

# EINFÜHRUNG IN DATENBANKEN ÜBUNGEN

PROF. DR. RER. NAT. ALEXANDER VÖß  
INFORM-PROFESSUR

FH AACHEN  
UNIVERSITY OF APPLIED SCIENCES

20.12.2021

# Übersicht

---

- Übung 0x01
- Übung 0x02
- Übung 0x03
- Übung 0x04
- Übung 0x04-Extra
- Übung 0x05
- Übung 0x06
- Übung 0x07
- Übung 0x08
- Übung 0x09
- Übung 0x0a
- Übung 0x0b

# ÜBUNG 0x01

# Übung 0x01

---

## Gute Modellierungen - schlechte Modellierungen

- Erstes Ziel der Übung ist es, die Diskussion über “gute” und “weniger gute” Datenbanken bzw. Entwürfe oder Modelle fortzuführen (zunächst noch unpräzise formuliert).
- Zweites Ziel ist die Funktionsweise von DataGrip zu verstehen und erste Schritte Richtung Modellierung zu unternehmen.  
Wer noch kein DataGrip inkl. erster "Kemper"-Datenbank installiert hat, sollte das möglichst schnell nachholen.
- Zuweilen haben wir die Teilaufgabe "Diskutieren Sie, wenn mgl., <Thema> mit anderen Matses". Gemeint sind natürlich hier andere Matses des Jahrgangs in einer virtuellen Übungsgruppe. Und ja, das ist ernst gemeint, denn das liefert Ihnen nochmal einen anderen Blick auf das <Thema> - je kontroverser je besser 😊.

# Übung 0x01

---

## Aufgabe 1

Sehen Sie sich die Daten einer CD Sammlung an, die wie folgt in einer Excel-Tabelle aufgelistet sind:

CD_Lied			
CD_ID	Album	Erscheinungsdatum	Titelliste
4711	Anastacia – Not That Kind	1999	{1. Not That Kind, 2. I'm Outta Love, 3. Cowboys & Kisses}
4712	Pink Floyd – Wish You Were Here	1964	{1. Shine On You Crazy Diamond}
4713	Anastacia – Freak of Nature	1999	{1. Paid my Dues}

- Das Feld **Album** enthält die Attributwertebereiche Interpret und Albumtitel.
- Das Feld **Titelliste** enthält eine Menge von Titeln.

- Diskutieren Sie, wenn mgl., Vor- und Nachteile dieser Tabelle mit anderen Matses.
- Überlegen Sie sich eine bessere Struktur, d.h. ggf. neue Tabellen und/oder Attribute (Spalten). Diskutieren Sie insbesondere, nach welchen Kriterien Sie die neue Struktur gewählt haben.
- Starten Sie DataGrip und legen Sie ein Schema "sammlung" und Ihre Tabellen mit ein paar Daten an (siehe Video).
- Weiterführende Anregungen zum Thema “Normalisierung” finden sich hier: [https://de.wikipedia.org/wiki/Normalisierung\\_\(Datenbank\)](https://de.wikipedia.org/wiki/Normalisierung_(Datenbank)) (als Ausblick).

# ÜBUNG 0x02

# Übung 0x02

---

## Aufgabe 2

Wir wollen weitere Schritte in Richtung einer "guten" Modellierung unternehmen und uns weiter mit den Tools beschäftigen.

- Überlegen Sie sich eine erste Tabellenstruktur, die Matses (Name und Matr.Nr.), Vorlesungen (Titel) und Kurswahlen (wer hört was) modelliert.
- Was für eine eindeutige Nr./Id geben Sie den Daten? Und wieso?
- Starten Sie DataGrip und legen Sie ein Schema "matsedienste" und Ihre Tabellen mit ein paar Daten an (siehe Video).
- Diskutieren Sie, wenn mgl., Vor- und Nachteile Ihres Entwurfes mit anderen Matses.

# Übung 0x02

---

## SQL-Praktikum

- Ziel dieser Übung ist es, die erforderlichen Übungsdaten im DataGrip zur Verfügung zu haben und die SQL-Übungsergänzungen zu beginnen. Beachten Sie das einführende Video dazu.

## Aufgabe 1

- Importieren Sie, soweit noch nicht geschehen, die Schemata `matse_mhist` und `kemper` aus dem Ilias in Ihre Datenbank. Achtung: das Schema zuerst anlegen!
- Laden Sie das SQL-Skript `db_ex_01.sql` ins DataGrip und führen Sie einen der ersten `USE` bzw. `SELECT`-Befehle aus. Klappt alles?
- Lesen und verstehen Sie das Kapitel “0x01 Select” in den Übungsergänzungen.
- Bearbeiten Sie die Aufgaben zum Thema “Select - Selektion und Projektion”.

# ÜBUNG 0x03

# Übung 0x03

---

## Aufgabe 1

Erörtern Sie mit anderen Matses die folgenden Begriffe und Themen und verstehen Sie die verwendeten Fachbegriffe. Finden oder formulieren Sie jeweils Beispiele.

- Phasen Datenbankentwurf,
- ANSI/SPARC-Architektur,
- Elemente des ER-Diagramms, Entität und Entitätstyp, Relation, Chen-Kardinalität.

# Übung 0x03

---

## Aufgabe 2

Folgende Anforderungen an ein Projekt der Landesverwaltung NRW sind gegeben.

Es gibt Städte, Stadtteile und Strassen. Städte haben einen Namen und ein Gründungsjahr, Stadtteile haben ebenfalls einen Namen und gehören zu einer Stadt. Strassen laufen durch genau einen Stadtteil und haben eine Bezeichnung.

- Modellieren Sie einen konzeptuellen Entwurf mittels eines einfachen ER-Diagramms.
- Ergänzen Sie jeden Entitatstyp um ein weiteres typisches Attribut.
- Ergänzen Sie weiter je Typ geeignete Schlüsselattribute.

Zusatzaufgabe:

- Erweitern bzw. modifizieren Sie das ER-Diagramm so, dass Strassen durch ein oder mehrere Stadtteile laufen können.
- Wie und wozu würde man eine Wohnadresse in Relation setzen?

