

Introduction to Data and Knowledge Engineering Sommersemester 2010



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Tutorium 3: Funktionale Abhängigkeiten und Normalformen

14. Mai 2010

Aufgabe 3.4 Multiple Choice

Welche der folgenden Aussagen ist richtig? Begründen Sie jede Ihrer Antworten oder geben Sie ein Gegenbeispiel an!

In welcher Beziehung stehen Normalformen zueinander?

- Ist eine Relation in 3 NF, so ist sie auch in 2 NF
- Ist eine Relation in 2 NF, so ist sie auch in 3 NF
- 2 NF und 3 NF bedingen sich überhaupt nicht gegenseitig: Ist eine Relation in 2 NF, so kann sie auch in 3 NF sein, ist eine Relation in 3 NF, so kann sie auch in 2 NF sein, muss es aber nicht.

Wann ist eine Relation in 1 NF?

- Eine Relation ist immer in 1 NF
- Eine Relation ist nicht immer in 1 NF
- Eine Relation ist in 1 NF, wenn jedes Attribut der Relation einen atomaren Inhalt hat.

Was bedeutet die Aussage „alle Nichtschlüsselattribute hängen vollständig vom Primärschlüssel ab“?

- Jedes Attribut, welches nicht der Schlüssel einer Relation ist, hängt von mindestens einem Teil des Primärschlüssels dieser Relation ab.
- Kein Attribut, welches nicht der Schlüssel einer Relation ist, hängt ausschließlich von einem Teil des Primärschlüssels dieser Relation ab.
- Jedes Attribut, welches nicht der Schlüssel einer Relation ist, ist nicht transitiv vom Primärschlüssel dieser Relation abhängig.
- Jedes Attribut, welches nicht der Schlüssel einer Relation ist, ist von allen Schlüsseln dieser Relation abhängig.

Wann ist eine Relation in 2 NF?

- Wenn der Primärschlüssel einer Relation aus genau einem Attribut besteht, von dem alle anderen Attribute abhängig sind, ist die Relation in 2 NF
- Wenn eine Relation höchstens so viele Attribute wie Schlüssel besitzt, ist sie in 2 NF
- Wenn alle Nichtschlüsselattribute einer Relation vollständig vom Primärschlüssel abhängen, ist diese Relation in 2 NF

Wann ist eine Relation in 3 NF?

- Wenn der Primärschlüssel einer Relation aus genau einem Attribut besteht, von dem alle anderen Attribute der Relation abhängig sind, ist die Relation in 3 NF.
- Wenn eine Relation höchstens ein Nichtschlüsselattribut besitzt, welches vollständig vom Primärschlüssel abhängig ist, ist sie in 3 NF.
- Wenn eine Relation in 2 NF ist und kein Nichtschlüsselattribut ausschließlich von anderen Nichtschlüsselattributen abhängig ist, ist die Relation in 3 NF.

Was ist eine Determinant (im Datenbankenbereich)?

- Ein Determinant ist ein besonders wichtiges Attribut in einer Relation.
- Ein Determinant ist eine Attributmenge, von der irgendein anderes Attribut vollständig abhängig ist.
- Ein Determinant ist ein anderer Begriff für Schlüsselkandidat einer Relation.

Wann ist eine Relation in BCNF?

- Wenn eine Relation in 3 NF ist, ist sie auch in BCNF.
- Wenn eine Relation in 3 NF ist und jeder Determinant ein Schlüsselkandidat der Relation ist, ist sie in 3 NF.
- Wenn eine Relation in 3 NF ist und nur einen Determinant besitzt, ist sie in 3 NF.

Aufgabe 3.5 1. Normalform

Ein Hotel in einem Schweizer Skigebiet sucht einen Mitarbeiter / eine Mitarbeiterin zur Besetzung der Rezeption. Die Daten der bisherigen Bewerber werden im Augenblick folgendermaßen gespeichert (Schlüsselattribute sind in grau gekennzeichnet, keine Fremdschlüssel):

Bewerbungen

ID	Vorname	Nachname	Geburtsdatum	Sprachen	Adresse
001	Anna	Müller	4.5.1980	Deutsch, Englisch, Französisch	Beethovenweg 83, 12345 Großstadt
002	Werner	Hagen	6.7.1975	Deutsch, Englisch	Hauptstraße 5, 34342 Kleindorf
003	Claire	Calmet	12.12.1989	Französisch	Boulevard Wilson 33, 06600 Antibes
004	Lisa	Schmidt	1.4.1977	Deutsch, Englisch, Spanisch	Berliner Straße 19, 45671 Anderstadt
005	Harry	Genius	4.3.1990	Deutsch, Englisch, Französisch, Spanisch, Schwedisch, Finnisch, Russisch	Bessunger Straße 44, 64285 Darmstadt

Ist das Model in 1 NF? Falls nein, wie könnte es verändert werden, damit 1 NF erfüllt wird? Tragen Sie ggf. zur Veranschaulichung auch die gegebenen Beispieldaten in Ihr verändertes Relationenschema ein.

