



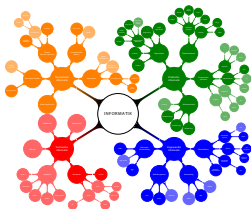
UNIVERSITÄT ZU LÜBECK  
INSTITUT FÜR  
THEORETISCHE INFORMATIK

# Kapitel 37

## Relationenalgebren

Wer ist Dollys Vater?

Vorlesung Einführung in die Informatik 2 vom 27. Mai 2014 von Till Tantau



## Lernziele von Kapitel 37

1. Konzept der Relation-Algebra kennen
2. Anfragen mittels Operationen formulieren können
3. Joins verstehen und anwenden können

# Gliederung von Kapitel 37

## 37.1 Einführung

37.1.1 Alles ist eine Tabelle

37.1.2 Relationen-Algebren

## 37.2 Operationen auf Relationen

37.2.1 Vereinigung und Schnitt

37.2.2 Selektion

37.2.3 Projektion

37.2.4 Kreuzprodukt

37.2.5 Join

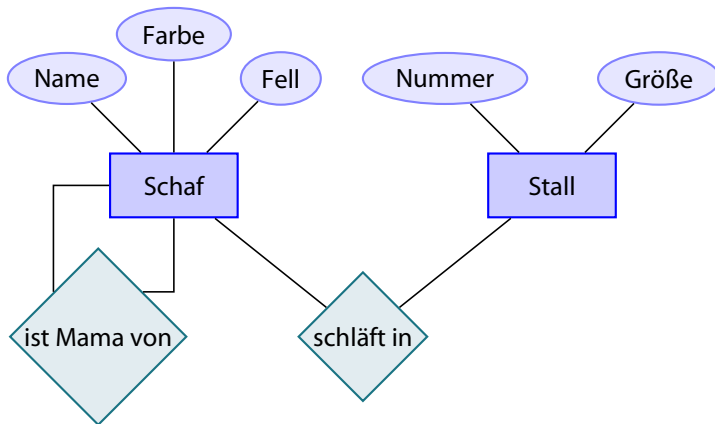
# Wiederholung: Ein Entity-Relationship-Diagramm.

## 37.1 Einführung

- Alles ist eine Tabelle
- Relationen-Algebren

## 37.2 Operationen auf Relationen

- Vereinigung und Schnitt
- Selektion
- Projektion
- Kreuzprodukt
- Join



## 37.1 Einführung

- Alles ist eine Tabelle  
Relationen-Algebren

## 37.2 Operationen auf Relationen

- Vereinigung und Schnitt
- Selektion
- Projektion
- Kreuzprodukt
- Join

- Ein E/R-Diagramm beschreibt die *Struktur* von Daten.
- Hinter jedem Entitätstyp (»Schaf« und »Stall«) steht eine Tabelle:
  - Die Spalten der Tabelle sind die Attribute dieses Typs.
  - Die Zeilen der Tabelle sind die Entitäten dieses Typs.
  - Eine Spalte bildet den primären Schlüssel.
- Hinter jedem Relationstyp (»ist Mama von«, »schläft in«) steht ebenfalls eine Tabelle:
  - Die Spalten der Tabelle sind die Primärattribute der Entitäten, die in Beziehung gesetzt werden.
  - Die Zeilen der Tabelle repräsentieren Beziehungen zwischen den in der Zeile angegebenen Entitäten.

### 37.1 Einführung

Alles ist eine Tabelle

► Relationen-Algebren

### 37.2 Operationen auf Relationen

Vereinigung und Schnitt

Selektion

Projektion

Kreuzprodukt

Join

- Die Tabellen für Entitäten und für Relationships werden zwar unterschiedlich interpretiert, *mathematisch* und *speichertechnisch* gibt es aber *keinen Unterschied*.
- Man konzentriert sich deshalb darauf, *Tabellen besonders effizient* zu verwalten.
- Statt *Tabellen* spricht man mathematisch von *Relationen*. In einer *relationalen Datenbank* werden (wenig überraschend) genau solche *Relationen* verwaltet.

Merke: Die Datenbank »weiß nichts« von dem E/R-Modell; aus ihrer Sicht wird eine Menge gleichberechtigter Tabellen verwaltet.

# Die Relationen-Algebra.

Was waren nochmal Relationen?

## Definition

Seien  $A_1$  bis  $A_n$  Mengen. Eine *Relation* auf diesen Mengen ist eine Teilmenge von  $A_1 \times \dots \times A_n$ .

