

# Verstehen der Reellen Zahlen

## Ein Weg zur KI

Hanns Sommer

### Einführung

Verständlich machen bedeutet, Vorstellungsbilder zu erzeugen, aus denen der intendierte Sachverhalt sichtbar wird. Der Begriff des Beweises in der Mathematik kann in diesem Sinne als die Verständlich-Machung von Sätzen der Mathematik verstanden werden. Er erfordert:

- Die Herstellung eines Verstehenshorizonts und
- das Erkennen des intendierten Sachverhalts in diesem.

Wir haben in Sommer (2023) diese Methode zum Verstehen von Aussagen der Zahlentheorie verwendet; um aber auch die reellen Zahlen ( $\mathcal{R}$ ) erfassen zu können, müssen wir sie erheblich erweitern. Diese Erweiterung wird uns einerseits die prinzipiellen Unterschiede zwischen dem Verstehenshorizont für die natürlichen Zahlen und demjenigen für die reellen Zahlen aufzeigen, und andererseits werden wir erkennen, dass damit genau das Verstehen erfasst wird, das wir zur Entwicklung und zum Einsatz von Werkzeugen der Künstlichen Intelligenz (KI) benötigen.

### Einführung in die Mengenlehre

Die Grundideen zum Verständnis der reellen Zahlen wurden von Georg Cantor mit der Mengenlehre bereitgestellt. Im Gegensatz zur Zahlentheorie, in der Unendliches als *Potentiales* aufgefasst wird, führt Cantor das *aktual Unendliche* in die Mathematik ein. Während im ersten Falle der Existenzbegriff auf das Sichtbare im finit konstruierten Bild beschränkt ist, postuliert man für das aktual Unendliche, dessen Existenz auch dann, wenn im Bilde nur eine Prozedur dargestellt werden kann, die dieses erzeugt, ohne dass es selbst konkret erscheint. Für die natürlichen Zahlen ( $\mathcal{N}$ ) ergibt dies, als potential unendliche Menge die Darstellung  $\mathcal{N} := 1, 2, 3, 4, \dots$ , wobei die Punkte in ein Unendliches verweisen, das nicht fassbar ist und dem als Nicht-Konkretisiertem kein weiterer Wert angefügt

werden kann. In der Auffassung von  $\mathcal{N}$  als aktual Unendlichem wird mit dem Bild  $\langle 1, 2, 3, 4, \dots \rangle$ ,  $\mathcal{N}$  als vollständig dargestellt vorgestellt, so dass man sich dieses Bild um die Buchstaben  $a, b, c, \dots, z$  erweitert vorstellen kann, zum Bild  $\langle 1, 2, 3, 4, \dots, a, b, c, \dots, z \rangle$ .

**Axiome der Mengenlehre.** Cantor erklärt eine Menge  $M$  als eine beliebige Zusammenfassung von Elementen zu einem Ganzen, wobei entscheidbar ist, ob ein Element  $e$  zu  $M$  gehört ( $e \in M$ ) oder ob  $e$  nicht zu  $M$  gehört ( $e \notin M$ ).

Für Mengen  $M, N$  wird definiert:

$$N \subseteq M \Leftrightarrow (m \in N \Rightarrow m \in M).$$

Die Mengenlehre wird beschrieben durch *Axiome zur Konstruktion von Mengen*: Sind  $M, N$  Mengen und ist  $\eta()$  eine Eigenschaft der Elemente, so existiert die Aussonderungsmenge:

$$M_\eta := \{m \in M \mid \eta(m) \text{ ist wahr}\}$$

und somit

$$M - N := \{m \in M \mid m \notin N\}$$

und

$$M \cap N := \{m \mid m \in M \text{ und } m \in N\},$$

die Vereinigungsmenge

$$M \cup N := \{m \mid m \in M \text{ oder } m \in N\},$$

die Paarmenge

$$M \times N$$

und die Potenzmenge

$$P(M) := \{N \mid N \subseteq M\}.$$

Eine besondere Rolle spielt das *Auswahlaxiom*, das die Auswahl eines Elements  $m$  aus beliebigen Mengen ( $M \neq \emptyset$ ) sichert:

*Auswahlaxiom.* Zu jeder Menge nicht-leerer Mengen  $\{M_i\}_{i \in I}$  existiert eine Menge  $\bar{M}$ , die jeweils genau ein Element  $m_i$  aus jeder dieser Mengen  $M_i$  enthält.

