Hochschule **RheinMain**
University of Applied Sciences
Wiesbaden Rüsselsheim

RELATIONALE ALGEBRA FÜR DATENBANKEN

Sommersemester 2022 - April & Mai

Letztes Update: 20. April 2022

Prof. Dr. Eva-Maria Iwer

Fachbereich Design Informatik Medien (DCSM)
Hochschule **RheinMain**

GLIEDERUNG

EINLEITUNG

EINLEITUNG

Edgar F. Codd

→ Ende der 1960er-Jahre entwickelte Edgar F. Codd am IBM Research Laboratory in San Jose die Grundlagen der heutigen relationalen Algebra

→ er stellte die Behauptung auf, dass das relationale Modell in vielen Aspekten dem Graphenmodell überlegen sei

→ Abhängigkeit vom Aufbau des jeweiligen Netzwerks

→ Bei benachbarten Daten, sehr kurze Antwortzeiten

→ Im Netzwerk stark verstreute Daten, kann die Wartezeit unzumutbar lang werden

EINLEITUNG

Edgar F. Codd

→ Die Datenbankentwickler mussten bei der Erstellung eines Netzwerkmodells von vorneherein sämtliche denkbaren Anfragen berücksichtigen, da nachträgliche Änderungen am Datenmodell nur noch sehr schwer umgesetzt werden konnten.

→ Idee: Relationen (Tabellen), die je nach Anfrage unterschiedlich miteinander verknüpft werden können

EINLEITUNG

Relationale Algebra

→ Grundlage jeder Datenbanksprache eines RDBMS

→ Wenn das Verständnis der Funktionsweise erlernt ist (Train your Brain), kann jede RDBMS-Sprache einfach erlernt werden.

UNTERSCHIEDE ZUR MATHEMATIK

VON DS ZU DB

Unterschiede

- → Berechnung im Bereich der Multimengen
- → Differenz-Zeichen ”–”
- → Bezeichnung Tupel steht für n-Tupel

MULTIMENGEN

VON DER MENGENTHEORIE ZUR MULTIMENGENTHEORIE

Mengentheorie

Duplikate sind nicht erlaubt.

Multimengen

Multimengen (bags) unterscheiden sich zu normalen Mengen (sets) nur durch die Eigenschaft, dass Duplikate erlaubt sind.

VON DER MENGENTHEORIE ZUR MULTIMENGENTHEORIE

Auswirkung auf bereits bekannte Operatoren

Die Operatoren arbeiten grundsätzlich genauso, jedoch gibt es folgende Besonderheiten:

→ Vereinigung $R \cup S$: Wenn ein Element in R a und in S b mal auftritt, dann tritt es in der Ergebnisrelation a+b mal auf.

→ Schnitt: $R \cap S$: Wenn ein Element in R a und in S b mal auftritt, dann tritt es in der Ergebnisrelation min(a,b) mal auf.

→ Differenz: $R - S$: Wenn ein Element in R a und in S b mal auftritt, dann tritt es in der Ergebnisrelation a-b mal auf, wenn a größer b, ansonsten 0 mal.

BEGRIFFE UND BEDEUTUNG

BEGRIFFE

Algebra

Eine Algebra ist eine formale Sprache, in der Operatoren mit dazugehörigen Regeln definiert werden.

relationale Algebra

Eine relationale Algebra ist eine Abfragesprache auf Relation, in der relationale Operationen mit entsprechenden Regeln definiert sind. Diese relationalen Operatoren führen eine oder mehrere Relationen in eine Ergebnisrelation über. Das Ergebnis einer relationalen Operation ist immer eine Relation, somit können wir darauf wieder einen relationalen Operator anwenden.

BEGRIFFE

Relation

Eine Relation ist eine zweidimensionale Tabelle, welche die Daten repräsentiert. Sie besteht aus einem Kopf und einem Rumpf und hat folgende Eigenschaften:

→ Tupel sind nicht geordnet.

→ Attribute sind nicht geordnet.

→ Alle Attribute sind atomar.

BEGRIFFE

Beispiel einer Relation

ID	name	firstName	address
123	Smith	Peter	123 1st St, Redmond
689	Meyer	Marc	123 2nd St, Seattle
117	Kim	Nathalie	17 5th St, Seattle

Tabelle: Relation Student

BEGRIFFE

Tuple

Jede Zeile präsentiert einen Studenten. Wir nennen jede Zeile ein Tupel.

Attribut

Jede Spalte präsentiert eine Eigenschaft von einem Studenten. Wir nennen das Attribut.

BEGRIFFE

Attribute

Die Spalten einer Relation werden Attribute genannt. In der Beispielrelation Student auf Folie 15 sind die Attribute:

→ ID

→ name

→ firstName

→ address

Attribute stehen am Kopf der Spalten und beschreiben die Bedeutung der Inhalte in der darunterliegenden Spalte.

BEGRIFFE

Tupel

→ Die Zeilen einer Relation heißen Tupel.

→ Die Außnahme ist hierbei die Kopfzeile, welche die Namen der Attribute enthält.

→ Ein Tupel hat für jedes Attribute einen Wert.

→ Wenn ein Tupel isoliert auftritt, soll die Reihenfolge des Relationsschema verwendet werden. Zum Beispiel: (689, Meyer, Marc, „123 2nd St, Seattle“)

BEGRIFFE

Relationsschema

Das Relationsschema ist der Name der Relation und die Menge der dazugehörigen Attribute. Das Schema für die Relation Student auf Folie 15 ist

Student(ID, name, firstName, address)

Die Attribute sind eine Menge und keine Liste. Das bedeutet, folgendes Relationsschema ist auch valide:

Student(name, address, firstName, ID)

Um ein vollständes Relationsschema zu erzeugen, muss noch zu jedem Attribute eine Domain (Gebiet) hinzugefügt werden. Das Ergebnis sähe dann folgendermaßen aus:

Student(ID:integer, name:string, firstName:string, address:string)

BEGRIFFE

Domain

Die Domain umfasst die Menge aller möglichen Werte für die Attribute. Jeder Attributwert in der Relation muss mit der Domain übereinstimmen.

BEGRIFFE

mögliche Domains

- → integer
- → real
- → string
- → boolean

BEGRIFFE

nicht erlaubte Domains

- → array
- → set
- → list
- → oder alle anderen Typen bei denen der Inhalt in kleinere valide Teile aufgeteilt werden kann.

BEGRIFFE

Kardinalität

Die Kardinalität beschreibt die Anzahl der Tupel in einer Relation. Die Relation Student auf Folie 15 hat eine Kardinalität von 3.

Grad

Grad beschreibt die Anzahl der Attribute in einer Relation. Der Grad der Relation Student auf Folie 15 ist 4.

BEGRIFFE

Zusammenfassung

Formaler Bezeichner	Informelle Bezeichnung
Relation	Tabelle
Tupel	eine Zeile (Reihe) einer Tabelle
Kardinalität	Anzahl der Zeilen einer Tabelle
Attribute	eine Spalte (Feld) einer Tabelle
Grad	Anzahl der Spalter einer Tabelle
Gebiet	Menge aller möglichen Werte

Tabelle: Begriffe in der relationalen Datenstruktur

MENGENOPERATIONEN

MENGENOPERATIONEN

Die drei bekanntesten Mengenoperationen sind Vereinigung, Schnitt und Differenz. Es handelt sich hier jeweils um binäre Operatoren. Folgende Konditionen müssen erfüllt sein, um die Operationen anzuwenden - hier beispielhaft an den Relationen R und S:

→ R und S müssen die gleichen Attributnamen, sowie Domains, besitzen.

