

# 2. Übungsblatt

## 3.0 VU Datenmodellierung

9. April 2014

### Allgemeines

In diesem Übungsteil sollten Sie Aufgabenstellungen aus den Bereich SQL und Normalformentheorie bearbeiten.

Lösen Sie die Beispiele eigenständig, denn in der Praxis (und bei der Prüfung) sind Sie auch auf sich alleine gestellt. Wir weisen Sie darauf hin, dass abgeschriebene Lösungen mit 0 Punkten beurteilt werden.

Geben Sie ein einziges PDF Dokument ab. Erstellen Sie Ihr Abgabedokument computerunterstützt. Wir akzeptieren keine gescannten handschriftlichen PDF-Dateien.

### Deadlines

<b>16.5.</b>	<b>06:55 Uhr</b>	Deadline für den Upload über den COURSEMANAGER
<i>24.5.</i>	<i>08:00 Uhr</i>	Feedback im COURSEMANAGER verfügbar
<i>25.5.</i>	<i>23:59 Uhr</i>	Reservierung eines Termins für das Tutorengespräch
<i>25.5.</i>	<i>23:59 Uhr</i>	Reservierung eines Termins für das Abgabegespräch

### Tutorengespräch

1. Sie müssen sich über den COURSEMANAGER zu einem Tutorengespräch anmelden. Bitte machen Sie das rechtzeitig, je später Sie sich anmelden, umso eingeschränkter ist das Terminangebot.
2. Wenn Sie kein Blatt abgegeben haben, sind Sie nicht zum Tutorengespräch zugelassen.
3. Sie kommen mit Ihrem Studierendenausweis zu der von Ihnen reservierten Zeit vorbei, und absolvieren das Tutorengespräch. Sie erhalten dabei wertvolles Feedback zu Ihrem Übungsblatt. Das Gespräch ist verpflichtend und dauert ca. 30 Minuten.

## Abgabegespräch

1. Sie müssen sich über den COURSEMANAGER zu einem Abgabegespräch anmelden. Bitte machen Sie das rechtzeitig, je später Sie sich anmelden, umso eingeschränkter ist das Terminangebot.
2. Wenn Sie kein Blatt abgegeben haben oder kein Tutorengespräch geführt haben, sind Sie nicht zum Abgabegespräch zugelassen.
3. Sie kommen mit Ihrem Studierendenausweis zu der von Ihnen reservierten Zeit vorbei, und absolvieren das Abgabegespräch. Stoffgebiet des Abgabegesprächs sind die mit dem Übungsblatt abgedeckten Themengebiete. Wir setzen voraus, dass Sie sich mit Ihrer korrigierten Abgabe auseinandergesetzt haben. Das Abgabegespräch hat Prüfungscharakter.
4. Sie absolvieren Ihr Abgabegespräch gemeinsam mit anderen KollegInnen. Das Gespräch dauert ca. 60 Minuten.
5. Sie können auf die Abgabe maximal 15 Punkte erreichen. Diese setzen sich wie folgt zusammen:  
**5 Punkte** auf das Übungsblatt  
**10 Punkte** auf das Abgabegespräch
6. Die Assistenten tragen die Punkte des Abgabegesprächs in den COURSEMANAGER ein und Sie sehen dort, wieviele Punkte Sie bekommen haben.
7. Falls Sie nicht zu Ihrem Gesprächstermin erscheinen, bekommen Sie automatisch 0 Punkte auf das Gespräch.

## SQL

### Aufgabe 1 (eSQL) [1.0 Punkte]

Lösen Sie alle 10 (inkl. Unterpunkte) unter

<http://minteka.dbai.tuwien.ac.at/eSQL-tutorial/>

zur Verfügung gestellten SQL-Aufgaben des aktuellen Übungskurses. Loggen Sie sich dabei mit dem Usernamen und dem Passwort ein, das Sie bereits vom COURSEMANAGER kennen. Der Abschlusstest der Übung wird über dieselbe Plattform abgewickelt. Es empfiehlt sich daher zusätzlich auch mit Beispielen aus den vorigen Semestern zu üben.

