



Übung zur Vorlesung *Grundlagen: Datenbanken im WS21/22*

Michael Jungmair, Josef Schmeißer, Moritz Sichert, Lukas Vogel (gdb@in.tum.de)
<https://db.in.tum.de/teaching/ws2122/grundlagen/>

Blatt Nr. 07

Hausaufgabe 1

Betrachten Sie das Relationenschema

Fahrplan: {[Linie, Verbund, von, nach, von GPS, nach GPS, Preis, #Fahrzeuge, Modus]}

mit der folgenden beispielhaften Ausprägung:

Linie	Verbund	von	nach	von GPS	nach GPS	Preis	#Fahrzeuge	Modus
U6	MVV	GF	G	0N 0W	1S 0W	1€	20	U-Bahn
U6	MVV	G	GH	1S 0W	2S 0W	1€	20	U-Bahn
U6	MVV	GH	FR	2S 0W	5S 0W	3€	20	U-Bahn
U3	MVV	MF	GI	8S 0W	9S 0W	1€	16	U-Bahn
690	MVV	GF	DI	0N 0W	1N 0W	1€	5	Bus
690	MVV	DI	NF	1N 0W	3N 1W	2€	5	Bus
690	MVV	NF	EH	3N 1W	5N 2W	2€	5	Bus
S1	MVV	NF	EH	3N 1W	5N 2W	3€	8	S-Bahn

- Bestimmen Sie die geltenden FDs.
- Bestimmen Sie die Kandidatenschlüssel.

Lösung:

a) Im Relationenschema gelten die folgenden funktionalen Abhängigkeiten:

- {Linie} → {#Fahrzeuge, Modus}
- {von} → {von GPS}
- {nach} → {nach GPS}
- {von GPS} → {von}
- {nach GPS} → {nach}
- {Linie, von, nach} → {Preis}
- ∅ → {Verbund}

Anhand der Ausprägung werden manche FDs nicht widerlegt, wie z.B. {#Fahrzeuge} → {Linie}. Bei dieser Musterlösung haben wir aber angenommen, dass es noch weitere Tupel in der Ausprägung geben könnte, die diese FD widerlegt.

Natürlich gelten auch alle anderen funktionalen Abhängigkeiten, die mit Hilfe der Armstrong-Axiome daraus hergeleitet werden können.

b) Die möglichen Kandidatenschlüssel sind:

- {Linie, von, nach}
- {Linie, von GPS, nach}
- {Linie, von, nach GPS}
- {Linie, von GPS, nach GPS}

Aus „Linie“ können die Attribute „#Fahrzeuge“ und „Modus“, aus „Linie“, „von“ und „nach“ das Attribut „Preis“ abgeleitet werden. Da „von“ und „von GPS“ sowie „nach“ und „nach GPS“ jeweils in beide Richtungen abgeleitet werden können, enthält der Kandidatenschlüssel alle möglichen Kombinationen.

Hausaufgabe 2

Betrachten Sie ein abstraktes Relationenschema $\mathcal{R} = \{A, B, C, D, E, F, G\}$ mit den FDs

1. $A \rightarrow BC$
2. $C \rightarrow DA$
3. $E \rightarrow ABC$
4. $F \rightarrow CD$
5. $CD \rightarrow BEF$

- a) Berechnen Sie die Attributhülle von A .
- b) Bestimmen Sie alle Kandidatenschlüssel.
- c) Bestimmen Sie zu den gegebenen FDs die kanonische Überdeckung.
- d) Überführen Sie die Relation in die dritte Normalform, indem Sie den Synthesealgorithmus anwenden.

Lösung:

Attributhülle von A

Berechnung der Attributhülle von A mit Hilfe des bekannten *AttrHülle*-Algorithmus.

Aufruf: $AttrHülle(FD, \{A\})$.

Schritt	betrachtete FD	Ergebnis
init		$\{A\}$
1.	$A \rightarrow BC$	$\{A, B, C\}$
2.	$C \rightarrow DA$	$\{A, B, C, D\}$
3.	$CD \rightarrow BEF$	$\{A, B, C, D, E, F\}$

Damit enthält die Attributhülle von A alle Attribute außer G .

Kandidatenschlüssel

Nach der vorherigen Berechnung enthält die Attributhülle von $\{A\}$ alle Attribute außer $\{G\}$. Damit ist $\{A, G\}$ ein Superschlüssel. $\{A, G\}$ ist auch minimal, da es keine FD gibt, die G enthält. Somit ist $\{A, G\}$ ein Kandidatenschlüssel.

