3 MENGEN, ALPHABETE, ABBILDUNGEN

Im Kapitel über Signale, Nachrichten, ... usw. haben wir auch über *Inschriften* gesprochen. Ein typisches Beispiel ist der Rosetta-Stein (Abb. 3.1), der für Jean-François Champollion *die* Hilfe war, um die Bedeutung ägyptischer Hieroglyphen zu entschlüsseln. Auf dem Stein findet man Texte in drei Schriften: in Hieroglyphen, in demotischer Schrift und auf Altgriechisch in griechischen Großbuchstaben.

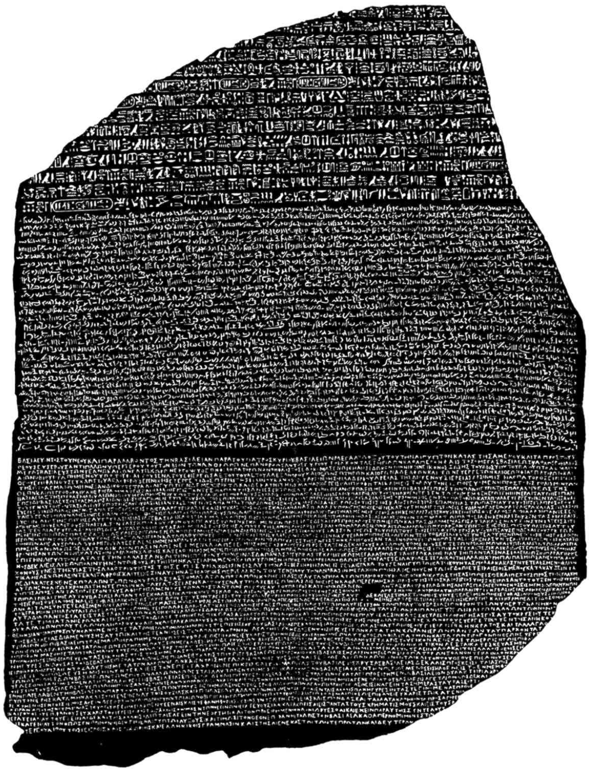

Abbildung 3.1: Der Rosetta-Stein, heute im Britischen Museum, London. Bildquelle: `http://www.gutenberg.org/ebooks/48649` (17.9.15)

Wir sind gewohnt, lange Inschriften aus (Kopien der) immer wieder gleichen Zeichen zusammenzusetzen. Zum Beispiel in europäischen Schriften sind das die Buchstaben, aus denen Wörter aufgebaut sind. Im asiatischen Raum gibt es Schriften mit mehreren Tausend Zeichen, von denen viele jeweils für etwas stehen, was wir als Wort bezeichnen würden.

Einen Vorrat an Zeichen, aus denen Texte zusammengesetzt werden, nennt man ein Alphabet. Das ist etwas präziser gesagt eine endliche Menge von Zeichen.

Daher werden wir in diesem Kapitel zuerst einige Anmerkungen zum Begriff der *Menge* machen. Anschließend werden wir insbesondere auf zwei wichtige Alphabete der Informatik zu sprechen kommen. Das wird dann auch gleich motivieren, warum wir uns mit *Relationen* und *Abbildungen* beschäftigen.

3.1 MENGEN

Jede Menge kann man sich als einen „Behälter“ vorstellen, der „Objekte“ enthält, zum Beispiel die Menge, die die Zahlen 1, 2 und 3 enthält und nichts sonst. Man sagt dann, dass 1 ein Element dieser Menge ist, und ebenso 2 und 3, aber keine anderen Objekte. Weder ist die 42 oder eine andere Zahl ein Element der Menge, noch andere „andersartige“ Objekte wie zum Beispiel der linke Schuh des Autors dieser Zeilen zu der Zeit, als er diese Zeilen schrieb.[1]

eine ganz wichtige allgemeine Bemerkung

An vielen Stellen in diesem Vorlesungsskript (und allgemein in der Mathematik und Informatik) steht man vor dem Problem, dass man über Dinge reden möchte, die nur (mehr oder weniger) mühsam (sprich: länglich) zu beschreiben sind. Statt dann immer wieder die längliche Beschreibung zu wiederholen, vergibt man eine *Namen* dafür. Zum Beispiel hätten wir oben sagen können:

Namen

> *„… zum Beispiel die Menge A, die die Zahlen* 1, 2 *und* 3 *…*
> *Man sagt dann, dass* 1 *ein Element von A ist …“*

Damit ist A schlicht eine Abkürzung für etwas Längeres. Es ist ein Name für etwas ganz bestimmtes.

In einem zweiten Schritt geht man bei der Benutzung von Namen noch weiter. Gehen wir davon aus, dass A der Name für obige Menge ist. Dann sagt man auch so etwas wie: *„Es sei x ein Objekt, das in A enthalten ist.“* Hier wird ein neuer Name eingeführt, nämlich „x“. Damit wird jetzt aber *kein* ganz bestimmtes Objekt bezeichnet, das in A enthalten ist. Vielmehr ist so etwas gemeint wie:

> *„Wir nehmen jetzt ein beliebiges Objekt her, das in A enthalten ist. Es ist gleichgültig, welches es ist. Wir nennen es x.“*

Wenn dann im weiteren Verlauf in einer Aussage von „x“ die Rede ist, ist das eine Art Platzhalter, und man sollte sich klar machen, dass man ihn durch jedes konkrete Objekt ersetzen kann, das in A enthalten ist.

