

Grundlagen Datenbanken: Übung 08

Tanmay Deshpande

Gruppe 20 & 21

ge94vem@mytum.de



QR-Code für die Folien



Wiederholung

Woche 08



Mehrwertige Abhängigkeiten (MVDs)

| R | | |
|-----------------------|------------------------|-------------------------|
| α A1 ... Ai | β Ai+1 ... Aj | γ Aj+1 ... An |
| a1 ... ai | ai+1 ... aj | aj+1 ... an |
| a1 ... ai | bi+1 ... bj | bj+1 ... bn |
| a1 ... ai | bi+1 ... bj | aj+1 ... an |
| a1 ... ai | ai+1 ... aj | bj+1 ... bn |

$\alpha \twoheadrightarrow \beta$ gilt genau dann wenn

- Wenn es zwei Tupel t1 und t2 mit gleichen α -Werten gibt
- Dann muss es auch zwei Tupel t3 und t4 geben mit
 - $t3.\alpha = t4.\alpha = t1.\alpha = t2.\alpha$
 - $t3.\beta = t1.\beta, t4.\beta = t2.\beta$
 - $t3.\gamma = t2.\gamma, t4.\gamma = t1.\gamma$

Jede FD ist auch eine MVD, aber nicht umgekehrt!

Mehrwertige Abhängigkeiten (MVDs)

- In anderen Worten, wenn $\alpha \twoheadrightarrow \beta$ gilt, bei zwei Tupeln mit gleichem α –Wert, kann man die β –Werte umtauschen und die resultierenden Tupel müssen auch in der Relation sein.
- Bsp: Wir wollen, dass $\text{Name} \twoheadrightarrow \text{Beer}$ gilt

| Name | Beer | Wine |
|-------|-------|-------|
| Alina | Dark | Red |
| Alina | Light | White |

- Wir müssen die Tupel (Alina, Dark, White) und (Alina, Light, Red) hinzufügen
- Die neue Relation sieht so aus:

| Name | Beer | Wine |
|-------|-------|-------|
| Alina | Dark | Red |
| Alina | Light | White |
| Alina | Dark | White |
| Alina | Light | Red |

Inferenzregeln für MVDs

- *Reflexivität*: $\beta \subseteq \alpha \Rightarrow \alpha \rightarrow \beta$
- *Verstärkung*: Sei $\alpha \rightarrow \beta$. Dann gilt $\gamma\alpha \rightarrow \gamma\beta$.
- *Transitivität*: Sei $\alpha \rightarrow \beta$ und $\beta \rightarrow \gamma$. Dann gilt $\alpha \rightarrow \gamma$.
- *Komplement*: $\alpha \twoheadrightarrow \beta$. Dann gilt $\alpha \twoheadrightarrow \mathcal{R} - \beta - \alpha$.
- *Mehrwertige Verstärkung*: Sei $\alpha \twoheadrightarrow \beta$ und $\delta \subseteq \gamma$. Dann gilt $\gamma\alpha \twoheadrightarrow \delta\beta$.
- *Mehrwertige Transitivität*: Sei $\alpha \twoheadrightarrow \beta$ und $\beta \twoheadrightarrow \gamma$. Dann gilt $\alpha \twoheadrightarrow \gamma - \beta$.
- *Verallgemeinerung*: Sei $\alpha \rightarrow \beta$. Dann gilt $\alpha \twoheadrightarrow \beta$.
- *Koaleszenz*: Sei $\alpha \twoheadrightarrow \beta$ und $\gamma \subseteq \beta$. Existiert ein $\delta \subseteq \mathcal{R}$, so daß $\delta \cap \beta = \emptyset$ und $\delta \rightarrow \gamma$, gilt $\alpha \rightarrow \gamma$.
- *Mehrwertige Vereinigung*: sei $\alpha \twoheadrightarrow \beta$ und $\alpha \twoheadrightarrow \gamma$. Dann gilt $\alpha \twoheadrightarrow \gamma\beta$.
- *Schnittmenge*: Sei $\alpha \twoheadrightarrow \beta$ und $\alpha \twoheadrightarrow \gamma$. Dann gilt $\alpha \twoheadrightarrow \beta \cap \gamma$.
- *Differenz*: Sei $\alpha \twoheadrightarrow \beta$ und $\alpha \twoheadrightarrow \gamma$. Dann gilt $\alpha \twoheadrightarrow \beta - \gamma$ und $\alpha \twoheadrightarrow \gamma - \beta$.

