

1. Übungsblatt (WS 2020) – Musterlösung

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 **05.11. 12:00Uhr** Upload der Abgabe über TUWEL
ab **18.11. 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 arbeiten an einem Startup, dass das Buchmarketing revolutionieren wird. Nachdem sie sich mehrere Millionen Euro an Funding gesichert haben beschliessen Sie das Projekt auch tatsächlich umzusetzen. Sie beginnning mit dem Entwurf der Datenbank auf Basis ihrer bisherigen Notizen.

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:

Das zentrale Element der Datenbank ist das Erfassen von Schriftstücken. Für jedes Schriftstück wird ein Titel (TITEL) und eine Seitenanzahl (SEITEN) gespeichert, wobei der Titel für jedes Schriftstück eindeutig ist. Zusätzlich wird zu jedem Schriftstück gespeichert ob es sich speziell um ein Sachbuch, ein Roman oder Lyrik handelt. Ein Schriftstück muss allerdings nicht unbedingt in eine der drei Kategorien fallen. Bei einem Roman wird dabei zusätzlich das Genre (GENRE) gespeichert. Sachbücher können über Romane handeln, allerdings maximal über 3 verschiedene Romane. Es ist natürlich auch möglich dass ein Sachbuch gar nicht über einen Roman handelt. Bücher werden von Autorinnen geschrieben. Eine Autorin hat immer mindestens ein Schriftstück geschrieben und ein Schriftstück hat mindestens eine Autorin. Eine Autorin wird eindeutig durch eine Kombination an Namen (NAME) und Geburtsdatum (GEBDAT) identifiziert.

Zu einem Schriftstück können verschiedene Editionen herausgegeben werden. Die Editionen eines Schriftstücks können untereinander an ihrer Editionsnummer (ED-NR) unterschieden werden. Zwischen verschiedenen Schriftstücken sind Editionsnummern allerdings nicht zwingend eindeutig. Zusätzlich wird einer Edition ein Jahr (JAHR) zugeordnet. Eine Edition hat mindestens eine Auflage. Die Auflagen einer Edition können mittels Auflagennummer (ANR) unterschieden werden. Weiters wird zu jeder Auflage gespeichert wer den Druck übernimmt (DRUCKEREI). Weiters sind manche Editionen auch Sondereditionen, welche zusätzlich zu einer normalen Edition noch einen Anlass (ANLASS) haben.

Editionen werden von einem Verlag herausgegeben. Einem Verlag ist ein Name (NAME) und ein Budget (BUDGET) zugeordnet, wobei der Name den Verlag eindeutig identifiziert sind. In jedem Verlag gibt es Abteilungen welche innerhalb des Verlags durch eine eindeutige Kombination von Zuständigkeitsbereich (BEREICH) und der Stadt der Niederlassung (STADT) identifiziert. In verschiedenen Verlagen können allerdings die gleichen Bereich/Stadt Kombinationen vorkommen. Abteilungen ohne einen Verlag gibt es nicht.

Ein Verlag bewirbt an einem konkreten Datum (DATUM), über bis zu 2 Marketing Kanäle, eine oder mehrere Zielgruppen. Ein Marketing Kanal hat einen Namen (KNAME) – über den der Kanal eindeutig identifiziert wird – und zugehörige Kosten (KOSTEN). Weiters kann ein Marketing Kanal unter anderem ein Soziales Netzwerk oder eine Zeitung sein. Bei der Zeitung die Auflage (AUFLAGE). Bei Sozialen Netzwerken wird noch die konkrete Plattform (PLATFORM) erfasst. Für eine Zielgruppe werden eine Bezeichnung (BEZ) und eine Altersgruppe (AGRP) erfasst, wobei die Bezeichnung für jede Zielgruppe eindeutig ist.

Zielgruppen nutzen Soziale Netzwerke und interessieren sich für mindestens ein Schriftstück. Weiters liebt jede Zielgruppe bis zu 5 Autorinnen. Leider hasst jede Zielgruppe auch mindestens eine (und beliebig viele) Autorinnen. Das bewegt Autorinnen dazu unter Pseudonymen zu arbeiten. Dazu wird erfasst welche Autorin ein Alias von welcher anderen Autorin ist.

Lösung: Siehe Abbildung 1.

