2. Übungsblatt

3.0 VU Datenmodellierung

15. Mai 2012

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

25.05. 06:55 Uhr	Upload über den CourseManager. Es werden keine verspäteten Abgaben akzeptiert.
02.06. 06:55 Uhr	Feedback im CourseManager verfügbar
03.06. 06:55 Uhr	Reservierung eines Termins für das Abgabegespräch

Abgabegespräch

1. Sie müssen sich über den CourseManager zu einem Abgabegespräch (=Prüfungsgespräch) anmelden. Bitte machen Sie das rechtzeitig, je später Sie sich anmelden, umso eingeschränkter ist das Terminangebot. Bis zur oben genannten Deadline garantieren wie Ihnen einen Termin.

2. Sie müssen mindestens einen Punkt auf das Übungsblatt bekommen. Wenn Sie weniger als einen Punkt auf das Blatt bekommen, oder kein Blatt abgegeben 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.

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. Um die Lehrveranstaltung positiv abzuschließen, brauchen Sie mindestens 1 Punkt auf das Übungsblatt und 1 Punkt auf das Abgabegespräch.

8. Falls Sie nicht zu Ihrem Gesprächstermin erscheinen, bekommen Sie automatisch 0 Punkte und damit ein negatives Zeugnis.

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 x, y und z, die wie folgt erstellt wurden:

```
CREATE TABLE x (
        xid INTEGER PRIMARY KEY
);

CREATE TABLE y (
        yid INTEGER PRIMARY KEY,
        xid INTEGER REFERENCES x (xid) ON DELETE CASCADE
);

CREATE TABLE z (
```

```
        zid INTEGER PRIMARY KEY,
        yid INTEGER REFERENCES y (yid) ON DELETE CASCADE
);
```

Die Ausprägungen der Relationen seien:

$$x = \{(1), (2), (3), (4), (5)\}$$
$$y = \{(1, 1), (2, 2), (3, 4), (4, 3)(5, 5)\}$$
$$z = \{(1, 1), (2, 3), (3, 4)\}$$

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

(a) `DELETE FROM` x `WHERE` xid $= 1$ und

(b) `DELETE FROM` x `WHERE` xid $= 3$ `OR` xid > 4

ergeben, jeweils auf die ursprüngliche Ausprägung angewendet.

Lösung:

(a) $x = \{(2), (3), (4), (5)\},\ y = \{(2, 2), (3, 4), (4, 3), (5, 5)\},\ z = \{(2, 3), (3, 4)\}$

(b) $x = \{(1), (2), (4)\},\ y = \{(1, 1), (2, 2), (3, 4)\},\ z = \{(1, 1), (2, 3)\}$

Aufgabe 3 (Allquantifizierung) *[0.5 Punkte]*
In einem all-quantifizierten Ausdruck wird z.B.: nach jenen Kunden gesucht, die bereits *alle* Filme ausgeliehen haben. Diesen Sachverhalt kann man in klassischer Prädikatenlogik erster Stufe wie folgt ausdrücken: $\varphi(k) \equiv \mathit{Kunde}(k) \wedge \forall f\ (\mathit{Film}(f) \rightarrow \mathit{leiht}(k, f))$ 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.

```
Kunde(kkey)
Film(pkey)
leiht(Kunde.kkey, Film.fkey)
```

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

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

```
SELECT k.kkey
FROM Kunde k JOIN leiht l ON l.kkey = k.kkey
GROUP BY k.kkey
HAVING COUNT(*) = (SELECT COUNT(*) FROM Film);
```

Der andere Ansatz nutzt die Äquivalenz $\forall x\ \varphi(x) \equiv \neg\exists x\ \neg\varphi(x)$. Wir wollen also all jene Kunden finden, für die es keinen Film gibt, den sie nicht ausgeborgt haben.

```
SELECT k.kkey
FROM kunde k
WHERE not exists (SELECT * FROM Film f
                  WHERE f.fkey NOT IN (SELECT l.fkey
                                       FROM leiht l
                                       WHERE l.kkey=k.kkey))
```

