2. Übungsblatt

3.0 VU Datenmodellierung

12. Mai 2015

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

5.6.	**06:55 Uhr**	Deadline für den Upload über den CourseManager
12.6.	08:00 Uhr	Feedback im CourseManager verfügbar
14.6.	23:59 Uhr	Reservierung eines Termins für das Tutorengespräch
14.6.	23:59 Uhr	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. Das Gespräch ist verpflichtend und dauert ca. 30 Minuten.

4. Nützen Sie das Tutorengespräch, um wertvolles Feedback zu Ihrem Übungsblatt zu erhalten: Gehen Sie vor dem Gespräch Ihre korrigierten Lösungen durch, identifizieren Sie Unklarheiten, und stellen Sie den Tutoren entsprechend **Fragen**. Trauen Sie sich zuzugeben, wenn Sie etwas nicht verstanden haben. Ihre Fragen haben keinerlei (negativen) Einfluss auf Ihre Note!

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 9 (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,
        a INTEGER
);

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

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

Die Ausprägungen der Relationen seien:

A = {(1,3), (2,6), (3,3), (4,4), (5,2)}

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

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

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

(a) DELTE `FROM` A `WHERE` a>3 und

(b) `DELETE FROM` A `WHERE` aid<=2 `OR` aid >=5

ergeben, jeweils auf die ursprüngliche Ausprägung angewendet.

Lösung:

(a) A = {(1,3), (3,3), (5,2))}
B = {(1,1), (2,3), (3,1), (5,5)}
C = {(1,3,5), (4,3,null), (5,1,null)}

(b) A = {(3,3), (4,4)}
B = {(2,3)}
C = {(1,3,null), (2,4,null), (3,4,null), (4,3,null))}

Aufgabe 3 (Allquantifizierung) *[0.5 Punkte]*
In einem all-quantifizierten Ausdruck wird zum Beispiel nach jenen ÜbersetzerInnen gesucht, die *alle* erforderlichen Sprachen beherrschen. Diesen Sachverhalt kann man in

klassischer Prädikatenlogik erster Stufe wie folgt ausdrücken: $\varphi(h) \equiv \mathit{Uebersetzerin}(u) \wedge \forall x\ (\mathit{Sprache}(s) \rightarrow \mathit{beherrscht}(u, s))$.

Erklären Sie wie der Allquantor entweder durch logische Äquivalenz oder durch Abzählen 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.

```
Uebsetzerin(ukey)
Sprache(skey)
beherrscht(Uebersetzerin.ukey, Sprache.skey)
```

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 u.ukey
FROM Uebersetzerin u JOIN beherrscht b on u.ukey=b.ukey
GROUP BY u.ukey
HAVING COUNT(*) = (SELECT COUNT(*) FROM Sprache);
```

Der andere Ansatz nutzt die Äquivalenz $\forall x\ \varphi(x) \equiv \neg\exists x\ \neg\varphi(x)$. Wir wollen also all jene ÜbersetzerInnen finden, für die es keine Sprache gibt, die sie nicht beherrschen.

