

# 1. Übungsblatt

## 3.0 VU Datenmodellierung

21. April 2015

### 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

- |       |           |  |
|-------|-----------|--|
| 8.5.  | 06:55 Uhr | Upload über den COURSEMANAGER                      |
| 15.5. | 08:00 Uhr | Feedback im COURSEMANAGER verfügbar                |
| 17.5. | 23:59 Uhr | Reservierung eines Termins für das Tutorengespräch |
| 17.5. | 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. Kommen Sie mit Ihrem Studierendenausweis zu der von Ihnen reservierten Zeit vorbei, und absolvieren Sie das Tutorengespräch. Das Gespräch ist verpflichtend und dauert ca. 30 Minuten.

4. Nützen Sie das Tutorengespräch, um wertvolles Feedback zu Ihrem Übungsblatt zu erhalten: Gehen Sie vor dem Gespräch Ihre korrigierten Lösungen durch, identifizieren Sie Unklarheiten, und stellen Sie den Tutoren entsprechend **Fragen**. Trauen Sie sich zuzugeben, wenn Sie etwas nicht verstanden haben. Ihre Fragen haben keinerlei (negativen) Einfluss auf Ihre Note!

## 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://wiki.gnome.org/Apps/Dia>, binaries unter <http://dia-installer.de>). Achtung: im Diagramm Editor ER auswählen!

Mitarbeiter besitzen eine Personalnummer (PNR), über die sie eindeutig identifiziert werden können, eine eindeutige Sozialversicherungsnummer (SVNR) und es werden zusätzlich Vorname (VNAME), Nachname (NNAME), Eintrittsdatum (EINTRITTSDAT) und Gehalt (GEHALT) gespeichert. Es gibt zwei Arten von Mitarbeitern: Organisatoren und Aufnahmeleiter. Bei den Aufnahmeleitern wird auch die Berufserfahrung (ERFAHRUNG) in Jahren gespeichert.

In jedem Studio arbeiten mehrere Mitarbeiter. Ein Mitarbeiter arbeitet in einem oder mehreren Studios. Jedes Studio wird von einem der Mitarbeiter geleitet. Ein Studio wird eindeutig identifiziert durch die Postleitzahl (PLZ) und der Straße (STRASSE). Es werden weiters Ort (ORT) und Hausnummer (HAUSNR) gespeichert. Innerhalb eines Studios gibt es mehrere Räume. Ein Raum ist identifiziert durch das Studio und die Raumnummer (RAUMNR). Ein Raum wird von einem Aufnahmeleiter gewartet, ein Aufnahmeleiter kann mehrere Räume warten. In jedem Raum gibt es ein oder mehrere Instrumente. Instrumente werden durch eine ID (INSTID) identifiziert und haben eine bestimmte Bezeichnung (BEZEICHNUNG).

Bei einem Künstler sind Vorname (VNAME), Nachname (NNAME) und Adresse bekannt. Ein Künstler ist eindeutig bestimmbar durch seinen Künstlernamen (KNAME). Von einem Künstler gibt es Alben, die über den Künstler und eine AlbumID (ALBUMID) gefunden werden können und haben einen Albumnamen (ANAMEN) und ein Genre (GENRE). Ein Album besteht aus mehreren Liedern, die durch das Album und eine LiedID (LIEDID) identifiziert sind und einen Liednamen (LNAME) haben. Ein Lied kann ein Remix mehrerer anderer Lieder sein.

Ein Aufnahmeleiter ist zuständig für die Aufnahme von Liedern. Diese finden in einem oder mehreren Räumen statt. Ein Aufnahmeleiter kann auch Künstler beraten. Organisatoren können Events veranstalten, an denen mehrere Künstler teilnehmen. Events werden durch einen Eventnamen (ENAME) und das Datum (DATUM) identifiziert und spielen einen Umsatz (UMSATZ) ein. Organisatoren schließen außerdem Verträge mit Firmen zu den Alben ab. Hier muss vermerkt werden, über wieviele Alben (ANZAHL) und zu welchem Preis pro Album (STUECKPREIS) die Verträge abgeschlossen wurden. Von den Abnehmern ist der eindeutige Firmenname (FIRMENNAME), die Kontonummer (KONTONR) und die Bankleitzahl (BLZ) bekannt.