Normalformentheorie

Aufgabe 4 (Armstrong Axiome) *[0.5 Punkte]*
Gegeben ist ein Relationenschema ABCDEF und zwei Mengen F_1 und F_2 von funktionalen Abhängigkeiten.

$$F_1 = \{AC \to AE, E \to BF, B \to AC, CE \to B, AC \to D\}$$
$$F_2 = \{AC \to AE, E \to BF, B \to AC, CE \to B, B \to DF\}$$

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

Lösung:
Ja, da $F_1^+ = F_2^+$. Es reicht aus, folgende zwei Eigenschaften zu zeigen:

$F_1 \subseteq F_2^+$: Mithilfe der Armstrong-Axiome zeigen wir, dass $\mathrm{AC} \to \mathrm{D} \in F_2^+$. Die restlichen FDs folgen trivial, da $F_1 \setminus \{\ \mathrm{AC} \to \mathrm{D}\ \} \subseteq F_2 \subseteq F_2^+$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{[\text{Gegeben}]}{\mathrm{AC} \to \mathrm{AE}}}{\mathrm{AC} \to \mathrm{E}}\,\text{Dekomp.} \quad \overset{[\text{Gegeben}]}{\mathrm{E} \to \mathrm{BF}}}{\mathrm{AC} \to \mathrm{BF}}\,\text{Trans.}}{\mathrm{AC} \to \mathrm{B}}\,\text{Dekomp.} \quad \overset{[\text{Gegeben}]}{\mathrm{B} \to \mathrm{DF}}}{\dfrac{\mathrm{AC} \to \mathrm{DF}}{\mathrm{AC} \to \mathrm{D}}\,\text{Dekomp.}}\,\text{Trans.}$$

$F_2 \subseteq F_1^+$: Analog zu ersten Fall. Es gilt $F_2 \setminus \{\ \mathrm{B} \to \mathrm{DF}\ \} \subseteq F_1 \subseteq F_1^+$. Es bleibt somit zu zeigen: $\mathrm{B} \to \mathrm{DF} \in F_1^+$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\overset{[\text{Gegeben}]}{\mathrm{AC} \to \mathrm{AE}}}{\mathrm{AC} \to \mathrm{E}}\,\text{Dekomp.} \quad \overset{[\text{Gegeben}]}{\mathrm{E} \to \mathrm{BF}}}{\mathrm{AC} \to \mathrm{BF}}\,\text{Trans.} \quad \overset{[\text{Gegeben}]}{\mathrm{AC} \to \mathrm{D}}}{\mathrm{AC} \to \mathrm{BDF}}\,\text{Vereinig.} \quad \overset{[\text{Gegeben}]}{\mathrm{B} \to \mathrm{AC}}}{\dfrac{\mathrm{B} \to \mathrm{BDF}}{\mathrm{B} \to \mathrm{DF}}\,\text{Dekomp.}}\,\text{Trans.}$$

Aufgabe 5 (Kanonische Überdeckung) *[0.5 Punkte]*
Bestimmen Sie die kanonische Überdeckung folgender Mengen funktionaler Abhängigkeiten über dem Relationenschema $ABCDEF$:

(a) $F^1 = \{F \to ADF,\ A \to E,\ D \to EG,\ DE \to E,\ G \to E,\ BCF \to A,\ G \to A\}$

(b) $F^2 = \{C \to B,\ DE \to DEG,\ A \to G,\ F \to DE,\ A \to B,\ C \to G,\ B \to FB\}$

Lösung:

(a) $F^1_C = \{A \to E,\ D \to G,\ F \to D,\ G \to A\}$

(b) $F^2_C = \{A \to B, C \to B, B \to F, F \to DE, DE \to G\}$

Aufgabe 6 (Schlüsselbestimmung) *[0.5 Punkte]*

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

$$\mathcal{R} = ABCDE$$
$$F = \{D \to B,\ BE \to C,\ A \to C,\ B \to D\}$$

Lösung:
ABE und ADE sind die einzigen Schlüssel. Die Menge der Superschlüssel:

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

(b) Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCDE$$
$$F = \{DE \to A, A \to B, B \to AC, E \to D\}$$

Erklären Sie, warum DE kein Schlüssel ist.

