Datenmodellierung

VU 184.685, WS 2015

Relationale Entwurfstheorie

Nysret Musliu, Sebastian Skritek

Institut für Informationssysteme
Technische Universität Wien

FAKULTÄT
FÜR INFORMATIK

Faculty of Informatics

Acknowledgments

Die Folien sind eine kleine Erweiterung der Folien von Katrin Seyr.

Überblick

Ziele

- Finetuning des relationalen Schemas
- Qualität eines Relationenschemas:
  - Einhaltung von Konsistenzbedingungen
  - Vermeidung von Redundanzen
- Bisher: Modellierung mittels Constraints
- Nun: zusätzlich Modellierung mittels Datenabhängigkeiten
  - Funktionale Abhängigkeiten
  - Inklusionsabhängigkeiten
  - Verbundabhängigkeiten

Ziele

- Grundlage: Funktionale Abhängigkeiten (FDs)
  - Definition
  - Schlüssel
  - Bestimmung
  - Hülle
  - Kanonische Überdeckung
- Normalformen als Gütekriterium
- Ggfls. Verbesserung eines Relationenschemas
  - Synthesealgorithmus
  - Dekomposition

Funktionale Abhängigkeiten

- Schema $\mathcal{R} = \{A, B, C, D, \ldots, H\}$
- Attribute: A, B, C, ... Attributmengen: α, β, ...
- Relation: R, Tupel: r, s, t, ... Projektion: $r.\alpha, t.\beta, \ldots$

Funktionale Abhängigkeiten

- Schema $\mathcal{R} = \{A, B, C, D, \ldots, H\}$
- Attribute: $A, B, C, \ldots$ Attributmengen: $\alpha, \beta, \ldots$
- Relation: R, Tupel: $r, s, t, \ldots$ Projektion: $r.\alpha, t.\beta, \ldots$

Definition (Funktionale Abhängigkeit)

Seien $\alpha \subseteq \mathcal{R}, \beta \subseteq \mathcal{R}$.
Eine Relation R erfüllt eine funktionale Abhängigkeit (FD) $\alpha \rightarrow \beta$ genau dann, wenn für alle Tupel $r, t \in R$ mit $r.\alpha = t.\alpha$ gilt: $r.\beta = t.\beta$.

Funktionale Abhängigkeiten

- Schema $\mathcal{R} = \{A, B, C, D, \ldots, H\}$
- Attribute: $A, B, C, \ldots$ Attributmengen: $\alpha, \beta, \ldots$
- Relation: R, Tupel: $r, s, t, \ldots$ Projektion: $r.\alpha, t.\beta, \ldots$

Definition (Funktionale Abhängigkeit)

Seien $\alpha \subseteq \mathcal{R}, \beta \subseteq \mathcal{R}$.
Eine Relation R erfüllt eine funktionale Abhängigkeit (FD) $\alpha \rightarrow \beta$ genau dann, wenn für alle Tupel $r, t \in R$ mit $r.\alpha = t.\alpha$ gilt: $r.\beta = t.\beta$.

$\alpha \rightarrow \beta$: wenn zwei Tupel die gleichen Werte für alle Attribute in α haben, so haben sie die gleichen Werte auch für alle Attribute in β.

Funktionale Abhängigkeiten

- Schema $\mathcal{R} = \{A, B, C, D, \dots, H\}$
- Attribute: $A, B, C, \dots$ Attributmengen: $\alpha, \beta, \dots$
- Relation: R, Tupel: $r, s, t, \dots$ Projektion: $r.\alpha, t.\beta, \dots$

Definition (Funktionale Abhängigkeit)

Seien $\alpha \subseteq \mathcal{R}, \beta \subseteq \mathcal{R}$.
Eine Relation R erfüllt eine funktionale Abhängigkeit (FD) $\alpha \rightarrow \beta$ genau dann, wenn für alle Tupel $r, t \in R$ mit $r.\alpha = t.\alpha$ gilt: $r.\beta = t.\beta$.

$\alpha \rightarrow \beta$: wenn zwei Tupel die gleichen Werte für alle Attribute in α haben, so haben sie die gleichen Werte auch für alle Attribute in β.

"Die α-Werte bestimmen die β-Werte eindeutig (funktional)"

Funktionale Abhängigkeiten

Beispiel

- Ein Familienstammbaum impliziert folgende FDs:
  - Kind $\rightarrow$ Vater, Mutter

Stammbaum

Kind	Mutter	Vater	Oma	Opa
Sofie	Sabine	Alfons	Linde	Lothar
Sofie	Sabine	Alfons	Lisa	Hubert
Niklas	Sabine	Alfons	Linde	Lothar
Niklas	Sabine	Alfons	Lisa	Hubert

Funktionale Abhängigkeiten

Beispiel

- Ein Familienstammbaum impliziert folgende FDs:
  - Kind $\rightarrow$ Vater, Mutter
  - Kind, Oma $\rightarrow$ Opa

Stammbaum

Kind	Mutter	Vater	Oma	Opa
Sofie	Sabine	Alfons	Linde	Lothar
Sofie	Sabine	Alfons	Lisa	Hubert
Niklas	Sabine	Alfons	Linde	Lothar
Niklas	Sabine	Alfons	Lisa	Hubert

Funktionale Abhängigkeiten

Beispiel

- Ein Familienstammbaum impliziert folgende FDs:
  - Kind $\rightarrow$ Vater, Mutter
  - Kind, Oma $\rightarrow$ Opa
  - Kind, Opa $\rightarrow$ Oma

Stammbaum

Kind	Mutter	Vater	Oma	Opa
Sofie	Sabine	Alfons	Linde	Lothar
Sofie	Sabine	Alfons	Lisa	Hubert
Niklas	Sabine	Alfons	Linde	Lothar
Niklas	Sabine	Alfons	Lisa	Hubert

Funktionale Abhängigkeiten

Beispiel

- Ein Familienstammbaum impliziert folgende FDs:
  - Kind $\rightarrow$ Vater, Mutter
  - Kind, Oma $\rightarrow$ Opa
  - Kind, Opa $\rightarrow$ Oma

Stammbaum

Kind	Mutter	Vater	Oma	Opa
Sofie	Sabine	Alfons	Linde	Lothar
Sofie	Sabine	Alfons	Lisa	Hubert
Niklas	Sabine	Alfons	Linde	Lothar
Niklas	Sabine	Alfons	Lisa	Hubert
Sofie	...	...	Linde	Willi

Funktionale Abhängigkeiten

Beispiel

- Schema $\mathcal{R} = \{A, B, C, D\}$ der Relation R
- Frage: sind folgende FDs auf der Relation R gültig oder nicht:

$$\{A\} \rightarrow \{B\}$$
$$\{C, D\} \rightarrow \{B\}$$
$$\{B\} \rightarrow \{C\}$$
$$\{A, B\} \rightarrow \{C\}$$
$$\{B, C\} \rightarrow \{A\}$$
$$\{B\} \rightarrow \{A\}$$

R			
A	*B*	*C*	*D*
a4	b2	c4	d3
a1	b1	c1	d1
a1	b1	c1	d2
a2	b2	c3	d2
a3	b2	c4	d3

Funktionale Abhängigkeiten

Beispiel

- Schema $\mathcal{R} = \{A, B, C, D\}$ der Relation R
- Frage: sind folgende FDs auf der Relation R gültig oder nicht:

$\{A\} \rightarrow \{B\}$ ja

$\{C, D\} \rightarrow \{B\}$

$\{B\} \rightarrow \{C\}$

$\{A, B\} \rightarrow \{C\}$

$\{B, C\} \rightarrow \{A\}$

$\{B\} \rightarrow \{A\}$

R			
A	*B*	*C*	*D*
a4	b2	c4	d3
a1	b1	c1	d1
a1	b1	c1	d2
a2	b2	c3	d2
a3	b2	c4	d3

Funktionale Abhängigkeiten

Beispiel

- Schema $\mathcal{R} = \{A, B, C, D\}$ der Relation R
- Frage: sind folgende FDs auf der Relation R gültig oder nicht:

FD	
$\{A\} \rightarrow \{B\}$	ja
$\{C, D\} \rightarrow \{B\}$	ja
$\{B\} \rightarrow \{C\}$	
$\{A, B\} \rightarrow \{C\}$	
$\{B, C\} \rightarrow \{A\}$	
$\{B\} \rightarrow \{A\}$	

R			
A	B	C	D
a4	b2	c4	d3
a1	b1	c1	d1
a1	b1	c1	d2
a2	b2	c3	d2
a3	b2	c4	d3

Funktionale Abhängigkeiten

Beispiel

- Schema $\mathcal{R} = \{A, B, C, D\}$ der Relation R
- Frage: sind folgende FDs auf der Relation R gültig oder nicht:

$\{A\} \rightarrow \{B\}$	ja
$\{C, D\} \rightarrow \{B\}$	ja
$\{B\} \rightarrow \{C\}$	nein
$\{A, B\} \rightarrow \{C\}$	
$\{B, C\} \rightarrow \{A\}$	
$\{B\} \rightarrow \{A\}$	

R			
A	*B*	*C*	*D*
a4	b2	c4	d3
a1	b1	c1	d1
a1	b1	c1	d2
a2	b2	c3	d2
a3	b2	c4	d3

Funktionale Abhängigkeiten

Beispiel

- Schema $\mathcal{R} = \{A, B, C, D\}$ der Relation R
- Frage: sind folgende FDs auf der Relation R gültig oder nicht:

$\{A\} \rightarrow \{B\}$	ja
$\{C, D\} \rightarrow \{B\}$	ja
$\{B\} \rightarrow \{C\}$	nein
$\{A, B\} \rightarrow \{C\}$	ja
$\{B, C\} \rightarrow \{A\}$	
$\{B\} \rightarrow \{A\}$	

R

A	*B*	*C*	*D*
a4	b2	c4	d3
a1	b1	c1	d1
a1	b1	c1	d2
a2	b2	c3	d2
a3	b2	c4	d3

Funktionale Abhängigkeiten

Beispiel

- Schema $\mathcal{R} = \{A, B, C, D\}$ der Relation R
- Frage: sind folgende FDs auf der Relation R gültig oder nicht:

FD	
$\{A\} \to \{B\}$	ja
$\{C, D\} \to \{B\}$	ja
$\{B\} \to \{C\}$	nein
$\{A, B\} \to \{C\}$	ja
$\{B, C\} \to \{A\}$	nein
$\{B\} \to \{A\}$	

R

A	*B*	*C*	*D*
a4	b2	c4	d3
a1	b1	c1	d1
a1	b1	c1	d2
a2	b2	c3	d2
a3	b2	c4	d3

Funktionale Abhängigkeiten

Beispiel

- Schema $\mathcal{R} = \{A, B, C, D\}$ der Relation R
- Frage: sind folgende FDs auf der Relation R gültig oder nicht:

FD	
$\{A\} \rightarrow \{B\}$	ja
$\{C, D\} \rightarrow \{B\}$	ja
$\{B\} \rightarrow \{C\}$	nein
$\{A, B\} \rightarrow \{C\}$	ja
$\{B, C\} \rightarrow \{A\}$	nein
$\{B\} \rightarrow \{A\}$	nein

R			
A	*B*	*C*	*D*
a4	b2	c4	d3
a1	b1	c1	d1
a1	b1	c1	d2
a2	b2	c3	d2
a3	b2	c4	d3

Einhaltung einer FD

Gesucht: Vorgehen zur Überprüfung, ob eine gegebene Relation R die FD $\alpha \rightarrow \beta$ erfüllt:

Einhaltung einer FD

Gesucht: Vorgehen zur Überprüfung, ob eine gegebene Relation R die FD $\alpha \to \beta$ erfüllt:

- Die folgende SQL-Abfrage liefert ein leeres Ergebnis:

```
select * from R r1, R r2
where r1.α = r2.α and r1.β ≠ r2.β;
```

Einhaltung einer FD

Gesucht: Vorgehen zur Überprüfung, ob eine gegebene Relation R die FD $\alpha \rightarrow \beta$ erfüllt:

- Die folgende SQL-Abfrage liefert ein leeres Ergebnis:

```
select * from R r1, R r2
where r1.α = r2.α and r1.β ≠ r2.β;
```

```
select * from R r1 where exists(
  select * from R r2 where r1.α = r2.α
                       and r1.β ≠ r2.β);
```

Einhaltung einer FD

Gesucht: Vorgehen zur Überprüfung, ob eine gegebene Relation R die FD $\alpha \to \beta$ erfüllt:

- Die folgende SQL-Abfrage liefert ein leeres Ergebnis:

```
select * from R r1, R r2
where r1.α = r2.α and r1.β ≠ r2.β;
```

```
select * from R r1 where exists(
  select * from R r2 where r1.α = r2.α
                      and r1.β ≠ r2.β);
```

- Für alle möglichen Werte c enhält das Ergebnis der Abfrage

$$\pi_\beta(\sigma_{\alpha=c}(R))$$

maximal ein Tupel

Einhaltung einer FD

Algorithmus zur Überprüfung, ob eine gegebene Relation R die FD $\alpha \rightarrow \beta$ erfüllt:

Input: $(R, \alpha \rightarrow \beta)$: Relation R und eine FD $\alpha \rightarrow \beta$

Output: ja, wenn FD erfüllt, nein sonst

Einhaltung $(R, \alpha \rightarrow \beta)$

- sortiere R nach den Werten von α
- falls alle Gruppen von Tupeln mit gleichen α Werten die gleichen β Werte aufweisen: output(ja) sonst output(nein)

Einhaltung einer FD

Algorithmus zur Überprüfung, ob eine gegebene
Relation R die FD $\alpha \rightarrow \beta$ erfüllt:

Input: $(R, \alpha \rightarrow \beta)$: Relation R und eine FD $\alpha \rightarrow \beta$

Output: ja, wenn FD erfüllt, nein sonst

Einhaltung $(R, \alpha \rightarrow \beta)$

- sortiere R nach den Werten von α
- falls alle Gruppen von Tupeln mit gleichen α Werten die gleichen β Werte aufweisen: output(ja) sonst output(nein)

Laufzeit von Einhaltung ist bestimmt durch Aufwand für die Sortierung - $O(nlogn)$

Bestimmung von FDs

Beispiel

Gegeben: Information über Professoren anhand folgender Attribute: Professoren:{[PersNr, Name, Rang, Raum, Ort, Straße, PLZ, Vorwahl, BLand, EW, Landesregierung]}
Frage: Welche funktionalen Abhängigkeiten lassen sich aufgrund der Semantik der zu modellierenden Miniwelt finden?

Bestimmung von FDs

Beispiel

Gegeben: Information über Professoren anhand folgender Attribute:
Professoren:{[PersNr, Name, Rang, Raum, Ort, Straße, PLZ, Vorwahl, BLand, EW, Landesregierung]}
Frage: Welche funktionalen Abhängigkeiten lassen sich aufgrund der Semantik der zu modellierenden Miniwelt finden?

- PersNr ist ein Kandidatenschlüssel:
  {PersNr} $\rightarrow$ {PersNr, Name, Rang, Raum, Ort, Straße, PLZ, Vorwahl, BLand, EW, Landesregierung}

Bestimmung von FDs

Beispiel

Gegeben: Information über Professoren anhand folgender Attribute:
Professoren:{[PersNr, Name, Rang, Raum, Ort, Straße, PLZ, Vorwahl, BLand, EW, Landesregierung]}
Frage: Welche funktionalen Abhängigkeiten lassen sich aufgrund der Semantik der zu modellierenden Miniwelt finden?

