



Skript zur Vorlesung
Datenbanksysteme I
Wintersemester 2010/2011

Kapitel 3: Die Relationale Algebra

Vorlesung: PD Dr. Matthias Schubert,
Übungen: Thomas Bernecker, Andreas Züfle

Skript © 2005 Christian Böhm

http://www.dbs.ifi.lmu.de/cms/Datenbanksysteme_I



Arbeiten mit Relationen

- Es gibt viele *formale* Modelle, um...
 - mit Relationen zu arbeiten
 - Anfragen zu formulieren
- Wichtigste Beispiele:
 - **Relationale Algebra**
 - **Relationen-Kalkül**
- Sie dienen als theoretisches Fundament für konkrete Anfragesprachen wie
 - SQL: Basiert i.w. auf der relationalen Algebra
 - QBE (= Query By Example) und Quel:
Basieren auf dem Relationen-Kalkül



Begriff Relationale Algebra

- Mathematik:
 - Algebra ist eine Operanden-Menge mit Operationen
 - Abgeschlossenheit: Werden Elemente der Menge mittels eines Operators verknüpft, ist das Ergebnis wieder ein Element der Menge
 - Beispiele
 - Natürliche Zahlen mit Addition, Multiplikation
 - Zeichenketten mit Konkatination
 - Boolesche Algebra: Wahrheitswerte mit \wedge , \vee , \neg
 - Mengen-Algebra:
 - Wertebereich: die Menge (*Klasse*) der Mengen
 - Operationen z.B. \cup , \cap , $-$ (Differenzmenge)



Begriff Relationale Algebra

- Relationale Algebra:
 - „Rechnen mit Relationen“
 - Was sind hier die Operanden? **Relationen (Tabellen)**
 - Beispiele für Operationen?
 - **Selektion von Tupeln nach Kriterien (z.B. *Gehalt* > 1000)**
 - **Kombination mehrerer Tabellen**
 - Abgeschlossenheit:
Ergebnis einer Anfrage ist immer eine (**neue**) Relation (oft ohne eigenen Namen)
 - Damit können einfache Terme der relationalen Algebra zu komplexeren zusammengesetzt werden



Grundoperationen

- 5 Grundoperationen der Relationalen Algebra:
 - Vereinigung: $R = S \cup T$
 - Differenz: $R = S - T$
 - Kartesisches Produkt (Kreuzprodukt): $R = S \times T$
 - Selektion: $R = \sigma_F(S)$
 - Projektion: $R = \pi_{A,B,\dots}(S)$
- Mit den Grundoperationen lassen sich weitere Operationen, (z.B. die Schnittmenge) nachbilden
- Manchmal wird die Umbenennung von Attributen als 6. Grundoperation bezeichnet



Vereinigung und Differenz

- Diese Operationen sind nur anwendbar, wenn die Schemata der beiden Relationen S und T übereinstimmen
- Die Ergebnis-Relation R bekommt Schema von S
- Vereinigung: $R = S \cup T = \{t \mid t \in S \vee t \in T\}$
- Differenz: $R' = S - T = \{t \mid t \in S \wedge t \notin T\}$
- Was wissen wir über die *Kardinalität* des Ergebnisses (Anzahl der Tupel von R)?
 - $|R| = |S \cup T| \leq |S| + |T|$
 - $|R'| = |S - T| \geq |S| - |T|$



Beispiel

Mitarbeiter:

Name	Vorname
Huber	Egon
Maier	Wolfgang
Schmidt	Helmut

Studenten:

Name	Vorname
Müller	Heinz
Schmidt	Helmut

Alle Personen, die Mitarbeiter oder Studenten sind:

Mitarbeiter \cup Studenten:

Name	Vorname
Huber	Egon
Maier	Wolfgang
Schmidt	Helmut
Müller	Heinz
Schmidt	Helmut

Duplikat-
Elimination!