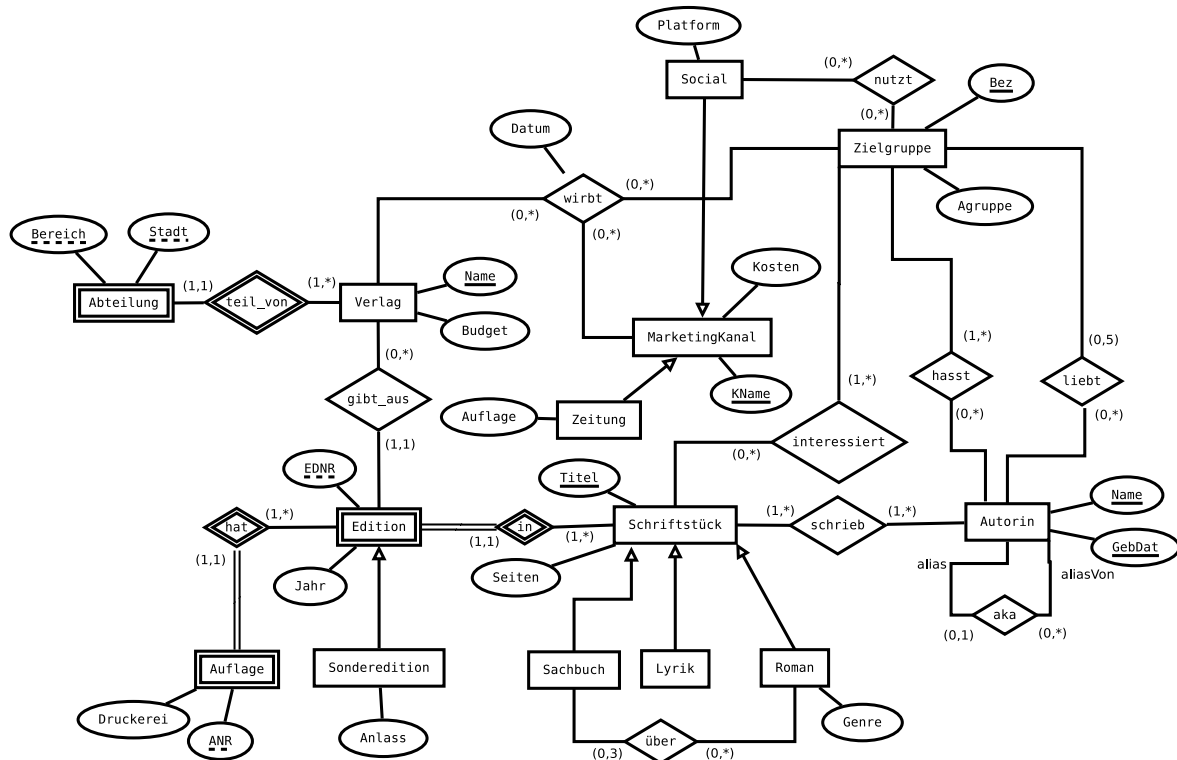


Abbildung 1: Lösung von Aufgabe 1

Achtung: Bei der mehrstelligen Beziehung zwischen Verlag, Marketing Kanal und Zielgruppe ist es (mit der (min,max) Notation) nicht möglich die Angabe im vollen Detail zu modellieren. Wie auch schon im Forum von einigen Kollegen diskutiert wurde, kann man auf Grund diverser Gründe (z.B. des Datums in der Beziehung) keine Schranken für die Entitäten setzen. Solch eine Situation ist in formaler Modellierung aller Art gängig und es ist daher immer wichtig die Grenzen von formalen Modellierungssprachen zu kennen.

In einigen Fällen ist nicht eindeutig aus der Angabe herauszulesen ob es sich um eine (0,*) oder (1,*) Beziehungen 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 2 dargestellte EER-Diagramm.