- PersNr ist ein Kandidatenschlüssel:
  {PersNr} $\rightarrow$ {PersNr, Name, Rang, Raum, Ort, Straße, PLZ, Vorwahl, BLand, EW, Landesregierung}
- Ortsnamen sind innerhalb eines Bundeslandes eindeutig:
  {Ort, BLand} $\rightarrow$ {EW, Vorwahl}

Bestimmung von FDs

Beispiel

- Die Postleitzahl identifiziert einen Ort, das Bundesland und die Einwohnerzahl:
  $\{\text{PLZ}\} \rightarrow \{\text{BLand, Ort, EW}\}$

Bestimmung von FDs

Beispiel

- Die Postleitzahl identifiziert einen Ort, das Bundesland und die Einwohnerzahl:
  $\{\text{PLZ}\} \rightarrow \{\text{BLand, Ort, EW}\}$
- Die Postleitzahl ändert sich innerhalb der Straße eines Ortes nicht:
  $\{\text{BLand, Ort, Straße}\} \rightarrow \{\text{PLZ}\}$

Bestimmung von FDs

Beispiel

- Die Postleitzahl identifiziert einen Ort, das Bundesland und die Einwohnerzahl:
  $\{\text{PLZ}\} \to \{\text{BLand, Ort, EW}\}$
- Die Postleitzahl ändert sich innerhalb der Straße eines Ortes nicht:
  $\{\text{BLand, Ort, Straße}\} \to \{\text{PLZ}\}$
- Landesregierung speichert die Partei des Ministerpräsidenten:
  $\{\text{BLand}\} \to \{\text{Landesregierung}\}$

Bestimmung von FDs

Beispiel

- Die Postleitzahl identifiziert einen Ort, das Bundesland und die Einwohnerzahl:
  $\{\text{PLZ}\} \rightarrow \{\text{BLand, Ort, EW}\}$
- Die Postleitzahl ändert sich innerhalb der Straße eines Ortes nicht:
  $\{\text{BLand, Ort, Straße}\} \rightarrow \{\text{PLZ}\}$
- Landesregierung speichert die Partei des Ministerpräsidenten:
  $\{\text{BLand}\} \rightarrow \{\text{Landesregierung}\}$
- In einem Raum kann nur ein Professor sitzen:
  $\{\text{Raum}\} \rightarrow \{\text{PersNr}\}$

Bestimmung von FDs

Gegeben: Menge F von FDs.

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name

Bestimmung von FDs

Gegeben: Menge F von FDs.

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name

Frage 1: Welche weiteren FDs können daraus abgeleitet werden?

Bestimmung von FDs

Gegeben: Menge F von FDs.

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name

Frage 1: Welche weiteren FDs können daraus abgeleitet werden?

Frage 2: Gegeben zusätzlich eine Menge γ an Attributen, welche Attribute kann ich aus γ aufgrund von F herleiten?

Hülle einer Attributmenge

Gegeben: Eine Menge γ von Attributen und F eine Menge von FDs.

Frage: Welche Attribute kann ich aus γ aufgrund von F herleiten?

Beispiel: $\{\text{Raum} \rightarrow \text{PersNr}, \text{PersNr} \rightarrow \text{Name}\}, \{\text{Raum}\}$
$\Rightarrow \{\text{PersNr}, \text{Name}\}$

Hülle einer Attributmenge

Gegeben: Eine Menge γ von Attributen und F eine Menge von FDs.

Frage: Welche Attribute kann ich aus γ aufgrund von F herleiten?

Beispiel: {Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name}, {Raum} $\Rightarrow$ {PersNr, Name}

Definition (Hülle einer Attributmenge)

Die Menge γ^+ der Attribute, die aus einer Menge γ von Attributen aufgrund einer Menge von FDs F hergeleitet werden können, nennt man die Hülle der Attributmenge γ.

Hülle einer Attributmenge

Berechnung mittels Algorithmus AttrHülle:

Input: (F, γ): Menge F von FDs und Menge γ von Attributen

Output: Menge der Attribute γ^+.

AttrHülle (F, γ)

$\gamma^+ = \gamma$

while $\exists(\alpha \to \beta) \in F$ mit $\alpha \subseteq \gamma^+$ und $\beta \nsubseteq \gamma^+$ **do**

$\quad \gamma^+ := \gamma^+ \cup \beta$

return(γ^+)

Hülle einer Attributmenge

Beispiel

Sei $F = \{RS \rightarrow T, U \rightarrow VX, RX \rightarrow W, T \rightarrow RU\}$

AttrHülle$(F, \{T\})$:

AttrHülle$(F, \{RS\})$:

Hülle einer Attributmenge

Beispiel

Sei $F = \{RS \rightarrow T, U \rightarrow VX, RX \rightarrow W, T \rightarrow RU\}$

AttrHülle$(F, \{T\})$: $\{R, T, U, V, W, X\}$

AttrHülle$(F, \{RS\})$:

Hülle einer Attributmenge

Beispiel

Sei $F = \{RS \to T, U \to VX, RX \to W, T \to RU\}$

AttrHülle$(F, \{T\})$: $\{R, T, U, V, W, X\}$

AttrHülle$(F, \{RS\})$: $\{R, S, T, U, V, W, X\}$

Schlüssel

Definition (Schlüssel)

$\gamma \subseteq \mathcal{R}$ ist ein Kandidatenschlüssel oder Schlüssel, falls die beiden folgenden Bedingungen erfüllt sind:

Schlüssel

Definition (Schlüssel)

$\gamma \subseteq \mathcal{R}$ ist ein Kandidatenschlüssel oder Schlüssel, falls die beiden folgenden Bedingungen erfüllt sind:

1. $\gamma \rightarrow \mathcal{R}$
2. γ ist minimal, d.h. für alle $A \in \gamma$: $(\gamma - \{A\}) \not\rightarrow \mathcal{R}$

Schlüssel

Definition (Schlüssel)

$\gamma \subseteq \mathcal{R}$ ist ein Kandidatenschlüssel oder Schlüssel, falls die beiden folgenden Bedingungen erfüllt sind:

1. $\gamma \to \mathcal{R}$
2. γ ist minimal, d.h. für alle $A \in \gamma$: $(\gamma - \{A\}) \not\to \mathcal{R}$

Im Relationenmodell: Auszeichnung eines Schlüssels als Primärschlüssel zur Verknüpfung von Tabellen mittels Primär- und Fremdschlüssel

Schlüssel

Definition (Schlüssel)

$\gamma \subseteq \mathcal{R}$ ist ein Kandidatenschlüssel oder Schlüssel, falls die beiden folgenden Bedingungen erfüllt sind:

1. $\gamma \to \mathcal{R}$
2. γ ist minimal, d.h. für alle $A \in \gamma$: $(\gamma - \{A\}) \not\to \mathcal{R}$

Im Relationenmodell: Auszeichnung eines Schlüssels als Primärschlüssel zur Verknüpfung von Tabellen mittels Primär- und Fremdschlüssel

Definition (Superschlüssel)

$\gamma \subseteq \mathcal{R}$ ist ein Superschlüssel (Oberschlüssel), falls $\gamma \to \mathcal{R}$

- keine Minimalität bei Superschlüsseln
- ein Superschlüssel ist eine Obermenge eines Schlüssels

Schlüsselbestimmung

Beispiel

Eine Stadt wird beschrieben durch ihren Namen (Name), das Bundesland (BLand) in dem sie liegt, die Vorwahl (VW) und die Einwohnerzahl (EW).

Frage: Welche FDs gelten in diesem Szenario?

Schlüsselbestimmung

Beispiel

Eine Stadt wird beschrieben durch ihren Namen (Name), das Bundesland (BLand) in dem sie liegt, die Vorwahl (VW) und die Einwohnerzahl (EW).
Frage: Welche FDs gelten in diesem Szenario?

- Ortsnamen sind innerhalb eines Bundeslandes eindeutig d.h. $\{$Name, BLand$\} \rightarrow$ VW, EW

Schlüsselbestimmung

Beispiel

Eine Stadt wird beschrieben durch ihren Namen (Name), das Bundesland (BLand) in dem sie liegt, die Vorwahl (VW) und die Einwohnerzahl (EW).

Frage: Welche FDs gelten in diesem Szenario?

- Ortsnamen sind innerhalb eines Bundeslandes eindeutig d.h. $\{\text{Name}, \text{BLand}\} \rightarrow \text{VW}, \text{EW}$
- Mehrere Städte können dieselbe Vorwahl haben, aber nur, wenn sie einen unterschiedlichen Namen haben d.h. $\{\text{Name}, \text{VW}\} \rightarrow \text{BLand}, \text{EW}$

Schlüsselbestimmung

Beispiel

Eine Stadt wird beschrieben durch ihren Namen (Name), das Bundesland (BLand) in dem sie liegt, die Vorwahl (VW) und die Einwohnerzahl (EW).

Frage: Welche FDs gelten in diesem Szenario?

- Ortsnamen sind innerhalb eines Bundeslandes eindeutig d.h. $\{$Name, BLand$\} \rightarrow$ VW, EW
- Mehrere Städte können dieselbe Vorwahl haben, aber nur, wenn sie einen unterschiedlichen Namen haben d.h. $\{$Name, VW$\} \rightarrow$ BLand, EW

Frage: Welche Kandidatenschlüssel ergeben sich aufgrund der gefundenen FDs?

Schlüsselbestimmung

Beispiel

Eine Stadt wird beschrieben durch ihren Namen (Name), das Bundesland (BLand) in dem sie liegt, die Vorwahl (VW) und die Einwohnerzahl (EW).

Frage: Welche FDs gelten in diesem Szenario?

- Ortsnamen sind innerhalb eines Bundeslandes eindeutig d.h. $\{$Name, BLand$\} \rightarrow$ VW, EW
- Mehrere Städte können dieselbe Vorwahl haben, aber nur, wenn sie einen unterschiedlichen Namen haben d.h. $\{$Name, VW$\} \rightarrow$ BLand, EW

Frage: Welche Kandidatenschlüssel ergeben sich aufgrund der gefundenen FDs?

- $\{$Name, BLand$\}$

Schlüsselbestimmung

Beispiel

Eine Stadt wird beschrieben durch ihren Namen (Name), das Bundesland (BLand) in dem sie liegt, die Vorwahl (VW) und die Einwohnerzahl (EW).

Frage: Welche FDs gelten in diesem Szenario?

- Ortsnamen sind innerhalb eines Bundeslandes eindeutig d.h. $\{\text{Name, BLand}\} \rightarrow \text{VW, EW}$
- Mehrere Städte können dieselbe Vorwahl haben, aber nur, wenn sie einen unterschiedlichen Namen haben d.h. $\{\text{Name, VW}\} \rightarrow \text{BLand, EW}$

Frage: Welche Kandidatenschlüssel ergeben sich aufgrund der gefundenen FDs?

- $\{\text{Name, BLand}\}$
- $\{\text{Name, VW}\}$

Schlüsselbestimmung mittels AttrHülle

Problem: Suche nach Kandidatenschlüsseln einer Relation R aufgrund der vorhandenen FDs.

Definition (Schlüssel - Wiederholung)

$\gamma \subseteq \mathcal{R}$ ist ein Kandidatenschlüssel oder Schlüssel, falls die beiden folgenden Bedingungen erfüllt sind:

1. $\gamma \rightarrow \mathcal{R}$
2. γ ist minimal, d.h. $\forall A \in \gamma : (\gamma - \{A\}) \not\rightarrow \mathcal{R}$

Schlüsselbestimmung mittels AttrHülle

Problem: Suche nach Kandidatenschlüsseln einer Relation R aufgrund der vorhandenen FDs.

Definition (Schlüssel - Wiederholung)

$\gamma \subseteq \mathcal{R}$ ist ein Kandidatenschlüssel oder Schlüssel, falls die beiden folgenden Bedingungen erfüllt sind:

1. $\gamma \rightarrow \mathcal{R}$
2. γ ist minimal, d.h. $\forall A \in \gamma : (\gamma - \{A\}) \not\rightarrow \mathcal{R}$

1. $\gamma \rightarrow \mathcal{R}$, wenn $\text{AttrHülle}(F, \gamma) = \mathcal{R}$

Schlüsselbestimmung mittels AttrHülle

Problem: Suche nach Kandidatenschlüsseln einer Relation R aufgrund der vorhandenen FDs.

Definition (Schlüssel - Wiederholung)

$\gamma \subseteq \mathcal{R}$ ist ein Kandidatenschlüssel oder Schlüssel, falls die beiden folgenden Bedingungen erfüllt sind:

1. $\gamma \rightarrow \mathcal{R}$
2. γ ist minimal, d.h. $\forall A \in \gamma : (\gamma - \{A\}) \not\rightarrow \mathcal{R}$

1. $\gamma \rightarrow \mathcal{R}$, wenn AttrHülle$(F, \gamma) = \mathcal{R}$
2. die Minimalität ist erfüllt, wenn für jedes Attribut A aus γ gilt: AttrHülle$(F, \gamma - \{A\}) \neq \mathcal{R}$

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \rightarrow BDAE\}$

- Zeitraubende Vorgehensweise: Durchprobieren aller einelementigen, aller zweielementigen, aller dreielementigen Schlüsselkandidaten mittels AttrHülle

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \rightarrow BDAE\}$

- Zeitraubende Vorgehensweise: Durchprobieren aller einelementigen, aller zweielementigen, aller dreielementigen Schlüsselkandidaten mittels AttrHülle
- Alternativ: Verwendung der folgenden Heuristik: Alle Attribute, die nicht auf der rechten Seite vorkommen, können mittels AttrHülle nicht hergeleitet werden und müssen daher im Schlüssel enthalten sein.

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \rightarrow BDAE\}$

- Zeitraubende Vorgehensweise: Durchprobieren aller einelementigen, aller zweielementigen, aller dreielementigen Schlüsselkandidaten mittels AttrHülle
- Alternativ: Verwendung der folgenden Heuristik: Alle Attribute, die nicht auf der rechten Seite vorkommen, können mittels AttrHülle nicht hergeleitet werden und müssen daher im Schlüssel enthalten sein.
- Hier: C und F kommen rechts nicht vor, daher folgender Versuch:

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \to BDAE\}$

- Zeitraubende Vorgehensweise: Durchprobieren aller einelementigen, aller zweielementigen, aller dreielementigen Schlüsselkandidaten mittels AttrHülle
- Alternativ: Verwendung der folgenden Heuristik: Alle Attribute, die nicht auf der rechten Seite vorkommen, können mittels AttrHülle nicht hergeleitet werden und müssen daher im Schlüssel enthalten sein.
- Hier: C und F kommen rechts nicht vor, daher folgender Versuch: AttrHülle$(\{C \to BDAE\}, CF) = \{C, F, B, D, A, E\}$

$$\Rightarrow$$

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \rightarrow BDAE\}$

- Zeitraubende Vorgehensweise: Durchprobieren aller einelementigen, aller zweielementigen, aller dreielementigen Schlüsselkandidaten mittels AttrHülle
- Alternativ: Verwendung der folgenden Heuristik: Alle Attribute, die nicht auf der rechten Seite vorkommen, können mittels AttrHülle nicht hergeleitet werden und müssen daher im Schlüssel enthalten sein.
- Hier: C und F kommen rechts nicht vor, daher folgender Versuch: AttrHülle$(\{C \rightarrow BDAE\}, CF) = \{C, F, B, D, A, E\}$

$$\Rightarrow$$

CF ist Schlüssel von $\mathcal{R}$

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \rightarrow BD, D \rightarrow AE, E \rightarrow CF, F \rightarrow E\}$
Heuristik hier nicht zielführend, da alle Attribute hergeleitet werden können.

