

```

17 string sInput;
18 int iLength, iN;
19 double dblTemp;
20 bool again = true;
21
22 while (again) {
23     iN = -1;
24     again = false;
25     getline(cin, sInput);
26     system("cls");
27     stringstream(sInput) >> dblTemp;
28     iLength = sInput.length();
29     if (iLength < 4) {
30         again = true;
31         continue;
32     }
33     if (sInput[iLength - 3] != '.') {

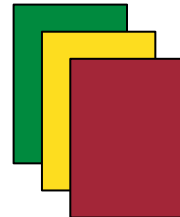
```

Programmierung von Systemen – 17 – DB-Entwurfstheorie

Matthias Tichy & Stefan Götz | SoSe 2020

Zurück zum Anfang

- Warum muss das alles so kompliziert sein?
- Warum so viele Tabellen?
- Warum diese Schlüssel?
- Warum überhaupt?
- Wieso? Weshalb? Warum?



Einfacher

TeileNr	Farben	KalkKosten	Preis	LiefName	LiefStadt	Bewertung
K28	neutral / rot / blau	2,60 / 2,60 / 2,90	2,60 / 2,70 / 2,70	Kalbe & Söhne	Augsburg	2
K28	neutral / rot / blau	2,60 / 2,60 / 2,90	2,40 / 2,50 / 2,50	Meier KG	Ulm	1
P18	rot / blau	2,30 / 2,50	2,20 / 2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

- Schlüssel?
 - TeileNr, Farben, LiefName?

Atomar

<u>TeileNr</u>	<u>Farbe</u>	KalkKosten	Preis	<u>LiefName</u>	LiefStadt	Bewertung
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

Schlüssel: TeileNr, Farbe, LiefName

Wirklich einfacher?

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

neues Teil, aber noch kein Lieferant

Wirklich einfacher?

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1
AR	rot	--	--	--	--	--

neues Teil, aber noch kein Lieferant

Einfügeanomalie

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1
AR	rot	--	--	--	--	--



neues Teil, aber noch kein Lieferant

weitere Probleme

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

Teil PZ nicht mehr benötigt

weitere Probleme

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1

Teil PZ nicht mehr benötigt → Lieferant auch weg!

Löschanomalie

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1



Teil PZ nicht mehr benötigt → Lieferant auch weg!

noch mehr Probleme

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

Meier KG zieht um nach Blaustein

noch mehr Probleme

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Blaustein	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

Meier KG zieht um nach Blaustein → Konsistenz!

Änderungsanomalie

<u>TeileNr</u>	<u>Farbe</u>	<u>KalkKosten</u>	<u>Preis</u>	<u>LiefName</u>	<u>LiefStadt</u>	<u>Bewertung</u>
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Blaustein	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1



Meier KG zieht um nach Blaustein → Konsistenz!

