TU München, Fakultät für Informatik
Lehrstuhl III: Datenbanksysteme
Prof. Alfons Kemper, Ph.D.

Übung zur Vorlesung *Grundlagen: Datenbanken* im WS23/24

Christoph Anneser, Michael Jungmair, Stefan Lehner, Moritz Sichert, Lukas Vogel
(gdb@in.tum.de)
https://db.in.tum.de/teaching/ws2324/grundlagen/

Blatt Nr. 08

Hausaufgabe 1

Gegeben sei die Relation $R : \{[A, B, C]\}$ mit $A \in \{1, 2, 3\}$, $B \in \{x, y, z\}$, $C \in \{7, 8, 9\}$. Außerdem ist die folgende Ausprägung gegeben:

A	B	C
1	x	8
1	y	9

Fügen Sie dieser Ausprägung möglichst wenige Tupel hinzu, sodass **alle** MVDs der Form $\alpha \twoheadrightarrow \beta$ gelten mit $\alpha \subseteq \{A, B, C\}$, $|\alpha| \geq 1$, $\beta \subseteq \{A, B, C\}$.

Lösung:

Relevant für diese Aufgabe sind die folgenden MVDs:

1. $A \twoheadrightarrow B$
2. $A \twoheadrightarrow C$ (Komplement zu 1.)
3. $B \twoheadrightarrow A$
4. $B \twoheadrightarrow C$ (Komplement zu 3.)
5. $C \twoheadrightarrow A$
6. $C \twoheadrightarrow B$ (Komplement zu 5.)

Alle anderen MVDs der Form $\alpha \twoheadrightarrow \beta$ (z.B. $AB \twoheadrightarrow C$ oder $A \twoheadrightarrow \emptyset$) sind trivial, d.h. sie gelten immer.

In der Ausprägung wird $A \twoheadrightarrow B$ (und damit auch $A \twoheadrightarrow C$) verletzt. Es müssen die Tupel $(1, x, 9)$ und $(1, y, 8)$ hinzugefügt werden. Dadurch erhält man die folgende Ausprägung:

A	B	C
1	x	8
1	y	9
1	x	9
1	y	8

Da diese Ausprägung für A ausschließlich den Wert 1 enthält, gelten die MVDs $B \twoheadrightarrow A$ und $C \twoheadrightarrow A$ hier bereits. Damit gelten auch die Komplemente $B \twoheadrightarrow C$ und $C \twoheadrightarrow B$.

Hausaufgabe 2

Geben Sie für jede der Normalformen 1NF, 2NF, 3NF, BCNF, 4NF jeweils eine Relation mit FDs an, sodass die Relation in der gewünschten Normalform ist (und in keiner höheren).

Lösung:

Für alle Normalformen betrachten wie die Relation $\mathcal{R} = \{A, B, C, D\}$.

- 1.NF:

  FDs:

  - $AB \to C$
  - $B \to D$

  Die Relation ist nicht mengenwertig, daher 1. NF. D ist lediglich von B abhängig, der Kandidatenschlüssel ist aber AB, weswegen D nicht voll funktional vom Kandidatenschlüssel abhängig ist, daher keine 2. NF.

- 2. NF:

  FDs:

  - $AB \to C$
  - $C \to D$

  Jedes Attribut der Relation ist voll funktional abhängig vom Kandidatenschlüssel AB, daher 2.NF. Das Attribut D ist transitiv und nicht direkt vom Kandidatenschlüssel abhängig, darum nicht 3. NF.

- 3. NF:

  FDs:

  - $BC \to AD$
  - $D \to C$

  Für jede FD gilt entweder, dass sie trivial ist, dass die linke Seite Superschlüssel ist oder dass die rechte Seite in einem Kandidatenschlüssel enthalten ist, daher 3. NF. Bei der BCNF fällt die dritte erlaubte Art von FD weg, daher müssen FDs trivial sein oder ihre linke Seite Superschlüssel. Da die zweite FD des Beispiels dies verletzt, ist die Relation nicht in BCNF und daher genau in 3. NF.

- BCNF:

  FDs:

  - $AB \to CD$
  - $BC \to AD$
  - $D \twoheadrightarrow C$

  BCNF, da die BCNF verletzende FD aus dem Beispiel für 3. NF entfernt wurde. Nicht 4. NF weil eine nicht triviale MVD gilt, deren linke Seite nicht Superschlüssel ist.

- 4. NF

  FDs:

  - $AB \to CD$
  - $BC \to AD$

Nach Entfernung der nicht trivialen MVD dann auch 4. NF.

Hausaufgabe 3

Bestimmen Sie alle Kandidatenschlüssel der Relation R. Wenden Sie den Dekompositionsalgorithmus an, um die Relation R in die BCNF zu zerlegen und unterstreichen Sie die Schlüssel der Teilrelationen des Endergebnisses.

$$R : \{[A, B, C, D, E, F]\}$$

FDs:

1. $B \to DA$
2. $DEF \to B$
3. $C \to EA$

Prüfen Sie als erstes, ob FD 1) für die Zerlegung geeignet ist und - falls dies der Fall ist - verwenden Sie diese im ersten Zerlegungsschritt. Für diese Aufgabe ist zu bedenken, dass die oben angegebenen FDs eine Charakterisierung der insgesamt geltenden FDs sind. Die Menge der geltenden FDs ist größer. Wieso? Wie muss dies beim Dekompositionsalgorithmus genutzt werden?

