

Das Problem der Intentionalität - Naturalistische Lösung oder meßtheoretische Auflösung?

Ansgar Beckermann

Zusammenfassung: Die Tatsache, daß viele mentale Zustände in dem Sinne intentional sind, daß sie einen repräsentationalen Inhalt haben, scheint auf den ersten Blick einen grundsätzlichen Unterschied zwischen mentalen und physischen Zuständen zu markieren. In den letzten Jahren ist allerdings (z.B. von Dretske, Fodor und Millikan) vielfach versucht worden zu zeigen, daß die Eigenschaft, einen repräsentationalen Inhalt zu haben, doch naturalisiert werden kann. Diese Versuche sind jedoch nicht nur alle mit jeweils typischen Schwierigkeiten konfrontiert, sie stehen auch so unvermittelt nebeneinander, daß eine begründete Entscheidung zwischen ihnen unmöglich scheint. In dem vorliegenden Aufsatz wird argumentiert, daß der Ansatz aller Versuche zur Naturalisierung von Intentionalität verfehlt ist, weil er von einer falschen Prämisse ausgeht, und daß es erfolgversprechender ist, das Problem der Intentionalität durch eine meßtheoretische Deutung intentionaler Termini aufzulösen. Dieser Deutung zufolge spielen die Inhalte, mit deren Hilfe wir intentionale Zustände charakterisieren, in etwa dieselbe Rolle wie die Zahlen, die wir bei der Charakterisierung physikalischer Größen verwenden.

Summary: The fact that many mental states are intentional in the sense that they have representational content, at first glance seems to mark a fundamental difference between mental and physical states. In recent years many authors (e.g., Dretske, Fodor, and Millikan) have nevertheless tried to show that the property of having representational content can yet be naturalized. These attempts, however, not only face a lot of difficulties. There is also the problem that it seems impossible to make a rational choice between them. In the present paper it is argued that all attempts to naturalize intentionality are misguided since they all rest on a false premis. The author claims that it is much more promising to dissolve the problem of intentionality by a measurement theoretic interpretation of intentional terms. According to this interpretation, the contents we use in characterizing intentional states are playing much the same role as the numbers we use in characterizing physical quantities.

I

((1)) Die philosophische Diskussion des Leib-Seele-Problems wird von einer immer wiederkehrenden Argumentationsfigur beherrscht. Wenn materialistisch oder naturalistisch gesinnte Philosophen behaupten, mentale Phänomene seien letzten Endes auch nur physische Phänomene, wenn auch vielleicht physische Phänomene einer besonders komplexen Art, entgegenen die Gegner des Naturalismus, dies sei unmöglich, da mentale Phänomene Eigenschaften hätten, die physische Phänomene aus prinzipiellen Gründen nicht haben könnten.

((2)) Auch in der gegenwärtigen Diskussion spielt ein solches Argument eine zentrale Rolle, ein Argument, das besonders auf Franz Brentano zurückgeht. Besonders prägnant hat Brentano dieses Argument im ersten Kapitel des zweiten Buches seiner "Psychologie vom empirischen Standpunkt" formuliert:

"Jedes psychische Phänomen ist durch das charakterisiert, was die Scholastiker des Mittelalters die intentionale (...) Inexistenz eines Gegenstandes genannt haben, und was wir, obwohl mit nicht ganz unzweideutigen Ausdrücken, die Beziehung auf einen Inhalt, die Richtung auf ein Objekt (...), oder die immanente Gegenständlichkeit nennen würden. In der Vorstellung ist etwas vorgestellt, in dem Urteile etwas

anerkannt oder verworfen, in der Liebe geliebt, in dem Hasse gehaßt, in dem Begehren begehrt usw. Diese intentionale Inexistenz ist den psychischen Phänomenen ausschließlich eigentümlich. Kein physisches Phänomen zeigt etwas Ähnliches. Und somit können wir die psychischen Phänomene definieren, indem wir sagen, sie seien solche Phänomene, welche intentional einen Gegenstand in sich enthalten." (Brentano 1924, 124f.)

((3)) Ein kritischer Punkt dieser Argumentation liegt sicher in ihrer uneingeschränkten Allgemeinheit. Denn es ist zweifelhaft, daß tatsächlich alle mentalen Phänomene im Sinne Brentanos intentional sind.¹ Wenn man sich unwohl fühlt, wenn man nervös, erfreut oder deprimiert ist, dann haben diese Zustände zwar in der Regel einen Grund oder Anlaß, aber sie sind nicht auf diesen Grund gerichtet, sie haben ihn nicht als intentionalen Gegenstand. Doch dies ist kein besonders folgenreicher Punkt. Auch wenn nicht alle mentalen Zustände, sondern eben nur die intentionalen Zustände wie Überzeugungen, Wünsche, Hoffnungen, Befürchtungen, usw. im Brentanoschen Sinne intentional sind, bleibt seine These ja brisant genug.

((4)) Es gibt allerdings noch einen zweiten unklaren Punkt. Wenn man von den zitierten Formulierungen ausgeht, dann sieht es zunächst so aus, als sei ein Phänomen intentional, wenn es in bestimmter Weise auf einen Gegenstand gerichtet

ist bzw. diesen Gegenstand in einer bestimmten nicht räumlichen Weise als Objekt in sich enthält. Doch in der neueren Diskussion ist man sich weitgehend einig darüber, daß die Intentionalität von Wünschen, Überzeugungen usw. nicht darin besteht, daß sie eine schwer begriffliche Relation zu (möglicherweise nicht einmal existierenden) Gegenständen beinhalten, sondern darin, daß sie Wahrheits- bzw. Erfüllungsbedingungen haben.² Wenn man dies akzeptiert, ist es zunächst naheliegend, Brentanos These folgendermaßen zu präzisieren.

BRENTANOS THESE (Version 1):

Intentionale Zustände haben einen repräsentationalen Inhalt, d.h. sie haben Wahrheits- oder Erfüllungsbedingungen. Physische Zustände können in diesem Sinne keinen repräsentationalen Inhalt haben. Also gibt es einen prinzipiellen Unterschied zwischen physischen und intentionalen Zuständen.

((5)) In dieser Form wäre Brentanos These jedoch sehr unplausibel. Denn was spricht dafür, daß physische Zustände keinen repräsentationalen Inhalt, also keine Wahrheits- bzw. Erfüllungsbedingungen haben können? Immerhin wissen wir doch, daß sprachliche Äußerungen einen repräsentationalen Inhalt haben. Und sprachliche Äußerungen sind, ob sie nun in Mustern von Schallwellen oder als Inskriptionen auf Papier realisiert sind, zunächst einmal physische Phänomene. Generell ist die These, daß physische Phänomene keine Wahrheits- oder Erfüllungsbedingungen haben können, also sicher nicht haltbar.

((6)) Mit einer solchen Argumentation würde man jedoch am Kern des Problems vorbeigehen. Denn entscheidend ist in diesem Zusammenhang nicht die Frage, ob physische Phänomene einen repräsentationalen Inhalt haben können, sondern die Frage, ob die Eigenschaft, einen repräsentationalen Inhalt bzw. Wahrheits- oder Erfüllungsbedingungen zu haben, eine Eigenschaft ist, die im Rahmen eines naturalistischen Weltbildes expliziert werden kann. Dies ist jedenfalls auch für J. Fodor die entscheidende Frage, der in seiner Problemanalyse mit Brentano völlig übereinstimmt. Wie Brentano geht Fodor davon aus: 1. daß es intentionale Zustände gibt, d.h. daß diese Zustände real sind; 2. daß intentionale Zustände dadurch charakterisiert sind, daß sie einen intentionalen bzw. repräsentationalen Inhalt haben; und 3. daß sich intentionale Zustände in diesem Merkmal von fast allen anderen Dingen in dieser Welt unterscheiden. Nur Symbole haben nach Fodor ebenfalls die Eigenschaft, etwas zu repräsentieren. Die Frage ist daher, wie es dazu kommt, daß sich intentionale Zustände und Symbole in dieser Hinsicht von allen anderen Dingen in der Welt unterscheiden?

“To a first approximation, symbols and mental states both have representational content. And nothing else does that belongs to the causal order: not rocks, or worms or trees or spiral nebulae ... It appears increasingly that the main joint business of the philosophy of language and the philosophy of mind is the problem of representation itself: the metaphysical question of the place of meaning in the world order. How can anything manage to be about anything; and why is it that only thoughts and symbols succeed?” (Fodor 1987, xi)

Brentanos Position im Hinblick auf diese Frage ist offenbar durch die Auffassung gekennzeichnet, daß eine Bedeutung oder einen repräsentationalen Inhalt zu haben etwas ist, was über den Bereich des Physischen hinausgeht. Und seine These läßt sich daher meiner Meinung nach am besten so formulieren:

BRENTANOS THESE (Version 2):

Intentionale Zustände haben einen repräsentationalen Inhalt, d.h. sie haben Wahrheits- oder Erfüllungsbedingungen. Die Eigenschaft, einen repräsentationalen Inhalt zu haben, ist jedoch nicht naturalisierbar, sie kann nicht im Rahmen einer naturalistischen Weltanschauung expliziert werden. Deshalb handelt es sich bei intentionalen Zuständen um Zustände, die über den Bereich des Physischen hinausgehen.

((7)) Es ist klar, daß in dieser These für jeden naturalistisch gesinnten Philosophen eine besondere Herausforderung liegt. Zumindest dann, wenn er nicht versucht ist, die ganze Intentionalitätsproblematik durch einen Sprung in den Eliminativismus zu erledigen. Denn Intentionalität im Sinne dieser These gehört sicher nicht zu den letzten und irreduziblen Grundeigenschaften des Universums wie etwa Masse und Ladung. Also kann man als naturalistisch gesinnter Philosoph nur Realist im Hinblick auf intentionale Zustände sein, wenn man zugleich Reduktionist ist.

“If the semantic and the intentional are real properties of things, it must be in virtue of their identity with (or maybe their supervenience on?) properties that are themselves neither intentional nor semantic. If aboutness is real, it must be really something else.” (Fodor 1987, 97)

Das Problem, das sich aus der Brentanoschen These ergibt, lautet daher:

DAS PROBLEM DER INTENTIONALITÄT:

Kann die Eigenschaft, einen repräsentationalen Inhalt zu haben, nicht doch naturalisiert werden? D.h., ist es - im Gegensatz zu Brentanos Überzeugung - nicht doch möglich, in nichtsemantischem und nichtintentionalem Vokabular notwendige und hinreichende Bedingungen dafür zu formulieren, daß ein (physischer) Zustand den repräsentationalen Inhalt *p* hat?

II

((8)) Viele analytische Philosophen haben in den letzten Jahren versucht, dieses Problem zu lösen. Es war fast so etwas wie das beherrschende Thema der Diskussion. Dennoch ist es bisher jedenfalls nicht gelungen, einen Lösungsvorschlag zu präsentieren, der auch nur weitgehende Zustimmung gefunden hätte. Stattdessen gibt es eine Vielzahl von zum Teil sehr divergierenden Ansätzen. Es gibt den informationstheoretisch inspirierten Ansatz Dretskes, Fodors Theorie, in der die Idee asymmetrisch voneinander abhängender Kausalbeziehungen die entscheidende Rolle spielt, und

schließlich gibt es den besonders von Millikan und Papineau vertretenen funktional-teleologischen Ansatz.³

((9)) Die Grundlage der Dretskeschen Überlegungen zur Bestimmung des Inhalts intentionaler Zustände bilden seine Definitionen der semantischen Information und des semantischen Inhalts. Die Grundidee seiner Definition der semantischen Information läßt sich in dem Satz zusammenfassen: Ein Signal s trägt genau dann die Information, daß e , wenn man aus s lernen kann, daß e . Oder mit anderen Worten: ein Signal s trägt genau dann die Information, daß e , wenn man aus dem Auftreten von s schließen kann, daß e der Fall ist. Bei dem Versuch, dieser Grundidee eine befriedigende Form zu geben, benutzt Dretske zwar die statistische Terminologie der mathematischen Informationstheorie. Aus vielen Bemerkungen geht aber hervor, daß er der Auffassung ist, daß ein Signal s die Information e nur dann tragen kann, wenn der Fall, daß s auftritt, ohne daß e auftritt, vollständig ausgeschlossen ist. Es scheint mir daher angemessener, den Begriff der semantischen Information folgendermaßen zu definieren:

(SI_{nf}) Ein Signal s trägt genau dann die Information e , wenn das Auftreten von e eine notwendige Bedingung für das Auftreten von s ist, d.h. wenn der Satz "Wenn s , dann e " ein wahrer nomologisch gestützter Konditionalsatz ist.