Bemerkungen:

- ▶ Das *Kreuzprodukt*  $A_1 \times \dots \times A_n$  enthält gerade alle Tupel, deren erste Komponente aus  $A_1$  stammt, deren zweite Komponente aus  $A_2$  und so weiter.
- ▶ Eine *Teilmenge* dieses Kreuzprodukts ist also eine Menge von Tupeln, wobei jedes jeweils ein Element aus  $A_1$  in Beziehung setzt mit einem Element aus  $A_2$  und gleichzeitig einem aus  $A_3$  und so weiter.
- ▶ Ist beispielsweise  $A_1$  die Menge aller Namen,  $A_2$  die Menge aller Farben und  $A_3$  die Menge aller Fellarten, so ist eine Relation auf  $A_1$  bis  $A_3$  gerade eine Tabelle für den Entitätstyp Schaf.
- ▶ Es gibt zwei Unterschiede zwischen Tabellen und Relationen: Eine Relation kann jedes Tupel nur einmal enthalten und die Reihenfolge der Tupel ist egal.

### 37.1 Einführung

Alles ist eine Tabelle

▶ Relationen-Algebren

### 37.2 Operationen auf Relationen

Vereinigung und Schnitt

Selektion

Projektion

Kreuzprodukt

Join

### 37.1 Einführung

Alles ist eine Tabelle

► Relationen-Algebren

### 37.2 Operationen auf Relationen

Vereinigung und Schnitt

Selektion

Projektion

Kreuzprodukt

Join

- In einer Datenbank werden viele Relationen gespeichert, eine pro Entitätstyp und Relationshipstyp.
- Solche Relationen lassen sich nun *verknüpfen*, um neue Relationen zu bilden.  
Beispielsweise kann man aus der Relation aller Schafe durch Filterung die Relation aller schwarzen Schafe gewinnen.
- Die Klasse aller Relationen zusammen mit Operationen wie »Filterung« heißt mathematisch etwas hochtrabend *Relationen-Algebra*.



## Definition

Seien  $A, B \subseteq A_1 \times \dots \times A_n$  Relationen auf denselben Mengen. Dann heißt  $A \cup B$  die *Vereinigung* von  $A$  und  $B$  und  $A \cap B$  der *Schnitt* von  $A$  und  $B$ .

## Beispiel

- Sei  $A$  die Relation, die nur schwarze Schafe enthält (also alle Tupel in der Schafstabelle, bei denen die Farbkomponente »schwarz« lautet).
- Sei  $B$  die Relation, die nur lockige Schafe enthält.
- Dann ist  $A \cap B$  die Relation, die nur lockige, schwarze Schafe enthält.
- Dann ist  $A \cup B$  die Relation, die alle Schafe enthält, die lockig oder schwarz sind.

### 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

### 37.2 Operationen auf Relationen

- Vereinigung und Schnitt
- Selektion
- Projektion
- Kreuzprodukt
- Join

## 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

## 37.2 Operationen auf Relationen

Vereinigung und Schnitt

- Selektion
- Projektion
- Kreuzprodukt
- Join

## Definition

Seien  $A \subseteq A_1 \times \cdots \times A_n$  eine Relation und  $P$  ein Prädikat auf Tupeln (also ein Test, der für jedes Tupel entweder wahr oder falsch ist). Dann ist die Selektion  $\sigma_P(A) \subseteq A_1 \times \cdots \times A_n$  die Relation, die nur diejenigen Tupel aus  $A$  enthält, für die  $P$  gilt.

## Beispiel

- $\sigma_{\text{Farbe}=\text{schwarz}}(\text{Schaf})$  ist die Relation aller schwarzen Schafe.
- $\sigma_{\text{Größe}>50}(\text{Stall})$  ist die Relation aller Ställe, die größer als 50 sind.
- Die folgenden Relationen sind gleich:
  1.  $\sigma_{\text{Farbe}=\text{schwarz}}(\text{Schaf}) \cap \sigma_{\text{Fell}=\text{lockig}}(\text{Schaf})$
  2.  $\sigma_{\text{Farbe}=\text{schwarz} \wedge \text{Fell}=\text{lockig}}(\text{Schaf})$

37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

37.2 Operationen auf  
Relationen

Vereinigung und Schnitt

- Selektion
- Projektion
- Kreuzprodukt
- Join

## Zur Übung

Sei  $A$  die Relation aller schwarzen Schafe,  $B$  die Relation aller lockigen Schafe und  $C$  die Relation aller kranken Schafe.

