

2. Übungsblatt (WS 2019) – Musterlösung

3.0 VU Datenmodellierung / 6.0 VU Datenbanksysteme

Informationen zum Übungsblatt

Allgemeines

In diesem Übungsteil sollen Sie Aufgabenstellungen aus den Bereichen SQL und Normalformtheorie bearbeiten.

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 06.11. 12:00 Uhr:** Bearbeiten der SQL-Aufgaben im Online Tool möglich
- bis 13.11. 12:00 Uhr:** Upload der Abgabe über TUWEL
- ab 22.11. 13:00 Uhr Korrektur und Feedback in TUWEL verfügbar

Tutorsprechstunden (freiwillig)

Rund eine Woche vor der Abgabedeadline bieten die TutorInnen Sprechstunden an. Falls Sie Probleme mit oder Fragen zum Stoff des Übungsblattes haben, es Verständnisprobleme mit den Beispielen oder technische Fragen gibt, kommen Sie bitte einfach vorbei. Die TutorInnen beantworten Ihnen gerne Ihre Fragen zum Stoff, oder helfen Ihnen bei Problemen weiter.

Ziel der Sprechstunden ist es, Ihnen beim **Verständnis des Stoffs** zu helfen, nicht, das Übungsblatt für Sie zu rechnen, oder die eigenen Lösungen vorab korrigiert zu bekommen.

Die Teilnahme ist vollkommen freiwillig — Termine und Orte der Tutorensprechstunden finden Sie in TUWEL.

Durchsprache der Übungsbeispiel (freiwillig)

In den Tagen nach Rückgabe der korrigierten Abgaben gibt es die Möglichkeit die Übungsbeispiele in kleineren Gruppen (max. 25 Personen) durchzusprechen. Jede dieser Gruppen wird von einer Assistentin/einem Assistenten geleitet. Der genaue Ablauf in einer Übungsgruppe kann variieren, und hängt auch von Ihren Wünschen und Fragen ab. Die grundsätzliche Idee ist es, die Beispiele durchzurechnen, und speziell auf Ihre Fragen und mögliche Unklarheiten einzugehen. Die (relativ) kleine Gruppengröße soll eine aktive Teilnahme ermöglichen. Daher ist es auch wichtig, dass Sie sich bereits im Vorfeld mit Ihrer korrigierten Abgabe auseinandersetzen, und Unklarheiten identifizieren. Trauen Sie sich, entsprechend Fragen zu stellen – keine Frage kann irgendeinen (negativen) Einfluss auf Ihre Note haben.

Die Teilnahme an so einer Gruppe ist absolut freiwillig. Um die Gruppengröße klein zu halten ist eine Anmeldung in TUWEL erforderlich. Termine und Orte finden Sie in TUWEL.

Weitere Fragen – TUWEL Forum

Sie können darüber hinaus das TUWEL Forum verwenden, sollten Sie inhaltliche oder organisatorische Fragen haben.

SQL

Aufgabe 1 (eSQL)	[5 Punkte]
-------------------------	------------

Lösen Sie in unserer Online-Übungsumgebung **die ersten 10 SQL-Aufgaben (Aufgabe 1-10)** des aktuellen Übungskurses. (*Hinweis:* Es ist möglich, dass wir zu Übungszwecken im Laufe der Zeit zusätzliche Aufgaben online stellen. Es steht Ihnen frei, diese ebenfalls zu lösen – Punkte für dieses Übungsblatt gibt es jedoch nur für die ersten 10 Aufgaben.)

Sie erreichen die Umgebung über TUWEL: Wählen Sie im Abschnitt “2.Übungsblatt” die Aktivität **eSQL Tool**. Sie benötigen kein weiteres Passwort, die Authentifizierung erfolgt über TUWEL.

Der verpflichtende SQL-Test wird über die selbe Plattform abgewickelt. Es empfiehlt sich daher zusätzlich auch mit Beispielen aus den vorigen Semestern zu üben.

Achtung!

**Abweichende Deadline für den Abschluss von Aufgabe 1:
Mittwoch, 6. November 2019, 12:00 Uhr!**

Normalformtheorie

Aufgabe 2 (Funktionale Abhängigkeiten)

[0.4 Punkte]

Geben ist ein Relationenschema

Angebot (Kaffee, Bohne, Land, Roestung, Verarbeitung, Typ)

mit der folgenden Ausprägung (nach Kaffee sortiert):

Angebot					
Kaffee	Bohne	Land	Röstung	Verarbeitung	Typ
Alice	Arabica	Brasilien	plus	natur	rein
Bob	Arabica	Ecuador	plus	nass	rein
Carol	Arabica	Peru	full	nass	mischung
Carol	Arabica	Honduras	full	natur	mischung
Carol	Robusta	Indien	plus	nass	mischung
Dan	Arabica	Brasilien	full	nass	mischung
Dan	Robusta	Indien	full	natur	mischung
Eve	Arabica	Indonesien	full	nass	rein
Faythe	Arabica	Äthiopien	plus	natur	mischung
Faythe	Arabica	Guatemala	plus	nass	mischung
Grace	Arabica	Äthiopien	normal	nass	rein
Rupert	Robusta	Indien	full	nass	rein

Überprüfen Sie für jede der unten angegebenen funktionalen Abhängigkeiten, ob sie auf der gegebenen Ausprägung gelten oder nicht. Geben Sie für jede FD die Antwort (ja/nein) an. Falls eine FD nicht erfüllt ist geben Sie außerdem ein entsprechendes Gegenbeispiel an. Wenn eine FD erfüllt ist geben Sie ein Tupel an, welches man der Ausprägung hinzufügen könnte um die FD zu verletzen.

(a) Land \rightarrow Bohne

Ja

Ein mögliches Tupel wäre

- (New, Robusta, Indonesien, full, natur, mischung),

wobei nur die Werte “Robusta” und “Indonesien” für eine Verletzung der FD notwendig sind. Die anderen Werte können frei gewählt werden.