## **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.

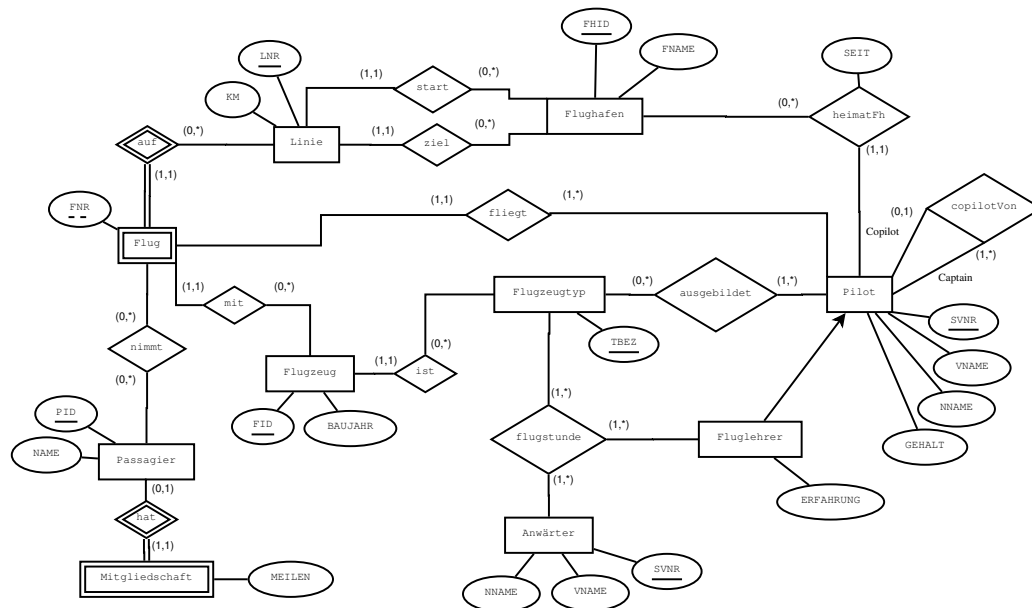


Abbildung 1: EER-Diagramm zu Aufgabe 3

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

Überführen Sie das EER-Diagramm aus Abbildung 1 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.

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

Gegeben ist das EER-Diagramm aus Abbildung 2.

Dazu wurde das folgende (fehlerhafte!) Relationenschema erstellt:

Personal (PID, VNAME, NNAME)  
 Pharmazeut (*Personal.PID*)  
 Referent (*Personal.PID*)  
 Probant (SVNR, NAME)  
 Labor (LABNR, FACHGEBIET)  
 Medikament (MNAME, *Labor.LABNR*)  
 Studie (SID, *Medikament.MNAME*, DAUER)  
 Kunde (KID, KNAME, BEREICH, *Referent.PID*)  
 vertritt (*Medikament.MNAME*, *Referent.PID*)  
 nehmen\_teil (*Studie.SID*, *Probant.SVNR*)  
 untersucht (*Probant.SVNR*, *Pharmazeut.PID*)  
 kauft (*Medikament.MNAME*, *Kunde.KID*)  
 überwacht (*Studie.SID*, *Medikament.MNAME*, *Pharmazeut.PID*)

Finden Sie mindestens 4 Fehler, die sich bei der Überführung des EER-Diagramms ins Relationenmodell eingeschlichen haben, und korrigieren Sie das Relationenschema

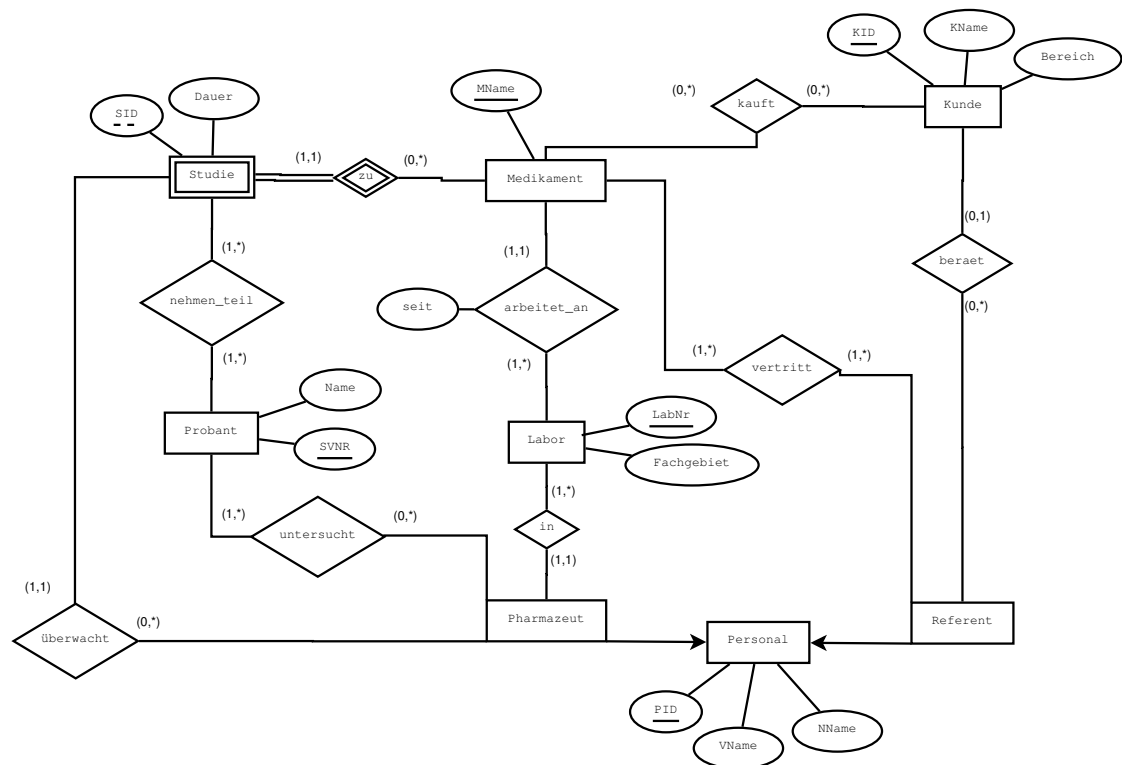


Abbildung 2: EER-Diagramm zu Aufgabe 4

entsprechend. Wie lautet das korrekte Relationenschema? Beachten Sie, dass Nullwerte nicht erlaubt sind.