Versuchen Sie, mit Hilfe von Vereinigung, Schnitt und/oder Selektion folgende Mengen zu beschreiben:

1. Die Menge aller schwarzen Schafe, die lockig und krank sind.
2. Die Menge aller gesunden lockigen Schafe.

## Definition

Seien  $A \subseteq A_1 \times \cdots \times A_n$  eine Relation und seien  $A_{i_1}$  bis  $A_{i_j}$  einige dieser Mengen. Dann entsteht die *Projektion*  $\pi_{A_{i_1}, \dots, A_{i_j}}(A) \subseteq A_{i_1} \times \cdots \times A_{i_j}$ , indem man »nur die Spalten«  $A_{i_1}$  bis  $A_{i_j}$  betrachtet.

## Beispiel

► Schaf =

Name	Farbe	Fell
Dolly	weiß	lockig
Flauschi	schwarz	wuschelig
Peter	schwarz	wuschelig
Flauschi	beige	glatt

►  $\pi_{\text{Farbe}, \text{Fell}}(\text{Schaf}) =$

Farbe	Fell
weiß	lockig
schwarz	wuschelig
beige	glatt

### 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

### 37.2 Operationen auf Relationen

Vereinigung und Schnitt  
Selektion  
► Projektion  
Kreuzprodukt  
Join

37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

37.2 Operationen auf  
Relationen

Vereinigung und Schnitt  
Selektion  
Projektion

► Kreuzprodukt  
Join

## Definition

Seien  $A \subseteq A_1 \times \cdots \times A_n$  und  $B \subseteq B_1 \times \cdots \times B_m$  Relationen. Dann ist das *Kreuzprodukt*  $A \times B \subseteq A_1 \times \cdots \times A_n \times B_1 \times \cdots \times B_m$  die Relation, die alle Kombinationen von Elementen aus  $A$  und Elementen aus  $B$  enthält.

- Beim Kreuzprodukt paart man alle Elemente der ersten Menge mit allen Elementen der zweiten Menge.
- Die Größe des Kreuzproduktes ist gerade das Produkt der Größen der Relationen  $A$  und  $B$ .

# Beispiel für das Kreuzprodukt.

Gegeben seien folgende Relationen:

Name	Farbe	Fell
Dolly	weiß	lockig
Flauschi	schwarz	wuschelig
Peter	schwarz	wuschelig

Schaf-Name	Stall-Nummer
Dolly	1
Peter	5

Ihr Kreuzprodukt lautet:

Name	Farbe	Fell	Schaf-Name	Stall-Nr.
Dolly	weiß	lockig	Dolly	1
Dolly	weiß	lockig	Peter	5
Flauschi	schwarz	wuschelig	Dolly	1
Flauschi	schwarz	wuschelig	Peter	5
Peter	schwarz	wuschelig	Dolly	1
Peter	schwarz	wuschelig	Peter	5

## 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

## 37.2 Operationen auf Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
► Kreuzprodukt  
Join

## 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

## 37.2 Operationen auf Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

► Join

- Nehmen wir an, wir wollen wissen, in *welchen Ställen schwarze Schafe schlafen*.

*Problem:* Diese Information liefert uns *weder* die Schafrelation *noch* die *Schläft-in*-Relation.

- Wir müssen diese beiden Tabellen deshalb *vereinigen*.

*Problem:* Das Kreuzprodukt setzt immer alles mit allem in Relation.

- Wir filtern deshalb nach solchen Paaren im Kreuzprodukt, bei denen der Schafname im ersten Teil des Tupels zum Schafname im zweiten Teil des Tupels passt.

*Problem:* Dies liefert große Tupel, wir wollen aber nur die Stall-Nummer wissen.

- Dann projizieren wir auf die Stall-Nummer.

$\pi_{\text{Stall-Nummer}}(\sigma_{\text{Name=Schaf-Name}}(\sigma_{\text{Farbe=schwarz}}(\text{Schaf}) \times \text{Schläft-in}))$ .

# Die Anfrage Schritt für Schritt.

Die Ausgangstabellen.

37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

37.2 Operationen auf  
Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

► Join

Schaf =

Name	Farbe	Fell
Dolly	weiß	lockig
Flauschi	schwarz	wuschelig
Peter	schwarz	wuschelig

Schläft-In =

Schaf-Name	Stall-Nr.
Dolly	1
Peter	5



# Die Anfrage Schritt für Schritt.

Die gefilterte Schaf-tabelle.