((10)) Der Informationsgehalt eines Signals s ist niemals eindeutig. Denn natürlich gibt es immer verschiedene notwendige Bedingungen für das Auftreten von s . Der semantische Inhalt eines Signals oder eines repräsentationalen Zustandes sollte aber eindeutig sein, und deshalb schlägt Dretske vor, den semantischen Inhalt eines Signals s als die spezifischste Information zu definieren, die dieser Zustand trägt, d.h. als die Information, die s in vollständig digitalisierter Form trägt.

(SI) Ein Signal s hat genau dann den semantischen Inhalt e , wenn s die Information e in vollständig digitalisierter Form trägt.⁴

((11)) Den Begriff der vollständig digitalisierten Information entwickelt Dretske in zwei Schritten. Seine Grundidee ist dabei, daß ein Signal s eine Information e in vollständig digitalisierter Form trägt, wenn diese Information in keiner anderen Information eingebettet ist. Zunächst führt Dretske deshalb den Begriff der Einbettung ein.

(E) Wenn ein Signal s die Information e trägt, dann ist in e die Information e' genau dann eingebettet, wenn e die Information e' trägt.⁵

Mit Hilfe dieses Begriffs kann der Begriff der vollständig digitalisierten Information dann folgendermaßen definiert werden:

(VDI) Ein Signal s trägt die Information e genau dann in vollständig digitalisierter Form, wenn s die Information e trägt und wenn s keine Information e' trägt, von der gilt: die Information e ist

in der Information e' eingebettet, aber nicht umgekehrt.⁶

((12)) Wenn man sich diese Definition noch einmal genauer ansieht, wird allerdings klar, daß sie nichts anderes besagt als die folgende viel durchsichtigere Formulierung:⁷

(VDI') Ein Signal s trägt die Information e genau dann in vollständig digitalisierter Form, wenn das Auftreten von e eine notwendige und hinreichende Bedingung für das Auftreten von s ist.

Und da das so ist, läßt sich auch die Definition des Begriffs des semantischen Inhalts vereinfachen.

(SI') Ein Signal s hat genau dann den semantischen Inhalt e , wenn e eine notwendige und hinreichende Bedingung für s ist.

((13)) Auch der Begriff des semantischen Inhalts ist jedoch noch nicht mit dem identisch, was man normalerweise den repräsentationalen Inhalt intentionaler Zustände nennt.⁸ Denn wenn s den semantischen Inhalt e hat, dann muß der Definition (SI') zufolge, wenn s der Fall ist, auch e der Fall sein. Der Fall einer Fehlrepräsentation, d.h. der Fall, in dem ein Zustand mit dem semantischen Inhalt e auftritt, obwohl e nicht der Fall ist, ist also schon aufgrund der angeführten Definitionen ausgeschlossen. Aber natürlich kann man etwas Falsches glauben oder von etwas Falschem überzeugt sein. In (1981) versucht Dretske eine Lösung für dieses Problem zu finden, indem er auf die Entstehung der Mechanismen Bezug nimmt, die dafür sorgen, daß ein System S auf bestimmte Ereignisse in der Außenwelt und nur auf diese mit einem bestimmten Zustand Z reagiert. Dabei denkt Dretske besonders an Fälle, in denen solche Mechanismen durch Prozesse entstehen, die als Lernprozesse aufgefaßt werden können.

((14)) Nehmen wir als Beispiel ein System S , das zu einem bestimmten Zeitpunkt t_0 noch nicht über einen Zustand verfügt, der verläßlich auf Rotkehlchen und nur auf Rotkehlchen reagiert. In einer Zeitspanne nach t_0 werden S eine Reihe von Vögeln gezeigt und zugleich wird S durch Belohnung oder eine andere Art von Training dazu gebracht, auf Rotkehlchen und nur auf Rotkehlchen mit dem Zustand Z zu reagieren. Wenn das System am Ende der Trainingsperiode keinen Fehler mehr macht, d.h. wenn es am Ende der Trainingsperiode den Zustand Z genau dann annimmt, wenn das gezeigte Beispiel ein Rotkehlchen ist, dann hat Dretske zufolge Z den repräsentationalen Inhalt, hier Rotkehlchen. Und das gilt auch, wenn das System zu einem späteren Zeitpunkt doch einmal Fehler machen sollte, also z.B. mit dem Zustand Z auf eine junge Nachtigall reagiert.

((15)) Die Probleme dieser Auffassung sind oft benannt worden.⁹ Soweit ich sehen kann, liegt ihr Hauptproblem in der Frage, wann der Lernprozeß beendet ist und welcher repräsentationale Inhalt am Ende erworben wurde. Nach Dretske ist der Lernprozeß beendet, wenn das System keinen Fehler mehr macht, d.h. auf alle vorgelegten Beispiele

korrekt reagiert. Später ist es jedoch durchaus möglich, daß es wieder Fehler macht. Aber wenn das System später tatsächlich einen Fehler macht, dann legt dies zumindest den Verdacht nahe, daß es auch schon während der Lernperiode auf ein entsprechendes Beispiel falsch reagiert hätte. Ein solches Beispiel wurde ihm nur de facto nicht vorgelegt. Wenn man die Bedingung für das Ende des Lernvorgangs jedoch so formuliert, daß das System nicht nur auf alle tatsächlich vorgelegten, sondern auf alle möglichen Beispiele korrekt reagiert, dann gibt es nur zwei Möglichkeiten: entweder ein solcher Endzustand wird tatsächlich nie erreicht oder das System bildet am Ende einen perfekten Indikator aus. In diesem Fall sind spätere Fehler aber völlig ausgeschlossen. Mit diesem Problem hängt ein zweites eng zusammen. Wenn der Lernvorgang abgeschlossen ist, falls das System auf alle vorgelegten Beispiele von Rotkehlchen und nur auf diese mit der Ausbildung von Z reagiert, welchen repräsentationalen Gehalt hat Z dann eigentlich? Jede endliche Menge von Beispielen instantiiert eine ganze Reihe verschiedener Eigenschaften. Wenn z.B. alle vorgelegten Beispiele von Rotkehlchen zwischen 14 und 15 cm groß waren, die Beispiele von nicht-Rotkehlchen dagegen entweder größer oder kleiner, hat Z dann den repräsentationalen Gehalt Rotkehlchen oder den repräsentationalen Gehalt 14-15 cm großer Vogel?¹⁰

((16)) J. Fodor geht bei seinem Versuch, eine naturalistische Analyse für den Inhalt repräsentationaler Zustände zu geben, zunächst von einer Definition aus, die er CCT (Crude Causal Theory) nennt:

(CCT) Ein Zustand Z in einem System S hat genau dann den repräsentationalen Inhalt e , wenn (1) alle e Z in S verursachen und (2) Z in S nur durch e verursacht wird.

Aber diese Definition ist natürlich unzureichend. Denn einerseits löst sie das Problem der Fehlrepräsentation nicht, und andererseits kann man in der Regel sicher nicht davon ausgehen, daß Zustände mit dem repräsentationalen Inhalt e von allen e und nur von e verursacht werden.

((17)) Wie kann man CCT also so verbessern, daß sich eine angemessenere Analyse ergibt? Fodors Antwort auf diese Frage hat zwei Teile, die sich jeweils auf den "alle"- und auf den "nur"-Teil von CCT beziehen. Sicher werden nicht alle e den Zustand Z verursachen. Wenn wir wieder an das Rotkehlchenbeispiel denken, dann wird Z z.B. nicht auftreten, wenn die Beleuchtungsverhältnisse ungünstig sind oder wenn sich das Rotkehlchen in einem anderen Raum befindet. Aber dieses Problem ist Fodor zufolge lösbar. Denn seiner Meinung nach ist es möglich, die optimalen Bedingungen M , unter denen alle e Z verursachen, ohne Benutzung intentionaler oder semantischer Termini allein in Begriffen der Psychophysik anzugeben (auch wenn wir dies de facto im Augenblick noch nicht können). CCT läßt sich daher folgendermaßen verbessern:

(CCT') Ein Zustand Z in einem System S hat genau dann den repräsentationalen Inhalt e , wenn (1)

unter optimalen Bedingungen alle e Z in S verursachen und (2) Z in S nur durch e verursacht wird.

Diese Verbesserung ist zwar auch, was den Teil (1) des Definiens angeht, noch nicht ganz ausreichend, da sie den Fall, daß e für S nicht beobachtbar ist, nicht hinreichend berücksichtigt. Aber ich will auf dieses Problem hier nicht weiter eingehen, weil die Grundidee Fodors auch durch diese Formulierung recht gut wiedergegeben wird.

((18)) Was den "nur"-Teil von CCT angeht, hat Fodor folgende Idee. Sicher wird es immer mal wieder vorkommen, daß der Zustand Z nicht nur durch ein Rotkehlchen, sondern z.B. durch eine Nachtigall verursacht wird. Aber die Tatsache, daß Z manchmal durch Nachtigallen verursacht wird, ist in asymmetrischer Weise abhängig von der Tatsache, daß Z in der Regel durch Rotkehlchen verursacht wird. Denn Nachtigallen verursachen Z nur, weil sie unter bestimmten Bedingungen wie Rotkehlchen aussehen. Falls Rotkehlchen Z nicht verursachen würden, würden daher auch Nachtigallen Z nicht verursachen. Wenn ich lernen würde, Nachtigallen auch unter diesen Bedingungen von Rotkehlchen zu unterscheiden, würden Rotkehlchen immer noch Z verursachen, Nachtigallen aber nicht mehr. Auch wenn Z manchmal durch eine Nachtigall verursacht wird, bedeutet das also nicht, daß Z nicht den Inhalt, hier Rotkehlchen, sondern den Inhalt, hier Rotkehlchen oder Nachtigall, hat, sofern nur die Tatsache, daß Z manchmal durch Nachtigallen verursacht wird, asymmetrisch von der Tatsache abhängt, daß Z in der Regel durch Rotkehlchen verursacht wird. Auf diese Weise kommt Fodor schließlich zur Formulierung seiner SLCCCT ("Slightly less crude causal theory of content"):

(SLCCCT) Ein Zustand Z in einem System S hat genau dann den repräsentationalen Inhalt e , wenn (1) unter optimalen Bedingungen alle e Z in S verursachen und wenn (2), falls in einigen Fällen e' Z in S verursachen, die Kausalbeziehung zwischen e' und Z asymmetrisch von der Kausalbeziehung zwischen e und Z abhängt.¹¹

((19)) Die Plausibilität dieses Vorschlags hängt offenbar entscheidend davon ab, ob die Idee einer asymmetrischen Abhängigkeit kausaler Beziehungen konsistent explizierbar ist. Und daran sind zumindest Zweifel möglich. Cummins¹² hat z.B. argumentiert, daß eine solche asymmetrische Abhängigkeit nur gegeben wäre, wenn man zeigen könnte, daß von den folgenden beiden kontrafaktischen Konditionalsätzen der erste wahr und der zweite falsch ist.

(1) Wenn Rotkehlchen Z nicht verursachen würden, würden auch Nachtigallen Z nicht verursachen.

(2) Wenn Nachtigallen Z nicht verursachen würden, würden auch Rotkehlchen Z nicht verursachen.

Aber wie will man das zeigen? Fodor scheint von folgendem

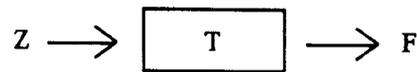
Bild auszugehen. Rotkehlchen verursachen den Zustand Z , weil sie wie Rotkehlchen aussehen und weil ein solches Aussehen Z verursacht. Nachtigallen schleichen sich sozusagen in diese kausale Beziehung ein, da sie in einigen Fällen ebenfalls wie Rotkehlchen aussehen und auf diese Weise Z verursachen. Wie könnte man nun zeigen, daß (2) falsch ist? Prinzipiell scheint es nur zwei Möglichkeiten zu geben, die Beziehung zwischen Nachtigallen und dem Zustand Z aufzubrechen. Entweder (i) Z wird nicht mehr durch rotkehlchenhaftes Aussehen verursacht; aber dann würden auch Rotkehlchen Z nicht mehr verursachen. Oder (ii) Nachtigallen haben nie mehr ein rotkehlchenhaftes Aussehen. Vielleicht lernen wir etwas über den Unterschied zwischen Rotkehlchen und Nachtigallen, was dazu führt, daß Nachtigallen für uns in keiner Situation mehr wie Rotkehlchen aussehen. In diesem Falle würden Rotkehlchen immer noch Z verursachen. (2) wäre also, wie erhofft, falsch.