### Aufgabe 2 (Kaskadierendes Löschen) [0.5 Punkte]

Gegeben sei eine Datenbank mit den Relationen  $A$ ,  $B$  und  $C$ , die wie folgt erstellt wurden:

```
CREATE TABLE A (  
    aid INTEGER PRIMARY KEY,  
    bid INTEGER REFERENCES B(bid) ON DELETE CASCADE
```

);

```
CREATE TABLE B (  
    bid INTEGER PRIMARY KEY,  
    attr INTEGER
```

);

```
CREATE TABLE C (  
    cid INTEGER PRIMARY KEY,  
    bid INTEGER REFERENCES B(bid) ON DELETE SET NULL
```

);

Die Ausprägungen der Relationen seien:

$$A = \{(1, 3), (2, 3), (3, 5), (4, 2), (5, 4)\}$$

$$B = \{(1, 1), (2, 2), (3, 3), (4, 1), (5, 1)\}$$

$$C = \{(1, 1), (2, 1), (3, 4), (4, 4), (5, 4)\}$$

Geben Sie die Ausprägungen an, die sich durch das Ausführen der Statements

(a) `DELETE FROM B WHERE bid < 3` und

(b) `DELETE FROM B WHERE attr = 1`

ergeben, jeweils auf die ursprüngliche Ausprägung angewendet.

**Lösung:**

(a)  $A = \{(1, 3), (2, 3), (3, 5), (5, 4)\}$

$$B = \{(3, 3), (4, 1), (5, 1)\}$$

$$C = \{(1, \text{null}), (2, \text{null}), (3, 4), (4, 4), (5, 4)\}$$

(b)  $A = \{(1, 3), (2, 3), (4, 2)\}$

$$B = \{(2, 2), (3, 3)\}$$

$$C = \{(1, \text{null}), (2, \text{null}), (3, \text{null}), (4, \text{null}), (5, \text{null})\}$$

**Aufgabe 3 (Allquantifizierung) [0.5 Punkte]**

In einem all-quantifizierten Ausdruck wird zum Beispiel nach jenen Fahrern gesucht, die *alle* verfügbaren Fahrzeuge lenken können/dürfen. Diesen Sachverhalt kann man in klassischer Prädikatenlogik erster Stufe wie folgt ausdrücken:  $\varphi(k) \equiv \text{Fahrer}(f) \wedge \forall z (\text{Fahrzeug}(z) \rightarrow \text{lenkt}(f, z))$ . Erklären Sie die beiden Methoden, durch die ein Allquantor in SQL ausgedrückt werden kann, zunächst allgemein und dann auch an dem oben angeführten einfachen Beispiel. Geben Sie hierbei die entsprechenden Abfragen in SQL-Syntax an.

Fahrer(fkey)

Fahrzeug(zkey)

lenkt(Fahrer.fkey, Fahrzeug.zkey)

**Lösung:**

Prinzipiell kann ein All-Quantor in SQL mit Hilfe von COUNT oder durch NOT EXISTS, NOT IN ausgedrückt werden.

Bei dem Ansatz mit COUNT werden zunächst alle Tupel in der betrachteten Domäne gezählt. Dann werden die Tupel in der betrachteten Domäne gezählt, die überdies die gewünschte Bedingung erfüllen. Stimmen die beiden Ergebnisse überein, erfüllen alle betrachteten Tupel die Bedingung. Eine Lösung wäre daher:

```
SELECT f.fkey
FROM Fahrer f JOIN lenkt l ON f.fkey = l.fkey
GROUP BY f.fkey
HAVING COUNT(*) = (SELECT COUNT(*) FROM Fahrzeug);
```

Der andere Ansatz nutzt die Äquivalenz  $\forall x \varphi(x) \equiv \neg \exists x \neg \varphi(x)$ . Wir wollen also all jene Fahrer finden, für die es kein Fahrzeug gibt, das sie nicht lenken können.

```
SELECT f.fkey
FROM Fahrer f
WHERE NOT EXISTS (SELECT * FROM Fahrzeug z
                  WHERE z.zkey NOT IN (SELECT l.zkey
                                       FROM lenkt l
                                       WHERE l.fkey=f.fkey))
```

**Normalformentheorie****Aufgabe 4 (Armstrong Axiome) [0.5 Punkte]**

- (a) Gegeben ist ein Relationenschema  $\mathcal{R} = ABCDE$  und zwei Mengen  $F_1$  und  $F_2$  von funktionalen Abhängigkeiten.

$$F_1 = \{A \rightarrow BC, C \rightarrow D, BD \rightarrow E, A \rightarrow D\}$$

$$F_2 = \{A \rightarrow BC, C \rightarrow D, AD \rightarrow E, BD \rightarrow E\}$$

Sind  $F_1$  und  $F_2$  äquivalent? Begründen Sie Ihre Antwort formal und dokumentieren Sie den Lösungsweg.