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \to BD, D \to AE, E \to CF, F \to E\}$
Heuristik hier nicht zielführend, da alle Attribute hergeleitet werden können.

- C: AttrHülle$(F_d, C) = \{C, B, D, A, E, F\} \Rightarrow C$ ist Schlüssel von $\mathcal{R}$

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \rightarrow BD, D \rightarrow AE, E \rightarrow CF, F \rightarrow E\}$
Heuristik hier nicht zielführend, da alle Attribute hergeleitet werden können.

C: AttrHülle$(F_d, C) = \{C, B, D, A, E, F\} \Rightarrow C$ ist Schlüssel von $\mathcal{R}$
Achtung: C kann aus E hergeleitet werden also:

E: AttrHülle$(F_d, E) = \{E, C, B, D, A, F\} \Rightarrow E$ ist auch Schlüssel von $\mathcal{R}$

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \to BD, D \to AE, E \to CF, F \to E\}$
Heuristik hier nicht zielführend, da alle Attribute hergeleitet werden können.

- C: AttrHülle$(F_d, C) = \{C, B, D, A, E, F\} \Rightarrow C$ ist Schlüssel von $\mathcal{R}$
  Achtung: C kann aus E hergeleitet werden also:
- E: AttrHülle$(F_d, E) = \{E, C, B, D, A, F\} \Rightarrow E$ ist auch Schlüssel von $\mathcal{R}$
  Achtung: E kann aus D oder F hergeleitet werden also:
- D: AttrHülle$(F_d, D) = \{D, A, E, C, F, B\} \Rightarrow D$ ist Schlüssel von $\mathcal{R}$,
- F: AttrHülle$(F_d, F) = \{F, E, C, B, D, A\} \Rightarrow F$ ist Schlüssel von $\mathcal{R}$

Schlüsselbestimmung mittels AttrHülle

Beispiel

$\mathcal{R} = \{ABCDEF\}, F_d = \{C \rightarrow BD, D \rightarrow AE, E \rightarrow CF, F \rightarrow E\}$
Heuristik hier nicht zielführend, da alle Attribute hergeleitet werden können.

- C: AttrHülle$(F_d, C) = \{C, B, D, A, E, F\} \Rightarrow C$ ist Schlüssel von $\mathcal{R}$
  Achtung: C kann aus E hergeleitet werden also:
- E: AttrHülle$(F_d, E) = \{E, C, B, D, A, F\} \Rightarrow E$ ist auch Schlüssel von $\mathcal{R}$
  Achtung: E kann aus D oder F hergeleitet werden also:
- D: AttrHülle$(F_d, D) = \{D, A, E, C, F, B\} \Rightarrow D$ ist Schlüssel von $\mathcal{R}$,
- F: AttrHülle$(F_d, F) = \{F, E, C, B, D, A\} \Rightarrow F$ ist Schlüssel von $\mathcal{R}$

Hülle von FDs

Gegeben: Menge F von FDs.

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name

Hülle von FDs

Gegeben: Menge F von FDs.

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name

Frage 1: Welche weiteren FDs können daraus abgeleitet werden?

Hülle von FDs

Gegeben: Menge F von FDs.

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name

Frage 1: Welche weiteren FDs können daraus abgeleitet werden?

Frage 2: Gegeben zusätzlich eine Menge γ an Attributen, welche Attribute kann ich aus γ aufgrund von F herleiten?

Hülle von FDs

Problem: F Menge von FDs. Welche weiteren FDs können daraus abgeleitet werden?

Hülle von FDs

Problem: F Menge von FDs. Welche weiteren FDs können daraus abgeleitet werden?

Beispiel: Raum $\to$ PersNr, PersNr $\to$ Name $\Rightarrow$ Raum $\to$ Name

Hülle von FDs

Problem: F Menge von FDs. Welche weiteren FDs können daraus abgeleitet werden?

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name $\Rightarrow$ Raum $\rightarrow$ Name

Definition (Hülle von FDs)

Die Menge aller aus F ableitbaren FDs wird Hülle F^+ von F genannt.

Hülle von FDs

Problem: F Menge von FDs. Welche weiteren FDs können daraus abgeleitet werden?

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name $\Rightarrow$ Raum $\rightarrow$ Name

Definition (Hülle von FDs)

Die Menge aller aus F ableitbaren FDs wird Hülle F^+ von F genannt.

Verständnishilfe aus der Mathematik: V eine Menge von Vektoren. Die Menge aller Vektoren, die aus V mittels Linearkombinationen erhalten werden: lineare Hülle von V.

Hülle von FDs

Problem: F Menge von FDs. Welche weiteren FDs können daraus abgeleitet werden?

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name $\Rightarrow$ Raum $\rightarrow$ Name

Definition (Hülle von FDs)

Die Menge aller aus F ableitbaren FDs wird Hülle F^+ von F genannt.

Verständnishilfe aus der Mathematik: V eine Menge von Vektoren. Die Menge aller Vektoren, die aus V mittels Linearkombinationen erhalten werden: lineare Hülle von V.

Berechnung der Hülle F^+ von F mittels der Armstrong Axiome (1974).

Hülle von FDs

Problem: F Menge von FDs. Welche weiteren FDs können daraus abgeleitet werden?

Beispiel: Raum $\rightarrow$ PersNr, PersNr $\rightarrow$ Name $\Rightarrow$ Raum $\rightarrow$ Name

Definition (Hülle von FDs)

Die Menge aller aus F ableitbaren FDs wird Hülle F^+ von F genannt.

Verständnishilfe aus der Mathematik: V eine Menge von Vektoren. Die Menge aller Vektoren, die aus V mittels Linearkombinationen erhalten werden: lineare Hülle von V.

Berechnung der Hülle F^+ von F mittels der Armstrong Axiome (1974).

Theorem

Die Armstrong Axiome sind vollständig (erzeugen alle implizierten FDs) und korrekt (erzeugen nur gültige FDs)

Die Armstrong Axiome

Reflexivität: Sei β eine Teilmenge von α ($\beta \subseteq \alpha$) dann gilt immer $\alpha \to \beta$. Insbesondere gilt $\alpha \to \alpha$.

Die Armstrong Axiome

Reflexivität: Sei β eine Teilmenge von α ($\beta \subseteq \alpha$) dann gilt immer $\alpha \to \beta$. Insbesondere gilt $\alpha \to \alpha$.

Verstärkung: Falls $\alpha \to \beta$ gilt, dann gilt auch $\alpha\gamma \to \beta\gamma$.

Die Armstrong Axiome

Reflexivität: Sei β eine Teilmenge von α ($\beta \subseteq \alpha$) dann gilt immer $\alpha \rightarrow \beta$. Insbesondere gilt $\alpha \rightarrow \alpha$.

Verstärkung: Falls $\alpha \rightarrow \beta$ gilt, dann gilt auch $\alpha\gamma \rightarrow \beta\gamma$.

Transitivität: Falls $\alpha \rightarrow \beta$, $\beta \rightarrow \gamma$ gelten, dann gilt auch $\alpha \rightarrow \gamma$.

Die Armstrong Axiome

Reflexivität: Sei β eine Teilmenge von α ($\beta \subseteq \alpha$) dann gilt immer $\alpha \rightarrow \beta$. Insbesondere gilt $\alpha \rightarrow \alpha$.

Verstärkung: Falls $\alpha \rightarrow \beta$ gilt, dann gilt auch $\alpha\gamma \rightarrow \beta\gamma$.

Transitivität: Falls $\alpha \rightarrow \beta$, $\beta \rightarrow \gamma$ gelten, dann gilt auch $\alpha \rightarrow \gamma$.

Zusätzliche Axiome, die die Herleitung der Hülle erleichtern:

Die Armstrong Axiome

Reflexivität: Sei β eine Teilmenge von α ($\beta \subseteq \alpha$) dann gilt immer $\alpha \to \beta$. Insbesondere gilt $\alpha \to \alpha$.

Verstärkung: Falls $\alpha \to \beta$ gilt, dann gilt auch $\alpha\gamma \to \beta\gamma$.

Transitivität: Falls $\alpha \to \beta$, $\beta \to \gamma$ gelten, dann gilt auch $\alpha \to \gamma$.

Zusätzliche Axiome, die die Herleitung der Hülle erleichtern:

Vereinigung: $\alpha \to \beta$ und $\alpha \to \gamma \Rightarrow \alpha \to \beta\gamma$.

Die Armstrong Axiome

Reflexivität: Sei β eine Teilmenge von α ($\beta \subseteq \alpha$) dann gilt immer $\alpha \to \beta$. Insbesondere gilt $\alpha \to \alpha$.

Verstärkung: Falls $\alpha \to \beta$ gilt, dann gilt auch $\alpha\gamma \to \beta\gamma$.

Transitivität: Falls $\alpha \to \beta$, $\beta \to \gamma$ gelten, dann gilt auch $\alpha \to \gamma$.

Zusätzliche Axiome, die die Herleitung der Hülle erleichtern:

Vereinigung: $\alpha \to \beta$ und $\alpha \to \gamma \Rightarrow \alpha \to \beta\gamma$.

Dekomposition: $\alpha \to \beta\gamma \Rightarrow \alpha \to \beta$ und $\alpha \to \gamma$

wichtig: kann immer erreichen, dass rechts nur ein Attribut steht.

Die Armstrong Axiome

Reflexivität: Sei β eine Teilmenge von α ($\beta \subseteq \alpha$) dann gilt immer $\alpha \rightarrow \beta$. Insbesondere gilt $\alpha \rightarrow \alpha$.

Verstärkung: Falls $\alpha \rightarrow \beta$ gilt, dann gilt auch $\alpha\gamma \rightarrow \beta\gamma$.

Transitivität: Falls $\alpha \rightarrow \beta$, $\beta \rightarrow \gamma$ gelten, dann gilt auch $\alpha \rightarrow \gamma$.

Zusätzliche Axiome, die die Herleitung der Hülle erleichtern:

Vereinigung: $\alpha \rightarrow \beta$ und $\alpha \rightarrow \gamma \Rightarrow \alpha \rightarrow \beta\gamma$.

Dekomposition: $\alpha \rightarrow \beta\gamma \Rightarrow \alpha \rightarrow \beta$ und $\alpha \rightarrow \gamma$

wichtig: kann immer erreichen, dass rechts nur ein Attribut steht.

Pseudotransitivität: $\alpha \rightarrow \beta$ und $\gamma\beta \rightarrow \delta \Rightarrow \alpha\gamma \rightarrow \delta$.

Die Armstrong Axiome

Beispiel

Herleitung der FD $\{\text{PLZ}\} \rightarrow \{\text{Landesregierung}\}$ aus den restlichen FDs im Beispielschema Professoren:

Es gelten: $\{\text{PLZ}\} \rightarrow \{\text{Bland, Ort, EW}\}$ und
$\{\text{BLand}\} \rightarrow \{\text{Landesregierung}\}$

Die Armstrong Axiome

Beispiel

Herleitung der FD $\{PLZ\} \rightarrow \{Landesregierung\}$ aus den restlichen FDs im Beispielschema Professoren:

Es gelten: $\{PLZ\} \rightarrow \{Bland, Ort, EW\}$ und
$\{BLand\} \rightarrow \{Landesregierung\}$

Dekomposition von $\{PLZ\} \rightarrow \{Bland, Ort, EW\}$:

$\{PLZ\} \rightarrow \{Bland\}$, $\{PLZ\} \rightarrow \{Ort\}$, $\{PLZ\} \rightarrow \{EW\}$

Die Armstrong Axiome

Beispiel

Herleitung der FD {PLZ} $\rightarrow$ {Landesregierung} aus den restlichen FDs im Beispielschema Professoren:

Es gelten: {PLZ} $\rightarrow$ {Bland, Ort, EW} und
{BLand} $\rightarrow${Landesregierung}

Dekomposition von {PLZ} $\rightarrow$ {Bland, Ort, EW}:

{PLZ} $\rightarrow$ {Bland}, {PLZ} $\rightarrow$ {Ort}, {PLZ} $\rightarrow$ {EW}

Transitivität von {PLZ} $\rightarrow$ {BLand}, {BLand} $\rightarrow$ {Landesregierung}:

{PLZ} $\rightarrow$ {*Landesregierung*}

Äquivalenz von Mengen von FDs

Definition (Äquivalenz von FDs)

Zwei Mengen F, G von FDs sind äquivalent ($F \equiv G$), wenn sie dieselbe Hülle besitzen, d.h. $F^+ = G^+$

Mathematik: zwei Mengen von Vektoren sind "äquivalent", wenn sie denselben Vektorraum aufspannen

Äquivalenz von Mengen von FDs

Definition (Äquivalenz von FDs)

Zwei Mengen F, G von FDs sind äquivalent $(F \equiv G)$, wenn sie dieselbe Hülle besitzen, d.h. $F^+ = G^+$

Mathematik: zwei Mengen von Vektoren sind "äquivalent", wenn sie denselben Vektorraum aufspannen

Natürlich gilt: $F \equiv G$ genau dann wenn

- $F \subseteq G^+$ und
- $G \subseteq F^+$.

Äquivalenz von Mengen von FDs

Definition (Äquivalenz von FDs)

Zwei Mengen F, G von FDs sind äquivalent ($F \equiv G$), wenn sie dieselbe Hülle besitzen, d.h. $F^+ = G^+$

Mathematik: zwei Mengen von Vektoren sind "äquivalent", wenn sie denselben Vektorraum aufspannen

Problem: gesucht ist eine möglichst knappe Darstellung von FDs ("Basis")

Äquivalenz von Mengen von FDs

Definition (Äquivalenz von FDs)

Zwei Mengen F, G von FDs sind äquivalent ($F \equiv G$), wenn sie dieselbe Hülle besitzen, d.h. $F^+ = G^+$

Mathematik: zwei Mengen von Vektoren sind "äquivalent", wenn sie denselben Vektorraum aufspannen

Problem: gesucht ist eine möglichst knappe Darstellung von FDs ("Basis")

Lösung: die Kanonische Überdeckung

Mathematik: Basis eines Vektorraumes

Kanonische Überdeckung

Definition (Kanonische Überdeckung)

F_C heißt kanonische Überdeckung einer Menge von FDs F, wenn die folgenden drei Kriterien erfüllt sind:

1. $F_C^+ = F^+$ ($F_C$ist äquivalent zu F)
2. In F_C existieren keine FDs, die überflüssige Attribute enthalten
3. Jede linke Seite einer FD in F_C ist einzigartig.

Kanonische Überdeckung

Definition (Kanonische Überdeckung)

F_C heißt kanonische Überdeckung einer Menge von FDs F, wenn die folgenden drei Kriterien erfüllt sind:

1. $F_C^+ = F^+$ (F_C ist äquivalent zu F)
2. In F_C existieren keine FDs, die überflüssige Attribute enthalten
3. Jede linke Seite einer FD in F_C ist einzigartig.

Theorem

Zu jeder Menge von funktionalen Abhängigkeiten F gibt es eine kanonische Überdeckung F_C.

Mathematik: Zu jeder Menge von Vektoren, die einen Vektorraum aufspannen, gibt es eine Basis.