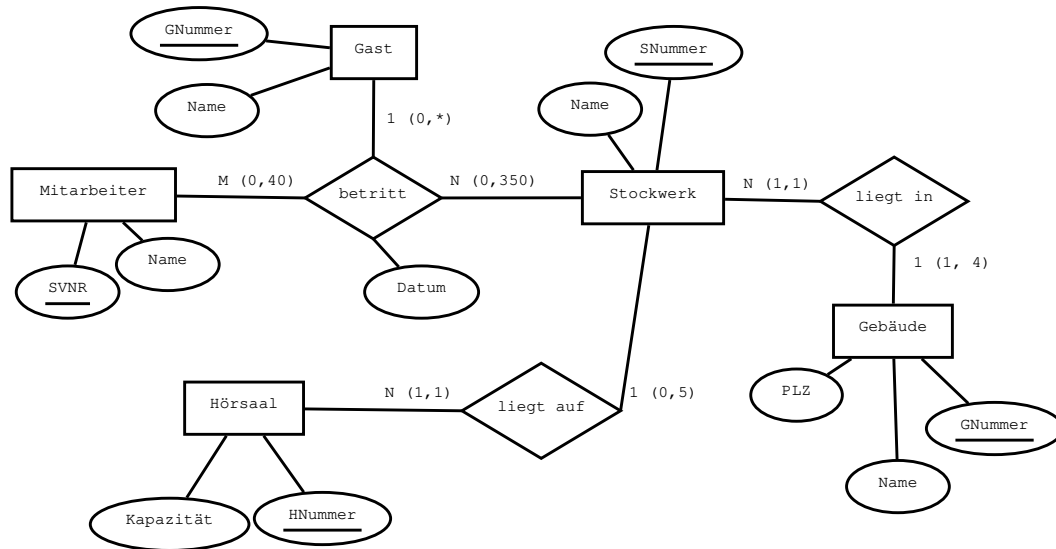


Abbildung 2: EER-Diagramm zu Aufgabe 2

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.

Lösung:

Ja, beide Notationen enthalten teilweise Informationen, welche durch die jeweils andere Notation nicht ausgedrückt werden können. Wir geben im folgenden jeweils nur ein Beispiel für jede Notation an - es gibt noch mehr:

- Lässt man beim Beziehungstyp **betritt** die *Funktionalitäten* weg, so kann durch die (min,max)-Notation nicht ausgedrückt werden, dass nur ein Gast pro Mitarbeiter und

Stockwerk das Stockwerk betreten darf. Ein konkretes Gegenbeispiel wäre, dass der Mitarbeiter mit SVNr 2000 das Stockwerk mit SNummer 1 und Gast mit GNummer 2 sowie mit Gast mit GNummer 3 betritt. Das ist nicht möglich, wenn man die Kardinalität berücksichtigt.

- Lässt man beim Beziehungstyp **liegt auf** die *(min,max)*-Notation weg, so wird durch die Angabe der Funktionalitäten nicht ausgedrückt, dass auf einem Stockwerk nicht mehr als 5 Hörsäle liegen können.

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

Überführen Sie das EER-Diagramm aus Abbildung 3 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.

