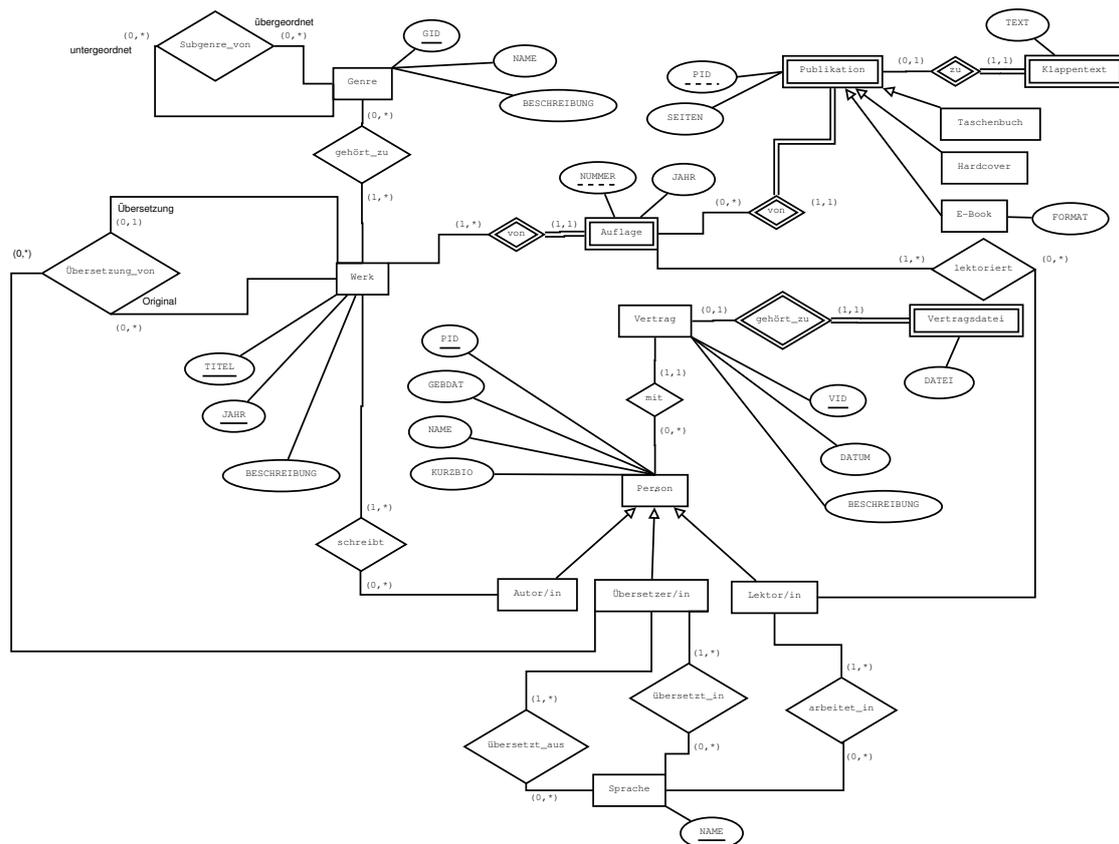


Abbildung 1: Lösung von Aufgabe 1

**Aufgabe 2 (EER-Beispiel selbst erstellen) [0.5 Punkte]**

Ihr Kollege Herr Mustermann hat Sie gebeten, ihm ein wenig Nachhilfe zu geben. Erstellen Sie für ihn ein kleines Übungsbeispiel zum Thema EER-Modellierung. Das Beispiel sollte mindestens 5 Entitäten, 5 Beziehungen, eine Generalisierung und eine schwache Entität enthalten. Vielleicht bringen Sie in dem Beispiel auch noch eine rekursive oder eine ternäre Beziehung unter. Sie erhalten die Punkte für die textuell ausformulierte Angabe sowie für die Musterlösung.

