

Strukturierte Repräsentation

»Gedanken ohne Inhalt sind leer«

Sommersemester 2025

05.03.2026

EINGEREICHT BEI

Friedrich-Schiller-Universität Jena

Fakultät für Mathematik und Informatik

Institut für Informatik

Seminar Neuronale Netze — Symbolverarbeitung — Kognition

Prof. Dr. Clemens Beckstein

Prof. Dr. Stefan Artmann

Johannes Mitschunas, M. Sc.

EINGEREICHT DURCH

Friedrich Answin Daniel Motz

daniel.motz@uni-jena.de

Inhaltsverzeichnis

1. Einleitung	1
2. Theory of Mind	2
3. Symbole und Bedeutung	4
3.1. Mentale Repräsentationen	4
3.2. Strukturierte Repräsentation	7
4. Abschlussbemerkung	8
Bibliografie	9

1. Einleitung

»Mentale Akrobaten« nennen sich Menschen, die innert weniger Sekunden die Reihenfolge von 52 Karten memorisieren und noch nach Tagen akkurat reproduzieren können. Was erstmal amüsant und sportlich ambitioniert scheint, erfordert Fähigkeiten, die Menschen täglich gebrauchen. Über Jahrtausende wurden glücklicherweise Medien zur Hirn-externen Speicherung von Informationen erfunden. Es war allerdings lange nötig, dies auch ohne »Behelf« zu schaffen.¹ Das Metrum, der Reim und auch eine Melodie dienten als *Gedächtnisstütze*. *Mnemotechniken* nennt man solche Methoden, mit deren Hilfe Bilder im Gedächtnis gezielt abgelegt und exakt abgerufen werden können. Selbst Mönche gebrauchten und lehrten sie zur Erinnerung von Gebeten [2]. Einige dieser Techniken sind auch heute noch bekannt, wie etwa der »Gedächtnispalast«. Uns erscheint es so, als wären wir bewusst in der Lage »virtuelle« Welten in unseren Gedanken zu erschaffen, die Entsprechungen mit realen Orten haben. Zusätzlich können wir diese Welten so manipulieren, dass wir sie zur Ablage von Informationen nutzen können. Informationen scheinen also nicht beliebig im Gedächtnis zu liegen, sondern in dekodierbare Strukturen übersetzt zu werden. Mit anderen Worten: wir kreieren, so kommt es uns zumindest vor, abrufbare, gedankliche Abbilder (realer) Gegenstände und Konzepte, die uns das Erinnern und Handeln ermöglichen.

Gedankliche Abbilder werden in verschiedenen Disziplinen diskutiert, darunter die der kognitive Psychologie, Neurologie, künstlichen Intelligenz (KI) und Philosophie. Der Begriff *strukturierte Repräsentation* findet sich vorwiegend in der Literatur der ingenieurmäßigen Realisierung der Theorien der Kognitionswissenschaften (etwa [3], [4], [5]), während sonst überwiegend von *mentalen Repräsentationen* die Rede ist (etwa [6], [7], [8]). Diese Begriffe beschreiben jedoch ähnliche Ideen aus verschiedenen Richtungen.

Längst geht es jedoch nicht mehr nur um Abbilder, sondern vielmehr um die Voraussetzungen für das, was man vulgär *denken* nennt. Einige Fragen müssen notwendigerweise geklärt werden, bevor mentale Repräsentationen im Kontext der Theorien, die versuchen, sie zu beantworten, überhaupt diskutiert werden können; zum Beispiel: Wie erhalten diese mentalen Bilder ihre *Bedeutung* und in welcher Relation stehen sie zu ihren *Inhalten*? Wie kann man im Allgemeinen mehr über mentale Repräsentation erfahren?