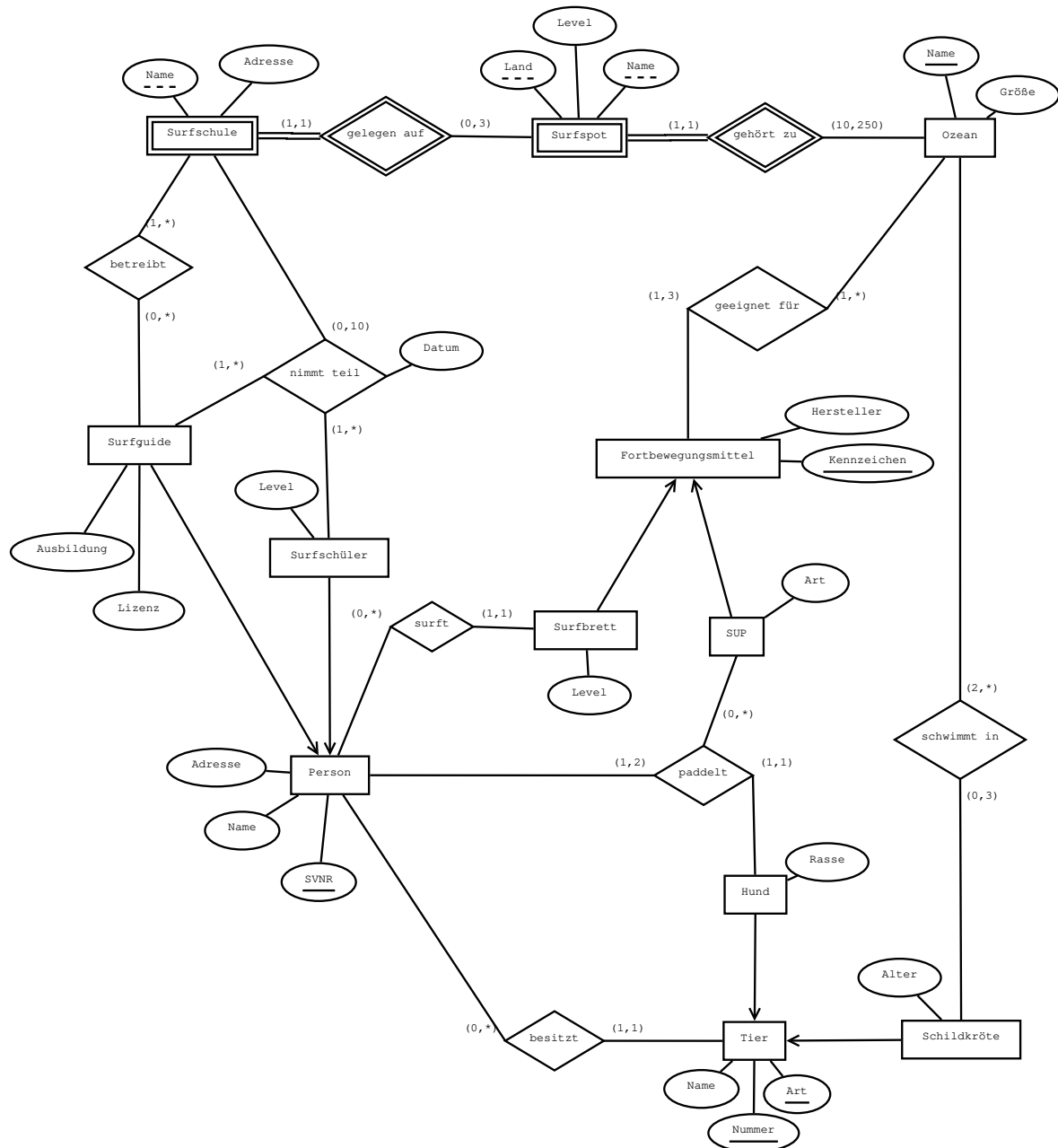


Abbildung 3: EER-Diagramm zu Aufgabe 3

Lösung:

Ozean	(<u>Name</u> , Größe)
Surfspot	(<u>Ozean_Name</u> : <i>Ozean.Name</i> , <u>Land</u> , <u>Name</u> , Level)
Surfschule	(<u>Ozean_Name</u> : <i>Surfspot.Ozean_Name</i> , <u>Spot_Land</u> : <i>Surfspot.Land</i> , <u>Spot_Name</u> : <i>Surfspot.Name</i> , <u>Schule_Name</u> , Adresse)
Person	(SVNR, Adresse, PName)
Surfschüler	(<u>SVNR</u> : <i>Person.SVNR</i> , Level)
Surfguide	(<u>SVNR</u> : <i>Person.SVNR</i> , Ausbildung, Lizenz)
Tier	(<u>Nummer</u> , <u>Art</u> , Name, Besitzer: <i>Person.SVNR</i>)
Hund	(<u>Nummer</u> : <i>Tier.Nummer</i> , <u>Art</u> : <i>Tier.Art</i> , Rasse, paddeltMit: <i>Person.SVNR</i> , paddelt: <i>SUP.Kennzeichen</i>)
Schildkröte	(<u>Nummer</u> : <i>Tier.Nummer</i> , <u>Art</u> : <i>Tier.Art</i> , Alter)
Fortbewegungsmittel	(<u>Kennzeichen</u> , Hersteller)
SUP	(<u>Kennzeichen</u> : <i>Fortbewegungsmittel.Kennzeichen</i> , Art)
Surfbrett	(<u>Kennzeichen</u> : <i>Fortbewegungsmittel.Kennzeichen</i> , Level, surft: <i>Person.SVNR</i>)
geeignet_für	(<u>Ozean_Name</u> : <i>Ozean.Name</i> , <u>Fortb_Kennzeichen</u> : <i>Fortbewegungsmittel.Kennzeichen</i>)
schwimmt_in	(<u>Ozean_Name</u> : <i>Ozean.Name</i> , <u>Sch_Nummer</u> : <i>Schildkröte.Nummer</i> , <u>Sch_Art</u> : <i>Schildkröte.Art</i>)
betreibt	(<u>Ozean_Name</u> : <i>Surfschule.Ozean_Name</i> , <u>Spot_Land</u> : <i>Surfschule.Spot_Land</i> , <u>Spot_Name</u> : <i>Surfschule.Spot_Name</i> , <u>Schule_Name</u> : <i>Surfschule.Schule_Name</i> , <u>SVNR</u> : <i>Surfguide.SVNR</i>)
nimmt_teil	(<u>Ozean_Name</u> : <i>Surfschule.Ozean_Name</i> , <u>Spot_Land</u> : <i>Surfschule.Spot_Land</i> , <u>Spot_Name</u> : <i>Surfschule.Spot_Name</i> , <u>Schule_Name</u> : <i>Surfschule.Schule_Name</i> , <u>Guide</u> : <i>Surfguide.SVNR</i> , <u>Schüler</u> : <i>Surfschüler.SVNR</i> , Datum)

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.