```
SELECT u.ukey
FROM Uebersetzerin u
WHERE NOT EXISTS(SELECT * FROM Sprache s
        WHERE s.skey NOT IN (SELECT b.skey
                             FROM beherrscht b
                             WHERE b.ukey=u.ukey));
```

Normalformentheorie

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

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

$$F_1 = \{A \rightarrow B, A \rightarrow A\}$$
$$F_2 = \{AB \rightarrow B\}$$

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

(b) Zeigen Sie, dass das Axiom der Dekomposition aus den grundlegenden Armstrong Axiomen hergeleitet werden kann.

(c) Gegeben ist ein Relationenschema $\mathcal{R} = ABCDE$ und eine Menge F_3 von funktionalen Abhängigkeiten:

$$F_3 = \{A \to B, B \to C, CD \to E, C \to ED\}$$

Zeigen Sie, dass $A \to E \in F_3^+$ gilt.

Lösung:

(a) Nein, F_1 und F_2 sind nicht äquivalent, da $F_1^+ \neq F_2^+$. Wir geben F_1^+ und F_2^+ vollständig an um zu zeigen, dass $F_1^+ \neq F_2^+$ gilt.

$$F_1^+ = \{A \to A, B \to B, AB \to AB, AB \to A, AB \to B, A \to B, A \to AB\}$$

Die FDs $A \to A$, $B \to B$, $AB \to AB$, $AB \to A$ und $AB \to B$ folgen direkt aus dem Axiom der Reflexivität. Die FD $A \to B$ kommt von F_1 und $A \to AB$ folgt aus dem Axiom der Vereinigung angewandt auf $A \to B$ und $A \to A$.

Man beachte, dass F_1^+ alle FDs enthält, die über $\mathcal{R} = AB$ möglich sind, mit Ausnahme der FDs $B \to A$ und $B \to AB$.

$$F_2^+ = \{A \to A, B \to B, AB \to AB, AB \to A, AB \to B\}$$

Hier folgen alle FDs direkt aus dem Axiom der Reflexivität. Die FD $A \to B$ is nicht in F_2^+, daher gilt $F_1^+ \neq F_2^+$.

(b) Das Axiom der Dekomposition lautet: Falls $\alpha \to \beta\gamma$ gilt, dann gilt auch $\alpha \to \beta$ und $\alpha \to \gamma$. Dies kann wie folgt hergeleitet werden:

$$\text{Trans.}\ \dfrac{\begin{array}{c}[\text{gegeben}]\\ \alpha \to \beta\gamma\end{array} \qquad \dfrac{}{\beta\gamma \to \gamma}\ \text{Refl.}}{\alpha \to \gamma}$$

$$\text{Trans.}\ \dfrac{\begin{array}{c}[\text{gegeben}]\\ \alpha \to \beta\gamma\end{array} \qquad \dfrac{}{\beta\gamma \to \beta}\ \text{Refl.}}{\alpha \to \beta}$$

(c) Wir leiten die FD $A \to E$ wie folgt aus F_3 ab.

$$\dfrac{\dfrac{\overset{[\text{gegeben}]}{A \to B} \qquad \overset{[\text{gegeben}]}{B \to C}}{A \to C}\,\text{Trans.} \qquad \dfrac{\overset{[\text{gegeben}]}{C \to ED}}{C \to E}\,\text{Dekomp.}}{A \to E}\,\text{Trans.}$$

Es gilt daher, dass $A \to E \in F_3^+$.

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) $F^1 = \{A \to B, A \to C, D \to EF, DEF \to AC, E \to F\}$

(b) $F^2 = \{AD \to CB, B \to E, E \to C, D \to E, BD \to C, AD \to E\}$

Lösung:

(a) $F_C^1 = \{A \to BC, D \to AE, E \to F\}$

(b) $F_C^2 = \{AD \to B, B \to E, E \to C, D \to E\}$

Aufgabe 6 (Schlüsselbestimmung) *[0.5 Punkte]*

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

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

Lösung:
Die einzigen Schlüssel sind AB, AD und AC. Die Menge der Superschlüssel:

$$\{AB, AC, AD, ABC, ABD, ACD, ABCD\}$$

(b) Gegeben sei folgendes Relationenschema samt funktionalen Abhängigkeiten:

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

Erklären Sie, warum CD und AB keine Schlüssel sind. Berechnen Sie weiters alle Schlüssel.

Lösung:
AB ist kein Schlüssel da $AB \not\to \mathcal{R}$. Das kann man unter anderem daran erkennen, dass $\mathsf{AttrHülle}(F, \{AB\}) = \{ABC\} \neq \mathcal{R}$. CD ist kein Schlüssel, da D bereits Schlüssel ist. CD ist daher nicht minimal. Der einzige Schlüssel ist D.

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

$$\mathcal{R} = ABCDEF$$
$$F = \{A \to D, BC \to FA, C \to DC, D \to F, D \to E, E \to D\}$$

