

Paradigmen im Kampf

Dreyfus, Wittgenstein und ein »Scheideweg«
in der Künstlichen Intelligenz

Seminararbeit
Erhard Konrad:
Grenzen der Künstlichen Intelligenz
Sommersemester 2000

Christian Siefkes
Frank Eckert

Technische Universität Berlin
Skr. FR 6-2
Franklinstr. 28/29, 10587 Berlin

christian@siefkes.net, freck@cs.tu-berlin.de
<http://www.siefkes.net/>

21. Dezember 2000
Version vom 28. April 2002

Inhalt

1	Einleitung	2
2	Beschreibung der beiden Forschungsprogramme	2
2.1	Symbolrepräsentation	3
2.2	Konnektionismus	4
3	Philosophischer Hintergrund	5
3.1	Reduktionistische Tradition	5
3.2	Holistische Kritik am Reduktionismus	8
3.2.1	Nicht alles lässt sich definieren	9
3.2.2	Nicht alles lässt sich klassifizieren	11
3.2.3	Es gibt keine kontextfreien einfachen Elemente	12
3.2.4	Geist ohne Körper?	13
4	Geschichte	13
4.1	Ähnliche Entwicklung in den 50er und 60er Jahren	13
4.2	Sieg der Symbolrepräsentation in den 70er Jahren	14
4.3	»Comeback« des Konnektionismus	15
5	Ausblick	16
	Literatur	18

1 Einleitung

Der Philosoph und KI-Kritiker Hubert Dreyfus sieht die Entwicklung der Künstlichen Intelligenz als geprägt durch – bisweilen erbitterte – Auseinandersetzungen zwischen den Anhängern¹ zweier grundlegend verschiedener Paradigmen (im Kuhn'schen Sinne) bzw. Forschungsprogramme (in der Terminologie Imre Lakatos' – wir verwenden die beiden Begriffe synonym). Beide behaupten, Computern intelligentes Verhalten beibringen zu können, verfolgen dabei aber sehr unterschiedliche Ansätze.

Auf der einen Seite steht der Ansatz der Symbolrepräsentation, der beansprucht, die Welt durch Formalisierung für Computer zugänglich zu machen. Sein Handwerkszeug ist die Logik, ihre Auffassung von Intelligenz sieht diese Schule in »Problemlösung« verwirklicht. Auf der anderen Seite steht der Konnektionismus, der die Funktion biologischer Neuronen simuliert und damit erreichen will, dass Computer gewissermassen von sich aus »lernen«. Dazu bedient sich diese Seite der Statistik.²

Historisch haben beide Schulen zu gewissen Zeiten beachtliche Erfolge in der praktischen Anwendung hervorgebracht, die einige ihrer Vertreter gerne als endgültigen Durchbruch des jeweiligen Forschungsprogramms gedeutet hätten. Ein solcher endgültiger Triumph einer Seite steht jedoch bis heute aus.

Beide Ansätze beruhen auf weitgehenden philosophischen Annahmen über die Welt und über biologische bzw. menschliche Intelligenz, ohne diese jedoch explizit herauszustellen. Ziel unserer Arbeit ist es, die beiden Positionen und insbesondere die zugrunde liegenden Vorstellungen deutlich zu machen. Dazu gehen wir nach einer Vorstellung der beiden entgegengesetzten Paradigmen ausführlich auf deren philosophische Hintergründe ein. Dabei beschäftigen wir uns besonders mit Ludwig Wittgenstein, der für beide Positionen von großer Relevanz ist. Abschließend geben wir einen kurzen Abriss der historischen Entwicklung beider Seiten.

2 Beschreibung der beiden Forschungsprogramme

1. Im Singular verwenden wir meist die weibliche Form, im Plural die kurze, die wir als geschlechtsneutral verstehen.

2. vgl. [Dreyfus 1994](#), 215 f

2.1 Symbolrepräsentation

Aus Sicht der Anhänger der Symbolrepräsentation entsteht intelligentes Verhalten durch die Formalisierung der Realität mittels Symbolen und den Umgang mit diesen Symbolen nach formalen Regeln. Die entscheidende Eigenschaft von Computern, die diese zu »intelligentem« Verhalten befähigt, ist somit ihre Fähigkeit zum Operieren mit solchen Symbolen. Diese Fähigkeit betonen unter anderem ab 1955 Allen Newell und Herbert Simon, die erkennen, dass Bitfolgen nicht nur Zahlen, sondern letztlich beliebige Dinge repräsentieren können. In Newells Worten »betrachteten die KI-Forscher Computer als Maschinen, die Symbole manipulieren. Die entscheidende Sache dabei war, wie sie sagten, daß beliebige Dinge als Symbole kodiert werden können, sogar Zahlen.«³

Programme lassen sich als Regeln über Beziehungen zwischen diesen Symbolen deuten; damit kann ein Computersystem »von sich aus« mittels logischer Schlüsse Fakten über die repräsentierten Objekte ableiten, die ihm noch nicht explizit bekannt sind.⁴

Für Newell und Simon sind das menschliche Gehirn und Computer beide physikalisch-symbolische Systeme, die zwar grundverschieden aufgebaut sind, deren Funktion sich aber auf einer abstrakten Ebene gemeinsam beschreiben lässt. Beide gelangen durch Operationen mit Symbolen zu intelligentem Verhalten:⁵

Die These vom physikalisch-symbolischen System: Ein physikalisch-symbolisches System verfügt über die notwendigen und hinreichenden Mittel für allgemeines intelligentes Verhalten. Mit »notwendig« meinen wir, daß jedes beliebige System, das allgemeines intelligentes Verhalten zeigt, sich bei seiner Analyse als physikalisch-symbolisches System herausstellen wird. Mit »hinreichend« meinen wir, daß jedes beliebige physikalisch-symbolische System hinreichender Größe weitergehend so organisiert werden kann, daß es allgemeines intelligentes Verhalten zeigt.⁶