Wie steht es nun mit (1)? Wieder scheint es nur zwei Möglichkeiten zu geben, die kausale Beziehung zwischen Rotkehlchen und dem Zustand Z aufzubrechen. Die erste Möglichkeit ist wieder (i), daß Z nicht mehr durch rotkehlchenhaftes Aussehen verursacht wird. In diesem Falle würden auch Nachtigallen Z nicht mehr verursachen, weil wir ja angenommen hatten, daß Nachtigallen, wenn sie Z verursachen, dies gerade auf dem Wege über ein rotkehlchenhaftes Aussehen tun. In diesem Fall wäre also (1) wahr; aber natürlich auch (2) - im Gegensatz zu dem, was wir angestrebt hatten. Also bleibt nur die zweite Möglichkeit: (iii) Selbst Rotkehlchen sehen nicht mehr wie Rotkehlchen aus. Vielleicht sterben sie aus oder verändern ihre äußere Gestalt in erheblichem Maße. Aber das würde an der Beziehung zwischen Nachtigallen und dem Zustand Z nichts ändern. In diesem Falle wäre (1) also falsch.

Um zu dem gewünschten Ergebnis zu kommen, daß (1) wahr und (2) falsch ist, müssen wir also verschiedene Standards verwenden. Wir brauchen (ii), um (2) falsch, und (i), um (1) wahr zu machen. Aber die möglichen Welten, in denen (ii) der Fall ist, sind nicht dieselben, in denen (i) der Fall ist. Mit anderen Worten, es gibt keine gemeinsame Interpretation, die (1) wahr und (2) falsch macht.¹³

((20)) Ein dritter Ansatz zur Festlegung des Inhaltes repräsentationaler Zustände, der besonders von Ruth Millikan und David Papineau¹⁴ entwickelt wurde, geht von teleologischen Überlegungen aus, Überlegungen zur biologischen Funktion dieser Zustände. Millikan zufolge kann man nur dann davon sprechen, daß ein Zustand Z in einem System S überhaupt ein repräsentationaler Zustand mit einem bestimmten Inhalt e ist, wenn dieser Zustand für das System die Funktion eines Zeichens hat, d.h. genauer: wenn es in S Teilsysteme gibt, die ihre biologische Funktion nur erfüllen können, wenn Z ein Indikator für e ist. Natürliche Zeichen - d.h. Zeichen im Sinne der Dretscheschen Definition (SInf) - sind daher per se noch keine Repräsentationen oder repräsentationale Zustände.

((21)) Die Grundidee des Millikanschen Ansatzes läßt sich mit folgendem Bild veranschaulichen:



Das Gesamtsystem S enthält ein Teilsystem T , dessen Verhalten von Z abhängt und dessen biologische Funktion F ist. Damit ist folgendes gemeint. F ist eine Wirkung von T , die für das Gesamtsystem positiv ist, und T ist genau deshalb als Teilsystem von S entstanden (oder wird genau deshalb als Teilsystem von S erhalten), weil es die Wirkung F hat. Konkret kann das z. B. heißen, daß T in gefährlichen Situationen eine Fluchtreaktion auslöst und daß T im Laufe der Evolution genau aus diesem Grund in S entstanden ist, weil Systeme mit dieser Fluchtreaktion überlebensfähiger waren als andere. Nun kann T seine Funktion, Fluchtreaktionen in die richtige Richtung auszulösen, aber nur erfüllen, wenn es darüber "informiert" ist, wann eine Gefahr vorliegt und aus welcher Richtung sie kommt. Diese Information zu liefern, ist die Aufgabe oder Funktion von Z . Aus dieser Überlegung ergibt sich die folgende Formulierung für den repräsentationalen Inhalt von Z :

(FI) Z hat im System S genau dann den repräsentationalen Inhalt e , wenn es ein Teilsystem T von S gibt, dessen Verhalten von Z abhängt, und wenn e im Normalfall ein Hauptfaktor dafür ist, daß T seine biologische Funktion erfüllt.¹⁵

((22)) Nehmen wir an, es sei T 's biologische Funktion, Fluchtbewegungen in jeweils die Richtung auszulösen, die der Richtung, aus der sich eine Gefahr nähert, entgegengesetzt ist. Nehmen wir weiter an, daß T eine Fluchtbewegung nach links auslöst, wenn sich S im Zustand Z befindet, und eine Fluchtbewegung nach rechts, wenn sich S im Zustand Z' befindet. Dann kann T seine biologische Funktion offenbar nur erfüllen, wenn Z im Normalfall dann auftritt, wenn sich eine Gefahr von rechts nähert, und Z' im Normalfall dann, wenn sich eine Gefahr von links nähert. Insofern ist es die Funktion von Z anzuzeigen, daß sich eine Gefahr von rechts nähert, und entsprechend die Funktion von Z' anzuzeigen, daß sich eine Gefahr von links nähert.

((23)) Dies ist in groben Zügen die Grundidee der Formulierung (FI). Aber warum ist in dieser Formulierung vom Normalfall und von einem Hauptfaktor die Rede? Viele biologische Systeme zeigen häufig auch dann Fluchtverhalten, wenn sie gar nicht wirklich in Gefahr sind. Sie verhalten sich nach dem Motto "Lieber einmal zu viel als einmal zu wenig weglaufen". Dies liegt im allgemeinen daran, daß die Indikatoren für Gefahren überempfindlich reagieren. Sie reagieren, könnte man sagen, nicht nur auf wirkliche Gefahren, sondern auch schon auf mögliche Anzeichen von Gefahren. Aus diesem Grund wird sich in unserem System S der Zustand Z möglicherweise schon dann ausbilden, wenn sich dem System von rechts ein dunkler Schatten nähert, auch wenn solch ein Schatten nur in einem Bruchteil aller Fälle von einer wirklichen Gefahr hervorgerufen wird. Dies bedeutet Millikan zufolge jedoch nicht, daß Z deshalb statt des repräsentationalen Inhalts, Gefahr von rechts, den Inhalt, Schatten von

rechts, hätte. Denn die biologische Funktion von *T* ist es, bei Gefahr ein Fluchtverhalten in die richtige Richtung auszulösen; weil es diese Wirkung hat, ist es im Laufe der Evolution in *S* entstanden und nicht, weil es bei jedem sich annähernden Schatten ein solches Fluchtverhalten auslöst. Der Normalfall ist also der, in dem sich wirklich eine Gefahr nähert; denn dies ist der Fall, für den *T* gedacht ist. Der Ausdruck "Normalfall" ist also normativ und nicht statistisch zu verstehen. (Möglicherweise gehen die meisten Fluchtbewegungen, die *T* auslöst, auf einen Fehlalarm zurück.) In Millikans eigenen Worten:

"My term 'normal' should be read normatively, historically, and relative to specific function. In the first instance, 'normal' applies to explanations. A 'normal explanation' explains the performance of a particular function, telling us how it was (typically) historically performed on those (perhaps rare) occasions when it was properly performed." (1989, 284)

((24)) Warum Hauptfaktor?¹⁶ Daß *T* im Normalfall seine biologische Funktion erfüllen kann, hängt in der Regel nicht nur von einem Faktor ab, sondern von einer Vielzahl von Faktoren. Wenn sich eine Gefahr von rechts nähert, ist eine Fluchtreaktion in die entgegengesetzte Richtung sicher biologisch vorteilhaft; aber das gilt natürlich immer. Denn eine Fluchtreaktion in die entgegengesetzte Richtung ist sicher nicht biologisch vorteilhaft, wenn sich in dieser Richtung ein großer Graben befindet, in den das System *S* fallen könnte, oder wenn in dieser Richtung noch größere Gefahren lauern. Dennoch ist die Funktion von *Z* nur anzuzeigen, daß sich eine Gefahr von rechts nähert, und nicht auch noch, daß sich links kein Graben befindet und keine andere Gefahr, usw. Der Grund dafür ist, daß zwischen der Bedingung, daß sich von rechts eine Gefahr nähert, und diesen anderen Bedingungen eine Asymmetrie besteht, die sich darin ausdrückt, daß es keinen Sinn macht, nach diesen anderen Umständen zu fragen, wenn jene Bedingung nicht erfüllt ist. Cummins erläutert diese Asymmetrie mit einem schönen Beispiel. Man kann, wenn man das Verhalten eines Pendels erklären will, zunächst die Wirkungen der Reibung und des Luftwiderstands vernachlässigen und sich ganz auf den Einfluß der Länge des Pendels und der Gravitationskraft konzentrieren. Man kann aber nicht umgekehrt vorgehen. D.h. man kann nicht zuerst nur die Wirkung von Reibung und Luftwiderstand betrachten und dabei den Einfluß der Länge des Pendels und der Gravitation völlig außer acht lassen. Denn auf diese Weise kommt man überhaupt nicht zu einem Ergebnis, nicht einmal zu einem nur näherungsweise richtigen.

III

((25)) Im letzten Abschnitt habe ich einen zugegebenermaßen sehr knappen Überblick über die Theorien von Dretske, Fodor und Millikan gegeben. Dabei wollte ich auf der einen Seite zunächst deutlich machen, mit welchen Mitteln heute im allgemeinen versucht wird, das Problem der Intentionalität zu lösen, und auf der anderen Seite zugleich auf die Probleme hinweisen, mit denen die verschiedenen Versuche konfrontiert sind. Aber diese Probleme sollen hier nicht im Vordergrund stehen. Wichtiger ist mir, daß bei der Darstel-

lung der Ansätze von Dretske, Fodor und Millikan zweierlei deutlich geworden ist. Erstens, daß bei der Zuschreibung intentionaler Inhalte¹⁷ offenbar eine Reihe durchaus verschiedener Aspekte - Kovariation, Kausalität, biologische Funktion - eine Rolle spielen, die nicht immer leicht miteinander zu vereinbaren sind. Und zweitens, daß alle diese Ansätze ohne weitere Begründung von der Voraussetzung ausgehen, daß es auf die Frage, welchen Inhalt ein intentionaler Zustand hat, eine klare und eindeutige Antwort gibt, d.h. daß es natürliche Bedingungen gibt, die den Inhalt intentionaler Zustände eindeutig festlegen. Die Tatsache, daß die verschiedenen Ansätze schon bei einfachen Beispielen zu völlig verschiedenen Ergebnissen kommen, zeigt meiner Meinung nach jedoch, daß gerade diese Voraussetzung außerordentlich zweifelhaft ist.

((26)) Dies wird an dem folgenden Beispiel besonders deutlich, das, nachdem es von Dretske in (1986) in die Diskussion eingeführt wurde, inzwischen schon eine gewisse Berühmtheit erlangt hat. Einige im Meer lebende Bakterien enthalten in ihrem Zellkörper kleine Magnete, sogenannte Magnetosome. Diese Magnetosome bewirken, daß sich die Bakterien immer parallel zum Magnetfeld der Erde ausrichten. In der nördlichen Hemisphäre führt daher jede Vorwärtsbewegung diese Bakterien nach unten in tieferes Wasser, da die magnetischen Feldlinien in dieser Hemisphäre nach unten zum geomagnetischen Nordpol gerichtet sind. Der Überlebenswert dieses magnetotaktischen Verhaltens ist nicht ohne weiteres offensichtlich. Aber bei den Bakterien, um die es hier geht, handelt es sich um sogenannte anaerobe Bakterien, die nur in sauerstofffreiem Wasser überleben können. Aus diesem Grunde wird allgemein angenommen, daß die Magnetosome die Funktion haben, die Bakterien von sauerstoffhaltigem Oberflächenwasser fernzuhalten. Festgestellt wurde jedenfalls, daß bei Bakterien der gleichen Art, die in der südlichen Hemisphäre leben, die Magnetosome genau umgekehrt gepolt sind. Diese Bakterien werden durch ihre Magnetosome daher in Richtung des geomagnetischen Südpols ausgerichtet, was dort, wo sie leben, den gleichen überlebensfördernden Effekt hat. Wenn man ein südliches Bakterium dagegegen in den Nordatlantik verpflanzt, bewegt es sich nach oben in sauerstoffhaltiges Wasser und zerstört sich damit selbst. Wie kann unter diesen Umständen die Antwort auf die Frage aussehen, was genau die Magnetosome mit ihrer Ausrichtung anzeigen? Die Richtung des magnetischen Feldes in ihrer unmittelbaren Umgebung? Die Richtung des Magnetfeldes der Erde? Oder die Richtung, in der sich sauerstofffreies Wasser befindet?

((27)) Wenn man Fodor folgt, liegt zunächst die erste Antwort nahe. Denn die Richtung der Magnetosome wird ausschließlich durch die Richtung des magnetischen Feldes in der unmittelbaren Umgebung der Bakterien verursacht. Wenn die Magnetosome die Richtung des magnetischen Feldes in ihrer unmittelbaren Umgebung anzeigen, stellt sich jedoch sofort das Problem der Fehlrepräsentation. Denn die Magnetosome können nicht anders, als sich in dieser Richtung auszurichten.