Diese Diskussionen finden sich sowohl in der Philosophie, als auch in der Biologie, Psychologie, Informatik und allem, was den Neurowissenschaften zuzuordnen ist. In der »Philosophy of Mind« und Semiotik wird die Frage der Zuweisung von Bedeutung unter dem *Intentionalitätsproblem* geführt. Es ist einerseits fraglich, ob und wie der Mensch intentionale Propositionen vertreten kann und, unter der Voraussetzung, dass er es kann, welche Anlagen dazu notwendig sind. Die Forschung ist eng verwandt mit der Diskussion über Bewusstsein und auch der Fähigkeit zu denken an sich. Im Versuch der Annäherung an und Nachbildung der menschlichen Fähigkeiten des Denken, erlangte der Zweig der KI (unter einigen)

¹Vor der ersten schriftlichen Fassung der Odyssee wurde sie wohl schon einige Zeit mündlich überliefert stammt womöglich nicht alleinig aus der Feder Homers [1].

den Rang eines adäquaten Modells zur Beschreibung (menschlicher) mentaler Phänomene. Dieser Ansatz wurde aufgrund einer mangelhaften Begründung vom Vorhandensein einer Intentionalität in solchen Systemen kritisiert (bspw. Searle, Dreyfus). Ein geläufiger Ansatz ist daher, sich über die Physiologie des menschlichen Gehirns anzunähern, wie es in den heutigen Neurowissenschaften Praxis ist, denn es sei denkbar, dass die Aktivierungen im Hirn (in Form elektromagnetischer Impulse) den Denkprozessen und mentalen Repräsentationen entsprechen und sie daher erklären könnten [9]. Einige Philosophen vertreten sogar die Auffassung, dass ausschließlich Neuronen in der Lage seien, die notwendigen Anlagen zu liefern. Überraschenderweise glauben auch namhafte Unternehmen im Bereich des maschinellen Lernens, dass sie durch das Beobachten der Aktivierungen in überwacht trainierten neuronalen Netzen großer Sprachmodelle Konzepte und sogar »mentale Prozesse« entdecken können, die erklären, wie und warum diese Modelle ihre Ausgaben (Outputs) erzeugen: »[...] like one of those fMRI brain images« [10], [11]. Unter Kognitionspsychologen hingegen ist es Praxis, sich über Introspektion, Sprache und Verhaltensbeobachtung anzunähern. Etwa werden Probanden gebeten, ihre *Gedanken* zu einem Lösungsprozess zu erläutern [7]. Hier lieferte die KI ebenfalls wichtige Anstöße [12], bspw. »Wie können KI-Systeme Sachverhalte oder Konzepte repräsentieren, »erlernen« oder speichern und Erkenntnisse erschließen; und wie sieht es dann vermutlich beim Menschen aus?«

Diese Diskussionen lassen sich schwerlich abdecken und sicherlich nicht alleinig mit diesen zwei Fragen, jedoch ist es aus praktikablen Gründen notwendig, sich auf die Bearbeitung zweier Fragen zu beschränken:

1. **Was versteht man unter einem Mind?** (Abschnitt 2)
2. **Was ist unter einer Repräsentation zu verstehen?** (Abschnitt 3)

2. Theory of Mind

Die notwendigen und hinreichenden Bedingungen dessen, was wir »denken«, Mind² und mentale Repräsentation nennen, werden bis heute diskutiert [13], [14], sowie gleichsam Möglichkeiten der ingenieurmäßigen Umsetzung [4], [5], [10]. Es konkurrieren diverse Positionen zu den Fragen, ob das menschliche Mind über kognitive Prozesse verfügt (was auch immer das konkret heißen möge) und ob sie über freien Willen verfügen (oder lediglich Spielball kausaler Zusammenhänge auf physikalischer Ebene sind). Es gibt, wie bereits aufgeführt, verschiedene Versuche sich der tatsächlichen Funktionsweise des Minds über das Gehirn anzunähern, gegeben, dass es tatsächlich darüber verfügt, was man zu finden gedenkt. Für die Zwecke dieser Arbeit wird unterstellt, dass das Mind seinen Grund in der physischen und sinnlich wahrnehmbaren Welt hat. Diese Annahme scheint mir zur Fortsetzung der Diskussion notwendig, da mir wenig Literatur und Wissen über die Welt, außerhalb der sinnlich wahrnehmbaren, zur Verfügung steht.

²Der Begriff »Mind« scheint geeigneter als Übersetzungen wie »Verstand« oder »Geist«, da sie im Dt. einen ungewollten Schwerpunkt setzen könnten.