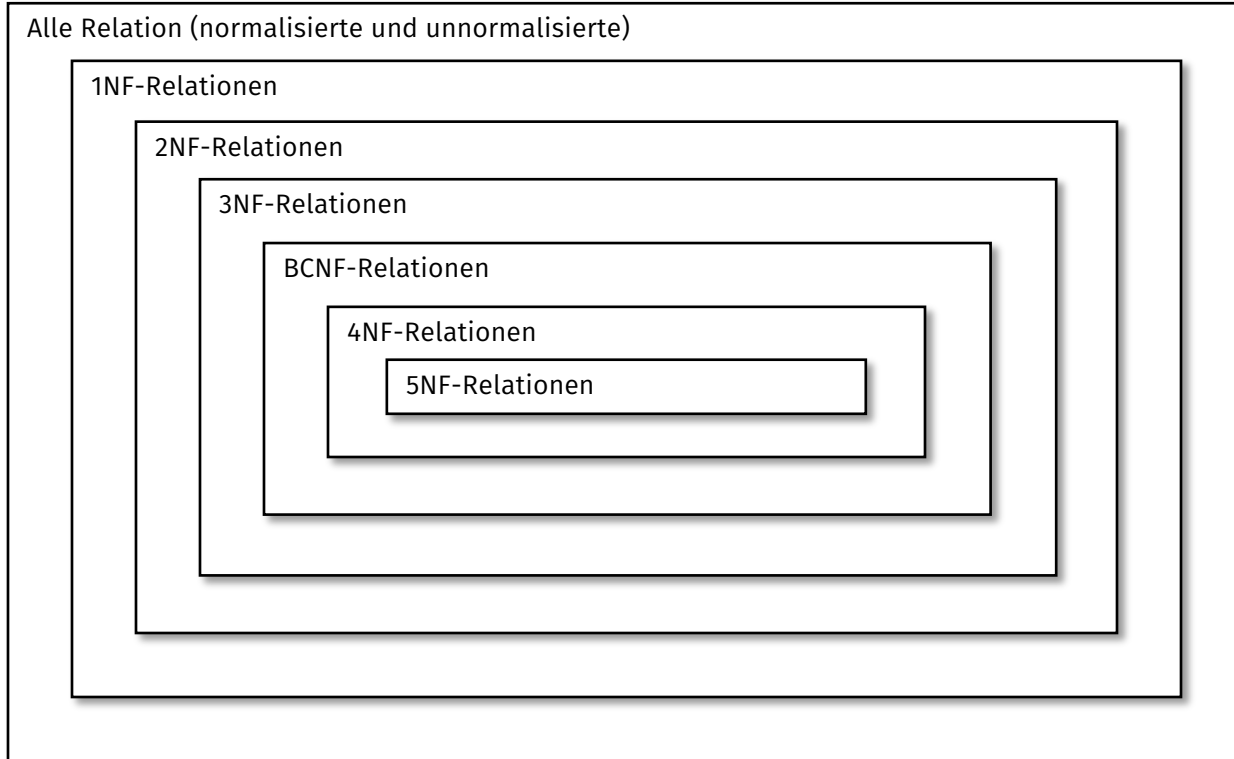
Fazit

- Darstellung der Information in DB nicht „egal“
- Aufstellung und Einhaltung von Entwurfsregeln erforderlich
- Im relationalen Kontext (aber Grundkonzepte übertragbar auf andere DB-Modelle):
 - Normalformen
 - Entwurfstheorie / Entwurfsregeln

Entwurfstheorie

- Forschungsfrage:
Welchen prinzipiellen Anforderungen müssen DB-Schemas genügen, damit Anomalien nicht auftreten?
- Von Codd (1971/72) drei Normalformen vorgeschlagen (1NF, 2NF, 3NF)
- Boyce-Codd-Normalform (BCNF) (ca. 1974) beseitigt einige Schwächen von 3NF
- viele weitere, spezielle Normalformen:
u.a. Fagin „neue“ Normalform (4NF) (1977) und „projection-join normal form“ (5NF)

Zusammenhänge



Wie geht man vor?

- Normalisierungsprozess führt zur Aufspaltung von Attributen und Tabellen
- Ausgehend von einer allumfassenden Tabelle Zerlegung nach gewissen Regeln
- Wichtig:
 - verlustfreie Zerlegung
 - abhängigkeitsbewahrende Zerlegung

Wichtige Hinweise

- Normalformen sind Regeln/Hinweise für guten DB-Entwurf und kein „Kochrezept“
- Normalisierung erfordert Wissen über die konkreten (realen) Zusammenhänge der Attribute (➔ nicht automatisierbar)
- Je weiter (2NF → 3NF → ...) normalisiert werden soll, desto mehr Hintergrundwissen über die Daten erforderlich

Wichtige Hinweise

- Normalisierung führt i.d.R. zu einer größeren Anzahl von Relationen
→ evtl. Performanz-Probleme
- Daher:
 - nicht ohne Not auf eine gute Normalisierung verzichten, aber auch nicht Normalisierung um der Normalisierung willen durchführen
 - nach Normalisierung Relationen mit gleichem Primärschlüssel wieder zusammenfassen

1. Normalform

Def.: Eine Relation ist in 1. NF genau dann, wenn alle ihre Attributwerte **atomar** sind und die Relation **keine Wiederholungsgruppen** enthält (und einen Primärschlüssel hat).

nicht in 1. NF:

TeileNr	Farben	KalkKosten	Preis	LiefName	LiefStadt	Bewertung
K28	neutral / rot / blau	2,60 / 2,60 / 2,90	2,60 / 2,70 / 2,70	Kalbe & Söhne	Augsburg	2
K28	neutral / rot / blau	2,60 / 2,60 / 2,90	2,40 / 2,50 / 2,50	Meier KG	Ulm	1

1. Normalform

Lösung:

- Aufspalten der Attributwerte in einzelne Zeilen

Beispiel für Wiederholungsgruppen:

- Telefon1, Telefon2, Telefon3

Aber:

- "Telefon, Fax, Mobil" könnte schon wieder intendiert sein

2. Normalform

Def.: Eine Relation ist in 2. NF genau dann, wenn (sie eine Relation 1. NF ist und) kein Nichtschlüsselattribut **funktional abhängig** von einer echten Teilmenge eines Schlüsselkandidaten ist.