- (b) Zeigen Sie, dass aus den Axiomen Reflexivität, Verstärkung und Pseudotransitivität alle anderen Armstrong Axiome hergeleitet werden können. (Hinweis: Es genügt zu zeigen, dass alle grundlegenden Axiome hergeleitet werden können.)
- (c) Zeigen Sie, dass das Axiom der Reflexivität nicht aus folgenden Axiomen hergeleitet werden kann: Verstärkung, Transitivität, Vereinigung, Dekomposition und Pseudotransitivität. (Hinweis: Welche FDs lassen sich aus der leeren Menge an FDs herleiten?)

**Lösung:**

- (a) Ja,  $F_1$  und  $F_2$  sind äquivalent, da  $F_1^+ = F_2^+$ . Wir zeigen dies indem wir zeigen, dass  $F_1^+ \subseteq F_2^+$  und  $F_2^+ \subseteq F_1^+$  gilt.

Für  $F_1^+ \subseteq F_2^+$  reicht es zu zeigen, dass  $F_1 \subseteq F_2^+$  gilt, also dass die Hülle von  $F_2$  alles enthält, was auch in  $F_1$  ist.

$F_1 \setminus F_2 = \{A \rightarrow D\}$ , es bleibt also zu zeigen, dass  $A \rightarrow D \in F_2^+$  gilt. Wir leiten  $A \rightarrow D$  aus  $F_2$  wie folgt her:

$$\frac{\begin{array}{c} \text{[Gegeben]} \\ A \rightarrow BC \\ \hline A \rightarrow C \end{array} \text{ Dekomp.} \quad \begin{array}{c} \text{[Gegeben]} \\ C \rightarrow D \end{array}}{A \rightarrow D} \text{ Trans.}$$

Daher gilt, dass  $F_1 \subseteq F_2^+$  wahr ist und aus der Definition der Hülle folgt, dass auch  $F_1^+ \subseteq F_2^+$  gilt.

Für  $F_2^+ \subseteq F_1^+$  reicht es zu zeigen, dass  $F_2 \subseteq F_1^+$  gilt, also dass die Hülle von  $F_1$  alles enthält, was auch in  $F_2$  ist.  $F_2 \setminus F_1 = \{AD \rightarrow E\}$ , es bleibt also zu zeigen, dass  $AD \rightarrow E \in F_1^+$  gilt. Wir leiten  $AD \rightarrow E$  aus  $F_1$  wie folgt her:

$$\frac{\begin{array}{c} \text{[Gegeben]} \\ A \rightarrow BC \\ \hline A \rightarrow B \end{array} \text{ Dekomp.} \quad \begin{array}{c} \text{[Gegeben]} \\ BD \rightarrow E \end{array}}{AD \rightarrow E} \text{ Pseudotrans.}$$

Daher gilt, dass  $F_2 \subseteq F_1^+$  ist und aus der Definition der Hülle folgt dann, dass  $F_2^+ \subseteq F_1^+$  ebenso gilt.

- (b) Es reicht, alle grundlegenden Armstrong Axiome herzuleiten, also Reflexivität, Verstärkung und Transitivität. Es fehlt daher nur noch das Axiom der Transitivität.

$$\frac{\begin{array}{c} \text{[Gegeben]} \\ \alpha \rightarrow \beta \end{array} \quad \begin{array}{c} \text{[Gegeben]} \\ \beta\delta \rightarrow \gamma \end{array}}{\alpha\delta \rightarrow \gamma} \text{ Pseudotrans.}$$

Sei nun  $\delta = \emptyset$ , dann gilt  $\beta\delta = \beta$  und  $\alpha\delta = \alpha$ . Obige Herleitung ist also gleich zu: Gegeben  $\alpha \rightarrow \beta$  und  $\beta \rightarrow \gamma$ , dann gilt  $\alpha \rightarrow \gamma$  (Axiom der Transitivität).

- (c) Sei  $\mathcal{R} = AB$  ein Relationenschema und  $F = \{\}$  die leere Menge an FDs. Mittels der Reflexivität, lässt sich unter anderem  $A \rightarrow A$  aus  $F$  herleiten.

Betrachten wir nun die restlichen Axiome, so fällt auf, dass jedes davon bereits existierende FDs voraussetzt um anwendbar zu sein. Aus  $F$  können (ohne Reflexivität) also keine weiteren FDs hergeleitet werden. Somit können die Axiome der Verstärkung, Transitivität, Vereinigung, Dekomposition und Pseudotransitivität die FD  $A \rightarrow A$  nicht aus  $F$  herleiten.

Daraus folgt, dass sich das Axiom der Reflexivität auch nicht aus diesen Axiomen herleiten lassen kann.

