

1. Datenmodelle

1.1 Das Entität-Relationship-Modell (Objekt-Beziehungs-Modell)

Daten enthalten Informationen über die Eigenschaften von Objekten.

Diejenigen Eigenschaften eines Objektes, die für eine bestimmte Aufgabe ausgewählt werden, nennt man **Attribute (Merkmale)**.

Um ein reales System im Modell darstellen zu können, benötigt man Verfahren, wodurch es möglich ist die Eigenschaften des Systems darzustellen, d.h. man arbeitet am Modell.

Bsp.: Teilchenmodell in der Physik

Atommodell in der Chemie

Kommunikationsmodell in Deutsch etc.

Das fertige Modell kann dann in geeigneter Form auf den Computer abgebildet werden. Diese Vorgehensweise ist schon vom Programmieren bekannt, wo durch die Erstellung von Struktogrammen der Entwurf eines Programmes erleichtert wird. Diese graphischen Darstellungsweisen sind meist genormt.

1.1.1 Entität-Typen (Objekt-Typen)

- Eine **Entität (Objekt)** ist ein reales oder abstraktes Objekt der realen Welt, das über **Attribute (Merkmale)** verfügt.

Bsp.: - Personen und Institutionen

- Sachen

- Vorgänge

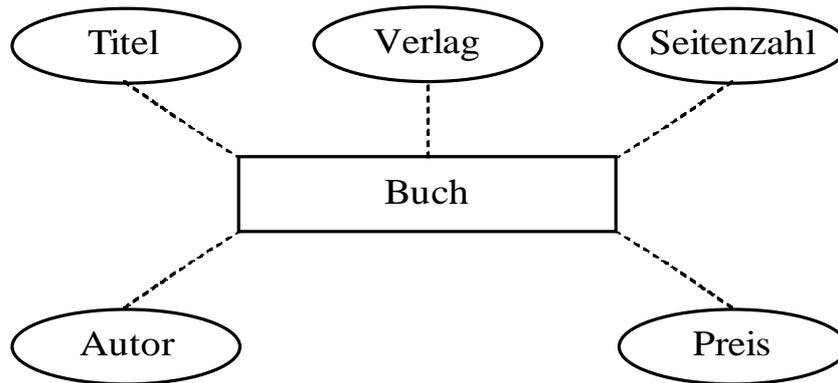
- Im Datenmodell wird eine Entität (Objekt) durch diese Attribute (Merkmale) beschrieben, d. h. zwei Entitäten (Objekte) müssen sich in mindestens einem Attribut (Merkmal) unterscheiden.
- Klassen gleichartiger Entitäten (Objekte) werden durch einen Entität (Objekt)-Typ beschrieben.

Bsp.: Schulklasse (Entität-Typ)

Schüler (Entität)

- Es werden zuerst die Entitäten (Objekte) definiert und dann werden die Beziehungen graphisch dargestellt.
Rechtecke für Entitäten, Ovale für Attribute und gestrichelte Linien (Kanten) zur Verbindung der Entitäten mit ihren Attributen.

Bsp.:



- Eine Teilmenge der Attributmenge (Merkmalsmenge), welche die Entitäten (Objekte) eindeutig identifiziert heißt **Schlüssel**.
- Oft wird hierfür ein künstliches Attribut (Merkmal) verwendet, da z. B. mit Nachname und Vorname eine Person noch nicht eindeutig identifiziert ist.

1.1.2 Beziehungen und Entität-Relationship-Diagramme

Die Beziehungen zwischen Entität (Objekt)-Typen werden in Entität-Relationship-Diagrammen (ER-Diagramm) als Rauten dargestellt. Die Beziehungen werden durch durchgezogene Linien (Kanten) angebunden.

- **1 : 1 - Relation** (Beziehung):



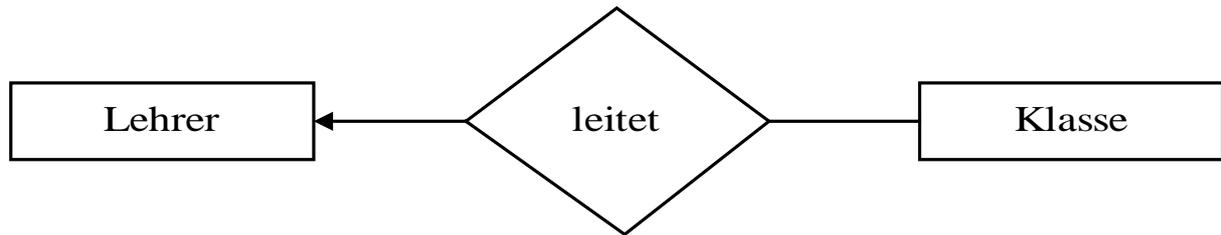
Der Schüler hat **genau eine** Schülernummer.