Sie formulieren dies im selben Artikel sogar noch schärfer: »Das Studium von Logik und Computern hat uns offenbart, daß Intelligenz auf physikalisch-symbolischen Systemen beruht.«⁷

3. [Newell 1983](#), 196 – zitiert nach Dreyfus

4. vgl. [Dreyfus 1994](#), 215 f

5. vgl. [Dreyfus 1994](#), 216 f

6. [Newell und Simon 1981](#), 41 – zitiert nach Dreyfus

7. [Newell und Simon 1981](#), 64 – zitiert nach Dreyfus

Da somit in diesem Ansatz sämtliches intelligente Verhalten auf der Möglichkeit beruht, die reale Welt, in der wir leben, in formalen Symbolen zu beschreiben, muss diese Möglichkeit genau untersucht werden. Auf welche philosophischen Traditionen sich die Vertreter dieser Schule bei dieser Annahme stützen und welche enormen Probleme darin liegen, werden wir in den Abschnitten 3.1 bzw. 3.2 beleuchten, wo es um den Reduktionismus bzw. die holistisch begründete Kritik daran geht.

2.2 Konnektionismus

Einen gänzlich anderen Ansatz zur Analyse und Nachbildung intelligenten Verhaltens verfolgt der Konnektionismus. Er untersucht nicht etwa die Art, wie verschiedene physikalische Systeme, nämlich das menschliche Gehirn und Computer, mit einer Repräsentation der Welt umgehen können, sondern beschäftigt sich direkt mit der physikalischen Struktur, auf der diese Fähigkeit beim Menschen beruht. Es soll ein Modell des menschlichen Gehirns geschaffen werden, das dann ähnliche Fähigkeiten wie dieses hervorbringt. Die praktische Realisierung solcher Modelle sind künstliche neuronale Netze. Dieser Ansatz erhebt nicht den Anspruch, das Verhalten des menschlichen Gehirns auf einer Gesamtebene verstehen und beschreiben zu können, es genügt ihm vielmehr zu verstehen, wie seine einzelnen Bestandteile, die Neuronen, funktionieren.⁸

Sämtliche Fähigkeiten eines künstlichen neuronalen Netzes beruhen auf der Vernetzung einer Vielzahl einzelner »Knoten«, die jeweils sehr einfach aufgebaut sind und deren Funktion der biologischer Neuronen ähnelt. Zwischen diesen Knoten wird ausschließlich rein numerische Information ohne weitere Interpretation ausgetauscht. Dabei hat jede einzelne Verbindung eine bestimmte Gewichtung, die bestimmt, wie groß der Einfluss genau dieser Verbindung ist. Einzelne Knoten empfangen von mehreren anderen Knoten deren Aktivität, d. h. Ausgangswert, und bestimmen daraus anhand der Gewichtungen und einer Aktivierungsfunktion ihre eigene Aktivität. Diese dient wiederum als Eingangssignal für andere Knoten. In der Regel werden bestimmte Knoten als Eingang und andere als Ausgang definiert.

Ziel eines solchen Netzes ist es dann, die korrekten Aktivitäten auf den Ausgangsknoten zu jeder Menge an Eingangswerten zu berechnen. Erreicht wird dies durch einen automatischen Trainingsvorgang, in dem die Gewichtungen der Verbindungen angepasst werden. Das Netz wird bei diesem Training mit typischen Daten konfrontiert und die Gewichtungen werden so angepasst, dass

8. vgl. [Dreyfus 1994](#), 217 f

die Fehlerrate, d. h. die Differenz zwischen den vom Netz ermittelten Ausgangswerten und den Sollwerten, sich verringert. Es sind eine Vielzahl von Algorithmen bekannt, die dieses Training bewirken.⁹

Wissen über einzelne Aspekte des Diskursbereiches wird in einem künstlichen neuronalen Netz in der Regel nicht durch einzelne Knoten oder Bereiche des Netzes repräsentiert. So wie es im menschlichen Gehirn keine »grandmother cells« gibt, die speziell für die Wiedererkennung der eigenen Großmutter zuständig wären, wird ein Konzept wie z. B. *Großmutter* auch in einem künstlichen neuronalen Netz durch verteilte Aktivitätsmuster repräsentiert. Daran beteiligte Knoten können auch in einer Vielzahl anderer Konzepte aktiv sein, die aus menschlicher Sicht keine einleuchtenden Gemeinsamkeiten besitzen. Das Netz schafft sich also gewissermaßen eigene Wissensrepräsentationen, die völlig von denen abweichen können, die wir im Alltag benutzen.¹⁰

Im Gegensatz zum symbolrepräsentierenden Ansatz entsteht im konnektionistischen Umgang mit einem Diskursbereich keine explizite Theorie desselben:

The theorist can thus begin with something short of a full-blooded theory of the domain. Instead of having formulated an elegant set of rules or axioms he or she may simply know how they wish to code the inputs and outputs, and have some hypothesis concerning the architecture of the system (number of hidden units, etc.) and the appropriate training instances and learning rules.¹¹

Dieser Ansatz geht viel weniger als der der Symbolrepräsentation auf klassische philosophische Traditionen zurück. Mit seinem theoretischen Hintergrund, dem Holismus, beschäftigen wir uns im Abschnitt 3.2.