(b) Bohne \rightarrow Land.

Nein

Es existiert eine große Menge an möglichen Gegenbeispielen, zwei Beispiele sind

- (Alice, Arabica, Brasilien, plus, natur, rein) und (Bob, Arabica, Ecuador, plus, nass, rein):
Arabica = Arabica aber *Brasilien \neq Ecuador*.
- (Eve, Arabica, Indonesien, full, nass, rein) und (Faythe, Arabica, Äthiopien, plus, natur, mischung):
Arabica = Arabica aber *Indonesien \neq Äthiopien*.
- Generell dienen sämtliche Paare an Tupeln mit den selben Werten für Bohne aber unterschiedlichen Werten für Land als mögliche Gegenbeispiele.

- (c) **Kaffee, Bohne → Land.** **Nein**
 Das einzige Gegenbeispiel in diesem Fall sind die beiden Tupel
- (Carol, Arabica, Peru, full, nass, Mischung) und
 (Carol, Arabica, Honduras, full, natur, Mischung):
(Carol, Arabica) = (Carol, Arabica), aber *Peru ≠ Honduras*.
- (d) **Land → Bohne, Verarbeitung.** **Nein**
 Es existieren wiederum mehrere Gegenbeispiele, zwei davon sind
- (Alice, Arabica, Brasilien, plus, natur, rein) und
 (Dan, Arabica, Brasilien, full, nass, Mischung):
Brasilien = Brasilien aber *Arabica, natur ≠ Arabica, nass*.
 - (Dan, Robusta, Indien, full, natur, Mischung) und
 (Rupert, Robusta, Indien, full, nass, rein):
Indien = Indien aber *(Arabica, natur) ≠ (Arabica, nass)*.
- (e) **Bohne, Land, Röstung → Kaffee, Verarbeitung.** **Nein**
 In diesem Fall besteht das einzige Gegenbeispiel aus
- (Rupert, Robusta, Indien, full, nass, rein) und
 (Dan, Robusta, Indien, full, natur, Mischung):
(Robusta, Indien, full) = (Robusta, Indien, full) aber *(Rupert, nass) ≠ (Dan, natur)*

Aufgabe 3 (Äquivalenz Funktionaler Abhängigkeiten)

[0.6 Punkte]

- (a) Gegeben ist ein Relationenschema $GHIJKL$ und zwei Mengen F_1 und F_2 von funktionalen Abhängigkeiten.

$$F_1 = \{HI \rightarrow GJ, G \rightarrow KL, IL \rightarrow JK, GK \rightarrow I\}$$

$$F_2 = \{HI \rightarrow GL, G \rightarrow KLI, IL \rightarrow JK, GKH \rightarrow H\}$$

Sind F_1 und F_2 äquivalent? Begründen Sie Ihre Antwort mit Hilfe der Hüllen der beiden Mengen an FDs und dokumentieren Sie den Lösungsweg.

Lösung:

Ja, $F_1 \equiv F_2$. Dies folgt aus $F_1^+ = F_2^+$ (d.h. die Hüllen der beiden Mengen an FDs sind identisch). Um die Gleichheit $F_1^+ = F_2^+$ zu zeigen werden wir nicht die beiden Hüllen F_1^+ und F_2^+ berechnen, sondern nur zeigen dass für jede FD $F \in F_1$ auch $F \in F_2^+$ gilt, und umgekehrt für jede FD $F \in F_2$ auch $F \in F_1^+$ gilt. Daraus folgt sofort $F_1^+ = F_2^+$.

Um wiederum zu zeigen, dass eine FD $\alpha \rightarrow \beta$ in F_i^+ (für $i \in \{1, 2\}$) liegt, berechnen wir die Attributhülle γ von α unter F_i . Das Ergebnis ergibt sich dann aus der Beziehung $F_i \models \alpha \rightarrow \text{AttrH}(\alpha, F_i)$.

Man könnte natürlich auch anders vorgehen, und $F_i \models \alpha \rightarrow \beta$ auf andere Arten zeigen, z.B. mittels der Armstrong Axiome.

$$F_1^+ \subseteq F_2^+:$$

$HI \rightarrow GJ$ Wir berechnen $AttrH(HI, F_2)$.

- Auf Grund der FD $HI \rightarrow GL$ erhalten wir $HIGL \in AttrH(HI, F_2)$.
- Auf Grund der FD $IL \rightarrow JK$ erhalten wir $HIGLJK \in AttrH(HI, F_2)$.

Nachdem die Attributhülle nun bereits sämtliche Attribute umfasst können wir die Berechnung abbrechen und erhalten

$AttrH(HI, F_2) = GHIJKL$. Daher ist $GJ \subseteq AttrH(HI, F_2)$, und daraus folgt $F_2 \models HI \rightarrow GJ$, d.h. $HI \rightarrow GJ \in F_2^*$.

$G \rightarrow KL$ Nachdem F_2 die FD $G \rightarrow KLI$ enthält folgt $G \rightarrow KL \in F_2^+$ sofort.

$IL \rightarrow JK$ Die FD befindet sich direkt in F_2 , es bleibt daher nichts zu zeigen.

$GK \rightarrow I$ $AttrH(GK, F_2) = GKLIJ$. Da $I \in AttrH(GK, F_2)$ gilt $GK \rightarrow I \in F_2^*$.

(Zur Berechnung der Attributhülle: Aus GK folgt wegen $G \rightarrow KLI$ dass $KLI \in AttrH(GK, F_2)$. Man könnte an dieser Stelle bereits abbrechen, da I bereits als Teil der Attributhülle identifiziert wurde. Berechnet man die komplette Hülle erhält man noch $J \in AttrH(GK, F_2)$ auf Grund der FD $IL \rightarrow JK$ und der Tatsache dass $IL \in AttrH(GK, F_2)$.)