Element einer Menge

Es sei A eine Menge und x ein Objekt. Dann heißt x *Element* von A, wenn es in A enthalten ist, und sonst nicht. Wir bezeichnen mit

$$x \in A$$

[1] eine blaue Sandale

die Aussage, die wahr ist, wenn x Element von A ist, und andernfalls falsch. Mit

$$x \notin A$$

bezeichnen wir das Gegenteil von $x \in A$, also die Aussage, die genau dann wahr ist, wenn x nicht Element von A ist. Je nachdem, für welche konkreten Werte die Namen x und A stehen, ist also immer eine der beiden Aussagen wahr und die andere falsch.

Anstelle von "Es sei x ein Element von A" (und "Es sei x ein Objekt so, dass $x \in A$ wahr ist") schreiben wir abkürzend "Es sei $x \in A$".

Jede Menge soll durch die Elemente, die sie enthält, eindeutig bestimmt sein: Für zwei Mengen A und B ist also genau dann $A = B$, wenn für jedes Objekt x gilt:

$$\text{wenn } x \in A, \text{ dann } x \in B \quad \text{und} \quad \text{wenn } x \in B, \text{ dann } x \in A \ .$$

Ist jedes Element einer Menge A auch Element einer Menge B, dann schreiben wir $A \subseteq B$ und sprechen davon, dass A *Teilmenge von* B sei und B *Obermenge* von A. Folglich gilt

Teilmenge
Obermenge

3.1 **Lemma.** Für beliebige Mengen A und B ist

$$\text{genau dann } A = B, \text{ wenn } A \subseteq B \text{ und } B \subseteq A \ .$$

Kleine endliche Mengen kann man angeben, indem man alle Elemente aufzählt. Wir schließen die Aufzählung aller Elemente einer Menge in geschweifte Klammern ein:

$$\{1, 2, 3, 4, 5, 42, \mathtt{A}, \mathtt{B}\} \ .$$

Die kleinste Menge enthält gar keine Elemente. Für sie schreibt man

$$\{\,\} \qquad \text{oder} \qquad \emptyset$$

und spricht von der *leeren Menge*.

leere Menge

Man beachte, dass man die gleiche Menge verschieden hinschreiben kann:

$$\{1, 2, 3\} = \{3, 2, 1\} = \{1, 1, 2, 3, 3, 2, 1, 3, 3, 3\} \ .$$

Unsere Definition der Gleichheit von Mengen verlangt nur, dass sie jeweils die gleichen Elemente enthalten müssen.

Bei großen oder gar unendlichen Mengen kann man nicht explizit alle Elemente einzeln aufführen. In solchen Fällen gehen wir anders vor. Zum einen setzen wir manche unendlichen Mengen einfach als bekannt voraus. Für die Menge der positiven ganzen Zahlen schreiben wir $\mathbb{N}_+$ und für die Menge der nichtnegativen ganzen Zahlen $\mathbb{N}_0$. Es ist also sozusagen

$$\mathbb{N}_+ = \{1,2,3,4,\dots\} \text{ und } \mathbb{N}_0 = \{0,1,2,3,4,\dots\} ,$$

wenn man einmal unterstellt, dass die Pünktchen von allen Lesern „richtig“ interpretiert werden.

set comprehension

Eine Möglichkeit auch unendliche Menge präzise zu spezifizieren bietet eine Notation, die im Englischen *set comprehension* heißt und für wir keine gute deutsche Bezeichnung kennen.[2] Dafür benötigt man für jedes Element x einer gewissen Grundmenge M eine Aussage $P(x)$, die wahr oder falsch ist. (Wenn $P(x)$ wahr ist, sagen wir auch, „$P(x)$ gilt“.) Dann bezeichnet $\{x \in M \mid P(x)\}$ die Menge derjenigen Elemente $x \in M$, für die die Aussage $P(x)$ wahr ist, d. h. für jedes $y \in M$ gilt:

$$y \in \{x \in M \mid P(x)\} \text{ genau dann, wenn } P(y) \text{ wahr ist .}$$

Man liest das zum Beispiel als *„die Menge aller $x \in M$, für die $P(x)$ wahr ist“*. Zum Beispiel ist

$$\{x \in \mathbb{N}_+ \mid x \text{ ist ohne Rest durch 2 teilbar}\}$$

die Menge aller positiven geraden Zahlen. Wenn die Grundmenge M aus dem Kontext eindeutig hervorgeht, schreiben wir abkürzend $\{x \mid P(x)\}$.

