6. Relationaler Datenbankentwurf

- Funktionale Abhängigkeiten
- Schema-Eigenschaften
- Transformationseigenschaften
- Entwurfsverfahren
- Mehrwertige Abhängigkeiten
- Weitere Abhängigkeiten

Relationaler Datenbankentwurf: Überblick

- Verfeinern des logischen Entwurfs
- Ziel: Vermeidung von Redundanzen durch Aufspalten von von Relationenschemata, ohne gleichzeitig
  - semantische Informationen zu verlieren (Abhängigkeitstreue)
  - die Möglichkeit zur Rekonstruktion der Relationen zu verlieren (Verbundtreue)
- Beispiel: $\leadsto$ □ □
- Redundanzvermeidung durch Normalformen (s.u.)

Funktionale Abhängigkeiten

- Funktionale Abhängigkeit einer Relation zwischen Attributmengen X und Y, wenn in jedem Tupel der Relation der Attributwert unter den X-Komponenten den Attributwert unter den Y-Komponenten festlegt
- Funktionale Abhängigkeit (kurz: FD, von functional dependency), Schreibweise: $X \rightarrow Y$
- Im Beispiel auf folgender Folie: $\texttt{ISBN} \rightarrow \texttt{Titel, Verlag}$
- Nicht: $\texttt{ISBN} \rightarrow \texttt{Autor, Stichwort}$
- Trivialerweise: $\texttt{ISBN} \rightarrow \texttt{ISBN}$
- Formal $\leadsto$ □ □

Bücher-Relation mit Redundanzen

Bücher

ISBN	Titel	Autor	Version	Stichwort	Verlagsname
0-8053-1753-8	Princ.of DBS	Elmasri	1,1989	RDB	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Navathe	1,1989	RDB	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Elmasri	2,1994	RDB	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Navathe	2,1994	RDB	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Elmasri	1,1989	Lehrbuch	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Navathe	1,1989	Lehrbuch	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Elmasri	2,1994	Lehrbuch	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Navathe	2,1994	Lehrbuch	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Elmasri	1,1989	ER	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Navathe	1,1989	ER	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Elmasri	2,1994	ER	Benj./Cumm.
0-8053-1753-8	Princ.of DBS	Navathe	2,1994	ER	Benj./Cumm.

Schlüssel als Spezialfall

- für Beispiel auf folgender Folie
  $\texttt{PANr} \rightarrow \texttt{Vorname, Nachname, PLZ, Ort, Straße,}$
  $\texttt{Hausnummer, Geburtsdatum}$
- Immer: $\texttt{PANr} \rightarrow \texttt{PANr}$, dann gesamtes Schema auf rechter Seite
- Wenn linke Seite minimal: Schlüssel
- Formal: Schlüssel X liegt vor, wenn für Relationenschema R FD $X \rightarrow R$ gilt und X minimal

Ziel des Datenbankentwurfs: alle gegebenen funktionalen Abhängigkeiten in "Schlüsselabhängigkeiten" umformen, ohne dabei semantische Information zu verlieren

Schlüssel im Beispiel

Personen

PANr	Vorname	Nachname	PLZ	Ort	Straße	HNr	Geb.datum
4711	Andreas	Heuer	18209	DBR	BHS	15	31.10.1958
5588	Gunter	Saake	39106	MD	STS	55	05.10.1960
6834	Michael	Korn	39104	MD	BS	41	24.09.1974
7754	Andreas	Möller	18209	DBR	RS	31	25.02.1976
8832	Tamara	Jagellovsk	38106	BS	GS	12	11.11.1973
9912	Antje	Hellhof	18059	HRO	AES	21	04.04.1970
9999	Christa	Loeser	69121	HD	TS	38	10.05.1969

Pers_Telefon

PANr	Telefon
4711	038203-12230
4711	0381-498-3401
4711	0381-498-3427
5588	0391-345677
5588	0391-5592-3800
9999	06221-400177

Ableitung von FDs

r

A	B	C
a_1	b_1	c_1
a_2	b_1	c_1
a_3	b_2	c_1
a_4	b_1	c_1

- genügt $A \rightarrow B$ und $B \rightarrow C$
- dann gilt auch $A \rightarrow C$
- nicht ableitbar $C \rightarrow A$ oder $C \rightarrow B$

Ableitung von FDs II

- Gilt für f über R **SAT**$_R(F) \subseteq$ **SAT**$_R(f)$, dann *impliziert* F die FD f (kurz: $F \models f$)
- obiges Beispiel:

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

Hüllenbildung F^+: $\leadsto \Box\ \Box$

Ableitungsregeln

- *gültig* (sound)
- *vollständig* (complete)
- *unabhängig* (independent) oder auch bzgl. $\subseteq$ minimal

Name		Regel		
R	Reflexivität	$\{\}$	$\Longrightarrow$	$X \rightarrow X$
A	Akkumulation	$\{X \rightarrow YZ, Z \rightarrow VW\}$	$\Longrightarrow$	$X \rightarrow YZV$
P	Projektivität	$\{X \rightarrow YZ\}$	$\Longrightarrow$	$X \rightarrow Y$

Membership-Problem

Kann eine bestimmte FD $X \rightarrow Y$ aus der vorgegebenen Menge F abgeleitet werden, d.h. wird sie von F impliziert?

