2. Übungsblatt (WS 2017) – Musterlösung

3.0 VU Datenmodellierung / 6.0 VU Datenbanksysteme

Informationen zum Übungsblatt

Allgemeines

In diesem Übungsteil sollen Sie Aufgabenstellungen aus den Bereichen SQL und Normalformentheorie 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 gescannten handschriftlichen PDF-Dateien.

Das Übungsblatt enthält 7 Aufgaben, auf welche Sie insgesamt 15 Punkte erhalten können.

Deadlines

bis 10.11.	**12:00 Uhr**	Upload der Abgabe über TUWEL
ab 27.11.	13:00 Uhr	Korrektur und Feedback in TUWEL verfügbar

Tutorensprechstunden (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 auseinander setzen, 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) *[5.0 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.

Normalformentheorie

Aufgabe 2 (Funktionale Abhängigkeiten) *[1 Punkte]*

(a) Geben ist ein Relationenschema

`Segelcrew(Name, Vegetarisch, Tätigkeit, besondereMerkmale, Segelerfahrung, Segelschein)`

mit der folgenden Ausprägung:

Segelcrew					
Name	Vegetarisch	Tätigkeit	besondereMerkmale	Segelerfahrung	Segelschein
Franz	ja	fischt	rote Schuhe	nein	nein
Chrisi	nein	fischt	grüne Jacke	viel	ja
Karo	nein	wäscht ab	gelbe Jacke	nein	nein
Resi	nein	kocht	rosa Pulli	wenig	nein
Jaki	nein	navigiert	gelbe Mütze	viel	ja
Lanu	nein	liest	rote Jacke	viel	ja
Caro	ja	kocht	Augenklappe	nein	nein
Steffi	nein	navigiert	gelbe Mütze	viel	ja
Ben	nein	fotographiert	Kamera	nein	nein

Überprüfen Sie für jede der untenstehenden Aussagen, ob sie in der angegebenen Ausprägung stimmt. Geben Sie für jede Aussage die Antwort (ja/nein) an. Falls eine Aussage nicht zutrifft geben Sie außerdem ein entsprechendes Gegenbeispiel an.

a) Es navigieren nur Leute mit gelber Mütze. **Ja**

b) Es gilt die Funktionale Abhängigkeit:
Segelerfahrung $\rightarrow$ Segelschein. **Ja**

c) Vegetarier fischen nicht. **Nein**

Gegenbeispiel: *Franz* ist Vegetarier, fischt aber.

d) Es gilt die Funktionale Abhängigkeit:
besondereMerkmale $\rightarrow$ Segelerfahrung. **Ja**

e) Es gilt die Funktionale Abhängigkeit:
Segelerfahrung→ besondereMerkmale. **Nein**

Gegenbeispiel: Ein mögliches Gegenbeispiel sind die erste und die dritte Zeile der Ausprägung. Beide haben den selben Wert *nein* für das Attribut `Segelerfahrung`, jedoch unterschiedliche Werte für das Attribut `besondereMerkmale`.

(b) Gegeben ein Relationenschema $\mathcal{R} = ABCDE$ und die Mengen $\mathcal{F}_1$ und $\mathcal{F}_2$ von Funktionalen Abhängigkeiten:

$$\mathcal{F}_1 = \{C \to DE, B \to C, D \to AC, E \to C\}$$
$$\mathcal{F}_2 = \{C \to EA, E \to D, B \to C, D \to C\}$$

Sind $\mathcal{F}_1$ und $\mathcal{F}_2$ äquivalent? Begründen Sie Ihre Antwort!

Lösung:
Ja, $\mathcal{F}_1$ und $\mathcal{F}_2$ sind äquivalent. Dafür genügt es zu zeigen, dass $\mathcal{F}_1$ in der Hülle von $\mathcal{F}_2$ enthalten ist, und umgekehrt.

