

	BG - Schwerpunkt Datenverarbeitungstechnik 13.1 – Datenbanken	Nr. 7-01
	Normalisierung Einführung	ag

Hinweis für alle folgenden Aufgaben

Ist in den folgenden Tabellen ein Attribut unterstrichen, so handelt sich um den Primärschlüssel. Dieser Schlüssel kann auch aus mehreren Attributen bestehen.

Wenn Sie Tabellen erstellen, so sind die Primärschlüssel zu unterstreichen. Eventuelle Fremdschlüssel sind durch das Anhängen einer Raute (#) kenntlich zu machen.

1. Aufgabe

Man spricht von einem unnormalisierten Datenbankschema, wenn so genannte Anomalien (manchmal auch Mutations-anomalien genannt) auftreten können. Anomalien haben schädliche Folgen für die Konsistenz einer Datenbank.

Es gibt drei Arten: **Einfügeanomalie, Änderungsanomalie, Löschanomalie**

Erläutern Sie die drei Anomalien unter Verwendung der folgenden (unnormalisierten) Datenbanktabelle.

FACH-LEHRER

<u>Lehrerkürzel</u>	<u>Lehrername</u>	<u>Fachname</u>	<u>FachleiterKürzel</u>	<u>FachhelferName</u>
Hei	Heitech	Info	Hei	Lotech
Mid	Middletech	Info	Hei	Lotech
Zwi	Zwitech	Info	Hei	Lotech
Jag	Jagger	Musik	Jag	Brahms
Sch	Schiller	Deutsch	Sch	Ratlos
Goe	Goethe	Deutsch	Sch	Ratlos

Anmerkung: 2 Attribute (unterstrichen) bilden zusammen den Primärschlüssel (weil ein Fach von zwei Lehrern unterrichtet werden kann und ein Lehrer meist zwei Fächer unterrichtet)

FACH-LEHRER

<u>Lehrerkürzel</u>	<u>Lehrername</u>	<u>Fachname</u>	<u>FachleiterKürzel</u>	<u>FachhelferName</u>
Hei	Heitech	Info	Hei	Lotech
Mid	Middletech	Info	Hei	Lotech
Zwi	Zwitech	Info	Hei	Lotech
Jag	Jagger	Musik	Jag	Brahms
Sch	Schiller	Deutsch	Schil	Ratlos
Goe	Goethe	Deutsch	Schil	Ratlos

Was kann man zu der Tabelle sagen?

Die Tabelle enthält viele Redundanzen.

Wie wirkt es sich auf die Datenintegrität aus, wenn ... ?

- **Einfügeanomalie**

Ein neues Fach Kunst soll eingeführt werden, jedoch ist noch kein Lehrer dafür verfügbar.

Das Fach kann nicht eingeführt werden - man muss warten bis ein Lehrer einsetzbar ist, weil das Attribut Lehrerkürzel zum Schlüssel gehört und immer einen Wert (NOT NULL) haben muss.

- **Löschanomalie**

Lehrer Jagger (Jag) wird pensioniert.

Wenn Jagger die Schule verlässt, dann muss das ganze Fach Musik gelöscht werden. Außerdem geht die Information verloren, dass Brahms Fachhelfer ist.

- **Änderungsanomalie**

Das Fach Info wird der Fachhelfer Lotech durch den Fachhelfer Heiend ersetzt.

Wenn Fiffikus statt Lotech Fachhelfer wird, dann müssen Änderungen in drei Datensätzen gemacht werden.

	BG - Schwerpunkt Datenverarbeitungstechnik 13.1 – Datenbanken	Nr. 7-01
	Normalisierung Einführung	ag

Beim Einfügen von Daten in eine Datenbank spricht man von einer **Einfüge-Anomalie (Insertion-Anomalie)**, wenn eine neue Entität eines Entitätstyps in die Datenbank nicht oder nur schwierig eingetragen werden kann, weil nicht zu allen Schlüssel-Attributen Werte vorliegen. So können beispielsweise Informationen nicht aufgenommen werden, da andere, zu diesem Zusammenhang bzw. Zeitpunkt (noch) uninteressante Angaben fehlen.

