4 Relationenalgebra

4.1 Relationen

- Der Begriff der Relation kann mathematisch präzise definiert werden.
- Für uns genügt es zu wissen: *Relation* = *Tabelle*
- Wir werden sehen, wie man mit Relationen, d. h. mit Tabellen sinnvoll operieren kann.

Tabellen (Repetition)

Merkmal oder Attribut

Tabellenname

MITARBEITER

Schlüsselmerkmal

ID	Name	Ort
5	Schweizer	Fribourg
64	Becker	Basel
28	Meier	Bern
13	Huber	Basel

Datensatz, Zeile, Tupel

Spalte

Datenwert

4.2 Mengenorientierte Operatoren

Übersicht

- Vereinigung von Tabellen ($R \cup S$)
- Durchschnitt von Tabellen ($R \cap S$)
- Differenz von Tabellen ($R \setminus S$)
- Kartesisches Produkt von Tabellen ($R \times S$)

Für die ersten drei Operatoren müssen die Tabellen vereinigungsverträglich sein.

Zwei Tabellen sind *vereinigungsverträglich*, wenn sie folgende Eigenschaften haben.

- Beide Tabellen haben die gleiche *Anzahl* Merkmale.
- Die Datentypen korrespondierender Spalten sind identisch.

Aufgabe

Sind die folgenden Tabellen vereinigungsverträglich?

WAHLPFLICHTFACH

WPF_ID	Titel	Lehrperson
41	Geschichte der Neuzeit	A. Meier
65	Die Mathematik der Antike	L. Euler
19	Volleyball II	S. Portmann

ABENDKURS

Kurs_ID	Titel	Kursleitung
108	Word für Fortgeschrittene	R. Ratlos
459	Metallbearbeitung	H. Phaistos
98	Advanced	T. Beutel

Ja, denn sie haben gleich viele Spalten und stimmen in den Datentypen entsprechender Spalten überein

4.2.1 Der Vereinigungsoperator (union)

Zwei vereinigungsverträgliche Tabellen R und S werden mengentheoretisch vereinigt, indem sämtliche Einträge aus R und sämtliche Einträge aus S in die Resultattabelle eingefügt werden. Gleichzeitig werden identische Datensätze eliminiert.

ADRESSEN1

ID	Name	Vorname
12	Meier	Andreas
47	Müller	Michael
39	Schmid	Susanne

ADRESSEN2

ID	Name	Vorname
65	Bischof	Maria
39	Schmid	Susanne

ADRESSEN3 = ADRESSEN1 $\cup$ ADRESSEN2

ID	Name	Vorname
12	Meier	Andreas
47	Müller	Michael
39	Schmid	Susanne
65	Bischof	Maria

4.2.2 Der Durchschnittsoperator (intersection)

Zwei vereinigungsverträgliche Tabellen R und S werden geschnitten, indem sämtliche Einträge, die sowohl in R als auch in S vorhanden sind, in die Resultattabelle aufgenommen werden.

ADRESSEN1

ID	Name	Vorname
12	Meier	Andreas
47	Müller	Michael
39	Schmid	Susanne

ADRESSEN2

ID	Name	Vorname
65	Bischof	Maria
39	Schmid	Susanne

ADRESSEN3 = ADRESSEN1 $\cap$ ADRESSEN2

ID	Name	Vorname
39	Schmid	Susanne

4.2.3 Der Subtraktionsoperator \ (difference)

Sind R und S zwei vereinigungsverträgliche Tabellen, so wird die Differenz $R \setminus S$ gebildet, indem man aus R sämtliche Einträge entfernt, die in S enthalten sind.

ADRESSEN1

ID	Name	Vorname
12	Meier	Andreas
47	Müller	Michael
39	Schmid	Susanne

ADRESSEN2

ID	Name	Vorname
65	Bischof	Maria
39	Schmid	Susanne

ADRESSEN3 = ADRESSEN1 \ ADRESSEN2

ID	Name	Vorname
12	Meier	Andreas
47	Müller	Michael

4.2.4 Das kartesische Produkt

- Unter dem kartesischen Produkt $R \times S$ versteht man die Menge aller möglichen Kombinationen aus Tupeln aus R mit Tupeln aus S.
- Für das kartesische Produkt müssen die betrachteten Tabellen *nicht* vereinigungsverträglich sein.

ADRESSEN

ID	Name	Vorname
12	Meier	Andreas
47	Müller	Michael
39	Schmid	Susanne

HOBBY

Hobby_ID	Bezeichnung
1	Wandern
2	Schwimmen

FREIZEIT = ADRESSEN × HOBBY

ID	Name	Vorname	*Hobby_ID*	Bezeichnung
12	Meier	Andreas	1	Wandern
12	Meier	Andreas	2	Schwimmen
47	Müller	Michael	1	Wandern
47	Müller	Michael	2	Schwimmen
39	Schmid	Susanne	1	Wandern
39	Schmid	Susanne	2	Schwimmen

4.3 Die relationenorientierten Operatoren

Übersicht

- Die relationenorientierten Operatoren ergänzen die mengenorientierten Operatoren

- Wie beim kartesischen Produkt werden *keine* vereinigungsverträglichen Tabellen vorausgesetzt.

