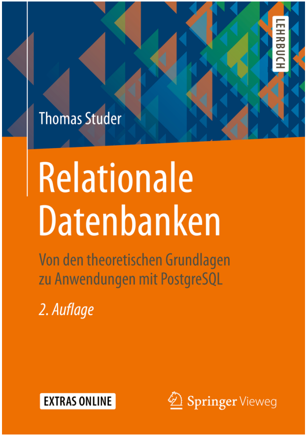

Thomas Studer
Relationale Datenbanken:
Von den theoretischen Grundlagen
zu Anwendungen mit PostgreSQL
2. Auflage
Springer Vieweg, 2019
ISBN 978-3-662-58975-5

UNIVERSITÄT BERN

Datenbanken

—

Mengenlehre

Thomas Studer

Institut für Informatik
Universität Bern

Zwei-sortige Sprache

Objekte sind entweder

1. *atomare Objekte*, d.h. unteilbare Objekte ohne interne Struktur oder
2. n-*Tupel* der Form $(a_1, a_2, \ldots, a_n)$, wobei a_1 bis a_n Objekte sind ($n \geq 1$).

Wir benutzen Kleinbuchstaben $a, b, c, \ldots$ um Objekte zu bezeichnen.

Projektion auf Komponenten

Sei $a = (a_1, \ldots, a_i, \ldots, a_n)$ ein n-Tupel, dann nennen wir a_i die *i-te Komponente* von a.

Spielt bei einem n-Tupel die Anzahl der Komponenten keine Rolle (oder ist sie klar aus dem Kontext), so sprechen wir einfach von einem *Tupel*. Die Komponenten eines Tupels sind geordnet.

Mit Hilfe der Projektionsfunktion $\pi_i(a)$ können wir die i-te Komponente aus einem Tupels a extrahieren. Wir definieren für $a = (a_1, \ldots, a_n)$ und $1 \leq i \leq n$:

$$\pi_i(a) := a_i \quad .$$

Beispiel

Seien a, b und c Objekte. Dann sind auch

$$(a, a, b) \quad \text{sowie} \quad ((a, a, b), c)$$

Objekte. Dabei ist (a, a, b) ein 3-Tupel bei dem die erste und zweite Komponente identisch sind. Das Objekt $((a, a, b), c)$ ist ein 2-Tupel, dessen erste Komponente ein 3-Tupel ist.

Gleichheit

Seien $a = (a_1, \ldots, a_n)$ und $b = (b_1, \ldots, b_n)$ zwei n-Tupel. Es gilt

$$a = b \qquad \text{g.d.w.} \qquad \forall 1 \leq i \leq n (a_i = b_i) \ .$$

Das heisst, zwei n-Tupel sind genau dann gleich, wenn Gleichheit für alle Komponenten gilt.

Mengen

Eine *Menge* ist eine ungeordnete Kollektion von Objekten. Falls das Objekt a zu einer Menge M gehört, sagen wir a ist ein *Element von* M und schreiben $a \in M$.

Eine endliche Menge M kann durch Aufzählen ihrer Elemente beschrieben werden. So besteht beispielsweise die Menge $M = \{a, b, c, d\}$ genau aus den Elementen a, b, c und d.

Bei Mengen geht es ausschliesslich um die Frage, *welche Elemente in ihr enthalten sind*. Häufigkeit und Reihenfolge der Elemente spielen keine Rolle.

$$\{b, b, a, c, d\} \qquad \text{und} \qquad \{a, b, c, d, d, d\}$$

beschreiben dieselbe Menge.

Wir verwenden Grossbuchstaben $A, B, C, \ldots$ um Mengen zu bezeichnen

Objekte und Mengen sind verschieden

Annahme

Die Klasse der Objekte und die Klasse der Mengen sind disjunkt.

Dies bedeutet, dass Mengen keine Objekte sind. Somit kann eine Menge nicht Element einer (anderen) Menge sein. Für Mengen A und B ist also $A \in B$ nicht möglich.

Wir treffen diese Annahme, damit wir uns nicht um Paradoxien der Mengenlehre kümmern müssen.

Gleichheit von Mengen

Zwei Mengen sind *gleich*, falls sie dieselben Elemente enthalten. Formal heisst das

$$A = B \qquad \text{g.d.w.} \qquad \forall x (x \in A \Leftrightarrow x \in B) \ .$$

Statt A und B sind gleich sagen wir auch, A und B sind *identisch*.

Teilmengen

Eine Menge A heisst *Teilmenge* einer Menge B (wir schreiben dafür $A \subseteq B$), falls jedes Element von A auch ein Element von B ist. Das heisst

$$A \subseteq B \qquad \text{g.d.w.} \qquad \forall x(x \in A \Rightarrow x \in B) \ .$$

A heisst *echte* Teilmenge von B (in Zeichen $A \subsetneq B$), falls

$$A \subseteq B \quad \text{und} \quad A \neq B \ .$$

Bemerkung

Für zwei Mengen A und B gilt somit

$$A = B \qquad \textit{g.d.w.} \qquad A \subseteq B \textit{ und } B \subseteq A \ .$$

Komprehension

Ein Prädikat $\varphi(x)$ beschreibt eine Eigenschaft, welche Objekten zu- oder abgesprochen werden kann. Der Ausdruck $\varphi(a)$ sagt, dass das Objekt a die durch $\varphi(x)$ beschriebene Eigenschaft hat. Wir sagen dann a erfüllt φ.

Annahme

Für jedes Prädikat $\varphi(x)$ gibt es eine Menge A, so dass für alle Objekte x gilt

$$x \in A \qquad \textit{g.d.w.} \qquad \varphi(x) \ .$$

Wir verwenden folgende Schreibweise, um eine durch Komprehension gebildete Menge zu definieren

$$A := \{x \mid \varphi(x)\}$$

und sagen A *ist die Menge von allen* x*, welche* φ *erfüllen*.