→ R und S müssen so geordnet werden, dass die Ordnung der Attribute in R und S gleich sind.

MENGENOPERATIONEN

Vereinigung

SYMBOL: $\cup$

DEFINITION: Aus den beiden Relationen R und S ($R \cup S$) wird eine neue Relation erzeugt, die alle Tupel enthält, die in mindestens einer der beiden Relationen vorkommt. Tupel die in beiden Relationen vorkommen werden so oft aufgenommen wie sie in der Summe vorkommen.

MATHEMATISCHE BESCHREIBUNG: $R \cup S = \{\text{t} \mid \text{t} \in \text{R oder t} \in S\}$

MENGENOPERATIONEN

Vereinigung

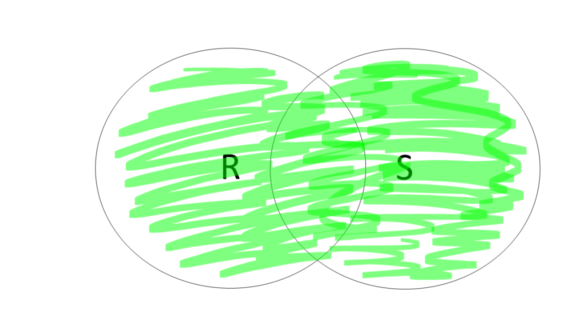


Abbildung: $R \cup S$

MENGENOPERATIONEN

$R \cup S$ (allgemein)

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{array} \qquad S = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_5 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{array}$$

Ergebnis:

$$R \cup S = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_1 & b_1 & c_1 \\ a_5 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{array}$$

MENGENOPERATIONEN

$R \cup S$ (leicht)

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & b \\ \hline 4 & 1 \\ 5 & 2 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \cup S = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 2 & 4 \\ 3 & 5 \\ 4 & 1 \\ 5 & 2 \\ 5 & 2 \end{array}$$

MENGENOPERATIONEN

Schnitt

SYMBOL: $\cap$

DEFINITION: Aus den beiden Relationen R und S ($R \cap S$) wird eine Relation erzeugt, die nur die Tupel enthält, die in beiden Relationen vorkommen.

Bei Mehrfachvorkommen der Tupel wird die geringere Anzahl der Tupel von den beiden Ausgangsrelationen in die Ergebnisrelation übernommen. MATHEMATISCHE BESCHREIBUNG: $R \cap S = \{$t | t $\in$ R and t $\in S\}$

SET OPERATIONS

Schnitt

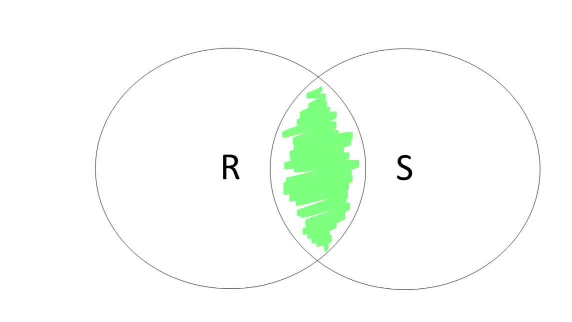


Abbildung: $R \cap S$

MENGENOPERATIONEN

$R \cap S$ (allgemein)

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_5 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_7 & b_4 & c_4 \end{array}$$

Ergebnis:

$$R \cap S = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{array}$$

MENGENOPERATIONEN

$R \cap S$ (leicht)

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & b \\ \hline 4 & 1 \\ 5 & 2 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \cap S = \begin{array}{cc} a & b \\ \hline 5 & 2 \end{array}$$

MENGENOPERATIONEN

Differenz

SYMBOL: $-$

DEFINITION: Aus den beiden Relationen ($R - S$) wird eine Relation erzeugt, welche die Tupel enthält, die in der Relation R, nicht aber in S vorkommen.

MATHEMATISCHE BESCHREIBUNG: $R - S = \{\text{t} \mid \text{t} \in$ R und nicht t $\in S\}$

MENGENOPERATIONEN

Differenz

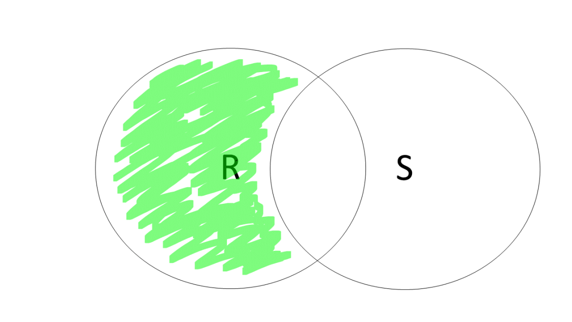


Abbildung: $R - S$

MENGENOPERATIONEN

Differenz

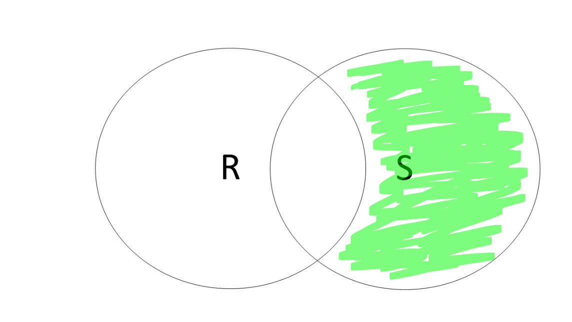


Abbildung: $S - R$

MENGENOPERATIONEN

$R - S$ (allgemein)

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_5 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_7 & b_4 & c_4 \end{array}$$

Ergebnis:

$$R - S = \begin{array}{ccc} a & b & c \\ \hline a_2 & b_2 & c_2 \\ a_4 & b_4 & c_4 \end{array} \qquad S - R = \begin{array}{ccc} a & b & c \\ \hline a_5 & b_2 & c_2 \\ a_7 & b_4 & c_4 \end{array}$$

MENGENOPERATIONEN

$R - S$ (leicht)

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & b \\ \hline 4 & 1 \\ 5 & 2 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R - S = \begin{array}{cc} a & b \\ \hline 2 & 4 \\ 3 & 5 \end{array} \qquad S - R = \begin{array}{cc} a & b \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

ENTFERNUNGSOPERATIONEN

ENTFERNUNGSOPERATIONEN

Allgemein

Die Entfernungsoperationen sind unäre Operationen, welche entweder Attribute (Projektion) oder bestimmte Tupel (Selektion) aus der Relation entfernen.

ENTFERNUNGSOPERATIONEN

Projektion

SYMBOL: π (pi)

DEFINITION: Aus der Relation R wird eine neue Relation erzeugt, die alle Tupel von R enthält, eingeschränkt auf die angegeben Auswahl von Attributen. Die Werte des Ausdrucks $\pi_{A_1,A_2,...,A_n}(R)$ ist die Relation die nur die Spalten für die Attribute $\pi_{A_1,A_2,...,A_n}$ von R besitzt. Das Schema der entstehenden Relation ist die Menge der Attribute $A_1, A_2, ..., A_n$, welche wir aus üblicherweise in der aufgeführten Ordnung zeigen.

