

# 1. Übungsblatt

## 3.0 VU Datenmodellierung

5. März 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>21.03.</b>	<b>06:55 Uhr</b>	Upload über den COURSEMANAGER
<i>29.03.</i>	<i>08:00 Uhr</i>	Feedback im COURSEMANAGER verfügbar
<i>30.03.</i>	<i>23:59 Uhr</i>	Reservierung eines Termins für das Tutorengespräch und 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.
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 Leiterin einer Fitnesscenterkette möchte, dass die Verwaltung in Zukunft mithilfe einer Datenbank erfolgt.

Sie möchte außerdem eine Filialsuche auf der Webseite der Kette integrieren. Deshalb wird der Standort jedes Fitnesscenters gespeichert. Dafür ist es notwendig, die Straße (Strasse), die Hausnummer (Nr), die Postleitzahl (Plz) und den Ort (Ort) des Standortes zu kennen. Ein Standort wird eindeutig durch die Straße und die Hausnummer identifiziert.

Ein Fitnesscenter wird eindeutig durch den Standort identifiziert. Es wird auch der Name (Name), die Telefonnummer (TelNr) und die Mailadresse (Email) gespeichert.

Von den Kunden eines Fitnesscenters werden folgende Daten gespeichert: Eine Kundennummer (KNR) dient der eindeutigen Identifikation eines Kunden. Außerdem werden der Vorname (VName), der Nachname (NName), das Geburtsdatum (GebDat), die Adresse (Adresse), die Telefonnummer (Telefon) und falls der Kunde ein Abonnement hat, wird das Ablaufdatum des Abos gespeichert.

Die Mitarbeiter eines Fitnesscenters werden eindeutig durch die Sozialversicherungsnummer (SVNR) identifiziert. Des Weiteren werden der Name (Name), die Telefonnummer (Telefon), die Adresse (Adresse) und das Eintrittsdatum (EintrittsDat) festgehalten. Es werden zwei Arten von Mitarbeitern unterschieden: die administrativen Mitarbeiter und die Trainer. Von den Mitarbeitern der Administration ist die Funktion (Funktion) bekannt, von den Trainern die Ausbildung (Ausbildung) und der Stundensatz (Stundensatz). Ein Fitnesscenter wird von genau einem Mitarbeiter geleitet umgekehrt kann ein Mitarbeiter nur ein Fitnesscenter leiten.

Die Räume werden über eine Nummer (RNr) eindeutig pro Fitnesscenter identifiziert. Zusätzlich haben die Räume einen Namen (Name) und eine Kapazität (Kapazitaet).

Ein Fitnesscenter stellt seinen Kunden verschiedene Geräte zur Verfügung. Jedes dieser Geräte soll in der Datenbank erfasst werden. Ein Gerät hat eine eindeutige Seriennummer (SNr), einen Namen (Name), eine Beschreibung (Beschreibung), ein Anschaffungsdatum (AnschaffDat) und eine Garantiezeit in Monaten (Garantie). Außerdem wird noch gespeichert, in welchem Raum sich das Gerät befindet.

Es werden verschiedene Kurse angeboten. Diese werden durch eine eindeutige Nummer (KNr) identifiziert. Zusätzlich werden die Beschreibung (Beschreibung) wie z.B.: "High Intensity Cardio Programm für Fortgeschrittene", ein Datum (Datum) und die Beginn- und Endzeit (Beginn und Ende) eines Kurses festgehalten. Ein Kurs wird in einem Raum unter der Leitung von ein bis mehreren Trainern abgehalten.

Die Kurse können eine hierarchische Organisation haben. Ein Kurs kann mehrere vorausgehende und nachfolgende Kurse haben. Es gibt aber auch Kurse, die keine vorausgehenden und nachfolgenden Kurse haben.

Die Kunden können an mehreren Kursen teilnehmen. Außerdem werden noch Einzeltrainings angeboten. Dazu ist bekannt, welche Trainer, welche Kunden mit welchem Schwerpunkt (Schwerpunkt) betreuen.