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:

$$F_c = \{A \to D, BC \to A, C \to D, D \to EF, E \to D\}$$

2. Erstelle ein Relationenschema für jedes Element in F_c und ordne geltende FDs zu:

Relationenschema	Geltende FDs
$\mathcal{R}_1 = AD$	$F_1 = \{A \to D\}$
$\mathcal{R}_2 = ABC$	$F_2 = \{BC \to A\}$
$\mathcal{R}_3 = CD$	$F_3 = \{C \to D\}$
$\mathcal{R}_4 = DEF$	$F_4 = \{D \to EF, E \to D\}$
$\mathcal{R}_5 = ED$	$F_5 = \{E \to D\}$

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

   Der Schlüssel BC ist in zumindest einem der erzeugten Teilschemata enthalten ($\mathcal{R}_2$). Es muss kein weiteres Schema erzeugt werden.

4. Das Schema $\mathcal{R}_5$ kann eliminiert werden, da $\mathcal{R}_5 \subseteq \mathcal{R}_4$ gilt.

Ergebnis (Schlüssel sind unterstrichen):

Relationenschema	Geltende FDs
$\mathcal{R}_1 = \underline{A}D$	$F_1 = \{A \to D\}$
$\mathcal{R}_2 = A\underline{BC}$	$F_2 = \{BC \to A\}$
$\mathcal{R}_3 = \underline{C}D$	$F_3 = \{C \to D\}$
$\mathcal{R}_4 = \underline{D}EF$	$F_4 = \{D \to EF, E \to D\}$

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

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

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

(a) in dritter Normalform und/oder

(b) in Boyce-Codd-Normalform

ist. Begründen Sie Ihre Antworten.

Lösung:
Die einzigen Schlüssel von $\mathcal{R}$ sind AC und AE. Wir wenden die Dekomposition an, damit alle FDs auf der rechten Seite nur noch ein Attribut haben und erhalten die folgenden FDs:

$$\{AC \to D, AC \to E, E \to C, ACD \to B, ED \to D\}$$

Für 3NF prüfen wir nun für jede dieser FDs, ob sie eine der drei Bedingungen erfüllt: $AC \to D$ hat links einen Superschlüssel, $AC \to E$ hat links einen Superschlüssel, $E \to C$ enthält rechts einen Teil eines Schlüssels, $ACD \to B$ hat links einen Superschlüssel, und $ED \to D$ ist trivial. Jede FD erfüllt daher eine der geforderten Eigenschaften und somit ist F in 3NF.

Für BCNF prüfen wir ob jede der FDs eine der zwei Bedingungen erfüllt: $AC \to D$ hat links einen Superschlüssel, $AC \to E$ hat links einen Superschlüssel, $E \to C$ hat links keinen Superschlüssel und ist auch nicht trivial, daher erfüllt $E \to C$ keine der beiden Bedingungen und somit ist F nicht in BCNF.

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

$$\begin{aligned} R &= ABCDEG \\ F &= \{B \to CD, ABC \to E, C \to G\} \end{aligned}$$

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:
Einziger Schlüssel von $\mathcal{R}$ ist AB. Die FDs $B \to CD$ und $C \to G$ 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 $B \to CD$. Zerlege $\mathcal{R}$ daher in:

$\mathcal{R}_1 = BCD$	$F_1 = \{B \to CD\}$	Schlüssel: B
$\mathcal{R}_2 = ABEG$	$F_2 = \{\}$	Schlüssel: $ABEG$

$\mathcal{R}_1$ ist nun in BCNF und $\mathcal{R}_2$ auch. Die Relationen der Zerlegung sind somit: $\mathcal{R}_1$ und $\mathcal{R}_2$. Da etwa die FD $ABC \to E$ verloren geht — das heisst $(ABC \to E) \notin (F_1 \cup F_2)^+$ — ist die Zerlegung nicht abhängigkeitstreu (des weiteren geht die FD $C \to G$ verloren).

Alternative Lösung wenn zuerst die FD $C \to G$ ausgewählt wird:

$\mathcal{R}_1 = CG$	$F_1 = \{C \to G\}$	Schlüssel: C
$\mathcal{R}_2 = ABCDE$	$F_2 = \{B \to CD, ABC \to E\}$	Schlüssel: AB

$\mathcal{R}_1$ ist in BCNF aber $\mathcal{R}_2$ noch nicht. Daher wird R_2 weiter zerlegt indem wir die einzige FD auswählen, welche die BCNF verletzt, nämlich $B \to CD$.

$\mathcal{R}_{21} = BCD$	$F_{21} = \{B \to CD\}$	Schlüssel: B
$\mathcal{R}_{22} = ABE$	$F_{22} = \{\}$	Schlüssel: ABE

Die Relationen der Zerlegung sind somit: $\mathcal{R}_1, \mathcal{R}_{21}$ und $\mathcal{R}_{22}$. Da die FD $ABC \to E$ verloren geht, ist die Zerlegung nicht abhängigkeitstreu.