In den Theorien, die annehmen, dass die Gründe für unser Verhalten physikalisch sind, gibt es zwei populäre Bewegungen: eine 1) **eliminativistisch-materialistische** oder auch reduktionistisch-physikalistische, die annimmt, dass alles Mentale auf physische Zusammenhänge reduziert werden kann (bspw. [15]) und eine 2) **nichtreduktionistische**, die mitunter annimmt, dass aufgrund des adäquaten Zusammenspiels auf physischer Ebene komplexeres Verhalten *emergiert*³, oder, dass mentale Zustände über ihre *Funktion* und *Rolle* definiert sind, unabhängig von der konkreten Implementation⁴ (Silikonchip, »Wetware« oder »Alien Goo«⁵) [12].⁶ Diese Arbeit befasst sich hauptsächlich mit der nichtreduktionistischen Theorie.

Premise: Brains cause minds.

Premise: Any system capable of causing minds must have powers equivalent to those of brains.

Premise: Syntax alone is not sufficient for causing minds.⁷

Conclusion: Any artificial system that could cause minds must have more than just formal syntactic processes; it must have something that replicates the powers of the brain.

nach John Searle, 1984⁸

Abbildung 1: Searles Analyse des Nichtreduktionismus in einer Debatte mit Professor Margaret Boden.

Da der eliminativistisch Materialismus die Existenz eines Minds als Phänomen über der Physiologie ablehnt und damit auch mentale Repräsentation als überholte Darstellung ansieht, ergibt seine Diskussion im Rahmen der Seminararbeit als solches in diesem Kontext wenig Sinn. Prinzipiell schließt der Materialismus nicht aus, dass ein Gehirn dann in einen Zustand gerät, der Verhalten verursacht, jedoch seien dies rein neuronale Phänomene, die auf dieser Ebene erklärt werden müssen [15].

Ein seit einiger Zeit vorherrschender Glaube ist, dass man etwas über das Mind lernen kann, wenn man sich an Berechnungsmodellen wie der Turingmaschine orientiert und das Mind ebenfalls als ein solches auffasst. Die daraus partiell motivierte Kognitionspsychologie sieht die Computer-Hirn Analogie einerseits als Tradition, das Mind mit dem zu vergleichen, was dem technischen Stand entspricht (Aristoteles mit einem Taubenschlag, später Mühlen, Telefonschalttafeln, ...); und heute sei es eben der Computer [7]. Andererseits ist man auch der Überzeugung, dass das Mind eben genau ein symbolverarbeitendes System ist und, dass sie intern exakt so verarbeitet werden: »It is clear that computers and people are both symbol systems in the sense that external and internal events are symbolised, and the resultant symbols are then manipulated in various ways« [7].

³gemeint ist die im engl. Emergentism genannte Strömung

⁴gemeint ist die im engl. Functionalism genannte Strömung

⁵Ein wörtliches Zitat von Searle

⁶den Ansatz der Supervenienz nach etwa Jaegwon Kim und Donald Davidson möchte ich zwar erwähnen, aber im Rahmen dieser Arbeit zur nichtreduktionistischen Theorie verschieben

⁷Searle entnimmt dies dem »Chinese Room«-Gedankenexperiment

⁸In [dieser Diskussion](https://youtu.be/b5a9NJ10iGQ), <https://youtu.be/b5a9NJ10iGQ>

Bekannte Gegenargumente zeigen das gänzliche Fehlen der Intentionalität in formalen Systeme auf (etwa [16], [14, S. 10]). Dieses Problem und auch das der Repräsentationen, wird im folgenden Abschnitt behandelt.

