

1. Übungsblatt

3.0 VU Datenmodellierung

13. März 2013

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

05.04. 06:55 Uhr Upload über den COURSEMANAGER
13.04. 12:00 Uhr Feedback im COURSEMANAGER verfügbar

Hinweise

1. Für dieses Übungsblatt werden maximal 5 Punkte vergeben.
2. Sie müssen mindestens einen Punkt auf die Abgabe bekommen. Wenn Sie weniger als einen Punkt auf das Blatt bekommen, oder kein Blatt abgegeben haben, sind Sie nicht zur ersten Leistungsüberprüfung zugelassen.
3. Im COURSEMANAGER ist ersichtlich, wieviele Punkte Sie auf die Abgabe bekommen haben. Weiters können Sie dort auf das Feedback zu Ihrer Abgabe zugreifen.

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. Eine Unterstützung bei der Erstellung von EER Diagrammen bietet das Tool dia (<http://live.gnome.org/Dia>). Achtung: im Diagramm Editor EER auswählen!

Für eine Firma, die internationale Projekte durchführt, soll eine Datenbank angelegt werden, um bei Schulungen und Projekten zukünftig einen besseren Überblick zu bekommen.

Die Firma ist in mehreren Stockwerken angesiedelt wobei jeder Stock eindeutig durch die Stocknummer (STOCK) identifiziert ist. In einem Stock befinden sich mehrere Zimmer mit einer pro Stock eindeutigen Zimmernummer (ZNR) und einer Traktbezeichnung (TRAKT) welche nur mit A, B oder C befüllt ist. Ein Zimmer kann ein Meetingraum oder ein Büro sein. Meetingräume haben eine Kapazität (KAP) und eine Info, ob der Meetingraum für Telefonkonferenzen nutzbar ist (TELKO). Bei einem Büro sind die Anzahl der Schreibtische bekannt (TISCHE) als auch die Anzahl der digitalen Telefone (VOIP). Eine Abteilung ist in einem oder auch mehreren Stockwerken angesiedelt, wobei von dieser der Name (NAME) und eine eindeutige Kennnummer (KNR) gespeichert werden soll. In einer Abteilung arbeiten verschiedene Teams an verschiedenen Projekten. Ein Team setzt sich aus mehreren Mitarbeitern zusammen, hier soll gespeichert werden seit wann (SEIT) der Mitarbeiter in dem Team ist. Das Team ist durch eine eindeutige Teamnummer (TEAMNR) gekennzeichnet. Zusätzlich hat es auch einen Namen (NAME). Ein Mitarbeiter hat eine eindeutige Personalnummer (PNR), eine eindeutige Sozialversicherungsnummer (SVNR), er hat eine Adresse (ADRESSE), eine Handynummer (MOBILE), ein Gehalt (GEHALT) und ein Eintrittsdatum (SEIT). Eine Abteilung wird von genau einem Mitarbeiter geleitet, das gleiche gilt für ein Team - hier wird genau ein Teamleiter festgelegt. Jeder Mitarbeiter kann in mehreren Teams mitwirken, jedoch maximal ein Team und maximal eine Abteilung leiten. Weiters ist anzugeben, wer der Vorgesetzte eines Mitarbeiters ist. Es gibt auch Schulungen, die ein Mitarbeiter besuchen kann. Eine Schulung ist eindeutig identifizierbar durch das Thema (THEMA) und das Datum (DATUM) an dem sie stattfindet. Außerdem wird auch ein interner Preis (PREIS) gespeichert. Es ist bekannt welche Mitarbeiter welche Schulungen besucht haben. Ein Projekt wird durch seine Projektnummer identifiziert (PNR), es hat einen Namen (NAME), eine Durchführungssprache (SPRACHE) und auch ein Projektbudget (BUDGET) zur Verfügung. Projekte werden für verschiedene Länder durchgeführt (z.B. wird das Projekt Maut4712 für Belgien, Russland und Australien realisiert). Ein Land wird durch den Ländercode (LCODE) identifiziert, eine festgelegte Verhandlungssprache (SPRACHE) ist ebenso bekannt. Es soll weiters gespeichert werden,

in welchen Projekten welche Mitarbeiter welche Aufgabenbereiche übernehmen. Ein Aufgabenbereich wird durch seine Bezeichnung (BEZEICHNUNG) identifiziert.