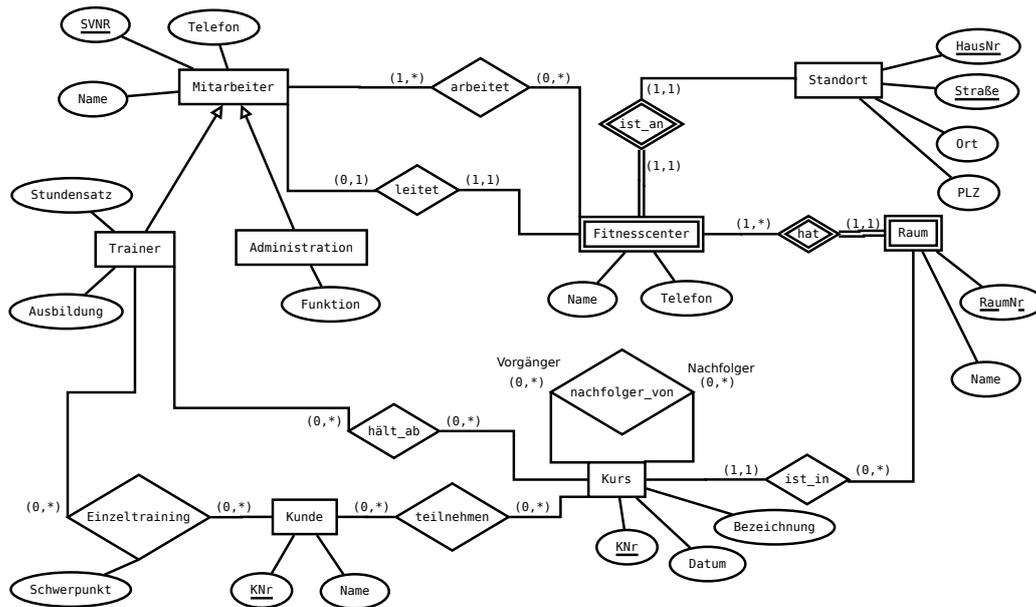


Abbildung 2: EER-Diagramm zu Aufgabe 3

### Aufgabe 3 (Überführung ins Relationenschema) [0.5 Punkte]

Überführen Sie das EER-Diagramm aus Abbildung 2 ins Relationenschema. Nullwerte sind nicht erlaubt. Verwenden Sie möglichst wenig Relationen. Unterstreichen Sie sämtliche Primärschlüssel, schreiben Sie die Fremdschlüssel kursiv und stellen Sie sicher, dass ein Fremdschlüssel eindeutig der passenden Relation zugeordnet werden kann.

#### Lösung:

Mitarbeiter(SVNR, Telefon, Name)  
 Trainer(Mitarbeiter.SVNR, Stundensatz, Ausbildung)  
 Administration(Mitarbeiter.SVNR, Funktion)  
 Standort(Straße, HausNr, Ort, PLZ, Name, Telefon, *Mitarbeiter.SVNR*)  
 Raum(Standort.Straße, Standort.HausNr, RaumNr, Name)  
 Kurs(KNr, Bezeichnung, Datum, *Raum.Straße*, *Raum.HausNr*, *Raum.RaumNr*)  
 Kunde(KNr, Name)  
 arbeitet(Mitarbeiter.SVNR, *Standort.Straße*, *Standort.HausNr*)  
 hält\_ab(Trainer.SVNR, Kurs.KNr)  
 nachfolger\_von(Vorgänger: Kurs.KNr, Nachfolger: Kurs.KNr)  
 teilnehmen(Kurs.KNr, Kunde.KNr)  
 Einzeltraining(Trainer.SVNR, Kunde.KNr, Schwerpunkt)

### Aufgabe 4 (Fehlersuche) [0.5 Punkte]

Gegeben ist das EER-Diagramm aus Abbildung 3.

Dazu wurde das folgende (fehlerhafte!) Relationenschema erstellt:

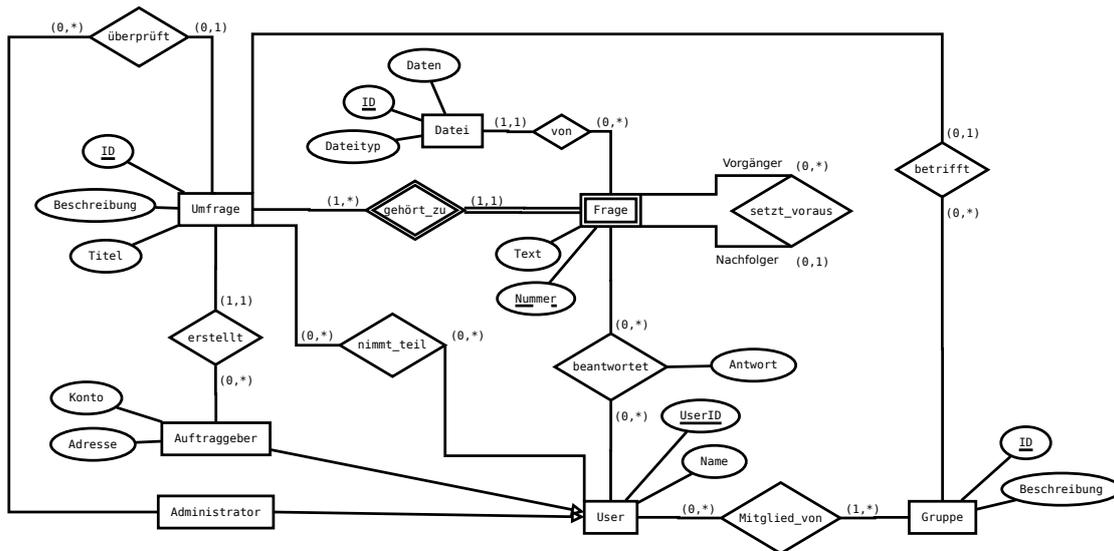


Abbildung 3: EER-Diagramm zu Aufgabe 4

User(UserID, Name)  
 Auftraggeber(User.UserID, Konto, Adresse)  
 Administrator(User.UserID)  
 Umfrage(ID, Beschreibung, Titel, User.UserID, Gruppe.ID)  
 Frage(Umfrage.ID, Nummer, Text, Datei.ID)  
 Datei(ID, Dateityp, Daten)  
 Gruppe(ID, Beschreibung)  
 überprüft(Umfrage.ID, Administrator.UserID)  
 nimmt\_teil(Umfrage.ID, UserID)  
 setzt\_voraus(Frage.ID, Frage.Nummer, Frage.ID, Frage.Nummer)  
 beantwortet(User.UserID, Frage.ID, Frage.Nummer, Antwort)  
 Mitglied\_von(User.UserID, Gruppe.ID)

Finden Sie mindestens 4 Fehler, die sich bei der Überführung des EER-Diagramms ins Relationenmodell eingeschlichen haben, und korrigieren Sie das Relationenschema entsprechend. Wie lautet das korrekte Relationenschema? Beachten Sie, dass NULL-Werte nicht erlaubt sind.

### Lösung:

Folgende Fehler wurden korrigiert:

- Bei der Beziehung `nimmt_teil` ist das Attribut `UserID` nicht als Fremdschlüssel gekennzeichnet.
- Die Beziehung `von` wurde nicht korrekt aufgelöst. Zu einer Frage kann es auch 0 oder mehrere Dateien geben. Daher können die IDs der Dateien nicht als Attribut in der Relation `Frage` gespeichert werden. Stattdessen sollte in der Relation

Datei ein Fremdschlüssel auf die Relation `Frage` mit den Attributen `ID` und `Nummer` gespeichert werden, da jede Datei zu genau einer Frage gehört.

- Die Beziehung `betrifft` wurde falsch aufgelöst. Die Kardinalität (0,1) bedeutet, dass es auch Umfragen geben kann, die keine Gruppe betreffen. Wenn man die `Gruppe.ID` als Fremdschlüssel in der Relation `Umfrage` speichert, können daher NULL-Werte entstehen. Richtig wäre es, eine eigene Relation `betrifft` mit der `Umfrage.ID` als Schlüssel einzuführen.
- `Mitglied_von` ist laut ER-Diagramm eine M:N-Beziehung. Ein User kann also bei mehr als einer Gruppe Mitglied sein. Daher darf der Schlüssel der Relation `Mitglied_von` nicht nur aus der `UserID` bestehen. Die ID der Gruppe muss auch zum Schlüssel gehören.
- Der Fremdschlüssel `UserID` in der Relation `Umfrage` sollte sich auf die Relation `Auftraggeber` und nicht `User` beziehen, da nicht alle User eine Umfrage erstellen dürfen.
- Die beiden Vorkommen der Schlüsselattribute `Frage.ID` und `Frage.Nummer` in der rekursiven Beziehung `setzt_voraus` müssen unterschiedliche Bezeichnungen haben.

