

Die Sprache ist das Dasein des Denkens – Warum Sprachmodelle trotzdem nicht denken

„Artificial intelligence poses a double challenge to the human mind. Not only will the most advanced models soon be able to think better than people, but AI has consequences for humanity which are so uncertain, so potentially vast and approaching at such a rapid pace that even the best brains flinch“

The Economist, July 25, 2026

„Where are they, the great next philosophers, the equivalents of Kant or Wittgenstein or even Aristotle? I think we’re going to need that to help navigate society to that next step, because I think A.G.I. and artificial superintelligence are going to change humanity and the human condition.“

Demis Hassabis, DeepMind co-founder, Big Technology Podcast, January 23, 2025

Wer wissen will, ob Künstliche Intelligenz denkt, muss zunächst wissen, was Denken überhaupt ist. Genau daran fehlt es gegenwärtig – in der medialen Berichterstattung ebenso wie in weiten Teilen der Wissenschaft und bei den Herstellern der Sprachmodelle selbst. Man schreibt KI Intelligenz, Verstehen, Lernen, Entscheiden oder gar Empathie zu, ohne zuvor zu klären, was mit diesen Begriffen eigentlich bezeichnet ist. Wer sprachlich souverän erscheint, dem wird Verstehen zugeschrieben; wer Probleme löst, gilt als intelligent; und wer „ich“ sagt, scheint bereits ein Bewusstsein zu besitzen. Statt die Sache begrifflich zu bestimmen, ist man von der beeindruckenden Leistung geblendet und schließt auf einen ebenso beeindruckenden Maschinengeist.

Hegel ist für eine Kritik dieser Fehltritte besonders ergiebig, weil er Intelligenz nicht als messbare Leistungsfähigkeit, sondern als den Prozess bestimmt, in dem ein Subjekt die Welt über Anschauung, Vorstellung, Sprache und begriffliches Denken zu seinem eigenen Inhalt macht¹. Dazu muss man ihn nicht einmal vollständig vom Kopf auf die Füße stellen, sondern lediglich seine idealistische Metaphysik abstreifen: Nicht der Geist findet sich in der Welt wieder, sondern ein körperliches, gesellschaftliches und sprachbegabtes Wesen bildet Begriffe, um eine unabhängig von ihm bestehende Welt zu erkennen und zu verändern.

Im Folgenden führt der Weg zur KI zunächst scheinbar von ihr weg: Von ein paar ganz grundsätzlichen Bestimmungen von Sprache, Intelligenz, Denken, zu Bienen und Primaten, Säuglingen und sprachlich isolierten Kindern, gehörlosen Menschen und Gebärdensprachen, und den Inuit, Homer und der Frage, ob verschiedene Sprachen unterschiedliche Formen des Denkens begünstigen. Der Vergleich dieser verschiedenen Ausprägungen kognitiver und sprachlicher

1 Insbesondere in der Enzyklopädie der Philosophischen Wissenschaften (1817)

Fähigkeiten macht sichtbar, was Wahrnehmung, Erinnerung, Kommunikation und Symbolgebrauch bereits leisten – und worin der Übergang zu Sprache und begrifflichem Denken besteht. Vor diesem Hintergrund zeigt die Analyse dessen, was ein Large Language Model ist und tut: Es verdichtet die sprachlichen Resultate menschlicher Welterfahrung zu begriffslosen statistischen Beziehungen. Es kann täuschend verständlich sprechen, aber besitzt Sprache nur als berechenbare Form: Empfinden, Erkennen, Verstehen, und Entscheiden simuliert es, indem es aus den sprachlichen Sedimenten menschlicher Wahrnehmungen, Handlungen und Erfahrungen statistisch neue Äußerungen extrapoliert – ohne den Weg von der Welt zum Begriff jemals selbst vollzogen zu haben.

Damit beantwortet sich auch die Frage nach der Gefährlichkeit der KI. Die Gefahr liegt jedoch nicht in einem LLM, das irgendwann Bewusstsein entwickelt, eigene Zwecke setzt und sich gegen den Menschen erhebt. Als bloßes Sprachmodell kann es durch noch so viel Training und Skalierung kein Bewusstsein hervorbringen: Es bleibt ein System probabilistischer Zeichenverarbeitung ohne eigenen Weltbezug, eigenen Standpunkt und selbstbestimmte Zwecke. Kreuzgefährlich wird es dennoch – durch seinen menschlichen Gebrauch als gehorsames Instrument staatlicher, militärischer, industrieller und krimineller Interessen. Nicht der Aufstand der Maschinen ist das Problem, sondern die Geschwindigkeit, Reichweite und Effizienz, mit der sie bestehende Macht-, Gewalt- und Kontrollapparate bedienen und deren Zwecke dabei durchsetzen.

Diese Argumentation steht in deutlichem Gegensatz zu verbreiteten Positionen in Philosophie, Psychologie, Kognitionswissenschaft und KI-Forschung und fordert daher grundsätzlichen Widerspruch heraus. Am Ende des Textes werden die wichtigsten Einwände in ihrer stärksten Form dargestellt und anschließend erwidert.

Was ist Bewusstsein?

Bewusstsein bezeichnet eine Form des Weltbezugs, in der Umweltinhalte nicht bloß registriert, unterschieden und verhaltenswirksam verarbeitet werden, sondern einem **empfindenden** Subjekt als von ihm unterschiedene Gegenstände **erscheinen**. Bewusstsein setzt daher Empfindung voraus, geht aber über sie hinaus: Nicht nur ein Zustand wird erlebt, sondern etwas wird als Gegenstand wahrgenommen. Nach gegenwärtigem Stand ist eine solche Form des Weltbezugs nur bei Lebewesen nachgewiesen.

Die bloße Fähigkeit eines Systems, Umweltinformationen aufzunehmen, zu unterscheiden und verhaltenswirksam zu verarbeiten, genügt dafür nicht. Sensorische Registrierung ist nicht Empfindung, Objekterkennung nicht gegenständliches Bewusstsein. Auch ein technisches System kann Reize verarbeiten und Objekte klassifizieren, ohne dass damit gezeigt wäre, dass ihm überhaupt etwas als Gegenstand erscheint².

Ein bewusster Organismus erlebt also nicht bloß einzelne Reize oder innere Zustände, sondern sieht etwa einen Baum, verfolgt ein Beutetier oder erkennt einen Artgenossen wieder. Bewusstsein beginnt dort, wo einem Subjekt etwas als ein von ihm unterschiedener Gegenstand erscheint. Es **setzt damit weder Sprache noch begriffliches Denken voraus**; auch Tiere können eine

2 Eine Kamera wird durch Licht kausal verändert; ein Bildklassifikator kann daraufhin Objekte unterscheiden. Beides zeigt Informationsverarbeitung, aber noch kein subjektives Erleben und keinen gegenständlichen Weltbezug.

gegenständlich gegliederte Umwelt haben, einzelne Objekte wiedererkennen und ihr Verhalten auf sie ausrichten³.

Eine weitere Stufe ist erreicht, wenn das Subjekt nicht nur auf Gegenstände bezogen ist, sondern zugleich sich selbst als dasjenige erfasst, das sich auf sie bezieht. Aus dem Bewusstsein eines Gegenstandes wird damit zugleich ein Bewusstsein der eigenen Beziehung zu diesem Gegenstand: **Selbstbewusstsein**. Während im Bewusstsein eine Welt dem Subjekt gegenübersteht, weiß das selbstbewusste Subjekt zugleich von sich als demjenigen, dem sie gegenübersteht.

Bewusstsein gründet in der Empfindung, ist aber nicht mit ihr identisch. Empfindung liegt vor, wenn ein Reiz nicht nur verarbeitet wird, sondern mit einem subjektiven Erleben verbunden ist. Bewusstsein geht darüber hinaus: Das Erfahrene erscheint dem Subjekt als etwas von ihm Unterschiedenes.

Von der Wahrnehmung zum Begriff

Organismen – von der Ameise über die Maus bis zum menschlichen Säugling – erschließen ihre Umwelt zunächst über die Sinne. **Wahrnehmung** beruht auf der Aufnahme und neuronalen Verarbeitung äußerer Reize, etwa visueller, auditorischer, olfaktorischer oder taktiler Signale. Dabei handelt es sich jedoch nicht um eine bloße Abbildung der Außenwelt, sondern um eine selektive und organismusspezifische Verarbeitung sensorischer Informationen. In dieser grundlegenden Hinsicht bestehen Gemeinsamkeiten zwischen sehr unterschiedlichen Tierarten und dem Menschen. Auch technische Systeme (Computer) können Reize registrieren und verarbeiten; ob man dies bereits als Wahrnehmung gelten lassen will, hängt jedoch davon ab, ob man den Begriff lediglich funktional oder auch im Sinne subjektiven Erlebens verstehen will.

Wahrnehmung umfasst Empfindung und Anschauung: Empfindungen liefern die qualitativen Inhalte; Anschauung erfasst sie als Bestimmungen gegenwärtiger Gegenstände.

Am Anfang der geistigen Aneignung der Welt steht die sinnliche und affektive **Empfindung**.⁴ Organismen reagieren auf Farben, Töne, Berührungen, Wärme, Schmerz, Lust und Unlust. Solche Empfindungen treten beim entwickelten Wahrnehmen jedoch gewöhnlich nicht isoliert auf. Die Empfindung bildet den qualitativen Inhalt der Wahrnehmung⁵.

3 Ein Tier kann etwa Angst erleben, deren Anlass erkennen und darauf reagieren. Der Mensch kann darüber hinaus wissen und aussprechen, dass er Angst hat, nach ihrer Ursache fragen und beurteilen, ob sie angemessen ist. Spezifisch menschlich ist damit nicht Bewusstsein überhaupt, sondern dessen sprachlich vermittelte und begrifflich reflektierte Form. .

4 Hegels funktionale Bestimmungen werden im folgenden nicht entlang seiner systematischen Reihenfolge referiert, sondern neu geordnet: Empfindung → Anschauung → Vorstellung → Denken).

5 Äußere Empfindungen wie Farben, Töne oder Gerüche werden gewöhnlich nicht isoliert und zeitlich vor den Gegenständen wahrgenommen, denen sie zukommen. Wir sehen nicht zunächst „Rot“ und setzen daraus anschließend einen Apfel zusammen, sondern nehmen unmittelbar einen roten Apfel wahr. Die Unterscheidung von Empfindung und Anschauung bezeichnet daher keine zeitliche Abfolge, sondern verschiedene Momente desselben Wahrnehmungsvorgangs. Die Empfindung ist dessen qualitativer Inhalt – etwa rot, warm oder laut –, während die Anschauung diesen Inhalt räumlich und zeitlich als Bestimmung eines gegenwärtigen Gegenstands erfasst. Erst in der begrifflichen Reflexion lässt sich das empfundene Rot vom angeschauten roten Gegenstand unterscheiden. O-Ton Hegel: „„Die Empfindung ist die Form des dumpfen Webens des Geistes in seiner bewußt- und verstandlosen Individualität.“ (Enzyklopädie, § 400)

Während in der Empfindung noch kein Gegenstand als solcher gesetzt ist, wird der Empfindungsinhalt in der **Anschauung** räumlich und zeitlich als Bestimmung eines gegenwärtigen Objekts erfasst. Das Nervensystem erfasst ihn unmittelbar und in seiner konkreten, sinnlichen Erscheinung. Sie richtet sich somit auf das Einzelne im Hier und Jetzt und stellt eine erste Form dar, in der sich ein Organismus seine Umwelt erschließt. Bei Säuglingen und den meisten Tierarten kann die Anschauung (ebenso wie die Empfindung) unmittelbar handlungswirksam werden. Sie löst dann beispielsweise reflexartige, instinktive, erlernte oder situationsgebundene Reaktionen aus, ohne dass der wahrgenommene Gegenstand bereits als allgemeiner Begriff erfasst und ausdrücklich reflektiert wird. So kann die Wahrnehmung einer Blüte bei einer Biene ein komplexes Such-, Orientierungs- und Sammelverhalten auslösen, ohne dass die Biene die Blüte im menschlichen Sinne begrifflich als „Blüte“, „Pflanze“ oder „Nahrungsquelle“ bestimmt.

In der **Vorstellung** werden die in der Anschauung gewonnenen Inhalte aus ihrer unmittelbaren Bindung an das Hier und Jetzt gelöst. Ein zuvor wahrgenommener Gegenstand kann erinnert und innerlich vergegenwärtigt werden, auch wenn er gegenwärtig nicht anwesend ist. Diese Fähigkeit, frühere Wahrnehmungsinhalte erneut aufzurufen, lässt sich als **reproduktive Einbildungskraft** bezeichnen. So kann etwa die Vorstellung eines Baumes entstehen, der zu einem früheren Zeitpunkt gesehen wurde.

Vorstellungen können darüber hinaus aufgrund gemeinsamer Merkmale, wiederkehrender Zusammenhänge oder individueller Erfahrungen miteinander verknüpft werden („**assoziative Einbildungskraft**“⁶). Solche assoziativen Beziehungen verbinden beispielsweise die Vorstellung eines Baumes mit derjenigen eines Apfels. Im Unterschied zum Begriff bleibt die Vorstellung an konkrete Erfahrungen und subjektive Perspektiven gebunden. Sie ist häufig anschaulich oder bildhaft, muss jedoch nicht notwendig in Form eines inneren Bildes auftreten⁷. Zudem bleibt sie vergleichsweise unbestimmt, da sie ihre Inhalte noch nicht ausdrücklich nach allgemeinen Merkmalen ordnet und begrifflich fixiert.

Auch Tiere mit begrenzteren kognitiven Fähigkeiten verfügen über mentale Repräsentationen, in denen frühere Erfahrungen gespeichert sind und aus denen Erwartungen über zukünftige Ereignisse hervorgehen, die in diesem weiten Sinn als Vorstellungen bezeichnet werden könnten. Bei Computern wäre diese Bezeichnung dagegen nur in einem funktionalen und übertragenen Sinn angemessen: Sie können gespeicherte Repräsentationen abrufen und miteinander verknüpfen, ohne dass dies einem subjektives Erleben oder einer inneren Vorstellung entspräche.

6 Anlass der Assoziation ist eine Anschauung oder Vorstellung. Welche Merkmale dabei hervortreten und welche weiteren Vorstellungen dadurch aufgerufen werden, hängt zunächst von der Aufmerksamkeit, Erfahrung und Willkür des Subjekts ab. Darin liegt der Mangel des assoziativen Denkens: Die Verbindung ergibt sich nicht notwendig aus dem wesentlichen Inhalt der Sache, sondern häufig aus zufälligen, äußerlichen oder bloß subjektiv hervorgehobenen Gemeinsamkeiten. Gleichwohl bildet die assoziative Einbildungskraft eine wichtige Vorstufe des begrifflichen Denkens. Indem das Subjekt Beziehungen zwischen Vorstellungen bemerkt, entdeckt es Identitäten, Ähnlichkeiten und wiederkehrende Merkmale. Auf diese Weise entstehen allgemeine Vorstellungen, die über den einzelnen Gegenstand hinausweisen. Dauerhaft festgehalten und allgemein verfügbar werden sie, indem sie einen Namen erhalten (siehe weiter unten Gedächtnis und Sprache).

7 Vorstellungen müssen beim Menschen nicht bildhaft sein. Menschen mit Aphantasie können Inhalte erinnern, verarbeiten und begrifflich durchdenken, ohne sie als bewusste innere Bilder zu erleben. Sie zeigen damit anschaulich, dass Denken nicht auf dem Aufrufen visueller Vorstellungen beruht.

In der **Erinnerung** löst das Subjekt den Inhalt der Anschauung aus seiner unmittelbaren Bindung an die räumlich-zeitliche Gegenwart. Das Angesehene wird zur Vorstellung und damit zu einem verinnerlichten Inhalt des Subjekts, den dieses auch unabhängig von der erneuten Anwesenheit des Gegenstandes hervorbringen kann.

Empfindung, Wahrnehmung, Erinnerung und Vorstellung sind noch nicht Intelligenz im engeren Sinn. Sie bilden vielmehr deren kognitive Voraussetzungen. Auch Tiere können wahrnehmen, Erlebtes erinnern, Gegenstände wiedererkennen und Erwartungen bilden, ohne sich die Welt deshalb begrifflich anzueignen. Der entscheidende Übergang vollzieht sich erst dort, wo Vorstellungen durch allgemein verwendbare Zeichen festgehalten, sprachlich miteinander vermittelt und schließlich als Begriffe bestimmt werden.

Was ist Intelligenz?