3 Philosophischer Hintergrund

3.1 Reduktionistische Tradition

Zum philosophischen Hintergrund der reduktionistischen Tradition gehören die Mathematiktheoretiker Gottlob Frege, Bertrand Russell und Alfred North Whitehead, deren Ideen sich wiederum zurückführen lassen auf René Descartes und Gottfried Wilhelm Leibniz. Für Descartes basiert das Verstehen auf dem

9. vgl. [Clark und Lutz 1992](#), 1 ff

10. vgl. [Clark und Lutz 1992](#), 3 f

11. [Clark und Lutz 1992](#), 2

Umgang mit *Repräsentationen*, die aus einfachen Elementen zusammengesetzt sind. Für Thomas Hobbes ist der Umgang mit diesen Repräsentationen ein rein syntaktischer Vorgang, ist Urteilen nichts anderes als *Rechnen*.¹² Leibniz träumt von einem universellen Alphabet, das aus diesen einfachen Elementen des Denkens besteht, und dessen Anwendung es ermöglichen würde, alle kritischen Fragen einfach auszurechnen und so eindeutig zu lösen.¹³

Letztlich prägt diese Tradition die gesamte westliche Philosophie, angefangen mit Sokrates und Platon. Betrachtet werden die »Tatsachen«, nicht jedoch die »Welt als solche«. Dabei liegt die Vorstellung zugrunde, dass Verstehen und Wissen im Bilden einer *Theorie* besteht, die Beziehungen zwischen objektiven, kontextfreien Elementen beschreibt. Wir verstehen die Welt, indem wir sie *formalisieren*.¹⁴

Platon erkennt nur Wissen, das man explizit ausdrücken kann, überhaupt als Wissen an. Was nicht klar ausgedrückt werden kann, ist kein Wissen, sondern nur Glaube und Gefühl. Köche, die sich auf Geschmack und Intuition verlassen, oder Dichter, die ihrer Inspiration folgen, verfügen seiner Meinung nach über kein Wissen.¹⁵

Für Immanuel Kant sind alle Begriffe eigentlich Regeln. Der Begriff *Hund* entspricht etwa der Regel: Was vier Beine hat, bellt und mit dem Schwanz wedelt, ist ein Hund. Diese Auffassung, der Geist folge bestimmten Regeln, wird immer weiter ausgebaut. Edmund Husserl sieht Begriffe als Regelhierarchien, wobei jede Regel wiederum in Unterregeln zerlegbar ist – zur Hunde-Regel gehört dann etwa eine Unterregel zum Erkennen eines Schwanzes.¹⁶

Diese Tradition wird von dem frühen Wittgenstein fortgeführt. In seinem *Tractatus logico-philosophicus*¹⁷ versucht er, die Regeln zu erfassen, die dem Denken zugrunde liegen, ohne die es keine sinnvollen Gedanken gibt. Den Sinn seines Werks fasst er im Vorwort so zusammen:

Was sich überhaupt sagen läßt, läßt sich klar sagen; und wovon man nicht reden kann, darüber muß man schweigen.

Das Buch will also dem Denken eine Grenze ziehen, oder vielmehr – nicht dem Denken, sondern dem Ausdruck der Gedanken:

12. vgl. [Dreyfus 1994](#), 216 f

13. vgl. [Dreyfus 1992](#), 69 f, 211

14. vgl. [Dreyfus 1994](#), 222 f

15. vgl. [Dreyfus 1992](#), 67 f

16. vgl. [Dreyfus und Dreyfus 1987](#), 22

17. [Wittgenstein 1921](#)

Denn um dem Denken eine Grenze zu ziehen, müssten wir beide Seiten dieser Grenze denken können (wir müssten also denken können, was sich nicht denken lässt).

Die Grenze wird also nur in der Sprache gezogen werden können und was jenseits der Grenze liegt, wird einfach Unsinn sein.¹⁸

Schon in den allerersten Sätzen des Textes wird sein reduktionistisches Weltbild sichtbar, das die Welt als Zusammensetzung von Bestandteilen sieht, die er *Tatsachen* nennt und die einzeln und unabhängig voneinander erfasst werden können:

1 Die Welt ist alles, was der Fall ist.

1.1 Die Welt ist die Gesamtheit der Tatsachen, nicht der Dinge.

1.11 Die Welt ist durch die Tatsachen bestimmt und dadurch, daß es *alle* Tatsachen sind. [...]

1.2 Die Welt zerfällt in Tatsachen.

1.21 Eines kann der Fall sein oder nicht der Fall sein und alles übrige gleich bleiben.

2 Was der Fall ist, die Tatsache, ist das Bestehen von Sachverhalten.¹⁹

Der menschliche Geist kann diese Tatsachen erfassen, er kann sich Bilder der Tatsachen machen, die Modelle der Wirklichkeit darstellen und dieser in ihrer logischen Struktur entsprechen.²⁰ Diese Bilder nennt Wittgenstein *Gedanken*, und »[d]ie Gesamtheit der wahren Gedanken sind ein Bild der Welt«²¹.