- **n : 1 - Relation** (Beziehung):



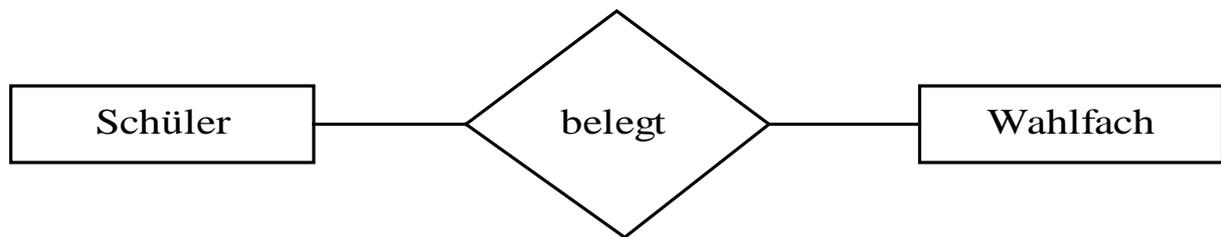
Ein Schüler gehört **höchstens** zu **einer** Klasse.

- **1 : m - Relation** (Beziehung):



Höchstens ein Lehrer leitet eine Klasse.

- **n : m - Relation** (Beziehung):

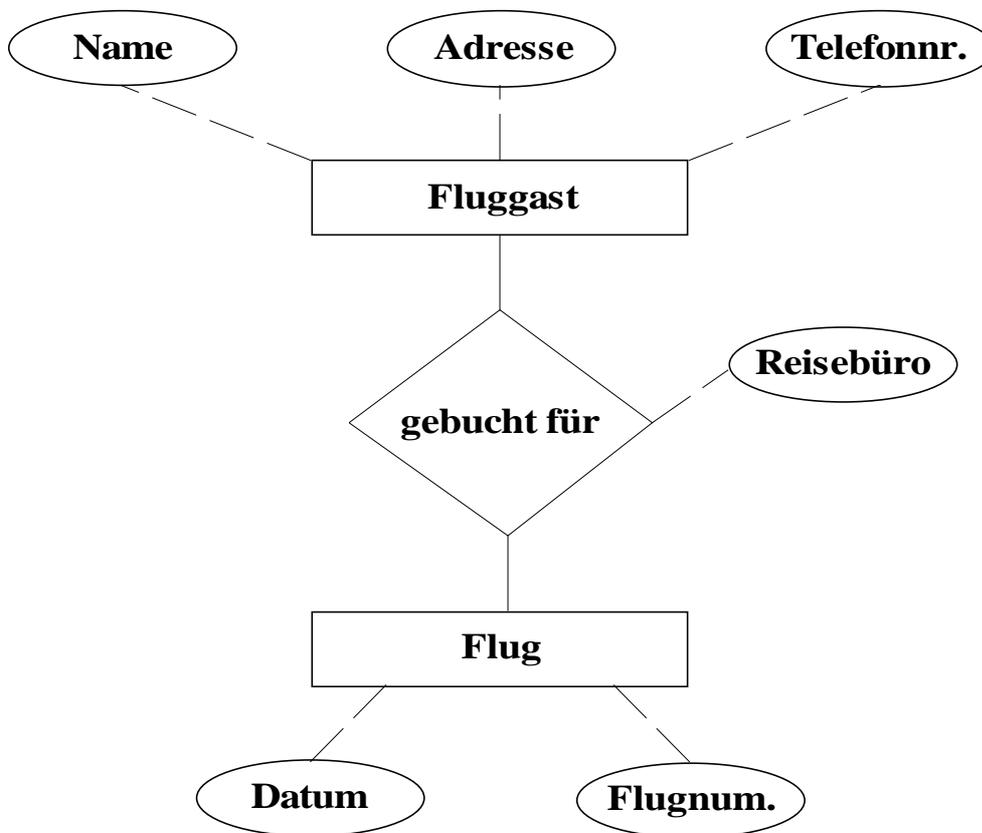


Verschiedene Schüler belegen verschiedene Wahlfächer.

Bsp.:

A large rectangular area filled with a grid of dashed lines, intended for the user to write an example of the relationships described above.

1.1.3 ER-Diagramm Luftdienst



1.2 Das relationale Datenmodell

1.2.1 Tabellen (Relationen)

- In **relationalen Datenbanken** werden alle Informationen in Form von Tabellen (**Relationen**) abgelegt.
- Die **Spaltenbezeichnungen** enthalten die **Attribute** (Merkmale).
- Die unter den Spaltenbezeichnungen aufgeführten Werte sind die **Attributwerte** (Merkmalswerte).
- In einer Zeile der Tabelle steht dann zu **jedem Attribut** (Merkmal) jeweils **ein Wert**.
- So eine Tabellenzeile stellt **einen Datensatz** dar, d. h. Merkmale werden zu Datensätze zusammengefaßt.

Für so eine Tabelle müssen die **beiden** Bedingungen gelten:

1. Jede Zeile der Tabelle tritt nur einmal auf, d. h. zwei Zeilen **unterscheiden** sich also **mindestens** in einem Attributwert (Merkmalswert).