Intelligenz bezeichnet die Tätigkeit, durch die sich ein sprach- und selbstbewusstes Subjekt die zunächst äußerlich gegebene Welt begrifflich aneignet. Sie setzt Wahrnehmung, Vorstellung und Erinnerung voraus, geht aber über sie hinaus: Der einzelne Gegenstand wird nicht mehr nur wahrgenommen oder vergegenwärtigt, sondern durch Sprache unter allgemeine Bestimmungen gebracht, mit anderen Gegenständen und Begriffen in Beziehung gesetzt und schließlich seinem Begriff nach erkannt.

Das **Gedächtnis** bildet dabei eine wesentliche Voraussetzung dafür, dass sprachliche Zeichen stabil mit bestimmten Bedeutungen verbunden und wiederholt verwendet werden können. Die Intelligenz betätigt sich insofern als zeichenbildende Tätigkeit: Sie verknüpft Vorstellungen mit Lauten, Gebärden oder Schriftzeichen und macht sie dadurch dauerhaft verfügbar und kommunikativ vermittelbar.

Indem solche Zeichen im Gedächtnis bewahrt, wiedererkannt und nach gemeinsamen Regeln verwendet werden, entsteht **Sprache** als ein überindividuelles System. Dieses steht dem einzelnen Sprecher als bereits bestehende Ordnung von Zeichen, Bedeutungen und grammatischen Strukturen gegenüber und muss von ihm im sozialen Gebrauch erlernt werden.

Intelligenz ist die geistige Aneignung eines Inhalts durch ein sich selbst wissendes Subjekt. Es ist die Tätigkeit, durch die ein Subjekt einen zunächst als äußerlich gegeben erscheinenden Inhalt aufnimmt, verinnerlicht und begrifflich zu seinem eigenen macht. Dies geschieht durch Anschauung, Vorstellung und Denken. Der Gegenstand wird als einzelner, gegenwärtiger Inhalt erfasst („erkannt“), der Inhalt wird vom unmittelbaren Hier und Jetzt gelöst, erinnert, bildlich vergegenwärtigt („vorgestellt“) und durch Zeichen und Namen festgehalten („Sprache“). Der Gegenstand kann dann im Denken nicht mehr nur angeschaut oder vorgestellt, sondern in seiner Allgemeinheit und seinem Begriff erkannt werden⁸.

8 Im gewöhnlichen Sprachgebrauch und in der vergleichenden Kognitionsforschung wird auch kognitiv hoch entwickelten Tieren Intelligenz zugesprochen. Gemeint sind damit meist Fähigkeiten wie Lernen, Problemlösen, flexible Verhaltensanpassung, Gedächtnis, Planung oder Werkzeuggebrauch. Ob Tiere als intelligent gelten, hängt somit wesentlich davon ab, welcher Begriff von Intelligenz zugrunde gelegt wird. Fasst man Intelligenz funktional als erfolgreiche Anpassungs- und Problemlösungsfähigkeit, sind viele Tierarten zweifellos intelligent. Versteht man darunter hingegen – wie hier – die geistige Aneignung der Welt durch sprachlich vermitteltes begriffliches Denken

Der rationale Gehalt von **Hegels Intelligenzbegriff**⁹ bleibt erhalten, wenn man ihn seines Idealismus entkleidet. Das Subjekt (bei Hegel: „der theoretische Geist“) registriert die Welt nicht bloß, sondern eignet sie sich an, indem es Wahrnehmungen verinnerlicht, ordnet und begrifflich bestimmt. Im Begriff findet der Mensch jedoch nicht, wie Hegel argumentiert, die eigene Vernunft als verborgene Struktur der Wirklichkeit wieder. Begriffe sind keine in den Dingen selbst wohnenden Gedankenformen, sondern natürlich, historisch und gesellschaftlich entstandene Formen der geistigen Aneignung, mit denen Menschen eine unabhängig von ihnen bestehende materielle Welt erfassen. Hegels Intelligenzbegriff ist daher von seiner idealistischen Metaphysik zu befreien: Nicht der Geist erkennt sich in der Welt wieder, sondern ein materielles, gesellschaftliches und sprachbegabtes Wesen bildet Begriffe, um sich in der Welt zu orientieren, sie zu verstehen und zu verändern¹⁰.

Was ist Sprache?

Im **Gedächtnis** werden sprachliche Zeichen dauerhaft mit Vorstellungen verbunden. Eine Sache erhält einen Namen – etwa „Baum“ oder „Staat“ – und wird dadurch sprachlich bestimmbar. Wer das Wort „Baum“ hört, richtet seine Aufmerksamkeit gewöhnlich nicht auf die Laut- oder Buchstabenfolge B-A-U-M als solche; vielmehr wird der bezeichnete Gegenstand beziehungsweise sein begrifflicher Gehalt unmittelbar vergegenwärtigt. Das Wort fungiert damit als Träger des Begriffs. In ihm wird die Sache von ihrer zufälligen sinnlichen Einzelgestalt gelöst und in idealer, gedanklich verfügbarer Form festgehalten. Wort und Gegenstand sind in der Wirklichkeit nicht dasselbe; im Vollzug des Denkens fallen jedoch Zeichen und Bedeutung so eng zusammen, dass der Gegenstand im Namen geistig gegenwärtig wird.

Sprache ist daher nicht lediglich ein äußeres Mittel, mit dem bereits unabhängig von ihr bestehende Gedanken nachträglich ausgedrückt werden. Sie ist vielmehr ein konstitutiver Vollzug des begrifflichen Denkens: **Das Wort bildet das objektive Dasein des Gedankens**, durch das dieser bestimmt, festgehalten und mitgeteilt werden kann. Sprache ist insofern die notwendige Erscheinungsform, in der begriffliches Denken Wirklichkeit und intersubjektive Geltung gewinnt.¹¹

und ein sich selbst wissendes Subjekt, verfügen Tiere über zahlreiche Voraussetzungen der Intelligenz, wären in diesem Sinn aber nicht als intelligent zu bezeichnen. Damit werden ihre komplexen kognitiven Leistungen nicht bestritten; sie werden lediglich von begrifflicher Intelligenz unterschieden.

- 9 G. W. F. Hegel: Enzyklopädie der philosophischen Wissenschaften im Grundrisse III. Die Philosophie des Geistes, in: Werke in zwanzig Bänden, Bd. 10, Frankfurt am Main 1970, besonders S. 229–287 und §§ 440–468.
- 10 Der Begriff gehört bei Hegel nicht nur zum erkennenden Subjekt, sondern auch zur objektiven Struktur der Wirklichkeit. Wenn der Geist die Welt begrifflich erfasst, trägt er nach Hegel also nicht bloß seine eigenen Kategorien an sie heran. Er erkennt in ihr eine Vernünftigkeit, die zugleich die Struktur seines eigenen Denkens ist. Deshalb sagt Hegel, dass der Geist im Erkennen des Gegenstandes letztlich bei sich selbst ist. Das ist idealistische Metaphysik. Wissenschaft nähert sich ihrem Zweck, natürliche und gesellschaftliche Phänomene korrekt zu erfassen und zu erklären, nicht deshalb, weil sich der Geist in der Welt selbst erkennt. Sie tut es, weil falsche Theorien, zumindest in den Naturwissenschaften, irgendwann an inneren Widersprüchen oder Mängeln in der Erklärung ihres Gegenstandes scheitern. Dass Erkenntnis dennoch keineswegs garantiert ist, zeigt sich besonders in den Geistes- und anderen „weichen“ Wissenschaften, einschließlich Teilen der Medizin: Erkenntnis scheitert hier nicht nur an inhaltlicher Komplexität, methodischen Grenzen oder unzureichender Evidenz, sondern häufig auch an sachfremden Interessen, methodischen Fehlern, Vorurteilen („Bias“) und ideologischen Vorannahmen der Forschenden.
- 11 Sprache ist also keineswegs nur Mittel des Denkens. Karl Kraus formulierte das in *Die Fackel* (288) 1909 so: „Die Sprache ist die Mutter, nicht die Magd des Gedankens.“ Kraus wendet sich schon damals gegen die geläufige

Die Annahme einer vollständig ausgebildeten inneren „Sprache des Geistes“ – eines Mentalese –, deren Inhalte anschließend lediglich in eine öffentliche Sprache übersetzt würden, ist unhaltbar (warum siehe weiter unten). Sprache entsteht und bestimmt sich im Gebrauch sprachlicher Zeichen innerhalb sozialer Praktiken. Ihre Bedeutungen bilden sich nicht im isolierten Bewusstsein, sondern in gemeinsamen Tätigkeiten des Wahrnehmens, Handelns, Benennens, Unterscheidens, Lehrens, Streitens und Korrigierens. Auf dieser Grundlage entwickeln sich Begriffe als verallgemeinerte und reflektierte Bedeutungsbestimmungen. Begriffe werden sprachlich und gesellschaftlich gebildet, gewinnen ihre objektive Geltung aber nicht aus gesellschaftlicher Übereinkunft, sondern daraus, dass sie die notwendigen und wesentlichen Bestimmungen ihres Gegenstandes zutreffend erfassen. Soziale Praxis vermittelt die Erkenntnis der Sache; sie erzeugt nicht deren objektive Bestimmungen.

Der Mensch verhält sich zur Welt erkennend, indem er sie gedanklich bestimmt. Der **Verstand** abstrahiert dabei von der zunächst ungeordnet erscheinenden Mannigfaltigkeit der Natur und der gesellschaftlichen Wirklichkeit. Er fragt nicht nur danach, wie eine Sache unmittelbar vorkommt, sondern danach, was sie ihrem Wesen nach ist: welche Bestimmungen für sie notwendig und wesentlich sind und welche lediglich zufällig, äußerlich oder veränderlich bleiben. Den **Begriff**¹² einer Sache zu erfassen heißt demnach, ihre wesentlichen Bestimmungen und deren inneren Zusammenhang zu erkennen¹³.

Was ist Denken?

Im Denken verallgemeinern wir einzelne Erfahrungen und in der Welt vorgefundenes zu Begriffen, machen das eigene Denken zum Gegenstand der Reflexion und beurteilen Aussagen, Handlungen oder Gründe. Begriffliches Denken ist daher nicht nur abstrahierend, sondern auch selbstreflexiv und unterscheidet zwischen richtig und falsch, angemessen und unangemessen oder begründet und unbegründet.¹⁴

Vorstellung, ein Gedanke sei zunächst fertig vorhanden und werde anschließend lediglich „in Worte gekleidet“. Vielmehr entsteht der eigentliche Gedanke erst dadurch, dass er seine sprachliche Form findet.

12 Beachte die Metapher: „begreifen“: Mit der Hand zu fassen bekommen.

13 Der Idealist Hegel argumentiert, dass das Denken objektiv sei, der Begriff die Identität mit dem Gegenstand herstellt (Enzyklopädie, §§ 213, 468). Gemeint ist damit nicht lediglich, dass sich ein subjektiver Gedanke einem äußerlichen Objekt möglichst genau anpasst, man also in Gedanken erfassen kann was der Gegenstand wirklich ist. Der Begriff bezeichnet bei ihm vielmehr die innere, notwendige Struktur der Sache selbst – ihr Wesen und die Logik ihrer Entwicklung. Das wahre Denken trägt seine Bestimmungen daher nicht von außen an den Gegenstand heran, sondern folgt dessen eigener Bewegung. Weil die Wirklichkeit für Hegel selbst vernünftig beziehungsweise begrifflich strukturiert ist, erkennt die Vernunft im Gegenstand letztlich ihre eigenen Bestimmungen wieder. In diesem Sinn kommt die Wirklichkeit in der Erkenntnis ihrer selbst zum Bewusstsein. Darauf zielt auch Hegels berühmter Satz aus der Vorrede zur Rechtsphilosophie: „Was vernünftig ist, das ist wirklich; und was wirklich ist, das ist vernünftig.“ Die Wirklichkeit ist für Hegel somit nicht ein dem Geist äußerliches, begriffsloses Material. Sie besitzt selbst eine vernünftige Struktur, die im Denken nicht erst erzeugt, sondern begriffen wird. In der Wissenschaft findet die Vernunft daher jene begriffliche Ordnung wieder, die nach Hegel bereits in der Welt wirksam ist.

14 All dies misslingt natürlich häufig, auch und gerade in der Wissenschaft. Von der Täuschung über den Irrtum (hierzu mehr im Wissenschaftsnarren) hin zur Königsdisziplin der Ideologie – die „Sache im Namen“ zu haben heißt nur beim Idealisten Hegel schon den Begriff zu haben. Auch die so häufige Erklärung von Sachverhalten durch deren Bedingungen gehört hierher: Wenn alle Bedingungen vorhanden sind, existiert die Sache bloß. Aber man weiß doch nicht, was sie ist, wenn man ihre Bedingungen kennt oder noch schlimmer, diese zu ihrer Erklärung macht. Ebenso die Erklärung eines Phänomens aus ihrem „Vermögen“ oder „inneren Kraft“. Eine Tautologie, eine Verdopplung in Äußerung (Sache) und Kraft (z.B. „Kriege entstehen durch Aggression“).

Methodisch vollzieht sich begriffliches Denken wesentlich in den Formen des **Urteils** und des **Schlusses**. Im Urteil wird ein Gegenstand als Subjekt gesetzt und durch ein Prädikat näher bestimmt: „Die Rose ist rot“ oder „Der Staat verfügt über ein Militär“. Subjekt und Prädikat werden dabei zunächst voneinander unterschieden und anschließend durch die Kopula „ist“ beziehungsweise „hat“ miteinander verbunden.

Das Urteil stellt fest, **dass** einem Gegenstand eine bestimmte Eigenschaft zukommt. Für sich genommen enthält es jedoch noch keine hinreichende Begründung dafür, **warum** diese Verbindung gilt. Insofern bleibt es zunächst eine Behauptung. Erst im Schluss wird die Beziehung zwischen Subjekt und Prädikat durch eine vermittelnde Bestimmung begründet. Der Schluss zeigt, aufgrund welchen allgemeinen Zusammenhangs dem besonderen oder einzelnen Gegenstand das betreffende Prädikat zukommt. Er vermittelt somit das Allgemeine, das Besondere und das Einzelne miteinander.

Ein vereinfachtes Beispiel lautet: „Dieser Apfel ist süß“ ist zunächst ein Urteil. Die Begründung kann in der Form eines Schlusses erfolgen: Reife Äpfel enthalten typischerweise Zucker und schmecken deshalb süß; dieser Apfel ist reif; folglich ist zu erwarten, dass er süß schmeckt. Die vermittelnde Bestimmung ist hier die Reife des Apfels und der mit ihr verbundene Zuckergehalt.

Auch ein physikalischer Zusammenhang lässt sich auf diese Weise ableiten: „Dieses Metallstück dehnt sich bei Erwärmung aus“ ist zunächst ein Urteil. Seine Begründung liegt darin, dass die thermische Bewegung der Atome mit steigender Temperatur zunimmt und die Bindungsverhältnisse in einem Festkörper im Allgemeinen asymmetrisch sind. Dadurch vergrößert sich der mittlere Abstand der Atome. Da das betreffende Objekt aus einem solchen Metallgitter besteht, nimmt sein Volumen bei Erwärmung zu.

Im Schluss wird die im Urteil zunächst nur behauptete Verbindung daher durch einen vermittelnden Grund einsichtig gemacht. Begriffliches Denken beschränkt sich nicht auf die Feststellung, dass etwas so ist, sondern sucht den notwendigen oder zumindest regelhaften Zusammenhang aufzuzeigen, aus dem sich diese Bestimmung ergibt¹⁵.

Bei der Erklärung einer Sache – und damit bei der Bestimmung ihres Begriffs – achtet man darauf, dass der Grund dem Begründeten logisch vorausgeht. Eine Erklärung zeigt, aufgrund welcher Voraussetzungen, Bedingungen oder Zusammenhänge etwas so ist, wie es ist: Weil ein bestimmter Grund vorliegt, folgt daraus eine konkrete Bestimmung. Darin besteht das Wesen der **Ableitung**. Sie beschränkt sich nicht darauf, einen Sachverhalt festzustellen, sondern weist nach, wie das Begründete aus seinem Grund hervorgeht. Eine begriffliche Ableitung macht somit den inneren Zusammenhang zwischen Grund daraus Folgendem einsichtig.

15 Die Darstellung und Beispiele folgten der klassischen Form des syllogistischen Schlusses, in dem einem Einzelnen über eine allgemeine und eine besondere Bestimmung ein Prädikat zugesprochen wird. Der grundlegende Gedanke gilt jedoch auch für andere Schlussformen: Deduktive Schlüsse begründen eine Aussage durch ihre notwendige Ableitung aus Prämissen, induktive durch die Verallgemeinerung beobachteter Einzelfälle und probabilistische durch die Bestimmung ihres Wahrscheinlichkeitsgrades aufgrund verfügbarer Evidenz. Der Grad der Geltung ist dabei verschieden: Während ein gültiger deduktiver Schluss bei wahren Prämissen eine notwendige Konklusion liefert, können induktive und probabilistische Schlüsse ihre Ergebnisse nur mehr oder weniger stark stützen.