Eine **Lösch-Anomalie (Delete-Anomalie)** entsteht, wenn durch das Löschen eines Datensatzes mehr Informationen als erwünscht verloren gehen. Sie entsteht, wenn ein Datensatz mehrere unabhängige Informationen enthält. Durch das Löschen der einen Information wird dann auch die andere gelöscht, obwohl diese noch benötigt wird.

Beim Ändern von Daten in einer Datenbank spricht man von einer **Änderungs-Anomalie (Update-Anomalie)**, wenn nicht alle (redundanten) Vorkommen einer Entität zugleich geändert werden. Dieses kann zu inkonsistenten Daten führen.

(siehe wikipedia)

2. Aufgabe

Nennen Sie die Definitionen der ersten drei Normalformen nach Codd.

(Es gibt noch weitere Normalformen, welche aber in der Praxis nur eine geringe Rolle spielen.)

Lösung auch für 3. Aufgabe und 5. Aufgabe

Die erste Normalform ist dann gegeben, wenn **alle Attributwerte atomar** sind.

Tabelle **ProjektMitarbeiter** (vor Normalisierung)

MNr	Name	Ort	Projekt
200	Huber	München	1, 7
207	Schmitz	Köln	5, 7, 11

Tabelle **ProjektMitarbeiter** (in erster Normalform)

MNr	ProjNr	Name	Ort
200	1	Huber	München
200	7	Huber	München
207	5	Schmitz	Köln
207	7	Schmitz	Köln
207	11	Schmitz	Köln

Die zweite Normalform ist dann gegeben, wenn

a) die erste Normalform gegeben ist und

b) **alle Nichtschlüsselattribute voll von jedem Schlüssel abhängig** sind.

Die Tabelle PROJEKTMITARBEITER (in 1. Normalform) enthält den zusammengesetzten Schlüssel (*MNr, ProjNr*). Die Attribute *Name* und *Ort* sind jedoch nicht voll vom Schlüssel sondern **nur von einem Teil des Schlüssels (*MNr*) abhängig**.

Folglich muss die Tabelle in zwei Teiltabellen zerlegt werden:

Tabelle **Mitarbeiter** (in 2. Normalform)

MNr	Name	Ort
200	Huber	München
207	Schmitz	Köln

Tabelle **ProjektMitarbeiter** (in 2. NF)

MNr	ProjNr
200	1
200	7
207	5
207	7
207	11

	BG - Schwerpunkt Datenverarbeitungstechnik 13.1 – Datenbanken	Nr. 7-01
	Normalisierung Einführung	ag

Die dritte Normalform ist dann gegeben, wenn
a) die zweite Normalform gegeben ist und
b) **kein Nichtschlüsselattribut von irgendeinem Schlüssel transitiv abhängig ist.**

Tabelle **Abteilungsmitarbeiter** (in 2. NF)

MNr	Name	Ort	Abteilung	Bezeichnung
200	Huber	München	A5	Verkauf
207	Schmitz	Köln	A2	Informatik
355	Berger	Frankfurt	A5	Verkauf

Aus der Tabelle ABTEILUNGSMITARBEITER (in 2. Normalform) geht hervor, dass das Attribut *Bezeichnung* transitiv (mittelbar) vom Attribut *Abteilung* abhängig ist.

Folglich muss die Tabelle in zwei Teiltabellen zerlegt werden, wobei das Attribut *Abteilung* in der Tabelle MITARBEITER als Fremdschlüssel fungiert:

Tabelle **Mitarbeiter** (in 3. Normalform)

MNr	Name	Ort	Abteilung
200	Huber	München	A5
207	Schmitz	Köln	A2
355	Berger	Frankfurt	A5

Tabelle **Abteilung** (in 3. NF)

AbtNr	Bezeichnung
A2	Informatik
A5	Verkauf

siehe auch

<http://www.oszhd.de/schule.de/gymnasium/faecher/informatik/datenbanken/normal/index.htm>

3. Aufgabe

Gegeben ist die folgende (unnormalisierte) Datenbanktabelle:

- Warum befindet sich die Tabelle nicht in der 1. Normalform?
- Überführen Sie die Tabelle in die 1. Normalform.