**Aufgabe 5 (Kanonische Überdeckung)** [0.5 Punkte]

Bestimmen Sie eine kanonische Überdeckung folgender Mengen funktionaler Abhängigkeiten über dem Relationenschema  $ABCDEFG$  und dokumentieren Sie den Lösungsweg.

$$(a) F^1 = \{ACD \rightarrow E, EF \rightarrow E, AD \rightarrow CG, DE \rightarrow F, B \rightarrow BC, AD \rightarrow F, BDC \rightarrow G\}$$

$$(b) F^2 = \{B \rightarrow BFG, ABG \rightarrow ADE, BG \rightarrow BD, B \rightarrow C, C \rightarrow GFC\}$$

**Lösung:**

$$(a) F_C^1 = \{AD \rightarrow CEG, B \rightarrow C, BD \rightarrow G, DE \rightarrow F\}$$

$$(b) F_C^2 = \{AB \rightarrow E, B \rightarrow CD, C \rightarrow FG\}$$

**Aufgabe 6 (Schlüsselbestimmung)** [0.5 Punkte]

(a) Bestimmen Sie für folgendes Relationenschema samt funktionalen Abhängigkeiten alle Schlüssel und alle Superschlüssel.

$$\mathcal{R} = ABCDE$$

$$F = \{A \rightarrow BC, C \rightarrow D, A \rightarrow D\}$$

**Lösung:**

Der einzige Schlüssel ist  $AE$ . Die Menge der Superschlüssel:

$$\{AE, ABE, ACE, ADE, ABCE, ABDE, ACDE, ABCDE\}$$

(b) Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCDEFG$$

$$F = \{A \rightarrow BC, C \rightarrow D, AC \rightarrow E, B \rightarrow A, E \rightarrow BC\}$$

Erklären Sie, warum  $C$  und  $ABFG$  keine Schlüssel sind. Berechnen Sie weiters alle Schlüssel.

**Lösung:**

$C$  ist kein Schlüssel, da  $\text{AttrHülle}(F, \{C\}) = \{CD\} \neq \mathcal{R}$ . Da nun aber  $AFG$  ein Schlüssel ist, ist  $ABFG$  nicht minimal und deshalb kein Schlüssel. Die Menge aller Schlüssel ist  $\{AFG, BFG, EFG\}$ .

**Aufgabe 7 (Synthesealgorithmus)** [0.5 Punkte]

Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCDEF$$

$$F = \{AB \rightarrow BE, AB \rightarrow D, EF \rightarrow C, E \rightarrow DE, C \rightarrow F\}$$

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

**Lösung:**

- Bestimmung der kanonischen Überdeckung:

$$F_c = \{AB \rightarrow E, EF \rightarrow C, E \rightarrow D, C \rightarrow F\}$$

- Erstelle Relationenschemata für jedes Element von  $F_c$ :

Relationenschema	Geltende FDs
$\mathcal{R}_1 = ABE$	$F_1 = \{AB \rightarrow E\}$
$\mathcal{R}_2 = CEF$	$F_2 = \{EF \rightarrow C, C \rightarrow F\}$
$\mathcal{R}_3 = DE$	$F_3 = \{E \rightarrow D\}$
$\mathcal{R}_4 = CF$	$F_4 = \{C \rightarrow F\}$

- Bestimmung aller Kandidatenschlüssel von  $\mathcal{R}$  bzgl.  $F_c$ :  $\{ABC, ABF\}$ .

Keiner der Schlüssel ( $ABC$  und  $ABF$ ) ist in einem der erzeugten Teilschemata enthalten. Deshalb erzeugen wir das Schema  $\mathcal{R}_\kappa = ABF$  mit  $F_\kappa = \emptyset$ .

- Das Schema  $\mathcal{R}_4$  kann eliminiert werden, da  $\mathcal{R}_4 \subseteq \mathcal{R}_2$  gilt.

Ergebnis (Schlüssel sind unterstrichen):

Relationenschema	Geltende FDs
$\mathcal{R}_1 = \underline{ABE}$	$F_1 = \{AB \rightarrow E\}$
$\mathcal{R}_2 = \underline{EFC}$	$F_2 = \{EF \rightarrow C, C \rightarrow F\}$
$\mathcal{R}_3 = \underline{ED}$	$F_3 = \{E \rightarrow D\}$
$\mathcal{R}_\kappa = \underline{ABF}$	$F_\kappa = \emptyset$

**Aufgabe 8 (Normalformen)** [0.5 Punkte]

Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCDE$$

$$F = \{CD \rightarrow AE, AB \rightarrow BDE, CE \rightarrow BD\}$$

Geben Sie an, ob  $\mathcal{R}$

- in dritter Normalform und/oder
- in Boyce-Codd-Normalform

ist. Begründen Sie Ihre Antworten.