Kanonische Überdeckung

- Die kanonische Überdeckung (F_c) einer Menge an FDs (F) muss die folgende drei Eigenschaften erfüllen:

1. $F_c \equiv F \Rightarrow F_c^+ \equiv F^+$

2. In F_c existieren keine überflüssige FDs. Das bedeutet:

$\forall A \in \alpha : \left(F_c - (\alpha \rightarrow \beta) \cup ((\alpha - A) \rightarrow \beta) \right)$ ist nicht äquivalent zu F_c

$\forall B \in \beta : \left(F_c - (\alpha \rightarrow \beta) \cup (\alpha \rightarrow (\beta - B)) \right)$ ist nicht äquivalent zu F_c

3. Jede linke Seite einer funktionalen Abhängigkeit in F_c ist einzigartig

- Gegeben seien die FDs:

A→C

ACD→B

AB→CE

Berechne die kanonische Überdeckung wie folgt:

- 1. Linksreduktion: $\forall A \in \alpha$ überprüfe, ob $\beta \subseteq AttrHülle(F, \alpha - A)$

A→C

AD→B

AB→CE

- 2. Rechtsreduktion: $\forall B \in \beta$ überprüfe, ob $B \subseteq AttrHülle(F - (\alpha \rightarrow \beta) \cup (\alpha \rightarrow (\beta - B)), \alpha)$

A→C

AD→B

AB→E

- 3. FDs der Form $\alpha \rightarrow \emptyset$ entfernen:

A→C

AD→B

AB→E

- 4. FDs zusammenfassen:

A→C

AD→B

AB→E

Normalisierung

- Warum brauchen wir es überhaupt? Um schlechte Relationenschemata zu verbessern
- Bsp:

| 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 | 5259 | Der Wiener Kreis | 2 |
| 2137 | Kant | C4 | 7 | 4630 | Die 3 Kritiken | 4 |

Update-Anomalien

- Sokrates zieht um, von Raum 226 in R. 338. Was passiert?

Einfüge-Anomalien

- Neue/r Prof ohne Vorlesungen?

Löschanomalien

- Letzte Vorlesung einer/s Profs wird gelöscht? Was passiert?

- Lösung: versuche eine „schlechte“ Relationenschema in verschiedenen Teilrelationen zu zerlegen, ohne irgendetwelche Informationen aus der ursprünglichen Relation zu verlieren (Verlustlosigkeit) oder irgendetwelche FDs zu verlieren (Abhängigkeitserhaltung)
- Algorithmen wie Syntheselalgorithmus und Dekompositionsalgorithmus helfen uns, diese Ziele zu erreichen und die Relationen in eine bessere Version (Normalform) zu bringen

Erste Normalform (1. NF)

- Nur atomare Attribute

Zweite Normalform (2. NF)

- 1.NF ist erfüllt, und...
- Jedes Nichtschlüssel-Attribut voll funktional abhängig von jedem Kandidatenschlüssel der Relation
- In anderen Worten, für alle Attribute b , die auf der rechten Seite einer FD sind, muss gelten:
 1. b ist nicht von einer echten Teilmenge eines Kandidatenschlüssels abhängig oder
 2. b ist Teil eines Kandidatenschlüssels

Dritte Normalform (3. NF)

- 2. NF ist erfüllt, und...
- Für jede FD der Form $\alpha \rightarrow \beta$ gelten eine der folgenden Bedingungen:
 1. die FD ist trivial ($\alpha \subseteq \beta$)
 2. α ist Superschlüssel
 3. Jedes Attribut in β ist in einem Kandidatenschlüssel enthalten

Boyce-Codd Normalform (BCNF)

- 3.NF ist erfüllt, und...
- Für jede FD der Form $\alpha \rightarrow \beta$ gelten eine der folgenden Bedingungen:
 1. die FD ist trivial ($\alpha \subseteq \beta$)
 2. α ist Superschlüssel

Vierte Normalform (4.NF)

- BCNF ist erfüllt, und...
- Für jede MVD der Form $\alpha \twoheadrightarrow \beta$ gelten eine der folgenden Bedingungen:
 1. die MVD ist trivial ($\alpha \subseteq \beta$)
 2. α ist Superschlüssel