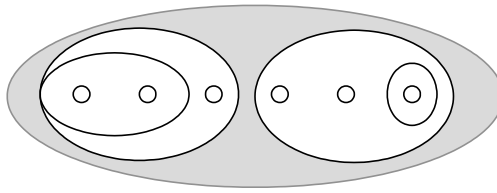


Abbildung 1. Ein System von Mengen

Die Auswahl der Elemente  $m_i$  wird dabei nicht konkret spezifiziert, sondern nur postuliert, und ist damit für unsere bildliche Anschauung nicht eindeutig. Wir können sagen: Durch eine Auswahl der Elemente durch das Auswahlaxiom wird eine *Sichtweise auf das Mengensystem*  $\{M_i\}_{i \in I}$  eingeführt, die dieses konkretisiert. Verschiedene Sichtweisen entsprechen unterschiedlichen Auffassungen.

Während es für das potential Unendliche nur eine einzige Charakterisierung gibt, war Cantors bedeutende Entdeckung, dass mit dem aktual Unendlichen mehrere Größen- oder Mächtigkeitsklassen existieren. Cantor definiert:

$M$  ist gleich mächtig zu  $N$  ( $M \equiv N$ ).  
 $\Leftrightarrow$  Es gibt eine Bijektion  $\Psi : M \rightarrow N$ .

**Eine Grundvorstellung zur Mengenlehre.** Analog zur Veranschaulichung der Aussagen über  $\mathcal{N}$  (Sommer, 2023), können wir uns vorstellen, dass wir ausgehend von einer Grundstruktur mit unendlichen vielen Elementen (z. B.  $\mathcal{N}$ ) diese mit den Konstruktionsaxiomen der Mengenlehre fortlaufend um neugebildete Mengen erweitern, wobei auch Mengen als Mengenelemente zugelassen sind. (Vgl. Abbildung 1, in der Umrandetes zu Mengen zusammengefasst ist.) Da wir diese Konstruktionsschritte *der Reihe nach* ausführen, entsteht, wie bereits Skolem festgestellt hat, ein Modell der Mengenlehre, dessen Objekte der Menge  $\mathcal{N}$  bijektiv zugeordnet werden können. (Außerdem gilt hier  $M \notin M$ .)

Um Aussagen in unserem Modell der Mengenlehre zu beweisen, benötigen wir konkrete Vorstellungen, in denen wir eine Bestätigung der Aussagen erkennen können. Zur Konkretisierung führen wir im Bild darstellbare Prozeduren ein, die aber kein konkretes Bild der beschriebenen Objekte liefern, da sie das Auswahlaxiom verwenden.

In der Mengenlehre ist in manchen Fällen auch erkennbar, dass ein Sachverhalt nicht im Bild darstellbar ist.

**Argument zur Nichtexistenz.** Es kann vorkommen, dass mit den Konstruktionschritten oder Prozeduren,

die zur Darstellung eines Objekts notwendig sind, im Bild eine zusätzliche Information mitentsteht, die der Existenz des Objekts widerspricht.

Mit diesem Argumentationsschema kann z. B. die Existenz einer Menge aller Mengen (der Allmenge) ausgeschlossen werden. Man betrachte Abbildung 1 mit den schwarz eingezeichneten, bisher erzeugten Mengen und füge nun anschließend, in diesem Bild, die grau gezeichnete Allmenge ein. Dann erkennt man sofort, dass die eingefügte Allmenge gar keine Allmenge ist, da sie sich selbst nicht als Element enthält.

Um die Mengen geordnet darstellen zu können, verwenden wir den

*Wohlordnungssatz.* Auf jeder Menge  $M$  kann eine totale Ordnung  $<$  so eingeführt werden, dass zu jeder Teilmenge von  $M$  bezüglich  $<$  ein kleinstes Element existiert.

