



Informationssysteme SS 2002

Übung 7

Beispiellösung

Aufgabe 1: Normalformen

Betrachten Sie die folgenden Relationen (Schlüssel unterstrichen):

$R(\underline{A}, B, C, D)$
 $S(\underline{E}, F, G, C, D)$

und die zusätzlich auf ihnen definierte Menge der Funktionalabhängigkeiten:

$\{ \begin{array}{l} CD \rightarrow A, \\ C \rightarrow B, \\ GC \rightarrow D, \\ G \rightarrow E \end{array} \}.$

Überprüfen Sie die nachfolgenden Aussagen auf ihre Richtigkeit und geben Sie eine kurze Begründung für Ihre Antwort:

Aus den Primärschlüsselangaben für R und S ergeben sich zwei weitere Funktionalabhängigkeiten:

$A \rightarrow BCD$
 $E \rightarrow FGCD$

a) R ist in 3NF

Nein, R ist nicht in 3NF. Wegen der Funktionalabhängigkeit (Abk.: FA) $\{CD \rightarrow A\}$ und da A ein Primärschlüssel ist, folgt, daß die Attributmenge $\{C,D\}$ ein Schlüsselkandidat ist. Da C keinen Schlüsselkandidaten enthält, und da B kein Schlüsselattribut ist, verletzt die FA $\{C \rightarrow B\}$ die 3NF-Bedingung.

b) S ist in BCNF

Ja, S ist in BCNF. Da E Schlüsselkandidat ist (wegen Primärschlüssel) und die FA $G \rightarrow E$ gilt, ist G ein weiterer Schlüsselkandidat für die Relation S.

Zur Kontrolle, ob eine Relation in BCNF ist, reicht es aus zu kontrollieren, ob die linken Seiten aller FA einen Schlüsselkandidaten enthalten. Wir brauchen also nicht alle implizierten Abhängigkeiten zu berücksichtigen, da durch die Armstrong Axiome die linken Seiten von FA entweder vergrößert werden oder von bereits bestehenden Fas übernommen werden.

Die gegebenen FA $F = \{GC \rightarrow D, G \rightarrow E, E \rightarrow FGCD\}$ erfüllen die BCNF-Bedingung.

Aufgabe 2: BCNF-Zerlegung, Abhängigkeitsbewahrung

Gegeben sei die Relation

R (Institut, Gebäude, Vorsteher, Telefon)

mit den Funktionalabhängigkeiten

$F = \{ \text{Institut} \rightarrow \text{Vorsteher} \\ \text{Vorsteher} \rightarrow \text{Telefon} \}$

a) Bestimmen Sie alle Schlüsselkandidaten.

Mit dem Algorithmus zur Bestimmung aller Schlüsselkandidaten einer Relation (siehe Skript) lässt sich leicht herausfinden, daß {Institut, Gebäude} der einzige Schlüsselkandidat ist.

b) In welcher Normalform befindet sich die Relation R?

Die Relation ist nicht in 3NF, da die FA $\text{Vorsteher} \rightarrow \text{Telefon}$ die 3NF-Bedingung verletzt. Somit ist die Relation auch nicht in BCNF.

c) Führen Sie eine BCNF-Zerlegung durch. Bestimmen Sie in jedem Zerlegungsschritt alle Schlüsselkandidaten und die Menge der Funktionalabhängigkeiten für jede der entstandenen Relationen. Prüfen Sie, ob die Zerlegung abhängigkeitsbewahrend ist.

Kurzschreibweise: $R(\underline{I}, \underline{G}, V, T)$

FA: $F = \{I \rightarrow V, V \rightarrow T\}$

Transitive Hülle von F: $F^+ = \{I \rightarrow V, V \rightarrow T, I \rightarrow T, \dots\}$

Bemerkung: Die trivialen Abhängigkeiten, die durch Reflexivität und Erweiterungen eingeführt werden, werden nicht extra aufgelistet.

Schlüsselkandidat: $\underline{I}, \underline{G}$

- eine mögliche Zerlegung:

Zerlegungsschritte:

1. Zerlegung von R entlang der FA $V \rightarrow T$:

- $R_1(\underline{V}, T)$ mit $V \rightarrow T \in F_1^+$

- $R_2(\underline{I}, \underline{G}, V)$ mit $I \rightarrow V \in F_2^+$

2. Zerlegung von R_2 entlang der FA $I \rightarrow V$:

- $R_{21}(I, V)$ mit $I \rightarrow V \in F_{21}^+$

- $R_{22}(I, G)$ nur mit trivialen FA

Es gilt: $\{I \rightarrow V, V \rightarrow T\} \subseteq \{F_1^+ \cup F_{21}^+ \cup F_{22}^+\}$
 $\rightarrow (F_1^+ \cup F_{21}^+ \cup F_{22}^+) \subseteq F^+$

Die Zerlegung ist also abhängigkeitsbewahrend.

- eine andere mögliche Zerlegung:

Zerlegungsschritte:

1. Zerlegung von R entlang der FA $I \rightarrow V$:
 - $R_1(\underline{I}, V)$ mit $\{I \rightarrow V\} \in F_1^+$
 - $R_2(\underline{I}, \underline{G}, T)$ mit $\{I \rightarrow T\} \in F_2^+$
2. Zerlegung von R_2 entlang der FA $I \rightarrow T$:
 - $R_{21}(\underline{I}, T)$ mit $\{I \rightarrow T\} \in F_{21}^+$
 - $R_{22}(\underline{I}, \underline{G})$ nur mit trivialen Abhängigkeiten

Wir können uns nun leicht überlegen, daß diese Zerlegung nicht abhängigkeitsbewahrend ist, da wir die Abhängigkeit $V \rightarrow T$ nicht ableiten können. Dies wird klar, wenn wir berücksichtigen, daß es keine Abhängigkeit gibt, bei der auf der linken Seite nur V vorkommt, und die nicht trivial ist.

Aufgabe 3: Relationaler Datenbankentwurf

Gegeben sei das relationale Schema EXAM (S, F, R, P) mit

EXAM ...Examen
 S ...Student
 F ...Fach
 R ...Rang
 P ...Prüfer

und den Funktionalabhängigkeiten

{ S, F \rightarrow P,
 R, F \rightarrow S,
 F, S \rightarrow R }.

Wir setzen hier voraus, daß jeder Student in einer Prüfung eine Platzziffer (Rang) bekommt, d.h. ausschließlich relativ bewertet wird. In einem Fach kann ein Rang nur einmal vergeben werden.

a) Beschreiben Sie die angegebenen Funktionalabhängigkeiten mit eigenen Worten.

{S, F \rightarrow P}: Ein Student hat in einem Fach genau einen Prüfer.

{R, F \rightarrow S}: In einem Fach kann eine Platzziffer (Rang) nur einmal an einen Studenten vergeben werden.

{F, S \rightarrow R}: Ein Student bekommt in einem Fach genau einen Rang.

b) Bestimmen Sie alle Schlüsselkandidaten in EXAM.

Die Schlüsselkandidaten der Relation EXAM sind {F, R} und {S, F}.

c) Stellen Sie fest, in welcher Normalform sich EXAM befindet und zerlegen Sie sie gegebenenfalls in BCNF-Relationen.

Die Relation EXAM ist in 3NF und nicht in BCNF.

d) Beurteilen Sie die Zerlegung von EXAM in EX (S, F, R) und PS (S, F, P) bezüglich:

- Verlustfreiheit

- Abhängigkeitsbewahrung
- Normalform
- sonstige Kriterien

Sollte weiter zerlegt werden? Wenn ja, wie?

Die Zerlegung ist verlustfrei, da entlang der FA $\{S, F \rightarrow R\}$ zerlegt wurde, sie ist abhängigkeitsbewahrend, und die resultierenden Relationen EX und PS sind beide in BCNF. Allerdings ist die Zerlegung unnötig. Es wird bzgl. Normalformen nichts erreicht, was nicht durch EXAM gegeben wäre; im Gegenteil, es wird unnötige „Schlüsselredundanz“ bzgl. S, F eingeführt. Diese „Schlüsselredundanz“ führt zu keinen Anomalien, erhöht aber potentiell die Speicherkosten.

e) Wie ändern sich die Ergebnisse von Teil b) bis d), falls zusätzlich gilt: $P \rightarrow F$, d.h. jeder Prüfer prüft nur ein Fach?

Zu b): Die Schlüsselkandidaten sind dann $\{S, F\}$, $\{R, F\}$, $\{P, R\}$ und $\{P, S\}$.

Zu c): Die Relation ist in 3NF aber nicht mehr in BCNF, da die FA $P \rightarrow F$ die BCNF-Bedingung verletzt. Diese Relation soll daher weiter zerlegt werden.

Zu d): Die Zerlegung ist verlustfrei und abhängigkeitsbewahrend.

Zerlegungsschritte:

1. Zerlegung von EXAM (schon gegeben):

- EX (S, F, R) mit $F_{ex} = \{S, F \rightarrow R, R, F \rightarrow S\}$, Schlüsselkandidaten sind $\{S, F\}$ und $\{R, F\}$ und die Relation ist in BCNF.
- PS (S, F, P) mit $F_{ps} = \{S, F \rightarrow P, P \rightarrow F\}$, Schlüsselkandidaten sind $\{S, F\}$ und $\{P, S\}$. Die Relation ist in 3NF aber nicht in BCNF, da die FA $P \rightarrow F$ die BCNF-Bedingung verletzt. Die Relation PS sollte daher weiter zerlegt werden.

2. Zerlegung von PS entlang der FA $P \rightarrow F$:

- $PS_1 (\underline{P}, F)$ mit $\{P \rightarrow F\} \in F_1^+$, in BCNF
- $PS_2 (\underline{P}, S)$ nur mit trivialen Abhängigkeiten, in BCNF

Die Zerlegung ist nicht abhängigkeitsbewahrend, da beim 2. Zerlegungsschritt die FA $S, F \rightarrow P$ verloren geht.

Aufgabe 4: Relationensynthese

Gegeben seien die folgenden Attribute einer simplifizierten pharmazeutischen Datenbank:

Arz: Arzneimittel
 Rez: Rezeptur (d.h. Zusammensetzung eines Arzneimittels)
 Pack: Packung (d.h. Packungsgröße)
 Herst: Hersteller (des Arzneimittels)
 Form: Verabreichungsform (z.B. Tabletten, Tropfen, usw.)
 Preis: Preis der Arznei
 Dosis: Verabreichungsdosis (z.B. 2 Tabletten täglich)
 Neben: Nebenwirkungen
 Pat: Patient
 Krank: Krankheit

Folgende Funktionalabhängigkeiten sollen gelten:

{
 Arz \rightarrow Rez,
 Pack, Arz, Form \rightarrow Preis,
 Arz, Dosis, Pat \rightarrow Neben,
 Arz \rightarrow Herst,
 Herst, Krank, Pat \rightarrow Arz, Dosis,
 Arz, Pat \rightarrow Dosis, Form,
 Herst, Arz \rightarrow Form }
}

- a) Machen Sie sich zunächst die Bedeutungen der Funktionalabhängigkeiten klar. Erläutern Sie, inwieweit diese der Realität entsprechen und wo sie ggf. revidiert oder ergänzt werden sollten.

Mit diesen Fas könnte ein Patient zu einer Krankheit nur eine einzige Arznei bekommen, soweit es einen Hersteller, z.B. Bayer, betrifft! Das ist unrealistisch, da Therapien oft mehrere Arzneien kombinieren.

- b) Führen Sie für die gegebenen Funktionalabhängigkeiten - unabhängig von Ihren Erwägungen bei a) - eine Relationensynthese durch.

1. nach Anwendung der Zerlegungsregel:

F' = {
 Arz \rightarrow Rez
 Pack, Arz, Form \rightarrow Preis
 Arz, Dosis, Pat \rightarrow Neben
 Arz \rightarrow Herst
 Herst, Krank, Pat \rightarrow Arz
 Herst, Krank, Pat \rightarrow Dosis
 Arz, Pat \rightarrow Dosis
 Arz, Pat \rightarrow Form
 Herst, Arz \rightarrow Form }
}

2. nach Entfernung redundanter Attribute auf den linken Seiten

F'' = {
 Arz \rightarrow Rez
 Pack, Arz \rightarrow Preis
 Arz, Pat \rightarrow Neben
 Arz \rightarrow Herst
 Herst, Krank, Pat \rightarrow Arz
 Herst, Krank, Pat \rightarrow Dosis
 Arz, Pat \rightarrow Dosis
 Arz, Pat \rightarrow Form
 Arz \rightarrow Form }
}

(wg. Arz \rightarrow Form und Pack, Arz, Form \rightarrow Preis)
(wg. Arz, Pat \rightarrow Dosis und Arz, Dosis, Pat \rightarrow Neben)
(wg. Arz \rightarrow Herst und Herst, Arz \rightarrow Form)

3. nach Entfernung redundanter FAs

$F''' = \{$
 $\text{Arz} \rightarrow \text{Rez}$
 $\text{Pack, Arz} \rightarrow \text{Preis}$
 $\text{Arz, Pat} \rightarrow \text{Neben}$
 $\text{Arz} \rightarrow \text{Herst}$
 $\text{Herst, Krank, Pat} \rightarrow \text{Arz}$
 $\text{Arz, Pat} \rightarrow \text{Dosis}$
 $\text{Arz} \rightarrow \text{Form} \quad \}$

redundante FAs sind:

$\text{Herst, Krank, Pat} \rightarrow \text{Dosis}$

$\text{Arz, Pat} \rightarrow \text{Form}$

wg. $\text{Herst, Krank, Pat} \rightarrow \text{Arz}$ und $\text{Arz, Pat} \rightarrow \text{Dosis}$

wg. $\text{Arz} \rightarrow \text{Form}$

F''' ist eine Überdeckung von F .

4. erzeugtes Datenbankschema:

Arzneien	(<u>Arz</u> , Rez, Herst, Form)	BCNF
Verschreibungen	(<u>Herst, Krank, Pat</u> , Arz)	nicht BCNF
Patienten	(<u>Arz, Pat</u> , Dosis, Neben)	BCNF
Preise	(<u>Pack, Arz</u> , Preis)	BCNF

Relation Verschreibungen ist in 3NF, aber nicht in BCNF:

Schlüsselkandidaten sind: Herst, Krank, Pat und Arz, Krank, Pat

Die FA $\text{Arz} \rightarrow \text{Herst}$ verletzt die BCNF-Bedingung, da Arz kein Schlüsselkandidat, sondern nur Schlüsselattribut ist.

Schlüssel des gesamten ursprünglichen Schemas ist z.B. $\text{Pack, Arz, Pat, Krank}$

Also wird eine zusätzliche Relation eingefügt:

Abrechnungen (Pack, Arz, Pat, Krank)