3. Symbole und Bedeutung

Ein naiver Symbolbegriff wird häufig mit der scholastischen Formel *aliquid stat pro aliquo* (»etwas steht für etwas anderes«) umschrieben. Damit ist die Grundidee gemeint, dass ein Zeichen stellvertretend für ein Anderes (möglicherweise einen Gegenstand, eine Person, ein Konzept) steht [17]. Jedoch greift diese einfache Stellvertreter-Relation zu kurz: Es ist nicht in der Sache des Zeichens an sich, welche Bedeutung es hat, da es eines interpretierenden Subjekts und eines Bedeutungszusammenhangs bedarf [18]. So definierte etwa Peirce ein Zeichen explizit als dreistellige Relation, nämlich »etwas, das für jemanden in irgendeiner Hinsicht für etwas steht«. Ein Symbol erhält seinen Inhalt folglich durch die Intentionalität des Mind: die mentale Fähigkeit, sich auf Gegenstände, Eigenschaften oder Zustände der Welt zu beziehen [18]. Während die Semiotik solche Bedeutungsrelationen analysiert, betrachtet die Informationstheorie Information zunächst ohne Bezug auf Semantik. Shannon formulierte 1948 Information als statistische Größe der Signalübertragung und betonte, dass die »semantischen Aspekte der Kommunikation [für das technische Problem] irrelevant« seien [19]. Gleichwohl haben Philosophen – etwa Fred Dretske – versucht, an Shannons Begriff anzuknüpfen, indem sie die in Signalen enthaltene Information als Basis für mentale Repräsentationen und Bedeutung heranziehen [20]. Allerdings reicht eine reine Kausal- oder Informationsbeziehung nicht aus, um die Semantik zu erklären: Baumringe, Rauch oder ein Thermometer tragen zwar Information über ihre Ursachen⁹, aber sie repräsentieren diese nicht im vollen Sinne. Erst wenn ein kognitives System solche Information in einem intentionalen, bedeutungstragenden Zustand verarbeitet, kann man von einer mentalen Repräsentation mit semantischem Gehalt sprechen, also von einem Symbol im anspruchsvolleren Sinn, das etwas für etwas steht und diesem »Stehen-für« oder »Aboutness« durch Intentionalität und Bezugnahme auf Bedeutungen Substanz verleiht [6], [14]. Genau diesen Mangel der Intentionalität in rein syntaktischen Systemen nahmen u.a. Searle¹⁰ und Dreyfus¹¹ zum Anlass ihrer Kritiken. Das Intentionalitätsproblem (für Menschen wie auch Maschinen) scheint bis heute noch nicht endgültig gelöst, und bildet auch ein Problem für die Diskussion mentaler Repräsentationen.

3.1. Mentale Repräsentationen

Einige Autoren zum gleichen Thema (insbesondere der Psychologie) fordern für eine (mentale) Repräsentation einen klaren Bezug von einem *systeminternen* zu einem *systemexternen*

⁹in der Semiotik »Index« genannt

¹⁰Unter anderem in [diesem Talk](#)

¹¹Unter anderem in [21]

Zustand [9, S. 2].¹² Es ist fraglich, wie dann eine (mentale) Repräsentation auf bspw. ein Konzept einzuordnen ist; sind sie systemextern, also sind Konzepte real? Und wenn man sich schon auf eine konkrete Sache (bspw. diese Kaffeetasse vor mir) bezieht; wie charakterisiert sich der Bezug zu einem realen Objekt? Ein erst kürzlich schriftlich publizierter Vorschlag [14] positioniert sich vorsichtig:

1. Construing a mental state as a representation does not presuppose a special, substantive relation – what we can call a representation relation – holding between the mental state and what the representation is about, for example, the nearby coffee cup, or Paris or Napoleon.
2. Content is not an essential property of mental states – that is, the very same kind of mental state might have had a different content, or no content at all.
3. The attribution of content to mental states is always pragmatically motivated. It serves to characterize, for certain uses and purposes, features of the mind that are not themselves intentional. [...] As I put it, content serves to gloss these features for the intended purposes.

Frances Egan, Deflating Mental Representation, 2025 [14, S. 2]

Mit diesem »deflationary account« (im Weiteren Deflationary Account on Mental Representations, kurz *DAMR* genannt) stellt Egan die traditionelle Vorstellung infrage, dass mentale Repräsentationen eine Art mysteriöse, substanzielle Verbindung zur Welt haben müssen. Ihre Position lässt sich so zusammenfassen:

- *Keine Notwendigkeit einer (substanzielle) Repräsentationsrelation.* Die Deutung eines mentalen Zustands als Repräsentation beruht nicht auf einer festen »Repräsentationsbeziehung« zwischen dem Zustand und seinem Gegenstand.¹³ Die Repräsentation ist eine Frage der Interpretation, nicht einer metaphysischen Verbindung.
- *Inhalt¹⁴ ist nicht essenziell und variabel.* Der spezifische Inhalt eines mentalen Zustands ist eine variable Eigenschaft. Derselbe »Typ« mentalen Zustands könnte einen anderen oder gar keinen Inhalt haben.
- *Die Zuschreibung eines Inhalts ist pragmatisch (und damit potenziell fehlerhaft bzw. nicht repräsentativ).* Wenn wir mentalen Zuständen einen Inhalt zuschreiben, tun wir dies aus praktischen Gründen; insbesondere, um bestimmte Merkmale mentaler Aktivität für unsere Erklärungszwecke hervorzuheben. Der Inhalt ist, wie Egan es formuliert, ein »Gloss« – eine Art nützlicher Kommentar, der an sich nicht-intentionale Prozesse für uns *handhabbar* macht.

Außerdem wird ein Vorschlag, den sie *external sortalism* nennt, gemacht, wie ein Gloss verliehen werden kann: Um diese mentalen Zustände mit einem Gloss zu versehen oder zu deuten, verwenden wir die Außenwelt als eine Art Gerüst oder Modell. Wir *sortieren* und *verstehen* interne Zustände, indem wir sie auf externe Objekte und Kategorien beziehen [14, u.a. S. 4].

¹²Und auch andere Autoren in der gleichen Reihe, Hrsg. Engelkamp & Pechmann. Die Ansichten entsprechen weitgehend dem robusten Realismus.

¹³konträr zur Position, die von [9] unter Kognitionspsychologen als Konsens postuliert wird

¹⁴original: content, womit womöglich eigentlich auf die Semantik angespielt wird

Der Gegensatz zwischen dem DAMR und robustem Realismus betrifft also vorrangig die Frage, wie der Gehalt mentaler Repräsentationen zu verstehen ist. Der robuste Realismus, der häufig mit kausalen oder informationstheoretischen Ansätzen verbunden wird, geht davon aus, dass mentale Repräsentationen einen bestimmten, objektiven Inhalt besitzen, der eine intrinsische, geistunabhängige Eigenschaft der entsprechenden computerähnlichen Zustände ist. Dieser Inhalt wird als eine substanzielle, naturalistisch beschreibbare Beziehung zwischen dem mentalen Zustand und der Welt verstanden.

Im Gegensatz dazu argumentiert Egan, dass die Zuschreibung von Inhalt – etwa die Annahme, eine Erfahrung repräsentiere „eine rote Tomate“ – lediglich eine pragmatisch motivierte Ergänzung oder ein nützliches heuristisches Mittel sei. Der Inhalt ist demnach keine grundlegende, intrinsische Eigenschaft des mentalen Zustands selbst, sondern eine sekundäre, extrinsische Zuschreibung, die wir aus erklärungs-technischen Gründen vornehmen. Dieser Gloss, also interpretative Überlagerung, dient dazu, die formalen, nicht-semantischen Mechanismen des Gehirns, mit der Welt des rationalen menschlichen Verhaltens, in Beziehung zu setzen.

Egan bietet außerdem an, dass unser Verständnis von Wahrnehmungserfahrungen über eine Abbildung zwischen zwei Strukturen funktioniert:

- *S1: Das Modell der Außenwelt.* Dies ist unsere stabile, konzeptuelle Karte der Welt, die im Laufe der Zeit durch sensorische und motorische Interaktionen aufgebaut wird.
- *S2: Wahrnehmungserfahrungen.* Dies sind unsere privaten, inneren Zustände.