# Übung 0x03

---

## Aufgabe 3

Starten Sie DataGrip (oder Ihre jeweilige Alternative) und implementieren Sie das Städteprojekt aus Aufgabe 2:

- Entwerfen Sie geeignete Tabellen und geben ein paar reale Werte dazu ein.
- Statten Sie die Tabellen mit Schlüsselattributen aus - welche wählen Sie?
- Versuchen Sie sich an geeigneten Relationen, orientieren Sie sich dazu an dem Ansatz aus dem Beispiel im Video (Fotoshooting) oder wählen Sie einen eigenen. Wir werden das Thema noch behandeln, es geht hier um eine erste Beschäftigung mit der Fragestellung, wie Relationen abgebildet werden können.

## Übung 0x03

---

### Aufgabe 4

Erörtern Sie mit anderen Matses die folgenden Fragen und wählen Sie für sich eine Vorgehensweise, die Sie im Folgenden konsequent durchhalten.

- Welche Namenskonvention wählen Sie für Tabellen und Attribute (klein/gross, Upper/LowerCamelCase, etc. und warum?)
- Benennen Sie Tabellen in Ein- oder Mehrzahl (Person vs. Personen)?
- Statten Sie jede Tabelle mit einem eigenen ‘id’-Feld aus oder benutzen Sie gegebene “natürliche” Attribute wie etwa eine PLZ oder Personalausweisnummer?

# ÜBUNG 0x04

# Übung 0x04

---

**Vorab: Das SQL-Praktikum wird nächste Woche fortgesetzt. Wir bleiben zunächst bei der Modellierung. Aufgabe 3 ist eine echte Prüfungsaufgabe, um auch einmal den Umfang und die Komplexität zu zeigen.**

## Aufgabe 1

Erörtern Sie mit anderen Matses die folgenden Begriffe/Themen:

- Chen-Notation vs. Min-Max-Notation vs. UML-Notation;
- n-stellige Relationen;
- Spezialisierungen und Generalisierungen.

# Übung 0x04

---

## Aufgabe 2

Modellieren Sie jeweils für folgende Situationen einen konzeptuellen Entwurf mittels eines ER-Diagramms, einmal in Chen- (i) und einmal in Min-Max-Notation (ii):

- (i) Ein Fluss mündet maximal in ein Meer. In ein Meer mündet mindestens ein Fluss, normalerweise münden dort aber mehrere Flüsse.
- (ii) Digitale Fotos einer Sammlung, jeweils mit Datum und Auflösung, sollen mit Stichworten versehen werden. Damit es nicht zu unübersichtlich wird, besitzt jedes Foto mindestens eins, aber max. 5 Stichworte. Es darf auch Stichworte ohne Fotos geben.

# Übung 0x04

---

## Aufgabe 3

Der GaiaZOO in Kerkrade möchte seine Tiere in einer Datenbank verwalten. Er beherbergt verschiedene Tierarten, die von Pflegern betreut werden. Alle Tiere einer Art befinden sich in genau einem ganz bestimmten Areal im Zoo. Dort können aber natürlich mehrere Tierarten leben. Aufgrund unterschiedlicher Allergien bekommt jedes Tier eine individuelle fest vorgegebene Futtermischung mit bestimmten Zutaten. Weiter gilt:

- Pfleger kümmern sich nicht um einzelne Tiere, sondern allgemein um alle Tiere einer Tierart. D.h. Ihnen sind bestimmte Tierarten zugeordnet. In der Regel sind das je Tierart mindestens zwei.
- Falls das letzte Tier einer Tierart sterben sollte, wird umgehend ein neues Tier dieser Art erworben. Dennoch kommt es vor, dass zeitweilig ein Gehege auch leer sein kann.
- Die Areale sind in genau zwei Kategorien eingeteilt: Innen- und Aussengehege. Für jedes Areal ist die Fläche in m<sup>2</sup> wichtig, aber die Innengehege besitzen ausserdem noch eine Temperaturvorgabe. Vereinzelt kommt es vor, dass Tiere innen und aussen leben können.
- Für die Zutaten einer individuellen Futtermischung eines Tieres ist gewünscht, sie in Prozent der Gesamtmenge einer Mahlzeit anzugeben. Jedes Tier bekommt verschiedene Mahlzeiten am Tag, wobei dann eine konkrete Mengenangabe in kg je Mahlzeit ausreicht, um die Mahlzeit zusammenzustellen. Mögliche Mahlzeiten sind Morgens 6 Uhr. Mittags 12 Uhr. Abends 18 Uhr und gegen Mitternacht 24 Uhr.

# Übung 0x04

---

## Aufgabe 3 - Fortsetzung

Die Aufgaben lauten:

- Modellieren Sie beschriebene Mini-Welt in einem ER-Diagramm und geben Sie die Kardinalitäten in der Chen-Notation an.
- Ihr Modell sollte u.A. dies abbilden können. **Das ist ein Test für Ihr Modell.**
  - Im Innengehege “Busch und Haufen” leben Igel bei 18°C, u.a. auch der Igel “Walter”. Er bekommt immer um 6 Uhr und um 18 Uhr jeweils 200g seines Futters, welches speziell für ihn aus 50% Haferflocken und 50% Weizenkleie zusammengesetzt ist.
  - Pfleger “Max” kümmert sich um die Igel und die Waldameisen (Fun Fact: Max bekommt morgens auch Haferflocken – aber das ist hier eigentlich gar nicht relevant... ).
- Optional: Implementieren Sie Ihren Entwurf in einem DBMS und hinterlegen Sie Beispieldaten.

# ÜBUNG 0x04 - EXTRA

# Übung 0x04-Extra

---

## Aufgaben

Aufgrund des Feiertages haben wir keinen neuen Vorlesungsstoff, daher üben wir im Bereich 'Modellierung' und haben das SQL-Praktikum etwas vorgezogen.

Nutzen Sie die Zeit gerne für Diskussionen um die Modelle. Wenn Sie am Ende mit den SQL-Aufgaben nicht fertig werden, ist das unproblematisch, da wir dort in der nächsten Woche weiter machen.

Ich hoffe, Sie hatten einen erholsamen Feiertag und ein schönes Halloween...