ENTFERNUNGSOPERATIONEN

Projektion

A	B	C	D
a	b	c	d
a	f	g	h
a	j	k	l
m	n	o	p
q	r	s	t
u	v	w	x
a	z	aa	bb

A	D
a	d
a	h
a	l
m	p
q	t
u	x
a	bb

Abbildung: $\pi_{A,D}(R)$

ENTFERNUNGSOPERATIONEN

$\pi(R)$ (allgemein)

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array}$$

Ergebnis:

$$\pi_{a,c} = \begin{array}{cc} a & c \\ \hline a_1 & c_1 \\ a_2 & c_2 \\ a_3 & c_3 \\ a_4 & c_4 \end{array}$$

ENTFERNUNGSOPERATIONEN

$\pi(R)$ **(leicht)**

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & b \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Result:

$$\pi_a(R) = \begin{array}{c} a \\ \hline 5 \\ 2 \\ 3 \end{array} \qquad \pi_b(S) = \begin{array}{c} b \\ \hline 1 \\ 2 \end{array}$$

ENTFERNUNGSOPERATIONEN

Selektion

SYMBOL: σ_C (sigma)

DEFINITION: Aus der Relation R wird eine neue Relation erzeugt, die nur diejenigen Tupel aus der Relation R enthält, die der angegebenen Bedingung (C) genügen. Es entsteht eine Teilmenge der Tupel von R. Die Bedingung C wird für jedes Tupel getestet, wenn das Ergebnis wahr ist, wird das Tupel in die Ergebnisrelation übernommen.

ENTFERNUNGSOPERATIONEN

Selektion

A	B	C	D
a	b	c	d
a	f	g	h
a	j	k	l
m	n	o	p
q	r	s	t
u	v	w	x
a	z	aa	bb

A	B	C	D
a	b	c	d
a	f	g	h
a	j	k	l
a	z	aa	bb

Abbildung: $\sigma_{A='a'}(R)$

ENTFERNUNGSOPERATIONEN

Bedingung in der Selektion

Die Bedingung C kann ein komplexer Ausdruck mit einer Vielzahl von Teilen sein. Folgende Operatoren können verwendet werden:

→ UND ($\wedge$)

→ ODER ($\vee$)

→ NICHT ($\neg$)

WAHRHEITSTABELLE

UND $\wedge$, ODER $\vee$ und Negation $\neg$

A	B	$A \wedge B$	$A \vee B$	$\neg A$
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	0

ENTFERNUNGSOPERATIONEN

$\sigma_C(R)$ allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array}$$

Ergebnis:

$$\sigma_{a='a_1'} = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \end{array}$$

ENTFERNUNGSOPERATIONEN

$\sigma_C(R)$ (einfach)

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & b \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$\sigma_{a=5}(R) = \begin{array}{cc} a & b \\ \hline 5 & 2 \end{array} \qquad \sigma_{b=2}(S) = \begin{array}{cc} a & b \\ \hline 5 & 2 \end{array}$$

ENTFERNUNGSOPERATIONEN

$\sigma_C(R)$ (schwieriger)

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 5 & 3 \end{array}$$

Ergbenis:

$$\sigma_{a=1 \text{ AND } c=3}(R) = \begin{array}{ccc} a & b & c \\ \hline 1 & 2 & 3 \\ 1 & 5 & 3 \end{array}$$

$$\sigma_{a=1 \text{ OR } c=5}(R) = \begin{array}{ccc} a & b & c \\ \hline 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 5 & 3 \end{array}$$

UMBENENNUNGSOPERATIONEN

UMBENNENUNGSOPERATIONEN

Rho

SYMBOL: ρ (rho)

DEFINITION: Mit dem Umbennenungsoperator $\rho_{S(A_1,A_2,1_3,...,A_n)}(R)$ wird eine neue Relation erstellt, welche genau die gleichen Tupel enthält, wie die Ausgangsrelation R, jedoch mit dem neuen Namen S. Außerdem, wurden die Attribute der Ergebnisrelation S folgendermaßen benannt $A_1, A_2, A_3, ..., A_n$. Die Attribute sind dabei in der Reihenfolge von links nach rechts geordnet. Wenn wir nur den Namen der Relation zu S umwandeln wollen, können wir die Attribute weglassen, und einfach $\rho_S(R)$ sagen.

UMBENNENUNG

rho (allgemein)

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array}$$

Ergebnis:

$$\rho_{S(g,h,i)}(R) = S = \begin{array}{ccc} g & h & i \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array}$$

UMBENNENUNG

rho (einfach)

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array}$$

Ergebnis:

$$\rho_S(R) = S = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array}$$

UMBENNENUNG

rho (einfach)

given:

$$R = \begin{array}{ccc} a & b & c \\ \hline 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 5 & 3 \end{array}$$

Result:

$$\rho_{S(c,b,a)}(R) = S = \begin{array}{ccc} c & b & a \\ \hline 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 5 & 3 \end{array}$$

ERWEITERTE OPERATIONEN

Erweiterte Projektion

SYMBOL: π_L (pi)
DEFINITION: Wir erweitern den Projektionsoperator, um einen Berechnungs- und Umbenennungskomponente. Die erweiterte Projektion, welcher mit $\pi_L(R)$ bezeichnet wird, hat im Index eine Liste, welche folgende Elemente besitzen kann:

1. Attribute von R
2. Einen Ausdruck $x \rightarrow y$, wobei x und y Namen für Attribute sind. Das Element $x \rightarrow y$ in der Liste L bedeutet, dass das Attribute X aus R projeziert wird und in y umbenannt wird.

ERWEITERTE OPERATIONEN

Erweiterte Projektion

3. Einen Ausdruck $E \rightarrow z$, wobei E ein Ausdruck ist, welche Attribute von R enthalten kann, sowie Konstanten, arithmethische Operaten und auch Zeichenoperatoren. z ist dabei wieder der neue Name des Ergebnisattributs. Zum Beispiel $a + b \rightarrow x$ als Listenelement representiert die summe der Werte der Attribute a und b, wobei das Ergebnisattribut x genannt wird. Das Element $c \| d \rightarrow e$ bedeutet, dass die Strings aus a und b zusammengefügt werden zu einem Attribute und dieses jetzt e heißt.

ERWEITERTE OPERATIONEN

Erweiterte Projektion einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 5 & 2 \end{array}$$

Ergebnis:

$$\pi_{a \rightarrow c, b}(R) = \begin{array}{cc} c & b \\ \hline 5 & 2 \\ 2 & 4 \\ 5 & 2 \end{array}$$

ERWEITERTE OPERATIONEN

Erweiterte Projektion schwieriger

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{array}$$

Ergebnis:

$$\pi_{a+b\rightarrow x, b-1\rightarrow y, b\rightarrow z}(R) = \begin{array}{ccc} x & y & z \\ \hline 3 & 1 & 2 \\ 4 & 2 & 3 \\ 5 & 3 & 4 \end{array}$$

DUPLIKATENTFERNER

ERWEITERTE OPERATIONEN

Duplikatentferner

SYMBOL: δ (Delta)
DEFINITION: Der Duplikatenentferner δ entfernt alle doppelten Tupel in einer Relation.