Band		
name	genre	founded
Cid Rim	TBD	2010
JSBL	Funk	2008
CSH	Indie	2010
Dorian Concept	Electronic	2006

Musician		
name	instrument	studio
Will	Guitar	False
Peter	Bass	True
Clemens	Drums	True
Dorian	Keyboard	True

partOf	
mname	bname
Clemens	JSBL
Clemens	Cid Rim
Will	CSH
Dorian	JSBL
Dorian	Dorian Concept

Song			
title	by	length	genre
Killer Whales	CSH	6:14	Indie
My Boy	CSH	2:52	Indie
Swnerve	Cid Rim	4:26	Electronic
Failures III	Dorian Concept	3:40	Jazz

Bestimmen Sie das Ergebnis der folgenden Anfragen über diesen Relationen.

(a)

$$\pi_{\text{title, genre}} \left(\pi_{\text{title, length}}(\text{Song}) \times ((\rho_{\text{mname} \leftarrow \text{name}}(\sigma_{\text{studio}=\text{True}}(\text{Musician})) \bowtie \text{partOf}) \bowtie \rho_{\text{bname} \leftarrow \text{name}}(\sigma_{\text{founded} \geq 2010}(\text{Band}))) \right)$$

Lösung:

q	
title	genre
Killer Whales	TBD
My Boy	TBD
Swnerve	TBD
Failures III	TBD

(b)

$$\{m.name, b.name \mid m \in \mathbf{Musician} \wedge b \in \mathbf{Band} \wedge m.instrument \neq \text{'Bass'} \wedge \\ \exists p \in \mathbf{partOf} (\\ b.name = p.bname \wedge m.name = p.mname \wedge \\ (\forall s \in \mathbf{Songs} (b.name = s.by \rightarrow s.genre = b.genre))) \}$$

Lösung:

q	
m.name	b.name
Clemens	JSBL
Dorian	JSBL
Will	CSH

Aufgabe 5 (Äquivalenzen)

[2 Punkte]

Gegeben sind die folgenden Relationenschemata $R(\underline{ABC})$, $S(\underline{BDE})$, $T(\underline{CDF})$ 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) $q_1: (\pi_C(T) - \sigma_{C>5}(R)) \cap \rho_{C \leftarrow B}(\pi_B(S))$ und
 $q_2: (\sigma_{C>5}(\pi_C(T) - \pi_C(R))) \cap \rho_{C \leftarrow B}(\pi_B(S))$
- (b) $q_3: S \bowtie (\pi_{CD}(T) \cap \rho_{D \leftarrow A}(\pi_{CA}(R)))$ und
 $q_4: (S \bowtie \pi_{CD}(T)) \cap (S \bowtie \rho_{D \leftarrow A}(\pi_{CA}(R)))$
- (c) $q_5: \pi_{AB}(\sigma_{A>B \vee A<B}(R)) \times \pi_{DF}(\sigma_{D>F \vee D<F}(T))$ und
 $q_6: \pi_{ABDF}(\sigma_{(A>B \wedge D>F) \vee (A<B \wedge D<F)}(R \times T))$

Lösung:**Aufgabe (a)****Nein**, q_1 und q_2 **sind nicht** äquivalent.

q_2 ist eine gültige Aussage der Relationalen Algebra, q_1 jedoch nicht. Der Grund ist, dass sich die Schemata von $\pi_C(T)$ (enthält nur das Attribut C) und $\sigma_{C>5}(R)$ (enthält alle Attribute aus R : A, B, C) unterscheiden, was bei Mengendifferenz nicht erlaubt ist.

Aufgabe (b)**Ja**, q_3 und q_4 **sind** äquivalent.

Das beruht auf der Äquivalenz $S \bowtie (T \cap R) \equiv ((S \bowtie T) \cap (S \bowtie R))$.

Aufgabe (c)

Nein, q_5 und q_6 sind nicht äquivalent.

Bei q_5 können alle Tupel, bei denen $A \neq B$ und $D \neq F$ sind, enthalten sein. Bei q_6 hingegen können nur Tupel der Form $A < B, D < F$ und $A > B, D > F$ im Ergebnis enthalten sein.