Da jede FD aus F_1 in der Hülle von F_2 enthalten ist, erhalten wir $F_1^* \subseteq F_2^*$.

$F_2^+ \subseteq F_1^+$:

$HI \rightarrow GL$ Wir berechnen $AttrH(HI, F_1)$.

- Auf Grund der FD $HI \rightarrow GJ$ erhalten wir $HIGJ \in AttrH(HI, F_1)$.
- Auf Grund der FD $G \rightarrow KL$ erhalten wir $HIGJKL \in AttrH(HI, F_1)$.

Nachdem die Attributhülle nun bereits sämtliche Attribute umfasst können wir die Berechnung abbrechen und erhalten

$AttrH(HI, F_1) = GHIJKL$. Daher ist $GL \subseteq AttrH(HI, F_1)$, und daraus folgt $F_1 \models HI \rightarrow GL$, d.h. $HI \rightarrow GL \in F_1^*$.

$G \rightarrow KLI$ Wir berechnen $AttrH(G, F_1)$.

- Auf Grund der FD $G \rightarrow KL$ erhalten wir $GKL \in AttrH(G, F_1)$.
- Auf Grund der FD $GK \rightarrow I$ erhalten wir $GKLI \in AttrH(G, F_1)$.

Nachdem wir nun bereits gezeigt haben dass KLI in der Attributhülle enthalten sind, könnten wir die Berechnung bereits abbrechen. Wir berechnen trotzdem weiter die gesamte Hülle.

- Auf Grund der FD $IL \rightarrow JK$ erhalten wir $GKLIJK \in AttrH(G, F_1)$.

Nachdem die Attributhülle nun bereits sämtliche Attribute umfasst können wir die Berechnung abbrechen.

Wir erhalten $AttrH(G, F_1) = GHIJKL$. Daher ist $KLI \subseteq AttrH(G, F_1)$, und daraus folgt $F_1 \models G \rightarrow KLI$, d.h. $G \rightarrow KLI \in F_1^*$.

$IL \rightarrow JK$ Die FD befindet sich direkt in F_1 , es bleibt daher nichts zu zeigen.

$GKH \rightarrow H$ Hier handelt es sich um eine triviale FD (die rechte Seite ist eine Teilmenge der linken Seite). Triviale FDs sind immer erfüllt, und in jeder Hülle enthalten.

Da jede FD aus F_1 in der Hülle von F_2 enthalten ist, erhalten wir $F_1^* \subseteq F_2^*$.

Da sowohl $F_1^* \subseteq F_2^*$ gilt, als auch $F_2^* \subseteq F_1^*$, erhalten wir $F_1^* = F_2^*$, und daher sind F_1 und F_2 äquivalent.

- (b) Betrachten Sie die Menge F_1 an funktionalen Abhängigkeiten aus Aufgabe a). Zeigen Sie mit Hilfe der Armstrong-Axiome dass $F_1 \models \{HI \rightarrow GL\}$ gilt (dokumentieren Sie den Lösungsweg).

Lösung:

Um zu zeigen dass $F_1 \models \{HI \rightarrow GL\}$ gilt, müssen wir die FD $HI \rightarrow GL$ mit Hilfe der Armstrong-Axiome aus F_1 ableiten. Eine mögliche solche Ableitung ist

$$\begin{array}{c}
 \text{Dekomposition} \quad \frac{\frac{[Gegeben] \quad HI \rightarrow GJ}{HI \rightarrow G} \quad \frac{[Gegeben] \quad HI \rightarrow GJ}{HI \rightarrow L}}{HI \rightarrow GL} \quad \frac{\frac{[Gegeben] \quad G \rightarrow KL}{GJ \rightarrow JKL} \quad \frac{[Gegeben] \quad GJ \rightarrow L}{GJ \rightarrow L}}{GJ \rightarrow L} \quad \begin{array}{l} \text{Verstärkung} \\ \text{Dekomposition} \\ \text{Transitivität} \\ \text{Vereinigung} \end{array}
 \end{array}$$

Aufgabe 4 (Kanonische Überdeckung)

[2 Punkte]

Bestimmen Sie eine kanonische Überdeckung der Mengen $\mathcal{F}_1, \mathcal{F}_2$ von funktionalen Abhängigkeiten über dem Relationenschema $\mathcal{R} = HIJKLMN$ und dokumentieren Sie den Lösungsweg.

- (a) $\mathcal{F}_1 = \{KLN \rightarrow HI, I \rightarrow KL, H \rightarrow HKI, ILM \rightarrow LM, KH \rightarrow LM, MN \rightarrow IL\}$
 (b) $\mathcal{F}_2 = \{JL \rightarrow HM, L \rightarrow I, HMI \rightarrow KN, JL \rightarrow KN, H \rightarrow L, HLM \rightarrow K, HM \rightarrow J, J \rightarrow N\}$

Lösung:

- (a) Die Kanonische Überdeckung lässt sich mithilfe der folgenden vier Schritten berechnen: (Die Lösung ist nicht eindeutig.)

- **Dekomposition:**

Mit der Regel der Dekomposition erhalten wir die Menge

$$\{KLN \rightarrow H, KLN \rightarrow I, I \rightarrow K, I \rightarrow L, H \rightarrow H, H \rightarrow K, H \rightarrow I, ILM \rightarrow L, ILM \rightarrow M, KH \rightarrow L, KH \rightarrow M, MN \rightarrow I, MN \rightarrow L\}$$

- **Linksreduktion:**

Wir müssen uns jetzt sukzessive die Frage stellen, ob für eine funktionale Abhängigkeit ein Attribut auf der linken Seite “überflüssig” ist, d.h. weggelassen werden kann so dass sich die resultierende Menge an FDs äquivalent zur ursprünglichen Menge an FDs ist.