*Routine zur Definition von  $<$ :*

Initialisierung:  $M_0 := M, M_1 = \emptyset$   
 DO WHILE ( $M_0 \neq \emptyset$ ):  
 Wähle mit dem Auswahlaxiom  $m \in M_0$  und definiere  $m_1 < m$  für alle  $m_1 \in M_1$ ,  
 Entferne  $m$  aus  $M_0$  und setze  $M_1 := M_1 \cup \{m\}$ ,  
 ENDDO

Aufgrund der Auffassung vom aktual Unendlichen wird mit dieser Routine die Totalordnung  $<$  auf ganz  $M$  eingeführt. Wir können uns – wegen des Wohlordnungssatzes – Mengen als Ketten ihrer Elemente vorstellen. Zum Beispiel:

$$\underbrace{(123 \dots 123 \dots)}_1 \underbrace{(123 \dots 123 \dots)}_2$$

$$\underbrace{(123 \dots 123 \dots)}_{3 \dots} \dots \dots 123 \dots$$

**Verschiedene Sichtweisen für Aussagen über die Grundvorstellung.** Wir diskutieren die Existenz einer bijektiven Abbildung  $\psi$  zwischen einer Menge  $M$  mit  $M \equiv \mathcal{N}$  und ihrer Potenzmenge  $P(M)$ .

$M$  können wir als  $\{m_1, m_2, \dots\}$  veranschaulichen, aber um Bilder herzustellen, aus denen wir obige Frage beantworten können, müssen zwei Prozeduren zur Repräsentation der nicht konkret vorstellbaren Objekte betrachtet werden:

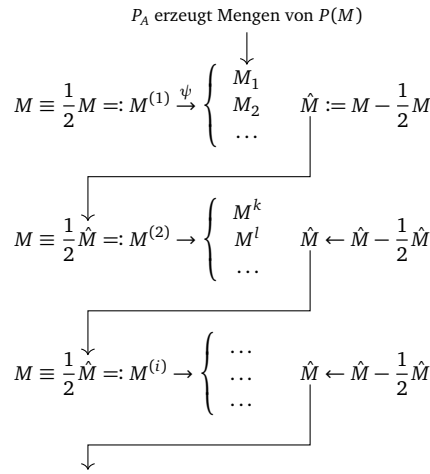


Abbildung 2. Konstruktion von  $\psi$

- Eine Prozedur  $P_\psi$  zur Herstellung einer injektiven Abbildung  $\psi : M \rightarrow P(M)$  und
- eine Prozedur  $P_A$  zur Auswahl der Elemente aus  $P(M)$ .

Wir nehmen zunächst  $P_A$  als konkret spezifiziert an und beschreiben damit die Bildung von  $P_\psi$  (unsere Konstruktion folgt derjenigen zum Banach-Tarski Paradoxon, 1924):  $\frac{1}{2}M$  sei die Menge, die durch Auswahl jedes zweiten Elements aus  $M$  entsteht. In Abbildung 2 ist die Prozedur  $P_\psi$  dargestellt.

**Erster Schritt.** Zunächst werden die Elemente von  $M^{(1)} := \frac{1}{2}M$  Elementen von  $P(M)$  zugewiesen, die von  $P_A$  erzeugt werden. Wegen  $M^{(1)} := \frac{1}{2}M \equiv M$  entspricht dies einer Menge der Mächtigkeit von  $M$ , der Elemente von  $P(M)$  zugewiesen wurden. Durch die Bezeichnungsänderung sind aber die Elemente aus  $\hat{M} := (M - \frac{1}{2}M)$  noch unverbraucht, d. h. noch nicht für die Zuweisung verwendet worden.

**Nächster Schritt.** Sei  $i'$  das Element, das in der Indizierungsmenge  $I$  auf  $i$  folgt.

In diesem Schritt werden die Elemente von  $M^{(i')} := \frac{1}{2}\hat{M}$  zugewiesen, wobei wegen  $M^{(i')} := \frac{1}{2}\hat{M} \equiv M$  wieder ein Block von der Mächtigkeit von  $M$  neuen Elementen aus  $P(M)$  zugewiesen wurde und eine Menge der Mächtigkeit  $M$  der Elemente von  $\hat{M} - \frac{1}{2}\hat{M}$  noch als unverbraucht zurückbleiben.