Lösung:
DE ist zwar ein Superschlüssel, aber kein Schlüssel, da DE nicht minimal ist. E alleine hingegen ist ein Schlüssel.

Aufgabe 7 (Synthesealgorithmus) *[0.5 Punkte]*
Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCDEFG$$
$$F = \{A \to CE,\ ADF \to G,\ BDF \to G,\ D \to E,\ C \to B,\ C \to F\}$$

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:

$$F_c = \{A \to CE,\ BDF \to G,\ C \to BF,\ D \to E\}$$

2. Erstelle Relationenschemata für jedes Element von F_c:

Relationenschema	Geltende FDs
$\mathcal{R}_1 = ACE$	$F_1 = \{A \to CE\}$
$\mathcal{R}_2 = BDFG$	$F_2 = \{BDF \to G\}$
$\mathcal{R}_3 = CBF$	$F_3 = \{C \to BF\}$
$\mathcal{R}_4 = DE$	$F_4 = \{D \to E\}$

3. Bestimmung aller Kandidatenschlüssel von $\mathcal{R}$ bzgl. F_c: AD.

   AD ist in keinem der erzeugten Teilschemata enthalten. Definiere deshalb das Schema $\mathcal{R}_\kappa = AD$ mit $F_\kappa = \emptyset$.

4. Kein Relationenschema ist zu eliminieren.

Ergebnis (Schlüssel sind unterstrichen):

Relationenschema	Geltende FDs
$\mathcal{R}_1 = \underline{A}CE$	$F_1 = \{A \to CE\}$
$\mathcal{R}_2 = \underline{BDF}G$	$F_2 = \{BDF \to G\}$
$\mathcal{R}_3 = \underline{C}BF$	$F_3 = \{C \to BF\}$
$\mathcal{R}_4 = \underline{D}E$	$F_4 = \{D \to E\}$
$\mathcal{R}_\kappa = \underline{AD}$	$F_\kappa = \emptyset$

Aufgabe 8 (Normalformen) *[0.5 Punkte]*
Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$\mathcal{R} = ABCD$$
$$F = \{A \to BD,\ AD \to C,\ BC \to AD\}$$

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

(a) in dritter Normalform ist,

(b) in Boyce-Codd-Normalform ist,

und begründen Sie Ihre Antworten.

Lösung:
A und BC sind die einzigen Schlüssel von $\mathcal{R}$. Da die linken Seiten aller FDs Superschlüssel sind, ist $\mathcal{R}$ in BCNF und daher auch in 3NF.

Aufgabe 9 (Dekompositionsalgorithmus) *[0.5 Punkte]*
Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

$$R = ABCD$$
$$F = \{AB \to C,\ AC \to D\}$$

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.

Lösung:
Der einzige Schlüssel von $\mathcal{R}$ ist AB. Die FD $AC \to D$ verletzt die BCNF, da sie nicht trivial und AC kein Superschlüssel ist. Zerlege $\mathcal{R}$ daher in:

$\mathcal{R}_1 = ACD$	$F_1 = \{AC \to D\}$	Schlüssel: AC
$\mathcal{R}_2 = ABC$	$F_2 = \{AB \to C\}$	Schlüssel: AB

$\mathcal{R}_1$ und $\mathcal{R}_2$ sind beide in BCNF und die Zerlegung ist abhängigkeitserhaltend, da bereits $F = F_1 \cup F_2$ gilt.