Alle Mitarbeiter ohne diejenigen, die auch Studenten sind:

Mitarbeiter – Studenten:

Name	Vorname
Huber	Egon
Maier	Wolfgang



Kartesisches Produkt

Wie in Kapitel 2 bezeichnet das Kreuzprodukt

$$R = S \times T$$

**die Menge aller möglichen Kombinationen
von Tupeln aus S und T**

- Seien a_1, a_2, \dots, a_s die Attribute von S und b_1, b_2, \dots, b_t die Attribute von T
- Dann ist $R = S \times T$ die folgende Menge (Relation):
 $\{(a_1, \dots, a_s, b_1, \dots, b_t) \mid (a_1, \dots, a_s) \in S \wedge (b_1, \dots, b_t) \in T\}$
- Für die Anzahl der Tupel gilt:

$$|S \times T| = |S| \cdot |T|$$



Beispiel

Mitarbeiter

PNr	Name	Vorname	Abteilung
001	Huber	Erwin	01
002	Mayer	Hugo	01
003	Müller	Anton	02

Abteilungen

ANr	Abteilungsname
01	Buchhaltung
02	Produktion

Mitarbeiter \times Abteilungen

PNr	Name	Vorname	Abteilung	ANr	Abteilungsname
001	Huber	Erwin	01	01	Buchhaltung
001	Huber	Erwin	01	02	Produktion
002	Mayer	Hugo	01	01	Buchhaltung
002	Mayer	Hugo	01	02	Produktion
003	Müller	Anton	02	01	Buchhaltung
003	Müller	Anton	02	02	Produktion

Frage: Ist dies richtig?



Selektion

- Mit der Selektion $R = \sigma_F(S)$ werden diejenigen Tupel aus einer Relation S ausgewählt, die eine durch die logische Formel F vorgegebene Eigenschaft erfüllen
- R bekommt das gleiche Schema wie S
- Die Formel F besteht aus:
 - Konstanten („Meier“)
 - Attributen: Als Name (PNr) oder Nummer (\$1)
 - Vergleichsoperatoren: $=, <, \leq, >, \geq, \neq$
 - Boole'sche Operatoren: \wedge, \vee, \neg
- Formel F wird für jedes Tupel von S ausgewertet



Beispiel

Mitarbeiter

PNr	Name	Vorname	Abteilung
001	Huber	Erwin	01
002	Mayer	Hugo	01
003	Müller	Anton	02

Alle Mitarbeiter von Abteilung 01:

$\sigma_{\text{Abteilung}=01}(\text{Mitarbeiter})$

PNr	Name	Vorname	Abteilung
001	Huber	Erwin	01
002	Mayer	Hugo	01

Kann jetzt die Frage von S. 9 beantwortet werden?



Beispiel

Mitarbeiter \times Abteilungen

PNr	Name	Vorname	Abteilung	ANr	Abteilungsname
001	Huber	Erwin	01	01	Buchhaltung
001	Huber	Erwin	01	02	Produktion
002	Mayer	Hugo	01	01	Buchhaltung
002	Mayer	Hugo	01	02	Produktion
003	Müller	Anton	02	01	Buchhaltung
003	Müller	Anton	02	02	Produktion

$\sigma_{\text{Abteilung}=\text{ANr}}(\text{Mitarbeiter} \times \text{Abteilungen})$

PNr	Name	Vorname	Abteilung	ANr	Abteilungsname
001	Huber	Erwin	01	01	Buchhaltung
002	Mayer	Hugo	01	01	Buchhaltung
003	Müller	Anton	02	02	Produktion