ERWEITERTE OPERATIONEN

Duplikatentferner allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{array}$$

Ergebnis:

$$\delta(R) = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{array}$$

ERWEITERTE OPERATIONEN

Duplikatentferner einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 5 & 2 \end{array}$$

Ergebnis:

$$\delta(R) = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \end{array}$$

SORTIERUNG

ERWEITERTE OPERATIONEN

Sortieren

SYMBOL: τ (Tau)

DEFINITION: Der Ausdruck $\tau_L(R)$, wobei R eine Relation und L eine Liste von Attributen von R ist, sortiert die Tupel in der Relation R nach der Ordnung welche durch L angegeben werden. Wenn L die Liste $A_1, A_2, ..., A_n$ ist, dann werden die Tupel von R als erstes nach ihrem Werten im Attribute A_1 geordnet. Wenn verschiedene Tupel den gleichen Wert haben, wird nach dem Wert in Attribute A_2 sortiert und dann immer so weiter. Wenn ein Tupel auch im Attribute A_n gleich ist, kann es beliebig sortiert werden.

In der relationalen Algebra wird immer aufwärts sortiert.

ERWEITERTE OPERATIONEN

Sortieren einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 5 & 2 \end{array}$$

Ergebnis:

$$\tau_a(R) = \begin{array}{cc} a & b \\ \hline 2 & 4 \\ 5 & 2 \\ 5 & 2 \end{array}$$

ERWEITERTE OPERATIONEN

Sortieren schwieriger

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{array}$$

Result:

$$\tau_{a,b}(R) = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{array}$$

KOMBINATIONSOPERATIONEN 1

KOMBINATIONSOPERATOREN

Allgemein

Die Kombinationsoperatoren verbinden verschiedene Relationen miteinander. Wir unterscheiden zwischen:

→ Kreuzprodukt bzw. Kartesische Produkt (Cartesian Product)

→ Verbund (Natural Join)

→ und einer Vielzahl von kombinierten Operatoren wie:

- → Theta Join
- → Semi Join
- → Antijoin

KOMBINATIONSOPERATOREN

Kreuzprodukt

SYMBOL: $\times$

DEFINITION: Das Kreuzprodukt von zwei Relationen R und S ($R \times S$) ist die Menge von Paaren welche durch die Kombination der Tupel der beiden Relationen entsteht. Das Kreuzprodukt wird oft auch Kartesisches Produkt genannt. Das Kreuzprodukt verknüpft das erste Tupel aus R mit dem ersten Tupel aus S, dann mit dem zweiten bis hin zum letzten aus S. Dieses Verfahren wiederholt sich für jedes weitere Tupel aus R.

MATHEMATISCHE BESCHREIBUNG: $R \times S = \{qt|$ q $\in$ R and t $\in S\}$

KOMBINATIONSOPERATOREN

Kreuzprodukt

Wenn Relationen miteinander verbunden werden, welche die gleichen Attributsnamen haben, muss der Relationsname vor dem Attribut angegeben werden, damit die Attributsnamen eindeutig sind.

Eigenschaften Kreuzprodukt

Die Kardinalität der Ergebnisrelation ist das Produkt der jeweiligen Kardinalitäten.
Der Grad der Ergebnisrelation ist die Summe der jeweiligen Grade der Ausgangsrelationen.

KOMBINATIONSOPERATOREN

Kreuzprodukt

A	B	C
a	b	c
d	e	f

D	E	F
g	h	i
j	k	l

A	B	C	D	E	F
a	b	c	g	h	i
a	b	c	j	k	l
d	e	f	j	k	l
d	e	f	g	h	i

Abbildung: $R \times S$

KOMBINATIONSOPERATOREN

Kreuzprodukt

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{array} \qquad S = \begin{array}{ccc} d & e & f \\ \hline d_1 & e_1 & f_1 \\ d_2 & e_2 & f_2 \end{array}$$

Ergebnis:

$$R \times S = \begin{array}{cccccc} a & b & c & d & e & f \\ \hline a_1 & b_1 & c_1 & d_1 & e_1 & f_1 \\ a_1 & b_1 & c_1 & d_2 & e_2 & f_2 \\ a_2 & b_2 & c_2 & d_1 & e_1 & f_1 \\ a_2 & b_2 & c_2 & d_2 & e_2 & f_2 \end{array}$$

KOMBINATIONSOPERATOREN

Kreuzprodukt - einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} c & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \times S = \begin{array}{cccc} a & b & c & d \\ \hline 5 & 2 & 4 & 1 \\ 5 & 2 & 5 & 2 \\ 2 & 4 & 4 & 1 \\ 2 & 4 & 5 & 2 \\ 3 & 5 & 4 & 1 \\ 3 & 5 & 5 & 2 \end{array}$$

KOMBINATIONSOPERATOREN

Kreuzprodukt - schwieriger

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \times S = \begin{array}{cccc} R.a & b & S.a & d \\ \hline 1 & 2 & 4 & 1 \\ 1 & 2 & 5 & 2 \\ 1 & 3 & 4 & 1 \\ 1 & 3 & 5 & 2 \\ 1 & 4 & 4 & 1 \\ 1 & 4 & 5 & 2 \end{array}$$

KOMBINATIONSOPERATOREN

Verbund

SYMBOL: $\bowtie$

DEFINITION: Aus den beiden Relationen R und S wird bei einem Verbund ($R \bowtie S$) eine Relation erzeugt, die aus allen Kombinationen der Tupel der beiden Relationen besteht. Es werden jedoch nur diejenigen Tupel aufgelistet, die in jeweils einem Attribut der beiden Relationen gleiche Werte besitzen. Dieses vergleichende Attribut wird nur einmal ausgegeben. Man nennt den Verbund auch Natural Join.

KOMBINATIONSOPERATOREN

Verbund

Das Ergebnis von $R \bowtie S$ enthält nur die Paare, welche bei gleichnamigen Attributen die gleichen Werte haben.
Sei $A_1, A_2, ..., A_n$ die Menge alle Attribute, welche sowohl in der Relation R, als auch in S auftreten, dann wird das Tupel r von R und das Tupel s aus S erfolgreich miteinander verbunden, wenn und nur wenn r und s die gleichen Werte in $A_1, A_2, ..., A_n$ haben.