Membership-Problem: “$X \rightarrow Y \in F^+$?”

Hülle einer Attributmenge X bzgl. F ist $X_F^* := \{A \mid X \rightarrow A \in F^+\}$

Das Membership-Problem kann nun durch das modifizierte Problem

Membership-Problem (2): “$Y \subseteq X_F^*$?”

in linearer Zeit gelöst werden

RAP-Algorithmus

1. Bestimme X, setze $X^* := X$ ($\boxed{\text{R}}$-Regel für X)
2. Gibt es FD $f_1 := X_1 \rightarrow Y_1 \in F$ mit $X_1 \subseteq X^*$?
3. Wenn ja, dann wird X^* gemäß $X^* := X^* \cup Y_1$ vergrößert ($\boxed{\text{A}}$-Regel).
4. Führe Schritt 2 und 3 so lange aus, bis X^* stabil (Hülle)
5. Ist $Y \subseteq X^*$, dann ist $X \rightarrow Y \in F^+$ ($\boxed{\text{P}}$-Regel)

Überdeckungen

F heißt *äquivalent* zu G

(oder: F *Überdeckung* von G; kurz: $F \equiv G$)

falls $F^+ = G^+$

Schema-Eigenschaften

Relationenschemata, Schlüssel und Fremdschlüssel so wählen, daß

1. alle Anwendungsdaten aus den Basisrelationen hergeleitet werden können,
2. nur semantisch sinnvolle und konsistente Anwendungsdaten dargestellt werden können und
3. die Anwendungsdaten möglichst nicht-redundant dargestellt werden.

Hier: Forderung 3

- Redundanzen innerhalb einer Relation: Normalformen
- globale Redundanzen: Minimalität

Update-Anomalien

Redundanzen in Basisrelationen unerwünscht:

- Belegen unnötigen Speicherplatz (eher unwichtig)
- Information redundant $\longrightarrow$ Änderung muß diese Information in allen ihren Vorkommen verändern (in relationalen Systemen nur schwer zu realisieren)

Beispiel **insert**-Anomalie:

Bücher

ISBN	Titel	Autor	Version	Stichwort	Verlagsname
0-8053-1753-8	Princ.of DBS	Elmasri	3,1996	RDB	Springer

in Bücher-Relation einfügen (FD, MVD verletzt; besser: auf Schlüsselabhängigkeiten zurückführen)

Erste Normalform

- führt zunächst Redundanzen ein
- Erste Normalform (1NF): nur atomare Attribute in Relationenschemata

Invnr	Titel	ISBN	Autoren
0007	Dr. No	3-125	James Bond
1201	Objektbanken	3-111	Heuer, Scholl

wäre in erster Normalform

Invnr	Titel	ISBN	Autor
0007	Dr. No	3-125	James Bond
1201	Objektbanken	3-111	Heuer
1201	Objektbanken	3-111	Scholl

Zweite Normalform

Zweite und weitere Normalformen: aufgrund der Struktur von Abhängigkeiten Redundanzen entdecken

- Zweite Normalform: Keine partiellen Abhängigkeiten zwischen einem Schlüssel und weiteren Nicht-Primattributen (Attribute, die nicht in einem Schlüssel auftauchen)
- *partielle Abhängigkeit* liegt vor, wenn ein Attribut funktional schon von einem *Teil* des Schlüssels abhängt

Zweite Normalform II

- Beispiel:

  `Invnr` $\rightarrow$ `Titel`

  und

  `Invnr, Autor` $\rightarrow$ `Invnr, Titel, ISBN, Autor`

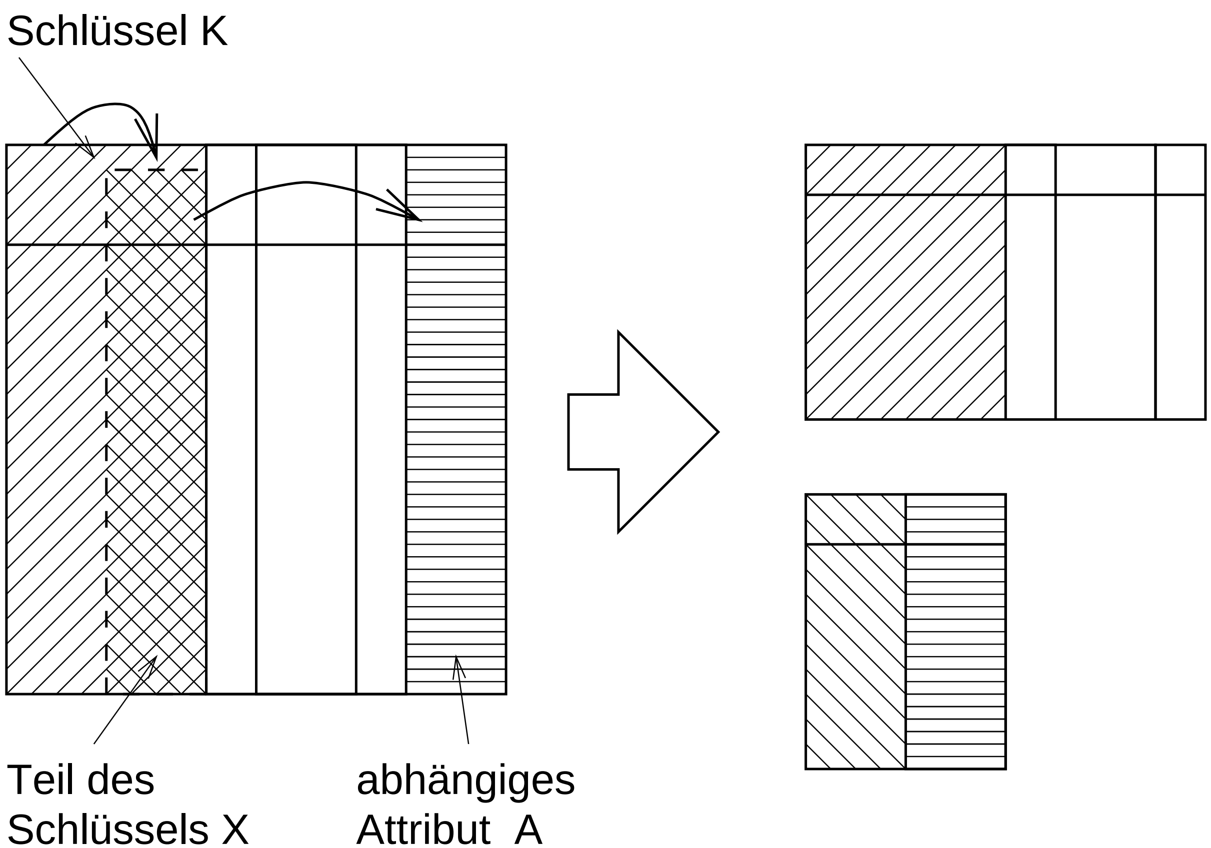
  `Invnr` und `Autor` zusammen Schlüssel

  `Titel` hängt aber allein von `Invnr` ab

- Zweite Normalform erreichen durch Elimination der rechten Seite der partiellen Abhängigkeit und Kopie der linken Seite (siehe nächste Folie)