Aufgabe 3.6 2. Normalform

Ein Fachgebiet bietet verschiedene Lehrveranstaltungen an. Jede dieser Lehrveranstaltungen hat einen Namen (VName) und einen eindeutigen Schlüssel (VID), welcher der Veranstaltung vom Prüfungssekretariat des Fachbereichs zugewiesen wurde. Das Fachgebiet erlaubt in jeder seiner Veranstaltungen generell, dass Gruppen von Studenten gemeinsam die Übungen bearbeiten. Die maximale Größe dieser Gruppen wird zunächst nicht festgelegt, da sie in jeder Veranstaltung anders sein kann. Jede dieser Gruppen bekommt hierfür eine innerhalb der Veranstaltung eindeutige Gruppennummer (GID). Studenten besitzen außerdem jeweils eindeutige Matrikelnummern (Matr), denen wiederum Vornamen und Nachnamen zugeordnet sind. Ein Student kann mehrere der angebotenen Lehrveranstaltungen besuchen, aber innerhalb einer Lehrveranstaltung nur einer Gruppe zugeordnet sein. In einer Datenbank sollen diese Zuordnungen gespeichert werden. In einem ersten Versuch entstand folgendes Relationen-Schema (Schlüsselattribute sind in grau gekennzeichnet, keine Fremdschlüssel), für das zur Veranschaulichung einigen Beispiel-Entitäten angegeben sind:

Veranstaltung

VID	VName	GID	Matr	Vorname	Nachname
345	Angewandte Physik 1	1	1234567	Arno	Schmidt
345	Angewandte Physik 1	1	7654321	Anna	Müller
345	Angewandte Physik 1	2	3344559	Helmut	Groß
345	Angewandte Physik 1	2	3366771	Kevin	Amundsen
675	Theoretische Physik 1	1	5829455	Jacqueline	Herr
675	Theoretische Physik 1	1	3452649	Angela	Bennett
675	Theoretische Physik 1	2	1234567	Arno	Schmidt
675	Theoretische Physik 1	2	3344559	Helmut	Groß

Welche Probleme können hier auftreten? Was kann beispielsweise passieren, wenn Jacqueline sich jetzt auch in eine Übungsgruppe der Veranstaltung Angewandte Physik 1 eintragen lassen möchte?

Ist das Modell (d.h. das Relationen-Schema zusammen mit den Funktionalen Abhängigkeiten) in 2 NF? Falls nein, wie müsste es verändert werden, damit 2 NF erfüllt wird? Tragen Sie ggf. zur Veranschaulichung auch die gegebenen Beispieldaten in Ihr verändertes Relationenschema ein.

Aufgabe 3.7 3. Normalform und Boyce-Codd-Normalform

Eine Supermarktkette besitzt Filialen ganz Deutschland, von denen jede eine eindeutige Filialnummer (FID) hat. Für jede Filiale wird die Adresse gespeichert (Straße, Hausnummer, PLZ und Stadt) sowie der Nettoeinkommensdurchschnitt der Einwohner, welche in dem PLZ-Bereich wohnen, in dem sich die Filiale befindet (E-AVG). Mit dieser Information soll das Warenangebot in der entsprechenden Filiale reguliert werden. Bisher werden diese Informationen in folgendem Relationen-Schema (Schlüsselattribute sind in grau gekennzeichnet, keine Fremdschlüssel), das mit hier einigen Beispiel-Instanzen gefüllt ist:

Filiale

FID	Straße	Haus	PLZ	Stadt	E-AVG
421	Hauptstraße	12	45632	Altstadt	1 500
326	Nebengasse	42	56723	Armestadt	450
798	Große Straße	63	40195	Mittelstadt	2 000
667	Kleine Straße	55	40195	Mittelstadt	2 000
592	Schöne Straße	4	59603	Reichstadt	10 000
533	Lindenstraße	1	55555	Reichstadt	9 000
444	Lange Straße	1044	59603	Reichstadt	10 000
495	Großer Weg	230	34896	Halbestadt	3 000
245	Mittelweg	45	46245	Anderestadt	4 000
405	Villenstraße	7	40696	Millonärsstadt	2 000 000

Was ist an dem Relationenschema nicht optimal?

Ist das Modell in 3 NF? Falls nein, wie müsste es verändert werden, damit 3 NF erfüllt wird? Tragen Sie ggf. zur Veranschaulichung auch die gegebenen Beispieldaten in Ihr verändertes Relationenschema ein.

Überprüfen Sie anschließend, ob Ihre Lösung auch BCNF erfüllt.