Das setzt voraus, dass es ein Modell der Außenwelt existiert. Dieses sei durch wiederholte Wahrnehmung und haptische Interaktion mit der Umwelt erlangt und stabilisiert worden. Es handele sich um einen Entwicklungsprozess, der in der frühen Kindheit entstehe [14, S. 118]. Sie verweist auf die Entwicklungspsychologie, um zu erklären, dass Kinder zunächst den Unterschied zwischen Erscheinung und Realität begreifen müssen. Bevor ein Kind das Konzept entwickelt, dass Objekte stabile Eigenschaften besitzen (z. B. dass ein Auto wirklich rot ist, auch wenn ein Filter es schwarz erscheinen lässt), fehlt ihm ein stabiles externes Bezugsrahmenmodell. Im Wesentlichen lernen wir zunächst, aus einzelnen Erfahrungen zu abstrahieren und zu verallgemeinern, um eine stabile, allozentrische (d. h. nicht selbstzentrierte) Vorstellung der Welt aufzubauen. Erst wenn dieses externe Modell (S1) etabliert ist, können wir es nutzen, um unsere privaten Erfahrungszustände (S2) zu kategorisieren und zu verstehen. Da diese beiden Bereiche (S1 und S2) eine ähnliche Struktur teilen, können wir das Modell der Außenwelt (S1) verwenden, um unsere Wahrnehmungserfahrungen (S2) zu konzeptualisieren und über sie zu reflektieren [14, S. 118-119].

Egan identifiziert wichtige Probleme der Philosophy of Mind, jedoch bietet ihre Theorie noch keine handfeste Lösung. Wie sie selbst bekennt, ist der *external sortalism* eine Abwandlung des Repräsentationalismus, der die guten Aspekte behält und die problematischen

Festlegungen aufweicht [14, S. 109]. Jedoch bleibt etwa, **wie ein Gloss festgelegt wird, und auch welcher**, unterbestimmt.

From these premises, it follows that the same »vehicle« could support different representational contents, depending on the pragmatic aims of researchers operating within different scientific or explanatory contexts.

Pereira & Barcellos, 2025, [22]

Egan argumentiert, dass dies eine komplexe heuristische Funktion sei. So wie wir etwa einen Schmerz als Repräsentation von »Gewebeschaden« auffassen, um seine typische Ursache hervorzuheben, fassen wir auch eine Wahrnehmungserfahrung mit Inhalt auf, um sie mit dem Zustand der Außenwelt zu verknüpfen – und zwar zu dem Zweck, Bewertungen vorzunehmen und Überzeugungen zu bilden [14, S. 141].

3.2. Strukturierte Repräsentation

Marcus schafft einen Vorschlag für Forderungen an Symbolmanipulation, die er unbedingt erfüllt sehen möchte und die, seiner Auffassung nach, einzeln genommen bewiesen und widerlegt werden können, ohne den Gehalt einer anderen zu berühren:

- The mind represents abstract relationships between variables.
- The mind has a system of recursively structured representations.
- The mind distinguishes between mental representations of individuals and mental representations of kinds.

Gary F. Marcus, [3, S. 3]

Die von ihm angeführten Vorschläge (etwa von Ramsey, Stich, Garon, 1990) erfüllen seine Anforderungen nicht [3, S. 92-93]. Stattdessen seien semantische Netze ein geeignetes Vorbild, weil sie theoretisch beliebig rekursive Repräsentationen (auch *productivity of thought*) ermöglichen. Eine weitere Forderung, die Marcus zwar nicht benennt, durch sein Beispiel semantischer Netze allerdings implizit fordert, ist die Systematizität¹⁵.

Marcus versucht, die von ihm identifizierten Anforderungen mit einem Ansatz, den er *Treelets* nennt, zu erfüllen. Um beliebig rekursive Repräsentationen in neuronalen Netzen zu realisieren, führt er einen gerichteten Graphen ein, dessen Struktur von einem semantischen Netz motiviert ist [3, S. 108]. Es bleibt allerdings offen, wie etwa die Begriffe für die Inhalte der Treelets gewählt werden sollen. Auch die Forderung, dass ein großer Vorrat leerer Zellen vorgehalten werden soll wirft die (von ihm nicht beantwortete) Frage auf, nach welchen Prinzipien die Zellen gefüllt werden.

Es ist aus verschiedenen Gründen naheliegend unüberwachte Lernverfahren einzusetzen. Diesen Ansatz findet man auch unter anderem bei [4], [5]. Insbesondere [4] scheint vielversprechend, da Systematizität und Produktivität (of Thought) als Problem benannt werden. Allerdings werden für das Optimieren der (Vektor-)Repräsentationen im Training »inductive biases« verwendet [4, S. 98].