Was bedeutet "funktional abhängig"?

Funktionale Abhängigkeit

<u>TeileNr</u>	<u>Farbe</u>	KalkKosten	Preis	<u>LiefName</u>	LiefStadt	Bewertung
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

LiefName bestimmt LiefStadt und Bewertung:
$$\text{LiefName} \xrightarrow{R} \{\text{LiefStadt}, \text{Bewertung}\}$$

Funktionale Abhängigkeit

Def.: Seien $X, Y \subseteq \text{sch}(R)$.

Eine **funktionale Abhängigkeit (FD)** in R ist erfüllt, wenn für alle Ausprägungen $\text{val}(R)$ gilt:
 $t_1, t_2 \in \text{val}(R) : t_1(X) = t_2(X) \Rightarrow t_1(Y) = t_2(Y)$.

- Sprechweise: "X bestimmt Y", "Y hängt von X ab"
- Folgerung: In jeder Relation mit Schlüssel S hängen alle Attribute von S (funktional) ab.

D.h. es gilt z.B.:

- $R(\underline{S}, A, B, \dots) \Rightarrow S \xrightarrow{R} A, S \xrightarrow{R} B, \dots$
- $R(\underline{S1}, \underline{S2}, A, B, \dots) \Rightarrow \{S1, S2\} \xrightarrow{R} A, \{S1, S2\} \xrightarrow{R} B, \dots$

Zurück zur 2NF

... kein Nichtschlüsselattribut **funktional abhängig** von einer echten Teilmenge eines Schlüsselkandidaten ist.

<u>TeileNr</u>	<u>Farbe</u>	KalkKosten	Preis	<u>LiefName</u>	LiefStadt	Bewertung
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

2. Normalform

nicht in 2. Normalform:

<u>TeileNr</u>	<u>Farbe</u>	KalkKosten	Preis	<u>LiefName</u>	LiefStadt	Bewertung
K28	neutral	2,60	2,60	Kalbe & Söhne	Augsburg	2
K28	rot	2,60	2,70	Kalbe & Söhne	Augsburg	2
K28	blau	2,90	2,70	Kalbe & Söhne	Augsburg	2
K28	neutral	2,60	2,40	Meier KG	Ulm	1
K28	rot	2,60	2,50	Meier KG	Ulm	1
K28	blau	2,90	2,50	Meier KG	Ulm	1
P18	rot	2,30	2,20	Damm Ltd.	Hamburg	1
P18	blau	2,50	2,20	Damm Ltd.	Hamburg	1
PZ	neutral	2,40	2,40	Otto GmbH	Köln	1

LiefStadt und Bewertung funktional abhängig von LiefName (echte Teilmenge von Schlüssel)

2. Normalform

■ Anmerkungen:

- Ein Nichtschlüssel-Attribut A in 2. NF heißt **voll funktional abhängig** von Schlüssel S.
- S wird auch als **Determinante** von A bezeichnet
- Bei 2NF-Relation alle Nichtschlüsselattribute voll funktional abhängig von (allen) Schlüsseln.
- Relationen mit nur einem Schlüsselattribut stets in 2NF
- Relationen, die nur aus Schlüsselattributen bestehen, ebenfalls stets in 2NF
- Relationen, die nicht in 2NF sind, führen in der Regel zu Anomalien

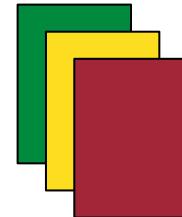
2. Normalform

Lösung:

- Aufteilung in zwei Relationen

<u>LiefNr</u>	LiefName	LiefStadt	Bewertung
506	Kalbe & Söhne	Augsburg	2
527	Meier KG	Ulm	1
545	Damm Ltd.	Hamburg	1
557	Otto GmbH	Köln	1

<u>TeileNr</u>	<u>Farbe</u>	KalkKosten	Preis	<u>LiefNr</u>
K28	neutral	2,60	2,60	506
K28	rot	2,60	2,70	506
K28	blau	2,90	2,70	506
K28	neutral	2,60	2,40	527
K28	rot	2,60	2,50	527
K28	blau	2,90	2,50	527
P18	rot	2,30	2,20	545
P18	blau	2,50	2,20	545
PZ	neutral	2,40	2,40	557



Nummern sind toll

In der Lieferanten-Tabelle wurde LiefNr als künstlicher Schlüssel eingeführt. Warum?