Aus strategischen Gründen möchte die Führungsebene die neue Datenbank nutzen, um die Meetingkultur ein wenig verbessern. Dazu ist es notwendig, projektspezifische Meetings in die Datenbank aufzunehmen. Hierfür werden Meetings gespeichert, die durch eine Meetingnummer (MNR) identifiziert werden, ein Datum (DATUM), und die Zeitspanne (VON), (BIS) werden ebenso vermerkt. Weiters soll es optional möglich sein, pro Meeting einen Link (LINK) zu vermerken, wo alle Unterlagen aufgelistet sind, damit diese später leichter aufzufinden sind. Ein Meeting wird für ein Projekt, in einem Meetingraum, mit mehreren Mitarbeitern abgehalten. Es gibt auch Meetings, die keinem Projekt zugeordnet sind.

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

Ihr Kollege Harald 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 folgendes EER-Diagramm (Abbildung 1) ins Relationenschema. Nullwerte sind nicht erlaubt. Beachten Sie hierbei, dass Sie möglichst wenige Relationen verwenden. 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 folgende EER-Diagramm (Abbildung 2).

Dazu wurde das folgende (fehlerhafte !) Relationenschema erstellt:

Adresse (AID, STRASSE, ORT)
Mitarbeiter (SVNR, NAME, POSITION, *Adresse.AID*)
Dienstplan (DATUM, VON, BIS)
Ruhepause (*Mitarbeiter.SVNR*, *Dienstplan.DATUM*, VON, BIS)
Kunde (KID, SEIT, *Adresse.AID*)
Firma (*Kunde.KID*, USTNR, EMAIL, FNAME)
Person (NAME, TELEFON)
betreut (*Mitarbeiter.SVNR*, *Kunde.KID*)
Typ (TID, BEZEICHNUNG)
Fahrzeug (FGNR, FARBE, MARKE)
ist_von (*Type.ID*, *Fahrzeug.FGNR*)
hat (*Kunde.KID*, *Fahrzeug.FGNR*)