Bei den FDs $KLN \rightarrow H$ und $KLN \rightarrow I$ ist dies nicht der Fall, hier kann daher nichts reduziert werden. Die nächsten FDs haben auf der linken Seite jeweils nur ein Attribut stehen, sind daher auch bereits linksreduziert. (Die FD $H \rightarrow H$ ist trivial und könnte daher sofort eliminiert werden; wir halten uns hier jedoch an den in der VO gebrachten Algorithmus und belassen die FD vorerst noch.)

Die FD $ILM \rightarrow L$ ist ebenfalls eine triviale FD. Wir könnten Sie daher entweder sofort eliminieren, oder zumindest ergibt sich ohne dass man die Attributhülle berechnen müsste dass I und M auf der linken Seite überflüssig sind. Die FD kann daher zu $L \rightarrow L$ reduziert werden. Die analoge Argumentation gilt für $ILM \rightarrow M$, welche zu $M \rightarrow M$ reduziert wird.

Die $KH \rightarrow L$ kann zu $H \rightarrow L$ reduziert werden: Die Berechnung der Attributhülle unter der aktuellen Menge $F_{cur} = \{KLN \rightarrow H, KLN \rightarrow I, I \rightarrow K, I \rightarrow L, H \rightarrow H, H \rightarrow K, H \rightarrow I, L \rightarrow L, M \rightarrow M, KH \rightarrow L, KH \rightarrow M, MN \rightarrow I, MN \rightarrow L\}$ an funktionalen Abhängigkeiten ergibt: Beginnend mit $\{H\}$ erhalten wir auf Grund der FD $H \rightarrow K$ dass $\{H, K\} \subseteq AttrH(H, F_{cur})$. Nun erhalten wir durch die FD $KH \rightarrow L$ dass $L \in AttrH(H, F_{cur})$, welches die zu überprüfende Eigenschaft darstellt.

Für die FD $KH \rightarrow M$ ergibt die analoge Argumentation (nur am Ende unter Verwendung der FD $KH \rightarrow M$ an Stelle von $KH \rightarrow L$) dass ebenfalls das K auf der linken Seite entfernt werden kann.

Die beiden letzten FDs $MN \rightarrow I$ und $MN \rightarrow L$ können nicht weiter reduziert werden, da die Hülle sowohl von M als auch von N jeweils nur M bzw. N enthält.

Somit ist die Menge F_r der funktionalen Abhängigkeiten nach der Linksreduktion gegeben als

$\{KLN \rightarrow H, KLN \rightarrow I, I \rightarrow K, I \rightarrow L, H \rightarrow H, H \rightarrow K, H \rightarrow I, L \rightarrow L, M \rightarrow M, H \rightarrow L, H \rightarrow M, MN \rightarrow I, MN \rightarrow L\}$

- Rechtsreduktion:

Bei der Rechtsreduktion muss man für jede funktionale Abhängigkeit überprüfen, ob das Entfernen der FD zu einer äquivalenten Menge an FDs führt, oder nicht. D.h. man muss überprüfen, ob die jeweilige FD aus den verbliebenen FDs folgt (abgeleitet werden kann) oder nicht.

Für $KLN \rightarrow H$ ist dies nicht der Fall, da dies die einzige (nicht triviale) FD ist welche H auf der rechten Seite stehen hat. Die FD $KLN \rightarrow I$ ist jedoch überflüssig, da $I \in AttrH(KLN, F_c)$ (sei $F_c = F_r \setminus \{KLN \rightarrow I\}$) gilt: Auf Grund der FD $KLN \rightarrow H$ gilt $H \in AttrH(KLN, F_c)$, woraus sich auf Grund von $H \rightarrow I$ die gewünschte Eigenschaft $I \in AttrH(KLN, F_c)$ ergibt. Wir erhalten $F_r = F_c$.

Die FD $I \rightarrow L$ ist nicht redundant, jedoch die triviale FD $H \rightarrow H$, welche entfernt wird.

Die FD $H \rightarrow K$ ist ebenfalls redundant: Wegen der FDs $H \rightarrow I$ und $I \rightarrow K$ erhalten wir $K \in AttrH(H, F_c)$ (für $F_c = F_r \setminus \{H \rightarrow K\}$). Die FD wird daher entfernt, und $F_r = F_c$.

Die FD $H \rightarrow I$ ist nicht redundant, jedoch die trivialen FDs $L \rightarrow L$ und $M \rightarrow M$, welche ebenfalls entfernt werden.

Die FD $H \rightarrow L$ ist redundant: Für $F_c = F_r \setminus \{H \rightarrow L\}$ erhalten wir auf Grund von $H \rightarrow I$ und $I \rightarrow L$ dass $L \in AttrH(H, F_c)$ gilt. Daher wird $H \rightarrow L$ entfernt und $F_r = F_c$.

Die FD $H \rightarrow M$ ist nicht redundant (die einzige FD mit M auf der rechten Seite), ebenso ist $MN \rightarrow I$ nicht redundant da ohne dieser FD die Hülle von MN nur MNL (aber nicht I) enthält.

Die FD $MN \rightarrow L$ ist schlussendlich wieder redundant, und zwar auf Grund der FDs $MN \rightarrow I$ und $I \rightarrow L$.

Somit ist die Menge der funktionalen Abhängigkeiten nach der Rechtsreduktion gegeben als

$$\{KLN \rightarrow H, I \rightarrow K, I \rightarrow L, H \rightarrow I, H \rightarrow M, MN \rightarrow I\}$$

- **Zusammenfassen:**

Als letzten Schritt müssen nur noch die Funktionalen Abhängigkeiten mit identischen linken Seiten zusammengefasst werden. Wir erhalten:

$$\{KLN \rightarrow H, I \rightarrow KL, H \rightarrow IM, MN \rightarrow I\}$$

(b) Wir berechnen wiederum die vier Schritte um eine kanonische Überdeckung zu erhalten (Die Lösung ist nicht eindeutig).

