### 2.3.2. Normalformen

Relationale Datenbanken müssen normalisiert werden. Damit werden Redundanzen und Anomalien (🡪 2.1.4.) verhindert. Redundanzen würden zu unbrauchbaren Datenbanken führen.

Oft ist es ausreichend, Datenbanken bis zur 3. Normalform zu bearbeiten.

ERSTE NORMALFORM: Eine Tabelle befindet sich in der ersten Normalform, wenn die Wertebereiche der Merkmale atomar sind.

Beispiel:

In einer Tabelle soll festgehalten werden, welche Fächer ein Schüler belegt und welche Zeugnisnoten er erhalten hat.

***Fachbelegung***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID\_Schüler | Name | ID\_Fach | Fachbezeichnung | Note |
| 1 | Willi Klein | 1 | Mathematik | gut |
| 2 | Deutsch | befriedigend |
| 2 | Udo Jung | 2 | Deutsch | gut |
| 3 | Englisch | ausreichend |

In dieser Tabelle sind:

* die Werte des Feldes „Name“ nicht atomar,
* jedem Schüler eine Liste von Fächern und Noten zugeteilt.

Um eine Tabelle in die 1. NF zu überführen, muss man

* Merkmale mit nicht atomaren Wertebereichen auf mehrere Merkmale aufteilen (Name 🡪 Vorname und Nachname)
* Listen sinngleicher Informationseinheiten entfernen, indem jeweils ein eigener Datensatz eingefügt wird
* den Schlüssel der Tabelle neu bestimmen.

***Fachbelegung***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID\_Schüler | ID\_Fach | Vorname | Nachname | Fach-bezeichnung | Note |
| 1 | 1 | Willi | Klein | Mathematik | gut |
| 1 | 2 | Willi | Klein | Deutsch | befriedigend |
| 2 | 2 | Udo | Jung | Deutsch | gut |
| 2 | 3 | Udo | Jung | Englisch | ausreichend |

Die Eindeutigkeit des Schlüssels ID\_Schüler geht verloren. Der Schlüssel der neuen Tabelle besteht aus einer Kombination der Merkmale ID-Schüler und ID-Fach.

ZWEITE NORMALFORM: Eine Tabelle befindet sich in der zweiten Normalform, wenn die erste Normalform erfüllt ist und alle Nichtschlüsselmerkmale vom zusammengesetzten Schlüssel voll funktional abhängig sind.

Ein Merkmal Y ist FUNKTIONAL ABHÄNGIG vom Merkmal X, wenn zu jedem möglichen Wert von S genau ein Wert aus Y existiert.

Man schreibt: X 🡪 Y

Die funktionale Abhängigkeit ist erfüllt für folgende Nichtschlüsselattribute

ID\_Schüler -🡪 Vorname

ID\_Schüler -🡪 Nachname

ID\_Fach 🡪 Fachbezeichnung

Es besteht aber keine funktionale Abhängigkeit zwischen ID\_Schüler und Note bzw. zwischen ID\_Fach und Note.

Kombiniert man aber beide Schlüssel, so entsteht eine eindeutige Zuordnung:

(ID\_Schüler, ID\_Fach) 🡪 Note

In solchen Fällen nennt man das Merkmal „Note“ voll funktional abhängig vom zusammengesetzten Gesamtschlüssel (ID\_Schüler, ID\_Fach).

Ein Merkmal Y ist VOLL FUNKTIONAL ABHÄNGIG vom einem aus X1 und X2 zusammengesetzten Schlüssel, wenn Y funktional abhängig vom Gesamtschlüssel, nicht aber von seinen Teilschlüsseln ist

Man schreibt: (X1, X2) ⇒ Y

Unser Beispiel zeigt: Wenn eine Tabelle einen zusammengesetzten Schlüssel besitzt, treten bei den Merkmalen redundante Datensätze auf, die bereits von einem Teilschlüssel funktional abhängig sind (Willi Klein doppelt eingetragen).