[YouTube - 2021 Full Halloween Show](#)

# Übung 0x04-Extra

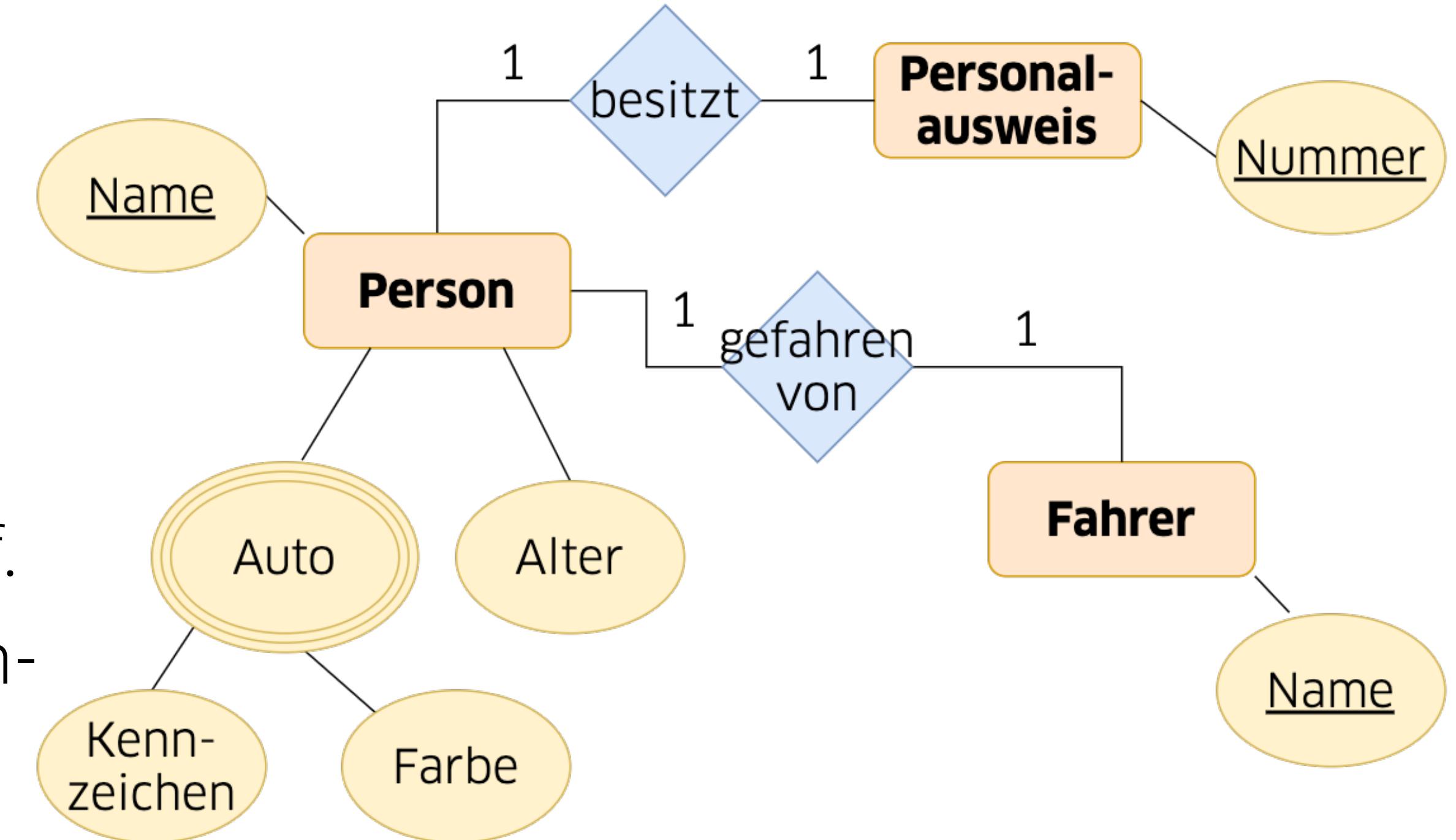
## Aufgabe 1

Diesen Entwurf bekommen Sie von Ihrem (neuen) Mitarbeiter. Er sollte modellieren:

- Eine Person besitzt einen Namen, ein Geb.datum, genau einen Personalausweis mit Nummer und eine Menge von Autos, jeweils mit Kennzeichen und Farbe.
- Jede Person darf genau eine andere Person angeben, die ihre Autos fahren darf.

Verbessern bzw. korrigieren Sie den Entwurf.

- Überlegen Sie sich einen ersten Implementationsentwurf. Wir werden diese Aufgabenstellung in der nächsten Woche systematisch für alle Fälle lösen.



## Übung 0x04-Extra

---

### Aufgabe 2 - Beispiel aus einer Prüfung

Geben Sie ein ER-Diagramm für die nachfolgend beschriebene Miniwelt an. Die Kardinalitäten sind in der Chen-Notation anzugeben. Kennzeichnen Sie die Schlüssel.

- In einem Bürogebäude sitzen Mitarbeiter mit Personalnummer und Namen an genau einem bestimmten Platz in Räumen mit eindeutiger Zimmernummer seit einem bestimmten Zeitpunkt. In den Zimmern sind Telefone mit eindeutiger Telefonnummer aufgestellt. Manchmal teilen sich die Personen des Zimmers einen Apparat, manchmal gibt es mehrere Telefone und sie stehen jeweils exklusiv einer bestimmten Person zur Verfügung und manchmal stehen Telefone auf leeren Schreibtischen und sind keiner Person zugeordnet. In jedem Fall ist aber jede Person telefonisch erreichbar. Ihr Modell soll u.a. folgendes beantworten können:
  - (i) In welchem Zimmer sitzt Frau Müller auf welchem Platz und seit wann?
  - (ii) Welche Telefonnummer hat sie und teilt sie sich den Apparat?
  - (iii) Wie viele Personen sitzen im Raum 123?

## Übung 0x04-Extra

---

### Aufgabe 3

Lesen und verstehen Sie das Kapitel “0x02 Join” in den SQL-Übungsergänzungen und bearbeiten Sie die Aufgaben dazu. Beachten Sie, dass es dazu ein SQL-Skript mit den vorgestellten SQL-Kommandos gibt ('db\_ex\_02'), so dass Sie nicht alles eingeben müssen.

Wie schon gesagt: Falls Sie aufgrund der Modellierung hier nicht fertig werden ist das kein Problem, wir werden diese Aufgabe in der nächsten Woche zu Ende führen.