Die Kombination aus Selektion und Kreuzprodukt heißt **Join**



Projektion

- Die Projektion $R = \pi_{A,B,\dots}(S)$ erlaubt es,
 - Spalten einer Relation auszuwählen
 - bzw. nicht ausgewählte Spalten zu streichen
 - die Reihenfolge der Spalten zu verändern
- In den Indizes sind die selektierten Attributnamen oder -Nummern ($\$1$) aufgeführt
- Für die Anzahl der Tupel des Ergebnisses gilt:

$$|\pi_{A,B,\dots}(S)| \leq |S|$$

Grund: Nach dem Streichen von Spalten können Duplikat-Tupel entstanden sein



Projektion: Beispiel

Mitarbeiter

PNr	Name	Vorname	Abteilung
001	Huber	Erwin	01
002	Mayer	Josef	01
003	Müller	Anton	02
004	Mayer	Maria	01

$\pi_{\text{Name, Abteilung}}(\text{Mitarbeiter}) = \dots$

Zwischenergebnis:

Name	Abteilung
Huber	01
Mayer	01
Müller	02
Mayer	01

Duplikate

Elimination →

Name	Abteilung
Huber	01
Mayer	01
Müller	02



Duplikatelimination

- Erforderlich nach...
 - Projektion
 - Vereinigung } „billige“ Basisoperationen, aber...
- Wie funktioniert Duplikatelimination?

```
for (int i = 0 ; i < R.length ; i++)  
  for (int j = 0 ; j < i ; j++)  
    if (R[i] == R[j])  
      // R[j] aus Array löschen
```
- Aufwand? $n=R.length$: $O(n^2)$
- Besserer Algorithmus mit Sortieren: $O(n \log n)$
 \Rightarrow An sich billige Operationen werden durch Duplikatelimination teuer



Beispiel-Anfragen

Gegeben sei folgendes Relationenschema:

Städte (SName: String, SEinw: Integer, Land: String)

Länder (LName: String, LEinw: Integer, Partei*: String)

* bei Koalitionsregierungen: jeweils eigenes Tupel pro Partei

- Bestimme alle Großstädte (≥ 500.000) und ihre Einwohner

$\pi_{SName, SEinw}(\sigma_{SEinw \geq 500.000}(Städte))$

- In welchem Land liegt die Stadt Passau?

$\pi_{Land}(\sigma_{SName=Passau}(Städte))$

- Bestimme die Namen aller Städte, deren Einwohnerzahl die eines beliebigen Landes übersteigt:

$\pi_{SName}(\sigma_{SEinw > LEinw}(Städte \times Länder))$



Beispiel-Anfragen

Gegeben sei folgendes Relationenschema:

Städte (SName: String, SEinw: Integer, Land: String)

Länder (LName: String, LEinw: Integer, Partei*: String)

* bei Koalitionsregierungen: jeweils eigenes Tupel pro Partei

- Finde alle Städtenamen in CDU-regierten Ländern

$$\pi_{\text{SName}}(\sigma_{\text{Land=LName}}(\text{Städte} \times \sigma_{\text{Partei=CDU}}(\text{Länder})))$$

oder auch:

$$\pi_{\text{SName}}(\sigma_{\text{Land=LName} \wedge \text{Partei=CDU}}(\text{Städte} \times \text{Länder}))$$

- Welche Länder werden von der SPD *allein* regiert

$$\pi_{\text{LName}}(\sigma_{\text{Partei=SPD}}(\text{Länder})) - \pi_{\text{LName}}(\sigma_{\text{Partei} \neq \text{SPD}}(\text{Länder}))$$



Abgeleitete Operationen

- Eine Reihe nützlicher Operationen lassen sich mit Hilfe der 5 Grundoperationen ausdrücken:

– Durchschnitt $R = S \cap T$

– Quotient $R = S \div T$

– Join $R = S \bowtie T$



Durchschnitt

- Idee: Finde gemeinsame Elemente in zwei Relationen (Schemata müssen übereinstimmen):

$$R' = S \cap T = \{t \mid t \in S \wedge t \in T\}$$

- Beispiel:
Welche Personen sind gleichzeitig Mitarbeiter und Student?