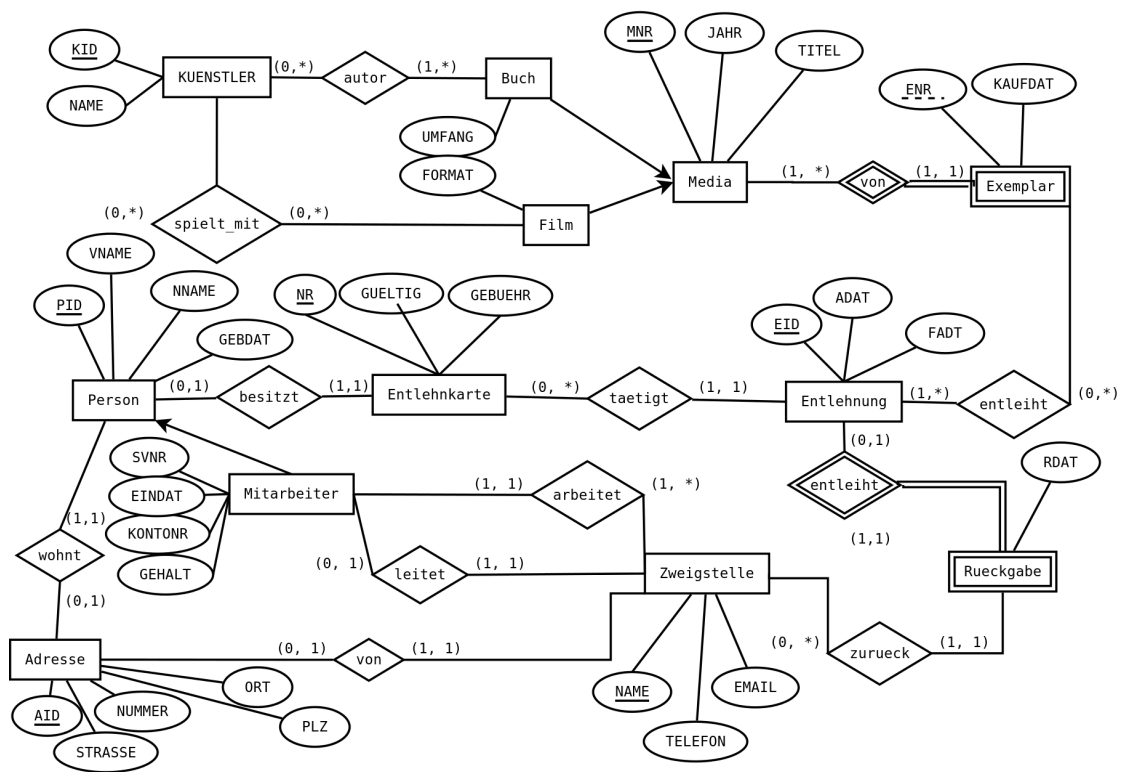


Abbildung 1: EER Diagramm zu Aufgabe 3

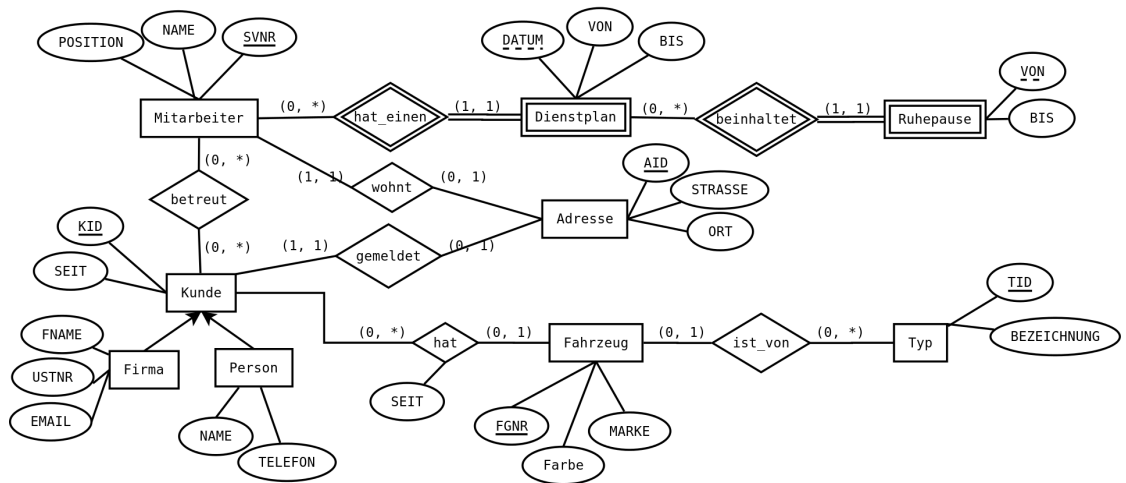


Abbildung 2: EER Diagramm zu Aufgabe 4

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.

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/OpenOffice/... Dokument einfügen. Zusätzlich sind die entsprechenden L^AT_EX 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 (σ , Π , \times , ρ , $-$, \cup) aus.

Aufgabe 6 (Äquivalenzen) [0.5 Punkte]