Korrigiertes Relationenmodell:

```
User(UserID, Name)
Auftraggeber(User.UserID, Konto, Adresse)
Administrator(User.UserID)
Umfrage(ID, Beschreibung, Titel, Auftraggeber.UserID)
Frage(Umfrage.ID, Nummer, Text)
Datei(ID, Dateityp, Daten, Frage.ID, Frage.Nummer)
Gruppe(ID, Beschreibung)
überprüft(Umfrage.ID, Administrator.UserID)
nimmt_teil(Umfrage.ID, User.UserID)
setzt_voraus(Nachfolger: Frage.ID, Nachfolger: Frage.Nummer, Vorgänger:
Frage.ID, Vorgänger: Frage.Nummer)
beantwortet(User.UserID, Frage.ID, Frage.Nummer, Antwort)
Mitglied_von(User.UserID, Gruppe.ID)
betrifft(Umfrage.ID, Gruppe.ID)
```

## 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 L<sup>A</sup>T<sub>E</sub>X Befehle vermerkt.

**Aufgabe 5 (Primitive Operatoren)** [0.5 Punkte]

Drücken Sie die Operationen Join ( $\bowtie$ ), Durchschnitt ( $\cap$ ), Semi-Join ( $\ltimes$ ) und Anti Semi-Join ( $\triangleleft$ ) mit Hilfe der primitiven Operatoren ( $\sigma$ ,  $\Pi$ ,  $\times$ ,  $\rho$ ,  $-$ ,  $\cup$ ) aus. Nehmen Sie an, dass zwei Relationen  $S$  und  $T$  gegeben sind. Hierbei umfasst das Schema von  $S$   $g + s$  viele Attribute ( $S_1, \dots, S_s, G_1, \dots, G_g$ ) und das Schema von  $T$   $g + t$  viele Attribute ( $T_1, \dots, T_t, G_1, \dots, G_g$ ).

Beim Mengendurchschnitt ( $\cap$ ) ist zu beachten, dass die Schemata der Relationen übereinstimmen müssen. Somit ist  $S \cap T$  kein gültiger Ausdruck in relationaler Algebra. Deshalb betrachten wir hier  $S' \cap T'$  mit  $S' := \Pi_{G_1, \dots, G_g}(S)$  und  $T' := \Pi_{G_1, \dots, G_g}(T)$ .

Beschreiben Sie weiters einen neuen, zusammengesetzten relationalen Operator in natürlicher Sprache und veranschaulichen Sie dessen Anwendung an einem Beispiel. Drücken Sie diesen Operator anschließend mit Hilfe der primitiven Operatoren aus. Der neue Operator soll aus zumindest drei primitiven Operatoren zusammengesetzt werden.

**Lösung:**

$$(a) S \bowtie T \equiv \Pi_{S.G_1, \dots, S.G_g, S_1, \dots, S_s, T_1, \dots, T_t}(\sigma_{S.G_1=T.G_1 \wedge \dots \wedge S.G_g=T.G_g}(S \times T))$$

$$(b) S' \cap T' \equiv S' - (S' - T')$$

$$(c) S \ltimes T \equiv \Pi_{T_1, \dots, T_t, T.G_1, \dots, T.G_g}(\sigma_{S.G_1=T.G_1 \wedge \dots \wedge S.G_g=T.G_g}(S \times T))$$

$$(d) S \triangleleft T \equiv T - \left( \Pi_{T.G_1, \dots, T.G_g, T_1, \dots, T_t}(\sigma_{S.G_1=T.G_1 \wedge \dots \wedge S.G_g=T.G_g}(S \times T)) \right)$$

**Aufgabe 6 (Äquivalenzen)** [0.5 Punkte]

Gegeben sind folgende relationale Ausdrücke über  $R(\underline{A}BC)$ ,  $S(\underline{D}E)$ ,  $T(\underline{B}D)$  und  $U(\underline{A}B)$ . Überprüfen Sie, ob die folgenden Ausdrücke in relationaler Algebra äquivalent sind. Falls ja, so erklären Sie warum. Falls nein, geben Sie eine Begründung an.

$$(a) r_1 : R \bowtie S \text{ und } r_2 : S \times R$$

$$(b) r_3 : \Pi_B(R) \bowtie \Pi_B(T) \text{ und } r_4 : \Pi_B(R \bowtie T)$$

$$(c) r_5 : \Pi_B(R) \bowtie \Pi_B(U) \text{ und } r_6 : \Pi_B(R \bowtie U)$$

$$(d) r_7 : \sigma_{A \neq 3 \vee B > 5}(R \bowtie U) \text{ und } r_8 : \sigma_{A \neq 3}(R \bowtie U) \cup \sigma_{B > 5}(R \bowtie U)$$

$$(e) r_9 : \sigma_{A \neq 3 \wedge B < 5}(R) \text{ und } r_{10} : \sigma_{A \neq 3}(R) \cup \sigma_{B < 5}(R)$$

$$(f) r_{11} : \Pi_D(S) - \Pi_D(T) \text{ und } r_{12} : \Pi_D(S - T)$$

**Lösung:**

(a) **Ja:** Weil  $R$  und  $S$  keine gemeinsamen Attribute haben degeneriert der Join zum Kreuzprodukt.