Diese Operatoren werden nachfolgende besprochen:

- Die Projektion einer Tabelle R auf ein Merkmal M
- Die Selektion einer Spalte aus einer Tabelle R anhand der Formel F
- Der Verbund zweier Tabellen R und S über das Prädikat P

Ein *Prädikat* ist eine Funktion, die einen Wahrheitswert (*wahr* oder *falsch*) zurückliefert.

4.3.1 Der Projektionsoperator $\pi_M(R)$

- Der Projektionsoperator bildet mit den in M angegebenen Merkmalsnamen die Tabelle R auf eine Teiltabelle ab.
- Die Merkmalsnamen dürfen in einer beliebigen Reihenfolge aufgelistet werden.

ADRESSEN

ID	Name	Vorname	PLZ	Ort
12	Meier	Andreas	6370	Stans
47	Müller	Michael	8000	Zürich
39	Schmid	Susanne	3000	Bern

$\pi_{\text{Vorname}}(ADRESSEN)$

Vorname
Andreas
Michael
Susanne

$\pi_{\text{Ort,Name}}(ADRESSEN)$

Ort	Name
Stans	Meier
Zürich	Müller
Bern	Schmid

4.3.2 Der Selektionsoperator $\sigma_F(R)$

- Unter $\sigma_F(R)$ versteht man alle Tupel aus R, welche die Selektionsbedingung F erfüllen.
- Eine Selektionsbedingung F besteht aus einer bestimmten Anzahl von Merkmalsnamen oder konstanten Werten, die durch Vergleichsoperatoren wie $<$, $>$ oder $=$ sowie durch logische Operatoren wie AND, OR und NOT miteinander kombiniert werden können.

ADRESSEN

ID	Name	Vorname	PLZ	Ort
12	Meier	Andreas	6370	Stans
47	Müller	Michael	8000	Zürich
39	Schmid	Susanne	3000	Bern
23	Meier	Stefan	6000	Luzern

$\sigma_{\text{Name=Meier}}(ADRESSEN)$

ID	Name	Vorname	PLZ	Ort
12	Meier	Andreas	6370	Stans
23	Meier	Stefan	6000	Luzern

$\sigma_{\text{Name=Meier AND PLZ<6100}}(ADRESSEN)$

ID	Name	Vorname	PLZ	Ort
23	Meier	Stefan	6000	Luzern

4.3.3 Der Verbundoperator $R \bowtie_P S$

- Der Verbundoperator der beiden Tabellen R und S über das Prädikat P ist die Menge aller Tupel aus dem kartesischen Produkt $R \times S$, die das Verbundprädikat P erfüllen.
- Das Verbundprädikat P enthält je ein Merkmal aus der Tabelle R und eines aus S. Diese beiden Merkmale werden durch die Vergleichsoperatoren $<$, $>$ und $=$ in Beziehung gesetzt, damit die Tabellen kombiniert werden können.
- Enthält das Verbundprädikat P den Vergleichsoperator $=$, so spricht man von einem *Gleichheitsverbund*. (engl. *equi-join*)
- Lässt man das Verbundprädikat weg ($P = \{\}$), so erhält man als Spezialfall das kartesische Produkt: $R \bowtie_{P=\{\}} S = R \times S$.

KUNDE

K_ID	Name	Ort
21312	Meier	Stans
5945	Müller	Zürich
45101	Schmid	Bern

EINKAUF

E_ID	Datum	K_ID	Betrag
135	3.9.2011	92415	123.70
136	4.9.2011	5945	95.00
137	7.9.2011	21312	69.35

KUNDE $\bowtie_{\text{KUNDE.K\_ID=EINKAUF.K\_ID}}$ EINKAUF

K_ID	Name	Ort	E_ID	Datum	K_ID	Betrag
21312	Meier	Stans	137	7.9.2011	21312	69.35
5945	Müller	Zürich	136	4.9.2011	5945	95.00

4.3.4 Der Verbundoperator $R \bowtie_P S$ (Fortsetzung)

KUNDE $\times$ EINKAUF

K_ID	Name	Ort	E_ID	Datum	K_ID	Betrag
21312	Meier	Stans	135	3.9.2011	92415	123.70
21312	Meier	Stans	136	4.9.2011	5945	95.00
21312	*Meier*	*Stans*	*137*	*7.9.2011*	*21312*	*69.35*
5945	Müller	Zürich	135	3.9.2011	92415	123.70
5945	*Müller*	*Zürich*	*136*	*4.9.2011*	*5945*	*95.00*
5945	Müller	Zürich	137	7.9.2011	21312	69.35
45101	Schmid	Bern	135	3.9.2011	92415	123.70
45101	Schmid	Bern	136	4.9.2011	5945	95.00
45101	Schmid	Bern	137	7.9.2011	21312	69.35

KUNDE $\bowtie_{\text{KUNDE.K\_ID=EINKAUF.K\_ID}}$ EINKAUF

K_ID	Name	Ort	E_ID	Datum	K_ID	Betrag
21312	*Meier*	*Stans*	*137*	*7.9.2011*	*21312*	*69.35*
5945	*Müller*	*Zürich*	*136*	*4.9.2011*	*5945*	*95.00*

Quelle

- Andreas Meier, *Relationale und Postrelationale Datenbanken*, Springer, 2007