**Lösung:** Siehe Abbildung 1.

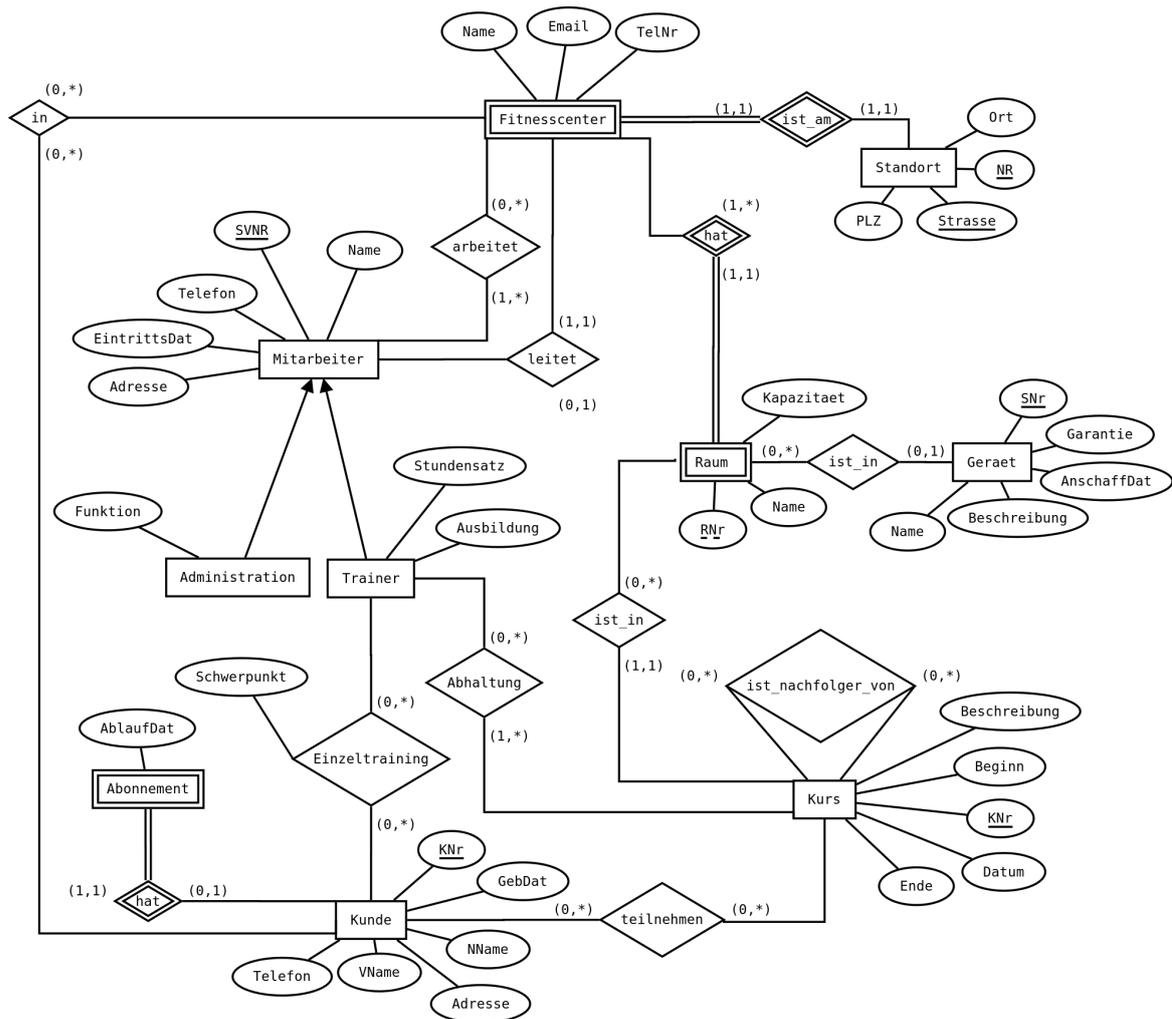


Abbildung 1: Lösung von Aufgabe 1

**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.

**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.

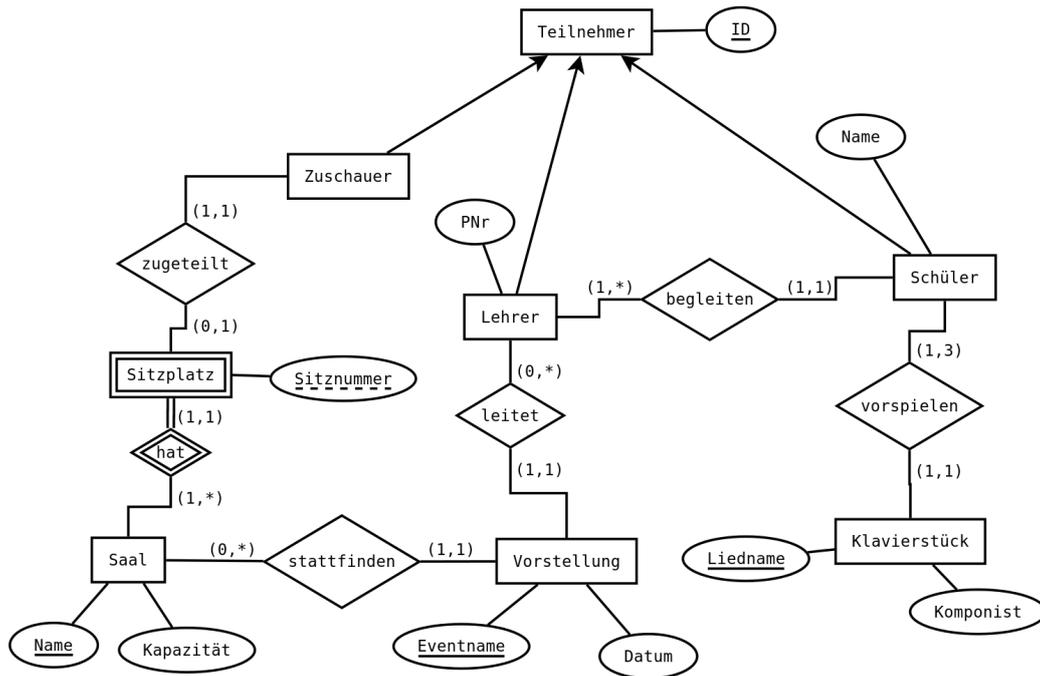


Abbildung 2: EER-Diagramm zu Aufgabe 3

**Lösung:**

Teilnehmer(ID)  
 Zuschauer(Teilnehmer.ID, Sitzplatz.Name, Sitzplatz.Sitznummer)  
 Lehrer(Teilnehmer.ID, PNr)  
 Schüler(Teilnehmer.ID, Name, Lehrer.ID)  
 Klavierstück(Liedname, Komponist, Schüler.ID)  
 Saal(Name, Kapazität)  
 Sitzplatz(Saal.Name, Sitznummer)  
 Vorstellung(Eventname, Datum, Saal.Name, Lehrer.ID)

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

Gegeben ist das EER-Diagramm aus Abbildung 3.

Dazu wurde das folgende (fehlerhafte !) Relationenschema erstellt:

Mitarbeiter(PNR, Anschrift, Name)  
 Abteilungsleiter(Mitarbeiter.PNR, Abteilung.AID)  
 Abteilung(AID, Name)  
 Artikel(ANR, Name, Preis, Abteilung.AID)  
 Rechnungen(RNR, Datum)  
 Kunde(ID)  
 Kundenkarte(ID, Name, Bonuspunkte, Anschrift)  
 Elektrogerät(ID, Bezeichnung, Garantie)

