

Übung 7: Relationaler Datenbankentwurf und Normalisierung

Aufgabe 1: Normalisierung

Gegeben folgende Relationen R und S in erster Normalform mit funktionalen Abhängigkeiten $F = F_G$:

$R = (A, B, C, D, E, F)$ mit
 $A, B \rightarrow C, D, E$
 $D \rightarrow F$ und
 $\{A, B\}$ ist Schlüssel.

$S = (V, W, X, Y, Z)$ mit
 $V \rightarrow W, Z$
 $W, Z \rightarrow V, X, Y$
 $Y \rightarrow Z$ und
 $\{V\}, \{W, Z\}, \{W, Y\}$ sind Schlüssel..

1. Zeigen Sie möglichst knapp, dass R nicht der dritten Normalform (3NF) entspricht.
Um diese Eigenschaft zu überprüfen, müssen wir zunächst die Schlüssel kennen. Einziger Schlüssel ist A, B , siehe letzte Übung. Um zu zeigen, dass eine Relation nicht der 3NF entspricht, reicht es dass ein nicht-Schlüssel Attribut transitiv vom Schlüssel abhängt. Das ist durch $D \rightarrow F$ der Fall. F ist nicht prim, D ist kein Schlüssel und F hängt transitiv ab vom Schlüssel A, B .
2. Zerlegen Sie R durch Dekomposition in Relationen die der dritten Normalform (3NF) entsprechen. Sind die entstandenen Relationen abhängigkeits- und verbundtreu?
Die funktionale Abhängigkeit $D \rightarrow F$ verletzt die 3NF Eigenschaft. Zerlegung:

$$\begin{aligned} R_1 &= \{D\}_F^+ = \{D, F\} \\ R_2 &= R - R_1 \cup \{D\} \\ &= \{A, B, C, D, E, F\} - \{D, F\} \cup \{D\} \\ &= \{A, B, C, D, E\}. \end{aligned}$$

Eine Dekomposition eines Schemas das nicht 3NF entspricht, ist immer abhängigkeits- und verbundtreu.

3. Zerlegen Sie R mit dem Syntheseverfahren. Sind die entstandenen Relationen abhängigkeits- und verbundtreu?
 - (a) Berechne Kanonische Überdeckung. Ist bereits gegeben: $F_G = F$
 - (b) Erzeuge für jede linke Seite eine Relation aus den Attributen der linken Seite und den Attributen die man mit der linken Seite erreichen kann:

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

- (c) Enthält eine der erzeugten Relationen den Schlüssel? Ja, R_1 enthält $\{A, B\}$.
- (d) Ist eine der Relationen vollständig in einer der anderen enthalten? Nein. Fertig.

Das Syntheseverfahren ist immer abhängigkeits- und verbundtreu.

4. Zeigen Sie möglichst knapp, dass S nicht der Boyce-Codd-Normalform (BCNF) entspricht.
Um zu zeigen, dass eine Relation nicht der BCNF entspricht, reicht es zu zeigen dass eine linke Seite einer funktionalen Abhängigkeit kein Superschlüssel ist. Das ist der Fall für $Y \rightarrow Z$, denn $\{Y\}_F^+ = \{Y, Z\} \neq S$.
5. Zerlegen Sie S durch Dekomposition in Relationen die der Boyce-Codd-Normalform (BCNF) entsprechen. Sind die entstandenen Relationen abhängigkeits- und verbundtreu? Ist die Zerlegung nützlich?

Die funktionale Abhängigkeit $Y \rightarrow Z$ verletzt die BCNF Eigenschaft. Zerlegung:

$$\begin{aligned} S_1 &= \{Y\}_F^+ = \{Y, Z\} \\ S_2 &= S - S_1 \cup \{Y\} \\ &= \{V, W, X, Y, Z\} - \{Y, Z\} \cup \{Y\} \\ &= \{V, W, X, Y\} \end{aligned}$$

Eine Dekomposition eines Schemas das nicht BCNF entspricht ist immer verbundtreu. Überprüfe Abhängigkeitstreue:

$$\begin{aligned} F(S_1) &= \{(Y \rightarrow Z)\} \\ F(S_2) &= \{(V \rightarrow W)\} \\ F &\neq F(S_1) \cup F(S_2), \text{ denn es ist} \\ F \cap (S_1 \cup S_2) &= \{(V \rightarrow Z), (W, Z \rightarrow V, X, Y)\} \end{aligned}$$

Die Zerlegung ist nicht abhängigkeits- und damit unbrauchbar.