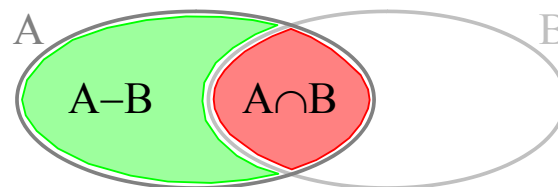
Mitarbeiter:	Name	Vorname	Studenten:	Name	Vorname
	Huber	Egon		Müller	Heinz
	Maier	Wolfgang		Schmidt	Helmut
	Schmidt	Helmut			

Mitarbeiter \cap Studenten:	Name	Vorname
	Schmidt	Helmut



Durchschnitt

- Implementierung der Operation „Durchschnitt“ mit Hilfe der Grundoperation „Differenz“:



- $A \cap B = A - (A - B)$
- Achtung!** Manche Lehrbücher definieren:
 - Durchschnitt ist Grundoperation
 - Differenz ist abgeleitete Operation(Definition gleichwertig, also genauso möglich)



Quotient

- Dient zur Simulation eines Allquantors
- Beispiel:

R_1	Programmierer	Sprache	R_2	Sprache
	Müller	Java		Basic
	Müller	Basic		C++
	Müller	C++		Java
	Huber	C++		
	Huber	Java		

- Welche Programmierer programmieren in **allen** Sprachen?

$R_1 \div R_2$	Programmierer
	Müller

- Umkehrung des kartesischen Produktes (daher: *Quotient*)



Join

- Wie vorher erwähnt:
Selektion über Kreuzprodukt zweier Relationen

– Theta-Join (Θ): $R \bowtie_{A \Theta B} S$

Allgemeiner Vergleich:

A ist ein Attribut von R und B ein Attribut von S

Θ ist ein beliebiges Prädikat, zum Beispiel einer der Operatoren: $=, <, \leq, >, \geq, \neq$

– Equi-Join: $R \bowtie_{A=B} S$

– Natural Join: $R \bowtie S$:

- Ein Equi-Join bezüglich aller gleichbenannten Attribute in R und S .
- Gleiche Spalten werden gestrichen (Projektion)



Join

- Implementierung mit Hilfe der Grundoperationen

$$R \bowtie_{A \Theta B} S = \sigma_{A \Theta B} (R \times S)$$

Gegeben sei folgendes Relationenschema:

Städte (SName: String, SEinw: Integer, Land: String)

Länder (LName: String, LEinw: Integer, Partei*: String)

* bei Koalitionsregierungen: jeweils eigenes Tupel pro Partei

- Finde alle Städtenamen in CDU-regierten Ländern

$$\pi_{\text{SName}} (\text{Städte} \bowtie_{\text{Land=LName}} \sigma_{\text{Partei=CDU}}(\text{Länder}))$$

- Bestimme die Namen aller Städte, deren Einwohnerzahl die eines beliebigen Landes übersteigt:

$$\pi_{\text{SName}} (\text{Städte} \bowtie_{\text{SEinw} > \text{LEinw}} \text{Länder})$$



SQL

- Die wichtigste Datenbank-Anfragesprache SQL beruht wesentlich auf der relationalen Algebra
- Grundform einer Anfrage*:

Projektion → **SELECT** <Liste von Attributnamen bzw. *>

Kreuzprodukt → **FROM** <ein oder mehrere Relationennamen>

Selektion → [**WHERE** <Bedingung>]

- Mengenoperationen:

SELECT ... FROM ... WHERE

UNION

SELECT ... FROM ... WHERE

* SQL ist **case-insensitive**: SELECT = select = SeLeCt



SQL

- Hauptunterschied zwischen SQL und rel. Algebra:
 - Operatoren bei SQL nicht beliebig schachtelbar
 - Jeder Operator hat seinen festen Platz
- Trotzdem:
 - Man kann zeigen, daß jeder Ausdruck der relationalen Algebra gleichwertig in SQL formuliert werden kann
 - Die feste Anordnung der Operatoren ist also keine wirkliche Einschränkung (Übersichtlichkeit)
 - Man sagt, SQL ist *relational vollständig*
- Weitere Unterschiede:
 - Nicht immer werden Duplikate eliminiert (Projektion)
 - zus. Auswertungsmöglichkeiten (Aggregate, Sortieren)