¹⁵Die Systematizität (im Kontext der Language of Thought Hypothesis) ist das Prinzip, dass unsere Fähigkeit, einen Gedanken zu denken – etwa »John liebt Mary« – impliziert, dass wir auch in der Lage sind, seine systematischen Varianten zu denken, wie zum Beispiel »Mary liebt John« (vgl. Fodor, Pylyshyn).

4. Abschlussbemerkung

Die Diskussion um mentale und strukturierte Repräsentationen zeigt, dass es sich dabei weniger um eine rein technische oder empirische, sondern vielmehr um eine konzeptuelle Herausforderung handelt. Weder ein rein materialistischer Zugang noch ein deflationärer Ansatz allein können das Phänomen vollständig erfassen. Während klassische Positionen wie der robuste Realismus versuchen, mentale Zustände über kausale Relationen zur Welt zu verankern, betont der deflationäre Ansatz, dass Bedeutungszuschreibungen stets pragmatisch und kontextabhängig sind.

Gedanken ohne Inhalt sind leer, Anschauungen ohne Begriffe sind blind.

Kant, Kritik der reinen Vernunft, 1787

Ansätze wie die Embodied Cognition deuten an, dass Repräsentation und Bedeutung nicht nur im Individuum, sondern im Zusammenspiel von Körper, Umwelt und sozialem Handeln verortet ist. Repräsentationen wären demnach keine statischen inneren Abbilder, sondern dynamische Muster des In-der-Welt-Seins. In diesem Sinne verschiebt sich der Fokus von der Frage, wie der Geist die Welt abbildet, zu der Frage, wie er in ihr wirksam ist.

Gleichwohl bleibt offen, ob das, was wir »Denken« nennen, überhaupt adäquat durch Repräsentationen beschrieben werden kann – oder ob Repräsentation selbst nur eine heuristische Metapher ist, die es uns erlaubt, komplexe Prozesse des Bewusstseins begrifflich zu ordnen. Künftige Forschung – sowohl in der Philosophy of Mind als auch in der künstlichen Intelligenz – wird zeigen müssen, ob diese Metapher trägt oder einer grundlegend anderen Beschreibung weichen muss.

Bibliografie

- [1] J. Foer, *Moonwalking with Einstein: The Art and Science of Remembering Everything*. New York: Penguin Press, 2011.
- [2] F. A. Yates, *The Art of Memory*. Penguin Books, 1978.
- [3] G. F. Marcus, *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. The MIT Press, 2001. doi: [10.7551/mitpress/1187.001.0001](https://doi.org/10.7551/mitpress/1187.001.0001).
- [4] A. Dittadi, „On the Generalization of Learned Structured Representations“. [Online]. Verfügbar unter: <https://arxiv.org/abs/2304.13001>
- [5] A. Kumar und P. Schrater, „Structured Representation“. [Online]. Verfügbar unter: <https://arxiv.org/abs/2505.12143>
- [6] J. Engelkamp und T. Pechmann, Hrsg., *Mentale Repräsentation*, 1. Aufl. Bern, Göttingen, Toronto, Seattle: Huber, 1993.
- [7] M. W. Eysenck, *Principles of Cognitive Psychology*, Reprint_1993, 1994 Aufl. in *Principles of Psychology*. Hove, UK and Hillsdale, USA: Lawrence Erlbaum Associates, 1993.
- [8] J. R. Searle, „Minds, Brains, and Programs“, *Behavioral and Brain Sciences*, Bd. 3, Nr. 3, S. 417–424, 1980, doi: [10.1017/S0140525X00005756](https://doi.org/10.1017/S0140525X00005756).
- [9] J. Engelkamp und T. Pechmann, „Kritische Anmerkungen zum Begriff der mentalen Repräsentation“, *Mentale Repräsentation*. Huber, S. 7–16, 1993.
- [10] Anthropic, „Interpretability: Understanding how AI models think“. Zugegriffen: 26. September 2025. [Online]. Verfügbar unter: <https://www.youtube.com/watch?v=fGKNUvivvnc>
- [11] J. Lindsey u. a., „On the Biology of a Large Language Model“, *Transformer Circuits*, März 2025, Zugegriffen: 27. September 2025. [Online]. Verfügbar unter: <https://transformer-circuits.pub/2025/attribution-graphs/biology.html>
- [12] E. N. Zalta und U. Nodelman, „The Computational Theory of Mind“, *The Stanford Encyclopedia of Philosophy*. 2024. Zugegriffen: 24. September 2025. [Online]. Verfügbar unter: <https://plato.stanford.edu/entries/computational-mind/>
- [13] A. Noë, „Can computers think? No. They can’t actually do anything“. [Online]. Verfügbar unter: <https://aeon.co/essays/can-computers-think-no-they-cant-actually-do-anything>
- [14] F. Egan, *Deflating Mental Representation*. The MIT Press, 2025. doi: [10.7551/mitpress/15666.001.0001](https://doi.org/10.7551/mitpress/15666.001.0001).
- [15] P. M. Churchland, „Eliminative Materialism and the Propositional Attitudes“, *The Journal of Philosophy*, Bd. 78, Nr. 2, S. 67–90, 1981, [Online]. Verfügbar unter: <http://www.jstor.org/stable/2025900>
- [16] J. R. Searle, „Is the Brain a Digital Computer?“, *Proceedings and Addresses of the American Philosophical Association*, Bd. 64, Nr. 3, S. 21–37, 1990, Zugegriffen: 3. Oktober 2025. [Online]. Verfügbar unter: <http://www.jstor.org/stable/3130074>