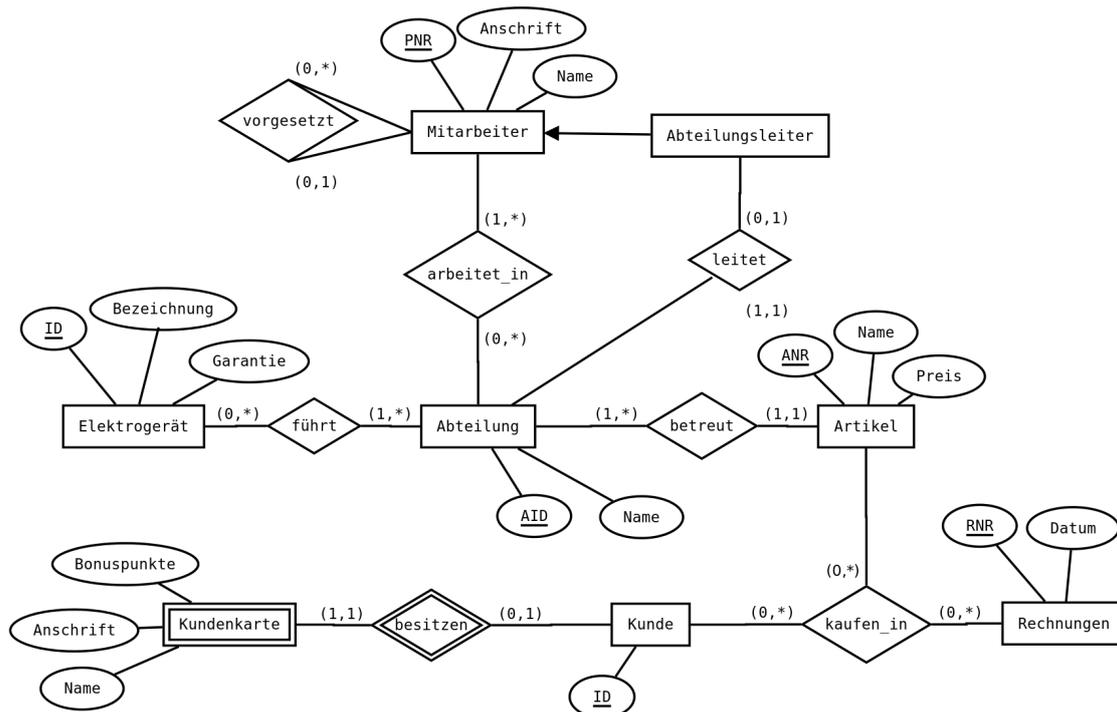


Abbildung 3: EER-Diagramm zu Aufgabe 4

vorgesetzt(*Mitarbeiter.PNR*, *Vorgesetzter: Mitarbeiter.PNR*)  
 arbeitet\_in(*Mitarbeiter.PNR*, *Abteilung.AID*)  
 kaufen\_in(*Kunde.ID*, *Rechnungen.RNR*, *Artikel.ANR*, Anzahl)  
 führt(*Abteilung.AID*)

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:

- Beim Entitätstyp **Abteilungsleiter** wurden die Attribute vom Entitätstyp **leitet** aufgenommen. Jedoch leitet nicht jeder Abteilungsleiter eine Abteilung, eine Abteilung muss aber von einem Abteilungsleiter geleitet werden. Somit würden NULL-Werte entstehen. Es muss das Schlüsselattribut des Entitätstyps **Abteilungsleiter** in **Abteilung** aufgenommen werden.
- Beim Entitätstyp **Kundenkarte** ist die Minimalität des Schlüssels verletzt, da die Attribute **Kunden.ID** und **Name** in Kombination nicht minimal sind. Man muss sich hier für **Kunden.ID** entscheiden, da im EER-Diagramm der Entitätstyp **Kundenkarte**

ein schwacher Entitätstyp ist, und deshalb das Attribut `Kunden.ID` zur eindeutigen Identifizierung notwendig ist. Ferner verweist das Attribut `ID` des Entitätstyps `Kundenkarte` nicht auf `Kunden`.

- Da es sich bei `arbeitet_in` um eine N:M Beziehung handelt, muss auch das Attribut `Abteilung.AID` Schlüssel sein.
- Die ternäre Beziehung `kaufen_in` wurde falsch aufgelöst. Der Schlüssel des Entitätstypen `Rechnungen` (`Rechnungen.RNR`) muss ebenfalls Teil des Primärschlüssels sein. Außerdem ist im EER-Diagramm das Attribut `Anzahl` nicht modelliert.
- Die Beziehung `führt` wurde falsch aufgelöst. Der Schlüssel des Entitätstypen `Elektrogerät` (`Elektrogerät.ID`) muss ebenfalls Teil des Primärschlüssels sein.