Vereinigung

Für zwei Mengen A und B ist die *Vereinigung* $A \cup B$ eine Menge, die durch die Forderung eindeutig bestimmt ist, dass für jedes Objekt x gilt:

$$x \in A \cup B \text{ genau dann, wenn } x \in A \text{ oder } x \in B .$$

Durchschnitt

Dabei meinen wir mit „oder“ das inklusive Oder: Es ist auch erlaubt, dass x Element beider Mengen ist. Der *Durchschnitt* $A \cap B$ ist eine Menge, die durch die Forderung eindeutig bestimmt ist, dass für jedes Objekt x gilt:

$$x \in A \cap B \text{ genau dann, wenn } x \in A \text{ und } x \in B .$$

Es sei noch einmal daran erinnert, dass sozusagen jedes Element einer Menge immer nur *„einmal* enthalten“ ist. Daher ist zum Beispiel

$$\{1,2,3\} \cup \{3,4,5\} = \{1,2,3,4,5\} .$$

disjunkte Mengen

Zwei Mengen, deren Durchschnitt die leere Menge ist, nennt man *disjunkt*.

Allgemein ist zum Beispiel für jede Menge A stets $A \cup A = A$. Diese und weitere Aussagen sind in den folgenden Lemmata zusammengefasst.

[2] Die deutsche Bezeichnung „Intensionale Mengennotation“ ist nicht weit verbreitet.

3.2 **Lemma.** Für jede Menge A ist $A \cup A = A$ und $A \cap A = A$.

3.3 **Lemma.** Für alle Mengen A und B ist $A \cup B = B \cup A$ und $A \cap B = B \cap A$.

3.4 **Lemma.** Für alle Mengen A und B ist $A \subseteq A \cup B$ und $A \cap B \subseteq A$.

3.5 **Lemma.** Für alle Mengen A, B und C ist $(A \cup B) \cup C = A \cup (B \cup C)$ und $(A \cap B) \cap C = A \cap (B \cap C)$.

Einer der wichtigsten Aspekte der Vorlesung „Grundbegriffe der Informatik" ist, dass Sie lernen, präzise zu formulieren und zu argumentieren. Denn ein Algorithmus muss nicht nur präzise aufgeschrieben werden. Man muss auch nachweisen können, dass es das tut, „was er soll". Dafür muss übrigens auch die Spezifikation präzise aufgeschrieben sein. Unter anderem aus diesem Grund werden wir in dieser Vorlesung viele Aussagen beweisen.

Die obigen Lemmata sind aber alle recht einfach. Deswegen beschränken wir uns hier auf einen Fall:

3.6 **Beweis. (erster Teil von Lemma 3.2)** Es sei A eine Menge. Um zu beweisen, dass $A = A \cup A$ ist, müssen wir nach Lemma 3.1 zeigen:

1. $A \subseteq A \cup A$
2. $A \cup A \subseteq A$

Sehen wir uns die beiden Fälle nacheinander an:

1. Wir müssen zeigen: Jedes Element $x \in A$ ist auch Element von $A \cup A$. Wenn man das ausführlich aufschreibt, ergibt sich in etwa folgendes:
   - *Es sei $x \in A$.*
     Wir beginnen mit der Voraussetzung. Und um zu zeigen, dass die Aussage für jedes Element gilt, argumentieren wir für ein „beliebiges" Element, für das keine weiteren Annahmen gemacht werden.
   - *Dann ist $x \in A$ oder $x \in A$.*
     Wir ziehen eine einfache Schlussfolgerung, der unser umgangssprachliches Verständnis von „oder" zugrunde liegt.
   - *Also ist $x \in A \cup A$.*
     Diese Schlussfolgerung ist gerade das, was per Definition von Mengenvereinigung gelten muss, damit die Behauptung stimmt.
2. Die umgekehrte Richtung ist genauso einfach: Wir müssen zeigen: Jedes Element $x \in A \cup A$ ist auch Element von A. Die Struktur der Argumentation ist ähnlich wie eben:

- *Es sei* $x \in A \cup A$.
  Wir beginnen wieder mit einem „beliebigen“ Element.
- *Dann ist* $x \in A$ *oder* $x \in A$.
  Das ergibt sich aufgrund der Definition von Mengenvereinigung.
- *Also ist* $x \in A$.

■

Mengendifferenz

Neben Vereinigung und Durchschnitt ist gelegentlich auch noch die *Mengendifferenz* $A \setminus B$ nützlich:

$$A \setminus B = \{x \in A \mid x \notin B\} .$$

Es ist zum Beispiel

$$\{1,2,3\} \setminus \{3,4,5\} = \{1,2\} .$$

In der nachfolgenden Abbildung 3.2 sind für zwei „allgemeine“ Mengen A und B die Lage einiger Teilmengen dargestellt. Es fehlt $A \cup B$; machen Sie sich klar, welcher Bereich des Bildes dazu gehört.

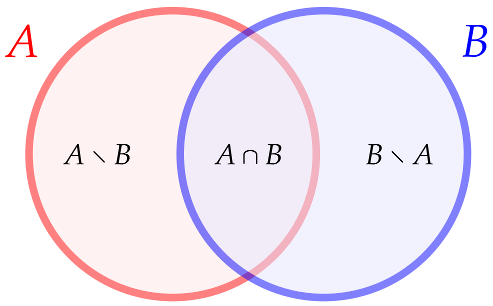

Abbildung 3.2: Teile zweier Mengen. Die linke, rote Kreisscheibe stelle eine Menge A dar und die rechte, blaue Kreisscheibe eine Menge B.