## 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 ( $\ltimes^c$ ) 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.

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

Gegeben sind folgende relationale Ausdrücke über  $R(\underline{ABC})$ ,  $S(\underline{DE})$ ,  $T(\underline{BD})$  und  $U(\underline{AB})$ . 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)  $r_1: \Pi_{A,B}(R \bowtie U)$  und  $r_2: \Pi_{A,B}(R) \bowtie \Pi_{A,B}(U)$
- (b)  $r_3: \sigma_{A>2}(\sigma_{B<3}(R \bowtie S))$  und  $r_4: \sigma_{B<3}(\sigma_{A>2}(R \times S))$
- (c)  $r_5: \sigma_{A=3 \wedge B<5}(R \bowtie T)$  und  $r_6: \sigma_{A=3}(R \bowtie T) \cup \sigma_{B<5}(R \bowtie T)$
- (d)  $r_7: R - (R \bowtie T)$  und  $r_8: R \triangleright T$
- (e)  $r_9: \Pi_A(R \bowtie U) \cup \Pi_A(R)$  und  $r_{10}: \Pi_A(R)$

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

Gegeben sind die Relationen  $R(\underline{ABCD})$  mit  $r$  Tupeln,  $S(\underline{BDE})$  mit  $s$  Tupeln und  $T(\underline{ACD})$  mit  $t$  Tupeln. Geben Sie die minimale bzw. maximale Größe (= Anzahl der Tupel) folgender Ausdrücke in relationaler Algebra an und begründen Sie Ihre Antwort.

- (a)  $r_1 : \Pi_{A,C,D}(R) \bowtie T, r = 7, t = 4$
- (b)  $r_2 : \Pi_A(T) \cup \Pi_A(R), t = 6, r = 9$
- (c)  $r_3 : \sigma_{E < 3}(S \bowtie T), s = 12, t = 3$
- (d)  $r_4 : T \bowtie \rho_{A \leftarrow D}(S), t = 5, s = 8$

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

Gegeben ist eine Radrennndatenbank in Form des folgenden Relationenschemas:

- R = Radrennen ( RID, Name, Jahr, Länge, Ort )
- F = Fahrer ( Vorname, Nachname, GebDat, Nationalität )
- N = nimmt\_teil ( Fahrer.Vorname, Fahrer.Nachname, Radrennen.RID, Platzierung )

Drücken Sie folgende Anfragen in relationaler Algebra aus.

- (a) Wir suchen die Namen jener Rennen, welche nach 1990 in Linz oder Graz stattgefunden haben, und an denen mindestens ein Italiener teilgenommen hat.
- (b) Wir suchen die Namen jener Fahrer, die noch nie ein Rennen mit mehr als 100km in Wien gefahren sind.
- (c) Betrachten Sie nun den folgenden Ausdruck in relationaler Algebra. Beschreiben Sie in eigenen Worten, welches Ergebnis dieser Ausdruck liefert.

$$\sigma_{\text{Nationalität}='Österreich'}(F) \bowtie \left( \sigma_{\text{Platzierung}=1}(N) \div \Pi_{RID}(\sigma_{\text{Ort}='Wien'}(R)) \right)$$

### 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 Domainkalküls. Geben Sie bei jeder der Anfragen an, in welchem Kalkül sie formuliert ist.

## 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) Wird der Wert jedes Tupels einer Relation durch ein einziges Attribut bereits eindeutig identifiziert, so muss dieses Attribut als Primärschlüssel dieser Relation gewählt werden.

- (b) Bei der Überführung des ER-Diagrammes in das Relationenschema müssen Relationen mit gleichem Schlüssel zusammengefasst werden wenn das Speichern von Nullwerten vermieden werden soll.
- (c) Der Domänenkalkül ist deklarativ, da nur spezifiziert wird, was die Tupel in der Ergebnismenge erfüllen müssen und nicht etwa wie die Auswertung erfolgen soll.
- (d) Der Divisionsoperator in der relationalen Algebra hat keine direkte Entsprechung in SQL. Abfragen in relationaler Algebra, welche den Divisionsoperator beinhalten, können daher im Allgemeinen nicht in SQL ausgedrückt werden.
- (e) Der Semi-Join  $L \bowtie R$  von  $L$  mit  $R$  liefert  $L$ , falls  $L$  und  $R$  keine gemeinsamen Attribute besitzen.