Ausgedrückt werden die Gedanken in *Sätzen*:²² »Der Gedanke ist der sinnvolle Satz«²³. Sätze bestehen aus *Namen*, die für einen Gegenstand stehen: »Der Name bedeutet den Gegenstand. Der Gegenstand ist seine Bedeutung. (A ist dasselbe Zeichen wie A.)«²⁴

Sätze haben einen festen Sinn, sie lassen sich eindeutig analysieren: »Es gibt eine und nur eine vollständige Analyse des Satzes.«²⁵ Sie stellen Bilder, Modelle der Wirklichkeit dar: »Der Satz ist ein Bild der Wirklichkeit. Der Satz ist ein Modell der Wirklichkeit, so wie wir sie uns denken.«²⁶

18. Wittgenstein 1921, 9

19. Wittgenstein 1921, 11 – Hervorhebung hier und im Folgenden immer im Original

20. vgl. Wittgenstein 1921, 14–16

21. Wittgenstein 1921, 17

22. vgl. Wittgenstein 1921, 17

23. Wittgenstein 1921, 25

24. Wittgenstein 1921, 19

25. Wittgenstein 1921, 20

26. Wittgenstein 1921, 26

Der Sinn eines Satzes liegt in seiner »Übereinstimmung und Nichtübereinstimmung mit den Möglichkeiten des Bestehens und Nichtbestehens der Sachverhalte.«²⁷ Die einfachste Art von Sätzen, *Elementarsätze* genannt, behauptet das Bestehen eines Sachverhalts (d. h. behauptet eine Tatsache). Ein Elementarsatz ist genau dann wahr, wenn der entsprechende Sachverhalt besteht.²⁸ Damit lässt sich die Welt komplett beschreiben: »Die Angabe aller wahren Elementarsätze beschreibt die Welt vollständig.«²⁹

Nicht nur die Wissenschaft besteht aus solchen eindeutigen Sätzen – »[d]ie Gesamtheit der wahren Sätze ist die gesamte Naturwissenschaft (oder die Gesamtheit der Naturwissenschaften)«³⁰ –, sondern das gesamte Denken: »Alles was überhaupt gedacht werden kann, kann klar gedacht werden. Alles, was sich aussprechen lässt, lässt sich klar aussprechen.«³¹ Was gedacht werden kann, kann auch ausgesprochen werden. Deshalb bedeuten »[d]ie Grenzen meiner Sprache [...] die Grenzen meiner Welt.«³²

Alle Probleme und Fragen lassen sich entweder rational lösen, oder sie existieren gar nicht, sind bloße Scheinprobleme. Es gibt nichts Metaphysisches, nichts »Höheres«, nichts, was der Vernunft nicht zugänglich wäre: »Zu einer Antwort, die man nicht aussprechen kann, kann man auch die Frage nicht aussprechen. *Das Rätsel* gibt es nicht. Wenn sich eine Frage überhaupt stellen lässt, so *kann* sie auch beantwortet werden. [...] Wovon man nicht sprechen kann, darüber muß man schweigen.«³³

3.2 Holistische Kritik am Reduktionismus

Anders als die Symbolrepräsentation basiert die holistische Tradition weniger auf der Philosophie als auf dem jungen Gebiet der Neurophysiologie. Hier sind die Vorstellungen von Donald O. Hebb über die Interaktion von Neuronen im Gehirn wichtig.³⁴

Philosophisch sind Martin Heidegger, Maurice Merleau-Ponty und der späte Wittgenstein relevant. Sie betonen die Wichtigkeit der Alltagspraxis und betrachten eine Theorie der Alltagswelt als unmöglich. In seinen *Philosophischen*

27. Wittgenstein 1921, 38

28. vgl. Wittgenstein 1921, 38 f

29. Wittgenstein 1921, 39

30. Wittgenstein 1921, 32

31. Wittgenstein 1921, 33

32. Wittgenstein 1921, 67

33. Wittgenstein 1921, 84 f

34. vgl. Dreyfus 1994, 217

*Untersuchungen*³⁵ kritisiert und widerlegt Wittgenstein seinen eigenen, 30 Jahre früher entstandenen *Tractatus*, nachdem sich seine Suche nach atomaren Fakten als vergeblich erwiesen hat.

Heidegger setzt sich mit seinem Lehrer Edmund Husserl kritisch auseinander und kommt zu dem Schluss, dass die Alltagswelt nicht in kontextfreie Elemente zerlegt werden kann. Der Kontext der Welt ist nicht nur Teil unseres Denkens, sondern Teil unseres Seins, unserer Sozialisation. Unser Alltags-Verständnis ist, wie Heidegger und Wittgenstein hervorheben, nicht zwangsläufig *Know-that* (Wissen), das sich in Regeln und Fakten umsetzen lässt, sondern ganzheitliches *Know-how* (Fähigkeiten).³⁶

Der späte Wittgenstein der *Philosophischen Untersuchungen* hat seine Meinung radikal geändert: »Seit ich nämlich vor 16 Jahren mich wieder mit Philosophie zu beschäftigen anfang, mußte ich schwere Irrtümer in dem erkennen, was ich in jenem ersten Buche [dem *Tractatus*] niedergelegt hatte.«³⁷ Er untersucht das Programm des *Tractatus* auf seine Durchführbarkeit und entdeckt eine Reihe von groben Mängeln.

Er weist darauf hin, dass Sprache nicht nur als System der Kommunikation, der Verständigung gebraucht wird. Deshalb charakterisiert das, was die Sprache als Kommunikationssystem charakterisiert, nicht unbedingt die Sprache allgemein. Um Missverständnisse und unzulässige Verallgemeinerungen zu vermeiden, sollte man also angeben, wenn man nur einen Spezialfall, wie etwa die Sprache als Kommunikationssystem, meint.³⁸

3.2.1 Nicht alles lässt sich definieren

Wie kommt man zu Definitionen? Definitionen durch Worte sind nur möglich, wenn es schon eine gemeinsame Sprache oder wenigstens einige schon definierte Worte gibt, die in der neuen Definition verwendet werden können; also lassen sich nicht *alle* Wörter so definieren.

Eine Abhilfe können *Hinweisdefinitionen* sein: man zeigt auf die Sache, die man meint. Hinweisdefinitionen sind jedoch nicht eindeutig: Deutet man auf zwei Nüsse und sagt dazu »zwei«, kann dies auch als Name der Nüsse verstanden werden, und deutet man auf eine Person und sagt dazu ihren Namen, kann dies auch als Farbname (etwa der Hautfarbe), als Name des Volks, dem die Person angehört, oder als Zahlwort (»eins«) verstanden werden.