Synthesealgorithmus

- Überführt die Relation in 3. NF
- Schritte:
 1. Kanonische Überdeckung berechnen
 2. Relationsschemata aus jeder FD der kanonischen Überdeckung formen
 3. Schlüssel hinzufügen
 4. Redundante Schemata eliminieren
- Verlustlos und abhängigkeitsbewahrend

Dekompositionsalgorithmus (BCNF)

- Überführt die Relation in BCNF
- Schritte:
 1. Initialisierung: Starte mit $Z = \{R\}$
 2. Solange es eine FD im Schema $R_i \in Z$ gibt, die BCNF verletzt:
 - * Zerlege R_i in $R_{i1} = \alpha \cup \beta$ und $R_{i2} = R_i - \beta$
 - * Entferne R_i aus Z und füge R_{i1} und R_{i2} ein
- Verlustlos aber **nicht** abhängigkeitsbewahrend

Dekompositionsalgorithmus (4.NF)

- Überführt die Relation in 4.NF
- Schritte:
 1. Initialisierung: Starte mit $Z = \{R\}$
 2. Solange es eine MVD im Schema $R_i \in Z$ gibt, die 4.NF verletzt:
 - * Zerlege R_i in $R_{i1} = \alpha \cup \beta$ und $R_{i2} = R_i - \beta$
 - * Entferne R_i aus Z und füge R_{i1} und R_{i2} ein
- Verlustlos aber **nicht** abhängigkeitsbewahrend

Aufgaben

Woche 08



Aufgabe 01

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

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

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

Lösungsvorschlag 01

- Relevant für diese Aufgabe sind die folgenden MVDs:

$$1. A \twoheadrightarrow B$$

$$2. A \twoheadrightarrow C$$

$$3. B \twoheadrightarrow A$$

$$4. B \twoheadrightarrow C$$

$$5. C \twoheadrightarrow A$$

$$6. C \twoheadrightarrow B$$

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

- Alle anderen MVDs der Form $AB \twoheadrightarrow C$ (mit zwei oder mehr Attribute auf der linken Seite) gelten immer und sind trivial
- Bei uns, sind die MVDs 1 und 2 verletzt. Die restlichen gelten bereits, da A immer 1 ist
- Wir fügen die Tupel (1, x, 9) und (1, y, 8) hinzu, damit sie gelten
- Unsere Tabelle sieht dann so aus:

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

Aufgabe 02

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

Lösungsvorschlag 02

- 1. NF:

- $AB \rightarrow C$
- $B \rightarrow D$

- 3.NF:

- $BC \rightarrow AD$
- $D \rightarrow C$

- 4.NF:

- $AB \rightarrow CD$
- $BC \rightarrow AD$

- 2. NF:

- $AB \rightarrow C$
- $C \rightarrow D$

- BCNF:

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

Aufgabe 03

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

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

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

Lösungsvorschlag 3a

- Gegeben das Schema $R = \{A, B, C, D, E, F, G\}$ und die FDs:

- $A \rightarrow BC$
- $C \rightarrow DA$
- $E \rightarrow ABC$
- $F \rightarrow CD$
- $CD \rightarrow BEF$

- Anhand der gegebenen FDs, kann man die Attributehülle von A wie folgt berechnen:

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

- Es enthält alle Attribute außer G

Lösungsvorschlag 3b

- Gegeben das Schema $R = \{A, B, C, D, E, F, G\}$ und die FDs:
 1. $A \rightarrow BC$
 2. $C \rightarrow DA$
 3. $E \rightarrow ABC$
 4. $F \rightarrow CD$
 5. $CD \rightarrow BEF$
- Kandidatenschlüssel der Relation sind: AG, CG, EG und FG

Lösungsvorschlag 3c

- Gegeben ist die Ausgangsmenge $F = \{A \rightarrow BC, C \rightarrow DA, E \rightarrow ABC, F \rightarrow CD, CD \rightarrow BEF\}$.
- Berechnung der kanonischen Überdeckung:
- Linksreduktion:

$$A \rightarrow BC$$

$$C \rightarrow DA$$

$$E \rightarrow ABC$$

$$F \rightarrow CD$$

$$C \rightarrow BEF$$

- Rechtsreduktion:

$$A \rightarrow C$$

$$C \rightarrow \emptyset$$

$$E \rightarrow A$$

$$F \rightarrow CD$$

$$C \rightarrow BEF$$

Lösungsvorschlag 3c (contd...)