Wir ändern nun die Bedeutung von  $\hat{M}$  und bezeichnen die Menge der unverbrauchten Elemente  $\hat{M} - \frac{1}{2}\hat{M}$  wieder mit  $\hat{M}$ ; anstelle des Folgeindex  $i'$  schreiben wir wieder  $i$ .

Mit den Bezeichnungsänderungen kann der *Nächste Schritt* erneut ausgeführt werden, womit wieder ein Block der Mächtigkeit von  $M$  Elementen aus  $P(M)$

Urbilder in  $M$  findet. Die Spezifikation von  $\psi$  stoppt, falls die Prozedur  $P_A$  keine weiteren Elemente mehr in  $P(M)$  findet, mit  $\psi(\hat{M}) \equiv P(M)$  für  $\hat{M} \subset M$ .

**Bemerkungen zur Prozedur  $P_\psi$ .** Jedes  $M^{(i)}$  und  $\hat{M}$  hat die Mächtigkeit von  $M$  bzw.  $\mathcal{N}$ , gleichgültig wie oft der ‚nächste Schritt‘ ausgeführt wird. Der Index  $i$  der Prozedur  $P_\psi$  hängt von  $P(M)$  bzw. von  $P_A$  ab. Andere Auswahlfunktionen könnten in  $P(M)$  weitere Elemente sichtbar machen. Wegen der vorausgesetzten Konstruierbarkeit der Objekte kann ja auch  $P(M)$  tatsächlich nicht ‚mehr‘ Darstellungsdetails aufweisen, als in  $\mathcal{N}$ -vielen Schritten erzeugbar sind.

Für die folgende Betrachtung nehmen wir an, es sei  $P_\psi$  vollständig spezifiziert vorgegeben, und wir suchen dazu passend  $P_A$ . Um die Elemente von  $P(M)$  zu beschreiben, verwenden wir deren charakteristische Funktionen  $\epsilon_i : M \rightarrow \{0, 1\}$ , d. h.  $M_i \subset M$  ist beschrieben durch  $M_i := \{m_j \in M \text{ mit } \epsilon_i(j) = 1\}$ .

In Abbildung 3 ist sowohl die Zuweisung  $\psi$  der Elemente  $m_i$  von  $M$  zu diesen charakteristischen Funktionen  $\epsilon_i$  angegeben, als auch die Spezifikation der Mengen  $M_i$  durch die Funktionen  $\epsilon_i$ .

Man erkennt sofort, dass die Teilmenge  $\bar{M} \subset M$  mit  $\bar{\epsilon}(i) := 1$  für  $\epsilon_i(i) = 0$  und  $\bar{\epsilon}(i) := 0$  für  $\epsilon_i(i) = 1$  keinem Element von  $P(M)$  zugewiesen wurde. Da für alle  $i$ :  $\bar{\epsilon}(i) \neq \epsilon_i(i)$  gilt.

Wird  $P_\psi$  als erste Prozedur vorgegeben, so ergibt die Teil-Darstellung von  $P_A$  mit Abbildung 3 eine neue Teilmenge von  $M$ , ohne Urbild bezüglich  $\psi$  in  $M$ .

Mit dem *Argument zur Nichtexistenz* erkennen wir daher, dass hier eine Bijektion  $\psi : M \rightarrow P(M)$  nicht existiert.

In Abbildung 3 entsteht  $P(M)$  mit einer Ordnungsstruktur, die von  $M$  durch  $\psi$  auf  $P(M)$  induziert wird

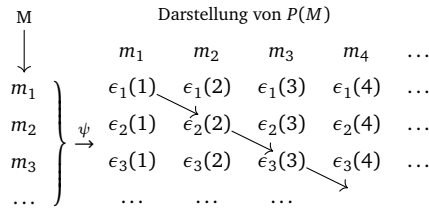


Abbildung 3. Diagonalisierung zur Konstruktion von  $\bar{M}$

und die verschieden ist von derjenigen in Abbildung 2. Die zusätzliche Struktur, die entsteht, wenn wir zuerst  $P_\psi$  vorgeben und dann erst  $P_A$  spezifizieren, ermöglicht das Diagonalisierungsargument für dieses  $P_A$ , auch in einer Vorstellungswelt, die nur aus abzählbar vielen Objekten besteht. Da  $P_A$  oder  $P_\psi$  aber nicht konkret spezifiziert sein müssen, sind deren Werte nicht notwendigerweise im Bild darstellbar und somit ist die Realisierung der Abbildungen 2 und 3 nicht garantiert. Unsere Überlegungen zeigen eine schwache Form des *Satzes von Paul Cohen* (2002):

Der Mächtigkeitbegriff ist durch die Axiome der Mengenlehre nicht eindeutig konkretisiert.