Gegenbeispiel

R			T			q_5				q_6			
<u>A</u>	B	C	<u>C</u>	<u>D</u>	F	A	B	D	F	A	B	D	F
2	1	5	0	5	1	2	1	5	1	2	1	5	1
1	2	5	7	3	8	2	1	3	8	1	2	3	8
						1	2	5	1				
						1	2	3	8				

Aufgabe 6 (Größenabschätzung)

[1.5 Punkte]

Gegeben sind die Relationenschemata $R(\underline{AB})$, $S(\underline{ABCD})$, und $T(\underline{ACE})$ 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) $q_1: \rho_{B \leftarrow C}(T) \bowtie \sigma_{A=4 \wedge B=2}(R)$ (mit $|R| = 5$ und $|T| = 4$)
- (b) $q_2: \pi_{EX}(\rho_{X \leftarrow A, Y \leftarrow C}(S) \times T) - \pi_{EX}(\rho_{X \leftarrow A}(T) \bowtie S)$ (mit $|S| = 7$ und $|T| = 4$)
- (c) $q_3: (\pi_A(\sigma_{A \neq 1}(R)) \cap \pi_A(\sigma_{A > 3}(T))) \cup \rho_{A \leftarrow C}(\pi_C(T) \bowtie \pi_C(S))$
(mit $|R| = 5$, $|S| = 3$ und $|T| = 3$)

Lösung:**Aufgabe (a)**

[Minimum: 4 | Maximum: 4]

Maximum = Minimum: 4

R		T			Ergebnis		
<u>A</u>	B	A	C	<u>E</u>	A	B	E
4	2	4	2	1	4	2	1
3	1	4	2	2	4	2	2
6	9	4	2	3	4	2	3
1	2	4	2	4	4	2	4
3	11						

Die wichtigste Beobachtung hier ist, dass man ganz Allgemein bei $T \bowtie R$ die Tupel von T immer erhalten bleiben. Es können also nie weniger als die 4 Tupel von T in unserem Ergebnis vorkommen. Die Umbenennung auf der Seite von T ändert an diesem Verhalten nichts, während die Selektion über R nur etwaige Tupel dort löscht. Man sieht schnell, dass die Tupel von R alle entfernt werden können um einfach das Minimum 4 zu erreichen.

Für das Maximum muss man weiters beobachten dass der Join auf die Attribute A und B joined. Durch die Selektion bei R bleibt aber nur maximal 1 Tupel an Werten für (A, B) übrig, nämlich $(4, 2)$. Der Join kann also nicht ein Tupel aus T mit mehr als einem Tupel aus R (nach der Selektion) verbinden. Es bleibt nur die Möglichkeit mit 0 oder einem Tupel aus R zu joinen. In beiden Fällen ergibt das insgesamt genau 4 Tupel im Ergebnis, da es sich um einen outer join handelt.

Aufgabe (b) [Minimum: 0 | Maximum: 28] Im Maximum bleiben nach der Projektion auf EX die vollen 28 Tupel aus dem Kreuzprodukt übrig indem man alle Werte der Spalte A von S unterschiedlich wählt. Dazu lässt sich noch (wie im Beispiel unten) die Ausprägung so wählen, dass der Semijoin rechts leer ist, also nichts durch die Differenz entfernt wird.

Beim Minimum kann man die Überlegung genau umgekehrt anstellen. Im Beispiel wird die Ausprägung so gewählt, dass das Kreuzprodukt nach der Projektion auf EX möglichst klein wird. (Es bleiben genau die Tupel $\{(1, 1), (2, 1), (3, 1), (4, 1)\}$ nach der Projektion.) Im Semijoin haben S und $\rho_{X \leftarrow A}(T)$ nur das Attribut C gemein. Wird dieses überall gleich gewählt, bleiben hier alle Tupel von T übrig und die Differenz wird entsprechend leer.

Minimum: 0

S			
A	B	C	D
1	9	2	1
1	9	2	2
1	9	2	3
1	9	2	4
1	9	2	5
1	9	2	6
1	9	2	7

T		
A	C	E
1	2	1
1	2	2
1	2	3
1	2	4

Ergebnis	
E	X

Maximum: 28

S			
A	B	C	D
1	9	2	1
2	9	2	2
3	9	2	3
4	9	2	4
5	9	2	5
6	9	2	6
7	9	2	7

T		
A	C	E
1	5	1
1	5	2
1	5	3
1	5	4

Aufgabe (c) [Minimum: 0 | Maximum: 6] Beim Minimum ist es relativ einfach, mit Hilfe der Selektionen, die Ausprägung so zu konstruieren, dass beide Terme der Vereinigung leer

sind. Dabei sollte aber trotzdem auf mögliche Schlüssel geachtet werden. Beachten Sie auch, dass der Join $\pi_C(T) \bowtie \pi_C(S)$ semantisch äquivalent zu einem Durchschnitt ist, da hier beide Operanden nur ein (namensgleiches) Attribut haben.

