

**Datenbank-Grundlagen**

**SS 2005**

**1b. Formalisierung des relationalen Datenmodells**  
**Prof. Dr. Stefan Böttcher**  
**Universität Paderborn**

**mit Material von**  
**Prof. Dr. Gregor Engels**

**Agenda:**  
 Begriffe Attribute und Domänen    Relationenschema und Relationen    Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen    SS2005    Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 1

**Das Relationenmodell (auch relationales Modell)**

Von Codd im Jahre 1970 eingeführt: Veranschaulichung eines Relationenschemas und einer Relation als Tabelle

□ Attributwerte eines Tupels im Standardrelationenmodell sind einfache vordefinierte Datenwerte (Integer, String, Datum, ...) → erste Normalform  
 □ im „Non-First Normal Form“-Modell = NF<sup>2</sup>-Modell kann ein Attributwert wieder eine Tabelle sein

Begriffe    Attribute und Domänen    Relationenschema und Relationen    Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen    SS2005    Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 2

**Zwei Relationen zur Darstellung von Personen**

Personen

PANr	Vorname	Nachname	PLZ	Ort	Straße	HNr	Geb.datum
4711	Andreas	Heuer	18209	DBR	BHS	15	31.10.1956
5588	Gunter	Saake	39106	MD	STS	55	05.10.1960
6834	Michael	Korn	39104	MD	BS	41	24.09.1974
7754	Andreas	Möller	18209	DBR	RS	31	25.02.1976
8832	Tamara	Jagellovek	38106	BS	GS	12	11.11.1973
9912	Antje	Heilhof	18059	HRO	AES	21	04.04.1970
9999	Christa	Loesser	69121	HD	TS	38	10.05.1969

Pers\_Telefon

PANr	Telefon
4711	038203-12230
4711	0381-498-3401
4711	0381-498-3427
5588	0391-345677
5588	0391-5592-3800
9999	06221-400177

Begriffe    Attribute und Domänen    Relationenschema und Relationen    Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen    SS2005    Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 3

**Begriffe des Relationenmodells**

Begriff	Informale Bedeutung
Attribut	Bezeichnung der Spalte einer Tabelle
Wertebereich	Mögliche Werte eines Attributs (auch Domäne)
Attributwert	Element eines Wertebereichs
Relationenschema	Menge von Attributen
Relation	Menge von Zeilen einer Tabelle
Tupel	Zeile einer Tabelle
Datenbankschema	Menge von Relationenschemata
Datenbank	Menge von Relationen (Basisrelationen)

Begriffe    Attribute und Domänen    Relationenschema und Relationen    Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen    SS2005    Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 4

**Begriffe des Relationenmodells II**

Begriff	Informale Bedeutung
Schlüssel	minimale Menge von Attributen, deren Werte ein Tupel einer Tabelle eindeutig identifizieren
Primärschlüssel	ein beim Datenbankentwurf ausgezeichnete Schlüssel
Fremdschlüssel	Attributmengende, die in einer anderen Relation Schlüssel ist
Fremdschlüsselbedingung	alle Attributwerte des Fremdschlüssels tauchen in der anderen Relation als Werte des Schlüssels auf

Begriffe    Attribute und Domänen    Relationenschema und Relationen    Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen    SS2005    Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 5

**Attribute und Domänen**

$U$  (Universum):  
 nichtleere endliche Menge aller Attribute eines Datenbankschemas

- $A \in U$  ein Attribut
- $D = \{D_1, \dots, D_m\}$  Menge nichtleerer endlicher Mengen, die Wertebereiche (domains) der Attribute
- $dom: U \rightarrow D$  totale Funktion, die jedem Attribut  $A \in U$  seine Domäne  $dom(A)$  zuordnet
- $\omega \in dom(A)$  Attributwert für  $A$

Begriffe    Attribute und Domänen    Relationenschema und Relationen    Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen    SS2005    Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 6

### Beispiele zu Attributen und Domänen

- $U = \{ PANr, Vorname, Nachname, \dots \}$
- $D = \{ \text{string}, \text{integer}, \dots \}$
- $dom(PANr) = \text{string}$
- $dom(Vorname) = \text{string}$
- $dom(HNr) = \text{integer}$
- $Gunter \in dom(Vorname)$