# ÜBUNG 0x05

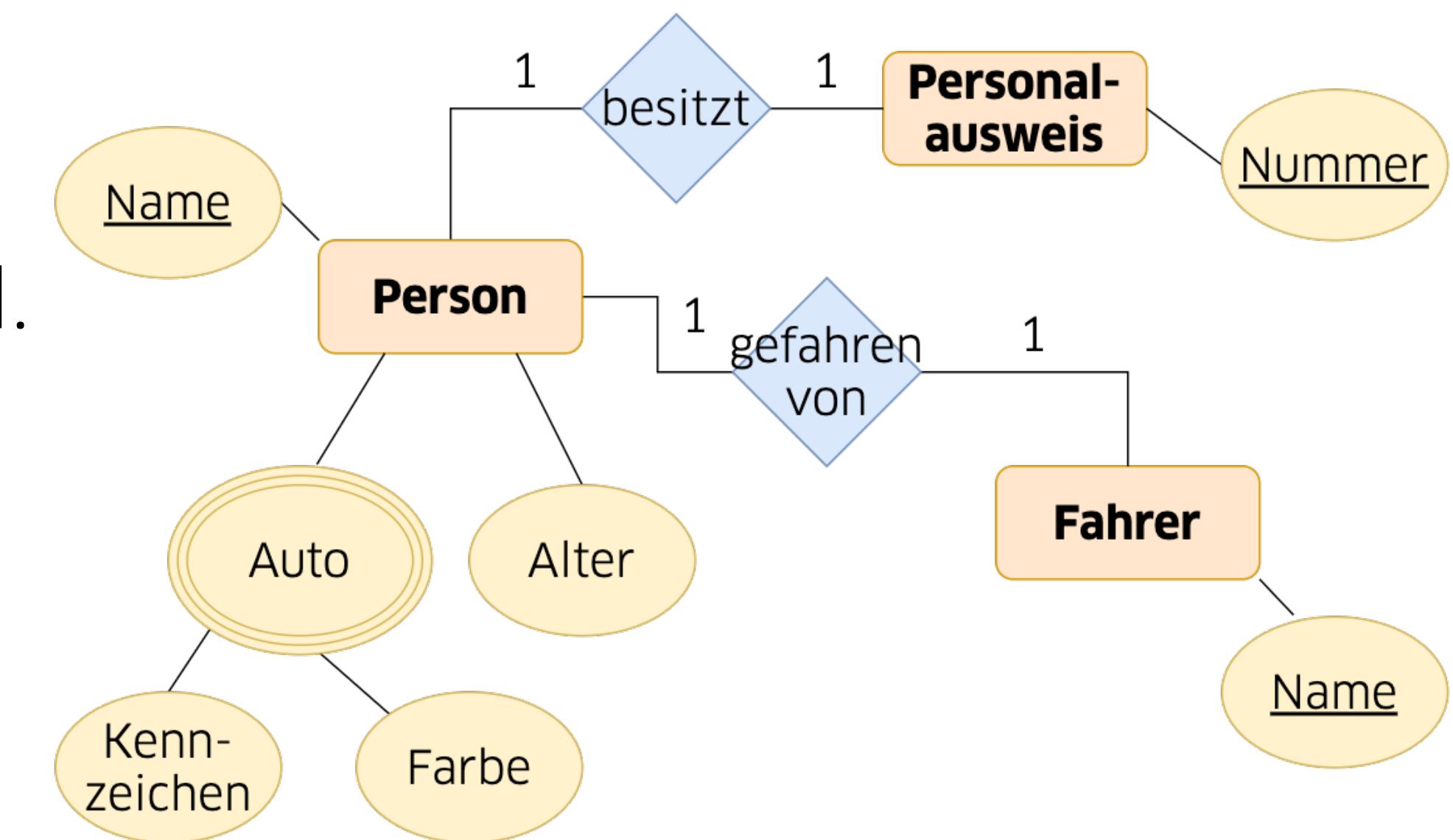
# Übung 0x05

---

## Aufgabe 1 - Ergänzung zu Aufgabe 1, Übung 0x04-Extra

Diese Modellierung haben Sie zuvor bekommen, sie verbessert (!) und sich einen ersten Implementationsentwurf überlegt.

- Vergleichen Sie Ihre Überführung in ein Datenbankschema mit den Strategien der Vorlesung aus Unit 0x05, Relationales Modell. Beachten Sie hierbei insbesondere die Umsetzung Ihrer Relationen und des mehrwertigen Attributs.
- Geben Sie alle resultierenden Tabellen (inkl. der Relationen) in der Form  $\{[a_1:T_1, \dots, a_n:T_n]\}$  an. Sie können die SQL-Datentypen verwenden.



## Übung 0x05

---

### Aufgabe 2 - Fortführung Aufgabe 3, Übung 0x04-Extra

- Lesen und verstehen Sie das Kapitel “0x02 Join” in den SQL-Übungsergänzungen und bearbeiten Sie die Aufgaben dazu. Beachten Sie, dass es dazu ein SQL-Skript mit den vorgestellten SQL-Kommandos gibt ('db\_ex\_02'), so dass Sie nicht alles eingeben müssen.
- Verstehen Sie dazu unbedingt die dazu passenden Kapitel zur Relationalen Algebra, Unit 0x05 (Selektion  $\sigma$ , Projektion  $\Pi$ , Kartesisches Produkt  $\times$ , Theta-Verbund/-Join  $\bowtie\Theta$ ).

## Übung 0x05

---

### Aufgabe 3 - WarmUp

Formulieren Sie folgenden Ausdruck als sql-Befehl (in matse\_mhist):

- $\sigma_{\text{Bezeichnung beginnt mit 'A')}(\Pi_{\text{Bezeichnung}}(\text{abteilung}))$

Formulieren Sie folgenden sql-Befehl in der relationalen Algebra.

```
SELECT P.bezeichnung, P.stueckpreis, W.bezeichnung  
FROM Produkt P JOIN Warengruppe W ON P.warengruppe_id=W.id  
WHERE P.stueckpreis>3.0;
```

# Übung 0x05

---

## Aufgabe 4

Das folgende Datenbankschema *Cinema* modelliert Kinofilme und örtliches -programm über die Relationen Orte, Filme und Programm:

```
Cinema (
    schema(Orte) = {[Kino, Adresse, Telefonnr.]},
    schema(Filme) = {[Titel, Regie, Schauspieler/in]},
    schema(Programm) = {[Kino, Titel, Zeit]}
```

)

### Beachten Sie bitte:

- Wir nehmen an, dass alle Attribute einen Text-Datentyp, z.B. Varchar, besitzen.
- Es geht hier *nicht* um die geschickte Modellierung des Schemas mit IDs, Relationen etc., sondern um *Aufgaben zur Relationalen Algebra*. Deswegen können wir auch ohne Bauchschmerzen mehrere Primärattribute verkraften...

