

1. Übungsblatt (WS 2021)

3.0 VU Datenmodellierung / 6.0 VU Datenbanksysteme

Informationen zum Übungsblatt

Allgemeines

In diesem Übungsteil entwerfen Sie eine kleine Datenbank mittels EER-Diagrammen, überführen ein EER-Diagramm in ein Relationenschema, und üben den Umgang mit der relationalen Algebra und dem Relationenkalkül.

Lösen Sie die Beispiele **eigenständig** (auch bei der Prüfung und vermutlich auch in der Praxis sind Sie auf sich alleine gestellt)! Wir weisen Sie darauf hin, dass sämtliche abgeschriebene Lösungen mit 0 Punkten beurteilt werden (sowohl das “Original” als auch die “Kopie”).

Geben Sie ein einziges PDF Dokument ab (max. 5MB). Erstellen Sie Ihr Abgabedokument computerunterstützt. Wir akzeptieren **keine PDF-Dateien mit handschriftlichen Inhalten**.

Das Übungsblatt enthält 8 Aufgaben, auf welche Sie insgesamt 15 Punkte erhalten können.

Deadlines

bis **04.11. 12:00Uhr** Upload der Abgabe über TUWEL
ab **17.11. 13:00Uhr** Korrektur und Feedback in TUWEL verfügbar

Weitere Fragen – TUWEL Forum

Sie können darüber hinaus das TUWEL Forum verwenden, sollten Sie inhaltliche oder organisatorische Fragen haben. **Posten Sie auf keinen Fall Ihre (partielle) Lösungen im Forum!**

Änderungen im Ablauf bzgl. COVID-19

Wegen der andauernden Sondersituation werden heuer keine Sprechstunden zum Übungsblatt stattfinden. Bitte wenden Sie sich stattdessen verstärkt an das TUWEL Forum wenn Sie Probleme damit haben die Angaben zu verstehen oder wenn Sie technische Schwierigkeiten haben.

Wenn möglich empfehlen wir Ihnen auch das Forum zur Diskussion mit Ihren Kommilitonen zu Nutzen. Ein gemeinsames Analysieren von Problemen hilft erfahrungsgemäß allen Beteiligten dabei den Stoff besser zu verstehen.

Aufgaben: EER-Diagramme

Aufgabe 1 (EER-Diagramm erstellen)

[3 Punkte]

Sie haben Ihre Wertpapiere über diverse Depots verstreut und somit den Überblick verloren. Die Portfoliomanagement-Tools die Sie nach einer Google-Suche fanden, entsprachen allerdings alle nicht Ihren Ansprüchen ... daher besteht nur eine einzige Lösung: Sie entwickeln Ihr eigenes. Unter anderem umfasst dies auch den Entwurf der relationalen Datenbank.

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, und Redundanzen sollen vermieden werden. 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!). Sie können das EER-Diagramm aber natürlich mit jeder beliebigen Software erstellen.

Beschreibung des zu modellierenden Sachverhalts:

Ein Broker hat einen eindeutigen Namen (NAME), einen API key (KEY) und eine endpoint URL (URL), wodurch die Daten automatisch importiert werden können.

Ein Depot hat einen Namen (NAME), welcher zusammen mit dem des Brokers das Depot identifiziert, sowie womöglich jährliche Fixkosten (KOSTEN).

Die Datenbank ermöglicht es, Wertpapiere einzutragen. Für jedes Wertpapier wird der Name (NAME) und eine Identifikationsnummer (ISIN) gespeichert. Es werden drei Arten von Wertpapieren unterstützt: Aktien, Fonds und Anleihen. Für Fonds werden zusätzlich die aktuellen jährlichen Kosten (TER) gespeichert. Wertpapiere werden in einem Depot gelagert und innerhalb des Depots durch die ISIN identifiziert - das bedeutet, dass man dasselbe Wertpapier potentiell in mehreren Depots besitzen kann.