Beim Maximum kann man beobachten, dass beide Terme der Vereinigung nicht größer als $|T|$ sein können (wegen der Semantik des Mengendurchschnitts). Wenn man weiters erkennt dass die beiden Terme in Bezug auf T komplett unabhängig voneinander sind, dann kann man einfach Ergebnis der Größe $2|T|$ erzeugen. Da die beiden einzelnen Terme nicht größer als $|T|$ sein können ist das auch garantiert das Maximum.

Minimum: 0

R		S				T		
<u>A</u>	B	A	B	<u>C</u>	<u>D</u>	A	C	<u>E</u>
1	1	0	0	5	0	1	1	1
2	2	0	0	6	0	2	2	2
3	3	0	0	7	0	3	3	2
4	4							
5	5							

Maximum: 6

R		S				T		
<u>A</u>	B	A	B	<u>C</u>	<u>D</u>	A	C	<u>E</u>
4	1	0	0	1	0	6	1	1
5	2	0	0	2	0	7	2	2
6	3	0	0	3	0	8	3	4
7	4							
8	5							

Ergebnis A: {1, 2, 3, 6, 7, 8}

Aufgabe 7 (Abfragesprachen)

[1 Punkt]

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

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

$$\pi_{A,C}(S \bowtie T)$$

in den Tupel- und den Domänenkalkül.

Lösung:

Domänenkalkül:

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

Tupelkalkül:

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

(b) Übersetzen Sie die Abfrage

$$\{x \mid \exists y, z ([x, y] \in R \wedge [y, x, z, 3] \in S)\}$$

in den Tupelkalkül und in Relationale Algebra.

Lösung:

Relationale Algebra:

$$\pi_{R.A}(R \bowtie_{R.A=S.B \wedge R.B=S.A} \sigma_{D=3}(S))$$

Tupelkalkül:

$$\{r.A \mid r \in R \wedge \exists s(s \in S \wedge r.A = s.B \wedge r.B = s.A \wedge s.D = 3)\}$$

(c) Übersetzen Sie die Abfrage

$$\{u, v \mid u \in R \wedge v \in T \wedge \exists w \in S(w.C > v.C)\}$$

in den Domänenkalkül und Relationale Algebra.

Lösung:

Relationale Algebra:

$$R \times \pi_{A,C,E}(\sigma_{X>C}(T \times \rho_{X \leftarrow C}(\pi_C(S))))$$

Domänenkalkül:

$$\{(a1, b, a2, c, e) \mid [a1, b] \in R \wedge [a2, c, e] \in T \wedge \exists as, bs, cs, ds([as, bs, cs, ds] \in S \wedge cs > c)\}$$

Aufgabe 8 (Formalisieren von Anfragen)

[4 Punkte]

Ein Unternehmen verwaltet Informationen über die Zutritte seiner Mitarbeiter in einer Datenbank mit folgendem Schema (Primärschlüssel sind unterstrichen, Fremdschlüssel sind kursiv geschrieben).

Mitarbeiter (PersNr, SVNr, MName)

Stockwerk (SName, Bezeichnung)

betritt (MitarbeiterNr: *Mitarbeiter.PersNr*, StockName: *Stockwerk.SName*, Datum)

verlässt (MitarbeiterNr: *Mitarbeiter.PersNr*, StockName: *Stockwerk.SName*, Datum)

Gebäude (GName, PLZ)

unterteiltIn(StockwerkN: *Stockwerk.Name*, GName: *Gebäude.GName*)

Hörsaal (HNummer, Name)

liegtIn (HörsaalNr: *Hörsaal.HNummer*, SName: *Stockwerk.SName*)

(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 den Namen, die Personalnummer und die Sozialversicherungsnummer aller Mitarbeiter an, die sich jetzt gerade auf Stockwerk mit Namen 'H1' befinden.