	FD aus $\mathcal{F}_1$	$\in \mathcal{F}_2^+$?	folgt aus Transitivität und folgenden FDs	
	$C \to D$	ja	$C \to E, E \to D$	
	$C \to E$	ja	$C \to E$	
$\mathcal{F}_1 \subseteq \mathcal{F}_2^+$?	$B \to C$	ja	$B \to C$	✓
	$D \to A$	ja	$D \to C, C \to A$	
	$D \to C$	ja	$D \to C$	
	$E \to C$	ja	$E \to D, D \to C$	

	FD aus $\mathcal{F}_2$	$\in \mathcal{F}_1^+$?	folgt aus Transitivität und folgenden FDs	
	$C \to E$	ja	$C \to E$	
	$C \to A$	ja	$C \to D, D \to A$	
$\mathcal{F}_2 \subseteq \mathcal{F}_1^+$?	$E \to D$	ja	$E \to C, C \to D$	✓
	$B \to C$	ja	$B \to C$	
	$D \to C$	ja	$D \to C$	

Hinweis: Für eine FD $\alpha \to \beta$ und eine Menge $\mathcal{F}_d$ von funktionalien Abhängigkeiten können Sie $\alpha \to \beta \in \mathcal{F}_d^+$ auch einfach mit Hilfe der Attributhülle bestimmen:
$\alpha \to \beta \in \mathcal{F}_d^+$ genau dann wenn $\beta \subseteq AttrH(\alpha, \mathcal{F})$.

Aufgabe 3 (Kanonische Überdeckung) *[2 Punkte]*

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

(a) $\mathcal{F}_1 = \{C \to A, CD \to ABFG, A \to B, E \to BCDE, B \to B, D \to AC, ABCG \to E\}$

(b) $\mathcal{F}_2 = \{DE \to C,\ C \to B,\ B \to F,\ ABE \to BA,\ A \to C,\ A \to F,\ CDE \to F\}$

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
  $\{C \to A, CD \to A, CD \to B, CD \to F, CD \to G, A \to B, E \to B, E \to C, E \to D, E \to E, B \to B, D \to A, D \to C, ABCG \to E\}$

- Linksreduktion:

  Wir müssen uns jetzt sukzessive die Frage stellen, ob für eine funktionale Abhängigkeit ein Attribut auf der linken Seite überflüssig ist. Bei $CD \to A$ kann D gestrichen werden, da bereits $C \to A$ gilt, und bei $CD \to B$, $CD \to F$ und $CD \to G$ kann auch jeweils C gestrichen werden, weil bereits $D \to C$ gilt. $E \to E$ und $B \to B$ sind trivial und können getrichen werden. Bei $ABCG \to E$ können A und B gestrichen werden; A weil bereits $C \to A$ gilt (und daher E in der Attributhülle von BCG unter den verbliebenen FDs liegt wegen $ABCG \to E$), B weil aus $C \to A$ und $A \to B$ folgt dass $C \to B$ gilt (und daher E in der Attributhülle von CG unter den verbliebenen FDs liegt wegen $BCG \to E$).

  Somit ist die Menge der funktionalen Abhängigkeiten nach der Linksreduktion gegeben als
  $\{C \to A, D \to B, D \to F, D \to G, A \to B, E \to B, E \to C, E \to D, D \to A, D \to C, CG \to E\}$.

- Rechtsreduktion:

  Bei der Rechtsreduktion muss man für jede funktionale Abhängigkeit überprüfen, ob sie noch gültig ist, selbst wenn man sie streicht.

  Für $D \to B$ ist das hier der Fall: Aus D folgt C, aus C folgt A und aus A folgt B, daraus kann man tranisitv $D \to B$ ableiten. Das gleiche gilt für $E \to B$, denn das lässt sich aus $E \to C$, $C \to A$ und $A \to B$ ableiten. $E \to C$ lässt sich ebenfalls streichen, weil es aus $E \to D$ und $D \to C$ folgt. Schlussendlich kann noch $D \to A$ entfernt werden, da auf Grund der FDs $D \to C$ und $C \to A$ gilt dass A in der Attributhülle von D unter den verbleibenden FDs liegt. Somit ist die Menge der funktionalen Abhängigkeiten nach der Rechtsreduktion gegeben als
  $\{C \to A, D \to F, D \to G, A \to B, E \to D, D \to C, CG \to E\}$.