Kardinalität

Die Anzahl der Elemente einer Menge A bezeichnet man auch als ihre *Kardinalität*, für die wir meist

$$|A| \qquad \text{(oder vereinzelt card } A\text{)}$$

schreiben werden. Für jede endliche Menge A ist das offensichtlich eine nichtnegative ganze Zahl. Für unendliche Mengen werden wir nicht definieren, was $|A|$ ist, und vermeiden darüber zu reden.

3.2 ALPHABETE

Unter einem *Alphabet* wollen wir eine endliche nichtleere Menge verstehen, deren Elemente wir *Zeichen* oder *Symbole* nennen. Was dabei genau „Zeichen" sind, wollen wir nicht weiter hinterfragen. Es seien einfach die elementaren Bausteine, aus denen Inschriften zusammengesetzt sind. Stellen Sie sich zum Beispiel einige Buchstaben des deutschen Alphabetes vor. Hier sind einfache Beispiele: *Alphabet*

- $A = \{\texttt{|}\}$
- $A = \{\texttt{a}, \texttt{b}, \texttt{c}\}$
- $A = \{\texttt{0}, \texttt{1}\}$
- Manchmal erfindet man auch Zeichen: $A = \{\texttt{1}, \texttt{0}, \rotatebox[origin=c]{180}{\texttt{1}}\}$
- $A = \{\texttt{0}, \texttt{1}, \texttt{2}, \texttt{3}, \texttt{4}, \texttt{5}, \texttt{6}, \texttt{7}, \texttt{8}, \texttt{9}, \texttt{A}, \texttt{B}, \texttt{C}, \texttt{D}, \texttt{E}, \texttt{F}\}$

Gelegentlich nimmt man aber auch einen etwas abstrakteren Standpunkt ein und sieht zum Beispiel jeden der folgenden „Kästen" als jeweils *ein* Zeichen eines gewissen Alphabetes an:

$\boxed{\texttt{int}}\ \boxed{\texttt{adams}}\ \boxed{\texttt{=}}\ \boxed{\texttt{42}}\ \boxed{\texttt{;}}$

So etwas wird Ihnen zum Beispiel in Vorlesungen über Übersetzerbau wieder begegnen.

Bereits in dieser Vorlesung werden wir im Kapitel über aussagenlogische Formeln annehmen, dass das zugrunde liegende Alphabet Symbole $\texttt{P}_1$, $\texttt{P}_2$, $\texttt{P}_3$, usw. enthält. Oder Sie stellen sich vor, man hätte nur die Symbole $\texttt{P}$ sowie die zehn Indexziffern $_0, _1, _2, \dots, _9$, aus denen die „eigentlichen" grundlegenden syntaktischen Bausteine zusammengesetzt sind.

3.2.1 Beispiel ASCII

Ein wichtiges Alphabet ist der sogenannte *ASCII*-Zeichensatz. Die Abkürzung steht für *American Standard Code for Information Interchange*. Diese Spezifikation umfasst insbesondere eine Liste von 94 „druckbaren" und einem „unsichtbaren" Zeichen, die man z. B. in Emails verwenden darf. Außerdem hat jedes Zeichen eine Nummer aus dem Bereich der natürlichen Zahlen zwischen 32 und 126. Die vollständige Liste findet man in Tabelle 3.1. Wie man dort sieht, fehlen in diesem Alphabet etliche Buchstaben aus nichtenglischen Alphabeten, wie zum Beispiel ä, ç, è, ğ, ñ, œ, ß, ů usw., von Kyrillisch, Japanisch und vielen anderen außereuropäischen Schriften ganz zu schweigen. *ASCII*

Auf ein Zeichen in Tabelle 3.1 sei ausdrücklich hingewiesen, nämlich das mit Nummer 32. Das ist das „Leerzeichen". Man gibt es normalerweise auf einer Rechnertastatur ein, indem man die extrabreite Taste ohne Beschriftung drückt. Auf dem Bildschirm wird dafür in der Regel *nichts* dargestellt. Damit man es trotz-

		40	(	50	2	60	<	70	F
		41	)	51	3	61	=	71	G
32	␣	42	*	52	4	62	>	72	H
33	!	43	+	53	5	63	?	73	I
34	"	44	,	54	6	64	@	74	J
35	#	45	-	55	7	65	A	75	K
36	$	46	.	56	8	66	B	76	L
37	%	47	/	57	9	67	C	77	M
38	&	48	0	58	:	68	D	78	N
39	'	49	1	59	;	69	E	79	O
80	P	90	Z	100	d	110	n	120	x
81	Q	91	[	101	e	111	o	121	y
82	R	92	\	102	f	112	p	122	z
83	S	93	]	103	g	113	q	123	{
84	T	94	^	104	h	114	r	124	|
85	U	95	_	105	i	115	s	125	}
86	V	96	`	106	j	116	t	126	~
87	W	97	a	107	k	117	u		
88	X	98	b	108	l	118	v		
89	Y	99	c	109	m	119	w		

Tabelle 3.1: Die „druckbaren“ Zeichen des ASCII-Zeichensatzes (einschließlich Leerzeichen)

dem sieht und um darauf aufmerksam zu machen, dass das ein Zeichen ist, ist es in der Tabelle als ␣ dargestellt.