Bsp.:

Nachname	Vorname	Gebdatum	Wohnort
...
Huber	Max	12.03.67	Garching
Huber	Max	31.06.54	Garching
...

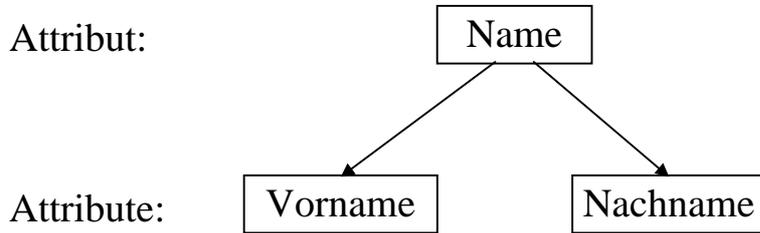
- Ergebnis einer Suche nicht eindeutig
 ⇒ wenn Information nicht eindeutig
 ⇒ dann wertlos
- Jeder Datensatz muß **ein Attribut** haben, welches **eindeutig** ist. (Ideal Fingerabdruck oder genetischer Code bei Personen).

Künstliches Attribut

- ideal positive, ganze Zahlen
- Kontonummer, Bankleitzahl, Schecknummer, Telefonnummer, PLZ, Kennwort (Geheimzahl bei EC-Karte)
- Das eindeutige Merkmal eines Datensatzes nennt man den **Schlüssel**. So, wie ein Türschlüssel zu einer ganz bestimmten Tür paßt, paßt das Schlüsselmerkmal zu einem ganz bestimmten Datensatz.

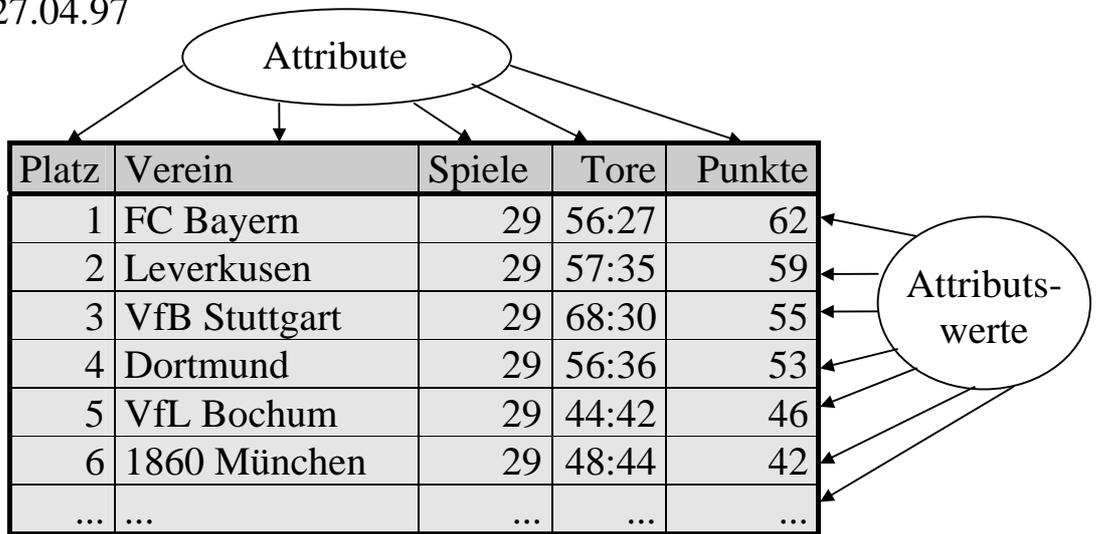
2. Alle Attributwerte (Merkmalswerte) sind nicht weiter zerlegbar, sie sind **atomar**.

Bsp.:



In der Tabelle treten keine Werte wie etwa Adressen auf, die selbst wieder aus Straße, Hausnummer, Postleitzahl und Ort bestehen.

Bsp.: Stand 27.04.97



Bsp.:

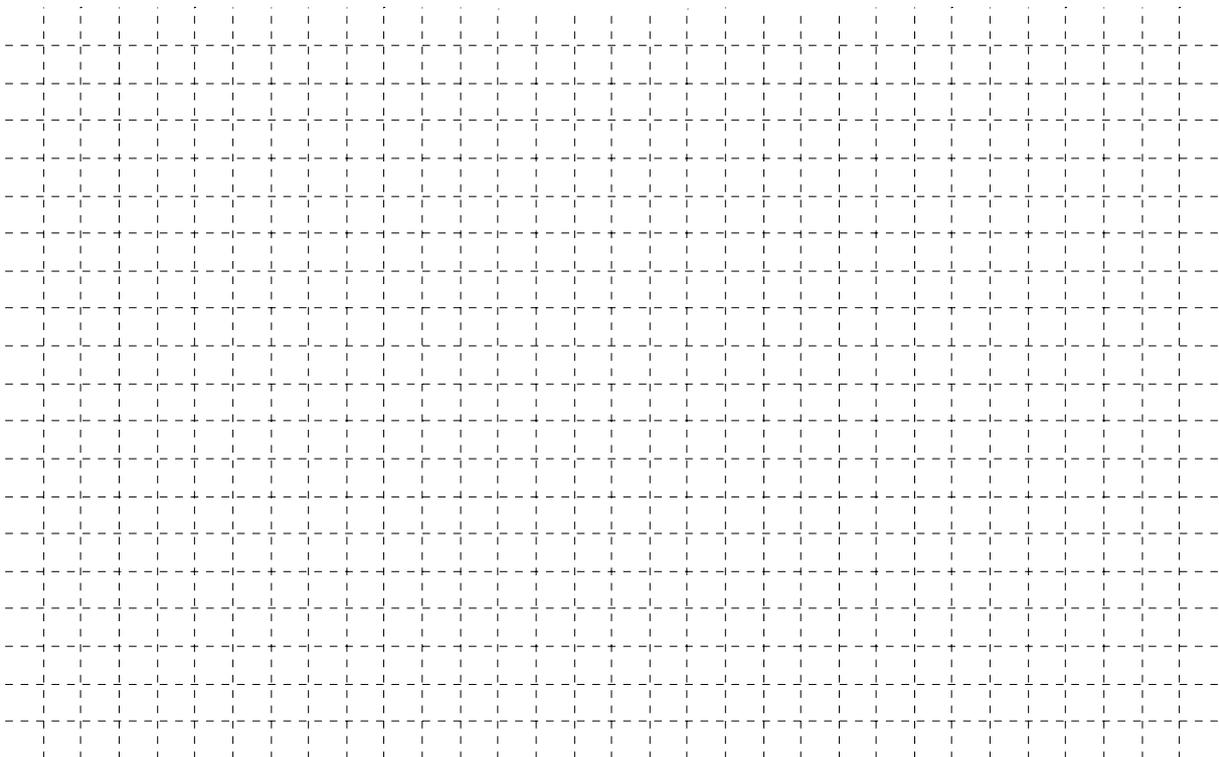
1.2.2 Vom ER-Modell zu den Tabellen

Die grafische Beschreibung eines Systems im ER-Modell liefert die Grundlage für eine Abbildung dieses Systems auf eine relationale Datenbank.

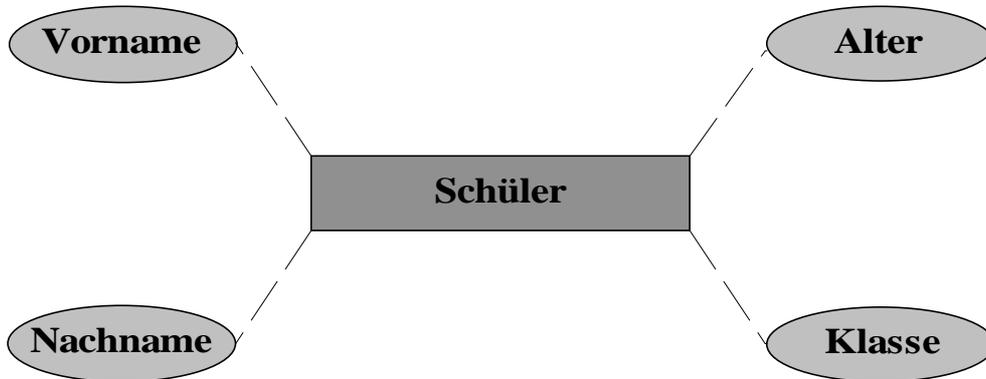
Man benötigt dafür Regeln, nach denen sich die Entität (Objekt)-Typen und Beziehungen des ER-Modells in Relationstabellen übertragen lassen.

- Beziehungen entsprechen den **Entität (Objekt)-Typen** des ER-Modells.
- Die **Attribute (Merkmale)** des Entität (Objekt)-Typs bilden auch die Attribute (Merkmale, Spalten) einer Tabelle.
- Einzelne Entitäten (Objekte), gekennzeichnet durch eine spezielle Kombination von Attributs (Merkmals-)werten, entsprechen einer Zeile (Datensatz) der Tabelle.
- Im ER-Modell werden **abstrakte** Zusammenhänge zwischen Systemteilen beschrieben, ohne daß **konkrete Daten** über einzelne Entitäten (Objekte) im Modell enthalten sind.
- In Tabellen dagegen spiegelt sich die abstrakte Struktur in den Attributen (Merkmalen) und **zusätzlich** sind die Einzelinformationen über alle beschriebenen Entitäten (Objekte) in der Tabelle enthalten.

Bsp:



Bsp.: Der Entität (Objekt)-Typ **Schüler** hat die Attribute (Merkmale) **Name, Vorname, Alter** und **Klasse**.



Dann entspricht diesem Typ eine Tabelle **Schüler** mit den Attributen (Merkmale) **Name, Vorname, Alter** und **Klasse**. Einzelne Entitäten (Objekte) dieses Typs werden durch Zeilen der Tabelle beschrieben.