Auf der anderen Seite scheint es aber doch ganz plausibel anzunehmen, daß tatsächlich etwas schief läuft, wenn wir die Richtung der Magnetosome durch einen Stabmagneten verändern und dadurch bewirken, daß sich die Bakterien in sauerstoffhaltiges Wasser bewegen und auf diese Weise selbst zerstören. D.h., in diesem Fall scheint es plausibel anzunehmen, daß die Magnetosome etwas Falsches anzeigen. Auch dieser Intuition kann man jedoch im Rahmen der Überlegungen Fodors gerecht werden. Denn Fodors Theorie ist nicht auf die erste Antwort festgelegt, sie ist auch mit der zweiten Antwort vereinbar. Fodor zufolge zeigt die Ausrichtung von Magnetosomen genau dann X an, wenn (1) unter optimalen Bedingungen X die Ausrichtung der Magnetosome verursacht und wenn (2) die Tatsache, daß in manchen Fällen Y die Ausrichtung der Magnetosome verursacht, asymmetrisch von der Tatsache abhängt, daß in der Regel X diese Ausrichtung verursacht. Der Fodorschen Theorie zufolge zeigt die Ausrichtung der Magnetosome daher nicht die Richtung des magnetischen Feldes in ihrer unmittelbaren Umgebung, sondern die Ausrichtung des Erdmagnetfeldes an, wenn (1) unter optimalen Bedingungen das Erdmagnetfeld die Ausrichtung der Magnetosome verursacht und wenn (2) die Tatsache, daß manchmal andere Magnetfelder die Ausrichtung von Magnetosomen verursachen, asymmetrisch von der Tatsache abhängt, daß in der Regel das Erdmagnetfeld diese Ausrichtung verursacht.

Die Klausel (1) bereitet in diesem Fall keine besonderen Schwierigkeiten. Denn damit diese Klausel erfüllt ist, genügt es zu sagen, daß keine optimalen Bedingungen vorliegen, wenn das Erdmagnetfeld durch andere Magnetfelder überlagert oder abgeschirmt wird. Mit der Klausel (2) scheint es dagegen etwas schwieriger zu sein. Denn um zu zeigen, daß auch diese Klausel erfüllt ist, muß man nachweisen, daß die Tatsache, daß andere Magnetfelder manchmal die Ausrichtung von Magnetosomen verursachen, asymmetrisch von der Tatsache abhängt, daß in der Regel das Erdmagnetfeld diese Ausrichtung verursacht. Auf den ersten Blick ist nicht ganz klar, wie man dabei vorgehen soll. Aber vielleicht kann man dieses Problem lösen, indem man plausibel macht, daß in den Bakterien, um die es geht, niemals Magnetosome entstanden wären, wenn die Ausrichtung dieser Magnetosome nicht in der Regel durch das Erdmagnetfeld verursacht würde.¹⁸

((28)) Wenn man Fodor folgt, kommen aber auf jeden Fall nur die ersten beiden Antworten in Frage. Denn sauerstofffreies Wasser liegt zwar in der Regel in der Richtung des geomagnetischen Nordpols. Aber mit der Verursachung der Ausrichtung von Magnetosomen hat es nichts zu tun. Für den funktional-teleologischen Ansatz Millikans spielt dieser kausale Aspekt jedoch überhaupt keine Rolle. Denn diesem Ansatz zufolge ist nicht interessant, wie ein repräsentationaler Zustand normalerweise verursacht wird, interessant ist vielmehr nur, was die Teilsysteme des Gesamtsystems, die in ihrem Verhalten durch diesen Zustand beeinflusst werden, benötigen, um ihre biologische Funktion erfüllen zu können. Die motorischen Systeme der Bakterien, um die es geht, sind jedoch im Laufe der Evolution eben deshalb entstanden, weil sie im Normalfall die Wirkung haben, diese Bakterien in

Richtung sauerstofffreien Wassers zu bewegen bzw. von sauerstoffhaltigem Wasser fernzuhalten. In dieser Wirkung liegt daher ihre biologische Funktion. Um diese Wirkung erzielen zu können, sind die motorischen Systeme jedoch darauf angewiesen, daß die Bakterien sich in die richtige Richtung - d.h. in die Richtung sauerstofffreien Wassers - ausrichten. Diese Arbeit wird offenbar von den Magnetosomen erledigt. Damit die motorischen Teilsysteme ihre Funktion erfüllen können, müssen sich die Magnetosome also in Richtung sauerstofffreien Wassers ausrichten. Und daher repräsentieren sie durch ihre Ausrichtung auch genau die Richtung sauerstofffreien Wassers und nichts anderes.

“What [the systems that react to the magnetosomes] need is only that the pull be in the direction of oxygen-free water at the time. ... they care not at all how it came about that the pull is in that direction; the magnetosomes that points toward oxygen-free water quite by accident and not in accordance with any normal explanation will do just as well as ones that points that way for normal reasons. (...) What the magnetosome represents then is univocal; it represents only the direction of oxygen-free water.” (Millikan 1989, 290f.)

((29)) Wie läßt sich dieser Streit zwischen Fodor und Millikan, der sich ja fast wie ein Glaubensstreit ausnimmt, entscheiden? Was spricht für die eine oder die andere Antwort? Was spricht überhaupt für die Annahme, daß es auf die Frage, was die Magnetosome in Dretskes Bakterien repräsentieren, eine und nur eine richtige Antwort gibt? Tatsache ist, daß die Ausrichtung dieser Magnetosome in der Regel sowohl mit der Richtung des Erdmagnetfeldes als auch mit der Richtung sauerstofffreien Wassers in engem Zusammenhang steht. Aber wie will man entscheiden, welcher dieser beiden Aspekte für die Antwort auf die Frage entscheidend ist, was die Magnetosome in Dretskes Bakterien repräsentieren?

Wenn man die verschiedenen Antworten auf diese Frage vergleicht, kann man sich meiner Meinung nach des folgenden Eindrucks nicht erwehren. Erstens: Bei der Zuschreibung repräsentationaler Inhalte¹⁹ spielen offenbar eine ganze Reihe von Faktoren eine Rolle, deren jeweiliger Beitrag in der Regel nicht genau abgegrenzt werden kann. Und zweitens: Es scheint überhaupt keine klaren Kriterien zu geben, die es ermöglichen würden zu entscheiden, welcher der vielen vorgeschlagenen Lösungsversuche der richtige ist. Mit anderen Worten: Alle sind in gewisser Weise plausibel; aber mit allen ist auch der Eindruck einer gewissen Willkür verbunden.

IV

((30)) Ich denke, daß einem spätestens nach dieser Diagnose der Verdacht kommen sollte, daß möglicherweise das ganze Programm der Naturalisierung des Intentionalen, so wie es von Dretske, Fodor, Millikan und vielen anderen betrieben wird, falsch konzipiert ist. Aber was genau ist falsch daran? Wo liegt der Fehler, der allen diesen Ansätzen gemeinsam ist? Das Fodorsche Naturalisierungsprogramm geht offenbar von folgenden Voraussetzungen aus. Erstens: Es gibt Überzeugungen, Wünsche, Befürchtungen usw., d.h. intentionale

Zustände sind real. Zweitens: Intentionale Zustände sind dadurch charakterisiert, daß sie einen repräsentationalen Inhalt haben. Drittens: Physische Zustände können daher nur dann intentionale Zustände sein, wenn es zu ihren Eigenschaften gehört, daß sie in einer spezifischen semantischen Relation zu einer Proposition p stehen. Und schließlich viertens: Semantische Eigenschaften dieser Art sind im Rahmen eines materialistischen Grundansatzes nur akzeptierbar, wenn sie sich naturalisieren lassen, d.h. wenn sich in nichtsemantischem und nichtintentionalem Vokabular notwendige und hinreichende Bedingungen dafür formulieren lassen, daß ein (physischer) Zustand den repräsentationalen Inhalt p hat.

((31)) Meiner Meinung nach liegt der entscheidende Fehler in der dritten Annahme, daß physische Zustände nur dann intentionale Zustände sein können, wenn es zu ihren Eigenschaften gehört, in einer spezifischen semantischen Relation zu einer Proposition p zu stehen. Soweit ich sehen kann, beruht diese Annahme auf einer Fehlinterpretation der Tatsache, daß wir zur Bezeichnung intentionaler Zustände sprachliche Ausdrücke verwenden, die in der Regel mit Hilfe von daß-Sätzen gebildet werden. Allerdings liegt diese Fehlinterpretation nahe. Denn wenn man sagt, daß ein System S glaubt, daß Hunde bellen, dann scheint man damit doch zu sagen, daß sich S in einem Zustand befindet, der erstens zu einem bestimmten Typ von intentionalen Zuständen gehört, nämlich zum Typ der Überzeugungen, und der zweitens einen bestimmten Inhalt hat, nämlich den Inhalt, daß Hunde bellen. Wie kann ein Zustand also die Überzeugung sein, daß Hunde bellen, wenn er nicht die Eigenschaft hat, diesen Inhalt zu haben, und worin anders soll diese Eigenschaft bestehen als darin, daß dieser Zustand in einer spezifischen semantischen Relation zu der Proposition, daß Hunde bellen, steht?

((32)) Doch diese Argumentation ist kurzschlüssig. Denn daraus, daß wir zur Bezeichnung bzw. Klassifizierung der Zustände eines Systems Ausdrücke wie "die Überzeugung, daß p " verwenden, folgt keineswegs, daß diese Bezeichnung auf einen Zustand nur dann zutreffen kann, wenn dieser Zustand die Eigenschaft hat, in einer spezifischen semantischen Relation zur Proposition p zu stehen. Dieser Punkt ist in den letzten Jahren besonders von Paul Churchland betont worden, der ebenso wie einige andere Autoren (Field, Stalnaker, Davidson²⁰) die These vertreten hat, daß das Vokabular, mit dem wir intentionale Zustände bezeichnen, am besten zu verstehen sei in Analogie zu dem Vokabular, mit dem wir über physikalische Größen wie Länge, Gewicht und Temperatur reden, und daß die Rolle, die daß-Sätze in diesem Vokabular spielen, in etwa der Rolle der Zahlausdrücke im Vokabular der metrischen Begriffe entspreche. Meiner Meinung nach muß man mit dieser Analogie zwar vorsichtig sein, da sie dazu verführen könnte, die Unterschiede zwischen diesen beiden Familien von Begriffen zu verwischen. Aber im Kern halte ich sie für richtig. Wenn man das Vokabular verstehen will, mit dem wir uns auf intentionale Zustände beziehen, sollte man daher am besten mit einer Analyse des Vokabulars der metrischen Begriffe beginnen.

((33)) Ausdrücke wie "x hat eine Länge von y cm", "x hat

eine Masse von y kg" und "x hat eine Temperatur von y Grad Celsius" sehen zwar auf den ersten Blick so aus, als würden sie Relationen bezeichnen, die zwischen raumzeitlichen Gegenständen und bestimmten Zahlen bestehen oder nicht bestehen. Doch dieser erste Eindruck ist irreführend. Denn tatsächlich faßt man die Teilausdrücke "hat eine Länge von ... cm", "hat eine Masse von ... kg" und "hat eine Temperatur von ... Grad Celsius" am besten als Funktoren mit einer Leerstelle auf, aus denen ein Prädikat entsteht, wenn man an der Leerstelle einen Ausdruck für eine (positive) reelle Zahl einsetzt.²¹ Wenn man sagt "Der Würfel a hat eine Masse von 2 kg", schreibt man a damit zwar eine bestimmte Eigenschaft zu, eine wirkliche Eigenschaft, die durchaus auch kausale Relevanz hat. Aus der Tatsache, daß wir bei der Zuschreibung dieser Eigenschaft einen Ausdruck gebrauchen, der einen Ausdruck für die Zahl 2 enthält, folgt jedoch nicht, daß diese Eigenschaft darin besteht, daß a in einer spezifischen Relation zu der Zahl 2 steht. Dies zeigt sich schon daran, daß wir dieselbe Eigenschaft auch mit Hilfe der Ausdrücke "hat eine Masse von 2000 g" oder "hat eine Masse von 70,55 Ounces" zuschreiben können. Wenn das so ist, stellt sich jedoch die Frage, warum wir bei der Zuschreibung von Masseeigenschaften ebenso wie bei der Zuschreibung von Längen- oder Temperatureigenschaften überhaupt Prädikate verwenden, die Zahlausdrücke als wesentlichen Bestandteil enthalten.

((34)) Ein Teil der Antwort auf diese Frage besteht darin, daß die verschiedenen Eigenschaften, die wir alle mit Ausdrücken der Form "hat eine Masse von ... kg" zuschreiben, eine Familie bilden, so wie etwa die Farbeigenschaften, blau zu sein oder grün zu sein oder rot zu sein, eine Familie oder Dimension bilden. Aber das ist noch keine ausreichende Antwort. Denn die verschiedenen Farbeigenschaften bezeichnen wir ja auch nicht mit Hilfe von Ausdrücken, die durch Anwendung eines gemeinsamen Funktors auf einen Zahlausdruck oder einen ähnlichen Indexausdruck entstehen. Worin besteht also der Unterschied zwischen Masseeigenschaften und Farbeigenschaften, der diesem Unterschied in der Bezeichnungsweise zugrundeliegt?

