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. 07

Hausaufgabe 1

Betrachten Sie das Relationenschema

Fahrplan: {[Linie, Verband, 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

a) Bestimmen Sie die geltenden FDs.

b) Bestimmen Sie die Kandidatenschlüssel.

Lösung:

a) Im Relationenschema gelten die folgenden funktionalen Abhängigkeiten:
   - {Linie} $\rightarrow$ {#Fahrzeuge, Modus}
   - {von} $\rightarrow$ {von GPS}
   - {nach} $\rightarrow$ {nach GPS}
   - {von GPS} $\rightarrow$ {von}
   - {nach GPS} $\rightarrow$ {nach}
   - {Linie, von, nach} $\rightarrow$ {Preis}
   - $\emptyset \rightarrow$ {Verbund}

   Anhand der Ausprägung werden manche FDs nicht widerlegt, wie z.B. {#Fahrzeuge} $\rightarrow$ {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. $A \to BC$
2. $C \to DA$
3. $E \to ABC$
4. $F \to CD$
5. $CD \to BEF$

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.	$A \to BC$	$\{A, B, C\}$
2.	$C \to DA$	$\{A, B, C, D\}$
3.	$CD \to BEF$	$\{A, B, C, D, E, F\}$

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 $\{A, G\}$ ist 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 $F = \{A \to BC, C \to DA, E \to ABC, F \to CD, CD \to BEF\}$.

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

   Einzige in Betracht kommende FD ist $CD \to BEF$.

   - 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.	$C \to DA$	$\{A, C, D\}$
2.	$CD \to BEF$	$\{A, B, C, D, E, F\}$

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

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

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

   Bisheriges Zwischenergebnis:

$$
\begin{aligned}
A &\to BC && (1)\\
C &\to DA && (2)\\
E &\to ABC && (3)\\
F &\to CD && (4)\\
C &\to BEF && (5)
\end{aligned}
$$

   Betrachte FD (1):

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

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

   Betrachte FD (2):

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

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

Betrachte FD (3)**:**

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

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

Betrachte FD (4)**:**

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

Damit bleibt FD (4) unverändert.

Betrachte FD (5)**:**

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

Damit bleibt FD (5) unverändert.

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

   Bisheriges Zwischenergebnis:

$$\begin{aligned} A &\to C \\ C &\to \emptyset \qquad (6) \\ E &\to A \\ F &\to CD \\ C &\to BEF \end{aligned}$$

   FD (6) wird eliminiert.

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

   Bisheriges Zwischenergebnis:

$$F_c = \begin{cases} A \to C \\ E \to A \\ F \to CD \\ C \to BEF \end{cases}$$

   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:

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

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. $A \to BC$
2. $B \to AC$

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

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

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

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

Hausaufgabe 4

Welche Studenten haben alle Vorlesungen, die sie haben prüfen lassen, auch tatsächlich vorher gehört? Formulieren Sie eine SQL-Anfrage, welche diese Studenten ausgibt.

Lösung:
Die Anforderung, dass die Studenten im Anfrage-Ergebnis alle Vorlesungen, die sie haben prüfen lassen auch tatsächlich gehört haben, lässt sich umschreiben zu: „Es darf keine Vorlesung geben, die geprüft wurde, zu der es aber keinen Eintrag in *hoeren* gibt."

```
select s.*
from Studenten s
where not exists (select * from pruefen p
                  where s.MatrNr = p.MatrNr
                    and not exists (select *
                                    from hoeren h
                                    where h.MatrNr = s.MatrNr
                                      and h.VorlNr = p.VorlNr));
```