SELECT

- Entspricht **Projektion** in der relationalen Algebra
- Aber: Duplikatelimination nur, wenn durch das Schlüsselwort **DISTINCT** explizit verlangt
- Syntax:
 - SELECT * FROM ... -- Keine Projektion
 - SELECT **A₁, A₂, ...** FROM ... -- Projektion ohne
-- Duplikatelimination
 - SELECT **DISTINCT** A₁, A₂, ... -- Projektion mit
-- Duplikatelimination
- Bei der zweiten Form kann die Ergebnis„*relation*“ also u.U. Duplikate enthalten
- Grund: Performanz



SELECT

- Bei den Attributen A_1, A_2, \dots lässt sich angeben...
 - Ein Attributname einer beliebigen Relation, die in der FROM-Klausel angegeben ist
 - Ein **skalarer Ausdruck**, der Attribute und Konstanten mittels arithmetischer Operationen verknüpft
 - Im Extremfall: Nur eine Konstante
 - Aggregationsfunktionen (siehe später)
 - Ein Ausdruck der Form A_1 **AS** A_2 :
 A_2 wird der neue Attributname (Spaltenüberschrift)

- Beispiel:

```
select pname
       preis*13.7603 as oespr,
       preis*kurs as usdpr,
       'US$' as currency
from   produkt, wahrungen....
```

pname	oespr	usdpr	currency
nagel	6.88	0.45	US\$
dübel	1.37	0.09	US\$
...			



FROM

- Enthält mindestens einen Eintrag der Form R_1
- Enthält die FROM-Klausel mehrere Einträge
 - FROM R_1, R_2, \dots
 so wird das kartesische Produkt gebildet:
 - $R_1 \times R_2 \times \dots$
- Enthalten zwei verschiedene Relationen R_1, R_2 ein Attribut mit gleichem Namen, dann ist dies in der SELECT- und WHERE-Klausel mehrdeutig
- Eindeutigkeit durch vorangestellten Relationennamen:


```
SELECT   Mitarbeiter.Name, Abteilung.Name, ...
FROM     Mitarbeiter, Abteilung
WHERE    ...
```



FROM

- Man kann Schreibarbeit sparen, indem man den Relationen temporär (innerhalb der Anfrage) kurze Namen zuweist (**Alias-Namen**):

```
SELECT    m.Name, a.Name, ...
FROM      Mitarbeiter m, Abteilung a
WHERE     ...
```

- Dies lässt sich in der SELECT-Klausel auch mit der Sternchen-Notation kombinieren:

```
SELECT    m.*, a.Name AS Abteilungsname, ...
FROM      Mitarbeiter m, Abteilung a
WHERE     ...
```

- Manchmal **Self-Join** einer Relation mit sich selbst:

```
SELECT    m1.Name, m2.Name, ...
FROM      Mitarbeiter m1, Mitarbeiter m2
WHERE     ...
```



WHERE

- Entspricht der **Selektion** der relationalen Algebra
- Enthält genau ein logisches Prädikat Θ (Funktion die einen booleschen Wert (wahr/falsch) zurück gibt).
- Idee: Alle Tupel des kartesischen Produktes aus der FROM-Klausel werden getestet, ob sie Θ erfüllen.
- Das logische Prädikat besteht aus
 - Vergleichen zwischen Attributwerten und Konstanten
 - Vergleichen zwischen verschiedenen Attributen
 - Vergleichsoperatoren: $=$, $<$, $<=$, $>$, $>=$, $<>$
 - Test auf *Wert undefiniert*: A_1 **IS NULL/IS NOT NULL**
 - Inexakter Stringvergleich: A_1 **LIKE** 'Datenbank%'
 - A_1 **IN** (2, 3, 5, 7, 11, 13)