Die Antwort auf diese Frage ist, soweit ich sehen kann, in den Überlegungen zur Grundlegung der Meßtheorie zu finden, wie sie besonders von Hempel (1952) und Suppes und Zinnes (1963) formuliert worden sind. Diese Überlegungen kann man kurz so zusammenfassen: Der Begriff der Masse ist ein metrisierbarer Begriff, weil es im Bereich B der Gegenstände, auf die wir diesen Begriff anwenden können, eine zweistellige empirische Vergleichsrelation R (konkret: die Relation des Aufwiegens auf einer Balkenwaage) und eine ebenfalls zweistellige Zusammenfügeoperation $*$ gibt, so daß die Struktur $\langle B, R, * \rangle$ der Struktur $\langle \mathbb{R}^+, \leq, + \rangle$ der positiven reellen Zahlen mit der Kleiner-Gleich-Relation und der Addition entspricht. D.h. genauer: Von der Menge B gibt es einen strukturerhaltenden Homomorphismus m in die Menge der positiven reellen Zahlen, also eine Abbildung, für die gilt:

$$(1) \quad R(a,b) \text{ genau dann, wenn } m(a) \leq m(b)$$

$$(2) \quad m(a*b) = m(a) + m(b).^{22}$$

((35)) Wenn wir bei der Bildung der Prädikatausdrücke, mit denen wir Masseeigenschaften zuschreiben, Ausdrücke für die Zahlen verwenden, die der Homomorphismus m den Elementen von B zuordnet, können wir daher aus dem Prädikatausdruck selbst ablesen, welchen Platz ein Gegenstand in der Struktur $\langle B, R, * \rangle$ einnimmt, und d.h. insbesondere, in welchen Relationen er zu anderen Gegenständen der Menge B steht. Wenn wir von drei Gegenständen a, b und c wissen, daß sie eine Masse von 2 kg, 3 kg bzw. 5 kg haben, dann wissen wir damit zugleich, daß b schwerer ist als a , daß c schwerer ist als a und schwerer als b und daß a und b zusammen gleich schwer sind wie c .²³ Und da viele Gesetze der Physik in quantitativen Termini formuliert sind, wissen wir dann z.B. auch, daß die Geschwindigkeit von a anderthalbmal so groß sein wird wie die Geschwindigkeit von b , wenn auf diese beiden Gegenstände dieselbe Kraft über denselben Zeitraum einwirkt.²⁴ Wenn wir von einem Gegenstand sagen können, daß er eine Masse von 2 kg hat, dann sind in dem Zahlausdruck "2" also Informationen über den Platz kodiert, den dieser Gegenstand in der Struktur aller Gegenstände der Menge B einnimmt.

V

((36)) Die Grundidee von Autoren wie Churchland, Field, Stalnaker und Davidson ist nun, daß Propositionen im Hinblick auf intentionale Zustände eine ähnliche Rolle spielen wie Zahlen im Hinblick auf physikalische Größen. Bzw., um die in diesem Zusammenhang angemessenere formale Redeweise zu verwenden: daß die daß-Sätze, mit deren Hilfe wir die Ausdrücke bilden, mit denen wir intentionale Zustände zuschreiben, eine ähnliche Rolle spielen wie die Zahlausdrücke, die wir bei der Bildung metrischer Begriffe gebrauchen. Wenn wir sagen, das System S hat die Überzeugung, daß p , dann bedeutet das dieser Idee zufolge nicht, daß wir S damit einen Zustand zuschreiben, der u.a. die Eigenschaft hat, in einer spezifischen semantischen Relation zu der Proposition p zu stehen. Vielmehr sagen wir damit, daß sich S in einem Zustand befindet, der zu anderen Zuständen von S und bestimmten Zuständen außerhalb von S in bestimmten Relationen steht. In den daß-Sätzen, die im intentionalen Vokabular eine so wesentliche Rolle spielen, sind also - ebenso wie in den Zahlausdrücken im Vokabular der metrischen Begriffe - Informationen darüber kodiert, welche Rolle ein Zustand in einem ganzen Netz von Zuständen spielt.

((37)) Ich denke, daß diese Analyse sehr plausibel ist. Allerdings ist es nicht leicht, sie im Detail auszuformulieren. Denn es gibt erhebliche Unterschiede zwischen den Tatsachen, die der Verwendung des intentionalen bzw. des quantitativen Vokabulars zugrundeliegen. Der erste Unterschied betrifft die Frage, auf welche Entitäten das quantitative bzw. das intentionale Vokabular überhaupt angewendet werden kann. Ausdrücke wie "hat eine Länge von ... m" oder "hat eine Masse von ... kg" werden in der Regel auf raum-zeitliche Gegenstände angewendet. Und auch das intentionale Vokabular scheint zunächst einmal auf Individuen, insbesondere auf Per-

sonen, anwendbar. Die von Churchland, Field, Stalnaker und Davidson vorgeschlagene Analogie macht jedoch nur dann einen Sinn, wenn das intentionale Vokabular in erster Linie nicht auf Gegenstände bzw. Personen, sondern auf Zustände von Gegenständen oder Personen angewendet wird. Statt an Ausdrücke wie "hat die Überzeugung, daß p " ist in diesem Zusammenhang daher eher an Ausdrücke wie "ist eine Überzeugung, daß p " oder "ist ein Wunsch, daß q " zu denken.

((38)) Der zweite Unterschied hängt mit dem ersten eng zusammen. Denn während bei den Strukturen, die der Verwendung metrischer Begriffe zugrundeliegen, empirische Relationen und Operationen zwischen Gegenständen - wie etwa die Relation des Aufwiegens oder die Operation des Zusammenlegens - die entscheidende Rolle spielen, kommen solche Relationen und Operationen zwischen Zuständen nicht in Betracht. Stattdessen werden in diesem Fall eher für Zustände spezifische Relationen wie die Kausalrelation eine wichtige Rolle spielen. Und soweit ich sehen kann, sind die meisten Autoren in der Tat der Meinung, daß die Kausalrelation in diesem Zusammenhang die zentrale Rolle spielt.

((39)) Die ersten beiden Unterschiede, die ich gerade angeführt habe, sind vielleicht nicht von so entscheidender Bedeutung. Aber der dritte Unterschied, auf den ich jetzt zu sprechen kommen möchte, ist doch recht gravierend. Denn er hat direkte Folgen für die Frage nach der Identität von intentionalen Zuständen. D.h., genauer gesagt, für die Frage, wann wir von zwei (physikalisch) verschiedenen Zuständen sagen, daß sie denselben intentionalen Zustand instantiiieren. Für physikalische Größen wie die Masse ist charakteristisch, daß die Vergleichsrelation R konnex ist. Alle Gegenstände der Grundmenge B lassen sich bzgl. dieser Relation vergleichen, d.h., für je zwei Elemente a und b von B gilt $R(a, b)$ oder $R(b, a)$ oder beides. Auf der Grundlage dieser Tatsache ergibt sich, daß zwei Gegenstände a und b genau dann dieselbe Masse haben, wenn sie in der Menge B bzgl. der Relation R und der Operation $*$ denselben Platz einnehmen, d.h., wenn für alle Elemente x von B gilt: $R(a, x)$ gdw $R(b, x)$ und $a * x \equiv b * x$. Und dies ist offenbar genau dann der Fall, wenn $a \equiv b$ gilt. Im Falle physikalischer Größen läßt sich die Situation also in etwa durch das folgende Diagramm darstellen.

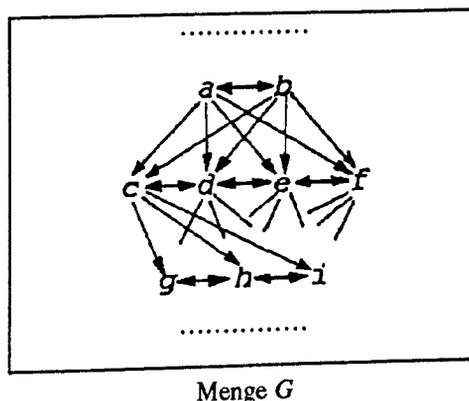


Abb. 1

((40)) Im Falle intentionaler Zustände liegen die Dinge jedoch anders. Denn zwei (physikalische) Zustände x und y , die dieselbe Überzeugung oder denselben Wunsch instantiieren, sind in der Regel Zustände verschiedener Systeme und daher nicht ohne weiteres miteinander vergleichbar. Für intentionale Zustände läßt sich die Situation daher besser durch das folgende Diagramm darstellen.

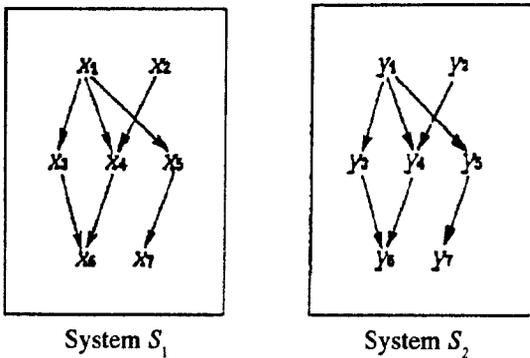


Abb. 2

Bei den intentionalen Zuständen haben wir es also nicht mit nur einer Menge von Zuständen und entsprechend mit nur einer Struktur, sondern mit einer im Prinzip unendlichen Anzahl von Mengen bzw. Strukturen zu tun. Und dies ist ein erheblicher Unterschied zum Fall der physikalischen Größen. Allerdings liegt es in diesem Fall nahe, die Identitätsbedingungen für intentionale Zustände so zu formulieren: Der Zustand x im System S_1 ist derselbe intentionale Zustand wie der Zustand y im System S_2 , wenn x in S_1 strukturell dieselbe Rolle spielt wie y in S_2 . Diese Formulierung setzt jedoch voraus, daß die Strukturen von S_1 und S_2 ihrerseits homomorph sind, d.h. daß es eine strukturerhaltende Abbildung der Zustände von S_1 auf die Zustände von S_2 gibt.

((41)) Schließlich gibt es noch einen vierten Unterschied. Ich hatte schon gesagt, daß wir physikalische Größen zwar mit Hilfe von Prädikaten zuschreiben, die Zahlausdrücke als wesentlichen Bestandteil enthalten, daß dies aber nicht bedeutet, daß die Gegenstände, auf die diese Prädikate zutreffen, in einer spezifischen Relation zu den entsprechenden Zahlen stehen. Wenn man sagt, daß dieser Tisch eine Masse von 10 kg hat, sagt man damit nicht, daß zwischen diesem Tisch und der Zahl 10 irgendeine Art von Relation besteht. Dies zeigt sich, wie gesagt, insbesondere daran, daß wir in diesem Zusammenhang grundsätzlich auch jeden anderen Zahlausdruck verwenden könnten, weil sogar bei extensiven Größen der Homomorphismus, der dem ganzen Verfahren zugrundeliegt, nur bis auf die Multiplikation mit einer positiven Zahl eindeutig ist. Aus dieser Tatsache ergibt sich, so könnte man sagen, eine spezifische Unbestimmtheit bei der Zuschreibung physikalischer Größen. Man kann nicht sagen, daß sie einen spezifischen "numerischen Inhalt" haben, d.h. daß sie nur mit Hilfe bestimmter Zahlausdrücke zugeschrieben werden können.

((42)) Auch hier liegen die Dinge bei den intentionalen Zu-

ständen in gewisser Weise anders. Denn für die Zuschreibung dieser Zustände spielen nicht nur die kausalen Relationen eine Rolle, die zwischen den Zuständen eines Systems bestehen, sondern auch Kausalbeziehungen zwischen diesen Zuständen und Zuständen außerhalb des Systems. So liegt es z.B. nahe, von einem Zustand x nur dann zu sagen, er sei eine Überzeugung, daß p , wenn er regelmäßig von solchen und nur von solchen Situationen verursacht wird, in denen p wahr ist. Und auf der anderen Seite liegt es ebenso nahe, von einem Zustand y nur dann zu sagen, er sei ein Wunsch, daß q , wenn dieser Zustand in der Regel Handlungen verursacht, die geeignet sind, q herbeizuführen. Insgesamt muß das in Abb. 2 gezeigte Diagramm daher noch einmal um diese nicht zum System selbst gehörenden Zustände ergänzt werden.