Berechnung Kanonische Überdeckung (1)

Die Berechnung der Kanonischen Überdechung ergibt sich direkt aus der Definition:

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite (Äquivalenz gewährleistet)

Berechnung Kanonische Überdeckung (1)

Die Berechnung der Kanonischen Überdechung ergibt sich direkt aus der Definition:

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite (Äquivalenz gewährleistet)
2. Kürze überflüssige Attribute wie folgt:

Berechnung Kanonische Überdeckung (1)

Die Berechnung der Kanonischen Überdechung ergibt sich direkt aus der Definition:

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite (Äquivalenz gewährleistet)
2. Kürze überflüssige Attribute wie folgt:
    1. Führe für jede FD $\alpha \to B \in F$ die Linksreduktion durch, also: kann aus α ein Attribut gekürzt werden, sodass das Ergebnis äquivalent zur ursprünglichen Menge von FDs bleibt?

Berechnung Kanonische Überdeckung (1)

Die Berechnung der Kanonischen Überdechung ergibt sich direkt aus der Definition:

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite (Äquivalenz gewährleistet)
2. Kürze überflüssige Attribute wie folgt:
   1. Führe für jede FD $\alpha \to B \in F$ die Linksreduktion durch, also: kann aus α ein Attribut gekürzt werden, sodass das Ergebnis äquivalent zur ursprünglichen Menge von FDs bleibt?
   2. Führe für jede (verbliebene) FD $\alpha \to B \in F$ die Rechtsreduktion durch, also: kann B gekürzt werden, sodass das Ergebnis äquivalent zur ursprünglichen Menge von FDs bleibt?

Berechnung Kanonische Überdeckung (1)

Die Berechnung der Kanonischen Überdechung ergibt sich direkt aus der Definition:

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite (Äquivalenz gewährleistet)
2. Kürze überflüssige Attribute wie folgt:
   1. Führe für jede FD $\alpha \to B \in F$ die Linksreduktion durch, also: kann aus α ein Attribut gekürzt werden, sodass das Ergebnis äquivalent zur ursprünglichen Menge von FDs bleibt?
   2. Führe für jede (verbliebene) FD $\alpha \to B \in F$ die Rechtsreduktion durch, also: kann B gekürzt werden, sodass das Ergebnis äquivalent zur ursprünglichen Menge von FDs bleibt?
3. Fasse mittels der Vereinigungsregel FDs zusammen (Äquivalenz gewährleistet).

Berechnung Kanonische Überdeckung (1)

Die Berechnung der Kanonischen Überdechung ergibt sich direkt aus der Definition:

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite (Äquivalenz gewährleistet)
2. Kürze überflüssige Attribute wie folgt:
   1. Führe für jede FD $\alpha \to B \in F$ die Linksreduktion durch, also: kann aus α ein Attribut gekürzt werden, sodass das Ergebnis äquivalent zur ursprünglichen Menge von FDs bleibt?
   2. Führe für jede (verbliebene) FD $\alpha \to B \in F$ die Rechtsreduktion durch, also: kann B gekürzt werden, sodass das Ergebnis äquivalent zur ursprünglichen Menge von FDs bleibt?
3. Fasse mittels der Vereinigungsregel FDs zusammen (Äquivalenz gewährleistet).

Damit in den Kürzungsschritten sichergestellt bleibt, dass nur äquivalente Ergebnisse erhalten werden, verwenden wir den Algorithmus AttrHülle

Berechnung Kanonische Überdeckung (2)

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite

Berechnung Kanonische Überdeckung (2)

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite
2. Kürze überflüssige Attribute wie folgt:
    1. Führe für jede FD $\alpha \to B \in F$ die Linksreduktion durch:

       $\forall A \in \alpha$: gilt $B \subseteq \text{AttrHülle}(F, \alpha - A)$ (ist A überflüssig?)
       wenn ja: ersetze in F $\alpha \to B$ durch $(\alpha - A) \to B$.

Berechnung Kanonische Überdeckung (2)

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite
2. Kürze überflüssige Attribute wie folgt:
   1. Führe für jede FD $\alpha \to B \in F$ die Linksreduktion durch:

      $\forall A \in \alpha$: gilt $B \subseteq$ AttrHülle$(F, \alpha - A)$ (ist A überflüssig?)
      wenn ja: ersetze in F $\alpha \to B$ durch $(\alpha - A) \to B$.

   2. Führe für jede (verbliebene) FD $\alpha \to B \in F$ die Rechtsreduktion durch:

      gilt $B \subseteq$ AttrHülle$(F - (\alpha \to B), \alpha)$ (ist B bzw. $\alpha \to B$ überflüssig?)
      wenn ja: streiche $\alpha \to B$

Berechnung Kanonische Überdeckung (2)

1. Zerlege alle FDs mittels Dekomposition auf der rechten Seite
2. Kürze überflüssige Attribute wie folgt:
    1. Führe für jede FD $\alpha \rightarrow B \in F$ die Linksreduktion durch:

       $\forall A \in \alpha$: gilt $B \subseteq$ AttrHülle$(F, \alpha - A)$ (ist A überflüssig?)
       wenn ja: ersetze in F $\alpha \rightarrow B$ durch $(\alpha - A) \rightarrow B$.

    2. Führe für jede (verbliebene) FD $\alpha \rightarrow B \in F$ die Rechtsreduktion durch:

       gilt $B \subseteq$ AttrHülle$(F - (\alpha \rightarrow B), \alpha)$ (ist B bzw. $\alpha \rightarrow B$ überflüssig?)
       wenn ja: streiche $\alpha \rightarrow B$

3. Fasse mittels der Vereinigungsregel FDs zusammen.

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \rightarrow B, B \rightarrow C, AB \rightarrow C\}$

1. Dekomposition nicht notwendig

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \rightarrow B, B \rightarrow C, AB \rightarrow C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $A \rightarrow B$: bereits reduziert
      $B \rightarrow C$: bereits reduziert

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, B \to C, AB \to C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

$$\begin{aligned} A \to B&: \text{ bereits reduziert} \\ B \to C&: \text{ bereits reduziert} \\ AB \to C&: \ C \in \text{AttrHülle}(F, A) \end{aligned}$$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, B \to C, AB \to C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

$A \to B$: bereits reduziert

$B \to C$: bereits reduziert

$AB \to C$: $C \in \text{AttrHülle}(F, A)$ ja

$\Rightarrow F := \{A \to B, B \to C, A \to C\}$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, B \to C, AB \to C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $A \to B$: bereits reduziert

      $B \to C$: bereits reduziert

      $AB \to C$: $C \in \text{AttrHülle}(F, A)$ ja

      $\Rightarrow F := \{A \to B, B \to C, A \to C\}$
   2. Rechtsreduktion:

      $A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, B \to C, AB \to C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $A \to B$: bereits reduziert

      $B \to C$: bereits reduziert

      $AB \to C$: $C \in \text{AttrHülle}(F, A)$ ja

      $\Rightarrow F := \{A \to B, B \to C, A \to C\}$
   2. Rechtsreduktion:

      $A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, B \to C, AB \to C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $A \to B$: bereits reduziert
      $B \to C$: bereits reduziert
      $AB \to C$: $C \in \text{AttrHülle}(F, A)$ ja
      $\Rightarrow F := \{A \to B, B \to C, A \to C\}$
   2. Rechtsreduktion:

      $A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein
      $B \to C$: $C \in \text{AttrHülle}(F \setminus B \to C, B)$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, B \to C, AB \to C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $A \to B$: bereits reduziert

      $B \to C$: bereits reduziert

      $AB \to C$: $C \in \text{AttrHülle}(F, A)$ ja

      $\Rightarrow F := \{A \to B, B \to C, A \to C\}$
   2. Rechtsreduktion:

      $A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein

      $B \to C$: $C \in \text{AttrHülle}(F \setminus B \to C, B)$ nein

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, B \to C, AB \to C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $A \to B$: bereits reduziert
      $B \to C$: bereits reduziert
      $AB \to C$: $C \in \text{AttrHülle}(F, A)$ ja
      $\Rightarrow F := \{A \to B, B \to C, A \to C\}$
   2. Rechtsreduktion:

      $A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein
      $B \to C$: $C \in \text{AttrHülle}(F \setminus B \to C, B)$ nein
      $A \to C$: $C \in \text{AttrHülle}(F \setminus A \to C, A)$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, B \to C, AB \to C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $A \to B$: bereits reduziert
      $B \to C$: bereits reduziert
      $AB \to C$: $C \in \text{AttrHülle}(F, A)$ ja
      $\Rightarrow F := \{A \to B, B \to C, A \to C\}$
   2. Rechtsreduktion:

      $A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein
      $B \to C$: $C \in \text{AttrHülle}(F \setminus B \to C, B)$ nein
      $A \to C$: $C \in \text{AttrHülle}(F \setminus A \to C, A)$ ja
      $\Rightarrow F := \{A \to B, B \to C\}$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \rightarrow B, B \rightarrow C, AB \rightarrow C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $$\begin{aligned} A \rightarrow B&: \text{ bereits reduziert} \\ B \rightarrow C&: \text{ bereits reduziert} \\ AB \rightarrow C&: C \in \text{AttrHülle}(F, A) \text{ ja} \\ &\Rightarrow F := \{A \rightarrow B, B \rightarrow C, A \rightarrow C\} \end{aligned}$$

   2. Rechtsreduktion:

      $$\begin{aligned} A \rightarrow B&: B \in \text{AttrHülle}(F \setminus A \rightarrow B, A) \text{ nein} \\ B \rightarrow C&: C \in \text{AttrHülle}(F \setminus B \rightarrow C, B) \text{ nein} \\ A \rightarrow C&: C \in \text{AttrHülle}(F \setminus A \rightarrow C, A) \text{ ja} \\ &\Rightarrow F := \{A \rightarrow B, B \rightarrow C\} \end{aligned}$$

3. Vereinigungsregel nicht anwendbar

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \rightarrow B, B \rightarrow C, AB \rightarrow C\}$

1. Dekomposition nicht notwendig
2. Kürzen
   1. Linksreduktion:

      $A \rightarrow B$: bereits reduziert
      $B \rightarrow C$: bereits reduziert
      $AB \rightarrow C$: $C \in \text{AttrHülle}(F, A)$ ja
      $\Rightarrow F := \{A \rightarrow B, B \rightarrow C, A \rightarrow C\}$
   2. Rechtsreduktion:

      $A \rightarrow B$: $B \in \text{AttrHülle}(F \setminus A \rightarrow B, A)$ nein
      $B \rightarrow C$: $C \in \text{AttrHülle}(F \setminus B \rightarrow C, B)$ nein
      $A \rightarrow C$: $C \in \text{AttrHülle}(F \setminus A \rightarrow C, A)$ ja
      $\Rightarrow F := \{A \rightarrow B, B \rightarrow C\}$
3. Vereinigungsregel nicht anwendbar
   $\Rightarrow F_C := \{A \rightarrow B, B \rightarrow C\}$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to BD, AC \to E, CD \to E, E \to A, D \to C\}$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to BD, AC \to E, CD \to E, E \to A, D \to C\}$

1. Dekomposition:
   $F = \{A \to B, A \to D, AC \to E, CD \to E, E \to A, D \to C\}$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to BD, AC \to E, CD \to E, E \to A, D \to C\}$

1. Dekomposition:
   $F = \{A \to B, A \to D, AC \to E, CD \to E, E \to A, D \to C\}$
2. Linksreduktion:

   $A \to B$: ok

   $A \to D$: ok

   $E \to A$: ok

   $D \to C$: ok

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to BD, AC \to E, CD \to E, E \to A, D \to C\}$

1. Dekomposition:
   $F = \{A \to B, A \to D, AC \to E, CD \to E, E \to A, D \to C\}$
2. Linksreduktion:

   $A \to B$: ok

   $A \to D$: ok

   $E \to A$: ok

   $D \to C$: ok

   $AC \to E$: $E \in \text{AttrHülle}(F, A)$ ja $\Rightarrow$
   $F = \{A \to B, A \to D, A \to E, CD \to E, E \to A, D \to C\}$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to BD, AC \to E, CD \to E, E \to A, D \to C\}$

1. Dekomposition:
   $F = \{A \to B, A \to D, AC \to E, CD \to E, E \to A, D \to C\}$
2. Linksreduktion:

   $A \to B$: ok

   $A \to D$: ok

   $E \to A$: ok

   $D \to C$: ok

   $AC \to E$: $E \in \text{AttrHülle}(F, A)$ ja $\Rightarrow$
   $F = \{A \to B, A \to D, A \to E, CD \to E, E \to A, D \to C\}$

   $CD \to E$: $E \in \text{AttrHülle}(F, C)$ nein
   $E \in \text{AttrHülle}(F, D)$ ja $\Rightarrow$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, A \to D, A \to E, D \to E, E \to A, D \to C\}$

3. Rechtsreduktion:

   $A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \rightarrow B, A \rightarrow D, A \rightarrow E, D \rightarrow E, E \rightarrow A, D \rightarrow C\}$

3. Rechtsreduktion:

$A \rightarrow B$: $B \in \text{AttrHülle}(F \setminus A \rightarrow B, A)$ nein
$A \rightarrow D$: $D \in \text{AttrHülle}(F \setminus A \rightarrow D, A)$ nein

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, A \to D, A \to E, D \to E, E \to A, D \to C\}$

3. Rechtsreduktion:

$A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein

$A \to D$: $D \in \text{AttrHülle}(F \setminus A \to D, A)$ nein

$A \to E$: $E \in \text{AttrHülle}(F \setminus A \to E, A)$ ja $\Rightarrow$

$F = \{A \to B, A \to D, D \to E, E \to A, D \to C\}$

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, A \to D, A \to E, D \to E, E \to A, D \to C\}$

3. Rechtsreduktion:

$A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein

$A \to D$: $D \in \text{AttrHülle}(F \setminus A \to D, A)$ nein

$A \to E$: $E \in \text{AttrHülle}(F \setminus A \to E, A)$ ja $\Rightarrow$

$F = \{A \to B, A \to D, D \to E, E \to A, D \to C\}$

$D \to E$: $E \in \text{AttrHülle}(F \setminus D \to E, D)$ nein

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, A \to D, A \to E, D \to E, E \to A, D \to C\}$

3. Rechtsreduktion:

$A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein

$A \to D$: $D \in \text{AttrHülle}(F \setminus A \to D, A)$ nein

$A \to E$: $E \in \text{AttrHülle}(F \setminus A \to E, A)$ ja $\Rightarrow$

$F = \{A \to B, A \to D, D \to E, E \to A, D \to C\}$

$D \to E$: $E \in \text{AttrHülle}(F \setminus D \to E, D)$ nein

$E \to A$: $A \in \text{AttrHülle}(F \setminus E \to A, E)$ nein

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \to B, A \to D, A \to E, D \to E, E \to A, D \to C\}$

3. Rechtsreduktion:

$A \to B$: $B \in \text{AttrHülle}(F \setminus A \to B, A)$ nein

$A \to D$: $D \in \text{AttrHülle}(F \setminus A \to D, A)$ nein

$A \to E$: $E \in \text{AttrHülle}(F \setminus A \to E, A)$ ja $\Rightarrow$
$F = \{A \to B, A \to D, D \to E, E \to A, D \to C\}$

$D \to E$: $E \in \text{AttrHülle}(F \setminus D \to E, D)$ nein

$E \to A$: $A \in \text{AttrHülle}(F \setminus E \to A, E)$ nein

$D \to C$: $C \in \text{AttrHülle}(F \setminus D \to C, D)$ nein

Berechnung Kanonische Überdeckung (2)

Beispiel

$F = \{A \rightarrow B, A \rightarrow D, A \rightarrow E, D \rightarrow E, E \rightarrow A, D \rightarrow C\}$

3. Rechtsreduktion:

$A \rightarrow B$: $B \in \text{AttrHülle}(F \setminus A \rightarrow B, A)$ nein

$A \rightarrow D$: $D \in \text{AttrHülle}(F \setminus A \rightarrow D, A)$ nein

$A \rightarrow E$: $E \in \text{AttrHülle}(F \setminus A \rightarrow E, A)$ ja $\Rightarrow$

$F = \{A \rightarrow B, A \rightarrow D, D \rightarrow E, E \rightarrow A, D \rightarrow C\}$

$D \rightarrow E$: $E \in \text{AttrHülle}(F \setminus D \rightarrow E, D)$ nein

$E \rightarrow A$: $A \in \text{AttrHülle}(F \setminus E \rightarrow A, E)$ nein

$D \rightarrow C$: $C \in \text{AttrHülle}(F \setminus D \rightarrow C, D)$ nein

4. Zusammenfassen: $F_C = \{A \rightarrow BD, D \rightarrow EC, E \rightarrow A\}$

Überblick Wh.