Wertpapiere werden in Orders gekauft und verkauft. Ein Order hat eine Timestamp (TIME), einen Preis per Stück (PREIS), eine Menge (NR) und eine Ordergebühr (GEB). Es wird durch eine Order ID (OId) vollständig identifiziert. Ein Order kann entweder ein Kauf- oder ein Verkauforder sein. Für Verkauforders wird zusätzlich gespeichert, wie viel Kapitalertragssteuer gezahlt oder rückerstattet wurde (KEST).

Ein Order findet immer an genau einer Börse statt. Für eine Börse wird der eindeutige Name (NAME) und die Zeitzone (TZ) gespeichert. Für jede Börse wird ebenfalls vermerkt, welche Broker die Börse unterstützen.

Wertpapiere schütten Dividenden aus. Vergangene Ausschüttungen werden in der Datenbank gespeichert. Es wird das Datum (DATUM), der Kurswert zum Zeitpunkt (KURS), sowie die Höhe per Stück (ERTRAG) gespeichert. Eine Dividendenausschüttung wird durch das Wertpapier und das Datum vollständig identifiziert.

Fonds sind haben eine besondere Eigenschaft - sie können andere Wertpapiere, also Aktien, Anleihen, oder auch Fonds, welche in der Datenbank gespeichert sind, enthalten. Dazu wird auch vermerkt, mit welcher Gewichtung (GEWICHT) diese zum Fond beitragen.

In einigen Fällen ist nicht eindeutig aus der Angabe herauszulesen ob es sich um eine (0, ★) oder (1, ★) Funktionalität handelt. In diesen Fällen ist natürlich beides korrekt. In praktischen Szenarien ist es allerdings ratsam sich um eine volle Spezifizierung solcher Fälle zu bemühen anstatt nur sinnvolle Annahmen zu treffen.

Aufgabe 2 (Semantik von EER Diagrammen)**[1 Punkt]**

Betrachten Sie das in Abbildung 1 dargestellte EER-Diagramm.

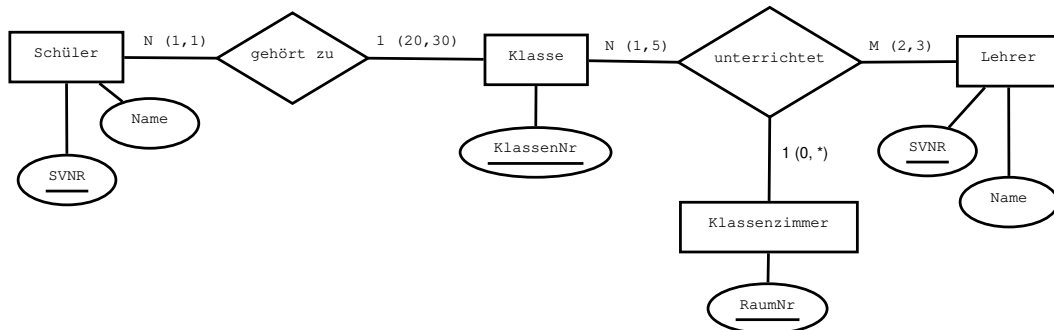


Abbildung 1: EER-Diagramm zu Aufgabe 2

In dem ER-Diagramm wird sowohl die Notation mittels Funktionalitäten, als auch die (min,max)-Notation verwendet.

(Anmerkung: 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.

- Geben Sie einen konkreten Beziehungstyp im Diagramm an, bei welchem das Weglassen einer Notation zum Verlust von Informationen führt.
- Für den gewählten Beziehungstyp führt das Weglassen welcher Notation zum Informationsverlust?
- Erklären Sie kurz in eigenen Worten, welche Information nicht mehr dargestellt werden kann.
- Geben Sie ein konkretes Beispiel für die verlorene Information an. D.h., für den von Ihnen gewählten Beziehungstyp, geben Sie eine Ausprägung an welche (mindestens) eine durch die weggelassene Notation ausgedrückte Bedingung verletzt, aber sämtliche Beschränkungen der verbleibenden Notation erfüllt.