Korrigiertes Relationenmodell:

```

Mitarbeiter(PNR, Anschrift, Name)
Abteilungsleiter(Mitarbeiter.PNR)
Abteilung(AID, Name, Abteilungsleiter.PNR)
Artikel(ANR, Name, Preis, Abteilung.AID)
Rechnungen(RNR, Datum)
Kunde(ID)
Kundenkarte(Kunde.ID, Name, Bonuspunkte, Anschrift)
Elektrogerät(ID, Bezeichnung, Garantie)
vorgesetzt(Mitarbeiter.PNR, Vorgesetzter: Mitarbeiter.PNR)
arbeitet_in(Mitarbeiter.PNR, Abteilung.AID)
kaufen_in(Kunde.ID, Rechnungen.RNR, Artikel.ANR)
führt(Abteilung.AID, Elektrogerät.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 ( $\rhd$ ) 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 ( $T_1, \dots, T_t, G_1, \dots, G_g$ ).

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) Beim Mengendurchschnitt ist zu beachten, dass die Schemata der Relationen übereinstimmen müssen. Somit ist  $S \cap T$  kein gültiger Ausdruck in relationaler Algebra. Deshalb betrachten wir  $S' := \Pi_{G_1, \dots, G_g}(S)$  und  $T' := \Pi_{G_1, \dots, G_g}(T)$ . Der Mengendurchschnitt kann nun folgendermaßen ausgedrückt werden:  $S' \cap T' \equiv S' - (S' - T')$
- (c)  $S \ltimes T \equiv \Pi_{S.G_1, \dots, S.G_g, S_1, \dots, S_s} (\sigma_{S.G_1=T.G_1 \wedge \dots \wedge S.G_g=T.G_g} (S \times T))$
- (d)  $S \triangleright T \equiv S - \left( \Pi_{S.G_1, \dots, S.G_g, S_1, \dots, S_s} (\sigma_{S.G_1=T.G_1 \wedge \dots \wedge S.G_g=T.G_g} (S \times T)) \right)$

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

Gegeben sind folgende relationale Ausdrücke über  $R(\underline{AB})$ ,  $S(\underline{CDE})$ ,  $T(\underline{ACD})$  und  $U(\underline{CD})$ . 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)  $r_1 : \sigma_{C=13} (\Pi_{CD}(T) \cap U)$  und  $r_2 : \sigma_{C=13} (U - (U - \Pi_{CD}(T)))$
- (b)  $r_3 : (T \bowtie S) - (\sigma_{C \leq 2}(S \bowtie T) \cap \sigma_{D=24}(T \bowtie S))$  und  $r_4 : \sigma_{C > 2 \wedge D \neq 24}(T \bowtie S)$
- (c)  $r_5 : \Pi_E (\sigma_{A \neq 1 \vee B > 3} (R \times S))$  und  $r_6 : \Pi_E (\sigma_{B > 3} (\Pi_{ABE} (R \times S)) \cup \sigma_{A \neq 1} (\Pi_{ABE} (S \times R)))$
- (d)  $r_7 : \Pi_{CD} ((T \cap T) \bowtie (S \cup S))$  und  $r_8 : \Pi_{CD} (S) \bowtie \Pi_{CD} (T)$
- (e)  $r_9 : \Pi_{DE} (S \bowtie T)$  und  $r_{10} : \Pi_E (T) \times \Pi_D (S)$
- (f)  $r_{11} : \Pi_{CD} (R \bowtie S) - (\Pi_{CD} (S) \bowtie T)$  und  $r_{12} : \Pi_{CD} ((R \bowtie S) - (\Pi_{CD} (S) \bowtie T))$

**Lösung:**

- (a) **Ja:** Hier wurde die Äquivalenz  $X \cap Y \equiv X - (X - Y)$  angewendet (siehe Aufgabe 5(b)).
- (b) **Nein:** Hier wurde versucht die Äquivalenz  $\overline{X \cup Y} \equiv \overline{X} \cap \overline{Y}$  bzw.  $X \cup Y \equiv \overline{\overline{X} \cap \overline{Y}}$  anzuwenden. Allerdings ist beim Umformulieren der Selektion ein Fehler unterlaufen. Wenn das Attribut  $C$  auf Integerwerte beschränkt ist und in  $r_4$  das Selektionsprädikat durch  $C > 2 \vee D \neq 24$  ersetzt wird, gilt die Äquivalenz wieder.
- (c) **Ja:** Hier kann die Selektion aufgebrochen werden. Das Hinzufügen der Projektion ist unkritisch, da alle in der Selektion vorkommenden oder später benötigten Attribute erhalten bleiben.
- (d) **Ja:** Diese Äquivalenz gilt, da  $X \cap X \equiv X$  und  $X \cup X \equiv X$  und weil bei der Projektion alle gemeinsamen Attribute erhalten bleiben.