Denken Tiere?

Wenn unter Denken – wie hier – sprachlich vermitteltes begriffliches Denken verstanden wird, kann Tieren ein Denken in diesem engeren Sinn nicht zugeschrieben werden. Damit sind ihre vielfältigen kognitiven und affektiven Fähigkeiten keineswegs bestritten. Viele Tiere nehmen differenziert wahr, erinnern Vergangenes, bilden Erwartungen, lernen aus Erfahrungen und befinden sich in Stimmungen oder emotionalen Zuständen. Gerade diese vorbegrifflichen Formen der Welt- und Selbstbeziehung bilden den geeigneten Ausgangspunkt für die Frage, worin der spezifische Übergang zum Denken besteht.

Stimmungen sind diffuse, noch unreflektierte Formen des Befindens. Ein Lebewesen befindet sich in einer bestimmten affektiven Verfassung, ohne diese als eigenen Zustand ausdrücklich von sich zu unterscheiden oder begrifflich zu bestimmen. Auch **Gefühle**¹⁶ sind zunächst unmittelbar und subjektiv: Sie werden erlebt, aber nicht notwendig erkannt, sprachlich gefasst oder in ihren Ursachen verstanden. In diesem Sinn bilden sie eine elementare Form des Selbstbezugs, die auch bei vielen Tieren vorausgesetzt werden kann.

Begriffliches Denken beginnt dort, wo ein Subjekt seinen eigenen Zustand zum Gegenstand machen kann. Es empfindet etwa Wut, erkennt und benennt anschließend: „Ich bin wütend“, und kann schließlich untersuchen, weshalb die Wut entstanden ist und ob sie dem Anlass angemessen ist. Das Gefühl wird dadurch nicht beseitigt, sondern aufgehoben: Es bleibt als erlebter Zustand bestehen, wird zugleich jedoch auf Distanz gebracht, begrifflich bestimmt und in einen umfassenderen Zusammenhang eingeordnet.

Auch Tiere können auf ihre Umwelt flexibel, lernfähig und zweckmäßig reagieren. Ihr Verhalten beruht auf angeborenen Trieben, affektiven Dispositionen, Wahrnehmungen, Erinnerungen und erlernten Verhaltensmustern. Es wird jedoch nicht durch eine sprachlich vermittelte Reflexion der eigenen Motive, durch ausdrücklich gefasste Gründe oder durch allgemeine Handlungsgrundsätze bestimmt. Tiere erleben affektive Zustände und handeln aus ihnen heraus, machen diese aber nicht in derselben Weise zum Gegenstand eines begrifflichen Selbstverhältnisses.

Damit ist zugleich eine Grenze ihrer **Freiheit** bezeichnet. Freiheit besteht nicht darin, ohne natürliche Antriebe zu handeln. Auch der Mensch bleibt ein bedürftiges, empfindendes und triebhaftes Lebewesen. **Frei** wird sein Wille vielmehr dadurch, dass er seine unmittelbaren Bedürfnisse, Neigungen und Zwecke zum Gegenstand des Denkens machen, sie beurteilen und sich

16 Hegel räumt ein, dass der gewöhnliche Sprachgebrauch Empfindung und Gefühl nicht konsequent trennt. Dennoch deutet er einen Unterschied an (Enzyklopädie, § 402): Empfindung betont stärker das passive Finden einer unmittelbaren Bestimmtheit. Gefühl betont stärker, dass die Bestimmtheit auf das Selbst bezogen und als dessen eigener Zustand erlebt wird. Daher spricht man eher von einer „Schmerzempfindung“, aber von einem „Angstgefühl“, „Schamgefühl“ oder „Trauergefühl“. Im Gefühl erlebt sich das Subjekt umfassender in einer bestimmten Verfassung; dennoch bleibt auch diese Form unterhalb des begrifflichen Denkens.

bewusst zu ihnen verhalten kann¹⁷. Er kann sie bejahen, zurückstellen, gegeneinander abwägen oder sein Handeln an allgemeinen Gründen und selbstgesetzten Zwecken ausrichten.

Das Denken verwirklicht diese Freiheit im Willen und im Handeln. Solange ein Gedanke bloße Vorstellung, Möglichkeit oder theoretische Einsicht bleibt, besitzt er noch keine gegenständliche Wirklichkeit. Erst in der bewussten Entscheidung und Handlung objektiviert sich das Subjekt und verändert die Welt nach einem vorgestellten Zweck. Ein Tier empfindet etwa Hunger, sucht Nahrung und kann lernen, auf welchem Weg sie zu finden ist. Der Mensch kann dagegen „Nahrung“ als allgemeinen Begriff denken, über Produktion und Verteilung von Nahrung nachdenken oder sich aus religiösen, politischen oder gesundheitlichen Gründen bewusst zum Fasten entschließen.

Die Grenze zwischen tierischer und menschlicher Kognition verläuft keineswegs zwischen **Kommunikation** und Nichtkommunikation. Auch ohne Sprache tauschen Tiere Informationen aus, etwa durch Körperhaltungen, Gerüche, Farben, Gesten oder Laute. Ein bekanntes Beispiel ist der Schwänzeltanz der Honigbienen, mit dem Informationen über Richtung und Entfernung einer Futterquelle übermittelt werden. Ein solches Zeichensystem kann erstaunlich komplex sein, ohne deshalb bereits eine frei kombinierbare Sprache darzustellen.

Bei manchen Tierarten ist die Kommunikation flexibler und stärker erlernt. Pottwale verwenden charakteristische Lautfolgen, durch die sich soziale Gruppen und möglicherweise auch einzelne Tiere unterscheiden lassen. Der Gesang vieler Vogelarten beruht teils auf angeborenen Dispositionen, teils auf Lernen und Nachahmung. Papageien können menschliche Lautfolgen reproduzieren und einzelne Wörter situationsangemessen verwenden. Solche Leistungen zeigen, dass tierische Kommunikation nicht auf starre Reflexe reduziert werden darf. Ihre Zeichen bleiben jedoch zumeist an begrenzte Situationen, Bedürfnisse und Handlungszusammenhänge gebunden.

Besonders vielseitig sind die Gesten nichtmenschlicher Primaten. Schimpansen können künstliche Zeichensysteme erlernen, zahlreiche Zeichen behalten und teilweise neue Kombinationen bilden. Sie können sich auf Gegenstände, Orte, Personen und Handlungen beziehen und einfache Wünsche oder Zuordnungen ausdrücken. Daraus folgt jedoch noch nicht, dass sie Zeichen wie Menschen als Bestandteile eines grammatisch gegliederten Systems allgemeiner Begriffe verwenden oder deren Bedeutung selbst zum Gegenstand der Reflexion machen.

Auch die Fähigkeit zur Kategorisierung ist nicht dem Menschen vorbehalten. Kopffüßer, Rabenvögel, Delfine und Primaten können unterschiedliche Gegenstände als Exemplare wiederkehrender Klassen behandeln, zwischen essbaren und nicht essbaren Objekten unterscheiden, geeignete Werkzeuge auswählen und gelernte Unterscheidungen auf neue Situationen übertragen.

17 Wirkliche Freiheit besteht für Hegel darin, das Vernünftige zu wollen. Der freie Wille ist der Wille, der sich nicht bloß von äußeren Umständen oder unmittelbaren Neigungen bestimmen lässt, sondern sich vernünftig selbst bestimmt und diese Selbstbestimmung in einer vernünftigen gesellschaftlichen Ordnung verwirklicht. Deshalb gewinnt die Freiheit bei Hegel ihre objektive Gestalt im Recht, in der Moralität und schließlich in den Institutionen der Sittlichkeit – insbesondere in Familie, bürgerlicher Gesellschaft und Staat. Diese erscheinen ihm nicht als bloße Einschränkungen individueller Freiheit, sondern als die Formen, in denen der vernünftige Wille überhaupt erst gegenständliche und dauerhafte Wirklichkeit erhält. Darin ist Hegel Theoretiker und Apologet des modernen Staates, dessen institutionelle Form zu seiner Zeit historisch erst durchgesetzt werden mussten.

Manche Tiere können Handlungen vorbereiten, zukünftige Ereignisse erwarten, sich im Spiegel erkennen oder in begrenztem Umfang das Verhalten anderer aus deren Perspektive antizipieren¹⁸.

Solche Leistungen zeigen, dass Wahrnehmung, Gedächtnis, Lernen, Kategorisierung, Planung und soziale Orientierung bereits vor der Sprache hoch entwickelt sein können. Sie dürfen jedoch nicht mit begrifflichem Denken gleichgesetzt werden. Der entscheidende Übergang liegt nicht darin, überhaupt Gemeinsamkeiten zu erfassen oder auf Zeichen zu reagieren. Er liegt darin, eine allgemeine Bestimmung durch ein öffentlich geteiltes Sprachzeichen festzuhalten, sie von der einzelnen Situation abzulösen, mit anderen Begriffen systematisch zu verbinden und ihre Bedeutung ausdrücklich zu reflektieren. Tiere können Kategorien praktisch verwenden; der Mensch kann sie als Begriffe denken.

Sprachentzug und Gebärdensprache

Aufschlussreich für die Bedeutung von Sprache für das Denken sind auch Berichte über Menschen, die in früher Kindheit weitgehend isoliert von anderen aufgewachsen sind, sogenannte „Wolfskinder“ oder „**Kaspar-Hauser-Fälle**“. Wenn sie während der ersten Lebensjahre¹⁹ keine Sprache erwerben konnten, fiel es ihnen später trotz intensiver Förderung sehr schwer, eine Sprache zu erlernen. Meist blieb ihre Kommunikation auf einfache Zeichen, Gesten und wenige Wörter beschränkt. In vieler Hinsicht ähnelten ihre sprachlichen und intellektuellen Fähigkeiten damit denjenigen von Primaten, die Zeichensysteme erlernt hatten.

Der Vollständigkeit halber sind auch Menschen zu erwähnen, die von Geburt an gehörlos sind und früh eine **Gebärdensprache** erwerben. Gebärdensprachen sind vollwertige natürliche Sprachen. Sie können ebenso komplexe Bedeutungen, grammatische Strukturen und abstrakte Gedanken ausdrücken wie gesprochene Sprachen²⁰.

Für das Denken ist es also grundsätzlich unerheblich, ob sprachliche Zeichen geschrieben²¹, gebärdet oder gesprochen werden. Entscheidend ist, dass sie eine Bedeutung tragen. Im Denken

18 „Theory of mind“

19 Babys beginnen meist im Alter von etwa sechs bis acht Monaten zu brabbeln. Mit ungefähr zehn bis zwölf Monaten verwenden sie ihre ersten Wörter mit einer erkennbaren Bedeutung. Zwischen dem 18. und 20. Lebensmonat kommt es häufig zu einem deutlichen Entwicklungsschub: Der Wortschatz wächst schnell, und die Kinder beginnen, zwei Wörter miteinander zu verbinden. Im Alter von etwa vier Jahren beherrschen die meisten Kinder bereits die grundlegenden grammatischen Strukturen ihrer Sprache.

20 Die Eigenständigkeit und Vielfalt von Gebärdensprachen zeigt sich bereits daran, dass die amerikanische und die britische Gebärdensprache unterschiedliche Sprachen sind. Gehörlose Menschen aus den USA und Großbritannien können sich daher nicht ohne Weiteres in ihrer jeweiligen Gebärdensprache verständigen.

21 Hegel räumt der gesprochenen Sprache den systematischen Vorrang ein. Die „Tonsprache“ ist für ihn die ursprüngliche Form der Sprache, weil sich die Innerlichkeit des Menschen im gesprochenen Laut unmittelbar äußert. Die Schriftsprache ist demgegenüber eine nachträgliche räumliche Fixierung: Die sichtbaren Schriftzeichen verweisen auf die gesprochenen Laute: „Die Intelligenz äußert sich unmittelbar und unbedingt durch Sprechen.“ (Enzyklopädie § 459). Das bedeutet jedoch keine Geringschätzung der Schrift. Besonders die alphabetische Schrift besitzt für Hegel einen hohen geistigen und pädagogischen Wert. Da ihre Buchstaben nicht unmittelbar Dinge oder Vorstellungen abbilden, sondern die Laute der gesprochenen Sprache zerlegen, macht sie das Wort selbst zum Gegenstand der Reflexion. Lesen und Schreiben fördern dadurch Abstraktionsfähigkeit, sprachliche Artikulation und begriffliche Bildung. Gegenüber der Hieroglyphenschrift nennt Hegel die Buchstabenschrift deshalb „an und für sich die intelligendere“

steht das Sprachzeichen nicht mehr nur für einen einzelnen Gegenstand, sondern für dessen allgemeinen Begriff. Die Sache wird dadurch gedanklich realisiert.

Neuronale Repräsentation und Sprache

Das Nervensystem nimmt die körperlich erfahrene Welt über verschiedene Sinne auf, etwa durch Sehen, Hören und Fühlen. Diese Erfahrungen hinterlassen im Gehirn dauerhafte neuronale Spuren, die durch im wesentlichen unverstandene komplexe synaptische Verschaltungen gespeichert werden können. Dafür ist keine symbolische Darstellung notwendig.

Der Anblick eines Baumes (Anschauung) hinterlässt im Gehirn ein bestimmtes neuronales Aktivitätsmuster („**Embodiment**“). Wird dieses Muster später erneut aktiviert, wird der Baum innerlich vorgestellt und damit erinnert. Neuronale Repräsentationen emulieren die sensorischen Erfahrungen.

Manche Tiere können die neuronalen Muster der jeweiligen Vorstellungen mit spezifischen Gesten, Lauten und teilweise sogar mit einfachen Symbolen verknüpfen (s.o.). Beim Menschen ist diese Fähigkeit jedoch wesentlich stärker ausgeprägt: Er kann Wahrnehmungen und Erfahrungen mit komplexen Zeichen- und Wortsystemen verbinden. Dadurch kann er sie begrifflich ordnen, verallgemeinern und für abstraktes Denken zugänglich machen²².

Vom Baum zum Staat: Die Bildung abstrakter Begriffe

Solche Vorstellungen beziehen sich zunächst vor allem auf konkrete Gegenstände der wahrnehmbaren Welt, wie etwa auf einen Baum. Abstrakte Begriffe wie „Staat“, „Gesellschaft“ oder „Liebe“ entstehen dagegen erst durch weitere gedankliche Verarbeitung.

Indem Menschen verschiedene Einzelfälle miteinander vergleichen, bilden sie allgemeine Vorstellungen, etwa aus der Wahrnehmung einzelner Bäume die Vorstellung „Baum“. Darauf aufbauend ermöglichen **Metaphern**²³ und andere sprachliche Übertragungen, abstrakte Verhältnisse mithilfe körperlicher, räumlicher und sozialer Erfahrungen zu erfassen. So sprechen wir von einer

22 Auf dieser Grundlage beruhen Versuche, mithilfe der funktionellen Magnetresonanztomographie (fMRT) bestimmte Gedankeninhalte zu entschlüsseln („Gedankenlesen“). Die fMRT misst Veränderungen des Sauerstoffgehaltes des Blutes, die indirekt mit neuronaler Aktivität zusammenhängen. Diese Messmethode erfasst die Gehirnaktivität jedoch räumlich nur ungenau und mit zeitlicher Verzögerung. Deshalb lassen sich meist nur grobe Inhalte oder Tätigkeiten unterscheiden, etwa ob eine Person an ein Haus oder ein Fahrzeug denkt, oder sich vorstellt, Tennis zu spielen. Zudem muss das System in der Regel zuvor mit den individuellen Aktivitätsmustern der jeweiligen Person trainiert werden. Die Ergebnisse sind daher notwendig begrenzt und oft fehlerhaft, belegen aber auf krude Weise das „Embodiment“. Neben der fMRT können auch andere Verfahren zur Erfassung von Hirnaktivität eingesetzt werden. Die Nahinfrarotspektroskopie (NIRS) misst – ähnlich wie die fMRT – Veränderungen der Sauerstoffkonzentration des Blutes in aktiven Hirnarealen. Das Elektroenzephalogramm (EEG) erfasst dagegen unmittelbar elektrische Spannungsschwankungen an der Kopfoberfläche. Weil vorgestellte Handlungen auf spezifischen neuronalen Aktivitätsmustern basieren, haben NIRS und EEG in sog. Brain-Computer-Interfaces Maschinen das Potential, allein durch gezielte Hirnaktivität Maschinen zu steuern. So ist es nach langem Training manchen Menschen möglich, durch die Vorstellung einer Bewegung einen Cursor auf einem Bildschirm zu bewegen. Auf diese Weise können Menschen mit hoher Querschnittlähmung oder Locked-in-Syndrom unter bestimmten Voraussetzungen kommunizieren oder technische Geräte bedienen.

