

Introduction to Data and Knowledge Engineering Sommersemester 2010



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Übung 3 Funktionale Abhängigkeiten und Normalformen

14. Mai 2010

Aufgabe 3.1 FDs Bestimmen 1

Aus den folgenden Aussagen soll ein einfaches Modell einer Universität erstellt werden. Dazu wollen wir die darin implizit enthaltenen Funktionalen Abhängigkeiten extrahieren.

- Ein Professor kann mehrere Vorlesungen halten und eine Vorlesung kann von mehreren Professoren gehalten werden.
- Es gibt zu jeder Vorlesung ein Foliensatz, unabhängig davon, welcher Professor sie hält. Verschiedene Vorlesungen verwenden nicht den gleichen Foliensatz. Jeder Foliensatz hat einen Professor als Autor.
- Die Studenten bekommen zur Vorlesung eine Note.

Da wir noch kein Modell haben, tun wir so als würde es nur einen Typ geben, in dem alle relevanten Konzepte durch je ein Attribut dargestellt sind. Das Modell besteht zunächst nur aus einer Menge von Attributen, also ein *schlüsselloses* Tabellen-Schema, das die Attribute aufzählt aber weder Primär- noch Fremdschlüssel festlegt.

Bennenen Sie die Attribute mit einzelnen Grossbuchstaben und verwenden Sie diese um die Funktionalen Abhängigkeiten anzugeben, die aus der obigen Beschreibung hervorgehen.

Aufgabe 3.2 FDs Bestimmen 2

Gegeben sei folgendes Tabelle-Schema ohne Angaben zu Primär- und Fremdschlüsseln, aber mit einigen schon eingefügten Datensätzen.

Dingsbums

A	B	C	D
a	1	q	r
m	1	b	8
a	1	l	r
b	3	m	0
b	2	p	0
c	3	v	r

Leider liegt keine Interpretation der Attribute vor. Leiten Sie die Menge der Funktionalen Abhängigkeiten direkt aus den vorhandenen Datensätzen ab.

FDs, die Sie auch durch die Anwendung von Regeln aus schon gefundenen FDs erhalten, können Sie weglassen. Z.B. lässt sich aus fast allen FDs durch Erweiterung, also durch die Regel $X \rightarrow Y \Rightarrow XZ \rightarrow Y$, eine zusätzliche FD herleiten.

Wählen Sie den einzigen Schlüsselkandidaten als Primärschlüssel. Ist das Modell (so weit Sie es aus den Daten sagen können) in 3NF?

Wenn nur ein Datensatz vorläge, z.B. die erste Zeile, welche FDs würden wir dann rauslesen? Was würde das für die Erfüllung der ersten drei Normalformen bedeuten?

Aufgabe 3.3 Äquivalenzklassen

Definition: Eine *Äquivalenzklasse* einer FD-Menge F ist eine Teilmenge von FDs in F , deren linke Seiten äquivalent sind. Eine Menge von Attributen einer FD-Menge sind *äquivalent*, wenn sie alle voneinander funktional abhängig sind.

In Aufgabe 2.3 haben wir eine für das schlüssellose Tabellen-Schema $R = (ABCDEFGHI)$ vorgegebene FD-Menge auf eine kleinere FD-Menge $F = \{B \rightarrow C, A \rightarrow DE, B \rightarrow F, F \rightarrow GH, D \rightarrow I\}$ reduziert.

Eine Methode, anhand einer FD-Menge das Schema in mehrere Schemata zu zerlegen, macht Folgendes. Man bestimmt die Äquivalenzklassen der FD-Menge und erstellt aus jeder dieser Teilmengen ein Schema, das die darin vorkommenden Attribute umfasst.

Bestimmen Sie die Äquivalenzklassen für die FD-Menge F , zerlegen sie R in die entsprechenden Teilschemata und bestimmen für jedes dieser Schemata die Schlüsselkandidaten.