- (e) **Nein:** Der Ausdruck  $r_{10}$  ist kein gültiger Ausdruck in relationaler Algebra, da die Projektion auf nicht im Schema enthaltene Attribute erfolgt. *Hinweis:* Selbst für den gültigen Ausdruck  $r'_{10} : \Pi_D(T) \times \Pi_E(S)$  gilt die Äquivalenz nicht: Die Projektion in  $r_9$  erfolgt erst nach dem Join, weshalb beim Join auf Gleichheit der Attribute  $C$  und  $D$  getestet wird. In  $r'_{10}$  wird hingegen das Kreuzprodukt berechnet.
- (f) **Nein:** Hier ist ebenfalls zu beachten, dass der Ausdruck  $r_{12}$  kein gültiger Ausdruck in relationaler Algebra ist, da bei der Mengendifferenz die beiden Relationen unterschiedliche Schemata haben ( $ABCDE$  und  $CD$ ).

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

Gegeben sind die Relationen  $R(ABCD)$  mit  $x$  Tupeln,  $S(BCD)$  mit  $y$  Tupeln und  $T(AEBF)$  mit  $z$  Tupeln. Weiters sind folgende Ausdrücke in relationaler Algebra gegeben:

$$r_1 : S - \Pi_{BCD}(\rho_{G \leftarrow B, C \leftarrow E}(T) \bowtie R) \quad r_2 : S - \Pi_{BCD}(\rho_{G \leftarrow B, C \leftarrow E}(T) \bowtie \Pi_{BD}(R))$$

$$r_3 : S - \Pi_{BCD}(\rho_{G \leftarrow B, C \leftarrow E}(T) \bowtie R)$$

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  $x = 7, y = 17, z = 12$ . **Lösung:** (min 10/max 17)
- (b) Betrachten Sie  $r_2$  unter der Annahme  $x = 7, y = 17, z = 12$ . **Lösung:** (min 0/max 17)

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  $x < y < z$  (und  $z \geq 1$ ). **Lösung:** (min 0/max  $y$ )
- (d) Betrachten Sie  $r_3$  unter der Annahme  $x < z < y$  (und  $z \geq 1$ ). **Lösung:** (min  $\max(0, y - x - z + 1)$ /max  $y$ )

**Beispiel:** Der relationale Ausdruck  $\Pi_{BCD}(R) \cup S$  hat unter der Annahme  $x < y$  zumindest  $y$ , jedoch höchstens  $x + y$  Tupel in der Ergebnismenge.

**Lösung:**

Aufgabe (a): Die beiden gejointen Relationen haben die Attribute AC gemeinsam. Im Idealfall findet jedes Tupel in  $R$  einen Joinpartner in der linken Relation (umgekehrt ist dies selbstverständlich nicht möglich). Im schlechtesten Fall wird kein einziger Joinpartner gefunden. Das heißt, das Ergebnis des Joins beinhaltet minimal 0 und maximal 7 Tupel.

Es ist möglich, dass alle 7 Tupel vom Ergebnis des Joins – projiziert auf die Attribute BCD – in  $S$  vorkommen. Werden nun diese 7 Tupel von  $S$  abgezogen, enthält die Ergebnisrelation minimal 10 Tupel. In der Ergebnisrelation sind maximal alle Tupel von  $S$  enthalten. Dies kommt vor, wenn entweder der Join 0 Tupel als Ergebnis liefert, oder kein Tupel vom Ergebnis des Joins in  $S$  vorkommt.

Aufgabe (b): Die rechte Seite des Joins enthält minimal 1 und maximal 7 Tupel. Da die beiden Relationen des Joins keine Attribute gemeinsam haben, degeneriert der Join zu einem Kreuzprodukt und liefert somit minimal 12 und maximal 84 Tupel.

Sind in diesen maximal 84 Tupeln (projiziert auf BCD) alle Tupel von  $S$  enthalten, ist kein Tupel in der Ergebnisrelation. Ist hingegen keines der Tupel vom Ergebnis des Joins (projiziert auf BCD) in  $S$ , so verbleiben 17 Tupel in der Ergebnisrelation.