Tabelle **ProjektMitarbeiter** (vor Normalisierung)

MNr	Name	Ort	Projekt
200	Huber	München	1, 7
207	Schmitz	Köln	5, 7, 11

	BG - Schwerpunkt Datenverarbeitungstechnik 13.1 – Datenbanken	Nr. 7-01
	Normalisierung Einführung	ag

4. Aufgabe

- Zeigen Sie, dass sich diese Tabelle nicht in der zweiten, wohl aber in der ersten Normalform befindet.
- Überführen Sie die Tabelle in die 2. Normalform.

CD-ID	Albumtitel	Interpret	Track	Titel
4711	Century Child	Nightwish	4	Ever Dream
4711	Century Child	Nightwish	9	The Phantom of the Opera
8421	Invisible Touch	Genesis	3	Land of Confusion

Verletzung der 2.NF: Der Primärschlüssel der Relation ist aus den Feldern CD_ID und Track zusammengesetzt. (Grundsätzlich darf ein Primärschlüssel aus mehreren Attributen bestehen, jedoch entsteht daraus im genannten Beispiel ein Konflikt.)

Die Felder Albumtitel und Interpret sind vom Feld CD_ID abhängig, aber nicht vom Feld Track. Dies (Punkt 2) verletzt die 2. Normalform, da die drei nicht-primären Attribute nicht nur von einem Teil des Schlüssels (hier CD_ID) abhängen dürfen. Wäre Punkt 1 nicht erfüllt, so könnte dies nicht passieren.

2.NF (sinngemäß)

CD				Lied		
CD_ID	Albumtitel	Interpret	Jahr der Gründung	CD_ID	Track	Titel
4711	Not That Kind	Anastacia	1999	4711	1	Not That Kind
4712	Wish You Were Here	Pink Floyd	1964	4711	2	I'm Outta Love
4713	Freak of Nature	Anastacia	1999	4711	3	Cowboys & Kisses
				4712	1	Shine On You Crazy Diamond
				4713	1	Paid my Dues

Das Attribut *CD_ID* aus der Tabelle *Lied* bezeichnet man als Fremdschlüssel, der auf den Primärschlüssel der Tabelle *CD* verweist. Zugleich stellen die Attribute *CD_ID* und *Track* den zusammengesetzten Primärschlüssel der Tabelle *Lied* dar.

5. Aufgabe

In welcher Normalform (keine/1./2./3.) befindet sich diese Tabelle? Begründen Sie Ihre Aussage. Führen Sie, falls erforderlich, eine Normalisierung bis zur 3. Normalform durch.

Tabelle **Abteilungsmitarbeiter**

MNr	Name	Ort	Abteilung	Bezeichnung
200	Huber	München	A5	Verkauf
207	Schmitz	Köln	A2	Informatik
355	Berger	Frankfurt	A5	Verkauf

	BG - Schwerpunkt Datenverarbeitungstechnik 13.1 – Datenbanken	Nr. 7-01
	Normalisierung Einführung	ag

6. Aufgabe

Eine Beraterfirma beschäftigt IT- und Finanzberater. In einer Datenbank sollen Angaben zu den Beratern und ihren jeweiligen Kunden gespeichert werden. So wird z. B. der Kunde Kai Mauer von drei verschiedenen Personen beraten (Claus Thaler, Bud Weiser und Hein Ecken). Die folgende Tabelle soll ein erster Entwurf sein (deshalb ohne Markierung eines Primärschlüssels), welcher in eine Datenbankstruktur umzusetzen ist.