- Formal: $\leadsto$ $\Box$ $\Box$

Veranschaulichung zweite Normalform



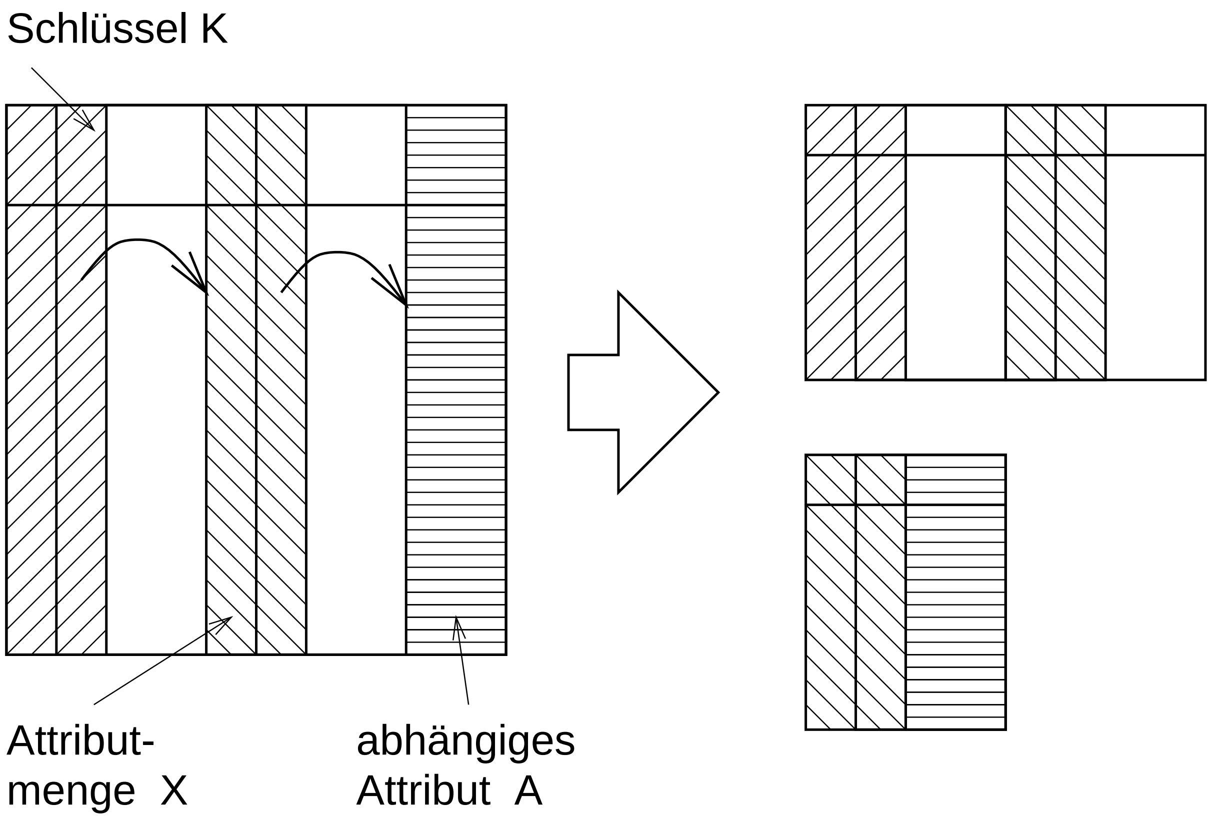
Dritte Normalform

- Dritte Normalform: Keine transitiven Abhängigkeiten zwischen einem Schlüssel und weiteren Nicht-Primattributen
- *transitive Abhängigkeit*: Schlüssel K bestimmt Attributmenge X funktional, diese aber auch eine Attributmenge Y

  transitive Abhängigkeit $K \rightarrow X \rightarrow Y$
- Beispiel: $\mathtt{PANr} \rightarrow \mathtt{PLZ}$ und $\mathtt{PLZ} \rightarrow \mathtt{Ort}$

  Information, daß zur `PLZ` '18209' der `Ort` 'DBR' gehört, ist redundant
- Dritte Normalform erreichen durch Elimination von Y und Kopie von X (siehe nächste Folie)
- Formal: $\leadsto \Box\ \Box$

Veranschaulichung dritte Normalform



Boyce-Codd-Normalform

- Nicht nur Nicht-Primattribute betrachten
- Im aktuellen Postleitzahlsystem der Deutschen Post innerhalb der Attribute

  $$\mathtt{PLZ},\ \mathtt{Ort},\ \mathtt{Straße},\ \mathtt{Hausnummer}$$

  folgende funktionalen Abhängigkeiten:

  $$\mathtt{Ort},\ \mathtt{Straße},\ \mathtt{Hausnummer} \rightarrow \mathtt{PLZ},$$
  $$\mathtt{PLZ} \rightarrow \mathtt{Ort}$$

  Schlüssel: `Ort, Straße, Hausnummer` und `PLZ, Straße, Hausnummer`

alle Attribute nun Primattribute: dritte Normalform

Boyce-Codd-Normalform II

- Trotzdem Redundanz:

$$\texttt{PLZ, Straße, Hausnummer} \rightarrow \texttt{PLZ} \rightarrow \texttt{Ort}$$

- partielle (oder transitive) Abhängigkeit

Boyce-Codd-Normalform (BCNF) definiert transitive Abhängigkeiten nicht nur über Nicht-Primattribute

Minimalität

- Global Redundanzen vermeiden
- andere Kriterien (wie Normalformen) mit möglichst wenig Schemata erreichen
- Beispiel: Attributmenge ABC, FD-Menge $\{A \rightarrow B, B \rightarrow C\}$
- Datenbankschemata in dritter Normalform:

$$S = \{(AB, \{A\}), (BC, \{B\})\}$$

$$S' = \{(AB, \{A\}), (BC, \{B\}), (AC, \{A\})\}$$

Redundanzen in S'