35. Wittgenstein 1953

36. vgl. Dreyfus 1994, 224–227

37. Wittgenstein 1953, 232

38. vgl. Wittgenstein 1953, 239

Bevor also Hinweisdefinitionen funktionieren können, muss also schon klar sein, welchen Platz in der Sprache das so zu Definierende einnehmen soll: ob es z. B. um eine Zahl, eine Farbe oder eine Länge geht. Diese Begriffe müssten also eingeführt werden, *bevor* man mit Hinweisdefinitionen beginnen kann – sie lassen sich nicht mit Hinweisen definieren. Versucht man sie aber mit Worten zu definieren, landet man wieder beim unendlichen Regress: jede Definition durch Worte muss aus Worten bestehen, die vorher definiert werden müssen. Aus diesem Regress kommen wir nur heraus, wenn wir die Idee aufgeben, *alle* Worte zu definieren.³⁹

Zudem sind Aussagen wie »jedes Wort der Sprache bezeichnet etwas« so allgemein, dass sie nichts aussagen.⁴⁰ Ein Ausdruck wie *Handgriff* kann ganz unterschiedliche Dinge bezeichnen; und die Dinge, die unter dem Begriff *Werkzeug* zusammengefasst werden, haben kaum etwas gemeinsam, als dass sie alle als *Werkzeug* bezeichnet werden. Definitionen sind häufig unmöglich oder künstlich, aufgesetzt:

Denke dir, jemand sagte: »*Alle* Werkzeuge dienen dazu, etwas zu modifizieren. So, der Hammer die Lage des Nagels, die Säge die Form des Bretts, etc.« – Und was modifiziert der Maßstab, der Leimtopf, die Nägel? – »Unser Wissen um die Länge eines Dings, die Temperatur des Leims, und die Festigkeit der Kiste.« – Wäre mit dieser Assimilation des Ausdrucks etwas gewonnen? –⁴¹

Ein anderes Beispiel: Was haben alle *Spiele* miteinander gemeinsam? Es gibt eine Reihe von Dingen, die viele Spiele gemeinsam haben, aber es gibt immer auch Ausnahmen. Die meisten Spiele sind unterhaltend, haben Gewinner und Verlierer, verlangen Glück und/oder Geschicklichkeit, aber keines dieser Charakteristika trifft auf alle Spiele zu – keines könnte als Definition des Begriffs *Spiel* dienen. Der Begriff des *Spiels* wird durch ein Netz von Ähnlichkeiten und Verwandtschaften gekennzeichnet; definieren lässt er sich nicht, ohne ihn zu verändern (bestimmte Spiele auszuschließen).⁴² Nicht durch eine Definition, sondern durch die Aufzählung von Beispielen, durch die Verwendung, kommen Wörter zu ihrer Bedeutung.⁴³

39. vgl. Wittgenstein 1953, 252 f

40. vgl. Wittgenstein 1953, 243 f

41. Wittgenstein 1953, 244

42. vgl. Wittgenstein 1953, 277 f

43. vgl. Wittgenstein 1953, 306

3.2.2 Nicht alles lässt sich klassifizieren

Die Unterscheidung zwischen Sätzen und Worten – essenziell für das logische System des *Tractatus*, wo Sätze Tatsachen und Worte Gegenständen entsprechen – ist fließend: Ist der Ausruf »Platte!« eines Bauarbeiters, der eine Platte braucht, ein Wort oder ein Satz? Fasst man ihn als Wort auf, hat er nicht die übliche Bedeutung des Wortes *Platte*, denn er ist ja zugleich ein Befehl. Man kann ihn als Satz versteht, als Verkürzung des Satzes »Bring mir eine Platte!«, den der Bauarbeiter eigentlich meint. Aber dieses *Meinen* muss ja nicht bedeuten, dass der Bauarbeiter den längeren Satz *denkt*, während er den kurzen ausspricht – die Bedeutung liegt also nicht unbedingt im Sprechen oder Denken.⁴⁴

Der Unterschied zwischen den Ausdrücken »Platte« (»Das ist eine Platte«) und »Platte!« (»Bring mir eine Platte!«) liegt in ihrer Rolle im *Sprachspiel*, d. h. in dem, was gemeint ist, im Kontext (oft wird der Unterschied zwar auch durch Tonfall und Gestik deutlich, das ist aber nicht immer so).⁴⁵

Lassen sich diese unterschiedlichen Verwendungen wenigstens ordentlich klassifizieren, indem alle Sätze z. B. in die drei Arten *Behauptung*, *Frage* und *Befehl* eingeteilt werden? Wittgenstein verneint dies: es gibt unzählige Arten von Sätzen, und neue Verwendungen können jederzeit entstehen:

[D]as Sprechen der Sprache [ist] ein Teil [...] einer Tätigkeit, oder einer Lebensform.

Führe dir die Mannigfaltigkeit der Sprachspiele⁴⁶ an diesen Beispielen, und anderen, vor Augen:

- Befehlen, und nach Befehlen handeln –
- Beschreiben eines Gegenstands nach dem Ansehen, oder nach Messungen –
- Herstellen eines Gegenstands nach einer Beschreibung (Zeichnung) –
- Berichten eines Hergangs –
- Über den Hergang Vermutungen anstellen – [...]
- Eine Geschichte erfinden; und lesen – [...]
- Einen Witz machen, erzählen –
- Ein angewandtes Rechenexempel lösen –

44. vgl. Wittgenstein 1953, 246

45. vgl. Wittgenstein 1953, 248

46. Wittgenstein verwendet den Ausdruck *Sprachspiele* für besondere Praktiken, die mit dem Gebrauch der Sprache zusammenhängen, z. B. das Nachsprechen der Worte der Lehrerin beim Sprachenlernen.