Aufgabe 3 (Überführung ins Relationenschema)**[2 Punkte]**

Überführen Sie das EER-Diagramm aus Abbildung 2 in ein Relationenschema. Nullwerte sind nicht erlaubt (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 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.

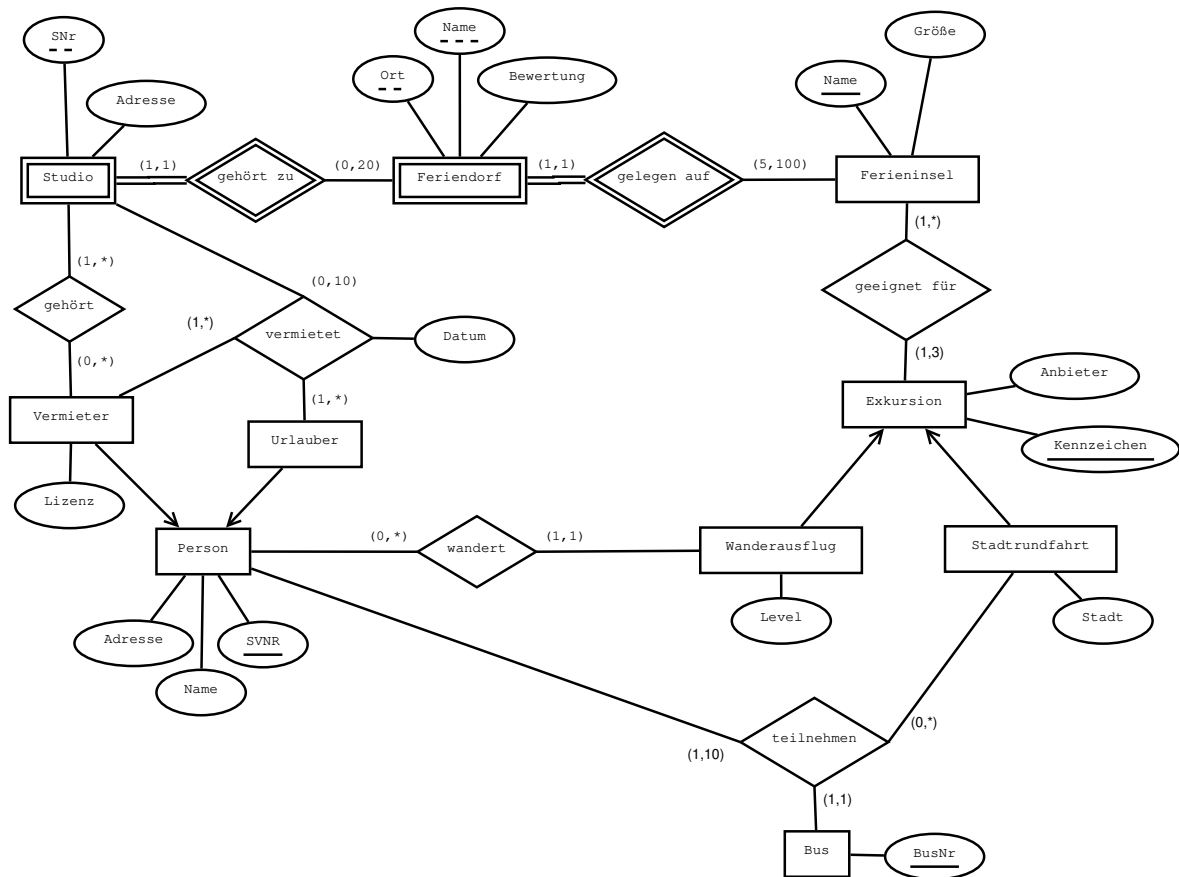


Abbildung 2: EER-Diagramm zu Aufgabe 3

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 L^AT_EX Befehle vermerkt.

Aufgabe 4 (Auswerten)

[0.5 Punkte]

Gegeben sind die folgenden vier Relationen.