3.2.2 Beispiel Unicode

Der Unicode Standard (siehe auch `http://www.unicode.org`) definiert mehrere Dinge. Das wichtigste und Ausgangspunkt für alles weitere ist eine umfassende Liste von Zeichen, die in der ein oder anderen der vielen heute gesprochenen Sprachen (z. B. in Europa, im mittleren Osten, oder in Asien) benutzt wird. Die Seite `http://www.unicode.org/charts/` vermittelt einen ersten Eindruck von der existierenden Vielfalt.

Das ist mit anderen Worten ein Alphabet, und zwar ein großes: es umfasst rund 100 000 Zeichen.

Der Unicode-Standard spezifiziert weitaus mehr als nur einen Zeichensatz.[3]

[3] Hinzu kommt etwa die Sortierreihenfolge von Buchstaben in verschiedenen Sprachen.

Für uns sind hier zunächst nur die beiden folgenden Aspekte wichtig:

1. Es wird eine große (aber endliche) Menge A_U von Zeichen festgelegt, und
2. eine Nummerierung dieser Zeichen, jedenfalls in einem gewissen Sinne.

Punkt 1 ist klar. Hinter der Formulierung von Punkt 2 verbirgt sich genauer folgendes: Jedem Zeichen aus A_U ist eine nichtnegative ganze Zahl zugeordnet, der auch sogenannte *Code Point* des Zeichens. Die Liste der benutzten Code Points ist allerdings nicht „zusammenhängend".

Jedenfalls liegt eine Beziehung zwischen Unicode-Zeichen und nichtnegativen ganzen Zahlen vor. Man spricht von einer Relation. (Wenn Ihnen die folgenden Zeilen schon etwas sagen: schön. Wenn nicht, gedulden Sie sich bis Abschnitt 3.3 wenige Zeilen weiter.)

Genauer liegt sogar eine Abbildung $f : A_U \to \mathbb{N}_0$ vor. Sie ist

- eine Abbildung, weil jedem Zeichen nur *eine* Nummer zugewiesen wird,
- injektiv, weil verschiedenen Zeichen verschiedene Nummern zugewiesen werden,
- aber natürlich nicht surjektiv (weil A_U nur endlich viele Zeichen enthält).

Entsprechendes gilt natürlich auch für den ASCII-Zeichensatz.

3.3 Relationen und Abbildungen

3.3.1 Paare, Tupel und kartesische Produkte

Es seien A und B zwei Mengen. Für Elemente $a \in A$ und $b \in B$ wird durch

$$(a, b)$$

etwas notiert, was man als *Paar*, *geordnetes Paar* oder *Tupel* bezeichnet. Der Teil a vor dem Komma heißt *erste Komponente* des Paares und der Teil b nach dem Komma *zweite Komponente*. *Paar* *Tupel*

Zwei Paare (a, b) und (x, y) sind genau dann gleich, wenn $a = x$ ist und $b = y$. Also sind $(1, 2)$ und $(2, 1)$ *verschiedene* Paare. Außerdem können erste und zweite Komponente eines Paares durchaus gleich sein. Auch $(1, 1)$ ist ein Paar mit zwei Komponenten.

Das *kartesische Produkt* der Mengen A und B, geschrieben $A \times B$, ist die Menge *aller* Paare (a, b) mit $a \in A$ und $b \in B$: *kartesisches Produkt*

$$A \times B = \{(a, b) \mid a \in A \text{ und } b \in B\} .$$

Hier haben Gebrauch gemacht von unserer Vereinbarung, dass man die Grundmenge (hier aller Paare) in einer set comprehension weglassen darf, wenn diese klar aus dem Kontext hervorgeht. Manchmal ist es bequem, mit Verallgemeinerungen von Paaren zu arbeiten, die nicht genau zwei Komponenten haben. Eine Möglichkeit besteht darin, diesen allgemeineren Fall auf den schon definierten einfachen Fall zurückzuführen: Seien etwa drei Mengen M_1, M_2 und M_3 gegeben. Dann ist $(M_1 \times M_2) \times M_3$ die Menge aller Paare der Form $((x_1, x_2), x_3)$, deren erste Komponente $(x_1, x_2) \in M_1 \times M_2$ ist, und deren zweite Komponente $x_3 \in M_3$ ist. Um beim Schreiben Klammern zu sparen und das Ganze so lesbarer zu machen, erlauben wir, dass man im Ausdruck $(M_1 \times M_2) \times M_3$ die Klammern weglassen darf und im Ausdruck $((x_1, x_2), x_3)$ nur die äußeren Klammern schreiben muss. Alle anderen dürfen weggelassen werden. Dann ist also sozusagen

$$M_1 \times M_2 \times M_3 = \{(x_1, x_2, x_3) \mid x_1 \in M_1 \text{ und } x_2 \in M_2 \text{ und } x_3 \in M_3\} .$$