## Übung 0x05

Cinema (

schema(Orte) = {[Kino, Adresse, Telefonnr.]},

schema(Filme) = {[Titel, Regie, Schauspieler/in]},

schema(Programm) = {[Kino, Titel, Zeit]}

)

### Aufgabe 4 - Fortsetzung

Interpretieren Sie folgende Ausdrücke der relationalen Algebra in natürlicher Sprache:

- $\Pi_{\text{Regie}}(\sigma_{\text{Schauspieler/in}=\text{Daniel Craig}}(\text{Filme}))$
- $\Pi_{\text{Schauspieler/in}}(\text{Filme} \bowtie_{\text{Filme.Titel}=\text{Programm.Titel}} \text{Programm})$
- $\Pi_{\text{Kino, Adresse}}(\text{Orte} \bowtie_{\text{Kino}=\text{Titel}} (\Pi_{\text{Titel}}(\sigma_{\text{Schauspieler/in}=\text{Marilyn Monroe}}(\text{Filme})) \setminus \Pi_{\text{Titel}}(\text{Programm})))$



Geben Sie folgende Fragen als Ausdrücke der relationalen Algebra an:

- In welchen Filmen (Titel) spielt Roger Moore mit?
- In welchen Kinos (Kino, Adresse) laufen Filme mit Scarlett Johansson?
- Welche Regie hat noch nie mit Chris Hemsworth gearbeitet?



# ÜBUNG 0x06

# Übung 0x06

---

## Aufgabe 1

Erörtern Sie mit anderen Matses die folgenden Fragen/Themen:

- Was für einen Sinn hat die “Relationale Algebra”?
- Grundoperationen der relationalen Algebra und ihre Umsetzung in SQL. Beachten Sie dazu auch die kleine Sammlung `db_ex_algebra.sql`.

# Übung 0x06

---

## Aufgabe 2

Bestimmen Sie mit Hilfe der Division ( $R \div S$ ) der relationalen Algebra aus der Liste aller Aufträge alle die Kunden, die Gas/Wasser im Januar und Elektro im Februar beauftragen möchten.

- Bestimmen Sie zunächst  $S$ .
- Bestimmen Sie  $R \div S$  durch “scharfes Hinsehen”.
- Bestimmen Sie schrittweise  $R \div S$  analog zu den Schritten in den Vorlesungsfolien.

Kunde	Gewerk	Zeitraum
Jochen	Gas/Wasser	Januar
Jochen	Elektro	Februar
Maxi	Gas/Wasser	Januar
Maxi	Putzer	Januar
Maxi	Elektro	Februar
Marie	Gas/Wasser	Februar
Kevin	Elektro	Februar

# Übung 0x06

---

## Aufgabe 3

Lesen und verstehen Sie das Kapitel “0x03 Group Functions” in den SQL-Übungs-ergänzungen und bearbeiten Sie die Aufgaben dazu.

Beachten Sie, dass es dazu ein SQL-Skript mit den vorgestellten SQL-Kommandos gibt, so dass Sie nicht alles eingeben müssen.

# ÜBUNG 0x07

# Übung 0x07

---

## Aufgabe 1

Erörtern Sie mit anderen Matses die folgenden Fragen/Themen:

- Was bedeuten “Funktionale Abhängigkeit”, “Schlüsselkandidat”, “Superschlüssel”?
- Welche Rolle spielen die “Armstrong-Axiome”?
- Was bedeutet die “Attributhülle”?

## Übung 0x07

---

### Aufgabe 2

Gegeben sei das folgende Relationenschema R. Welche funktionalen und nicht-trivialen (d.h. ohne z.B.  $AB \rightarrow B$ ) Abhängigkeiten könnten für R vorliegen?

<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
a1	b1	c1	d1
a1	b2	c2	d2
a1	b3	c1	d1
a2	b1	c1	d3

Tipp: Zählt man z.B.  $AB \rightarrow C$ ,  $AB \rightarrow D$  und  $AB \rightarrow CD$  auch als drei Abhängigkeiten, dann sollten es insgesamt 16 Stück sein...

## Übung 0x07

---

### Aufgabe 3

Zu einer Relation R sind funktionale Abhängigkeiten (FD) gegeben.

Errechnen Sie jeweils die Attributhülle  $a^+$  von  $a$  und bestimmen Sie mind. einen Schlüsselkandidaten mit Hilfe der Berechnung der Attributhülle.

(i)  $R = \{ A, B, C, D, E, F \}$ ,  $FD = \{ A \rightarrow D; B \rightarrow CE; BC \rightarrow F; C \rightarrow BE; D \rightarrow E; E \rightarrow F \}$   
 $a = A$ .

(ii)  $R = \{ A, B, C, D, E, F \}$ ,  $FD = \{ E \rightarrow F; F \rightarrow A; ABD \rightarrow E \}$   
 $a = BD$ .

(iii)  $R = \{ A, B, C, D \}$ ,  $FD = \{ A \rightarrow B; C \rightarrow B; D \rightarrow ABC; AC \rightarrow D \}$   
 $a = C$ .

Tipp: Schlüsselkandidaten können Sie unter <https://normalizer.db.in.tum.de/index.py> überprüfen. Aber bitte zuerst selber probieren!

## Übung 0x07

---

### Aufgabe 4

Leiten Sie mittels der drei Armstrong-Axiome *Reflexivität*, *Transitivität* und *Verstärkung* die *Dekompositionsregel* her.