KOMBINATIONSOPERATOREN

Verbund

A	B	C
a	b	c
d	e	f

A	E	F
a	h	i
a	k	l

A	B	C	E	F
a	b	c	h	i
a	b	c	k	l

Abbildung: $R \bowtie S$

KOMBINATIONSOPERATOREN

Verbund - allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \\ a_4 & e_4 & f_4 \end{array}$$

Ergebnis:

$$R \bowtie S = \begin{array}{ccccc} a & b & c & e & f \\ \hline a_1 & b_1 & c_1 & e_1 & f_1 \\ a_2 & b_2 & c_2 & e_2 & f_2 \\ a_3 & b_3 & c_3 & e_3 & f_3 \\ a_4 & b_4 & c_4 & e_4 & f_4 \end{array}$$

KOMBINATIONSOPERATOREN

Verbund - einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \bowtie S = \begin{array}{ccc} a & b & d \\ \hline 5 & 2 & 2 \end{array}$$

KOMBINATIONSOPERATOREN

Verbund - schwieriger

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 5 & 3 \end{array} \qquad S = \begin{array}{ccc} b & d & a \\ \hline 2 & 2 & 1 \\ 2 & 4 & 2 \\ 2 & 5 & 1 \end{array}$$

Ergebnis:

$$R \bowtie S = \begin{array}{cccc} a & b & c & d \\ \hline 1 & 2 & 3 & 2 \\ 1 & 2 & 3 & 5 \end{array}$$

KOMBINATIONSOPERATOREN

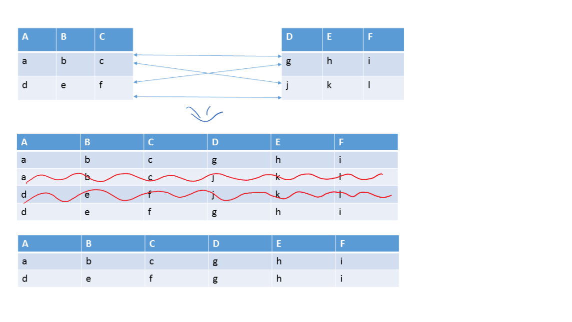
Theta Join

SYMBOL: $\bowtie_\theta$

DEFINITION: Der Theta Join von zwei Relationen R und S ($R \bowtie_\theta S$) sind die Tupel, welche nach einem Kreuzprodukt beider Relationen, eine Bedingung θ erfüllen. Mathematisch bedeutet dies: $R \bowtie_\theta S = \sigma_\theta(R \times S)$

KOMBINATIONSOPERATOREN

Theta Join



A	B	C
a	b	c
d	e	f

D	E	F
g	h	i
j	k	l

A	B	C	D	E	F
a	b	c	g	h	i
~~a~~	~~b~~	~~c~~	~~j~~	~~k~~	~~l~~
~~d~~	~~e~~	~~f~~	~~j~~	~~k~~	~~l~~
d	e	f	g	h	i

A	B	C	D	E	F
a	b	c	g	h	i
d	e	f	g	h	i

Abbildung: $R \bowtie_{d=“g“} S$

KOMBINATIONSOPERATOREN

Theta Join allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{array} \qquad S = \begin{array}{ccc} d & e & f \\ \hline d_1 & e_1 & f_1 \\ d_2 & e_2 & f_2 \\ d_3 & e_3 & f_3 \end{array}$$

Ergebnis:

$$R \bowtie_{a=“a_1“ \vee e=“e_2“} S = \begin{array}{cccccc} a & b & c & d & e & f \\ \hline a_1 & b_1 & c_1 & d_1 & e_1 & f_1 \\ a_1 & b_1 & c_1 & d_2 & e_2 & f_2 \\ a_1 & b_1 & c_1 & d_3 & e_3 & f_3 \\ a_2 & b_2 & c_2 & d_2 & e_2 & f_2 \\ a_3 & b_3 & c_3 & d_2 & e_2 & f_2 \end{array}$$

KOMBINATIONSOPERATOREN

Theta Join einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} c & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \bowtie_{a=5} S = \begin{array}{cccc} a & b & c & d \\ \hline 5 & 2 & 4 & 1 \\ 5 & 2 & 5 & 2 \end{array}$$

KOMBINATIONSOPERATOREN

Theta Join schwieriger

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \bowtie_{S.a=5} S = \begin{array}{cccc} R.a & b & S.a & d \\ \hline 1 & 2 & 5 & 2 \\ 1 & 3 & 5 & 2 \\ 1 & 4 & 5 & 2 \end{array}$$

KOMPLEXE ABFRAGEN

KOMPLEXE ALGEBRA

Und nun bringen wir alles zusammen

Bis jetzt haben wir die Operatoren nur einzeln betrachtet, aber natürlich ist es auch möglich diese zu kombinieren in einer Abfrage.

Viele Operatoren

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array}$$

Ergebnis:

$$\pi_a(\sigma_{b=2}(R)) = \begin{array}{c} a \\ \hline 5 \end{array}$$

KOMPLEXE ALGEBRA

Viele Operatoren

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} b & c \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$\rho_{T(a,b,c)}(\pi_{R.b,S.b,c}(R \times S)) \cup (R \bowtie S) = \begin{array}{ccc} a & b & c \\ \hline 2 & 4 & 1 \\ 2 & 5 & 2 \\ 4 & 4 & 1 \\ 4 & 5 & 2 \\ 5 & 4 & 1 \end{array}$$

KOMPLEXE ALGEBRA

Hohe Komplexität

Je mehr Operatoren in einer Abfrage verwendet werden, um so schwieriger ist sie zu verstehen. Um Missverständnisse vorzubeugen, gibt es zwei verschieden Presentationsmöglichkeiten:

1. Expression Tree: Eine visuelle Präsentation einer Abfrage
2. Lineare Notation: Eine strukturierte Art und Weise eine Abfrage darzustellen

KOMPLEXE ALGEBRA

Expression Tree

Expressions Trees werden vom Boden aus bearbeitet. Dabei wird der Operatorknoten an seinen Kindern (Ursprungsrelationen bzw. Zwischenrelationen) ausgeführt. Durch die Abarbeitung von untern nach oben ist immer sichergestellt, dass die entsprechenden Relationen zur Verfügung stehen, wenn wir sie benötigen.

KOMPLEXE ALGEBRA

Expression Tree

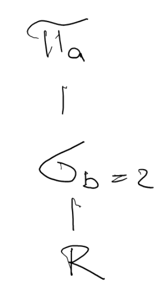


Abbildung: Beispiel der Seite 90

KOMPLEXE ALGEBRA

Expression Tree

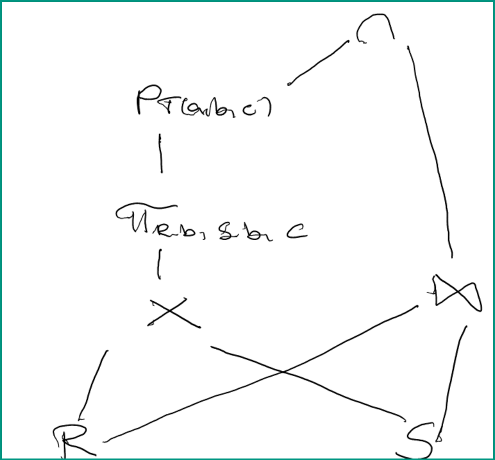


Abbildung: Beispiel der Seite 91

KOMPLEXE ALGEBRA

Lineare Notation

Die Lineare Notation hat folgende Regeln:

1. Jede Zeile beginnt mit einem Relationsname und dahinter in Klammern die Attribute der Relation. Die Domain muss nicht genannt werden. Der Name der finalen Relation ist Antwort.
2. Nach dem Relationsname kommt das Anweisungssymbol :=
3. Auf der rechten Seite des Anweisungssymbol befindet sich die eigentliche Abfrage. Dabei soll aber pro Zeile nur ein Operator benutzt werden.
4. Die letzte Relation heißt Answer oder Result und muss nicht (aber kann) die Attribute enthalten.