Aufgabe (c) und (d): Betrachten wir zunächst den linken äußeren Join. Für die minimale Anzahl an Tupel im Ergebnis dieses Joins, nehmen wir an, dass kein Tupel auf der linken Seite einen Joinpartner in  $R$  findet. Der äußere linke Join erhält jedoch alle Tupel auf der linken Seite und füllt die neuen Attribute, die durch die rechte Relation hinzukommen, mit NULL-Werten auf. Dies bedeutet, dass zumindest  $z$  Tupel enthalten sind. Um die maximale Anzahl an Tupel im Ergebnis des Joins zu erreichen, nehmen wir an, dass ein Tupel (nennen wir es  $t$ ) auf der linken Seite mit jedem Tupel in  $R$  joinen kann (ergibt  $x$  Tupel). Die restlichen  $z - 1$  Tupel der linken Seite des linken äußeren Joins können keinen Joinpartner in  $R$  finden, da die Joinattribute AC vor der Umbenennung Schlüsselattribute in  $T$  sind und wir die Relation  $R$  so gewählt haben, dass jedes Tupel in  $R$  mit dem Tupel  $t$  joint. Insgesamt ergibt das minimal  $z$  und maximal  $x + z - 1$  Tupel für das Ergebnis des linken äußeren Joins.

Die maximale Anzahl  $y$  an Tupel in der Ergebnisrelation für (c) und (d) erhalten wir, indem wir annehmen, dass kein Tupel aus  $S$  entfernt wird.

Minimum für (c): Die rechte Seite der Mengendifferenz enthält maximal  $x + z - 1$  Tupel. Es kann nun sein, dass jedes Tupel in  $S$  auch auf der rechten Seite der Mengendifferenz vorkommt (da  $y < z$ ). Somit enthält die Ergebnisrelation minimal 0 Tupel.

Minimum für (d): Die rechte Seite der Mengendifferenz enthält maximal  $x + z - 1$  Tupel. Nun müssen wir zwei Fälle unterscheiden: (1) sind  $x$  und  $z$  genügend groß ( $x + z - 1 \geq y$ ) und wurden die Tupel in den Relationen entsprechend gewählt, so enthält die Ergebnisrelation keine Tupel (analog zu Unteraufgabe (c)); (2) gilt jedoch  $y > x + z - 1$ , so sind zumindest  $y - x - z + 1$  Tupel in der Ergebnisrelation (bei entsprechender Wahl der Tupel in den Relationen). Somit enthält die Ergebnisrelation zumindest das Maximum der beiden Möglichkeiten, das heißt zumindest  $\max(0, y - x - z + 1)$  Tupel.

#### **Zusatzfragen für Wissbegierige:**

Wie würde die Lösung für die Unteraufgabe (d) ohne der Annahme  $z \geq 1$  lauten? Warum würde die präsentierte Lösung in dem Fall nicht (ganz) richtig sein? (Hinweis: Betrachten Sie den Sonderfall  $x = z = 0$  mit  $y \geq 1$ .)

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

Gegeben ist eine Ausstellungsdatenbank in Form des folgenden Relationenschemas:

- A = Ausstellung ( AID, Titel, VonJahr, BisJahr, Ort)

- L = Aussteller ( AName, Anschrift )
- K = Kuenstler ( *Aussteller.AName*, GebDat, Nationalitaet )
- S = Sammlung ( *Aussteller.AName*, Gruendungsjahr, Groesse )
- U = stellt \_ aus ( *Ausstellung.AID*, *Aussteller.AName* )

Drücken Sie folgende Anfragen in relationaler Algebra aus.

- Wir suchen jene großen, vor 1990 gegründeten Sammlungen mit über 100 Werken, die in Innsbruck und Salzburg ausgestellt haben.
- Wir suchen die Geburtsdaten jener nicht-österreichischen Künstler, die noch nie in Graz oder Linz ausgestellt haben.
- Betrachten Sie nun den folgenden Ausdruck in relationaler Algebra. Beschreiben Sie in eigenen Worten, welches Ergebnis dieser Ausdruck liefert.

$$\sigma_{\text{Groesse} < 20}(S) \bowtie \left( U \div \Pi_{AID}(\sigma_{\text{Ort}='Wien' \wedge \text{vonJahr} \leq 2010 \wedge \text{bisJahr} \geq 2010}(A)) \right)$$

**Lösung:**

- $[\sigma_{\text{Gruendungsjahr} < 1990 \wedge \text{Groesse} > 100}(S) \bowtie (U \bowtie \sigma_{\text{Ort}='Innsbruck'}(A))] \bowtie (U \bowtie \sigma_{\text{Ort}='Salzburg'}(A))$