Unnormalisierte Tabelle

Kunden-ID	KundenName	Berater-ID	BeraterName	Aufgabe	Stundenlohn	Stunden	Total
1	Kai Mauer	1	Claus Thaler	IT-Berater	50,00	3	150,00
		2	Bud Weiser	Finanzberater	45,00	5	225,00
		4	Hein Ecken	Finanzberater	30,00	8	240,00
2	Karl S. Ruhe	2	Bud Weiser	Finanzberater	45,00	4	180,00
3	Hans A. Rostock	1	Claus Thaler	IT-Berater	50,00	2	100,00
4	Roman Tisch	1	Claus Thaler	IT-Berater	50,00	10	500,00
		2	Bud Weiser	Finanzberater	45,00	5	225,00
		4	Hein Ecken	Finanzberater	30,00	5	150,00
5	Timo Beil	3	Paul Aner	IT-Berater	60,00	12	720,00

Die in den folgenden Teilaufgaben zu erstellenden Tabellen sind vollständig mit allen relevanten Attributen und Attributwerten (z. B. "Finanzberater") anzugeben. Verwenden Sie dafür MS-EXCEL (o. ä.).

- Überführen Sie die Tabelle in die 1. Normalform und legen Sie den Primärschlüssel fest. Sollten die bereits vorhandenen Attribute für die Bildung des Primärschlüssels geeignet sein, so sind auch nur diese zu verwenden.
- Wandeln Sie die Tabelle aus a) in die 3. Normalform um. Ihr Lösungsweg muss auch die 2. Normalform enthalten.

1. Normalform:							
Kunden-ID	KundenName	Berater-ID	BeraterName	Aufgabe	Stundenlohn	Stunden	Total
1	Kai Mauer	1	Claus Thaler	IT-Berater	50,00	3	150,00
1	Kai Mauer	2	Bud Weiser	Finanzberater	45,00	5	225,00
1	Kai Mauer	4	Hein Ecken	Finanzberater	30,00	8	240,00
2	Karl S. Ruhe	2	Bud Weiser	Finanzberater	45,00	4	180,00
3	Hans A. Rostock	1	Claus Thaler	IT-Berater	50,00	2	100,00
4	Roman Tisch	1	Claus Thaler	IT-Berater	50,00	10	500,00
4	Roman Tisch	2	Bud Weiser	Finanzberater	45,00	5	225,00
4	Roman Tisch	4	Hein Ecken	Finanzberater	30,00	5	150,00
5	Timo Beil	3	Paul Aner	IT-Berater	60,00	12	720,00

2. Normalform:

Tabelle Kunden:

<u>Kunden-ID</u>	<u>KundenName</u>
1	Kai Mauer
2	Karl S. Ruhe
3	Hans A. Rostock
4	Roman Tisch
5	Timo Beil

Tabelle Berater:

<u>Berater-ID</u>	<u>BeraterName</u>	<u>Aufgabe</u>	<u>Stundenlohn</u>
1	Claus Thaler	IT-Berater	50,00
2	Bud Weiser	Finanzberater	45,00
3	Paul Aner	IT-Berater	60,00
4	Hein Ecken	Finanzberater	30,00

Tabelle Arbeit:

<u>Kunden-ID #</u>	<u>Berater-ID #</u>	<u>Stunden</u>	<u>Total</u>
1	1	3	150,00
1	2	5	225,00
1	4	8	240,00
2	2	4	180,00
3	1	2	100,00
4	1	10	500,00
4	2	5	225,00
4	4	5	150,00
5	3	12	720,00

3. Normalform:

Tabelle Kunden:

<u>Kunden-ID</u>	<u>KundenName</u>
1	Kai Mauer
2	Karl S. Ruhe
3	Hans A. Rostock
4	Roman Tisch
5	Timo Beil

Tabelle Berater:

<u>Berater-ID</u>	<u>BeraterName</u>	<u>Aufgaben-ID #</u>	<u>Stundenlohn</u>
1	Claus Thaler	1	50,00
2	Bud Weiser	2	45,00
3	Paul Aner	1	60,00
4	Hein Ecken	2	30,00

Tabelle Aufgabe:

<u>Aufgaben-ID</u>	<u>Bezeichnung</u>
1	IT-Berater
2	Finanzberater

Tabelle Arbeit:

<u>Kunden-ID #</u>	<u>Berater-ID #</u>	<u>Stunden</u>
1	1	3
1	2	5
1	4	8
2	2	4
3	1	2
4	1	10
4	2	5
4	4	5
5	3	12