- Zusammenfassen:

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

  $\{C \to A, D \to CFG, A \to B, E \to D, CG \to E\}$.

(b)
- Dekomposition:
  $\{DE \to C,\ C \to B,\ B \to F,\ ABE \to B, ABE \to A,\ A \to C,\ A \to F,\ CDE \to F\}$

- Linksreduktion:

  Die FD $ABE \to B$ kann zu $B \to B$ reduziert werden. Da die FD trivial ist, kann sie gleich gestrichen werden. Das gleiche gilt für $ABE \to A$. In $CDE \to F$ kann C gestrichen werden, weil $DE \to C$ gilt (und daher, wegen der FD $CDE \to F$, gilt dass F in der Attributhülle von DE unter den verbliebenden FDs enthalten ist). Nach Streichung von doppelten FDs erhält man nach der Linksreduktion somit

$$\{DE \to C,\ C \to B,\ B \to F,\ A \to C,\ A \to F, DE \to F\}$$

- Rechtsreduktion

  $DE \to F$ kann gestrichen werden, weil $DE \to F$ aus $DE \to C$, $C \to B$ und $B \to F$ abgeleitet werden kann.
  $A \to F$ kann gestrichen werden, weil $A \to F$ aus $A \to C$, $C \to B$ und $B \to F$ abgeleitet werden kann.

  Dann erhält man: $\{DE \to C,\ C \to B,\ B \to F,\ A \to C\}$

- Zusammenfassen:

  $\{DE \to C,\ C \to B,\ B \to F,\ A \to C\}$

Aufgabe 4 (Schlüsselbestimmung) *[2 Punkte]*

Bestimmen Sie für die folgenden Relationenschemata samt Funktionalen Abhängigkeiten alle Schlüssel und alle Superschlüssel.

(a) $\mathcal{R} = (Nachname, Vorname, GebDat, PLZ, Ort)$

$\mathcal{F} = \{\{GebDat, Vorname\} \to \{Nachname, PLZ\},$
$\{PLZ\} \to \{Ort\}, \{Nachname, GebDat\} \to \{Vorname\}\}$

Lösung:
Die Schlüssel sind $\{GebDat, Vorname\}$ und $\{GebDat, Nachname\}$.
Die Menge der Superschlüssel: { {Nachname, Gebdat},{Vorname, Gebdat}, {Nachname, Vorname, Gebdat}, {Nachname, Gebdat, PLZ}, {Nachname, Gebdat, Ort},{Nachname, Gebdat, PLZ, Ort}, {Vorname, Gebdat, PLZ}, {Vorname, Gebdat, Ort}, {Nachname, Vorname, Gebdat, PLZ}, {Nachname, Vorname, Gebdat, Ort}, {Vorname, Gebdat, PLZ, Ort}, {Nachname, Vorname, Gebdat, PLZ, Ort} }

(b) $\mathcal{R} = ABCDEF$
$\mathcal{F} = \{BEF \to CD, CDE \to AF, AE \to B, FD \to A\}$

Lösung:
Die Schlüssel sind AEF,BEF, DEF und CDE.
Die Menge der Superschlüssel:

{ AEF, BEF, CDE, DEF, ABEF, ACDE, ACEF, ADEF, BCDE, BCEF, BDEF, CDEF, ABCDE, ABCEF, ABDEF, ACDEF, BCDEF, ABCDEF }

Aufgabe 5 (Normalformen) *[1 Punkte]*

Gegeben ist jeweils ein Relationenschema $\mathcal{R}$ samt einer Menge $\mathcal{F}$ an dazugehörigen Funktionalen Abhängigkeiten.
Überprüfen Sie ob $\mathcal{R}$

- in dritter Normalform ist,

- in Boyce-Codd-Normalform ist,

und begründen Sie Ihre Antworten.