(b) **Ja:**  $R$  und  $T$  haben nur das Attribut  $B$  gemeinsam, die Ausdrücke auf beiden Seiten liefern daher alle Werte für  $B$ , die sowohl in  $R$  als auch in  $T$  enthalten sind.

- (c) **Nein:** Die gemeinsame Attribute von  $R$  und  $U$  sind  $A$  und  $B$ , der Join in  $r_6$  liefert daher nur die Tupel wo sowohl  $A$  als auch  $B$  passen, während  $r_5$  nur auf Werten von  $B$  joint.
- (d) **Ja:** Die Selektionsbedingung in  $r_7$  enthält ein ODER, welches auch als Vereinigung ausgedrückt werden kann wie in  $r_8$ .
- (e) **Nein:** Hier enthält die Selektion ein UND, welches nicht mittels Vereinigung ausgedrückt werden kann. *Hinweis:* Man kann die Selektion aber aufbrechen zu  $r'_{10} : \sigma_{A \neq 3}(\sigma_{B < 5}(R))$ . Dann ist  $r'_{10}$  äquivalent zu  $r_9$ .
- (f) **Nein:** Der Ausdruck  $r_{12}$  ist kein gültiger Ausdruck in relationaler Algebra, da bei der Mengendifferenz die beiden Relationen die gleichen Schemata haben müssen, aber die Schemata von  $S$  und  $T$  sind unterschiedlich ( $DE$  bzw.  $BD$ ).

**Aufgabe 7 (Größenabschätzung) [0.5 Punkte]**

Gegeben sind die Relationen  $R(\underline{A}BCD)$  mit  $x$  Tupeln,  $S(\underline{B}DE)$  mit  $y$  Tupeln und  $T(\underline{A}CD)$  mit  $z$  Tupeln. Weiters sind folgende Ausdrücke in relationaler Algebra gegeben:

$$\begin{aligned}
 r_1 &: \Pi_B(\rho_{B \leftarrow C}(T)) \cup \Pi_B(R \bowtie S) & r_2 &: \Pi_A(S \bowtie T) \bowtie \Pi_A(R) \\
 r_3 &: \Pi_A(R \bowtie T) \cup \Pi_A(S \bowtie T) & r_4 &: \Pi_A(\rho_{A \leftarrow B}(S) \bowtie T) \cup \Pi_A(R \bowtie T)
 \end{aligned}$$

Geben Sie die minimale bzw. maximale Größe (= Anzahl der Tupel) der durch die folgenden Annahmen entstehenden Relationen an und begründen Sie Ihre Antwort.

- (a) Betrachten Sie  $r_1$  unter der Annahme  $x = 9, y = 12, z = 4$ . **Lösung:** (min 1/max 13)
- (b) Betrachten Sie  $r_2$  unter der Annahme  $x = 9, y = 12, z = 4$ . **Lösung:** (min 0/max 4)

In den folgenden Unteraufgaben sind keine konkreten Werte für  $x, y$  und  $z$  gegeben. Berechnen Sie abhängig von den Variablen  $x, y$  und  $z$  die minimale und maximale Anzahl der Tupel in der Ergebnismenge und begründen Sie Ihre Antwort.

- (c) Betrachten Sie  $r_3$  unter der Annahme  $x \leq y \leq z$ . **Lösung:** (min 0/max  $z$ )
- (d) Betrachten Sie  $r_4$  unter der Annahme  $1 \leq x \leq y \leq z$ . **Lösung:** (min 1/max  $x + y$ )

**Lösung:**

Aufgabe (a):  $\Pi_B(R \bowtie S)$  hat minimal 0 Tupel (kein Tupel aus  $R$  findet einen Joinpartner) und maximal 9 Tupel (jedes Tupel aus  $R$  findet genau eine Partner aus  $S$ ).  $\Pi_B(\rho_{B \leftarrow C}(T))$  hat minimal 1 Tupel (alle Tupel in  $T$  haben den gleichen  $C$  Wert) und maximal 4 Tupel (alle Tupel in  $T$  haben unterschiedliche  $C$  Werte).  $\Pi_B(\rho_{B \leftarrow C}(T)) \cup \Pi_B(R \bowtie S)$  hat somit minimal 1 Tupel und maximal  $4 + 9 = 13$  Tupel.