- Entferne FDs der Form $\alpha \rightarrow \emptyset$:
A \rightarrow C
E \rightarrow A
F \rightarrow CD
C \rightarrow BEF
- FDs zusammenfassen:
Keine FDs müssen zusammengefasst werden
- Kanonische Überdeckung:
A \rightarrow C
E \rightarrow A
F \rightarrow CD
C \rightarrow BEF

Lösungsvorschlag 3d

- Die kanonische Überdeckung kann in 3. NF nach dem Synthesealgorithmus überführt werden:

Relationenschema erstellen:

$$R_1 = \{\underline{A}, C\}$$

$$R_2 = \{\underline{E}, A\}$$

$$R_3 = \{\underline{F}, C, D\}$$

$$R_4 = \{\underline{C}, B, E, F\}$$

- Keine der Relationen enthält einen Kandidatenschlüssel. Wir fügen die folgende Relation hinzu:

$$R_\kappa = \{\underline{A}, G\}$$

- Es gibt keine überflüssige Relationen. Wir haben das gewünschte Ergebnis

Aufgabe 04

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

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

FDs:

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

Lösungsvorschlag 4

- Gegeben:
 1. $B \rightarrow DA$
 2. $DEF \rightarrow B$
 3. $C \rightarrow EA$
- Starte mit $Z = \{R\}$
- $B \rightarrow DA$ verletzt BCNF, R muss zerlegt werden
- | | |
|---|---|
| • R1: $\{[A, \underline{B}, D]\}$ mit FDs: $B \rightarrow DA$ | • R2: $\{[\underline{B}, C, E, \underline{F}]\}$ mit FDs: $C \rightarrow E$ |
|---|---|
- R1 erfüllt BCNF. R2 nicht, deshalb muss es weiter zerlegt werden
- | | |
|--|---|
| • R2.1: $\{[\underline{C}, E]\}$ mit FDs: $C \rightarrow E$ | • R2.2: $\{[\underline{B}, C, \underline{F}]\}$ mit FDs: keine |
|--|---|
- Ergebnis = R1: $\{[A, \underline{B}, D]\}$, R2.1: $\{[\underline{C}, E]\}$, R2.2: $\{[\underline{B}, C, \underline{F}]\}$

Aufgabe 05

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

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

$$AB \rightarrow CDE$$

$$B \twoheadrightarrow D$$

$$C \twoheadrightarrow DE$$

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

Lösungsvorschlag 5

- Gegeben: $AB \rightarrow CDE$
 $B \twoheadrightarrow D$
 $C \twoheadrightarrow DE$
- Starte mit $Z = \{R\}$
- $B \twoheadrightarrow D$ verletzt 4.NF, R muss zerlegt werden
- R1: $\{\underline{B}, D\}$ mit FDs oder MVDs: keine
 R2: $\{\underline{A}, B, C, E\}$ mit FDs oder MVDs: $AB \rightarrow CE, C \twoheadrightarrow AB$
- R1 erfüllt 4.NF. R2 nicht, deshalb muss es weiter zerlegt werden
- R2.1: $\{\underline{A}, B, C\}$ mit FDs oder MVDs: $AB \rightarrow C$
 R2.2: $\{\underline{C}, E\}$ mit FDs oder MVDs: keine
- Ergebnis: R1: $\{\underline{B}, D\}$, R2.1: $\{\underline{A}, B, C\}$ und R2.2: $\{\underline{C}, E\}$

Lösungsvorschlag 5 (contd...)

- Gegeben: $AB \rightarrow CDE$
 $B \twoheadrightarrow D$
 $C \twoheadrightarrow DE$
- Starte mit $Z = \{R\}$
- $C \twoheadrightarrow DE$ verletzt 4.NF, R muss zerlegt werden
- R1: $\{[\underline{C}, \underline{D}, \underline{E}] \}$ mit FDs oder MVDs:
keine
- R2: $\{[\underline{A}, \underline{B}, C] \}$ mit FDs oder MVDs:
 $AB \rightarrow C$
- Die beiden Relationen erfüllen schon 4.NF
- Ergebnis: R1: $\{[\underline{C}, \underline{D}, \underline{E}] \}$ und R2: $\{[\underline{A}, \underline{B}, C] \}$