(a) $\mathcal{R} = \{$ Nachname, Vorname, GebDat, PLZ, Ort $\}$
$\mathcal{F} = \{\{GebDat, Vorname\} \rightarrow \{Nachname, PLZ\},$
$\{PLZ\} \rightarrow \{Ort\}, \{Nachname, GebDat\} \rightarrow \{Vorname\}\}$

Lösung:
$\mathcal{R}$ is weder in dritter noch in Boyce-Codd-Normalform.

Begründung: Die Schlüssel sind $\{GebDat, Vorname\}$ und $\{GebDat, Nachname\}$. Die drei Anforderungen für Funktionale Abhängigkeiten der Form $\alpha \rightarrow B$, die man benötigt um ein Relationenschema $\mathcal{R}$ auf dritte Normalform bzw. Boyce-Codd Normalform zu untersuchen, 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 Bedingungen 1,2 oder 3 erfüllt ist. Um dies zu untersuchen wenden wir die Dekomposition an, und kennzeichnen für jede FD, welche Bedingungen sie erfüllt:

$$\mathcal{F} = \{ \underbrace{\{\text{GebDat, Vorname}\} \rightarrow \{\text{Nachname}\}}_{2,3}, \underbrace{\{\text{GebDat, Vorname}\} \rightarrow \{\text{PLZ}\}}_{2}, \underbrace{\{\text{PLZ}\} \rightarrow \{\text{Ort}\}}_{-}, \underbrace{\{\text{Nachname, GebDat}\} \rightarrow \{\text{Vorname}\}}_{2,3} \}$$

Nachdem eine FD keine der drei Bedingungen erfüllt ist das Schema nicht in dritter Normalform.

- Falls die Normalformen nicht erfüllt sind erklären Sie, zu welchen Problem das im konkreten Fall führen kann.

  Lösung:
  Die Information welcher Ort zu einer PLZ gehört wird möglicherweise redundant gespeichert: wenn eine PLZ öfters als einmal in einer Ausprägung vorkommt, wird der dazugehörige, eindeutige, Ort ebenfalls jedes mal mitgespeichert. Das führt dazu, dass man bei vielen Einfüge/Änderungsoperationen darauf aufpassen muss, die FD $PLZ \rightarrow Ort$ nicht zu verletzen. (Bedenken Sie in dem Zusammenhang nicht nur unterschiedliche Ortsangaben, sondern z.B. unterschiedliche Schreibweisen des selben Ortes).

- Durch welche generellen Modifikationen könnte man diese Probleme beheben, und welche konkreten Algorithmen würden Sie verwenden, um diese Modifikationen durchzuführen?

Lösung:
Wenn man das Relationenschema auf mehrere Schemata aufsplittet, kann man vermeiden Informationen mehrfach zu speichern. Mit dem Synthesealgorithmus kann man eine Form erreichen, die in dritter Normalform ist. Mit den Dekompositionsalgorithmus erreicht man eine Form in BCNF (und damit auch in 3NF).

- Falls $\mathcal{R}$ diese Eigenschaft nicht schon erfüllt, bringen Sie $\mathcal{R}$ in eine Form, die sowohl in der dritten Normalform, als auch in Boyce-Codd-Normalform ist. Sie müssen keinen konkreten Algorithmus verwenden, sondern können selbst probieren, wie eine sinnvolle Form Ihrer Meinung nach aussehen sollte. Begründen Sie Ihre Antwort.

Lösung:
Problematisch ist die Speicherung der Information welcher Ort zu einer PLZ gehört an mehreren Stellen. Man kann dies einfach beheben in dem man statt der einen Relation zwei Relationen mit folgenden Schemata erstellt:

$\mathcal{R}_1 = \{Nachname, Vorname, GebDat, PLZ\}$ und
$\mathcal{R}_2 = \{PLZ, Ort\}$.

Jeder Relation kann man dann die jeweiligen Funktionalen Abhängigkeiten zuordnen und leicht überprüfen, ob die dritte Normalform und BCNF erfüllt ist.