Schema-Eigenschaften: Zusammenfassung

Kennung	Schemaeigenschaft	Kurzcharakteristik
	1NF	nur atomare Attribute
	2NF	keine partielle Abhängigkeit eines Nicht-Primattributes von einem Schlüssel
S 1	3NF	keine transitive Abhängigkeit eines Nicht-Primattributes von einem Schlüssel
	BCNF	keine transitive Abhängigkeit eines Attributes von einem Schlüssel
S 2	Minimalität	minimale Anzahl von Relationenschemata, die die anderen Eigenschaften erfüllt

Transformationseigenschaften

- Erreichen von Normalformen durch Zerlegung von Relationenschemata
- Dabei beachten:

  1. nur semantisch sinnvolle und konsistente Anwendungsdaten darstellen (*Abhängigkeitstreue*)
  2. alle Anwendungsdaten sollen aus Basisrelationen hergeleitet werden können (*Verbundtreue*)

Abhängigkeitstreue

- Abhängigkeitstreue ist folgende Forderung:
  - Allgemein: Menge der erfaßten Abhängigkeiten äquivalent zur Menge der im System darstellbaren Abhängigkeiten (etwa Schlüssel und Fremdschlüssel)
  - Hier spezieller: Menge der FDs äquivalent zur Menge der Schlüsselabhängigkeiten

Abhängigkeitstreue: Beispiel

- Attribute:
  `PLZ (P), Ort (O), Straße(S), Hausnummer(H)`
- funktionale Abhängigkeiten F:
  $\mathtt{OSH} \rightarrow \mathtt{P}$, $\mathtt{P} \rightarrow \mathtt{O}$
- Datenbankschema S:
  $(\mathtt{OSHP}, \{\mathtt{OSH}\})$
- Menge der zugehörigen Schlüsselabhängigkeiten:
  $\{ \mathtt{OSH} \rightarrow \mathtt{OSHP} \}$
  nicht äquivalent zu F; S ist nicht *abhängigkeitstreu*

Abhängigkeitstreue formal

- $S = \{(R_1, \mathcal{K}_1), \ldots, (R_p, \mathcal{K}_p)\}$ lokal erweitertes Datenbankschema, F Menge lokaler Abhängigkeiten
- S *charakterisiert vollständig* F (oder: ist *abhängigkeitstreu* bezüglich F) genau dann, wenn

$$F \equiv \{K \rightarrow R \mid (R, \mathcal{K}) \in S, K \in \mathcal{K}\}$$

Verbundtreue

- Originalrelation soll aus zerlegten Relationen mit natürlichem Verbund zurückgewonnen werden können
- Beispiel:
  Relationenschema $R = ABC$ in $R_1 = AB$ und $R_2 = BC$ zerlegt; ist bei

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

  nicht verbundtreu, bei

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

  verbundtreu.

Verbundtreue II

- *Kriterium:* Attributmenge im Schnitt der entstandenen Relationenschemata (hier: B) bestimmt eines der beiden Relationenschemata (hier: BC) funktional (ist also Schlüssel)

Beispielrelationen zur Verbundtreue

- Originalrelation:

A	B	C
1	2	3
4	2	5

- Dekomposition:

A	B
1	2
4	2

B	C
2	3
2	5

- Verbund (nicht verbundtreu):

A	B	C
1	2	3
4	2	5
1	2	5
4	2	3

Beispielrelationen zur Verbundtreue II

- Originalrelation:

A	B	C
1	2	3
4	2	3

- Dekomposition:

A	B
1	2
4	2

B	C
2	3

- Verbund (verbundtreu):

A	B	C
1	2	3
4	2	3

Verbundtreue formal

- *Definition:*
  Dekomposition einer Attributmenge X in $X_1, \ldots, X_p$ mit $X = \bigcup_{i=1}^{p} X_i$ heißt *verbundtreu* ($\pi \bowtie$-treu, lossless) bezüglich einer Menge von Abhängigkeiten G über X genau dann, wenn

$$\forall r \in \mathbf{SAT}_X(G) : \pi_{X_1}(r) \bowtie \cdots \bowtie \pi_{X_p}(r) = r$$

  gilt

- *Einfaches Kriterium für zwei Relationenschemata:*
  Dekomposition von X in X_1 und X_2 ist verbundtreu bzgl. F, wenn $X_1 \cap X_2 \rightarrow X_1 \in F^+$ oder $X_1 \cap X_2 \rightarrow X_2 \in F^+$

Verbundtreue formal II

- *Allgemeineres Kriterium:*
  G Menge funktionaler Abhängigkeiten

$$\exists i \in \{1, \ldots, p\} : X_i \rightarrow X \in G^+ \quad \Longrightarrow$$