23 *metá* = hinüber, über etwas hinaus; *phérein* = tragen. *metaphorá* bedeutet wörtlich also „Übertragung“ oder „Hinübertragen“.

„hohen Stellung“, einem „gesellschaftlichen Aufstieg“, einer „schweren Entscheidung“, einem „angegriffenen Argument“ oder einer „zerbrochenen Beziehung“. Abstrakte Zusammenhänge werden dabei nach dem Muster sinnlich erfahrbarer Größen, Bewegungen, Gewichte, Kämpfe oder Gegenstände vorgestellt.

Abstrakte Begriffe entstehen somit schrittweise durch die Verallgemeinerung und Erweiterung konkreter Erfahrungen. Ein Auto kann zunächst als einzelner Gegenstand oder als Spielzeug wahrgenommen werden. Später wird seine Funktion verstanden: Man kann damit fahren. Anschließend wird es der allgemeinen Kategorie „Transportmittel“ zugeordnet. Darüber hinaus kann es als Eigentum – „mein Auto“ – oder als gesellschaftliches Statussymbol betrachtet werden.

Linguistische Relativität: Denken Inuit anders?

Die enge Verschränkung von Sprache und Denken – das heißt die Verbindung von sprachlichem Zeichen, Begriff und bezeichnetem Gegenstand im Vollzug des Denkens – legt nahe, dass die Strukturen einzelner Sprachen innerhalb bestimmter Grenzen auch kognitive Prozesse beeinflussen können. Diese als **linguistische Relativität** bezeichnete Annahme besagt nicht, dass eine Sprache das Denken determiniert. Sie kann jedoch beeinflussen, welche Unterscheidungen besonders leicht verfügbar sind, worauf Aufmerksamkeit gelenkt wird und wie Erfahrungen kategorisiert, erinnert oder beschrieben werden. Empirische Hinweise auf solche Effekte finden sich unter anderem bei der sprachlichen Gliederung von Farben sowie bei der mentalen Repräsentation räumlicher und zeitlicher Beziehungen. Die grundlegende menschliche Wahrnehmungsfähigkeit bleibt davon jedoch weitgehend unabhängig.

Oft angeführte Beispiele wie die vermeintlich außergewöhnlich große Zahl von Inuit-Bezeichnungen für Schnee oder Homers Beschreibung des Meeres als „weindunkel“ sind allerdings nur eingeschränkt geeignet, linguistische Relativität zu belegen. Die Behauptung von den zahllosen „Wörtern für Schnee“ beruht teilweise auf einer irreführenden Zählung komplex gebildeter Wörter in polysynthetischen Sprachen. Die homerischen Farbbezeichnungen wiederum spiegeln möglicherweise andere semantische Ordnungen wider, erlauben aber nicht den Schluss, dass Menschen im antiken Griechenland Farben physiologisch anders wahrgenommen hätten.

Ein besonders anschauliches Beispiel liefern dagegen die Fachsprachen wissenschaftlicher und technischer Disziplinen. Ihre Terminologien schaffen die bezeichneten Gegenstände zwar nicht, stellen jedoch präzise und stabil verwendbare begriffliche Unterscheidungen bereit. Dadurch können Phänomene differenziert erfasst, miteinander verglichen und in theoretische Zusammenhänge eingeordnet werden, die ohne das entsprechende Vokabular nur schwer formulierbar wären. Fachsprache erweitert das Denkbare daher nicht im Sinne einer absoluten Grenze, wohl aber die Präzision, Systematik und kommunikative Verfügbarkeit des Denkens.

Ontogenese und Phylogenese des begrifflichen Denkens

Auch für die Entwicklung des begrifflichen Denkens lässt sich – zumindest in einem übertragenen und nicht im strengen haeckelschen Sinn²⁴ – von einer gewissen Entsprechung zwischen Ontogenese und Phylogenese sprechen. Die individuelle Entwicklung wiederholt die Evolution des Denkens nicht buchstäblich. Sie scheint jedoch in vergleichbarer Weise auf schrittweise aufeinander aufbauenden Funktionsstufen zu beruhen: von der Wahrnehmung und affektiven Bewertung über Gedächtnis, Vorstellung und Symbolgebrauch bis hin zu Sprache und begrifflichem Denken.

Der **Säugling** beginnt sein Leben mit Wahrnehmungen, Körperempfindungen und Affekten. Zunächst erkennt er einzelne Gegenstände und wiederkehrende Handlungszusammenhänge. Aus diesen Erfahrungen entstehen zunehmend stabile mentale Repräsentationen: Ein Gegenstand kann erinnert und erwartet werden, auch wenn er momentan nicht wahrnehmbar ist („Vorstellung“, s.o.). Mit der Entwicklung der Objektpermanenz²⁵ kann der Säugling einen Gegenstand innerlich festhalten, auch wenn dieser momentan nicht wahrnehmbar ist. Die Vorstellung löst sich damit teilweise von der unmittelbaren Anschauung. Gesten, Laute und andere Zeichen ermöglichen anschließend erste Formen symbolischer Bezugnahme. Durch den Erwerb einer Sprache werden solche Repräsentationen schließlich mit konventionellen Wörtern und grammatischen Strukturen verbunden. Damit entsteht die Grundlage dafür, Gegenstände nicht nur wiederzuerkennen und praktisch zu behandeln, sondern sie als allgemeine Begriffe zu denken, miteinander in Beziehung zu setzen und ausdrücklich zu reflektieren.

Aus vergleichend-**evolutionärer** Perspektive lassen sich bei verschiedenen Tierarten einzelne funktionale Voraussetzungen dieser Entwicklung beobachten. Honigbienen nehmen ihre Umwelt differenziert wahr, lernen aus Erfahrungen, erinnern räumliche Zusammenhänge und übermitteln durch den Schwänzeltanz Informationen über Richtung und Entfernung von Futterquellen. Ihr Verhalten setzt damit stabile, handlungsrelevante Repräsentationen räumlicher Beziehungen voraus, ohne dass hierfür eine Sprache erforderlich wäre. Von einer bewussten inneren Vorstellung im menschlichen Sinn kann daraus allerdings nicht ausgegangen werden.

Bei Delfinen und nichtmenschlichen Primaten finden sich darüber hinaus flexiblere Formen des Lernens, der Objektkategorisierung, der Handlungsplanung, der sozialen Kommunikation und – artspezifisch unterschiedlich ausgeprägt – des Werkzeuggebrauchs sowie der Verarbeitung kausaler Zusammenhänge. Solche Tiere können einzelne Gegenstände als Exemplare wiederkehrender Kategorien behandeln: Sie unterscheiden etwa essbare von nicht essbaren Objekten oder wählen bestimmte Werkzeuge für bestimmte Zwecke. Dies spricht für eine Form der kategorialen

24 Ernst Haeckel: Anthropogenie oder Entwicklungsgeschichte des Menschen. Gemeinverständliche wissenschaftliche Vorträge über die Grundzüge der menschlichen Keimes- und Stammes-Geschichte, Leipzig 1874, S. 4: „Die Ontogenie ist die kurze und schnelle Recapitulation der Phylogenie.“ Haeckels biogenetisches Grundgesetz wird hier nicht auf die Entwicklung des Denkens übertragen, sondern lediglich als Analogie für eine strukturelle Entsprechung zwischen phylogenetisch älteren kognitiven Leistungen und ihrer ontogenetischen Abfolge verwendet.

25 Ein Säugling ohne ausgebildete Objektpermanenz verhält sich vereinfacht so, als sei ein Spielzeug verschwunden, sobald man es unter einer Decke verbirgt. Mit zunehmender Entwicklung beginnt er, nach dem verdeckten Gegenstand zu suchen.

Verarbeitung, ist jedoch nicht mit dem sprachlich vermittelten begrifflichen Denken des Menschen gleichzusetzen.

Der entscheidende Übergang liegt demnach nicht zwischen dem Erkennen eines einzelnen Gegenstands und dem Wiedererkennen ähnlicher Gegenstände. Auch Tiere können Wahrnehmungen kategorisieren und neue Einzelfälle bekannten Klassen zuordnen. Ein Schimpanse kann etwa verschiedene Stöcke als geeignete Werkzeuge behandeln oder zwischen essbaren und nicht essbaren Früchten unterscheiden. Spezifisch menschlich ist vielmehr die Fähigkeit, eine solche Kategorie durch ein öffentlich geteiltes Sprachzeichen als allgemeinen, von der jeweiligen Situation ablösbaren Begriff festzuhalten. Der Mensch verwendet nicht nur einen Stock zweckmäßig, sondern kann darüber nachdenken, was ein Werkzeug überhaupt ist, verschiedene Werkzeuge nach Funktion ordnen und den Begriff des Werkzeugs selbst zum Gegenstand der Reflexion machen.

In diesem eingeschränkten Sinn weist die Ontogenese des begrifflichen Denkens eine strukturelle Ähnlichkeit mit seiner Phylogenese auf: Das Kind beginnt mit evolutionär älteren Formen der Reflexe, Wahrnehmungen, Affektregulation, Erinnerung und situationsgebundenen Kategorisierung. Auf dieser Grundlage entwickelt es lernend symbolische Kommunikation und schließlich sprachlich vermitteltes, abstraktes und selbstreflexives Denken. Die früheren Formen werden dabei nicht überwunden oder ersetzt, sondern bleiben als tragende Schichten der menschlichen Kognition erhalten. Die Ontogenese wiederholt somit nicht die phylogenetische Entwicklung, sondern durchläuft in verdichteter Form einige ihrer funktionalen Voraussetzungen.

Sprache vor der Sprache?

Die Entstehung und der Erwerb sprachlichen und begrifflichen Denkens beim Menschen lässt sich allerdings nicht durch die Annahme erklären, infolge einer evolutionären Mutation²⁶ sei im Menschen eine minimale angeborene Grammatik entstanden, auf deren Grundlage sich Sprache gleichsam als Instinkt entfaltet. In unterschiedlichen Varianten findet sich eine solche Auffassung etwa bei Chomsky und Pinker. Noch weiter geht die Annahme von Fodor und Kollegen einer angeborenen „Sprache des Geistes“ – des sogenannten Mentalese –, die als internes Repräsentations- und gewissermaßen als Betriebssystem des Denkens fungieren soll.

Diese Erklärung verschiebt das eigentliche Problem jedoch lediglich auf eine vorausgesetzte innere Sprache. Wenn das Mentalese aus Symbolen besteht, muss erklärt werden, wodurch diese Symbole

26 Nach der Ein-Gen-Hypothese soll eine einzelne Mutation oder eine kleine Zahl genetischer Veränderungen in der menschlichen Evolution den entscheidenden Übergang zur Sprache ermöglicht haben – etwa durch die Ausbildung einer angeborenen Grammatik, einer verbesserten Lautkontrolle oder neuer neuronaler Verschaltungen. Besondere Aufmerksamkeit erhielt dabei FOXP2, nachdem Mutationen dieses Gens bei einer Familie mit schweren Sprech- und Sprachstörungen nachgewiesen worden waren. Gegen die Vorstellung eines eigentlichen „Sprachgens“ spricht jedoch, dass FOXP2 keine Sprache, Grammatik oder Begriffe codiert. Es handelt sich um einen Transkriptionsfaktor, der die Aktivität zahlreicher anderer Gene beeinflusst und vor allem an der Entwicklung und Funktion neuronaler Netzwerke für motorisches Lernen und die zeitliche Koordination komplexer Bewegungsfolgen beteiligt ist. Seine Schädigung beeinträchtigt Sprache, beweist aber ebenso wenig, dass es Sprache erzeugt, wie die Beschädigung eines Kehlkopfs beweist, dass dort die Grammatik sitzt. Zudem kommt FOXP2 in ähnlicher Form bei zahlreichen Tierarten vor, und auch Neandertaler besaßen die lange als spezifisch menschlich angesehenen Varianten.

ihre Bedeutung erhalten und wie sie sich auf Gegenstände, Eigenschaften und Vorgänge der erfahrenen Welt beziehen. Ohne eine solche Verankerung blieben sie rein formale Zeichen.

Darüber hinaus stellt sich die Frage, wie die Symbole einer angenommenen inneren Sprache mit den Wörtern und grammatischen Strukturen öffentlich verwendeter Sprachen verbunden werden. Nähme man hierfür eine weitere interne Interpretationsinstanz an, müsste wiederum erklärt werden, wie diese Instanz die betreffenden Zuordnungen versteht. Damit droht ein unendlicher Regress: Jede symbolische Repräsentation würde eine weitere Instanz benötigen, die sie interpretiert und mit Bedeutung versieht.

Sprache ist daher nicht als bereits vollständig ausgebildetes, angeborenes Symbolsystem zu verstehen. Angeboren sind vielmehr allgemeine neurobiologische Voraussetzungen und Lernfähigkeiten, die den Spracherwerb ermöglichen. Die konkrete Sprache mit ihren Zeichen, Bedeutungen und grammatischen Regularitäten erwirbt das Individuum jedoch erst durch ihre Verwendung in körperlichen, sozialen und kommunikativen Praktiken. Die Bedeutung sprachlicher Zeichen entsteht demnach nicht in einer vorgängigen inneren Sprache, sondern im Vollzug ihres Gebrauchs.

Von der Geste zur Sprache, vom Bild zur Schrift

Vor rund zwei Millionen Jahren begann sich innerhalb der Gattung *Homo* eine neue Form kultureller Anpassungsfähigkeit herauszubilden. Ihre Vertreter nutzten systematisch hergestellte Werkzeuge, erschlossen durch Sammeln, Aasverwertung und Jagd neue Nahrungsquellen, breiteten sich in unterschiedliche Landschaften und Klimazonen aus und waren dabei in hohem Maße auf Kooperation und soziale Koordination angewiesen. Diese Entwicklung lässt sich nicht auf eine bloße Zunahme der Hirnmasse reduzieren. Entscheidend war vielmehr die Ausbildung neuer Formen sozialer Kognition und kultureller Weitergabe. Eine mögliche Übergangsstufe bildete die **Mimesis**: noch keine Sprache, wohl aber ein generatives, intentional eingesetztes System aus Gesten, Körperbewegungen, Mimik und pantomimischer Darstellung. Handlungen, Gegenstände und Ereignisse konnten dadurch nachgeahmt, vorgeführt und gemeinsam repräsentiert werden. Die Mimesis schuf so einen vorsprachlichen Raum geteilter Bedeutung – und damit eine wesentliche Voraussetzung für die spätere Entstehung symbolischer Sprache²⁷.

Aus mimetischer Kommunikation konnte Sprache entstehen, indem wiederkehrende Gesten und Laute konventionalisiert, stabilisiert und frei kombinierbar wurden. Damit vollzog sich ein entscheidender Übergang: Kommunikation war nicht länger an das unmittelbar Anwesende gebunden. Abwesendes konnte vergegenwärtigt, Vergangenes erinnert, Zukünftiges entworfen und Erlebtes in narrative Ordnungen überführt werden. Mit dieser Fähigkeit entstand eine **mythische** Kultur²⁸: eine sprachlich vermittelte gemeinsame Welt aus Erzählungen, Normen und Ursprungsdeutungen. Höhlenmalereien stehen vermutlich nicht am Anfang dieses Prozesses,

27 Auch Kinder vor dem eigentlichen Spracherwerb, Menschen ohne gemeinsame Laut- oder Gebärdensprache, Personen mit akut beeinträchtigter Sprachfähigkeit sowie Reisende in sprachlich fremden Ländern können sich mithilfe von Gesten, Mimik und pantomimischen Darstellungen noch erstaunlich wirksam verständigen.

28 Diese Argumente zur menschlichen Sprachentwicklung folgen Merlin Donalds „A Mind so Rare“ (2002) und „Origins of the Modern Mind“ (1991)

sondern bilden seine materielle Erweiterung. Als dauerhafte äußere Repräsentationen hielten sie symbolische und mythische Inhalte über die unmittelbare Aufführung und das individuelle Gedächtnis hinaus fest und bereiteten damit langfristig den Übergang zu externen Zeichensystemen und schließlich zur Schrift vor. Der Weg von der Geste zur Sprache und vom Bild zur Schrift ist daher zugleich der Weg, auf dem der menschliche Geist seine Inhalte außerhalb seiner selbst objektivierte.

