Übung zur Vorlesung *Grundlagen: Datenbanken* im WS17/18

Harald Lang, Linnea Passing (gdb@in.tum.de)
http://www-db.in.tum.de/teaching/ws1718/grundlagen/

Blatt Nr. 08

Hausaufgabe 1

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 2

Betrachten Sie ein abstraktes Relationenschema $\mathcal{R} = \{A, B, C, D, E, F\}$ 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: $\mathit{AttrHülle}(FD, \{A\})$.

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.

Kandidatenschlüssel

$\{A\}$ ist nach der vorherigen Berechnung (Attributhülle von A) ein Superschlüssel. Da $\{A\}$ außerdem minimal ist, ist $\{A\}$ ein Kandidatenschlüssel. Da man aus $\{C\}$ und $\{E\}$ direkt A folgern kann, handelt es sich hier ebenfalls um Superschlüssel und da sie einelementig sind (also minimal sind) auch um Kandidatenschlüssel. Aus $\{F\}$ wiederum kann C und somit A gefolgert werden. Damit ist $\{F\}$ analog zu oben auch ein Kandidatenschlüssel.

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

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

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

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:

$$A \to BC \quad (1)$$
$$C \to DA \quad (2)$$
$$E \to ABC \quad (3)$$
$$F \to CD \quad (4)$$
$$C \to BEF \quad (5)$$

Betrachte FD (1)**:**

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

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

Betrachte FD (2)**:**

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

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

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:

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

R_1 enthält einen der Kandidatenschlüssel (sogar zwei: nämlich $\{A\}$ und $\{C\}$), so dass kein zusätzliches Schema erstellt werden muss. Keines der Relationenschemata ist in einem anderen Schema enthalten, so dass nichts eliminiert werden kann.

Hausaufgabe 3

Gegeben sei ein erweitertes Universitätsschema mit den folgenden zusätzlichen Relationen *StudentenGF* und *ProfessorenF*:

$$\begin{array}{rcl} \text{StudentenGF} & : & \{[\underline{\text{MatrNr : integer}}, \text{Name : varchar(20)}, \text{Semester : integer}, \\ & & \;\text{Geschlecht : char}, \text{FakName : varchar(20)}]\} \\ \text{ProfessorenF} & : & \{[\underline{\text{PersNr : integer}}, \text{Name : varchar(20)}, \text{Rang : char(2)}, \\ & & \;\text{Raum : integer}, \text{FakName : varchar(20)}]\} \end{array}$$

Die erweiterten Tabellen sind auch in der Webschnittstelle angelegt.

(a) Ermitteln Sie den Männeranteil an den verschiedenen Fakultäten in SQL!

(b) Ermitteln Sie in SQL die Studenten, die alle Vorlesungen ihrer Fakultät hören. Geben Sie zwei Lösungen an, höchstens eine davon darf auf Abzählen basieren.

Lösung:

(a)
```
with
FakTotal as (
  select FakName, count(*) as Total
  from StudentenGF
  group by FakName),
FakMaenner as (
  select FakName,count(*) as Maenner
  from StudentenGF
  where Geschlecht = 'M'
  group by FakName)
select FakTotal.FakName, (case when Maenner is null
   then 0 else Maenner end)/(total*1.0)
from FakTotal left outer join FakMaenner
on FakTotal.FakName = FakMaenner.FakName
```

Wir müssen beachten, dass nicht jede Fakultät Männer beherbergt, weswegen diese Fakultäten (in der Standardausprägung im SQL Interface ist dies für Theologie der Fall) dann aus dem Ergebnis herausfallen würden. Aus diesem Grund verwenden wir einen `LEFT OUTER JOIN` um die Zahl der Männer und die Zahl der Studenten insgesamt zu verbinden, wodurch auch die Theologie Fakultät im Ergebnis enthalten ist, auch wenn es keine Männer gibt.

Das `CASE`-Konstrukt dient in der oberen Anfrage dazu, den NULL Wert, die durch den Left Join für die Anzahl der Männer entstehen, wenn es keine Männer gibt, durch die Zahl 0 zu ersetzen. Alternativ ist dies möglich, indem man `COALESCE(maenner,0)/(total*1.0)` verwendet.

Alternativ können wir das **case**-Konstrukt verwenden, um die Anzahl der Männer an den jeweiligen Fakultäten zu ermitteln. Den Männeranteil erhalten wir dann, indem wir die Anzahl der Männer durch die Gesamtanzahl der Studenten an der Fakultät teilen.

```
select FakName,
 (sum(case when Geschlecht = 'M' then 1.00 else 0.00
    end)) / count(*)
from StudentenGF
group by FakName
```

(b) Wir fordern hier, dass es keine Vorlesung an der Fakultät des Studenten (d.h. von einem Professor der gleichen Fakultät gelesen) geben darf, die vom Studenten nicht gehört wird.

```
select s.*
from StudentenGF s
where not exists (select *
                  from Vorlesungen v, ProfessorenF p
                  where v.gelesenVon = p.PersNr
                    and p.FakName = s.FakName
                    and not exists
                       (select *
                        from hoeren h
                        where h.VorlNr = v.VorlNr
                          and h.MatrNr = s.MatrNr));
```

Alternativ:

```
select * from StudentenGF s
where
(select count(*)
from Vorlesungen v, ProfessorenF p
where v.gelesenVon = p.PersNr and p.FakName = s.FakName
   )
=
(select count(*)
from hoeren h, Vorlesungen v, ProfessorenF p
where h.MatrNr = s.MatrNr and h.VorlNr = v.VorlNr and p
   .PersNr = v.gelesenVon and p.FakName = s.FakName)
```