  Die Dekomposition von X in $X_1, \ldots, X_p$ ist verbundtreu bezüglich G

- minimale Teilmenge von X_i: *Universalschlüssel*
- Beispiel erster Fall: AC der einzige Universalschlüssel, in keinem Relationenschema enthaltenn
- Beispiel zweiter Fall: Universalschlüssel A

Transformationseigenschaften: Zusammenfassung

Kennung	Transformationseigenschaft	Kurzcharakteristik
T 1	Abhängigkeitstreue	alle gegebenen Abhängigkeiten sind durch Schlüssel repräsentiert
T 2	Verbundtreue	die Originalrelationen können duch den Verbund der Basisrelationen wiedergewonnen werden

Entwurfsverfahren

- Universum $\mathcal{U}$ und FD-Menge F gegeben
- lokal erweitertes Datenbankschema $S = \{(R_1, \mathcal{K}_1), \ldots, (R_p, \mathcal{K}_p)\}$ berechnen mit

T	**1**	S charakterisiert vollständig F
S	**1**	S ist in 3NF bezüglich F
T	**2**	Dekomposition von $\mathcal{U}$ in $R_1, \ldots, R_p$ ist verbundtreu bezüglich F
S	**2**	Minimalität, d.h.

$\nexists S' : S'$ erfüllt $\boxed{\text{T}}\boxed{1}$, $\boxed{\text{S}}\boxed{1}$, $\boxed{\text{T}}\boxed{2}$ und $|S'| < |S|$

Entwurfsverfahren II

- Datenbankschemata schlecht entworfen, wenn nur eins dieser vier Kriterien nicht erfüllt
- Beispiel: $S = \{(AB, \{A\}), (BC, \{B\}), (AC, \{A\})\}$ erfüllt $\boxed{\text{T}}\boxed{1}$, $\boxed{\text{S}}\boxed{1}$ und $\boxed{\text{T}}\boxed{2}$ bezüglich $F = \{A \rightarrow B, B \rightarrow C, A \rightarrow C\}$

  in dritter Relation AC-Tupel redundant oder inkonsistent
- korrekt: $S' = \{(AB, \{A\}), (BC, \{B\})\}$

BCNF und Abhängigkeitstreue unvereinbar

- Attribute `PLZ (P), Ort (O), Straße(S), Hausnummer(H)`
- funktionale Abhängigkeiten F

$$\texttt{OSH} \rightarrow \texttt{P},\ \texttt{P} \rightarrow \texttt{O}$$

- Datenbankschema S

$$(\texttt{OSHP}, \{\texttt{OSH, PSH}\})$$

- `PSH` auch Schlüssel, da $\texttt{PSH} \rightarrow \texttt{OSHP}$ mit `PSH` minimal
- Schema in 3NF, da alle Attribute Primattribute

BCNF und Abhängigkeitstreue unvereinbar II

- Schema nicht in BCNF, da

$$\{\, \mathtt{PSH} \rightarrow \mathtt{P} \rightarrow \mathtt{O} \,\}$$

  transitive Abhängigkeit des Primattributs $\mathtt{O}$

- Jede Zerlegung von $\mathtt{OSHP}$ zerstört Abhängigkeit

$$\mathtt{OSH} \rightarrow \mathtt{P}$$

- Abhängigkeitstreue nicht gewährleistet

Dekomposition: Start

Start: initiales Relationenschema R mit allen Attributen und einer von erfaß-ten Abhängigkeiten implizierten Schlüsselmenge

- Attributmenge $\mathcal{U}$ und eine FD-Menge F
- suche alle $K \rightarrow \mathcal{U}$ mit K minimal, für die $K \rightarrow \mathcal{U} \in F^+$ gilt ($\mathcal{K}(F)$)
- $(\mathcal{U}, \mathcal{K}(F))$ initiales Relationenschema

Dekomposition: Normalisierung

Normalisierungsschritt: falls $K \rightarrow X \rightarrow Y$, aus R Attributmenge Y eliminieren und mit X in ein neues Relationenschema stecken