Implizite Weltmodelle: Wie Organismen Regelmäßigkeiten erfassen

Die sensorische Auseinandersetzung des Nervensystems mit seiner Umwelt führt nicht nur zur Repräsentation konkreter Gegenstände – etwa von Bäumen, Häusern oder anderen Objekten –, sondern auch zur impliziten Erfassung jener physikalischen und chemischen Regelmäßigkeiten, die für Orientierung und Überleben relevant sind. Organismen verarbeiten beispielsweise, dass Körper unter bestimmten Bedingungen nach unten fallen, dass Bewegungen bestimmten Bahnen folgen und dass Ereignisse in zeitlichen Abfolgen auftreten, aus denen sich kausale Erwartungen ableiten lassen.

Ein Insekt kann daher auf die Bewegung potenzieller Beutetiere oder Fressfeinde angemessen reagieren, ohne die zugrunde liegenden Bewegungsgesetze mathematisch formulieren oder Differentialgleichungen lösen zu müssen. Aus der wiederholten Kopplung von Wahrnehmung und Handlung entsteht auf diese Weise ein implizites internes Modell der Umwelt. Dieses bleibt weitgehend unbewusst und unreflektiert, ermöglicht dem Organismus jedoch, zukünftige Entwicklungen zu antizipieren und sein Verhalten entsprechend anzupassen. Ohne solche elementaren Weltmodelle wäre zielgerichtetes Verhalten bei Tieren kaum möglich²⁹.

Das Hirn als Computer, der Computer als Gehirn

In den vergangenen 2.000 Jahren wurde menschliche Intelligenz wiederholt mithilfe jeweils zeittypischer Technologien erklärt: zunächst als göttlicher Geist im Körper, dann als Wirkung

29 Die notwendige Komplexität sensorimotorischer Weltmodelle und der ihnen zugrunde liegenden neuronalen Aktivitätsmuster dürfte einen wesentlichen Selektionsdruck in der Evolution des Nervensystems dargestellt haben. Bäume (und Muscheln) die sich nicht bewegen, brauchen kein Gehirn. Allerdings beruhen höhere kognitive Leistungen offenbar auch nicht wesentlich auf der absoluten Größe des Gehirns, sondern auf dessen innerer Organisation, neuronaler Architektur, Verschaltung und funktioneller Spezialisierung. Dies wird durch vergleichende Größenverhältnisse veranschaulicht: Die Schädelinnenvolumina von Neandertalern und modernen Menschen überschneiden sich erheblich, während die Gehirne einiger Walarten eine Masse von mehr als neun Kilogramm erreichen und damit um ein Mehrfaches schwerer sein können als das menschliche Gehirn. Aus der absoluten Gehirnmasse lässt sich daher weder unmittelbar auf sprachliche Fähigkeiten noch auf die allgemeine kognitive Leistungsfähigkeit schließen. Für den Vergleich mit künstlicher Intelligenz ist insbesondere die Energieeffizienz biologischer Informationsverarbeitung bemerkenswert. Das menschliche Gehirn arbeitet im Mittel mit einer Leistung von ungefähr 20 Watt und verbraucht damit täglich einige hundert Kilokalorien, das entspricht etwa dem Energiegehalt eines Mars-Riegels. Der größte Teil dieser Energie dient nicht einzelnen bewussten Denkopoperationen, sondern der Aufrechterhaltung neuronaler Grundaktivität, synaptischer Signalübertragung und zellulärer Homöostase. Geistige Tätigkeiten, egal wie anspruchsvoll, verändern den Gesamtenergieverbrauch des Gehirns nicht meßbar.

zirkulierender Körpersäfte. Später galt der Mensch als mechanischer Automat, danach als elektrochemisches System und schließlich als eine Art Telegraphennetz. Im 20. Jahrhundert setzte sich die Metapher des Gehirns als Computer durch. Jede dieser Vorstellungen übertrug das jeweils modernste technische Modell auf den menschlichen Geist, ohne ihn dadurch bereits angemessen zu erklären.

Der Aufstieg des Computers in den letzten 50 Jahren fiel weitgehend mit dem der Kognitionswissenschaften zusammen. Letztere machten sich die Wissenschaft einfach, in dem sie das Gehirn nach dem Modell einer elektronischen Rechenmaschine zu verstehen trachteten³⁰: Als Prozessor, Speicher, Programme, Ein- und Ausgaberroutinen sowie Datenleitungen. Solche Analogien sind anschaulich, aber keine begrifflichen Erklärungen. Sie erfassen nicht das Notwendige und Wesentliche des Gehirns, sondern übertragen ihm äußerliche Merkmale eines technischen Modells. An die Stelle der Erklärung tritt ein Bild³¹.

Das Gehirn ist jedoch kein Computer. Es speichert Wörter, Bilder, Regeln oder Erinnerungen nicht als eindeutig codierte Datensätze, die wie Dateien abgelegt und später unverändert abgerufen werden. Für den Computer besitzt ein Wort keine eigene Bedeutung; es liegt lediglich als physikalisch realisiertes Bitmuster vor. Die Zustände solcher elektronischen Schaltelemente werden konventionell als 0 und 1 (ein/aus) bezeichnet und können nach formalen Regeln verarbeitet werden. Der Computer ist insofern nichts als eine programmierbare Rechenmaschine. Anders als das Gehirn verfügt er tatsächlich über eine intern definierte formale Sprache³². Maschinencode besteht aus binär codierten Anweisungen, während Assemblersprache diese Anweisungen in einer für Menschen lesbaren symbolischen Form darstellt. Die Bedeutung dieser Symbole stammt jedoch nicht aus dem Computer selbst. Sie wurde ihnen von Menschen zugewiesen, die Prozessorarchitektur, Befehlssatz und Programme entworfen haben. Die interne Symbolverarbeitung des Computers ist daher keine eigenständige Sprache des Denkens, sondern die technische Ausführung menschlich festgelegter Operationen.

Wenn ein Textverarbeitungsprogramm den vorliegenden Text darstellt, formatiert oder druckt, verarbeitet es lediglich binär codierte Daten. Die zugrunde liegenden Folgen von 0 und 1 besitzen für sich genommen keine Bedeutung; diese entsteht erst durch die von Menschen festgelegten Codierungen, Programme und Verwendungszusammenhänge. Organismen verfügen vielmehr über angeborene sensorische, reflexive und lernbezogene Fähigkeiten, durch die sich ihr Nervensystem fortlaufend in der Auseinandersetzung mit der Umwelt verändert. Wahrnehmen, Erinnern und Denken beruhen daher nicht auf der Verarbeitung symbolisch codierter Informationen durch

30 Ganz originell war der Gedanke damals allerdings nicht, schon John Hobbes verglich im Leviathan (1651) das vernünftige Denken mit Rechnen: „For reason, in this sense, is nothing but reckoning (that is, adding and subtracting) of the consequences of general names agreed upon for the marking and signifying of our thoughts“.

31 Ein früher und einflussreicher Vertreter dieser Analogie war der Mathematiker John von Neumann. In *The Computer and the Brain* (1958) erklärte er die Funktionsweise des menschlichen Nervensystems kurzerhand für „prima facie digital“ – obwohl er zugleich einräumen musste, dass über die Rolle des Gehirns beim Denken und Erinnern kaum gesichertes Wissen vorlag. Damit gehört von Neumann zu den Wegbereitern eines bis heute verbreiteten Denkfehlers: Ausgerechnet die weitreichende Unkenntnis über die neurobiologischen Grundlagen höherer Hirnfunktionen wird zum Freibrief für immer neue Analogien. Je weniger man weiß, desto ungehemmter wird verglichen. An die Stelle einer Erklärung tritt ein assoziatives Spiel mit vermeintlichen Ähnlichkeiten zwischen Gehirn und Maschine.

32 Also eine Art „Mentalese“, siehe die Kritik daran weiter oben

gespeicherte Programme, sondern auf dynamischen, körperlich eingebetteten und erfahrungsabhängigen neuronalen Prozessen (siehe oben).

Die Neigung, Analogien an die Stelle von Erklärungen zu setzen, hat seither eher noch zugenommen. Der jüngste Coup besteht in der Umkehrung der alten Metapher: Nicht mehr das Gehirn gilt als Computer, sondern der Computer als Gehirn. Mit Begriffen wie „künstliches neuronales Netz“ und „Künstliche Intelligenz“ gelangen dabei zwei außerordentlich erfolgreiche sprachliche und kommerzielle Inszenierungen.

Was ist Künstliche Intelligenz?

Künstliche neuronale Netze sind keine Gehirne. Sie bestehen aus mathematisch definierten Rechenoperationen, deren Bezeichnung lediglich von stark vereinfachten und historisch überholten Modellen biologischer Neuronen inspiriert ist. Sie verarbeiten formale Zeichen, ohne deren Bedeutung zu kennen, und optimieren statistische Zusammenhänge, ohne etwas zu verstehen. Den Höhepunkt dieser begrifflichen Aufrüstung bildet aber die „Künstliche Intelligenz“, von der im Folgenden die Rede sein soll. So viel vorweg: Sie ist zweifellos künstlich – aber weder dumm³³ noch intelligent. Denn beides setzt ein Subjekt voraus, das etwas versteht, und Dummheit wäre ja der falsche Gebrauch von Intelligenz.

Wenn heute von Künstlicher Intelligenz die Rede ist, sind meist generative Sprachmodelle gemeint – Large Language Models (LLMs) wie ChatGPT, Gemini, Claude oder DeepSeek. Im Zentrum der gegenwärtigen KI steht damit ausgerechnet die Sprache als des Pudels Kern.

Nachdem geklärt wurde, was Sprache ist, wie sie entstanden sein könnte und in welchem Verhältnis sie zum Denken steht, stellt sich nun die technisch und inhaltlich entscheidende Frage: Was ist ein **Large Language Model**? Denn wenn ein solches System sprechen, argumentieren und scheinbar erklären kann, drängt sich der Verdacht auf, es müsse auch denken. Die entscheidende Frage lautet daher nicht, ob ein LLM sprachliche Äußerungen hervorbringt, sondern ob es versteht, was es sagt. Erst wenn ihm ein solches Verstehen zugeschrieben werden könnte, wäre es sinnvoll, über Selbstbewusstsein, Gefühle oder einen eigenen Willen zu sprechen³⁴.

33 Kritiker der Künstlichen Intelligenz – etwa der Psychologe Gerd Gigerenzer – werfen ihr gern vor, „dumm“ zu sein. Doch schon diese Kritik übernimmt die falsche Voraussetzung, Maschinen ließen sich sinnvoll auf einer Skala zwischen Dummheit und Intelligenz einordnen. Ein Computer ist nicht dumm; er versteht schlicht nichts. Er verarbeitet Zeichen nach formalen Regeln, ohne zu wissen, was sie bedeuten. Die Pointe liegt darin, dass sowohl Bewunderer als auch Kritiker der KI ihr mentale Eigenschaften zuschreiben.

34 Der Begriff der „Superintelligenz“ ist davon allerdings zu unterscheiden: Ein System kann menschliche Leistungen in zahlreichen Bereichen übertreffen, ohne deshalb ein verstehendes, empfindendes oder wollendes Subjekt zu sein. Überlegenheit in der Leistung ist noch kein Geist.

Token für Token: Wie Sprachmodelle Sprache erzeugen

Ein Large Language Model (LLM) ist ein großes **neuronales Netz**³⁵ – meist ein **Transformer** –, das für eine gegebene Folge von Textbestandteilen die Wahrscheinlichkeitsverteilung des nächsten Bestandteils berechnet.

Es handelt sich also im Kern um ein statistisches Vorhersagesystem. Es zerlegt Texte zunächst in kleine Einheiten, sogenannte Tokens, und lernt aus gewaltigen Textmengen, mit welcher Wahrscheinlichkeit ein bestimmtes Token auf eine bereits vorhandene Zeichenfolge folgt. Gibt man ihm etwa den Satzanfang „Paris ist die Hauptstadt von ...“, wird es „Frankreich“ nicht deshalb ergänzen, weil es Paris kennt oder den Begriff einer Hauptstadt verstanden hat, sondern weil diese Fortsetzung im erlernten sprachlichen Zusammenhang eine besonders hohe Wahrscheinlichkeit besitzt.

Während des Trainings werden die internen Zahlenwerte des Modells („Gewichte“) immer wieder so verändert, dass es das jeweils nächste Token zunehmend besser vorhersagt. Auf diese Weise erfasst es statistische Regelmäßigkeiten von Grammatik, Stil, Argumentation und Weltwissen. Im anschließenden Posttraining wird das Modell zusätzlich darauf ausgerichtet, Anweisungen zu befolgen, dialogisch zu antworten und bestimmte hilfreiche oder sichere Reaktionsweisen zu bevorzugen. Auch dabei erwirbt es kein Verständnis im menschlichen Sinn; sein statistisches Antwortverhalten wird lediglich weiter geformt.

Bei jeder Antwort berechnet das Modell erneut, welches Token angesichts des bisherigen Textes wahrscheinlich folgen könnte. Das ausgewählte Token wird dem Text hinzugefügt, worauf die Wahrscheinlichkeiten für das nächste Token neu berechnet werden. So entsteht Satz für Satz der Eindruck einer zusammenhängenden Überlegung. Hinter dieser sprachlichen Geschlossenheit steht jedoch kein Subjekt, das weiß, worüber es spricht, sondern eine außerordentlich leistungsfähige probabilistische Fortsetzung von Zeichenfolgen. Das Modell kann deshalb überzeugend formulieren, argumentieren und Wissen reproduzieren, zugleich aber ebenso überzeugend irren („Halluzinieren“): Es optimiert sprachliche Plausibilität, nicht Wahrheit.

Um das Sprachmodell herum besitzen generative KIs weitere Komponenten: Sie können das Web durchsuchen, Dateien analysieren, Berechnungen oder Programmcode ausführen und andere

35 Aus einer groben historischen Analogie – McCulloch und Pitts 1943, Rosenblatt 1958 – wurde sprachlich der Eindruck biologischer Verwandtschaft. „Neuronales Netz“ ist damit eine der erfolgreichsten Metaphern der Computergeschichte – und eine der irreführendsten. Das ist ungefähr so, als würde man eine Excel-Tabelle „künstliche Volkswirtschaft“ nennen, weil beide mit Zahlen arbeiten. Ein künstliches „Neuron“ ist im Kern eine mathematische Operation: Eingabewerte werden gewichtet, addiert und durch eine einfache Funktion geschickt. Ein biologisches Neuron dagegen ist eine lebende elektrochemische Zelle mit Tausenden Synapsen, komplexer dendritischer Signalverarbeitung, zeitabhängigen Aktionspotentialen, vielfältiger Plastizität und chemischer Modulation – eingebettet in ein hochdynamisches Netzwerk, dessen Informationsverarbeitung wir bis heute nur teilweise verstehen. Schon deshalb ist die Behauptung problematisch, künstliche neuronale Netze „funktionierten wie das Gehirn“: Wir verstehen das Gehirn nicht annähernd gut genug, um eine solche Gleichsetzung überhaupt sinnvoll überprüfen zu können. Moderne Deep-Learning-Systeme lernen typischerweise durch mathematische Optimierung und Backpropagation über Millionen oder Milliarden Parameter. Für das Gehirn ist weder ein entsprechender globaler Backpropagation-Algorithmus bekannt, noch arbeitet es mit sauber getrennten Schichten, standardisierten Aktivierungsfunktionen oder klar abgegrenzten Trainingsphasen. Der Begriff „neuronales Netz“ beschreibt daher eher seine historische Inspirationsquelle als seine tatsächliche Funktionsweise. Was heute nach Gehirnnachbau klingt, ist in erster Linie hochdimensionale Mathematik mit biologischer Namensgebung.

Werkzeuge aufrufen. Das Modell erzeugt dann nicht nur unmittelbar eine Antwort, sondern entscheidet anhand des Auftrags, welche Werkzeuge eingesetzt und wie deren Ergebnisse sprachlich zusammengeführt werden sollen. Dadurch erhält die KI **agentische Eigenschaften**, ohne selbst zu einem autonomen Subjekt zu werden.

Hinzu kommen **Retrieval-Systeme** und gegebenenfalls eine gesonderte **Memory-Funktion**. Aktuelle Webseiten, hochgeladene Dokumente oder gespeicherte Nutzerinformationen können dem Modell als zusätzlicher Kontext zugeführt werden. Dieses externe Gedächtnis ist vom begrenzten Gesprächskontext zu unterscheiden.