Aus einer Sprache in eine andere übersetzen –
Bitten, Danken, Fluchen, Grüßen, Beten.⁴⁷

3.2.3 Es gibt keine kontextfreien einfachen Elemente

Aus welchen einfachen Bestandteilen setzt sich die Realität zusammen? Wittgenstein beantwortet diese Frage mit einer Gegenfrage:

Was sind die einfachen Bestandteile eines Sessels? – Die Stücke Holz, aus denen er zusammengesetzt ist? Oder die Moleküle, oder die Atome? – »Einfach« heißt: nicht zusammengesetzt. Und da kommt es darauf an: in welchem Sinne »zusammengesetzt«? Es hat gar keinen Sinn von den »einfachen Bestandteilen des Sessels schlichtweg« zu reden.⁴⁸

Es braucht also immer einen Kontext, der für eine bestimmte Bedeutung sorgt – es gibt keine Bedeutung an sich.

Das Wort »zusammengesetzt« (und also das Wort »einfach«) wird von uns in einer Unzahl verschiedener, in verschiedener Weise miteinander verwandten, Arten benützt. ([...] Ist diese Strecke von 2 cm einfach, oder besteht sie aus zwei Teilstrecken von je 1 cm? Aber warum nicht aus einem Stück von 3 cm Länge und einem, in negativem Sinn angesetzten, Stück von 1 cm?)

Auf die *philosophische* Frage: »Ist das Gesichtsbild dieses Baumes zusammengesetzt, und welches sind seine Bestandteile?« ist die richtige Antwort: »Das kommt drauf an, was du unter »zusammengesetzt« verstehst.« (Und das ist natürlich keine Beantwortung, sondern eine Zurückweisung der Frage.)⁴⁹

Die Analyse eines Satzes bzw. Dinges und seine Zerlegung in seine Bestandteile bringt nicht unbedingt weiter. Da ein Besen aus Stiel und Bürste besteht, kann man die Aufforderung »Bring mir den Besen!« zerlegen in: »Bring mir den Besenstiel und die Bürste, die an ihm steckt!« *Dieser* Befehl dürfte jedoch eine eher irritierte Antwort hervorrufen: »Willst du den Besen haben? Und warum drückst du das so sonderbar aus?«⁵⁰

47. Wittgenstein 1953, 250

48. Wittgenstein 1953, 264

49. Wittgenstein 1953, 265

50. vgl. Wittgenstein 1953, 274

Unser Verständnis der Sprache wie der Welt basiert auf unseren Erfahrungen, es lässt sich nicht aus diesem Kontext lösen:

Ich sehe ein Bild: es stellt einen alten Mann dar, der auf einen Stock gestützt einen steilen Weg aufwärts geht. – Und wie das? Konnte es nicht auch so aussehen, wenn er in dieser Stellung die Straße herunterrutschte? Ein Marsbewohner würde das Bild vielleicht so beschreiben. Ich brauche nicht zu erklären, warum *wir* es nicht so beschreiben.⁵¹

3.2.4 Geist ohne Körper?

Ein anderer Punkt, den die Reduktionisten gern übersehen, ist das Verhältnis von Körper und Geist und die Rolle, die der Körper für Intelligenz und Denken spielen kann. Für die reduktionistischen Denker wie Platon und Descartes ist der Körper eher ein Hindernis für die geistigen Tätigkeiten, keineswegs eine unentbehrliche Vorbedingung. Aber gerade die »simplen« Tätigkeiten, wo die Sinne und körperliche Aktivitäten wichtig sind – etwa die Bewegung im Raum, das Erkennen von Mustern oder auch das Führen von Gesprächen –, bereiten der KI die größten Probleme, während bei rein geistigen Aktivitäten wie Schachspielen die Computer sehr viel eher mit Menschen konkurrieren können.⁵²

Merleau-Ponty argumentiert, dass beim Lernen einer Aktivität – etwa Autofahren, Tanzen, oder eine Fremdsprache (aus-)sprechen – ein Teil der Fähigkeit schließlich vom Körper verinnerlicht wird. Nur Anfänger müssen bewusst expliziten Regeln folgen; sobald man eine Fähigkeit sicher zu beherrschen beginnt, verlieren diese Regeln ihre Bedeutung – der Körper weiß von alleine, was zu tun ist. Die Rolle des Körpers wird am deutlichsten, wo Sinnestätigkeiten wie Schmecken und Berühren, aber auch Hören und Sehen nötig sind.⁵³

4 Geschichte

4.1 Ähnliche Entwicklung in den 50er und 60er Jahren

Die Entwicklung der beiden Paradigmen beginnt ähnlich: In den 50er Jahren verzeichnen beide bemerkenswerte frühe Erfolge, die zu großen Hoffnungen

51. Wittgenstein 1953, 309

52. vgl. Dreyfus 1992, 235–237

53. vgl. Dreyfus 1992, 248 f

Anlass geben. Allen Newell und Herbert Simon gelingt es 1956, mit Hilfe symbolischer Repräsentationen einfache Rätsel zu lösen und Theoreme zu beweisen. Frank Rosenblatt setzt die holistischen Ideen im Konzept des *Perceptron* um. 1956 kann ein Perceptron die Ähnlichkeit von Mustern erkennen. Die Vertreter beider Paradigmen sind der Meinung, den ersten Schritt zur Nachbildung menschlicher kognitiver Fähigkeiten auf Computern gemacht zu haben.