- $\mathcal{R} = (R, \mathcal{K})$ und F über R gegeben
- Falls $\mathcal{R}$ in 3NF ist: fertig
- Sonst: existiert für Schlüssel K

$$K \rightarrow Y, Y \not\rightarrow K, Y \rightarrow A, A \notin KY$$

Wähle dann

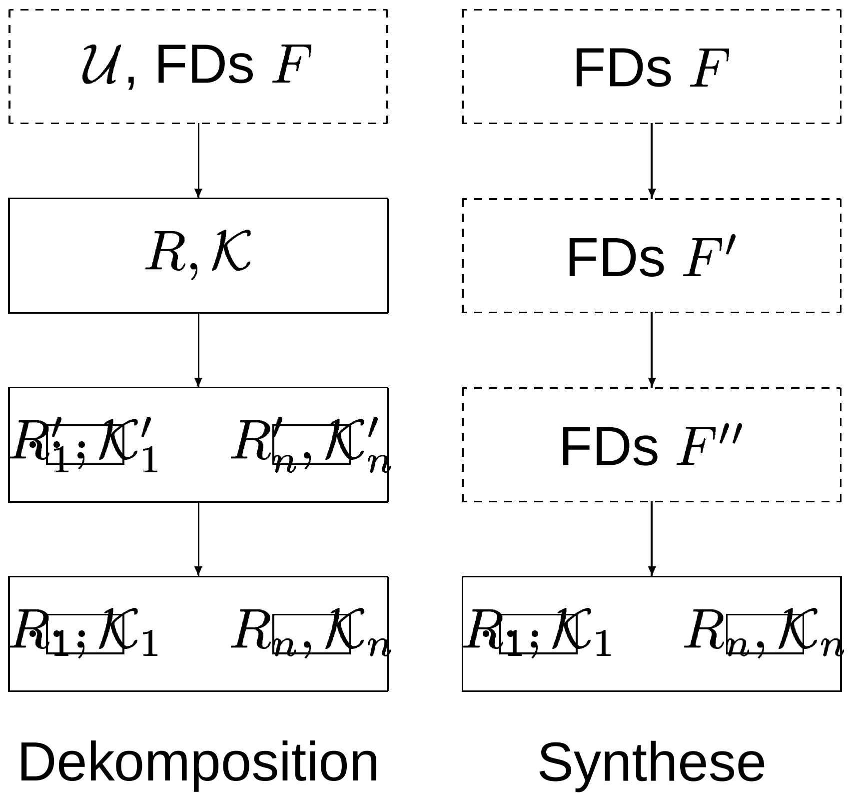
$$R_1 := R - A \qquad R_2 := YA$$

$$\mathcal{R}_1 := (R_1, \mathcal{K}) \qquad \mathcal{R}_2 := (R_2, \mathcal{K}_2 = \{Y\})$$

Dekomposition: Normalisierung II

- Vorteile: 3NF, Verbundtreue
- Nachteile: restliche Kriterien nicht, reihenfolgeabhängig, NP-vollständig (Schlüsselsuche)

Vergleich Dekomposition — Synthese



Syntheseverfahren

- Prinzip: Synthese formt Original-FD-Menge F in resultierende Menge von Schlüsselabhängigkeiten G so um, daß $F \equiv G$ gilt
- “Abhängigkeitstreue” im Verfahren verankert
- 3NF und Minimalität wird auch erreicht, reihenfolgeunabhängig
- Zeitkomplexität: quadratisch

Syntheseverfahren (grob)

- Eliminiere Redundanzen durch Entfernen überflüssiger FDs und Attribute.

  FDs aus F überflüssig, wenn *redundant*

  d.h. $F - \{f\} \equiv F$, dann f redundant

  Attribute aus F überflüssig, wenn unwesentlich (siehe Formalia 2NF)

- Fasse FDs zu “Äquivalenzklassen”

  FDs in eine Klasse, die gleiche oder äquivalente linke Seiten haben; pro Äquivalenzklasse ein Relationenschema

- Trick Verbundtreue: Original-FD-Menge F um $\mathcal{U} \rightarrow \delta$ erweitern, δ Dummy-Attribut, das nach Synthese entfernt wird

Mehrwertige Abhängigkeiten

- Mehrwertige Abhängigkeit (kurz: MVD, multivalued dependency) $X \rightarrow\!\!\!\rightarrow Y$
- innerhalb einer Relation r wird einem Attributwert von X eine Menge von Y-Werten zugeordnet, unabhängig von den Werten der restlichen Attribute von r
- Beispiel: In `Bücher`

  $\mathtt{ISBN} \rightarrow\!\!\!\rightarrow \mathtt{Autor}$
  $\mathtt{ISBN} \rightarrow\!\!\!\rightarrow \mathtt{Version}$
  $\mathtt{ISBN} \rightarrow\!\!\!\rightarrow \mathtt{Stichwort}$