Manufacturer		
name	country	founded
Cube	Germany	1993
Specialized	USA	1974
KTM	Austria	1992

Model		
name	year	cost
Attention	2022	900
Reaction C:62 Pro	2020	1800
Turbo Levo	2021	5600
Rockhopper Expert 29	2020	1200

producedBy	
by	mname
Cube	Attention
Cube	Reaction C:62 Pro
Specialized	Turbo Levo
Specialized	Rockhopper Expert 29

Sale		
mname	date	payment
Reaction C:62 Pro	2020-09-09	card
Reaction C:62 Pro	2021-01-03	cash
Reaction C:62 Pro	2021-06-08	card
Rockhopper Expert 29	2020-04-06	card
Turbo Levo	2021-06-08	card

Bestimmen Sie das Ergebnis der folgenden Anfragen über diesen Relationen.

(a)

$$\sigma_{sname=name} \left(\rho_{sname \leftarrow name} \left(\pi_{name, year}(\text{Model}) \bowtie \rho_{name \leftarrow mname}(\text{Sale}) \right) \times \left(\pi_{name, country} \left(\sigma_{founded > 1990}(\text{Manufacturer}) \right) \bowtie \rho_{name \leftarrow by}(\text{producedBy}) \right) \right)$$

(b)

$$\{m.country, s.date \mid m \in \text{Manufacturer} \wedge s \in \text{Sale} \wedge \exists p \in \text{producedBy} (p.by = m.name \wedge \exists o \in \text{Model} (o.name = p.mname \wedge o.cost < 2000))\}$$

Aufgabe 5 (Äquivalenzen)

[2 Punkte]

Gegeben sind die folgenden Relationenschemata $R(\underline{ABC})$, $S(\underline{BDE})$, $T(\underline{ADF})$ sowie untenstehende Paare q_i, q_j an Ausdrücken der relationalen Algebra. Für alle diese Paare:

- 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 NULL-Werte in den Ausprägungen verboten sind.
- Begründen Sie Ihre Antwort mit einer kurzen **Erklärung**.
- Falls die beiden Ausdrücke *nicht* äquivalent 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) $q_1: \sigma_{A>3 \vee D<4}(R \bowtie S)$ und
 $q_2: \sigma_{A>3}(R \bowtie S) \cap \sigma_{D<4}(R \bowtie S)$
- (b) $q_3: (\sigma_{A<3}(\pi_A(T) - \pi_A(R))) \cup \rho_{A \leftarrow E} \pi_E(S)$ und
 $q_4: ((\pi_A(T) - \sigma_{A<3}(R)) \cup \rho_{A \leftarrow E} \pi_E(S))$
- (c) $q_5: \pi_{AB}(((\pi_D(T) - \pi_D(S)) \bowtie T) \bowtie S) \cap \pi_{AB}(R)$ und
 $q_6: \pi_{AB}(R) - \rho_{B \leftarrow D}(\pi_{AD}(T \bowtie S) \cup \rho_{D \leftarrow B}(\pi_{AB}(R)))$

Aufgabe 6 (Größenabschätzung)

[1.5 Punkte]

Gegeben sind die Relationenschemata $R(\underline{ABCD})$, $S(\underline{ABCDE})$, und $T(\underline{ACDE})$ sowie je eine Ausprägung für jedes dieser Schemata, wobei $|R|$ Tupel in der Ausprägung für R vorhanden sind, $|S|$ Tupel in jener für S , und $|T|$ Tupel in jener für T .

- Geben Sie die minimale bzw. maximale Größe (= Anzahl der Tupel) folgender Ausdrücke unter Annahme der angegebenen Werte für $|R|$, $|S|$, $|T|$ in relationaler Algebra an.
- Begründen Sie Ihre Antwort.
- Geben Sie sowohl für das Minimum als auch für das Maximum jeweils **konkrete Ausprägungen** mit der angegebenen Anzahl von Tupeln an ($|R|$ für R , $|S|$ für S , $|T|$ für T), auf welchen die Anfrage tatsächlich die von Ihnen berechnete minimale/maximal Anzahl an Tupeln enthält. Geben Sie jeweils außerdem das Ergebnis der Anfrage für Ihre Ausprägungen an.