Da man aus $\{C\}$ und $\{E\}$ direkt A folgern kann, handelt es sich bei $\{C, G\}$ und $\{E, G\}$ ebenfalls um Superschlüssel. Sie sind ebenfalls minimal und damit Kandidatenschlüssel.

Aus $\{F\}$ wiederum kann C und somit A gefolgert werden. Damit ist $\{F, G\}$ analog zu oben auch ein Kandidatenschlüssel.

$\{B, G\}$ und $\{D, G\}$ sind dagegen keine Kandidatenschlüssel. $\{B, G\}$ ist nicht einmal Superschlüssel. $\{C, D, G\}$ wäre zwar ein Superschlüssel, allerdings kein Kandidatenschlüssel, da nicht minimal.

Kandidatenschlüssel sind: $\{A, G\}, \{C, G\}, \{E, G\}, \{F, G\}$.

Kanonische Überdeckung

Gegeben ist die Ausgangsmenge $F = \{A \rightarrow BC, C \rightarrow DA, E \rightarrow ABC, F \rightarrow CD, CD \rightarrow BEF\}$.

1. Führe für jede FD $\alpha \rightarrow \beta \in F$ die Linksreduktion durch.

Einzig in Betracht kommende FD ist $CD \rightarrow BEF$.

- Ist C überflüssig?
 $AttrHülle(F, \{D\}) = \{D\} \not\supseteq \{B, E, F\}$
- Ist D überflüssig?
 $AttrHülle(F, \{C\}) =$

Schritt	betrachtete FD	Ergebnis
init		$\{C\}$
1.	$C \rightarrow DA$	$\{A, C, D\}$
2.	$CD \rightarrow BEF$	$\{A, B, C, D, E, F\}$

$$\{C\}^+ = \{A, B, C, D, E, F\} \supseteq \{B, E, F\}$$

Damit kann $CD \rightarrow BEF$ zu $C \rightarrow BEF$ reduziert werden.

2. Führe für jede (verbliebene) FD $\alpha \rightarrow \beta$ die Rechtsreduktion durch.

Bisheriges Zwischenergebnis:

$$A \rightarrow BC \tag{1}$$

$$C \rightarrow DA \tag{2}$$

$$E \rightarrow ABC \tag{3}$$

$$F \rightarrow CD \tag{4}$$

$$C \rightarrow BEF \tag{5}$$

Betrachte FD (1):

- Ist B überflüssig?
 $B \in AttrHülle(F - FD (1) \cup (A \rightarrow C), A)$, da $A \rightarrow C \rightarrow BEF$.
- Ist C überflüssig?
 $C \notin AttrHülle(F - FD (1) \cup (A \rightarrow \emptyset), A)$.

Damit erhält man für FD (1): $A \rightarrow C$.

Betrachte FD (2):

- Ist D überflüssig?
 $D \in AttrHülle(F - FD (2) \cup (C \rightarrow A), C)$, da $C \rightarrow BEF, F \rightarrow CD$.
- Ist A überflüssig?
 $A \in AttrHülle(F - FD (2) \cup (C \rightarrow \emptyset), C)$, da $C \rightarrow BEF, E \rightarrow ABC$.

Damit erhält man für FD (2): $C \rightarrow \emptyset$.

Betrachte FD (3):

- Ist A überflüssig?
 $A \notin \text{AttrHülle}(F - \text{FD (3)} \cup (E \rightarrow BC), E)$.
- Ist B überflüssig?
 $B \in \text{AttrHülle}(F - \text{FD (3)} \cup (E \rightarrow AC), E)$, da $E \rightarrow AC, C \rightarrow BEF$.
- Ist C überflüssig?
 $C \in \text{AttrHülle}(F - \text{FD (3)} \cup (E \rightarrow A), E)$, da $E \rightarrow A, A \rightarrow C$.

Damit erhält man für FD (3): $E \rightarrow A$.

Betrachte FD (4):

- Ist C überflüssig?
 $C \notin \text{AttrHülle}(F - \text{FD (4)} \cup (F \rightarrow D), F)$.
- Ist D überflüssig?
 $D \notin \text{AttrHülle}(F - \text{FD (4)} \cup (F \rightarrow C), F)$.

Damit bleibt FD (4) unverändert.

Betrachte FD (5):

- Ist B überflüssig?
 $B \notin \text{AttrHülle}(F - \text{FD (5)} \cup (C \rightarrow EF), C)$.
- Ist E überflüssig?
 $E \notin \text{AttrHülle}(F - \text{FD (5)} \cup (C \rightarrow BF), C)$.
- Ist F überflüssig?
 $F \notin \text{AttrHülle}(F - \text{FD (5)} \cup (C \rightarrow BE), C)$.

Damit bleibt FD (5) unverändert.

3. *Entferne die FDs der Form $\alpha \rightarrow \emptyset$.*

Bisheriges Zwischenergebnis:

$$\begin{array}{l}
 A \rightarrow C \\
 C \rightarrow \emptyset \\
 E \rightarrow A \\
 F \rightarrow CD \\
 C \rightarrow BEF
 \end{array} \tag{6}$$

FD (6) wird eliminiert.

4. *Fasse mittels der Vereinigungsregel FDs der Form $\alpha \rightarrow \beta_1, \dots, \alpha \rightarrow \beta_n$ zusammen.*

Bisheriges Zwischenergebnis:

$$F_c = \left\{ \begin{array}{l} A \rightarrow C \\ E \rightarrow A \\ F \rightarrow CD \\ C \rightarrow BEF \end{array} \right.$$

Es werden keine FDs vereinigt, da es keine zwei FDs mit gleicher linker Seite gibt.

F_c ist eine kanonische Überdeckung zur Ausgangsmenge F .

Dritte Normalform

Bestimmen der kanonischen Überdeckung siehe oben. Für jede funktionale Abhängigkeit aus der kanonischen Überdeckung wird ein Relationenschema erstellt:

$$\begin{aligned}R_1 &= \{\underline{A}, C\} \\R_2 &= \{\underline{E}, A\} \\R_3 &= \{\underline{E}, C, D\} \\R_4 &= \{\underline{C}, B, E, F\}\end{aligned}$$

Kein Schema enthält einen Kandidatenschlüssel. Also muss ein zusätzliches erstellt werden. Beispielsweise $R_\kappa = \{\underline{A}, G\}$.

Keines der Relationenschemata ist in einem anderen Schema enthalten, sodass nichts eliminiert werden kann.

Hausaufgabe 3

Ist die kanonische Überdeckung F_c einer Menge F von funktionalen Abhängigkeiten eindeutig? Begründen Sie Ihre Antwort oder finden Sie ein Gegenbeispiel.

Lösung:

Die kanonische Überdeckung F_c zu einer Menge von funktionalen Abhängigkeiten F ist nicht eindeutig.

Begründung: Im Algorithmus zur Bestimmung der kanonischen Überdeckung ist nicht festgelegt, in welcher Reihenfolge die FDs bearbeitet werden.

Als Beispiel seien folgende funktionale Abhängigkeiten gegeben:

1. $A \rightarrow BC$
2. $B \rightarrow AC$

Wird die erste FD in der Rechtsreduktion zuerst abgearbeitet, ergibt sich:

$$F_c = \{A \rightarrow B, B \rightarrow AC\}$$

Wird die zweite FD in der Rechtsreduktion zuerst abgearbeitet, erhält man hingegen:

$$F_c = \{A \rightarrow BC, B \rightarrow A\}$$

Hausaufgabe 4

Gegeben sei eine Relation

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

die schon sehr viele Daten enthält (Millionen Tupel). Sie „vermuten“, dass folgendes gilt:

- a) AB ist ein Superschlüssel der Relation
- b) $DE \rightarrow B$

Formulieren Sie SQL-Anfragen, die Ihre Vermutungen bestätigen oder widerlegen.

Lösung:

- a) Durch Gruppierung nach A und B kann anhand der Anzahl der Tupel ermittelt werden, ob hier eine Verletzung der Schlüsseleigenschaft vorliegt. Werden also mindestens zwei Tupel mit den gleichen Werten für A und B als Ergebnis ausgegeben, so bildet AB keinen Schlüssel der Relation, ist das Ergebnis der Anfrage jedoch leer, so ist AB ein Superschlüssel.

```
select A, B
from R
group by A, B
having count(*) > 1;
```

- b) In diesem Fall muss nur gelten, dass für alle Tupel, die gleiche Werte in D und E besitzen, auch die Werte für das Attribut B gleich sind. D.h. wenn nach D und E gruppiert wird, muss die Anzahl der verschiedenen Werte für B kleiner oder gleich 1 sein. Es gilt wieder, dass das Ergebnis der Anfrage alle Tupel enthält, die die Vermutung verletzen. Ist das Ergebnis leer, so gilt $DE \rightarrow B$.

```
select D, E
from R
group by D, E
having count(distinct B) > 1;
```