- MVD formal: $\leadsto \square \; \square$

Mehrwertige Abhängigkeiten II

- Schwierigkeiten bei MVDs:
  $R = \{\,\texttt{Student, Fach, Vorlesung}\,\}$
  mit FD $\texttt{Student} \rightarrow \texttt{Fach}$ und MVD $\texttt{Fach} \rightarrow\!\!\!\rightarrow \texttt{Vorlesung}$
  - zusätzliches Attribut ($\texttt{SWS}$)
  - zusätzliche FD $\texttt{Vorlesung} \rightarrow \texttt{SWS}$
  - unsinnige FD $\texttt{Fach} \rightarrow \texttt{SWS}$ ableitbar
- MVD falsch spezifiziert
- $\texttt{Fach} \rightarrow\!\!\!\rightarrow \texttt{Vorlesung}$ (Werte zum Attribut $\texttt{Vorlesung}$ unabhängig von den Werten aller restlichen Attribute?)
- Sinnvoll $\texttt{Student} \rightarrow \texttt{Fach}$, $\texttt{Vorlesung} \rightarrow \texttt{SWS}$ und MVD $\texttt{Fach} \rightarrow\!\!\!\rightarrow \texttt{Vorlesung}, \texttt{SWS}$

Vierte Normalform

Name	Kind	Hobby
James Bond	Hugo	Autos
James Bond	Egon	Autos
James Bond	Hugo	Action
James Bond	Egon	Action
James Bond	Hugo	Klettern
James Bond	Egon	Klettern

$\mathtt{Name} \twoheadrightarrow \mathtt{Kind}, \mathtt{Name} \twoheadrightarrow \mathtt{Hobby}$

vierte Normalform durch Elimination der rechten Seite einer der beiden mehrwertigen Abhängigkeiten, linke Seite mit dieser rechten Seite in neue Relation kopiert

Name	Kind
James Bond	Hugo
James Bond	Egon

Name	Hobby
James Bond	Autos
James Bond	Action
James Bond	Klettern

Transformationseigenschaften bei MVDs

"Unabhängigkeit" der Attributmengen Y und Z voneinander: pro X-Wert bildet kartesisches Produkt der Y- und Z-Werte den YZ-Wert

$$X \rightarrow\!\!\!\rightarrow Y \quad \Longleftrightarrow \quad \forall X - \text{Werte } x : \pi_{YZ}(\sigma_{X=x}(r)) = \pi_Y(\sigma_{X=x}(r)) \bowtie \pi_Z(\sigma_{X=x}(r))$$

genau für alle $r \in \textbf{SAT}_R(\{X \rightarrow\!\!\!\rightarrow Y\})$ gilt verbundtreue Dekomposition

$$r = \pi_{XY}(r) \bowtie \pi_{XZ}(r), \qquad (1)$$

Weitere Abhängigkeiten

Verbundabhängigkeiten

- Kann man R ohne Informationsverlust in $R_1, \ldots, R_p$ auftrennen: $\bowtie [R_1, \ldots, R_p]$ Verbundabhängigkeit (kurz: JD, join dependency)
- Für `Bücher` gilt

$$\bowtie [\ \{\texttt{ISBN, Titel, Verlag}\},\ \{\texttt{ISBN, Autor}\},\ \{\texttt{ISBN, Stichwort}\},\ \{\texttt{ISBN, Version}\}\ ]$$

Weitere Abhängigkeiten

- Formal: Verbundabhängigkeit ist für $S = \{R_1, \ldots, R_p\}$ über $\mathcal{U}$ Ausdruck der Form $\bowtie [R_1, \ldots, R_p]$; $r(\mathcal{U})$ *genügt* der JD $\bowtie [R_1, \ldots, R_p]$ genau dann, wenn $r = \pi_{R_1}(r) \bowtie \cdots \bowtie \pi_{R_p}(r)$ gilt
- JDs sind Verallgemeinerung von MVDs: für $\mathcal{U} = XYZ$ und $r(\mathcal{U})$

$$r \text{ genügt } X \rightarrow\!\!\!\rightarrow Y \quad \Longleftrightarrow \quad r \text{ genügt } \bowtie [XY, XZ]$$

beziehungsweise für $S = \{R_1, R_2\}$ über $\mathcal{U}$

$$r \text{ genügt } \bowtie [R_1, R_2] \quad \Longleftrightarrow \quad r \text{ genügt } R_1 \cap R_2 \rightarrow\!\!\!\rightarrow R_1$$

Inklusionsabhängigkeiten

Verallgemeinerung von Fremdschlüsseln

- auf der rechten Seite einer Fremdschlüsselabhängigkeit nicht unbedingt den Primärschlüssel einer Relation: Inklusionsabhängigkeit (kurz: IND, von inclusion dependency)
- X-Werte in einer Relation $r_1(R_1)$ kommen auch als Y-Werte in einer Relation $r_2(R_2)$ vor: Inklusionsabhängigkeit $R_1[X] \subseteq R_2[Y]$