KOMPLEXE ALGEBRA

Lineare Notation

gegeben:

$$\pi_a(\sigma_{b=2}(R))$$

Ergebnis:

$$A(a,b) := \sigma_{b=2}(R)$$
$$Answer(a) := \pi_a(A)$$

KOMPLEXE ALGEBRA

Lineare Notation

gegeben:

$$\rho_{T(a,b,c)}(\pi_{R.b,S.b,c}(R \times S)) \cup (R \bowtie S)$$

Ergebnis:

$$\begin{aligned} A(a, R.b, S.b, c) &:= R \times S \\ B(R.b, S.b, c) &:= \pi_{R.b,S.b,c}(A) \\ T(a, b, c) &:= \rho_{T(a,b,c)}(B) \\ C(a, b, c) &:= R \bowtie S \\ Answer(a, b, c) &:= T \cup C \end{aligned}$$

KOMBINATIONSOPERATIONEN 2

KOMBINATIONSOPERATOREN

Links-Semi-Join

SYMBOL: $\ltimes$

DEFINITION: Das Ergebnis des Links-Semi-Joins von zwei Relationen R and S ($R \ltimes S$) ist die Menge von Tupeln, welche in R sind und für die es in S ein Tupel gibt, welche den gleichen Wert in den gemeinsamen Attributen hat. Mathematisch bedeutet dies: $R \ltimes S = \pi_{alle-Attribte-von-R}(R \bowtie S)$. Es wird also ein Verbund von den beiden Relationen gebildet und dann nur die Attribute von R projeziert.

KOMBINATIONSOPERATOREN

Links-Semi-Join allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \\ a_5 & e_3 & f_3 \end{array}$$

Ergebnis:

$$R \ltimes S = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{array}$$

KOMBINATIONSOPERATOREN

Links-Semi-Join einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \ltimes S = \begin{array}{cc} a & b \\ \hline 5 & 2 \end{array}$$

KOMBINATIONSOPERATOREN

Rechts-Semi-Join

SYMBOL: $\rtimes$

DEFINITION: Das Ergebnis des Rechts-Semi-Joins von zwei Relationen R and S ($R \rtimes S$) ist die Menge von Tupeln, welche in S sind und für die es in R ein Tupel gibt, welche den gleichen Wert in den gemeinsamen Attributen hat. Mathematisch bedeutet dies: $R \rtimes S = \pi_{alle-Attribte-von-S}(R \bowtie S)$. Es wird also ein Verbund von den beiden Relationen gebildet und dann nur die Attribute von S projeziert.

KOMBINATIONSOPERATOREN

Rechts-Semi-Join

A	B	C
a	b	c
d	e	f

A	E	F
a	h	i
a	k	l
x	y	z

A	E	F
a	h	i
a	k	l

Abbildung: $R \rtimes S$

KOMBINATIONSOPERATOREN

Rechts-Semi-Join allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \\ a_5 & e_5 & f_5 \end{array}$$

Ergebnis:

$$R \rtimes S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \end{array}$$

KOMBINATIONSOPERATOREN

Rechts-Semi-Join einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \rtimes S = \begin{array}{cc} a & d \\ \hline 5 & 2 \end{array}$$

KOMBINATIONSOPERATOREN

Rechts-Anti-Join

SYMBOL: $\triangleleft$

DEFINITION: Das Ergebnis des Rechts-Anti-Joins von zwei Relationen R and S ($R \triangleleft S$) ist die Menge von Tupeln, welche in S sind und für die es in R KEIN Tupel gibt, welche den gleichen Wert in den gemeinsamen Attributen hat. Mathematisch bedeutet dies: $R \triangleleft S = S - \pi_{alle-Attribte-von-S}(R \bowtie S)$. Es wird die Differenz der Ausgangsrelation S und des Rechts-Semi-Joins berechnet.

KOMBINATIONSOPERATOREN

Rechts-Anti-Join

A	B	C
a	b	c
d	e	f

A	E	F
a	h	i
a	k	l
x	y	z

A	E	F
x	y	z

Abbildung: $R \triangleleft S$

KOMBINATIONSOPERATOREN

Rechts-Anti-Join allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \\ a_5 & e_5 & f_5 \end{array}$$

Ergebnis:

$$R \triangleleft S = \begin{array}{ccc} a & e & f \\ \hline a_5 & e_5 & f_5 \end{array}$$

KOMBINATIONSOPERATOREN

Rechts-Anti-Join einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \lhd S = \begin{array}{cc} a & d \\ \hline 4 & 1 \end{array}$$

KOMBINATIONSOPERATOREN

Links-Anti-Join

SYMBOL: $\triangleright$

DEFINITION: Das Ergebnis des Links-Anti-Joins von zwei Relationen R and S ($R \triangleright S$) ist die Menge von Tupeln, welche in R sind und für die es in S KEIN Tupel gibt, welche den gleichen Wert in den gemeinsamen Attributen hat. Mathematisch bedeutet dies: $R \triangleright S = R - \pi_{alle-Attribte-von-R}(R \bowtie S)$. Es wird die Differenz der Ausgangsrelation R und des Links-Semi-Joins berechnet.

KOMBINATIONSOPERATOREN

Links-Anti-Join

A	B	C
a	b	c
d	e	f

A	E	F
a	h	i
a	k	l
x	y	z

A	B	C
d	e	f

Abbildung: $R \triangleright S$

KOMBINATIONSOPERATOREN

Links-Anti-Join allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \\ a_5 & e_5 & f_5 \end{array}$$

Ergebnis:

$$R \triangleright S = \begin{array}{ccc} a & b & c \\ \hline a_4 & b_4 & c_4 \end{array}$$

KOMBINATIONSOPERATOREN

Links-Anti-Join einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \triangleright S = \begin{array}{cc} a & b \\ \hline 2 & 4 \\ 3 & 5 \end{array}$$

GRUPPIEREN

ERWEITERTE OPERATOREN

Aggregation

DEFINITION: Diese Operatoren werden verwendet, um die Werte in einer Spalte zu summieren oder aggregieren. Die Standard-Operatoren sind:

ERWEITERTE OPERATOREN

Aggregation

1. SUM produziert die Summe der Spaltenwerte mit nummerischen Werten.
2. AVG berechnet den nummerischen Durchschnitt der Spalte. Nur anwendbar wenn die Spalte nummerische Werte hat.
3. MIN und MAX, angewendet bei nummerischen Werten, gibt die kleinsten oder größten Werte einer Spalte aus. Wenn die Spalten strings enthält, ist die Ausgabe der alphabetisch erste oder letzte Wert.
4. COUNT ist der einzige Aggratations Operator, welche sowohl auf Spalten, als auch Relationen angewendet werden kann. Die Ausgabe ist die Anzahl der Zeilen der Relation/Spalte.

ERWEITERTE OPERATOREN

Aggregation

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 5 & 2 \end{array}$$

Ergebnis:

1. SUM(R.b) = 8
2. AVG(R.a) = 4
3. MIN(R.b) = 2
4. MAX(R.a) = 5
5. COUNT(R.a) = COUNT(R.b) = COUNT(R) = 3

ERWEITERTE OPERATIONEN

Gruppierung

SYMBOL: γ (Gamma)

DEFINITION: Die meiste Zeit wollen wir Aggregationen nicht auf die ganze Spalte anwenden, sondern wir benötigen Informationen über Tupel die in einem oder mehreren Werten gleich sind. Wir müssen Gruppen nach Attributwerten bilden und in diesen Gruppen Aggregationen anwenden.