6. Zerlegen Sie S mit dem Syntheseverfahren. Sind die entstandenen Relationen abhängigkeits- und verbundtreu?

- (a) Berechne Kanonische Überdeckung. Ist bereits gegeben: $F_C = F$
- (b) Erzeuge für jede linke Seite eine Relation aus den Attributen der linken Seite und den Attributen die man mit der linken Seite erreichen kann:

$$\begin{aligned} S_1 &= \{V, W, Z\} \\ S_2 &= \{V, W, X, Y, Z\} \\ S_3 &= \{Y, Z\} \end{aligned}$$

- (c) Enthält eine der erzeugten Relationen den Schlüssel? Ja, S_2 enthält $\{V\}$, $\{W, Z\}$ und $\{W, Y\}$.
- (d) Ist eine der Relationen vollständig in einer der anderen enthalten? Ja. Entferne S_1 und S_3

Das Syntheseverfahren ist immer abhängigkeits- und verbundtreu, allerdings hat uns das Verfahren keine Verbesserung gebracht.

Anmerkung: In der Musterlösung der letzten Übung wurde eine alternative Lösung für das Cover F_C für S angegeben. Wird diese alternative Lösung für das Syntheseverfahren herangezogen, ergibt sich folgende Zerlegung:

$$\begin{aligned} S_1 &= \{V, W, X, Y\} \\ S_2 &= \{W, Z, V\} \\ S_3 &= \{Y, Z\} \end{aligned}$$

Diese ist ebenfalls abhängigkeits- und verbundtreu und besser als die andere Lösung zu S , da keine Relation mit allen ursprünglichen Attributen erzeugt wurde.

Aufgabe 2: Vereinfachtes Syntheseverfahren

Gegeben folgende Relation R und funktionale Abhängigkeiten F :

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

1. Offensichtlich entspricht R nicht der zweiten Normalform. Der einzige Schlüssel ist $\{A, D\}$ und alle Primattribute hängen nur von einer echten Teilmenge des Schlüssels ab. An welcher Stelle des Syntheseverfahrens wird eine Schwierigkeit mit der Laufzeit auftreten?

Nach erster Zerlegung in

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

$$R_2 = \{D, E, F\}$$

beinhaltet keine der entstandenen Relationen einen Schlüssel, denn es ist $\{R_1\}_F^+ \neq R$ und $\{R_2\}_F^+ \neq R$. Somit müssen wir den Schlüssel suchen und das kostet uns im schlimmsten Fall exponentiell viel Laufzeit. In diesem Fall haben wir Glück und unsere Heuristik liefert uns die passenden Schlüsselattribute.

2. Wenn Sie das vereinfachte Syntheseverfahren mit der zusätzlichen Regel $A, B, C, D, E, F \rightarrow \delta$ anwenden würden, was ist die Verbesserung und was ist als Ergebnis zu erwarten?

Wird die zusätzliche Regel $A, B, C, D, E, F \rightarrow \delta$ eingefügt, wird sie sich zu $A, D \rightarrow \delta$ verkürzen. Dadurch entsteht eine neue Relation R_3 mit $\{A, B\}_F^+ = R$. Damit haben wir das Syntheseverfahren erfolgreich angewandt ohne einen Schlüssel zu berechnen.

3. Führen Sie das vereinfachte Syntheseverfahren für R durch.

(a) Berechne Kanonische Überdeckung.

i. $F_C = F$

ii. $F_C = \text{SPLITTING}(F_C) =$

$$A \rightarrow B$$

$$A \rightarrow C$$

$$D \rightarrow E$$

$$D \rightarrow F$$

$$A, B, C, D, E, F \rightarrow \delta$$

iii. Minimiere linke Seiten

$$\delta \in \{A, D\}_F^+ = \{R\}$$

$$\Rightarrow F_C = F_C - \{A, B, C, D, E, F \rightarrow \delta\} \cup \{A, D \rightarrow \delta\}$$

iv. Entferne überflüssige FDs: keine.

v. Zusammenfassung linker Seiten

$$A \rightarrow B, C$$

$$D \rightarrow E, F$$

$$A, D \rightarrow \delta$$

- (b) Erzeuge für jede linke Seite eine Relation aus den Attributen der linken Seite und den Attributen die man mit der linken Seite erreichen kann (ohne δ):

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

$$R_2 = \{D, E, F\}$$

$$R_3 = \{A, D\}$$

(c) Enthält eine der erzeugten Relationen den Schlüssel? Ja, $R_3: \{A, D\}_F^+ = R$.

(d) Ist eine der Relationen vollständig in einer der anderen enthalten? Nein. Fertig.

Zusatzaufgabe: Vergleich Synthese und Dekomposition

Gegeben folgende Relationen S_1 und S_2 in erster Normalform mit funktionalen Abhängigkeiten:

$S_1 = (A, P, H, R, O, D, I, T, E)$ mit
 $R \rightarrow O, D$
 $O \rightarrow A, H, P, R$ und
 $\{I, T, E, R\}, \{I, T, E, O\}$ sind Schlüssel.