- **Dekomposition:**

Durch Anwendung der Dekompositionsregel erhalten wir die Menge

$$\{JL \rightarrow H, JL \rightarrow M, L \rightarrow I, HMI \rightarrow K, HMI \rightarrow N, JL \rightarrow K, JL \rightarrow N, H \rightarrow L, HLM \rightarrow K, HM \rightarrow J, J \rightarrow N\}$$

- **Linksreduktion:**

Wir müssen uns jetzt sukzessive die Frage stellen, ob für eine funktionale Abhängigkeit ein Attribut auf der linken Seite “überflüssig” ist, d.h. weggelassen werden kann so dass sich die resultierende Menge an FDs äquivalent zur ursprünglichen Menge an FDs ist.

Die erste FD welche auf der linken Seite redundante Attribute enthält ist $HMI \rightarrow K$, und zwar kann I entfernt werden. Die Berechnung der Attributhülle von HM unter den FDs ergibt folgendes: Auf Grund der FD $HM \rightarrow J$ ist J Teil der Hülle, und auf Grund von $H \rightarrow L$ ist L in der Hülle enthalten. Durch JL und die FD $JL \rightarrow K$ ergibt sich nun, dass K in der Hülle von HM enthalten ist, und daher ist I redundant und kann entfernt werden.

Das selbe Argument (nur mit $JL \rightarrow N$ anstatt $JL \rightarrow K$ als letzter FD) ergibt auch dass $HMI \rightarrow N$ zu $HM \rightarrow N$ reduziert werden kann.

Auf Grund der FD $J \rightarrow N$ kann auch die FD $JL \rightarrow N$ zu $J \rightarrow N$ reduziert werden. Ebenso kann $HLM \rightarrow K$ wegen $HM \rightarrow K$ zu $HM \rightarrow K$ reduziert werden. (Da wir mit einer Menge an funktionalen Abhängigkeiten arbeiten können keine Duplikate auftreten, d.h. die FDs $HLM \rightarrow K$ und $JL \rightarrow N$ werden effektiv gelöscht.)

Nachdem keine der verbleibenden FDs weiter reduziert werden kann ist die Menge der funktionalen Abhängigkeiten nach der Linksreduktion gegeben als

$$\{JL \rightarrow H, JL \rightarrow M, L \rightarrow I, HM \rightarrow K, HM \rightarrow N, JL \rightarrow K, J \rightarrow N, H \rightarrow L, HM \rightarrow J\}$$

- **Rechtsreduktion:**

Bei der Rechtsreduktion muss man für jede funktionale Abhängigkeit überprüfen, ob das Entfernen der FD zu einer äquivalenten Menge an FDs führt, oder nicht. D.h. man muss überprüfen, ob die jeweilige FD aus den verbliebenen FDs folgt (abgeleitet werden kann) oder nicht.

$JL \rightarrow H$ ist die einzige FD mit H auf der rechten Seite, und muss daher erhalten bleiben. Genauso $JL \rightarrow M$ und $L \rightarrow I$ für M bzw. I auf der rechten Seite.

Die FD $HM \rightarrow K$ hingegen ist redundant: Mittels $HM \rightarrow J$ (HMJ), $H \rightarrow L$ ($HMJL$) und $JL \rightarrow K$ ($HMJLK$) erhalten wir, dass K auch ohne der FD $HM \rightarrow K$ Teil der Hülle von HM ist.

Die FD $HM \rightarrow N$ ist redundant wegen $HM \rightarrow J$ und $J \rightarrow N$.

Die verbleibende Menge an FDs kann nicht weiter reduziert werden (die rechte Seite jeder FD ist eindeutig).

Somit ist die Menge der funktionalen Abhängigkeiten nach der Rechtsreduktion gegeben als

$$\{JL \rightarrow H, JL \rightarrow M, L \rightarrow I, JL \rightarrow K, J \rightarrow N, H \rightarrow L, HM \rightarrow J\}$$

• **Zusammenfassen:**

Als letzten Schritt müssen nur noch die Funktionalen Abhängigkeiten mit identischen linken Seiten zusammengefasst werden:

$$\{JL \rightarrow KHM, L \rightarrow I, J \rightarrow N, H \rightarrow L, HM \rightarrow J\}$$

Eine mögliche alternative Lösung ist

$$\{JL \rightarrow HM, L \rightarrow I, J \rightarrow N, H \rightarrow L, HM \rightarrow JK\}$$

Aufgabe 5 (Schlüsselbestimmung)

[2 Punkte]

Bestimmen Sie für die folgenden Relationenschemata samt Funktionalen Abhängigkeiten alle Schlüssel und alle Superschlüssel (*Hinweis:* Für Aufgabe a) listen Sie bitte die Menge der Superschlüssel auf. Für Aufgabe b) müssen Sie **nicht sämtliche Superschlüssel niederschreiben**; es genügt wenn die Menge eindeutig hervorgeht.)

(a) $\mathcal{R} = FGHIJ$

$$F = \{FG \rightarrow GI, FJ \rightarrow I, F \rightarrow G, FI \rightarrow J, H \rightarrow J, I \rightarrow F\}$$

Lösung:

Die Schlüssel sind FH und HI . Die Menge der Superschlüssel ist

$$\{FH, HI, FGH, FHI, FHJ, GHI, HIJ, FGHI, FHIJ, FGHJ, GHIJ, FGHIJ\}$$

(b) $\mathcal{R} = EFGHIJK$

$$F = \{GI \rightarrow EF, E \rightarrow K, FJ \rightarrow GE, FE \rightarrow I, IK \rightarrow KJH\}$$

Lösung:

Die Menge der Schlüssel ist

$$\{GI, FJ, EF, FIK\}.$$

Die Menge der Superschlüssel ist

$\{EF, EFG, EFGH, EFGHI, EFGHIJ, EFGHIJK, EFGHIK, EFGHJ, EFGHJK, EFGHK, EFGI, EFGIJ, EFGIJK, EFGIK, EFGJ, EFGJK, EFGK, EFH, EFHI, EFHIJ, EFHIJK, EFHIK, EFHJ, EFHJK, EFHK, EFI, EFIJ, EFIJK, EFIK, EFJ, EFJK, EFK, EGHI, EGHIJ, EGHIJK, EGHIK, EGI, EGIJ, EGIJK, EGIK, FGHI, FGHIJ, FGHIJK, FGHIK, FGHJ, FGHJK, FGI, FGIJ, FGIJK, FGIK, FGJ, FGJK, FHIJ, FHIJK, FHIK, FHJ, FHJK, FIJ, FIJK, FIK, FJ, FJK, GHI, GHIJ, GHIJK, GHIK, GI, GIJ, GIJK, GIK\}$

Aufgabe 6 (Normalformen)

[1 Punkt]

Gegeben ist jeweils ein Relationenschema \mathcal{R} samt einer Menge \mathcal{F} an dazugehörigen Funktionalen Abhängigkeiten. TODO: Schlüssel mit angeben?

Überprüfen Sie ob \mathcal{R}

- in dritter Normalform ist,
- in Boyce-Codd-Normalform ist,

und begründen Sie Ihre Antworten.

Lösung:

Die drei Eigenschaften, mit welcher für eine Menge funktionaler Abhängigkeiten der Form $\alpha \rightarrow B$ entschieden werden kann, ob sich das Relationenschema \mathcal{R} in dritter Normalform bzw. Boyce-Codd Normalform befindet, sind wie folgt:

1. $B \in \alpha$, d.h. die FD ist trivial;
2. α ist Superschlüssel von \mathcal{R} ;
3. das Attribut B ist in einem Schlüssel von \mathcal{R} enthalten.

Dabei ist ein Relationenschema \mathcal{R} in Boyce-Codd Normalform, wenn für jede Funktionale Abhängigkeit Bedingungen 1 oder 2 erfüllt ist, und \mathcal{R} ist in dritter Normalform wenn für jede Funktionale Abhängigkeit eine der Bedingungen 1, 2 oder 3 erfüllt ist.

- (a) $\mathcal{R} = ABCD$,
 $\mathcal{F} = \{AB \rightarrow C, C \rightarrow A, B \rightarrow DBA, AC \rightarrow BD\}$

Lösung:

Wir bestimmen zuerst die Schlüssel von \mathcal{R} . Diese sind: B und C .

Um die Bedingungen zu untersuchen wenden wir zuerst die Dekomposition an, und kennzeichnen anschließend für jede FD, welche Bedingungen sie erfüllt:

$$\mathcal{F} = \left\{ \underbrace{AB \rightarrow C}_{2,3}, \underbrace{C \rightarrow A}_2, \underbrace{B \rightarrow D}_2, \underbrace{B \rightarrow B}_{1,2,3}, \underbrace{B \rightarrow A}_2, \right. \\ \left. \underbrace{AC \rightarrow B}_{2,3}, \underbrace{AC \rightarrow D}_2 \right\}.$$

Nachdem jede FD entweder Eigenschaft 1 oder 2 erfüllt, befindet sich \mathcal{R} in BCNF (und daher auch in 3NF).

(b) $\mathcal{R} = ABCDEF$

$$F = \{ACD \rightarrow BC, BDF \rightarrow CE, EF \rightarrow ABD, AEB \rightarrow CD, ABC \rightarrow BF, \\ AD \rightarrow AC, ACF \rightarrow CF\}$$

Die Schlüssel sind ebenfalls gegeben als EF , BDF , ABE und DA **Lösung:**

Um die Bedingungen zu untersuchen wenden wir die Dekomposition an, und kennzeichnen für jede FD, welche Bedingungen sie erfüllt:

$$\mathcal{F} = \left\{ \underbrace{ACD \rightarrow B}_{2,3}, \underbrace{ACD \rightarrow C}_{1,2}, \underbrace{BDF \rightarrow C}_2, \underbrace{BDF \rightarrow E}_{2,3}, \underbrace{EF \rightarrow A}_{2,3}, \right. \\ \left. \underbrace{EF \rightarrow B}_{2,3}, \underbrace{EF \rightarrow D}_{2,3}, \underbrace{ABE \rightarrow C}_2, \underbrace{ABE \rightarrow D}_{2,3}, \underbrace{ABC \rightarrow B}_{1,3}, \right. \\ \left. \underbrace{ABC \rightarrow F}_3, \underbrace{AD \rightarrow A}_{1,2}, \underbrace{AD \rightarrow C}_2, \underbrace{ACF \rightarrow C}_1, \underbrace{ACF \rightarrow F}_{1,3} \right\}.$$

Nachdem jede FD eine der drei Eigenschaften erfüllt, eine FD jedoch nur Eigenschaft (3) befindet sich das Schema im 3NF, jedoch nicht in BCNF.

Aufgabe 7 (Synthesealgorithmus)

[2 Punkte]

Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCDEFGH$$

$$F = \{BCE \rightarrow AG, AB \rightarrow FH, BG \rightarrow C, A \rightarrow CH, AF \rightarrow B, D \rightarrow EF\}$$

Gesucht ist eine verlustlose und abhängigkeitserhaltende Zerlegung in dritter Normalform. Wenden Sie hierzu den Synthesealgorithmus an und dokumentieren Sie das Ergebnis der einzelnen Schritte. Bestimmen Sie alle Schlüssel von \mathcal{R} und allen Relationen der Zerlegung.**Lösung:**

1. Kanonischen Überdeckung bestimmen:

$$\mathcal{F}_c = \{BCE \rightarrow AG, AB \rightarrow F, BG \rightarrow C, A \rightarrow CH, AF \rightarrow B, D \rightarrow EF\}$$