Index vom γ Operator ist eine Liste L von Elementen, wobei jedes Element nur eines der folgenden Eigenschaften besitzt:

ERWEITERTE OPERATIONEN

Gruppierung

1. Ein Attribut ohne eine Aggregation, ist ein Gruppierungs-Attribute. Nach diesen Attributen werden die Tupel gruppiert.
2. Ein Attribute mit einer Aggregation gibt an, was in den jeweiligen Gruppen berechnet werden soll. Wenn hinter der Aggregation ein Pfeil ist, so ist der Wert dahinter, der neue Name des Berechungsattributs.

ERWEITERTE OPERATIONEN

Gruppierung

Die Ergebnisrelation des Ausdrucks $\gamma_L(R)$ ist folgendermaßen aufgebaut:

1. Als erstes werden die Tupel in Gruppen eingeteilt. Jedes Tupel in einer Gruppen hat einen bestimmten Wert in den Gruppierungs-Attributen der Liste L. Wenn es keine Gruppierungs-Attribute gibt, ist die ganze Relation in einer Gruppe.
2. Für jede Gruppe, wird ein Ergebnistupel geworden, welche folgende Teile enthält:
   2.1 Die Gruppierungs-Attribut-Werte für die Gruppe und
   2.2 die Aggregationen aus der Liste L, welche über alle Tupel der Gruppe berechnet werden.

ERWEITERTE OPERATIONEN

Gruppierung einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 5 & 2 \end{array}$$

Ergebnis:

$$\gamma_{a,sum(b)}(R) = \begin{array}{cc} a & b \\ \hline 2 & 4 \\ 5 & 4 \end{array}$$

ERWEITERTE OPERATIONEN

Gruppierung schwieriger

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 5 & 3 \\ 1 & 5 & 3 \\ 2 & 4 & 5 \end{array}$$

Ergebnis:

$$\gamma_{b,c,MAX(a)\to d}(R) = \begin{array}{ccc} b & c & d \\ \hline 2 & 3 & 1 \\ 4 & 5 & 2 \\ 5 & 3 & 1 \end{array}$$

OUTERJOIN

ERWEITERTE OPERATIONEN

Links Outer Join

Symbol: $\ltimes_o$

DEFINITION: Das Ergebnis des Links Outer Joins von den zwei Relationen R und S ($R \ltimes_o S$) ist das Ergebnis des Verbundes von R und S ($R \bowtie S$) und alle Tupel von R für die es in der Relation S kein Tupel mit den gleichen Werten in den gemeinsamen Attributen gibt.

Die Tupel, welche nicht mit einem Tupel aus der anderen Relation kombiniert werden können, nennt man dangling Tupels (hängende Tupel). Die Werte für die nicht-gemeinsamen Attribute der rechten Relation werden mit NULL oder $\perp$ aufgefüllt.

ERWEITERTE OPERATIONEN

Links Outer Join

A	B	C
a	b	c
d	e	f

A	E	F
a	h	i
a	k	l
x	y	z

A	B	C	E	F
a	b	c	h	i
a	b	c	k	l
d	e	f	NULL	NULL
~~x~~	~~NULL~~	~~NULL~~	~~y~~	~~z~~

A	B	C	E	F
a	b	c	h	i
a	b	c	k	l
d	e	f	NULL	NULL

Abbildung: $R \ltimes_o S$

ERWEITERTE OPERATIONEN

Links Outer Join allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \\ a_5 & e_5 & f_5 \end{array}$$

Ergebnis:

$$R \ltimes_o S = \begin{array}{ccccc} a & b & c & e & f \\ \hline a_1 & b_1 & c_1 & e_1 & f_1 \\ a_2 & b_2 & c_2 & e_2 & f_2 \\ a_3 & b_3 & c_3 & e_3 & f_3 \\ a_4 & b_4 & c_4 & \perp & \perp \end{array}$$

ERWEITERTE OPERATIONEN

Links Outer Join einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \ltimes_o S = \begin{array}{ccc} a & b & d \\ \hline 5 & 2 & 2 \\ 2 & 4 & \perp \\ 3 & 5 & \perp \end{array}$$

ERWEITERTE OPERATIONEN

Rechts Outer Join

Symbol: $\rtimes_o$

DEFINITION: Das Ergebnis des Links Outer Joins von den zwei Relationen R und S ($R \rtimes_o S$) ist das Ergebnis des Verbundes von R und S ($R \bowtie S$) und alle Tupel von S für die es in der Relation R kein Tupel mit den gleichen Werten in den gemeinsamen Attributen gibt.

Die Tupel, welche nicht mit einem Tupel aus der anderen Relation kombiniert werden können, nennt man dangling Tupels (hängende Tupel). Die Werte für die nicht-gemeinsamen Attribute der linken Relation werden mit NULL oder $\perp$ aufgefüllt.

ERWEITERTE OPERATIONEN

Rechts Outer Join

A	B	C
a	b	c
d	e	f

A	E	F
a	h	i
a	k	l
x	y	z

A	B	C	E	F
a	b	c	h	i
a	b	c	k	l
~~d~~	~~e~~	~~f~~	~~NULL~~	~~NULL~~
x	NULL	NULL	y	z

A	B	C	E	F
a	b	c	h	i
a	b	c	k	l
x	NULL	NULL	y	z

Abbildung: $R \rtimes_o S$

ERWEITERTE OPERATIONEN

Rechts Outer Join allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \\ a_5 & e_5 & f_5 \end{array}$$

Ergebnis:

$$R \rtimes_o S = \begin{array}{ccccc} a & b & c & e & f \\ \hline a_1 & b_1 & c_1 & e_1 & f_1 \\ a_2 & b_2 & c_2 & e_2 & f_2 \\ a_3 & b_3 & c_3 & e_3 & f_3 \\ a_5 & \perp & \perp & e_5 & f_5 \end{array}$$

ERWEITERTE OPERATIONEN

Rechts Outer Join einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \rtimes_o S = \begin{array}{ccc} a & b & d \\ \hline 5 & 2 & 2 \\ 4 & \perp & 1 \end{array}$$

ERWEITERTE OPERATIONEN

Full Outer Join

Symbol: $⟗_o$

DEFINITION: Das Ergebnis des Full Outer Join der zwei Relationen R und S ($R ⟗_o S$) kombiniert die Ergebnisse des Links und Rechts Outer Joins.

Das Full Outer Join von R und S ist ähnlich zu dem Verbund von R und S (R Join S). Zu den komibierten Tupels werden noch alle Tupels von R und S hinzugefügt, welche nicht mit einem Tupel aus der anderen Relation kombiniert werden können. Der nicht bekannte Wert bei den nicht komibinierten Tupeln wird mit NULL or $\perp$ belegt.

