

# Introduction to Data and Knowledge Engineering Sommersemester 2010



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

Tutorium 4: Funktionale Abhängigkeiten - Syntheseverfahren

21. Mai 2010

## Aufgabe 4.4 Terminologie wichtiger Begriffe im Kontext des Relationalen Modells

Machen Sie sich nochmal mit den in der Vorlesung eingeführten Begriffen vertraut. Die Begriffe werden nicht von allen Leuten exakt gleich verwendet, aber in den vielen Fällen sind sich die Leute darüber einig was Sie bedeuten.

Seien die folgenden Begriffe gegeben:

Attribut, Attributwert, Datenbank, Datenbankschema, Fremdschlüssel, Fremdschlüsselbedingung, Relation, Relationenschema, Tupel, Schlüssel, Primärschlüssel, Abhängigkeitstreue, Verbundtreue, Wertebereich

Weisen sie jedem dieser Begriffe diejenige Beschreibung zu welche am besten passt.

Begriff	Informelle Beschreibung
	Spalte einer Tabelle
	Mögliche Werte eines Attributes (auch Domäne)
	Element eines Wertebereichs
	Menge von Attributen
	Menge von Zeilen (Datensätze) einer Tabelle
	Zeile einer Tabelle
	Menge von Relationenschemata
	Menge von Relationen (Basisrelationen)
	Minimale Menge von Attributen, deren Werte eine Zeile einer Tabelle eindeutig identifizieren
	Ein beim Datenbankentwurf ausgezeichneter Schlüssel
	Attributmenge, die in einer anderen Relation ein Schlüssel ist
	Alle Attributwerte des Fremdschlüssels tauchen in der anderen Relation als Werte des Schlüssels auf
	Alle gegebenen Abhängigkeiten sind durch Schlüssel repräsentiert
	Die Originalrelationen können durch den Verbund der Basisrelationen wiedergewonnen werden

---

## Aufgabe 4.5 Vereinfachtes Syntheseverfahren

---

Ein Datenbank-Admin hat von einem Unternehmen den Auftrag erhalten, eine existierende, komplexe Datenbank hinsichtlich besserer Strukturierung zu optimieren. Unter den zahllosen Tabellen innerhalb der Datenbank, hat er eine besondere Tabelle in Augenschein genommen, welche geradezu prädestiniert ist, optimiert zu werden.

Der DB-Admin erinnert sich an jene Tage seines Studiums an der TUD, als er in einer DKE Vorlesung ein Verfahren erlernt hatte, welches ermöglicht Relationen-Schemata algorithmisch zu zerlegen um damit höhere Normalformen zu erreichen. Die Tabelle um die es geht, hat folgendes Schema (ohne Angabe der Schlüssel):

TicketID	TransactionNr	CustomerOp	AccountMngrOp	ItemTitle	ItemID
⋮	⋮	⋮	⋮	⋮	⋮

Zwecks Lesbarkeit vereinfachen wir das unübersichtliche Relationen-Schema zu der folgenden abstrahierte Form:

A	B	C	D	E	F
⋮	⋮	⋮	⋮	⋮	⋮

Weiterhin sei bekannt, dass innerhalb der Tabelle folgende funktionale Abhängigkeiten existieren:

- $AE \rightarrow BD$ ,
- $B \rightarrow D$ ,
- $CD \rightarrow A$ ,
- $CE \rightarrow DF$ ,
- $DFB \rightarrow CE$

Gehen Sie nun wie folgt vor, bestimmen Sie zunächst sämtliche Schlüsselkandidaten, welche für die gegebene Relationenschema in Frage kommen, sowie die höchste vorliegende Normalform (max. 3 NF).

Wenden Sie anschließend das vereinfachte Syntheseverfahren inklusive mit dummy-FD an, wie auf dem Aufgabenblatt 4 Übung angegeben ist. Überlegen Sie sich, wozu die dummy-FD gebraucht wird. Prüfen Sie im Anschluß darauf, ob das Ergebniss die 3 NF erfüllt. Ist dies nicht der Fall, so geben Sie bitte an, weshalb nicht und wie sich dies reparieren lässt.

Erklären Sie wie die Begriffe „Aufspalten von Relationen“, „Ausgangsrelation“ sowie „Informationsverlust“ miteinander zusammenhängen.