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 7

### Relationenschemata und Relationen

- $R \subseteq U$  ist ein Relationenschema
- Relation  $r$  über  $R = \{A_1, \dots, A_n\}$  (kurz:  $r(R)$ ) ist endl. Menge von Abbildungen  $t: R \rightarrow \bigcup_{i=1}^n D_i$ 
  - $t \in r$  ist ein Tupel der Relation
  - für  $A \in R$  bezeichnet  $t(A) \in dom(A)$  den Attributwert von  $t$  an der Stelle  $A$
- $REL(R) := \{r \mid r(R)\}$  Menge aller möglichen Relationen über Relationenschema  $R$

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 8

### Beispiele zu Schemata und Relationen

- $Pers\_Telefon = \{PANr, Telefon\} \subseteq U$
- $r(Pers\_Telefon)$ :
  - $t1: \{PANr, Telefon\} \rightarrow \text{string} \cup \text{integer} \cup \dots$
  - $t1(PANr) = 4711 \in dom(PANr) = \text{integer}$
  - $t1(Telefon) = 038203-12230 \in dom(Telefon) = \text{string}$
  - $t2: \{PANr, Telefon\} \rightarrow \text{string} \cup \text{integer} \cup \dots$
  - $t2(PANr) = 4711$
  - $t2(Telefon) = 0381-498-3401$

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 9

### Vergleich mit klassischer Definition der Relation als Teilmenge des kartesischen Produktes

- Relation als Teilmenge des kartesischen Produktes
- $r_1 \subseteq dom(PANr) \times dom(Vorname) \times dom(Nachname)$
  - $r_2 \subseteq dom(PANr) \times dom(Nachname) \times dom(Vorname)$
- $r_1$  und  $r_2$  sind verschiedene Relationen bei Definition mittels kartesischem Produkt, aber gleich bei der Definition als Menge von Abbildungen!

- Unterschied in der Identifizierung von Attributwerten
- durch Position (Tupel als Element des Kart. Produkt, siehe oben)
  - durch Namen (Tupel als Abbildung, siehe Folie 13/14)?

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 10

### Vergleich mit klassischer Definition der Relation als Teilmenge des kartesischen Produktes

- $r_1 \subseteq dom(PANr) \times dom(Vorname) \times dom(Nachname)$ 

PANr	Vorname	Nachname
4711	Andreas	Heuer
5588	Gunter	Saake
0007	Andy	Schür

 $t1 = (4711, 'Andreas', 'Heuer')$   
 $t2 = (5588, 'Gunter', 'Saake')$   
 $t3 = (0007, 'Andy', 'Schür')$
  - $r_2 \subseteq dom(PANr) \times dom(Nachname) \times dom(Vorname)$ 

PANr	Nachname	Vorname
4711	Heuer	Andreas
5588	Saake	Gunter
0007	Schür	Andy

 $t1 = (4711, 'Heuer', 'Andreas')$   
 $t2 = (5588, 'Saake', 'Gunter')$   
 $t3 = (0007, 'Schür', 'Andy')$
- als kartesisches Produkt definierte Relationen  $r_1$  und  $r_2$  sind verschieden

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 11

### Vergleich mit klassischer Definition der Relation als Teilmenge des kartesischen Produktes

- | PANr | Vorname | Nachname |
|------|---------|----------|
| 4711 | Andreas | Heuer    |
| 5588 | Gunter  | Saake    |
| 6834 | Michael | Korn     |
- | PANr | Nachname | Vorname |
|------|----------|---------|
| 4711 | Heuer    | Andreas |
| 5588 | Saake    | Gunter  |
| 6834 | Korn     | Michael |
- Relationen  $r_1$  und  $r_2$  bestehen aus Tupeln (Abbildungen)  
 $t_1, t_2, t_3: \{PANr, Vorname, Nachname\} \rightarrow \text{string} \cup \text{integer}, \dots$  mit
- $t_1(PANr) = 4711, t_1(Vorname) = 'Andreas', t_1(Nachname) = 'Heuer'$   
 $t_2(PANr) = 5588, t_2(Vorname) = 'Gunter', t_2(Nachname) = 'Saake'$   
 $t_3(PANr) = 6834, t_3(Vorname) = 'Michael', t_3(Nachname) = 'Korn'$

→ als Menge von Abbildungen definierte Relationen  $r_1$  und  $r_2$  sind gleich

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 12

### Datenbankschema und Datenbank

- Menge von Relationenschemata  
 $S := \{R_1, \dots, R_p\}$  ist ein Datenbankschema
- Datenbank über  $S$  ist Menge von Relationen  $d := \{r_1, \dots, r_p\}$ ,  
 wobei  $r_i(R_i)$
- Datenbank  $d$  über  $S$  wird notiert als  $d(S)$
- Relation  $r \in d$  wird Basisrelation genannt  
 (im Gegensatz zu abgeleiteten Relationen)