ERWEITERTE OPERATIONEN

Full Outer Join

A	B	C
a	b	c
d	e	f

A	E	F
a	h	i
a	k	l
x	y	z

A	B	C	E	F
a	b	c	h	i
a	b	c	k	l
d	e	f	NULL	NULL
x	NULL	NULL	y	z

Abbildung: $R \bowtie_o S$

ERWEITERTE OPERATIONEN

Full Outer Join allgemein

gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{array} \qquad S = \begin{array}{ccc} a & e & f \\ \hline a_1 & e_1 & f_1 \\ a_2 & e_2 & f_2 \\ a_3 & e_3 & f_3 \\ a_5 & e_5 & f_5 \end{array}$$

Ergebnis:

$$R \bowtie_o S = \begin{array}{ccccc} a & b & c & e & f \\ \hline a_1 & b_1 & c_1 & e_1 & f_1 \\ a_2 & b_2 & c_2 & e_2 & f_2 \\ a_3 & b_3 & c_3 & e_3 & f_3 \\ a_4 & b_4 & c_4 & \perp & \perp \\ a_5 & \perp & \perp & e_5 & f_5 \end{array}$$

ERWEITERTE OPERATIONEN

Full Outer Join einfach

gegeben:

$$R = \begin{array}{cc} a & b \\ \hline 5 & 2 \\ 2 & 4 \\ 3 & 5 \end{array} \qquad S = \begin{array}{cc} a & d \\ \hline 4 & 1 \\ 5 & 2 \end{array}$$

Ergebnis:

$$R \bowtie_o S = \begin{array}{ccc} a & b & d \\ \hline 5 & 2 & 2 \\ 2 & 4 & \perp \\ 3 & 5 & \perp \\ 4 & \perp & 1 \end{array}$$

NULL VALUE

NULL VALUE

Was bedeutet NULL?

→ die Abwesenheit eines Wertes, weil keiner existiert

→ die Abwesenheit, da man den Wert (noch) nicht kennt

Der Nullwert ist verschieden von der Zahl 0, da diese einen Wert, also eine Information repräsentiert. Sprachlich wird der Nullwert von der algebraischen Zahl 0 durch die Aussprache unterschieden, indem es [’nʌl] gesprochen wird, um Missverständnisse zu vermeiden

NULL VALUE

Der Wahrheitswert Unbekannt

Bis jetzt haben wir angenommen, dass das Ergebnis eines Vergleiches entweder Wahr oder Falsch ist, und wir können diese Wahrheitswerte miteinander kombinieren (siehe Folie 49).

Wir haben jetzt gesehen, dass es einen NULL wer gibt. Damit kann das Ergebnis des Vergleichs einen dritten Wahrheitswert annehmen: UNBEKANNT.

NULL VALUE

Wahrheitstabelle bei Drei-Wahrheitswerten

Die Regel ist ziemlich einfach:

→ Wahr ist 1

→ Falsch ist 0

→ Unbekannt ist $\frac{1}{2}$ (Irgendwas zwischen 1 und 0)

Bei den Kombinationen von zwei Werten gilt das Folgende:

→ UND - das Mininum der zwei Werte ist das Ergebnis

→ ODER - das Maximum der zwei Werte ist das Ergebnis

→ NEGIERUNG - Wenn der Wert v ist, ist das Ergebnis 1-v

WAHRHEITSTABELLE BEI DREI-WAHRHEITSWERTEN

UND $\wedge$, ODER $\vee$ und Negation $\neg$

A	B	$A \wedge B$	$A \vee B$	$\neg A$
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	0
0	$\frac{1}{2}$	0	$\frac{1}{2}$	1
$\frac{1}{2}$	0	0	$\frac{1}{2}$	$\frac{1}{2}$
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

TEILUNGSOPERATION

TEILUNGSOPERATOR

Division

Symbol: $\div$

DEFINITION: Die Division setzt voraus, dass die Relation R mindestens alle Attribute (att) der Relation S enthält. Das Ergebnis von (R $\div$ S) sind alle Tupel aus R, deren Werte in den Attributen att mit den Werten aus S übereinstimmen, welche jedoch nur die Attribute außer att enthält. Anders betrachtet, so wie die Mulitiplikation zur Division im arithemitschen Algebra ist, ist das Kreuzprodukt($\times$) zur Division ($\div$) in der relationalen Algebra.

TEILUNGSOPERATOR

Division

Zur Berechung von $R \div S$ müssen folgeden Schritte ausgeführt werden:

1. Berechnung aller möglichen Paarungen
2. Entfernung der bestehenden Paare
3. Entfernen der Nicht-Antworten von den möglichen Antworten

TEILUNGSOPERATOR

Division und Kreuzprodukt

Kreuzprodukt:
gegeben:

$$T = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 3 & 4 \end{array} \qquad S = \begin{array}{c} c \\ \hline 5 \\ 6 \end{array}$$

Ergebnis:

$$R = S \times T = \begin{array}{ccc} a & b & c \\ \hline 1 & 2 & 5 \\ 1 & 2 & 6 \\ 3 & 4 & 5 \\ 3 & 4 & 6 \end{array}$$

TEILUNGSOPERATOR

Division und Kreuzprodukt

Division:
gegeben:

$$R = \begin{array}{ccc} a & b & c \\ \hline 1 & 2 & 5 \\ 1 & 2 & 6 \\ 3 & 4 & 5 \\ 3 & 4 & 6 \end{array} \qquad S = \begin{array}{c} c \\ \hline 5 \\ 6 \end{array}$$

Ergebnis:

$$T = R \div S = \begin{array}{cc} a & b \\ \hline 1 & 2 \\ 3 & 4 \end{array}$$

TEILUNGSOPERATOR

Divisionsbeispiel - 1

Aufgabe:Welche Person (SSN) hat in jeder Bank des Landes ein Konto. Oder anders gesagt: Wenn es eine Bank im Land X gibt, dann hat die Person ein Konto dort.

gegeben:

$$person = \begin{array}{c} SSN \\ \hline 123 \\ 467 \\ 896 \\ 563 \end{array} \quad accountInfo = \begin{array}{ccc} SSN & AccNo & bank \\ \hline 123 & 1111 & BoA \\ 467 & 1222 & BoA \\ 123 & 333 & Chase \\ 896 & 444 & Chase \\ 123 & 555 & USB \\ 467 & 666 & USB \end{array}$$

Gesucht: accountInfo $\div$ person

TEILUNGSOPERATOR

Divisionsbeispiel - 2

Idee: Finde alle Werte die NICHT zur Antwort gehören und entferne sie von der Liste der möglichen Antworten.

Schritt 1: Berechne ein Kreuzprodukt von jeder möglichen Personen und jeder möglichen Bank.

Schritt 2: Entferne alle bestehenden (wahren) Daten. Das Ergebnis ist eine Tabelle mit allen unwahren Kombinationen.

Schritt 3: Entferne alle Personen, welche in der Ergebnistabelle von Schritt 2 auftauchen. Diese können nicht im Ergebnis enthalten sein. Die restlichen Personen, welche NICHT in der Ergebnistabelle von Schritt 2 auftauchen, sind die gesuchten Personen.

EIGENSCHAFTEN DER RELATIONALEN OPERATOREN

EIGENSCHAFTEN

Kommutativgesetz

$R \cup S = S \cup R$

$R \cap S = S \cap R$

$R \times S = S \times R$

$R \bowtie S = S \bowtie R$

EIGENSCHAFTEN

Assoziativgesetz

$R \cup (S \cup T) = (R \cup S) \cup T$
$R \cap (S \cap T) = (R \cap S) \cap T$
$R \times (S \times T) = (R \times S) \times T$