### Die reellen Zahlen

Das Verstehen der reellen Zahlen ( $\mathcal{R}$ ) und ihre *Andersartigkeit* zu den natürlichen Zahlen ( $\mathcal{N}$ ) ist mit den Grundvorstellungen zur Mengenlehre möglich. Die Elemente von  $\mathcal{R}$  sind Grenzwerte in  $\mathcal{Q}$ , die mit verschiedenen Anschauungsweisen entstehen. Jedes  $r \in \mathcal{R}$  basiert auf einer nicht notwendigerweise konkretisierten Auswahlfunktion. Die Axiome von  $\mathcal{R}$  fixieren daher nur die Menge der Aussagen für die so konstruierten Objekte, nicht aber deren Konstruktion selbst. Beispielfhaft betrachten wir hier nur zwei der vielen Konstruktionsmöglichkeiten für reelle Zahlen, es gibt viele weitere, häufig über den Begriff der Cauchy-Folgen (Weiss, 2015):

(A)  $\mathcal{R} \equiv \mathcal{D}\mathcal{S} :=$  Menge der Dedekind'schen Schnitte in den rationalen Zahlen ( $\mathcal{Q}$ ), wobei ein Element  $x$  in  $\mathcal{D}\mathcal{S}$  beschrieben wird durch:  $x \equiv (\underline{U}, \bar{U})$  mit  $\underline{U}, \bar{U} \subset \mathcal{Q}$  und für alle  $y \in \underline{U}, z \in \bar{U}$  gilt:  $y < z$ .

- Jede der folgenden Auswahlfunktionen aus Intervallen in  $\mathcal{Q}$  ergibt ein  $x$ :

(Initialisierung:)

Wähle  $q_{u,1} \in \mathcal{Q}$  und  $q_{o,1} \in (q_{u,1}, q_{u,1} + 1)$ .

(Schritt  $i \rightarrow i + 1$ :)

Wähle:  $q_{u,i+1} \in (\max\{q_{u,i}, q_{o,i} - \frac{1}{i}\}, q_{o,i})$  und  $q_{o,i+1} \in (\max\{q_{u,i+1}, q_{o,j} - \frac{1}{i}\}, q_{o,j})$

Definiere:  $\underline{U} := \bigcup_{i=1}^{\infty} \{q \in \mathcal{Q} \mid q < q_{u,i}\}$ , und  $\bar{U} := \bigcup_{i=1}^{\infty} \{q \in \mathcal{Q} \mid q > q_{o,i}\}$ .

(B)  $\mathcal{R} \equiv \mathcal{D}\mathcal{B} :=$  Menge der unendlichen Dualbrüche, wobei ein Dualbruch beschrieben wird durch:  $\pm n, \epsilon_1 \epsilon_2 \epsilon_3 \dots$  mit  $n \in \mathcal{N} \cup \{0\}$  und  $\epsilon_i \in \{0, 1\}$ , und für kein  $j \in \mathcal{N}$  gilt  $1 = \epsilon_j = \epsilon_{j+1} = \epsilon_{j+2} = \dots$

- Die Elemente von  $\mathcal{D}\mathcal{B}$  entstehen durch eine Auswahl aus  $\{0, 1\}$  für jede Ziffer, wobei  $\dots 0111\dots$  ersetzt wird durch  $\dots 1000$ .

Aus Darstellung (B) erkennen wir direkt:

$\mathcal{R}$  ist gleich mächtig zu  $P(\mathcal{N})$  oder den Abbildungen von  $\mathcal{N}$  nach  $\{0, 1\}$ , denn die Menge der ausgeschlossenen Dualbrüche, zu denen ein  $j \in \mathcal{N}$  existiert mit  $1 = \epsilon_j = \epsilon_{j+1} = \epsilon_{j+2} = \dots$ , ist gleich mächtig zu  $\mathcal{Q}$  bzw.  $\mathcal{N}$ .