2. Alle Kandidatenschlüssel von \mathcal{R} bzgl. \mathcal{F}_c bestimmen:

$$\{BDG, AD, BCD\}.$$

3. Relationenschemata für jedes Element von \mathcal{F}_c erstellen:

Relationenschema	Geltende FDs
$\mathcal{R}_1 = ABCGE$	$\mathcal{F}_1 = \{BCE \rightarrow AG, BG \rightarrow C, A \rightarrow C\}$
$\mathcal{R}_2 = ABF$	$\mathcal{F}_2 = \{AB \rightarrow F, AF \rightarrow B\}$
$\mathcal{R}_3 = BCG$	$\mathcal{F}_3 = \{BG \rightarrow C\}$
$\mathcal{R}_4 = ACH$	$\mathcal{F}_4 = \{A \rightarrow CH\}$
$\mathcal{R}_5 = ABF$	$\mathcal{F}_5 = \{AF \rightarrow B, AB \rightarrow F\}$
$\mathcal{R}_6 = DEF$	$\mathcal{F}_6 = \{D \rightarrow EF\}$

4. Schema eliminieren welche in anderen Schemata enthalten sind:

Das Schema \mathcal{R}_3 ist im Schema \mathcal{R}_1 enthalten und muss daher eliminiert werden. Die Schemata \mathcal{R}_2 und \mathcal{R}_5 sind identisch, daher wird \mathcal{R}_5 ebenfalls eliminiert.

Relationenschema	Geltende FDs
$\mathcal{R}_1 = ABCGE$	$\mathcal{F}_1 = \{BCE \rightarrow AG, BG \rightarrow C, A \rightarrow C\}$
$\mathcal{R}_2 = ABF$	$\mathcal{F}_2 = \{AB \rightarrow F, AF \rightarrow B\}$
$\mathcal{R}_4 = ACH$	$\mathcal{F}_4 = \{A \rightarrow CH\}$
$\mathcal{R}_6 = DEF$	$\mathcal{F}_6 = \{D \rightarrow EF\}$

5. Testen ob Kandidatenschlüssel in einem der Teilschema enthalten ist:

Keiner der Schlüssel ist in einem Teilschema enthalten. Es muss ein neues Relationenschema $\mathcal{R}_{\mathcal{K}}$ erstellt werden, wir wählen den Schlüssel AD :

Relationenschema	Geltende FDs
$\mathcal{R}_{\mathcal{K}} = AD$	$\mathcal{F}_{\mathcal{K}} = \emptyset$

6. Ergebnis (je ein Schlüssel ist unterstrichen):

Relationenschema	Geltende FDs
$\mathcal{R}_1 = \underline{A}BCGE$	$\mathcal{F}_1 = \{BCE \rightarrow AG, BG \rightarrow C, A \rightarrow C\}$
$\mathcal{R}_2 = \underline{A}BF$	$\mathcal{F}_2 = \{AB \rightarrow F, AF \rightarrow B\}$
$\mathcal{R}_4 = \underline{A}CH$	$\mathcal{F}_4 = \{A \rightarrow CH\}$
$\mathcal{R}_6 = \underline{D}EF$	$\mathcal{F}_6 = \{D \rightarrow EF\}$
$\mathcal{R}_{\mathcal{K}} = \underline{A}\underline{D}$	$\mathcal{F}_{\mathcal{K}} = \emptyset$

Aufgabe 8 (Dekompositionsalgorithmus)

[2 Punkte]

Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten und allen Schlüsseln:

$$\mathcal{R} = ABCDEF$$

$$\mathcal{F} = \{F \rightarrow C, D \rightarrow C, BDE \rightarrow AC, ABD \rightarrow DEF, BD \rightarrow E, EF \rightarrow AEB, AD \rightarrow CEF\}$$

Schlüssel: DEF, BD, AD

Gesucht ist eine verlustlose Zerlegung in Boyce-Codd-Normalform. Wenden Sie hierzu den Dekompositionsalgorithmus an und dokumentieren Sie das Ergebnis der einzelnen Schritte. Bestimmen Sie alle Schlüssel von allen Relationen der Zerlegung. Ist die Zerlegung abhängigkeiterhaltend? Wenn die Zerlegung nicht abhängigkeiterhaltend ist, geben Sie an welche der Abhängigkeiten in \mathcal{F} verloren gegangen ist.

Hinweis: Bestimmen Sie bei jeder Zerlegung die jeweilige Hülle an FDs!

Lösung:

Die drei FDs

- $F \rightarrow C$,

- $D \rightarrow C$, und
- $EF \rightarrow AEB$

verletzen die BCNF (sind nicht trivial und die linke Seite beinhaltet keinen Schlüssel), die restlichen FDs erfüllen die BCNF. Es gibt daher drei Möglichkeiten für die Zerlegung. Wir zeigen hier alle drei Möglichkeiten auf.

Hinweis: Wir geben in jeder Hülle $\mathcal{F}_i^+[\mathcal{R}_j]$ nur die nicht-trivialen, linksreduzierten Abhängigkeiten an.

- Die FD $F \rightarrow C$ wird gewählt. Damit erhalten wir das Schema:

$$\begin{array}{ll} \mathcal{R}_1 = CF & \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{F \rightarrow C\} \\ & \text{Schlüssel: } F \\ \mathcal{R}_2 = ABDEF & \mathcal{F}_2 = \mathcal{F}^+[\mathcal{R}_2] = \{AD \rightarrow EF, EF \rightarrow AB, BD \rightarrow A\} \\ & \text{Schlüssel: } AD, BD, DEF \end{array}$$

Das Schema \mathcal{R}_1 ist nun in BCNF, nicht jedoch das Schema \mathcal{R}_2 , da die FD $EF \rightarrow AB$ wiederum die Bedingungen der BCNF verletzt. Daher muss \mathcal{R}_2 weiter zerlegt werden. Dabei wird nach der einzigen FD zerlegt, welche die BCNF verletzt.