Aufgabe (b):  $\Pi_A(R)$  hat minimal 1 und maximal 9 Tupel.  $S \bowtie T$  hat mindestens 0 und maximal 48 Tupel.  $\Pi_A(S \bowtie T)$  hat dann minimal 0 und maximal 4 Tupel, da auf das Schlüsselattribut von  $T$  projiziert wird.  $\Pi_A(S \bowtie T) \bowtie \Pi_A(R)$  hat somit minimal 0 Tupel (falls bereits  $\Pi_A(S \bowtie T)$  0 Tupel hat), und maximal 4 Tupel (die maximal 4 Tupel der linken Seite bleiben wegen dem linken äußeren Join erhalten).

Aufgabe (c):  $R \bowtie T$  hat minimal 0 Tupel und maximal  $x$  Tupel. (Jedes Attribut von  $T$  ist auch ein Attribut von  $R$ , daher kann jedes Tupel aus  $R$  maximal einen Joinpartner in  $T$  finden.)  $\Pi_A(R \bowtie T)$  hat minimal 0 Tupel und maximal  $x$  Tupel.  $S \bowtie T$  hat minimal 0 Tupel und maximal  $y \cdot z$  Tupel (jedes Tupel aus  $S$  findet  $z$  Joinpartner in  $T$ ).  $\Pi_A(S \bowtie T)$  hat minimal 0 Tupel und maximal  $z$  Tupel (da auf den Schlüssel von  $T$  projiziert wird).  $\Pi_A(R \bowtie T) \cup \Pi_A(S \bowtie T)$  hat somit minimal 0 und maximal  $z$  Tupel (die  $z$  Tupel der rechten Seite, es können keine weiteren Tupel dazukommen, da jedes der  $x$  Tupel der linken Seite bereits in der rechten Seite enthalten ist).

Aufgabe (d):  $\rho_{A \leftarrow B}(S) \bowtie T$  hat minimal 0 und maximal  $y$  Tupel (da  $A$  Schlüssel in  $T$  ist kann jedes Tupel in  $S$  maximal einen Joinpartner in  $T$  finden).  $\Pi_A(\rho_{A \leftarrow B}(S) \bowtie T)$  hat ebenfalls minimal 0 und maximal  $y$  Tupel.  $R \bowtie T$  hat genau  $x$  Tupel, da alle Tupel der linken Seite jedenfalls erhalten bleiben und jedes Tupel der linken Seite maximal einen Joinpartner aus  $T$  finden kann.  $\Pi_A(R \bowtie T)$  hat minimal 1 Tupel (falls alle  $A$  Attribute in  $R$  den gleichen Wert haben) und maximal  $x$  Tupel (falls alle  $A$  Attribute in  $R$  unterschiedliche Werte haben).  $\Pi_A(\rho_{A \leftarrow B}(S) \bowtie T) \cup \Pi_A(R \bowtie T)$  hat somit minimal 1 Tupel und maximal  $x + y$  Tupel.

### Aufgabe 8 (Formalisieren von Anfragen) [0.5 Punkte]

Gegeben ist eine Personaldatenbank in Form des folgenden Relationenschemas:

M = Mitarbeiter(PNR, Vorname, Nachname, Wohnort)  
A = Abteilung(AID, Name, *Mitarbeiter.PNR*)  
I = arbeitet\_in(*Mitarbeiter.PNR*, *Abteilung.AID*, Zeitraum)

Dabei ist davon auszugehen, dass die in der Relation A gespeicherte PNR die PNR des Abteilungsleiters ist.

Drücken Sie folgende Anfragen in relationaler Algebra aus.