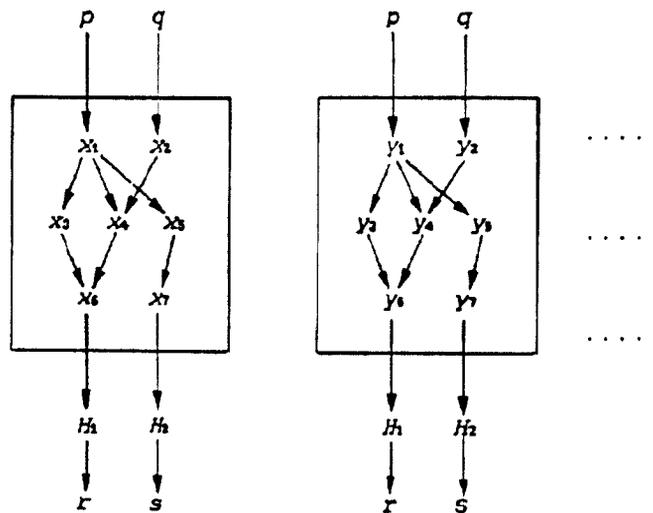


Abb. 3

((43)) Vor dem Hintergrund dieses Diagramms können wir nun zu der entscheidenden Frage kommen, welche Struktur die Zustände eines Systems bilden müssen, damit es möglich ist, diese Zustände so auf die Menge der Propositionen abzubilden, daß es sinnvoll wird, diese Menge zur Indexmenge für entsprechende Prädikate zu machen. Im Zusammenhang mit dieser Frage ist häufig die Idee formuliert worden, daß die Kausalbeziehungen zwischen den Zuständen eines Systems den semantischen Beziehungen und insbesondere den logischen Folgerungsbeziehungen zwischen den korrespondierenden Propositionen entsprechen sollten. Meines Wissens ist diese Idee jedoch nie bis ins Detail ausbuchstabiert worden. D.h., es gibt in diesem Bereich nichts, was den Repräsentationstheoremen im Bereich der metrischen Begriffe entspräche.²⁵ Insbesondere ist nie die Frage diskutiert worden, welche Eigenschaften die Kausalbeziehungen zwischen den Zuständen eines Systems haben müssen, damit es möglich ist, diese Zustände homomorph in die Menge der Propositionen abzubilden. Fodors Überlegungen zu einer Sprache des Geistes, derzufolge Denken im syntaktischen Umformen strukturierter Repräsentationen besteht, sind hier vielleicht noch am einschlägigsten. Aber Fodor selbst hat seine Überlegungen niemals im Zusammenhang mit der hier zur Debatte stehenden Frage gesehen. Letzten Endes kann

man also nur den Mangel an entsprechenden Überlegungen konstatieren. Vielleicht können wir die Repräsentationsproblematik hier für den Augenblick jedoch zunächst einmal einklammern bzw. als gelöst voraussetzen und den genannten Vorschlag einfach so formulieren:

(KT) Es ist möglich und sinnvoll, die physikalischen Zustände von Systemen mit intentionalen Prädikaten zu bezeichnen, d.h. mit Prädikaten, die aus der Anwendung eines Funktors wie "ist eine Überzeugung x " oder "ist ein Wunsch y " auf einen daß-Satz entstehen, wenn es eine strukturerhaltende Abbildung f dieser Zustände in die Menge der Propositionen gibt, d.h. eine Abbildung, für die gilt:

- (i) Zustand x_1 verursacht Zustand x_2 genau dann, wenn die Proposition $f(x_1)$ die Proposition $f(x_2)$ logisch impliziert.

Diesen Vorschlag möchte ich, da er sich im wesentlichen an den Überlegungen der klassischen Meßtheorie orientiert, die Klassische Theorie (KT) nennen.

((44)) Offensichtlich leidet diese KT jedoch an einer ganzen Reihe von Schwächen. Erstens ist die in ihr implizit enthaltene Rationalitätsannahme viel zu streng.²⁶ Denn die Bedingung (i) impliziert, daß eine Überzeugung $\langle p \rangle$ eine Überzeugung $\langle q \rangle$ ²⁷ nur dann verursachen kann, wenn q logisch aus p folgt. Aber natürlich soll es auch möglich sein, daß eine Überzeugung $\langle p \rangle$ eine Überzeugung $\langle q \rangle$ verursacht, wenn p nur gute Gründe für den Glauben an q liefert, ohne q deshalb auch schon logisch zu implizieren. Und schließlich sollen auch irrationale Überzeugungen möglich sein, d.h. auch Überzeugungen, die durch andere Überzeugungen verursacht werden, die keine guten Gründe für sie darstellen.

((45)) Die zweite Schwäche der KT liegt darin, daß sie die Kausalbeziehungen völlig außer acht läßt, die zwischen den internen Zuständen eines Systems und bestimmten Zuständen in seiner Umgebung bzw. bestimmten Handlungen des Systems bestehen. Und diese Kausalbeziehungen spielen, wie wir schon gesehen hatten, bei der Zuschreibung intentionaler Inhalte offenbar ebenfalls eine große Rolle.

((46)) Mit dieser Schwäche hängt die dritte eng zusammen. Der KT zufolge ist der Inhalt intentionaler Zustände außerordentlich unbestimmt. Denn wenn es eine Abbildung f gibt, die die Bedingung (i) erfüllt, dann gibt es unendlich viele verschiedene Abbildungen dieser Art. Nehmen wir z.B. an, daß f die internen Zustände eines Systems S auf eine Teilmenge P der Propositionen abbildet, deren Elemente durch die daß-Sätze "daß p_1 ", "daß p_2 ", "daß p_3 " usw. bezeichnet werden (bzw. durch die Ausdrücke " $\langle p_1 \rangle$ ", " $\langle p_2 \rangle$ ", " $\langle p_3 \rangle$ " usw.). Wenn wir in diesen Ausdrücken überall den Namen "Konrad Adenauer" durch "Willy Brandt" ersetzen, erhalten wir eine Menge von Ausdrücken, die die Elemente einer anderen Teilmenge P' von Propositionen bezeichnen. Offenbar bestehen aber zwischen den Elementen von P' dieselben logischen

Beziehungen wie zwischen den entsprechenden Elementen von P . Wenn wir die Funktion f so wählen, daß sie den internen Zuständen von S die entsprechenden Elemente von P' zuordnet, erfüllt also auch f die Bedingung (i).

((47)) Die vierte Schwäche der KT besteht schließlich darin, daß sie keine vernünftige Unterscheidung verschiedener Typen intentionaler Zustände ermöglicht, also z.B. keine Unterscheidung von Wünschen und Überzeugungen.²⁸ Denn dieser Theorie zufolge ordnet die Funktion f jedem internen Zustand eine Proposition p_i zu, ohne uns zu sagen, ob es sich bei diesem Zustand um den Wunsch, daß p_i , die Überzeugung, daß p_i , oder um noch einen anderen intentionalen Zustand mit dem Inhalt p_i handelt. An dieser Stelle wird im übrigen auch deutlich, daß die KT bestenfalls im Hinblick auf Überzeugungen eine gewisse Plausibilität in Anspruch nehmen kann. Denn während es zumindest nicht unplausibel ist anzunehmen, daß jemand, der von $\langle p \rangle$ und von \langle wenn p , dann $q \rangle$ überzeugt ist, auch von $\langle q \rangle$ überzeugt sein sollte, macht eine entsprechende Annahme im Hinblick auf Wünsche wenig Sinn.

VI

((48)) Ich denke, die aufgezählten Schwächen zeigen, daß man nicht sehr weit kommt, wenn man sich bei der Analyse intentionalen Vokabulars zu sehr an der Homomorphismusidee der klassischen Meßtheorie orientiert. Zum Glück gibt es aber eine Alternative, die im wesentlichen auf Ideen von Brian Loar basiert.²⁹ Auch diese Alternative geht von der Annahme aus, daß wir die internen Zustände eines Systems genau dann mit Hilfe des intentionalen Vokabulars bezeichnen können, wenn sie zueinander und zu bestimmten Zuständen außerhalb des Systems in bestimmten kausalen Beziehungen stehen. Aber sie entwickelt diese Idee auf eine andere, eher an funktionalistischen Überlegungen orientierte Weise.

((49)) Ausgangspunkt ist dabei die Beobachtung, daß Kausalbeziehungen häufig nicht nur zwischen bestimmten einzelnen intentionalen Zuständen, sondern zwischen allen Zuständen bestehen, deren Inhalte die gleiche logische Form haben. Mit anderen Worten: Es ist nicht nur so, daß die Überzeugung, daß Hans blond und 1.80 m groß ist, die Überzeugung verursacht, daß Hans 1.80 m groß ist; es gilt vielmehr generell, daß Überzeugungen, deren Inhalt die logische Form $p \ \& \ q$ hat, die entsprechende Überzeugung $\langle q \rangle$ verursachen. Diese Tatsache führt dazu, daß man bei der Formulierung der Gesetze, die diesen Kausalbeziehungen zugrunde liegen, über die propositionalen Inhalte quantifizieren kann. Für die gerade angeführte Kausalbeziehung könnte ein entsprechendes Gesetz daher lauten:

- (1) $(x)(p)(q)$ (Wenn x die Überzeugung $\langle p \ \& \ q \rangle$ hat, dann hat x auch die Überzeugung $\langle q \rangle$)

Und ein Gesetz, das Wünsche und Überzeugungen in nahegelegener Weise miteinander verbindet, könnte in etwa lauten:

- (2) $(x)(p)(q)$ (Wenn x den Wunsch $\langle p \rangle$ und die Überzeugung $\langle p$ nur, wenn $q \rangle$ hat, dann hat x auch den Wunsch $\langle q \rangle$)

Gesetze dieser Art verknüpfen aber nicht nur intentionale Zustände miteinander, sondern auch Zustände außerhalb des Systems mit seinen intentionalen Zuständen und diese Zustände ihrerseits mit möglichen Handlungen des Systems. Unter optimalen Wahrnehmungsbedingungen z.B. sollte jede wahrgenommene Situation auch zu einer entsprechenden Überzeugung führen. Und wenn entsprechende Bedingungen vorliegen, sollte ein System die Handlungen auch ausführen, die es auszuführen wünscht. Neben die Gesetze (1) und (2) treten deshalb außerdem Gesetze wie

- (3) $(x)(p)$ (Wenn $\langle p \rangle$ wahr ist und die Bedingung C_1 vorliegt, dann hat x die Überzeugung $\langle p \rangle$)

und

- (4) $(x)(p)$ (Wenn $x \langle p \rangle$ wünscht und die Bedingung C_2 vorliegt, dann wird $\langle p \rangle$ wahr)

Diese Gesetze sind sicher nicht die einzigen, die für intentionale Zustände charakteristisch sind, und sie sind sicher auch nicht optimal formuliert. Um des Arguments willen möchte ich hier jedoch davon ausgehen, daß die Gesetze (1) - (4) plausibel sind und daß sie zusammen eine akzeptable Theorie intentionaler Zustände darstellen.

((50)) Für die zuvor angestellten meßtheoretischen Überlegungen war die Frage kennzeichnend, ob sich die Menge der Zustände eines Systems S in strukturhaltender Weise in die Menge der Propositionen abbilden läßt. Diese Frage können wir nun in etwas abgewandelter Form wieder aufnehmen, wobei es jedoch aus technischen Gründen sinnvoll ist, nicht Abbildungen von der Menge der Zustände eines Systems in die Menge der Propositionen, sondern umgekehrt Abbildungen von der Menge der Propositionen in die Menge der Zustände eines Systems zu betrachten.³⁰ Denn dann können wir diese Frage so formulieren: Läßt sich die Menge der Propositionen so auf die Menge der Zustände eines Systems S abbilden, daß die Zustände, die auf diese Weise bestimmten Propositionen zugeordnet werden, die Gesetze (1) - (4) erfüllen?

((51)) Wie ist das genau zu verstehen? Zunächst einmal müssen wir zumindest von zwei Funktionen f und g ausgehen, um das zuvor schon angesprochene Problem zu lösen, z.B. Wünsche und Überzeugungen voneinander unterscheiden zu können. Nehmen wir also an, f und g seien zwei solche Funktionen, die die Menge der Propositionen auf verschiedene Teilmengen der internen Zustände des Systems S abbilden. Die Funktion g ordnet der Proposition $\langle S$ wird reich \rangle also den Zustand x_1 und der Proposition $\langle S$ spart \rangle den Zustand x_2 zu und die Funktion f der Proposition $\langle S$ wird nur reich, wenn S spart \rangle den Zustand x_3 . In diesem Fall können wir offenbar fragen, ob die Zustände x_1 und x_3 tatsächlich den Zustand x_2 verursachen, so wie es dem Gesetz (2) entsprechen

würde. Wenn wir so fragen, betrachten wir jedoch wieder nur einen Einzelfall und der ist für sich genommen wenig aussagekräftig. Wirklich interessant ist dagegen die Frage, ob eine entsprechende Kausalbeziehung für alle Argumente von f und g gilt, die die gleiche logische Form haben wie die Propositionen in dem gerade angesprochenen Beispiel. D.h., ob für alle Propositionen $\langle p \rangle$ und $\langle q \rangle$ gilt, daß die Zustände, die die Funktion g der Proposition $\langle p \rangle$ und die Funktion f der Proposition $\langle p$ nur, wenn $q \rangle$ zuordnet, jeweils den Zustand verursachen, den die Funktion g der Proposition $\langle q \rangle$ zuordnet. Denn wenn das so ist, dann ist die dem Gesetz (2) entsprechende gesetzesartige Aussage

- (2') $(p)(q)$ (Wenn S in $g(\langle p \rangle)$ und in $f(\langle p$ nur, wenn $q \rangle)$ ist, dann ist S in $g(\langle q \rangle)$)

wahr, und dann sind die Zustände von S , die den Wertebereich der Funktion f bilden, daher plausible Kandidaten für Überzeugungszustände und die Zustände von S , die den Wertebereich der Funktion g bilden, plausible Kandidaten für Wunschzustände.