Lösung:

- Dekompositionsalgorithmus:
  - Starte mit $Z := \{R\}$.
  - R in BCNF? - Nein, $B \to DA$ verletzt die BCNF.
    - Zerlegung anhand FD $B \to DA$, da $\{B\}$ kein Superschlüssel:
      $R_1 : \{[A, B, D]\}$ mit den FDs $F_1 = \{B \to DA\}$,
      $R_2 : \{[B, C, E, F]\}$ mit den FDs $F_2 = \{C \to E\}$, FD (2) geht verloren und FD (3) geht "teilweise" verloren: wenn $C \to AE$ gilt, dann gilt auch $C \to A$ und $C \to E$ (Dekompositionsregel), aber lediglich $C \to E$ bleibt erhalten.
      $Z := \{R_1, R_2\}$
  - R_1 in BCNF? - Ja.
  - R_2 in BCNF? - Nein, $C \to E$ verletzt die BCNF.
    - Zerlegung anhand FD $C \to E$, da $\{C\}$ kein Superschlüssel:
      $R_{2.1} : \{[C, E]\}$ mit den FDs $F_{2.1} = \{C \to E\}$,
      $R_{2.2} : \{[B, C, F]\}$ mit ausschließlich trivialen FDs.
      $Z := \{R_1, R_{2.1}, R_{2.2}\}$
  - $R_{2.1}$ in BCNF? - Ja.
  - $R_{2.2}$ in BCNF? - Ja.
- Ergebnis:

$$\begin{array}{rcl} R_1 & : & \{[A, \underline{B}, D]\} \\ R_{2.1} & : & \{[\underline{C}, E]\} \\ R_{2.2} & : & \{[\underline{B}, \underline{C}, \underline{F}]\} \end{array}$$

Im Allgemeinen ist eine gegebene FD-Menge weder minimal noch vollständig. Die angegebenen FDs enthalten also möglicherweise Redundanzen einerseits und andererseits werden triviale Abhängigkeiten i.d.R. nicht explizit mit angegeben. Bei der Ausführung des Dekompositionsalgorithmus müssen jedoch alle *geltenden* FDs betrachtet werden, die sich mit Hilfe der Axiome von Armstrong herleiten lassen (F^+). So gilt im obigen Beispiel in R_2 die FD $C \to E$, obwohl diese nicht explizit angegeben war.

Hausaufgabe 4

Überführen Sie das folgende Schema verlustlos in die 4. NF:

$$R : \{[A, B, C, D, E]\}$$

$$\begin{aligned} AB &\to CDE \\ B &\twoheadrightarrow D \\ C &\twoheadrightarrow DE \end{aligned}$$

Beachten Sie, dass es zwei mögliche Lösungen gibt. Geben Sie beide an!

Lösung:

Dieses Schema halt AB als einzigen Kandidatenschlüssel. Dementsprechend verletzen beide MVDs die Bedingungen der 4. NF. Somit können beide zur Zerlegung verwendet werden. In diesem Schema führt das dann zu zwei verschiedenen Ergebnissen.

1. Lösung:

Für die erste Lösung verwenden wir die MVD $B \twoheadrightarrow D$ zur Dekomposition. Damit erhalten wir also:

$$\begin{aligned} \mathcal{R}_1 &= \{\underline{B, D}\} \\ \mathcal{R}_2 &= \{\underline{A, B}, C, E\} \end{aligned}$$

Für $\mathcal{R}_1$ gelten jetzt keine (nicht-trivialen) FDs oder MVDs mehr, also ist es in 4. NF. Für $\mathcal{R}_2$ gelten $AB \to CE$ und $C \twoheadrightarrow AB$. Letztere ergibt sich aus der MVD $C \twoheadrightarrow DE$ mit Hilfe der Komplementregel. Also muss nun $\mathcal{R}_2$ Anhand von dieser MVD aufgeteilt werden. Daraus ergeben sich:

$$\begin{aligned} \mathcal{R}_{2.1} &= \{\underline{A, B}, C\} \\ \mathcal{R}_{2.2} &= \{\underline{C, E}\} \end{aligned}$$

Für $\mathcal{R}_{2.1}$ gilt nur noch $AB \to C$. Für $\mathcal{R}_{2.2}$ gelten keine (nicht-trivialen) FDs oder MVDS. Damit sind beide in der 4. NF.

Das Ergebnis ist also: $\{\mathcal{R}_1, \mathcal{R}_{2.1}, \mathcal{R}_{2.2}\}$.

2. Lösung:

Für die zweite Lösung verwenden wir die MVD $C \twoheadrightarrow DE$ zur Dekomposition. Damit erhalten wir:

$$\begin{aligned} \mathcal{R}_1 &= \{\underline{C, D, E}\} \\ \mathcal{R}_2 &= \{\underline{A, B}, C\} \end{aligned}$$

Für $\mathcal{R}_1$ gelten keine (nicht-trivialen) FDs oder MVDs mehr. Für $\mathcal{R}_2$ gilt $AB \rightarrow C$. Damit sind beide in der 4. NF und eine weitere Zerlegung ist nicht notwendig.

Das Ergebnis ist also: $\{\mathcal{R}_1, \mathcal{R}_2\}$.