- Gesucht sind die Namen (Vorname und Nachname) der Abteilungsleiter von allen Abteilungen, in denen schon mindestens ein Mitarbeiter gearbeitet hat, der in Wien wohnt.
- Gesucht sind die Namen (Vorname und Nachname) aller Mitarbeiter, die weder in Linz noch in Salzburg wohnen, keine Abteilungsleiter sind und noch nie in der Abteilung „Controlling“ gearbeitet haben.
- Betrachten Sie nun den folgenden Ausdruck in relationaler Algebra. Beschreiben Sie in eigenen Worten, welches Ergebnis dieser Ausdruck liefert.

$$\Pi_{\text{Vorname, Nachname}} \left( M \times \left( \Pi_{\text{PNR, AID}}(I) \div \Pi_{\text{AID}}(A \times \sigma_{\text{Wohnort}='Wien'}(M)) \right) \right)$$

**Lösung:**

(a)  $\Pi_{\text{Vorname, Nachname}} \left( M \times (A \times \Pi_{\text{AID}}(I \times \sigma_{\text{Wohnort}='Wien'}(M))) \right)$

(b) Es gibt viele mögliche Lösungen, zwei davon sind:

$$\Pi_{\text{Vorname, Nachname}} \left( \sigma_{\text{Wohnort} \neq \text{'Linz'} \wedge \text{Wohnort} \neq \text{'Salzburg'}}(M) \times \left( \Pi_{\text{PNR}}(M) - \left( \Pi_{\text{PNR}}(A) \cup \Pi_{\text{PNR}}(I \times \sigma_{\text{Name}='Controlling'}(A)) \right) \right) \right)$$

$$\Pi_{\text{Vorname, Nachname}} \left( \sigma_{\text{Wohnort} \neq \text{'Linz'} \wedge \text{Wohnort} \neq \text{'Salzburg'}}(M) \times \left( \left( \Pi_{\text{PNR}}(M) - \Pi_{\text{PNR}}(A) \right) - \Pi_{\text{PNR}}(I \times \sigma_{\text{Name}='Controlling'}(A)) \right) \right)$$

(c) Gesucht sind die Namen aller Mitarbeiter, die schon in allen Abteilungen gearbeitet haben, deren Abteilungsleiter in Wien wohnen.

**Aufgabe 9 (Relationenkalkül) [0.5 Punkte]**

Übersetzen Sie die Anfragen, die Sie in Aufgabe 8 (a) und Aufgabe 8 (b) in relationaler Algebra formuliert haben, in sichere Ausdrücke des Tupel- und Domänenkalküls. Geben Sie bei jeder der Anfragen an, in welchem Kalkül sie formuliert ist.

**Lösung:**

(a) Tupelkalkül:  $\left\{ [t.\text{Vorname}, t.\text{Nachname}] \mid t \in M \wedge \exists a \in A (t.\text{PNR} = a.\text{PNR} \wedge \exists i \in I (a.\text{AID} = i.\text{AID} \wedge \exists m \in M (i.\text{PNR} = m.\text{PNR} \wedge m.\text{Wohnort} = \text{'Wien'})) \right\}$

Domänenkalkül:  $\left\{ [v, n] \mid \exists p, w ([p, v, n, w] \in M \wedge \exists a, an ([a, an, p] \in A \wedge \exists p', z ([p', a, z] \in I \wedge \exists v', n', w' ([p', v', n', w'] \in M \wedge w' = \text{'Wien'}))) \right\}$

(b) Tupelkalkül:  $\left\{ [t.\text{Vorname}, t.\text{Nachname}] \mid t \in M \wedge t.\text{Wohnort} \neq \text{'Linz'} \wedge t.\text{Wohnort} \neq \text{'Salzburg'} \wedge \neg \exists a \in A (a.\text{PNR} = t.\text{PNR}) \wedge \neg \exists i \in I (i.\text{PNR} = t.\text{PNR} \wedge \exists a \in A (a.\text{AID} = i.\text{AID} \wedge a.\text{Name} = \text{'Controlling'})) \right\}$

$$\text{Domänenkalkül: } \left\{ [v, n] \mid \begin{aligned} &\exists p, w \left( [p, v, n, w] \in M \wedge w \neq \text{'Linz'} \wedge w \neq \text{'Salzburg'} \wedge \right. \\ &\quad \neg \exists a, n' ([a, n', p] \in A) \wedge \\ &\quad \neg \exists a, z ([p, a, z] \in I \wedge \\ &\quad \left. \exists n', p' ([a, n', p'] \in A \wedge n' = \text{'Controlling'}) \right) \end{aligned} \right\}$$

## Gesamter Stoff

### Aufgabe 10 (Falsche Merksätze?) [0.5 Punkte]

Ihr Kollege Karl Konfus erklärt einige wichtige Zusammenhänge, die er sich für die Prüfung merken will. Leider hat er manche Dinge falsch verstanden. Finden Sie die Fehler, begründen Sie warum manche Erklärungen richtig oder falsch sind, und stellen Sie die falschen Aussagen richtig.