*Der Gleichheitsoperator wird **nicht** etwa wie in Java verdoppelt



WHERE

- Innerhalb eines Prädikates: Skalare Ausdrücke:
 - Numerische Werte/Attribute mit **+**, **-**, *****, **/** verknüpfbar
 - Strings: **char_length**, Konkatination **||** und **substring**
 - Spezielle Operatoren für Datum und Zeit
 - Übliche Klammernsetzung.
- Einzelne Prädikate können mit **AND**, **OR**, **NOT** zu komplexeren zusammengefasst werden
- Effizientere Ausführung möglich mit Index



WHERE

- Inexakte Stringsuche: A_1 **LIKE** 'Datenbank%'
 - bedeutet: Alle Datensätze, bei denen Attribut A_1 mit dem Präfix *Datenbank* beginnt.
 - Entsprechend: A_1 **LIKE** '%Daten%'
 - In dem Spezialstring hinter LIKE ...
 - % steht für einen beliebig belegbaren Teilstring
 - _ steht für ein einzelnes frei belegbares Zeichen

- Beispiel:

Alle Mitarbeiter, deren
Nachname auf 'er' endet:

**select * from mitarbeiter
where name like '%er'**

Mitarbeiter

PNr	Name	Vorname	ANr
001	Huber	Erwin	01
002	Mayer	Josef	01
003	Müller	Anton	02
004	Schmidt	Helmut	01



Join

- Normalerweise wird der Join wie bei der relationalen Algebra als Selektionsbedingung über dem kartesischen Produkt formuliert.
- Beispiel: Join zwischen Mitarbeiter und Abteilung
select * from Mitarbeiter m, Abteilungen a **where** m.ANr = a.ANr
- In neueren SQL-Dialekten auch möglich:
 - **select * from** Mitarbeiter m **join** Abteilungen a **on** a.ANr=m.ANr
 - **select * from** Mitarbeiter **join** Abteilungen **using** (ANr)
 - **select * from** Mitarbeiter **natural join** Abteilungen

Nach diesem Konstrukt können mit einer WHERE-Klausel weitere Bedingungen an das Ergebnis gestellt werden.



Beispiel (Wdh. S. 12)

select * from Mitarbeiter m, Abteilungen a...

PNr	Name	Vorname	m.ANr	a.ANr	Abteilungsname
001	Huber	Erwin	01	01	Buchhaltung
001	Huber	Erwin	01	02	Produktion
002	Mayer	Hugo	01	01	Buchhaltung
002	Mayer	Hugo	01	02	Produktion
003	Müller	Anton	02	01	Buchhaltung
003	Müller	Anton	02	02	Produktion

...**where** m.ANr = a.ANr

PNr	Name	Vorname	m.ANr	a.ANr	Abteilungsname
001	Huber	Erwin	01	01	Buchhaltung
002	Mayer	Hugo	01	01	Buchhaltung
003	Müller	Anton	02	02	Produktion

Kann jetzt die Frage von S. 9 beantwortet werden?



Beispiele:

- Gegeben sei folgendes Datenbankschema:
 - Kunde (KName, KAdr, Kto)
 - Auftrag (KName, Ware, Menge)
 - Lieferant (LName, LAdr, Ware, Preis)
- Welche Lieferanten liefern Mehl oder Milch?

```
select distinct LName
from Lieferant
where Ware = 'Mehl' or Ware = 'Milch'
```
- Welche Lieferanten liefern irgendetwas, das der Kunde Huber bestellt hat?

```
select distinct LName
from Lieferant l, Auftrag a
where l.Ware = a.Ware and KName = 'Huber'
```



Beispiele (Self-Join):

Kunde (<u>KName</u> , KAdr, Kto)
Auftrag (<u>KName</u> , <u>Ware</u> , Menge)
Lieferant (<u>LName</u> , LAdr, <u>Ware</u> , Preis)

- Name und Adressen aller Kunden, deren Kontostand kleiner als der von Huber ist

```
select k1.KName, k1.Adr
from Kunde k1, Kunde k2
where k1.Kto < k2.Kto and k2.KName = 'Huber'
```
- Finde alle Paare von Lieferanten, die eine gleiche Ware liefern

```
select distinct L1.Lname, L2.LName
from Lieferant L1, Lieferant L2
where L1.Ware=L2.Ware and L1.LName<L2.LName
```

 ?