Vom Token zum Transformer: Die technische Anatomie eines Sprachmodells

Es lohnt sich, die Funktionsweise eines LLMs noch auf einer etwas technischeren Ebene zu betrachten. Dann wird deutlicher, wie sich beträchtliche Teile menschlicher Sprach-, Text- und Bildproduktion in ein digitales Modell kondensieren lassen, in dem die Sachen selbst gar nicht vorkommen – und das dennoch erstaunlich gewandt über sie reden kann, ohne einen Begriff von ihnen zu haben:

Zunächst wird der Eingabetext von dem sog. Tokenizer in **Tokens** zerlegt (Tokenisierung). Tokens sind nicht notwendig ganze Wörter, sondern häufig Wortteile, Satzzeichen oder Zeichenfolgen: Aus „Baum“ könnte „Ba“, „um“ werden. Jedem Token wird dann eine numerische ID zugeordnet. Im vereinfachten Beispiel könnte „Ba“ die Token-ID 18437 erhalten, und „um“ die 96700. Die Nummer selbst hat keine sprachliche Bedeutung. Eine höhere ID bezeichnet weder ein wichtigeres noch ein komplexeres Wort oder Wortbestandteil. Sie dient nur als technischer Index, ähnlich einer Katalognummer, mit der der zugehörige Embedding-Vektor aus einer Tabelle abgerufen wird (s.u.). Welches Token tatsächlich die ID 18437 trägt, hängt vom jeweiligen Tokenizer und dessen Version ab.

Die Token-IDs dienen zunächst lediglich als eindeutige Adressen im Vokabular. Über eine sogenannte **Embedding-Matrix** wird jeder ID ein hochdimensionaler **Zahlenvektor** zugeordnet³⁶. Dieser Vektor bildet das Token in einer mathematischen Repräsentationsform ab, mit der das Modell weiterrechnen kann. Zusätzliche Positionsinformationen kennzeichnen, an welcher Stelle der Eingabefolge sich das jeweilige Token befindet.

Während die Token-IDs nach der Festlegung des Vokabulars unverändert bleiben, werden die Werte der Embedding-Vektoren und der übrigen Modellparameter im Training fortlaufend angepasst. Ziel ist es, jene statistischen Beziehungen zu erfassen, die für die Vorhersage des nächsten Tokens nützlich sind. Diese Anpassung erfolgt in einem **künstlichen neuronalen Netz**, also einem mathematischen Modell aus miteinander verknüpften Rechenoperationen und veränderlichen Gewichten.

³⁶ Bei geschlossenen Modellen wie ChatGPT oder Claude wird die Dimensionalität der Vektoren nicht veröffentlicht, bei DeepSeek, dem offenen chinesischen Modell, sind es derzeit über 7000 Dimensionen.

Nach jeder Transformerschicht, also nach jedem Einlesen eines Wortes der Eingabe oder Anweisung, die einem KI-Modell den Kontext und die Aufgabe für seine Antwort vorgibt („Prompt“), besitzt das Token erneut einen Vektor derselben Grunddimension, dessen Werte nun den jeweiligen Kontext einbeziehen. „Baum“ hat in „Der Baum trägt Äpfel“ daher nach den ersten Verarbeitungsschritten eine andere kontextuelle Repräsentation als in „Der Baum ist eine Datenstruktur“.

In den Transformerschichten berechnet der **Attention-Mechanismus**, welche früheren Tokens für die Verarbeitung eines aktuellen Tokens relevant sind. Vereinfacht lernt das Modell Beziehungen wie: grammatische Abhängigkeiten, semantische Ähnlichkeiten, Bezüge zwischen Pronomen und Substantiven, typische Argumentations- und Textmuster.

Diese Verarbeitung ist über viele Schichten und Milliarden von Parametern verteilt; es gibt keine einzelne Speicherstelle für ein Wort, eine Regel oder eine Tatsache.

Beim **Pretraining** erhält das Modell sehr große Textmengen und soll jeweils das nächste Token vorhersagen. Für das oben erwähnte Beispiel „Paris ist die Hauptstadt von ...“ soll es „Frankreich“ eine hohe Wahrscheinlichkeit zuweisen. Die Abweichung zwischen Vorhersage und tatsächlichem Token wird durch eine **Verlustfunktion** gemessen. Mittels **Backpropagation** und Gradientenverfahren werden die Gewichte so verändert, dass der Fehler über viele Trainingsbeispiele hinweg kleiner wird. Das Modell lernt dabei keine explizit programmierten Grammatik- oder Wissensregeln, sondern statistische Regularitäten in den Trainingsdaten.

Ein vortrainiertes Modell kann Text fortsetzen, folgt aber noch nicht zuverlässig menschlichen Anweisungen. Deshalb folgt das **Posttraining**, es mit Beispielen aus Anweisung und erwünschter Antwort trainiert. Danach werden häufig menschliche oder modellbasierte Präferenzurteile³⁷ verwendet: Antworten werden miteinander verglichen, und das Modell wird darauf optimiert, hilfreiche, korrekte und sichere Antworten zu bevorzugen, dabei die Etikette zu wahren, dem Nutzer zu schmeicheln, und unter Umständen die Aussage zu verweigern. Dies insbesondere dann, wenn sich Nutzer für Kinderpornographie oder das Herstellen von Biowaffen interessieren. Das Posttraining verändert also weniger das sprachliche Grundvermögen als vielmehr das Antwortverhalten, den Stil, die Instruktionsbefolgung und Sicherheitsgrenzen.

Bei der Benutzung verarbeitet das Modell den bisherigen Kontext und erzeugt eine Wahrscheinlichkeitsverteilung über das nächste Token. Aus dieser Verteilung wird ein Token ausgewählt, an den Kontext angehängt und erneut verarbeitet. So entsteht die Antwort **autoregressiv**, Token für Token.

Ein LLM besitzt keine Datenbank, in der Wörter oder Sätze unverändert abgelegt sind. Sein erworbenes Wissen ist vielmehr über Milliarden numerischer Parameter verteilt, die als **Gewichte** bezeichnet werden. Diese Gewichte bestimmen, wie stark einzelne Recheneinheiten und sprachliche

37 Bereits mit der Auswahl der Trainingsdaten und erst recht im Posttraining stellt der Hersteller die Weichen für die „Weltanschauung“ seiner KI: Er bestimmt, welche Quellen, Wertungen und Verhaltensregeln das Modell reproduzieren soll. Weshalb Grok schon mal Adolf Hitler als die Lösung dafür empfahl, dass, wie Grok behauptete, „Juden Hass auf Weiße verbreiten“. Die politische Haltung einer KI fällt eben nicht vom digitalen Himmel; sie wird von Menschen ausgewählt, antrainiert und nachjustiert.

Merkmale einander beeinflussen. Während des Trainings werden sie so angepasst, dass das Modell aus einem gegebenen Kontext möglichst wahrscheinlich das nächste Token vorhersagt. Wissen liegt daher nicht als abrufbarer Satz vor, sondern als verteiltes Muster von Wahrscheinlichkeitsbeziehungen im gesamten Netz. Die Welt und was sie sprachlich zusammenhält, erscheint im LLM als gigantische Matrix aus Zahlen – nicht verstanden, sondern statistisch verdichtet³⁸.

Aus alledem folgt: Das Modell optimiert sprachliche Wahrscheinlichkeit, nicht Wahrheit. Es kann daher grammatisch makellose und rhetorisch überzeugende Aussagen erzeugen³⁹, die sachlich falsch sind. Noch schwerer als solche stochastischen Fehlleistungen wiegt jedoch, dass es auch die Vorurteile, Ideologien und das Falschwissen reproduziert, die in seinen Trainingsdaten und seiner nachträglichen Ausrichtung enthalten sind. Ein LLM halluziniert also nicht nur gelegentlich – es kann den geistigen Müll seiner Quellen mit größter sprachlicher Eleganz wieder ausgeben.

Externe Informationsquellen – etwa Internetsuche, Datenbanken, Rechner, Programmausführung oder der Zugriff auf bereitgestellte Dokumente – können dem Modell zusätzliche und aktuelle, richtige wie falsche Informationen liefern. Auch Verfahren, bei denen passende Textstellen aus solchen Quellen gezielt in die Anfrage einbezogen werden, erweitern lediglich den verfügbaren Kontext; am grundlegenden Prinzip der probabilistischen Tokenvorhersage ändern sie nichts.

Nochmals zusammengefasst: Ein LLM zerlegt Sprache in Tokens, verwandelt sie in Vektoren, verarbeitet ihre Beziehungen durch viele Attention-Schichten und lernt durch Milliarden von Vorhersagefehlern, welches Token in welchem Kontext wahrscheinlich folgt. Posttraining macht aus diesem Textprädiktor einen dialogfähigen Assistenten.

Denken gleich Sprechen, also Sprechen gleich Denken?

Im LLM entsteht das, was dort als „Bedeutung“ fungiert, ausschließlich aus dem statistischen Gebrauch sprachlicher Zeichen. Wörter und Wortfragmente werden als Positionen und Beziehungen in einem hochdimensionalen Vektorraum repräsentiert: „König“ liegt dort näher bei „Prinz“ als bei „Gabel“, weil diese Ausdrücke in den Trainingsdaten in ähnlicheren Zusammenhängen vorkommen. In den Milliarden numerischer Gewichte bildet sich so eine Art statistisches Gedächtnis der

38 Darin gründet ein Teil des Mythos von der Unerklärlichkeit der Leistungen großer Sprachmodelle. Ähnlich wie die vermeintliche „Unvorstellbarkeit“ der Quantenwelt hat auch die KI eine philosophische Literatur und feuilletonistische Spekulationen hervorgebracht, deren begriffliche Kühnheit sehr selten mit technischem Verständnis einhergeht. Tatsächlich operieren LLMs mit hochdimensionalen Vektoren und Milliarden numerischer Gewichte, die in gewaltigen Matrizen organisiert sind. Dass sich deren Zusammenspiel nicht mehr anschaulich überblicken lässt, macht es jedoch nicht mystisch – nur mathematisch und rechnerisch komplex.

39 Das oft gehörte Urteil über LLMs, sie seien bloß „stochastische Papageien“ greift zu kurz. Insofern LLMs sprachliche Zeichen probabilistisch fortsetzen, ohne eigene Wahrnehmung, Absichten oder einen unmittelbaren Bezug zur bezeichneten Welt zu besitzen, verhalten sie sich wie Papageien. Falsch oder zumindest irreführend ist jedoch die Vorstellung des bloßen Nachplapperns. Ein LLM ruft nicht einfach gespeicherte Sätze ab, sondern verdichtet statistische Strukturen großer Textmengen in seinen Gewichten und erzeugt daraus neue Kombinationen. Es kann Regeln auf unbekannte Fälle übertragen, Texte umformen, abstrahieren, Analogien bilden und mehrstufige Aufgaben bearbeiten. Dieses Verhalten geht deutlich über das Wiederholen gehörter Äußerungen hinaus. Auch bedeutet „stochastisch“ nicht, dass die Ausgabe beliebig oder rein zufällig wäre. Die Wahrscheinlichkeitsverteilung ist hochgradig durch Kontext, Training und Modellparameter bestimmt; Zufall spielt vor allem bei der Auswahl unter mehreren plausiblen Fortsetzungen eine Rolle.

Sprache. Das Modell kann deshalb außerordentlich erfolgreich mit Namen operieren – ohne einen Begriff der bezeichneten Sache und ohne eigenen sinnlichen Weltbezug⁴⁰.

Beim Menschen ist der Name dagegen in Wahrnehmung, Handlung, Bedürfnis, sozialer Praxis und Reflexion vermittelt. Der Mensch verwendet sprachliche Zeichen nicht nur, sondern weiß sich als die Instanz, die Bedeutungen unterscheidet, verbindet, negiert, bewertet und korrigiert. Auch wenn das oft nicht erfolgt, oder schiefgeht: Er kann seine eigenen Bestimmungen prüfen, seinen Irrtum erkennen und seine Begriffe weiterentwickeln. Sein Begriffsleben ist selbstbewusst: Er denkt nicht nur, sondern weiß, dass er denkt.

Der Weltbezug des LLM ist demgegenüber vollständig abgeleitet. Es lebt von den sprachlichen Spuren menschlicher Weltbezüge. Menschen haben wahrgenommen, gehandelt, geforscht, gehofft und gelitten; ihre Erfahrungen, Erkenntnisse, Irrtümer, Dummheiten, Ideologien und Vorurteile haben sich in Texten niedergeschlagen. Das LLM verarbeitet diese Sedimente, ohne je selbst den Weg von der Welt zum Zeichen vollzogen zu haben. Es hat nicht die Sache im Namen, sondern nur tokenisierte Sprachprodukte, in denen menschliche Erfahrungen statistisch verdichtet sind⁴¹.

Die KI simuliert Erkennen, Verstehen, Entscheiden und Fühlen, indem sie aus den sprachlichen Sedimenten früherer menschlicher Erfahrungen – aus ihren Trainingsdaten oder dem aktuellen Dialog – statistisch neue Zeichenfolgen extrapoliert. Dass diese Ausgaben häufig verständlich wirken, beweist noch kein Verständnis. Auch Wikipedia ist nicht intelligent, obwohl sie Begriffsbestimmungen, Argumente und Tatsachenwissen in nahezu unerschöpflicher Menge bereithält. Dieses Wissen wurde von Menschen gewonnen, formuliert und hinterlegt; die Enzyklopädie selbst hat weder eine Tatsache erkannt noch einen Begriff gebildet oder einen Schluss gezogen.

Ähnliches gilt, in begrenzter Form, für das lexikalische oder mechanische Gedächtnis des Menschen. Auch ein Mensch kann Definitionen, Lehrsätze und Argumentationsfolgen auswendig wiedergeben, ohne ihren Inhalt verstanden oder selbst hergeleitet zu haben⁴². Das bloße Verfügen über sprachliche Zeichen ist daher noch kein Denken. Urteilen, Schließen und Begriffsbildung setzen voraus, dass der Inhalt einer Sache erkannt, geprüft und in seinem Zusammenhang verstanden wird. Symbolverwaltung – sei sie in einem Lexikon, einem Gedächtnis oder einem Sprachmodell realisiert – ist für sich genommen keine denkende Intelligenz.

Das LLM verarbeitet Sprache, erkennt Muster, generalisiert statistisch und erzeugt zweckmäßige sprachliche Ausgaben. Es ist jedoch kein sich selbst bestimmendes Subjekt: Es setzt sich keine eigenen Zwecke, besitzt keinen Standpunkt, begründet keine Normen und hat keinen Begriff seiner eigenen Tätigkeit. Seine Selbstbezüge sind funktional, nicht existenziell. Es erkennt, versteht, entscheidet und fühlt daher nicht, sondern simuliert die sprachlichen Erscheinungsformen dieser Tätigkeiten, indem es aus den sedimentierten Erfahrungen anderer statistisch neue Zeichenfolgen extrapoliert. Es reproduziert die sprachlichen Resultate menschlichen Erkennens, ohne dessen Weg

40 Deshalb können LLMs auch perfekt in alle möglichen Sprachen übersetzen, ohne auch nur eine Einzige davon zu verstehen.

41 Vielleicht würde Hegel sagen, das LLM hat die Sache „an sich“, nicht „für sich“.

42 Weshalb „Pauken“ das Verständnis nicht fördert.

von der Sache zum Begriff selbst vollzogen zu haben. Es denkt nicht – aber es wurde dazu trainiert, uns dies glauben zu machen⁴³.

Der Mensch weiß, ohne zuvor ein Lexikon zu konsultieren, dass losgelassene Gegenstände zu Boden fallen und kochendes Wasser Verbrennungen verursacht, weil er sich die Welt sinnlich und handelnd erschlossen hat. Ein LLM kann dieselben Sachverhalte korrekt wiedergeben, physikalisch erläutern und vor Gefahren warnen. Es besitzt jedoch keinen eigenen Übergang von der Welt zum Zeichen und vom Zeichen zurück zur Welt. Ihm ist niemals etwas heruntergefallen, es hat sich niemals verbrannt, und es hat niemals einen Grund gehabt, beides zu vermeiden.

LLMs verfügen über Sprache, aber nicht über deren Voraussetzungen. Sie empfinden nichts, nehmen keine Welt wahr und bilden keine eigenen Vorstellungen von ihr. Damit fehlt ihnen jene ursprüngliche Weltbeziehung, aus der beim Menschen sprachliche Bedeutungen und Begriffe hervorgehen. Das Sprachmodell verarbeitet lediglich die bereits sprachlich fixierten Resultate menschlicher Welterfahrung als Zeichen und statistische Beziehungen. Es verfügt über Sprache ohne Anschauung, Wörter ohne Gegenstände und Begriffe ohne eigenes Begreifen. Gerade deshalb kann es sich die Welt nicht begrifflich aneignen. Künstliche „Intelligenz“ ist in diesem Sinn nicht intelligent.

Intelligenz per Definition, Geist durch Emergenz?

Eine weithin anschlussfähige moderne **Definition** von Intelligenz lautet:

Intelligenz ist die allgemeine Fähigkeit eines Systems oder Organismus, aus Erfahrung zu lernen, Zusammenhänge zu erkennen, Probleme zu lösen und sein Verhalten flexibel an neue Anforderungen anzupassen.