- Angenommen, es ist eine Relation **A** mit Schema  $\underline{CD}$  und eine Relation **B** mit Schema  $\underline{CE}$  gegeben. Der Ausdruck  $A \cup B$  in der relationalen Algebra liefert alle Tupel, die in mindestens einer der beiden Relationen vorkommen.
- Mit der (min, max)-Notation kann man in ER-Diagrammen Bedingungen an die Kardinalitäten ausdrücken, die in der M:N-Notation nicht ausdrückbar sind.
- Variablen im Relationenkalkül stehen, wie der Name schon sagt, immer für Relationen. Wenn ich ausdrücken will, dass es in der Relation  $R$  ein Tupel gibt, das für das Attribut  $A$  den Wert 0 hat, kann ich beispielsweise schreiben  $\exists R(R.A = 0)$ .
- Wenn in einer Relation **R** ein Fremdschlüsselattribut vorkommt, ist dieses Attribut immer Teil des Schlüssels und gehört daher im Relationenschema unterstrichen.
- Angenommen, **B** ist ein schwacher Entitätstyp, der von dem übergeordneten Entitätstyp **A** abhängt. Dann ist bei der Beziehung zwischen **A** und **B** auf der Seite von **B** nur die Kardinalität (1, 1) möglich.

### Lösung:

- Falsch:** Die Mengenoperationen  $\cup, \cap$  und  $-$  sind in der relationalen Algebra nur dann anwendbar, wenn beide beteiligten Relationen das gleiche Schema haben. Der Ausdruck  $A \cup B$  ist mit den angegebenen Schemata also überhaupt nicht wohlgeformt.
- Richtig:** Zum Beispiel kann man mit der M:N-Notation nicht modellieren, dass bei einer Beziehung zwischen den Entitätstypen **A** und **B** jeder **A**-Entität mindestens eine **B**-Entität zugeordnet sein muss, wobei keine weiteren Einschränkungen gelten.
- Falsch:** Die Variablen stehen im Tupelkalkül für einzelne Tupel einer Relation und im Domänenkalkül für einzelne Attributwerte, die in einer Relation vorkommen. Variablen für Relationen gibt es aber in keiner Variante des Relationenkalküls. Deshalb ist es auch falsch, z.B.  $\exists R(\dots)$  oder  $\forall R(\dots)$  zu schreiben, wenn  $R$  der Name einer

Relation ist. Eine richtige Formel für das angegebene Beispiel im Tupelkalkül wäre z.B.  $\exists t \in R(t.A = 0)$ .

- (d) **Falsch:** Fremdschlüsselattribute heißen deshalb so, weil sie in einer *anderen* Relation zum Schlüssel gehören. Ein Fremdschlüsselattribut, das in  $R$  vorkommt, muss aber nicht unbedingt zum Schlüssel von  $R$  gehören. Unterstrichen wird es nur, wenn es Teil des Schlüssels von  $R$  ist.
- (e) **Richtig:** Jede  $B$ -Entität muss von einer eindeutigen  $A$ -Entität abhängen. Angenommen,  $A$  hat den Schlüssel  $C$  und  $B$  hat das zusätzliche schwache Schlüsselattribut  $D$ . Wenn die Kardinalität auf der Seite von  $B$   $(1, n)$  mit  $n > 1$  oder  $(1, *)$  wäre, dann könnte es zwei  $B$ -Entitäten mit gleichem  $C$ -Wert geben. Da die  $D$ -Werte nicht global eindeutig sein müssen, wäre es dann möglich, dass die beiden Entitäten auch den gleichen  $D$ -Wert haben und durch den Schlüssel nicht eindeutig identifiziert werden. Wenn die Kardinalität auf der Seite von  $B$   $(0, n)$  für irgendein  $n$  oder  $(0, *)$  wäre, könnte es  $B$ -Entitäten geben, die für das Schlüsselattribut  $C$  gar keinen Wert haben. Also wäre wieder keine eindeutige Identifizierung durch den Schlüssel möglich. Der einzige mögliche Wert ist also  $(1, 1)$ .