**Lösung:**

$CD$ ,  $CE$  und  $ABC$  sind die (einzigen) Schlüssel von  $\mathcal{R}$ . Für die FDs  $CD \rightarrow AE$  und  $CE \rightarrow BD$  gilt, dass  $CD$  bzw.  $CE$  Superschlüssel von  $\mathcal{R}$  sind.  $AB \rightarrow BDE$  kann mittels Dekompositionsregel in die folgenden drei FDs zerlegt werden: (1) Die FD  $AB \rightarrow B$  ist trivial. (2) Die rechte Seite der FD  $AB \rightarrow D$  ist in Schlüssel  $CD$  enthalten. (3) Die rechte Seite der FD  $AB \rightarrow E$  ist in Schlüssel  $CE$  enthalten. Daher ist  $\mathcal{R}$  in 3NF.  $\mathcal{R}$  ist nicht in BCNF, da  $AB \rightarrow D$  weder trivial, noch die linke Seite  $AB$  ein Superschlüssel ist.

**Aufgabe 9 (Dekompositionsalgorithmus) [0.5 Punkte]**

Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$R = ADEGH$$

$$F = \{A \rightarrow DE, E \rightarrow G, A \rightarrow G\}$$

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 abhängigkeiterhaltend? Wenn die Zerlegung nicht abhängigkeiterhaltend ist, geben Sie an, welche Abhängigkeiten verloren gegangen sind.

**Lösung:**

Einziger Schlüssel von  $\mathcal{R}$  ist  $AH$ . Alle FDs verletzen die BCNF, da sie nicht trivial sind und deren linke Seite kein Superschlüssel ist. Nun wählen wir eine der FDs aus, die die BCNF verletzt. Wir entscheiden uns hier für  $A \rightarrow DE$ . Zerlege  $\mathcal{R}$  daher in:

$\mathcal{R}_1 = ADE$	$F_1 = \{A \rightarrow DE\}$	Schlüssel: $A$
$\mathcal{R}_2 = AGH$	$F_2 = \{A \rightarrow G\}$	Schlüssel: $AH$

$\mathcal{R}_1$  ist nun in BCNF, aber  $\mathcal{R}_2$  noch nicht. Daher wird  $\mathcal{R}_2$  weiter zerlegt, indem wir die einzige geltende FD auswählen:

$\mathcal{R}_{21} = AG$	$F_{21} = \{A \rightarrow G\}$	Schlüssel: $A$
$\mathcal{R}_{22} = AH$	$F_{22} = \emptyset$	Schlüssel: $AH$

Die Relationen der Zerlegung sind somit:  $\mathcal{R}_1, \mathcal{R}_{21}$  und  $\mathcal{R}_{22}$ . Da die FD  $E \rightarrow G$  verloren geht (das heisst  $(E \rightarrow G) \notin (F_1 \cup F_{21} \cup F_{22})^+$ ), ist die Zerlegung nicht abhängigkeiterhaltend.

Alternative Lösung wenn zuerst die FD  $E \rightarrow G$  ausgewählt wird:

$\mathcal{R}_1 = EG$	$F_1 = \{E \rightarrow G\}$	Schlüssel: $E$
$\mathcal{R}_{21} = ADE$	$F_{21} = \{A \rightarrow DE\}$	Schlüssel: $A$
$\mathcal{R}_{22} = AH$	$F_{22} = \emptyset$	Schlüssel: $AH$

Hier ist zu beachten, dass in diesem Fall die Zerlegung abhängigkeiterhaltend ist. Denn während  $E \rightarrow G \in F_1$  und  $A \rightarrow DE \in F_{21}$  gilt, ist auch  $A \rightarrow G$  in  $(F_1 \cup F_{21} \cup F_{22})^+$



enthalten. Der Grund hierfür ist, dass  $A \rightarrow G$  mittels Transitivität aus  $A \rightarrow E$  und  $E \rightarrow G$  hergeleitet werden kann.

Alternative Lösung wenn zuerst die FD  $A \rightarrow G$  ausgewählt wird:

$\mathcal{R}_1 = AG$	$F_1 = \{A \rightarrow G\}$	Schlüssel: $A$
$\mathcal{R}_{21} = ADE$	$F_{21} = \{A \rightarrow DE\}$	Schlüssel: $A$
$\mathcal{R}_{22} = AH$	$F_{22} = \emptyset$	Schlüssel: $AH$

Diese Zerlegung ist nicht abhängigkeitsstreu, da die FD  $E \rightarrow G$  verloren geht.