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 13

### Beispiele: Datenbankschema und Datenbank

- $S = \{Personen, Pers\_Telefon\}$
- $d(S) = \{r(Personen), r(Pers\_Telefon)\}$

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 14

### Lokale Integritätsbedingungen (1)

Attributmengemenge  $K := \{B_1, \dots, B_k\} \subseteq R$  (Relationenschema) heißt  
**identifizierende Attributmengemenge** für  $r(R)$ , falls\*:

$$\forall t_1, t_2 \in r : t_1 \neq t_2 \Rightarrow \exists B \in K : t_1(B) \neq t_2(B)$$

**Schlüssel:** eine minimal identifizierende Attributmengemenge  
**Primärschlüssel:** ein ausgezeichnete Schlüssel

\* für  $B \subseteq R$  bezeichnet  $t(B)$  die Einschränkung auf die  $B$ -Werte von  $t$

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 15

### Lokale Integritätsbedingungen (2)

**Lokale Integritätsbedingung:** eine Abbildung  
 $b : \{r \mid r(R)\} \rightarrow \{true, false\}$  von der Menge aller Relationen über  $R$  auf die Wahrheitswerte.

**Bemerkung:** Schlüssel ist Spezialfall einer lokalen Integritätsbedingung

**Lokale Integritätsbedingungen:** Menge  $B$  aller lokalen Integritätsbedingungsabbildungen  $b : \{r \mid r(R)\} \rightarrow \{true, false\}$

**Lokal erweitertes Datenbankschema:** ein DB-Schema mit lokalen Integritätsbedingungen  $S = \{(R_1, B_1), \dots, (R_p, B_p)\}$

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 16

### Globale Integritätsbedingungen

Eine **Fremdschlüsselbedingung** für eine Relation  $r_1(R_1)$  ist ein Ausdruck  $R_1(X) \rightarrow R_2(Y)$  mit  $X \subseteq R_1$  und  $Y \subseteq R_2$

- $X$  heißt dann **Fremdschlüssel** von  $R_1$  auf  $Y$  in  $R_2$ .

Eine Datenbank  $d$  genügt der Fremdschlüsselbedingung  $R_1(X) \rightarrow R_2(Y)$  falls

- in  $d$  existiert Relation  $r_2(R_2)$  mit Primärschlüssel  $Y$
- es gilt  $\{t(X) \mid t \in r_1\} \subseteq \{t(Y) \mid t \in r_2\}$

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 17

### Beispiele für Schlüssel und Fremdschlüssel

**Schlüssel:**

- $\{Vorname, Nachname\}$  und  $\{PANr\}$  sind u.a. Schlüssel für Relation *Personen*
- nur  $\{PANr, Telefon\}$  ist Schlüssel für Relation *Pers\_Telefon*

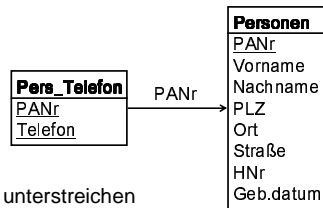
**Fremdschlüssel:**

- $PANr$  von *Pers\_Telefon* ist ein Fremdschlüssel auf  $PANr$  von *Personen*
- Fremdschlüsselbedingung  $Pers\_Telefon(PANr) \rightarrow Personen(PANr)$   
 $\{t(PANr) \mid t \in r(Pers\_Telefon)\} \subseteq \{t(PANr) \mid t \in r(Personen)\}$   
 $\{4711, 5588, 9999\} \subseteq \{4711, 5588, 6834, 7754, \dots\}$
- $PANr$  von *Personen* ist kein Fremdschlüssel auf  $PANr$  von *Pers\_Telefon*

Begriffe Attribute und Domänen Relationenschema und Relationen Datenbankschema, Datenbanken und Integritätsbedingungen  
 Datenbank Grundlagen SS2005 Kapitel 1/Teil b - Formalisierung des relationalen Datenmodells / 18

Datenbank-Grundlagen SS 2005, Kapitel 1 -  
Teil b: Formalisierung des Relationalen  
Datenmodells

Graphische Notation für Primär- und Fremdschlüssel



- Primärschlüssel unterstreichen  
Im Beispiel: Primärschlüssel {*PANr*} und {*PANr*, *Telefon*}
- Fremdschlüssel durch markierte Pfeile darstellen;  
verweisen auf Primärschlüssel des Zielschemas  
Im Beispiel: *Pers\_Telefon*(*PANr*) → *Personen*(*PANr*)