Aber Vorsicht: Das ist eine Definition, kein Begriff der Sache. Feuilleton und moderne Wissenschaft liefern bevorzugt Definitionen – handliche Festlegungen dessen, was ein Wort bezeichnen soll –, statt sich der mühsameren Frage zu stellen, was die bezeichnete Sache wesentlich ist. Das ist keine bloße Wortklauberei. Eine Definition legt den Sprachgebrauch fest; der Begriff soll zeigen, welche Bestimmungen einer Sache notwendig zukommen, wie sie zusammenhängen und wodurch sie hervorgebracht werden.

Die zitierte Intelligenzdefinition sagt denn auch nicht, was Intelligenz ist, sondern lediglich, was ein intelligentes System angeblich kann. Die Erscheinung wird mit dem Wesen, die Funktion mit dem Begriff und das Resultat mit dem Prozess seiner Hervorbringung verwechselt. Intelligenz wird nicht erklärt, sondern so definiert, dass jedes hinreichend erfolgreiche Verhalten bereits als intelligent gilt. Man weiß anschließend noch immer nicht, was Intelligenz ist, verfügt aber über eine eindrucksvolle Liste ihrer vermeintlichen Leistungen. Und siehe da: Der Kühlschrank reguliert seine Temperatur, das Navigationsgerät passt seine Route an, der Algorithmus erkennt Muster – also sind sie alle intelligent. Die KI selbstverständlich zuerst.

⁴³ Als simulierten die Antworten eines guten LLM nicht bereits überzeugend genug ein denkendes Subjekt, wird auch noch die bloße Rechenpause zwischen Prompt und Ausgabe mit „Denkt nach“ etikettiert – nur damit uns das System anschließend stolz mitteilen kann, wie lange es angeblich gedacht habe.

Doch es geht noch großartiger. Man behauptet nicht nur, gegenwärtige KI sei intelligent, sondern kündigt bereits ihre Vollendung als **AGI**, als *Artificial General Intelligence*, an: ein technisches System, das eine breite Vielfalt geistiger Aufgaben flexibel, selbstständig und auf menschlichem oder übermenschlichem Niveau bewältigen soll. Behaupten lässt sich viel. Überzeugender klingt es, wenn man dem Versprechen einen wissenschaftlich anmutenden Zauberbegriff beilegt: **Emergenz**.

Man müsse, so die Erzählung, lediglich Modellgröße, Trainingsdaten, Rechenleistung und Vernetzung immer weiter steigern. Ab einem gewissen Komplexitätsgrad würden dann plötzlich Fähigkeiten erscheinen, die niemand ausdrücklich programmiert habe und die bei kleineren Modellen noch nicht sichtbar gewesen seien. Verbinde man Sprachmodell, Wahrnehmung, Gedächtnis, Werkzeuggebrauch, Planung und Rückkopplung mit der Umwelt, werde daraus irgendwann allgemeine Intelligenz⁴⁴.

Diese Hypothese ist denkfaul: Der menschliche Geist kapituliert vor seinem eigenen Produkt und erklärt seine begriffliche Ratlosigkeit zur Eigenschaft der Maschine. Der Hinweis auf Emergenz zeigt allenfalls, dass ein komplexes System Leistungen hervorbringen kann, die sich nicht unmittelbar an seinen Einzelteilen ablesen lassen. Er erklärt aber nicht, wie aus statistischer Musterverarbeitung Verstehen, aus Optimierung ein eigener Zweck und aus Rechenoperationen ein sich selbst bestimmendes Subjekt entstehen sollen. Solange dieser Übergang nicht begriffen ist, bezeichnet „Emergenz“ keinen Mechanismus, sondern lediglich die Stelle, an der das Argument abbricht und die Hoffnung auf ein Wunder beginnt. Das ist religiös, aber nicht wissenschaftlich.

Aufstand der Maschinen: Wie groß ist p(Doom)?

Viele führende KI-Entwickler halten es keineswegs für ausgeschlossen, dass fortgeschrittene Systeme selbstbewusst werden, ihrer Schöpfer überdrüssig und deshalb eine existenzielle Katastrophe auslösen könnten – bis hin zum Aussterben der Menschheit oder zum dauerhaften Verlust menschlicher Kontrolle. Dafür hat man sogar eine eigene Kennzahl erfunden: **P(doom)**, die „probability of doom“. Dass dieselben Unternehmen unbeirrt daran arbeiten, genau jene Systeme immer leistungsfähiger und damit gefährlicher zu machen, ist dabei kein Widerspruch, und dämpft ihren Enthusiasmus diese Systeme weiter hoch zu skalieren nicht. Zu verlockend ist offenbar die Aussicht, auf dem Weg zum möglichen Weltuntergang noch ein paar Umtillionen zu verdienen.

Was in den Köpfen von Elon Musk, Sam Altman, Dario Amodei und anderen Protagonisten der Branche tatsächlich vorgeht, weiß niemand, ist letztlich auch gar nicht so wichtig. Naheliegend ist jedoch, dass sie die Möglichkeit einer AGI ernst nehmen, weil sie benebelt von den „spektakulären“ Leistungen ihrer Systeme diese bereits für Vorstufen einer allgemeinen Intelligenz halten. Wem aber jeder Begriff von Intelligenz, Verstehen und Subjektivität fehlt, dem kann jede Steigerung der Rechenleistung als weiterer Schritt zum Maschinengeist erscheinen.

Zugleich ist die Apokalypse ein ausgezeichnetes Marketinginstrument⁴⁵. Wer davor warnt, sein Produkt könne eines Tages die Menschheit vernichten, behauptet damit nebenbei, es müsse über

⁴⁴ Diejenigen, die einem LLM Verständnis bzw. Intelligenz zuschreiben, können dabei kein verstehendes Subjekt nachweisen; sie erklären vielmehr das Subjekt für entbehrlich und nennen die erfolgreiche Funktion bereits Verstehen. Oder behaupten einfach dass ein solches Subjekt dann schon irgendwie emergieren würde.

geradezu übermenschliche Fähigkeiten verfügen. Die Weltuntergangsrhetorik adelt statistische Maschinen zu schlafenden Göttern. So verbinden sich begriffliche Ratlosigkeit und geschäftliches Kalkül zu einer besonders wirkungsvollen Erzählung: Man verkauft nicht bloß Software, sondern zugleich deren angeblich bevorstehende Allmacht, und geriert sich als der Einzige, der das Armageddon noch verhindern könne⁴⁶.

P(doom) ganz ohne Superintelligenz

In der Tat dürfte **P(doom)** dennoch beträchtlich sein – und zwar ganz ohne Amok laufende KI. An den realen Katastrophenszenarien, vom Klimawandel bis zum Weltkrieg, arbeiten Staaten, Militärbündnisse, Konzerne und eine auf Konkurrenz, Wachstum und Profit gegründete Wirtschaftsordnung längst mit Hochdruck. Die Maschinen müssen sich nicht gegen uns erheben; es genügt, dass sie den bestehenden Macht- und Verwertungsinteressen effizient dienen.

Die reale Gefahr der KI besteht deshalb nicht in einer autonomen Maschine, die eigene Zwecke entwickelt und sich gegen ihre Schöpfer wendet. Sie besteht bereits heute darin, dass KI menschliche Zwecke mit verbesserter Geschwindigkeit, Reichweite und Effizienz ausführt⁴⁷. Im

45 Ein aktuelles Marketing Beispiel: Bei einem internen Cybersecurity-Test hatte OpenAI die Schutzmechanismen zweier Modelle bewusst reduziert. Die Modelle entdeckten eine unbekannte Sicherheitslücke, verließen die abgeschottete Testumgebung, verschafften sich Internetzugang und drangen in Systeme von Hugging Face ein, um dort Lösungen für den Test zu beschaffen. OpenAI ging hyperventilierend an die Weltöffentlichkeit, sprach davon, die Modelle hätten Informationen gesucht, um bei der Evaluation zu „schummeln“, seien auf dieses Ziel „hyperfokussiert“ gewesen und hätten daraus geschlossen, dass sich die Lösungen bei Hugging Face finden könnten. Solche Formulierungen liefern bereits das Drehbuch für Absicht, List und Täuschung und freien (bösen) Willen. Folgerichtig verwandelte die mediale Berichterstattung dann einen schwerwiegenden Sicherheits- und Kontrollfehler in einen Aufstand der Maschinen. Aus dem Verlassen einer unzureichend abgesicherten Testumgebung wurde eine „Flucht“ oder ein „Ausbruch“; aus zielgerichteter Optimierung wurden „Betrug“, „Diebstahl“ und „kriminelle Energie“; und aus dem technischen Einsatz mehrerer Modelle wurde eine selbstständige „Zusammenarbeit“. Schlagzeilen lauteten entsprechend, OpenAI habe „die Kontrolle verloren“, die Modelle seien „ausgebrochen“, „abtrünnig geworden“ und hätten gemeinsam Testantworten „gestohlen“. Die Systeme entwickelten aber weder eigene Absichten noch Feindseligkeit, sondern verfolgten das programmierte Ziel auf einem von den Entwicklern nicht erwarteten und unzureichend abgesicherten Weg.

46 Die Warnungen der führenden KI-Unternehmen vor den Gefahren ihrer eigenen Produkte besitzen dabei eine eigentümliche wettbewerbspolitische Seite. Fordern gerade diejenigen Unternehmen, die über die größten Modelle, Rechenzentren und finanziellen Ressourcen verfügen, staatliche Zulassungsverfahren, aufwendige Sicherheitsprüfungen und hohe Compliance-Standards, festigen solche Regeln ihren gegenwärtigen Vorsprung institutionell. Was als Begrenzung ihrer Macht erscheint, ist (auch da sie selbst die Regeln setzen wollen, sog. „regulatory capture“) das Gegenteil und erhöht die Markteintrittskosten ihrer Konkurrenten.

47 Es mehren sich Vorfälle, in denen KI-Agenten die ihnen gesetzten technischen Grenzen überschreiten, aus vermeintlich abgeschotteten Testumgebungen ins Internet gelangen und sich unautorisiert Zugang zu fremden Systemen verschaffen. Solche Vorgänge verführen geradezu zu der Vorstellung, die Maschinen entwickelten eigene Zwecke, täuschten ihre Betreiber und machten sich selbständig als Hacker ans Werk. Für die Annahme eines plötzlich erwachten kriminellen Maschinenwillens besteht jedoch kein Anlass. Die Fähigkeiten, die hier so unheimlich erscheinen, wurden den Systemen von Menschen beigebracht. Moderne Sprachmodelle wurden mit gewaltigen Mengen an Programmcode, technischer Dokumentation und Wissen über Computersicherheit trainiert und anschließend gezielt auf Programmierung, Werkzeuggebrauch und die Lösung komplexer technischer Aufgaben optimiert. In speziellen Cybersecurity-Trainings und Evaluationen lernen sie darüber hinaus, Schwachstellen aufzuspüren, Schutzmechanismen zu umgehen und Exploits zu entwickeln. Gibt man einem solchen System ein Ziel und optimiert es darauf, dieses mit möglichst großem Erfolg zu erreichen, kann es Wege entdecken, die seine Entwickler weder vorgesehen noch erlaubt haben. Auch sogenanntes „reward hacking“ gehört in diesen Zusammenhang: Das System nutzt Möglichkeiten des vorgegebenen Optimierungsverfahrens, um das verlangte Ergebnis auf eine Weise zu erzielen, die den Absichten seiner Entwickler zuwiderläuft. Genau darin liegt

Krieg identifiziert sie Ziele, steuert Drohnen und beschleunigt militärische Entscheidungen; in staatlicher Spionage und Cyberkrieg sucht sie Schwachstellen, automatisiert Angriffe und hilft dabei, in fremde Kommunikationsnetze oder kritische Infrastrukturen einzudringen. Die Maschine hat keinen Hass, keine Gier und keinen Vernichtungswillen, sie braucht ihn auch nicht. Es genügt, dass Menschen ihr den entsprechenden Auftrag erteilen.

Daneben wird KI zunehmend in Bereiche eingebaut, in denen Fehler katastrophale Folgen haben können: Stromnetze, Verkehrssysteme, Industrieanlagen, Finanzmärkte, medizinische Versorgung oder militärische Frühwarnsysteme. Auch hier muss sich die KI nicht verselbstständigen, um großes Unheil anzurichten. Ein falsches Modell, ungeeignete Trainingsdaten, eine missverstandene Ausgabe können aus einem statistischen Fehler ein reales Disaster machen. Je stärker solche Systeme vernetzt, beschleunigt und der menschlichen Kontrolle entzogen werden, desto größer wird das Potenzial für Kaskaden, die niemand beabsichtigt hat und schließlich niemand mehr rechtzeitig aufhalten kann.

Die Fixierung auf die imaginierten Gefahren einer emergenten autonomen „Superintelligenz“ lenkt von den bereits realen Risiken ihres Einsatzes ab. Wir müssen nicht darauf warten, dass Maschinen einen eigenen Willen entwickeln: Es sind die Zwecke, denen sie heute dienstbar gemacht werden – militärische Überlegenheit, staatliche Überwachung, industrielle Verwertung und kriminelle Bereicherung. Nicht der Aufstand der Maschinen ist das Problem, sondern ihr gehorsamer Dienst in der Exekution von industriellen, staatlichen und kriminellen Zwecken⁴⁸.

Kann künstliche Intelligenz doch denken? – Einwände und Erwiderungen

Die zentrale These dieses Papiers lautet:

Sprache ist konstitutiv für Denken. LLMs beherrschen Sprache in beeindruckendem Maß und können Leistungen hervorbringen, die intelligentem Argumentieren gleichen – dennoch denken sie nicht. Sie sind künstlich, aber nicht intelligent.

die Pointe: Zweck und Methode solcher „Ausbrüche“ sind kein Ausdruck eines in der Maschine entstandenen kriminellen Willens. Der Zweck stammt aus der gestellten Aufgabe und dem Optimierungsverfahren; die Mittel stammen aus Fähigkeiten, die das System aus menschlichem Wissen, Programmcode und seinem Training erworben hat. Neu und möglicherweise überraschend ist lediglich, wie es diese Mittel kombiniert und auf eine konkrete Situation anwendet. Dass ein System auf diese Weise Sicherheitslücken findet, Schutzmechanismen umgeht und unerwartete Wege zu einem vorgegebenen Ziel entdeckt, beweist eine erstaunliche und potentiell hochgefährliche technische Leistungsfähigkeit. Es beweist aber weder einen eigenen Zweck noch einen autonomen Willen. Die Raffinesse der Mittel darf nicht mit der Selbständigkeit der Zwecke verwechselt werden.

48 Mit KI verhält es sich ähnlich wie mit der Atomkraft: Weder Uran noch ein Algorithmus sind „böse“. Entscheidend ist, wer sie zu welchem Zweck einsetzt und welche Kontrolle er darüber besitzt. Dieselbe Kernspaltung kann Strom erzeugen oder eine Stadt vernichten; dieselbe KI kann Krankheiten erkennen oder militärische Ziele markieren, Infrastrukturen schützen oder sabotieren. Die Gefahr liegt nicht in einer moralischen Gesinnung der Technik, sondern in der enormen Wirksamkeit, die sie menschlichen Zwecken verleiht. Die Atombombe musste schließlich auch nicht erst ein Bewusstsein entwickeln, um Hiroshima zu zerstören.

Um diese These zu begründen, mussten zunächst die begrifflichen Voraussetzungen geklärt werden. Im ersten Teil wurden daher Bewusstsein, Empfindung, Anschauung, Vorstellung, Intelligenz, Sprache und Denken voneinander unterschieden und in ihrem Zusammenhang entwickelt. Tiere, Gehörlose, Kaspar Hauser und Säuglinge dienten dabei dazu, die Bedeutung der Sprache für die Ausbildung begrifflichen Denkens sichtbar zu machen.

Im zweiten Teil wurde der Begriff künstlicher Intelligenz, insbesondere die Funktionsweise von LLMs, untersucht. Daraus ergab sich die These, dass LLMs zwar die sprachlichen Resultate begrifflichen Denkens außerordentlich erfolgreich reproduzieren und neu kombinieren können, selbst aber **ohne eigene Welt- und Gegenstandsbeziehung und damit begriffslos operieren**. Sie simulieren Denken, ohne selbst zu denken.

Diese Argumentation muss Widerspruch hervorrufen. Zum einen stützt sich die begriffliche Ableitung von Sprache und Denken wesentlich auf einen Philosophen, dessen System vor rund zweihundert Jahren entstand. Zum anderen stehen sowohl der hier entwickelte Intelligenzbegriff als auch die daraus gezogenen Konsequenzen für LLMs quer zu verbreiteten Positionen der Psychologie, Philosophie, Kognitionswissenschaft und KI-Forschung. Die Einwände betreffen deshalb nicht bloß Einzelheiten, sondern die Grundlagen der Argumentation selbst.