# ÜBUNG 0x08

## Übung 0x08

---

### Aufgabe 1

Gegeben sind die folgenden Relationen  $R=\{A,B,C,D,E,\dots\}$  mit jeweils einer Menge FD von funktionalen Abhängigkeiten. Bestimmen Sie jeweils die *kanonische Überdeckung FD* und *alle Schlüsselkandidaten*:

- (a)  $FD=\{A \rightarrow BE, AE \rightarrow BD, F \rightarrow CD, CD \rightarrow BEF, CF \rightarrow B\}.$
- (b)  $FD=\{AB \rightarrow CDE, BC \rightarrow EF, BEF \rightarrow A, D \rightarrow AC, F \rightarrow E\}.$
- (c)  $FD=\{A \rightarrow CD, AB \rightarrow C, BC \rightarrow E, AB \rightarrow E\}.$
- (d)  $FD=\{F \rightarrow ADF, A \rightarrow E, D \rightarrow EG, DE \rightarrow E, G \rightarrow E, BCF \rightarrow A, G \rightarrow A\}.$
- (e)  $FD=\{C \rightarrow B, DE \rightarrow DEG, A \rightarrow G, F \rightarrow DE, A \rightarrow B, C \rightarrow G, B \rightarrow FB\}.$
- (f)  $FD=\{AE \rightarrow BCD, A \rightarrow CD, ADE \rightarrow C\}.$
- (g)  $FD=\{AE \rightarrow BCD, B \rightarrow E, ABD \rightarrow DE\}.$
- (h)  $FD=\{ADE \rightarrow BE, B \rightarrow ACD, D \rightarrow AC, ACE \rightarrow BD, C \rightarrow B\}.$

## Übung 0x08

---

### Aufgabe 1 - Fortsetzung

Beantworten Sie für alle Teilaufgaben ebenfalls die Frage:

- Ist jedes nicht-prim Attribut voll funktional abhängig von allen Schlüsseln?

Hier <https://normalizer.db.in.tum.de/index.py> können Sie überprüfen, ob Ihr Ergebnis stimmt und welche Attribute sich haben entfernen lassen.

## Übung 0x08

---

### Anmerkung

Bei den Schlüsselkandidaten zuerst die rechten Seiten der funktionalen Abhängigkeiten überprüfen, ob Attribute fehlen bzw. nicht gefolgt werden. Die gehören in den Schlüsselkandidaten, da man sie nicht folgern kann (im Video diskutiert).

Die kanonische Überdeckung berechnet man am besten genauso wie im Algorithmus beschrieben. Einfach der Reihe nach die Attribute in der Links- und Rechtsreduktion über die Attributhülle testen und ggf. die FD vereinfachen.

Beachten Sie dazu auch die Anmerkungen aus der Vorlesung.

# ÜBUNG 0x09

# Übung 0x09

---

## Kommentar zur Übung vorab

- Wir werden dieses Thema auch in der nächsten Woche noch üben! Es ist zwar sehr technisch, aber ist m.E. ganz gut lernbar und damit auch ein guter Punkteliereant...
- In der kommenden Übung werden wir noch ein paar Lösungen kommentieren, aber zunächst sollten Sie selbst etwas Gefühl für die Zerlegungen bekommen.
- Ja, es sind viele Teilaufgaben. Wenn Sie ein Unterthema verstanden haben, z.B. die Ermittlung der Schlüssel, können Sie gerne, etwa mit Hilfe des DB-Normalizers, diese ermitteln lassen und sich auf die Algorithmen konzentrieren.  
DB-Normalizer: <https://normalizer.db.in.tum.de/index.py>
- Ich würde mich darüber freuen, wenn Sie bei der Evaluation mitmachen, konstruktiv kommentieren, Verbesserungen vorschlagen oder auch loben. Es ist im Moment nicht leicht, Feedback zu bekommen. Sie können bei Bedarf auch immer per (privater) Mail fragen oder die Live-Session am Montag nutzen.

## Übung 0x09

---

### Aufgabe 1

Gegeben sind die folgenden Relationen  $R=\{A,B,C,D,E,\dots\}$  mit jeweils einer Menge FD von funktionalen Abhängigkeiten (siehe auch Aufgabe 8.1).

- (a)  $FD=\{A \rightarrow BE, AE \rightarrow BD, F \rightarrow CD, CD \rightarrow BEF, CF \rightarrow B\}$ .
- (b)  $FD=\{AB \rightarrow CDE, BC \rightarrow EF, BEF \rightarrow A, D \rightarrow AC, F \rightarrow E\}$ .
- (c)  $FD=\{A \rightarrow CD, AB \rightarrow C, BC \rightarrow E, AB \rightarrow E\}$ .
- (d)  $FD=\{F \rightarrow ADF, A \rightarrow E, D \rightarrow EG, DE \rightarrow E, G \rightarrow E, BCF \rightarrow A, G \rightarrow A\}$ .
- (e)  $FD=\{C \rightarrow B, DE \rightarrow DEG, A \rightarrow G, F \rightarrow DE, A \rightarrow B, C \rightarrow G, B \rightarrow FB\}$ .
- (f)  $FD=\{AE \rightarrow BCD, A \rightarrow CD, ADE \rightarrow C\}$ .
- (g)  $FD=\{AE \rightarrow BCD, B \rightarrow E, ABD \rightarrow DE\}$ .
- (h)  $FD=\{ADE \rightarrow BE, B \rightarrow ACD, D \rightarrow AC, ACE \rightarrow BD, C \rightarrow B\}$ .

## Übung 0x09

---

### Aufgabe 1 - Fortsetzung

- Prüfen Sie, in welcher Normalform das jeweilige Schema vorliegt (1NF, 2NF, 3NF oder BCNF).
- Überführen Sie es jeweils mit dem Synthesealgorithmus in 3NF, falls notwendig, und
- mit dem Dekompositionsalgorithmus in BCNF, falls notwendig. Denken Sie hier daran, dass Sie die neuen Schemata ebenfalls testen und ggf. weiter zerlegen müssen.

Falls Sie Zwischenergebnisse aus Aufgabe 8.1. vorliegen haben (Schlüssel, kanonische Überdeckung), können Sie diese natürlich nutzen!

Hier <https://normalizer.db.in.tum.de/index.py> können Sie überprüfen, ob Ihr Ergebnis stimmt.

# Übung 0x09

---

## Aufgabe 2

Suchen Sie sich interessante Beispiele im DB-Normalizer. Irgendetwas, was Ihnen auffällt, was überraschend war, was bei Ihnen nochmal zu einem Erkenntnisgewinn geführt hat. Nutzen Sie dazu die Möglichkeit, zufällige Aufgaben zu generieren.