Anfang der 60er Jahre erscheinen somit beide Paradigmen vielversprechend; beide machen sich aber auch durch übertriebene Versprechungen angreifbar. Beide haben sich an »Spielzeugproblemen« bewährt, während unklar ist, ob und wieweit sie sich auf die Komplexität der wirklichen Welt übertragen lassen.⁵⁴

4.2 Sieg der Symbolrepräsentation in den 70er Jahren

Zehn Jahre später sieht das Bild völlig anders aus: das holistische Paradigma ist praktisch völlig verschwunden, während das reduktionistische Programm die KI-Forschung prägt.

Eine wesentliche Ursache für diesen Umbruch sieht Dreyfus in dem Wirken von Marvin Minsky und Seymour Papert, zwei bedeutenden Vertretern der Symbolrepräsentation. Sie kritisieren nicht nur die tatsächliche Umsetzung des konkurrierenden Paradigmas im Perceptron, sondern vor allem auch dessen philosophische Wurzeln, den Holismus. Dabei sind sie alles andere als zurückhaltend:

Our discussions will include some rather sharp criticisms of earlier work in this area. Perceptrons have been widely publicized as “pattern recognition” or “learning” machines and as such have been discussed in a large number of books, journal articles, and voluminous “reports.” Most of this writing [...] is without scientific value and we will not usually refer by name to the works we criticize.⁵⁵

Mit diesem »philosophische[n] Kreuzzug«⁵⁶ sind sie ungemein erfolgreich, die neuronalen Netze gelten allgemein als erledigt. Mit der inhaltlichen Kritik allein lässt sich dies nicht erklären, da beiden Paradigmen hoffnungsvolle Anfänge und problematische Entwicklungsperspektiven gemeinsam sind.

54. vgl. [Dreyfus 1994](#), 218 f

55. [Minsky und Papert 1969](#), 4

56. [Dreyfus 1994](#), 220 f

Vielmehr scheint der Erfolg von Minskys und Paperts Angriff auf den philosophischen Vorstellungen zu basieren, die sie mit den meisten KI-Forschern teilen, dem »quasi-religiöse[n] philosophische[n] Vorurteil gegen den Holismus«⁵⁷. Dazu trägt die Vorstellung bei, dass Denken und Mustererkennung zwei getrennte Bereiche darstellen und dass das Denken die wichtigere davon ist – wiederum ein Ergebnis der reduktionistischen philosophischen Tradition.⁵⁸

4.3 »Comeback« des Konnektionismus

Doch das Paradigma der Symbolrepräsentation kann sich seines Siegs nicht lange freuen, denn es kommt mit der Formalisierung des Alltagswissens nicht voran. Eine Reihe von Datenstrukturen wie Minskys *Frames* oder Roger Schranks *Scripts* werden erprobt, funktionieren aber nur in eng begrenzten Gebieten. Das ganze Programm der symbolrepräsentierenden KI kommt zum Stillstand:

Das Problem *Alltags-Verstand* hat die KI davon abgehalten, mit der Verwirklichung von Simons Voraussage 1965, daß »Maschinen innerhalb von zwanzig Jahren fähig sein werden, jede Arbeit des Menschen zu verrichten«, auch nur zu beginnen. In der Tat hat das Problem *Alltags-Verstand* jeden Fortschritt in der theoretischen KI der letzten beiden Dekaden blockiert.⁵⁹

Die Einwände, die Heidegger und der späte Wittgenstein aus theoretischen Erwägungen heraus gemacht haben⁶⁰, werden so durch die Praxis bestätigt: »Die rationalistische Tradition ist [...] einem empirischen Test unterzogen worden, und dieser schlug fehl.«⁶¹

Nachdem sich so das reduktionistische Forschungsprogramm mehr und mehr als degeneriert (im Sinne von Lakatos) erweist, wird das unterdrückte holistische Forschungsprogramm wieder attraktiver. Zwei Veröffentlichungen von John Hopfield zu neuronalen Netzen 1982 und 1984 beleben das wissenschaftliche Interesse neu. 1983 wird zum ersten Mal ein konnektionistisches Programm wieder öffentlich gefördert, und zwar von der bedeutenden US-amerikanischen *DARPA (Defense Advanced Research Projects Agency)*. 1986 stoßen die von David Rumelhart und James McClelland herausgegebenen Bücher über *Parallel Distributed Processing* auf enormes Interesse, 1987 veranstaltet die *IEEE (Institute*

57. Dreyfus 1994, 221

58. vgl. Dreyfus 1994, 219–222

59. Dreyfus 1994, 226

60. Vgl. Abschnitt 3.2 auf Seite 8.

61. Dreyfus 1994, 228

of *Electrical and Electronics Engineers*) die erste einer Reihe von Konferenzen über neuronale Netze. Der Konnektionismus wird zum dominanten Gebiet der KI.⁶²

Allerdings sieht sich der Konnektionismus seit seinem »Comeback« ähnlichen Problemen ausgesetzt wie die Symbolrepräsentation zuvor: die Skalierbarkeit ist ein Problem; während neuronale Netze im Kleinen gut funktionieren, macht die Erweiterung auf größere Wissensbereiche Probleme. Die optimistischen Antworten der Forscher – noch klappt es nicht, aber das werden wir bald haben – erinnert verdächtig an den Umgang mit Kritik der symbolrepräsentierenden KI in den 60er Jahren.⁶³

Wie weit das konnektionistische Paradigma es bringen wird, ist daher fraglich: »Das Modellieren neuronaler Netze sollte einfach die selbe Chance bekommen, wirklich zu scheitern, wie sie der symbolische Ansatz gehabt hat.«⁶⁴

5 Ausblick

Wissenschaftstheoretisch betrachtet, ist diese Auseinandersetzung konkurrierender Ideen ein interessantes Beispiel für den Kampf verschiedener Paradigmen um die Vorherrschaft auf einem wissenschaftlichen Gebiet. Hier zeigt sich, wie einerseits ein Paradigma mit primär ideellen Begründungen – dem »philosophischen Kreuzzug« – eine Zeitlang die Oberhand gewinnen kann; andererseits der Mangel an Erfolg in der Praxis die Suche nach (in diesem Falle, die Wiederentdeckung von) Alternativen hervorruft.

Dreyfus stellt fest: »Im Licht dieser Sackgasse erscheint die klassische, auf Symbolen beruhende KI mehr und mehr als ein perfektes Beispiel für das, was Imre Lakatos ein degenerierendes Forschungsprogramm genannt hat.«⁶⁵ Fragt man sich, wie berechtigt seine Vermutung ist, dass dem Konnektionismus ein ähnliches Schicksal blühen könnte, wird der Blick auf die philosophischen Hintergründe der beiden Paradigmen besonders interessant.

Die Symbolrepräsentation ist deutlich enger an die Tradition des Reduktionismus gebunden als der Konnektionismus an den Holismus. Dies liegt daran, dass der Holismus zwar eine fundierte Kritik an den reduktionistischen Vorstellungen liefert, aber kaum mit ausgearbeiteten Alternativmodellen aufwarten kann, die sich in der KI umsetzen ließen.

62. vgl. [Hecht-Nielsen 1990](#), 18 f

63. vgl. [Dreyfus 1994](#), 228 f

64. [Dreyfus 1994](#), 229

65. [Dreyfus 1994](#), 227

Insofern verdankt der Konnektionismus dem Holismus vor allem die Erkenntnis darüber, was *nicht* geht, welche Ansätze man besser vermeiden sollte. Der konkrete Alternativansatz des Konnektionismus ergibt sich dagegen nicht direkt aus der Philosophie, sondern aus der Neurophysiologie. Doch ob dieser wirklich zu einer Nachbildung des menschlichen Geistes führen kann, ist offen. Zumal auch der konnektionistische Ansatz im Allgemeinen nicht alle Hinweise des Holismus beachtet: Wenn zum »Geist« ein Körper gehört, sind die typischen neuronalen Netze genauso zum Scheitern verurteilt wie die Symbolrepräsentation.

In der Praxis haben beide Paradigmen ihre Berechtigung, wenn es nicht darum geht, den menschlichen Verstand komplett nachzubilden, sondern Lösungen für bestimmte Probleme zu finden. Es scheint wenig wahrscheinlich, dass in Zukunft eines der beiden vorgestellten Paradigmen endgültig über das andere triumphieren könnte. Auch wenn viele philosophische Annahmen der beiden Seiten miteinander unvereinbar scheinen, stellen sich in der praktischen Anwendung dieser Ideen meist ganz andere Fragen: In jedem einzelnen Anwendungsfall sollte geprüft werden, welcher Ansatz erfolgsversprechender ist.

Als ergiebig könnte es sich erweisen, die beiden Paradigmen zu verbinden. Möglicherweise arbeitet das Gehirn je nach Bedarf mal mit Problemlösungsstrategien, mal mit Mustererkennung, entsprechend dem Bild der linken und rechten Gehirnhälfte.⁶⁶ Dreyfus meint dazu jedoch: »Es ist noch zu früh, über eine Kombination der beiden Ansätze überhaupt nachzudenken, da bis jetzt keiner der beiden genug erreicht hat, um ein solides Fundament zu besitzen.«⁶⁷

66. vgl. [Dreyfus 1994](#), 229

67. [Dreyfus 1994](#), 229

Literatur

- Clark, Andy und Rudi Lutz (1992). Introduction. In *Connectionism in Context*, Hg. Andy Clark und Rudi Lutz, S. 1–15. Springer, London.
- Dreyfus, Hubert L. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. MIT Press, Cambridge, MA.
- Dreyfus, Hubert L. (1994). Den Geist konstruieren vs. Das Gehirn modellieren: Die KI kehrt zu einem Scheideweg zurück. In *Die maschinelle Kunst des Denkens*, Hg. Günther Cyranek und Wolfgang Coy, S. 215–230. Vieweg, Wiesbaden.
- Dreyfus, Hubert L. und Stuart E. Dreyfus (1987). *Künstliche Intelligenz. Von den Grenzen der Denkmaschine und dem Wert der Intuition*. Rowohlt, Reinbek bei Hamburg.
- Hecht-Nielsen, Robert (1990). *Neurocomputing*. Addison-Wesley, Reading, MA.
- Minsky, Marvin und Seymour Papert (1969). *Perceptrons*. MIT Press, Cambridge, MA.
- Newell, Allen (1983). Intellectual Issues in the History of Artificial Intelligence. In *The Study of Information: Interdisciplinary Messages*, Hg. F. Machlup und U. Mansfield, S. 196 ff. Wiley, New York.
- Newell, Allen und Herbert Simon (1981). Computer Science as Empirical Inquiry: Symbols and Search. In *Mind Design*, Hg. J. Haugeland, S. 41 ff. MIT Press, Cambridge, MA.
- Wittgenstein, Ludwig (1921). *Tractatus logico-philosophicus*. In *Werkausgabe*, Bd. 1, S. 7–85. Suhrkamp, Frankfurt am Main 1984.
- Wittgenstein, Ludwig (1953). *Philosophische Untersuchungen*. In *Werkausgabe*, Bd. 1, S. 225–580. Suhrkamp, Frankfurt am Main 1984.