- Ziele
- Funktionale Abhängigkeiten
  - Definitionen
  - Kanonische Überdeckung
- Entwurfstheorie und Zerlegung
  - "Schlechte" Relationenschemata
  - Zerlegung von Relationenschemata
  - Kriterien für eine "sinnvolle" Zerlegung
- Normalformen (1., 2., 3., Boyce-Codd)
  - Normalisierung durch Synthesealgorithmus
  - Normalisierung durch Dekomposition

Entwurfstheorie und Zerlegung

- "Schlechte" Relationenschemata
- Zerlegung von Relationenschemata
- Kriterien für eine "sinnvolle" Zerlegung
    - Verlustlosigkeit
    - Abhängigkeitstreue

“Schlechte” Relationenschemata

ProfVorl						
PersNr	Name	Rang	Raum	<u>VorlNr</u>	Titel	SWS
2125	Sokrates	C4	226	5041	Ethik	4
2125	Sokrates	C4	226	5049	Mäeutik	2
2125	Sokrates	C4	226	4052	Logik	4
...	...	...	...	...	...	...
2132	Popper	C3	52	5295	Der Wiener Kreis	2
2137	Kant	C4	7	4630	Die drei Kritiken	4

"Schlechte" Relationenschemata

ProfVorl

PersNr	Name	Rang	Raum	VorlNr	Titel	SWS
2125	Sokrates	C4	226	5041	Ethik	4
2125	Sokrates	C4	226	5049	Mäeutik	2
2125	Sokrates	C4	226	4052	Logik	4
...	...	...	...	...	...	...
2132	Popper	C3	52	5295	Der Wiener Kreis	2
2137	Kant	C4	7	4630	Die drei Kritiken	4

Update-Anomalie: Sokrates zieht um

"Schlechte" Relationenschemata

ProfVorl

PersNr	Name	Rang	Raum	VorlNr	Titel	SWS
2125	Sokrates	C4	226	5041	Ethik	4
2125	Sokrates	C4	226	5049	Mäeutik	2
2125	Sokrates	C4	226	4052	Logik	4
...	...	...	...	...	...	...
2132	Popper	C3	52	5295	Der Wiener Kreis	2
2137	Kant	C4	7	4630	Die drei Kritiken	4

Update-Anomalie: Sokrates zieht um

Lösch-Anomalie: Bsp: "Die 3 Kritiken" fällt weg

"Schlechte" Relationenschemata

ProfVorl

PersNr	Name	Rang	Raum	VorlNr	Titel	SWS
2125	Sokrates	C4	226	5041	Ethik	4
2125	Sokrates	C4	226	5049	Mäeutik	2
2125	Sokrates	C4	226	4052	Logik	4
...	...	...	...	...	...	...
2132	Popper	C3	52	5295	Der Wiener Kreis	2
2137	Kant	C4	7	4630	Die drei Kritiken	4

Update-Anomalie: Sokrates zieht um

Lösch-Anomalie: Bsp: "Die 3 Kritiken" fällt weg

Einfügeanomalie: Bsp: Curie ist neu und liest noch keine Vorlesungen (hat keinen Schlüssel?!)

Zerlegung von Relationenschemata (1)

- Anomalien beruhen auf der Tatsache, dass nicht zusammenpassende Informationen zusammen gespeichert wurden.

Zerlegung von Relationenschemata (1)

- Anomalien beruhen auf der Tatsache, dass nicht zusammenpassende Informationen zusammen gespeichert wurden.
- Intuitiv:
  - alle Informationen zu Professoren werden in einer Relation gespeichert,
  - alle Informationen zu Vorlesungen werden in einer Relation gespeichert,
  - die Information über die Verknüpfung beider Relationen wird in einer Relation gespeichert.

Zerlegung von Relationenschemata (1)

- Anomalien beruhen auf der Tatsache, dass nicht zusammenpassende Informationen zusammen gespeichert wurden.
- Intuitiv:
  - alle Informationen zu Professoren werden in einer Relation gespeichert,
  - alle Informationen zu Vorlesungen werden in einer Relation gespeichert,
  - die Information über die Verknüpfung beider Relationen wird in einer Relation gespeichert.
- Lösung: Zerlegung des Schemas in Teilschemata

Zerlegung von Relationenschemata (1)

- Anomalien beruhen auf der Tatsache, dass nicht zusammenpassende Informationen zusammen gespeichert wurden.
- Intuitiv:
  - alle Informationen zu Professoren werden in einer Relation gespeichert,
  - alle Informationen zu Vorlesungen werden in einer Relation gespeichert,
  - die Information über die Verknüpfung beider Relationen wird in einer Relation gespeichert.
- Lösung: Zerlegung des Schemas in Teilschemata
- Frage: was ist eine “sinnvolle” Zerlegung?

Zerlegung von Relationenschemata

Beispiel

Zerlegung von ProfVorl(PersNr, Name, Rang, Raum, VorlNr, Titel, SWS) in Teilschemata

Zerlegung von Relationenschemata

Beispiel

Zerlegung von ProfVorl(PersNr, Name, Rang, Raum, VorlNr, Titel, SWS) in Teilschemata

Versuch 1: Zerlegung in:

- Prof(PersNr, Name, Rang, Raum)
- Vorlesung(VorlNr, Titel, SWS)

Zerlegung von Relationenschemata

Beispiel

Zerlegung von ProfVorl(PersNr, Name, Rang, Raum, VorlNr, Titel, SWS) in Teilschemata

Versuch 1: Zerlegung in:

- Prof(PersNr, Name, Rang, Raum)
- Vorlesung(VorlNr, Titel, SWS)

Problem: wer hält welche LVA ?

Ursache: Teilschemata können nicht mehr korrekt verknüpft werden

Zerlegung von Relationenschemata

Beispiel

Zerlegung von ProfVorl(PersNr, Name, Rang, Raum, VorlNr, Titel, SWS) in Teilschemata

Versuch 1: Zerlegung in:

- Prof(PersNr, Name, Rang, Raum)
- Vorlesung(VorlNr, Titel, SWS)

Problem: wer hält welche LVA ?

Ursache: Teilschemata können nicht mehr korrekt verknüpft werden

Versuch 2: Zerlegung in:

- Prof(PersNr, Name, Raum, VorlNr, Titel, SWS)
- Titel(Name, Rang)

Zerlegung von Relationenschemata

Beispiel

Zerlegung von ProfVorl(PersNr, Name, Rang, Raum, VorlNr, Titel, SWS) in Teilschemata

Versuch 1: Zerlegung in:

- Prof(PersNr, Name, Rang, Raum)
- Vorlesung(VorlNr, Titel, SWS)

Problem: wer hält welche LVA ?

Ursache: Teilschemata können nicht mehr korrekt verknüpft werden

Versuch 2: Zerlegung in:

- Prof(PersNr, Name, Raum, VorlNr, Titel, SWS)
- Titel(Name, Rang)

Problem: Titel von Professoren gleichen Namens

Ursache: die FD $\{\text{PersNr}\} \rightarrow \{\text{Rang}\}$ ist verlorengegangen

Zerlegung von Relationenschemata (2)

- Es gibt zwei Korrektheitskriterien für die Zerlegung von Relationenschemata:

  Verlustlosigkeit (Verlust von Information)
  Abhängigkeitstreue (Verlust von Metainformation)

Zerlegung von Relationenschemata (2)

- Es gibt zwei Korrektheitskriterien für die Zerlegung von Relationenschemata:

  Verlustlosigkeit (Verlust von Information)
  Abhängigkeitstreue (Verlust von Metainformation)

Definition (Verlustlosigkeit - intuitiv)

Die in der Ausprägung R des Schemas $\mathcal{R}$ enthaltenen Informationen müssen aus den Ausprägungen $R_1, \ldots R_n$ der neuen Schemata $\mathcal{R}_1, \ldots \mathcal{R}_n$ rekonstruierbar sein.

Zerlegung von Relationenschemata (2)

- Es gibt zwei Korrektheitskriterien für die Zerlegung von Relationenschemata:

  Verlustlosigkeit (Verlust von Information)
  Abhängigkeitstreue (Verlust von Metainformation)

Definition (Verlustlosigkeit - intuitiv)

Die in der Ausprägung R des Schemas $\mathcal{R}$ enthaltenen Informationen müssen aus den Ausprägungen $R_1, \ldots R_n$ der neuen Schemata $\mathcal{R}_1, \ldots \mathcal{R}_n$ rekonstruierbar sein.

Definition (Abhängigkeitstreue - intuitiv)

Die auf $\mathcal{R}$ geltenden funktionalen Abhängigkeiten müssen auf die Schemata $\mathcal{R}_1, \ldots \mathcal{R}_n$ übertragbar sein.

Zerlegung von Relationenschemata (3)

Definition (Verlustlosigkeit)

Gegeben ein Schema $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$, mit den Ausprägungen R, $R_1 := \pi_{\mathcal{R}_1}(R)$ und $R_2 := \pi_{\mathcal{R}_2}(R)$
Die Zerlegung von $\mathcal{R}$ in $\mathcal{R}_1$ und $\mathcal{R}_2$ ist verlustlos, falls für jede mögliche (gültige) Ausprägung R von $\mathcal{R}$ gilt:

$$R = R_1 \bowtie R_2$$

Zerlegung von Relationenschemata (3)

Definition (Verlustlosigkeit)

Gegeben ein Schema $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$, mit den Ausprägungen R, $R_1 := \pi_{\mathcal{R}_1}(R)$ und $R_2 := \pi_{\mathcal{R}_2}(R)$
Die Zerlegung von $\mathcal{R}$ in $\mathcal{R}_1$ und $\mathcal{R}_2$ ist verlustlos, falls für jede mögliche (gültige) Ausprägung R von $\mathcal{R}$ gilt:

$$R = R_1 \bowtie R_2$$

Theorem (Hinreichende Bedingung für Verlustlosigkeit)

Eine Zerlegung von R in R_1 und R_2 ist verlustlos, wenn die Joinattribute in einer der Teilrelationen Schlüssel sind.

Zerlegung von Relationenschemata

Beispiel (Verlustlose Zerlegung 1)

Zerlegung von ProfVorl(PersNr, Name, Rang, VorlNr, Titel, SWS) in mehrere Teilschemata:

ProfVorl					
PersNr	Name	Rang	VorlNr	Titel	SWS
2125	Sokrates	C4	5041	Ethik	4
2125	Sokrates	C4	5049	Mäeutik	2
2125	Sokrates	C4	4052	Logik	4
...	...	...	...	...	...
2132	Popper	C3	5295	Der Wiener Kreis	2
2137	Kant	C4	4630	Die drei Kritiken	4

Zerlegung von Relationenschemata

Beispiel (Verlustlose Zerlegung 1)

Versuch 1:

Professoren		
<u>PersNr</u>	Name	Rang
2125	Sokrates	C4
2132	Popper	C3
2137	Kant	C4
...	...	...

Vorlesungen		
<u>VorlNr</u>	Titel	SWS
5041	Ethik	4
5049	Mäeutik	2
4052	Logik	4
5295	Der Wiener Kreis	2
4630	Die drei Kritiken	4
...	...	...

Frage: gilt ProfVorl $=$ Professoren $\bowtie$ Vorlesungen ?

Zerlegung von Relationenschemata

Beispiel (Verlustlose Zerlegung 1)

Versuch 1:

Professoren		
PersNr	Name	Rang
2125	Sokrates	C4
2132	Popper	C3
2137	Kant	C4
...	...	...

Vorlesungen		
VorlNr	Titel	SWS
5041	Ethik	4
5049	Mäeutik	2
4052	Logik	4
5295	Der Wiener Kreis	2
4630	Die drei Kritiken	4
...	...	...

Frage: gilt ProfVorl $=$ Professoren $\bowtie$ Vorlesungen ?

Nein: die Verknüpfung mittels Join wird mangels Joinattributs zum Karthesischen Produkt

Zerlegung von Relationenschemata

Beispiel (Verlustlose Zerlegung 1)

Versuch 2:

Professoren

PersNr	Name	Rang
2125	Sokrates	C4
2132	Popper	C3
2137	Kant	C4
...	...	...

Vorlesungen

VorlNr	Titel	SWS	PersNr
5041	Ethik	4	2125
5049	Mäeutik	2	2125
4052	Logik	4	2125
5295	Der Wiener Kreis	2	2132
4630	Die drei Kritiken	4	2137
...	...	...	...

Frage: gilt ProfVorl $=$ Professoren $\bowtie$ Vorlesungen ?

Zerlegung von Relationenschemata

Beispiel (Verlustlose Zerlegung 1)

Versuch 2:

Professoren

<u>PersNr</u>	Name	Rang
2125	Sokrates	C4
2132	Popper	C3
2137	Kant	C4
...	...	...

Vorlesungen

<u>VorlNr</u>	Titel	SWS	PersNr
5041	Ethik	4	2125
5049	Mäeutik	2	2125
4052	Logik	4	2125
5295	Der Wiener Kreis	2	2132
4630	Die drei Kritiken	4	2137
...	...	...	...

Frage: gilt ProfVorl = Professoren $\bowtie$ Vorlesungen ?

Ja: das Joinattribut PersNr ist Schlüssel in Professoren

$\rightarrow$ hinreichende Bedingung ist erfüllt $\rightarrow$ Zerlegung ist

Zerlegung von Relationenschemata

Beispiel (Verlustlose Zerlegung 2)

Zerlege Relation Eltern: {[Vater, Mutter, <u>Kind</u>]} in zwei Teilrelationen:

$\mathcal{R}_1$ Väter: {[Vater, <u>Kind</u>]}

$\mathcal{R}_2$ Mütter: {[Mutter, <u>Kind</u>]}

Frage: ist die Zerlegung verlustlos?

Zerlegung von Relationenschemata

Beispiel (Verlustlose Zerlegung 2)

Zerlege Relation Eltern: {[Vater, Mutter, <u>Kind</u>]} in zwei Teilrelationen:

- $\mathcal{R}_1$ Väter: {[Vater, <u>Kind</u>]}
- $\mathcal{R}_2$ Mütter: {[Mutter, <u>Kind</u>]}

Frage: ist die Zerlegung verlustlos?

Ja: das Joinattribut von Väter mit Mütter ist Kind $\Rightarrow$ Verlustlosigkeit ist garantiert

Zerlegung von Relationenschemata

Beispiel (Verlustlose Zerlegung 2)

Zerlege Relation Eltern: {[Vater, Mutter, <u>Kind</u>]} in zwei Teilrelationen:

$\mathcal{R}_1$ Väter: {[Vater, <u>Kind</u>]}

$\mathcal{R}_2$ Mütter: {[Mutter, <u>Kind</u>]}

Frage: ist die Zerlegung verlustlos?