Zu den reellen Zahlen gibt es also verschiedene Möglichkeiten der Konkretisierung und diese *Nichteindeutigkeit der Konkretisierung* von  $\mathcal{R}$  hat die Konsequenz, dass der Nachweis von unterschiedlichen Aussagen über  $\mathcal{R}$  häufig unterschiedliche Darstellungen von  $\mathcal{R}$  erforderlich macht.

Betrachten wir als Beispiel den

*Zwischenwertsatz.* Sei  $f : [-1, 1] \rightarrow [-1, 1]$  stetig mit  $f(-1) = -1$  und  $f(1) = 1$ , dann existiert  $x \in (0, 1)$  mit  $f(x) = 0$ .

Wird  $\mathcal{R}$  mittels  $\mathcal{D}\mathcal{S}$  dargestellt, so ergibt  $x = (\underline{U}, \bar{U})$  mit  $\underline{U} := \{y \in \mathcal{Q} \mid y < -1\}$  oder für alle  $z \in [-1, 1]$  mit  $z < y$  gilt:  $f(z) < 0$  und  $\bar{U} := \mathcal{Q} - \underline{U}$  einen Wert  $x$  mit  $f(x) = 0$ .

In  $\mathcal{D}\mathcal{B}$  wäre dagegen die Lösung  $x$  selbst approximativ nicht spezifizierbar, da für eine nicht konkretisierte Funktion  $f$  nicht entschieden werden kann, ob diese nicht bereits vor dem Erreichen eines Schnittpunkts die Achse berührt hat.

Während  $x$  in der Sprache  $\mathcal{D}\mathcal{S}$  niedergeschrieben werden kann, ist dies in  $\mathcal{D}\mathcal{B}$  unmöglich.

Die Nichteindeutigkeit der Konkretisierung von  $\mathcal{R}$  überträgt sich auf Theorien, die auf  $\mathcal{R}$  begründet sind, z. B. auf die Integrationstheorie, mit dem

Lebesgue-Integral und dem dazu inkompatiblen Cauchy-Hauptwert-Integral, oder auf die *Theorie der partiellen Differentialgleichungen*. Da letztere in der Physik besonders wichtig ist, wollen wir uns ein Beispiel aus der Hydrodynamik genauer ansehen:

Das Verhalten einer inkompressiblen Strömung wird beschrieben durch die Navier-Stokes'schen Gleichungen:

$$\underbrace{\frac{\partial \vec{u}}{\partial t}}_I = -\underbrace{(\vec{u}\nabla)\vec{u}}_{II} + \underbrace{\frac{1}{\rho}\nabla p}_{III} + \underbrace{\nu\nabla^2\vec{u}}_{IV} \quad (1)$$

( $\vec{u}$  beschreibt die Strömungsgeschwindigkeit,  $\rho$  die Dichte,  $p$  den Druck und  $\nu$  die Zähigkeit der Flüssigkeit.)

In anschaulichen Begriffen bedeutet diese Gleichung:

- I. die Änderung der Strömungsgeschwindigkeit  $\vec{u}$  in einem Punkt  $\vec{x}$  wird bewirkt durch
- II. den Transport von Teilen der Strömung aus Nachbarpunkten zum Punkt  $\vec{x}$  und
- III. der Druck-Wirkung auf die Strömungsteile im Punkt  $\vec{x}$  und
- IV. der Difussion oder Vermischung der Teilchengeschwindigkeiten in einer kleinen Umgebung des Punktes  $\vec{x}$ .

Die Navier-Stokes-schen Gleichungen sind damit eine mathematische Darstellung lokal beschreibbarer Phänomene in der Flüssigkeit oder das Ergebnis einer *lokalen Sichtweise*. Sie sind mit der Operator-Splitting-Methode oder direkt durch Verwendung der angegebenen Veranschaulichung und der Lattice-Boltzmann-Methode approximativ lösbar.

Die Lösungen der Navier-Stokes'schen Gleichung zeigen aber nicht alles, was wir über turbulente Strömungen wissen können. Kolmogoroff hat zur Herleitung seines Wirbelzerfallsgesetzes eine *globale Sichtweise* verwendet. Er betrachtet Wirbel als *Selbststabilisierende Strukturen* und postuliert ein *Analogie-Gesetz*:

Große Wirbel verhalten sich gegenüber mittelgroßen Wirbeln völlig analog wie mittelgroße Wirbel sich gegenüber kleinen Wirbeln verhalten.

Kolmogoroff's Herleitung des Wirbelzerfallsgesetzes aus dem Analogie-Gesetz macht uns nochmals die *Bedeutung der Sichtweisen* für die Theorien, die auf  $\mathcal{R}$  basieren, deutlich. Wir werden im Folgenden feststellen, dass auch zur Entwicklung und zum Einsatz der Werkzeuge der Künstlichen Intelligenz die Entwicklung von Sichtweisen eines der wichtigsten Hilfsmittel darstellt.

### Verstehenshorizonte der KI

In der Mengenlehre und in den Theorien, die auf  $\mathcal{R}$  basieren werden Aussagen dadurch begründet, indem sie mittels geeigneter Sichtweisen in Korrespondenz gesetzt werden zu finit Konzipiertem, das dann diese Begründung liefert (vgl. Abbildung 4). Paul Cohen (2002) schreibt:

The only reality we truly comprehend is that of our own experience ... The laws of the infinite are extrapolations of our experience with the finite.

Durch die Erschaffung geeigneter Anschauungen wird sichergestellt, dass mit der Prozedur von Abbildung 4 das Bestätigung findet, was entweder mathematisch interessante Theorien liefert, oder dem Beobachteten, in der realen Welt entspricht. Das Verstehen der Mathematik erlaubt verschiedene Auffassungen. Wir können  $\mathcal{R}$  als ‚größer‘ wie  $\mathcal{N}$  oder als ‚andersartig‘ zu  $\mathcal{N}$  auffassen. Im ersten Falle erhalten wir Cantors Mächtigkeitsordnung und im zweiten Falle ein Verständnis dafür, warum Aussagen der Theorien, die auf  $\mathcal{R}$  basieren, mehrere Sichtweisen zu ihrer Begründung benötigen.

Die Konzipierung eines Verstehenshorizonts ist aber auch ein Grundproblem der KI. Dies wird im Beitrag von Schorcht et al. (2024) verdeutlicht. Um sicher einen Text in eine Bilddarstellung umzusetzen, müssten wir in der Lage sein, ChatGPT mitzuteilen, in welchem Verstehenshorizont dieses Bild eingebettet sein soll. Dafür aber gibt es keine eindeutig spezifizierte Kommunikationsmöglichkeit, denn die Spezifikation einer solchen würde die Generalität von ChatGPT stark einschränken. Wir müssen daher, zu jedem *Werkzeug mit Genereller Künstlicher Intelligenz*, auf einer übergeordneten Ebene einen Verstehenshorizont schaffen,

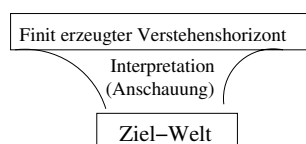


Abbildung 4. Ziel-Welt sind Mengen,  $\mathcal{R}$  oder eine Computer-Hardware

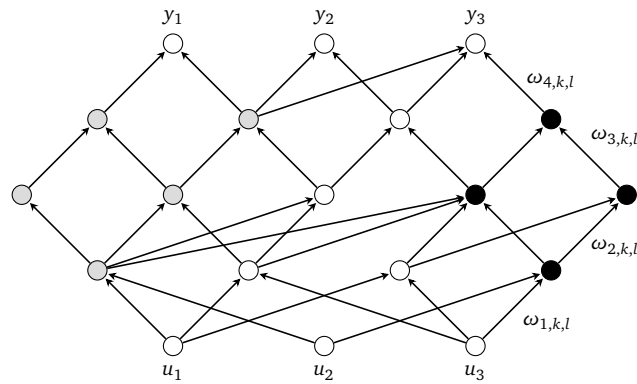


Abbildung 5. Neuronales Netz aufgefasst als Gruppen

der uns das Werkzeug verstehbar macht, ohne dieses vollständig zu erfassen. Das was wir erfassen, bleibt immer auch von unserer Sichtweise abhängig.

Beispielhaft wollen wir diese Methode am Verstehen der Lernfähigkeit eines Neuronales Netzes darstellen (Sommer & Kroll, 2023). Das Netz erhält Lernbeispiele

$$\{(\vec{u}_j, \vec{y}_j) \text{ mit } j = 1, \dots, K\}, \text{ mit } \vec{u}_j \in \mathcal{R}^{n_0}$$

und  $\vec{y}_j \in \mathcal{R}^{n_L}$ .

Für die Werte,  $n_{ik}$ , die in den Knoten des Netzes gespeichert werden, gilt das Gleichungssystem:

$$n_{(i+1),k} = f \left( \sum_{l=1}^{\rho_i} \omega_{i,k,l} \times n_{i,l} + \alpha_{i+1,k} \right),$$

wobei  $f$  und  $\rho_i \in \mathcal{N}$  geeignet vorgegeben werden.

Lernen bedeutet, dass das Netz durch Auffinden geeigneter Leitwerte  $\omega_{i,k,l}$ , eine Funktion  $(\vec{u}_j \rightarrow \vec{y}_{Netz,j})$  realisiert, die  $\vec{y}_{Netz,j} \approx y_j$  erfüllt. Das Lernen des Netzes wird üblicherweise mit dem Back-Propagation-Algorithmus durchgeführt.

Um den Lernprozess des Netzes anschaulich zu erfassen, betrachten wir Abbildung 5 für das Lernbeispiel  $(\vec{u}_i, \vec{y}_i)$  mit der folgenden Sichtweise:

- Das Netz wird aufgefasst als Menge unterschiedlicher Knoten-Gruppen.
- Der Lernalgorithmus wählt die grauen Akteure aus, die den größten Beitrag bezüglich aller Gruppen zur Darstellung des Lernbeispiels  $(\vec{u}_i, \vec{y}_i)$  leisten.
- Die Zuleitungen zur gewählten grauen Gruppe werden für  $\vec{u}_i$  verstärkt und diejenigen zu den anderen schwarzen Gruppen abgeschwächt.
- die Anpassung der ausgewählten grauen Gruppe an das Lernbeispiel  $(\vec{u}_i, \vec{y}_i)$  und an zu diesem ähnliche Beispiele wird verbessert, wohingegen die anderen schwarzen Gruppen zum Erlernen anderer Lernbeispiele frei bleiben.

Unsere Argumentation zeigt, dass die Fähigkeit zur Entwicklung von Sichtweisen, nach dem Vorbild der Mathematik, für die KI unverzichtbar ist. Wissen, auch dasjenige das die KI liefert, ist nicht absolut, es muss bezüglich einem Verstehenshorizont begründet sein. Aber welche Denkweisen der Mathematik machen die KI verstehbar? Unser Beitrag soll zur Diskussion dieser Frage anregen.

*Danksagung.* Mein Dank gilt Vincent J. Sommer und Friedrich T. Sommer für hilfreiche Diskussionen.

#### Literatur

- Banach, S. & Tarski, A. (1924). Sur la décomposition des ensembles de points en parties respectivement congruentes. *Fund. Math.*, 6, 244–277.
- Cohen, P. (2002). The discovery of forcing. *Rocky Mountain Journal of Mathematics*, 32(4), 1071–1100.
- Schorcht, S, Baumanns, L., Buchholtz, N., Huget, J., Peters, F. & Pohl, M. (2024). Lernt die KI nun Sehen und Zeichnen? Chancen und Herausforderungen der Bildgenerierung und Bildinterpretation mit ChatGPT für die mathematische Forschung. *Mitteilungen der GDM*, 116, 22–29.
- Sommer, H. (2023). Verstehen der Zahlen. *Mitteilungen der GDM*, 115, 39–43.
- Sommer, H. & Kroll, A. (2023). Zum Einsatz von KI-Methoden zur Lösung von Problemen der Regelungstechnik. *Proceedings: 33. Workshop Computational Intelligence*. Berlin, 23.–24. November 2023.
- Weiss, I. (2015). *The real numbers – A survey of constructions*. [arXiv:1506.03467v1](https://arxiv.org/abs/1506.03467v1)

Hanns Sommer, Universität Kassel  
[hjsommer@uni-kassel.de](mailto:hjsommer@uni-kassel.de)