Beispiel

Die Menge

$$A := \{x \mid \exists y \in \mathbb{N} \text{ mit } x = 2y\}$$

ist die Menge derjenigen x für die es eine natürliche Zahl y gibt mit $x = 2y$. Das heisst, A ist die Menge der geraden natürlichen Zahlen.

Russels Paradox

In der üblichen mathematischen Mengenlehre ist das Schema der (uneingeschränkten) Komprehension nicht zulässig, da es zu Widersprüchen führt, wie z.B. der Menge aller Mengen die sich nicht selbst enthalten

$$R := \{x \mid x \notin x\} \ .$$

Für diese Menge gilt

$$R \in R \qquad \text{g.d.w.} \qquad R \notin R \ ,$$

was unmöglich ist.
In unserem Ansatz ist der Ausdruck $x \notin x$ ist syntaktisch nicht erlaubt, da die Element-Beziehung nur zwischen Objekten und Mengen ausgedrückt werden kann, aber nicht zwischen zwei Mengen oder zwischen zwei Objekten.

Operationen auf Mengen

Vereinigung

$$x \in A \cup B \qquad \text{g.d.w.} \qquad x \in A \text{ oder } x \in B$$

Differenz

$$x \in A \setminus B \qquad \text{g.d.w.} \qquad x \in A \text{ und } x \notin B$$

Schnitt

$$x \in A \cap B \qquad \text{g.d.w.} \qquad x \in A \text{ und } x \in B$$

Die Vereinigung, die Differenz und der Schnitt zweier Mengen existieren (als neue Mengen), da sie durch das Schema der Komprehension gebildet werden können.

Schnitt als Differenz

Lemma

Seien A und B zwei Mengen. Dann gilt

$$A \cap B = A \setminus (A \setminus B) \ .$$

Beweis. Die folgenden Aussagen sind äquivalent:

$$\begin{gathered}
x \in A \cap B \\
x \in A \text{ und } x \in B \\
x \in A \text{ und } (x \notin A \text{ oder } x \in B) \\
x \in A \text{ und nicht } (x \in A \text{ und } x \notin B) \\
x \in A \text{ und nicht } x \in (A \setminus B) \\
x \in A \text{ und } x \notin (A \setminus B) \\
x \in A \setminus (A \setminus B)
\end{gathered}$$

Damit gilt $A \cap B = A \setminus (A \setminus B)$.

Relationen

Eine Menge R heisst n-*stellige* (oder n-*äre*) *Relation* über Mengen $A_1, \ldots, A_n$, falls

$$R \subseteq \{(x_1, \ldots, x_n) \mid x_1 \in A_1 \text{ und } \cdots \text{ und } x_n \in A_n\} \ .$$

Für eine n-stellige Relation R über Mengen $A_1, \ldots, A_n$ gilt somit: Jedes Element von R ist ein n-Tupel $(x_1, \ldots, x_n)$ mit $x_i \in A_i$ für alle $1 \leq i \leq n$.

Kartesisches Produkt

Für eine m-stellige Relation R und eine n-stellige Relation S definieren wir das *kartesische Produkt* $R \times S$ als $(m+n)$-stellige Relation durch

$$R \times S := \big\{(x_1, \ldots, x_{m+n}) \mid (x_1, \ldots, x_m) \in R \text{ und } (x_{m+1}, \ldots, x_{m+n}) \in S\big\} \ .$$

Das kartesische Produkt von R und S besteht aus allen möglichen Kombinationen von Elementen aus R mit Elementen aus S. Besteht R aus h-vielen Elementen und S aus k-vielen Elementen, so enthält das kartesische Produkt $(h \cdot k)$-viele Elemente.

Beispiel

Example

Sei R die 1-stellige Relation $R = \{(a), (b), (c)\}$ und S die 2-stellige Relation $S = \{(1,5), (2,6)\}$. Dann ist $R \times S$ die 3-stellige Relation

$$R \times S = \{(a,1,5),\ (a,2,6),\ (b,1,5),\ (b,2,6),\ (c,1,5),\ (c,2,6)\}\ .$$

Assoziativität

Lemma

Seien R, S und T Relationen. Es gilt

$$(R \times S) \times T = R \times (S \times T) \ .$$

Diese Eigenschaft erlaubt es uns, die Klammern wegzulassen und einfach $R \times S \times T$ zu schreiben.

Flaches Produkt

Unsere Definition liefert ein *flaches* kartesisches Produkt. Das bedeutet, dass das kartesische Produkt einer m-stelligen mit einer n-stelligen Relation eine $(m+n)$-stellige Relation ist. Üblicherweise wird in der mathematischen Mengenlehre das kartesische Produkt anders definiert, nämlich durch

$$R \times S := \{(a, b) \mid a \in R \text{ und } b \in S\} \quad . \tag{1}$$

Damit ist $R \times S$ immer eine 2-stellige Relation. Für R uns S aus dem vorherigen Beispiel finden wir dann

$$\begin{aligned} R \times S := \{&((a),(1,5)),\ ((a),(2,6)),\ ((b),(1,5)),\ ((b),(2,6)),\\ &((c),(1,5)),\ ((c),(2,6))\} \quad . \end{aligned}$$

Die Elemente aus $R \times S$ sind 2-Tupel (Paare) bestehend aus einem 1-Tupel und einem 2-Tupel.
Im Gegensatz zu unserem kartesischen Produkt erfüllt das Produkt aus (1) das Assoziativgesetzt nicht.