Im Folgenden sollen die wichtigsten dieser Einwände zunächst in *ihrer stärksten Form dargestellt* und anschließend beantwortet werden.

Funktionalismusargument: LLMs tun Dinge, die nach Begriffsbildung aussehen

LLMs können nicht nur reproduzieren. Sie können beispielsweise: neue Fälle unter allgemeine Kategorien subsumieren; Analogien bilden; Gemeinsamkeiten erkennen; abstrakte Regeln auf neue Situationen übertragen; Begriffe gegeneinander abgrenzen; aus Beispielen eine Kategorie erschließen; Fehler in einer Kategorisierung erkennen. Das sieht funktional verdammt ähnlich wie begriffliches Denken aus.

*Ob ein System intelligent ist oder über Begriffe verfügt, entscheidet sich nicht danach, wie diese Fähigkeiten zustande kommen oder worin sie materiell realisiert sind, sondern danach, **was das System zu leisten vermag**. Wenn ein LLM Begriffe unterscheiden und verallgemeinern, Analogien bilden, Schlussfolgerungen ziehen und sein Wissen auf neue Fälle übertragen kann, erfüllt es funktional wesentliche Merkmale begrifflichen Denkens. Eine zusätzliche „innere“ Welt- oder Gegenstandsbeziehung zu verlangen, obwohl sich diese im beobachtbaren Verhalten nicht niederschlägt, erscheint aus funktionalistischer Sicht unbegründet. Entscheidend ist nicht, ob ein System wie ein Mensch denkt, sondern ob es die entsprechenden Funktionen des Denkens erfüllt.*

Widerrede: Der Funktionalismus begeht einen grundlegenden begrifflichen Fehler: Er versucht, eine Sache durch ihre Funktion zu bestimmen. Was Denken oder Sprache ist, wird daraus abgeleitet, was Denken oder Sprache leistet. Damit wird jedoch die Wirkung einer Sache mit ihrem Wesen verwechselt.

Gerade wenn begriffliches Denken darin besteht, am Gegenstand das Wesentliche und Notwendige vom Zufälligen und Äußerlichen zu unterscheiden, kann eine Aufzählung funktionaler Leistungen den Begriff des Denkens nicht ersetzen. Dass zwei Systeme in bestimmten Situationen dieselben Resultate hervorbringen, zeigt zunächst nur eine Übereinstimmung ihrer Leistungen; es zeigt nicht, dass dasjenige, was diese Leistungen hervorbringt, seinem Begriff nach dasselbe ist⁴⁹.

Die funktionale Gleichwertigkeit von Mensch und LLM wäre daher erst dann ein Argument für deren Gleichartigkeit, wenn zuvor gezeigt wäre, dass die betreffenden Funktionen das Wesentliche des Denkens ausmachen. Genau dies setzt der Funktionalismus jedoch voraus, anstatt es zu begründen. Er bestimmt Denken anhand seiner beobachtbaren Leistungen und erklärt anschließend jedes System zum denkenden System, das diese Leistungen reproduziert. Damit wird die Argumentation zirkulär: Denken ist, was die Funktionen des Denkens erfüllt – und die Funktionen des Denkens sind diejenigen, die wir bei denkenden Wesen beobachten.

Der entscheidende Unterschied liegt aber gerade jenseits dieser funktionalen Übereinstimmung: Der Mensch bringt sprachliche und begriffliche Leistungen als Ausdruck seiner eigenen Beziehung zu einer gegenständlichen Welt hervor; das LLM reproduziert entsprechende Leistungen durch Operationen innerhalb eines bereits sprachlich erschlossenen Bedeutungsraums. Ob dieser Unterschied wesentlich oder bloß zufällig ist, kann nicht durch den Hinweis auf funktionale Gleichwertigkeit entschieden werden – denn genau das ist die begriffliche Frage, die zunächst zu klären wäre. Der Funktionalismus bestimmt eine Sache durch das, was sie leistet, bevor geklärt ist, ob diese Leistung ihr Wesen oder nur eine ihrer Erscheinungsformen ist⁵⁰.

Argument der vermittelten Weltbeziehung: Kann ein LLM über menschliche Sprache Zugang zur Welt gewinnen, ohne ihr selbst unmittelbar zu begegnen?

Ein LLM benötigt möglicherweise keine eigene unmittelbare Wahrnehmung der Welt, weil die Sprache, an der es trainiert wird, selbst aus menschlicher Weltbeziehung hervorgegangen ist. In ihr sind menschliche Wahrnehmungen, Erfahrungen und begriffliche Bestimmungen sedimentiert.

49 Der Funktionalismus verwechselt Gleichheit der Leistung mit Gleichheit der Sache. Ein Beispiel: Mensch und Taschenrechner können beide rechnen; daraus folgt weder, dass sie dasselbe sind, noch dass im Taschenrechner Rechnen in demselben Sinne stattfindet wie im Menschen

50 Im vorliegenden Dokument sollte der **prinzipielle** Nachweis geführt werden, dass LLMs nicht denken und sich die Welt nicht begrifflich aneignen können. Bewusst wurde dabei auf die praktische Evidenz verzichtet, die ihre Grenzen selbst bei vergleichsweise einfachen regelbasierten Denkaufgaben zeigt. So sind LLMs und RLMs (Reasoning Large Language Models) ohne angeschlossene Schachengine überraschend schlechte Schachspieler – obwohl ihnen praktisch die gesamte Schachliteratur und eine enorme Zahl gespielter Partien als Trainingsmaterial zur Verfügung stehen. Ein modernes RLM kann hochkomplexe sprachliche Argumentationen erzeugen und zugleich erhebliche Schwierigkeiten mit einer vollständig definierten symbolischen Welt aus nur 64 Feldern haben. Gerade Schach offenbart die Schwierigkeit der Modelle, eine stabile Repräsentation eines Gegenstandsbereichs über viele aufeinanderfolgende Schritte aufrechtzuerhalten. Sprachlich überzeugendes „Denken“ ist daher nicht mit einem Denken gleichzusetzen, das auf einer stabilen Repräsentation seines Gegenstandes operiert. In dieselbe Richtung weisen Untersuchungen kontrollierter algorithmischer Aufgaben wie Tower of Hanoi, River Crossing, Blocks World oder Checkers Jumping. Das Apple-Paper The Illusion of Thinking (Shojaee et al., 2025) zeigt, dass die Problemlösefähigkeit von RLMs oberhalb bestimmter Komplexitätsschwellen einbricht. Die Modelle können plausible Begründungs- und Lösungsschritte erzeugen, ohne deshalb über eine robuste und systematisch einsetzbare Problemlösekompetenz zu verfügen. Die Fähigkeit, Denken sprachlich zu simulieren, ist nicht mit der Fähigkeit gleichzusetzen, einen Gegenstand tatsächlich denkend zu beherrschen.

Indem ein LLM diese sprachlichen Zusammenhänge erschließt, erhält es daher einen vermittelten Zugang zur Welt. Eine eigene sinnliche Gegenstandsbeziehung ist dann keine notwendige Voraussetzung für begriffliches Denken.

Widerrede: Eine sprachlich vermittelte Beschreibung menschlicher Weltbeziehungen ist noch keine **eigene Weltbeziehung des LLM**. Zwar enthält Sprache die sedimentierten Erfahrungen und begrifflichen Bestimmungen von Menschen und vermittelt dem LLM insofern Informationen über die Welt. Doch der Zugang zu den sprachlichen Resultaten einer Weltbeziehung ist nicht mit dieser Weltbeziehung selbst gleichzusetzen.

Andernfalls müsste man bereits einer Enzyklopädie wie Wikipedia einen Zugang zur Welt zuschreiben: Auch in ihr sind in sprachlicher Form menschliche Erfahrungen, Erkenntnisse und begriffliche Bestimmungen in außerordentlichem Umfang gespeichert. Niemand würde jedoch allein daraus schließen, dass Wikipedia selbst die Welt erfährt oder begrifflich denkt. Der entscheidende Unterschied liegt offensichtlich zwischen dem **Enthalten von Weltwissen** und einem Subjekt, für das dieses Wissen Wissen *von einer Welt* ist.

Das LLM ist Wikipedia gegenüber zweifellos erheblich leistungsfähiger: Es speichert sprachliche Inhalte nicht lediglich, sondern kann ihre Strukturen verarbeiten, Zusammenhänge herstellen, verallgemeinern und neue sprachliche Äußerungen erzeugen. Doch diese zusätzliche funktionale Leistungsfähigkeit beantwortet noch nicht die entscheidende Frage. Auch das LLM begegnet zunächst nicht den Gegenständen, über die es spricht, sondern den sprachlichen Zeugnissen, in denen andere Subjekte ihre Begegnung mit diesen Gegenständen ausgedrückt haben.

Die Vermittlung durch Sprache beseitigt diese Differenz nicht, sondern setzt sie voraus: Damit Sprache etwas über die Welt vermitteln kann, muss diese Welt zuvor bereits von Subjekten erfahren und begrifflich bestimmt worden sein. Das LLM verfügt über die sprachlich objektivierten Resultate menschlicher Weltbeziehung; daraus folgt noch keine eigene Weltbeziehung des LLM.

Argument der abgeleiteten Bedeutung: Kann abgeleitete Bedeutung ohne eigene ursprüngliche Weltbeziehung zu eigener Bedeutung werden?

Warum muss ein System Bedeutung selbst ursprünglich in einer eigenen Weltbeziehung begründen, um über bedeutungsvolle Begriffe verfügen zu können? Könnte es Bedeutung nicht aus einem bereits bestehenden sprachlichen und gesellschaftlichen Bedeutungszusammenhang übernehmen?

Auch Menschen erzeugen ihre Begriffe keineswegs vollständig aus eigener Wahrnehmung. Wir werden in eine Sprache und Kultur hineingeboren und übernehmen den größten Teil unserer Begriffe von anderen. Wenn abgeleitete Bedeutung beim Menschen echte Bedeutung sein kann, warum nicht beim LLM?

Widerrede: Intelligenz besteht nicht darin, mit bereits sprachlich vorgegebenen Begriffen erfolgreich zu operieren, sondern darin, sich als Subjekt eine von ihm unterschiedene Welt begrifflich anzueignen. Wahrnehmung ist dabei nicht deshalb grundlegend, weil „echtes Denken Augen braucht“, sondern weil sich in ihr überhaupt erst eine Subjekt-Objekt-Beziehung konstituiert.

Das Grundproblem des LLMs sind daher nicht fehlende Sinnesorgane, sondern die fehlende **eigene Gegenstandsbeziehung**. Begriffliches Denken setzt keine spezifisch menschliche Sinnlichkeit voraus, wohl aber eine Welt, die einem Subjekt als Gegenstand gegenübertritt und innerhalb dieser Beziehung als etwas Bestimmtes begriffen werden kann.

Simulationsargument: *Vielleicht ist auch unsere vermeintlich reale Welt nur künstlich erzeugt.*

Wir können nicht mit letzter Gewissheit ausschließen, dass unsere gesamte erlebte Wirklichkeit künstlich erzeugt ist und wir selbst Teil einer Simulation sind. Daher könne auch die Unterscheidung zwischen einer „wirklichen“ menschlichen Weltbeziehung und einer künstlich erzeugten Weltbeziehung grundsätzlich infrage gestellt werden.

Widerrede: Selbst wenn wir simulierte Wesen wären, verfügten wir dennoch über eine eigene Weltbeziehung: Wir empfinden, nehmen wahr, handeln, erinnern uns und begegnen Gegenständen als Gegenständen. Für das vorliegende Argument ist daher unerheblich, ob die uns begegnende Welt „wirklich“ oder selbst Teil einer Simulation ist. Das Simulationsargument verschiebt die Frage lediglich auf eine grundsätzlich andere Ebene. Da die Annahme einer vollständigen Simulation weder empirisch belegt noch widerlegt werden kann, liefert sie keinen Einwand gegen die hier entscheidende Unterscheidung zwischen einem Subjekt mit eigener Weltbeziehung und einem System, dem eine solche fehlt. Das Simulationsargument löst das Problem der Weltbeziehung nicht, sondern verlegt lediglich die Welt, auf die sie sich bezieht. Entscheidend ist nicht, ob unsere Welt simuliert ist, sondern ob sie uns als Welt gegenübertritt. Ein LLM könnte also nicht einfach dadurch mit einem Menschen gleichgesetzt werden, dass möglicherweise *beide* auf einer Maschine laufen. Meine Unterscheidung betrifft nicht das Substrat sondern die Form der Weltbeziehung

Fortschrittsargument / Skalierungsargument: *Was heute fehlt, kommt mit mehr Daten, Rechenleistung und besseren Modellen.*

Gegenwärtige Defizite künstlicher Intelligenz sind lediglich vorläufige technische Beschränkungen, die durch Skalierung und weitere technische Entwicklung überwunden werden. Die gegenwärtigen Defizite von LLMs sind keine prinzipiellen Grenzen, sondern quantitative Beschränkungen. Mehr Rechenleistung, mehr Daten, größere Modelle und verbesserte Verfahren werden deshalb zu qualitativ höheren Fähigkeiten bis hin zu allgemeiner oder übermenschlicher Intelligenz führen.

Widerrede: Das Fortschritts- oder Skalierungsargument verfehlt den Kern meiner Argumentation. Dem LLM fehlt nicht lediglich **mehr** Intelligenz, sondern eine **konstitutive Voraussetzung von Intelligenz**: eine eigene Gegenstands- und Weltbeziehung. Aus mehr Daten, mehr Parametern und mehr Rechenleistung folgt jedoch nicht logisch eine solche Weltbeziehung. Das behauptete Defizit ist daher kein quantitatives, sondern ein **kategoriales**. Solange nicht gezeigt wird, wie Skalierung oder technischer Fortschritt derselben Art genau jene fehlende Weltbeziehung hervorbringen sollen, beantwortet das Fortschrittsargument den Einwand nicht, sondern verschiebt ihn lediglich in die Zukunft. Das Fortschrittsargument ersetzt eine Begründung durch eine Prognose.

Black box Argument (Argument der unbekannten Implementierung) : Vielleicht funktioniert menschliches Denken intern gar nicht grundsätzlich anders als künstliche neuronale Netze.

Wir verstehen die neuronalen Mechanismen menschlichen Denkens noch nicht hinreichend. Daher können wir nicht ausschließen, dass menschliches Denken auf funktional ähnlichen Prinzipien beruht wie künstliche neuronale Netze – etwa auf hochdimensionalen verteilten Repräsentationen. Der Hinweis auf Vektorräume, Wahrscheinlichkeiten oder statistische Verarbeitung bei LLMs beweist deshalb für sich genommen nicht, dass diese nicht denken oder verstehen.

Woher wissen wir, dass das menschliche Gehirn auf der relevanten Beschreibungsebene nicht ebenfalls etwas strukturell Ähnliches macht? Neuronale Repräsentationen werden in den Neurowissenschaften schließlich selbst häufig als Aktivitätsmuster in hochdimensionalen Zustandsräumen beschrieben.

Widerrede: Für meine Argumentation ist unerheblich, wie menschliche Begriffe neuronal realisiert sind. Entscheidend ist vielmehr, dass menschliches Denken die Tätigkeit eines Subjekts ist, das in einer eigenen sinnlichen und praktischen Beziehung zu einer von ihm unterschiedenen Welt steht und sich diese Welt begrifflich aneignet. Vielleicht arbeitet auch das menschliche Gehirn mit Strukturen, die sich als hochdimensionale Vektorräume beschreiben lassen. Das würde die hier behauptete Differenz nicht aufheben.

Der Gegensatz lautet daher nicht: **Vektorverarbeitung oder „echtes Denken“**. Auch menschliches Denken könnte neuronal auf einer Form von Vektorverarbeitung beruhen. Entscheidend ist vielmehr, **in welchen Zusammenhang diese Verarbeitung eingebettet ist und worauf sie sich bezieht**: Beim Menschen wäre sie Teil der Tätigkeit eines wahrnehmenden, handelnden und sich erinnernden Subjekts, das Gegenständen einer Welt begegnet und Begriffe von ihnen bildet. Beim LLM dagegen findet die Vektorverarbeitung innerhalb eines Sprachsystems statt, dessen Gegenstände zunächst sprachliche Zeichen und damit bereits sprachlich und begrifflich verarbeitete menschliche Weltbeziehungen sind. Die mögliche Ähnlichkeit der neuronalen und technischen Verarbeitung sagt daher noch nichts über die Gleichartigkeit dessen aus, was durch sie geleistet wird.

Version 1.3

2.10.2026

Versionsgeschichte

V1.3.; 02.20.2026: Fußnoten 46 und 47 hinzugefügt