-
- [17] G. Harras, T. Herrmann, und J. Grabowski, „Aliquid stat pro aliquo — aber wie?“, *Bedeutung · Konzepte. Bedeutungskonzepte*. in Psycholinguistische Studien. VS Verlag für Sozialwissenschaften, Opladen, S. 9–19, 1996. doi: [10.1007/978-3-322-86544-1_1](https://doi.org/10.1007/978-3-322-86544-1_1).
- [18] M. Bratman, „Intentionality“, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, 2023. Zugegriffen: 23. September 2025. [Online]. Verfügbar unter: <https://plato.stanford.edu/entries/intentionality/>
- [19] O. Lombardi, F. Holik, und L. Vanni, „What is Shannon Information?“, *Philosophy of Science Archive*, Nr. 10911/1, 2019, [Online]. Verfügbar unter: https://philsci-archive.pitt.edu/10911/1/What_is_Shannon_Information.pdf
- [20] D. Pitt, „Mental Representation“, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, 2022. Zugegriffen: 23. September 2025. [Online]. Verfügbar unter: <https://plato.stanford.edu/archives/fall2022/entries/mental-representation/>
- [21] H. L. Dreyfus, *What Computers Still Can?T Do: A Critique of Artificial Reason*. MIT Press, 1992.
- [22] R. H. de Sá Pereira und V. M. Barcellos, „Review of Egan F. Deflating Mental Representation. Cambridge, Ma: MIT Press, 192, 2025.“, *Manuscrito*, Bd. 48, Nr. 4, S. 1–10, 2025, [Online]. Verfügbar unter: <https://philarchive.org/rec/PERROE-5>

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende schriftliche Hausarbeit (Seminararbeit, Belegarbeit) selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe. Die Stellen der Arbeit, die anderen Werken wörtlich oder sinngemäß entnommen sind, wurden in jedem Fall unter Angabe der Quellen (einschließlich des World Wide Web und anderer elektronischer Text- und Datensammlungen) kenntlich gemacht. Dies gilt auch für beigegebene Zeichnungen, bildliche Darstellungen, Skizzen und dergleichen. Ich versichere weiter, dass die Arbeit in gleicher oder ähnlicher Fassung noch nicht Bestandteil einer Prüfungsleistung oder einer schriftlichen Hausarbeit (Seminararbeit, Belegarbeit) war. Mir ist bewusst, dass jedes Zuwiderhandeln als Täuschungsversuch zu gelten hat, aufgrund dessen das Seminar oder die Übung als nicht bestanden bewertet und die Anerkennung der Hausarbeit als Leistungsnachweis/Modulprüfung (Scheinvergabe) ausgeschlossen wird. Ich bin mir weiter darüber im Klaren, dass das zuständige Lehrprüfungsamt/Studienbüro über den Betrugsversuch informiert werden kann und Plagiate rechtlich als Straftatbestand gewertet werden.

Ort: **Unterschrift:**

Datum: