

AG Datenbanken und Informationssysteme

Wintersemester 2006 / 2007

Prof. Dr.-Ing. Dr. h. c. Theo Härder
Fachbereich Informatik
Technische Universität Kaiserslautern



<http://www.dvs.informatik.uni-kl.de>

3. Übungsblatt

Für die Übung am Donnerstag, **16. November 2006**,
von 15:30 bis 17:00 Uhr in 13/222.

Aufgabe 1: Hülle einer Attributmenge

Sei F eine Menge von FAen und X eine Teilmenge von \mathbf{R} . Die Hülle einer Attributmenge X , genannt X^+ , kann auch durch den Algorithmus $\text{AttrHülle}(F, X)$ wie folgt berechnet werden:

Schritt 0: $X_0 = X$
Schritt $k+1$: $X_{k+1} = X_k \cup \{Z \mid Y \rightarrow Z \in F \text{ und } Y \subseteq X_k\}$
Wenn $X_{k+1} = X_k$, dann beende den Algorithmus und setze
 $\text{AttrHülle}(F, X) = X_k$,
sonst fahre mit dem nächsten Schritt von k fort.

Ermitteln Sie nun mit Hilfe des obigen Algorithmus jeweils die Attributhülle der beiden folgenden Mengen $X = \{A, E\}$ und $Y = \{A, B\}$ mit der FA-Menge $G = \{A \rightarrow D, AB \rightarrow E, BI \rightarrow E, CD \rightarrow I, E \rightarrow C\}$. Dokumentieren Sie die Zwischenergebnisse und die dabei benutzten FAen in den einzelnen Schritten.

Lösung:

Schritt _i	X_i	Benutzte FAs
0	AE	keine
1	$AE \cup DC$	$A \rightarrow D, E \rightarrow C$
2	$AEDC \cup DCI$	$A \rightarrow D, E \rightarrow C, CD \rightarrow I$
3	$AEDCI \cup DCI$	$A \rightarrow D, E \rightarrow C, CD \rightarrow I$

Es ergibt sich somit die Hülle $X^+ = \{A, C, D, E, I\}$

Schritt _i	Y _i	Benutzte FAs
0	AB	keine
1	AB ∪ DE	A → D, AB → E
2	ABDE ∪ DEC	A → D, AB → E, E → C
3	ABDEC ∪ DECI	A → D, AB → E, E → C, CD → I
4	ABDECI ∪ DECI	A → D, AB → E, E → C, CD → I, BI → E

Es ergibt sich somit die Hülle $Y^+ = \{A, B, C, D, E, I\}$

Aufgabe 2: Bestimmung von Superschlüsseln und Schlüsselkandidaten

Man kann mit dem Algorithmus CLOSURE aus der Vorlesung überprüfen, ob die Attributmeng X einen Superschlüssel eines Relationenschemas **R** bzgl. einer FA-Menge F bildet. Dazu wendet man den Algorithmus auf X und F an, um X^+ zu bestimmen. Nur wenn $X^+ = \mathbf{R}$ ergibt, dann ist X ein Superschlüssel von **R**.

Gegeben seien $\mathbf{R} = \{PNR, PNAME, FACH, MATNR, NAME, GEB, ADR, NOTE, PDAT, FBNR, FBNAME\}$ und eine FA-Menge

$$F = \{ PNR \rightarrow PNAME, FACH, \\ MATNR \rightarrow NAME, GEB, ADR, FBNR, \\ NAME, GEB, ADR \rightarrow MATNR, \\ PNR, MATNR, FBNR \rightarrow NOTE, PDAT, \\ FBNR \rightarrow FBNAME \}.$$

Zeigen Sie mit Hilfe von CLOSURE, dass die Attributmeng {PNR, MATNR} sogar ein Schlüsselkandidat in **R** ist. Dokumentieren Sie bei jeder Anwendung von CLOSURE die Zwischenergebnisse und die benutzten FAs in den einzelnen Schritten.

Lösung:

Zuerst muss mit CLOSURE({PNR, MATNR}, F) gezeigt werden, dass {PNR, MATNR} ein Superschlüssel in **R** ist:

Schritt _i	OLDDEP _i	NEWDEP _i	Benutzte FAs
0	{}	{PNR, MATNR}	keine
1	{PNR, MATNR}	{PNR, MATNR} ∪ {PNAME, FACH} ∪ {NAME, GEB, ADR, FBNR} ∪ {MATNR} ∪ {NOTE, PDAT} ∪ {FBNAME}	{PNR} → {PNAME, FACH} {MATNR} → {NAME, GEB, ADR, FBNR} {NAME, GEB, ADR} → {MATNR} {PNR, MATNR, FBNR} → {NOTE, PDAT} {FBNR} → {FBNAME}

Schritt _i	OLDDEP _i	NEWDEP _i	Benutzte FAs
2	{PNR, MATNR, PNAME, FACH, NAME, GEB, ADR, FBNR, NOTE, PDAT, FBNAME}	{PNR, MATNR, PNAME, FACH, NAME, GEB, ADR, FBNR, NOTE, PDAT, FBNAME} \cup {PNAME, FACH} \cup {NAME, GEB, ADR, FBNR} \cup {MATNR} \cup {NOTE, PDAT} \cup {FBNAME}	{PNR} \rightarrow {PNAME, FACH} {MATNR} \rightarrow {NAME, GEB, ADR, FBNR} {NAME, GEB, ADR} \rightarrow {MATNR} {PNR, MATNR, FBNR} \rightarrow {NOTE, PDAT} {FBNR} \rightarrow {FBNAME}

$CLOSURE(\{PNR, MATNR\}, F) = \{PNR, MATNR\}^+ = \mathbf{R}$. Hieraus folgt, dass {PNR, MATNR} ein Superschlüssel ist.

Um zu zeigen, dass {PNR, MATNR} ein Schlüsselkandidat in **R** ist, wendet man CLOSURE wieder an. Dazu darf jede echte Teilmenge von {PNR, MATNR} ({PNR} und {MATNR}) keinen Superschlüssel in **R** bilden. Damit zeigt man die Minimalität von {PNR, MATNR}.

Schritt _i	OLDDEP _i	NEWDEP _i	Benutzte FAs
0	{}	{PNR}	keine
1	{PNR}	{PNR} \cup {PNAME, FACH}	{PNR} \rightarrow {PNAME, FACH}
2	{PNR, PNAME, FACH}	{PNR, PNAME, FACH} \cup {PNAME, FACH}	{PNR} \rightarrow {PNAME, FACH}

Somit ist $CLOSURE(\{PNR\}, F) = \{PNR\}^+ = \{PNR, PNAME, FACH\} \subsetneq \mathbf{R}$.

Schritt _i	OLDDEP _i	NEWDEP _i	Benutzte FAs
0	{}	{MATNR}	keine
1	{MATNR}	{MATNR} \cup {NAME, GEB, ADR, FBNR} \cup {MATNR} \cup {FBNAME}	{MATNR} \rightarrow {NAME, GEB, ADR, FBNR} {NAME, GEB, ADR} \rightarrow {MATNR} {FBNR} \rightarrow {FBNAME}
2	{MATNR, NAME, GEB, ADR, FBNR, FBNAME}	{MATNR, NAME, GEB, ADR, FBNR, FBNAME} \cup {NAME, GEB, ADR, FBNR} \cup {MATNR} \cup {FBNAME}	{MATNR} \rightarrow {NAME, GEB, ADR, FBNR} {NAME, GEB, ADR} \rightarrow {MATNR} {FBNR} \rightarrow {FBNAME}

Somit ist $CLOSURE(\{MATNR\}, F) = \{MATNR\}^+ = \{MATNR, NAME, GEB, ADR, FBNR, FBNAME\} \subsetneq \mathbf{R}$.

Damit ist gezeigt, dass {PNR, MATNR} ein Schlüsselkandidat in **R** ist.

Aufgabe 3: Normalisierung von Relationen

Gegeben sei folgende unnormalisierte Relation:

ABT (ANR, AName, Budget,
 PERS (PNR, Name, Büro-Nr, Tel-Nr, GEH-HISTORIE (Datum, Gehalt)),
 BÜRO (Büro-Nr, Fläche, TEL (Tel-Nr)))

mit den relationenwertigen Attributen PERS, GEH-HISTORIE, BÜRO und TEL.

Gegeben sind folgende funktionalen Abhängigkeiten:

ANR → AName, Budget

PNR → Name, ANR, Büro-Nr, Tel-Nr

Tel-Nr → Büro-Nr

Büro-Nr → ANR, Fläche

PNR, Datum → Gehalt

Überführen Sie die Relation schrittweise in 1NF, 2NF und 3NF.

Lösung:

Aus dem Aufgabentext ergeben sich noch die beiden FAen:

Überführung in 1NF (Relationen mit atomaren Attributwerten mittels „copy down the key“):

ABT1 (ANR, AName, Budget)

PERS1 (ANR, PNR, Name, Büro-Nr, Tel-Nr)

GEH-HIST1 (ANR, PNR, Datum, Gehalt)

BÜRO1 (ANR, Büro-Nr, Fläche)

TEL1 (ANR, Büro-Nr, Tel-Nr)

Überführung in 2NF (jedes Nicht-Primärattribut voll funktional von jedem Schlüsselkandidaten abhängig):

Momentan sind die gewählte Primärschlüssel von BÜRO1, GEH-HIST1, TEL1 und PERS1 nicht minimal:

"Zerlegung"/Umformung von PERS1 in
 PERS2 (PNR, ANR, Name, Büro-Nr, Tel-Nr)

"Zerlegung"/Umformung von GEH-HIST in
 GEH-HIST2 (PNR, Datum, Gehalt)

[PERS_IN_ABT2 (PNR, ANR)]

"Zerlegung"/Umformung von BÜRO1 in
 BÜRO2 (Büro-Nr, ANR, Fläche)

"Zerlegung"/Umformung von TEL1 in
 TEL2 (Tel-Nr, Büro-Nr, ANR)

- ABT1 ist bereits in 2NF.

- Relation PERS_IN_ABT2 ist in PERS2 bereits abgebildet und damit redundant

Überführung in 3NF (Entfernen transitiver Abhängigkeiten):

Transitive **Abhängigkeitskette** in PERS1: PNR → Tel-Nr → Büro-Nr → ANR.

Zerlegung von PERS2 in:

PERS3 (PNR, Name, Tel-Nr)

TEL_IN_BÜRO3 (Tel-Nr, Büro-Nr)

BÜRO_IN_AB3a (Büro-Nr, ANR)

Zerlegung von TEL2 in:

TEL3 (TelNr, BüroNr)

BÜRO_IN_AB3b (Büro-Nr, ANR)

- BÜRO2, ABT1 und GEH-HIST2 sind bereits in 3NF.

- Relationen TEL_IN_BÜRO3 und TEL3 sind redundant.

- Relationen BÜRO_IN_AB3a und BÜRO_IN_AB3b sind redundant.

ACHTUNG: Die Kette von transitiven Abhängigkeiten in PERS2 kann durch verschiedene Teilschritte zerlegt werden. Hierdurch kann zusätzlich z.B. die Relation TEL_IN_AB3 (TelNr, ANR) entstehen. Diese Lösung wäre auch ein Schema in 3NF. Die Aufnahme der Relation sollte jedoch vermieden werden, weil die in ihr abgebildete Abhängigkeit bereits durch die Relationen TEL_IN_BÜRO3 und BÜRO_IN_AB3 abgebildet ist.

Ergebnisschema in 3NF:

ABT1 (ANR, AName, Budget)

PERS3 (PNR, Name, Tel-Nr)

GEH-HIST2 (PNR, Datum, Gehalt)

BÜRO2 (Büro-Nr, ANR, Fläche)

BÜRO_IN_AB3 (Büro-Nr, ANR)

TEL_IN_BÜRO3 (TelNr, BüroNr)

Aufgabe 4: Normalisierung von Relationen in BCNF

Gegeben seien $\mathbf{R} = \{\text{VERKÄUFER, KUNDE, DATUM, ARTIKEL, HERSTELLER, LIEFERADR, PREIS}\}$ mit folgender Relation:

BESTELLUNG (VERKÄUFER, KUNDE, DATUM, ARTIKEL, HERSTELLER,
LIEFERADR, PREIS)

und eine FA-Menge

$$F = \{ \text{KUNDE} \rightarrow \text{VERKÄUFER}, \\ \text{KUNDE, DATUM} \rightarrow \text{LIEFERADR, ARTIKEL}, \\ \text{DATUM, ARTIKEL, HERSTELLER} \rightarrow \text{PREIS}, \\ \text{ARTIKEL} \rightarrow \text{HERSTELLER}, \\ \text{LIEFERADR} \rightarrow \text{KUNDE} \}.$$

Normalisieren Sie die Relation BESTELLUNG bzgl. F, bis alle Relationen in BCNF sind. Ist die Zerlegung in BCNF-Relationen sinnvoll? Begründen Sie Ihre Antwort.

Lösung:

Sei $V = \text{VERKÄUFER}$, $K = \text{KUNDE}$, $D = \text{DATUM}$, $A = \text{ARTIKEL}$, $L = \text{LIEFERADR}$, $P = \text{PREIS}$,
 $H = \text{HERSTELLER}$

und $F = \{K \rightarrow V, KD \rightarrow LA, DAH \rightarrow P, A \rightarrow H, L \rightarrow K\}$

a) Man bestimmt zuerst alle Schlüsselkandidaten in B:

Schlüsselkandidaten sind KD und DL. Denn für KD gilt $\{KD\}^+ = \mathbf{R}$ und $\{K\}^+ = \{KV\} \not\rightarrow \mathbf{R}$ bzw. $\{D\}^+ = \{D\} \not\rightarrow \mathbf{R}$. Analog für DL gilt $\{DL\}^+ = \mathbf{R}$ und $\{L\}^+ = \{LKV\} \not\rightarrow \mathbf{R}$ bzw. $\{D\}^+ = \{D\} \not\rightarrow \mathbf{R}$.

b) Demnach ergeben sich die Nicht-Primärattribute:

Die Nicht-Primärattribute sind A, H, P und V (alle Attribute, die nicht zu einem Schlüsselkandidaten gehören).

Man untersucht BESTELLUNG (V, K, D, A, H, L, P) bzgl. der Normalformen:

c) BESTELLUNG ist in 1NF, da alle Attribute atomare Werte besitzen.

d) BESTELLUNG ist nicht in 2NF, weil sowohl $K \rightarrow V$ als auch $L \rightarrow V$, d.h. das Nicht-Primärattribut V ist von KD bzw. DL nicht voll funktional abhängig.

Daher zerlegt man BESTELLUNG in die 2NF Relationen:

(1) BESTELLUNG1 (K, D, A, H, L, P) mit KD und DL als Schlüsselkandidaten sowie A, H und P als Nicht-Primärattribute.

(2) KUNDE (K, V) mit K als Schlüsselkandidat und V als Nicht-Primärattribut.
KUNDE ist in BCNF (einziger Determinant K ist auch Schlüsselkandidat).

e) Man prüft, ob BESTELLUNG1 (K, D, A, H, L, P) in 3NF:

BESTELLUNG1 ist nicht in 3NF, da H transitiv abhängig von KD ist ($KD \rightarrow A \rightarrow H$) und P transitiv abhängig von KD ($KD \rightarrow LA$, da $A \rightarrow H$ gilt $KD \rightarrow DAH \rightarrow P$). Daher überführt man BESTELLUNG1 in folgende Relationen in 3NF:

(1) BESTELLUNG2 (K, D, A, L) mit KD und DL als Schlüsselkandidaten sowie A als Nicht-Primärattribut.

(2) PREIS (D, A, P) mit DA als Schlüsselkandidat und P als Nicht-Primärattribut.
PREIS ist in BCNF (einziger Determinant DA ist Schlüsselkandidat).

- (3) ARTIKEL (A, H) mit A als Schlüsselkandidat und H als Nicht-Primärattribut.
ARTIKEL ist in BCNF (einziger Determinant A ist Schlüsselkandidat).
- f) Schließlich prüft man noch, ob BESTELLUNG2 (K, D, A, L) in BCNF ist:
Sie ist nicht in BCNF, weil der Determinant L ($L \rightarrow K$) kein Schlüsselkandidat ist. Die Relation wird in folgende BCNF-Relationen zerlegt werden:
- (1) BESTELLUNG3 (K, D, A) mit KD als Schlüsselkandidat und A als Nicht-Primärattribut
 - (2) LIEFERADRESSE (L, K) mit L als Schlüsselkandidat und K als Nicht-Primärattribut.

Die Zerlegung von BESTELLUNG2 in BCNF ist allerdings nicht sinnvoll (zu streng), weil $KD \rightarrow L$ dann nicht bewahrt werden kann und extern geprüft werden muss.