Beispiele (Self-Join)

Lieferant*

Müller	Mehl
Müller	Haferfl
Bäcker	Mehl

Ohne Zusatzbedingung:

Müller	Mehl	Müller	Mehl
Müller	Mehl	Bäcker	Mehl
Müller	Haferfl	Müller	Haferfl
Bäcker	Mehl	Müller	Mehl
Bäcker	Mehl	Bäcker	Mehl

Nach Projektion:

Müller	Müller	}	L1.LName = L2.LName
Müller	Bäcker		
Bäcker	Müller		
Bäcker	Bäcker		

L1.LName > L2.LName



UNION, INTERSECT, EXCEPT

- Üblicherweise werden mit diesen Operationen die Ergebnisse zweier SELECT-FROM-WHERE-Blöcke verknüpft:

```
select * from Mitarbeiter where name like 'A%'
union -- Vereinigung mit Duplikatelimination
select * from Studenten where name like 'A%'
```

- Bei neueren Datenbanksystemen ist auch möglich:

```
select * from Mitarbeiter union Studenten where ...
```

- Genauso bei:

- Durchschnitt: **INTERSECT**
- Differenz: **EXCEPT**
- Vereinigung **ohne** Duplikatelimination: **UNION ALL**



UNION, INTERSECT, EXCEPT

- Die **relationale Algebra** verlangt, daß die beiden Relationen, die verknüpft werden, das **gleiche** Schema besitzen (Namen und Wertebereiche)
- **SQL** verlangt nur **kompatible Wertebereiche**, d.h.:
 - beide Wertebereich sind **character** (Länge usw. egal)
 - beide Wertebereiche sind Zahlen (Genauigkeit egal)
 - oder beide Wertebereiche sind gleich



UNION, INTERSECT, EXCEPT

- Mit dem Schlüsselwort **corresponding** beschränken sich die Operationen automatisch auf die **gleich benannten** Attribute
- Beispiel (aus *Datenbanken kompakt*):

R_1 :

A	B	C
1	2	3
2	3	4

R_2 :

A	C	D
2	2	3
5	3	2

R_1 union R_2 : R_1 union corresponding R_2 :

A	B	C
1	2	3
2	3	4
2	2	3
5	3	2

A	C
1	3
2	4
2	2
5	3



Änderungs-Operationen

- Bisher: Nur *Anfragen* an das Datenbanksystem
- Änderungsoperationen modifizieren den Inhalt eines oder mehrerer Tupel einer Relation
- Grundsätzlich unterscheiden wir:
 - **INSERT**: Einfügen von Tupeln in eine Relation
 - **DELETE**: Löschen von Tupeln aus einer Relation
 - **UPDATE**: Ändern von Tupeln einer Relation
- Diese Operationen sind verfügbar als...
 - **Ein-Tupel-Operationen**
z.B. die Erfassung eines neuen Mitarbeiters
 - **Mehr-Tupel-Operationen**
z.B. die Erhöhung aller Gehälter um 2.1%



Die UPDATE-Anweisung

- Syntax:
update *relation*
set $attribut_1 = ausdruck_1$
[, ... ,
 $attribut_n = ausdruck_n$]*
[**where** *bedingung*]
- Wirkung:
In allen Tupeln der Relation, die die Bedingung erfüllen (falls angegeben, sonst in allen Tupeln), werden die Attributwerte wie angegeben gesetzt