Ja: das Joinattribut von Väter mit Mütter ist Kind $\Rightarrow$ Verlustlosigkeit ist garantiert

Bemerkung: Die Zerlegung von Eltern ist zwar verlustlos, aber auch unnötig, da die Relation in sehr gutem Zustand (Normalform) ist

Zerlegung von Relationenschemata (3)

Definition (Abhängigkeitstreue)

Gegeben ein Schema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$.
Die Zerlegung von $\mathcal{R}$ in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist abhängigkeitstreu, wenn

$$F_{\mathcal{R}} \equiv (F_{\mathcal{R}_1} \cup \cdots \cup F_{\mathcal{R}_n}) \quad \text{bzw.} \quad F_{\mathcal{R}}^+ = (F_{\mathcal{R}_1} \cup \cdots \cup F_{\mathcal{R}_n})^+$$

Zerlegung von Relationenschemata (3)

Beispiel (Abhängigkeitstreue Zerlegung)

Gegeben ein Relationenschema PLZverzeichnis:{[Straße, Ort, BLand, PLZ]} und folgenden Annahmen:

- Orte werden durch Namen und Bundesland eindeutig identifiziert.
- PLZgebiete gehen nicht über Ortsgrenzen hinweg $\Rightarrow$ {PLZ} $\rightarrow$ {Ort, BLand}
- Innerhalb einer Straße ändert sich die PLZ nicht $\Rightarrow$ {Straße, Ort, BLand} $\rightarrow$ {PLZ}.

Zerlegung von Relationenschemata (3)

Beispiel (Abhängigkeitstreue Zerlegung)

Gegeben ein Relationenschema PLZverzeichnis:{[Straße, Ort, BLand, PLZ]} und folgenden Annahmen:

- Orte werden durch Namen und Bundesland eindeutig identifiziert.
- PLZgebiete gehen nicht über Ortsgrenzen hinweg $\Rightarrow$ {PLZ} $\rightarrow$ {Ort, BLand}
- Innerhalb einer Straße ändert sich die PLZ nicht $\Rightarrow$ {Straße, Ort, BLand} $\rightarrow$ {PLZ}.

Zerlegung in:

$\mathcal{R}_1$ Straßen: {[<u>PLZ, Straße</u>]}

$\mathcal{R}_2$ Orte: {[<u>PLZ</u>, Ort, BLand]} mit FD {PLZ} $\rightarrow$ {Ort, BLand}

Zerlegung von Relationenschemata (3)

Beispiel (Abhängigkeitstreue Zerlegung)

Gegeben ein Relationenschema PLZverzeichnis:{[Straße, Ort, BLand, PLZ]} und folgenden Annahmen:

- Orte werden durch Namen und Bundesland eindeutig identifiziert.
- PLZgebiete gehen nicht über Ortsgrenzen hinweg $\Rightarrow$ {PLZ} $\rightarrow$ {Ort, BLand}
- Innerhalb einer Straße ändert sich die PLZ nicht $\Rightarrow$ {Straße, Ort, BLand} $\rightarrow$ {PLZ}.

Zerlegung in:

$\mathcal{R}_1$ Straßen: {[<u>PLZ, Straße</u>]}

$\mathcal{R}_2$ Orte: {[<u>PLZ</u>, Ort, BLand]} mit FD {PLZ} $\rightarrow$ {Ort, BLand}

ABER: {Straße, Ort, BLand} $\rightarrow$ {PLZ} ist verloren gegangen.
(Zerlegung ist verlustlos, da PLZ Schlüssel in Orte)

Zerlegung von Relationenschemata (3)

Beispiel (Abhängigkeitstreue Zerlegung)

PLZVerzeichnis

Ort	BLand	Straße	PLZ
Frankfurt	Hessen	Goethestraße	60313
Frankfurt	Hessen	Schillerstraße	60437
Frankfurt	Brandenburg	Goethestraße	15234

Zerlegung von Relationenschemata (3)

Beispiel (Abhängigkeitstreue Zerlegung)

PLZVerzeichnis

Ort	BLand	Straße	PLZ
Frankfurt	Hessen	Goethestraße	60313
Frankfurt	Hessen	Schillerstraße	60437
Frankfurt	Brandenburg	Goethestraße	15234

Straßen

PLZ	Straße
60313	Goethestraße
60437	Schillerstraße
15234	Goethestraße
15235	Goethestraße

Orte

Ort	BLand	PLZ
Frankfurt	Hessen	60313
Frankfurt	Hessen	60437
Frankfurt	Brandenburg	15234
Frankfurt	Brandenburg	15235

Normalformen

- 1. Normalform
  - NF^2 Relationen
- 2. Normalform
- 3. Normalform
  - Zerlegung in 3NF
  - Synthesealgorithmus
- Boyce-Codd Normalform
  - Zerlegung in BCNF
  - Dekompositionsalgorithmus

1. Normalform

Definition (1. Normalform)

Ein Relationenschema $\mathcal{R}$ ist in 1. Normalform, wenn die Domänen von $\mathcal{R}$ atomar sind.

1. Normalform

Definition (1. Normalform)

Ein Relationenschema $\mathcal{R}$ ist in 1. Normalform, wenn die Domänen von $\mathcal{R}$ atomar sind.

Eltern

Vater	Mutter	Kinder
Johann	Martha	{Else, Lucie}
Heinz	Martha	{Cleo}

nicht in 1NF

Eltern

Vater	Mutter	Kind
Johann	Martha	Else
Johann	Martha	Lucie
Heinz	Martha	Cleo

in 1NF

NF^2 Relationen

- Non-First Normal Form Relationen
- Attribut einer Relation kann selbst eine Relation sein
- im SQL2003 Standard enthalten

NF^2 Relationen

- Non-First Normal Form Relationen
- Attribut einer Relation kann selbst eine Relation sein
- im SQL2003 Standard enthalten

Eltern			
Vater	Mutter	Kinder	
		KName	KAlter
Johann	Martha	Else	5
		Lucie	3
Johann	Maria	Theo	3
		Josef	1
Heinz	Martha	Cleo	9

2. Normalform

- Intuitiv: Relationenschema $\mathcal{R}$ verletzt die 2. Normalform, wenn in der Relation Informationen über mehr als ein Konzept modelliert werden.
- Nur von historischem Interesse, da es immer noch zu Anomalien kommen kann.

3. Normalform

Definition (3. Normalform)

Ein Relationenschema $\mathcal{R}$ ist in dritter Normalform, wenn für jede auf $\mathcal{R}$ geltende funktionale Abhängigkeit der Form $\alpha \rightarrow B$ mit $\alpha \subseteq \mathcal{R}$ und $B \in \mathcal{R}$ eine der folgenden drei Bedingungen gilt:

1. $B \in \alpha$, d.h., die FD ist trivial
2. α ist Superschlüssel von $\mathcal{R}$.
3. das Attribut B ist in einem der Schlüssel von $\mathcal{R}$ enthalten

3. Normalform

Beispiel (Check auf 3NF 1)

$$\begin{aligned} \mathcal{R} &= (ABCDEF), \\ F &= \{C \to BDAE\} = \{C \to B, C \to D, C \to A, C \to E\}. \end{aligned}$$

3. Normalform

Beispiel (Check auf 3NF 1)

$$\mathcal{R} = (ABCDEF),$$
$$F = \{C \rightarrow BDAE\} = \{C \rightarrow B, C \rightarrow D, C \rightarrow A, C \rightarrow E\}.$$

Der Schlüssel von $\mathcal{R}$ ist: CF

3. Normalform

Beispiel (Check auf 3NF 1)

$$\mathcal{R} = (ABCDEF),$$
$$F = \{C \to BDAE\} = \{C \to B, C \to D, C \to A, C \to E\}.$$

Der Schlüssel von $\mathcal{R}$ ist: CF

$(\mathcal{R}, F)$ ist in 3NF, wenn für jede FD eine der drei NF-Bedingungen gilt:

1. Keine FD ist trivial: Bedingung 1 gilt nicht

3. Normalform

Beispiel (Check auf 3NF 1)

$$\mathcal{R} = (ABCDEF),$$
$$F = \{C \to BDAE\} = \{C \to B, C \to D, C \to A, C \to E\}.$$

Der Schlüssel von $\mathcal{R}$ ist: CF

$(\mathcal{R}, F)$ ist in 3NF, wenn für jede FD eine der drei NF-Bedingungen gilt:

1. Keine FD ist trivial: Bedingung 1 gilt nicht
2. Ist die Attributmenge auf der linken Seite der FDs ein Superschlüssel von $\mathcal{R}$? Nein, Bedingung 2 gilt für keine FD.

3. Normalform

Beispiel (Check auf 3NF 1)

$$\mathcal{R} = (ABCDEF),$$
$$F = \{C \to BDAE\} = \{C \to B, C \to D, C \to A, C \to E\}.$$

Der Schlüssel von $\mathcal{R}$ ist: CF

$(\mathcal{R}, F)$ ist in 3NF, wenn für jede FD eine der drei NF-Bedingungen gilt:

1. Keine FD ist trivial: Bedingung 1 gilt nicht
2. Ist die Attributmenge auf der linken Seite der FDs ein Superschlüssel von $\mathcal{R}$? Nein, Bedingung 2 gilt für keine FD.
3. Ist das Attribut auf der rechten Seite der FDs in einem der Schlüssel von $\mathcal{R}$ enthalten? Nein, Bedingung 3 gilt für keine FD.

3. Normalform

Beispiel (Check auf 3NF 1)

$$\mathcal{R} = (ABCDEF),$$
$$F = \{C \to BDAE\} = \{C \to B, C \to D, C \to A, C \to E\}.$$

Der Schlüssel von $\mathcal{R}$ ist: CF

$(\mathcal{R}, F)$ ist in 3NF, wenn für jede FD eine der drei NF-Bedingungen gilt:

1. Keine FD ist trivial: Bedingung 1 gilt nicht
2. Ist die Attributmenge auf der linken Seite der FDs ein Superschlüssel von $\mathcal{R}$? Nein, Bedingung 2 gilt für keine FD.
3. Ist das Attribut auf der rechten Seite der FDs in einem der Schlüssel von $\mathcal{R}$ enthalten? Nein, Bedingung 3 gilt für keine FD.

Mindestens eine FD ($C \to B$) verletzt die 3NF $\to$ $\mathcal{R}$ ist nicht in

3. Normalform

Beispiel (Check auf 3NF 2)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E\}$$

3. Normalform

Beispiel (Check auf 3NF 2)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \rightarrow B, C \rightarrow D, D \rightarrow A, D \rightarrow E, E \rightarrow C, E \rightarrow F, F \rightarrow E\}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, F, D

3. Normalform

Beispiel (Check auf 3NF 2)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \rightarrow B, C \rightarrow D, D \rightarrow A, D \rightarrow E, E \rightarrow C, E \rightarrow F, F \rightarrow E\}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, F, D

1. Keine FD ist trivial: Bedingung 1 gilt nicht

3. Normalform

Beispiel (Check auf 3NF 2)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \rightarrow B, C \rightarrow D, D \rightarrow A, D \rightarrow E, E \rightarrow C, E \rightarrow F, F \rightarrow E\}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, F, D

1. Keine FD ist trivial: Bedingung 1 gilt nicht
2. Ist die Attributmenge auf der linker Seite der FDs ein Superschlüssel von $\mathcal{R}$? OK für $C \rightarrow B, C \rightarrow D, D \rightarrow A, D \rightarrow E, E \rightarrow C, E \rightarrow F, F \rightarrow E$.

3. Normalform

Beispiel (Check auf 3NF 2)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E\}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, F, D

1. Keine FD ist trivial: Bedingung 1 gilt nicht
2. Ist die Attributmenge auf der linker Seite der FDs ein Superschlüssel von $\mathcal{R}$? OK für $C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E$.
3. Ist das Attribut auf der rechten Seite der FDs in einem der Schlüssel von $\mathcal{R}$ enthalten? OK für $C \to D, D \to E, E \to CF, F \to E$, gilt aber nicht für $C \to B, D \to A$

3. Normalform

Beispiel (Check auf 3NF 2)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E\}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, F, D

1. Keine FD ist trivial: Bedingung 1 gilt nicht
2. Ist die Attributmenge auf der linker Seite der FDs ein Superschlüssel von $\mathcal{R}$? OK für $C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E$.
3. Ist das Attribut auf der rechten Seite der FDs in einem der Schlüssel von $\mathcal{R}$ enthalten? OK für $C \to D, D \to E, E \to CF, F \to E$, gilt aber nicht für $C \to B, D \to A$

Keine FD verletzt die 3NF $\Rightarrow$ in 3NF

(In diesem Fall schon nach Schritt 2 geklärt.)

3. Normalform

Beispiel

ProfVorl $\{[\text{PersNr, Name, Rang, Raum, VorlNr, Titel, SWS}]\}$

$$F = \{\text{PersNr} \rightarrow \text{Name Rang Raum},$$
$$\text{VorlNr} \rightarrow \text{PersNr Name Rang Raum VorlNr Titel SWS}\}$$

ProfVorl						
PersNr	Name	Rang	Raum	<u>VorlNr</u>	Titel	SWS
2125	Sokrates	C4	226	5041	Ethik	4
2125	Sokrates	C4	226	5049	Mäeutik	2
2125	Sokrates	C4	226	4052	Logik	4
...	...	...	...	...	...	...
2132	Popper	C3	52	5295	Der Wiener Kreis	2
2137	Kant	C4	7	4630	Die drei Kritiken	4

Nicht in 3. Normalform

Zerlegung in 3NF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in dritter Normalform, wenn alle $\mathcal{R}_i$ in dritter Normalform sind.

Zerlegung in 3NF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in dritter Normalform, wenn alle $\mathcal{R}_i$ in dritter Normalform sind.

- Gegeben Relationenschema $\mathcal{R}$ mit FDs F.
- Gesucht: Zerlegung in Teilschemata $\mathcal{R}_1 \ldots \mathcal{R}_n$, für die gilt:
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist verlustlos,
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist abhängigkeitstreu,
  - alle $\mathcal{R}_1 \ldots \mathcal{R}_n$ sind in dritter Normalform.

Zerlegung in 3NF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in dritter Normalform, wenn alle $\mathcal{R}_i$ in dritter Normalform sind.

- Gegeben Relationenschema $\mathcal{R}$ mit FDs F.
- Gesucht: Zerlegung in Teilschemata $\mathcal{R}_1 \ldots \mathcal{R}_n$, für die gilt:
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist verlustlos,
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist abhängigkeitstreu,
  - alle $\mathcal{R}_1 \ldots \mathcal{R}_n$ sind in dritter Normalform.
- Lösung: Synthesealgorithmus

Der Synthesealgorithmus

1. Bestimme kanonische Überdeckung F_c zu F.

Der Synthesealgorithmus

1. Bestimme kanonische Überdeckung F_c zu F.
2. Für jede funktionale Abhängigkeit $\alpha \rightarrow \beta \in F_c$:
   - Erstelle ein Relationenschema $\mathcal{R}_\alpha := \alpha \cup \beta$
   - Ordne $\mathcal{R}_\alpha$ die FDs $F_\alpha := \{\alpha' \rightarrow \beta' \in F_c \mid \alpha' \cup \beta' \subseteq \mathcal{R}_\alpha\}$ zu.