Tripel

Die Elemente solcher Mengen nennt man auch *Tripel*. Wir wollen diese Vorgehensweise verallgemeinern und für jede positive ganze Zahl n so etwas wie $M_1 \times M_2 \times \dots \times M_n$ definieren, aber *ohne* die etwas vagen Pünktchen. Dazu definieren wir erstens für jedes $n \in \mathbb{N}_+$ eine Menge

$$\mathbb{P}_n = \{i \in \mathbb{N}_+ \mid i \leq n\} ,$$

und zweitens induktiv für jedes $n \in \mathbb{N}_+$ und alle Mengen M_i, mit $i \in \mathbb{P}_n$, eine Menge, die wir *kartesisches Produkt von M_i, mit $i \in \mathbb{P}_n$*, nennen und in der Form

$$\bigtimes_{i \in \mathbb{P}_n} M_i$$

notieren, wie folgt:

$$\bigtimes_{i \in \mathbb{P}_1} M_i = M_1 ,$$

$$\text{für jedes } n \in \mathbb{N}_+\text{: } \bigtimes_{i \in \mathbb{P}_{n+1}} M_i = \left(\bigtimes_{i \in \mathbb{P}_n} M_i \right) \times M_{n+1} .$$

induktive Definition

Das ist eine sogenannte *induktive Definition*. Man definiert

- zunächst explizit, was die Notation im Fall $n = 1$ bedeuten soll, nämlich $\bigtimes_{i \in \mathbb{P}_1} M_i = M_1$, (es hat sich einfach als nützlich erwiesen, das gerade so festzulegen), und dann
- wie man $\bigtimes_{i \in \mathbb{P}_{n+1}} M_i$ ausrechnen kann, wenn man schon $\bigtimes_{i \in \mathbb{P}_n} M_i$ kennt.

Man kann nun z. B. ausrechnen:

$$\begin{aligned}
\bigtimes_{i\in\mathbb{P}_2} M_i &= \left(\bigtimes_{i\in\mathbb{P}_1} M_i\right) \times M_2 \\
&= M_1 \times M_2 \\
\text{dann} \bigtimes_{i\in\mathbb{P}_3} M_i &= \left(\bigtimes_{i\in\mathbb{P}_2} M_i\right) \times M_3 \\
&= (M_1 \times M_2) \times M_3 \\
\text{dann} \bigtimes_{i\in\mathbb{P}_4} M_i &= \left(\bigtimes_{i\in\mathbb{P}_3} M_i\right) \times M_4 \\
&= ((M_1 \times M_2) \times M_3) \times M_4
\end{aligned}$$

und so weiter. Es sei noch einmal daran erinnert, dass wir zur Vereinfachung die Klammern hätten weglassen dürfen. Man gewinnt den Eindruck, dass für jede positive ganze Zahl n eindeutig festgelegt ist, was $\bigtimes_{i\in\mathbb{P}_n} M_i$ sein soll. Das ist in der Tat so. Machen Sie sich das in diesem Beispiel klar!

Allgemein ist es ein Faktum der Mathematik, dass eine Menge M, die nur positive ganze Zahlen enthält und die Eigenschaften erfüllt, dass

1. $1 \in M$ ist und
2. für jedes $n \in \mathbb{N}_0$ gilt, dass wann immer $n \in M$ ist, dann auch $n + 1 \in M$,

gerade die Menge $\mathbb{N}_+$ ist. (Wenn man die postiven ganzen Zahlen axiomatisch definiert, ist das sogar gerade eine der Forderungen.) Und wenn M nichtnegative ganzen Zahlen enthält und man mit $0 \in M$ beginnen kann, dann ist $M = \mathbb{N}_0$. Daraus ergibt sich zum einen die Möglichkeit, mit Hilfe induktiver Definitionen überhaupt sinnvolle Festlegungen zu treffen, zum anderen fußt darauf das Beweisprinzip der sogenannten *vollständigen Induktion*, auf das wir in späteren Kapiteln ausführlich eingehen und zurückgreifen werden.

Die Elemente eines kartesischen Produkts von n Mengen nennt man auch *n-Tupel*. Falls n hinreichend klein ist, schreiben wir z. B. auch $M_1 \times M_2 \times M_3$ statt $\bigtimes_{i\in\mathbb{P}_3} M_i$. *n-Tupel*

Sind Mengen M_i, mit $i \in \mathbb{P}_n$ und $n \in \mathbb{N}_+$, alle gleich einer Menge M, dann nennen wir $\bigtimes_{i\in\mathbb{P}_n} M_i$ *n-faches kartesisches Produkt der Menge M mit sich selbst* und notieren dieses kurz als M^n. Mit anderen Worten, $M^n = \bigtimes_{i\in\mathbb{P}_n} M$.

3.3.2 Relationen und Abbildungen

Die Beziehung zwischen den Unicode-Zeichen in A_U und nichtnegativen ganzen Zahlen kann man durch die Angabe aller Paare (a, n), für die $a \in A_U$ ist und n

der zu a gehörenden Code Point, vollständig beschreiben. Für die Menge U aller dieser Paare gilt also $U \subseteq A_U \times \mathbb{N}_0$.

Relation *binäre Relation* *ternäre Relation*

Eine Teilmenge $R \subseteq A \times B$ heißt auch eine *Relation*. Manchmal sagt man noch genauer *binäre* Relation; und manchmal noch genauer „von A in B". Das kann man auf mehr als zwei Faktoren verallgemeinern: Eine *ternäre* Relation ist zum Beispiel eine Teilmenge $R \subseteq A \times B \times C$.