- Wir zerlegen nach $EF \rightarrow AB$ zu

$$\begin{array}{lll} \mathcal{R}_1 = CF & \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{F \rightarrow C\} & \text{Schlüssel: } F \\ \mathcal{R}_{2,1} = ABEF & \mathcal{F}_{2,1} = \mathcal{F}_2^+[\mathcal{R}_{2,1}] = \{EF \rightarrow AB\} & \text{Schlüssel: } EF \\ \mathcal{R}_{2,2} = DEF & \mathcal{F}_{2,2} = \mathcal{F}_2^+[\mathcal{R}_{2,2}] = \{\} & \text{Schlüssel: } DEF \end{array}$$

Nun erfüllen alle Relationenschemata die BCNF, es ist daher kein weiteres zerlegen notwendig.

Diese Zerlegung ist nicht abhängigkeiterhaltend. Von den gegebenen FDs gehen $D \rightarrow C$, $BDE \rightarrow AC$, $ABD \rightarrow DEF$, $BD \rightarrow E$ und $AD \rightarrow CEF$ verloren.

- Die FD $D \rightarrow C$ wird gewählt. Damit erhalten wir das Schema:

$$\begin{array}{ll} \mathcal{R}_1 = CD & \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{D \rightarrow C\} \\ & \text{Schlüssel: } D \\ \mathcal{R}_2 = ABDEF & \mathcal{F}_2 = \mathcal{F}^+[\mathcal{R}_2] = \{AD \rightarrow EF, EF \rightarrow AB, BD \rightarrow A\} \\ & \text{Schlüssel: } AD, BD, DEF \end{array}$$

\mathcal{R}_2 ist nun ident zu \mathcal{R}_2 zuvor. Wir zerlegen \mathcal{R}_2 daher wie zuvor in $\mathcal{R}_{2,1}$ und $\mathcal{R}_{2,2}$ und erhalten das Schema

$$\begin{array}{lll} \mathcal{R}_1 = CD & \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{D \rightarrow C\} & \text{Schlüssel: } D \\ \mathcal{R}_{2,1} = ABEF & \mathcal{F}_{2,1} = \mathcal{F}_2^+[\mathcal{R}_{2,1}] = \{EF \rightarrow AB\} & \text{Schlüssel: } EF \\ \mathcal{R}_{2,2} = DEF & \mathcal{F}_{2,2} = \mathcal{F}_2^+[\mathcal{R}_{2,2}] = \{\} & \text{Schlüssel: } DEF \end{array}$$

in BCNF.

Die Zerlegung ist nicht abhängigkeiterhaltend. Von den gegebenen FDs gehen $F \rightarrow C$, $BDE \rightarrow AC$, $ABD \rightarrow EF$, $BD \rightarrow E$ und $AD \rightarrow CEF$ verloren.

- Die FD $EF \rightarrow AEB$ wird gewählt. Damit erhalten wir das Schema:

$$\begin{array}{ll} \mathcal{R}_1 = ABEF & \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{EF \rightarrow AB\} \\ & \text{Schlüssel: } EF \\ \mathcal{R}_2 = CDEF & \mathcal{F}_2 = \mathcal{F}^+[\mathcal{R}_2] = \{F \rightarrow C, D \rightarrow C\} \\ & \text{Schlüssel: } DEF \end{array}$$

Das Schema \mathcal{R}_1 ist nun in BCNF, nicht jedoch das Schema \mathcal{R}_2 , da beide FDs – $F \rightarrow C$ und $D \rightarrow C$ – die Bedingungen der BCNF nicht erfüllen. Das Schema muss also weiter zerlegt werden, wobei zwei Möglichkeiten zur Auswahl stehen.

- Wir zerlegen nach $F \rightarrow C$ zu

$$\begin{array}{lll} \mathcal{R}_1 = ABEF & \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{EF \rightarrow AB\} & \text{Schlüssel: } EF \\ \mathcal{R}_{2,1} = CF & \mathcal{F}_{2,1} = \mathcal{F}_2^+[\mathcal{R}_{2,1}] = \{F \rightarrow C\} & \text{Schlüssel: } F \\ \mathcal{R}_{2,2} = DEF & \mathcal{F}_{2,2} = \mathcal{F}_2^+[\mathcal{R}_{2,2}] = \{\} & \text{Schlüssel: } DEF \end{array}$$

Nun erfüllen alle Relationenschemata die BCNF, es ist daher kein weiteres zerlegen notwendig.

Diese Zerlegung ist nicht abhängigkeiterhaltend. Von den gegebenen FDs gehen $D \rightarrow C$, $BDE \rightarrow AC$, $ABD \rightarrow EF$, $BD \rightarrow E$ und $AD \rightarrow CEF$ verloren.

- Wir zerlegen nach $D \rightarrow C$ zu

$$\begin{array}{lll} \mathcal{R}_1 = ABEF & \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{EF \rightarrow AB\} & \text{Schlüssel: } EF \\ \mathcal{R}_{2,1} = CD & \mathcal{F}_{2,1} = \mathcal{F}_2^+[\mathcal{R}_{2,1}] = \{D \rightarrow C\} & \text{Schlüssel: } D \\ \mathcal{R}_{2,2} = DEF & \mathcal{F}_{2,2} = \mathcal{F}_2^+[\mathcal{R}_{2,2}] = \{\} & \text{Schlüssel: } DEF \end{array}$$

Nun erfüllen alle Relationenschemata die BCNF, es ist daher kein weiteres zerlegen notwendig.

Diese Zerlegung ist nicht abhängigkeiterhaltend. Von den gegebenen FDs gehen $F \rightarrow C$, $BDE \rightarrow AC$, $ABD \rightarrow EF$, $BD \rightarrow E$ und $AD \rightarrow CEF$ verloren.