- (a) $q_1: (\rho_{A \leftarrow C, B \leftarrow D}(\pi_{C,D}(R)) \bowtie S) \bowtie R$ (mit $|R| = 4$ und $|S| = 5$)
- (b) $q_2: ((\pi_{A,C}(R) - \pi_{A,C}(S)) - \pi_{A,C}(T)) \cup \rho_{A \leftarrow D, C \leftarrow E}(\pi_{D,E}(T))$ (mit $|R| = 5$, $|S| = 4$ und $|T| = 3$)
- (c) $q_3: (\pi_{CDE}(\sigma_{A=1 \vee A \neq 1}(T)) \cap \pi_{CDE}(S)) \bowtie \sigma_{A<10}(R)$
(mit $|R| = 5$, $|S| = 5$ und $|T| = 1$)

Aufgabe 7 (Abfragesprachen)

[1 Punkt]

Gegeben sind die Relationenschemata $R(\underline{A}B)$, $S(\underline{A}\underline{B}\underline{C}\underline{D})$, und $T(\underline{A}\underline{C}\underline{E})$.

In den folgenden Aufgaben ist eine Abfrage in einer der Sprachen aus der Vorlesung gegeben. Die Abfrage soll in die jeweils beiden anderen Sprachen übersetzt werden.

- (a) Übersetzen Sie die Abfrage

$$\pi_{A,B}(R \ltimes (\pi_{A,B}(R) - \pi_{A,B}(\sigma_{C \geq 5}(S))))$$

in den Tupel- und den Domänenkalkül.

- (b) Übersetzen Sie die Abfrage

$$\{[a, c] \mid \exists b, d([a, b, c, d] \in S \wedge [c, a, 5] \in T)\}$$

in den Tupelkalkül und in Relationale Algebra.

- (c) Übersetzen Sie die Abfrage

$$\{[s.A, s.C] \mid s \in S \wedge \exists t(t \in T \wedge t.A = s.A \wedge t.C = s.C)\}$$

in den Domänenkalkül und Relationale Algebra.

Aufgabe 8 (Formalisieren von Anfragen)

[4 Punkte]

Eine Hausverwaltung speichert Informationen zu den verwalteten Wohnhäusern in einer Datenbank mit folgendem Schema (Primärschlüssel sind unterstrichen, Fremdschlüssel sind kursiv geschrieben).

Haus (HNr, Adresse, Bezirk)
 Eigentümer (Name, GebDat, IBAN)
 Wohnung (HausNr: *Haus.HNr*, Top, Stock, Größe, Nutzwert,
 EigentümerName: *Eigentümer.Name*, EigentümerGebDat: *Eigentümer.GebDat*)
 BKZahlung (ZNr, Datum, Betrag,
 WohnungHausNr: *Wohnung.HausNr*, WohnungTop: *Wohnung.Top*)
 Versicherung(Versicherer, VertragNr, Haus: *Haus.HNr*)
 Schaden (Versicherer: *Versicherung.Versicherer*,
 VertragNr: *Versicherung.VertragNr*, SchadenNr, Beschreibung)
 betrifft (Versicherer: *Schaden.Versicherer*, VertragNr: *Schaden.VertragNr*,
 SchadenNr: *Schaden.SchadenNr*, WohnungTop, *Wohnung.top*)

(Sie dürfen im Folgenden gerne passende (eindeutige) Abkürzungen sowohl für die Relationen- als auch die Tabellennamen verwenden.)

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

- (a) Geben Sie Name und Geburtsdatum aller Eigentümer aus welche zwei Wohnungen im Bezirk '1030' besitzen, welche im selben Haus und Stockwerk sind.

- (b) Geben Sie die Namen alle Eigentümer der Wohnungen aus die im Haus mit HNr 99 in einem Schadensfall im 2. Stock involviert waren.
- (c) Geben Sie alle Adressen und Bezirke der Häuser aus welche Wohnungen haben die mindestens im 5. Stock oder höher sind.