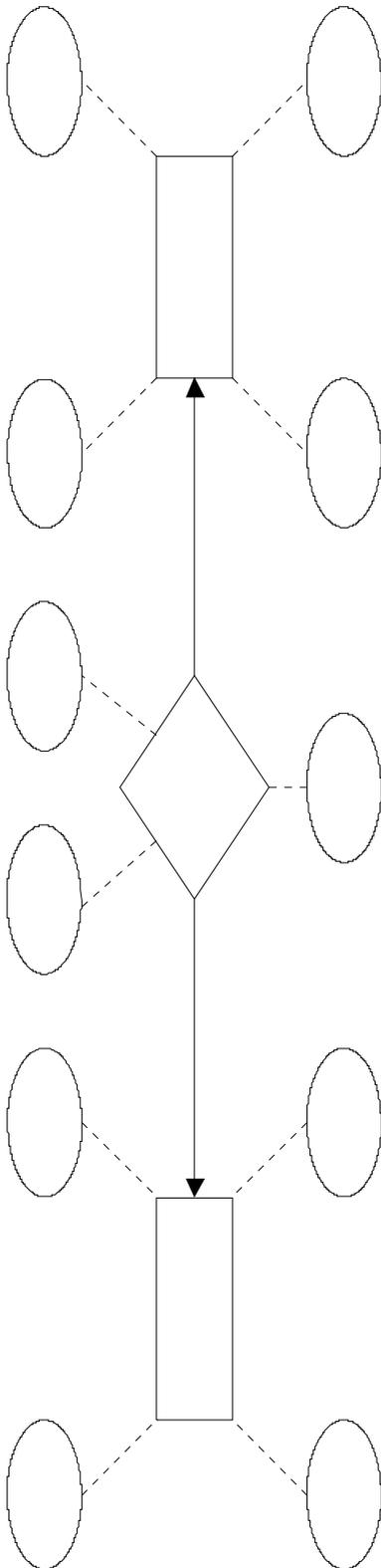
Relation **Schüler**

Name	Vorname	Alter	Klasse
Huber	Petra	16	9b
Schmidt	Xaver	14	8c
...

Zusammenfassung:

- Entitäts-Typ: Schulklasse
- Entität: Schüler
- Attribute: Nachname, Vorname, Alter, ...
- Attributswerte: Huber, Erich, 16, ...

2. Konkrete Anwendung



2.1 Zusammenfassung von Datensätzen zu Tabellen

In einer Tabelle befinden sich Objekte, die unter dem Gesichtspunkt ihrer gespeicherten Merkmale zusammengehören.

- Gemeinsame Merkmale eines Objektes werden zu einem Datensatz zusammengefaßt
- Gleich strukturierte Datensätze bilden eine Tabelle

2.2 Wahl eines Datenverwaltungsprogrammes

Für die Definition der Datensatzstruktur und die Verwaltung der Daten wird ein relationales Datenbanksystem verwendet. Dieses besitzt folgende Eigenschaften:

- Die **Zeilen** der Tabelle stellen je einen **Datensatz** dar. Die **Spalten** der Tabelle enthalten die einzelnen **Merkmale**.
- Eine **Selektion** von Datensätzen erfolgt durch die **Auswahl von Zeilen** der Tabelle, für die eine bestimmte **Selektionsbedingung** erfüllt ist.
- Eine **Projektion** der Tabelle erfolgt durch eine **Auswahl von Spalten** mit den gewünschten **Merkmalen**.
- Durch (für den Benutzer unsichtbare) Zeiger lassen sich **Feldinhalte** aus **mehreren** Tabellen **zusammenführen**.

2.3 Normalisieren der Datenbasis

Die Art und Weise, wie die gewünschten Merkmale in verschiedene Tabellen aufgeteilt und zu Datensätzen zusammengefaßt werden, entscheidet über die Bearbeitungsmöglichkeiten der Daten.

Allgemein gilt:

- Es ist nicht möglich, ein einzelnes Datenfeld zu löschen, es muß immer der gesamte Datensatz gelöscht werden \Rightarrow Verlust von Information.
 - Das Schlüsselfeld kennzeichnet immer den gesamten Datensatz und nicht ein einzelnes Datenfeld
 - Es können immer nur ganze Datensätze in eine Tabelle eingefügt werden. Ein neuer Datensatz ist immer ein neues Objekt in der Tabelle. Mit einem neuen Datensatz kann nicht bei bereits erfaßten Objekten ein weiteres Merkmal hinzugefügt werden.
-

Beim Betrieb einer Datenbank kann es durch falsche Strukturen der Tabellen zu Komplikationen kommen. Um dies zu verhindern verwendet man die Normalisierungsregeln.

2.3.1 Erste Normalform

SNR	Nachname	Vorname	Klasse	Datum	BNR1	Titel1	Autor1	Preis1	BNR2	Titel2	Autor2
12	Huber	Max	10a	23.03.97	321	Das ..	Go..	14,80	657	Li..	D..
34	Maier	Steffi	10c	08.04.97	845	Heu..	Sch..	12,30	734	Ko..	A..
...