Überführung einer Tabelle in die zweite Normalform

|  |  |
| --- | --- |
| Bestimme alle Nichtschlüsselattribute, die bereits von einem Teilschlüssel funktional abhängig sind.  Diese Merkmale sind nicht voll funktional abhängig vom Gesamtschlüssel und müssen aus der Tabelle entfernt werden. | ID\_Schüler -🡪 Vorname  ID\_Schüler -🡪 Nachname  ID\_Fach 🡪 Fachbezeichnung |
| Bilde aus den Teilschlüsseln und allen von ihr funktional abhängigen Nichtschlüsselmerkmalen eigene Tabellen. | ***Schüler***   |  |  |  | | --- | --- | --- | | ID\_Schüler | Vorname | Nachname | | 1 | Willi | Klein | | 2 | Udo | Jung |   ***Fach***   |  |  | | --- | --- | | ID\_Fach | Fachbezeichnung | | 1 | Mathematik | | 2 | Deutsch | | 3 | Englisch | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Entferne aus der ursprünglichen Tabelle Fachbelegung alle nicht voll funktional abhängigen Nichtschlüsselmerkmale. | |  |  |  | | --- | --- | --- | | ID\_Schüler | ID\_Fach | Note | | 1 | 1 | gut | | 1 | 2 | befriedigend | | 2 | 2 | gut | | 2 | 3 | ausreichend | |

DRITTE NORMALFORM: Eine Tabelle befindet sich in der dritten Normalform, wenn die zweite Normalform erfüllt ist und kein Nichtschlüsselmerkmal vom Schlüssel transitiv abhängig ist.

Beispiel:

In einer Tabelle sollen die Namen der Schüler und die Namen der Klassenlehrer erfasst werden.

***Schüler***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID\_Schüler | Nachname | Vorname | ID\_Klassenlehrer | Klassenlehrername |
| 1 | Ernst | Uwe | 1 | Moll |
| 2 | Müller | Kai | 1 | Moll |
| 3 | Specht | Eva | 2 | Kurp |
| 4 | Groß | Ute | 2 | Kurp |

Diese Tabelle befindet sich bereits in der ersten Normalform (atomare Merkmale, keine Listen) und in der zweiten Normalform (1. NF erfüllt und keine zusammengesetzten Schlüssel).

Es existieren folgende funktionale Abhängigkeiten:

ID\_Schüler 🡪 Nachname

ID\_Schüler 🡪 Vorname

ID\_Schüler 🡪 ID\_Klassenlehrer

ID\_Klassenlehrer 🡪 Klassenlehrername

ID\_Schüler 🡪 Klassenlehrername

Hier zeigt sich folgende funktionale Abhängigkeit:

ID\_Schüler 🡪 ID\_Klassenlehrer 🡪 Klassenlehrername

Wenn man aus „S bestimmt A“ und „A bestimmt B“ folgern kann, dass zwangläufig auch „S bestimmt B“ gilt, dann ist TRANSITIVITÄT gegeben.

Dies kann genau dann zu Redundanzen führen, wenn das zweite Merkmal (ID\_Klassenlehrer) kein mögliches Schlüsselmerkmal der Tabelle ist. Das wäre genau dann der Fall, wenn es eine funktionale Abhängigkeit des Merkmales 1 (ID\_Schüler) vom Merkmal 2 (ID\_Klassenlehrer) gäbe.

Ein Merkmal B ist TRANSITIV ABHÄNGIG vom Merkmal S, wenn es ein Merkmal A gibt, so dass gilt:

B ist funktional Abhängig von A A 🡪 B

A ist funktional abhängig von S S 🡪 A

S ist nicht funktional abhängig von A A ↛ S

Überführung einer Tabelle in die dritte Normalform:

1. Bestimme alle Nichtschlüsselmerkmale, die transitiv vom Schlüssel (ID\_Schüler) abhängig sind.

***Schüler***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID\_Schüler | Nachname | Vorname | ID\_Klassenlehrer | Klassenlehrername |
| 1 | Ernst | Uwe | 1 | Moll |
| 2 | Müller | Kai | 1 | Moll |
| 3 | Specht | Eva | 2 | Kurp |
| 4 | Groß | Ute | 2 | Kurp |

1. Bilde aus dem transitiv abhängigen Nichtschlüsselmerkmal (Klassenlehrername) und dem Nichtschlüsselattribut, von dem es funktional abhängig ist (ID\_Klassenlehrer), eine eigene Tabelle.

**Klassenlehrer**

|  |  |
| --- | --- |
| ID\_Klassenlehrer | Klassenlehrername |
| 1 | Moll |
| 2 | Kurp |

1. Entferne aus der ursprünglichen Tabelle (Schüler) alle transitiv abhängigen Nichtschlüsselmerkmale.

***Schüler***

|  |  |  |  |
| --- | --- | --- | --- |
| ID\_Schüler | Nachname | Vorname | ID\_Klassenlehrer |
| 1 | Ernst | Uwe | 1 |
| 2 | Müller | Kai | 1 |
| 3 | Specht | Eva | 2 |
| 4 | Groß | Ute | 2 |