Tauschen Sie sich mit anderen Matses darüber aus. Wenn Sie eine wirklich verrückte oder interessante Normalisierungsaufgabe gefunden haben, können Sie sie mir auch gerne mailen: [a.voss@fh-aachen.de](mailto:a.voss@fh-aachen.de)

# ÜBUNG 0x0A

# Übung 0x0a

---

## Aufgabe 1

- Setzen Sie sich mit Matses Ihres Vertrauens zusammen und üben Sie diese Art der Aufgabe, bis Sie es entweder können oder mir oder den betreuenden Kollegen Fragen schicken oder stellen zu Punkten, die unklar sind.  
Nutzen Sie <https://normalizer.db.in.tum.de/index.py>, um interessante Aufgaben zu generieren.

Die Aufgabenstellung lautet wie in Aufgabe 0x09.1, also kurz:

- Schlüsselkandidaten?
- Welche Normalform liegt vor (1NF, 2NF, 3NF oder BCNF)?
- Synthesealgorithmus (für 3NF) oder Dekompositionsalgorithmus (für BCNF).

# Übung 0x0a

---

## Aufgabe 1 - Anmerkungen

- Bei der BCNF taucht immer wieder die berechtigte Frage auf, welche impliziten Abhängigkeiten müssen denn zwingend erkannt werden? Zum Teil werden in dem Tool auch "unnötige" funktionale Abhängigkeiten angezeigt, die aus den anderen abgeleitet werden und streichbare Attribute enthalten oder trivial sind.
- Wir sind primär an denen interessiert, die nicht trivial aber wichtig sind, weil sie ein Verständnis der Aufgabe belegen. Dazu nun ein Beispiel mit Abhängigkeiten und einer unverbindlichen Einschätzung, welche man sehen sollte.
- $R=\{A,B,C,D,E\}$  mit  $FDs=\{ AD \rightarrow B, A \rightarrow CD, C \rightarrow D, B \rightarrow C \}$ . Es gilt:
  - $AD \rightarrow B$  verletzt BCNF und folglich ergibt die erste Zerlegung:  
 $R_1=\{ABD\}$  mit zugehörigen  $FD_1s=\{ AD \rightarrow B, A \rightarrow D, A \rightarrow B, B \rightarrow D \}$ .
  - Diese werden nun kommentiert.

## Übung 0x0a

---

### Aufgabe 1 - Fortsetzung Anmerkungen

- Aufgabe:  $R=\{A,B,C,D,E\}$  mit FDs={  $AD \rightarrow B$ ,  $A \rightarrow CD$ ,  $C \rightarrow D$ ,  $B \rightarrow C$  }. Es gilt:  
 $AD \rightarrow B$  verletzt BCNF, Zerlegung lautet  $R_1=\{ABD\}$ ,  $FD_{1S}=\{ AD \rightarrow B, A \rightarrow D, A \rightarrow B, B \rightarrow D \}$ .
  - $AD \rightarrow B$  sollte in jedem Fall angegeben werden, denn das ist die ursprüngliche BCNF-verletzende Abhängigkeit.
  - $A \rightarrow D$  sollte angegeben werden, denn das bedeutet, dass Ihnen klar ist, dass  $A \rightarrow CD$  auch  $A \rightarrow D$  umfasst.
  - $A \rightarrow B$  ist tricky. Hier würde ich eher Extrapunkte vergeben für die, die es sehen und nichts abziehen, falls man diese übersieht. Warum gilt diese? Aus  $A \rightarrow CD$  folgt  $A \rightarrow D$  (siehe Punkt vorher) und mit  $A \rightarrow D$  gilt auch  $AD \rightarrow B$ , also auch  $A \rightarrow B$ .
  - $B \rightarrow D$  sollte angegeben werden, denn das prüft das Verständnis für einfache transitive Abhängigkeiten:  $B \rightarrow C$  und  $C$  nicht in  $\{ABD\}$ , aber  $C \rightarrow D$ , also  $B \rightarrow D$ .
- Ich hoffe, die Idee hinter der Einschätzung ist einigermassen verständlich.

# Übung 0x0a

## Beispiel

$$R = \{ A, B, C, D, E \}$$

← Achtung  $\checkmark$   
evtl. kommt  
ein Attribut in  
den FDs nicht  
vor → nicht  
vergessen  $\checkmark$

$$\text{FDs: } \{ \begin{array}{l} ACD \rightarrow B \\ D \rightarrow BC \\ BC \rightarrow ABC \end{array} \}$$

Untersuchung  
nicht  
vollständig?

- Scharfes Hinzeichen: - wg.  $D \rightarrow BC$  ist  $ACD \rightarrow B$  komplett überflüssig
- in  $BC \rightarrow ABC$  ist  $\rightarrow BC$  trivial  
also äquivalent zu  $BC \rightarrow A$

- Diese Äquivalenzen könnte man direkt ausnutzen, aber Achtung: Fehleranfällig  $\checkmark$
- Im Zweifel stets Links- und Rechtsreduktion,  
falls gefragt (3NF oder Synthesealg.)
- Diese Überlegung sollte bei der Reduktion herauskommen.

Aber zunächst: Welche NF liegt vor?

⇒ Schlüsselkandidat  $\checkmark$

(1)

### Schlüssel:

z.B. gesucht  $K$  mit  $K \subseteq R$

$$\text{und } K^+ = \text{Attr. Menge (FDs, 4)} = R$$

- Scharfes Hinzeichen: sowohl  $D$  als auch  $E$  können rechts in FDs nicht vor  
⇒ sie können nie gefolgt werden  
⇒ müsste mind. in  $K$

Reicht das schon?

$$DE^+ : DE \rightarrow DE \rightarrow BCDE \rightarrow ABCDE \\ = R \quad \checkmark$$

Also:  $DE$  notwendig und hinreichend,  
es gibt kein weiterer Schlüsselkand., nur  
noch Superschlüssel, z.B.  $ADE$

### NF:

- Erste NF liegt vor, da  $A, B, C, \dots$  als automatisch vorausgesetzt werden  $\checkmark$

### (2NF)

### Zweite NF:

- Welche Attribute sind prim?  $DE$   
" nicht-prim?  $ABC$  → nur die interessant für Test

→ Skript S.299: Lernen

- wg.  $D \rightarrow BC$  ist z.B.  $B$  nur von  $D$ , also nicht voll-funktional von  $DE$ ,  
abhängig ⇒ keine 2NF  $\checkmark$
- keine 3NF, 2NF (2)

# Übung 0x0a

## Beispiel

Wie kommt du aufgabestellung?