((52)) Um diese Überlegung weiter zu präzisieren, ist es zunächst nötig, auch für die Gesetze (1), (3) und (4) entsprechende gesetzesartige Aussagen zu formulieren:

- (1') $(p)(q)$ (Wenn S in $f(\langle p \& q \rangle)$ ist, dann ist S in $f(\langle q \rangle)$)

- (3') (p) (Wenn $\langle p \rangle$ wahr ist und die Bedingung C_1 vorliegt, dann ist S in $f(\langle p \rangle)$)

- (4') (p) (Wenn S in $g(\langle p \rangle)$ ist und die Bedingung C_2 vorliegt, dann wird $\langle p \rangle$ wahr)

Wenn wir die Konjunktion von (1') - (4') T nennen, dann ist T genau genommen keine Aussage, sondern eine Aussagefunktion, da in ihr die Variablen f und g zunächst frei vorkommen. Man kann daher die Frage stellen, ob es Funktionen gibt, die diese Aussagefunktion erfüllen. Und dies ist tatsächlich die in diesem Zusammenhang entscheidende Frage. Denn während die Grundidee der an der klassischen Meßtheorie orientierten Klassischen Theorie besagt, daß es genau dann möglich und sinnvoll ist, die internen Zustände eines Systems S mit Hilfe des intentionalen Vokabulars zu bezeichnen, wenn es eine homomorphe Abbildung dieser Zustände in die Menge der Propositionen gibt, lautet die Grundidee der an den Überlegungen Loars orientierten Alternativen Theorie:

- (AT) Es ist möglich und sinnvoll, die physikalischen Zustände eines Systems S mit intentionalen Prädikaten zu bezeichnen, d.h. mit Prädikaten, die aus der Anwendung eines Faktors wie "ist eine Überzeugung x " oder "ist ein Wunsch y " auf einen daß-Satz entstehen, wenn es Abbildungen f und g von der Menge der Propositionen in die Menge der internen Zustände von S gibt, die die Aussagefunktion T erfüllen.

Die Alternative Theorie verzichtet also auf die für den klassischen Ansatz charakteristische Idee der homomorphen Abbildung zugunsten einer ähnlichen, aber eher an funktionalistischen Ansätzen orientierten Grundidee.

VII

((53)) Im letzten Abschnitt hat sich gezeigt, daß es - trotz der zuvor angesprochenen Schwierigkeiten - tatsächlich doch möglich ist, eine meßtheoretische Analyse intentionaler Prädikate auch im Detail durchzuführen. Damit hat sich die Grundidee dieser Analyse bewährt, eine Idee, in der, um es noch einmal zu wiederholen, zweierlei enthalten ist. Erstens der eher formale Punkt, daß man Ausdrücke wie "hat die Überzeugung, daß p " oder "ist eine Überzeugung, daß p " als Prädikate auffassen sollte, die durch Anwendung entsprechender Funktoren auf daß-Sätze entstehen, d.h. auf Ausdrücke, die die Elemente einer bestimmten Indexmenge bezeichnen. Die Bedeutung dieser formalen Analyse wird jedoch erst deutlich, wenn man sieht, wie sie mit dem zweiten inhaltlichen Punkt zusammenhängt. Denn dieser Punkt besagt: Wenn wir von einem System S sagen, es habe die Überzeugung, daß p , dann bedeutet das nicht, daß wir S damit einen Zustand zuschreiben, der die Eigenschaft hat, in einer spezifischen semantischen Relation zu der Proposition p zu stehen. Vielmehr sagen wir auf diese Weise, daß sich S in einem Zustand befindet, der mit den anderen Zuständen von S und mit bestimmten Zuständen außerhalb von S in bestimmter Weise kausal verbunden ist. In dem entsprechenden daß-Satz sind also Informationen darüber kodiert, welche Rolle dieser Zustand in einer ganzen Struktur von Zuständen spielt.

((54)) Wenn diese Analyse richtig ist, bedeutet das aber eine Auflösung des Problems der Intentionalität. Denn dieses Problem war durch die Frage charakterisiert, ob die Eigenschaft, einen repräsentationalen Inhalt zu haben, nicht doch naturalisiert werden kann, d.h. ob es nicht doch möglich ist, in nicht-semantischem und nichtintentionalem Vokabular notwendige und hinreichende Bedingungen dafür zu formulieren, daß ein (physischer) Zustand den repräsentationalen Inhalt p hat.

Wenn jedoch die meßtheoretische Analyse zutrifft, dann verwenden wir intentionale Begriffe gar nicht dazu, Systemen Zustände zuzuschreiben, zu deren Eigenschaften es gehört, in einer spezifischen semantischen Relation zu einer Proposition zu stehen, und dann entfällt damit die Grundvoraussetzung des Problems der Intentionalität. Mit anderen Worten, wenn die meßtheoretische Analyse zutrifft, dann stellt sich die Frage nach der Naturalisierbarkeit der semantischen Eigenschaften intentionaler Zustände einfach nicht mehr.

Anmerkungen

1. Vgl. dagegen Brentanos Argumentation (1924, 126-128).

2. An dieser Stelle scheint es mir sinnvoll, auf ein mögliches Mißverständnis hinzuweisen, das sich aus der Tatsache ergeben könnte, daß die Ausdrücke "Intention" und "intentional" im Deutschen im allgemeinen im

Sinne von "Absicht" bzw. "absichtlich" gebraucht werden. Der Ausdruck "Intentionalität", so wie er von Brentano (im Rückgriff auf den Sprachgebrauch einiger mittelalterlicher Philosophen) in die philosophische Diskussion eingeführt wurde, ist jedoch ein *terminus technicus* der Philosophie, dessen Bedeutung mit diesem herkömmlichen Sinn von "Intention" und "intentional" nur sehr wenig zu tun hat. Intentionalität im Sinne Brentanos bedeutet nicht Absichtlichkeit, sondern Gerichtetheit auf ein Objekt (im Englischen häufig mit "aboutness" übersetzt). Intentional in diesem Sinne sind Phänomene, wenn sie auf etwas gerichtet sind bzw. wenn sie Wahrheits- bzw. Erfüllungsbedingungen haben. Absichten und Wünsche gehören daher zwar ebenfalls zur Gruppe der intentionalen Zustände, da auch sie ein Objekt bzw. Erfüllungsbedingungen haben, aber sie bilden nur eine Art von intentionalen Zuständen neben vielen anderen wie etwa Überzeugungen, Hoffnungen, Befürchtungen, usw. D.h. sie haben im Hinblick auf das Problem der Intentionalität keine Sonderstellung. Vgl. zu diesem Punkt z.B. Searle (1983, ch. I, sec. i).

3. In dieser Aufzählung scheint auf den ersten Blick der interpretations-theoretische Ansatz von Haugeland und Cummins zu fehlen. Dieser hat jedoch eine größere Ähnlichkeit mit den meßtheoretischen Überlegungen, auf die ich unten ausführlich eingehen werde, als mit den hier aufgeführten klassischen Versuchen der Naturalisierung der Intentionalität (vgl. Cummins 1989, 169 ch. 9 n.7). Er gehört aus diesem Grund nicht in diese Gruppe.

4. Vgl. Dretske (1981), S. 185.

5. Aufgrund dieser Formulierung ist klar, daß gilt: Wenn ein Signal s die Information e trägt und wenn in der Information e die Information e' eingebettet ist, dann trägt s auch die Information e' . Denn in diesem Fall ist e eine notwendige Bedingung für s und e' eine notwendige Bedingung für e , also auch e' eine notwendige Bedingung für s .

6. Bei dieser Formulierung habe ich versucht, die verschiedenen Definitionen Dretskes in (1981, 137, 177, 185 und 260) auf einen Nenner zu bringen.

7. Dies läßt sich folgendermaßen beweisen. Daß s die Information e trägt, besagt, daß e eine notwendige Bedingung für s ist. Und wenn s keine Information e' trägt, von der gilt: die Information e ist in der Information e' eingebettet, aber nicht umgekehrt, dann bedeutet das, daß es kein e' gibt, für das gilt:

(i) e' ist eine notwendige Bedingung für s und e ist eine notwendige Bedingung für e' , aber e' ist keine notwendige Bedingung für e .

Dies ist aber äquivalent zu der Forderung, daß e eine hinreichende Bedingung für s ist. Denn wenn es ein e' gibt, das die Bedingung (i) erfüllt, dann ist e keine hinreichende Bedingung für s . Und wenn e keine hinreichende (aber eine notwendige Bedingung) für s ist, dann gibt es eine Bedingung e' , die die Bedingung (i) erfüllt.

Dies ergibt sich aus der folgenden Überlegung. Wenn e keine hinreichende (aber eine notwendige Bedingung) für s ist, dann kommt es zumindest manchmal vor, daß e auftritt, während s nicht der Fall ist. Nennen wir diesen Fall e'' , dann gilt per definitionem: $e'' \leftrightarrow (\neg s \ \& \ e)$. Betrachten wir nun die Bedingung $(\neg e'' \ \& \ e)$. Offenbar gilt: $(\neg e'' \ \& \ e) \rightarrow e$. Und, da e eine notwendige Bedingung für s ist, gilt auch: $s \rightarrow (\neg e'' \ \& \ e)$. $(\neg e'' \ \& \ e)$ ist also eine notwendige Bedingung für s und e eine notwendige Bedingung für $(\neg e'' \ \& \ e)$. Aber $(\neg e'' \ \& \ e)$ ist keine notwendige Bedingung für e . Denn wenn das so wäre, müßte gelten: $e \rightarrow (\neg e'' \ \& \ e)$. Hieraus würde aber folgen $e \rightarrow (s \vee \neg e)$, und hieraus $e \rightarrow s$. D.h., wenn $(\neg e'' \ \& \ e)$ eine notwendige Bedingung für e wäre, dann wäre e eine hinreichende Bedingung für s im Widerspruch zur Voraussetzung. Wenn e eine notwendige, aber keine hinreichende Bedingung für s ist, gibt es also eine Bedingung - nämlich $(\neg e'' \ \& \ e)$ -, die die Bedingung (i) erfüllt.

Zu diesem Resultat kann man auch einfacher gelangen. Denn wenn e eine notwendige, aber keine hinreichende Bedingung für s ist, dann erfüllt auch s selbst die Bedingung (i). In diesem Fall benutzt man aber offenbar die Tatsache, daß aufgrund der Definitionen Dretskes s immer auch die Information s trägt. Dretske hat in einem Briefwechsel jedoch angedeutet, daß es sinnvoll wäre, diesen Fall auszuschließen.

8. Vgl. Dretske (1981, 190).

9. Fodor (1987, 102-104), Cummins (1989, 67-69).

10. In (1986) entwickelt Dretske zur Überwindung des Problems der Fehlrepräsentation eine alternative Definition des Inhalts repräsentationaler Zustände, die wegen ihrer funktional-teleologischen Komponenten schon einen Schritt in Richtung der Überlegungen von Millikan und Papineau darstellt.

11. Dies ist nicht genau die Formulierung Fodors, der unter Berücksichtigung der Probleme, die im Absatz zuvor angesprochen wurden, vielmehr folgende Formulierung vorschlägt (1987, 126):

(SLCCT) Ein Zustand Z in einem System S hat genau dann den repräsentationalen Inhalt e , wenn (1) alle $e \in Z$ in S verursachen, die in S psychophysische Spuren hervorrufen, zu denen S in einer psychophysisch optimalen Beziehung steht, und wenn (2), falls nicht- $e \in Z$ in S verursachen, diese Tatsache asymmetrisch davon abhängt, daß in der Regel $e \in Z$ in S verursachen.

12. Siehe Cummins (1989, 58ff.).

13. In Beckermann (erscheint demnächst) habe ich zu zeigen versucht, daß man Fodors Theorie trotz dieser Einwände vielleicht doch retten kann.