$\sigma_{\text{Farbe}=\text{schwarz}}(\text{Schaf}) =$

Name	Farbe	Fell
Flauschi	schwarz	wuschelig
Peter	schwarz	wuschelig

Schläft-In =

Schaf-Name	Stall-Nr.
Dolly	1
Peter	5

37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

37.2 Operationen auf  
Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

► Join

# Die Anfrage Schritt für Schritt.

Das Kreuzprodukt.

37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

37.2 Operationen auf  
Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

► Join

$\sigma_{\text{Farbe=schwarz}}(\text{Schaf}) \times \text{Schläft-In} =$

Name	Farbe	Fell	Schaf-Name	Stall-Nr.
Flauschi	schwarz	wuschelig	Dolly	1
Flauschi	schwarz	wuschelig	Peter	5
Peter	schwarz	wuschelig	Dolly	1
Peter	schwarz	wuschelig	Peter	5

# Die Anfrage Schritt für Schritt.

## Die Selektion.

### Kapitel 37 Relationenalgebren

#### 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

#### 37.2 Operationen auf Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

► Join

$$\sigma_{\text{Name=Schaf-Name}}(\sigma_{\text{Farbe=schwarz}}(\text{Schaf}) \times \text{Schläft-In}) =$$

Name	Farbe	Fell	Schaf-Name	Stall-Nr.
Peter	schwarz	wuschelig	Peter	5

# Die Anfrage Schritt für Schritt.

## Die Projektion.

### Kapitel 37 Relationenalgebren

#### 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

#### 37.2 Operationen auf Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

► Join

$$\pi_{\text{Stall}}(\sigma_{\text{Name=Schaf-Name}}(\sigma_{\text{Farbe=schwarz}}(\text{Schaf}) \times \text{Schläft-In})) =$$

Stall-Nr.
5

37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

37.2 Operationen auf  
Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

► Join

## Zur Übung

Beschreiben Sie mit Hilfe der bisherigen Operationen folgende Relation: Sie soll alle Paare enthalten von einem Schafsnamen zusammen mit der Größe des Stalls, in dem das Schaf schläft. (Wem das zu einfach ist: Alle Paare von Schafsnamen zusammen mit der Größe des Stalls, in dem die Mutter des Schafs schläft.)

- ▶ Es kommt sehr oft vor, dass man zwei Relationen derart »zusammenfügen« möchte, dass nur diejenigen Tupel des Kreuzproduktes betrachtet werden sollen, bei denen zwei bestimmte Attribute gleich sind.
- ▶ Man nennt dies einen *Join* der Relationen *anhand* zweier Attribute.

## 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

## 37.2 Operationen auf Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

▶ Join

37-22

## Definition

Sei  $A \subseteq A_1 \times \cdots \times A_n$  und  $B \subseteq B_1 \times \cdots \times B_m$ . Sei  $X = A_i = B_j$  für geeignete  $i$  und  $j$ . Dann ist der *Join*  $\theta_X(A, B)$  die Menge  $\sigma_{\text{alte } i\text{-Komponente} = \text{alte } j\text{-Komponente}}(A \times B)$ .

## Beispiel

Die Anfrage von vorhin lässt sich nun kürzer schreiben:

$$\pi_{\text{Stall}}(\theta_{\text{Name}}(\sigma_{\text{Farbe}=\text{schwarz}}(\text{Schaf}), \text{Schläft-In})).$$

## Relationen

Eine *Relation* auf Mengen  $A_1$  bis  $A_n$  ist eine Teilmenge des Kreuzprodukts dieser Mengen. Sie »entspricht« einer Tabelle, mit  $n$  Spalten und so vielen Zeilen wie es Tupel in der Relation gibt.

## Operationen auf Relationen

Seien  $A, B \subseteq A_1 \times \cdots \times A_n$  Relationen.

- ▶  $A \times B$  ist das *Kreuzprodukt* der Relationen.
- ▶  $A \cup B$  ist die *Vereinigung* der Relationen.
- ▶  $A \cap B$  ist der *Schnitt* der Relationen.
- ▶  $\sigma_P(A)$  ist die *Selektion* anhand eines Prädikats  $P$ .
- ▶  $\pi_{A_i}(A)$  ist die *Projektion* auf ein Attribut.
- ▶  $\theta_{A_i}(A, B)$  ist der *Join* der Relationen anhand eines Attributs.

### 37.1 Einführung

Alles ist eine Tabelle  
Relationen-Algebren

### 37.2 Operationen auf Relationen

Vereinigung und Schnitt  
Selektion  
Projektion  
Kreuzprodukt

▶ Join