- $\Pi_{\text{GebDat}}(\sigma_{\text{Nationalitaet} \neq \text{'Österreich'}}(K) - (K \bowtie (U \bowtie \sigma_{\text{Ort}='Graz' \vee \text{Ort}='Linz'}(A))))$
- Gesucht werden jene Sammlungen mit weniger als 20 Kunstwerken, die im Jahr 2010 in allen Ausstellungen in Wien ausgestellt haben.

**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:**

- Tupelkalkül:  $\left\{ s \mid s \in S \wedge s.\text{Groesse} > 100 \wedge s.\text{Gruendungsjahr} < 1990 \wedge \exists i_1 \in U (i_1.\text{AName} = s.\text{AName} \wedge \exists i_2 \in A (i_2.\text{AID} = i_1.\text{AID} \wedge i_2.\text{Ort} = \text{'Innsbruck'})) \wedge \exists u_1 \in U (u_1.\text{AName} = s.\text{AName} \wedge \exists u_2 \in A (u_2.\text{AID} = u_1.\text{AID} \wedge u_2.\text{Ort} = \text{'Salzburg'})) \right\}$

$$\text{Domänenkalkül: } \left\{ [n, j, g] \mid [n, j, g] \in S \wedge j < 1990 \wedge g > 100 \wedge \exists a_1, t_1, v_1, b_1, o_1 ([a_1, n] \in U \wedge [a_1, t_1, v_1, b_1, o_1] \in A \wedge o_1 = \text{'Innsbruck'}) \wedge \exists a_2, t_2, v_2, b_2, o_2 ([a_2, n] \in U \wedge [a_2, t_2, v_2, b_2, o_2] \in A \wedge o_2 = \text{'Salzburg'}) \right\}$$

(b) Tupelkalkül:  $\left\{ [k.\text{GebDat}] \mid k \in K \wedge k.\text{Nationalitaet} \neq \text{'Österreich'} \wedge \nexists u \in U \left( u.\text{AName} = k.\text{AName} \wedge \exists a \in A \left( a.\text{AID} = u.\text{AID} \wedge (a.\text{Ort} = \text{'Graz'} \vee a.\text{Ort} = \text{'Linz'}) \right) \right) \right\}$

Domänenkalkül:  $\left\{ [g] \mid \exists n, t \left( [n, g, t] \in K \wedge t \neq \text{'Österreich'} \wedge \nexists a, t, v, b, o \left( [a, n] \in U \wedge [a, t, v, b, o] \in A \wedge (o = \text{'Graz'} \vee o = \text{'Linz'}) \right) \right) \right\}$

## 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.

- (a) Jeder Entität muss ein Schlüssel zugeordnet werden. Hierbei ist es immer möglich einen künstlichen Schlüssel zu erzeugen.
- (b) Bei der Überführung des ER-Diagrammes in das Relationenschema werden bei einer Generalisierung alle Untertypen in den Obertyp integriert, um weniger Relationen zu erzeugen. Das heißt, die Attribute der Untertypen und des Obertyps werden gemeinsam in einer einzigen Relation gespeichert.
- (c) Im Gegensatz zu SQL ist der Domänenkalkül nicht deklarativ, da Variablen an Domänen gebunden werden.
- (d) Jeder noch so komplizierte Ausdruck in relationaler Algebra lässt sich in einen äquivalenten SQL-Ausdruck umwandeln.
- (e) Der Selektionsoperator eliminiert jene Tupel, die das übergebene Selektionsprädikat erfüllen.

### Lösung:

- (a) **Richtig:** Für jede Entität muß ein Schlüssel spezifiziert werden, was auch immer möglich ist. Wenn es sinnvoll ist, kann aber auch ein künstlicher Schlüssel erzeugt werden.
- (b) **Falsch:** Diese Vorgehensweise führt (fast immer) zu NULL-Werten, und ist deshalb nicht sinnvoll.
- (c) **Falsch:** Der Domänenkalkül ist wie auch SQL deklarativ, da nur spezifiziert wird, was die Tupel in der Ergebnismenge erfüllen müssen und nicht etwa wie die Auswertung erfolgen soll.

- (d) **Richtig:** Dies ist möglich, da SQL relational vollständig ist.
- (e) **Falsch:** Die Aussage müsste lauten: Der Selektionsoperator eliminiert jene Tupel, die das übergebene Selektionsprädikat *nicht* erfüllen.