Lösung:

Relationale Algebra: (Achtung: um genau zu sein müsste man hier noch alle Attribute umbenennen, damit man sie joinen kann.)

$$\pi_{MName, PersNr, SVNR}(\sigma_{StockName='H1'}((betr_ - beendet) \bowtie Mitarbeiter))$$

beendet ist definiert als:

$$\pi_{MitarbeiterNr, StockName, Datum}(\sigma_{Datum \leq EndDatum}(betr_ \bowtie \rho_{EndDatum \leftarrow Datum}(verl_)))$$

Tupelkalkül:

$$\{[m.MName, m.PersNr, m.SVNR] \mid m \in Mitarbeiter \wedge \exists b \in betr_ \wedge \\ m.PersNr = b.MitarbeiterNr \wedge \neg \exists v \in verl_ (v.MitarbeiterNr = b.MitarbeiterNr \wedge \\ v.StockName = b.StockName \wedge v.Datum \geq b.Datum \wedge b.StockName = 'H1')\}$$

Domänenkalkül:

$$\{[name, persNr, svnr] \mid ([name, persNr, svnr] \in Mitarbeiter \wedge \\ \exists startDatum, stockName ([persNr, stockName, startDatum] \in betr_ \wedge \neg \exists endDatum (\\ [persNr, stockName, endDatum] \in verl_ \wedge startDatum \leq endDatum \wedge stockName = 'H1')))\}$$

- (b) Es befinden sich gerade Studenten in Hörsaal mit dem Name 'Gödel'. Geben Sie die Mitarbeiter aus, die gerade auf einem Stockwerk sind, auf welchem sich auch diese Studenten aufhalten.

Lösung:

Relationale Algebra:

$$\pi_{PersNr}(\sigma_{Hörsaal.Name='Gödel'}((betr_ - beendet) \bowtie Mitarbeiter) \bowtie liegtIn \bowtie Hörsaal)$$

beendet ist definiert als:

$$\pi_{MitarbeiterNr, StockName, Datum}(\sigma_{Datum \leq EndDatum}(betr_ \bowtie \rho_{EndDatum \leftarrow Datum}(verl_)))$$

Tupelkalkül:

$$\{[m.PersNr] \mid m \in Mitarbeiter \wedge \exists b \in betr_ \wedge \exists l \in liegtIn \wedge \exists h \in Hörsaal \wedge$$

$$\begin{aligned} & m.PersNr = b.MitarbeiterNr \wedge \neg \exists v \in \text{verlässt} (\\ & v.MitarbeiterNr = b.MitarbeiterNr \wedge v.StockName = b.StockName \wedge v.Datum \geq b.Datum) \wedge \\ & b.StockName = l.SName \wedge l.HörsaalNr = h.HNummer \wedge h.Name = 'Gödel' \} \end{aligned}$$

Domänenkalkül:

$$\begin{aligned} & \{ [persNr] \mid \exists name, svnr ([name, persNr, svnr] \in \text{Mitarbeiter} \wedge \\ & \exists startDatum, stockName ([persNr, stockName, startDatum] \in \text{betritt} \wedge \\ & \neg \exists endDatum ([persNr, stockName, endDatum] \in \text{verlässt} \wedge startDatum \leq endDatum)) \wedge \\ & \exists hNr ([hNr, stockName] \in \text{liegtIn} \wedge [hNr, hName] \in \text{Hörsaal} \wedge hName = 'Gödel')) \} \end{aligned}$$

(c) Geben Sie alle Mitarbeiter aus, die an einem Tag in verschiedenen Gebäuden waren.

Lösung:

Relationale Algebra:

$$\begin{aligned} & \pi_{PersNr} (\sigma_{Datum = verglDatum \wedge GName \neq verglGeb} ((\text{betritt} \bowtie \text{unterteiltIn}) \bowtie \\ & (\rho_{verglGeb \leftarrow GName, verglDatum \leftarrow Datum} \text{verlässt} \bowtie \text{unterteiltIn}))) \end{aligned}$$

Tupelkalkül:

$$\begin{aligned} & \{ [b.MitarbeiterNr] \mid b \in \text{betritt} \wedge \exists u1 \in \text{unterteiltIn} \wedge \exists u2 \in \text{unterteiltIn} \wedge \exists v \in \text{verlässt} \wedge \\ & b.Datum = v.Datum \wedge b.MitarbeiterNr = v.MitarbeiterNr \wedge \\ & b.StockName = u1.StockwerkN \wedge v.StockName = u2.StockwerkN \wedge u1.GName \neq u2.GName \} \end{aligned}$$

Domänenkalkül:

$$\begin{aligned} & \{ [mitarbNr] \mid ([mitarbNr, stockName, datum] \in \text{betritt} \wedge \\ & \exists gName ([stockName, gName] \in \text{unterteiltIn} \wedge \\ & \exists stockV ([mitarbNr, stockV, datum] \in \text{verlässt} \wedge stockName \neq stockV \wedge \\ & \exists gNameV ([stockV, gNameV] \in \text{unterteiltIn} \wedge gNameV \neq gName) \} \end{aligned}$$