$S_2 = (A, P, H, R, O, D, I, T, E)$ mit
 $R \rightarrow O$
 $O \rightarrow A, P, H, R, D$ und
 $\{I, T, E, R\}, \{I, T, E, O\}$ sind Schlüssel.

- Zerlegen Sie S_1 und S_2 durch Dekomposition in Relationen der Boyes-Codd Normalform (BCNF). Geht das?

Offensichtlich entsprechen weder S_1 noch S_2 der BCNF, da nicht alle Attribute von einem Schlüssel abhängen. Das ist z.B. für H nicht der Fall, da dieses Attribut nur von einem Teil der Schlüssel abhängt. Damit verletzen die Regeln

$$R \rightarrow O, D \in S_1$$

$$R \rightarrow O \in S_2$$

die BCNF. Nach dem Dekompositionsverfahren können S_1 und S_2 zerlegt werden in

$$\begin{aligned} S_{1a} &= \{R\}_F^+ D \\ &= \{R, O, D, A, H, P, R\} \\ S_{1b} &= S_1 - S_{1a} \cup \{R\} \\ &= \{I, T, E, R\} \end{aligned}$$

$$\begin{aligned} S_{2a} &= \{O\}_F^+ D \\ &= \{O, A, P, H, R, D\} \\ S_{2b} &= S_2 - S_{2a} \cup \{O\} \\ &= \{I, T, E, O\} \end{aligned}$$

Dekomposition ist immer Verbundtreu. Sicherheitshalber nochmal die Probe:

$$\{I, T, E, R\}_F^+ D = \{A, P, H, R, O, D, I, T, E\}$$

$$\{I, T, E, O\}_F^+ D = \{A, P, H, R, O, D, I, T, E\}$$

Die Abhängigkeitsreue ist ebenfalls gegeben, denn es sind

$$\begin{aligned} R \rightarrow O, D &\in FD_{S_{1a}} \\ O \rightarrow A, H, P, R &\in FD_{S_{1a}} \end{aligned}$$

$$\begin{aligned} R \rightarrow O &\in FD_{S_{2a}} \\ O \rightarrow A, H, P, R, D &\in FD_{S_{2a}} \end{aligned}$$

Die Dekomposition liefert und außerdem BCNF-konforme Relationen, denn in S_{1a} hängen alle Attribute von den Schlüsseln $\{R\}$ und $\{O\}$ ab, in S_{1b} hängen alle Attribute von dem Schlüssel $\{I, T, E, R\}$ ab, in S_{2a} hängen alle Attribute von den Schlüsseln $\{R\}$ und $\{O\}$ ab und in S_{2b} hängen alle Attribute von dem Schlüssel $\{I, T, E, O\}$ ab.

- Dekomponieren Sie die Relationenschema mittels des Syntheseverfahrens

$$\begin{aligned} S_{1a} &= R, O, D \\ S_{1b} &= O, A, H, P, R \end{aligned}$$

$$\begin{aligned} S_{2a} &= R, O \\ S_{2b} &= O, A, H, P, R, D \end{aligned}$$

Ist ein Schlüssel der ursprünglichen Relationen S_1 und S_2 in den Attributen der neuen Relationenschema enthalten? Nein, füge daher noch eine zusätzliche Relation hinzu, z.B.:

$$S_{1c} = I, T, E, O$$

$$S_{2c} = I, T, E, R$$

- Führen Sie das vereinfachte Syntheseverfahren durch.

Füge zusätzlich die Regel $A, P, H, R, O, D, I, T, E \rightarrow \delta$ zu den Funktionalen Abhängigkeiten hinzu. Diese Regel vereinfacht sich z.B. zu

$$I, T, E, O \rightarrow \delta$$

$$I, T, E, R \rightarrow \delta$$

Damit ergibt sich jeweils eine zusätzliche Relation

$$S_{1c} = I, T, E, O$$

$$S_{2c} = I, T, E, R$$

- Welches ist jetzt das bessere Verfahren: Dekomposition, Synthese oder vereinfachtes Syntheseverfahren?
Beim Dekompositionsverfahren entstehen nur zwei Relationen die zudem noch der BCNF entsprechen, das ist also das bessere Ergebnis. Beim Syntheseverfahren entstehen drei Relationen und für eine ist es sogar notwendig die Schlüssel zu kennen, was wieder eine hohe Laufzeitkomplexität nach sich zieht. Beim vereinfachten Syntheseverfahren ist die Laufzeitkomplexität geringer, es entstehen aber ebenfalls drei Relationen.