Die durch Unicode definierte Menge $U \subseteq A_U \times \mathbb{N}_0$ hat „besondere" Eigenschaften, die nicht jede Relation hat. Diese (und ein paar andere) Eigenschaften wollen wir im folgenden kurz aufzählen und allgemein definieren:

1. Zum Beispiel gibt es für jedes Zeichen $a \in A_U$ (mindestens) ein $n \in \mathbb{N}_0$ mit $(a, n) \in U$.

   linkstotal

   Allgemein nennt man eine Relation $R \subseteq A \times B$ *linkstotal*, wenn für jedes $a \in A$ (mindestens) ein $b \in B$ existiert mit $(a, b) \in R$.
2. Für kein Zeichen $a \in A_U$ gibt es mehrere $n \in \mathbb{N}_0$ mit der Eigenschaft $(a, n) \in U$.

   rechtseindeutig

   Allgemein nennt man eine Relation $R \subseteq A \times B$ *rechtseindeutig*, wenn es für kein $a \in A$ zwei $b_1 \in B$ und $b_2 \in B$ mit $b_1 \neq b_2$ gibt so, dass sowohl $(a, b_1) \in R$ als auch $(a, b_2) \in R$ ist.
3. Relationen, die linkstotal und rechtseindeutig sind, kennen Sie auch unter anderen Namen: Man nennt sie *Abbildungen* oder auch *Funktionen* und man schreibt dann üblicherweise $R : A \to B$. Es heißt dann A der *Definitionsbereich* und B der *Zielbereich* der Abbildung.

   Abbildung *Funktion* *Definitionsbereich* *Zielbereich*

   Gelegentlich ist es vorteilhaft, sich mit Relationen zu beschäftigen, von denen man nur weiß, dass sie rechtseindeutig sind. Sie nennt man manchmal *partielle Abbildungen*. (Bei ihnen verzichtet man also auf die Linkstotalität.)

   partielle Abbildung
4. Außerdem gibt es bei Unicode keine zwei verschiedene Zeichen a_1 und a_2, denen der gleiche Code Point zugeordnet ist.

   linkseindeutig

   Eine Relation $R \subseteq A \times B$ heißt *linkseindeutig*, wenn für alle $(a_1, b_1) \in R$ und alle $(a_2, b_2) \in R$ gilt:
   $$\text{wenn } a_1 \neq a_2, \text{ dann } b_1 \neq b_2 .$$
5. *injektiv* Eine Abbildung, die linkseindeutig ist, heißt *injektiv*.
6. Der Vollständigkeit halber definieren wir auch gleich noch, wann eine Relation $R \subseteq A \times B$ *rechtstotal* heißt: wenn für jedes $b \in B$ ein $a \in A$ existiert, für das $(a, b) \in R$ ist.

   rechtstotal
7. *surjektiv* Eine Abbildung, die rechtstotal ist, heißt *surjektiv*.
8. *bijektiv* Eine Abbildung, die sowohl injektiv als auch surjektiv ist, heißt *bijektiv*.

Im Laufe der Vorlesung wird es immer wieder notwendig sein, Abbildungen zu definieren. Dazu gehört als erstes immer die Angabe von Definitionsbereich und

© worsch&wacker 2008–2015

Zielbereich (denn davon ist ja zum Beispiel unter Umständen abhängig, ob eine Abbildung injektiv ist). Wir notieren das wie oben erwähnt in der Form

$$f \colon A \to B \ .$$

In dieser Situation gibt es für jedes $a \in A$ *genau ein* $b \in B$ mit $(a, b) \in f$. Man schreibt dann üblicherweise $f(a) = b$ und nennt b den *Wert* oder *Funktionswert* von f an der Stelle a. Wenn für jedes Element a des Definitionsbereichs der Funktionswert $f(a)$ spezifiziert werden soll, schreiben wir das oft in einer der Formen

Funktionswert

$$\begin{aligned} x &\mapsto \text{Ausdruck, in dem } x \text{ vorkommen darf,} \\ (x, y) &\mapsto \text{Ausdruck, in dem } x \text{ und } y \text{ vorkommen dürfen,} \end{aligned}$$

falls der Definitionsbereich von f eine Menge bzw. das kartesische Produkt zweier Mengen ist. (Wenn der Definitionsbereich kartesisches Produkt von mehr als zwei Mengen ist, gehen wir analog vor.) Im ersten Fall wird ein Name (hier x) für den Wert eingeführt, auf den die Abbildung angewendet wird. Im zweiten Fall wird angenommen, dass der Wert ein Paar ist, es werden sogar Namen (x und y) für seine „Bestandteile" eingeführt und es wird darauf vertraut, dass allen klar ist, welche Bestandteile gemeint sind.

Auf der rechten Seite der Definition von Funktionswerten erlauben wir uns die Freiheit, nicht nur einfache arithmetische Ausdrücke zu notieren, sondern z. B. auch Fallunterscheidungen. Hier ist ein einfaches Beispiel:

$$f \colon \mathbb{N}_0 \to \mathbb{N}_0, n \mapsto \begin{cases} n/2, & \text{falls } n \text{ gerade,} \\ 3n + 1, & \text{falls } n \text{ ungerade.} \end{cases}$$

Man beachte, dass die Namen die in der Definition der Abbildung verwendet werden, völlig irrelevant sind. Genau die gleiche Abbildung wie eben wird definiert, wenn man schreibt:

$$g \colon \mathbb{N}_0 \to \mathbb{N}_0, x \mapsto \begin{cases} x/2, & \text{falls } x \text{ gerade,} \\ 3x + 1, & \text{falls } x \text{ ungerade.} \end{cases}$$