Mehrfachmerkmale

Problem: Ein Schüler leiht gleichzeitig mehrere Bücher aus.

Lösung: Die Merkmale BNR, Titel, Autor und Preis müssen mehrmals vorhanden sein. Solche Merkmale werden als **Mehrfachmerkmale** bezeichnet.

- Die Schülernummer bildet zu den Merkmalen **Titel**, **Autor** und **Preis** eine Relation (Abhängigkeit) **Schüler - Buch**, eine **1 : n** Beziehung, d. h. ein Schüler kann mehrere Bücher ausleihen.
- Wird die Relation umgedreht, so hat man die Relation **Buch - Schüler**, eine **1 : 1** Beziehung, d. h. jedes Buch kann nur von einem Schüler ausgeliehen werden.
- Diese Relation liefert dann eine Tabelle ohne Mehrfachmerkmalen, d. h. ein Schüler kann beliebig viele Bücher ausleihen.
- Jede Tabelle kann beliebig viele Datensätze haben, d. h. es gibt keine Beschränkung mehr.

BNR	SNR	Nachname	Vorname	Klasse	Datum	Titel	Autor	Preis
454	462	Huber	Karin	10b	23.03.97	Dr...	Go...	14,80
215	874	Schmidt	Petra	10c	22.03.97	Ki...	Sch...	35,90
365	265	Müller	Robert	10a	07.04.97	Op...	Le...	12,90
784	265	Müller	Robert	10a	07.04.97	Zu...	Bö...	26,70
852	265	Müller	Robert	10a	07.04.97	Wer...	Le...	39,90
926	462	Huber	Karin	10b	23.03.97	Be...	Ma...	9,80

1. Normalform:

In den Datensätzen gibt es keine Mehrfachmerkmale.

2.3.2 Zweite Normalform

Anomalien beim Betrieb der Tabelle *Ausleihe* in der 1. Normalform

- **Anomalie beim Ändern von Datensätzen**

In der Tabelle wird der Schüler mit den Merkmalen **Nachname**, **Vorname** und **Klasse** bei jedem Buch das er ausleiht gespeichert. Ändert sich zum Ende des Schuljahres die Klasse, weil der Schüler vorrückt, muß in **jedem** Datensatz der **Merkmalswert** für die Klasse geändert werden.

Wird diese Änderung nicht bei jedem Datensatz konsequent durchgeführt, kann es zu **Inkonsistenz** der Daten kommen, d. h. der Informationsgehalt wird mehrdeutig und damit wertlos.

Da die Daten mehrfach gespeichert sind kommt es zu **Redundanzen** der Daten, d. h. diese Redundanzen können bei einer Änderung der Daten zu einer Anomalie führen.

- **Anomalie beim Einfügen von Datensätzen**

Nicht ausgeliehene Bücher lassen sich in der Tabelle nicht erfassen, da der Kombinationsschlüssel **BNR/SNR** unvollständig wäre. Dieses Problem wäre zu lösen, in dem man der Bücherei eine Pseudoschülernummer gibt und somit die Bücherei auf diese Nummer alle restlichen Bücher ausleiht. Dabei könnten dann allerdings nicht die Merkmale Nachname, Vorname und Klasse gesperrt werden, wodurch es möglich ist, daß eine Schüler ein Buch unter der Pseudoschülernummer ausleiht. Das würde dann wieder eine Inkonsistenz der Daten bedeuten.

- **Anomalie beim Entfernen von Datensätzen**

Wird ein Buch von einem Schüler zurückgeben, müßte das Datum gelöscht werden, da die Ausleihe beendet ist. Datensätze können aber nur als Gesamtes gelöscht werden und nicht nur einzelne Merkmalswerte, d. h. es werden alle Merkmalswerte gelöscht, auch die, die man noch braucht.

⇒ Es dürfen sich nur solche Merkmale in einem Datensatz befinden, welche so voneinander abhängig sind, daß sie gemeinsam gelöscht werden können.

⇒ Die Merkmale **Nachname**, **Vorname** und **Klasse** hängen logisch nur vom Schlüssel **SNR** ab, aber nicht vom Primärschlüssel **BNR**.

BNR	SNR	Nachname	Vorname	Klasse	Datum	Titel	Autor	Preis
-----	-----	----------	---------	--------	-------	-------	-------	-------

Diese Merkmale hängen **nicht** vom Schlüssel BNR ab.

2. Normalform:

Jedes Merkmal des Datensatzes, die Schlüsselmerkmale ausgenommen, muß vom gesamten Kombinationsschlüssel abhängen.

Um die 2. Normalform zu erreichen, werden die von der **BNR** unabhängigen Merkmale in eine eigene Tabelle *Schueler* ausgelagert. Die beiden Tabellen werden über das Schlüsselmerkmal **SNR** miteinander verbunden.