Gegeben sind folgende relationale Ausdrücke über $R(\underline{BCA})$, $S(\underline{BD})$ und $T(\underline{AC})$. 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) $r_1 : \sigma_{C \neq 2}(S \bowtie T)$ und $r_2 : S \bowtie \sigma_{C \neq 2}(T)$
- (b) $r_3 : \Pi_C \left(\sigma_{C \geq 50} \left(\Pi_{BC}(R) \cap \sigma_{C > 30}(\Pi_{BC}(R)) \right) \right)$ und $r_4 : \sigma_{C \geq 50}(\Pi_C(R))$
- (c) $r_5 : R \bowtie T$ und $r_6 : T \bowtie \Pi_{BC}(R)$
- (d) $r_7 : \sigma_{C \neq 12 \wedge \neg A = 13}(\Pi_{AC}(R) \cup T)$ und $r_8 : \sigma_{A \neq 13}(\sigma_{C \neq 12}(T \cup \Pi_{AC}(R)))$
- (e) $r_9 : \sigma_{C=12}(T) \cap \sigma_{A > 30}(\Pi_{AC}(R))$ und $r_{10} : \sigma_{A > 30 \wedge C=12}(\Pi_{AC}(R) \cup T)$
- (f) $r_{11} : \Pi_{AD}(R \bowtie S)$ und $r_{12} : \Pi_A(R) \bowtie \Pi_D(S)$

Aufgabe 7 (Größenabschätzung) [0.5 Punkte]

Gegeben sind die Relationen $R(\underline{ABD})$ mit 11 Tupeln, $S(\underline{CDE})$ mit 3 Tupeln und $T(\underline{BEA})$ mit 8 Tupeln.

Geben Sie die minimale bzw. maximale Größe (= Anzahl der Tupel) der durch die folgenden Ausdrücke entstehenden Relationen an:

- (a) $\Pi_{AE}(T) \cup \Pi_{AE}(R \bowtie S)$
- (b) $S \bowtie T$
- (c) $\Pi_E(T \times \sigma_{B < 5}R) - \Pi_E S$

Aufgabe 8 (Formalisieren von Anfragen) [0.5 Punkte]

Gegeben ist das folgende Relationenschema:

- P = Pizza(PizzaName, Kalorien, Preis, Besonderheit)
- K = Kunde(KundenNr, Name, Wohnort, Adresse, Alter)
- S = Sonderangebot(PizzaName, Monat, Jahr, Rabatt)
- B = bestellt(KundenNr, PizzaName, Tag, Monat, Jahr)

Drücken Sie folgende Anfragen in relationaler Algebra aus.

- Wir suchen die Kundennummer aller jungen Leute (unter 30) aus Wien, die eine Pizza mit mehr als 5000 Kalorien bestellt haben, während diese als Sonderangebot angeboten wurde.
- Gesucht sind Kundennummer, Preis, und Kalorien aller Bestellungen von vegetarischen Pizzen (Besonderheit='vegetarisch'), welche noch nie im Sonderangebot waren.

Betrachten Sie nun den folgenden Ausdruck in relationaler Algebra. Beschreiben Sie in eigenen Worten, welches Ergebnis dieser Ausdruck liefert.

$$(c) \sigma_{\text{Wohnort} \neq \text{'Wien'}} \left(K \times \left(\Pi_{\text{PizzaName, KundenNr}}(B) \div \Pi_{\text{PizzaName}} \left(\sigma_{\text{Jahr}=2012}(S) \right) \right) \right)$$

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 Harald Heast 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 falsch sind, und stellen Sie die falschen Aussagen richtig.

- Wenn ich ein ER Diagramm in ein Relationenschema überführe, muß ich für jede Generalisierung einen neuen künstlichen Schlüssel erzeugen.
- Weiters haben rekursive N:M Beziehungen den Vorteil, daß die Schlüsselattribute im Relationenschema nur einmal gespeichert werden müssen, womit Speicherplatz gespart wird.

- (c) Alle Operatoren in relationaler Algebra können mit Hilfe der Basisoperatoren – sogar ohne Kreuzprodukt “ \times ” – ausgedrückt werden. Dadurch kann man komplexere Ausdrücke einfach mit Hilfe dieser wenigen Operatoren formulieren.
- (d) Ein Schlüssel muss immer minimal sein. Das heißt ein Schlüssel darf keine überflüssigen Attribute enthalten.
- (e) Der Begriff eines *sicheren Ausdrucks* lässt sich für den Tupelkalkül einfacher definieren als für relationale Algebra.