$\xrightarrow{\text{Überführung in 3NF}}$  mit Synthesealg. ?  $\Rightarrow$  kanonische Überdeckung  $FD^c$

$BCNF$  mit Dekomposition alg. ?  $\Rightarrow$  direkt

Erinnerung: Welche 3NF und BCNF geben, da 2NF verletzt

Synthesealg.: Skript S. 307:  $FD^c$  gehört  
 ~ Linkss- und Rechtsreduktion  
 $\Rightarrow$  Skript S. 269

Linksreduktion: Bsp.

$\alpha \rightarrow \beta$  sei  $\overbrace{ACD}^{\alpha} \rightarrow \overbrace{B}^{\beta}$ ; Überdeckk., ob  $\overbrace{C}^a$  überflüssig ist, d.h.

gilt  $\beta \in \text{Attr. Hülle } (\bar{F})$ ,  $\alpha - a$  ?

also  $B \subseteq AD^+$  ?

wg.  $AD \rightarrow AD \rightarrow AD \setminus C$ , also Ja  $\checkmark$   
 $\Rightarrow C$  überfl.

Ersche  $ACD \rightarrow B$  durch  $AD \rightarrow B$

Analog: A überflüssig, D nicht überfl.  
 $\Rightarrow$  bleibt  $D \rightarrow B$

Rechtsreduktion: Bsp.

$\alpha \rightarrow \beta$  sei  $\overbrace{BC}^{\alpha} \rightarrow \overbrace{A \setminus C}^{\beta}$ ; Überdeckk., ob  $\overbrace{B}^b$  überfl., d.h.

gilt  $\beta \in \text{Attr. Hülle } (\bar{F}) \setminus \alpha \rightarrow \beta \cup \alpha \rightarrow \beta - b, c$   
 $\Rightarrow$  Attr. Hülle  $(\{D \rightarrow B, D \rightarrow BC, BC \rightarrow AC\}, BC)$  ?

wg.  $BC \rightarrow BC \xrightarrow{AD^c}$ , also Ja (trivial)  
 $\Rightarrow B$  überfl.

Ersche  $BC \rightarrow A \setminus C$  durch  $BC \rightarrow AC$ , etc...

Achtung: - Die Folgerung ist trivial, es geht hier um  
 die Verz. der der Schritte v  
 - Normalerweise beginnt man vorne...

Bsp.  $\alpha \rightarrow \beta$  sei  $D \rightarrow B$

formal: gilt  $B \in \text{Attr. Hülle } (\{D \rightarrow BC, BC \rightarrow A \setminus C\}, D)$  ?

wg.  $D \rightarrow BC$  Ja, also

Ersche  $D \rightarrow B$  durch  $D \rightarrow \emptyset$  ... sinnlos!  
 $\Rightarrow$  man kann es direkt weglassen

Links- und Rechtsred.:

$$FD^c = \{D \rightarrow BC, BC \rightarrow A\}$$

$\Rightarrow$  Synthesealg. (S. 307)

$$\begin{aligned} R_1 &= \{BC \setminus D\} & FD_1 &= \{D \rightarrow BC\}, D \text{ schüssig} \\ R_2 &= \{A \setminus BC\} & FD_2 &= \{BC \rightarrow A\}, BC \text{ schüssig} \\ R_3 &= \{DE\} & \text{wg. } DE \text{ nirgends enthalten} \end{aligned}$$

# Übung 0x0a

Dekaputslösung: Skript S. 321  
Kriterien BCNF S. 319

## Beispiel

$$FD = \{ ACD \rightarrow B, D \rightarrow BC, BC \rightarrow ABC \}, \text{ Schloss DE}$$

$A \rightarrow B : ACD \rightarrow B$   
↑ weder trivial, noch superschlüssel  
 $\Rightarrow$  verletzt BCNF

$$\begin{aligned} R_1 &= A \cup B = A \underline{B C D} & FD_1 &= \{ ACD \rightarrow B, D \rightarrow BC, \\ R_2 &= R - B = A \underline{C D E} & BC \rightarrow ABC, \dots \} & \text{Schloss D} \end{aligned}$$
$$FD_2 = \{ D \rightarrow BC, \dots \} \text{ Schloss DE}$$

AchLg: Weitere Abhängigkeiten sind  
mgl., d.h. nur weil B gestrichen wurde,  
fallen nicht alle Abh. weg?

$\Rightarrow$  jetzt jeweils einzeln für  $R_1, FD_1$  und  $R_2, FD_2$

$R_1 : BC \rightarrow A$  verletzt BCNF

$$\sim R_{11} = A \underline{BC} \quad FD_{11} = \{ BC \rightarrow ABC \}$$
$$R_{12} = BC \underline{D} \quad FD_{12} = \{ D \rightarrow BC \}, \text{ etc.}$$

$\Rightarrow$  je nach Wahl der betrachteten Abh. unterschiedl. Ergebnis.

$\Rightarrow$  Abh. auf Wiederholung generiert „merkwürdige“ Abh. (redundanz)

$\Rightarrow$  BCNF auf mglst. einfache FDs  
sinnvoll, d.h. kanonische Wiederholung nicht notw.,  
aber „Schloss Minze“ hilft.

# Übung 0x0a

---

## Aufgabe 2

Lesen und verstehen Sie das Kapitel “0x04 Subselects” in den SQL-Übungsergänzungen und bearbeiten Sie die Aufgaben dazu. Beachten Sie, dass es dazu ein SQL-Skript mit den vorgestellten SQL-Kommandos gibt, so dass Sie nicht alles eingeben müssen.

# ÜBUNG 0x0B

# Übung 0x0b

---

## Aufgabe 1

Lesen und verstehen Sie das Kapitel “0x05 Schemas und Tabellen” in den SQL-Übungsergänzungen und bearbeiten Sie die Aufgaben dazu. Beachten Sie, dass es dazu ein SQL-Skript mit den vorgestellten SQL-Kommandos gibt, so dass Sie nicht alles eingeben müssen.

## Zum Schluß

**Ich wünsche Ihnen ein besinnliches Weihnachtsfest  
und einen guten Rutsch ins Jahr 2022.  
Bleiben Sie gesund!**