Mitunter ist die folgende Verallgemeinerung der Schreibweise $f(a)$ nützlich. Ist $f \colon A \to B$ und $M \subseteq A$, so sei

$$f(M) = \{b \in B \mid \text{ es gibt ein } a \in M \text{ mit } f(a) = b\} \ .$$

Das ist mühsam (zu schreiben und) zu lesen, weshalb wir folgende verallgemeinerte Notation für *set comprehensions* erlauben:

$$\text{statt } \{b \in B \mid \text{es gibt ein } a \text{ mit } f(a) = b \text{ und } P(a)\}$$
$$\text{kurz } \{f(a) \mid P(a)\} \, .$$

Damit ist $f(M) = \{f(a) \mid a \in M\}$. Für eine Abbildung $f\colon A \to B$ heißt $f(A)$ auch das *Bild* der Abbildung.

Bild

3.4 MEHR ZU MENGEN

Elemente einer Menge können ihrerseits wieder Mengen sein. Zum Beispiel enthält die Menge

$$M = \{\ 1,\ \{2,3\},\ 4,\ 5,\ 6,\ \{7,8,9\},\ \{\}\ \}$$

sieben (!) Elemente, von denen drei ihrerseits wieder Mengen sind, nämlich $\{2,3\}$, $\{7,8,9\}$ und die leere Menge $\{\}$. Man beachte auch, dass die Menge $\{\ \{\}\ \}$ ein Element enthält (die leere Menge), also selbst offensichtlich *nicht* leer ist!

Potenzmenge

Zu jeder Menge M definiert man die *Potenzmenge* als die Menge *aller* Teilmengen von M. Wir schreiben dafür 2^M, einige andere Autoren benutzen die Notation $\mathfrak{P}(M)$. Also:

$$2^M = \{A \mid A \subseteq M\} \, .$$

Betrachten wir als Beispiel $M = \{1,2,3\}$. Dann ist

$$2^{\{1,2,3\}} = \{\{\},\ \{1\},\ \{2\},\ \{3\},\ \{1,2\},\ \{1,3\},\ \{2,3\},\ \{1,2,3\}\} \, .$$

Allgemein kann man also immer statt $A \subseteq M$ auch $A \in 2^M$ schreiben; und gelgentlich tut man das auch.

Zu zwei beliebigen Mengen I und M betrachten wir nun eine Abbildung

$$f\colon I \to 2^M \, .$$

Wenn es Ihnen die Sache erleichtert, stellen Sie sich vor, es sei z. B. $I = \mathbb{P}_2 = \{1,2\}$ oder $I = \mathbb{N}_0$. Es ist jedenfalls für jedes $i \in I$ der Funktionswert $f(i)$ eine Element von 2^M, also eine Teilmenge von M. Statt $f(i)$ schreiben wir im Folgenden M_i. Wenn $I = \{1,2\}$ ist, dann sind also einfach zwei Mengen M_1 und M_2 festgelegt. Und wenn $I = \mathbb{N}_0$ ist, dann hat man eine sogenannte *abzählbar unendliche* Folge von Mengen (genauer Teilmengen von M) M_0, M_1, M_2, und so weiter.

Vereinigung
Durchschnitt

Als Verallgemeinerung von Durchschnitt und Vereinigung zweier Mengen definieren wir die *Vereinigung* $\bigcup_{i \in I} M_i$ und den *Durchschnitt* $\bigcap_{i \in I} M_i$ aller M_i so:

$$\bigcup_{i \in I} M_i = \{x \mid \text{es gibt ein } i \in I \text{ so, dass } x \in M_i\} \ ,$$
$$\bigcap_{i \in I} M_i = \{x \mid \text{für jedes } i \in I \text{ ist } x \in M_i\} \ .$$

Machen Sie sich bitte klar, dass im Fall $I = \{1, 2\}$ gilt:

$$\bigcup_{i \in I} M_i = M_1 \cup M_2 \text{ und}$$
$$\bigcap_{i \in I} M_i = M_1 \cap M_2 \ .$$

ZUSAMMENFASSUNG UND AUSBLICK

In diesem Kapitel wurde zunächst auf den Begriff der *Menge* eingegangen und es wurden einfache Operationen mit Mengen definiert.

Dann wurde der Begriff *Alphabet* eingeführt, der im weiteren Verlauf dieser Vorlesung und des Informatikstudiums noch an vielen Stellen eine Rolle spielen wird. Wer mehr über Schriften wissen möchte, findet zum Beispiel über die WWW-Seite `http://www.omniglot.com/writing/` (6.10.2015) weitere Informationen.

Als wichtige technische Hilfsmittel wurden die Begriffe *binäre Relation*, sowie *injektive*, *surjektive* und *bijektive Abbildungen* definiert.

Viele Aussagen hatten noch den Nachteil, dass sie umgangssprachlich und daher relativ weitschweifig formuliert waren. Nachdem im nächsten Kapitel die Begriffe *Wort* und *formale Sprache* eingeführt worden sein werden, werden wir uns daher anschließend ein erstes Mal mit *Aussagenlogik* und danach mit *Prädikatenlogik* (erster Stufe) beschäftigen.