14. Siehe bes. Millikan (1984; 1989) und Papineau (1985; 1988).

15. Vgl. zu dieser Formulierung Cummins (1989, 76).

16. Bei den folgenden Überlegungen stütze ich mich stark auf Cummins (1989, 78ff.).

17. Der Ausdruck "Inhaltszuschreibung" ist im Zusammenhang mit den Theorien Dretskes, Fodors und Millikans eigentlich unangemessen. Denn diese Theorien gehen gerade davon aus, daß Inhalte nicht von uns zugeschrieben, sondern durch objektive (in diesem Fall natürliche) Bedingungen festgelegt werden. Der meßtheoretische Ansatz, den ich im folgenden erläutern werde, gehört demgegenüber zur Gruppe der Theorien, die davon ausgehen, daß Inhaltszuschreibungen nicht durch objektive Tatsachen determiniert sind, sondern auf Interpretationen beruhen. (Diesen Hinweis verdanke ich Peter Lanz.)

18. Vgl. Beckermann (1991, 95).

19. Zur Verwendung des Ausdrucks "Zuschreibung" in diesem Zusammenhang vgl. oben Anm. 17.

20. Vgl. Churchland (1979, 100-107), Field (1980, 114), Stalnaker (1984, 9ff.), Davidson (1974, 147; 1989, 9ff.; erscheint demnächst). Vgl. auch Dennett (1982, 123ff.; 1987b, 208) und Matthews (1990). Field scheint der erste gewesen zu sein, der die Auffassung vertreten hat, die meßtheoretische Deutung intentionaler Prädikate könne möglicherweise zu einer Lösung des Problems der Intentionalität führen.

Eine Auseinandersetzung mit dem meßtheoretischen Ansatz findet sich in Lanz (1987, 95-127); auf die genannte Literatur geht Peter Lanz auch ausführlich in einem neuen noch unveröffentlichten Manuskript ein.

21. Diese Auffassung ist von P. Churchland in (1979, 100ff.) ausführlich erläutert worden. Zum ersten Mal erwähnt wird diese Möglichkeit meines Wissens in Quine (1970, Abs. 2.11).

22. Hauptthemen der Meßtheorie sind das sogenannte Repräsentations- und das Eindeutigkeitsproblem. Beim Repräsentationsproblem geht es um die Frage der Existenz strukturerhaltender Homomorphismen, d.h. genauer um die Frage, welche Bedingungen die Relation R und die Operation $*$ erfüllen müssen, damit sich beweisen läßt, daß solche Homomorphismen existieren. Beim Eindeutigkeitsproblem geht es dann darüberhinaus um die Frage, inwieweit diese Homomorphismen eindeutig sind.

Für extensive Größen wie Masse und Länge z.B. hat die Meßtheorie gezeigt, daß es eine bis auf die Multiplikation mit einer positiven reellen Zahl eindeutige Abbildung gibt, die den Bedingungen (1) und (2) genügt, wenn

die Relation R und die Operation $*$ folgende Bedingungen erfüllen (" $a \equiv b$ " ist dabei eine Abkürzung für " $R(a,b)$ und $R(b,a)$ "):

- | | | |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| (1) | $R(a,b)$ oder $R(b,a)$ | (Konnektivität) |
| (2) | Wenn $R(a,b)$ und $R(b,c)$, dann $R(a,c)$ | (Transitivität) |
| (3) | $(a * b) * c \equiv a * (b * c)$ | (Assoziativität) |
| (4) | $a * b \equiv b * a$ | (Kommutativität) |
| (5) | Wenn $R(a,b)$, dann $R(a * c, b * c)$ | (Monotonie) |
| (6) | Wenn $R(a,b)$, dann gibt es ein x aus B mit $b \equiv a * x$ | (Existenz der Differenz) |
| (7) | $R(a, a * b)$ | |
| (8) | Wenn $R(a,b)$, dann gibt es eine natürliche Zahl n mit $R(b, na)$ (dabei ist $n a$ rekursiv definiert durch $1a = a$ und $(n+1)a = na * a$) | |

23. Daß a schwerer ist als b , soll hier heißen, daß a b aufwiegt, aber nicht b a , und daß a gleich schwer ist wie b , soll heißen, daß a b aufwiegt und b a aufwiegt.

24. Genauer gesagt, der Geschwindigkeitszuwachs.

25. Dieser Punkt wird in Lanz (1987) mehrfach betont. Die Überlegungen im nächsten Abschnitt zeigen jedoch, daß man dies auch nicht unbedingt erwarten konnte.

26. Diese Kritik wurde in ähnlicher Form schon in Lanz (1987, 117ff.) formuliert.

27. Spitze Klammern dienen hier - als Alternative zu daß-Sätzen - zur Erzeugung von Ausdrücken, die Propositionen bezeichnen. Der Ausdruck " $\langle p \rangle$ " bezeichnet also die Proposition, daß p .

28. Auch zu diesem Punkt findet sich eine ähnliche Kritik in Lanz (1987, 113ff.). Lanz sieht eine weitere Disanalogie zwischen der Meßtheorie und der intentionalen Psychologie darin, daß sich die empirischen Relationen, die zwischen Gegenständen bestehen müssen, damit bestimmte ihrer Eigenschaften metrisierbar sind, ohne Gebrauch von Zahlen spezifizieren lassen, während entsprechende Relationen zwischen intentionalen Zuständen nicht ohne den Bezug auf Propositionen spezifizierbar sind. (Vgl. 1987, 107ff.)

29. Vgl. Loar (1981). Eine kurze und präzise Zusammenfassung der Ideen Loars findet sich in Schiffer (1986) und (1987, ch. 2).

30. Diese Abbildungen können auch mehrdeutig sein.

Literatur

Beckermann, A. (1991) "Der endgültige Todestoß für den Repräsentationalismus? Eine Replik auf Andreas Kemmerlings Artikel 'Mentale Repräsentation'". *Kognitionswissenschaft* 2, 91-98.

Brentano, F. (1924) *Psychologie vom empirischen Standpunkt*. Hrsgg. von O. Kraus. Leipzig: Meiner Verlag.

Churchland, P.M. (1979) *Scientific Realism and the Plasticity of Mind*. Cambridge: Cambridge University Press.

Cummins, R. (1989) *Meaning and Mental Representation*. Cambridge, Mass.: MIT Press.

Davidson, D. (1974) "Belief and the Basis of Meaning". *Synthese* 27, 309-323. Wiederabgedr. in: Davidson, D. *Inquiries into Truth and Meaning*. Oxford. Clarendon Press 1984, 141-154.

--- (1989) "What is Present to the Mind?". *Grazer Philosophische Studien* 36, 3-18.

--- (erscheint demnächst) "Reply to Lanz", in: Bieri, P. und Stöcker, R. xxx, Berlin/New York: Walter de Gruyter.

Dennett, D. (1982) "Beyond Belief", in: Woodfield, A. (ed.), *Thought and Object*. Oxford: Clarendon Press, 1-95. Wiederabgedr. in: Dennett (1987a), 117-202.

- (1987a) *The Intentional Stance*. Cambridge, Mass.: MIT Press.
- (1987b) "About Aboutness", in: Dennett (1987a), 203-211.
- Dretske, F. (1981) *Knowledge and the Flow of Information*. Oxford: Basil Blackwell.
- (1986) "Misrepresentation", in Bogdan, R.J. *Belief - Form, Content, and Function*. Oxford: Clarendon Press, 17-36.
- Field, H. (1980) "Postscript zu 'Mental Representation'", in: Block, N. (ed.) *Readings in the Philosophy of Psychology. Vol. 2*. Cambridge, Mass.: MIT Press, 112-114.
- Fodor, J. (1987) *Psychosemantics*. Cambridge, Mass.: MIT Press.
- (1991) *A Theory of Content and Other Essays*. Cambridge, Mass.: MIT Press.
- Haugeland, J. (1981) "Semantic Engines", in: Haugeland, J. (ed.) *Mind Design*. Cambridge, Mass.: MIT Press, 1-34.
- (1985) *Artificial Intelligence. The Very Idea*. Cambridge, Mass.: MIT Press.
- Hempel, C.G. (1952) *Fundamentals of Concept Formation in Empirical Science*. Chicago
- Lanz, P. (1987) *Menschliches Handeln zwischen Kausalität und Rationalität*. Frankfurt/M.: Athenäum.
- Loar, B. (1981) *Mind and Meaning*. Cambridge: Cambridge University Press.
- Mathews, R. (1990) "The Measure of Mind", Report No. 57/1990 Research Group on MIND AND BRAIN, ZiF (Bielefeld).
- Milikan, R. (1984) *Language, Thought, and Other Biological Categories*. Cambridge, Mass.: MIT Press.
- (1989) "Biosemantics". *Journal of Philosophy* 86, 281-297.
- Quine, W.V.O. (1970) *Philosophy of Logic*. Englewood Cliffs, NJ: Prentice-Hall.
- Papineau, D. (1985) "Representation and Explanation". *Philosophy of Science* 51, 550-572.
- (1988) *Reality and Representation*. Oxford: Basil Blackwell.
- Schiffer, S. (1986) "Functionalism and Belief", in: Brand, M. and Harnish, R. (eds.) *The Representation of Knowledge and Belief*. Tuscon, S. 127-159.
- (1987) *Remnants of Meaning*. Cambridge, Mass.: MIT Press.
- Searle, J. (1983) *Intentionality. An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press. Dt. Übers.: *Intentionalität. Eine Abhandlung zur Philosophie des Geistes*. Frankfurt/M.: Suhrkamp 1987.
- Stalnaker, R. (1984) *Inquiry*. Cambridge, Mass.: MIT Press.
- Suppes, P. and J. Zinnes (1963) "Basic Measurement Theory", in: Luce, R.D. et. al. (eds.) *Handbook of Mathematical Psychology, Bd. I*. New York.

Adresse

Prof. Dr. Ansgar Beckermann, Universität Göttingen, Philosophisches Seminar, Platz der Göttinger Sieben 5, D(W)-3400 Göttingen

Kritik

Ist der meßtheoretische Ansatz hinreichend bestimmt?

Godehard Brüntrup

((1)) Der von Prof. Beckermann (B.) präzise entwickelte meßtheoretische Ansatz (mA) ist komplex in die Debatte des Leib-Seele-Problemes verwoben. Jeder "Zug" auf diesem "Spielfeld" bringt eine Explosion von Folgeproblemen mit sich. Um der Konsequenzen überhaupt ansichtig zu werden, muß man sich der Topologie des Spielfeldes vergewissern und Interdependenzen aufdecken. Im vorliegenden Falle erweist sich diese Aufgabe aber als durchaus komplex, da man den mA sehr verschieden interpretieren kann. Es ist bemerkenswert, daß Philosophen, die auf sehr entfernten Positionen dieses Spielfeldes stehen, jeweils Sympathien mit dem mA bekunden. Eine meiner Fragen ist, ob nicht der mA genau in dem Maße an philosophischer Relevanz verliert, wie er sich scheinbar problemlos von so verschiedenen Autoren wie Davidson und den Churchlands inkorporieren läßt ((36))? In diesem Sinne sind die folgenden kurzen Bemerkungen weniger als eine Kritik, sondern mehr als eine Anfrage mit folgendem Inhalt gedacht: Kann die Leistungsfähigkeit des mA nicht erst wirklich abgeschätzt werden, wenn seine Einbettung in das jeweilige epistemologische und ontologische Umfeld explizit thematisiert wird? Wenn man den mA unter dieser Rücksicht auf seine philosophischen Affinitäten untersucht, so scheint er mir *prima facie* gegenläufige Tendenzen in sich zu vereinen, weil er einerseits Anleihen bei nicht-realistischen Theorien in der Philosophie des Geistes macht (Abstraktionismus, Fiktionalismus), andererseits einen funktionalistisch abgesicherten Realismus behauptet. Ich will diese Fragen hier - so gut es auf diesem engen Raum geht - etwas weiter präzisieren.

((2)) Eine wesentliche Unterbestimmtheit des Ansatzes sehe ich in der Frage des Bezuges zwischen den physischen Zuständen des Systems und dem intentionalen Vokabular. Auch B. arbeitet in den Abschnitten ((37)) - ((42)) heraus, daß die feinkörnige Aufklärung der Instantiierung von im Meßraum spezifizierten intentionalen Zuständen in bestimmten physischen Zuständen des kognitiven Systemes viel schwieriger durchzuführen ist als bei dem Fall der Temperatur von Körpern. Es soll damit nicht unterstellt werden, daß letzteres Problem philosophisch harmlos ist. Nicht umsonst propagiert ein Vertreter des mA (H. Field) nicht nur "thoughts without content", sondern auch "physics without numbers". Aber das philosophische und empirische Verständnis der Instantiierung und Realisation intentionaler Prädikate ist aus den genannten Gründen ungleich komplexer (kausale Relevanz der Zustände, multiple Instantiierung in verschiedenen Systemen, Einbezug externer Kausalbeziehungen). Zunächst wird die Frage nach der Realisierung der intentionalen Prädikate beim sogenannten "Repräsentationsproblem"