<u>LiefNr</u>	LiefName	LiefStadt	Bewertung
506	Kalbe & Söhne	Augsburg	2
527	Meier KG	Ulm	1
545	Damm Ltd.	Hamburg	1
557	Otto GmbH	Köln	1

Nummern sind toll

Namen sind Schall und Rauch

aus Performanzgründen (Zahlen sind besser zu verarbeiten als Strings)



künstlicher Schlüssel

Bemerkungen zu künstlichem Schlüsselattribut

- Bei manchen Datenbanken tatsächlich Performanzgewinn durch Zahlen
- Künstlicher Schlüssel zu empfehlen, wenn andere Schlüsselkandidaten nicht voll unter Kontrolle
Beispiele:
 - Lieferantennamen beliebig, daher evtl. auch zwei Lieferanten mit gleichem Namen in gleichem Ort.
→ keine Kontrolle
 - TeileNr kann willkürlich vergeben werden, daher kann Benutzer sicherstellen, dass keine zwei unterschiedlichen Teile gleich heißen. → Kontrolle

künstlicher Schlüssel

Bemerkungen zu künstlichem Schlüsselattribut

- In der Praxis sehr häufig angewendet
- Relation automatisch in 2NF
- Typischerweise hochzählende Nummer
- Problem bei Umsetzung:
 - aktuell höchster Wert bei Einfügen neuer Tupel evtl. nicht bekannt
 - ➔ 1. Auslesen, 2. Einfügen, 3. Hoffen, dass keiner schneller war
- Besser: **Autoinkrement**-Spalte

3. Normalform

Def.:

Relation R ist in 3NF genau dann, wenn kein ***Nichtschlüsselattribut*** **transitiv** vom Schlüssel **abhängt**.

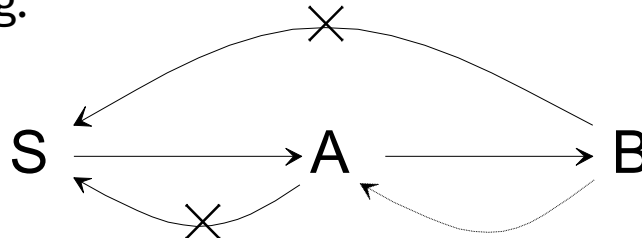
Was bedeutet "transitiv abhängig"?

Transitive Abhängigkeit

Def.: Ein Attribut B in R ist **transitiv abhängig** vom Schlüssel S, wenn

1. $B \notin S$
2. Eine Menge A von Attributen in R existiert ($B \notin A$), so dass gilt:
 $S \rightarrow A, A \rightarrow B, A \not\rightarrow S$

Visualisierung:



Transitive Abhängigkeit

Bemerkungen:

- Wegen $A \nrightarrow S$ kann A nicht Schlüssel in R sein, d.h. Attributwerte von A können in R mehrfach auftreten
- Da aber $A \rightarrow B$, redundante Einträge von B-Werten (\rightarrow Anomalien)
- Beispiel für Tabelle **nicht** in 3NF:

<u>Wettbewerb</u>	<u>Jahr</u>	Gewinner	Geburtstag
ACM	2013	Bert	26.07.
RoboCup	2015	Krümelmonster	02.11.
CaroloCup	2012	Ernie	28.01.
MatheOlympiade	2010	Bert	26.07.

3. Normalform

Def.:

Relation R ist in 3NF genau dann, wenn kein ***Nichtschlüsselattribut*** **transitiv** vom Schlüssel **abhängt**.

Alternative Darstellung:

Def.:

Relation R ist in 3NF genau dann, wenn für alle $Y \subset \text{sch}(R)$, $A \notin Y$, A *Nichtschlüsselattribut* in R gilt:

$Y \rightarrow A \Rightarrow Y$ ist Schlüssel oder enthält einen Schlüssel von R

Zerlegungen

- Erhaltung einer Normalform durch Zerlegung
- Aber: Zerlegung sollte nicht willkürlich geschehen
- gewünschte Eigenschaften:
 - **verlustfrei**
mindestens die gleiche Tupelmengen wie vor der Zerlegung
 - **informationserhaltend**
Ursprüngliche Relation sollte rekonstruiert werden können
- Gegenbeispiel:
Teile(TeileNr, Farbe, KalkKosten, Bestand, MinBestand) zerlegt in
 - *Teile1*(TeileNr, Farbe) und
 - *Teile2*(KalkKosten, Bestand, MinBestand)macht wenig Sinn!

Relationen zusammenfassen

- Relationen mit identischem Primärschlüssel können evtl. zusammengefasst werden
- keine entitätsfremden Informationen zusammenfassen!