     Informell: aus jeder FD wird ein eigenes Schema (Bedingung für Abhängigkeitstreue erfüllt) - kanonische Überdeckung wichtig, um nicht zu viele oder zu große Teilschemata zu generieren

Der Synthesealgorithmus

1. Bestimme kanonische Überdeckung F_c zu F.
2. Für jede funktionale Abhängigkeit $\alpha \rightarrow \beta \in F_c$:
   - Erstelle ein Relationenschema $\mathcal{R}_\alpha := \alpha \cup \beta$
   - Ordne $\mathcal{R}_\alpha$ die FDs $F_\alpha := \{\alpha' \rightarrow \beta' \in F_c \mid \alpha' \cup \beta' \subseteq \mathcal{R}_\alpha\}$ zu.

     Informell: aus jeder FD wird ein eigenes Schema (Bedingung für Abhängigkeitstreue erfüllt) - kanonische Überdeckung wichtig, um nicht zu viele oder zu große Teilschemata zu generieren
3. Enthält eines der in Schritt 2. erzeugten Teilschemata einen Schlüssel von $\mathcal{R}$ bzgl. $F_c \Rightarrow$ fertig

   andernfalls: wähle einen Schlüssel $\kappa \in \mathcal{R}$ aus und definiere folgendes zusätzliche Schema: $\mathcal{R}_\kappa := \kappa$ mit $F_\kappa := \emptyset$

   Informell: erzeuge ein Schema zum Verknüpfen der Teilschemata (Bedingung für Verlustfreiheit erfüllt)

Der Synthesealgorithmus

1. Bestimme kanonische Überdeckung F_c zu F.
2. Für jede funktionale Abhängigkeit $\alpha \to \beta \in F_c$:
   - Erstelle ein Relationenschema $\mathcal{R}_\alpha := \alpha \cup \beta$
   - Ordne $\mathcal{R}_\alpha$ die FDs $F_\alpha := \{\alpha' \to \beta' \in F_c \mid \alpha' \cup \beta' \subseteq \mathcal{R}_\alpha\}$ zu.

   Informell: aus jeder FD wird ein eigenes Schema (Bedingung für Abhängigkeitstreue erfüllt) - kanonische Überdeckung wichtig, um nicht zu viele oder zu große Teilschemata zu generieren
3. Enthält eines der in Schritt 2. erzeugten Teilschemata einen Schlüssel von $\mathcal{R}$ bzgl. $F_c \Rightarrow$ fertig
   andernfalls: wähle einen Schlüssel $\kappa \in \mathcal{R}$ aus und definiere folgendes zusätzliche Schema: $\mathcal{R}_\kappa := \kappa$ mit $F_\kappa := \emptyset$

   Informell: erzeuge ein Schema zum Verknüpfen der Teilschemata (Bedingung für Verlustfreiheit erfüllt)
4. Eliminiere die in einem anderen Schema $\mathcal{R}_{\alpha'}$ enthaltenen Schemata $\mathcal{R}_\alpha$.

   Informell: kürze überflüssige Schemata

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 1)

$\mathcal{R} =$
$\{PersNr, Name, Rang, Raum, Ort, Straße, PLZ, Vorwahl, BLand, EW, Landesregierung\} = \{P, N, R, Z, O, S, Plz, V, B, E, L\}$ und die FDs von früher.

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 1)

$\mathcal{R} =$
$\{PersNr, Name, Rang, Raum, Ort, Straße, PLZ, Vorwahl, BLand, EW, Landesregierung\} = \{P, N, R, Z, O, S, Plz, V, B, E, L\}$ und die FDs von früher.

1. Kanonische Überdeckung:
   $P \rightarrow NRZOSB$, $Z \rightarrow P$, $SBO \rightarrow Plz$, $OB \rightarrow EV$, $B \rightarrow L$, $Plz \rightarrow BO$

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 1)

$\mathcal{R} =$
$\{PersNr, Name, Rang, Raum, Ort, Straße, PLZ, Vorwahl, BLand, EW, Landesregierung\} = \{P, N, R, Z, O, S, Plz, V, B, E, L\}$ und die FDs von früher.

1. Kanonische Überdeckung:
   $P \to NRZOSB$, $Z \to P$, $SBO \to Plz$, $OB \to EV$, $B \to L$, $Plz \to BO$
2. Generierung der Teilschemata und Zuordnung aller FDs:
   $\{PNRZOSB\}$ es gelten: $P \to NRZOSB$ und $Z \to P$
   $\{ZP\}$ es gelten: $Z \to P$ und $P \to Z$
   $\{SBOPlz\}$ es gelten: $SBO \to Plz$ und $Plz \to BO$
   $\{OBEV\}$ es gilt: $OB \to EV$
   $\{BL\}$ es gilt: $B \to L$
   $\{PlzBO\}$ es gilt: $Plz \to BO$

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 1)

3. Enthält eines der Teilschemata einen Schlüssel von $\mathcal{R}$ bezüglich F? Ja: P war Schlüssel und ist in $\{PNRZOSB\}$ enthalten $\Rightarrow$ fertig

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 1)

3. Enthält eines der Teilschemata einen Schlüssel von $\mathcal{R}$ bezüglich F? Ja: P war Schlüssel und ist in $\{PNRZOSB\}$ enthalten $\Rightarrow$ fertig
4. Schemaelimination:

$$\{ZP\} \text{ ist schon in } \{PNRZOSB\} \text{ enthalten} \Rightarrow \text{kürzen}$$
$$\{PlzBO\} \text{ ist schon in } \{SBOPlz\} \text{ enthalten} \Rightarrow \text{kürzen}$$

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 1)

3. Enthält eines der Teilschemata einen Schlüssel von $\mathcal{R}$ bezüglich F? Ja: P war Schlüssel und ist in $\{PNRZOSB\}$ enthalten $\Rightarrow$ fertig
4. Schemaelimination:

$\{ZP\}$ ist schon in $\{PNRZOSB\}$ enthalten $\Rightarrow$ kürzen
$\{PlzBO\}$ ist schon in $\{SBOPlz\}$ enthalten $\Rightarrow$ kürzen

5. Ergebnis:

Professoren: {[PersNr, Name, Rang, Raum, Ort, Straße, BLand]}
PLZListe: {[Straße, Ort, BLand,PLZ]}
Städte: {[Ort, BLand, EW, Vorwahl]}
Regierung: {[BLand, Landesregierung]}

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 2)

$\mathcal{R} = (ABCDEF), F = \{A \rightarrow EC, BC \rightarrow F, D \rightarrow B\}$

Frage: in 3 NF?

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 2)

$\mathcal{R} = (ABCDEF), F = \{A \rightarrow EC, BC \rightarrow F, D \rightarrow B\}$

Frage: in 3 NF? Schlüssel: AD

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 2)

$\mathcal{R} = (ABCDEF), F = \{A \rightarrow EC, BC \rightarrow F, D \rightarrow B\}$

Frage: in 3 NF? Nein. Schlüssel: AD

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 2)

$\mathcal{R} = (ABCDEF), F = \{A \to EC, BC \to F, D \to B\}$

Frage: in 3 NF? Nein. Schlüssel: AD

1. ist schon kanonisch

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 2)

$\mathcal{R} = (ABCDEF), F = \{A \to EC, BC \to F, D \to B\}$
Frage: in 3 NF? Nein. Schlüssel: AD

1. ist schon kanonisch
2. $\mathcal{R}_1 = (AEC), \mathcal{R}_2 = (BCF), \mathcal{R}_3 = (DB)$
   keine (nichttrivialen) Zuordnungen möglich

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 2)

$\mathcal{R} = (ABCDEF), F = \{A \to EC, BC \to F, D \to B\}$
Frage: in 3 NF? Nein. Schlüssel: AD

1. ist schon kanonisch
2. $\mathcal{R}_1 = (AEC), \mathcal{R}_2 = (BCF), \mathcal{R}_3 = (DB)$
   keine (nichttrivialen) Zuordnungen möglich
3. Schlüssel von $\mathcal{R}$: AD hinzufügen von $\mathcal{R}_4 = (AD)$

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 2)

$\mathcal{R} = (ABCDEF), F = \{A \to EC, BC \to F, D \to B\}$
Frage: in 3 NF? Nein. Schlüssel: AD

1. ist schon kanonisch
2. $\mathcal{R}_1 = (AEC), \mathcal{R}_2 = (BCF), \mathcal{R}_3 = (DB)$
   keine (nichttrivialen) Zuordnungen möglich
3. Schlüssel von $\mathcal{R}$: AD hinzufügen von $\mathcal{R}_4 = (AD)$
4. Nichts ist zu eliminieren.

Der Synthesealgorithmus

Beispiel (Synthesealgorithmus 2)

$\mathcal{R} = (ABCDEF), F = \{A \to EC, BC \to F, D \to B\}$

Frage: in 3 NF? Nein. Schlüssel: AD

1. ist schon kanonisch
2. $\mathcal{R}_1 = (AEC), \mathcal{R}_2 = (BCF), \mathcal{R}_3 = (DB)$
   keine (nichttrivialen) Zuordnungen möglich
3. Schlüssel von $\mathcal{R}$: AD hinzufügen von $\mathcal{R}_4 = (AD)$
4. Nichts ist zu eliminieren.

$$\mathcal{R} = (AEC) \cup (BCF) \cup (DB) \cup (AD)$$

Der Synthesealgorithmus

Beispiel

Warum ist eine kanonische Überdeckung notwendig?

ProfVorl {[PersNr, Name, Rang, Raum, VorlNr, Titel, SWS]}

$$F_1: \{\text{PersNr} \rightarrow \text{Name Rang Raum}, \\ \text{VorlNr} \rightarrow \text{PersNr Name Rang Raum VorlNr Titel SWS}\}$$

Der Synthesealgorithmus

Beispiel

Warum ist eine kanonische Überdeckung notwendig?

ProfVorl {[PersNr, Name, Rang, Raum, VorlNr, Titel, SWS]}

F_1 : {PersNr → Name Rang Raum,
VorlNr → PersNr Name Rang Raum VorlNr Titel SWS}

F_2 : {PersNr → Name, PersNr → Rang, PersNr → Raum,
VorlNr → PersNr, VorlNr → Name, VorlNr → Rang,
VorlNr → Raum, VorlNr → VorlNr, VorlNr → Titel,
VorlNr → SWS}

Ist 3 NF genug?

Beispiel (Dekompositionsalgorithmus 1)

$$\mathcal{R} = \text{Städte}\{[\text{Ort, BLand, LH, EW}]\}$$
$$\text{F} = \{\text{Ort BLand} \rightarrow \text{EW, BLand} \rightarrow \text{LH, LH} \rightarrow \text{BLand}\}$$

Schlüssel:

Ist 3 NF genug?

Beispiel (Dekompositionsalgorithmus 1)

$$\mathcal{R} = \text{Städte}\{[\text{Ort, BLand, LH, EW}]\}$$
$$F = \{\text{Ort BLand} \rightarrow \text{EW}, \text{BLand} \rightarrow \text{LH}, \text{LH} \rightarrow \text{BLand}\}$$

Schlüssel: {Ort, BLand}, {Ort, LH}

Ist 3 NF genug?

Beispiel (Dekompositionsalgorithmus 1)

$$\mathcal{R} = \text{Städte}\{[\text{Ort, BLand, LH, EW}]\}$$
$$F = \{\text{Ort BLand} \rightarrow \text{EW, BLand} \rightarrow \text{LH, LH} \rightarrow \text{BLand}\}$$

Schlüssel: {Ort, BLand}, {Ort, LH} also in 3NF

Städte

Ort	BLand	LH	EW
o1	b1	lh1	1
o2	b1	lh1	2
o3	b1	lh1	3
o4	b2	lh2	1

Boyce-Codd Normalform

Definition (Boyce-CoddNormalform)

Ein Relationenschema $\mathcal{R}$ ist in Boyce-Codd Normalform, wenn für jede auf $\mathcal{R}$ geltende funktionale Abhängigkeit der Form $\alpha \rightarrow B$ mit $\alpha \subseteq \mathcal{R}$ und $B \in \mathcal{R}$ eine der folgenden zwei Bedingungen gilt:

1. $B \in \alpha$, d.h., die FD ist trivial
2. α ist Superschlüssel von $\mathcal{R}$.

Boyce-Codd Normalform

Definition (Boyce-CoddNormalform)

Ein Relationenschema $\mathcal{R}$ ist in Boyce-Codd Normalform, wenn für jede auf $\mathcal{R}$ geltende funktionale Abhängigkeit der Form $\alpha \to B$ mit $\alpha \subseteq \mathcal{R}$ und $B \in \mathcal{R}$ eine der folgenden zwei Bedingungen gilt:

1. $B \in \alpha$, d.h., die FD ist trivial
2. α ist Superschlüssel von $\mathcal{R}$.

Theorem (Normalformen Zusammenhang)

Die Normalformen hängen wie folgt zusammen:

$$BCNF \subseteq 3NF \subseteq 2NF \subseteq 1NF$$

Boyce-Codd Normalform

Beispiel (Check auf BCNF 1)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E\}$$

Boyce-Codd Normalform

Beispiel (Check auf BCNF 1)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \rightarrow B, C \rightarrow D, D \rightarrow A, D \rightarrow E, E \rightarrow C, E \rightarrow F, F \rightarrow E\}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, D, F

Boyce-Codd Normalform

Beispiel (Check auf BCNF 1)

$$\begin{aligned} \mathcal{R} &= (ABCDEF) \\ F &= \{C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E\} \end{aligned}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, D, F

$(\mathcal{R}, F)$ ist in BCNF, wenn für jede FD eine der zwei NF-Bedingungen gilt:

1. Keine FD ist trivial: Bedingung 1 gilt nicht

Boyce-Codd Normalform

Beispiel (Check auf BCNF 1)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E\}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, D, F
$(\mathcal{R}, F)$ ist in BCNF, wenn für jede FD eine der zwei NF-Bedingungen gilt:

1. Keine FD ist trivial: Bedingung 1 gilt nicht
2. Ist die Attributmenge auf der linken Seite der FDs ein Superschlüssel von $\mathcal{R}$? Ja, Bedingung 2 gilt für alle FDs.

Boyce-Codd Normalform

Beispiel (Check auf BCNF 1)

$$\mathcal{R} = (ABCDEF)$$
$$F = \{C \to B, C \to D, D \to A, D \to E, E \to C, E \to F, F \to E\}$$

Die Schlüssel von $\mathcal{R}$ sind: C, E, D, F

$(\mathcal{R}, F)$ ist in BCNF, wenn für jede FD eine der zwei NF-Bedingungen gilt:

1. Keine FD ist trivial: Bedingung 1 gilt nicht
2. Ist die Attributmenge auf der linken Seite der FDs ein Superschlüssel von $\mathcal{R}$? Ja, Bedingung 2 gilt für alle FDs.