Tabelle: AUSLEIHE

BNR	SNR	Datum	Titel	Autor	Preis
454	462	23.03.97	Dr...	Go...	14,80
215	874	22.03.97	Ki...	Sch...	35,90
365	265	07.04.97	Op...	Le...	12,90
784	265	07.04.97	Zu...	Bö...	26,70
852	265	07.04.97	Wer...	Le...	39,90
926	462	23.03.97	Be...	Ma...	9,80

**Tabelle: SCHUELER**

SNR	Nachname	Vorname	Klasse
462	Huber	Karin	10b
874	Schmidt	Petra	10c
265	Müller	Robert	10a

Die beiden Tabellen bilden jetzt eine Datenbasis in der **2. Normalform**. Es gibt keine Redundanzen bei den Schülerdaten mehr, lediglich bei der **SNR** ist eine Redundanz entstanden, welche aber zur Verbindung der beiden Tabellen benötigt wird.

Die Daten in der Tabelle *Schueler* können nun unabhängig von der Tabelle *Ausleihe* bearbeitet werden und sind nicht mehr redundant gespeichert, d. h. eine Änderung ist also stets nur an einem Datensatz durchzuführen.

Bsp.: Man muß den Merkmalswert für die Klasse bei einem Schüler nur noch einmal ändern.

2.3.3 Dritte Normalform

Die Merkmale **Titel**, **Autor** und **Preis** hängen aber weiterhin nicht nur vom Primärschlüssel **BNR** ab, sondern auch vom Merkmal **Datum** und durch dieses Merkmal aber auch noch zur Tabelle *Schueler*. Man nennt so einen Sachverhalt **transitive Abhängigkeit**.

Anomalien beim Betrieb der Tabelle in der 2. Normalform

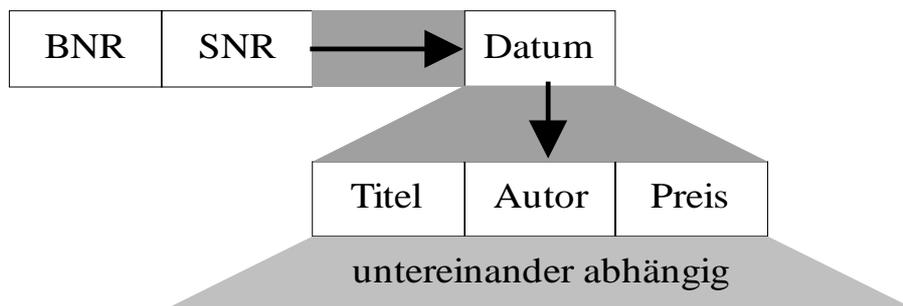
- **Anomalie beim Einfügen von Datensätzen**

Es können immer noch nicht alle Bücher der Bücherei erfaßt werden, sondern nur die ausgeliehenen Bücher.

- **Anomalie beim Entfernen von Datensätzen**

Wird ein Buch zurückgegeben, so werden neben dem Ausleihdatum auch die Daten für das Buch gelöscht.

⇒ Die Merkmale **Titel**, **Autor** und **Preis** hängen voneinander ab und bilden eine eigene Relation.



Die Tabelle *Ausleihe* hat eine transitive Abhängigkeit über das Merkmal **Datum** und die Merkmale **Titel**, **Autor** und **Preis** sind voneinander abhängig.

3. Normalform:

Mit Ausnahme der Schlüsselmerkmale sind alle Merkmale des Datensatzes voneinander unabhängig.

Es wird nun noch eine dritte Tabelle benötigt, welche die Merkmale **Titel**, **Autor** und **Preis** erfaßt. Dies sei die Tabelle *Buecher*. Die Tabellen *Ausleihe* und *Buecher* werden mit dem Schlüssel **BNR** verbunden.

Tabelle: *AUSLEIHE*

BNR	SNR	Datum
454	462	23.03.97
215	874	22.03.97
365	265	07.04.97
784	265	07.04.97
852	265	07.04.97
926	462	23.03.97

Tabelle: *BUECHER*

BNR	Titel	Autor	Preis
454	Dr...	Go...	14,80
215	Ki...	Sch...	35,90
365	Op...	Le...	12,90
784	Zu...	Bö...	26,70
852	Wer...	Le...	39,90
926	Be...	Ma...	9,80

Tabelle: *SCHUELER*

SNR	Nachname	Vorname	Klasse
462	Huber	Karin	10b
874	Schmidt	Petra	10c
265	Müller	Robert	10a
265	Müller	Robert	10a
265	Müller	Robert	10a
462	Huber	Karin	10b

Die drei Tabellen *Ausleihe*, *Buecher* und *Schueler* sind über die Schlüsselmerkmale **BNR** und **SNR** miteinander verbunden.