Auf der Relation $\mathcal{R}_1$ sind die Funktionalen Abhängigkeiten der Menge $\mathcal{F}$ gültig. Die Schlüssel der Relation sind $\{GebDat, Vorname\}$ und $\{GebDat, Nachname\}$ und es ist wieder gekennzeichnet, welche Bedingungen erfüllt werden.

$$\mathcal{F} = \{ \underbrace{\{\text{GebDat, Vorname}\} \rightarrow \{\text{Nachname}\}}_{2,3}, \underbrace{\{\text{GebDat, Vorname}\} \rightarrow \{\text{PLZ}\}}_{2}, \underbrace{\{\text{Nachname, GebDat}\} \rightarrow \{\text{Vorname}\}}_{2,3}\}$$

$\mathcal{R}_1$ befindet sich also sowohl in BCNF, als auch in dritter Normalform.

Auf der Relation $\mathcal{R}_2$ ist nur die Funktionale Abhängigkeit PLZ $\rightarrow$ Ort gültig. Diese erfüllt die zweite Bedingung, da PLZ Schlüssel von $\mathcal{R}_2$ ist.

Daraus folgt, dass sich auch $\mathcal{R}_2$ sowohl in BCNF, als auch in dritter Normalform befindet.

(b) $\mathcal{R} = ABCDEFGH$
$\mathcal{F} = \{A \rightarrow DFH, DH \rightarrow E, FH \rightarrow CGA, EB \rightarrow AF, AD \rightarrow BF\}$

(Sie brauchen die zusätzlichen Fragen aus Punkt (a) hier nicht zu beantworten.)

Lösung:
Die Schlüssel sind A, FH, EB, DHB. Die drei Anforderungen für Funktionale Abhängigkeiten der Form $\alpha \rightarrow B$, die man benötigt um ein Relationenschema $\mathcal{R}$ auf dritte Normalform bzw. Boyce-Codd Normalform zu untersuchen, 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 Bedingungen 1,2 oder 3 erfüllt ist. Um dies zu untersuchen wenden wir die Dekomposition an, und kennzeichnen für jede FD, welche Bedingungen sie erfüllt:

$$\begin{aligned}\mathcal{F} = \{ & \underbrace{A \to D}_{2,3}, \underbrace{A \to F}_{2,3}, \underbrace{A \to H}_{2,3}, \\ & \underbrace{DH \to E}_{3}, \underbrace{FH \to CGA}_{2}, \\ & \underbrace{EB \to AF}_{2}, \underbrace{AD \to B}_{2,3}, \underbrace{AD \to F}_{2,3}\}.\end{aligned}$$

Für jede FD trifft entweder Bedingung 1,2 oder 3 zu. Außerdem gibt es eine FD ($DH \to E$), für die Bedingung 1 und 2 nicht zutreffen. Deswegen ist $\mathcal{R}$ in dritter Normalform, aber nicht in BCNF.

Aufgabe 6 (Synthesealgorithmus) *[2 Punkte]*

Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCDEF$$
$$\mathcal{F} = \{BC \to A, A \to D, BE \to DF, EF \to C, E \to F, A \to B, B \to D\}$$

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. Bestimmung der kanonischen Überdeckung:

$$\mathcal{F}_c = \{BC \to A, E \to CF, A \to B, B \to D\}$$

2. Erstelle Relationenschemata für jedes Element von $\mathcal{F}_c$:

Relationenschema	Geltende FDs
$\mathcal{R}_1 = ABC$	$\mathcal{F}_1 = \{BC \to A, A \to B\}$
$\mathcal{R}_2 = CEF$	$\mathcal{F}_2 = \{E \to CF\}$
$\mathcal{R}_3 = AB$	$\mathcal{F}_3 = \{A \to B\}$
$\mathcal{R}_4 = BD$	$\mathcal{F}_4 = \{B \to D\}$

3. Bestimmung aller Kandidatenschlüssel von $\mathcal{R}$ bzgl. $\mathcal{F}_c$: AE, BE.

4. Keiner der Schlüssel ist in einem Teilschema anthalten. Es muss ein neues Relationenschema $\mathcal{R}_{\mathcal{K}}$ erstellt werden:

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

5. Das Schema $\mathcal{R}_3$ ist im Schema $\mathcal{R}_1$ enthalten und muss daher eliminiert werden.

Ergebnis (Schlüssel sind unterstrichen):

Relationenschema	Geltende FDs
$\mathcal{R}_1 = A\underline{BC}$	$\mathcal{F}_1 = \{BC \to A, A \to B\}$
$\mathcal{R}_2 = C\underline{E}F$	$\mathcal{F}_2 = \{E \to CF\}$
$\mathcal{R}_4 = \underline{B}D$	$\mathcal{F}_4 = \{B \to D\}$
$\mathcal{R}_{\mathcal{K}} = \underline{AE}$	$\mathcal{F}_{\mathcal{K}} = \emptyset$

Aufgabe 7 (Dekompositionsalgorithmus) *[2 Punkte]*

Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCDEF$$
$$\mathcal{F} = \{AB \to BCD, D \to EF, C \to AD\}$$

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 $\mathcal{R}$ und allen Relationen der Zerlegung. Ist die Zerlegung abhängigkeitserhaltend? Wenn die Zerlegung nicht abhängigkeitserhaltend ist, geben Sie an, welche Abhängigkeiten verloren gegangen sind. *Hinweis*: Bestimmen Sie bei jeder Zerlegung die jeweilige Hülle an FDs!

Lösung:

Die Schlüssel von $\mathcal{R}$ sind AB und BC. Doe erste FD erfüllt die BCNF, aber die anderen beiden FDs nicht. Da sie nicht trivial sind, und keine der Attributmengen D oder C ein Superschlüssel ist. Es gibt zwei Möglichkeiten für eine Zerlegung:

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

- Die FD $D \to EF$ wird gewählt. Damit erhalten wir:

  $\mathcal{R}_1 = DEF$ $\quad \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{D \to EF\}$ $\quad$ Schlüssel: D

  $\mathcal{R}_2 = ABCD$ $\quad \mathcal{F}_2 = \mathcal{F}^+[\mathcal{R}_2] = \{AB \to CD, C \to AD\}$ $\quad$ Schlüssel: AB, BC

  R_2 ist nicht in BCNF, muss also weiter zerlegt werden zu:

  $\mathcal{R}_1 = DEF$ $\quad \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{D \to EF\}$ $\quad$ Schlüssel: D

  $\mathcal{R}_{2,1} = ACD$ $\quad \mathcal{F}_{2,1} = \mathcal{F}_2^+[\mathcal{R}_{2,1}] = \{C \to AD\}$ $\quad$ Schlüssel: C

  $\mathcal{R}_{2,2} = BC$ $\quad \mathcal{F}_{2,2} = \mathcal{F}_2^+[\mathcal{R}_{2,2}] = \emptyset$ $\quad$ Schlüssel: BC

  Nun erfüllen alle Relationenschemata die BCNF, wir müssen daher nicht weiter zerlegen.

  Diese Zerlegung ist nicht abhängigkeitserhaltend, da z.B. $AB \to CDEF$ und $BC \to EF$ verloren gehen.

- Die FD $C \to AD$ wird gewählt. Daraus ergibt sich

  $\mathcal{R}_1 = ACD$ $\quad \mathcal{F}_1 = \mathcal{F}^+[\mathcal{R}_1] = \{C \to AD\}$ $\quad$ Schlüssel: C

  $\mathcal{R}_2 = BCEF$ $\quad \mathcal{F}_2 = \mathcal{F}^+[\mathcal{R}_2] = \{BC \to EF\}$ $\quad$ Schlüssel: BC

Beide Relationenschemata erfüllen die BCNF, wir müssen daher nicht weiter zerlegen.

Diese Zerlegung ist nicht abhängigkeitserhaltend, da z.B. $D \rightarrow EF$ und $AB \rightarrow CDEF$ verloren gehen.