Keine FD verletzt die BCNF $\Rightarrow$ in BCNF

Boyce-Codd Normalform

Beispiel (Check auf BCNF 2)

$\mathcal{R}$ = Städte:{Ort, BLand, LH, EW}

F = {Ort, BLand → EW, BLand → LH, LH → BLand}

Boyce-Codd Normalform

Beispiel (Check auf BCNF 2)

$\mathcal{R}$ = Städte:{Ort, BLand, LH, EW}

F = {Ort, BLand $\rightarrow$ EW, BLand $\rightarrow$ LH, LH $\rightarrow$ BLand}

Schlüssel: {Ort, BLand} bzw. {Ort, LH}

Boyce-Codd Normalform

Beispiel (Check auf BCNF 2)

$\mathcal{R}$ = Städte:{Ort, BLand, LH, EW}

F = {Ort, BLand $\rightarrow$ EW, BLand $\rightarrow$ LH, LH $\rightarrow$ BLand}

Schlüssel: {Ort, BLand} bzw. {Ort, LH}

1. keine FD ist trivial

Boyce-Codd Normalform

Beispiel (Check auf BCNF 2)

$\mathcal{R}$ = Städte:{Ort, BLand, LH, EW}

F = {Ort, BLand → EW, BLand → LH, LH → BLand}

Schlüssel: {Ort, BLand} bzw. {Ort, LH}

1. keine FD ist trivial
2. ist die Attributmenge auf der linken Seite der FDs ein Superschlüssel? Nein für BLand → LH, LH → BLand

Boyce-Codd Normalform

Beispiel (Check auf BCNF 2)

$\mathcal{R}$ = Städte:{Ort, BLand, LH, EW}

F = {Ort, BLand $\rightarrow$ EW, BLand $\rightarrow$ LH, LH $\rightarrow$ BLand}

Schlüssel: {Ort, BLand} bzw. {Ort, LH}

1. keine FD ist trivial
2. ist die Attributmenge auf der linken Seite der FDs ein Superschlüssel? Nein für BLand $\rightarrow$ LH, LH $\rightarrow$ BLand

$\Rightarrow \mathcal{R}$ nicht in BCNF

Boyce-Codd Normalform

Beispiel (Check auf BCNF 2)

$\mathcal{R}$ = Städte:{Ort, BLand, LH, EW}

F = {Ort, BLand $\rightarrow$ EW, BLand $\rightarrow$ LH, LH $\rightarrow$ BLand}

Schlüssel: {Ort, BLand} bzw. {Ort, LH}

1. keine FD ist trivial
2. ist die Attributmenge auf der linken Seite der FDs ein Superschlüssel? Nein für BLand $\rightarrow$ LH, LH $\rightarrow$ BLand

$\Rightarrow \mathcal{R}$ nicht in BCNF

Aber: in 3NF, da Bedingung 2 (3NF): auf der rechten Seite steht ein Schlüsselattribut erfüllt ist.

Zerlegung in BCNF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in Boyce-Codd Normalform, wenn alle $\mathcal{R}_i$ in Boyce-Codd Normalform sind.

Zerlegung in BCNF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in Boyce-Codd Normalform, wenn alle $\mathcal{R}_i$ in Boyce-Codd Normalform sind.

- Gegeben Relationenschema $\mathcal{R}$ mit FDs F.
- Gesucht: Zerlegung in Teilschemata $\mathcal{R}_1 \ldots \mathcal{R}_n$, für die gilt:
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist verlustlos,
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist abhängigkeitstreu,
  - alle $\mathcal{R}_1 \ldots \mathcal{R}_n$ sind in Boyce-Codd Normalform.

Zerlegung in BCNF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in Boyce-Codd Normalform, wenn alle $\mathcal{R}_i$ in Boyce-Codd Normalform sind.

- Gegeben Relationenschema $\mathcal{R}$ mit FDs F.
- Gesucht: Zerlegung in Teilschemata $\mathcal{R}_1 \ldots \mathcal{R}_n$, für die gilt:
    - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist verlustlos,
    - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist abhängigkeitstreu,
    - alle $\mathcal{R}_1 \ldots \mathcal{R}_n$ sind in Boyce-Codd Normalform.
- nur die BCNF garantiert völlige Anomaliefreiheit

Zerlegung in BCNF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in Boyce-Codd Normalform, wenn alle $\mathcal{R}_i$ in Boyce-Codd Normalform sind.

- Gegeben Relationenschema $\mathcal{R}$ mit FDs F.
- Gesucht: Zerlegung in Teilschemata $\mathcal{R}_1 \ldots \mathcal{R}_n$, für die gilt:
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist verlustlos,
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist abhängigkeitstreu,
  - alle $\mathcal{R}_1 \ldots \mathcal{R}_n$ sind in Boyce-Codd Normalform.
- nur die BCNF garantiert völlige Anomaliefreiheit
- Problem: eine verlustlose Zerlegung in BCNF ist immer möglich, aber die Abhängigkeitstreue ist nicht immer erzielbar

Zerlegung in BCNF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in Boyce-Codd Normalform, wenn alle $\mathcal{R}_i$ in Boyce-Codd Normalform sind.

- Gegeben Relationenschema $\mathcal{R}$ mit FDs F.
- Gesucht: Zerlegung in Teilschemata $\mathcal{R}_1 \ldots \mathcal{R}_n$, für die gilt:
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist verlustlos,
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist abhängigkeitstreu,
  - alle $\mathcal{R}_1 \ldots \mathcal{R}_n$ sind in Boyce-Codd Normalform.
- nur die BCNF garantiert völlige Anomaliefreiheit
- Problem: eine verlustlose Zerlegung in BCNF ist immer möglich, aber die Abhängigkeitstreue ist nicht immer erzielbar
- in der Praxis: 3NF wegen Abhängigkeitstreue

Zerlegung in BCNF

Theorem

Ein Relationenschema $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ ist in Boyce-Codd Normalform, wenn alle $\mathcal{R}_i$ in Boyce-Codd Normalform sind.

- Gegeben Relationenschema $\mathcal{R}$ mit FDs F.
- Gesucht: Zerlegung in Teilschemata $\mathcal{R}_1 \ldots \mathcal{R}_n$, für die gilt:
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist verlustlos,
  - Zerlegung in $\mathcal{R}_1 \ldots \mathcal{R}_n$ ist abhängigkeitstreu,
  - alle $\mathcal{R}_1 \ldots \mathcal{R}_n$ sind in Boyce-Codd Normalform.
- nur die BCNF garantiert völlige Anomaliefreiheit
- Problem: eine verlustlose Zerlegung in BCNF ist immer möglich, aber die Abhängigkeitstreue ist nicht immer erzielbar
- in der Praxis: 3NF wegen Abhängigkeitstreue
- Lösung: Dekompositionsalgorithmus (ohne Abhängigkeitstreue)

Dekompositionsalgorithmus

- Starte mit $Z = \{\mathcal{R}\}$
- Solange es ein Relationenschema $\mathcal{R}_i$ in Z gibt, das die BCNF verletzt d.h. es gibt eine FD $\alpha \to \beta$ in $\mathcal{R}_i$ mit $\beta \not\subseteq \alpha$ und $\neg(\alpha \to \mathcal{R}_i)$:

Dekompositionsalgorithmus

- Starte mit $Z = \{\mathcal{R}\}$
- Solange es ein Relationenschema $\mathcal{R}_i$ in Z gibt, das die BCNF verletzt d.h. es gibt eine FD $\alpha \to \beta$ in $\mathcal{R}_i$ mit $\beta \not\subseteq \alpha$ und $\neg(\alpha \to \mathcal{R}_i)$:

  Wähle eine solche FD $\alpha \to \beta$ aus und zerlege wie folgt:

  - $\mathcal{R}_{i_1} := (\alpha \cup \beta)$ und $\mathcal{R}_{i_2} := \mathcal{R}_i - \beta$
  - Entferne $\mathcal{R}_i$ aus Z und füge $\mathcal{R}_{i_1}$ und $\mathcal{R}_{i_2}$ ein:

$$Z := (Z - \{\mathcal{R}_i\}) \cup \{\mathcal{R}_{i_1}\} \cup \{\mathcal{R}_{i_2}\}$$

Dekompositionsalgorithmus

- Starte mit $Z = \{\mathcal{R}\}$
- Solange es ein Relationenschema $\mathcal{R}_i$ in Z gibt, das die BCNF verletzt d.h. es gibt eine FD $\alpha \to \beta$ in $\mathcal{R}_i$ mit $\beta \not\subseteq \alpha$ und $\neg(\alpha \to \mathcal{R}_i)$:
  Wähle eine solche FD $\alpha \to \beta$ aus und zerlege wie folgt:
  - $\mathcal{R}_{i_1} := (\alpha \cup \beta)$ und $\mathcal{R}_{i_2} := \mathcal{R}_i - \beta$
  - Entferne $\mathcal{R}_i$ aus Z und füge $\mathcal{R}_{i_1}$ und $\mathcal{R}_{i_2}$ ein:

$$Z := (Z - \{\mathcal{R}_i\}) \cup \{\mathcal{R}_{i_1}\} \cup \{\mathcal{R}_{i_2}\}$$

$\mathcal{R}_i$

$\alpha \to \beta$ und weitere FDs

$\mathcal{R}_{i_1} = (\alpha \cup \beta)$
$\{\alpha' \to \beta' \mid (\alpha' \cup \beta') \subseteq \mathcal{R}_{i_1}\}$

$\mathcal{R}_{i_2} = (\mathcal{R}_i - \beta)$
$\{\alpha' \to \beta' \mid (\alpha' \cup \beta') \subseteq \mathcal{R}_{i_2}\}$

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 1)

$$\mathcal{R} = \text{Städte}\{[\text{Ort, BLand, LH, EW}]\}$$
$$\text{F} = \{\text{Ort BLand} \rightarrow \text{EW}, \text{BLand} \rightarrow \text{LH}, \text{LH} \rightarrow \text{BLand}\}$$

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 1)

$$\mathcal{R} = \text{Städte}\{[\text{Ort, BLand, LH, EW}]\}$$

$$\text{F} = \{\text{Ort BLand} \rightarrow \text{EW, BLand} \rightarrow \text{LH, LH} \rightarrow \text{BLand}\}$$

Schlüssel: {Ort, BLand}, {Ort, LH}

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 1)

$\mathcal{R}$ = Städte{[Ort, BLand, LH, EW]}

F = {Ort BLand → EW, BLand → LH, LH → BLand}

Schlüssel: {Ort, BLand}, {Ort, LH}

- $Z := \mathcal{R}$
- BLand → LH verletzt die BCNF

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 1)

$\mathcal{R}$ = Städte{[Ort, BLand, LH, EW]}

F = {Ort BLand $\rightarrow$ EW, BLand $\rightarrow$ LH, LH $\rightarrow$ BLand}

Schlüssel: {Ort, BLand}, {Ort, LH}

- $Z := \mathcal{R}$
- BLand $\rightarrow$ LH verletzt die BCNF
  - $\mathcal{R}_1 :=$ (BLand, LH) und $\mathcal{R}_2 :=$ (Ort, BLand, EW)

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 1)

$\mathcal{R}$ = Städte{[Ort, BLand, LH, EW]}

F = {Ort BLand → EW, BLand → LH, LH → BLand}

Schlüssel: {Ort, BLand}, {Ort, LH}

- $Z := \mathcal{R}$
- BLand → LH verletzt die BCNF
  - $\mathcal{R}_1$:= (BLand, LH) und $\mathcal{R}_2$:= (Ort, BLand, EW)
  - Z := {BLand, LH} ∪ {Ort, BLand, EW}

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 1)

$\mathcal{R}$ = Städte{[Ort, BLand, LH, EW]}

F = {Ort BLand $\rightarrow$ EW, BLand $\rightarrow$ LH, LH $\rightarrow$ BLand}

Schlüssel: {Ort, BLand}, {Ort, LH}

- $Z := \mathcal{R}$
- BLand $\rightarrow$ LH verletzt die BCNF
  - $\mathcal{R}_1$:= (BLand, LH) und $\mathcal{R}_2$:= (Ort, BLand, EW)
  - Z := {BLand, LH} $\cup$ {Ort, BLand, EW}
- die Zerlegung ist verlustfrei und abhängigkeitstreu

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 2)

$$\mathcal{R} = \text{PLZListe}\{[\text{Straße, Ort, BLand, PLZ}]\}$$
$$F = \{\text{PLZ} \rightarrow \text{Ort BLand, Straße Ort BLand} \rightarrow \text{PLZ}\}$$

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 2)

$$\mathcal{R} = \text{PLZListe}\{[\text{Straße, Ort, BLand, PLZ}]\}$$
$$F = \{\text{PLZ} \rightarrow \text{Ort BLand, Straße Ort BLand} \rightarrow \text{PLZ}\}$$

Schlüssel:

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 2)

$\mathcal{R}$ = PLZListe{[Straße, Ort, BLand, PLZ]}

F = {PLZ $\rightarrow$ Ort BLand, Straße Ort BLand $\rightarrow$ PLZ}

Schlüssel: {Straße, Ort, BLand}, {PLZ, Straße},

- $Z := \mathcal{R}$

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 2)

$$\mathcal{R} = \text{PLZListe}\{[\text{Straße, Ort, BLand, PLZ}]\}$$
$$\text{F} = \{\text{PLZ} \rightarrow \text{Ort BLand, Straße Ort BLand} \rightarrow \text{PLZ}\}$$

Schlüssel: {Straße, Ort, BLand}, {PLZ, Straße},

- $Z := \mathcal{R}$
- {PLZ $\rightarrow$ Ort BLand} verletzt die BCNF

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 2)

$$\mathcal{R} = \text{PLZListe}\{[\text{Straße, Ort, BLand, PLZ}]\}$$
$$\text{F} = \{\text{PLZ} \rightarrow \text{Ort BLand}, \text{Straße Ort BLand} \rightarrow \text{PLZ}\}$$

Schlüssel: {Straße, Ort, BLand}, {PLZ, Straße},

- $Z := \mathcal{R}$
- {PLZ → Ort BLand} verletzt die BCNF
  - $\mathcal{R}_1 :=$ (PLZ, Ort, BLand) und $\mathcal{R}_2 :=$ (PLZ, Straße)

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 2)

$\mathcal{R}$ = PLZListe{[Straße, Ort, BLand, PLZ]}

F = {PLZ → Ort BLand, Straße Ort BLand → PLZ}

Schlüssel: {Straße, Ort, BLand}, {PLZ, Straße},

- $Z := \mathcal{R}$
- {PLZ → Ort BLand} verletzt die BCNF
  - $\mathcal{R}_1 :=$ (PLZ, Ort, BLand) und $\mathcal{R}_2 :=$ (PLZ, Straße)
  - $Z :=$ Orte:{PLZ, Ort, BLand} ∪ Straßen:{PLZ, Straße}

Dekompositionsalgorithmus

Beispiel (Dekompositionsalgorithmus 2)

$$\mathcal{R} = \text{PLZListe}\{[\text{Straße, Ort, BLand, PLZ}]\}$$
$$F = \{\text{PLZ} \rightarrow \text{Ort BLand}, \text{Straße Ort BLand} \rightarrow \text{PLZ}\}$$

Schlüssel: {Straße, Ort, BLand}, {PLZ, Straße},

- $Z := \mathcal{R}$
- {PLZ → Ort BLand} verletzt die BCNF
  - $\mathcal{R}_1 :=$ (PLZ, Ort, BLand) und $\mathcal{R}_2 :=$ (PLZ, Straße)
  - $Z :=$ Orte:{PLZ, Ort, BLand} ∪ Straßen:{PLZ, Straße}
- die Zerlegung ist verlustfrei aber die Abhängigkeit {Straße Ort BLand → PLZ} ist verloren gegangen

Lernziele

- Was sind FDs?
  - Wann ist eine FD erfüllt, wie kann man das überprüfen?
- Was sind (Super)Schlüssel?
  - Wie kann ich sie erkennen/überprüfen?
- Was ist die Attributhülle/Hülle von FDs?
  - Wie kann man sie berechnen?
- Wann sind zwei Mengen von FDs äquivalent?
- Was sind die Armstrong Axiome?
  - Wozu braucht man sie, welche gibt es?
- Was ist die kanonische Überdeckung?
  - Wie kann man sie berechnen?
- Was für Arten von Anomalien gibt es?
- Was ist eine verlustlose/abhängigkeitstreue Zerlegung?
- Welche Normalformen gibt es?
  - Wann sind sie erfüllt, und wie kann man sie berechnen?