*falsch in *Heuer&Saake*: Zuweisungen müssen durch Kommata getrennt werden



Die UPDATE-Anweisung

- UPDATE ist i.a. eine Mehrtuple-Operation
- Beispiel:
update Angestellte
set Gehalt = 6000
- Wie kann man sich auf ein einzelnes Tupel beschränken?
Spezifikation des Schlüssels in WHERE-Bedg.
- Beispiel:
update Angestellte
set Gehalt = 6000
where PNr = 7



Die UPDATE-Anweisung

- Der alte Attribut-Wert kann bei der Berechnung des neuen Attributwertes herangezogen werden
- Beispiel:
Erhöhe das Gehalt aller Angestellten, die weniger als 3000,-- € verdienen, um 2%
update Angestellte
set Gehalt = Gehalt * 1.02
where Gehalt < 3000
- UPDATE-Operationen können zur Verletzung von Integritätsbedingungen führen:
Abbruch der Operation mit Fehlermeldung.



Die DELETE-Anweisung

- Syntax:
delete from *relation*
[**where** *bedingung*]
- Wirkung:
 - Löscht alle Tupel, die die Bedingung erfüllen
 - Ist keine Bedingung angegeben, werden *alle* Tupel gelöscht
 - Abbruch der Operation, falls eine Integritätsbedingung verletzt würde (z.B. Fremdschlüssel ohne *cascade*)
- Beispiel: Löschen aller Angestellten mit Gehalt 0
delete from Angestellte
where Gehalt = 0



Die INSERT-Anweisung

- Zwei unterschiedliche Formen:
 - Einfügen konstanter Tupel (Ein-Tupel-Operation)
 - Einfügen berechneter Tupel (Mehr-Tupel-Operation)
- Syntax zum Einfügen konstanter Tupel:
insert into *relation* (*attribut₁*, *attribut₂*,...) **values** (*konstante₁*, *konstante₂*, ...)
- oder:
insert into *relation* **values** (*konstante₁*, *konstante₂*, ...)



Einfügen konstanter Tupel

- Wirkung:
Ist die optionale Attributliste hinter dem Relationennamen angegeben, dann...
 - können unvollständige Tupel eingefügt werden:
Nicht aufgeführte Attribute werden mit NULL belegt
 - werden die Werte durch die Reihenfolge in der Attributliste zugeordnet
- Beispiel:
insert into Angestellte (Vorame, Name, PNr)
values ('Donald', 'Duck', 678)

PNr	Name	Vorname	ANr
678	Duck	Donald	NULL



Einfügen konstanter Tupel

- Wirkung:
Ist die Attributliste *nicht* angegeben, dann...
 - können unvollständige Tupel nur durch explizite Angabe von NULL eingegeben werden
 - werden die Werte durch die Reihenfolge in der DDL-Definition der Relation zugeordnet
(mangelnde Datenunabhängigkeit!)
- Beispiel:
insert into Angestellte
values (678, 'Duck', 'Donald', NULL)

PNr	Name	Vorname	ANr
678	Duck	Donald	NULL



Einfügen berechneter Tupel

- Syntax zum Einfügen berechneter Tupel:
insert into *relation* [(*attribut*₁ , ...)]
(**select ... from ... where ...**)
- Wirkung:
 - Alle Tupel des Ergebnisses der SELECT-Anweisung werden in die Relation eingefügt
 - Die optionale Attributliste hat dieselbe Bedeutung wie bei der entsprechenden Ein-Tupel-Operation
 - Bei Verletzung von Integritätsbedingungen (z.B. Fremdschlüssel nicht vorhanden) wird die Operation nicht ausgeführt (Fehlermeldung)



Einfügen berechneter Tupel

- Beispiel:
Füge alle Lieferanten in die Kunden-Relation ein (mit Kontostand 0)
 - Datenbankschema:
 - Kunde (KName, KAdr, Kto)
 - Lieferant (LName, LAdr, Ware, Preis)
- insert into Kunde**
(select distinct LName, LAdr, 0 from Lieferant)