Wenn sich die Datenbasis in der **3. Normalform** befindet hat das folgende Vorteile:

- In der Tabelle *Buecher* können alle Bücher der Bücherei aufgenommen werden, auch wenn sie nicht ausgeliehen sind.
- Die Schüler werden in der separaten Tabelle *Schueler* verwaltet werden und die Datensätze können geändert werden, ohne das es zu Inkonsistenzen kommt.
- Die Tabelle *Ausleihe* enthält nur Daten, welche nach der Rückgabe eines Buches ohne Datenverlust gelöscht werden können.

3. Die programmtechnische Ausführung

3.1 Anlegen der Tabelle *Schueler*

Datensatzformat der dB-Datei : H:\DATEN\SCHUELER
Anzahl der Datensätze : 16
Datum der letzten Aktualisierung : 27.04.97

Feld	Feldname	Typ	Länge	Dez	Index
1	SNR	Zeichen	4		J
2	NACHNAME	Zeichen	20		N
3	VORNAME	Zeichen	20		N
4	KLASSE	Zeichen	4		N
*	Gesamt	*	49		

SNR	NACHNAME	VORNAME	KLASSE
1011	Huber	Robert	10A
1100	Meier	Ute	10D
1014	Schmidt	Josef	10B
1210	Kaufmann	Harald	10C
1001	Haller	Ferdinand	10D
1007	Amann	Klaus	10A
1023	Renauer	Johanna	10D
1065	Frey	Martina	10B
1083	Thaler	Xaver	10C
1109	Turba	Alexander	10A
1134	Bauer	Sabine	10B
1287	Schmidinger	Karl	10D
1089	Bauer	Max	10B
1010	Schmied	Hermann	10A
1025	Müller	Alexandra	10C
1400	Steiner	Jasmin	10D

3.2 Anlegen der Tabelle *Buecher*

Datensatzformat der dB-Datei : H:\DATEN\BUECHER
 Anzahl der Datensätze : 17
 Datum der letzten Aktualisierung : 27.04.97

Feld	Feldname	Typ	Länge	Dez	Index
1	BNR	Zeichen	3		J
2	BUCHTITEL	Zeichen	30		N
3	BUCHAUTOR	Zeichen	20		N
4	BUCHPREIS	Numerisch	6	2	N
*	Gesamt	*	60		

BNR	BUCHTITEL	BUCHAUTOR	BUCHPREIS
011	Faust	Goethe	22.80
043	Die Räuber	Schiller	19.80
098	Nathan der Weisse	Lessing	43.90
029	Die Physiker	Dürrenmatt	10.80
022	Schulreport	Struck	14.90
432	Gelächter von außen	Graf	18.90
864	Ende einer Dienstfahrt	Böll	9.80
584	Die Fliegen	Satre	5.80
874	Informatik	Goos	48.00
254	Physik der Materie	Stierstadt	98.00
257	Der kleine Erziehungsberater	Hacke	19.80
120	Borland Delphi 2	Doberenz	98.00
130	Der Ehrliche ist der Dumme	Wickert	38.00
547	Das Lexikon der Zahlen	Wells	14.80
999	Einführung in die Elektronik	Pütz	38.90
736	Computer Englisch	Schulze	22.80
455	English for Runaways	Heygen	7.80

3.3 Anlegen der Tabelle *Ausleihe*

Datensatzformat der dB-Datei : H:\DATEN\AUSLEIHE
Anzahl der Datensätze : 5
Datum der letzten Aktualisierung : 27.04.97

Feld	Feldname	Typ	Länge	Dez	Index
1	BNR	Zeichen	3		N
2	SNR	Zeichen	4		N
3	DATUM	Datum	8		N
*	Gesamt	*	16		

BNR	SNR	DATUM
011	1065	23.04.97
098	1134	13.04.97
022	1007	12.02.97
432	1109	07.04.97
455	1109	07.04.97

3.4 Arbeiten mit Tabellen

Tabelle Ausleihe		
BNR	SNR	DATUM
011	1065	23.04.97
098	1134	13.04.97
022	1007	12.02.97
432	1109	07.04.97
455	1109	07.04.97

Tabelle Buecher			
BNR	BUCHTITEL	BUCHAUTOR	BUCHPREIS
011	Faust	Goethe	22.80
043	Die Räuber	Schiller	19.80
098	Nathan der Weisse	Lessing	43.90
029	Die Physiker	Dürrenmatt	10.80
022	Schulreport	Struck	14.90
432	Gelächter von außen	Graf	18.90
864	Ende einer Dienstfahrt	Böll	9.80
736	Computer Englisch	Schulze	22.80
455	English for Runaways	Heygen	7.80

Tabelle Schueler			
SNR	NACHNAME	VORNAME	KLASSE
1011	Huber	Robert	10A
1100	Meier	Ute	10D
1014	Schmidt	Josef	10B
1210	Kaufmann	Harald	10C
1001	Haller	Ferdinand	10D
1007	Amann	Klaus	10A
1023	Renauer	Johanna	10D
1065	Frey	Martina	10B
1083	Thaler	Xaver	10C
1109	Turba	Alexander	10A
1134	Bauer	Sabine	10B
1287	Schmidinger	Karl	10D